Technische Universität München

TUM School of Computation, Information and Technology

Robust Transmitter Design for FDD Massive Antenna Systems with Limited Channel Knowledge

Donia Ben Amor

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades einer

Doktorin der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz:		Prof. Dr. Andreas Herkersdorf
Prüfende der Dissertation:	1.	Prof. Dr.-Ing. Wolfgang Utschick
	2.	Prof. Rodrigo de Lamare

Die Dissertation wurde am 20.11.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 11.07.2025 angenommen.

1 Acknowledgments

First of all, I would like to express my deepest gratitude to Professor Wolfgang Utschick for providing me with the opportunity to pursue my doctoral degree at the Chair of Methods of Signal Processing and for his continuous guidance and support on both a technical and personal level.

I would also like to sincerely thank Dr. Michael Joham for being a great supervisor and mentor, and for constantly pushing me to grow and improve. I am deeply grateful for his valuable discussions, steady support, and constructive feedback.

My special thanks go to Professor Bruno Clerckx for giving me the opportunity to spend a research stay at his chair at Imperial College London in October 2021, which was a highly enriching experience.

Additionally, I would like to thank all colleagues from the MSV research group for creating such a friendly and collaborative work environment. In particular, I am very thankful to Hela Jedda, from whom I learned a great deal and with whom I shared my first days in the office. I would also like to thank my former office mate and now dear friend, Valentina Rizzello, for the enjoyable time we spent together and for her kind help in reviewing this thesis.

I am grateful to all the students I had the pleasure to supervise—Marc, Florian, Dominik, Colin, Orestis, Srikar, Yuxi, Yatish, Sadaf, and Bassem—for their commitment, curiosity, and valuable contributions. Working with them has been an enriching experience that taught me as much as I hope I was able to teach in return.

My heartfelt thanks go to my parents, whose unconditional love, sacrifices, and faith in me have been my greatest motivation. I am equally thankful to my siblings, Marwa, Youssef, Mohamed, and Mondher, for their constant encouragement and support throughout the years.

I would also like to express my sincere gratitude to my best friend, Wafa, for her unwavering friendship.

Last but certainly not least, my deepest appreciation goes to my partner, Fahmi, for his patience, optimism, and endless support through both the good and challenging moments. His cheerfulness and belief in me have been a continuous source of strength and joy.

Abstract

In this doctoral thesis, we investigate robust transmit strategies for massive multiple-input-single-output (MISO) systems in frequency-division duplex (FDD) mode, addressing the challenges of incomplete channel state information (CSI) at the base station. Massive MISO systems promise enhanced spectral and energy efficiency; however, these advantages are contingent upon the accurate channel knowledge at the transmitter. Our research develops various precoding techniques that effectively manage limited CSI at the base station, offering a balance between computational complexity and system throughput.

We start by analyzing zero-forcing precoding under conditions of incomplete CSI, revealing surprising performance characteristics in high signal-to-noise ratio (SNR) environments, particularly regarding the impact of pilot design on achievable degrees of freedom, i.e., the number of interference-free served users. In scenarios with highly limited pilot numbers, we explore statistical precoding approaches that leverage second-order channel statistics to enhance performance while reducing complexity. Our work also includes the design of iterative precoding algorithms based on various optimization frameworks, such as minorization maximization (MM) and augmented weighted average mean square error (AWAMSE) approaches, which facilitate effective precoder design for downlink sum rate maximization. We extend these techniques to rate-splitting multiple access (RSMA) setups, demonstrating improved robustness against CSI inaccuracies. Furthermore, we address the joint optimization of pilot and precoder using deep learning techniques based on graph neural networks (GNNs), illustrating the superiority of the learned pilots over conventional orthogonal ones.

Overall, this study provides valuable insights and practical solutions for advancing massive MISO systems in FDD mode, emphasizing the critical need for efficient transmit design methods that can adapt to limited or imperfect CSI conditions while maintaining high system performance.

Contents

2 Introduction and Motivation

In the rapidly evolving field of wireless communications, massive multiple-input-multiple-output (MIMO) systems have emerged as a key technology for enhancing both spectral efficiency and energy efficiency [1], [2], [3] and are expected to play a crucial role in future 6G networks. This technology, which utilizes a large number of antenna elements at the base station, enables the simultaneous service of multiple users within the same time-frequency resource. As a result, it dramatically increases network capacity [4].

The potential benefits of massive MIMO are evident, but its practical implementation faces several challenges, particularly in frequency-division duplexing (FDD) systems [5]. FDD, which utilizes separate frequency bands for uplink and downlink transmissions, is the preferred duplexing mode in many existing cellular networks due to its advantages in coverage and compatibility with legacy systems [6]. However, implementing massive MIMO in FDD systems introduces unique challenges, especially in acquiring accurate channel state information (CSI) at the base station.

In FDD systems, the channel reciprocity that simplifies CSI acquisition in time-division-duplex (TDD) systems is not present [7]. This is because the uplink-downlink frequency gap exceeds the channel coherence bandwidth. Consequently, the base station must send pilots in the downlink and depend on subsequent CSI feedback from users [8] to acquire knowledge about the downlink channel. This inevitable feedback can lead to significant overhead, particularly as the number of antennas increases [9]. Such overhead not only consumes valuable spectral resources but also introduces latency, which can degrade system performance, especially in fast-fading environments with short coherence time. Moreover, the limited feedback bandwidth often results in quantization errors, further compromising the accuracy of the CSI available at the base station [10]. The challenges of FDD massive MIMO systems underscore the critical need for efficient precoder design methods that can cope with limited or imperfect CSI.

2.1 Related Works

While non-linear precoding techniques are generally more effective, they come with high computational complexity, which makes them challenging to implement in large-scale systems [11]. On the other hand, linear precoders offer a trade-off between performance and complexity, and early works demonstrated the potential of zero-forcing [12] and minimum mean-square-error (MMSE) [13] precoding in handling interference when perfect CSI is available at the transmitter.

Later works have considered the more practical scenario of partial CSI and investigated effective precoding approaches that can balance complexity and performance in this setting. [14] and [15] explored robustness in spectral and energy efficiency, mainly in TDD systems, while [16] focused on improving precoding for FDD systems with limited feedback. More recently, [17] and [18] proposed to leverage statistical CSI to boost the performance in imperfect CSI conditions. Exploiting second-order channel statistics can help with interference mitigation and power control at the transmitter with reduced CSI requirements. [19] and [20] developed statistical precoding frameworks, which model the effect of channel covariance matrices on the achievable rates and interference patterns.

In addition to statistical approaches, iterative algorithms for designing effective precoding strategies have gained significant attention [21]. Designing the optimal precoder for downlink sum rate maximization is known to be an NP-hard optimization problem [22], even in the perfect CSI case. The key idea of the iterative approaches is to smartly introduce sets of auxiliary variables, i.e., to augment the dimension of the space of interest. Despite the increased number of variables to be optimized, the reformulated optimization problem is more tractable, as it can be solved more easily for each set of variables (including the precoders) once the other sets of variables are fixed. A widely known method is the iterative weighted MMSE (IWMMSE) algorithm [23], [24], which leverages the MSE-rate relationship to reformulate the sum rate maximization problem as a three-block sum weighted MSE minimization problem that is convex in each of its blocks. Further works proposed accelerated versions of the IWMMSE algorithm [25], while others used model-based deep unfolding to mimic its alternating behavior [26].

Another approach based on the so-called Lagrangian dual transform and fractional programming (FP) was introduced in [27] for the perfect CSI case and operates directly on the instantaneous sum rate maximization problem. It utilizes the Lagrangian duality theory and a multi-ratio optimization framework denoted by the quadratic transform to develop an iterative algorithm where the precoder update is given in quasi-closed form.

Moreover, [28] proposed to apply the minorization maximization (MM) method [29] to solve the weighted sum rate maximization problem for a perfect CSI scenario featuring a reconfigurable intelligent surface. This method consists of iteratively solving a sequence of cleverly established surrogate problems. Recently, [21] showed that a connection between these algorithms can be found despite the difference in the underlying solution approach.

While all of the above-mentioned precoding methods were designed for spatial division multiple access (SDMA), rate splitting multiple access (RSMA) has been proven to enhance the system throughput, as it combines the advantages of SDMA and multi-casting, [30], [31]. The benefits of RSMA have been observed in scenarios with perfect CSI, e.g., in [32], [33], and more importantly, when only limited channel knowledge is available at the base station [34], [35]. Specifically, downlink RS is based on splitting each user's message into a private part and a common part that is decoded by all users, and therefore, it allows the users to partially decode the interference and partially treat it as noise. The potential of RS to enhance the degrees of freedom can be exploited by optimizing the resource allocation and transmission strategy [36]. RS is less sensitive to CSI inaccuracies than traditional precoding methods, as it leverages the common stream to offset errors in the private streams. By reducing the dependency on high-accuracy CSI, RS maintains stable performance under limited feedback and can manage the feedback-related constraints typical in massive MIMO FDD systems. One-layer RS, where the users are equipped with one successive interference cancellation (SIC) unit, offers a trade-off between complexity and notable throughput improvements [37], [38]. In the RS setup, the power allocation across the common and private streams can be dynamically adjusted based on users' channel conditions, optimizing performance despite limited CSI, especially in high-signal-to-noise-ratio (SNR) environments when the system is interference-limited.

Designing an efficient and robust precoding technique is essential for enhancing the system throughput, whereby the quality of the channel knowledge at the base station plays a central role. Alongside the precoder, the design of the downlink pilots used for channel sounding can influence the overall system performance. [39], [40] proposed to exploit channel sparsity and to use compressed sensing techniques to reduce the pilot overhead. A further study in [41] leveraged deep learning models to bypass explicit channel estimation and jointly design the pilots and precoder for a massive multiple-input-single-output (MISO) FDD system.

2.2 Key Contributions

This dissertation presents several novel approaches to address the challenges of massive MISO FDD systems. The key contributions are summarized as follows:

1. A comprehensive analysis of zero-forcing precoding under incomplete CSI, revealing unexpected performance characteristics in high-power regimes.

2. Development of a statistical precoding method that utilizes channel second-order statistics, reducing the computational overhead while outperforming the zero-forcing precoder in scenarios with highly limited pilot resources.

3. Introduction of efficient iterative precoding algorithms, including the augmented weighted average MSE (AWAMSE) and MM methods, which leverage instantaneous and long-term statistical channel information.

4. Extension of the proposed transmit strategies to the RSMA setup, demonstrating notable performance gains in high-power regimes with low computational complexity.

5. An innovative deep-learning framework based on graph neural networks (GNNs) for joint precoder and pilot design, achieving superior performance compared to conventional orthogonal pilots.

2.3 Outline

The thesis is organized as follows:

- Chapter 3 lays the background for the considered FDD MISO system, where the users are equipped with a single antenna. Here, we establish a model for both the training and data transmission phases of the downlink for the previously mentioned MA schemes, namely SDMA and RSMA. Moreover, depending on the assumptions regarding the channel knowledge at the base station, we derive performance metrics that will serve as utility functions for later system optimization.

- In Chapter 4, we present optimization frameworks that will be used in subsequent chapters to solve the precoder design problem for downlink sum rate maximization. These techniques include the Lagrangian dual transform and the

FP method [42], the IWMMSE approach [23], and the MM framework. Furthermore, we establish the so-called "rescale and optimize" approach, inspired by [43], which will become useful in several optimization problems.

- In Chapter 5, we analyze the zero-forcing precoder under incomplete CSI conditions and show the rather surprising result that combining this precoder with the Least Squares (LS) channel estimate yields an interference-free transmission in the asymptotic high-SNR regime under certain conditions on the downlink training pilots [44]. Although the linear MMSE (LMMSE) channel estimate exhibits a better estimation quality, applying it for zero-forcing fails to eliminate the inter-user interference. We additionally show that the number of deployed pilots in the system presents an upper bound for the achievable degrees of freedom.

- In Chapter 6, a statistical precoding approach is presented [45]. We show that leveraging the channel second-order statistics for precoder design can reduce the related computational complexity and offer performance gains in the highly limited channel knowledge case where even the zero-forcing precoder based on LS fails to manage the interference. Furthermore, we discuss our findings in [46] concerning the extension of the statistical precoder to the RSMA setup.

- Chapter 7 is concerned with different iterative precoding algorithms, namely the MM approach [47], the AWAMSE approach [48] and its extension to the RSMA setup [49]. These methods are established based on the so-called training-based lower bound, which makes use of the instantaneous channel knowledge as well as the channel second-order statistics. For the optimization procedure, we adopt several techniques from Chapter 4.

- While the previous chapters assume fixed orthogonal pilots, Chapter 8 addresses the pilot design problem. Our work [50] delivered a deep-learning framework for joint pilot and precoder optimization based on GNNs.

2.4 Notation

$(\bullet)^{\mathrm{T}}$	Transpose of a matrix
$(\bullet)^{\mathrm{H}}$	Hermitian of a matrix
$(\bullet)^{-1}$	Inverse of a matrix
$(\bullet)^{*}$	Complex Conjugate
$(\bullet)^{\dagger}$	Pseudo-inverse of a matrix
$\| \bullet \|_{\mathrm{F}}$	Frobenius-norm of a matrix
tr$(\bullet)$	trace of a matrix
vec$(\bullet)$	Vectorization operator
diag$(a_1, \ldots, a_K)$	Diagonal matrix having $a_1, \ldots, a_K$ as diagonal elements
blkdiag$(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_K)$	Block diagonal matrix having $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_K$ on the main diagonal
$\boldsymbol{I}_K$	Identity matrix of dimension K
$\mathbf{1}_K$	All-ones vector of dimension K
$\boldsymbol{e}_k$	kth column of the identity matrix
$\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C})$	Complex-valued circularly-symmetric Gaussian distribution with zero-mean and covariance matrix $\boldsymbol{C}$
$\mathbb{E}[\bullet]$	Expectation Operator
$\mathbb{E}[x\|y]$	Conditional expectation of x given y
$\Re\{\bullet\}$	Real part of a complex-valued number
$\angle(\bullet)$	Phase of a complex-valued number
$\| \bullet \|$	Magnitude of a complex-valued number
$\otimes$	Kronecker product
$\odot$	Hadamard product
$\boldsymbol{a} \perp \boldsymbol{b}$	Vector $\boldsymbol{a}$ is orthogonal to vector $\boldsymbol{b}$
$[\boldsymbol{X}]_{i,j}$	Element of matrix $\boldsymbol{X}$ in ith row and jth column

3 System Model and Setup

This work primarily focuses on the downlink in a single-cell setup, operating in FDD mode, where the base station is equipped with several antennas $M \gg 1$ and serves K single-antenna users.

3.1 Channel Model

The transmission channel between the base station and some generic user k is denoted by $\boldsymbol{h}_k \in \mathbb{C}^M$, where the mth element corresponds to the channel response between the mth base station antenna and the kth user. Under the block fading assumption, the channel response is considered static during the so-called coherence block of T time-frequency symbols. Furthermore, we assume the following model for the transmission channel

$$\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_k) \tag{3.1}$$

corresponding to a correlated Rayleigh fading model, where $\boldsymbol{C}_k$ denotes the channel covariance matrix. On the one hand, the microscopic fading effects are modeled by the Gaussian distribution, and the channel takes a new realization from this distribution at each coherence interval. On the other hand, the macroscopic effects are described by the covariance matrix. Here, large eigenvalues of $\boldsymbol{C}_k$ correspond to channel directions with strong spatial correlation.
We assume that the second-order statistics are perfectly known to the base station. The interested reader can refer to [51] and [45] for an insight into covariance matrix estimation in the FDD setup.

3.2 Downlink Training Phase

The data transmission in the uplink and downlink operations occurs over two different frequency bands in the FDD setup. The gap between the uplink and downlink carrier frequencies is usually higher than the channel coherence bandwidth, leading to a non-reciprocity between the uplink and downlink channels. Thus, channel estimates acquired in the uplink cannot be directly used for beamforming design, as it is usually

done in TDD systems. Instead, the base station sends known downlink pilots and requests the users to feed back their observations/estimates, as depicted in Fig. 3.1. For the sake of simplicity, we assume that the feedback of the channel measurements from the users to the base station takes place via an analog feedback link [8].

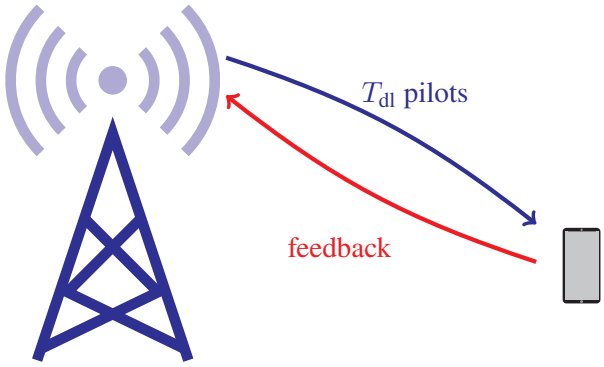

Fig. 3.1: Illustration of the downlink channel training phase

To train the M antennas at the base station, the number of downlink pilots T_{dl} should be at least equal to M. However, due to the large number of base station antennas and the limited channel coherence interval, dedicating M out of the T time-frequency symbols for channel probing would leave no room for payload data transmission. Therefore, we consider the case where $T_{\text{dl}} < M$, leading to incomplete channel knowledge at the base station. During the downlink training phase, the base station sends T_{dl} orthogonal unit-norm pilots from each antenna, and the $M \times T_{\text{dl}}$ pilot matrix is denoted by $\boldsymbol{\Phi}$. Due to the orthogonality of the pilots, it holds that $\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi} = \boldsymbol{I}_{T_{\text{dl}}}$ but in general $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}} \neq \boldsymbol{I}_M$, if $T_{\text{dl}} < M$.

After the downlink training phase, the T_{dl} channel observations at each user k are collected in $\boldsymbol{y}_k$ as follows

$$\boldsymbol{y}_k = \boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k} \tag{3.2}$$

where $\boldsymbol{n}_{1,k} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma_{1,k}^2 \boldsymbol{I}_{T_{\text{dl}}})$ stands for the downlink training noise with the normalized variance $\sigma_{1,k}^2 = \frac{1}{P_{\text{dl}}}$. Here, P_{dl} denotes the downlink transmit power.
The covariance matrix of $\boldsymbol{y}_k$ denoted by $\boldsymbol{C}_{\boldsymbol{y}_k}$ is given by

$$\boldsymbol{C}_{\boldsymbol{y}_k} = \boldsymbol{\Phi}^{\text{H}}\boldsymbol{C}_k\boldsymbol{\Phi} + \sigma_{1,k}^2 \boldsymbol{I}_{T_{\text{dl}}}. \tag{3.3}$$

Throughout this thesis, we will assume an interference-free noiseless feedback link, and the channel observations in (3.2) are available at the base station after the feedback phase [51].

In the case of a noisy feedback link, an additive white Gaussian noise portion $\boldsymbol{n}_{2,k}$ can be considered, whose variance depends on the transmit power P_{fb} the users employ for the feedback transmission [51]. Thus, the following observation is received at the base station

$$\tilde{\boldsymbol{y}}_k = \boldsymbol{y}_k + \boldsymbol{n}_{2,k} = \boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k} + \boldsymbol{n}_{2,k} \tag{3.4}$$

where $\boldsymbol{n}_{2,k} \sim (\boldsymbol{0}, \sigma_{2,k}^2 \boldsymbol{I}_{T_{\text{dl}}})$ and $\sigma_{2,k}^2 = \frac{1}{P_{\text{fb}}}$.

3.2.1 Channel Estimation

In the following, depending on the precoding approach, the base station uses the channel observations directly for beamforming or explicitly computes channel estimates based on the received measurements. In the latter case, we consider two linear estimators: the LS and the LMMSE estimators.

In the case of explicit channel estimation, the channel of user k can be decomposed as follows [52]

$$\boldsymbol{h}_k = \hat{\boldsymbol{h}}_k + \tilde{\boldsymbol{h}}_k \tag{3.5}$$

where $\hat{\boldsymbol{h}}_k$ and $\tilde{\boldsymbol{h}}_k$ denote the channel estimate and the zero-mean estimation error, respectively.

3.2.1.1 Least Squares Channel Estimator

Based on the channel observation vector $\boldsymbol{y}_k$, the base station can compute the LS channel estimate [53] $\hat{\boldsymbol{h}}_k^{\text{LS}}$ of user k according to

$$\hat{\boldsymbol{h}}_k^{\text{LS}} = \underset{\hat{\boldsymbol{h}}_k}{\arg\min} \, \|\boldsymbol{y}_k - \boldsymbol{\Phi}^{\text{H}}\hat{\boldsymbol{h}}_k\|^2 = (\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}})^{\dagger}\boldsymbol{\Phi}\boldsymbol{y}_k. \tag{3.6}$$

Since $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}$ defines an orthogonal projection matrix (recall that $\boldsymbol{\Phi}$ is an orthogonal matrix), it holds that $(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}})^{\dagger} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}$ and the LS estimate is given by

$$\hat{\boldsymbol{h}}_k^{\text{LS}} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi}\boldsymbol{y}_k = \boldsymbol{\Phi}\boldsymbol{y}_k. \tag{3.7}$$

The second equality follows from the property $\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi} = \boldsymbol{I}_{T_{\text{dl}}}$.

The LS estimator does not require prior statistical knowledge of the channel but depends on the used transmit power. We note that due to the suboptimality of the LS estimator (in the MSE sense), the LS estimate and the corresponding estimation error are correlated, i.e.,

$$\mathbb{E}[\hat{\boldsymbol{h}}_k^{\text{LS}}\tilde{\boldsymbol{h}}_k^{\text{LS,H}}] = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}(\boldsymbol{C}_k\boldsymbol{P}_{\boldsymbol{\Phi}} - \sigma_{1,k}^2\boldsymbol{I}_{T_{\text{dl}}}) \tag{3.8}$$

where $\boldsymbol{P}_{\boldsymbol{\Phi}} = \boldsymbol{I}_M - \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}$.
The covariance matrix of the LS estimation error can be written as

$$\boldsymbol{C}_{\mathrm{err},k}^{\mathrm{LS}} = \mathbb{E}[\tilde{\boldsymbol{h}}_k^{\mathrm{LS}} \tilde{\boldsymbol{h}}_k^{\mathrm{LS,H}}] = \boldsymbol{P}_{\boldsymbol{\Phi}} \boldsymbol{C}_k \boldsymbol{P}_{\boldsymbol{\Phi}} + \sigma_{1,k}^2 \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}} \tag{3.9}$$

where we used the fact that the estimation error is zero-mean.

3.2.1.2 Linear Minimum Mean Square Error Estimator

If, in addition to the observations $\boldsymbol{y}_k$, the channel statistics are known to the base station, the latter can compute a better quality channel estimate than the LS estimate, namely the LMMSE channel estimate given by [7]

$$\hat{\boldsymbol{h}}_k^{\mathrm{LMMSE}} = \boldsymbol{G}_k^* \boldsymbol{y}_k \tag{3.10}$$

where the matrix $\boldsymbol{G}_k^*$ is the solution of the following optimization problem

$$\boldsymbol{G}_k^* = \underset{\boldsymbol{G}_k}{\arg\min}\, \mathbb{E}[\|\boldsymbol{h}_k - \boldsymbol{G}_k \boldsymbol{y}_k\|^2] = \boldsymbol{C}_k \boldsymbol{\Phi} \boldsymbol{C}_{\boldsymbol{y}_k}^{-1}.$$

The LMMSE estimate of the kth user's channel hence reads as

$$\hat{\boldsymbol{h}}_k^{\mathrm{LMMSE}} = \boldsymbol{C}_k \boldsymbol{\Phi} \boldsymbol{C}_{\boldsymbol{y}_k}^{-1} \boldsymbol{y}_k. \tag{3.11}$$

Contrary to the LS estimate, the LMMSE estimate is uncorrelated with its corresponding estimation error [14]. This property will be used later when deriving some utility functions for precoder design.
The error covariance matrix is computed according to

$$\boldsymbol{C}_{\mathrm{err},k}^{\mathrm{LMMSE}} = \mathbb{E}[\tilde{\boldsymbol{h}}_k^{\mathrm{LMMSE}} \tilde{\boldsymbol{h}}_k^{\mathrm{LMMSE,H}}] = \boldsymbol{C}_k - \boldsymbol{G}_k^* \boldsymbol{\Phi}^{\mathrm{H}} \boldsymbol{C}_k \tag{3.12}$$

where we again used the zero-mean property of the estimation error.

3.3 Downlink Data Transmission Phase

In this work, we will consider two different schemes for the downlink data transmission phase, namely SDMA and RSMA. In Section 3.3.1, we will start by considering the more commonly used MA scheme, namely SDMA, and derive the related spectral efficiency expressions. The discussion on RSMA will be provided in Section 3.3.2, where several derivations from the SDMA scheme can be easily extended.

3.3.1 SDMA Scheme

We consider a linearly precoded SDMA scheme, where each user's message W_k is encoded into a single stream s_k, and the encoded messages are then transmitted to the users in the downlink using linear precoders. The users are, therefore, separated using the precoders, also called spatial beamformers, hence the name of this MA scheme.

The transmit vector at the base station antennas is given by the superposition of the precoded data streams and reads as

$$\boldsymbol{x} = \sum_{j=1}^{K} \boldsymbol{p}_j s_j \tag{3.13}$$

where the data streams $s_j \sim \mathcal{N}_{\mathbb{C}}(0,1), j = 1, \ldots, K$ are assumed to be mutually independent. $\boldsymbol{p}_j$ stands for the beamforming or precoding vector corresponding to the jth user.

The transmit vector and, hence, the precoders are subject to the transmit power constraint, namely

$$\mathbb{E}[\|\boldsymbol{x}\|^2] = \sum_{j=1}^{K} \|\boldsymbol{p}_j\|^2 \leq P_{\text{dl}} \tag{3.14}$$

where the expectation is w.r.t. the distribution of the data signals and P_{dl} denotes again the downlink transmit power.

The precoder maps the data symbol to the transmit antennas and determines its transmit directivity (see Fig. 3.2).

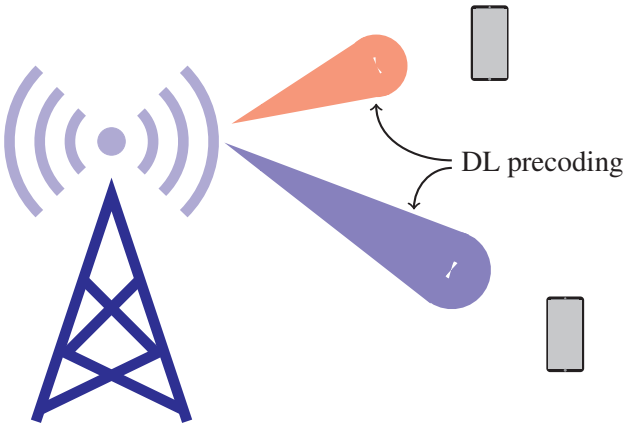

Fig. 3.2: Illustration of the downlink data transmission phase

The signal received at the kth user antenna is hence written as

$$r_k = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{x} + v_k. \tag{3.15}$$

Here, $v_k \sim \mathcal{N}_{\mathbb{C}}(0,1)$ is the additive white Gaussian noise experienced at the kth user's antenna, which is independent of the data signals s_j.
Using (3.13), the received signal is equivalently given by

$$r_k = \boldsymbol{h}_k^{\mathrm{H}} \sum_{j=1}^{K} \boldsymbol{p}_j s_j + v_k. \tag{3.16}$$

In the following chapters, we will study different strategies for designing the transmit precoding vectors. To this end, our ultimate goal is to maximize the spectral efficiency, i.e., the average number of bits of information that can be reliably transmitted over the considered channel.
If we denote the achievable data rate of user k by R_k, our optimization problem reads as

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} R_k = \sum_{k=1}^{K} \log_2(1 + \mathrm{SINR}_k) \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{dl}} \tag{3.17}$$

where SINR_k denotes the signal-to-interference-plus-noise ratio (SINR) of user k and we defined the precoding matrix $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K]$ as the collection of all users' precoders.

Depending on the assumptions about the channel knowledge at the transmitter, different expressions for the achievable rate per user can be derived and then employed as a utility function for precoder design.

3.3.1.1 Spectral Efficiency for Perfect Tx-CSI

If the random channel realization $\boldsymbol{h}_k$ is perfectly known at the transmitter, then the received signal in (3.16) can be decomposed as follows

$$r_k = \underbrace{\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k}_{h_{k,\mathrm{eff}}} s_k + \underbrace{\sum_{j \neq k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j s_j}_{z_k} + v_k. \tag{3.18}$$

Here, r_k corresponds to the output of a discrete memoryless interference channel with input s_k, random channel response $h_{k,\mathrm{eff}}$, interference term z_k and noise v_k.

According to (A.2), the instantaneous achievable rate of a generic user k is given by

$$R_k^{\mathrm{inst}} = \log_2 \left(\frac{|h_{k,\mathrm{eff}}|^2}{\mathbb{E}[|z_k|^2 | u] + \sigma_k^2} \right) = \log_2 \left(1 + \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k|^2}{\sum_{j \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1} \right) \tag{3.19}$$

and the average achievable rate reads as

$$R_k^{\text{CSI}} = \mathbb{E}[R_k^{\text{inst}}] = \mathbb{E}\left[\log_2\left(1 + \frac{|\boldsymbol{h}_k^{\text{H}}\boldsymbol{p}_k|^2}{\sum_{j\neq k}|\boldsymbol{h}_k^{\text{H}}\boldsymbol{p}_j|^2 + 1}\right)\right] \tag{3.20}$$

where the expectation is w.r.t. the channel realizations.

If perfect CSI is available, the base station can adapt its precoding strategy depending on the current channel realization such that the instantaneous achievable rate (3.19) is maximized.

3.3.1.2 Spectral Efficiency for Incomplete Tx-CSI

As previously discussed in Section 3.2.1, we assume in this work that the base station has only partial knowledge of the channel, which is represented by the estimates $\{\hat{\boldsymbol{h}}_k\}$. If these estimates are used naively as if they were the true channel, and (3.19) is relied upon to design the precoders, this would result in an overestimation of the achievable rates and, eventually, poor performance.

Therefore, we use the error model introduced in (3.5), which takes into account the channel estimation errors [52]. This enables us to rewrite the received signal at user k as follows

$$r_k = \underbrace{\hat{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_k}_{h_{k,\text{eff}}} s_k + \underbrace{\tilde{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_k s_k + \sum_{j\neq k}\boldsymbol{h}_k^{\text{H}}\boldsymbol{p}_j s_j}_{z_k} + v_k. \tag{3.21}$$

Note that the so-obtained decomposition is similar to (3.18), expect that the desired signal is now not only received over the known channel response characterized by the effective channel $h_{k,\text{eff}} = \hat{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_k$, but also over $\tilde{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_k$ which depends on the unknown estimation error $\tilde{\boldsymbol{h}}_k$. We denote the set of known channel estimates by $u = \{\hat{\boldsymbol{h}}_i\}$.

In order to provide an expression for the achievable rate based on (A.2), we need to compute the conditional variance of the interference signal z_k

$$\begin{aligned}
\mathbb{E}\left[|z_k|^2|u\right] &= \mathbb{E}\left[|v_k|^2|\{\hat{\boldsymbol{h}}_i\}\right] \\
&= \mathbb{E}\left[|s_k|^2\right]\mathbb{E}\left[|\tilde{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_k|^2|\{\hat{\boldsymbol{h}}_i\}\right] + \sum_{j\neq k}\mathbb{E}\left[|s_j|^2\right]\mathbb{E}\left[|\boldsymbol{h}_k^{\text{H}}\boldsymbol{p}_j|^2|\{\hat{\boldsymbol{h}}_i\}\right] \\
&= \boldsymbol{p}_k^{\text{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_k + \sum_{j\neq k}\left(\boldsymbol{p}_j^{\text{H}}\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_j + \boldsymbol{p}_j^{\text{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_j\right) \\
&= \sum_{j=1}^{K}\boldsymbol{p}_j^{\text{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_j + \sum_{j\neq k}|\hat{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{p}_j|^2
\end{aligned} \tag{3.22}$$

where $\boldsymbol{C}_{\text{err},k} = \mathbb{E}[\tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^{\mathrm{H}}]$ denotes the error covariance matrix of user k. We additionally used the fact that the input symbols are zero-mean, unit-variance, and mutually independent.
Note that in the third line of (3.22), we assumed that the estimation error is independent of the channel estimate. This assumption holds true if LMMSE channel estimation is performed, as discussed earlier in this chapter.

The instantaneous achievable spectral efficiency can now be written as

$$R_k^{\text{inst}} = \log_2 \left(1 + \frac{|\hat{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{p}_k|^2}{\sum_{j=1}^{K} \boldsymbol{p}_j^{\mathrm{H}} \boldsymbol{C}_{\text{err},k} \boldsymbol{p}_j + \sum_{j \neq k} |\hat{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1} \right) \tag{3.23}$$

and yet again, the average achievable rate is given by $\mathbb{E}[R_k^{\text{inst}}]$, where the expectation is taken w.r.t. to the channel realizations.

Next, we will provide an alternative expression for the achievable spectral efficiency that holds for every channel estimation scheme. The expression is extensively used in the massive MIMO literature and is mostly referred to as the hardening bound [54]. To this end, it is assumed that only the mean value of the effective channel $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k$ is available at the transmitter, i.e., the deterministic average precoded channel $\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]$ is known to the base station. This assumption is based on the so-called channel hardening effect, which refers to the phenomenon where the random fluctuations of the channel gain $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k$ decrease as the number of base station antennas M increases, hence resulting in an approximately deterministic effective channel [4].
In this case, we can decompose the received signal r_k from (3.16) as follows

$$r_k = \underbrace{\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]}_{h_{k,\text{eff}}} s_k + \underbrace{\left(\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k - \mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k] \right) s_k + \sum_{j \neq k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j s_j}_{z_k} + v_k. \tag{3.24}$$

The model (3.24) matches the output of a discrete memoryless interference channel presented in Appendix A.1 with the deterministic channel response $h_{k,\text{eff}} = \mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]$. Furthermore, the interference z_k is zero-mean and independent of the input s_k due to

$$\begin{aligned} \mathbb{E}[z_k s_k^*] &= \mathbb{E}\Big[\Big((\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k - \mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]) s_k + \sum_{j \neq k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j s_j \Big) s_k^* \Big] \\ &= \mathbb{E}\Big[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k - \mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k] \Big] \mathbb{E}[|s_k|^2] \\ &= 0 \end{aligned} \tag{3.25}$$

where we used the mutual independence of the input signals $s_j, j = 1, \ldots, K$.

The variance of the interference term is given by

$$\mathbb{E}[|z_k|^2] = \sum_{j=1}^{K} \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2] - |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]|^2 \tag{3.26}$$

and according to (A.1), the "effective" achievable spectral efficiency can be written as

$$R_k^{\text{eff}} = \log_2 \left(1 + \frac{|\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]|^2}{\sum_{j=1}^{K} \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2] - |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k]|^2 + 1} \right). \tag{3.27}$$

The expectations in (3.27) can be computed in closed form for certain precoding schemes such that the standard matched filter and the statistical precoder that will be discussed later in Chapter 6.

3.3.2 RSMA Scheme

RSMA consists of splitting each user's message into common and private parts. For simplicity's sake, we here focus on one-layer RS where the user's message W_k is divided into one common part $W_{\mathrm{c},k}$ and a private part $\boldsymbol{W}_{\mathrm{p},k}$. The common parts of all users' messages are combined to form a super common message W_{c}. The resulting $K+1$ messages, i.e., the super common message W_{c} and the K private messages $W_{\mathrm{p},1}, \ldots, W_{\mathrm{p},K}$, are encoded independently into the streams s_{c}, $s_{\mathrm{p},1}, \ldots, s_{\mathrm{p},K}$ (see Fig. 3.3(a)). The transmission of the private streams takes place via the private precoders $\boldsymbol{p}_{\mathrm{p},k}$ as is the case of SDMA. The common stream is sent to the users using the common precoder $\boldsymbol{p}_{\mathrm{c}}$.

Contrary to the SDMA scheme, where each user decodes its own message and treats the interference from other users as noise, in the RSMA scheme, each user first decodes the common message and retrieves an estimate thereof, i.e., $\hat{W}_{\mathrm{c},k}$ (see Fig. 3.3(b)). This means that a part of the interference is decoded at the receiver. Subsequently, each user performs SIC in order to subtract the common part from the received signal and then decodes its own private message $\hat{W}_{\mathrm{p},k}$. The combination of $\hat{W}_{\mathrm{c},k}$ and $\hat{W}_{\mathrm{p},k}$ yields than an estimate of the kth user message $\hat{W}_k$.

