THÈSE DE DOCTORAT DE
SORBONNE UNIVERSITÉ

Spécialité : Electroniques et Communications

Ecole doctorale : EDITE ED 130

Présentée par

Imène GHAMNIA

Pour obtenir le grade de

DOCTEUR de SORBONNE UNIVERSITE

Rate Balancing Methods for Multi-User MIMO Systems with Perfect or Partial CSIT

---

soutenue le 15 juin 2021, devant le jury composé de :

Prof. Dirk SLOCK	EURECOM, France	Examinateur, Directeur de thèse
Dr. Yi YUAN-WU	Orange Labs, France	Co-encadrant de thèse
Prof. Mérouane DEBBAH	CentraleSupélec, France	Examinateur, Président du Jury
Prof. Ghaya REKAYA	Telecom Paris, France	Rapporteur
Prof. Tharm RATNARAJAH	University of Edinburgh, Royaume-Uni	Rapporteur
Prof. Laura COTTATELLUCCI	Friedrich-Alexander University, Allemagne	Examinateur
Prof. Marios KOUNTOURIS	EURECOM, France	Examinateur
Dr. Raphaël VISOZ	Orange Labs, France	Examinateur

Titre : Méthodes d'équilibrage du débit pour les systèmes MIMO multi-utilisateurs avec connaissance parfaite ou partielle du canal.

Mots-clés : Coordination des interférences intercellulaires, Coordinated Beamforming (CoBF), Multi-User MIMO, équilibrage du débit, connaissance partielle du canal

Résumé : Avec la progression de l'utilisation des smartphones, les modèles de systèmes ont rapidement évolué pour répondre aux besoins croissants en terme de capacité dans les réseaux sans fil. En effet, les progrès technologiques ont été considérables, depuis les communications point à point mono-utilisateur et mono-antenne jusqu'aux réseaux cellulaires multi-cellules et multi-antennes. Depuis la 3G, la technologie MIMO (multiple-input multiple-output) pour les communications sans fil est désormais intégrée aux normes de la large bande sans fil. L'ajout de plusieurs antennes, tant du côté de l'émetteur que du côté du récepteur, permet le multiplexage spatial (c'est-à-dire l'envoi simultané de plusieurs flux de données), qui permet d'augmenter les débits de données, et l'exploitation de la diversité spatiale, qui permet d'améliorer considérablement la qualité des liaisons. MIMO Multi-Utilisateurs (MU) a été un sujet bien étudié dans le domaine des communications sans fil en raison du grand potentiel qu'il offre pour améliorer le débit du système. Par conséquent, différents critères de conception pour les communications MIMO MU ont été étudiés dans la littérature. La plupart des conceptions de liaisons descendantes prennent en compte les problèmes d'optimisation de la capacité totale de tous les utilisateurs. D'autre part, la principale limitation des communications sans fil modernes est l'interférence (intracellulaire et intercellulaire) due à la réutilisation des fréquences. Ainsi, dans un scénario MIMO MU, lors de l'optimisation de l'efficacité globale, l'allocation de puissance se concentre sur les bons canaux, c'est-à-dire que les utilisateurs soumis à une forte interférence (e.g., les utilisateurs en bordure de cellule) sont délaissés. Il en résulte une répartition inéquitable de puissance entre les utilisateurs. Pour pallier ce problème, différentes notions d'équité sont introduites, comme l'équité max-min, l'équité pondérée ou l'équité proportionnelle.

Dans cette thèse, nous nous concentrons sur l'équité max-min pondérée. En particulier, nous étudions le problème de l'équilibrage du débit pondéré par utilisateur dans un système MIMO multi-cellules MU. Nous abordons ce dernier dans le cadre d'une formulation conjointe du problème de beamforming et d'allocation de puissance, visant à satisfaire l'exigence d'équité. Dans la première partie, nous considérons la connaissance parfaite du canal pour résoudre le problème. Dans ce cas, nous maximisons le débit minimum pondéré via i) la dualité liaison montante/descendante et ii) la dualité Lagrangienne. Dans la deuxième partie, nous considérons la connaissance partielle du canal. Nous optimisons le problème d'équilibrage de débit ergodique via i) l'erreur quadratique moyenne pondérée (EQM) en exploitant la relation débit - EQM, et ii) deux approximations du débit estimé comme le débit de signal et de puissance d'interférence estimés (ESIP) au niveau du flux et du signal reçu. Par ailleurs, nous proposons une stratégie d'efficacité énergétique au moyen des approches d'équilibrage des débits proposées.

Title: Rate Balancing Methods for Multi-User MIMO Systems with Perfect or Partial CSIT.

Keywords: Inter-cell interference coordination (ICIC), Coordinated Beamforming (CoBF), MIMO multi-utilisateurs, Rate Balancing, Partial CSIT

Abstract: With the rise in smartphone usage, the system models have rapidly evolved to meet ever-growing needs for capacity in wireless networks. Indeed, there have been large advances in technology, from single-user single-antenna point-to-point communications to multi-cell multi-antenna cellular networks. In fact, multiple-input multiple-output (MIMO) technology for wireless communications is now incorporated into wireless broadband standards since 3G. Adding multiple antennas at both the transmitter and the receiver sides enables spatial multiplexing (i.e. sending multiple data streams simultaneously), which allows to increase data rates, and spatial diversity exploitation, which allows to greatly improve link quality. The multi-user MIMO downlink (so-called Broadcast Channel (BC)) has been a well investigated subject in wireless communications because of the high potential it offers in improving the system throughput. Therefore, different design criteria for multi-user MIMO communication have been investigated in the literature. Most of the downlink designs consider optimization problems w.r.t. the sum-capacity of all users. On the other hand, the major bottleneck of modern wireless communication is the interference (intracell and intercell) due to frequency reuse. Thus, in a multi-user MIMO scenario, when optimizing the overall efficiency, the power allocation is focused on the good channels, i.e., users that are subject to strong interference (e.g. cell-edge users) are neglected. The result is an unfair distribution of rate among users. In order to avoid this effect, different fairness notions have been introduced, like max-min fairness, weighted fairness, or proportional fairness.

In this thesis, we focus on the weighted max-min fairness. In particular, we study the (weighted) user rate balancing problem in a multi-cell multi-user MIMO system. We address this problem by joint beamforming and power allocation optimization, aiming to satisfy the fairness requirements. In the first part, we consider perfect knowledge of the channel to solve the problem. Therein, we maximize the minimum (weighted) rate via *i)* weighted user Mean Square Error (MSE) uplink/downlink duality and *ii)* Lagrangian duality. In the second part, we consider partial knowledge of the channel. We optimize the ergodic rate balancing problem via *i)* weighted expected MSE by exploiting the rate – MSE relation, and *ii)* two approximations of the expected rate as the Expected Signal and Interference Power (ESIP) rate at the stream level and the received signal level. Furthermore, we study the transmit power minimization problem with fixed user-rate targets and provide a strategy exploiting the proposed rate balancing approaches.

"Live as if you were to die tomorrow. Learn as if you were to live forever."

M. Gandhi

Acknowledgments

This PhD was a rich journey for me. A journey that I have not accomplish alone. I would like to express my deepest and sincere thanks to:

... **Dirk Slock**, my supervisor, for his never-ending guidance, inspiring thoughts, boundless help and assistance; I am honored to do my thesis under his supervision and fortunate to learn from his constructive criticism and scientific rigor; to **Yi Yuan-Wu**, my co-supervisor, for always being there to advise me, for her constant technical guidance and kind attitude; to **Raphaël Visoz** and **Ghaya Rekaya**, my jury members who have closely followed the progress of my thesis since the very first year, for their positive comments to my work; to **Mérouane Debbah**, for his constant support, enthusiasm and encouragement over the years, and for taking part of the jury members; to **Tharm Ratnarajah**, **Laura Cottatellucci** and **Marios Kountouris**, who complete the jury members, for the time they will spend to evaluate my works.

... all the members of Orange Labs RIDE team, namely, the team manager **Pierre Dubois**, for providing an adequate working environment and support; my dear colleagues/friends, **Romain**, **Rita**, **Nathalie**, **Ayat**, **Ali**, **Fred**, **Thierry**, **Ahmad**, **Meriem**, **Stefan**, and **Amel**, for the constructive exchanges and the memorable time we spent together; to my colleagues in EURECOM, namely, **Christo**, **Roya**, **Flavio** and **Chandan** for being supportive.

... all my friends, who always believed in me, for their encouragement and support, especially, **Hakim**, **Samira**, **Imen**, **Wissem**, **Amel**, **Houda**, **Nassima**, **Amina**, and **Hayet**; to my second family here in France, namely, **tatie Zoubida**, **Yasmine**, **Youssra**, **Karim**, **Yasser**, **Mira**, **Yassmine**, and **Wissal**; to all my friends from **TPMP group** for all the fun and positive vibes we shared together.

Last but not least, to my parents, my everything, for their advice, encouragement and moral support throughout my existence; to my brother **Adel** and his little family, for always motivating me and being there for me; to my wonderful sisters **Chahrazed** and **Karima**, for their good influence and positiveness; to all my relatives.

Thank you all.

Contents

Acronyms

BC	Broadcast Channel.
BF	Beamforming.
BS	Base Station.
CSI	Channel State Information.
CSIT	Channel State Information at the Transmitter.
DL	Downlink.
DPC	Dirty Paper Coding.
EMSE	Expected Mean Square Error.
ESIP	Expected Signal and Interference Power.
GSM	Global System for Mobile communication.
i.i.d.	independent and identically distributed.
IBC	Interfering Broadcast Channel.
IFC	Interfering Channel.
LMMSE	Linear Minimum Mean Square Error.

ma-MIMO	Massive MIMO.
MAC	Medium Access Control.
MIMO	Multiple-Input Multiple-Output.
MISO	Multiple-Input Single-Output.
MMR	Max-Min Rate.
MMSE	Minimum Mean Square Error.
MSE	Mean Square Error.
MU	Multi-User.
PDP	Power Delay Profile.
PF	Proportional Fair.
PM	power minimization.
QoE	Quality-of-Experience.
QoS	Quality-of-Service.
R-ESIP	Received signal level ESIP.
Rx	Receive.
S-ESIP	Stream level ESIP.
SE	Spectral Efficiency.
SIMO	Single-Input Multiple-Output.
SINR	Signal-to-Interference-plus-Noise Ratio.
SOCP	Second Order Cone Programming.

TDD	Time-Division Duplexing.
Tx	Transmit.
UE	User Equipement.
UL	Uplink.
ULA	Uniform Linear Array.
WEMSE	Weighted Expected Mean Square Error.
WSMSE	Weighted Sum Mean Square Error.
WSR	Weighted Sum Rate.

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Motivation and Related Works

1.1 Motivation

In the past twenty years, we have seen a tremendous growth in the demand for wireless data rate, and this trend is predicted to continue in the future. Several methods are proposed to satisfy the ever growing demand of data rates, including the following

- cell densification: putting the access points closer one to another;
- increasing bandwidth for transmission: the introduction of new high frequency bands in 5G $\geq 6Ghz$, allowing the introduction of millimeter-waves (mmWaves) communications;
- increasing the Spectral Efficiency (SE).

Massive Multiple-Input Multiple-Output (MIMO) has become a key solution to increase the spectral efficiency of wireless cellular systems [1]. In fact, MIMO technology for wireless communications is now incorporated into wireless broadband standards since 3G. The basic idea behind MIMO technology is that the more antennas the transmitter and the receiver are equipped with, the more the available signal paths between them will appear, the better the performance in terms of data rate and energy efficiency the system will get [2–4].

In downlink communications, the Base Station (BS) with multiple transmit antennas serves multiple users within the same time and frequency resource block. Therefore, proper resource allocation is needed to fully harvest the gain in spectral and energy efficiency; for example: user scheduling, subcarrier allocation, **power allocation and precoder (receiver) design**. The latter represents the most important aspect to enhance the performance of the system in the physical layer, and can be combined with frequency subcarrier allocation and user scheduling to further boost the performance.

Power allocation optimizatiogn in wireless networks has been an important research problem for decades, dating back to single-antenna wireless systems. In fact, power control schemes have been studied with different utility functions in the literature. In particular, the power minimization problem with target Signal-to-Interference-plus-Noise Ratio (SINR) constraints. Actually, the target SINRs are usually set according to the application. For example in Global System for Mobile communication (GSM) the target SINRs are set to ensure that the quality of a voice call is within an acceptable range. However, with the evolution of the wireless networks providing mobile data, the target SINRs are no longer easy to determine in a way to satisfy every user. Therefore, the SINRs and the power allocation have to be found simultaneously.

The power allocation problem can be formulated as a maximization of some utility in terms of data rate, by interchanging the SINRs and SE. Depending on the chosen utility function, we can achieve different points on the Pareto optimal boundary. In other words, we cannot increase the rate of any of the actif users without lowering the rate of the other users [5]. The two most commonly used utility functions are *i*) weighted max-min fairness, or balancing problem and *ii*) weighted sum problem. Let us consider the following example:
A base station transmits data to User Equipement (UE) under constant radio conditions that can be classified as either *good* or *bad*. The ones suffering from bad radio conditions require more radio resources (e.g., bandwidth, transmission power) than the others to get the same bit rate. With weighted sum utility maximization, the overall bit rate is maximized by allocating all resources to those mobiles having a good radio channel. However, such an allocation is not *fair* to UEs with bad radio conditions, see Figure 1.1. In this thesis, we focus our interest on the latters by considering max-min fairness optimization problem [6], which is used to provide the same quality-of-service for all users according to their priorities and make this value as large as possible. The weighted max-min fairness problem, also called balancing problem, can be expressed for different objectives other than SINR, such as the MSE and user rate.

Apart from power control, a BS, when equipped with multiple antennas, creates more degrees of freedom for resource allocation. By transmit and receive beamforming, the signal of a particular user can be strengthened which brings array gain, meanwhile interference can be suppressed [7]. Therefore the beamformer design can be very beneficial in increasing the users data rates. The optimization of the beamformer and power allocation simultaneously is thus a problem that attracts a lot of interest. In fact, with a fixed beamformer, the power allocation problem reamins the same as in power allocation problem which does not consider beamformers. However, the beamformer itself contains the power allocation parameters, therefore iterative solutions are proposed to solve the problem by first fixing the power allocation and optimize the beamformer [8], then update the power allocation with the known methods.

In this thesis, we focus on linear precoders for multi-cell Multi-User (MU) MIMO scenarios w.r.t. rate balancing. The multi-user MIMO for Downlink (DL) has been a well investigated subject in wireless communications because of the high potential it offers in improving the system throughput. Information theory has shown that the capacity of MU-MIMO channels could be achieved through Dirty Paper Coding (DPC) [9–11]. However, DPC is difficult to implement and computationally complex. Suboptimal linear

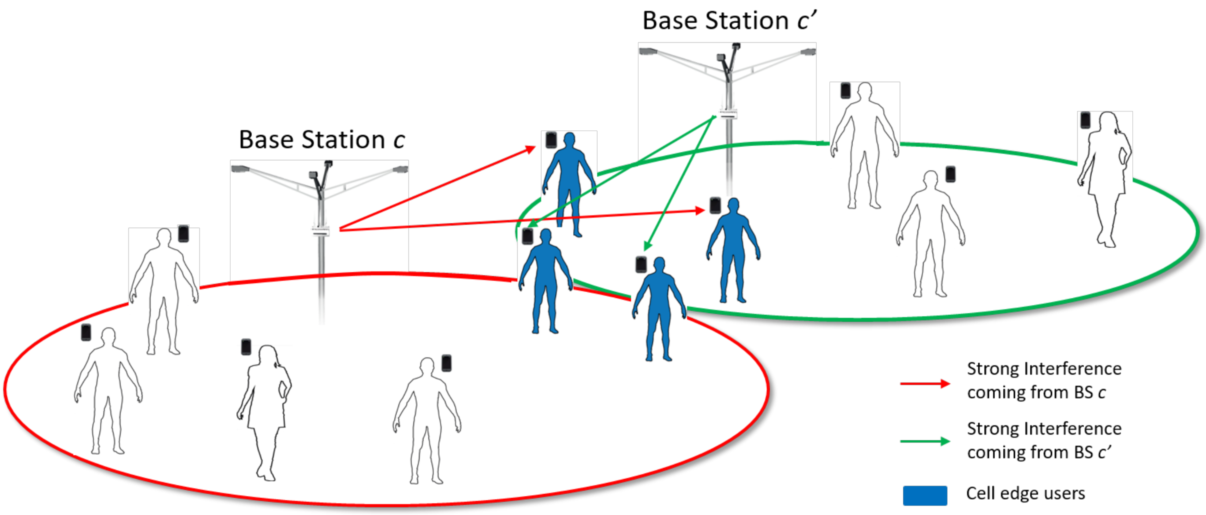

Figure 1.1: Cell edge users suffering from strong intercell interference.

Beamforming (BF) algorithms exist and can be divided into two main categories: the iterative [12–16] and the closed form (CF) solutions [17–21]. In this work, we focus on user rate balancing in a way to maximize the minimum per user (weighted) rate in the network. This balancing problem is studied in [22] without providing an explicit precoder design. We provide here a solution via the relation between user rate (summed over its streams) and a Weighted Sum Mean Square Error (WSMSE). But also another ingredient is required: the exploitation of scale factor that can be freely chosen in the weights for the weighted rate balancing. User-wise rate balancing outperforms user-wise MSE balancing or streamwise rate (or MSE/SINR) balancing when the streams of any MIMO user are quite unbalanced.

1.2 Related Works

Joint beamformer design and power allocation is a problem that attracts a lot of interest.