Given the common precoder $\boldsymbol{p}_{\mathrm{c}}$ and the private precoders $\boldsymbol{p}_{\mathrm{p},i}, i = 1 \ldots K$, the transmit vector $\boldsymbol{x}$ can now be written as the superposition of the precoded common and private data streams

$$\boldsymbol{x} = \boldsymbol{p}_{\mathrm{c}} s_{\mathrm{c}} + \sum_{i=1}^{K} \boldsymbol{p}_{\mathrm{p},i} s_{\mathrm{p},i} \tag{3.28}$$

where $s_{\mathrm{c}} \sim \mathcal{N}_{\mathbb{C}}(0,1)$ and $s_{\mathrm{p},i} \sim \mathcal{N}_{\mathbb{C}}(0,1), i = 1, \ldots, K$.
Due to the limited power budget available at the base station, the transmit vector $\boldsymbol{x}$

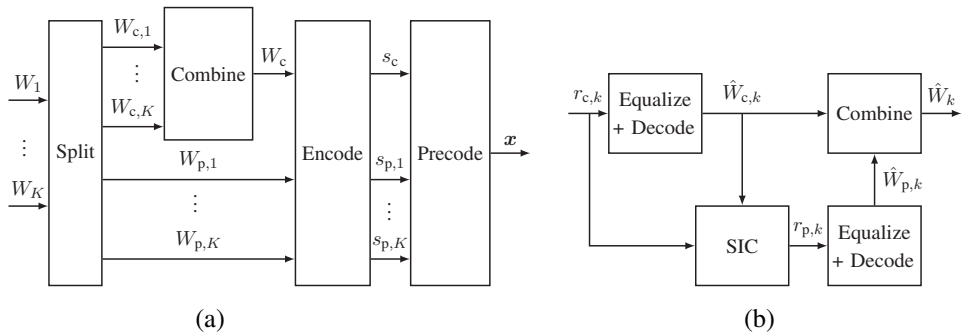

Fig. 3.3: System model during the data transmission phase for the one-layer RS setup (a): at the transmitter, (b): at the kth user (©IEEE 2023)

must satisfy the average transmit power constraint

$$\mathbb{E}[\|\boldsymbol{x}\|^2] = \|\boldsymbol{p}_{\mathrm{c}}\|^2 + \sum_{i=1}^{K} \|\boldsymbol{p}_{\mathrm{p},i}\|^2 \leq P_{\mathrm{dl}}. \tag{3.29}$$

Recall that the signal received at user k is given by (see (3.15))

$$r_k = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{x} + v_k. \tag{3.30}$$

Inserting the expression of the transmit vector (3.28), yields an expression for the received signal for the common message at user k

$$r_{\mathrm{c},k} = r_k = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}} s_{\mathrm{c}} + \sum_{i=1}^{K} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i} s_{\mathrm{p},i} + v_k. \tag{3.31}$$

We assume that user k knows perfectly the effective scalar channel $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}$, which can be achieved by sending dedicated beamformed pilot symbols in the downlink. This means the base station sends precoded pilots to the users, who can then estimate the combination of the channel with the precoders given by the scalars $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}$ and $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i}$. The knowledge of $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}$ at the user's side is necessary to be able to perform perfect SIC, i.e., in order to decode the common message and subtract it entirely from the received signal.

After performing perfect SIC, we can write the received signal for the private message $r_{\mathrm{p},k}$ as follows

$$r_{\mathrm{p},k} = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},k} s_{\mathrm{p},k} + \sum_{i \neq k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i} s_{\mathrm{p},i} + v_k. \tag{3.32}$$

In such an RS setup, we deal with two kinds of messages: common and private messages. We, therefore, differentiate between the common rate $R_{\mathrm{c},k}$ and the private rate $R_{\mathrm{p},k}$ for each user k. Since all users have to be able to decode the common stream, the common rate R_{c} should not be larger than the minimum common rate of all users, i.e.,

$$R_{\mathrm{c}} = \min_{k} R_{\mathrm{c},k}. \tag{3.33}$$

Our goal is to design the precoding vectors $\boldsymbol{p}_{\mathrm{c}}$ and $\boldsymbol{p}_{\mathrm{p},k}, k = 1, \ldots, K$ such that the achievable sum rate is maximized. This can be formulated as the following optimization problem.

$$\begin{aligned} \max_{\substack{R_{\mathrm{c}}, \boldsymbol{p}_{\mathrm{c}}, \\ \boldsymbol{p}_{\mathrm{p},1}, \ldots, \boldsymbol{p}_{\mathrm{p},K}}} \quad & R_{\mathrm{c}} + \sum_{i=1}^{K} R_{\mathrm{p},i} \quad \text{s.t.} \quad R_{\mathrm{c}} \leq R_{\mathrm{c},k}, \forall k \\ & \|\boldsymbol{p}_{\mathrm{c}}\|^2 + \sum_{i=1}^{K} \|\boldsymbol{p}_{\mathrm{p},i}\|^2 \leq P_{\mathrm{dl}}. \end{aligned} \tag{3.34}$$

Now, depending on the assumptions of the channel knowledge and the precoder design choice at the transmitter, different expressions for the achievable rates will be used, and the derivations follow the same lines as the SDMA case. The corresponding discussion will be considered in the following chapters whenever RS is employed.

4 Methodology

In this chapter, we will explore three powerful optimization techniques that are particularly effective for addressing precoder design challenges in wireless communications: the Lagrangian dual transform combined with FP [27], [42], the IWMMSE algorithm [23], and the MM algorithm [21]. These approaches share a common foundation in the principle of iterative optimization and the introduction of auxiliary variables, which expands the optimization space in a structured manner. The key idea behind these techniques is to reformulate complex optimization problems into more tractable forms by introducing carefully chosen auxiliary variables. The expanded problem can be decomposed into subproblems, each involving a subset of variables. This structure enables the use of block coordinate ascent/descent methods [55]. Here, the update of each variable or block of variables in each iteration can be computed efficiently, often in closed form. Under mild conditions, these algorithms are guaranteed to converge to a stationary point of the original problem, typically a local optimum.

The mentioned techniques provide systematic frameworks for developing efficient algorithms to solve complex precoder design problems. While the introduced auxiliary variables can be computed in closed form in nearly all cases, the precoder optimization step usually consists of a quadratically constrained quadratic program (QCQP) commonly solved via interior-point methods. To further reduce the computational complexity, we present by the end of this chapter the so-called ”rescaling” approach, firstly introduced in [43] and subsequently applied in [25] and [26] to the IWMMSE approach. The technique enables a relaxation of the underlying optimization problem to an unconstrained one, yielding a closed-form solution for the precoder, which has solely to be rescaled after final convergence to meet the underlying constraint.

4.1 Lagrangian Dual Transform and Fractional Programming

The Lagrangian dual transform, when combined with FP techniques such as the quadratic transform, is particularly effective for problems involving logarithmic and ratio-based utility functions. This approach transforms the original problem into a sequence of more manageable subproblems, resulting in convex optimization steps in

each iteration.

The procedure begins with considering a generic optimization problem involving maximizing a sum of logarithmic functions. By introducing auxiliary variables and applying Lagrangian duality, we can reformulate the problem into a more tractable form. This transformation allows us to separate the logarithmic terms from the original optimization variables, making the problem easier to solve.
After applying the Lagrangian dual transform, we end up with an objective function containing a sum of ratios. To address this, we can use FP techniques, specifically via the quadratic transform that will be reviewed in Section 4.1.2.

4.1.1 Lagrangian Dual Transform

Consider the following generic optimization problem

$$\max_{\boldsymbol{x}} \sum_{k=1}^{K} \log\left(1+\frac{N_k(\boldsymbol{x})}{D_k(\boldsymbol{x})}\right), \quad \text{s.t.} \quad \boldsymbol{x}\in\mathcal{X} \tag{4.1}$$

where $\mathcal{X}$ denotes an abstract constraint set.

We start by introducing a set of K auxiliary variables

$$\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_K]^{\mathrm{T}} \tag{4.2}$$

and rewrite (4.1) as follows

$$\begin{aligned} \max_{\boldsymbol{x}} \max_{\boldsymbol{\gamma}} \sum_{k=1}^{K} \log\left(1+\gamma_k\right) \quad \text{s.t.} \quad & \gamma_k \le \frac{N_k(\boldsymbol{x})}{D_k(\boldsymbol{x})},\ k=1,\ldots,K \\ & \boldsymbol{x}\in\mathcal{X}. \end{aligned} \tag{4.3}$$

Let us introduce a collection of the Lagrangian multipliers associated with the K inequality constraints in (4.3)

$$\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_K]^{\mathrm{T}}. \tag{4.4}$$

As explained in [42], the inner optimization problem in (4.3) is convex; hence, strong duality holds. The solution to its dual problem denoted by $(\boldsymbol{\lambda}^{\text{opt}}, \boldsymbol{\gamma}^{\text{opt}})$ is a saddle point that satisfies the first-order optimality condition of the Lagrangian function

$$\mathcal{L}_1(\boldsymbol{\gamma}, \boldsymbol{\lambda}) = \sum_{k=1}^{K} \log(1+\gamma_k) + \sum_{k=1}^{K} \lambda_k \left(\frac{N_k(\boldsymbol{x})}{D_k(\boldsymbol{x})} - \gamma_k\right). \tag{4.5}$$

Hence, the optimal λ_k^{opt} and γ_k^{opt} are respectively given by

$$\lambda_k^{\text{opt}} = \frac{1}{1+\gamma_k^{\text{opt}}}, \tag{4.6}$$

$$\gamma_k^{\text{opt}} = \frac{N_k(\boldsymbol{x})}{D_k(\boldsymbol{x})}. \tag{4.7}$$

Inserting (4.6) into the Lagrangian function (4.5) and combining it with the optimization problem w.r.t. $\boldsymbol{x}$, we can formulate the objective function denoted by $f_1(\boldsymbol{x}, \boldsymbol{\gamma})$ as follows

$$f_1(\boldsymbol{x}, \boldsymbol{\gamma}) = \sum_{k=1}^{K} \log(1+\gamma_k) - \sum_{k=1}^{K} \gamma_k + \sum_{k=1}^{K} \frac{(1+\gamma_k)N_k(\boldsymbol{x})}{N_k(\boldsymbol{x}) + D_k(\boldsymbol{x})}. \tag{4.8}$$

Our new optimization problem now reads as

$$\max_{\boldsymbol{x},\boldsymbol{\gamma}} f_1(\boldsymbol{x}, \boldsymbol{\gamma}), \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{X}. \tag{4.9}$$

Note that for fixed $\boldsymbol{x}$, the auxiliary variable γ_k is simply given by (4.7).

The new objective function in (4.8) combines logarithmic terms, linear terms, and ratios of the original functions. The transformed problem (4.9) is now in a form that's more amenable to optimization techniques, particularly FP.
For instance, if we now fix $\boldsymbol{\gamma}$, the only term of the objective in (4.8) that depends on $\boldsymbol{x}$ is the third one. This term consists of a sum-of-ratios to which the *quadratic transform* [27] can be applied by introducing a set of auxiliary variables for each of the K ratios. This transform enables decoupling the numerator and denominator of each fraction while ensuring that the original and transformed objectives are the same in the optimum.

4.1.2 Fractional Programming via Quadratic Transform

Let us denote by β_k the auxiliary variable associated with the kth fraction in (4.8). We then formulate the following optimization problem

$$\max_{\boldsymbol{x},\boldsymbol{\gamma},\boldsymbol{\beta}} f_2(\boldsymbol{x}, \boldsymbol{\gamma}, \boldsymbol{\beta}) \quad \text{s.t.} \quad \boldsymbol{x} \in \mathcal{X} \tag{4.10}$$

where $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_K]^{\mathrm{T}}$ and the objective function $f_2(\boldsymbol{x}, \boldsymbol{\gamma}, \boldsymbol{\beta})$ is given by

$$\begin{aligned} f_2(\boldsymbol{x}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = &\sum_{k=1}^{K} \log(1+\gamma_k) - \sum_{k=1}^{K} \gamma_k \\ &+ \sum_{k=1}^{K} (1+\gamma_k)\Big(2\Re\{\beta_k S_k(\boldsymbol{x})\} - |\beta_k|^2\big(N_k(\boldsymbol{x}) + D_k(\boldsymbol{x})\big)\Big) \end{aligned} \tag{4.11}$$

where we assumed that the numerator is given by

$$N_k(\boldsymbol{x}) = |S_k(\boldsymbol{x})|^2. \tag{4.12}$$

Again, for fixed $\boldsymbol{x}$ and $\boldsymbol{\gamma}$, each of the auxiliary variables β_k can be computed in closed form

$$\beta_k^{\text{opt}} = \frac{S_k(\boldsymbol{x})^*}{N_k(\boldsymbol{x}) + D_k(\boldsymbol{x})}. \tag{4.13}$$

Finally, it remains to solve the following optimization problem w.r.t. $\boldsymbol{x}$ for fixed $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$

$$\begin{aligned} &\max_{\boldsymbol{x}} \sum_{k=1}^{K} (1+\gamma_k)\Big(2\Re\{\beta_k S_k(\boldsymbol{x})\} - |\beta_k|^2\big(N_k(\boldsymbol{x}) + D_k(\boldsymbol{x})\big)\Big) \\ &\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X}. \end{aligned} \tag{4.14}$$

Assuming that $N_k(\boldsymbol{x})$ and $D_k(\boldsymbol{x})$ are quadratic functions in $\boldsymbol{x}$ and that the constraint set $\mathcal{X}$ is a convex set, (4.14) can be solved using off-the-shelf convex optimization approaches.

4.2 Iterative Weighted Minimum Mean Square Error Framework

The IWMMSE algorithm is tailored for precoder optimization, as it transforms the typically non-convex sum rate maximization (3.17) into an equivalent block-wise convex weighted sum MSE minimization. The key idea is introducing sets of auxiliary variables, namely the weighting factors and the receive filters. As their names already reveal, these variables are interpretable in terms of the communication system parameters as opposed to the previously discussed Lagrangian dual transform and FP method, where the auxiliary variables are introduced based on a purely mathematical consideration of the problem. The IWMMSE method can even be seen as a particular case of the Lagrangian dual transform and FP approach, as discussed later in this section.

Consider the multi-user MISO setup discussed in Chapter 3. For the IWMMSE approach, we introduce the first set of auxiliary variables consisting of the receive filters $g_k, k = 1, \ldots, K$. At each user k, the filter g_k is applied to the received signal r_k (see (3.16)) to obtain an estimate $\hat{s}_k$ of the data signal s_k

$$\hat{s}_k = g_k r_k = g_k \boldsymbol{h}_k^{\mathrm{H}} \sum_{j=1}^{K} \boldsymbol{p}_j s_j + g_k v_k. \tag{4.15}$$

Hence, we can define the MSE of the data signal for user k as

$$\varepsilon_k = \mathbb{E}[|s_k - \hat{s}_k|^2] \tag{4.16}$$

$$= 1 - 2\Re\{g_k \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k\} + |g_k|^2 (\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1). \tag{4.17}$$

The optimal receive filter that minimizes the MSE ε_k is given by

$$g_k^{\text{MMSE}} = \frac{\boldsymbol{p}_k^{\mathrm{H}} \boldsymbol{h}_k}{\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1}. \tag{4.18}$$

Inserting (4.18) into the MSE expression in (4.16) yields the following MMSE expression

$$\varepsilon_k^{\text{MMSE}} = 1 - \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k|^2}{\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1} = \frac{1}{1 + \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k|^2}{\sum_{j \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1}}. \tag{4.19}$$

Note that in (4.19), a relationship between the MMSE of user k and the corresponding SINR is established, namely

$$\varepsilon_k^{\text{MMSE}} = \frac{1}{1 + \text{SINR}_k} \tag{4.20}$$

and hence, the rate of user k can be expressed as

$$R_k = -\log_2(\varepsilon_k^{\text{MMSE}}). \tag{4.21}$$

The sum rate maximization problem (3.17) can be formulated as a minimization of the sum MSE as follows

$$\min_{\boldsymbol{P}} \sum_{k=1}^{K} \log_2(\varepsilon_k^{\text{MMSE}}) \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}}. \tag{4.22}$$

Before we proceed to solve (4.22), we state the following lemma from [56]

Lemma 1. *For any $E > 0$ and $u > 0$, it holds that*

$$\log_2 E = \min_u uE - \log_2(u) + \delta \tag{4.23}$$

with some constant δ.

The idea behind Lemma 1 is to introduce an auxiliary variable u, which allows a reformulation of the objective function as an optimization problem, where the function E is externalized from the logarithm. Moreover, u can be found in closed form.

We now apply Lemma 1 to ease the solution of (4.22). To this end, we introduce the so-called augmented weighted MSE (WMSE) ξ_k with the corresponding weight u_k defined as

$$\xi_k = u_k \varepsilon_k - \log_2(u_k). \tag{4.24}$$

Note that for fixed precoders and using the MMSE filters, the optimal weight u_k that minimizes ξ_k is obtained as

$$u_k^{\text{MMSE}} = \frac{1}{\varepsilon_k^{\text{MMSE}}}. \tag{4.25}$$

Inserting this result into (4.24) yields the augmented weighted MMSE (WMMSE) ξ_k^{MMSE} which is closely related to the rate of user k according to

$$\xi_k^{\text{MMSE}} = 1 - R_k. \tag{4.26}$$

The augmented WMSE minimization problem can be thus formulated as

$$\min_{\boldsymbol{G},\boldsymbol{U},\boldsymbol{P}} \sum_{k=1}^{K} \xi_k \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\text{F}}^2 \leq P_{\text{dl}} \tag{4.27}$$

where we defined the diagonal matrices containing the receive filters and the weights as follows

$$\boldsymbol{G} = \text{diag}(g_1, \ldots, g_K), \tag{4.28}$$

$$\boldsymbol{U} = \text{diag}(u_1, \ldots, u_K). \tag{4.29}$$

The variables g_k and u_k corresponding to user k can be found in closed form according to (4.18) and (4.25), respectively, and are decoupled from other users' variables.

By fixing $\boldsymbol{G}$ and $\boldsymbol{U}$ and ignoring all constant terms w.r.t. to the precoders, we obtain the following optimization problem

$$\begin{aligned} \min_{\boldsymbol{P}} \; & \sum_{k=1}^{K} u_k \left(-2\Re\{g_k \boldsymbol{h}_k^{\text{H}} \boldsymbol{p}_k\} + |g_k|^2 \Big(\sum_{j=1}^{K} |\boldsymbol{h}_k^{\text{H}} \boldsymbol{p}_j|^2 + 1 \Big) \right) \\ \text{s.t.} \; & \|\boldsymbol{P}\|_{\text{F}}^2 \leq P_{\text{dl}}. \end{aligned} \tag{4.30}$$

(4.30) is a QCQP and can thus be solved using standard convex solvers.

4.2.1 Relationship between IWMMSE and FP for Precoding

By setting the optimization variable $\boldsymbol{x}$ in (4.1) to be the precoding matrix $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K]$, and defining

$$N(\boldsymbol{P}) = |S(\boldsymbol{P})|^2 = |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k|^2, \quad D(\boldsymbol{P}) = \sum_{j \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1 \tag{4.31}$$

we can apply the Lagrangian dual transform and FP to solve the sum rate maximization problem in (3.17). The auxiliary variable γ_k is obtained following (4.7)

$$\gamma_k^{\text{opt}} = \frac{N(\boldsymbol{P})}{D(\boldsymbol{P})} = \text{SINR}_k. \tag{4.32}$$

We here notice the first similarity to the IWMMSE approach, namely that

$$u_k^{\text{MMSE}} = 1 + \gamma_k^{\text{opt}}. \tag{4.33}$$

Second, the optimal auxiliary variable β_k^{opt} is given based on (4.13) by

$$\beta_k^{\text{opt}} = \frac{S(\boldsymbol{P})^*}{N(\boldsymbol{P}) + D(\boldsymbol{P})} = \frac{\boldsymbol{p}_k^{\mathrm{H}} \boldsymbol{h}_k}{\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1} \tag{4.34}$$

which coincides with the MMSE receive filter in (4.18).

Finally, based on (4.10), the optimization problem w.r.t. the precoding matrix can be written as

$$\begin{aligned} &\max_{\boldsymbol{P}} \sum_{k=1}^{K} (1 + \gamma_k) \left(2\Re\{\beta_k \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k\} - |\beta_k|^2 \Big(\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1 \Big) \right) \\ &\text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}}. \end{aligned} \tag{4.35}$$

Obviously, (4.35) is equivalent to (4.30); therefore, both approaches boil down to the same optimization steps. However, we must emphasize that the FP framework is more versatile, and alternative ways of applying the quadratic transform can lead to different optimization problems. Therefore, the IWMMSE approach can be viewed as a particular case of the FP framework with interpretable auxiliary variables. The interested reader might refer to [21] for a deeper discussion about the connection between the two frameworks.

4.3 Minorization Maximization

Contrary to the previously presented frameworks that operate directly on the underlying sum rate maximization problem or the equivalently transformed one, the MM approach constructs a sequence of carefully chosen surrogate functions that minorize (lower bound) the original objective. This surrogate function is maximized in each iteration, which is typically easier to optimize than the original function. The MM approach offers flexibility in constructing surrogate functions, allowing tailored algorithms to exploit problem-specific structures.

Consider, for instance, the following generic maximization problem

$$\max_{x} f(x) \quad \text{s.t.} \quad x \in \mathcal{X}. \tag{4.36}$$

A surrogate function $\ell(x, \bar{x})$ for f at iterate $\bar{x}$ has to satisfy the following two conditions [21]

$$\ell(x, \bar{x}) \leq f(x), \forall x \in \mathcal{X}, \tag{4.37}$$

$$\ell(\bar{x}, \bar{x}) = f(\bar{x}), \forall \bar{x} \in \mathcal{X}. \tag{4.38}$$

While the first condition in (4.37) translates to the fact the surrogate function serves as a lower bound for the underlying objective, the second condition in (4.38) ensures that each iterate, the surrogate function coincides with the original function.

Since we are mainly considering the sum rate maximization problem, i.e., solving (3.17), we start by stating the following proposition that will help us provide a surrogate function for the rate expression

Proposition 1. *[28, Section III.B]*
For some $x \in \mathbb{C}$ and $z > 0$, it holds that

$$\log_2(1+\frac{|x|^2}{z}) \geq \log_2(1+\frac{|\bar{x}|^2}{\bar{z}}) - \frac{|\bar{x}|^2}{\bar{z}} + 2\Re\{\frac{\bar{x}^*}{\bar{z}}x\} - \frac{|\bar{x}|^2}{\bar{z}(\bar{z}+|\bar{x}|^2)}(z+|x|^2) \tag{4.39}$$

with equality for $(x, z) = (\bar{x}, \bar{z})$.

By substituting x and z by $\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k$ and $\sum_{j\neq k} |\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2+1$, respectively, the surrogate function $\ell_k(\boldsymbol{P}, \bar{\boldsymbol{P}})$ for the kth user's rate at iterate $\bar{\boldsymbol{P}}$ is given by

$$\ell_k(\boldsymbol{P}, \bar{\boldsymbol{P}}) = R_k(\bar{\boldsymbol{P}}) - \mathrm{SINR}_k(\bar{\boldsymbol{P}}) + 2\Re\{b_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k\} - a_k\Big(\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2 + 1\Big). \tag{4.40}$$

We here defined the auxiliary variables a_k and b_k as follows

$$a_k = \frac{\text{SINR}_k(\bar{\boldsymbol{P}})}{\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \bar{\boldsymbol{p}}_j|^2 + 1}, \tag{4.41}$$

$$b_k = \frac{\bar{\boldsymbol{p}}_k^{\mathrm{H}} \boldsymbol{h}_k}{\sum_{j \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \bar{\boldsymbol{p}}_j|^2 + 1}. \tag{4.42}$$

The sum rate maximization problem (3.17) can now be replaced by the maximization of the sum of the K surrogate functions corresponding to the users' rates, i.e.,

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} \ell_k(\boldsymbol{P}, \bar{\boldsymbol{P}}) \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}}. \tag{4.43}$$

If we ignore all the terms in (4.40) that are independent of $\boldsymbol{P}$, we end up with the following optimization problem

$$\begin{aligned} &\max_{\boldsymbol{P}} \sum_{k=1}^{K} 2\Re\{b_k \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k\} - a_k \Big(\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1 \Big) \\ &\text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}} \end{aligned} \tag{4.44}$$

which is again a QCQP w.r.t. the precoders for fixed a_k and b_k and can be solved using standard off-the-shelf solvers.

The optimization procedure starts with a feasible initial iterate $\bar{\boldsymbol{P}}$, then the auxiliary variables a_k and b_k are updated according to (4.41) and (4.42), respectively. Subsequently, the precoding matrix $\boldsymbol{P}$ is obtained by solving the optimization problem in (4.44).

4.3.1 Connection of the MM Approach to FP and IWMMSE

Note that the precoder optimization problem (4.44) based on the MM framework resembles the one given by (4.14), formulated according to the FP approach. One can show that they are essentially equivalent. For instance, if we use the update rules for the auxiliary variables γ_k and β_k respectively given by (4.32) and (4.34), we have the following relationships to the auxiliary variables a_k and b_k

$$\begin{aligned} (1+\gamma_k)\beta_k &= (1 + \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k|^2}{\sum_{j \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1}) \frac{\boldsymbol{p}_k^{\mathrm{H}} \boldsymbol{h}_k}{\sum_{j=1}^{K} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1} \\ &= \frac{\boldsymbol{p}_k^{\mathrm{H}} \boldsymbol{h}_k}{\sum_{j \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_j|^2 + 1} \\ &= b_k \end{aligned} \tag{4.45}$$

$$(1+\gamma_k)|\beta_k|^2 = (1+\frac{|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k|^2}{\sum_{j\neq k}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2+1})\frac{|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k|^2}{(\sum_{j=1}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2+1)^2} \tag{4.46}$$
$$= a_k.$$

4.4 Rescaling Approach

We have shown that the presented methods for precoder design are equivalent, and the most computationally demanding step consists of solving the QCQP to obtain the precoding matrix.

To this end, recall that we need to solve the following optimization problem for the precoding matrix $\boldsymbol{P}$

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} 2\Re\{b_k\boldsymbol{t}_k^{\mathrm{H}}\boldsymbol{p}_k\} - a_k\Big(\sum_{j=1}^{K}\boldsymbol{p}_j^{\mathrm{H}}\boldsymbol{T}_k\boldsymbol{p}_j+1\Big) \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}} \tag{4.47}$$

where we introduced the generic vector $\boldsymbol{t}_k \in \mathbb{C}^M$ and the non-negative definite matrix $\boldsymbol{T}_k$.

One approach to solve (4.47) is the Lagrangian multiplier method [21], where the solution of the precoder $\boldsymbol{p}_k$ is inferred from the first optimality condition of the Lagrangian function and is given by

$$\boldsymbol{p}_k(\lambda) = \left(\sum_{j=1}^{K} a_j\boldsymbol{T}_j + \lambda\boldsymbol{I}_M\right)^{-1} b_k^*\boldsymbol{t}_k. \tag{4.48}$$

Here, λ denotes the Lagrangian multiplier associated with the power constraint in (4.47) and can be obtained according to

$$\min\left\{\lambda \geq 0 : \|\boldsymbol{P}(\lambda)\|_{\mathrm{F}}^2 \leq P_{\text{dl}}\right\}. \tag{4.49}$$

Solving (4.49) requires a line search, e.g., via bisection [57], and can be computationally expensive since matrix inversion has to be performed in every search step.

Inspired by the scaling idea in [43] and its application to the IWMMSE framework [25], we propose the following general approach to recast (4.47) as an unconstrained optimization problem, where the solution is given in closed-form.

To this end, we first introduce a common scaling of the auxiliary variables a_k and b_k and define the scaled variables a_k' and b_k' in the following manner

$$a_k' = \alpha^{-2}a_k, \quad b_k' = \alpha^{-1}b_k \tag{4.50}$$

with some scaling factor $\alpha > 0$.
The optimization problem (4.47) can now be rewritten as

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} 2\alpha^{-1}\Re\{b_k \boldsymbol{t}_k^{\mathrm{H}} \boldsymbol{p}_k\} - \alpha^{-2} a_k \Big(\sum_{j=1}^{K} \boldsymbol{p}_j^{\mathrm{H}} \boldsymbol{T}_k \boldsymbol{p}_j + 1 \Big) \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{dl}}. \tag{4.51}$$

The solution for the precoder $\boldsymbol{p}_k$ can be expressed as a function of the scaling factor α by considering the stationarity condition

$$\boldsymbol{p}_k(\alpha) = \left(\sum_{j=1}^{K} a_j \boldsymbol{T}_j + \delta \boldsymbol{I}_M \right)^{-1} \alpha b_k^* \boldsymbol{t}_k \tag{4.52}$$

with $\delta = \lambda \alpha^2$, where λ denotes again the Lagrangian multiplier associated with the power constraint.

Now, the scaling factor can be inferred from the equality of the power constraint [25]

$$\alpha(\delta) = \sqrt{\frac{P_{\mathrm{dl}}}{\sum_{j=1}^{K} |b_j|^2 \boldsymbol{t}_j^{\mathrm{H}} \boldsymbol{X}(\delta)^{-2} \boldsymbol{t}_j}} \tag{4.53}$$

with $\boldsymbol{X}(\delta) = \sum_{j=1}^{K} a_j \boldsymbol{T}_j + \delta \boldsymbol{I}_M$.
This choice of α ensures that the obtained precoder solution in (4.52) fulfills the power constraint, and the optimization problem (4.51) can thus be recast as the following unconstrained optimization problem w.r.t. δ by inserting (4.52) and (4.53) into (4.51)

$$\begin{aligned} \max_{\delta} & \sum_{k=1}^{K} 2\Re\{|b_k|^2 \boldsymbol{t}_k^{\mathrm{H}} \boldsymbol{X}(\delta)^{-1} \boldsymbol{t}_k\} \\ & - \sum_{k=1}^{K} a_k \Big(\sum_{j=1}^{K} |b_j|^2 \boldsymbol{t}_j^{\mathrm{H}} \boldsymbol{X}(\delta)^{-1} \boldsymbol{T}_k \boldsymbol{X}(\delta)^{-1} \boldsymbol{t}_j + \frac{1}{P_{\mathrm{dl}}} \sum_{j=1}^{K} |b_j|^2 \boldsymbol{t}_j^{\mathrm{H}} \boldsymbol{X}(\delta)^{-2} \boldsymbol{t}_j \Big). \end{aligned} \tag{4.54}$$

To solve (4.54), we consider the stationarity condition

$$\begin{aligned} & -2 \sum_{k=1}^{K} |b_k|^2 \boldsymbol{t}_k^{\mathrm{H}} \boldsymbol{X}(\delta)^{-2} \boldsymbol{t}_k \\ & + 2 \sum_{j=1}^{K} |b_j|^2 \boldsymbol{t}_j^{\mathrm{H}} \left(\boldsymbol{X}(\delta)^{-2} - \delta \boldsymbol{X}(\delta)^{-3} \right) \boldsymbol{t}_j + 2 \sum_{j=1}^{K} |b_j|^2 \boldsymbol{t}_j^{\mathrm{H}} \boldsymbol{X}(\delta)^{-3} \boldsymbol{t}_j \left(\frac{1}{P_{\mathrm{dl}}} \sum_{k=1}^{K} a_k \right) \\ & \stackrel{!}{=} 0 \end{aligned} \tag{4.55}$$

where we used

$$\frac{\partial \boldsymbol{A}(\delta)^{-1}}{\partial \delta} = -\boldsymbol{A}(\delta)^{-1} \frac{\partial \boldsymbol{A}(\delta)}{\partial \delta} \boldsymbol{A}(\delta)^{-1}. \tag{4.56}$$

This yields the optimal δ

$$\delta^{\text{opt}} = \frac{1}{P_{\text{dl}}} \sum_{k=1}^{K} a_k. \tag{4.57}$$

The optimal unscaled precoder is, therefore, given by

$$\boldsymbol{p}_k = \left(\sum_{j=1}^{K} a_j \boldsymbol{T}_j + \Big(\frac{1}{P_{\text{dl}}} \sum_{i=1}^{K} a_i \Big) \boldsymbol{I}_M \right)^{-1} b_k^* \boldsymbol{t}_k. \tag{4.58}$$

In a nutshell, once the auxiliary variables a_k and b_k are computed, these are fixed, and the precoder is calculated in closed form according to (4.58). After the final convergence is reached, the precoder solution is scaled with α to satisfy the power constraint.

5 Zero-Forcing Precoding

Zero-forcing is a suboptimal yet simple linear precoding approach that was shown to be asymptotically optimal in the high-SNR regime [58], [59]. However, this behavior is only guaranteed in case of perfect CSI at the transmitter. In our setup, the base station has incomplete knowledge about the channel, and therefore, the zero-forcing asymptotic behavior for imperfect CSI has to be investigated.

In essence, zero-forcing precoding creates orthogonal beams for each user in order to eliminate inter-user interference. This is achieved by choosing the precoding vector of each user to be orthogonal to the channel vectors of the remaining users. The goal here is to maximize the system throughput by serving several users simultaneously in an interference-free manner. However, an accurate CSI is required at the transmitter in order to suppress the interference effectively. In this chapter, we show that even with incomplete CSI, the zero-forcing precoder based on the LS channel estimate achieves asymptotically full degrees of freedom under certain conditions on the used pilots [44].

5.1 Zero-Forcing for Perfect CSI

By defining the channel matrix as the collection of the users' channels

$$\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K] \tag{5.1}$$

the zero-forcing precoding matrix can be expressed as a function of the channel matrix as follows [12]

$$\boldsymbol{P}^{\text{zf}}(\boldsymbol{H}) = \alpha(\boldsymbol{H}^{\text{H}})^{\dagger} = \alpha\boldsymbol{H}(\boldsymbol{H}^{\text{H}}\boldsymbol{H})^{-1}. \tag{5.2}$$

In (5.2), we introduced the scaling factor

$$\alpha^2 = \frac{P_{\text{dl}}}{\text{tr}((\boldsymbol{H}^{\text{H}}\boldsymbol{H})^{-1})} \tag{5.3}$$

which normalizes the precoding matrix, such that the power constraint is fulfilled, i.e., $\text{tr}(\boldsymbol{P}^{\text{zf}}\boldsymbol{P}^{\text{zf,H}}) = P_{\text{dl}}$.
Note that the matrix $\boldsymbol{H}^{\text{H}}\boldsymbol{H}$ should be full-rank in order for the zero-forcing precoding matrix in (5.2) to exist. This is, in general, the case for $K \leq M$.

If we now collect the received signals of all users (see (3.16)) as follows

$$\boldsymbol{r} = \boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{s} + \boldsymbol{v} \tag{5.4}$$

with $\boldsymbol{s} = [s_1, \ldots, s_K]^{\mathrm{T}}$ and $\boldsymbol{v} = [v_1, \ldots, v_K]^{\mathrm{T}}$, and use the zero-forcing precoder, we obtain an interference-free transmission link, i.e.,

$$\boldsymbol{r} = \alpha \underbrace{\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}}_{\boldsymbol{I}_K}\boldsymbol{s} + \boldsymbol{v} = \alpha\boldsymbol{s} + \boldsymbol{v}. \tag{5.5}$$

This is achieved since we constructed the zero-forcing precoder based on the true channel matrix $\boldsymbol{H}$.
In the perfect CSI case, the sum rate achieved with zero-forcing precoding asymptotically approaches the sum rate obtained with the optimal precoding strategy, namely dirty paper coding for high transmit powers [60].

In the following, we will consider the imperfect CSI case from our setup and analyze the asymptotic behavior of the zero-forcing precoder computed based on the channel estimate (see Section 3.2.1) instead of the true channel as in (5.2).

5.2 Zero-Forcing for Imperfect CSI

5.2.1 Two-User Case

To simplify the asymptotic analysis, we first focus on the two-user case, i.e., $K = 2$. The results are generalizable for $K > 2$, and a proof for the case $K > 2$ with the LS channel estimate is presented in the next subsection.

We study the asymptotic behavior of the system in the high-power regime, i.e., for $P_{\text{dl}} \to \infty$. To understand the interference behavior, we focus on the combination of one user's channel $\boldsymbol{h}_k$ with another user's precoding vector $\boldsymbol{p}_j, j \neq k$ given by $\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j$. The received signal at user k for the two-user case reads as (cf. (3.16))

$$r_k = \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k s_k + \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j s_j + v_k. \tag{5.6}$$

We say that the interference is suppressed when $\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j = 0$ for $j \neq k$; otherwise, the system is interference-limited.

Before we proceed with our analysis, we write the precoding matrix $\boldsymbol{P}^{\text{zf}}$ as a

function of the estimated channel matrix $\hat{\boldsymbol{H}} = [\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2]$ as follows

$$\begin{aligned}\boldsymbol{P}^{\text{zf}}(\hat{\boldsymbol{H}}) &= [\boldsymbol{p}_1, \boldsymbol{p}_2] = \alpha \hat{\boldsymbol{H}}(\hat{\boldsymbol{H}}^{\text{H}}\hat{\boldsymbol{H}})^{-1} \\ &= \beta[\hat{\boldsymbol{h}}_1, \hat{\boldsymbol{h}}_2] \begin{bmatrix} \|\hat{\boldsymbol{h}}_2\|^2 & -\hat{\boldsymbol{h}}_1^{\text{H}}\hat{\boldsymbol{h}}_2 \\ -\hat{\boldsymbol{h}}_2^{\text{H}}\hat{\boldsymbol{h}}_1 & \|\hat{\boldsymbol{h}}_1\|^2 \end{bmatrix} \\ &= \beta \begin{bmatrix} \hat{\boldsymbol{h}}_1\|\hat{\boldsymbol{h}}_2\|^2 - \hat{\boldsymbol{h}}_2\hat{\boldsymbol{h}}_2^{\text{H}}\hat{\boldsymbol{h}}_1 & \hat{\boldsymbol{h}}_2\|\hat{\boldsymbol{h}}_1\|^2 - \hat{\boldsymbol{h}}_1\hat{\boldsymbol{h}}_1^{\text{H}}\hat{\boldsymbol{h}}_2 \end{bmatrix}\end{aligned} \tag{5.7}$$

where we defined

$$\beta = \alpha / \det(\hat{\boldsymbol{H}}^{\text{H}}\hat{\boldsymbol{H}}) = \alpha/(\|\hat{\boldsymbol{h}}_1\|^2\|\hat{\boldsymbol{h}}_2\|^2 - |\hat{\boldsymbol{h}}_1^{\text{H}}\hat{\boldsymbol{h}}_2|^2) \tag{5.8}$$

and $\alpha^2 = P_{\text{dl}}/\text{tr}((\hat{\boldsymbol{H}}^{\text{H}}\hat{\boldsymbol{H}})^{-1})$ is again a scaling factor used to normalize the precoding matrix such that the power constraint is satisfied (cf. (5.3)).

5.2.1.1 Asymptotic behavior of zero-forcing based on LS estimate

Recall that the LS channel estimate of the kth user is given by (see (3.7))

$$\boldsymbol{h}_k^{\text{LS}} = \boldsymbol{\Phi}\boldsymbol{y}_k = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k + \boldsymbol{\Phi}\boldsymbol{n}_{1,k}.$$

Hence, we can write the LS estimated channel matrix as

$$\hat{\boldsymbol{H}}^{\text{LS}} = \boldsymbol{\Phi}\boldsymbol{Y}. \tag{5.9}$$

Here, the matrix $\boldsymbol{Y}$ denotes the collection all users' channel observations (see (3.2)) and is given by

$$\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K] = \boldsymbol{\Phi}^{\text{H}}\boldsymbol{H} + \boldsymbol{N} \tag{5.10}$$

with $\boldsymbol{N} = [\boldsymbol{n}_{1,1}, \ldots, \boldsymbol{n}_{1,K}]$.

The zero-forcing precoding matrix based on the LS estimate can be written as

$$\boldsymbol{P}^{\text{zf}}(\hat{\boldsymbol{H}}^{\text{LS}}) = \alpha\hat{\boldsymbol{H}}^{\text{LS}}(\hat{\boldsymbol{H}}^{\text{LS,H}}\hat{\boldsymbol{H}}^{\text{LS}})^{-1} = \alpha\boldsymbol{\Phi}\boldsymbol{Y}(\boldsymbol{Y}^{\text{H}}\boldsymbol{Y})^{-1} \tag{5.11}$$

where we used the orthogonality of the pilot matrix, i.e., $\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi} = \boldsymbol{I}_{T_{\text{dl}}}$.
Note that $\boldsymbol{P}^{\text{zf}}(\hat{\boldsymbol{H}}^{\text{LS}})$ exists if $\boldsymbol{Y}^{\text{H}}\boldsymbol{Y}$ is invertible, which is the case if $T_{\text{dl}} \geq K$. Thus, the following analysis is only valid for the latter case since otherwise, the inverse in (5.11) is not defined.

First, recall that the downlink training noise is inversely proportional to the downlink transmit power (see Section 3.2). Hence, the "asymptotic" LS channel estimate for user k is given by

$$\hat{\tilde{\boldsymbol{h}}}_k = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k, \quad \text{for } P_{\text{dl}} \to \infty. \tag{5.12}$$

This can be seen by writing the noise vector $\boldsymbol{n}_{1,k}$ in the observation vector $\boldsymbol{y}_k$ as

$$\boldsymbol{n}_{1,k} = \sigma_{1,k}\tilde{\boldsymbol{n}}_k = \frac{1}{\sqrt{P_{\text{dl}}}}\tilde{\boldsymbol{n}}_k$$

with $\tilde{\boldsymbol{n}}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I}_{T_{\text{dl}}})$. For $P_{\text{dl}} \to \infty$, the noise term can, therefore, be neglected.

Let us now consider the interference caused by user 1 on user 2 in the asymptotic limit. To this end, we compute $\boldsymbol{h}_2^{\text{H}}\boldsymbol{p}_1$ for infinite P_{dl} using the result from (5.7). As for the channel estimate $\hat{\boldsymbol{h}}_k$, we deploy the asymptotic LS estimate $\hat{\tilde{\boldsymbol{h}}}_k$ in (5.12) leading to

$$\begin{aligned}
\boldsymbol{h}_2^{\text{H}}\boldsymbol{p}_1 &= \beta\boldsymbol{h}_2^{\text{H}}(\hat{\tilde{\boldsymbol{h}}}_1\|\hat{\tilde{\boldsymbol{h}}}_2\|_2 - \hat{\tilde{\boldsymbol{h}}}_2\hat{\tilde{\boldsymbol{h}}}_2^{\text{H}}\hat{\tilde{\boldsymbol{h}}}_1) \\
&= \beta\boldsymbol{h}_2^{\text{H}}(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_1\boldsymbol{h}_2^{\text{H}}\boldsymbol{\Phi}\underbrace{\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi}}_{\boldsymbol{I}_{T_{\text{dl}}}}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_2 - \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_2\boldsymbol{h}_2^{\text{H}}\boldsymbol{\Phi}\underbrace{\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi}}_{\boldsymbol{I}}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_1) \\
&= \beta(\boldsymbol{h}_2^{\text{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_1\boldsymbol{h}_2^{\text{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_2 - \boldsymbol{h}_2^{\text{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_2\boldsymbol{h}_2^{\text{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_1) \\
&= 0.
\end{aligned}$$

Similarly, we can prove that $\boldsymbol{h}_1^{\text{H}}\boldsymbol{p}_2 = 0$ for $P_{\text{dl}} \to \infty$.
Based on these results, we have shown that zero-forcing precoding based on the LS estimates leads asymptotically to an interference-free transmission.

5.2.1.2 Asymptotic behavior of zero-forcing based on LMMSE estimate

Now, we study the system's asymptotic behavior when the LMMSE estimates are used to compute the zero-forcing precoding matrix. Recall that the LMMSE channel estimate of user k is given by

$$\hat{\boldsymbol{h}}_k^{\text{LMMSE}} = \boldsymbol{C}_k\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\text{H}}\boldsymbol{C}_k\boldsymbol{\Phi} + \sigma_{1,k}^2\boldsymbol{I}_M)^{-1}\boldsymbol{h}_k.$$

Note that for $P_{\text{dl}} \to \infty$, $\sigma_{1,k}^2 \to 0$ and thus the LMMSE channel estimate can be written as

$$\hat{\tilde{\boldsymbol{h}}}_k = \boldsymbol{C}_k\underbrace{\boldsymbol{\Phi}\left(\boldsymbol{\Phi}^{\text{H}}\boldsymbol{C}_k\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^{\text{H}}}_{\boldsymbol{C}_{\Phi k}}\boldsymbol{h}_k = \boldsymbol{C}_k\boldsymbol{C}_{\Phi k}\boldsymbol{h}_k. \tag{5.13}$$

We again consider the term responsible for the interference gain $\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{p}_1$ as $P_{\mathrm{dl}} \to \infty$ using the channel estimates $\hat{\hat{\boldsymbol{h}}}_k$ in (5.13) and the precoder expression (5.7)

$$\begin{aligned}\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{p}_1 =&\beta\boldsymbol{h}_2^{\mathrm{H}}(\hat{\hat{\boldsymbol{h}}}_1\|\hat{\hat{\boldsymbol{h}}}_2\|^2 - \hat{\hat{\boldsymbol{h}}}_2\hat{\hat{\boldsymbol{h}}}_2^{\mathrm{H}}\,\hat{\hat{\boldsymbol{h}}}_1)\\ =&\beta\boldsymbol{h}_2^{\mathrm{H}}(\boldsymbol{C}_1\boldsymbol{C}_{\Phi 1}\boldsymbol{h}_1\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{C}_{\Phi 2}\boldsymbol{C}_2\boldsymbol{C}_2\boldsymbol{C}_{\Phi 2}\boldsymbol{h}_2 - \boldsymbol{C}_2\boldsymbol{C}_{\Phi 2}\boldsymbol{h}_2\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{C}_{\Phi 2}\boldsymbol{C}_2\boldsymbol{C}_1\boldsymbol{C}_{\Phi 1}\boldsymbol{h}_1)\\ =&\beta\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{C}_{\Phi 2}\boldsymbol{C}_2(\boldsymbol{C}_2\boldsymbol{C}_{\Phi 2}\boldsymbol{h}_2\boldsymbol{h}_2^{\mathrm{H}} - \boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{C}_2\boldsymbol{C}_{\Phi 2}\boldsymbol{h}_2\boldsymbol{I}_M)\boldsymbol{C}_1\boldsymbol{C}_{\Phi 1}\boldsymbol{h}_1\\ =&\beta\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{C}_{\Phi 2}\boldsymbol{C}_2\Big(\boldsymbol{C}_2\boldsymbol{C}_{\Phi 2}\boldsymbol{h}_2\boldsymbol{h}_2^{\mathrm{H}} - \mathrm{tr}(\boldsymbol{C}_2\boldsymbol{C}_{\Phi 2}\boldsymbol{h}_2\boldsymbol{h}_2^{\mathrm{H}})\boldsymbol{I}_M\Big)\boldsymbol{C}_1\boldsymbol{C}_{\Phi 1}\boldsymbol{h}_1\\ \neq&0.\end{aligned}$$

Obviously, the term $\boldsymbol{h}_2^{\mathrm{H}}\boldsymbol{p}_1$ does not converge to 0 when $P_{\mathrm{dl}} \to \infty$. Therefore, there exists a residual interference when zero-forcing precoding based on LMMSE channel estimation is used.

5.2.2 Generalization of the Asymptotic Behavior for $K > 2$

Using the compact formulation of the LS channel estimate in (5.9), we can show for the general case $K > 2$, that the zero-forcing precoder based on the LS estimate leads to an interference-free transmission in the high-power limit. Since the downlink training noise variance is inversely proportional to the downlink transmit power, the matrix $\boldsymbol{Y}$ of all channel observations (see (5.10)) converges to $\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{H}$ for $P_{\mathrm{dl}} \to \infty$. Thus, it holds for the zero-forcing precoding matrix when $P_{\mathrm{dl}} \to \infty$ that

$$\boldsymbol{P}^{\mathrm{zf,asy}} = \alpha\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{H}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{H})^{-1}. \tag{5.14}$$

Now, we consider the effective channel given by the combination of the channel matrix with the precoding matrix, i.e.,

$$\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{zf,asy}} = \alpha\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{H}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{H})^{-1} = \alpha\boldsymbol{I}_K. \tag{5.15}$$

It is clear from (5.15), that the effective channel asymptotically converges to a scaled identity, which provides a generalization of the behavior discussed in the previous subsection for $K = 2$. For $P_{\mathrm{dl}} \to \infty$, the zero-forcing precoder based on the LS estimate ensures an interference-free transmission if $T_{\mathrm{dl}} \geq K$.

5.3 Numerical Results

In this section, we assess the system performance in terms of the downlink achievable sum rate when zero-forcing precoding is used. Here, we demonstrate the validity of the previously derived analytical results. We denote by

- ZF-genie: the zero-forcing precoder computed based on the true channel. This serves as a performance benchmark for the other approaches.
- ZF-LS: the zero-forcing precoder computed based on the LS channel estimate.
- ZF-LMMSE: the zero-forcing precoder determined based on LMMSE channel estimate.

To this end, we evaluate the performance based on channel covariance matrices obtained by fitting a Gaussian mixture model (GMM) [61] to the downlink observations of channels generated according to the QUAsi Deterministic RadIo Channel GenerAtor (QuaDRiGa) [62] of an urban macro-cell non-line of sight scenario. Here, a single carrier with a carrier frequency of 3.5 GHz is considered. The base station is equipped with antennas in a uniform linear array (ULA) configuration, is placed at a height of 20m, and covers a 120°-sector. The distances between the users and the base station are randomly drawn between 35m and 500m.

For a setup with K users, we select uniformly at random K out of the $N_c = 64$ GMM components. The channels are then generated based on the K randomly selected covariance matrices according to (3.1).
For the Monte-Carlo runs, we consider $N_{\text{cov}} = 100$ users' constellations (random covariance matrix selections) with $N_{\text{ch}} = 100$ channel realizations per constellation.

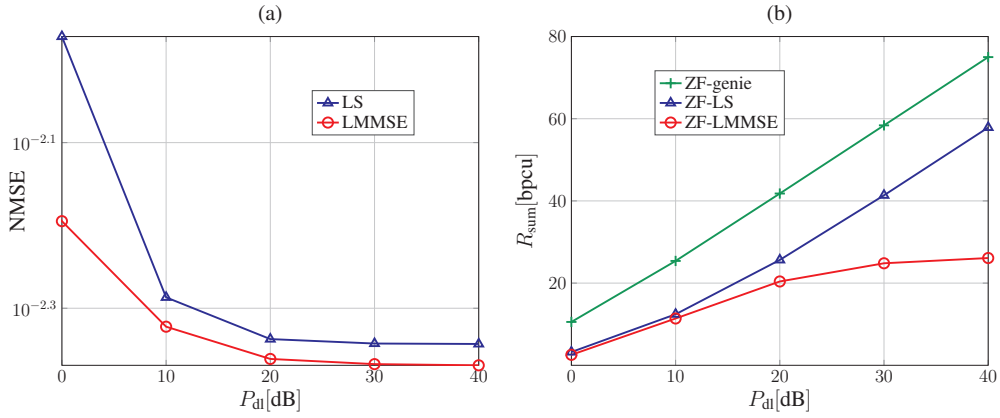

Fig. 5.1: NMSE of the LS and LMMSE channel estimators (a) and achievable sum rate of zero-forcing precoder based on true channel, LS and LMMSE channel estimates (b) for $M = 32$, $K = 5$, $T_{\text{dl}} = 16$ and sub-DFT pilot matrix

We assume a scenario with $M = 32$ antennas and $K = 5$ users and a varying number of pilots. Here, we define the pilot matrix as the first T_{dl} columns of the

$M \times M$ discrete Fourier Transform (DFT) matrix, which we refer to as the sub-DFT matrix.

We first consider the case where $T_{\text{dl}} = 16 = M/2$. In Fig. 5.1(a), we show the normalized MSE (NMSE) of the channel given by

$$\text{NMSE} = \mathbb{E}\left[\frac{\|\hat{\boldsymbol{H}} - \boldsymbol{H}\|_{\text{F}}^2}{\|\boldsymbol{H}\|_{\text{F}}^2}\right] \tag{5.16}$$

where the expectation is evaluated using Monte Carlo runs for the different users' constellations and their corresponding channel realizations. Clearly, the LMMSE channel estimate exhibits a better quality in terms of the NMSE over the whole power range. Interestingly, both estimates have different error floors due to pilot contamination among the channel coefficients. Furthermore, both estimators show a saturation of the NMSE in the high-power regime. This is because the systematic error caused by insufficient pilots does not vanish when the downlink transmit power is increased.

Although intuitively, one would expect that designing the precoder based on a better quality channel estimate would lead to a better system throughput, the results in Fig. 5.1(b) contradict this intuition. Here, we observe that the achievable rate obtained with zero forcing based on the LMMSE channel estimate saturates due to residual interference, as shown through our asymptotic analysis. Additionally, the interference-free behavior of the zero-forcing precoder based on the LS channel estimate is confirmed, as the rate curve is not saturating and exhibits the same high-SNR slope as the perfect CSI case.

We now plot the achievable sum rate using different pilot numbers and zero-forcing precoder based on the LS and LMMSE channel estimates in Fig. 5.2(a) and (b), respectively. On the one hand, we can observe in Fig. 5.2(a) that the approach based on LS eliminates the inter-user interference completely in the high-power regime for $T_{\text{dl}} > K$. In this case, the rate curves exhibit the same slope as in the full pilot case, and hence, the full degrees of freedom are achieved, with a degradation in performance in the form of an offset, as the number of downlink pilots is reduced. Only in the case $T_{\text{dl}} < K$, the corresponding rate curve saturates in the high-power region, which is compliant with our analysis in the previous section.

On the other hand, as soon as the number of downlink pilots falls below the number of antennas, the rate curves for zero-forcing based on the LMMSE channel estimates exhibit a saturation at high transmit power values due to the residual interference. It can be seen in Fig. 5.2(b) that already for $T_{\text{dl}} = 31$ the high-SNR slope is remarkably smaller than in the full pilot case ($T_{\text{dl}} = 32$).

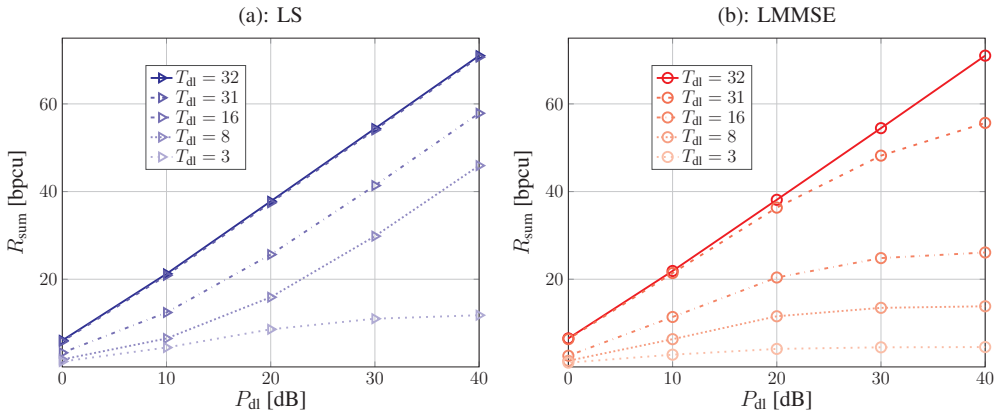

Fig. 5.2: Achievable sum rate versus P_{dl} when zero-forcing precoding is applied based on LS (a) and LMMSE (b) channel estimates for different numbers of pilots (©IEEE 2023)

5.4 Interpretation and Discussion

The asymptotic interference-free behavior of the zero-forcing precoder based on the LS channel estimate has been shown to hold under the condition $T_{\text{dl}} \geq K$. This implies that the maximal number of users that could be served interference-free in the high-power regime is upper-bounded by the number of employed pilots.

Having this in mind, we can gain an intuition about the obtained behavior. First, the LS channel estimate can be seen as a global channel estimator in the sense that for each user, the corresponding observation is mapped to the channel estimate using the pilot matrix, which is common for all users.
Recall that in the limit $P_{\text{dl}} \to \infty$, the LS channel estimate of user k is given by

$$\hat{\boldsymbol{h}}_k^{\text{LS}} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k. \tag{5.17}$$

This means, the channel estimate $\hat{\boldsymbol{h}}^{\text{LS}}$ is a projection of the true channel $\boldsymbol{h}_k$ using the orthogonal projector $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\text{H}}$, which is common for all users.
Now, the zero-forcing precoder that is based on the LS channel estimate is operating on the T_{dl}-dimensional subspace spanned by the columns of the pilot matrix, which eases the task of interference suppression due to the common subspace in which the channel estimates of all users reside.

The picture is different when it comes to the LMMSE channel estimate. For $P_{\text{dl}} \to \infty$, recall that the latter converges to

$$\hat{\boldsymbol{h}}_k^{\text{LMMSE}} = \boldsymbol{C}_k\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\text{H}}\boldsymbol{C}_k\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k. \tag{5.18}$$

The kth user channel is mapped to the channel estimate with a user-specific matrix, which is beneficial in terms of the estimation quality but impairs the interference mitigation using zero-forcing precoding due to the different subspaces it has to handle.

Let's suppose we use a global channel estimator based on the channel covariance matrices of all users. One possible way is to define the sum covariance matrix

$$\boldsymbol{C} = \sum_{k=1}^{K} \boldsymbol{C}_k \tag{5.19}$$

and the asymptotic global LMMSE channel estimate for user k is hence given by

$$\hat{\boldsymbol{h}}_k^{\text{gl.LMMSE}} = \boldsymbol{C}\boldsymbol{\Phi}(\boldsymbol{\Phi}^{\text{H}}\boldsymbol{C}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{h}_k. \tag{5.20}$$

As can be seen in Fig. 5.3(a), the considered global channel estimator exhibits a worse

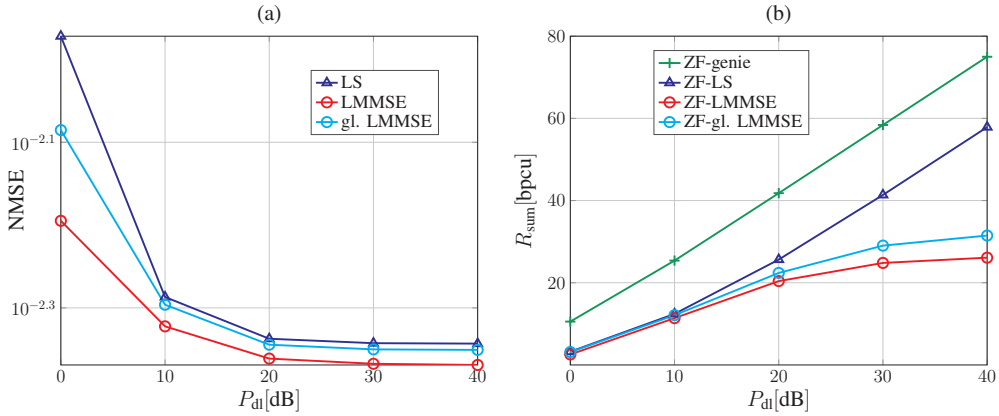

Fig. 5.3: NMSE of the LS, LMMSE and global LMMSE channel estimators (a) and average achievable sum rate of zero-forcing precoder based on the different channel estimates (b) for $M = 32$, $K = 5$, $T_{\text{dl}} = 16$ and sub-DFT pilot matrix

estimation quality compared to the conventional LMMSE estimator, but the error floor is different from LS as well. Although it shows a remarkable performance gain in terms of the achievable rate when zero-forcing is applied (see Fig.5.3(b)), saturation is still present in the high-power regime, and an interference-free behavior is not attained.

One further step is to adapt the pilot matrix to match the structure of the underlying channels. This could be achieved by defining the columns of $\boldsymbol{\Phi}$ as the T_{dl} eigenvectors corresponding to the dominant eigenvalues of $\boldsymbol{C}$ (referred to by "Dom-EVs").

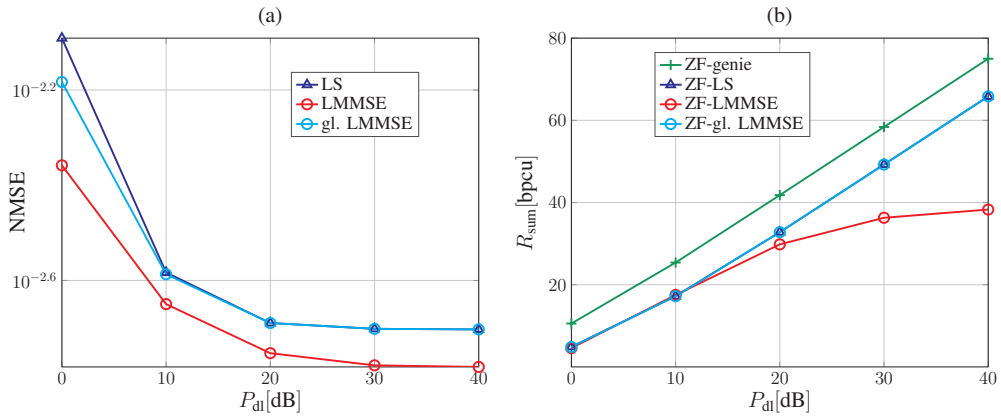

Fig. 5.4: NMSE of the LS, LMMSE, and global LMMSE channel estimators (a) and achievable sum rate of zero-forcing precoder based on the different channel estimates (b) for $M = 32$, $K = 5$, $T_{\text{dl}} = 16$ and "Dom-EVs" pilot matrix

One can easily see in Fig. 5.4 that the global LMMSE estimator now converges to the performance of the LS estimator not only in terms of the NMSE but also in terms of the achievable rate. Here, the asymptotic interference-free behavior is achieved by tuning the pilot matrix depending on the channel covariance matrices, which demonstrates how crucial pilot design can be for the system. The latter will be further discussed in Chapter 8.

Now, the question that has to be raised is, what if the asymptotic interference-free condition, i.e., $T_{\text{dl}} \geq K$, is not satisfied? We have seen that the zero-forcing precoder based on LS suffers from residual interference, and alternative precoding strategies have thus to be investigated.

6 Statistical Precoding

In the case of highly limited pilot resources, even the zero-forcing precoder based on LS is unable to handle the inter-user interference (see Chapter 5). In this chapter, we leverage the channel second-order statistics for transmit strategy design and establish a statistical precoding approach that can outperform the zero-forcing method. The proposed precoder exhibits a structure similar to the standard matched filter, allowing for a closed-form expression of the achievable spectral efficiency. As opposed to the conventional matched filter, the proposed precoder additionally exploits the channel structure captured by the covariance matrices. Consequently, the whole precoder optimization and the related computational complexity are pushed to the less frequent covariance variation [18], [45].

Despite the performance gains, the statistical precoder still exhibits a saturation of the achievable rate in the high-SNR region. We, therefore, propose to boost the performance by implementing one-layer RS in the system. Such an MA scheme allows for partial decoding of the interference, hence enhancing the system's degrees of freedom at high transmit powers [46]. Inspired by the works [63] and [64], we additionally develop a low-complexity precoding method for the RS setup.

6.1 Statistical Precoding

We propose a statistical precoder having the following form [45]

$$\boldsymbol{p}_k = \boldsymbol{A}_k \boldsymbol{y}_k \tag{6.1}$$

where $\boldsymbol{y}_k$ is the observation vector received at the base station via feedback in the uplink (see (3.2)). The $\boldsymbol{A}_k \in \mathbb{C}^{M \times T_{\text{dl}}}$ is a deterministic transformation matrix that depends on the channel statistics and the pilot matrix.
We also refer to the precoder in (6.1) as the bilinear precoder [18] since the received signal from (3.16) can be written as follows

$$r_k = \boldsymbol{h}_k^{\text{H}} \boldsymbol{A}_k \boldsymbol{y}_k s_k + \sum_{j \neq k} \boldsymbol{h}_k^{\text{H}} \boldsymbol{A}_j \boldsymbol{y}_j s_j + v_k. \tag{6.2}$$

Clearly, r_k is not only linear in the data symbol s_k (as is commonly the case) but also in the observation vector $\boldsymbol{y}_k$ and hence the bilinear terminology.

We shall note that the statistical precoder can alternatively be computed as

$$\boldsymbol{p}_k = \tilde{\boldsymbol{A}}_k \hat{\boldsymbol{h}}_k \tag{6.3}$$

where $\hat{\boldsymbol{h}}_k$ denotes the channel estimate (see Section 3.2.1) and $\tilde{\boldsymbol{A}}_k \in \mathbb{C}^{M \times M}$. In this case, the statistical precoder can be seen as a generalization of the standard matched filter with a statistical pre-filter $\tilde{\boldsymbol{A}}_k$.
Since the deterministic matrices are our design variables, adopting the precoder structure in (6.3) would lead to M^2 optimization variables compared to MT_{dl} if we use the precoder given by (6.1). Hence, to reduce the computational complexity, we will use the precoder given by (6.1) due to the assumption $T_{\text{dl}} < M$. Note that in this case, the intermediate step of explicitly computing the channel estimate can be circumvented, and the observation vector is directly mapped into the precoding vector using the statistical matrix $\boldsymbol{A}_k$.

6.1.1 Problem Formulation

The transformation matrices $\boldsymbol{A}_k$ have to be optimized so that the spectral efficiency is maximized. Due to the special structure of the proposed statistical precoder, (3.27) is well-suited as a figure-of-merit since all the expectation terms can be computed in closed form.

Using (3.27), the effective achievable rate of user k is evaluated as

$$R_k = \log_2 \left(1 + \frac{|\text{tr}(\boldsymbol{A}_k \boldsymbol{\Phi}^{\text{H}} \boldsymbol{C}_k)|^2}{\sum_{j=1}^{K} \text{tr}(\boldsymbol{A}_j \boldsymbol{C}_{\boldsymbol{y}_j} \boldsymbol{A}_j^{\text{H}} \boldsymbol{C}_k) + 1} \right) \tag{6.4}$$

where $\boldsymbol{C}_{\boldsymbol{y}_j}$ denotes the covariance matrix of $\boldsymbol{y}_j$ (see (3.3)).
With the vectorized transformation $\boldsymbol{a}_k = \text{vec}(\boldsymbol{A}_k)$, we can rewrite the effective achievable rate (6.4) as follows

$$R_k = \log_2 \left(1 + \frac{|\boldsymbol{q}_k^{\text{H}} \boldsymbol{a}_k|^2}{\sum_{j=1}^{K} \boldsymbol{a}_j^{\text{H}} \boldsymbol{Q}_{j,k} \boldsymbol{a}_j^{\text{H}} + 1} \right) \tag{6.5}$$

where we introduced the following variables

$$\boldsymbol{q}_k = (\boldsymbol{\Phi}^{\text{T}} \otimes \boldsymbol{I}_M) \boldsymbol{c}_k, \; \boldsymbol{c}_k = \text{vec}(\boldsymbol{C}_k), \tag{6.6}$$

$$\boldsymbol{Q}_{j,k} = \boldsymbol{C}_{\boldsymbol{y}_j}^{\text{T}} \otimes \boldsymbol{C}_k. \tag{6.7}$$

The detailed derivations of (6.4) and (6.5) are provided in Appendix A.2.

Due to the ergodic nature of transmission (the data symbols transmission spanning a long sequence of channel instances), we can formulate a long-term power constraint as follows

$$\mathbb{E}[\sum_{j=1}^{K} \|\boldsymbol{p}_j\|^2] \leq P_{\text{dl}}. \tag{6.8}$$

Note that the expectation here is taken w.r.t. the channels.
Given the statistical precoder in (6.1), the expectation in (6.8) can be expressed as a function of the vectorized transformation matrices as follows (see Appendix A.2 for the derivation)

$$\mathbb{E}[\sum_{j=1}^{K} \|\boldsymbol{p}_j\|^2] = \sum_{j=1}^{K} \boldsymbol{a}_j^{\mathrm{H}} \boldsymbol{F}_j \boldsymbol{a}_j = \boldsymbol{a}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a} \leq P_{\text{dl}} \tag{6.9}$$

where the optimization variable $\boldsymbol{a} = [\boldsymbol{a}_1^{\mathrm{T}}, \ldots, \boldsymbol{a}_K^{\mathrm{T}}]^{\mathrm{T}}$ contains the vectorized transformation matrices of all users and we additionally defined

$$\boldsymbol{F}_j = \boldsymbol{C}_{\boldsymbol{y}_j}^{\mathrm{T}} \otimes \boldsymbol{I}_M, \tag{6.10}$$

$$\boldsymbol{F} = \text{blkdiag}(\boldsymbol{F}_1, \ldots, \boldsymbol{F}_K). \tag{6.11}$$

Now, the maximization of the effective downlink sum rate under the average power constraint can be formulated as the following optimization problem

$$\max_{\boldsymbol{a}} \sum_{k=1}^{K} \log_2 \left(1 + \frac{|\boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{a}_k|^2}{\sum_{j=1}^{K} \boldsymbol{a}_j^{\mathrm{H}} \boldsymbol{Q}_{j,k} \boldsymbol{a}_j^{\mathrm{H}} + 1} \right) \quad \text{s.t.} \quad \boldsymbol{a}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a} \leq P_{\text{dl}}. \tag{6.12}$$

Note that the SINR of user k depends not only on the transformation vector $\boldsymbol{a}_k$ but also on the transformation vectors $\boldsymbol{a}_j$ of all other users $j \neq k$. This results in a coupling of the individual users' rates, which makes the optimization problem (6.12) non-convex and hence intractable.

6.1.2 Solution Approaches

In the following, we will present two approaches to solve (6.12) for the transformation vector $\boldsymbol{a}$. The first method uses the Lagrangian dual transform and FP to reformulate the underlying optimization problem and solve it in three alternating steps (see Section 4.1). The second approach is based on the uplink-downlink SINR duality [65]. To this end, an appropriate dual uplink system is constructed whereby hidden convexity properties

of (6.12) can be unveiled. The dual transformation vectors can be found in closed form in the uplink setup, and the optimization problem in the dual uplink boils down to a power allocation problem. Due to the uplink-downlink duality, the downlink transformation vectors can be easily inferred from the dual uplink solution via a linear relationship.

6.1.2.1 Statistical Precoding via Fractional Programming

The Lagrangian dual transform and FP approach reviewed in Section 4.1 will be used to solve (6.12). Our optimization variable is $\boldsymbol{a}$, and the numerator and denominator of the user's k SINR are respectively given by

$$N_k(\boldsymbol{a}) = |S_k(\boldsymbol{a})|^2 = |\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_k|^2, \tag{6.13}$$

$$D_k(\boldsymbol{a}) = \sum_{j=1}^{K} \boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{Q}_{j,k}\boldsymbol{a}_j + 1. \tag{6.14}$$

For fixed $\boldsymbol{a}$, the auxiliary variables γ_k and β_k are respectively updated according to (cf. (4.7) and (4.13))

$$\gamma_k = \frac{N_k(\boldsymbol{a})}{D_k(\boldsymbol{a})}, \tag{6.15}$$

$$\beta_k = \frac{S_k(\boldsymbol{a})^*}{N_k(\boldsymbol{a}) + D_k(\boldsymbol{a})}. \tag{6.16}$$

We now need to solve the optimization problem w.r.t. $\boldsymbol{a}$ for fixed $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ (cf. (4.14))

$$\begin{aligned} &\max_{\boldsymbol{a}} \sum_{k=1}^{K}(1+\gamma_k)\Big(2\Re\{\beta_k\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_k\} - |\beta_k|^2(|\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{a}_k|^2 + \sum_{j=1}^{K}\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{Q}_{j,k}\boldsymbol{a}_j + 1)\Big) \\ &\text{s.t.} \quad \boldsymbol{a}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{a} \leq P_{\text{dl}}. \end{aligned} \tag{6.17}$$

Due to the convexity of (6.17), the Karush-Kuhn-Tucker (KTT) conditions are not only necessary but also sufficient for optimality.
By introducing the Lagrangian multiplier λ associated with the power constraint in (6.17), we obtain the following expression for $\boldsymbol{a}_k$ from the first optimality condition

$$\boldsymbol{a}_k(\lambda) = \beta_k^*(1+\gamma_k)\Big(|\beta_k|^2\boldsymbol{q}_k\boldsymbol{q}_k^{\mathrm{H}} + \sum_{i=1}^{K}|\beta_i|^2(1+\gamma_i)\boldsymbol{Q}_{k,i} + \lambda\boldsymbol{F}_k\Big)^{-1}\boldsymbol{z}_k. \tag{6.18}$$

The solution for $\boldsymbol{a}_k$ stated in (6.18) is given in a quasi-closed form since the optimal Lagrangian multiplier λ still needs to be found via line search such that the power

constraint is fulfilled, i.e.,

$$\lambda^\star = \min\{\lambda \geq 0 : \boldsymbol{a}(\lambda)^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}(\lambda) \leq P_{\text{dl}}\}. \tag{6.19}$$

The one-dimensional search (e.g., via bisection [57]) requires the computation of the inverse of an $M \times T_{\text{dl}}$ matrix in each iteration. It can thus be costly, especially in a setup with a massive number of base station antennas, i.e., $M \gg 1$.
The line search can be circumvented by applying the rescaling approach discussed in Section 4.4. The constrained optimization problem in (6.17) is thereby transformed into an equivalent unconstrained one by including the power constraint into the objective function itself as follows

$$\begin{aligned}\max_{\boldsymbol{a}} \sum_{k=1}^{K} (1+\gamma_k) \Big(& 2\Re\{\beta_k \boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{a}_k\} - |\beta_k|^2 (|\boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{a}_k|^2 \\ & + \sum_{j=1}^{K} \boldsymbol{a}_j^{\mathrm{H}} \boldsymbol{Q}_{j,k} \boldsymbol{a}_j + \frac{1}{P_{\text{dl}}} \sum_{j=1}^{K} \boldsymbol{a}_j^{\mathrm{H}} \boldsymbol{F}_j \boldsymbol{a}_j) \Big).\end{aligned} \tag{6.20}$$

Hence, a closed-form expression for the unconstrained transformation vector $\boldsymbol{a}_k^{\text{unconst}}$ can be found by considering the first-order optimality condition of (6.20)

$$\boldsymbol{a}_k^{\text{unconst}} = \beta_k^* (1+\gamma_k) \Big(|\beta_k|^2 \boldsymbol{q}_k \boldsymbol{q}_k^{\mathrm{H}} + \sum_{i=1}^{K} |\beta_i|^2 (1+\gamma_i) (\boldsymbol{Q}_{k,i} + \frac{1}{P_{\text{dl}}} \boldsymbol{F}_k) \Big)^{-1} \boldsymbol{z}_k. \tag{6.21}$$

The solution obtained in (6.21) needs to be rescaled after convergence so that the power constraint is satisfied with equality, where the scaling factor α is given by

$$\alpha^2 = \frac{P_{\text{dl}}}{\boldsymbol{a}^{\text{unconst},\mathrm{H}} \boldsymbol{F} \boldsymbol{a}^{\text{unconst}}} \tag{6.22}$$

and the final solution for the transformation vector reads as

$$\boldsymbol{a}_k^\star = \alpha \boldsymbol{a}_k^{\text{unconst}}. \tag{6.23}$$

The alternating steps to solve the optimization problem in (6.12) using fractional programming are summarized in Algorithm 1.