1.2.1 Joint Design for max WSR

One important and commonly used utility function is the weighted sum performance. This problem is proposed to maximize the total system throughput, while considering weights to provide some fairness between the different users. The problem is expressed as maximizing the Weighted Sum Rate (WSR) [23–25]. The latter has been extensively studied, in MIMO downlink, via the corresponding MSE minimization problem [26–30]. In fact, using the well-known inverse relation between the SINR and MSE [31], the rate optimization problem can be equivalently formulated as an MSE minimization problem. In [26], WSR maximization problem was considered for MIMO Broadcast Channel (BC) via weighted Minimum Mean Square Error (MMSE). Therein, an alternating optimization algorithm, based on well-known transmit/receive MMSE designs, was proposed for finding a local weighted sum-rate optimum. Actually, while the WSMSE minimization problem is not jointly convex with respect to the transmit and receive beamformers, it is convex when either the transmitters or receivers are fixed. Performing alternating optimization of the transmit and receive beamformers is used to exploit the latter biconvex structure. An extension of [26] to interfering BC was handled in [27] and further in [28] where practical signaling issues for the interfering BC were studied.

The sum-MSE (unweighted) minimization problem was considered in [32] to study the optimality conditions of the uplink-downlink power allocation in detail. Various alternative convex formulations using techniques of linear matrix inequalities (LMIs) and Second Order Cone Programming (SOCP) were developed in [33] to minimize the sum-MSE under the total power constraint. However, the obtained results do not hold in general for weighted sum-MSE optimization. This prevents applying these techniques straightforwardly to, e.g., WMMSE methods. Still, the MSE duality has been successfully exploited also in the WSR maximization problem [34,35]. Another approach was proposed in [36] to maximize the WSR for MIMO BC. Therein, the MSE minimization approach via first order approximation of the non-convex objective function is used to design the beamformer. Then, the non-convex weighted MSE minimization problem is solved using a successive convex approximation of the Lagrangian function.

1.2.2 Joint Design for Balancing

Most of balancing optimization problems are non-convex and can not be solved directly. Despite that, several works over the litterature have developped optimal solutions [37,38]. For instance, [39] solved the max-min problem by a sequence of SOCP. Also, [40] showed that a semidefinite relaxation is tight for the problem, and the optimal solution can be constructed from the solution to a reformulated semidefinite program. In [41], the authors proposed an algorithm based on fixed-point that alternates between power update and beamformer updates, and the nonlinear Perron-Frobenius theory was applied to prove the convergence of the algorithm.

Another way to solve balancing optimization problems is to convert the problem from the DL channel to its' equivalent Uplink (UL) channel, by exploiting the uplink-downlink *duality*. Doing so, the transformed problem has better mathematical structure and convexity in the uplink, thus, the computational complexity of the original problem can be reduced [11], [8]. The uplink-downlink duality has been widely used to design optimal transmit and receive filters that ensure faireness requirements w.r.t the SINR [8, 42–46], the MSE [47–49], and the user rate [27].

Actually, optimization problems associated with MIMO systems are more complicated in the downlink channels, due to the joint design of transmit-receive filters and the coupled structure of the transmit filter along with power allocation. Most of these problems are non-convex and can not be solved directly. One way to overcome these difficulties is to convert the problem from the downlink channel to its' equivalent uplink channel, by exploiting the uplink-downlink *duality*. Doing so, the transformed problem has better mathematical structure and convexity in the uplink, thus, the computational complexity of the original problem can be reduced [11], [8].

The objective being to equalize all user SINRs, the SINR balancing problem is of particular interest because it is directly related to common performance measures like system capacity and bit error rates. Maximizing the minimum user SINR in the uplink can be done straightforwardly since the beamformers can be optimized individually and SINRs are only coupled by the users' transmit powers. In contrast, downlink optimization is generally a nontrivial task because the user SINRs depend on all optimization variables and have to be optimized jointly. Downlink transmitter optimization for single antenna receivers with constraint of the total transmit power is comprehensively studied in [8] and [44] where algorithmic solutions for maximizing the minimal user SINR are proposed. This SINR balancing technique has been extended to an underlay cognitive radio networks with transmit power and interference constraints in [45], [46].

Another well-known duality is the stream-wise MSE duality where it has been shown that the same MSE values are achievable in the downlink and the uplink with the same transmit power constraint. This MSE duality has been exploited to solve various minimum mean square error (MMSE) based optimization problems [47], [48]. In [49], three levels of MSE dualities have been established between MIMO BC and MIMO Medium Access Control (MAC) with the same transmit power constraint and these dualities have been exploited to reduce the computational complexity of the sum-MSE and weighted sum-MSE minimization problems in a MIMO BC. In [48], an iterative algorithm has been proposed to balance the capacity between users in a multi-user MIMO system by using stream-wise MSE duality but with a single transmit power constraint. However, this algorithm can not be applied to solve capacity balancing problem with multiple linear transmit covariance constraints.

1.3 Thesis Outline

This thesis contains four parts. Part I being the introduction, **Chapter 2** follows to provide the definition of the considered problem and some useful theoretical background.

In Part II, we focus on user rate balancing problem assuming perfect knowledge of the channel state information at the transmitter. In particular:

In Chapter 3, we consider the user rate balancing problem for a multi-user MIMO single cell (broadcast channel), under a total transmit power constraint. We use the rate - MSE relation to transform the problem into minimizing the maximum weighted-matrix MSE. The latter allows us to enable MSE duality between the downlink and its dual uplink channels. The results are extended to interfering broadcast channel, i.e., the multi-cell multi-user case, with total sum power constraint.

In Chapter 4, we study the maximization of the minimum user rate for multi-cell

MIMO system, considering per cell power constraints. In particular, we exploit the rate - MSE relation, formulate the balancing operation as constraints leading to Lagrangians in optimization duality, allowing to transform rate balancing into weighted MSE minimization with Perron Frobenius theory. The Lagrange multipliers for the multiple power constraints can be formulated as a single weighted power constraint in which the weighting can be optimized via subgradient methods, leading to the satisfaction with equality of all power constraints.

In Chapter 5, we address the total transmit power minimization problem subject to per user rate targets. In particular, we exploit user rate balancing optimization to derive an iterative solution based for the power minimization problem. Actually, minimizing the total transmit power problem w.r.t. individual rate requirements is a variation of user rate balancing problem w.r.t. total transmit power.

In the third part of the thesis, namely in Part III, we deal with user rate balancing under a more realistic channel model with partial CSIT. In particular:

In Chapter 6, we give an overview of balancing works with partial knowledge of the channel state information at the transmitter. Then, we detail the considered channel model.

In Chapter 7, we study the ergodic user rate balancing, which corresponds to maximizing the minimum (weighted) per user expected rate in the network. We consider a multi-cell multi-user MIMO system with per cell power constraints and partial CSIT. The latter combines both channel estimates and channel (error) covariance information. In particular, we introduce a novel extension to partial CSIT of Chapter 4, which considers the rate - MSE link to reformulate the user rate balancing problem into WMSE balancing problem. In partial CSIT, the latter becomes maximizing an expected rate lower bound in terms of expected MSE, leading to a Weighted Expected MSE (WEMSE) balancing problem.

In Chapter 8, we exploit a better approximation of the expected rate as the Expected Signal and Interference Power (ESIP) rate, based on an original minorizer for every individual rate term. The latter minorizer is appart from the global criterion, thus, the transmit beamformers can be optimigzed in parallel, which is interesting for distributed algorithms. Besides, this ESIPrate approach does not require the introduction of receive beamformers (no processing needed at the user side). We study the ESIPrate approach within two approximations: *i)* Received signal level ESIP (R-ESIP) and *ii)* Stream level ESIP (S-ESIP). Also, we optimize the total transmit power minimization.

Finally, Part IV includes our conclusions and future works related to work presented in this thesis

1.4 Contributions

The results obtained during the course of this PhD are published in the following

- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Transmit filter optimization Methods for Multi-cell MIMO user rate balancing with partial CSIT,” journal paper, ongoing.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Multi-Cell MIMO Power Minimization via Rate Balancing with Partial CSIT,” 2021 IEEE Global Communications Conference: Wireless Communications, submitted.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Multi-Cell MIMO User Rate Balancing with Power Method Generalized Eigenvectors,” 2021 IEEE Global Communications Conference: Communication Theory, submitted.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Multi-cell MIMO user rate balancing with partial CSIT: SESIP vs. RESIP,” in 22nd IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2021), submitted.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Multi-cell MIMO user rate balancing with partial CSIT,” in IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), Online, 2021.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”MIMO User Rate Balancing In Multicell Networks with Per Cell Power Constraints,” in IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 2020.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Rate Balancing for multi-user Multicell Downlink MIMO Systems,” in 27th European Signal Processing Conference, (EUSIPCO), Coruna, Spain, Sept 2019.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu, ”Multi-User MIMO Max-Min user rate via Weighted MSE Balancing,” in 16th International Symposium on Wireless Communication Systems (ISWCS), Oulu, Finland, Aug 2019.
- I. Ghamnia, D. Slock, and Y. Yuan-Wu,”Rate Balancing for multi-user MIMO Systems,” in IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, Jul 2019.

Chapter 2

Problem Definition and Background Theory

2.1 Overview

In wireless communication networks, transmit power is one of the most important degrees of freedom to control the rate performance. Actually, interference is the limiting factor for most cellular communication links, due to the broadcast nature of wireless communications. To overcome this, allocating properly the transmit power from the BS and the mobile terminals is of vast importance. This problem has been extensively studied in the past, for both the uplink and downlink transmission. In this chapter, we review the evolution of power allocation problem in terms of max-min fairness utility from its original form to the new formulations we developed in this thesis.

The reminder of the chapter is organized as follows. In Section 2.2, we provide a generalized formulation of the studied problem. Then, we present in Section 2.3 some theoretical material used in this thesis; namely, the UL-DL duality, Perron Frobenius theory, the link between user rate and MSE expressions, and, WSR maximization approach. Finally, we conclude in Section 2.4.

2.2 Problem Definition

The weighted max-min fairness problem is used to provide the same Quality-of-Service (QoS) to all users in the cell. We aim to serve all active users with equal weighted SE according to their respective priorities and optimize this value as large as possible. In MISO system, the considered max-min fairness is w.r.t. user rates $r_k = \log(1 + \mathrm{SINR}_k)$. We can then write the objective as a maximization of $\min_k r_k$. For conventional networks, the weighted max-min fairness problem can be formulated as follows

$$\begin{aligned} \max_{\{p_k\},t} \quad & t \\ \text{s.t.} \quad & p_k \leq P_k, \ \ \forall k, \\ & p_k \geq 0 \ \ \forall k, \\ & r_k \geq t \ \ \forall k, \end{aligned} \tag{2.1}$$

where p_k denotes the power for user $1 \leq k \leq K$, and P_k the power constraint.

For a fixed t, we can solve the problem (2.1) using methods for power minimization problem by omitting the power constraints. Then, the optimal t can be found via bisection search such that the power constraints are satisfied. Other advanced methods are possible to solve (2.1) efficiently in a distributed manner. For instance the Perron-Frobenius theory can applied in the case that effect of noise is ignored [50]. Also, (2.1) can be solved using non-linear Perron-Frobenius theory [51] and the FastLipschitz optimization approach [52, 53].

With the introduction of multiple antennas at the transmitter, the optimal joint beamformer design and power allocation problem can be formulated as

$$\begin{aligned} \max_{\{\boldsymbol{g}_k\},\{p_k\}} \ \min_k \quad & r_k \\ \text{s.t.} \quad & p_k \leq P_k, \ \ \forall k, \\ & p_k \geq 0 \ \ \forall k, \end{aligned} \tag{2.2}$$

where $\boldsymbol{g}_k$ is the transmit beamformer for user k and user rate is now defined as follows

$$r_k = \log\big(1 + \frac{p_k \mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \mathbf{g}_k}{\sum_{l \neq k} p_l \mathbf{g}_l^H \mathbf{h}_{kl}^H \mathbf{h}_{kl} \mathbf{g}_l + \sigma^2},\big) \tag{2.3}$$

where the channel from BS to user k is denoted as $\boldsymbol{h}_k$, and σ^2 is the noise variance.

Now, consider a multi-user MIMO system, the user rate balancing problem becomes as follows

$$\begin{aligned} \max_{\{\boldsymbol{G}_k\},\{\boldsymbol{P}_k\}} \ \min_k \quad & r_k \\ \text{s.t.} \quad & \mathrm{tr}(\boldsymbol{P}_k) \leq P_k \ \ \forall k \end{aligned} \tag{2.4}$$

where r_k is the kth user-rate

$$r_k = \ln\det\Big(\boldsymbol{I}+\boldsymbol{H}_k\boldsymbol{G}_k\boldsymbol{P}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\big(\sigma_n^2\boldsymbol{I}+\sum_{j\neq k}\boldsymbol{H}_k\boldsymbol{G}_j\boldsymbol{P}_j\boldsymbol{G}_j^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\big)^{-1}\Big) \tag{2.5}$$

and $\boldsymbol{G}_k$ and $\boldsymbol{P}_k$ denote, respectively, the transmit beamforming matrix and the diagonal non-negative matrix of transmission stream powers for user k.

Power control schemes for max-min fairness problems in general [54–60], and in particular for the one defined in (2.4) w.r.t. user rate, provide fairly the same quality of service for all users, which is a highly desirable feature in future systems. However, from a network-wide point of view, optimizing this balancing problem leads to a scalability issue since the performance is limited by the weakest user. In fact, for a considerable number of cells and active users, the probability of having a user with an extremely poor channel gets higher due to shadow fading. Thus, all users in the network would suffer from the weak channel of the worst user.

We want to provide fairness to the weak users in the network, without penalizing all active users. For that, we formulate the user balancing problem w.r.t. some user weight r_k°, rewriting (2.4) as follows

$$\begin{aligned} \max_{\{\boldsymbol{G}_k\},\{\boldsymbol{P}_k\}} \ \min_{k} \ & r_k/r_k^\circ \\ \text{s.t.} \ & \boldsymbol{P}_k \leq P_k \forall k \end{aligned} \tag{2.6}$$

The introduction of the weights r_k° to the balancing problem changes the optimization from providing the same rate to providing the same *relative* rate to all users. These weights can be considered as user priorities, referring to the type of the requested service, or can be exploited to assess the achievable user rate during time. In practical networks, the latters should be updated, after every scheduling decision, according to the channels quality and former achieved rates. In this thesis, however, we assume the priorities r_k° are fixed and/or determined from higher layers of the network, then we provide solutions to the per user (weighted) rate balancing problem.

2.3 Theoretical Background

Precoding (or digital beamforming) techniques are processing techniques that exploits transmit diversity in order to transmit one or multiple spatially directive signals. The precoder matrix represents a function of the estimated channel such that a directive signal (or a beam) is processed before transmission in a way to cancel the interference from other user's signals. Downlink communications within one cell (one base station) are refered to as Broadcast Channel (BC) [10, 61]: the BS transmits independent signals to its uncoordinated receivers. When considering multi-cell multi-user network, we are in an Interfering Broadcast Channel (IBC) where receivers suffer not only by the intra-cell, but also by the inter-cell interference [62–65]. Another model is considered in the literature: the Interfering Channel (IFC), wherein multiple BSs are serving one UE each [38, 66–68].

Let us consider an IFC (Figure 2.1), with K pairs of multiantenna BSs and single antenna User Equipment (UE). Each BS b_k is equipped with M_{b_k} antennas. The Transmit (Tx) beamformer and Receive (Rx) filter applied at the k-th user in DL and UL transmissions are denoted as $\mathbf{g}_k$ and $\tilde{\mathbf{g}}_k$, respectively. The Rx signal at the k-th UE in the DL phase, and the output of Rx filter at the b_k-th BS in the UL phase are expressed, respectively, as follows

$$y_k = \mathbf{h}_{kb_k}\mathbf{g}_k s_k + \sum_{\substack{l=1\\l\neq k}}^{K} \mathbf{h}_{kb_l}\mathbf{g}_l s_l + n_k \qquad \tilde{r}_{b_k} = \tilde{\mathbf{g}}_k \tilde{\mathbf{h}}_{kb_k}\tilde{s}_k + \sum_{\substack{l=1\\l\neq k}}^{K} \tilde{\mathbf{g}}_k \tilde{\mathbf{h}}_{kb_l}\tilde{s}_l + \tilde{\mathbf{g}}_k \tilde{n}_k$$

where $\mathbf{h}_{kb_l}$ and $\tilde{\mathbf{h}}_{b_l k}$ refer to the channel between the transmitting BS b_l and UE k in DL communication and the transmitting user k and BS b_l in the UL, respectively. s_k is the transmitted symbol and n_k represents the additive noise. We refer to the respective quantities in the UL by $\tilde{(\cdot)}$.

Hereafter, we provide some useful theoretical background to ensure good comprehension to interested readers, who may not be experts in the topics discussed here.

2.3.1 UL/DL Duality

In the following, we introduce the UL/DL duality for the aforedescribed Multiple-Input Single-Output (MISO) DL IFC.

MISO DL IFC

The SINR for the DL channel is expressed as

$$\mathrm{SINR}_k^{DL} = \frac{p_k \mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk}\mathbf{g}_k}{\sum_{l\neq k} p_l \mathbf{g}_l^H \mathbf{h}_{kl}^H \mathbf{h}_{kl}\mathbf{g}_l + \sigma^2} \tag{2.7}$$

with p_k being the Tx power allocated to the k-th user. It can be observed that the downlink SINRs (2.7) are coupled both by the transmission powers and the beamforming vectors, which makes a direct optimization very difficult.