Algorithm 1 Statistical Precoding Algorithm via FP

1: Initialize $\boldsymbol{a}$ with a feasible solution, such that $\boldsymbol{a}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{a} = P_{\mathrm{dl}}$
2: **while** No convergence is reached **do**
3: Compute γ_k according to (6.15)
4: Update β_k using following (6.16)
5: Compute the transformation vector $\boldsymbol{a}_k^{\mathrm{unconst}}$ according to (6.21)
6: **end while**
7: Scale $\boldsymbol{a}_k^{\mathrm{unconst}}$ to obtain the feasible final solution $\boldsymbol{a}_k^{\star}$ (see (6.23))

6.1.2.2 Statistical Precoding via Uplink-Downlink SINR Duality

An alternative way to solve (6.12) is to leverage the uplink-downlink SINR duality [65], [66], whereby an appropriate dual uplink system is constructed. To this end, the channels in the dual uplink are the same as in the downlink. The uplink-downlink duality is established by assuming that an equalizer $\boldsymbol{g}_k$ is applied to the received vector at the base station to retrieve an estimate of the kth user signal.
Similar to the precoder, the receive filter or equalizer $\boldsymbol{g}_k$ has the following form

$$\boldsymbol{g}_k = \boldsymbol{T}_k \boldsymbol{y}_k \tag{6.24}$$

where $\boldsymbol{T}_k$ denotes the dual transformation matrix to be optimized and $\boldsymbol{y}_k$ is the channel observation of user k (see (3.2)).

The received vector at the base station in such an uplink system is given by

$$\boldsymbol{r}^{\mathrm{dual}} = \sum_{j=1}^{K} \sqrt{\mu_j}\boldsymbol{h}_j s_j + \boldsymbol{v}^{\mathrm{dual}} \tag{6.25}$$

where $\mu_j \geq 0$ is the power of the jth user and $\boldsymbol{v}^{\mathrm{dual}} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I}_M)$ denotes the additive white Gaussian noise at the receive antennas.

The data signal estimate $\hat{s}_k$ of user k is obtained by applying the equalizer $\boldsymbol{g}_k$ to the received vector $\boldsymbol{r}^{\mathrm{dual}}$

$$\hat{s}_k = \boldsymbol{g}_k^{\mathrm{H}} \boldsymbol{r}^{\mathrm{dual}} = \sum_{j=1}^{K} \sqrt{\mu_j}\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{h}_j s_j + \boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{v}^{\mathrm{dual}}. \tag{6.26}$$

Similar to the bound on the downlink effective SINR in (3.27), we can formulate the effective dual uplink SINR as follows [18]

$$\mathrm{SINR}_k^{\mathrm{dual}} = \frac{\mu_k |\mathbb{E}[\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{h}_k]|^2}{\sum_{j=1}^{K} \mu_j \mathbb{E}[|\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{h}_j|^2] - \mu_k |\mathbb{E}[\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{h}_k]|^2 + \mathbb{E}[\|\boldsymbol{g}_k\|^2]}. \tag{6.27}$$

Due to the structure of the equalizer $\boldsymbol{g}_k$ in (6.24), all the expectation terms in (6.27) can be computed in closed form, and the effective dual uplink SINR thus reads as

$$\text{SINR}_k^{\text{dual}} = \frac{\mu_k |\boldsymbol{q}_k^{\text{H}} \boldsymbol{t}_k|^2}{\sum_{j=1}^{K} \mu_j \boldsymbol{t}_k^{\text{H}} \boldsymbol{Q}_{k,j} \boldsymbol{t}_k + \boldsymbol{t}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{t}_k} \tag{6.28}$$

where $\boldsymbol{t}_k = \text{vec}(\boldsymbol{T}_k)$ denotes the vectorized transformation matrix of user k.

We will now show that there exists a one-to-one relationship between the effective downlink and uplink SINRs. To this end, we state the following theorem

Theorem 1. *For the downlink SINR of some generic user k in the form*

$$\mathit{SINR}_k = \frac{\boldsymbol{a}_k^{\text{H}} \boldsymbol{q}_k \boldsymbol{q}_k^{\text{H}} \boldsymbol{a}_k}{\sum_j \boldsymbol{a}_j^{\text{H}} \boldsymbol{Q}_{j,k} \boldsymbol{a}_j + 1} \tag{6.29}$$

and the corresponding dual uplink SINR

$$\mathit{SINR}_k^{\mathit{dual}} = \frac{\mu_k \boldsymbol{t}_k^{\text{H}} \boldsymbol{q}_k \boldsymbol{q}_k^{\text{H}} \boldsymbol{t}_k}{\sum_j \mu_j \boldsymbol{t}_k^{\text{H}} \boldsymbol{Q}_{k,j} \boldsymbol{t}_k + \boldsymbol{t}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{t}_k} \tag{6.30}$$

with positive semi-definite matrix $\boldsymbol{Q}_{k,j}$ and positive definite matrix $\boldsymbol{F}_k$ the following holds

For any uplink transformation vector $\boldsymbol{t}_k$ and uplink powers μ_k such that $\sum_k \mu_k = P_{dl}$, we can find $\boldsymbol{a}_k$ such that $\sum_k \boldsymbol{a}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{a}_k = P_{dl}$ and $\mathit{SINR}_k = \mathit{SINR}_k^{dual}$ for all k. Conversely, for any downlink transformation vector $\boldsymbol{a}_k$ with $\sum_k \boldsymbol{a}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{a}_k = P_{dl}$, we can find $\boldsymbol{t}_k$ and μ_k such that $\sum_k \mu_k = P_{dl}$ and $\mathit{SINR}_k = \mathit{SINR}_k^{dual}$ for all k.

Proof. The proof is analogous to the one presented in [65] and is provided in Appendix A.4. □

Note that the effective dual uplink SINR of user k only depends on that user's transformation vector $\boldsymbol{t}_k$ and is also independent of a scaling of $\boldsymbol{t}_k$. Hence, a closed-form solution for $\boldsymbol{t}_k$ can be obtained for fixed uplink powers μ_k

$$\boldsymbol{t}_k^{\star} = \Big(\sum_{j=1}^{K} \mu_j \boldsymbol{Q}_{k,j} + \boldsymbol{F}_k \Big)^{-1} \boldsymbol{q}_k = \boldsymbol{X}_k^{-1} \boldsymbol{q}_k \tag{6.31}$$

where we introduced the matrix

$$\boldsymbol{X}_k = \sum_{j=1}^{K} \mu_j \boldsymbol{Q}_{k,j} + \boldsymbol{F}_k. \tag{6.32}$$

Inserting the solution stated in (6.31) into the dual uplink SINR expression (6.28) yields

$$\text{SINR}_k^{\text{dual},\star} = \mu_k \boldsymbol{q}_k^{\text{H}} \boldsymbol{X}_k^{-1} \boldsymbol{q}_k = \mu_k d_k \tag{6.33}$$

where we defined the coefficient $d_k = \boldsymbol{q}_k^{\text{H}} \boldsymbol{X}_k^{-1} \boldsymbol{q}_k$.

Now, what remains is to optimize the uplink powers such that the effective dual sum rate is maximized, i.e.,

$$\max_{\mu_1,\ldots,\mu_K} R_{\text{sum}}^{\text{dual},\star} = \max_{\mu_1,\ldots,\mu_K} \sum_{k=1}^{K} \log_2(1+\mu_k d_k) \quad \text{s.t.} \quad \sum_{k=1}^{K} \mu_k \leq P_{\text{dl}}. \tag{6.34}$$

We have to emphasize that the coefficient d_k depends on the uplink powers $\mu_k, k = 1,\ldots,K$, and therefore, an iterative approach has to be employed to find the optimal uplink powers.

To this end, we employ an iterative local optimum search inspired by the work [67]. The idea is to quickly converge to a locally optimal power allocation starting from an initial value of the powers.

First, we consider the first optimality condition of (6.34) by computing the derivative of $R_{\text{sum}}^{\text{dual},\star}$ w.r.t. the power of an individual user k

$$\frac{\partial R_{\text{sum}}^{\text{dual},\star}}{\partial \mu_k} = \frac{1}{\log 2} \left(\frac{d_k}{1+\mu_k d_k} - \sum_{i=1}^{K} \frac{\mu_i e_{i,k}}{1+\mu_i d_i} \right) \tag{6.35}$$

where we introduced the scalar $e_{i,k}$ given by

$$e_{i,k} = \boldsymbol{q}_i^{\text{H}} \boldsymbol{X}_i^{-1} \boldsymbol{Q}_{i,k} \boldsymbol{X}_i^{-1} \boldsymbol{q}_i. \tag{6.36}$$

Now, we construct an equivalent interference-free channel $\tilde{\boldsymbol{D}} \in \mathbb{C}^{K\times K}$, whose sum rate can be written as

$$R_{\text{sum}}^{\text{eq}} = \sum_{k=1}^{K} \log_2(1+\mu_k \tilde{d}_k) \tag{6.37}$$

where $\tilde{d}_k$ denotes the kth diagonal element of $\tilde{\boldsymbol{D}}$. The derivative of $R_{\text{sum}}^{\text{eq}}$ w.r.t. μ_k is given by

$$\frac{\partial R_{\text{sum}}^{\text{eq}}}{\partial \mu_k} = \frac{1}{\log 2} \frac{\tilde{d}_k}{1+\mu_k \tilde{d}_k}. \tag{6.38}$$

The optimum power allocation for the constructed channel is the well-known waterfilling solution over the channel gains $\tilde{d}_k$ [68].

By enforcing the derivatives in (6.35) and (6.38) to be the same, we ensure that the KKT conditions of both problems are fulfilled because the same constraint holds and

the derivatives are identical. Since equating the derivatives might not always result in a non-negative value of the $\tilde{d}_k$, we set $\tilde{d}_k$ to 0 whenever this occurs.

The overall approach to finding the downlink transformation vectors $\boldsymbol{a}_k$ using the uplink-downlink SINR duality is summarized in Algorithm 2. Once the power allocation problem is solved, the dual transformation vectors can be computed according to (6.31) and finally $\boldsymbol{a}_k$ is inferred from $\boldsymbol{t}_k$ based on the duality transform as described in A.4. Here, N_{iter} denotes the maximum number of iterations, and $\delta \in [0, 1]$ is a

Algorithm 2 Statistical Precoding Algorithm via UL-DL Duality

Initialize $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_K]^{\mathrm{T}}$
for $n = 1$ to N_{iter} **do**
  for $k = 1$ to K **do**
    $\boldsymbol{X}_k \leftarrow \sum_j \mu_j \boldsymbol{Q}_{k,j} + \frac{1}{P_{\text{dl}}} \boldsymbol{F}_k$
    $d_k \leftarrow \boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{X}_k^{-1} \boldsymbol{q}_k$
  end for
  for $k = 1$ to K **do**
    $\ell_k \leftarrow \left(\frac{d_k}{1+\mu_k d_k} - \sum_{i=1}^{K} \frac{\mu_i e_{i,k}}{1+\mu_i d_i} \right)^{-1} - \mu_k$
    if $\ell_k < 0$ **then**
      $\ell_k \leftarrow \infty$
    end if
  end for
  Compute $\boldsymbol{\mu}_{\text{tmp}}$ via waterfilling over the K channels with noise variances $\ell_1, \ldots, \ell_K$
  $\boldsymbol{\mu} \leftarrow \delta \boldsymbol{\mu} + (1 - \delta) \boldsymbol{\mu}_{\text{tmp}}$
end for
Determine the transformation vectors $\boldsymbol{t}_k$ according to (6.31)
Find $\boldsymbol{a}_k$ based on $\boldsymbol{t}_k$ according to the linear transform in (A.35)

smoothing factor used to eliminate oscillations of the algorithm [67].

6.1.3 Numerical Results

In this section, we evaluate the performance of the statistical precoder obtained using the solution approaches presented in Section 6.1.2 and compare it to the standard matched filter (MF) and zero-forcing (ZF) precoder, where the latter two are based on the LS channel estimate (3.7) and normalized to satisfy the power constraint. We denote by

- Stat-FP: the statistical precoder computed according to the FP method from Section 6.1.2.1.

- Stat-UL-DL: the statistical precoder determined based on the uplink-downlink SINR duality approach in Section 6.1.2.2.

We assume a setup with $M = 32$, $K = 8$, and a varying number of orthogonal Gaussian pilots, which are generated at random for every covariance constellation. We additionally focus on the highly limited channel knowledge case with $T_{\text{dl}} < K$ since this is where the zero-forcing precoder, also based on the LS channel estimate, shows a poor performance. The results of the average achievable sum rate in bits per channel use (bpcu) are computed according to the ergodic spectral efficiency expression in (3.20). The expectation is evaluated using Monte Carlo simulations for $N_{\text{cov}} = 100$ covariance constellations and $N_{\text{ch}} = 100$ channel realizations per constellation. The covariance matrices and corresponding channels are generated in the same manner as discussed in Section 5.3.

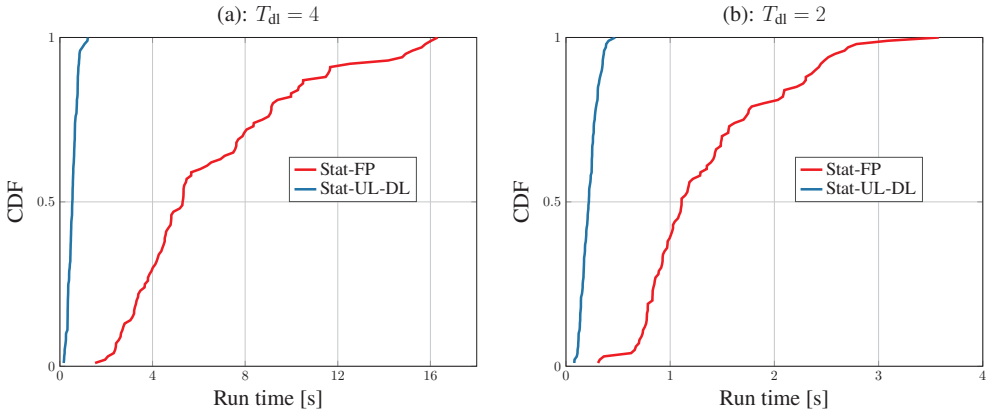

Fig. 6.1: CDF plots of the run time of the two statistical approaches at $P_{\text{dl}} = 40$ dB for $M = 32$, $K = 8$, and different T_{dl} values.

In Fig. 6.1, we show the cumulative distribution functions (CDFs) of the run times (in seconds) of the two statistical approaches evaluated for the $N_{\text{cov}} = 100$ constellations at a transmit power of $P_{\text{dl}} = 40$ dB for two different pilot numbers $T_{\text{dl}} = 4$ (Fig. 6.1(a)) and $T_{\text{dl}} = 2$ (Fig. 6.1(b)). It can be easily seen that in both cases, the Stat-UL-DL approach based on the uplink-downlink SINR duality shows a much lower run time than the Stat-FP method. This is because the precoder optimization problem in the former is reduced to a power allocation problem. The more

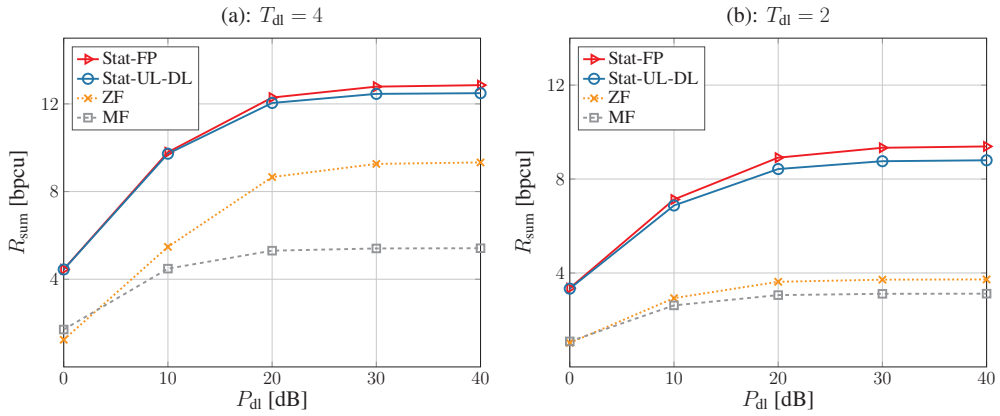

Fig. 6.2: Achievable sum rate versus P_{dl} for $M = 32$, $K = 8$ and different T_{dl} values.

time-consuming approach based on FP exhibits a slight advantage in the achievable throughput, especially in the high-power region, as seen in Fig. 6.2(b). Furthermore, the two statistical approaches outperform the conventional zero-forcing and matched filter precoders, and the gap becomes more significant for $T_{\text{dl}} = 2$. Despite this performance gain, the statistical precoders exhibit, however, a saturation of the sum rate in the high-power region because the system is interference-limited, and no interference management takes place. Hence, in order to boost performance, especially in this regime, we make use of RS, which is known to enhance the system's degrees of freedom [46].

6.2 Statistical Precoding with Rate Splitting

For the sake of simplicity, we assume that one-layer RS is implemented, i.e., only one layer of common messages for the K users is used as discussed in Section 3.3.2

Our aim is to design the common precoder $\boldsymbol{p}_{\text{c}}$ and the private precoders $\boldsymbol{p}_{\text{p},k}$, $k = 1, \ldots, K$ following the bilinear structure from Section 6.1, i.e.,

$$\boldsymbol{p}_{\text{c}} = \boldsymbol{A}_{\text{c}}\boldsymbol{y}, \tag{6.39}$$

$$\boldsymbol{p}_{\text{p},k} = \boldsymbol{A}_{\text{p},k}\boldsymbol{y}_k. \tag{6.40}$$

In (6.39), we introduced the vector $\boldsymbol{y}$ of all users' channel observations

$$\boldsymbol{y} = [\boldsymbol{y}_1^{\text{T}}, \boldsymbol{y}_2^{\text{T}}, \ldots, \boldsymbol{y}_K^{\text{T}}]^{\text{T}} \tag{6.41}$$

where the $\boldsymbol{y}_k$ are defined in (3.2).
The $\boldsymbol{A}_{\text{c}}$ and $\boldsymbol{A}_{\text{p},k}$ in (6.39) and (6.40), respectively, are the deterministic transformation

matrices which depend solely on the second-order channel statistics and the pilot matrix $\boldsymbol{\Phi}$. These matrices are, again, our design variables.
The structure of the common precoder in (6.39) implies that the latter has to be designed as the superposition of "individual precoders", i.e.,

$$\boldsymbol{p}_{\mathrm{c}} = \sum_{k=1}^{K} \boldsymbol{A}_{\mathrm{c},k} \boldsymbol{y}_k \tag{6.42}$$

with $\boldsymbol{A}_{\mathrm{c}} = [\boldsymbol{A}_{\mathrm{c},1}, \boldsymbol{A}_{\mathrm{c},2}, \ldots, \boldsymbol{A}_{\mathrm{c},K}]$. This choice is motivated by the multicast setup discussion in [69].

Recall that the transmit vector $\boldsymbol{x}$ must satisfy the average transmit power constraint

$$\mathbb{E}[\|\boldsymbol{x}\|^2] = \|\boldsymbol{p}_{\mathrm{c}}\|^2 + \sum_{i=1}^{K} \|\boldsymbol{p}_{\mathrm{p},i}\|^2 \leq P_{\mathrm{dl}}. \tag{6.43}$$

As we aim at optimizing the precoders based on the channel second-order statistics, we again consider the long-term power constraint by taking the expectation w.r.t. the channels over the expressions in (6.43). Inserting the expressions of the precoders (6.39) and (6.40) into the relaxed power constraint yields

$$\mathbb{E}[\|\boldsymbol{p}_{\mathrm{c}}\|^2] + \sum_{i=1}^{K} \mathbb{E}[\|\boldsymbol{p}_{\mathrm{p},i}\|^2] = \mathrm{tr}\left(\boldsymbol{A}_{\mathrm{c}} \boldsymbol{C}_{\boldsymbol{y}} \boldsymbol{A}_{\mathrm{c}}^{\mathrm{H}}\right) + \sum_{i=1}^{K} \mathrm{tr}\left(\boldsymbol{A}_{\mathrm{p},i} \boldsymbol{C}_{\boldsymbol{y}_i} \boldsymbol{A}_{\mathrm{p},i}^{\mathrm{H}}\right) \leq P_{\mathrm{dl}} \tag{6.44}$$

where we introduced $\boldsymbol{C}_{\boldsymbol{y}}$, the covariance matrix of $\boldsymbol{y}$ given by

$$\boldsymbol{C}_{\boldsymbol{y}} = \mathrm{blkdiag}\left(\boldsymbol{C}_{\boldsymbol{y}_1}, \boldsymbol{C}_{\boldsymbol{y}_2}, \ldots, \boldsymbol{C}_{\boldsymbol{y}_K}\right) \tag{6.45}$$

and $\boldsymbol{C}_{\boldsymbol{y}_k}$ is defined as the covariance matrix of $\boldsymbol{y}_k$ (see (3.3)). The block-diagonal structure of $\boldsymbol{C}_{\boldsymbol{y}}$ follows from the independence of the users' channel observations.
We now introduce the vectorized common and private transformation vectors

$$\boldsymbol{a}_{\mathrm{c}} = \mathrm{vec}(\boldsymbol{A}_{\mathrm{c}}), \quad \boldsymbol{a}_{\mathrm{p},i} = \mathrm{vec}(\boldsymbol{A}_{\mathrm{p},i}) \tag{6.46}$$

and $\boldsymbol{a}_{\mathrm{p}} = [\boldsymbol{a}_{\mathrm{p},1}^{\mathrm{T}}, \ldots, \boldsymbol{a}_{\mathrm{p},K}^{\mathrm{T}}]^{\mathrm{T}}$ containing the vectorized forms of the transformation matrices $\boldsymbol{A}_{\mathrm{p},1}, \ldots, \boldsymbol{A}_{\mathrm{p},K}$ of the private precoders.
The power constraint can be hence formulated as follows

$$\boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{c}} + \sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}} \boldsymbol{F}_i \boldsymbol{a}_{\mathrm{p},i} = \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{c}} + \boldsymbol{a}_{\mathrm{p}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{p}} \leq P_{\mathrm{dl}}. \tag{6.47}$$

where we applied the identities in (A.15) and (A.16) and used the matrices $\boldsymbol{F}_i$ and the matrix $\boldsymbol{F}$ defined in (6.10) and (6.11), respectively.

6.2.1 Problem Formulation

To optimize the transformation matrices $\boldsymbol{A}_{\mathrm{c}}$ and $\boldsymbol{A}_{\mathrm{p},k}$, we again reside to the spectral efficiency expression in (3.27), since it yields a closed form expression when using the statistical precoding approach.

Considering the RS setup, we deal with two kinds of messages: common and private. We, therefore, differentiate between the common rate $R_{\mathrm{c},k}$ and the private rate $R_{\mathrm{p},k}$ for each user k.

Recall that the received signal for the common message at user k is given by (see (3.31))

$$r_{\mathrm{c},k} = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}} s_{\mathrm{c}} + \sum_{i=1}^{K} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i} s_{\mathrm{p},i} + v_k. \tag{6.48}$$

We here assume that user k knows the effective channel coefficient $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}$ perfectly and is, therefore, able to perform perfect SIC. Hence, after decoding the common message, the latter can be entirely subtracted from the received signal. We thus obtain the received signal for the private message at user k as follows (see (3.32))

$$r_{\mathrm{p},k} = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},k} s_{\mathrm{p},k} + \sum_{i \neq k} \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i} s_{\mathrm{p},i} + v_k. \tag{6.49}$$

We shall emphasize that the assumptions on the channel knowledge at the transmitter (the base station) and the users are different. While the effective channel coefficients $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}$ and $\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i}$ are assumed to be perfectly known at the users, the base station only knows their statistics in the form of their respective means, i.e., $\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}]$ and $\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i}]$.

Under such assumptions, we can employ the spectral efficiency expression given in (3.27) to write the effective common and private SINRs of user k as follows

$$\begin{aligned} \mathrm{SINR}_{\mathrm{c},k} &= \frac{|\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}]|^2}{\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}|^2] - |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}]|^2 + \sum_{i=1}^{K} \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i}|^2] + 1}, \\ \mathrm{SINR}_{\mathrm{p},k} &= \frac{|\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},k}]|^2}{\sum_{i=1}^{K} \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i}|^2] - |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},k}]|^2 + 1}. \end{aligned} \tag{6.50}$$

After inserting the expressions for the statistical precoders (see (6.39) and (6.40)) into (6.50), we obtain the following common and private SINRs of user k

$$\mathrm{SINR}_{\mathrm{c},k} = \frac{|\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}|^2}{\boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{Q}_k \boldsymbol{a}_{\mathrm{c}} + |\boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{p},k}|^2 + \sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}} \boldsymbol{Q}_{i,k} \boldsymbol{a}_{\mathrm{p},i} + 1}, \tag{6.51}$$

$$\mathrm{SINR}_{\mathrm{p},k} = \frac{|\boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{p},k}|^2}{\sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}} \boldsymbol{Q}_{i,k} \boldsymbol{a}_{\mathrm{p},i} + 1}. \tag{6.52}$$

The variables $\boldsymbol{z}_k$, $\boldsymbol{q}_k$, $\boldsymbol{Q}_k$ and $\boldsymbol{Q}_{i,k}$ are respectively defined in (A.29), (A.17), (A.33), and (A.18).
The derivation of the common SINR expression is provided in Appendix A.3, whereas the private SINR is obtained in the same manner as explained in Appendix A.2 for the non-RS setup.

Now that we derived the SINR expressions and the power constraint, we can formulate the downlink sum rate maximization problem as follows

$$\begin{aligned} \max_{R_\mathrm{c},\boldsymbol{a}_\mathrm{c},\boldsymbol{a}_\mathrm{p}} \; R_\mathrm{c} + \sum_{k=1}^{K} R_{\mathrm{p},k} \quad \text{s.t.} \quad & R_\mathrm{c} \leq R_{\mathrm{c},k}, \forall k \\ & \boldsymbol{a}_\mathrm{c}^\mathrm{H} \boldsymbol{F} \boldsymbol{a}_\mathrm{c} + \boldsymbol{a}_\mathrm{p}^\mathrm{H} \boldsymbol{F} \boldsymbol{a}_\mathrm{p} \leq P_\mathrm{dl} \end{aligned} \tag{6.53}$$

where the common and private rates are respectively given by

$$R_{\mathrm{c},k} = \log_2(1 + \mathrm{SINR}_{\mathrm{c},k}), \quad R_{\mathrm{p},k} = \log_2(1 + \mathrm{SINR}_{\mathrm{p},k}). \tag{6.54}$$

The first constraint in (6.53) ensures that the common rate is not larger than the minimum common rate of all users, which allows the common message to be decoded by all users.

The optimization problem in (6.53) is non-convex and hence intractable. In the following, we will present different approaches to tackle this problem.

6.2.2 Statistical Rate Splitting Precoding Approach via Joint Optimization

One possible way to solve (6.53) is to apply the Lagrangian dual transform from Section 4.1 twice and jointly optimize the common and private transformation vectors. To this end, we define the following numerators and denominators

$$\begin{aligned} N_{\mathrm{c},k} &= |S_{\mathrm{c},k}|^2 = |\boldsymbol{z}_k^\mathrm{H} \boldsymbol{a}_\mathrm{c}|^2, \\ D_{\mathrm{c},k} &= \boldsymbol{a}_\mathrm{c}^\mathrm{H} \boldsymbol{Q}_k \boldsymbol{a}_\mathrm{c} + |\boldsymbol{q}_k^\mathrm{H} \boldsymbol{a}_{\mathrm{p},k}|^2 + \sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^\mathrm{H} \boldsymbol{Q}_{i,k} \boldsymbol{a}_{\mathrm{p},i} + 1, \\ N_{\mathrm{p},k} &= |S_{\mathrm{p},k}|^2 = |\boldsymbol{q}_k^\mathrm{H} \boldsymbol{a}_{\mathrm{p},k}|^2, \\ D_{\mathrm{c},k} &= \sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^\mathrm{H} \boldsymbol{Q}_{i,k} \boldsymbol{a}_{\mathrm{p},i} + 1. \end{aligned}$$

As discussed in Section 4.1, we need to introduce the following sets of auxiliary variables

$$\begin{aligned} \boldsymbol{\gamma}_\mathrm{c} &= [\gamma_{\mathrm{c},1}, \ldots, \gamma_{\mathrm{c},K}]^\mathrm{T}, & \boldsymbol{\gamma}_\mathrm{p} &= [\gamma_{\mathrm{p},1}, \ldots, \gamma_{\mathrm{p},K}]^\mathrm{T}, \\ \boldsymbol{\beta}_\mathrm{c} &= [\beta_{\mathrm{c},1}, \ldots, \beta_{\mathrm{c},K}]^\mathrm{T}, & \boldsymbol{\beta}_\mathrm{p} &= [\beta_{\mathrm{p},1}, \ldots, \beta_{\mathrm{p},K}]^\mathrm{T}. \end{aligned}$$

The first two sets $\boldsymbol{\gamma}_\mathrm{c}$ and $\boldsymbol{\gamma}_\mathrm{p}$ are used to externalize the optimization variables from the logarithm function, while $\boldsymbol{\beta}_\mathrm{c}$ and $\boldsymbol{\beta}_\mathrm{p}$ serve to decouple the numerators and denominators of the multiple ratios.

The transformed optimization problem is hence formulated as follows (cf. (4.10) and (4.11))

$$\begin{aligned}\max_{\substack{\boldsymbol{a}_\mathrm{c},\boldsymbol{a}_\mathrm{p},\\ \boldsymbol{\gamma}_\mathrm{c},\boldsymbol{\gamma}_\mathrm{p},\\ \boldsymbol{\beta}_\mathrm{c},\boldsymbol{\beta}_\mathrm{p}}} \min_{k}\ & \log_2(1+\gamma_{\mathrm{c},k}) - \gamma_{\mathrm{c},k} + (1+\gamma_{\mathrm{c},k})\left(2\Re\{\beta_{\mathrm{c},k}S_{\mathrm{c},k}\} - |\beta_{\mathrm{c},k}|^2(N_{\mathrm{c},k}+D_{\mathrm{c},k})\right) \\ & + \sum_{i=1}^{K} \log_2(1+\gamma_{\mathrm{p},i}) - \gamma_{\mathrm{p},i} + (1+\gamma_{\mathrm{p},i})\left(2\Re\{\beta_{\mathrm{p},i}S_{\mathrm{p},i}\} - |\beta_{\mathrm{p},i}|^2(N_{\mathrm{p},i}+D_{\mathrm{p},i})\right) \\ & \text{s.t.} \quad \boldsymbol{a}_\mathrm{c}^\mathrm{H}\boldsymbol{F}\boldsymbol{a}_\mathrm{c} + \boldsymbol{a}_\mathrm{p}^\mathrm{H}\boldsymbol{F}\boldsymbol{a}_\mathrm{p} \leq P_\mathrm{dl}.\end{aligned} \tag{6.55}$$

For fixed $\boldsymbol{a}_\mathrm{c}$ and $\boldsymbol{a}_\mathrm{p}$, the optimal auxiliary variables are given by

$$\begin{aligned}\gamma_{\mathrm{c},k}^{\mathrm{opt}} &= \frac{N_{\mathrm{c},k}}{D_{\mathrm{c},k}}, & \gamma_{\mathrm{p},k}^{\mathrm{opt}} &= \frac{N_{\mathrm{p},k}}{D_{\mathrm{p},k}}, \\ \beta_{\mathrm{c},k}^{\mathrm{opt}} &= \frac{S_{\mathrm{c},k}^*}{N_{\mathrm{c},k}+D_{\mathrm{c},k}}, & \beta_{\mathrm{p},k}^{\mathrm{opt}} &= \frac{S_{\mathrm{p},k}^*}{N_{\mathrm{p},k}+D_{\mathrm{p},k}}.\end{aligned}$$

Then, once these auxiliary variables are fixed, the optimization problem (6.55) is convex in the transformation vectors $\boldsymbol{a}_\mathrm{c}$ and $\boldsymbol{a}_\mathrm{p}$ and can be thus solved using standard convex solvers such as CVX [70]. Due to the high number of variables to be jointly optimized, this solution approach is not computationally efficient, and alternative ways to tackle (6.53) will be proposed in the sequel.

6.2.3 Statistical Rate Splitting Precoding via Three-Step Optimization

To reduce the computational cost, we propose a solution approach similar to [71], which mainly consists of decomposing (6.53) into three parts, where each subproblem can be solved efficiently [46].

To this end, we introduce α_c as the power fraction allocated to the common stream transmission. Hence, the power constraint in (6.47) can be reformulated as the combination of the following two constraints

$$\begin{aligned}\boldsymbol{a}_\mathrm{c}^\mathrm{H}\boldsymbol{F}\boldsymbol{a}_\mathrm{c} &\leq \alpha_\mathrm{c}P_\mathrm{dl}, \\ \boldsymbol{a}_\mathrm{p}^\mathrm{H}\boldsymbol{F}\boldsymbol{a}_\mathrm{p} &\leq (1-\alpha_\mathrm{c})P_\mathrm{dl}.\end{aligned} \tag{6.56}$$

(I) First, assuming a fixed power fraction α_{c}, we formulate the private rate optimization problem

$$\max_{\boldsymbol{a}_{\mathrm{p}}} \sum_{k=1}^{K} R_{\mathrm{p},k} \quad \text{s.t.} \quad \boldsymbol{a}_{\mathrm{p}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{p}} \leq (1-\alpha_{\mathrm{c}}) P_{\mathrm{dl}} \tag{6.57}$$

where we implicitly assume that the private precoder optimization does not affect the common rate. This is, in general, not the case since the common SINR depends on the private precoders as well. However, this assumption makes the optimization problem (6.53) more tractable by decoupling the private and the common rate optimization.

(II) Then, for fixed private precoders and for a given power fraction α_{c}, we recast the common precoder optimization problem as a max-min beamforming problem

$$\max_{\boldsymbol{a}_{\mathrm{c}}} \min_{k} R_{\mathrm{c},k} \quad \text{s.t.} \quad \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{c}} \leq \alpha_{\mathrm{c}} P_{\mathrm{dl}}. \tag{6.58}$$

Since the rate expression $R_{\mathrm{c},k}$ is monotonically increasing in $\mathrm{SINR}_{\mathrm{c},k}$, we can reformulate (6.58) as a max-min problem of the common SINR, that is,

$$\max_{\boldsymbol{a}_{\mathrm{c}}} \min_{k} \mathrm{SINR}_{\mathrm{c},k} \quad \text{s.t.} \quad \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{c}} \leq \alpha_{\mathrm{c}} P_{\mathrm{dl}}. \tag{6.59}$$

(III) Finally, the optimization over the power fraction α_{c} allocated to the common stream transmission has to be performed. Finding an optimal power fraction would require joint optimization over the precoders so that the overall sum rate is maximized. To circumvent the complexity of such an approach, we opt for a line search method that returns the scalar α_{c}, which maximizes the overall rate.

6.2.3.1 Private Precoder Optimization

The private precoder optimization boils down to the sum rate maximization problem considered in Section 6.1.1, and we again make use of the FP framework presented in Section 4.1 to solve the underlying problem.

Note that (6.57) has exactly the same form as (6.12), where $\boldsymbol{a}$, SINR_k and P_{dl} are respectively replaced by $\boldsymbol{a}_{\mathrm{p}}$, $\mathrm{SINR}_{\mathrm{p},k}$ and $(1-\alpha_{\mathrm{c}})P_{\mathrm{dl}}$. Therefore, the solution approach to the private precoder optimization problem (6.57) follows the same three steps described in Section 6.1.2.1 and is summarized in Algorithm 3.

Algorithm 3 Private Precoder Optimization Algorithm

1: Initialize $\boldsymbol{a}_{\mathrm{p}}$ with a feasible solution
2: **while** No convergence is reached **do**
3: Compute γ_k according to (6.15)
4: Find β_k based on (6.16)
5: Update the transformation vectors $\boldsymbol{a}_{\mathrm{p},k}$ according to (6.21)
6: **end while**
7: Scale the transformation vector $\boldsymbol{a}_{\mathrm{p}}$ to satisfy the power constraint (see (6.23))

6.2.3.2 Common Precoder Optimization

To solve the common precoder optimization in (6.59), we again exploit the quadratic transform [27] presented in Section 6.1.2.1 to reformulate the optimization problem by introducing a set of auxiliary variables.
The transformed optimization problem is then tackled using two approaches: One is based on an interior-point solver, e.g., CVX [70]. The other approach is an iterative algorithm inspired by [69]. It consists of iteratively updating the common transformation vector while ensuring that the minimum common SINR does not decrease after each update, which gives rise to the name *"Common SINR increasing algorithm"*.

First, we introduce a slack variable t such that the minimization w.r.t. the user index is recast as a constraint and where the resulting optimization problem is equivalent to the original max-min problem. We hence obtain the following optimization problem

$$\begin{aligned} \max_{\boldsymbol{a}_{\mathrm{c}},t} t \quad \text{s.t.} \quad & \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{a}_{\mathrm{c}} \leq \alpha_{\mathrm{c}}P_{\mathrm{dl}} \\ & \mathrm{SINR}_{\mathrm{c},k} \geq t, \forall k. \end{aligned} \tag{6.60}$$

To decouple the numerator and the denominator of the common SINR, we apply the quadratic transform and introduce an auxiliary variable η_k for each of the K common SINRs. This yields the following optimization problem equivalent to (6.60)

$$\begin{aligned} \max_{\boldsymbol{a}_{\mathrm{c}},t,\boldsymbol{\eta}} t \quad \text{s.t.} \quad & \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{a}_{\mathrm{c}} \leq \alpha_{\mathrm{c}}P_{\mathrm{dl}} \\ & 2\Re\{\eta_k^*\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{a}_{\mathrm{c}}\} - |\eta_k|^2(\boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{Q}_k\boldsymbol{a}_{\mathrm{c}} + \sigma_k^2) \geq t, \forall k \end{aligned} \tag{6.61}$$

where $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K]^{\mathrm{T}}$ denotes the collection of the K auxiliary variables.
We additionally defined $\sigma_k^2 = |\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_{\mathrm{p},k}|^2 + \sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}}\boldsymbol{Q}_{i,k}\boldsymbol{a}_{\mathrm{p},i} + 1$, which obviously does not depend on the common transformation vector $\boldsymbol{a}_{\mathrm{c}}$.

For fixed $\boldsymbol{a}_\mathrm{c}$, the optimal auxiliary variables η_k^{opt} can be determined by setting the derivative of the left-hand side of the SINR constraint in (6.61) to zero, which yields the following result

$$\eta_k^{\text{opt}} = \frac{\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_\mathrm{c}}{\boldsymbol{a}_\mathrm{c}^{\mathrm{H}} \boldsymbol{Q}_k \boldsymbol{a}_\mathrm{c} + \sigma_k^2}. \tag{6.62}$$

Plugging this expression into (6.61) leads to the same optimization problem in (6.60).

If we fix the auxiliary variables η_k, we end up with the following optimization problem

$$\begin{aligned} \max_{\boldsymbol{a}_\mathrm{c}, t} \; t \quad \text{s.t.} \quad & \boldsymbol{a}_\mathrm{c}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_\mathrm{c} \leq \alpha_\mathrm{c} P_\mathrm{dl} \\ & 2\Re\{\eta_k^* \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_\mathrm{c}\} - |\eta_k|^2 (\boldsymbol{a}_\mathrm{c}^{\mathrm{H}} \boldsymbol{Q}_k \boldsymbol{a}_\mathrm{c} + \sigma_k^2) \geq t, \forall k \end{aligned} \tag{6.63}$$

which is a QCQP and can be solved via CVX [70]. Alternatively, we propose an iterative approach inspired by [69] that aims to increase the common SINR in each of its iterations.

Consider again the optimization problem in (6.61) where the auxiliary variables η_k are kept constant. If we denote by $\ell(n)$ the user with the least common SINR after iteration n, the update rule for the common transformation vector $\boldsymbol{a}_\mathrm{c}$ at iteration $n+1$ is given by

$$\boldsymbol{a}_\mathrm{c}^{(n+1)} = (1-u)\boldsymbol{a}_\mathrm{c}^{(n)} + v \boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}} \boldsymbol{a}_\mathrm{c}^{(n),\perp} \tag{6.64}$$

where $u \in [0,1]$ denotes the step size and hence controls how much from the previous update is included in the current update, and $\boldsymbol{Q}_k$ is defined in (A.33). The length of the step size is adapted during the algorithm to speed up the convergence.

The second term in (6.64) involves different variables. v is a complex scalar that we have to tune together with the vector $\boldsymbol{a}_\mathrm{c}^{(n),\perp}$ for the update in (6.64) to ensure that the minimum common SINR at least does not decrease in step $n+1$ compared to step n and that the power constraint is satisfied. Applying $\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}$ (the inverse of the square root matrix of $\boldsymbol{Q}_{\ell(n)} = \boldsymbol{Q}_{\ell(n)}^{\frac{1}{2},\mathrm{H}} \boldsymbol{Q}_{\ell(n)}^{\frac{1}{2}}$) to $\boldsymbol{a}_\mathrm{c}^{(n),\perp}$ simplifies the tuning of the parameters v and $\boldsymbol{a}_\mathrm{c}^{(n),\perp}$ as it is presented in Appendix A.5.

Together with the conditions implied by the power constraint, the parameters v and $\boldsymbol{a}_\mathrm{c}^{(n),\perp}$ are given by (see Appendix A.5 for the derivation)

$$v = \sqrt{\alpha_\mathrm{c} P_\mathrm{dl}(2u-u^2)} \mathrm{e}^{-\mathrm{j}\angle\left(\boldsymbol{t}_\mathrm{c}^{(n),\mathrm{H}} \boldsymbol{a}_\mathrm{c}^{(n),\perp}\right)}, \tag{6.65}$$

$$\boldsymbol{a}_\mathrm{c}^{(n),\perp,\prime} = \boldsymbol{T}^{(n)} \boldsymbol{t}_\mathrm{c}^{(n)}, \tag{6.66}$$

$$\boldsymbol{a}_\mathrm{c}^{(n),\perp} = \frac{\boldsymbol{a}_\mathrm{c}^{(n),\perp,\prime}}{\|\boldsymbol{F}^{\frac{1}{2}} \boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}} \boldsymbol{a}_\mathrm{c}^{(n),\perp,\prime}\|_2} \tag{6.67}$$

where $\boldsymbol{t}_{\mathrm{c}}^{(n)} = \eta_{\ell(n)} \boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2},\mathrm{H}} \boldsymbol{z}_{\ell(n)} - |\eta_{\ell(n)}|^2 (1-u) \boldsymbol{Q}_{\ell(n)}^{\frac{1}{2}} \boldsymbol{a}_{\mathrm{c}}^{(n)}$.
The matrix $\boldsymbol{T}^{(n)}$ in (6.66) is the orthogonal projector onto the nullspace of the vector $\boldsymbol{t}^{(n)} = \boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2},\mathrm{H}} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}^{(n)}$ and is given by

$$\boldsymbol{T}^{(n)} = \boldsymbol{I}_{MT_{\mathrm{dl}}K} - \frac{\boldsymbol{t}^{(n)} \boldsymbol{t}^{(n),\mathrm{H}}}{\|\boldsymbol{t}^{(n)}\|_2^2}. \tag{6.68}$$

Suppose that $\ell(n)$ is the user with the least common SINR at the beginning of iteration n. Then, we compute the parameters v and $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$ according to (6.65) and (6.67), respectively. Using these parameters, we can find the temporary update $\boldsymbol{a}_{\mathrm{c}}^{\mathrm{temp}}$ with the current step size. This update might lead to an SINR of some other user k smaller than $\gamma_{\mathrm{c}}^{\mathrm{min}} = \mathrm{SINR}_{\mathrm{c},\ell(n)}$. If this is the case, then the update is not successful, and the step size is halved. Otherwise, the latter is increased by a factor of 2, and the update is performed as long as it does not drop the resulting minimum common SINR below $\gamma_{\mathrm{c}}^{\mathrm{min}}$.

The common precoder optimization is summarized in Algorithm 4. Note that in step 4, the common transformation vector $\boldsymbol{a}_{\mathrm{c}}$ can be either found via solving (6.63) using CVX or following the iterative common SINR increasing procedure summarized in Algorithm 5.

Algorithm 4 Common Precoder Optimization Algorithm

1: Initialize $\boldsymbol{a}_{\mathrm{c}} = \sqrt{\alpha_{\mathrm{c}} P_{\mathrm{dl}}} \frac{\boldsymbol{a}}{\|\boldsymbol{F}^{\frac{1}{2}} \boldsymbol{a}\|_2}$ with $\boldsymbol{a} = [\boldsymbol{a}_1^{\mathrm{T}}, \ldots, \boldsymbol{a}_K^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{a}_i = \mathrm{vec}(\boldsymbol{\Phi})$
2: **while** no convergence is reached **do**
3: Compute η_k according to (6.62)
4: Run the iterative common SINR increasing algorithm presented in Algorithm 5
5: **end while**

6.2.3.3 Power Fraction Optimization

For both the private and the common precoder optimization problems, we assumed that the power fraction allocated to the common stream is fixed. Now, in order to find the power fraction that maximizes the overall sum rate, we consider an outer loop for the power fraction optimization and perform a line search in the interval $[0, 1]$. The inner loop consists of the first two optimization problems stated before, namely the private precoder optimization and the common precoder optimization.

To find $\alpha_{\mathrm{c}} \in [0, 1]$, the power fraction allocated to the common stream, we perform a one-dimensional search using the golden section method, a simple approach to find

Algorithm 5 Common SINR Increasing Algorithm

Initialize: $u_{\max} \leftarrow 0.2$, $u \leftarrow u_{\max}$, $n \leftarrow 1$, $N_{\max} \leftarrow 100$
while $n \leq N_{\max}$ **do**
    $n \leftarrow n + 1$
    $\ell(n) \leftarrow \arg\min_{k} \text{SINR}_{\text{c},k}$ (see (6.51))
    $\gamma_{\text{c}}^{\min} \leftarrow \text{SINR}_{\text{c},\ell(n)}$
    Compute $\boldsymbol{a}_{\text{c}}^{(n),\perp}$ according to (6.67)
    Compute v according to (6.65)
    Find $\boldsymbol{a}_{\text{c}}^{\text{temp}} \leftarrow (1-u)\boldsymbol{a}_{\text{c}}^{(n)} + v\boldsymbol{a}_{\text{c}}^{(n),\perp}$
    Calculate the temporary minimum common SINR $\gamma_{\text{c}}^{\text{temp}}$
    if $\gamma_{\text{c}}^{\text{temp}} > \gamma_{\text{c}}^{\min}$ **then**
        $\boldsymbol{a}_{\text{c}}^{(n+1)} \leftarrow \boldsymbol{a}_{\text{c}}^{\text{temp}}$
        if $u \leq u_{\max}/2$ **then**
            $u \leftarrow 2u$
        end if
    else
        $\boldsymbol{a}_{\text{c}}^{(n+1)} \leftarrow \boldsymbol{a}_{\text{c}}^{(n)}$
        $u \leftarrow u/2$
    end if
end while

the maximum/minimum of a certain function over some interval $[a, b]$ by efficiently shrinking the search interval after each step [72].

For our setup, the function of interest is the sum rate $R_{\text{sum}}^{\text{RS}} = R_{\text{c}} + \sum_{k=1}^{K} R_{p,k}$, i.e., the objective of (6.53). The search interval is given by $[a, b] = [0, 1]$ and the method operates as follows. First, two intermediate values $\alpha_{\text{c}1}$ and $\alpha_{\text{c}2}$ are found according to

$$\alpha_{\text{c}1} = a + (1 - g)(b - a) \tag{6.69}$$

$$\alpha_{\text{c}2} = a + g(b - a) \tag{6.70}$$

where g is the inverse of the golden ratio $\frac{\sqrt{5}+1}{2}$ which gives rise to the name of this method.

Now, for the given $\alpha_{\text{c}1}$ and $\alpha_{\text{c}2}$, the two potential values of α_{c}, we run the private precoder optimization approach described in Sec. 6.2.3.1 for a fixed common precoder. We then proceed with solving the common precoder optimization problem following the approach presented in Sec. 6.2.3.2.
We hence obtain two values for the sum rate $R_{\text{sum},1} = R_{\text{sum}}(\alpha_{\text{c}1})$ and $R_{\text{sum},2} = R_{\text{sum}}(\alpha_{\text{c}2})$ corresponding to $\alpha_{\text{c}1}$ and $\alpha_{\text{c}2}$.
If $R_{\text{sum},1} > R_{\text{sum},2}$, then the search interval is reduced and the new boundaries are given by a and $b = \alpha_{\text{c}2}$, otherwise the new interval boundaries are $a = \alpha_{\text{c}1}$ and b.
Given the new search interval, we can recompute $\alpha_{\text{c}1}$ and $\alpha_{\text{c}2}$ according to (6.69) and (6.70) and then go through the previously described steps until the size of the interval given by $|b - a|$ drops below a certain threshold or the maximum number of iterations is reached.

6.2.4 Low-Complexity Statistical Rate Splitting Precoding Approach via Power Optimization

Inspired by the approach proposed in [64] for a TDD setup and to further reduce the complexity related to the computation of the transformation matrices, one strategy is to fix the direction of the precoders and solely optimize the power allocation among the common and private streams.

To this end, the precoded transmit vector $\boldsymbol{x}$ can be rewritten as

$$\boldsymbol{x} = \sqrt{\mu_{\text{c}}}\boldsymbol{p}_{\text{c}} s_{\text{c}} + \sum_{i=1}^{K} \sqrt{\mu_{\text{p},i}}\boldsymbol{p}_{\text{p},i} s_{\text{p},i} \tag{6.71}$$

where μ_{c} and $\mu_{\text{p},i}$ are the powers used for transmitting the common stream and the ith private stream, respectively, and must satisfy the total transmit power

$$\mu_{\text{c}} + \sum_{i=1}^{K} \mu_{\text{p},i} \leq P_{\text{dl}}. \tag{6.72}$$

Additionally, it holds for the precoders that

$$\mathbb{E}[\|\boldsymbol{p}_{\mathrm{c}}\|^2] = 1, \quad \mathbb{E}[\|\boldsymbol{p}_{\mathrm{p},i}\|^2] = 1,\ i = 1, \ldots, K. \tag{6.73}$$

The received signal for the common message at user k is given by

$$r_{\mathrm{c},k} = \sqrt{\mu_{\mathrm{c}}}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{c}}s_{\mathrm{c}} + \sum_{i=1}^{K}\sqrt{\mu_{\mathrm{p},i}}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{p},i}s_{\mathrm{p},i} + v_k. \tag{6.74}$$

Assuming perfect CSI at the receiver, i.e., that $\sqrt{\mu_{\mathrm{c}}}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{c}}$ is known at the kth user, perfect SIC can be performed and the received signal for the private message can be written as

$$r_{\mathrm{p},k} = \sum_{i=1}^{K}\sqrt{\mu_{\mathrm{p},i}}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{p},i}s_{\mathrm{p},i} + v_k. \tag{6.75}$$

We aim to maximize the system throughput by optimizing the lower bound on the spectral efficiency (3.27) as it can be expressed in closed form for the considered statistical precoding approach.
In an analogous manner to the derivations in Appendix A.2, we can write the lower bounds on the common and private SINRs as a function of the transformation vectors and the power factors

$$\mathrm{SINR}_{\mathrm{c},k} = \frac{\mu_{\mathrm{c}}|\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{a}_{\mathrm{c}}|^2}{\mu_{\mathrm{c}}\boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{Q}_k\boldsymbol{a}_{\mathrm{c}} + \mu_{\mathrm{p},k}|\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_{\mathrm{p},k}|^2 + \sum_{i=1}^{K}\mu_{\mathrm{p},i}\boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}}\boldsymbol{Q}_{i,k}\boldsymbol{a}_{\mathrm{p},i} + 1}, \tag{6.76}$$

$$\mathrm{SINR}_{\mathrm{p},k} = \frac{\mu_{\mathrm{p},k}|\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_{\mathrm{p},k}|^2}{\sum_{i=1}^{K}\mu_{\mathrm{p},i}\boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}}\boldsymbol{Q}_{i,k}\boldsymbol{a}_{\mathrm{p},i} + 1}. \tag{6.77}$$

To reduce the computational complexity related to the precoder design, we assume that the vectorized (unnormalized) transformations are given as follows

$$\boldsymbol{a}_{\mathrm{c}} = \sum_{i=1}^{K} w_i\boldsymbol{z}_i, \tag{6.78}$$

$$\boldsymbol{a}_{\mathrm{p},k} = \boldsymbol{q}_k, k = 1, \ldots, K. \tag{6.79}$$

The choice for the private transformation is motivated by the matched filter precoder based on instantaneous channel knowledge (the precoder, in the latter case, is simply the channel estimate). Here, we consider the filter matched to the channel statistical information. Note that the numerator of the private SINR in (6.77) has a similar form to the one of the SINR when perfect CSI is available, where $\boldsymbol{q}_k$ and $\boldsymbol{a}_{\mathrm{p},k}$ are analogous

to the channel and the precoding vectors, respectively.
As for the common transformation, we consider in (6.78) a weighted matched-filter-like approach, where the weights $w_i, i = 1, \ldots, K$, have to be optimized.

In order to fulfill the constraints in (6.73), we additionally must normalize the transformations $\boldsymbol{a}_{\mathrm{c}}, \boldsymbol{a}_{\mathrm{p},i}, \forall i$. Using the bilinear precoder structure (see (6.39) and (6.40)), we can rewrite the constraints in (6.73) as a function of the deterministic transformations as follows

$$\mathbb{E}[\|\boldsymbol{p}_{\mathrm{c}}\|^2] = \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{a}_{\mathrm{c}} = 1, \tag{6.80}$$

$$\mathbb{E}[\|\boldsymbol{p}_{\mathrm{p},i}\|^2] = \boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}} \boldsymbol{F}_i \boldsymbol{a}_{\mathrm{p},i} = 1, i = 1, \ldots, K \tag{6.81}$$

with $\boldsymbol{F}_i$ and $\boldsymbol{F}$ respectively defined in (6.10) and (6.11).

By imposing the structure of the deterministic transformations in (6.78) and (6.79), we only need to optimize the power coefficients $\mu_{\mathrm{c}}, \mu_{\mathrm{p},1}, \ldots, \mu_{\mathrm{p},K}$ and the weighting factors $w_1, \ldots, w_K$. To this end, we consider an alternating procedure, where the weights are optimized in one step, and the powers are updated in the following step. This approach is motivated by the fact that the private SINRs are independent of the common precoder $\boldsymbol{p}_{\mathrm{c}}$ and thus also $\boldsymbol{a}_{\mathrm{c}}$.

Hence, we can optimize the weights for fixed power allocation such that the minimum common SINR is maximized (we aim at maximizing the common rate, i.e., the minimum common rate among all users to ensure the decodability of the common stream by all users). As for the power allocation, we formulate the sum rate maximization problem and propose to solve it using the Lagrangian dual transform and the quadratic transform as described in Section 4.1. In the following, we will present the solution approach in detail.

6.2.4.1 Optimizing the Weighting Factors

As mentioned earlier, we propose to find the weights $w_1, \ldots, w_K$ such that the minimum common SINR is maximized. First, by defining the vector of all weights $\boldsymbol{w} = [w_1, \ldots, w_K]^{\mathrm{T}}$, we can rewrite the common transformation $\boldsymbol{a}_{\mathrm{c}}$ from (6.78) in a compact form as

$$\boldsymbol{a}_{\mathrm{c}} = \boldsymbol{Z}\boldsymbol{w}, \quad \text{with } \boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_K]. \tag{6.82}$$

Thus, the optimization problem for finding $\boldsymbol{w}$ can be formulated as

$$\max_{\boldsymbol{w}} \min_{k} \mathrm{SINR}_{\mathrm{c},k} \quad \text{s.t.} \quad \boldsymbol{w}^{\mathrm{T}} \boldsymbol{Z}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{Z} \boldsymbol{w} = 1. \tag{6.83}$$

By introducing a slack variable t, writing the common SINR as function of $\boldsymbol{w}$ and relaxing the constraint in (6.83), we obtain the following optimization problem

$$\begin{aligned} \max_{\boldsymbol{w},t} t \quad \text{s.t.} \quad & \boldsymbol{w}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{Z}\boldsymbol{w} \leq 1 \\ & \frac{\mu_{\mathrm{c}}|\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{Z}\boldsymbol{w}|^2}{\mu_{\mathrm{c}}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Q}_k\boldsymbol{Z}\boldsymbol{w}+\sigma_k^2} \geq t, \forall k \end{aligned} \tag{6.84}$$

where we defined $\sigma_k^2 = \mu_{\mathrm{p},k}|\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_{\mathrm{p},k}|^2 + \sum_{i=1}^{K}\mu_{\mathrm{p},i}\boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}}\boldsymbol{Q}_{i,k}\boldsymbol{a}_{\mathrm{p},i} + 1$, which does not depend on $\boldsymbol{a}_{\mathrm{c}}$ and therefore also not on $\boldsymbol{w}$.
Due to the structure of the common SINR, (6.84) is not convex. Hence, we again make use of the quadratic transform by introducing an auxiliary variable for each fraction. Let η_k be the auxiliary variable for the kth SINR, then (6.84) is reformulated as follows

$$\begin{aligned} \max_{\boldsymbol{w},t,\boldsymbol{\eta}} t \quad \text{s.t.} \quad & \boldsymbol{w}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{Z}\boldsymbol{w} \leq 1 \\ & 2\sqrt{\mu_{\mathrm{c}}}\Re\{\eta_k^*\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{Z}\boldsymbol{w}\} - |\eta_k|^2\left(\mu_{\mathrm{c}}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Q}_k\boldsymbol{Z}\boldsymbol{w}+\sigma_k^2\right) \geq t, \forall k. \end{aligned} \tag{6.85}$$

Note that by fixing $\boldsymbol{w}$, the auxiliary variable η_k can be computed in closed form according to

$$\eta_k^{\text{opt}} = \frac{\sqrt{\mu_{\mathrm{c}}}\boldsymbol{z}_k^{\mathrm{H}}\boldsymbol{Z}\boldsymbol{w}}{\mu_{\mathrm{c}}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Q}_k\boldsymbol{Z}\boldsymbol{w}+\sigma_k^2}. \tag{6.86}$$

For fixed η_k, (6.85) boils down to a convex optimization problem that can be solved using interior-point methods to obtain the optimal $\boldsymbol{w}$. We then need to alternate between updating η_k according to (6.86) and solving (6.85) for fixed η_k until convergence is reached.

6.2.4.2 Power Allocation

In the second step of our alternating procedure, we optimize the power coefficients $\boldsymbol{\mu} = [\mu_{\mathrm{c}}, \mu_{\mathrm{p},1}, \ldots, \mu_{\mathrm{p},K}]^{\mathrm{T}}$. To this end, we first formulate the downlink sum rate maximization problem based on the SINRs

$$\begin{aligned} \max_{\boldsymbol{\mu},R_{\mathrm{c}}} R_{\mathrm{c}} + \sum_{k=1}^{K}\log_2(1+\text{SINR}_{\mathrm{p},k}) \quad \text{s.t.} \quad & R_{\mathrm{c}} \leq \log_2(1+\text{SINR}_{\mathrm{c},k}),\ \forall k \\ & \mathbf{1}_K^{\mathrm{T}}\boldsymbol{\mu} \leq P_{\text{dl}}. \end{aligned} \tag{6.87}$$

To solve the non-convex optimization problem in (6.87), we again use the Lagrangian dual transform and the FP framework presented in Section 4.1. The introduced auxiliary variables are again given in closed form, whereas the optimization problem

w.r.t. $\boldsymbol{\mu}$ has to be solved using interior-point methods by leveraging its convexity. Note that here, the number of variables is drastically reduced compared to (6.55). While the proposed low-complexity approach tries to optimize $2(K+1)$ variables per iteration, the previously discussed three-step method needs to solve for $2MT_{\text{dl}}K+1$ variables.

6.2.5 Numerical Results

To assess the performance of the proposed statistical precoding approaches for the RS setup, we again consider $N_{\text{cov}} = 100$ covariance constellations and $N_{\text{ch}} = 100$ channel realization per covariance selection. Here, we focus on the limited channel knowledge case with $T_{\text{dl}} < K$ and evaluate the achievable sum rate results for the scenario described in Section 6.1.3, i.e., $M = 32$, $K = 8$, and randomly generated orthogonal Gaussian pilots for the following precoding approaches in the RS setup

- RS-3step method denotes the three-step statistical RS precoding approach presented in Section 6.2.3.
- RS-LowComp approach refers to the low-complexity statistical RS precoding strategy from Section 6.2.4.

We additionally compare these RS-based approaches to their non-RS counterparts in order to quantify the gain of RS. To this end, we consider

- noRS-FP method: this refers to the statistical precoder proposed in Section 6.1.2.1 based on the FP framework.
- noRS-LowComp approach: we here evaluate the low-complexity algorithm from Section 6.2.4 for the non-RS setup. Since no common stream transmission takes place here, the optimization of the weighting factors is omitted, and only the power allocation among the users is performed.

For all iterative methods, we set the maximum number of iterations to $N_{\text{iter}} = 100$. For the low complexity approaches, we initialize with equal power allocation and identical weights in the RS case.

In Fig.6.3, we can observe the gain of implementing RS in the system in both cases $T_{\text{dl}} = 4$ and $T_{\text{dl}} = 2$, especially in the high-power regime starting from $P_{\text{dl}} = 15$ dB. Unlike the non-RS methods, where the sum rate saturates at high transmit powers (as previously discussed in this chapter), the RS-enabled approaches maintain linear growth as transmit power increases.

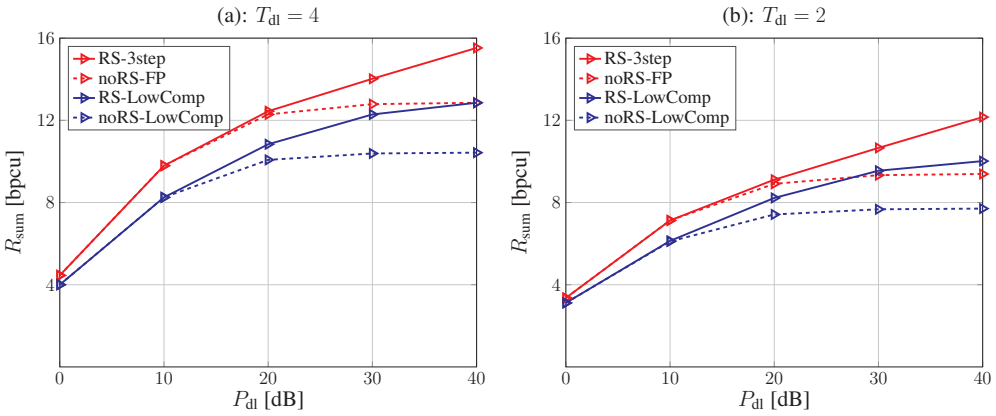

Fig. 6.3: Achievable sum rate versus P_{dl} for $M = 32$, $K = 8$ and different T_{dl} values in the RS setup.

Moreover, the RS-3step approach demonstrates a steeper slope at high SNR, outperforming the RS-LowComp method. This difference is expected due to the reduced degrees of freedom available in RS-LowComp for precoder design.

T_{dl} \ P_{dl}	0 dB	20 dB	40 dB
4 pilots	79.5	37.1	18.5
2 pilots	9.8	6.1	3.2

Table 6.1: Ratio of the average run time of the RS-3step approach relative to the average run time of the RS-LowComp algorithm for $M = 32$, $K = 8$, $T_{\text{dl}} = 4/2$ and various transmit powers

However, the RS-3step method's superior performance comes at the cost of significantly higher computation time, as shown in Table 6.1. The table provides the ratio of the average run time required for the RS-3step approach to converge relative to that of the RS-LowComp algorithm, averaged over the $N_{\text{cov}} = 100$ covariance constellations (since optimization occurs within the covariance coherence interval rather than per channel realization). Notably, the low-complexity approach is at least 18 times faster than the RS-3step approach for $T_{\text{dl}} = 4$ and at least 3 times faster for $T_{\text{dl}} = 2$. Hence, while the RS-3step method achieves higher sum rates, especially in high-SNR regimes, the RS-LowComp approach provides a much more computationally efficient alternative, making it preferable when lower run times are a priority.

7 AWAMSE and MM Precoding

In Chapter 5, we observed that the zero-forcing precoder based on the LMMSE channel estimate suffered from residual interference, although the LMMSE estimator exhibits a better estimation quality than the LS estimator due to exploiting the channel second-order statistics. This shows that naively applying conventional precoding approaches and taking the channel estimate as if it were the true channel can be potentially detrimental to the system performance, irrespective of the channel estimation quality.

In this chapter, we design the precoding vectors such that the training-based lower bound on the achievable rate is maximized. The derived lower bound takes into account the incomplete channel knowledge at the transmitter and is hence aware of channel estimation errors. To this end, we develop two optimization techniques. The first one relates the rate lower bound to an augmented weighted average MSE (AWAMSE) and solves the so-obtained optimization problem in an alternating fashion due to the introduced auxiliary variables. The proposed iterative AWAMSE precoding approach is highly accelerated due to the closed-form updates in each iteration [48]. Its computational efficiency is compared to the adapted stochastic IWMMSE (SIWMMSE) algorithm, where we propose circumventing the solution approach based on interior-point methods by providing an analytical expression for the precoder. The second technique is based on the MM technique presented in Section 4.3, which we here apply to the rate lower bound instead of the instantaneous rate [47]. Again, we provide a computationally efficient optimization procedure due to the analytical expressions of the auxiliary variables and, most importantly, the precoder.

Later in this chapter, we further extend the proposed AWAMSE to the RSMA setup in order to enhance the system performance in the very limited channel knowledge case and show the potential gains of implementing the RS scheme numerically [49].

7.1 Training-based Rate Lower Bound

We aim to solve the following throughput maximization problem based on the training-based lower bound on the SINR derived in Section 3.3.1.2.

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} \bar{R}_k = \max_{\boldsymbol{P}} \sum_{k=1}^{K} \log_2(1 + \overline{\text{SINR}}_k) \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\text{F}}^2 \leq P_{\text{dl}} \tag{7.1}$$

where the SINR of user k is given by (see (3.23))

$$\overline{\text{SINR}}_k = \frac{|\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_k|^2}{\sum_{j=1}^{K} \boldsymbol{p}_j^{\text{H}} \boldsymbol{C}_{\text{err},k} \boldsymbol{p}_j + \sum_{j \neq k} |\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_j|^2 + 1}. \tag{7.2}$$

Recall that we here adopt the error model for the channel $\boldsymbol{h}_k = \hat{\boldsymbol{h}}_k + \tilde{\boldsymbol{h}}_k$, where $\hat{\boldsymbol{h}}_k$ denotes the channel estimate of user k assumed to be known at the transmitter and $\boldsymbol{C}_{\text{err},k}$ stands for the error covariance matrix that can be found based on the channel statistics.

The SINR lower bound defined in (7.2) is based on the assumption that LMMSE channel estimation is performed, and thus, the error covariance matrix is computed according to (3.12).

Due to the dependence of the SINR of user k not only on its own precoding vector $\boldsymbol{p}_k$ but also on all other users' precoders $\boldsymbol{p}_j, j \neq k$, finding the optimal precoder in (7.1) is a non-trivial task. In the following, we present two iterative optimization techniques to efficiently solve (7.1), namely the AWAMSE and MM.

7.2 Augmented Weighted Average MSE

For the AWAMSE method, we formulate an equivalent optimization problem to (7.1) that is easier to handle. This equivalence is based on establishing a connection between the previously stated lower bound on the rate and an AWAMSE, which we will define in the sequel.

First, we introduce receive filters and weighting factors for each of the users that will serve as auxiliary variables, which allow us to unveil hidden convexity properties of the sum rate maximization problem as previously discussed in Section 4.2. The main difference here is that we introduce an average MSE that can be related to the considered SINR lower bound.

We first assume that a receive filter g_k is applied at each user k, which delivers an estimate $\hat{s}_k$ of the data signal s_k

$$\hat{s}_k = g_k r_k = g_k \hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_k s_k + g_k \tilde{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_k s_k + g_k \sum_{j \neq k} \boldsymbol{h}_k^{\text{H}} \boldsymbol{p}_j s_j + g_k n_k. \tag{7.3}$$

The MSE of the data signal s_k is given by $\varepsilon_k = \mathbb{E}\left[|\hat{s}_k - s_k|^2\right]$, where the expectation is taken w.r.t. the distribution of the data and the noise signals.
Using (7.3), the MSE can be evaluated as follows

$$\begin{aligned}\varepsilon_k &= \mathbb{E}\left[|\hat{s}_k - s_k|^2\right] \\ &= |g_k|^2|\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k|^2 + |g_k|^2|\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k|^2 + |g_k|^2\boldsymbol{p}_k^{\mathrm{H}}\tilde{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k + |g_k|^2\boldsymbol{p}_k^{\mathrm{H}}\hat{\boldsymbol{h}}_k\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k \\ &\quad - 2\Re\{g_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} - 2\Re\{g_k\tilde{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} + |g_k|^2\sum_{j\neq k}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2 + |g_k|^2 + 1.\end{aligned}$$

To establish a rate-MSE relationship, we define the average MSE (AMSE) as follows

$$\bar{\varepsilon}_k = \mathbb{E}[\varepsilon_k] \tag{7.4}$$

where the expectation is now taken w.r.t. the channel estimates.
By exploiting the independence of the LMMSE channel estimate $\hat{\boldsymbol{h}}_k$ and the estimation error $\tilde{\boldsymbol{h}}_k$ [7], $\bar{\varepsilon}_k$ is given by

$$\bar{\varepsilon}_k = |g_k|^2\sum_{j=1}^{K}|\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_j|^2 + |g_k|^2\sum_{j=1}^{K}\boldsymbol{p}_j^{\mathrm{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_j - 2\Re\{g_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} + |g_k|^2 + 1 \tag{7.5}$$

where we used the fact that the estimation error is zero-mean and has the covariance matrix $\boldsymbol{C}_{\text{err},k}$.
For fixed precoders, the AMSE minimizing receive filter is obtained by considering the first optimality condition of (7.5), which yields

$$g_k^{\text{MMSE}} = \boldsymbol{p}_k^{\mathrm{H}}\hat{\boldsymbol{h}}_k(D_k + 1)^{-1} \tag{7.6}$$

with

$$D_k = \sum_{j=1}^{K}|\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_j|^2 + \boldsymbol{p}_j^{\mathrm{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_j. \tag{7.7}$$

Inserting (7.6) into (7.5) results in

$$\bar{\varepsilon}_k^{\text{MMSE}} = 1 - |\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k|^2(D_k + 1)^{-1}. \tag{7.8}$$

Note that $\bar{\varepsilon}_k^{\text{MMSE}}$ is related to the rate lower bound $\bar{R}_k$ of user k according to

$$\bar{R}_k = \log_2(1 + \overline{\text{SINR}}_k) = -\log_2(\bar{\varepsilon}_k^{\text{MMSE}}). \tag{7.9}$$

If we additionally introduce a weighting factor u_k associated with the kth user's MSE, we can define the AWAMSE of user k as

$$\begin{aligned}\bar{\xi}_k &= u_k\bar{\varepsilon}_k - \log_2 u_k \\ &= u_k\big(1 - 2\Re\{g_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} + |g_k|^2(1 + \sum_{j=1}^{K}|\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_j|^2 + \boldsymbol{p}_j^{\mathrm{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_j)\big) - \log_2 u_k.\end{aligned} \tag{7.10}$$

Minimizing $\bar{\xi}_k$ w.r.t. the receive filter g_k and the weight u_k for fixed precoders, yields the following relationship to the rate lower bound

$$\bar{\xi}_k^{\text{MMSE}} = 1 - \bar{R}_k. \tag{7.11}$$

Here, the optimal receive filter is the MMSE receive filter given by (7.6) and the optimal weight reads as

$$u_k^{\text{MMSE}} = 1/\bar{\varepsilon}_k^{\text{MMSE}} = (1 - \hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k g_k^{\text{MMSE}})^{-1}. \tag{7.12}$$

In summary, the throughput maximization problem in (7.1) can be equivalently reformulated as the minimization of the AWAMSE due to (7.11)

$$\min_{\boldsymbol{G},\boldsymbol{U},\boldsymbol{P}} \sum_{k=1}^{K}\bar{\xi}_k \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}} \tag{7.13}$$

where we again collected the weights and the receive filters in the matrices $\boldsymbol{U}$ and $\boldsymbol{G}$, respectively

$$\boldsymbol{U} = \text{diag}(u_1, \ldots, u_K), \tag{7.14}$$

$$\boldsymbol{G} = \text{diag}(g_1, \ldots, g_K). \tag{7.15}$$

7.2.1 AWAMSE Algorithm

Due to the introduced sets of auxiliary variables, namely the received filters and the weights, the optimization problem (7.13) consists of three blocks, and therefore, an iterative solution approach is adopted, where each of the three blocks is updated in every iteration while the remaining two are fixed until convergence is reached.