Imposing a set of DL SINR constraints at each mobile station: $\mathrm{SINR}_k^{DL} = \gamma_k$, we obtain in matrix notation the following

$$\mathbf{\Phi p} + \boldsymbol{\sigma} = \mathbf{D p} \tag{2.8}$$

with:

$$[\mathbf{\Phi}]_{ij} = \begin{cases} \mathbf{g}_j^H \mathbf{h}_{ib_j}^H \mathbf{h}_{ib_j}\mathbf{g}_j = |\mathbf{h}_{ib_j}\mathbf{g}_j|^2, & j \neq i \\ 0, & j = i \end{cases}$$

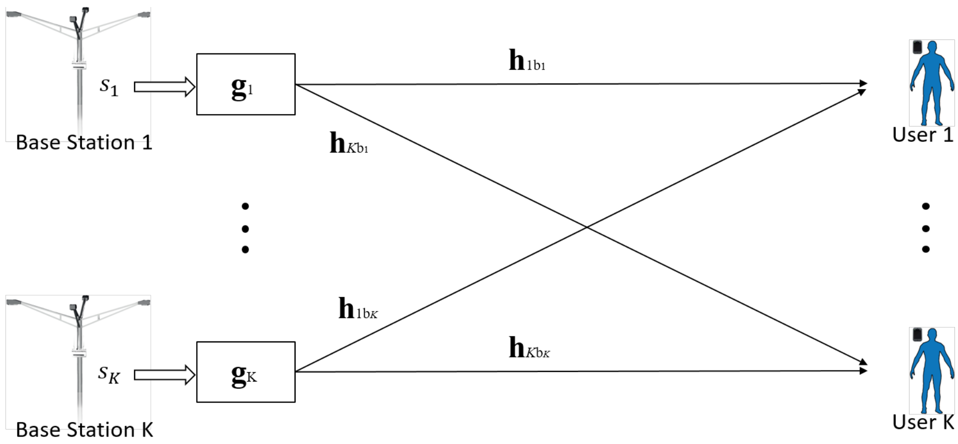

Figure 2.1: MISO Interfering Channel

$$\mathbf{D} = \mathrm{diag}\{\frac{|\mathbf{h}_{1b_1}\mathbf{g}_1|^2}{\gamma_1}, \ldots, \frac{|\mathbf{h}_{Kb_K}\mathbf{g}_K|^2}{\gamma_K}\}\ , \quad \boldsymbol{\sigma} = \sigma^2\mathbf{1}\,.$$

We can determine the TX power solving (2.8) w.r.t. $\mathbf{p}$ obtaining

$$\mathbf{p} = (\mathbf{D} - \mathbf{\Phi})^{-1}\boldsymbol{\sigma} \tag{2.9}$$

Feasible ($\mathbf{p} > 0$) if $\max_i |\lambda_i(\mathbf{D}^{-1}\mathbf{\Phi})| < 1$, where $\lambda_i(\boldsymbol{X})$ denotes the eigenvalue λ_i of matrix $\boldsymbol{X}$.

SIMO UL IFC

Now consider the UL scenario, with the same targets γ_k, see Figure 2.2 the Single-Input Multiple-Output (SIMO) IFC.

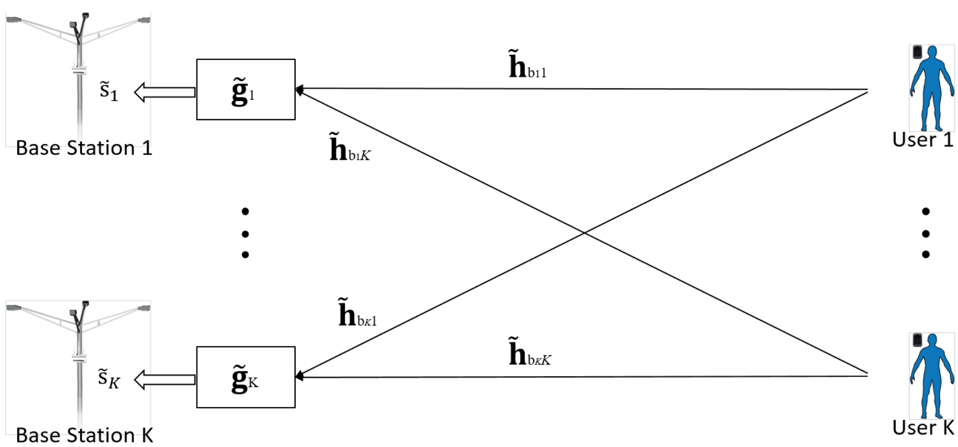

Figure 2.2: SIMO Interfering Channel

We assume in the following that the UL channel is the reciprocal of the DL one, i.e., both channels are described by the same covariances or $\tilde{\mathbf{h}}_{b_i j} = \mathbf{h}^H_{j b_i}$ and $\tilde{\mathbf{g}}_i = \mathbf{g}^H_i$. This approach is justified for Time-Division Duplexing (TDD) systems, which use the same carrier frequency for uplink and downlink. However, the following results are not restricted to TDD systems. The assumption of channel reciprocity only serves the purpose of comparing uplink and downlink achievable regions. The UL/DL duality will be introduced, afterward, between the actual DL and a virtual UL channel.

The SINR for the UL channel can be written as follows

$$\mathrm{SINR}_k^{UL} = \frac{q_k \mathbf{g}_k^H \mathbf{h}_{kb_k}^H \mathbf{h}_{kb_k} \mathbf{g}_k}{\mathbf{g}_k^H (\sum_{l \neq k} q_l \mathbf{h}_{lb_k}^H \mathbf{h}_{lb_k} + \sigma^2 \mathbf{I}) \mathbf{g}_k} \tag{2.10}$$

which represents the Rayleigh quotient, with q_k denoting the UL transmit power at UE k. We can see that these uplink SINRs are only coupled by the transmission powers. For a given power allocation, the beamformers which maximize (2.10) are well known [69].

In fact, the optimal UL Rx beamformer, $\hat{\mathbf{g}}_k$, is obtained as follows

$$\hat{\mathbf{g}}_k = \arg\max_{\mathbf{g}_k} \mathrm{SINR}_k^{UL} = \arg\max_{\mathbf{g}_k} \frac{q_k \mathbf{g}_k^H \mathbf{h}_{kb_k}^H \mathbf{h}_{kb_k} \mathbf{g}_k}{\mathbf{g}_k^H \boldsymbol{Q}_k(\boldsymbol{q}) \mathbf{g}_k} \tag{2.11}$$

where $\boldsymbol{Q}_k(\boldsymbol{q}) = \sum_{l \neq k} [\boldsymbol{q}]_l \mathbf{h}_{lb_k}^H \mathbf{h}_{lb_k} + \sigma^2 \mathbf{I}$, and collecting the individual UL transmit powers in the vector $\boldsymbol{q} = [q_1 \dots q_K]$. The matrices $\boldsymbol{Q}_k$ are nonsingular and symmetric, thus (2.11) is solved by the dominant generalized eigenvectors of the matrix pairs $(q_k \mathbf{h}_{kb_k}^H \mathbf{h}_{kb_k}, \, \boldsymbol{Q}_k(\boldsymbol{q}))$, i.e., $\hat{\mathbf{g}}_k = V_{max}(q_k \mathbf{h}_{kb_k}^H \mathbf{h}_{kb_k}, \, \boldsymbol{Q}_k(\boldsymbol{q})), 1 \leq k \leq K$. In case of $\mathrm{rank}\{\mathbf{h}_{kb_k}^H \mathbf{h}_{kb_k}\} = 1, 1 \leq k \leq K$, this is equivalent to the scaled MMSE beamforming solution [69]

$$\hat{\mathbf{g}}_k z = (\sum_{l \neq k} q_l \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \sigma^2 \mathbf{I})^{-1} \mathbf{h}_{kb_k}^H.$$

Generalized Eigenvectors

- Consider $N \times N$ matrices $\mathbf{A} = \mathbf{A}^H \geq 0$, $\mathbf{B} = \mathbf{B}^H > 0$.
- Generalized eigen vectors V_i and eigen values λ_i

  $$\mathbf{A}\, V_i = \lambda_i\, \mathbf{B}\, V_i \;,\; \mathbf{B}^{-1} \mathbf{A}\, V_i = \lambda_i\, V_i \;,\; \det(\mathbf{A} - \lambda\, \mathbf{B}) = 0$$

  $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0$, $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1, \dots, \lambda_N\}$.
  Note that the V_i in $\boldsymbol{V} = [V_1 \cdots V_N]$ are not orthogonal.
- Let $\mathbf{B} = \mathbf{B}^{1/2} \mathbf{B}^{H/2}$ and ($\boldsymbol{U}^H \boldsymbol{U} = \boldsymbol{I}$, $\boldsymbol{\Gamma}$ diagonal) eigen decomposition $\mathbf{B}^{-1/2} \mathbf{A}\, \mathbf{B}^{-H/2} = \boldsymbol{U} \boldsymbol{\Gamma} \boldsymbol{U}^H$. Then

  $$\boldsymbol{V} = \mathbf{B}^{-H/2} \boldsymbol{U} \;,\; \boldsymbol{V}^H \mathbf{A}\, \boldsymbol{V} = \boldsymbol{\Gamma} \;,\; \boldsymbol{V}^H \mathbf{B}\, \boldsymbol{V} = \boldsymbol{\Xi} = \boldsymbol{I} \;,\; \mathbf{A}\, \boldsymbol{V} = \mathbf{B}\, \boldsymbol{V}\, \boldsymbol{\Gamma}$$

  so $\boldsymbol{\Lambda} = \boldsymbol{\Gamma}$. If we normalize $diag(\boldsymbol{V}^H \boldsymbol{V}) = \boldsymbol{I}$, then $\boldsymbol{\Xi} \neq \boldsymbol{I}$ and $\boldsymbol{\Lambda} = \boldsymbol{\Xi}^{-1} \boldsymbol{\Gamma}$. The non-singular $\boldsymbol{V}$ *simultaneously diagonalizes* $\mathbf{A}$, $\mathbf{B}$.

- Rayleigh quotient is defined as follows

$$V_1 = V_{max}(\mathbf{A}, \mathbf{B}) = \arg\max_{V} \frac{V^H \mathbf{A}\, V}{V^H \mathbf{B}\, V}\ ,\ \lambda_1 = \mathrm{eig}_{max}(\mathbf{A}, \mathbf{B}) = \max_{V} \frac{V^H \mathbf{A}\, V}{V^H \mathbf{B}\, V}$$

Duality

Imposing the same SINR constraints also in the UL, i.e., $SINR_k^{UL} = SINR_k^{DL} = \gamma_k$ it is possible to rewrite that constraints as

$$\tilde{\mathbf{\Phi}}\mathbf{q} + \boldsymbol{\sigma} = \mathbf{D}\mathbf{q}$$

with

$$[\tilde{\mathbf{\Phi}}]_{ij} = \begin{cases} \mathbf{g}_i^H \mathbf{h}_{jb_i}^H \mathbf{h}_{jb_i} \mathbf{g}_i = |\mathbf{h}_{ji}\mathbf{g}_i|^2, & j \neq i \\ 0, & j = i \end{cases}$$

$$\mathbf{D} = \mathrm{diag}\{\frac{|\mathbf{h}_{1b_1}\mathbf{g}_1|^2}{\gamma_1}, \ldots, \frac{|\mathbf{h}_{Kb_K}\mathbf{g}_K|^2}{\gamma_K}\}.$$

The power vector can be found as follows

$$\mathbf{q} = (\mathbf{D} - \tilde{\mathbf{\Phi}})^{-1}\boldsymbol{\sigma} \tag{2.12}$$

Comparing the definition we can see that $\tilde{\mathbf{\Phi}} = \mathbf{\Phi}^T$. This implies that there exists a duality relationship between the DL MISO and UL SIMO IFCs.

We can extend the results for UL-DL duality for MAC/BC [8] to the MISO/SIMO IFC, which state that targets $\gamma_1, \ldots, \gamma_K$ are jointly feasible in UL and DL if and only if the spectral radius ρ of the weighted coupling matrix satisfies $\rho(\mathbf{D}^{-1}\mathbf{\Phi}) < 1$. Also, both UL and DL have the same SINR feasible region under a sum-power constraint, i.e., target SINRs are feasible in the DL if and only if the same targets are feasible in the UL:

$$\sum_i q_i = \mathbf{1}^T\mathbf{q} = \sigma^2\mathbf{1}^T(\mathbf{D} - \mathbf{\Phi}^T)^{-1}\mathbf{1} = \sigma^2\mathbf{1}^T(\mathbf{D} - \mathbf{\Phi})^{-1}\mathbf{1} = \sum_i p_i \tag{2.13}$$

Using these results it is possible to extend some BF design techniques used in the BC [8] to the MISO IFC: *i)* Max-Min SINR (SINR Balancing) and *ii)* Power minimization under SINR constraints.

2.3.2 Perron Frobenius Theory

We shall now consider the following weighted SINR balancing formulation

$$\max_{\mathbf{p}} \min_{k} \frac{\mathrm{SINR}_k^{DL}}{\gamma_k} \Leftrightarrow \min_{\mathbf{p}} \max_{k} \frac{\gamma_k}{\mathrm{SINR}_k^{DL}} \tag{2.14}$$

$$\stackrel{(a)}{=} \min_{\mathbf{p}} \max_{k} \frac{1}{p_k}[\mathbf{D}^{-1}\mathbf{\Phi}\mathbf{p} + \mathbf{D}^{-1}\boldsymbol{\sigma}]_k \tag{2.15}$$

where (a) follows from (2.8) and $\mathbf{p} = [p_1 \dots p_K]$. Consider that the total power is constrained by P, i.e., $\mathbf{1}^T\mathbf{p} = P$, reparameterize $\mathbf{p} = \frac{P}{\mathbf{1}^T\mathbf{p}'}\,\mathbf{p}'$ which satisfies the power constraint for any $\mathbf{p}'$ and rename $\mathbf{p}'$ as $\mathbf{p}$. Then, the SINR balancing becomes as follows

$$\min_{\mathbf{p}} \max_k \frac{[\mathbf{\Lambda}\,\mathbf{p}]_k}{p_k} \quad \text{with} \;\; \mathbf{\Lambda} = \mathbf{D}^{-1}[\mathbf{\Phi} + \frac{\sigma^2}{P}\,\mathbf{1}\mathbf{1}^T]$$

Classical nonnegative vectors and matrices: Note that the optimal powers should satisfy the power constraint with equality. Furthermore, at the optimum we shall have th equality

$$\frac{\text{SINR}_k^{DL}}{\gamma_k} = \frac{1}{\Delta}\,,\, \forall k \quad \leftrightarrow \quad \mathbf{\Lambda}\,\mathbf{p} = \Delta\,\mathbf{p}$$

since otherwise the user with higher SINR can lower its power, reducing interference to the user with the lowest SINR, which then increases. So, $\mathbf{p}$ is an eigen vector of $\mathbf{\Lambda}$ with eigen value Δ.

Now, for a non-negative matrix $\mathbf{\Lambda}$, the eigenvalue of the largest magnitude is positive, and its corresponding eigenvector $\mathbf{p}$ can be chosen to be non-negative. For a non-negative matrix, the non-negative eigen vector corresponding to the eigenvalue of the largest norm is positive and so is the corresponding eigen value Δ. There is only one such positive eigen pair which is called Perron Frobenius.

Actually, without physical motivation for the SINR equality, the Collatz-Wielandt formula for the Perron-Frobenius eigen pair is

$$\Delta = \min_{\mathbf{p}} \max_k \frac{[\mathbf{\Lambda}\,\mathbf{p}]_k}{p_k} = \text{eig}_{max}(\mathbf{\Lambda}),\; \mathbf{p}' = \arg\min_{\mathbf{p}} \max_k \frac{[\mathbf{\Lambda}\,\mathbf{p}]_k}{p_k} = V_{max}(\mathbf{\Lambda})$$

The non-smooth optimization criterion $\min_{\mathbf{p}} \max_k \frac{\gamma_k}{\text{SINR}_k^{DL}}$ can be transformed into a smooth problem

$$\max_{\boldsymbol{\lambda}} \min_{\mathbf{p}} \{\, t + \sum_k \lambda_k(\frac{\gamma_k}{\text{SINR}_k^{DL}} - t)\,\}$$

where the $\lambda_k \geq 0$ are the Lagrange multipliers for the constraints

$$\frac{\gamma_i}{\text{SINR}_i^{DL}} \leq t = \max_k \frac{\gamma_k}{\text{SINR}_k^{DL}}.$$

This leads to the Donsker-Varadhan-Friedland formula

$$\Delta = \text{eig}_{max}(\mathbf{\Lambda}) = \max_{\boldsymbol{\lambda}:\sum_k \lambda_k = 1} \min_{\mathbf{p}} \sum_k \lambda_k \frac{[\mathbf{\Lambda}\,\mathbf{p}]_k}{p_k}$$

A related formula is the Rayleigh quotient ($\lambda_k = \frac{p_k\,q_k}{\mathbf{q}^T\mathbf{p}}$)

$$\Delta = \max_{\mathbf{q}} \min_{\mathbf{p}} \frac{\mathbf{q}^T\mathbf{\Lambda}\,\mathbf{p}}{\mathbf{q}^T\mathbf{p}} \;\Rightarrow\; \mathbf{\Lambda}\,\mathbf{p} = \Delta\,\mathbf{p}\,,\; \mathbf{q}^T\mathbf{\Lambda} = \Delta\,\mathbf{q}^T$$

for which $\boldsymbol{q}$ and $\boldsymbol{p}$ are the left and right Perron Frobenius eigen vectors.

The beamformer optimization now only requires the dual UL powers q_k from either:

- the left Perron Frobenius eigen vector $\mathbf{q}$, via Lagrangian duality, in particular $\mathbf{g}'_k = \arg\min_{\mathbf{g}_k} [\mathbf{q}^T \boldsymbol{\Lambda}]_k$,
- by UL/DL duality from $\mathbf{p}$ and $\text{SINR}_k^{UL}(\mathbf{q}) = \text{SINR}_k^{DL}(\mathbf{p})$,
- by iterating interference functions, which correspond to the power method for finding a largest eigen vector iteratively.

2.3.3 Rate-MSE Relation

In the following, we consider a MIMO IBC with C cells with a total of K users, single stream each. System-wide user numbering: the $N_{rk} \times 1$ Rx signal at user k in cell b_k is

$$\boldsymbol{y}_k = \underbrace{\boldsymbol{H}_{k,b_k}\, \boldsymbol{g}_k\, s_k}_{\text{signal}} + \underbrace{\sum_{\substack{i \neq k \\ b_i = b_k}} \boldsymbol{H}_{k,b_k}\, \boldsymbol{g}_i\, s_i}_{\text{intracell interf.}} + \underbrace{\sum_{j \neq b_k} \sum_{i: b_i = j} \boldsymbol{H}_{k,j}\, \boldsymbol{g}_i\, s_i}_{\text{intercell interf.}} + \boldsymbol{v}_k$$

where s_k = intended (white, unit variance) scalar signal stream, $\boldsymbol{H}_{k,b_k} = N_{rk} \times N_{tb_k}$ channel from BS b_k to user k. BS b_k serves $K_{b_k} = \sum_{i:b_i=b_k} 1$ users. Noise whitened signal representation is considered: $\boldsymbol{v}_k \sim \mathcal{CN}(0, I_{N_{rk}})$. The $N_{tb_k} \times 1$ spatial *Tx filter* or *beamformer* (BF) is $\boldsymbol{g}_k$.

Treating interference as noise, user k will apply a linear *Rx filter* $\boldsymbol{f}_k$ to maximize the signal power (diversity) while reducing any residual interference that would not have been (sufficiently) suppressed by the BS Tx. The Rx filter output is $\widehat{s}_k = \boldsymbol{f}_k^H \boldsymbol{y}_k$

$$\begin{aligned}
\widehat{s}_k &= \boldsymbol{f}_k^H \boldsymbol{H}_{k,b_k}\, \boldsymbol{g}_k\, s_k + \sum_{i=1,\neq k}^{K} \boldsymbol{f}_k^H \boldsymbol{H}_{k,b_i}\, \boldsymbol{g}_i\, s_i + \boldsymbol{f}_k^H \boldsymbol{v}_k \\
&= \boldsymbol{f}_k^H \boldsymbol{h}_{k,k}\, s_k + \sum_{i \neq k} \boldsymbol{f}_k^H \boldsymbol{h}_{k,i}\, s_i + \boldsymbol{f}_k^H \boldsymbol{v}_k
\end{aligned}$$

where $\boldsymbol{h}_{k,i} = \boldsymbol{H}_{k,b_i}\, \boldsymbol{g}_i$ is the channel-Tx cascade vector.

Hereafter, we introduce the relation between the user rate and user weighted-MSE. Assuming Gaussian signaling, the achievable rate for user k is given as

$$\begin{aligned}
r_k &= \log\det\Big(1 - \boldsymbol{g}_k^H \boldsymbol{H}_{k,b_k}^H \boldsymbol{R}_k^{-1} \boldsymbol{H}_{k,b_k} \boldsymbol{g}_k\Big) \qquad (2.16) \\
\boldsymbol{R}_k &= \boldsymbol{R}_{\overline{k}} + \boldsymbol{H}_{k,b_k} \boldsymbol{g}_k \boldsymbol{g}_k^H \boldsymbol{H}_{k,b_k}^H\ , \\
\boldsymbol{R}_{\overline{k}} &= \sum_{i \neq k} \boldsymbol{H}_{k,b_i} \boldsymbol{g}_i \boldsymbol{g}_i^H \boldsymbol{H}_{k,b_i}^H + I_{N_{rk}}\ ,
\end{aligned}$$

where $\boldsymbol{R}_k$, $\boldsymbol{R}_{\overline{k}}$ represent the total, interference plus noise Rx covariance matrices, respectively, at UE k.

For a general Rx filter $\boldsymbol{f}_k$ we have the MSE $e_k(\boldsymbol{f}_k, \boldsymbol{g})$ defined as follows, given $\boldsymbol{g} = \{\boldsymbol{g}_k\}$

$$\begin{aligned} e_k(\boldsymbol{f}_k, \boldsymbol{g}) &= \mathbb{E}\Big\{|\hat{s}_k - s_k|^2\Big\} \\ &= (1 - \boldsymbol{f}_k^H \boldsymbol{H}_{kb_k} \boldsymbol{g}_k)(1 - \boldsymbol{g}_k^H \boldsymbol{H}_{kb_k}^H \boldsymbol{f}_k) + \sum_{i \neq k} \boldsymbol{f}_k^H \boldsymbol{H}_{kb_i} \boldsymbol{g}_i \boldsymbol{g}_i^H \boldsymbol{H}_{kb_i}^H \boldsymbol{f}_k + ||\boldsymbol{f}_k||^2 \quad (2.17) \\ &= 1 - \boldsymbol{f}_k^H \boldsymbol{H}_{kb_k} \boldsymbol{g}_k - \boldsymbol{g}_k^H \boldsymbol{H}_{kb_k}^H \boldsymbol{f}_k + \sum_{i} \boldsymbol{f}_k^H \boldsymbol{H}_{kb_i} \boldsymbol{g}_i \boldsymbol{g}_i^H \boldsymbol{H}_{kb_i}^H \boldsymbol{f}_k + ||\boldsymbol{f}_k||^2, \quad (2.18) \end{aligned}$$

assuming $\mathbb{E}\{|s_k|^2\} = 1$.

We notice that r_k can be expressed as a function of the error covariance e_k after MMSE receive filtering. The MMSE receive filter at user k is given as

$$\begin{aligned} \boldsymbol{f}_k^{\mathrm{MMSE}} &= \underset{\boldsymbol{f}_k}{\operatorname{argmin}}\, \mathbb{E}\{|\boldsymbol{f}_k^H \boldsymbol{y}_k - \mathrm{s}_k|^2\} \quad (2.19) \\ &= (\boldsymbol{g}_k^H \boldsymbol{H}_{kb_k}^H \boldsymbol{H}_{kb_k} \boldsymbol{g}_k + \boldsymbol{R}_{\bar{k}})^{-1} \boldsymbol{H}_{kb_k} \boldsymbol{g}_k \quad (2.20) \end{aligned}$$

The per user MSE e_k given that the MMSE-receive filter is applied can be written as

$$\begin{aligned} e_k(\boldsymbol{f}_k^{\mathrm{MMSE}}, \boldsymbol{g}) &= \mathbb{E}\Big\{|\boldsymbol{f}_k^{\mathrm{MMSE}H} \boldsymbol{y}_k - s_k|^2\Big\} \\ &= (1 + \boldsymbol{g}_k^H \boldsymbol{H}_{kb_k}^H \boldsymbol{R}_k^{-1} \boldsymbol{H}_{kb_k} \boldsymbol{g}_k)^{-1}. \quad (2.21) \end{aligned}$$

Given (2.16) and (2.21) the rate for user k can be written as

$$r_k = \log\det(e_k^{-1}). \quad (2.22)$$

2.3.4 WSR Maximization

Maximizing the weighted sum rate of the MIMO IBC system is expressed as follows

$$[\boldsymbol{g}_1^{\mathrm{WSR}} \dots \boldsymbol{g}_K^{\mathrm{WSR}}] = \underset{\boldsymbol{g}}{\operatorname{argmax}} \sum_{k=1}^{K} u_k r_k \quad (2.23)$$

with u_k being the per user weight (or priority): $0 \leq u_k \leq 1, \forall k$. The latter definition allow us to either exclude some user rate from the objective function, or reduce the problem to sum rate maximization, by assigning $u_k = 0$, or $u_k = 1, \forall k$, respectively.

In fact, the WSR maximization problem is a non-convex and complicated function of $\boldsymbol{g}$. We discuss in the following the different methods to solve it.

From WSR to WSMSE

Using the rate-MSE relation discussed previously, the WSR problem in (2.23) can be transformed into a *weighted sum MSE* problem, [26]. The resulting augmented cost function is as follows

$$\text{WSMSE}(\boldsymbol{g}, \boldsymbol{f}, w) = \sum_{k=1}^{K} u_k(w_k\, e_k(\boldsymbol{f}_k, \boldsymbol{g}) - \ln w_k) + \lambda(\sum_{k=1}^{K} ||\boldsymbol{g}_k||^2 - P)$$

where λ is the Lagrange multiplier and P denotes the total Tx power constraint.

After optimizing over the aggregate auxiliary Rx filters $\boldsymbol{f} = \{\boldsymbol{f}_k\}_{1\leq k\leq K}$ and weights w, we get the WSR back,

$$\min_{\boldsymbol{f},w} \text{WSMSE}(\boldsymbol{g}, \boldsymbol{f}, w) = -\text{WSR}(\boldsymbol{g}) + \overbrace{\sum_{k=1}^{K} u_k}^{\text{constant}}$$

With *alternating optimization*, we can solve simple quadratic or convex functions, namely

$$\begin{aligned}
&\min_{w_k} \text{WSMSE} \;\Rightarrow\; w_k = 1/e_k \\
&\min_{\boldsymbol{f}_k} \text{WSMSE} \;\Rightarrow \boldsymbol{f}_k = (\sum_i \boldsymbol{H}_k \boldsymbol{g}_i \boldsymbol{g}_i^H \boldsymbol{H}_k^H + I_{N_{r_k}})^{-1} \boldsymbol{H}_k \boldsymbol{g}_k \\
&\min_{\boldsymbol{g}_k} \text{WSMSE} \;\Rightarrow \boldsymbol{g}_k = (\sum_i u_i w_i \boldsymbol{H}_i^H \boldsymbol{f}_i \boldsymbol{f}_i^H \boldsymbol{H}_i + \lambda I_{N_t})^{-1} \boldsymbol{H}_k^H \boldsymbol{f}_k u_k w_k
\end{aligned}$$

Using UL/DL duality, the optimal Tx filter $\boldsymbol{g}_k$ is of the form of a MMSE linear Rx for the dual UL, in which λ plays the role of Rx noise variance and $u_k w_k$ plays the role of stream variance.

The optimal Lagrange multiplier λ can be found by

i) Bisection: line search on $\sum_{k=1}^{K} ||\boldsymbol{g}_k||^2 - P = 0$ [22].

ii) Analytical update as in [70, 71]: exploiting $\sum_k \boldsymbol{g}_k^H \frac{\partial \text{WSMSE}}{\partial \boldsymbol{g}_k^*} = 0$.

iii) Reparameterizing the BF to satisfy the power constraint [20]:

$$\boldsymbol{g}_k = \sqrt{\frac{P}{\sum_{i=1}^{K} ||\boldsymbol{g}_i^{'}||^2}}\, \boldsymbol{g}_k^{'},$$

with $\boldsymbol{g}_k^{'}$ now unconstrained

$$\text{SINR}_k = \frac{|\boldsymbol{f}_k \boldsymbol{H}_k \boldsymbol{g}_k^{'}|^2}{\sum_{i=1,\neq k}^{K} |\boldsymbol{f}_k \boldsymbol{H}_k \boldsymbol{g}_i^{'}|^2 + \frac{1}{P} ||\boldsymbol{f}_k||^2 \sum_{i=1}^{K} ||\boldsymbol{g}_i^{'}||^2} \;.$$

Note that *ii)* and *iii)* are equivalent, and lead to the same Lagrange multiplier expression obtained in [26] on the basis of a *heuristic* that was introduced in [72] as was pointed out in [70].

Alternating Minorization Approach

Another way to solve the problem in (2.23) is to use a classical difference of concave functions (DC) programming approach as in [73, 74]. Actually, the WSR maximization problem is non concave because of interference, the solution discussed here proposes to isolate the signal of interest from the sum rate of the rest of the signals.

Considering MU multi-stream MIMO IBC notation, let $\boldsymbol{Q}_k = \boldsymbol{G}_k\boldsymbol{G}_k^H$ be the transmit covariance for user k. The WSR maximization problem can be written as

$$\mathrm{WSR} = \sum_{k=1}^{K} u_k[\ln\det(\boldsymbol{R}_k) - \ln\det(\boldsymbol{R}_{\overline{k}})]$$

with $\boldsymbol{R}_k = \boldsymbol{H}_{k,b_k}\boldsymbol{Q}_k\boldsymbol{H}_{k,b_k}^H + \boldsymbol{R}_{\overline{k}}$, $\boldsymbol{R}_{\overline{k}} = \boldsymbol{H}_k(\sum_{i\neq k}\boldsymbol{Q}_i)\boldsymbol{H}_{k,b_i}^H + \boldsymbol{I}_{N_{r\,k}}$.

Consider the dependence of WSR on $\boldsymbol{Q}_k$ alone, we can write

$$\mathrm{WSR} = u_k\ln\det(\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{R}_k) + \mathrm{WSR}_{\overline{k}}\,,\ \ \mathrm{WSR}_{\overline{k}} = \sum_{i=1,\neq k}^{K} u_i\ln\det(\boldsymbol{R}_{\overline{i}}^{-1}\boldsymbol{R}_i)$$

where $\ln\det(\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{R}_k)$ is concave in $\boldsymbol{Q}_k$ and $WSR_{\overline{k}}$ is convex in $\boldsymbol{Q}_k$.

Since a linear function is simultaneously convex and concave, consider the first order Taylor series expansion in $\boldsymbol{Q}_k$ around $\widehat{\boldsymbol{Q}}$ (i.e. all $\widehat{\boldsymbol{Q}}_i$) with e.g. $\widehat{\boldsymbol{R}}_i = \boldsymbol{R}_i(\widehat{\boldsymbol{Q}})$, then

$$\begin{aligned}
\mathrm{WSR}_{\overline{k}}(\boldsymbol{Q}_k,\widehat{\boldsymbol{Q}}) &\approx \mathrm{WSR}_{\overline{k}}(\widehat{\boldsymbol{Q}}_k,\widehat{\boldsymbol{Q}}) - \mathrm{tr}\{(\boldsymbol{Q}_k - \widehat{\boldsymbol{Q}}_k)\widehat{\mathbf{A}}_k\},\\
\widehat{\mathbf{A}}_k = -\left.\frac{\partial\mathrm{WSR}_{\overline{k}}(\boldsymbol{Q}_k,\widehat{\boldsymbol{Q}})}{\partial\boldsymbol{Q}_k}\right|_{\widehat{\boldsymbol{Q}}_k,\widehat{\boldsymbol{Q}}} &= \sum_{i=1,\neq k}^{K} u_i\boldsymbol{H}_i^H(\widehat{\boldsymbol{R}}_{\overline{i}}^{-1} - \widehat{\boldsymbol{R}}_i^{-1})\boldsymbol{H}_i.
\end{aligned}$$

Note that the linearized (tangent) expression for $WSR_{\overline{k}}$ constitutes a lower bound for it. Now, dropping constant terms, reparameterizing $Q_k = \boldsymbol{G}_k\boldsymbol{G}_k^H$ and performing this linearization for all users, we obtain the following Lagrangian

$$\begin{aligned}
\mathrm{WSR}(\boldsymbol{G},\widehat{\boldsymbol{G}}) =& \sum_{k=1}^{K} u_k\ln\det(\boldsymbol{I} + \boldsymbol{G}_k^H\boldsymbol{H}_{k,b_k}^H\widehat{\boldsymbol{R}}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k)\\
&- \sum_{k=1}^{K}\mathrm{tr}\{\boldsymbol{G}_k^H(\widehat{\mathbf{A}}_k + \lambda_{b_k}\boldsymbol{I})\boldsymbol{G}_k\} + \sum_{c=1}^{C}\lambda_c P_c.
\end{aligned}$$

The gradient (w.r.t. $\boldsymbol{G}_k$) of this concave WSR lower bound is actually still the same as that of the original WSR or of the WSMSE criteria, allowing the following generalized eigenvector interpretation

$$\boldsymbol{H}_{k,b_k}^H \widehat{\boldsymbol{R}}_{\overline{k}}^{-1} \boldsymbol{H}_{k,b_k} \boldsymbol{G}_k = (\widehat{\mathbf{A}}_k + \lambda_{b_k} \boldsymbol{I}) \boldsymbol{G}_k \frac{1}{u_k} (I + \boldsymbol{G}_k^H \boldsymbol{H}_{k,b_k}^H \widehat{\boldsymbol{R}}_{\overline{k}}^{-1} \boldsymbol{H}_{k,b_k} \boldsymbol{G}_k)$$

or hence

$$\boldsymbol{G}_k^{'} = V_{max}(\boldsymbol{H}_{k,b_k}^H \widehat{\boldsymbol{R}}_{\overline{k}}^{-1} \boldsymbol{H}_{k,b_k}, \widehat{\mathbf{A}}_k + \lambda_{b_k} \boldsymbol{I}).$$

Introducing $\boldsymbol{P}_k \geq 0$ and substituting $\boldsymbol{G}_k = \boldsymbol{G}_k^{'} \boldsymbol{P}_k^{\frac{1}{2}}$, the Lagrange multipliers λ_c, for all c, are adjusted to satisfy the power constraints $\sum_{k:b_k=c} \text{tr}\{\boldsymbol{P}_k\} = P_c$.