For fixed $\boldsymbol{G}$ and $\boldsymbol{U}$, we obtain the following optimization problem w.r.t. the precoding matrix $\boldsymbol{P}$

$$\begin{aligned}&\min_{\boldsymbol{P}} \sum_{k=1}^{K} 2\Re\{u_k g_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} - u_k|g_k|^2\Big(\sum_{j=1}^{K}\boldsymbol{p}_j^{\mathrm{H}}(\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_{\text{err},k})\boldsymbol{p}_j + 1\Big) \\ &\text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}}.\end{aligned} \tag{7.16}$$

The so-obtained optimization problem exhibits the same form as (4.47), where we have the following substitutions

$$\begin{aligned} b_k &= u_k g_k, \quad a_k = u_k |g_k|^2, \\ \boldsymbol{t}_k &= \hat{\boldsymbol{h}}_k, \qquad \boldsymbol{T}_k = \hat{\boldsymbol{h}}_k \hat{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_{\mathrm{err},k}. \end{aligned} \tag{7.17}$$

Thus, we can follow the rescaling procedure described in Section 4.4 to infer the unconstrained solution for the kth user's precoder (see (4.58))

$$\boldsymbol{p}_k^{\text{unconst}} = \left(\sum_{j=1}^{K} u_j |g_j|^2 (\hat{\boldsymbol{h}}_j \hat{\boldsymbol{h}}_j^{\mathrm{H}} + \boldsymbol{C}_{\mathrm{err},j}) + \left(\frac{1}{P_{\mathrm{dl}}} \sum_{i=1}^{K} u_i |g_i|^2 \right) \boldsymbol{I}_M \right)^{-1} u_k g_k^* \hat{\boldsymbol{h}}_k. \tag{7.18}$$

Due to the introduction of the scaling factor α (see (4.53)), the computation of the receive filter g_k and the weighting factor u_k needs to be adapted as follows

$$g_k^{\text{unconst}} = \boldsymbol{p}_k^{\text{unconst,H}} \hat{\boldsymbol{h}}_k \left(D_k + \alpha^{-2} \right)^{-1} \tag{7.19}$$

$$u_k^{\text{unconst}} = \left(1 - |\hat{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{p}_k^{\text{unconst}}|^2 \left(D_k + \alpha^{-2} \right)^{-1} \right)^{-1} \tag{7.20}$$

with α given by

$$\alpha = \frac{\sqrt{P_{\mathrm{dl}}}}{\| \boldsymbol{P}^{\text{unconst}} \|_{\mathrm{F}}} \tag{7.21}$$

and D_k is similarly defined to (7.7) as a function of $\boldsymbol{P}^{\text{unconst}}$

$$D_k = \sum_{j=1}^{K} \boldsymbol{p}_j^{\text{unconst,H}} (\hat{\boldsymbol{h}}_k \hat{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_{\mathrm{err},k}) \boldsymbol{p}_j^{\text{unconst}}. \tag{7.22}$$

Algorithm 6 AWAMSE Algorithm

1: Initialize $\boldsymbol{P}$ with a feasible solution
2: **while** no convergence is reached **do**
3: Compute the receive filter g_k according to (7.19)
4: Compute the weighting factor u_k according to (7.20)
5: Calculate the unconstrained precoding matrix $\boldsymbol{P}^{\text{unconst}}$ based on (7.18)
6: **end while**
7: Scale the unconstrained precoding matrix with α to satisfy the power constraint (see (7.21))

The overall iterative AWAMSE precoding algorithm to solve (7.13) is summarized in Algorithm 6. Here, the final step ensures that the obtained precoder satisfies the power constraint by scaling with the factor α (see (7.21)).

7.3 MM Approach

As discussed in Section 4.3, the MM approach starts by carefully choosing a surrogate function that defines a lower bound to the underlying objective to be maximized [21]. This surrogate function is easier to handle and is maximized instead of the original function.

To this end, we exploit Proposition 1 to define a surrogate function to the objective formulated in (7.1) at iterate $\overline{\boldsymbol{P}}$. We here substitute x and z by $\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k$ and $\sum_{j=1}^{K}\boldsymbol{p}_j^{\mathrm{H}}\boldsymbol{C}_{\text{err},k}\boldsymbol{p}_j + \sum_{j\neq k}|\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k|^2 + 1$, respectively and obtain

$$f_k(\boldsymbol{P},\overline{\boldsymbol{P}}) = \bar{R}_k(\overline{\boldsymbol{P}}) - \overline{\text{SINR}}_k(\overline{\boldsymbol{P}}) + 2\Re\{b_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} - a_k\Big(\sum_{j=1}^{K}\boldsymbol{p}_j^{\mathrm{H}}(\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_{\text{err},k})\boldsymbol{p}_j + 1\Big) \tag{7.23}$$

where the scalars a_k and b_k are defined as follows

$$a_k = \frac{\overline{\text{SINR}}_k(\overline{\boldsymbol{P}})}{\sum_{j=1}^{K}\bar{\boldsymbol{p}}_j^{\mathrm{H}}(\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_{\text{err},k})\bar{\boldsymbol{p}}_j + 1}, \tag{7.24}$$

$$b_k = \frac{\overline{\text{SINR}}_k(\overline{\boldsymbol{P}})}{\hat{\boldsymbol{h}}_k^{\mathrm{H}}\bar{\boldsymbol{p}}_k}. \tag{7.25}$$

For a given iterate $\overline{\boldsymbol{P}}$ and thus fixed a_k and b_k, we end up with the following optimization problem w.r.t. $\boldsymbol{P}$

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} 2\Re\{b_k\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_k\} - a_k\Big(\sum_{j=1}^{K}\boldsymbol{p}_j^{\mathrm{H}}\boldsymbol{T}_k\boldsymbol{p}_j + 1\Big) \quad \text{s.t.} \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\text{dl}} \tag{7.26}$$

where the matrix $\boldsymbol{T}_k$ is given by

$$\boldsymbol{T}_k = \hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_{\text{err},k}. \tag{7.27}$$

(7.26) is a QCQP and can be solved using the Lagrangian multiplier method. Here, a line search, e.g., bisection, can be performed to find the optimal Lagrangian multiplier in each iteration [21]. Alternatively, by noticing that the optimization problem (7.26) exhibits exactly the same form as (4.47), where $\boldsymbol{t}_k$ is substituted by $\hat{\boldsymbol{h}}_k$, we can apply the rescaling approach from Section 4.4 to infer an analytical expression for the unconstrained precoder $\boldsymbol{p}_k$ (see (4.58))

$$\boldsymbol{p}_k^{\text{unconst}} = \left(\sum_{j=1}^{K} a_j\boldsymbol{T}_j + \Big(\frac{1}{P_{\text{dl}}}\sum_{i=1}^{K} a_i\Big)\boldsymbol{I}_M\right)^{-1} b_k^*\boldsymbol{t}_k. \tag{7.28}$$

The overall procedure of the MM approach is summarized in Alg. 7. The last step consists of rescaling the final solution to meet the power constraint, with the scaling factor α given by

$$\alpha = \frac{\sqrt{P_{\text{dl}}}}{\|\boldsymbol{P}^{\text{unconst}}\|_{\text{F}}}. \tag{7.29}$$

Algorithm 7 MM Algorithm

1: Initialize $\overline{\boldsymbol{P}}$ with a feasible solution
2: **while** no convergence is reached **do**
3: Compute the scalars a_k and b_k according to (7.24) and (7.25), respectively
4: Calculate the unconstrained precoding matrix $\boldsymbol{P}^{\text{unconst}}$ based on (7.28)
5: **end while**
6: Scale the unconstrained precoding matrix with α to satisfy the power constraint (see (7.29))

7.4 Adapted SIWMMSE Approach

Contrary to the proposed AWAMSE and MM methods, which operate on the training-based rate lower bound, the SIWMMSE approach aims at solving the stochastic average rate maximization problem given by

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} \tilde{R}_k \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\text{F}}^2 \leq P_{\text{dl}} \tag{7.30}$$

where $\tilde{R}_k = \mathbb{E}_{\boldsymbol{H}|\hat{\boldsymbol{H}}}[R_k|\hat{\boldsymbol{H}}]$ defines the average rate of user k and is a performance measure over the error distribution for a given channel estimate [73].

The optimization problem formulated in (7.30) is challenging due to its stochastic nature. To make it tractable, first, the average rate is approximated using the sample average approximation (SAA) [74]. Secondly, a connection between the approximated rates and the weighted MSE is derived, revealing the hidden convexity characteristics of the original rate maximization problem.

For the SAA approach, N_{s} channel samples are generated using known channel statistics. Assuming correlated channels (and not i.i.d. channels as in [73]), we suggest to draw the nth channel sample of user k according to [49]

$$\boldsymbol{h}_k^{(n)} = \hat{\boldsymbol{h}}_k + \boldsymbol{C}_{\text{err},k}^{\frac{1}{2}} \tilde{\boldsymbol{e}}_k^{(n)}. \tag{7.31}$$

Here $\hat{\boldsymbol{h}}_k$ is the LMMSE channel estimate known at the base station, $\tilde{\boldsymbol{e}}_k^{(n)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I}_M)$ is the randomly generated error vector, and $\boldsymbol{C}_{\text{err},k}^{\frac{1}{2}}$ denotes the square root matrix of the error covariance matrix $\boldsymbol{C}_{\text{err},k}$ (see (3.12)).

If we denote by $R_k^{(n)}$ the instantaneous rate evaluated for the nth channel sample, i.e.,

$$R_k^{(n)} = R_k^{\text{inst}}(\boldsymbol{H}^{(n)}) = \frac{|\boldsymbol{h}_k^{(n),\text{H}} \boldsymbol{p}_k|^2}{\sum_{j \neq k} |\boldsymbol{h}_k^{(n),\text{H}} \boldsymbol{p}_j|^2 + 1} \tag{7.32}$$

then, the approximated average rate is given by

$$\tilde{R}_k(N) = \frac{1}{N_\text{s}} \sum_{n=1}^{N_\text{s}} R_k^{(n)}. \tag{7.33}$$

We, hence, can formulate the following SAA rate maximization problem

$$\max_{\boldsymbol{P}} \sum_{k=1}^{K} \tilde{R}_k^{(N_\text{s})} \quad \text{s.t.} \quad \|\boldsymbol{P}\|_\text{F}^2 \leq P_\text{dl}. \tag{7.34}$$

To solve (7.34), similarly to Section 4.2, two sets of receive filters $g_k^{(n)}$ and weights $u_k^{(n)}$ are introduced for each channel sample n in order to reformulate the approximated sum rate maximization problem in (7.34) as the minimization of the approximated augmented weighted sum MSE.
For user k, the approximated augmented weighted MSE is given by

$$\tilde{\xi}_k^{(N_\text{s})} = \frac{1}{N_\text{s}} \sum_{n=1}^{N_\text{s}} \xi_k^{(n)} \tag{7.35}$$

where $\xi_k^{(n)}$ is the augmented weighted MSE (see (4.24)) evaluated for the nth channel sample $\boldsymbol{h}_k^{(n)}$ and the nth receive filter and weight, i.e., $g_k^{(n)}$ and $u_k^{(n)}$, respectively.

The overall optimization problem reads as

$$\min_{\boldsymbol{G},\boldsymbol{U},\boldsymbol{P}} \sum_{k=1}^{K} \tilde{\xi}_k^{(N_\text{s})} \quad \text{s.t.} \quad \|\boldsymbol{P}\|_\text{F}^2 \leq P_\text{dl}. \tag{7.36}$$

(7.36) is again solved in an alternating fashion, where for fixed precoding matrix $\boldsymbol{P}$, the receive filters $\boldsymbol{G}$ and $\boldsymbol{U}$ are computed for each of the N_s channel samples, unlike the perfect CSI case. This means $g_k^{(n)}$ and $u_k^{(n)}$ are found according to (4.18) and (4.25), where $\boldsymbol{h}_k$ is substituted by the nth channel sample $\boldsymbol{h}_k^{(n)}$. After the Monte Carlo

run, the precoding matrix is updated using the averages of the computed variables as described in the sequel.

In order to simplify the optimization problem w.r.t. the precoding matrix $\boldsymbol{P}$, we first introduce the following variables for nth channel sample

$$\begin{aligned} t_k^{(n)} &= u_k^{(n)} |g_k^{(n)}|^2, & v_k^{(n)} &= \log_2(u_k^{(n)}), \\ \boldsymbol{f}_k^{(n)} &= u_k^{(n)} g_k^{(n),*} \boldsymbol{h}_k^{(n)}, & \boldsymbol{\Psi}_k^{(n)} &= t_k^{(n)} \boldsymbol{h}_k^{(n)} \boldsymbol{h}_k^{(n),\mathrm{H}}. \end{aligned} \tag{7.37}$$

If we now consider the average of all the variables over the N_s channel samples, then we can formulate the precoder optimization problem as follows

$$\begin{aligned} \min_{\boldsymbol{P}} \; & \sum_{k=1}^{K} \sum_{j=1}^{K} \boldsymbol{p}_j^{\mathrm{H}} \tilde{\boldsymbol{\Psi}}_k \boldsymbol{p}_j + \tilde{t}_k - 2\Re\{\tilde{\boldsymbol{f}}_k^{\mathrm{H}} \boldsymbol{p}_k\} + \tilde{u}_k - \tilde{v}_k \\ \text{s.t.} \; & \sum_{k=1}^{K} \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{dl}} \end{aligned} \tag{7.38}$$

where $\tilde{a} = \frac{1}{N_\mathrm{s}} \sum_{n=1}^{N_\mathrm{s}} a^{(n)}$ denotes the average of some variable a over the N_s channel samples.

The optimization problem (7.38) is convex in the precoding matrix $\boldsymbol{P}$ and can be solved using interior-point methods or the Lagrangian multiplier approach. The latter requires a line search for the optimal Lagrangian multiplier associated with the power constraint. To circumvent this, we again apply the rescaling idea proposed in Section 4.4 to recast (7.38) as an unconstrained optimization problem in the following manner

$$\min_{\boldsymbol{P}} \sum_{k=1}^{K} \sum_{j=1}^{K} \boldsymbol{p}_j^{\mathrm{H}} \tilde{\boldsymbol{\Psi}}_k \boldsymbol{p}_j + \frac{\|\boldsymbol{P}\|_{\mathrm{F}}^2}{P_{\mathrm{dl}}} \tilde{t}_k - 2\Re\{\tilde{\boldsymbol{f}}_k^{\mathrm{H}} \boldsymbol{p}_k\}. \tag{7.39}$$

The precoding matrix can thus be updated in closed form as

$$\boldsymbol{P}^{\text{unconst}} = \left(\sum_{j=1}^{K} \tilde{\boldsymbol{\Psi}}_j + \left(\frac{1}{P_{\mathrm{dl}}} \sum_{j=1}^{K} \tilde{t}_j \right) \boldsymbol{I}_M \right)^{-1} \tilde{\boldsymbol{F}}_k \tag{7.40}$$

where we introduced the matrix $\tilde{\boldsymbol{F}}_k = [\tilde{\boldsymbol{f}}_1, \ldots, \tilde{\boldsymbol{f}}_K]$.
Upon convergence, the final solution has to be rescaled in order to satisfy the power constraint.

The overall procedure of the adapted SIWMMSE is summarized in Algorithm 8.

Algorithm 8 SIWMMSE Algorithm

1: Initialize $\boldsymbol{P}$ with a feasible solution
2: **while** no convergence is reached **do**
3: **for** $n = 1$ to N_{s} **do**
4: Compute $g_k^{(n)}$ and $u_k^{(n)}$ according to (4.18) and (4.25), respectively
5: **end for**
6: Compute the average of the variables in (7.37) over the N_{s} channel samples
7: Calculate the unconstrained precoding matrix $\boldsymbol{P}^{\mathrm{unconst}}$ based on (7.40)
8: **end while**
9: Scale the unconstrained precoding matrix to satisfy the power constraint

7.5 Numerical Results

To assess the potential of the proposed AWAMSE and MM precoding approaches, we again consider a scenario with incomplete channel knowledge at the base station where the LMMSE channel estimate is available. This matches our assumptions for the lower bound derivation in Section 7.1. We generate the channels according to $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_k)$, whereby the covariance matrix $\boldsymbol{C}_k$ is obtained from a random selection of a GMM component (see Section 5.3). Here, we average the results over $N_{\mathrm{cov}} = 100$ random selections of the covariance matrices (corresponding to different user constellations) and $N_{\mathrm{ch}} = 100$ channel realizations per setup. The pilot matrix is given by the sub-DFT matrix, i.e., the first T_{dl} columns of the $M \times M$ DFT matrix.

We first compare the AWAMSE approach to the following methods

- IWMMSE-inst [23] precoder: this refers to the "naively" applied standard IWMMSE algorithm, which operates on the instantaneous sum rate and treats the channel estimate as if it were the true channel.
- MMSE precoder [15]: here the precoder is given by

$$\boldsymbol{P}^{\mathrm{MMSE}} = \alpha \left(\hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^{\mathrm{H}} + \boldsymbol{C}_{\mathrm{err}} + \frac{M}{P_{\mathrm{dl}}} \boldsymbol{I}_M \right)^{-1} \hat{\boldsymbol{H}} \tag{7.41}$$

  where $\boldsymbol{C}_{\mathrm{err}} = \sum_{k=1}^{K} \boldsymbol{C}_{\mathrm{err},k}$ and the scaling factor α is used for power normalization. The MMSE precoder performs a regularization using the error covariance matrices and thus takes CSI errors into account. We use this precoder as initialization for all iterative methods in the following.

We first plot the achievable sum rate versus the downlink transmit power P_{dl} for a scenario with $M = 32$, $K = 8$, and varying numbers of training pilots in Fig.7.1. For

both cases, i.e., $T_{\text{dl}} = 4$ and $T_{\text{dl}} = 8$, the IWMMSE-inst approach shows comparable performance to AWAMSE at low power values, but the gap starts to increase starting from 10 dB. The IWMMSE-inst is unable to mitigate the inter-user interference in the high-power regime since it ignores the CSI errors during the precoder design and thus overestimates the achievable sum rates. This leads to a degradation of the actual achievable sum rate at high transmit powers. Our proposed AWAMSE approach exhibits, on the contrary, a good performance in the medium to high power regime and shows a high-SNR slope of approximately 1.9 and 1.2 for $T_{\text{dl}} = 8$ and $T_{\text{dl}} = 4$, respectively, compared to 2.45 for the perfect CSI case. The latter curve is obtained by running the standard IWMMSE using the actual channel and serves here only as a performance benchmark.
Furthermore, the MMSE precoder achieves a relatively decent performance for $T_{\text{dl}} = 8 = K$ with a high-SNR slope of approximately 1.34. However, it shows a saturation of the achievable sum rate in the high-power regime for $T_{\text{dl}} = 4 < K$ and exhibits a significant gap to the AWAMSE approach. Lastly, the zero-forcing (ZF) precoder yields the worst performance among all methods, as it fails to manage the interference, as discussed in Chapter 5.

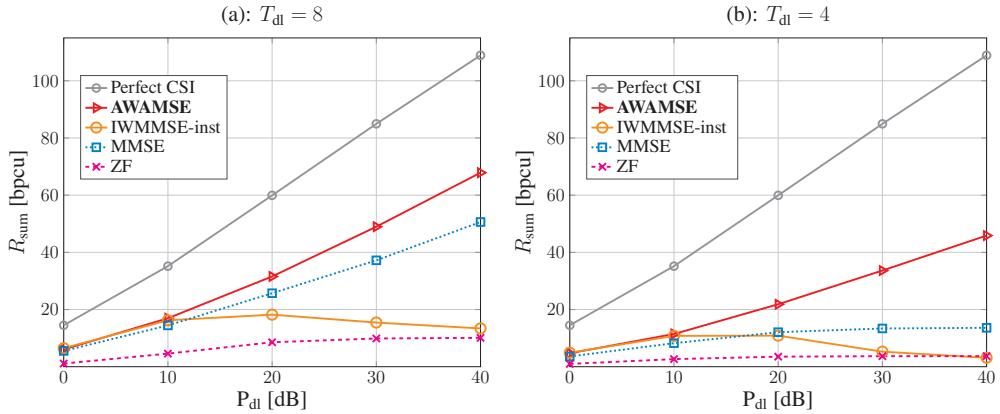

Fig. 7.1: Achievable sum rate versus P_{dl} for $M = 32$, $K = 8$ and different T_{dl} values using various precoding strategies

We now consider the performance of the proposed MM approach based on the rate lower bound, which we denote by MM-LB. This precoder is compared to the AWAMSE approach as well as to its two other counterparts based on the MM procedure, namely

- MM-inst [21] precoder: In this case, the MM approach is applied to the instantaneous rate, as discussed in Section 4.3, where the channel estimate is

used instead of the true channel.

- MMbisec-LB precoder: Here, a line search via bisection is performed before each precoder update in order to infer the optimal Lagrangian multiplier λ^{opt} associated with the power constraint. The precoder is then found according to

$$\boldsymbol{p}_k = \left(\sum_{j=1}^{K} a_j \boldsymbol{T}_j + \lambda^{\text{opt}} \boldsymbol{I}_M \right)^{-1} b_k^* \hat{\boldsymbol{h}}_k. \tag{7.42}$$

In Fig. 7.2(a), we plot the average achievable sum rate for $M = 32$, $K = 8$ and $T_{\text{dl}} = 4$. The MM-LB approach performs almost the same as the AWAMSE precoder. This is due to the connection between both methods discussed in Section 4.3.1. As previously observed with the IWMMSE-inst precoder, the MM-inst approach shows poor performance since it neglects the channel estimation error and overestimates the achievable rate.

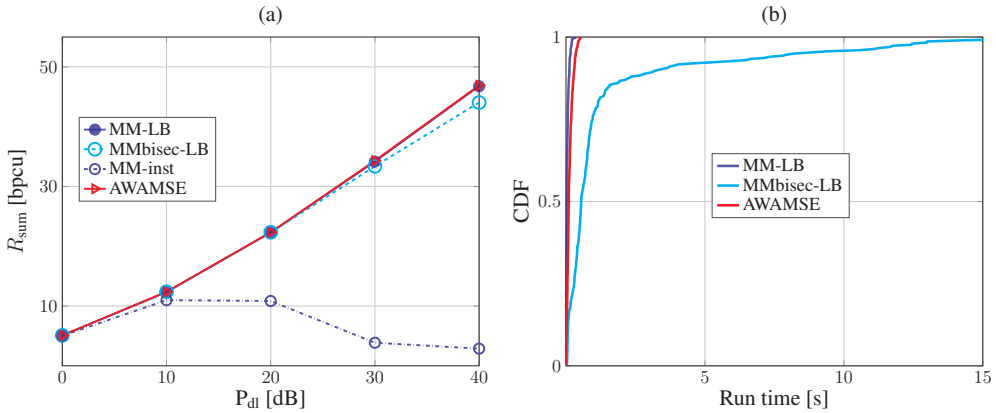

Fig. 7.2: Performance of the MM approach compared to AWAMSE for $M = 32$, $K = 8$ and $T_{\text{dl}} = 4$ in terms of (a) achievable sum rate versus P_{dl} and (b) run time at $P_{\text{dl}} = 30$ dB

For the MMbisec-LB approach and specifically the bisection method, we used the search interval $[0, 10^4]$ and $N_{\text{iter}} = 10^3$ as the maximum number of iterations. One can observe that the MMbisec-LB shows comparable performance to the former approaches at low to medium power values. However, we notice a slight degradation in the high-power regime, which can be explained by the limited accuracy of the bisection method [75], especially when the transmit power increases. Furthermore, the necessary line search via bisection leads to the highest computational complexity

compared to the other approaches. This can be observed in Fig. 7.2(b), where we plot the CDF of the run times (in seconds) of the different precoding methods at a transmit power of $P_{\text{dl}} = 30$ dB. While the run times needed for convergence for the MM-LB and AWAMSE approaches are comparable, they are negligible compared to the MMbisec-LB approach due to the analytical expressions and the circumvention of the line search.

Next, we consider the behavior of the proposed AWAMSE approach to the more computationally demanding SIWMMSE method revised in Section 7.4. For the latter, we use $N_{\text{s}} = 100$ Monte Carlo runs inside each iteration and set the maximum number of iterations to 100. As can be observed in Fig.7.3(a), both algorithms exhibit approximately the same performance for both cases $T_{\text{dl}} = 8$ and $T_{\text{dl}} = 4$, with a slight advantage of the SIWMMSE. Additionally, the power allocation patterns are very similar, as depicted in Fig.7.3(b). Here, we plot the ratio of the power assigned to each user to the total transmit power P_{dl} of 40 dB. Both algorithms serve the same users, namely $1, 3, 4$ and 8, and the number of active users is equal to $T_{\text{dl}} = 4$. This again confirms our conjecture in Chapter 5 that the number of training pilots defines the limit for the degrees of freedom achieved at high SNR.

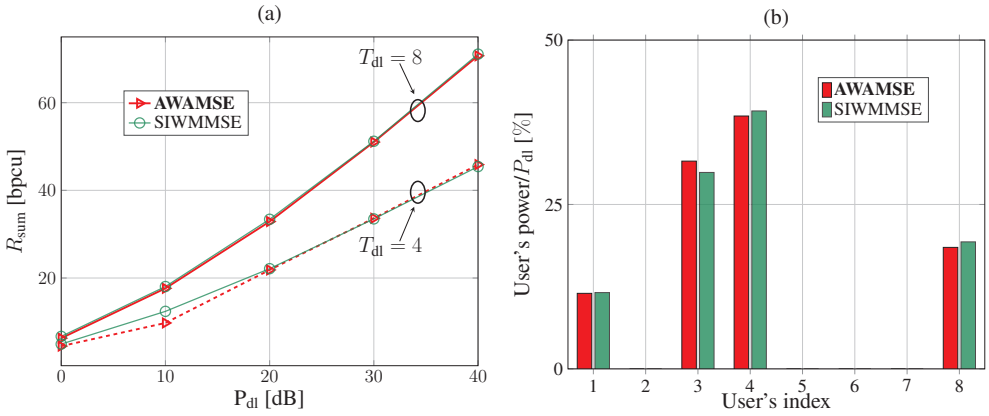

Fig. 7.3: Comparison of the AWAMSE and SIWMMSE algorithms for $M = 32$ and $K = 8$ in terms of (a) achievable sum rate versus P_{dl} using different pilot numbers and (b) power allocation at $P_{\text{dl}} = 20$ dB and $T_{\text{dl}} = 4$ (©IEEE 2024)

Despite the slight advantage of the SIWMMSE compared to AWAMSE in achievable throughput, our proposed approach is more computationally efficient. This is demonstrated by Fig. 7.4, where the empirical CDFs of the number of iterations and run times needed until convergence are plotted. The results are obtained for $T_{\text{dl}} = 4$

and at a transmit power of 20 dB. AWAMSE does not only require fewer iterations than SIWMMSE to reach convergence (see Fig. 7.4(a)), but it also shows a negligibly low run time, as can be observed in Fig. 7.4(b). This is because the SIWMMSE includes a Monte Carlo run inside each and every loop in order to average the variables over the channel samples prior to the precoder update step.

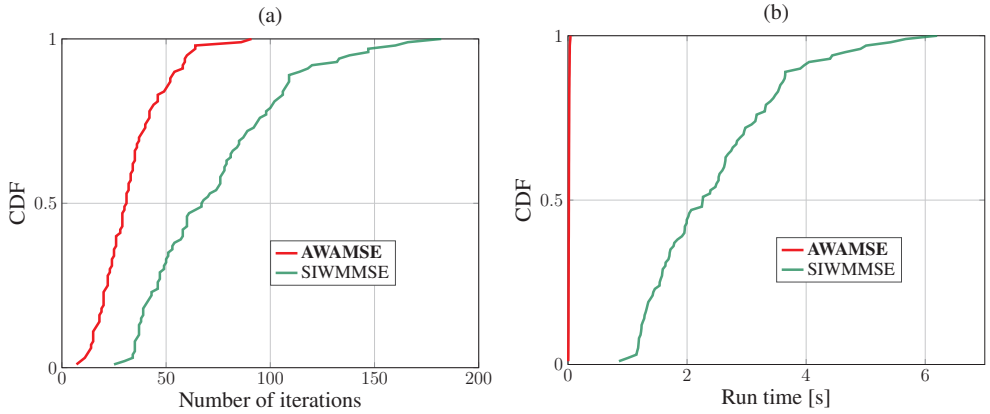

Fig. 7.4: CDF plots of the number of iterations (a) and the run time (b) for $M = 32$, $K = 8, T_{\text{dl}} = 4, P_{\text{dl}} = 20$ dB and the precoding algorithms AWAMSE and SIWMMSE (©IEEE 2024)

7.6 AWAMSE Rate Splitting Precoding Approach

We observed in Section 7.5 that the AWAMSE approach was efficient and competitive with state-of-the-art methods. However, the gap to the perfect CSI case was significant, especially when the number of users exceeded the number of pilots used for channel training. Hence, we propose implementing one-layer RS to boost the system performance in such a challenging scenario, and we formulate the AWAMSE minimization problem for the RS-based system [49].

Assuming incomplete channel knowledge at the base station, we again rely on the training-based SINR lower bound as in Section 7.1. To this end, the received signal at user k (see (3.31)) is decomposed as follows

$$r_{\text{c},k} = \underbrace{\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{c}} s_{\text{c}}}_{\text{desired part}} + \underbrace{\tilde{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{c}} s_{\text{c}} + \sum_{j=1}^{K} (\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{p},j} + \tilde{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{p},j}) s_{\text{p},j} + v_k}_{\text{interference+noise}} \tag{7.43}$$

where $\hat{\boldsymbol{h}}_k$ denotes again the LMMSE channel estimate known at the base station and $\tilde{\boldsymbol{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_{\text{err},k})$ is the channel estimation error.

The common SINR of user k can be lower bounded similarly to Section 3.3.1.2 as follows

$$\overline{\text{SINR}}_{\text{c},k} = \frac{|\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{c}}|^2}{\boldsymbol{p}_{\text{c}}^{\text{H}} \boldsymbol{C}_{\text{err},k} \boldsymbol{p}_{\text{c}} + \sum_{j=1}^{K} \boldsymbol{p}_{\text{p},j}^{\text{H}} \boldsymbol{C}_{\text{err},k} \boldsymbol{p}_{\text{p},j} + |\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{p},j}|^2 + 1}. \tag{7.44}$$

We here again assume that perfect SIC is performed and therefore, the already decoded common part at the receiver can be entirely subtracted from the received signal, and the private SINR of user k can be lower bounded as follows [49]

$$\overline{\text{SINR}}_{\text{p},k} = \frac{|\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{p},k}|^2}{\sum_{j=1}^{K} \boldsymbol{p}_{\text{p},j}^{\text{H}} \boldsymbol{C}_{\text{err},k} \boldsymbol{p}_{\text{p},j} + \sum_{j \neq k} |\hat{\boldsymbol{h}}_k^{\text{H}} \boldsymbol{p}_{\text{p},j}|^2 + 1}. \tag{7.45}$$

Thus, the common and private lower bounds are respectively given by

$$\overline{R}_{\text{c},k} = \log_2(1 + \overline{\text{SINR}}_{\text{c},k}), \quad \overline{R}_{\text{p},k} = \log_2(1 + \overline{\text{SINR}}_{\text{p},k}). \tag{7.46}$$

We aim to maximize the lower bound on the training-based sum rate by designing the common precoder $\boldsymbol{p}_{\text{c}}$ and the private precoding matrix $\boldsymbol{P}_{\text{p}} = [\boldsymbol{p}_{\text{p},1}, \ldots, \boldsymbol{p}_{\text{p},K}]$, i.e.,

$$\max_{\boldsymbol{p}_{\text{c}}, \boldsymbol{P}_{\text{p}}} \min_{k} \overline{R}_{\text{c},k} + \sum_{j=1}^{K} \overline{R}_{\text{p},j} \quad \text{s.t.} \quad \|\boldsymbol{p}_{\text{c}}\|^2 + \|\boldsymbol{P}_{\text{p}}\|_{\text{F}}^2 \leq P_{\text{dl}}. \tag{7.47}$$

Here, the inner minimization ensures the decodability of the common message for all receivers, and the resulting common rate is thus given by the minimum common rate among all users.

The throughput maximization problem in (7.47) is involved, and hence, we will formulate an AWAMSE minimization problem similar to the non-RS setup, which helps to reveal hidden convexity properties of (7.47) and is thus easier to tackle. In a later section, we will present an efficient heuristic solution approach to solve the underlying AWAMSE minimization problem in the RS setup.

7.6.1 Problem Formulation

Similarly to Section 7.1, we aim to establish a relationship between the common and private SINR lower bounds in (7.44) and (7.45) to their respective AMSEs, where the averaging takes place w.r.t. the channel estimates.

By introducing common and private receive filters $g_{\mathrm{c},k}$ and $g_{\mathrm{p},k}$, the common and private AMSEs are respectively given by

$$\bar{\varepsilon}_{\mathrm{c},k} = 1 - 2\Re\{g_{\mathrm{c},k}\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{c}}\} + |g_{\mathrm{c},k}|^2(D_{\mathrm{c},k}+1), \tag{7.48}$$

$$\bar{\varepsilon}_{\mathrm{p},k} = 1 - 2\Re\{g_{\mathrm{p},k}\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{p},k}\} + |g_{\mathrm{p},k}|^2(D_k+1) \tag{7.49}$$

where we again used D_k defined in (7.7) and introduced the variable $D_{\mathrm{c},k}$ given by

$$D_{\mathrm{c},k} = |\hat{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{p}_{\mathrm{c}}|^2 + \boldsymbol{p}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{err},k}\boldsymbol{p}_{\mathrm{c}} + D_k. \tag{7.50}$$

Optimizing the AMSEs w.r.t. the receive filters yields

$$g_{\mathrm{c},k}^{\mathrm{MMSE}} = \boldsymbol{p}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{h}_k(D_{\mathrm{c},k}+1)^{-1}, \tag{7.51}$$

$$g_{\mathrm{p},k}^{\mathrm{MMSE}} = \boldsymbol{p}_{\mathrm{p},k}^{\mathrm{H}}\boldsymbol{h}_k(D_k+1)^{-1}. \tag{7.52}$$

Inserting these solutions into the respective AMSE expressions gives rise to the following relationships between the common and private AMSEs and the rate lower bounds

$$\overline{R}_{\mathrm{c},k} = -\log_2(\bar{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}), \tag{7.53}$$

$$\overline{R}_{\mathrm{p},k} = -\log_2(\bar{\varepsilon}_{\mathrm{p},k}^{\mathrm{MMSE}}). \tag{7.54}$$

At this step, we have already established a connection between the AMSEs and the rates; directly optimizing the AMSEs is, however, still involved due to the log function. We, therefore, need to introduce another two sets of auxiliary variables to externalize the AMSEs from the log function and ease the optimization.
To this end, we consider common and private weights for each user $u_{\mathrm{c},k}$ and $u_{\mathrm{p},k}$ such that a common and a private augmented weighted AMSE can be defined in the following manner

$$\overline{\xi}_{\mathrm{c},k} = u_{\mathrm{c},k}\bar{\varepsilon}_{\mathrm{c},k} - \log_2(u_{\mathrm{c},k}), \tag{7.55}$$

$$\overline{\xi}_{\mathrm{p},k} = u_{\mathrm{p},k}\bar{\varepsilon}_{\mathrm{p},k} - \log_2(u_{\mathrm{p},k}). \tag{7.56}$$

Using the optimal common and private weights given by

$$u_{\mathrm{c},k}^{\mathrm{MMSE}} = \frac{1}{\bar{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}}, \quad u_{\mathrm{p},k}^{\mathrm{MMSE}} = \frac{1}{\bar{\varepsilon}_{\mathrm{p},k}^{\mathrm{MMSE}}} \tag{7.57}$$

we obtain the following AWAMSE-rate relationships

$$\overline{\xi}_{\mathrm{c},k}^{\mathrm{MMSE}} = 1 - \overline{R}_{\mathrm{c},k}, \quad \overline{\xi}_{\mathrm{p},k}^{\mathrm{MMSE}} = 1 - \overline{R}_{\mathrm{p},k}. \tag{7.58}$$

Therefore, the sum rate maximization problem in (7.47) can be recast as an AWAMSE minimization problem

$$\min_{\boldsymbol{G},\boldsymbol{U},\boldsymbol{P}} \max_{k} \overline{\xi}_{\mathrm{c},k} + \sum_{j=1}^{K} \overline{\xi}_{\mathrm{p},j} \quad \text{s.t.} \quad \|\boldsymbol{p}_{\mathrm{c}}\|^2 + \|\boldsymbol{P}_{\mathrm{p}}\|_{\mathrm{F}}^2 \leq P_{\mathrm{dl}} \tag{7.59}$$

where we defined the overall precoding matrix

$$\boldsymbol{P} = [\boldsymbol{p}_{\mathrm{c}}, \boldsymbol{P}_{\mathrm{p}}] \tag{7.60}$$

and the collection of the common and private receive filters and weights as follows

$$\begin{aligned} \boldsymbol{G} &= [[g_{\mathrm{c},1}, \ldots, g_{\mathrm{c},K}]^{\mathrm{T}}, [g_{\mathrm{p},1}, \ldots, g_{\mathrm{p},K}]^{\mathrm{T}}], \\ \boldsymbol{U} &= [[g_{\mathrm{c},1}, \ldots, u_{\mathrm{c},K}]^{\mathrm{T}}, [u_{\mathrm{p},1}, \ldots, u_{\mathrm{p},K}]^{\mathrm{T}}]. \end{aligned} \tag{7.61}$$

Again, (7.59) is a QCQP w.r.t. the precoders when the receive filters and weights are fixed. Instead of using interior-point solvers to find the optimal precoders, we propose in the following section a computationally efficient heuristic approach where the beamformer update is given in an analytical form.

7.6.2 AWAMSE Rate Splitting Precoder Design

In Section 7.2.1, we made use of the rescaling approach in order to infer a closed-form precoder solution for the non-RS setup. A similar derivation can be conducted to adapt this approach to the underlying RS case. Here, the scaling factor α is given by

$$\alpha = \frac{\sqrt{P_{\mathrm{dl}}}}{\|\boldsymbol{P}^{\mathrm{unconst}}\|_{\mathrm{F}}} \tag{7.62}$$

where $\boldsymbol{P}^{\mathrm{unconst}}$ denotes the final unconstrained solution of the overall precoding matrix.