Let us define $\boldsymbol{\Sigma}_k^{(1)} = \boldsymbol{G}_k^{'H} \boldsymbol{H}_k^H \widehat{\boldsymbol{R}}_{\overline{k}}^{-1} \boldsymbol{H}_k \boldsymbol{G}_k^{'}$ and $\boldsymbol{\Sigma}_k^{(2)} = \boldsymbol{G}_k^{'H} \widehat{\mathbf{A}}_k \boldsymbol{G}_k^{'}$. The advantage of this formulation is that it allows straightforward power adaptation, yielding

$$\text{WSR} = \sum_{c=1}^{C} \lambda_c P_c + \sum_{k=1}^{K} [u_k \ln\det(\boldsymbol{I} + \boldsymbol{P}_k \boldsymbol{\Sigma}_k^{(1)}) - \text{tr}\{\boldsymbol{P}_k (\boldsymbol{\Sigma}_k^{(2)} + \lambda_{b_k} \boldsymbol{I})\}]$$

which leads to the following *interference leakage aware water filling* (jointly for the $\boldsymbol{P}_k$ and λ_{b_k})

$$\boldsymbol{P}_k(l,l) = \left(u_k (\boldsymbol{\Sigma}_k^{(2)}(l,l) + \lambda_{b_k} \boldsymbol{I})^{-1} - \boldsymbol{\Sigma}_k^{-(1)}(l,l) \right)^+ \ , \quad \sum_{k:b_k=c} \text{tr}\{\boldsymbol{P}_k\} = P_c \ ,$$

for all l s.t. $\boldsymbol{\Sigma}_k^{(i)} \geq 0$ with $i = \{1, 2\}$, and $z^+ = \max(0, z)$. Note also that as with any alternating optimization procedure, there are many updating schedules possible, with different impact on convergence speed. The quantities to be updated are $\boldsymbol{G}_k^{'}$, $\boldsymbol{P}_k$ and λ_c. The advantage of the DC approach is that it works for any number of streams/user d_k, by simply taking more or less eigenvectors. In other words, we can take the $d_k^{\max}$ max eigenvectors of the eigenmatrix $\boldsymbol{G}_k^{'}$. We mean by the max eigenvectors, the eigenvectors corresponding to the highest eigenvalues. The waterfilling then automatically determines (at each iteration) how many streams can be sustained.

2.4 Closing Remarks

In this chapter, we presented an overview of the evolution of power allocation problem formulation with respect to max-min fairness utility, leading to the max-min per user (weighted) rate problem considered in this thesis. We also provided the needed theoretical background to ensure good comprehension in the next chapters.

Part II

Rate Balancing with Perfect CSIT

Chapter 3

Rate Balancing via MSE UL/DL duality with total power constraint: BC

3.1 Overview

In this chapter, we focus on transceiver optimization problem w.r.t. rate balancing. The problem is to maximize the minimum weighted rate (the ratio of rate and given priority). User-wise rate is considered, i.e., total rate of each user summed over its streams. The problem is studied in a multi-user MIMO BC under a total power constraint. This optimization problem is here solved in an alternating manner by exploiting matrix-weighted MSE uplink/downlink duality with proven convergence to a local optimum. The MSE duality [47] plays an important role, since it ensures that the same MMSE can be achieved in both links. This guarantees the convergence in each iteration.

The rest of this chapter is organized as follows. Section 3.2 details the system model under consideration. The problem formulation in downlink communications and its respective dual problem are given in Section 3.3. We provide the algorithmic solution based on UL/DL duality in 3.4. Numerical results for validation are presented in Section 3.5. We finally conclude in Section 3.6.

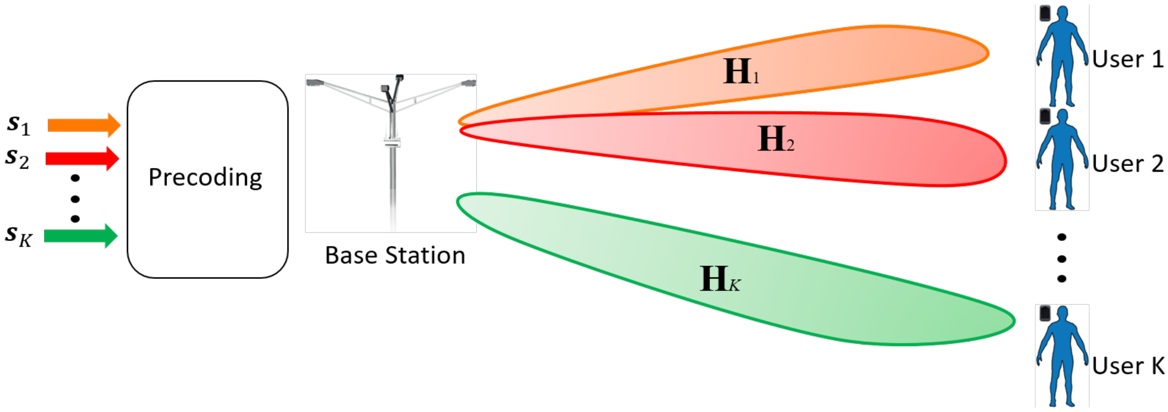

Figure 3.1: Single cell broadcast channel system model.

3.2 System Model

The considered network is a multi-user MIMO DL system, (see Figure 3.1). We focus on a Base Station (BS) of M transmit antennas serving K users of each N_k antennas, ($k = 1, ..., K$ is the users' index). The channel between the kth user and the BS is denoted by $\boldsymbol{H}_k^{\mathrm{H}} \in \mathbb{C}^{M \times N_k}$, and $\boldsymbol{H}^{\mathrm{H}} = [\boldsymbol{H}_1^{\mathrm{H}}, ..., \boldsymbol{H}_K^{\mathrm{H}}]$ is the overall channel matrix.

We assume zero-mean white Gaussian noise $\boldsymbol{n}_k \in \mathbb{C}^{N_k \times 1}$ with distribution $\mathcal{CN}(0, \sigma_n^2 \boldsymbol{I})$ at the kth user. We assume independent unity-power transmit symbols $\mathbf{s} = [\mathbf{s}_1^{\mathrm{T}} \dots \mathbf{s}_K^{\mathrm{T}}]^{\mathrm{T}}$, i.e., $\mathbb{E}\big[\mathbf{s}\mathbf{s}^{\mathrm{H}}\big] = \boldsymbol{I}$, where $\mathbf{s}_k \in \mathbb{C}^{d_k \times 1}$ is the data vector to be transmitted to the kth user, with d_k being the number of streams allowed by user k. The latter are transmitted using the transmit filtering matrix $\boldsymbol{\mathcal{G}} = \boldsymbol{G}\boldsymbol{P}^{1/2} \in \mathbb{C}^{M \times N_d}$, composed of the beamforming matrix $\boldsymbol{G} = [\boldsymbol{G}_1 \dots \boldsymbol{G}_K] = [\boldsymbol{g}_1 \dots \boldsymbol{g}_{N_d}]$ with normalized columns $\|\boldsymbol{g}_i\|_2 = 1$ and the diagonal non-negative DL power allocation $\boldsymbol{P}^{1/2} = \mathrm{blkdiag}\{\boldsymbol{P}_1^{1/2}, \dots, \boldsymbol{P}_K^{1/2}\}$ where $\mathrm{diag}(\boldsymbol{P}_k) \in \mathbb{R}_+^{d_k \times 1}$ contains the transmission powers and $N_d = \sum_{k=1}^K d_k$ is the total number of streams. The total transmit power is limited, i.e., $\mathrm{tr}\big(\boldsymbol{P}\big) \leq P_{\max}$, see Figure 3.2.

Similarly, the receive filtering matrix for each user is defined as $\boldsymbol{\mathcal{F}}_k^{\mathrm{H}} = \boldsymbol{P}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$, composed of beamforming matrix $\boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$ and the diagonal matrices $\boldsymbol{\beta}_k$ contain scaling factors which ensure that the columns of $\boldsymbol{F}_k^{\mathrm{H}}$ have unit norm. We define $\boldsymbol{\beta} = \mathrm{blkdiag}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_K\} = \mathrm{diag}\{[\beta_1 \dots \beta_{N_d}]\}$ and $\boldsymbol{F} = \mathrm{blkdiag}\{\boldsymbol{F}_1, \dots, \boldsymbol{F}_K\} = [\boldsymbol{f}_1 \dots \boldsymbol{f}_{N_d}]$ with normalized per-stream receivers, i.e., $\|\boldsymbol{f}_i\|_2 = 1$.

The MSE per stream $\varepsilon_i^{\mathrm{DL}}$ between the decision variable $\hat{s}_i$ and the transmit data symbol s_i is defined as follows

$$\begin{aligned}\varepsilon_i^{\mathrm{DL}} = \mathbb{E}\Big\{|\hat{s}_i - s_i|^2\Big\} &= \beta_i^2/p_i\boldsymbol{f}_i^{\mathrm{H}}\boldsymbol{H}\Big(\sum_{j=1}^{N_d} p_j\boldsymbol{g}_j\boldsymbol{g}_j^{\mathrm{H}}\Big)\boldsymbol{H}^{\mathrm{H}}\boldsymbol{f}_i \\ &- 2\beta_i\mathrm{Re}\big\{\boldsymbol{f}_i^{\mathrm{H}}\boldsymbol{H}\boldsymbol{g}_i\big\} + \sigma_n^2\beta_i^2/p_i + 1, \forall i \in \{1, ..., N_d\}.\end{aligned} \tag{3.1}$$

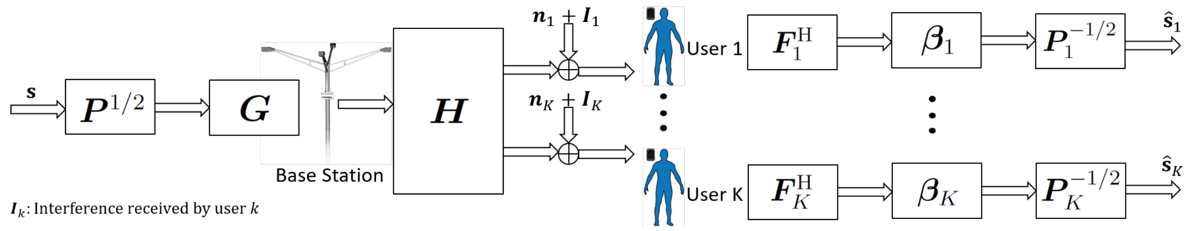

Figure 3.2: Equivalent downlink channel.

3.3 Problem Formulation

In this work, we aim to solve the weighted user-rate max-min optimization problem under a total transmit power constraint, i.e., the user rate balancing problem expressed as follows

$$\max_{\{\boldsymbol{G},\boldsymbol{P},\boldsymbol{F},\boldsymbol{\beta}\}} \min_{k} \; r_k/r_k^{\circ} \quad \text{s.t.} \;\; \mathrm{tr}\big(\boldsymbol{P}\big) \leq P_{\max} \tag{3.2}$$

where r_k is the kth user-rate

$$r_k = \ln\det\Big(\boldsymbol{I}+\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}(\sigma_n^2\boldsymbol{I}+\sum_{j\neq k}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_j\boldsymbol{\mathcal{G}}_j^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}})^{-1}\Big) \tag{3.3}$$

and r_k° is the rate scaling factor for user k. However, the problem presented in (3.2) is complex and can not be solved directly.

Lemma 1. *The rate of user k in (3.3) can also be represented as*

$$r_k = \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k} \big[\ln\det\big(\boldsymbol{W}_k\big) - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k^{\mathrm{DL}}\big) + d_k\big]. \tag{3.4}$$

where

$$\begin{aligned}\mathbf{E}_k^{\mathrm{DL}} &= \mathbb{E}\Big[(\hat{\mathbf{s}}_k - \mathbf{s}_k)(\hat{\mathbf{s}}_k - \mathbf{s}_k)^{\mathrm{H}}\Big] \\ &= (\boldsymbol{I}-\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k)(\boldsymbol{I}-\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k)^{\mathrm{H}} + \sum_{j\neq k}\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{H}_j\boldsymbol{\mathcal{G}}_j\boldsymbol{\mathcal{G}}_j^{\mathrm{H}}\boldsymbol{H}_j^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k + \sigma_n^2\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k\end{aligned} \tag{3.5}$$

is the kth-user DL MSE matrix between the decision variable $\hat{\mathbf{s}}_k$ and the transmit signal $\mathbf{s}_k$, and $\boldsymbol{W} = \{\boldsymbol{W}_k\}_{1\leq k\leq K}$ are auxiliary weight matrix variables with optimal solution $\boldsymbol{W}_k = \big(\mathbf{E}_k^{\mathrm{DL}}\big)^{-1}$ and $\boldsymbol{\mathcal{F}}_k = (\sigma_n^2\boldsymbol{I} + \sum_{j=1}^{K}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_j\boldsymbol{\mathcal{G}}_j^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}})^{-1}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k$, [27].

Proof. Appendix A.1. □

Now considering both (3.2) and (3.4), and introducing $t = \min_k r_k/r_k^{\circ}$, we have $\forall k$

$$\begin{aligned} & r_k/(t\,r_k^{\circ}) \geq 1 \text{ or } r_k/r_k^{\circ} \geq t \\ \overset{(a)}{\Longleftrightarrow}\; & \ln\det\big(\boldsymbol{W}_k\big) + d_k - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k^{\mathrm{DL}}\big) \geq tr_k^{\circ} \\ \Longleftrightarrow\; & \frac{\mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k^{\mathrm{DL}}\big)}{\ln\det\big(\boldsymbol{W}_k\big) + d_k - tr_k^{\circ}} \overset{(b)}{=} \frac{\epsilon_{w,k}^{\mathrm{DL}}}{\xi_k} \leq 1 \end{aligned} \tag{3.6}$$

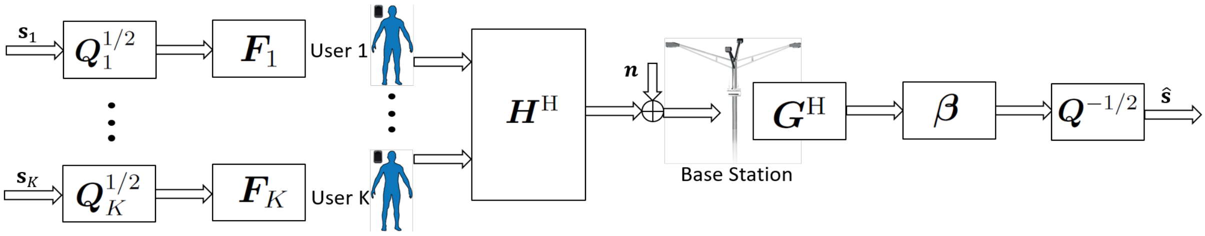

Figure 3.3: Dual uplink channel.

where (a) follows from (3.4) (with optimal $\boldsymbol{W}_k$) and (b) from $\epsilon_{w,k}^{\mathrm{DL}} = \mathrm{tr}\big(\boldsymbol{W}_k \mathbf{E}_k^{\mathrm{DL}}\big)$, the matrix-weighted MSE (WMSE), and $\xi_k = \ln\det\big(\boldsymbol{W}_k\big) + d_k - r_k^{\triangle}$ the WMSE requirement, with $r_k^{\triangle} = tr_k^{\circ}$ the individual rate target, i.e. $r_k \geq r_k^{\triangle}$. What we exploit here is a scale factor t that can be chosen freely in the rate weights r_k^{o} in (3.2), to transform the rate weights r_k° into target rates $r_k^{\triangle} = tr_k^{\circ}$, which at the same time allows to interpret the WMSE weights ξ_k as target WMSE values.

Doing so, the initial rate balancing optimization problem (3.2) can be transformed into a matrix-weighted MSE balancing problem expressed as follows

$$\begin{aligned} \min_{\{\boldsymbol{G},\boldsymbol{P},\boldsymbol{F},\boldsymbol{\beta}\}} \max_{k} \ & \epsilon_{w,k}^{\mathrm{DL}}/\xi_k \\ \text{s.t. } & \mathrm{tr}\big(\boldsymbol{P}\big) \leq P_{\max}, \end{aligned} \tag{3.7}$$

which needs to be complemented with an outer loop in which $\boldsymbol{W}_k = \big(\mathbf{E}_k^{\mathrm{DL}}\big)^{-1}$, $t = \min_k r_k/r_k^{\circ}$, $r_k^{\triangle} = tr_k^{\circ}$ and $\xi_k = d_k + r_k - r_k^{\triangle}$ get updated.

The problem in (3.7) is still difficult to be handled directly. In the next sections, we solve the problem via UL and DL MSE duality. To this aim, we model an equivalent UL-DL channel plus transceivers pair by separating the filters into two parts: a matrix with unity-norm columns and a scaling matrix [75]. Then, the UL and DL are proved to share the same MSE by switching the role of the normalized filters in the UL and DL. Doing so, an algorithmic solution can be derived for the optimization problem (3.7).

3.3.1 Dual UL Channel

In the equivalent UL model represented in Figure 3.3, we switch between the role of the normalized transmit and receive filters. In fact, $\boldsymbol{F}_k \boldsymbol{Q}_k^{1/2}$ is the kth transmit filter and $\boldsymbol{Q}^{-1/2}\boldsymbol{\beta}\boldsymbol{G}^{\mathrm{H}}$ is a multi-user receive filter, where $\boldsymbol{Q} = \mathrm{blkdiag}\big\{\boldsymbol{Q}_1, ..., \boldsymbol{Q}_K\big\}$ with $\mathrm{diag}(\boldsymbol{Q}_k) \in \mathbb{R}_+^{d_k \times 1}$ being the UL power allocation.

Although the quantities $\boldsymbol{H}, \boldsymbol{G}, \boldsymbol{F}$ and $\boldsymbol{\beta}$ are the same, the UL power allocation $\boldsymbol{q} = [q_1 \ldots q_{N_d}]^{\mathrm{T}} = \mathrm{diag}(\boldsymbol{Q})$ may differ from the DL allocation $\boldsymbol{p} = [p_1 \ldots p_{N_d}]^{\mathrm{T}} = \mathrm{diag}(\boldsymbol{P})$, both verifying the same sum power constraint $\|\boldsymbol{p}\|_1 = \|\boldsymbol{q}\|_1 \leq P_{\max}$.

The corresponding UL per stream MSE $\varepsilon_i^{\mathrm{UL}}$ is given by

$$\varepsilon_i^{\mathrm{UL}} = \beta_i^2/q_i \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \big(\sum_{j=1}^{N_d} q_j \boldsymbol{f}_j \boldsymbol{f}_j^{\mathrm{H}} \big) \boldsymbol{H} \boldsymbol{g}_i - 2\beta_i \mathrm{Re}\{\boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i\} + \sigma_n^2 \beta_i^2/q_i + 1, \forall i. \tag{3.8}$$

3.3.2 MSE Duality

With the equivalent DL channel and its dual UL, it has been shown that the same per stream MSE values are achieved in both links, i.e., $\boldsymbol{\varepsilon}^{\mathrm{UL/DL}} = \mathrm{diag}\big\{[\varepsilon_1^{\mathrm{UL/DL}} \dots \varepsilon_{N_d}^{\mathrm{UL/DL}}]\big\} = \mathrm{diag}\big\{[\varepsilon_1 \dots \varepsilon_{N_d}]\big\} = \boldsymbol{\varepsilon}$ [75].

The UL and DL power allocation, obtained by solving the MSE expressions as in (3.8) for UL w.r.t. the powers, are given by (see Appendix A.2)

$$\boldsymbol{q} = \sigma_n^2 (\boldsymbol{\varepsilon} - \boldsymbol{D} - \boldsymbol{\beta}^2 \boldsymbol{\Psi})^{-1} \boldsymbol{\beta}^2 \mathbf{1}_{N_d} \tag{3.9}$$

and

$$\boldsymbol{p} = \sigma_n^2 (\boldsymbol{\varepsilon} - \boldsymbol{D} - \boldsymbol{\beta}^2 \boldsymbol{\Psi}^T)^{-1} \boldsymbol{\beta}^2 \mathbf{1}_{N_d} \tag{3.10}$$

respectively, where the diagonal matrix $\boldsymbol{D}$ is defined as

$$[\boldsymbol{D}]_{ii} = \beta_i^2 \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i \boldsymbol{f}_i^{\mathrm{H}} \boldsymbol{H} \boldsymbol{g}_i - 2\beta_i \mathrm{Re}\{\boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i\} + 1$$

and

$$[\boldsymbol{\Psi}]_{ij} = \begin{cases} \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_j \boldsymbol{f}_j^{\mathrm{H}} \boldsymbol{H} \boldsymbol{g}_i, & i \neq j \\ 0, & i = j. \end{cases}$$

In fact, the MSE duality allows to optimize the transceiver design by switching between the virtual UL and actual DL channels. The optimal receive filtering matrices in both UL and DL are MMSE filters and given by

$$\boldsymbol{G}_k \boldsymbol{\beta}_k \boldsymbol{Q}_k^{-1/2} = \big(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{Q} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{H} + \sigma_n^2 \boldsymbol{I} \big)^{-1} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k \boldsymbol{Q}_k^{1/2} \tag{3.11}$$

and

$$\boldsymbol{F}_k \boldsymbol{\beta}_k \boldsymbol{P}_k^{-1/2} = \big(\boldsymbol{H}_k \boldsymbol{G} \boldsymbol{P} \boldsymbol{G}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} + \sigma_n^2 \boldsymbol{I} \big)^{-1} \boldsymbol{H}_k \boldsymbol{G}_k \boldsymbol{P}_k^{1/2} \tag{3.12}$$

respectively.

3.4 Algorithmic Solution via UL-DL MSE Duality

In this section, the problem (3.7) with respect to the matrix weighted user-MSE is studied. First, we start by the UL power allocation strategies. Then, the joint optimization will follow given the MSE duality. In fact, the MSE duality opens up a way to obtain optimal MMSE receiver designs in (3.11) and (3.12). The DL matrix weighted user-MSE optimization problems can be solved by optimizing the weighted MSE values of the dual UL system. The latter can be formulated as

$$\min_{\{\boldsymbol{G},\boldsymbol{F},\boldsymbol{W}\}} \max_k \ \epsilon_{w,k}^{\mathrm{UL}}/\xi_k \quad \text{s.t.} \ \operatorname{tr}\big(\boldsymbol{Q}\big) \le P_{\max} \tag{3.13}$$

where $\epsilon_{w,k}^{\mathrm{UL}} = \operatorname{tr}\big(\boldsymbol{W}_k \mathbf{E}_k^{\mathrm{UL}}\big)$, and

$$\begin{aligned}\mathbf{E}_k^{\mathrm{UL}} &= (\boldsymbol{I} - \boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{Q}_k^{1/2})(\boldsymbol{I} - \boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{Q}_k^{1/2})^{\mathrm{H}} \\ &+ \sum_{j\neq k} \boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_j^{\mathrm{H}}\boldsymbol{F}_j\boldsymbol{Q}_j\boldsymbol{F}_j^{\mathrm{H}}\boldsymbol{H}_j\boldsymbol{G}_k\boldsymbol{\beta}_k\boldsymbol{Q}_k^{-1/2} + \sigma_n^2\boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{G}_k\boldsymbol{\beta}_k\boldsymbol{Q}_k^{-1/2}.\end{aligned} \tag{3.14}$$

Then, based on the equivalent UL/DL channel pair, we derive a general framework for joint DL MSE design. First, in the *UL channel*, we find the globally optimal powers $\boldsymbol{Q}$ according to the optimization problem under consideration; then, we update the UL receivers as MMSE filters (3.11) and we compute the associated per stream MSE values $\varepsilon_i^{\mathrm{UL}}$, $\forall i$. Second, in the *DL channel*, we find the DL power allocation $\boldsymbol{P}$ which achieves the same UL MSE values; and we update the DL receivers as MMSE filters (3.12). Finally, we update $\boldsymbol{W}_k$.

The matrix weighted per user MSE can be expressed as follows

$$\begin{aligned}&\epsilon_{w,k}^{\mathrm{UL}} = \operatorname{tr}\big(\boldsymbol{W}_k\boldsymbol{E}_k^{\mathrm{UL}}\big) \\ &= \operatorname{tr}\big(\boldsymbol{W}_k\big) + \operatorname{tr}\big(\boldsymbol{W}_k\boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_k\boldsymbol{Q}_k\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{G}_k\boldsymbol{\beta}_k\boldsymbol{Q}_k^{-1/2}\big) \\ &- 2\mathrm{Re}\Big\{\operatorname{tr}\big(\boldsymbol{Q}_k^{1/2}\boldsymbol{W}_k\boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_k\big)\Big\} + \sigma_n^2\operatorname{tr}\big(\boldsymbol{W}_k\boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{G}_k\boldsymbol{\beta}_k\boldsymbol{Q}_k^{-1/2}\big) \\ &+ \sum_{j\neq k}\operatorname{tr}\big(\boldsymbol{W}_k\boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_j^{\mathrm{H}}\boldsymbol{F}_j\boldsymbol{Q}_j\boldsymbol{F}_j^{\mathrm{H}}\boldsymbol{H}_j\boldsymbol{G}_k\boldsymbol{\beta}_k\boldsymbol{Q}_k^{-1/2}\big), \forall k.\end{aligned} \tag{3.15}$$

We define $\boldsymbol{Q}_k = \tilde{q}_k\bar{\boldsymbol{Q}}_k$ where $\operatorname{tr}\big(\bar{\boldsymbol{Q}}_k\big) = 1$ and $\tilde{q}_k$ is the individual power of the kth user. Then, the transmit covariance matrix $\boldsymbol{R}_k = \boldsymbol{F}_k\boldsymbol{Q}_k\boldsymbol{F}_k^{\mathrm{H}}$ can be written as $\boldsymbol{R}_k = \tilde{q}_k\bar{\boldsymbol{R}}_k$ with $\operatorname{tr}\big(\bar{\boldsymbol{R}}_k\big) = 1$. Thus, the matrix weighted MSE $\epsilon_{w,k}$ becomes a function of $\tilde{\boldsymbol{q}} = [\tilde{q}_1, ..., \tilde{q}_K]^T$

$$\epsilon_{w,k}^{\mathrm{UL}} = a_k + \tilde{q}_k^{-1}\sum_{j\neq k}\tilde{q}_j b_{kj} + \tilde{q}_k^{-1}c_k\sigma_n^2, \forall k \tag{3.16}$$

where

$$\begin{aligned}a_k &= \operatorname{tr}\big(\boldsymbol{W}_k\big) + \operatorname{tr}\big(\boldsymbol{W}_k\bar{\boldsymbol{Q}}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\bar{\boldsymbol{R}}_k\boldsymbol{H}_k\boldsymbol{G}_k\boldsymbol{\beta}_k\bar{\boldsymbol{Q}}_k^{-1/2}\big) \\ &\quad - 2\mathrm{Re}\Big\{\operatorname{tr}\big(\boldsymbol{Q}_k^{1/2}\boldsymbol{W}_k\boldsymbol{Q}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_k\big)\Big\}, \\ b_{kj} &= \operatorname{tr}\big(\boldsymbol{W}_k\bar{\boldsymbol{Q}}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_j^{\mathrm{H}}\bar{\boldsymbol{R}}_j\boldsymbol{H}_j\boldsymbol{G}_k\boldsymbol{\beta}_k\bar{\boldsymbol{Q}}_k^{-1/2}\big)\end{aligned}$$

and

$$c_k = \operatorname{tr}\big(\boldsymbol{W}_k\bar{\boldsymbol{Q}}_k^{-1/2}\boldsymbol{\beta}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{G}_k\boldsymbol{\beta}_k\bar{\boldsymbol{Q}}_k^{-1/2}\big)$$

.

Actually, problem (3.13) always has a global minimizer $\tilde{\boldsymbol{q}}^{\text{opt}}$ characterized by the following equations:

$$\Delta^{\text{UL}} = \frac{\epsilon_{w,k}^{\text{UL}}(\boldsymbol{q}^{\text{opt}})}{\xi_k}, \quad \forall k, \tag{3.17}$$

$$\|\boldsymbol{q}^{\text{opt}}\|_1 = P_{\text{max}} \tag{3.18}$$

where Δ^{UL} is the minimum balanced matrix-weighted user MSE.

We aim to form an eigensystem by combining (3.17) and (3.18). For that, we rewrite (3.16) as

$$\boldsymbol{\epsilon}_w^{\text{UL}}\tilde{\boldsymbol{q}} = \boldsymbol{A}\tilde{\boldsymbol{q}} + \sigma_n^2\boldsymbol{C}\mathbf{1}_K \tag{3.19}$$

$$\text{where } \boldsymbol{\epsilon}_w^{\text{UL}} = \text{diag}\{\epsilon_{w,1}^{\text{UL}}, \ldots, \epsilon_{w,K}^{\text{UL}}\},$$

$$\boldsymbol{C} = \text{diag}\big\{c_1, \ldots, c_K\big\}, \tag{3.20}$$

$$\text{and } [\boldsymbol{A}]_{kj} = \begin{cases} b_{kj}, & k \neq j \\ a_k, & k = j. \end{cases} \tag{3.21}$$

Now, we define $\boldsymbol{\xi} = \text{diag}\big\{[\xi_1 \ldots \xi_K]\big\}$ and multiply both sides by $\boldsymbol{\xi}^{-1}$ to have

$$\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w^{\text{UL}}\tilde{\boldsymbol{q}} = \boldsymbol{\xi}^{-1}\boldsymbol{A}\tilde{\boldsymbol{q}} + \sigma_n^2\boldsymbol{\xi}^{-1}\boldsymbol{C}\mathbf{1}_K. \tag{3.22}$$

From (3.17), we have $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w^{\text{UL}}(\tilde{\boldsymbol{q}}^{\text{opt}}) = \Delta^{\text{UL}}\boldsymbol{I}$. Thus, (3.22) becomes

$$\Delta^{\text{UL}}\tilde{\boldsymbol{q}} = \boldsymbol{\xi}^{-1}\boldsymbol{A}\tilde{\boldsymbol{q}} + \sigma_n^2\boldsymbol{\xi}^{-1}\boldsymbol{C}\mathbf{1}_K. \tag{3.23}$$

From (3.18), we can reparameterize $\tilde{\boldsymbol{q}} = \frac{P_{\text{max}}}{\mathbf{1}_K^{\text{T}}\boldsymbol{q}'}\boldsymbol{q}'$ where $\boldsymbol{q}'$ is unconstrained. This allows to rewrite (3.23) as [43]

$$\boldsymbol{\Lambda}\tilde{\boldsymbol{q}}' = \Delta^{\text{UL}}\tilde{\boldsymbol{q}}', \ \boldsymbol{\Lambda} = \boldsymbol{\xi}^{-1}\boldsymbol{A} + \frac{\sigma_n^2}{P_{\text{max}}}\boldsymbol{\xi}^{-1}\boldsymbol{C}\mathbf{1}_K\mathbf{1}_K^{\text{T}} \tag{3.24}$$

It can be observed that Δ^{UL} is an eigenvalue of the non-negative extended coupling matrix $\boldsymbol{\Lambda}$. However, not all eigenvalues represent physically meaningful values. In particular, $\tilde{\boldsymbol{q}}^{\text{opt}} > 0$ and $\Delta^{\text{UL}} > 0$ must be fulfilled.

It is known that for any non-negative irreducible real matrix $\boldsymbol{X}$ with spectral radius $\rho(\boldsymbol{X})$, there exists a unique vector $\boldsymbol{q} > 0$ and $\lambda_{\text{max}}(\boldsymbol{X}) = \rho(\boldsymbol{X})$ such that $\boldsymbol{X}q = \lambda_{\text{max}}(\boldsymbol{X})\boldsymbol{q}$. The uniqueness of $\lambda_{\text{max}}(\boldsymbol{\Lambda})$ also follows from immediately from the function $\Delta^{\text{UL}}(P_{\text{max}})$ being strictly monotonically decreasing in P_{max}. This rules out the existence of two different balanced levels with the same sum power. Hence, the balanced level is given by

$$\Delta^{\text{UL, opt}} = \lambda_{\text{max}}(\boldsymbol{\Lambda}). \tag{3.25}$$

Therefore, the optimal power allocation $\tilde{\boldsymbol{q}}'$ is the principal eigenvector of the matrix $\boldsymbol{\Lambda}$ in (3.24). As noted in [8], we have in fact

$$\lambda_{\max}(\boldsymbol{\Lambda}) = \min_{\tilde{\boldsymbol{p}}} \max_{\tilde{\boldsymbol{q}}} \frac{\tilde{\boldsymbol{p}}^H \boldsymbol{\Lambda} \tilde{\boldsymbol{q}}}{\tilde{\boldsymbol{p}}^H \tilde{\boldsymbol{q}}} = \max_{\tilde{\boldsymbol{p}}} \min_{\tilde{\boldsymbol{q}}} \frac{\tilde{\boldsymbol{p}}^H \boldsymbol{\Lambda} \tilde{\boldsymbol{q}}}{\tilde{\boldsymbol{p}}^H \tilde{\boldsymbol{q}}} \tag{3.26}$$

where in [8] $\tilde{\boldsymbol{p}}$ was said to have no particular meaning but actually can be shown to relate to the DL powers. So, the proposed algorithm provides in the inner loop an alternating optimization of (3.26) w.r.t. $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{q}}$, $\boldsymbol{F}$, $\boldsymbol{G}$ [8], [75]. If we take for $\tilde{\boldsymbol{p}}$ the K standard basis vectors, then we get

$$\lambda_{\max}(\boldsymbol{\Lambda}) = \min_{\tilde{\boldsymbol{q}}} \max_{k} \frac{\left(\boldsymbol{\Lambda}\tilde{\boldsymbol{q}}\right)_k}{\tilde{\boldsymbol{q}}_k} \tag{3.27}$$

which from (3.17), (3.22), (3.24) can be seen to be exactly the WMSE balancing problem we want to solve.

3.4.1 Algorithm

The proposed optimization framework is summarized hereafter in Table 3.1. Superscripts $(.)^{(n)}$ and $(.)^{(tmp)}$ denote the n^{th} iteration and a temporary value, respectively. This algorithm is based on a double loop. The inner loop solves the WMSE balancing problem in (3.7) whereas the outer loop iteratively transforms the WMSE balancing problem into the original rate balancing problem in (3.2).

3.4.2 Proof of Convergence

In case the rate weights r_k° would not satisfy $r_k \geq r_k^\circ$, this issue will be rectified by the scale factor t after one iteration (of the outer loop). It can be shown that $t = \min_k \frac{r_k^{(m)}}{r_k^{\circ(m-1)}} \geq 1$. By contradiction, if this was not the case, it can be shown to lead to $\frac{\operatorname{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}} > 1,\ \forall k$ and hence $\Delta^{(m)} > 1$. But we have

$$\begin{aligned}\Delta^{(m)} &= \frac{\operatorname{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}},\ \forall k, = \max_k \frac{\operatorname{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}} \\ &\overset{(a)}{<} \max_k \frac{\operatorname{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m-1)}\right)}{\xi_k^{(m-1)}} = \max_k \frac{d_k}{\xi_k^{(m-1)}} \overset{(b)}{<} 1\,.\end{aligned} \tag{3.28}$$

Let $\mathbf{E} = \{\mathbf{E}_k, k = 1, ..., K\}$ and $f^{(m)}(\mathbf{E}) = \max_k \frac{\operatorname{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k\right)}{\xi_k^{(m-1)}}$. Then (a) is due to the fact that the algorithm in fact performs alternating minimization of $f^{(m)}(\mathbf{E})$ w.r.t. $\boldsymbol{G}$, $\boldsymbol{F}$, $\tilde{\boldsymbol{q}}$ and hence will lead to $f^{(m)}(\mathbf{E}^{(m)}) < f^{(m)}(\mathbf{E}^{(m-1)})$. On the other hand, (b) is due to $\xi_k^{(m-1)} = d_k + r_k^{(m-1)} - r_k^{\circ(m-1)} > d_k$, for $m \geq 3$.