The optimization problem (7.59) can be reformulated as an unconstrained one since the choice of α in (7.62) ensures the fulfillment of the power constraint

$$\begin{aligned} \min_{\boldsymbol{G},\boldsymbol{U},\boldsymbol{P}} \max_{k} u_{\mathrm{c},k} \left[1 - 2\Re\{g_{\mathrm{c},k} \hat{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}\} + |g_{\mathrm{c},k}|^2 (\alpha^{-2} + D_{\mathrm{c},k}) \right] - \log_2(u_{\mathrm{c},k}) \\ \sum_{j=1}^{K} u_{\mathrm{p},j} \left[1 - 2\Re\{g_{\mathrm{p},j} \hat{\boldsymbol{h}}_j^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},j}\} + |g_{\mathrm{p},j}|^2 (\alpha^{-2} + D_j) \right] - \log_2(u_{\mathrm{p},j}). \end{aligned} \tag{7.63}$$

Even for fixed $\boldsymbol{G}$ and $\boldsymbol{U}$, (7.63) is still involved due to the inner multicast problem and can only be solved using interior-point methods, which can be computationally

inefficient.
Hence, we propose a more efficient heuristic approach. To this end, we focus on the precoder optimization since this is the most challenging step. The receive filters and weights can be computed in closed form, as stated later in this section.

For given receive filters $\boldsymbol{G}$ and weights $\boldsymbol{U}$, let us assume that the user index for the common AWAMSE is fixed to some k_c and thus the inner maximization over the users can be dropped. The resulting optimization problem reads as

$$\begin{aligned}\min_{\boldsymbol{P}}\ & u_{\mathrm{c},k_\mathrm{c}}\left[1-2\Re\{g_{\mathrm{c},k_\mathrm{c}}\hat{\boldsymbol{h}}_{k_\mathrm{c}}^{\mathrm{H}}\boldsymbol{p}_\mathrm{c}\}+|g_{\mathrm{c},k_\mathrm{c}}|^2(\alpha^{-2}+D_{\mathrm{c},k_\mathrm{c}})\right]-\log_2(u_{\mathrm{c},k_\mathrm{c}})\\ & \sum_{j=1}^{K}u_{\mathrm{p},j}\left[1-2\Re\{g_{\mathrm{p},j}\hat{\boldsymbol{h}}_j^{\mathrm{H}}\boldsymbol{p}_{\mathrm{p},j}\}+|g_{\mathrm{p},j}|^2(\alpha^{-2}+D_j)\right]-\log_2(u_{\mathrm{p},j}).\end{aligned} \tag{7.64}$$

Now, every choice of $k_\mathrm{c}\in\{1,\ldots,K\}$ yields an analytical expression for the common and private precoders. The corresponding solutions are computed using the first optimality condition of (7.64).
For a given user index k_c, we obtain

$$\boldsymbol{p}_\mathrm{c}(k_\mathrm{c})=(\boldsymbol{A}+\boldsymbol{B})^{-1}u_{\mathrm{c},k_\mathrm{c}}g_{\mathrm{c},k_\mathrm{c}}^{*}\hat{\boldsymbol{h}}_{k_\mathrm{c}}, \tag{7.65}$$

$$\boldsymbol{p}_{\mathrm{p},j}(k_\mathrm{c})=(\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C})^{-1}u_{\mathrm{p},j}g_{\mathrm{p},j}^{*}\hat{\boldsymbol{h}}_j \tag{7.66}$$

where we introduced the following matrices

$$\boldsymbol{A}=u_{\mathrm{c},k_\mathrm{c}}|g_{\mathrm{c},k_\mathrm{c}}|^2(\hat{\boldsymbol{h}}_{k_\mathrm{c}}\hat{\boldsymbol{h}}_{k_\mathrm{c}}^{\mathrm{H}}+\boldsymbol{C}_{\mathrm{err},k_\mathrm{c}}+\frac{1}{P_\mathrm{dl}}\boldsymbol{I}_M), \tag{7.67}$$

$$\boldsymbol{B}=\sum_{i=1}^{K}u_{\mathrm{p},i}|g_{\mathrm{p},i}|^2\frac{1}{P_\mathrm{dl}}\boldsymbol{I}_M, \tag{7.68}$$

$$\boldsymbol{C}=\sum_{i=1}^{K}u_{\mathrm{p},i}|g_{\mathrm{p},i}|^2(\hat{\boldsymbol{h}}_i\hat{\boldsymbol{h}}_i^{\mathrm{H}}+\boldsymbol{C}_{\mathrm{err},i}). \tag{7.69}$$

The overall objective can be evaluated for every possible choice of $k_\mathrm{c}\in\{1,\ldots,K\}$ according to

$$\overline{\xi}(k_\mathrm{c})=\sum_{j=1}^{K}\overline{\xi}_{\mathrm{p},j}(\boldsymbol{P}(k_\mathrm{c}))+\max_{k}\overline{\xi}_{\mathrm{c},k}(\boldsymbol{P}(k_\mathrm{c})). \tag{7.70}$$

Suppose that updating the precoding matrix for every possible user index does not improve the objective compared to the previous iteration; then, our algorithm stops, and we adopt the precoder from the previous iteration as the final solution. Otherwise,

we select the user index $k_{\min}$, which yields the lowest sum AWAMSE among all possible user indices, i.e.,

$$k_{\min} = \arg\min_{k_{\mathrm{c}}} \overline{\xi}(k_{\mathrm{c}}) \tag{7.71}$$

and the precoder is updated as $\boldsymbol{P}(k_{\min})$.

Algorithm 9 AWAMSE-RS Algorithm

Initialize $\boldsymbol{P}^{\text{unconst}} \leftarrow \boldsymbol{P}^{(0)}$, $\overline{\xi}_{\min} \leftarrow \infty$ and $n \leftarrow 1$
while No convergence is reached **do**
  Update $g_{\mathrm{c},k}$ and $g_{\mathrm{p},k}$ according to (7.72) and (7.73)
  Compute $u_{\mathrm{c},k}$ and $u_{\mathrm{p},k}$ according to (7.74) and (7.75)
  for $k_{\mathrm{c}} = 1$ to K **do**
    Determine $\boldsymbol{P}(k_{\mathrm{c}})$ using (7.65) and (7.66)
    Compute $\overline{\xi}(k_{\mathrm{c}}) \leftarrow \sum_j \overline{\xi}_{\mathrm{p},j} + \max_k \overline{\xi}_{\mathrm{c},k}$
  end for
  if $\min_{k_{\mathrm{c}}} \overline{\xi}(k_{\mathrm{c}}) > \overline{\xi}_{\min}$ **then**
    $\boldsymbol{P}^{\text{unconst}} \leftarrow \boldsymbol{P}^{(n-1)}$
    break
  else
    $k_{\min} \leftarrow \arg\min_{k_{\mathrm{c}}} \overline{\xi}(k_{\mathrm{c}})$
    $\overline{\xi}_{\min} \leftarrow \overline{\xi}(k_{\min})$
    $\boldsymbol{P}^{(n)} \leftarrow \boldsymbol{P}(k_{\min})$
    $\boldsymbol{P}^{\text{unconst}} \leftarrow \boldsymbol{P}^{(n)}$
  end if
  $n \leftarrow n + 1$
end while
Scale $\boldsymbol{P}^{\text{unconst}}$ with α to satisfy the power constraint

Now, for fixed unconstrained precoder $\boldsymbol{P}^{\text{unconst}}$, and due to the reformulation in (7.63), the receive filters and weights are computed according to

$$g_{\mathrm{c},k} = \boldsymbol{p}_{\mathrm{c}}^{\text{unconst,H}} \hat{\boldsymbol{h}}_k (D_{\mathrm{c},k} + \alpha^{-2})^{-1}, \tag{7.72}$$

$$g_{\mathrm{p},k} = \boldsymbol{p}_{\mathrm{p},k}^{\text{unconst,H}} \hat{\boldsymbol{h}}_k (D_k + \alpha^{-2})^{-1}, \tag{7.73}$$

$$u_{\mathrm{c},k} = (1 - |\hat{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}^{\text{unconst}}|^2 (D_{\mathrm{c},k} + \alpha^{-2})^{-1})^{-1}, \tag{7.74}$$

$$u_{\mathrm{p},k} = (1 - |\hat{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},k}^{\text{unconst}}|^2 (D_k + \alpha^{-2})^{-1})^{-1}. \tag{7.75}$$

Our procedure can be summarized as follows. We start by initializing the precoding matrix with a feasible solution. Then, the receive filters and weights are updated according to (7.72)-(7.75). Subsequently, the search for the user index yielding the smallest sum AWAMSE is performed. These steps are repeated until the absolute change of the objective attains a certain threshold or the maximum number of iterations is reached.
We shall emphasize that we here ensure the decodability of the common message in each update of the precoder by considering the overall objective based on the maximum common AWAMSE among all users. After convergence, the obtained precoding matrix must be scaled using α (see (7.62)).

The proposed AWAMSE-RS precoder optimization procedure is presented in Algorithm 9. Note that our algorithm is versatile and can also be applied to alternative lower bounds, e.g., in our work [76], we propose to combine the statistical precoder from Chapter 6 with the presented heuristic approach.

7.6.3 Numerical Results

In this section, we demonstrate the advantage of implementing one-layer RS by comparing the proposed AWAMSE-RS precoding approach to its counterpart without RS, as presented in Section 7.2.1 and to the MMSE precoder from [15] (see (7.41)). We additionally use as performance benchmarks the following two alternating optimization methods for the RS setup

1. SIWMMSE-RS [73]: as discussed for the non-RS case in Section 7.4, the authors suggest approximating the ergodic rates using the SAA method. The ergodic rate maximization problem is then reformulated as an augmented weighted MSE minimization problem that is solved in an alternating fashion. While the receive filters and weights are given in closed form, finding an analytical expression for the precoder in the RS setup is difficult, and interior-point methods are used to obtain the solution. We here deploy CVX [70] as an optimization toolbox for the simulations [73].

2. SDR-RS [34]: This approach directly solves the sum rate lower bound maximization problem formulated in (7.47) via an alternating procedure. To this end, a set of auxiliary variables is introduced, and then semi-definite relaxation (SDR) [77] and concave-convex-procedure (CCCP) [29] are applied in order to obtain a solution for the precoder and the introduced auxiliary variables. Furthermore, and because of applying SDR, the optimal precoder leading to the best sum rate is found via random vector search.

Our AWAMSE-RS approach and the SIWMMSE-RS method are initialized as follows. We set the initial common precoder to $\sqrt{\alpha_{\mathrm{c}} P_{\mathrm{dl}}}\boldsymbol{u}$, where $\boldsymbol{u}$ denotes the left singular vector of $\hat{\boldsymbol{H}}$ which corresponds to its largest singular value. The $\alpha_{\mathrm{c}} \in [0, 1]$ is the power fraction we allocate for the common stream transmission and is simply set to 0.5. We initialize the private precoding matrix using the normalized MMSE precoder, i.e.,

$$\boldsymbol{P}_{\mathrm{p}} = \sqrt{(1-\alpha_{\mathrm{c}})P_{\mathrm{dl}}}\frac{\boldsymbol{P}^{\mathrm{MMSE}}}{\|\boldsymbol{P}^{\mathrm{MMSE}}\|_{\mathrm{F}}}. \tag{7.76}$$

For the SDR-RS method, the auxiliary variables are randomly initialized as suggested by the authors, and $N_{\mathrm{rd}} = 1000$ random vectors are used for the randomization step. For the SIWMMSE-RS method, $N_{\mathrm{s}} = 100$ samples are used for the Monte Carlo runs during optimization.

The performance of the different algorithms is evaluated for a system with $M = 16$ and $K = 5$. We focus on the more challenging case where the number of users exceeds the number of training pilots. The results are averaged using 100 random channel realizations, where we again generate the kth user channel according to $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_k)$. Here, $\boldsymbol{C}_k$ corresponds to a randomly selected GMM component (see Section 5.3).

We first assess the convergence behavior of the considered approaches. In Table 7.1, we display the average run times needed until convergence for different transmit power values. Due to interior-point methods, the SIWMMSE-RS and SDR-RS methods exhibit remarkably higher run times than the proposed AWAMSE-RS approach by a factor of at least 100 at all power levels. This is because our algorithm relies on analytical expressions and is, hence, computationally more efficient despite the required user search step.

Method \ P_{dl}	0 dB	20 dB	40 dB
SIWMMSE-RS	21.9s	111.6s	96.6s
SDR-RS	62.9s	82.23s	116.1s
AWAMSE-RS	**0.033**s	**0.094**s	**0.143**s

Table 7.1: Average run time in seconds (s) of the RS algorithms for $M = 16$, $K = 5$, $T_{\mathrm{dl}} = 3$ and different transmit powers (©IEEE 2024)

Additionally, we empirically evaluate the CDFs for the number of iterations and the run times needed until convergence of the considered iterative approaches at a transmit power of $P_{\mathrm{dl}} = 20$ dB. The corresponding results are displayed in Fig. 7.5.

While the SDR-RS approach needs the least number of iterations to reach convergence, as can be observed from Fig. 7.5(a), AWAMSE-RS exhibits a negligibly small run time compared to the other two methods (see Fig. 7.5(b)). This highlights once more the efficiency of our proposed approach in terms of computational complexity.

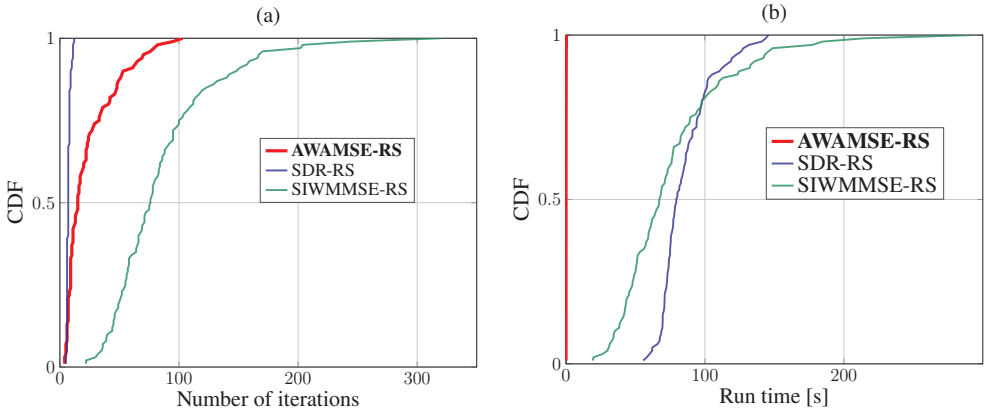

Fig. 7.5: CDF plots of the number of iterations (a) and the run time (b) needed until convergence at $P_{\text{dl}} = 20$ dB for the different RS algorithms (©IEEE 2024)

Next, we evaluate the achievable system throughput for the different precoding approaches. In Fig.7.6(a), we show the results for $T_{\text{dl}} = 3$ training pilots. All RS-based approaches exhibit almost the same performance for the entire power range and show a high-SNR slope of approximately 0.97. The gap to the non-RS methods increases drastically when the transmit power grows. The MMSE precoder almost saturates in this regime and is outperformed by the AWAMSE, having a high-SNR slope of about 0.47. These results demonstrate the ability of RS to boost the system performance when the base station has highly limited channel knowledge, especially in the high-power regime.

If we now consider a more challenging scenario with only $T_{\text{dl}} = 2$ pilots, all methods experience a performance degradation while the RS methods still deliver approximately the same throughput, as can be seen in Fig.7.6(b). Here, the benefits of implementing RS become increasingly apparent compared to the previous scenario. At a transmit power of 20 dB, the RS techniques outperform AWAMSE by 5.8 bpcu. This performance gap widens significantly as the power increases, reaching a substantial gap of about 17 bpcu at $P_{\text{dl}} = 40$ dB.

Finally, the power allocation obtained with the RS approaches at $P_{\text{dl}} = 40$ dB is displayed in Fig. 7.7. We here denote by "Com" and "UEk" the common stream and

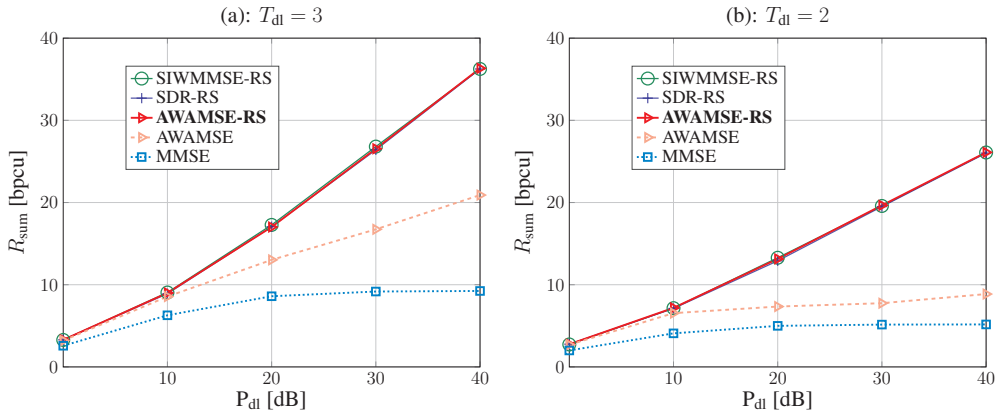

Fig. 7.6: Achievable sum rate versus P_{dl} for $M = 16$ and $K = 5$ using different RS precoding strategies (©IEEE 2024)

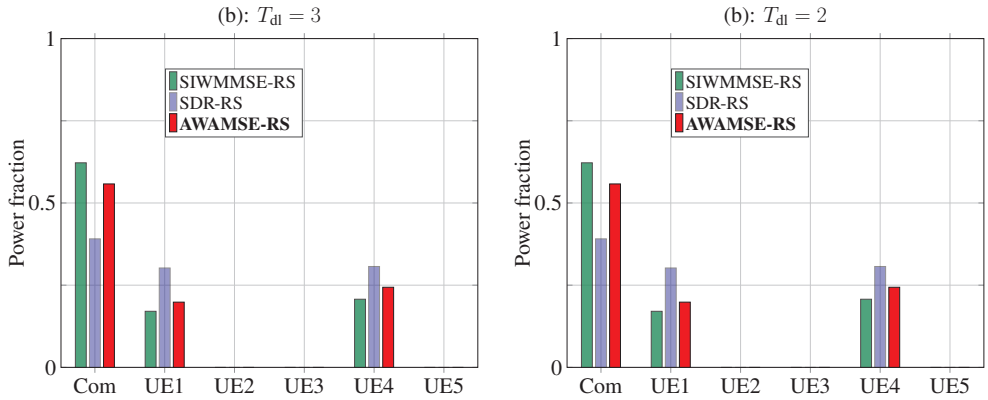

Fig. 7.7: Power allocation pattern for $P_{\text{dl}} = 40$ dB using different RS-based precoding strategies (©IEEE 2024)

the kth user, respectively. All three methods exhibit the same power allocation pattern for both $T_{\text{dl}} = 3$ (Fig. 7.7(a)) and $T_{\text{dl}} = 2$ (Fig. 7.7(b)). This confirms the comparable behavior in terms of the achievable sum rate discussed earlier. Additionally, apart from the common stream, we can observe that only T_{dl} users are active. This result was observed earlier in this chapter in the non-RS setup. It holds also for the RS case that the maximum number of active users is upper bounded by T_{dl}.

8 Joint Pilot and Precoder Design

So far, we have considered precoder design as a strategy to maximize system throughput, treating the training pilots as fixed, such as the truncated DFT matrix or orthogonal Gaussian pilots. However, as demonstrated in Chapter 5, the choice of the pilot matrix is crucial and can significantly affect the system performance. For instance, designing the pilots based on channel covariance matrices can enhance achievable throughput (see Section 5.3). This approach is intuitively appealing since it accounts for the underlying channel structure, but it does not necessarily yield the best possible performance.

Joint design of pilots and precoders in a model-based manner is a challenging task. Hence, we propose adopting a data-driven approach that delivers not only the sum-rate-maximizing precoder but also the pilot matrix. We will demonstrate that the resulting pilots are remarkably beneficial for previously presented precoding strategies, such as the zero-forcing and IWMMSE precoding methods based on the LS channel estimate [50].

8.1 End-to-End Precoding and Pilot Design

Our goal is to maximize the system throughput subject to the transmit power constraint by designing not only the precoders but also the pilot matrix. The underlying optimization problem thus reads as

$$\max_{\boldsymbol{P},\boldsymbol{\Phi}} \sum_{k=1}^{K} R_k \quad \text{s.t.} \quad \|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{dl}} \tag{8.1}$$

where R_k denotes the instantaneous achievable rate of user k (see (3.19)).

We propose an end-to-end (E2E) learning-based approach using an Edge-Graph-Attention-Network (Edge-GAT). In Fig. 8.1, we show the block diagram for the proposed approach. Note that this pipeline corresponds to the offline training phase of the model. In the inference phase, we assume that the base station has incomplete channel knowledge, and hence, only the channel observations $\boldsymbol{y}_k$ (see (3.2)) are used as input to the network.

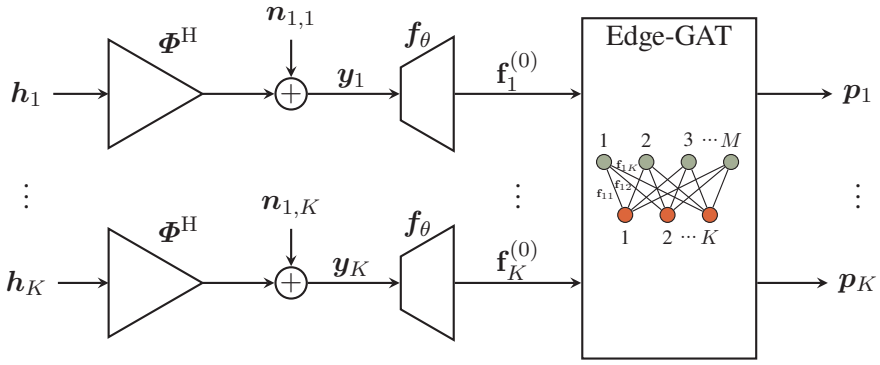

Fig. 8.1: Block diagram of the E2E Edge-GAT model (©IEEE 2024))

The pilot matrix $\boldsymbol{\Phi}$ is modeled as a complex fully connected layer without bias [78], with a normalization layer to ensure that the columns of $\boldsymbol{\Phi}$ are unit-norm. Prior to the Edge-GAT, we introduce a feature extraction layer $\boldsymbol{f}_\theta$, which is the same for all users to maintain scalability. Here, θ denotes the learnable parameters, and $\boldsymbol{f}_\theta$ is modeled as a multi-layer perceptron, which maps the observation vector $\boldsymbol{y}_k$ of user k to the feature vector $\mathbf{f}_k$. The latter is the input to the Edge-GAT network, whose output is the precoding vectors.

As depicted inside the Edge-GAT block in Fig. 8.1, the considered multi-user MISO system is modeled as a bipartite graph, where the M base station antennas and the K single-antenna users are each represented by a set of vertices. Furthermore, we initially associate each edge (m, k) connecting antenna m to user k with the feature vector $\mathbf{f}_{m,k}^{(0)}$. The Edge-GAT network consists of L layers over which the precoding matrix elements are learned. We assume that the initial feature vector $\mathbf{f}_{m,k}^{(0)}$ corresponds to the channel coefficient between antenna m and user k.

At layer ℓ of the Edge-GAT network, we define the update equation of the feature vector for edge (m, k) as follows

$$\mathbf{f}_{m,k}^{(\ell)} = \sigma\left(\boldsymbol{S}^{(\ell)}\mathbf{f}_{m,k}^{(\ell-1)} + \alpha \sum_{\substack{i=1 \\ i\neq m}}^{M} \boldsymbol{T}^{(\ell)}\mathbf{f}_{i,k}^{(\ell-1)} + \beta \sum_{\substack{j=1 \\ j\neq k}}^{K} \boldsymbol{a}_{k,j}^{(\ell-1)} \odot \boldsymbol{v}_{m,j}^{(\ell-1)}\right). \tag{8.2}$$

Here, the parameters in $\boldsymbol{v}_{m,k}^{(\ell)}$ depend on the feature vector $\mathbf{f}_{m,k}^{(\ell)}$ according to

$$\boldsymbol{v}_{m,k}^{(\ell)} = \boldsymbol{V}^{(\ell)}\mathbf{f}_{m,k}^{(\ell)}. \tag{8.3}$$

The weights $\boldsymbol{a}_{k,j}^{(\ell)}$ in (8.2) are found using linear attention [79] (and hence the name of

the network) as follows

$$\boldsymbol{a}_{k,j}^{(\ell)} = \sum_{m=1}^{M} \boldsymbol{q}_{m,k}^{(\ell)} \odot \boldsymbol{k}_{m,j}^{(\ell)} / M \tag{8.4}$$

where the queries and keys are respectively obtained as a function of the feature vector $\mathbf{f}_{m,k}^{(\ell)}$ as

$$\boldsymbol{q}_{m,k}^{(\ell)} = \boldsymbol{Q}^{(\ell)} \mathbf{f}_{m,k}^{(\ell)}, \quad \boldsymbol{k}_{m,k}^{(\ell)} = \boldsymbol{K}^{(\ell)} \mathbf{f}_{m,k}^{(\ell)}. \tag{8.5}$$

The α and β are scalar coefficients, and σ denotes the activation function, e.g., rectified linear unit (ReLu) [80]. The matrices $\boldsymbol{S}^{(\ell)}$, $\boldsymbol{T}^{(\ell)}$, $\boldsymbol{V}^{(\ell)}$, $\boldsymbol{Q}^{(\ell)}$ and $\boldsymbol{K}^{(\ell)}$ are the learnable parameters of layer ℓ. Note that these matrices are independent of the system dimensions, i.e., the number of base station antennas M and the number of users K, making the proposed precoding approach scalable. This means the network trained for a specific configuration M and K can be used for a different configuration with $M' \neq M$ and $K' \neq K$.
After the final layer L of the Edge-GAT, we introduce a normalization layer in order to satisfy the power constraint.

We train the model depicted in Fig. 8.1 end-to-end, with the sum rate being the negative loss. After the training phase, the learned pilot matrix $\boldsymbol{\Phi}$ is used for channel sounding. Note that the pilot matrix is learned jointly with the precoders to maximize the system throughput, which aligns with the overall goal.
In the online phase, we provide the trained network with the observation vectors $\boldsymbol{y}_k$ of each user k and obtain the precoding vector $\boldsymbol{p}_k$.

8.2 Numerical Results

To assess the performance of the proposed precoding approach and especially of the learned pilot matrix, we consider an urban micro-cell scenario where the channels are generated using QuaDRiGa [62]. Here, the base station is equipped with $M = 64$ antennas and serves $K = 20$ single-antenna users. Furthermore, we again assume incomplete channel knowledge, i.e., $T_{\text{dl}} < M$. All other simulation parameters are collected in Table 8.1.

We compare the end-to-end Edge-GAT precoding approach denoted by E2E Edge-GAT to the zero-forcing (ZF) precoder and to the standard IWMMSE precoding approach presented in Section 4.2. The latter two are fed with the LS channel estimate as input (see (3.7)). Additionally, we evaluate the performance of these approaches when the learned pilot matrix is used instead of the sub-DFT matrix.

Parameter	**Value**
Cell radius	150 m
Base station height	10 m
Number of training samples	48K
Number of validation samples	6K
Number of testing samples	5K-30K
Batch size	100
Learning rate	0.001
Optimizer	Adam
Number of hidden layers	10
Number of hidden neurons	64
α	0.1/M
β	0.1

Table 8.1: Simulation parameters for the E2E Edge-GAT model

In Fig. 8.2(a) and (b), we show the achievable sum rate of the different precoding methods versus the downlink transmit power P_{dl} for $T_{\text{dl}} = 16$ and $T_{\text{dl}} = 8$, respectively. In Fig. 8.2(a), the IWMMSE and the ZF approach exhibit a performance boost when the learned pilot matrix is used during channel probing compared to the sub-DFT case. This is because the learned pilot matrix is obtained by jointly optimizing the precoders such that the sum rate is maximized. Specifically, IWMMSE with the learned $\boldsymbol{\Phi}$ achieves a gain of almost 8 dB in the high-SNR region while maintaining the same high-SNR slope as its counterpart with the sub-DFT matrix. Interestingly, the ZF precoder shows a slight saturation of the sum rate with the learned $\boldsymbol{\Phi}$ in the high-SNR region. This behavior can be explained by the fact that there was no structure imposed on the pilot matrix during training, and the latter could potentially be non-orthogonal, leading to this loss in the degrees of freedom. Furthermore, the E2E Edge-GAT method achieves a performance comparable to the IWMMSE approach, which uses the learned pilot matrix.

Considering now a more challenging scenario with $T_{\text{dl}} = 8 < K$ as shown in Fig. 8.2, we observe that all methods experience a performance degradation compared to the previous case with $T_{\text{dl}} = 16$. Additionally, the performance gap between the methods based on the sub-DFT pilot matrix and the learned one diminishes. Lastly, the E2E Edge-GAT approach still competes with the IWMMSE approach but shows a slightly larger gap in the high-power regime.

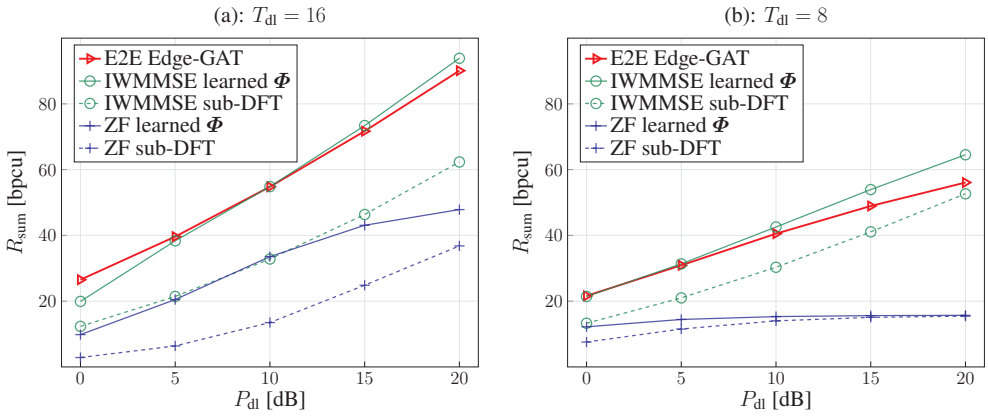

Fig. 8.2: Achievable sum rate versus P_{dl} using different precoding strategies (©IEEE 2024)

Finally, we consider the power allocation of the E2E Edge-GAT and the IWMMSE approach, which uses the learned $\boldsymbol{\Phi}$ for the case $T_{\text{dl}} = 8 < K$ and $P_{\text{dl}} = 20$ dB.

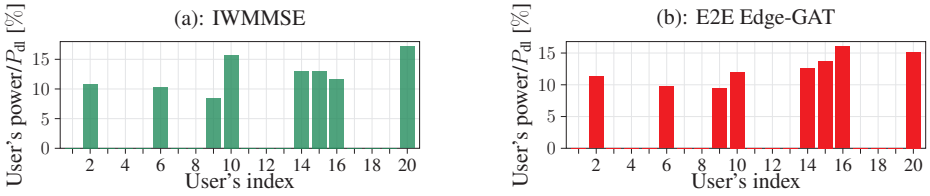

Fig. 8.3: Power allocation of the IWMMSE and the E2E Edge-GAT for $P_{\text{dl}} = 20$ dB and $T_{\text{dl}} = 8$ (©IEEE 2024)

We have observed in the previous chapters that the degrees of freedom or the number of active users is upper-bounded by T_{dl} in the high-power region [44]. This is again the case for both precoding strategies, as can be observed in Fig. 8.3 where the number of active users equals $T_{\text{dl}} = 8$. Furthermore, the power allocation patterns are similar, i.e., both methods decide to serve the same users with a comparable distribution of power. This is particularly interesting for the E2E Edge-GAT approach since this pure learning-based method does not have any related prior knowledge during the training phase.

9 Conclusion

In this dissertation, we provided an in-depth examination of different precoding techniques for multi-user massive MISO systems operating in FDD mode with incomplete CSI. We tackled significant challenges in modern wireless communication systems, particularly in the context of 5G and future technologies, where efficient use of the spectrum and strong performance amid channel uncertainties are crucial.

Key findings include the unexpected result that zero-forcing precoding combined with LS channel estimation can achieve interference-free transmission in high-SNR regimes under incomplete CSI. Here, we show that the number of training pilots defines the achievable degrees of freedom. The interference-free behavior of the zero-forcing precoder combined with LS holds, therefore, only when the number of pilots exceeds the number of users. Given highly limited pilot resources, we developed alternative strategies to outperform the zero-forcing precoder and cope with the incomplete CSI. For instance, we leveraged long-term channel statistics to design statistical precoders, reducing computational overhead by requiring optimization only when channel covariance matrices change. Moreover, our proposed AWAMSE and MM precoding methods have demonstrated superior performance to the zero-forcing and MMSE precoders and were shown to be competitive with existing, more computationally demanding state-of-the-art approaches.

One of the major contributions of this work is the integration of RSMA with various precoding techniques. This approach has demonstrated significant potential in enhancing spectral efficiency and robustness, showcasing RSMA's ability to handle inter-user interference, especially in the high-SNR regime, unlike its non-RS counterparts.

Another key contribution is our innovative deep-learning framework for jointly designing pilots and precoders using GNNs. Optimized pilot signals have outperformed conventional orthogonal pilots, paving the way for new possibilities in improving system performance through intelligent pilot design.

Looking ahead, several promising research directions emerge from this work:

- The presented precoding approaches, along with their RS extensions, were proposed for a single-cell scenario; they can be, however, extended to multi-cell

scenarios to tackle challenges such as inter-cell interference and coordinated multi-point transmission.

- As wireless systems evolve to higher frequency bands, the characteristics of the channels and the challenges associated with CSI acquisition change significantly. Adapting our precoding techniques for millimeter-wave and future terahertz communication systems could pave the way for ultra-high-speed, short-range communications.
- Another exciting direction is to integrate the proposed precoding techniques with other emerging technologies, such as full-duplex systems and reconfigurable intelligent surfaces.

In conclusion, the techniques and insights developed in this dissertation advance theoretical understanding and practical solutions for future wireless communication networks. These contributions form a solid foundation for further research and innovation, driving progress towards 6G and beyond and enabling more efficient, robust, and adaptive wireless systems.

Appendix A

A.1 Spectral Efficiency in a Discrete Memoryless Interference Channel

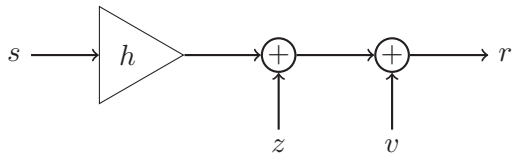

Fig. A.1: Discrete memoryless interference channel

In Fig. A.1, we depict the model for a discrete memoryless interference channel with input s, output r, channel response h, interference z, and independent Gaussian noise v. Assuming that the interference has zero mean and is uncorrelated with the input and the channel, an expression for the achievable spectral efficiency reads as

$$R = \log_2 \left(\frac{|h|^2}{\mathbb{E}[|z|^2] + 1} \right) \tag{A.1}$$

where the noise variance is assumed to be 1.

The derivation of (A.1) is based on three eventually suboptimal assumptions. First, the input distribution is assumed to be Gaussian, which is not necessarily optimal. Second, we assume that the input s is estimated from r using LMMSE estimation. Lastly, the estimation error is assumed to be Gaussian distributed, corresponding to the worst-case noise scenario.

Note that (A.1) holds for a deterministic channel response h [14]. If we now consider h a random variable, then we need to consider the so-called ergodic rate. To this end, we assume that the transmission of the input signal spans a long sequence of channel realizations.

Similar to the previous case with a deterministic channel response, an expression for the ergodic achievable rate can be written as

$$\bar{R} = \mathbb{E}[\log_2 \left(\frac{|h|^2}{\mathbb{E}[|z|^2|h,u] + 1} \right)]. \tag{A.2}$$

Here, u denotes a random variable that influences the interference term, and the outer expectation is taken w.r.t. these realizations. The term inside the expectation in (A.2) corresponds to the instantaneous rate for a given channel instance. The derivation of (A.2) follows the same lines of the deterministic channel response case, except that the interference needs to be conditionally independent of the input signal given h and u.

A.2 Derivation of the Rate Expression and Power Constraint for the Statistical Precoder

A.2.1 Derivation of the Effective SINR Expression

Recall that the effective SINR of user k is given by (cf. (3.27))

$$\mathrm{SINR}_k = \frac{|\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k]|^2}{\sum_{j=1}^{K}\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2] - |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k]|^2 + 1}. \tag{A.3}$$

Now we compute each term of (A.3) individually for $\boldsymbol{p}_k = \boldsymbol{A}_k\boldsymbol{y}_k$ and $\boldsymbol{y}_k = \boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k}$, where $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_k)$ and $\boldsymbol{n}_{1,k} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \frac{1}{P_{\mathrm{dl}}}\boldsymbol{I}_{T_{\mathrm{dl}}})$ are assumed to be independent.