Table 3.1: User Rate Balancing via UL/DL Duality Algorithm

1. initialize: $\boldsymbol{F}_k^{\mathrm{H}(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})$, $\bar{\boldsymbol{Q}}^{(0,0)} = \frac{P_{\max}}{N_d}\boldsymbol{I}$, $m = n = 0$ and $n_{\max}, m_{\max}$ and fix $r_k^{\circ(0)}$
2. compute UL receive filter $\boldsymbol{G}^{(0,0)}$ and $\boldsymbol{\beta}^{(0,0)}$ with (3.11)
3. set $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}$ and $\xi_k^{(0)} = d_k$
4. find optimal user power allocation $\tilde{\boldsymbol{q}}^{(0,0)}$ by solving (3.24) and compute $\boldsymbol{Q}_k^{(0,0)} = \tilde{q}_k^{(0,0)}\bar{\boldsymbol{Q}}_k^{(0,0)}$
5. **repeat**

   5.1 **repeat**

   $n \leftarrow n + 1$

   UL channel:

   - update $\boldsymbol{G}^{(n,m-1)}$ and $\boldsymbol{\beta}^{(tmp,tmp)}$ with (3.11)
   - compute the MSE values $\boldsymbol{\varepsilon}^{\mathrm{UL},(n)}$ with (3.8)

   DL channel:

   - compute $\boldsymbol{P}^{(n,m-1)}$ with (3.10)
   - update $\boldsymbol{F}^{(n,m-1)}$ and $\boldsymbol{\beta}^{(tmp,tmp)}$ with (3.12)
   - compute the MSE values $\boldsymbol{\varepsilon}^{\mathrm{DL}(n)}$ with (3.1)

   UL channel:

   - compute $\boldsymbol{Q}^{(tmp,tmp)}$ with (3.9) and $\bar{\boldsymbol{Q}}_k^{(n,m-1)} = \boldsymbol{Q}_k^{(tmp,tmp)}/\mathrm{tr}\big(\boldsymbol{Q}_k^{(tmp,tmp)}\big)$
   - find optimal user power allocation $\tilde{\boldsymbol{q}}^{(n,m-1)}$ by solving (3.24) and compute $\boldsymbol{Q}_k^{(n,m-1)} = \tilde{q}_k^{(n,m-1)}\bar{\boldsymbol{Q}}_k^{(n,m-1)}$

   5.2 **until** required accuracy is reached or $n \geq n_{\max}$

   5.3 $m \leftarrow m + 1$

   5.4 update $\boldsymbol{W}_k^{(m)} = (\mathbf{E}_k^{\mathrm{UL}(m)})^{-1}$, $r_k^{(m)} = \ln\det(\boldsymbol{W}_k^{(m)})$, $t = \min_k \frac{r_k^{(m)}}{r_k^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$, and $\xi_k^{(m)} = d_k + r_k^{(m)} - r_k^{\circ(m)}$

   5.5 do $n \leftarrow 0$ and set $(.)^{(n_{\max},m-1)} \to (.)^{(0,m)}$ in order to re-enter the inner loop
6. **until** required accuracy is reached or $m \geq m_{\max}$

Hence, $t \geq 1$. Of course, during the convergence $t > 1$. The increasing rate targets $r_k^{\circ(m)}$ constantly catch up with the increasing rates $r_k^{(m)}$. Now, the rates are upper bounded by the single user MIMO rates (using all power), and hence the rates will converge and the sequence t will converge to 1. That means that for at least one user k, $r_k^{(\infty)} = r_k^{\circ(\infty)}$. The question is whether this will be the case for all users, as is required for rate balancing. Now, the WMSE balancing leads at every outer iteration m to $\frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}} = \Delta^{(m)}, \forall k$. At convergence, this becomes $\frac{d_k}{\xi_k^{(\infty)}} = \Delta^{(\infty)}$ where $\xi_k^{(\infty)} = d_k + r_k^{(\infty)} - r_k^{\circ(\infty)}$. Hence, if we have convergence because for one user k_∞ we arrive at $r_{k_\infty}^{(\infty)} = r_{k_\infty}^{\circ(\infty)}$, then this implies $\Delta^{(\infty)} = 1$ which implies $r_k^{(\infty)} = r_k^{\circ(\infty)}, \forall k$. Hence, the rates will be maximized and balanced.

Remark 1. *In fact, the algorithm also converges with $n_{\max} = 1$, i.e., with only a single loop.*

3.5 Results

In this section, we numerically illustrate the performance of the proposed algorithm. The simulations are obtained under a channel modeled as follows : $\boldsymbol{H}_k^{\mathrm{H}} = \boldsymbol{B}_k \boldsymbol{\mathcal{U}}_k \boldsymbol{A}_k$ where $\boldsymbol{B}_k, \boldsymbol{A}_k$ are of dimensions $(M \times N_k)$ and $(N_k \times N_k)$ respectively, and have independent and identically distributed (i.i.d.) elements distributed as $\mathcal{CN}(0,1)$; $\boldsymbol{\mathcal{U}}_k = \mu \boldsymbol{U}_k$, with the normalization parameter $\mu = \left(\mathrm{trace}\left(\boldsymbol{U}_k\right)\right)^{-1/2}$ and $\boldsymbol{U}_k = \mathrm{diag}\left\{1, \alpha, \alpha^2, \ldots, \alpha^{N_k - 1}\right\}$ ($\alpha \in \mathbb{R}$ being a scalar parameter). This model allows to control the rank profile of the MIMO channels. For all simulations, we fix $\alpha = 0.3$ and take 1000 channel realisations and $n_{\max} = 20$. The algorithm converges after 4-5 (or 13-15) iterations of m at SNR $= \frac{P_{\max}}{\sigma_n^2}$ =10dB (or 30dB).

Figure 3.4 plots the minimum achieved per user rate using *i)* our max-min user rate approach with equal user priorities and *ii)* the user MSE balancing approach [75], as a function of the Signal to Noise Ratio (SNR). We observe that our approach outperforms significantly the unweighted MSE balancing optimization, and the gap gets larger with more streams. Note that we observe the same behavior with the classical i.i.d. channel $\boldsymbol{H}_k^{\mathrm{H}} = \boldsymbol{B}_k$, but with a smaller gap (e.g., for 15dB, $\frac{\min_k r_k(\text{weighted-MSE})}{\min_k r_k(\text{unweighted-MSE})} = 1.05$ instead of 1.18 with $M = 6, N_k = d_k = 2$ in Figure 3.4).

In Figure 3.5, we illustrate how rate is distributed among users according to their priorities represented by the rate targets r_k°. We can see that, using the min-max weighted MSE approach, the rate is equally distributed between the users with equal user priorities, i.e., $r_k^\circ = r_1^\circ\ \forall k$, whereas with different user priorities, the rate differs from one user to another accordingly. Furthermore, the Sum Rate (SR) reaches its maximum when user priorities are equal, as the channel statistics are identical for each user.

These results are extended to multi-user IBC case in [76]. Therein, we consider a multi-cell case and solve the user rate balancing problem using diagonal weight matrices by diagonalizing the user signal error covariance matrices, which allows to link the per

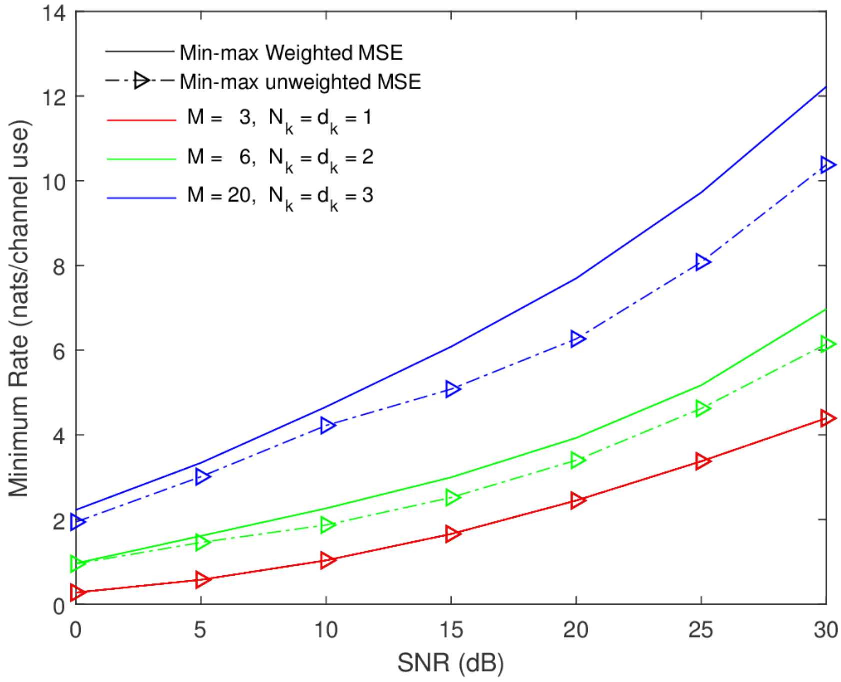

Figure 3.4: Minimum rate in the system vs. SNR: $K = 3$.

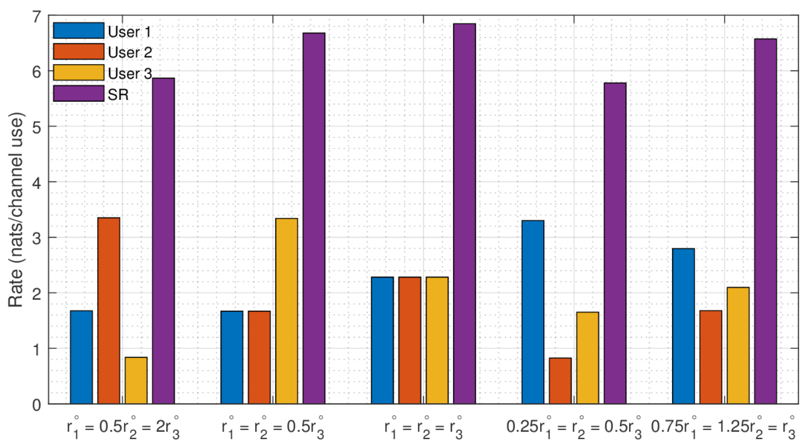

Figure 3.5: Rate distribution among users: $K = 3$, SNR= 10 dB, $M = 6, N_k = d_k = 2$.

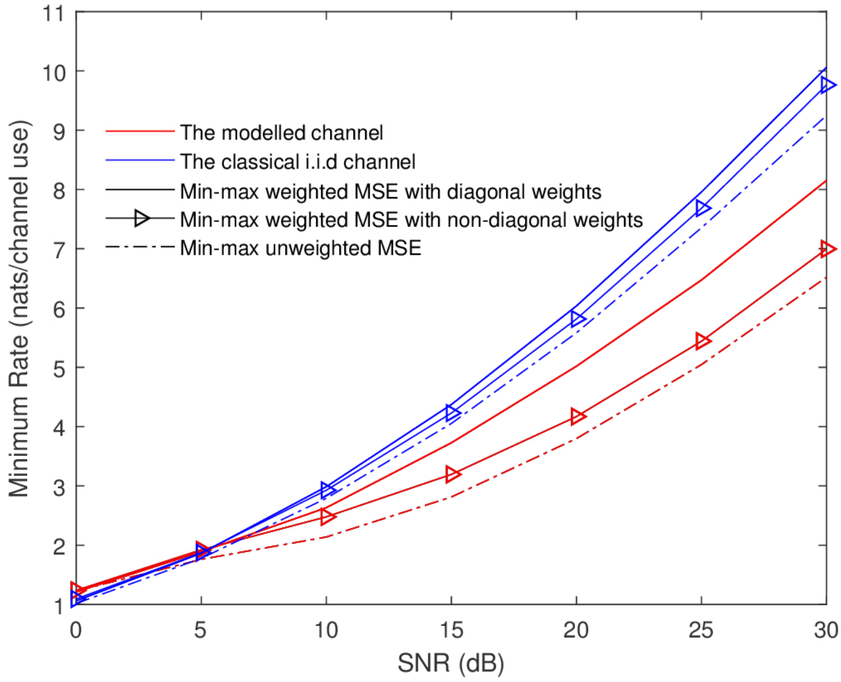

Figure 3.6: Minimum rate in the system vs. SNR: Channel comparison for $K = 3, M = 6, N_k = d_k = 2$.

stream and per user power allocation problems. In fact, define a modified transmit uplink filter as

$$\breve{\boldsymbol{F}}_k \breve{\boldsymbol{Q}}_k^{1/2} = \boldsymbol{F}_k \boldsymbol{Q}_k^{1/2} \mathbf{V}_k \, , \; \breve{\mathbf{E}}_k^{\mathrm{UL}} = \boldsymbol{\Sigma}_k \, , \tag{3.29}$$

where $\boldsymbol{V}_k$ is given by the eigenvalue decomposition $\mathbf{E}_k^{\mathrm{UL}} = \boldsymbol{V}_k \boldsymbol{\Sigma}_k \boldsymbol{V}_k^{\mathrm{H}}$. This operation allows us to diagonalize $\{\mathbf{E}_k^{\mathrm{UL}}, \boldsymbol{W}_k\}$ and does not change the user rates [26], but changes the identity of the streams (layers) of a user and the power distribution over them.

In Figure 3.6, we plot the minimum achieved rate as in Figure , using the classical i.i.d. Gaussian channel and diagonal matrix weights. We observe the same behavior with the classical i.i.d. Gaussian channel, but with a smaller gap. Also, we can see that the balanced rate obtained using diagonal $\{\boldsymbol{W}_k\}$ outperforms the balanced rate derived with non-diagonal weight matrices.

In Figure 3.7, we illustrate how rate is distributed among users according to their priorities represented here by $r_{k_c}^{\circ}$ for multi-cell case. We denote by C, K_c, M_c, the number of cells and number of users and transmit antennas per cell; while N_{k_c} and d_{k_c} denote the number of receive antennas and streams allowed per user k_c, respectively.

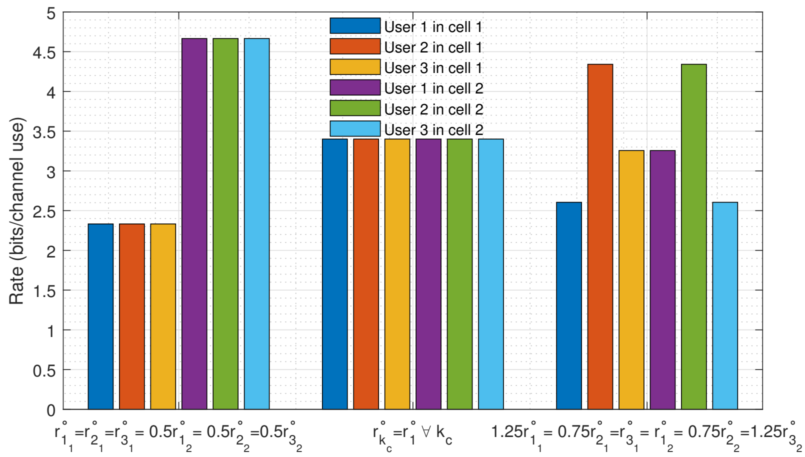

Figure 3.7: Rate distribution among users: SNR= 10 dB, $C = 2, K_c = 3, M_c = 12, N_{k_c} = d_{k_c} = 2$.

3.6 Closing Remarks

In this chapter, we optimized the rate distribution over the streams of a user, within the rate balancing of the users for a single cell multi-user BC. In this regard, we proposed an iterative algorithm to balance the rate between the users in a MIMO system. The latter was derived by transforming the max-min rate optimization problem into a min-max weighted MSE optimization problem to enable MSE duality. Numerical comparisons between the proposed weighted rate balancing approach and unweighted MSE balancing were provided.

Chapter 4

Rate Balancing via Lagrangian duality with per cell power constraints: IBC

4.1 Overview

In the previous chapter, we have investigated the performance improvements of weighted matrix MSE balancing under total power constraints via UL-DL duality. In this chapter, we move to per cell power constraints case, considering the problem of user rate balancing for the downlink transmission of multi-user multi-cell MIMO systems. Due to the multiple streams per user, user rate balancing involves both aspects of balancing and sum rate optimization. We exploit the rate MSE relation, formulate the balancing operation as constraints leading to Lagrangians in optimization duality, allowing to transform rate balancing into weighted MSE minimization with Perron Frobenius theory. The Lagrange multipliers for the multiple power constraints can be formulated as a single weighted power constraint in which the weighting can be optimized to lead to the satisfaction with equality of all power constraints. Actually, various problem formulations are possible, including single cell full power transmission leading to a dual norm optimization problem, and per cell rate balancing which breaks the balancing constraint between cells.

The rest of this chapter is organized as follows. Section 4.2 provides the details of the considered multi-cell IBC. In section 4.3, we introduce the Lagrangian duality to solve the user rate balancing problem. Simulation results are provided in Section 4.4 to validate the proposed algorithms and demonstrate their performance improvement over e.g. unweighted MSE balancing. Finally, conclusions are given in Section 4.5.

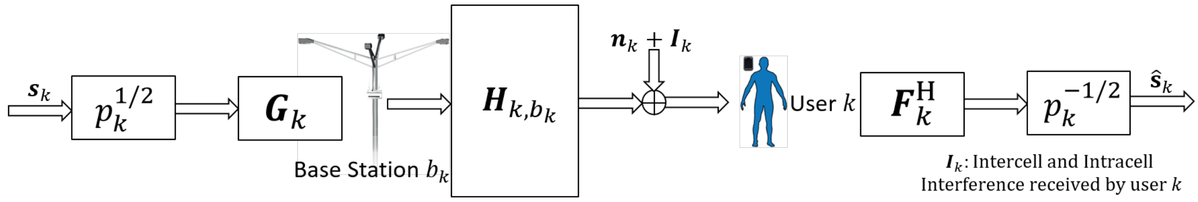

Figure 4.1: Downlink channel of user k in cell b_k.