First, the numerator is evaluated according to

$$\begin{aligned}
|\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k]|^2 &= |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{y}_k]|^2 \\
&= |\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k})]|^2 \\
&= |\mathrm{tr}(\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k})])|^2 \\
&= |\mathbb{E}[\mathrm{tr}(\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k + \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{n}_{1,k})]|^2 \\
&= |\mathbb{E}[\mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}} + \boldsymbol{A}_k\boldsymbol{n}_{1,k}\boldsymbol{h}_k^{\mathrm{H}})]|^2 \\
&= |\mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\mathbb{E}[\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}] + \boldsymbol{A}_k\underbrace{\mathbb{E}[\boldsymbol{n}_{1,k}\boldsymbol{h}_k^{\mathrm{H}}]}_{\boldsymbol{0}})|^2 \\
&= |\mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{C}_k)|^2.
\end{aligned} \tag{A.4}$$

We now concentrate on computing $\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k|^2]$

$$\begin{aligned}
\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k|^2] &= \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{y}_k|^2] \\
&= \mathbb{E}[\boldsymbol{y}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{y}_k] \\
&= \mathbb{E}[(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k})^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k + \boldsymbol{n}_{1,k})] \\
&= \mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k] + \mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{n}_{1,k}] \\
&+ \mathbb{E}[\boldsymbol{n}_{1,k}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{h}_k] + \mathbb{E}[\boldsymbol{n}_{1,k}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{n}_{1,k}] \qquad \text{(A.5)} \\
&= \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_k|^2] + \frac{1}{P_{\mathrm{dl}}}\mathrm{tr}(\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{C}_k\boldsymbol{A}_k). \qquad \text{(A.6)}
\end{aligned}$$

To compute the first term of (A.6), we use [[81], *Lemma 2*] and obtain therefore

$$\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_k|^2] = |\mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{C}_k)|^2 + \mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{C}_k\boldsymbol{\Phi}\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{C}_k) + \frac{1}{P_{\mathrm{dl}}}\mathrm{tr}(\boldsymbol{A}_k^{\mathrm{H}}\boldsymbol{C}_k\boldsymbol{A}_k). \tag{A.7}$$

Next, we consider the term $\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2]$ for $j \neq k$. Following the same steps done in (A.5), we obtain

$$\begin{aligned}\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{p}_j|^2] &= \mathbb{E}[\boldsymbol{h}_j^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_j\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{h}_j] + \mathbb{E}[\boldsymbol{h}_j^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_j\boldsymbol{n}_{1,j}] \\ &\quad + \mathbb{E}[\boldsymbol{n}_{1,j}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_j\boldsymbol{h}_j] + \mathbb{E}[\boldsymbol{n}_{1,j}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_j\boldsymbol{n}_{1,j}] \\ &= \mathrm{tr}(\mathbb{E}[\boldsymbol{h}_j\boldsymbol{h}_j^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_j\boldsymbol{\Phi}^{\mathrm{H}}]) + \mathrm{tr}(\mathbb{E}[\boldsymbol{n}_{1,j}\boldsymbol{n}_{1,j}^{\mathrm{H}}\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{A}_j]) \\ &= \mathrm{tr}(\boldsymbol{C}_j\boldsymbol{\Phi}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{C}_k\boldsymbol{A}_j\boldsymbol{\Phi}^{\mathrm{H}}) + \frac{1}{P_{\mathrm{dl}}}\mathrm{tr}(\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{C}_k\boldsymbol{A}_j)\end{aligned} \tag{A.8}$$

where we used the independence of $\boldsymbol{h}_k$ and $\boldsymbol{h}_j$ for $k \neq j$.

The effective SINR can be thus written as

$$\mathrm{SINR}_k = \frac{|\mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{C}_k)|^2}{\sum_{j=1}^{K}\mathrm{tr}(\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{C}_k) + 1}. \tag{A.9}$$

Now, let $\boldsymbol{a}_k = \mathrm{vec}(\boldsymbol{A}_k) \in \mathbb{C}^{MT_{\mathrm{dl}}\times 1}$ and $\boldsymbol{c}_k = \mathrm{vec}(\boldsymbol{C}_k) \in \mathbb{C}^{M^2\times 1}$. We can reformulate (A.9) as follows

$$\mathrm{SINR}_k = \frac{|\mathrm{tr}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{C}_k)|^2}{\sum_{j=1}^{K}\mathrm{tr}(\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{C}_k) + 1} \tag{A.10}$$

$$= \frac{|\mathrm{tr}(\boldsymbol{C}_k^{\mathrm{H}}\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}})|^2}{\sum_{j=1}^{K}\mathrm{tr}(\boldsymbol{A}_j^{\mathrm{H}}\boldsymbol{C}_k\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}) + 1} \tag{A.11}$$

$$= \frac{|\boldsymbol{c}_k^{\mathrm{H}}\mathrm{vec}(\boldsymbol{A}_k\boldsymbol{\Phi}^{\mathrm{H}})|^2}{\sum_{j=1}^{K}\boldsymbol{a}_j^{\mathrm{H}}\mathrm{vec}(\boldsymbol{C}_k\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}) + 1} \tag{A.12}$$

$$= \frac{|\boldsymbol{c}_k^{\mathrm{H}}(\boldsymbol{\Phi}^{\star}\otimes\boldsymbol{I}_M)\boldsymbol{a}_k|^2}{\sum_{j=1}^{K}\boldsymbol{a}_j^{\mathrm{H}}(\boldsymbol{C}_{\boldsymbol{y}_j}^{\mathrm{T}}\otimes\boldsymbol{C}_k)\boldsymbol{a}_j + 1} \tag{A.13}$$

$$= \frac{|\boldsymbol{q}_k^{\mathrm{H}}\boldsymbol{a}_k|^2}{\sum_{j=1}^{K}\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{Q}_{j,k}\boldsymbol{a}_j + 1} \tag{A.14}$$

where we used in (A.11) the property $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$ and that for a covariance matrix $\boldsymbol{C}_k = \boldsymbol{C}_k^{\mathrm{H}}$. In (A.11) and (A.12), we applied respectively the identities [82]

$$\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{B}) = \mathrm{vec}(\boldsymbol{A})^{\mathrm{H}}\mathrm{vec}(\boldsymbol{B}) \tag{A.15}$$

and

$$\text{vec}(\boldsymbol{AXB}) = (\boldsymbol{B}^{\text{T}} \otimes \boldsymbol{A})\text{vec}(\boldsymbol{X}). \tag{A.16}$$

Lastly, we introduced in (A.14) the following variables

$$\boldsymbol{q}_k = (\boldsymbol{\Phi}^{\text{T}} \otimes \boldsymbol{I}_M)\boldsymbol{c}_k, \tag{A.17}$$

$$\boldsymbol{Q}_{j,k} = \boldsymbol{C}_{\boldsymbol{y}_j}^{\text{T}} \otimes \boldsymbol{C}_k. \tag{A.18}$$

A.2.2 Derivation of the Power Constraint

Recall that the long-term power constraint is formulated as follows

$$\mathbb{E}[\|\boldsymbol{x}\|^2] = \mathbb{E}[\sum_{j=1}^{K} \|\boldsymbol{p}_j\|^2] \leq P_{\text{dl}}. \tag{A.19}$$

Using $\boldsymbol{p}_k = \boldsymbol{A}_k\boldsymbol{y}_k$, we obtain

$$\mathbb{E}[\sum_{j=1}^{K} \|\boldsymbol{p}_j\|^2] = \sum_{j=1}^{K} \mathbb{E}[\text{tr}(\boldsymbol{p}_j\boldsymbol{p}_j^{\text{H}})] = \sum_{j=1}^{K} \mathbb{E}[\text{tr}(\boldsymbol{A}_j\boldsymbol{y}_j\boldsymbol{y}_j^{\text{H}}\boldsymbol{A}_j^{\text{H}})] \tag{A.20}$$

$$= \sum_{j=1}^{K} \text{tr}(\boldsymbol{A}_j\mathbb{E}[\boldsymbol{y}_j\boldsymbol{y}_j^{\text{H}}]\boldsymbol{A}_j^{\text{H}}) = \sum_{j=1}^{K} \text{tr}(\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}\boldsymbol{A}_j^{\text{H}}). \tag{A.21}$$

Applying the identities given by (A.15) and (A.16) yields

$$\sum_{j=1}^{K} \text{tr}(\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}\boldsymbol{A}_j^{\text{H}}) = \sum_{j=1}^{K} \boldsymbol{a}_j^{\text{H}}\text{vec}(\boldsymbol{A}_j\boldsymbol{C}_{\boldsymbol{y}_j}) = \sum_{j=1}^{K} \boldsymbol{a}_j^{\text{H}}(\boldsymbol{C}_{\boldsymbol{y}_j}^{\text{T}} \otimes \boldsymbol{I}_M)\boldsymbol{a}_j \tag{A.22}$$

$$= \sum_{j=1}^{K} \boldsymbol{a}_j^{\text{H}}\boldsymbol{F}_j\boldsymbol{a}_j \leq P_{\text{dl}} \tag{A.23}$$

where we used the vectorized transformation matrix $\boldsymbol{a}_j = \text{vec}(\boldsymbol{A}_j)$ and introduced the matrix

$$\boldsymbol{F}_j = \boldsymbol{C}_{\boldsymbol{y}_j}^{\text{T}} \otimes \boldsymbol{I}_M. \tag{A.24}$$

A.3 Derivation of the Effective Common SINR Expression for the RS Statistical Precoding Approach

In order to derive a compact closed-form expression for the common SINR of user k, we first define the vectors $\boldsymbol{h}$ and $\boldsymbol{n}$, where we stack the channel vectors and noise

vectors of all users, respectively, such that

$$\begin{aligned} \boldsymbol{h} &= [\boldsymbol{h}_1^{\mathrm{T}}, \ldots, \boldsymbol{h}_K^{\mathrm{T}}]^{\mathrm{T}}, \qquad & \boldsymbol{n} &= [\boldsymbol{n}_1^{\mathrm{T}}, \ldots, \boldsymbol{n}_K^{\mathrm{T}}]^{\mathrm{T}}, \\ \boldsymbol{h}_k &= \boldsymbol{E}_k \boldsymbol{h}, & \boldsymbol{n}_k &= \boldsymbol{E}_k \boldsymbol{n} \end{aligned} \tag{A.25}$$

where $\boldsymbol{E}_k = \boldsymbol{e}_k^{\mathrm{T}} \otimes \boldsymbol{I}_M$ is an $M \times KM$-selection matrix.
The vector $\boldsymbol{y}$ of all users' channel observations can be therefore written as

$$\boldsymbol{y} = \underbrace{(\boldsymbol{I}_K \otimes \boldsymbol{\Phi}^{\mathrm{H}})}_{\boldsymbol{D}} \boldsymbol{h} + \boldsymbol{n} \tag{A.26}$$

whose covariance matrix is given by

$$\boldsymbol{C}_{\boldsymbol{y}} = \boldsymbol{D}\boldsymbol{C}_{\boldsymbol{h}}\boldsymbol{D}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{n}} \tag{A.27}$$

with $\boldsymbol{C}_{\boldsymbol{h}} = \mathbb{E}[\boldsymbol{h}\boldsymbol{h}^{\mathrm{H}}]$ and $\boldsymbol{C}_{\boldsymbol{n}} = \mathbb{E}[\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}]$ being the block diagonal covariance matrices of $\boldsymbol{h}$ and $\boldsymbol{n}$, respectively.
Following similar steps to Appendix A.2, we can evaluate the expectation inside the numerator of the common SINR as follows

$$\mathbb{E}[\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}] = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{h}} \boldsymbol{E}_k^{\mathrm{T}} \boldsymbol{A}_{\mathrm{c}} \boldsymbol{D}) = \boldsymbol{c}_{\boldsymbol{h}}^{\mathrm{H}} (\boldsymbol{D}^{\mathrm{T}} \otimes \boldsymbol{E}_k^{\mathrm{T}}) \boldsymbol{a}_{\mathrm{c}} = \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{c}} \tag{A.28}$$

where we used the vectorized forms $\boldsymbol{c}_{\boldsymbol{h}} = \mathrm{vec}(\boldsymbol{C}_{\boldsymbol{h}})$ and $\boldsymbol{a}_{\mathrm{c}} = \mathrm{vec}(\boldsymbol{A}_{\mathrm{c}})$, and defined the vector $\boldsymbol{z}_k$ as follows

$$\boldsymbol{z}_k = (\boldsymbol{D}^* \otimes \boldsymbol{E}_k) \boldsymbol{c}_{\boldsymbol{h}}. \tag{A.29}$$

Additionally, we have

$$\begin{aligned} \mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{c}}|^2] &= \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{h}} \boldsymbol{E}_k^{\mathrm{T}} \boldsymbol{A}_{\mathrm{c}} \boldsymbol{D}) \\ &\quad + \mathrm{tr}(\underbrace{\boldsymbol{E}_k \boldsymbol{C}_{\boldsymbol{h}} \boldsymbol{E}_k^{\mathrm{H}}}_{\boldsymbol{C}_k} \boldsymbol{A}_{\mathrm{c}} (\boldsymbol{D}\boldsymbol{C}_{\boldsymbol{h}}\boldsymbol{D}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{n}}) \boldsymbol{A}_{\mathrm{c}}^{\mathrm{H}}) && \text{(A.30)} \\ &= |\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}|^2 + \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} (\boldsymbol{C}_{\boldsymbol{y}}^{\mathrm{T}} \otimes \boldsymbol{C}_k) \boldsymbol{a}_{\mathrm{c}} && \text{(A.31)} \\ &= |\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}|^2 + \boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{Q}_k \boldsymbol{a}_{\mathrm{c}} && \text{(A.32)} \end{aligned}$$

where we introduced the matrix $\boldsymbol{Q}_k$ as

$$\boldsymbol{Q}_k = \boldsymbol{C}_{\boldsymbol{y}}^{\mathrm{T}} \otimes \boldsymbol{C}_k. \tag{A.33}$$

Using the result for $\mathbb{E}[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_{\mathrm{p},i}|^2]$ given in (A.8) (replacing $\boldsymbol{p}_i$ with $\boldsymbol{p}_{\mathrm{p},i}$) and after vectorization, the common SINR of user k can finally be written in its compact form as

$$\mathrm{SINR}_{\mathrm{c},k} = \frac{|\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}|^2}{\boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}} \boldsymbol{Q}_k \boldsymbol{a}_{\mathrm{c}} + |\boldsymbol{q}_k^{\mathrm{H}} \boldsymbol{a}_{\mathrm{p},k}|^2 + \sum_{i=1}^{K} \boldsymbol{a}_{\mathrm{p},i}^{\mathrm{H}} \boldsymbol{Q}_{i,k} \boldsymbol{a}_{\mathrm{p},i} + 1}. \tag{A.34}$$

A.4 Proof of the Uplink Downlink SINR Duality

We focus here on the uplink to downlink transform. The interested reader may refer to [65] for more details and for the converse proof.
Given the dual uplink transformation vector $\boldsymbol{t}_k$ and the power $\mu_k \geq 0$ such that $\sum_k \mu_k = P_{\text{dl}}$ with some dual uplink SINR, then the downlink transformation vector $\boldsymbol{a}_k$ can be found using the following linear relationship

$$\boldsymbol{a}_k = \sqrt{\alpha}_k \boldsymbol{t}_k \tag{A.35}$$

where the coefficients α_k are given by solving the system of equations $\text{SINR}_k = \text{SINR}_k^{\text{dual}}$ for all k:

$$\frac{\alpha_k \boldsymbol{t}_k^{\text{H}} \boldsymbol{q}_k \boldsymbol{q}_k \boldsymbol{t}_k}{\sum_j \alpha_j \boldsymbol{t}_j^{\text{H}} \boldsymbol{Q}_{j,k} \boldsymbol{t}_j + 1} \stackrel{!}{=} \frac{\mu_k \boldsymbol{t}_k^{\text{H}} \boldsymbol{q}_k \boldsymbol{q}_k \boldsymbol{t}_k}{\sum_j \mu_j \boldsymbol{t}_k^{\text{H}} \boldsymbol{Q}_{k,j} \boldsymbol{t}_k + \boldsymbol{t}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{t}_k}. \tag{A.36}$$

After rearranging and simplifying the terms, we obtain

$$\mu_k = \alpha_k (\boldsymbol{t}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{t}_k + \sum_{j \neq k} \mu_j \boldsymbol{t}_k^{\text{H}} \boldsymbol{Q}_{k,j} \boldsymbol{t}_k) - \mu_k \sum_{j \neq k} \alpha_j \boldsymbol{t}_j^{\text{H}} \boldsymbol{Q}_{j,k} \boldsymbol{t}_j \tag{A.37}$$

which can be written in matrix-vector notation as follows

$$\boldsymbol{\mu} = \boldsymbol{J} \boldsymbol{\alpha} \tag{A.38}$$

where $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_K]^{\text{T}}$ and $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]^{\text{T}}$.
The entry of the matrix $\boldsymbol{J}$ in the kth row and ith column reads as

$$[\boldsymbol{J}]_{k,i} = \begin{cases} \boldsymbol{t}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{t}_k + \sum_{j \neq k} \mu_j \boldsymbol{t}_k^{\text{H}} \boldsymbol{Q}_{k,j} \boldsymbol{t}_k, & \text{for } k = i \\ -\mu_k \boldsymbol{t}_j^{\text{H}} \boldsymbol{Q}_{j,k} \boldsymbol{t}_j, & \text{otherwise.} \end{cases} \tag{A.39}$$

Note that the matrix $\boldsymbol{J}$ is column-wise diagonally dominant with positive diagonal entries and negative off-diagonal entries. Therefore, $\boldsymbol{\alpha} = \boldsymbol{J}^{-1} \boldsymbol{\mu}$ is non-negative. Furthermore, it holds that

$$P_{\text{dl}} = \boldsymbol{1}^{\text{T}} \boldsymbol{\mu} = \boldsymbol{1}^{\text{T}} \boldsymbol{J} \boldsymbol{\alpha} = \sum_k \alpha_k \boldsymbol{t}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{t}_k = \sum_k \boldsymbol{a}_k^{\text{H}} \boldsymbol{F}_k \boldsymbol{a}_k. \tag{A.40}$$

Hence, the sum power constraint is preserved.

A.5 Iterative Common SINR Increasing Method

For fixed step size u, the parameters v and $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$ in (6.64) are chosen such that:

1. The power constraint $\boldsymbol{a}_{\mathrm{c}}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{a}_{\mathrm{c}} \leq \alpha_{\mathrm{c}}P_{\mathrm{dl}}$ is satisfied with equality at iteration $n+1$ for $\boldsymbol{a}_{\mathrm{c}} = \boldsymbol{a}_{\mathrm{c}}^{(n+1)}$.

$$\begin{aligned}&\boldsymbol{a}_{\mathrm{c}}^{(n+1),\mathrm{H}}\boldsymbol{F}\boldsymbol{a}_{\mathrm{c}}^{(n+1)}\\&= (1-u)^2\alpha_{\mathrm{c}}P_{\mathrm{dl}} + |v|^2\|\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}\|_2^2\\&+ 2(1-\alpha)\Re\{\beta\boldsymbol{a}_{\mathrm{c}}^{(n),\mathrm{H}}\boldsymbol{F}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}\} \stackrel{!}{=} \alpha_{\mathrm{c}}P_{\mathrm{dl}}.\end{aligned}$$

   If we set $\|\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}\|_2^2$ to 1, which corresponds to a normalization of the vector $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$, and $\Re\{\beta\boldsymbol{a}_{\mathrm{c}}^{(n),\mathrm{H}}\boldsymbol{F}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}\}$ to zero, by imposing the orthogonality constraint

$$\boldsymbol{a}_{\mathrm{c}}^{(n),\perp} \perp \boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2},\mathrm{H}}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{a}_{\mathrm{c}}^{(n)} \tag{A.41}$$

   the magnitude of v is given by

$$|v| = \sqrt{\alpha_{\mathrm{c}}P_{\mathrm{dl}}(2u-u^2)}. \tag{A.42}$$

   since $\boldsymbol{a}_{\mathrm{c}}^{(n),\mathrm{H}}\boldsymbol{F}\boldsymbol{a}_{\mathrm{c}}^{(n)} = \alpha_{\mathrm{c}}P_{\mathrm{dl}}$.

2. The left-hand-side of the SINR constraint in (6.61) is maximized. When inserting the update in (6.64), the latter is given by

$$\begin{aligned}&2\Re\{\eta_{\ell(n)}^{*}\boldsymbol{z}_{\ell(n)}^{\mathrm{H}}\boldsymbol{a}_{\mathrm{c}}^{(n+1)}\} - |\eta_{\ell(n)}|^2\left(\boldsymbol{a}_{\mathrm{c}}^{(n+1),\mathrm{H}}\boldsymbol{Q}_{\ell(n)}\boldsymbol{a}_{\mathrm{c}}^{(n+1)} + \sigma_{\ell(n)}^2\right) && \text{(A.43)}\\=&2(1-u)\Re\{\eta_{\ell(n)}^{*}\boldsymbol{z}_{\ell(n)}^{\mathrm{H}}\boldsymbol{a}_{\mathrm{c}}^{(n)}\} && \text{(A.44)}\\&-|\eta_{\ell(n)}|^2\left((1-u)^2\boldsymbol{a}_{\mathrm{c}}^{(n),\mathrm{H}}\boldsymbol{Q}_{\ell(n)}\boldsymbol{a}_{\mathrm{c}}^{(n)} + \sigma_{\ell(n)}^2\right) && \text{(A.45)}\\&-|\eta_{\ell(n)}|^2|v|^2\boldsymbol{a}_{\mathrm{c}}^{(n),\perp,\mathrm{H}}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2},\mathrm{H}}\boldsymbol{Q}_{\ell(n)}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}\boldsymbol{a}_{\mathrm{c}}^{(n),\perp} && \text{(A.46)}\\&+2|v|\Re\left\{\mathrm{e}^{\mathrm{j}\angle v}\boldsymbol{t}_{\mathrm{c}}^{(n),\mathrm{H}}\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}\right\} && \text{(A.47)}\end{aligned}$$

   where we used for the last term (A.47) that $v = |v|\mathrm{e}^{\mathrm{j}\angle v}$. Furthermore, we defined the vector $\boldsymbol{t}_{\mathrm{c}}^{(n)}$ such that

$$\boldsymbol{t}_{\mathrm{c}}^{(n)} = \eta_{\ell(n)}\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2},\mathrm{H}}\boldsymbol{z}_{\ell(n)} - |\eta_{\ell(n)}|^2(1-u)\boldsymbol{Q}_{\ell(n)}^{\frac{1}{2}}\boldsymbol{a}_{\mathrm{c}}^{(n)}. \tag{A.48}$$

The terms in (A.44) and (A.45) depend only on $\boldsymbol{a}_{\mathrm{c}}^{(n)}$, the common transformation vector from iteration n, and are therefore not taken into account when maximizing the whole expression.

Due to the transformation applied to $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$ via the matrix $\boldsymbol{Q}_{\ell(n)}^{-\frac{1}{2}}$, we can see that the third term in (A.46) simplifies to $|\eta_{\ell(n)}|^2|v|^2\|\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}\|^2$. Since the norm of $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$ is already fixed by the transmit power constraint, the only term we have to maximize is the one given by (A.47). Here, the maximization is done w.r.t. $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$ and $\angle v$, the phase of v (the magnitude is already fixed due to the power constraint, cf. (A.42)).

For fixed $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$, $\angle v$ is chosen such that the real part in (A.47) is maximized and thus it is given by

$$\angle v = -\angle \boldsymbol{t}_{\mathrm{c}}^{(n),\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}^{(n),\perp}. \tag{A.49}$$

With this choice for $\angle v$, the real part in (A.47) can be written as

$$|\boldsymbol{t}_{\mathrm{c}}^{(n),\mathrm{H}} \boldsymbol{a}_{\mathrm{c}}^{(n),\perp}| \tag{A.50}$$

which is maximized by choosing $\boldsymbol{a}_{\mathrm{c}}^{(n),\perp}$ as "co-linear" as possible to $\boldsymbol{t}_{\mathrm{c}}^{(n)}$.

Acronyms

5G	Fifth Generation cellular wireless standard
6G	Sixth Generation cellular wireless standard
AMSE	Average Mean Square Error
AWAMSE	Augmented Weighted Average Mean Square Error
CCCP	ConCave Convex Procedure
CDF	Cumulative Distribution Function
CSI	Channel State Information
DFT	Discrete Fourier Transform
DL	Downlink
E2E	End-to-End
FDD	Frequency Division Duplex
FP	Fractional Programming
GAT	Graph ATtention network
GMM	Gaussian Mixture Model
GNN	Graph Neural Network
IWMMSE	Iterative Weighted Minimum Mean Square Error
KKT	Karush-Kuhn-Tucker
LB	Lower Bound
LMMSE	Linear Minimum Mean Square Error
LS	Least Squares

MA	Multiple Access
MF	Matched Filter
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MM	Minorization Maximization
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NMSE	Normalized Mean Square Error
QCQP	Quadratically Constrained Quadratic Program
QuaDRiGa	QUAsi Deterministic RadIo Channel GenerAtor
ReLu	Rectified Linear unit
RS	Rate Splitting
RSMA	Rate Splitting Multiple Access
SAA	Sample Average Approximation
SDMA	Spatial Division Multiple Access
SDR	Semi-Definite Relaxation
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise Ratio
SIWMMSE	Stochastic Iterative Weighted Minimum Mean Square Error
SNR	Signal-to-Noise-Ratio
TDD	Time Division Duplex
Tx	Transmitter
UL	Uplink
ULA	Uniform Linear Array
WMMSE	Weighted Minimum Mean Square Error
WMSE	Weighted Mean Square Error
ZF	Zero Forcing

List of Figures

List of Tables

List of Algorithms

Bibliography

[1] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. Poor, *MIMO Wireless Communications*. Cambridge university press, 2007.

[2] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[3] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five Disruptive Technology Directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[4] T. L. Marzetta, E. Larsson, H. Yang, and H. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016. [Online]. Available: https://books.google.de/books?id=Be08DQAAQBAJ

[5] P. W. C. Chan, E. S. Lo, R. R. Wang, E. K. S. Au, V. K. N. Lau, R. S. Cheng, W. H. Mow, R. D. Murch, and K. B. Letaief, "The Evolution Path of 4G Networks: FDD or TDD?" *IEEE Communications Magazine*, vol. 44, no. 12, pp. 42–50, 2006.

[6] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable Rates of FDD Massive MIMO Systems With Spatial Channel Correlation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2868–2882, 2015.

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

[8] 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, 2010.

[9] T. L. Marzetta, "How Much Training is Required for Multiuser MIMO?" in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, 2006, pp. 359–363.

[10] M. Kobayashi, N. Jindal, and G. Caire, "Training and Feedback Optimization for Multiuser MIMO Downlink," *IEEE Transactions on Communications*, vol. 59, no. 8, 2011.

[11] B. M. Hochwald and S. ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, 2003.

[12] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, 2004.

[13] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication—Part I: Channel Inversion and Regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, 2005.

[14] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*, 2017, vol. 11, no. 3-4.

[15] ——, "Massive MIMO Has Unlimited Capacity," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 574–590, 2018.

[16] T. Yoo and A. Goldsmith, "Capacity of Fading MIMO Channels with Channel Estimation Error," in *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, vol. 2, 2004, pp. 808–813 Vol.2.

[17] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing - The Large-Scale Array Regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[18] D. Neumann, T. Wiese, M. Joham, and W. Utschick, "A Bilinear Equalizer for Massive MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 14, pp. 3740–3751, 2018.

[19] J. Hoydis, S. t. Brink, and M. Debbah, "Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[20] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels under Limited

Feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, 2012.

[21] Z. Zhang, Z. Zhao, K. Shen, D. P. Palomar, and W. Yu, "Discerning and Enhancing the Weighted Sum-Rate Maximization Algorithms in Communications," 2023, https://arxiv.org/pdf/2311.04546.pdf.

[22] Z. Luo and S. Zhang, "Dynamic Spectrum Management: Complexity and Duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, 2008.

[23] Q. Shi, M. Razaviyayn, Z. 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, 2011.

[24] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted Sum-Rate Maximization Using Weighted MMSE for MIMO-BC Beamforming Design," in *2009 IEEE International Conference on Communications*, 2009, pp. 1–6.

[25] X. Zhao, S. Lu, Q. Shi, and Z. Luo, "Rethinking WMMSE: Can Its Complexity Scale Linearly With the Number of BS Antennas?" *IEEE Transactions on Signal Processing*, vol. 71, pp. 433–446, 2023.

[26] Q. Hu, Y. Cai, Q. Shi, K. Xu, G. Yu, and Z. Ding, "Iterative Algorithm Induced Deep-Unfolding Neural Networks: Precoding Design for Multiuser MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1394–1410, 2021.

[27] K. Shen and W. Yu, "Fractional Programming for Communication Systems—Part I: Power Control and Beamforming," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2616–2630, 2018.

[28] Z. Zhang and Z. Zhao, "Rate Maximizations for Reconfigurable Intelligent Surface-Aided Wireless Networks: A Unified Framework via Block Minorization-Maximization," 2021, https://arxiv.org/abs/2105.02395.

[29] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2017.

[30] 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 Communications Magazine*, vol. 54, no. 5, pp. 98–105, 2016.

[31] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-Splitting Multiple Access: Fundamentals, Survey, and Future Research Trends," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2073–2126, 2022.

[32] A. R. Flores, R. C. de Lamare, and B. Clerckx, "Linear Precoding and Stream Combining for Rate Splitting in Multiuser MIMO Systems," *IEEE Communications Letters*, vol. 24, no. 4, pp. 890–894, 2020.

[33] G. Zhou, Y. Mao, and B. Clerckx, "Rate-Splitting Multiple Access for Multi-Antenna Downlink Communication Systems: Spectral and Energy Efficiency Tradeoff," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 4816–4828, 2022.

[34] J. An, O. Dizdar, B. Clerckx, and W. Shin, "Rate-Splitting Multiple Access for Multi-Antenna Broadcast Channel with Imperfect CSIT and CSIR," in *IEEE Annu. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2020, pp. 1–7.

[35] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A Rate Splitting Strategy for Massive MIMO With Imperfect CSIT," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4611–4624, 2016.

[36] O. Dizdar, Y. Mao, W. Han, and B. Clerckx, "Rate-Splitting Multiple Access: A New Frontier for the PHY Layer of 6G," in *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, 2020, pp. 1–7.

[37] H. Joudeh and B. Clerckx, "Robust Transmission in Downlink Multiuser MISO Systems: A Rate-Splitting Approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6227–6242, 2016.

[38] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A Hierarchical Rate Splitting Strategy for FDD Massive MIMO under Imperfect CSIT," in *2015 IEEE 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, 2015, pp. 80–84.

[39] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured Compressive Sensing-Based Spatio-Temporal Joint Channel Estimation for FDD Massive MIMO," *IEEE Transactions on Communications*, vol. 64, no. 2, p. 601 – 617, 2016.

[40] N. Shalavi, M. Atashbar, and M. Mohassel Feghhi, "Downlink channel estimation of FDD based massive MIMO using spatial partial-common sparsity modeling," *Physical Communication*, vol. 42, p. 101138, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1874490720302159

[41] F. Sohrabi, K. M. Attiah, and W. Yu, "Deep Learning for Distributed Channel Feedback and Multiuser Precoding in FDD Massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4044–4057, 2021.

[42] K. Shen and W. Yu, "Fractional Programming for Communication Systems—Part II: Uplink Scheduling via Matching," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2631–2644, 2018.

[43] M. Joham, K. Kusume, M. Gzara, W. Utschick, and J. Nossek, "Transmit Wiener Filter for the Downlink of TDDDS-CDMA Systems," in *IEEE Seventh International Symposium on Spread Spectrum Techniques and Applications,*, vol. 1, 2002, pp. 9–13 vol.1.

[44] D. Ben Amor, M. Joham, and W. Utschick, "Asymptotic Behavior of Zero-Forcing Precoding based on Imperfect Channel Knowledge for Massive MISO FDD Systems," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 1500–1505.

[45] ——, "Bilinear Precoding for FDD Massive MIMO System with Imperfect Covariance Matrices," in *WSA 2020; 24th International ITG Workshop on Smart Antennas*, 2020, pp. 1–6.

[46] ——, "Rate Splitting in FDD Massive MIMO Systems Based on the Second Order Statistics of Transmission Channels," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1351–1365, 2023.

[47] ——, "Robust Precoding for FDD MISO Systems via Minorization Maximization," 2024, accepted for publication in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, https://arxiv.org/abs/2411.02615.

[48] ——, "Highly Accelerated Weighted MMSE Algorithms for Designing Precoders in FDD Systems with Incomplete CSI," in *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, 2024, pp. 1–5.

[49] ——, "An Efficient Rate Splitting Precoding Approach in Multi-User MISO FDD Systems," in *2024 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2024, pp. 1852–1857.

[50] V. Rizzello, D. Ben Amor, M. Joham, and W. Utschick, "Scalable Multi-User Precoding and Pilot Optimization with Graph Neural Networks," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 2956–2961.

[51] M. Barzegar Khalilsarai, S. Haghighatshoar, X. Yi, and G. Caire, "FDD Massive MIMO via UL/DL Channel Covariance Extrapolation and Active Channel Sparsification," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 121–135, 2019.

[52] 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, 2003.

[53] M. Biguesh and A. B. Gershman, "Downlink Channel Estimation in Cellular Systems with Antenna Arrays at Base Stations Using Channel Probing with Feedback," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 9, p. 963649, 2004. [Online]. Available: https://doi.org/10.1155/S1110865704403023

[54] M. Medard, "The Effect upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 933–946, 2000.

[55] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, Massachusetts: Athena Scientific, 1999.

[56] J. Pan, W.-K. Ma, X. Xia, and Y. Tian, "WMMSE-based multiuser MIMO beamforming: A practice-oriented design and LTE system performance evaluation," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2015, pp. 435–439.

[57] B. L. Fox and D. M. Landi, "Searching for the Multiplier in One-Constraint Optimization Problems," *Operations Research*, vol. 18, no. 2, pp. 253–262, 1970. [Online]. Available: http://www.jstor.org/stable/168682

[58] 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, 2003.

[59] A. Wiesel, C. Y. Eldar, and S. Shamai, "Zero-Forcing Precoding and Generalized Inverses," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4409–4418, 2008.

[60] J. Lee and N. Jindal, "Dirty Paper Coding vs. Linear Precoding for MIMO Broadcast Channels," in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, 2006, pp. 779–783.

[61] M. Koller, B. Fesl, N. Turan, and W. Utschick, "An Asymptotically MSE-Optimal Estimator Based on Gaussian Mixture Models," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4109–4123, 2022.

[62] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D Multi-Cell Channel Model With Time Evolution for Enabling Virtual Field Trials," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, pp. 3242–3256, 2014.

[63] M. Bjelkanovic, D. Ben Amor, M. Joham, and W. Utschick, "A Low Complexity Rate-Splitting Bilinear Precoder for Massive MIMO," in *WSA & SCC 2023*, 2023, pp. 1–6.

[64] A. Mishra, Y. Mao, C. K. Thomas, L. Sanguinetti, and B. Clerckx, "Mitigating Intra-Cell Pilot Contamination in Massive MIMO: A Rate Splitting Approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3472–3487, 2023.

[65] M. Joham, A. Gründinger, A. Pastore, J. R. Fonollosa, and W. Utschick, "Rate Balancing in the Vector BC with Erroneous CSI at the Receivers," in *2013 47th Annual Conference on Information Sciences and Systems (CISS)*, 2013, pp. 1–6.

[66] D. Ben Amor, F. Strasser, M. Joham, and W. Utschick, "Uplink Downlink Duality for Multi-Cell Massive MISO FDD Systems with per Base Station Power Constraints," in *WSA 2021; 25th International ITG Workshop on Smart Antennas*, 2021, pp. 1–6.

[67] D. Schmidt, M. Joham, R. Hunger, and W. Utschick, "Near Maximum Sum-Rate Non-Zero-Forcing Linear Precoding with Successive User Selection," in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, 2006, pp. 2092–2096.

[68] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[69] R. Hunger, D. Schmidt, M. Joham, A. Schwing, and W. Utschick, "Design of Single-Group Multicasting-Beamformers," in *2007 IEEE International Conference on Communications*, 2007, pp. 2499–2505.

[70] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 2.1," http://cvxr.com/cvx, 2014.

[71] C. Kaulich, M. Joham, and W. Utschick, "Efficient Rate Splitting Method Using Successive Beamforming Techniques," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, 2020, pp. 1–6.

[72] J. Kiefer, "Sequential Minimax Search for a Maximum," *Proceedings of the American Mathematical Society*, vol. 4, no. 3, pp. 502–506, 1953. [Online]. Available: http://www.jstor.org/stable/2032161

[73] H. Joudeh and B. Clerckx, "Sum-Rate Maximization for Linearly Precoded Downlink Multiuser MISO Systems With Partial CSIT: A Rate-Splitting Approach," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4847–4861, 2016.

[74] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, ser. MPS-SIAM Series on Optimization. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2009.

[75] T. Kumar and K. Suresh, "Direct Lagrange Multiplier Updates in Topology Optimization Revisited," 2020.

[76] S. Syed, D. Ben Amor, M. Joham, and W. Utschick, "Bilinear Precoder Based Efficient Rate Splitting Method in FDD Systems," 2024, https://arxiv.org/abs/2411.02242.

[77] Z. Luo, W. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[78] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep Complex Networks," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2018.

[79] S. Liu, J. Guo, and C. Yang, "Multidimensional Graph Neural Networks for Wireless Communications," *IEEE Transactions on Wireless Communications*, vol. 23, no. 4, pp. 3057–3073, 2024.

[80] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[81] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353–4368, 2015.

[82] J. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp. 772–781, 1978.