4.2 System Model and Problem Formulation

We consider a MIMO system with C cells. Each cell c has one base station (BS) of M_c transmit antennas serving K_c users, with total number of users $\sum_c K_c = K$. We refer to the BS of user $k \in \{1, \ldots, K\}$ by b_k. Each user has N_k antennas. The channel between the kth user and the BS in cell j is denoted by $\boldsymbol{H}_{k,j}^{\mathrm{H}} \in \mathbb{C}^{M_j \times N_k}$. We consider zero-mean white Gaussian noise $\boldsymbol{n}_k \in \mathbb{C}^{N_k \times 1}$ with distribution $\mathcal{CN}(0, \sigma_n^2 \boldsymbol{I})$ at the kth user.

We assume independent unity-power transmit symbols $\mathbf{s}_c = [\mathbf{s}_{K_{1:c-1}+1}^{\mathrm{T}} \ldots \mathbf{s}_{K_{1:c}}^{\mathrm{T}}]^{\mathrm{T}}$, i.e., $\mathbb{E}[\mathbf{s}_c \mathbf{s}_c^{\mathrm{H}}] = \boldsymbol{I}$, where $\mathbf{s}_k \in \mathbb{C}^{d_k \times 1}$ is the data vector to be transmitted to the kth user, with d_k being the number of streams allowed by user k and $K_{1:c} = \sum_{i=1}^{c} K_i$. The latter is transmitted using the transmit filtering matrix $\boldsymbol{\mathcal{G}}_c = [\boldsymbol{\mathcal{G}}_{K_{1:c-1}+1} \ldots \boldsymbol{\mathcal{G}}_{K_{1:c}}] \in \mathbb{C}^{M_c \times N_c}$, with $\boldsymbol{\mathcal{G}}_k = p_k^{1/2} \boldsymbol{G}_k$, $\boldsymbol{G}_k$ being the beamforming matrix, p_k is non-negative downlink power allocation of user k and $N_c = \sum_{k:b_k=c} d_k$ is the total number of streams in cell c. Each cell is constrained with $P_{\max,c}$, i.e., the total transmit power in c is limitted such that $\sum_{k:b_k=c} p_k \leq P_{\max,c}$. The received signal at user k in cell b_k is

$$\boldsymbol{y}_k = \underbrace{\boldsymbol{H}_{k,b_k} \boldsymbol{\mathcal{G}}_k \boldsymbol{s}_k}_{\text{signal}} + \underbrace{\sum_{\substack{i \neq k \\ b_i = b_k}} \boldsymbol{H}_{k,b_k} \boldsymbol{\mathcal{G}}_i \boldsymbol{s}_i}_{\text{intracel interf.}} + \underbrace{\sum_{j \neq b_k} \sum_{i:b_i=j} \boldsymbol{H}_{k,j} \boldsymbol{\mathcal{G}}_i \boldsymbol{s}_i}_{\text{intercell interf.}} + \boldsymbol{n}_k \tag{4.1}$$

Similarly, the receive filtering matrix for each user k is defined as $\boldsymbol{\mathcal{F}}_k^{\mathrm{H}} = p_k^{-1/2} \boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$, composed of beamforming matrix $\boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$. The received filter output is $\hat{\boldsymbol{s}}_k = \boldsymbol{\mathcal{F}}_k^{\mathrm{H}} \boldsymbol{y}_k$. Figure 4.1 illustrates the described model.

4.2.1 Problem Formulation

In this work, we aim to solve the weighted user-rate max-min optimization problem under per cell total transmit power constraint, i.e., the user rate balancing problem expressed as follows

$$\begin{aligned} &\max_{\boldsymbol{G},p} \min_{k} \; r_k / r_k^{\circ} \\ &\text{s.t.} \sum_{k:b_k=c} p_k \leq P_{\max,c}, 1 \leq c \leq C \end{aligned} \tag{4.2}$$

where r_k is the kth user-rate

$$r_k = \ln\det\Big(\boldsymbol{I}+\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_k}^{\mathrm{H}}\Big) = \ln\det\Big(\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{R}_k\Big)\,, \tag{4.3}$$

$$\boldsymbol{R}_{\overline{k}} = \sigma_n^2\boldsymbol{I}+\sum_{l\neq k}\boldsymbol{H}_{k,b_l}\boldsymbol{\mathcal{G}}_l\boldsymbol{\mathcal{G}}_l^{\mathrm{H}}\boldsymbol{H}_{k,b_l}^{\mathrm{H}}\,, \tag{4.4}$$

$$\boldsymbol{R}_k = \boldsymbol{R}_{\overline{k}}+\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_k}^{\mathrm{H}}\,, \tag{4.5}$$

$\boldsymbol{R}_{\overline{k}}$ and $\boldsymbol{R}_k$ are the interference plus noise and total received signal covariances, and r_k° is the rate priority (weight) for user k. However, the problem presented in (4.2) is complex and can not be solved directly.

Lemma 2. *The rate of user k in (4.3) can also be represented as [27]*

$$r_k = \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k}\big[\ln\det\big(\boldsymbol{W}_k\big) - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k\big) + d_k\big]. \tag{4.6}$$

where

$$\begin{aligned}\mathbf{E}_k &= \mathbb{E}\Big[(\hat{\mathbf{s}}_k-\mathbf{s}_k)(\hat{\mathbf{s}}_k-\mathbf{s}_k)^{\mathrm{H}}\Big]\\ &= (\boldsymbol{I}-\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k)(\boldsymbol{I}-\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k)^{\mathrm{H}} + \sum_{l\neq k}\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_l}\boldsymbol{\mathcal{G}}_l\boldsymbol{\mathcal{G}}_l^{\mathrm{H}}\boldsymbol{H}_{k,b_l}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k + \sigma_n^2\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k\end{aligned} \tag{4.7}$$

is the kth-user downlink MSE matrix between the decision variable $\hat{\mathbf{s}}_k$ and the transmit signal $\mathbf{s}_k$, and $\{\boldsymbol{W}_k\}_{1\leq k\leq K}$ are auxiliary weight matrix variables with optimal solution $\boldsymbol{W}_k^{\mathrm{opt}} = \big[\mathbf{E}_k\big]^{-1}$ and the optimal receivers are

$$\boldsymbol{\mathcal{F}}_k = \boldsymbol{R}_k^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k. \tag{4.8}$$

Now consider both (4.2) and (4.6), and let us introduce $\xi_k = \ln\det\big(\boldsymbol{W}_k\big) + d_k - r_k^{\triangle}$, the WMSE requirement, with target rate $r_k^{\triangle}$. Assume that we shall be able to concoct an optimization algorithm that ensures that at all times and for all users the matrix-weighted MSE (WMSE) satisfies $\epsilon_{w,k} = \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k\big) \leq d_k$ and $\ln\det\big(\boldsymbol{W}_k\big) \geq r_k^{\triangle}$ or hence $\xi_k \geq d_k$. This leads $\forall k$ to

$$\begin{aligned}\frac{\epsilon_{w,k}}{\xi_k} \leq 1 \iff \ln\det\big(\boldsymbol{W}_k\big) + d_k - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k\big) \geq r_k^{\triangle}\\ \stackrel{(a)}{\Longrightarrow} r_k/r_k^{\triangle} \geq 1\end{aligned} \tag{4.9}$$

where (a) follows from (4.6). To get to (4.9), what we can exploit in (4.2) is a scale factor t that can be chosen freely in the rate weights r_k° in (4.2). We shall take $t = \min_k r_k/r_k^{\circ}$, which allows to transform the rate weights r_k° into target rates $r_k^{\triangle} = tr_k^{\circ}$, and at the same time allows to interpret the WMSE weights ξ_k as target WMSE values.

Doing so, the initial rate balancing optimization problem (4.2) can be transformed into a matrix-weighted MSE balancing problem expressed as follows

$$\begin{aligned}&\min_{\boldsymbol{G},p,\boldsymbol{\mathcal{F}}}\max_{k}\ \epsilon_{w,k}/\xi_k\\ \text{s.t.}\ &\sum_{k:b_k=c}p_k\le P_{\max,c},1\le c\le C\end{aligned}\tag{4.10}$$

which needs to be complemented with an outer loop in which $\boldsymbol{W}_k=\big(\mathbf{E}_k\big)^{-1}$, $t=\min_k r_k/r_k^{\circ}$, $r_k^{\triangle}=tr_k^{\circ}$ and $\xi_k=d_k+r_k-r_k^{\triangle}$ get updated. The problem in (7.13) is still difficult to be handled directly.

4.3 Proposed Solution

In this section, the problem (4.10) with respect to the matrix weighted user-MSE is studied. The per user matrix weighted MSE (WMSE) can be expressed as follows

$$\begin{aligned}\epsilon_{w,k}&=\mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k\big)\\&=\mathrm{tr}\big(\boldsymbol{W}_k\big)-\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_k}^{\mathrm{H}}\boldsymbol{F}_k\big)-\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\big)\\&+p_k^{-1}\sum_{l=1}^{K}p_l\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{H}_{k,b_l}\boldsymbol{G}_l\boldsymbol{G}_l^{\mathrm{H}}\boldsymbol{H}_{k,b_l}^{\mathrm{H}}\boldsymbol{F}_k\big)+\sigma_n^2p_k^{-1}\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{F}_k\big).\end{aligned}\tag{4.11}$$

Define the diagonal matrix $\boldsymbol{D}$ of signal WMSE contributions

$$\begin{aligned}[\boldsymbol{D}]_{ii}&=\mathrm{tr}\big(\boldsymbol{W}_i\big)-\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{G}_i^{\mathrm{H}}\boldsymbol{H}_{i,b_i}^{\mathrm{H}}\boldsymbol{F}_i\big)-\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{H}_{i,b_i}\boldsymbol{G}_i\big)\\&+\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{H}_{i,b_i}\boldsymbol{G}_i\boldsymbol{G}_i^{\mathrm{H}}\boldsymbol{H}_{i,b_i}^{\mathrm{H}}\boldsymbol{F}_i\big),\end{aligned}\tag{4.12}$$

and the matrix of weighted interference powers

$$[\boldsymbol{\Psi}]_{ij}=\begin{cases}\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{H}_{i,b_j}\boldsymbol{G}_j\boldsymbol{G}_j^{\mathrm{H}}\boldsymbol{H}_{i,b_j}^{\mathrm{H}}\boldsymbol{F}_i\big), & i\neq j\\ 0, & i=j.\end{cases}\tag{4.13}$$

We can rewrite (4.11) as, with $\boldsymbol{p}=[p_1\cdots p_K]^{\mathrm{T}}$

$$\epsilon_{w,i}=[\boldsymbol{D}]_{ii}+p_i^{-1}[\boldsymbol{\Psi}\boldsymbol{p}]_i+\sigma_n^2p_i^{-1}\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{F}_i\big)\tag{4.14}$$

Collecting all user WMSEs in a vector $\boldsymbol{\epsilon}_w=\mathrm{diag}(\epsilon_{w,1},\ldots,\epsilon_{w,K})$, we get

$$\boldsymbol{\epsilon}_w\mathbf{1}_K=\mathrm{diag}(\boldsymbol{p})^{-1}\left[(\boldsymbol{D}+\boldsymbol{\Psi})\mathrm{diag}(\boldsymbol{p})\mathbf{1}_K+\boldsymbol{\sigma}\right]\tag{4.15}$$

where $\boldsymbol{\sigma}$ is a $(K\times 1)$vector defined as

$$[\boldsymbol{\sigma}]_i=\sigma_n^2\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{F}_i\big).\tag{4.16}$$

By multiplying both sides of (4.15) with diag($\boldsymbol{p}$), we get

$$\boldsymbol{\epsilon}_w \boldsymbol{p} = (\boldsymbol{D} + \boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\sigma}. \tag{4.17}$$

Let $\boldsymbol{\xi} = \mathrm{diag}(\xi_1, \ldots, \xi_K)$, then

$$\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w\, \boldsymbol{p} = \boldsymbol{\xi}^{-1}(\boldsymbol{D} + \boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\xi}^{-1}\boldsymbol{\sigma}. \tag{4.18}$$

Actually, problem (4.10) always has a global minimizer $\boldsymbol{p}$ characterized by the equality $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w(\boldsymbol{p}) = \Delta\boldsymbol{I}$, i.e.,

$$\Delta\, \boldsymbol{p} = \boldsymbol{\xi}^{-1}(\boldsymbol{D} + \boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\xi}^{-1}\boldsymbol{\sigma}. \tag{4.19}$$

Now, consider the following problem

$$\begin{aligned} &\max_{\boldsymbol{G},p,\boldsymbol{\mathcal{F}}} \min_{k} \quad r_k/r_k^{\circ} \\ &\text{s.t.} \ \sum_{c=1}^{C} \theta_c \boldsymbol{c}_c^{\mathrm{T}} \boldsymbol{p} \leq \sum_{c=1}^{C} \theta_c P_{\max,c} \end{aligned} \tag{4.20}$$

where $\boldsymbol{c}_c$ is a column vector with $\boldsymbol{c}_c(j) = 1$ for $K_{1:c-1} + 1 \leq j \leq K_{1:c}$, and 0 elsewhere. This problem formulation is a relaxation of (4.2), and $\boldsymbol{\theta} = [\theta_1 \cdots \theta_C]^{\mathrm{T}}$can be interpreted as the weights on the individual power constraints in the relaxed problem. The power constraint in (4.20) can be interpreted as a single weighted power constraint

$$(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}})\, \boldsymbol{p} \leq \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max} \tag{4.21}$$

with $\boldsymbol{C}_C = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_C] \in \mathbb{R}_+^{K_{1:C} \times C}$ and $\boldsymbol{p}_{\max} = [P_{\max,1} \cdots P_{\max,C}]^{\mathrm{T}}$. Reparameterize $\boldsymbol{p} = \frac{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}}\boldsymbol{p}'}\boldsymbol{p}'$ where now $\boldsymbol{p}'$ is unconstrained, which allows us to write (4.19) as follows (rewriting $\boldsymbol{p}'$ as $\boldsymbol{p}$)

$$\Delta\, \boldsymbol{p} = \boldsymbol{\Lambda}\, \boldsymbol{p} \text{ with } \boldsymbol{\Lambda} = \boldsymbol{\xi}^{-1}(\boldsymbol{D} + \boldsymbol{\Psi}) + \frac{1}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}\boldsymbol{\xi}^{-1}\boldsymbol{\sigma}\, \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}}. \tag{4.22}$$

Now with (4.22), the WMSE balancing problem of (4.10) becomes

$$\min_{\boldsymbol{p}} \max_{k} \frac{\epsilon_{w,k}}{\xi_k} = \min_{\boldsymbol{p}} \max_{k} \frac{[\boldsymbol{\Lambda}\, \boldsymbol{p}]_k}{p_k} \tag{4.23}$$

According to the Collatz–Wielandt formula [77, Chapter 8], the above expression corresponds to the Perron-Frobenius (maximal) eigenvalue Δ of $\boldsymbol{\Lambda}$ and the optimal $\boldsymbol{p}$ is the corresponding Perron-Frobenius (right) eigenvector

$$\boldsymbol{\Lambda}\boldsymbol{p} = \Delta\, \boldsymbol{p}. \tag{4.24}$$

Note that this implies the equality $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w = \Delta\, \boldsymbol{I}$ as announced in (4.19).

4.3.1 Algorithmic Solution via Lagrangian Duality

The max-min weighted user rate optimization problem (4.2) can be reformulated as

$$\begin{aligned} &\min_{t,\boldsymbol{G},\boldsymbol{p}} -t \\ &\text{ s.t. } \; t\, r_k^\circ - r_k \leq 0,\; \boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c} \leq 0\,. \end{aligned} \tag{4.25}$$

Introducing Lagrange multipliers to augment the cost function with the constraints leads to the Lagrangian

$$\begin{aligned} &\max_{\lambda',\mu} \min_{t,\boldsymbol{G},\boldsymbol{p}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k'(t\, r_k^\circ - r_k) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \tag{4.26}$$

Integrating the result (4.6), we get a modified Lagrangian

$$\begin{aligned} &\max_{\lambda',\mu} \min_{t,\boldsymbol{G},\boldsymbol{p},\boldsymbol{\mathcal{F}},\boldsymbol{W}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k'(\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k) - \xi_k) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \tag{4.27}$$

From (4.20), we get $\mu_c = \mu\theta_c$. Introducing $\lambda_k = \lambda_k'\xi_k$, we can rewrite (with some abuse of notation since actually $\min_{\boldsymbol{W}}$ continues to apply to $\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k) - \xi_k(\boldsymbol{W}_k)$)

$$\begin{aligned} &\max_{\lambda,\mu} \min_{t,\boldsymbol{G},\boldsymbol{p},\boldsymbol{F},\boldsymbol{W}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k(\frac{\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k)}{\xi_k} - 1) + \mu\sum_c \theta_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \tag{4.28}$$

We shall solve this saddlepoint condition for $\mathcal{L}$ by alternating optimization. As far as the dependence on $\lambda, \boldsymbol{G}, \boldsymbol{p}, \boldsymbol{\mathcal{F}}$ is concerned, we have (omitting the power constraint)

$$\max_{\lambda} \min_{\boldsymbol{G},\boldsymbol{p},\boldsymbol{\mathcal{F}}} \sum_k \lambda_k \frac{\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k)}{\xi_k} \tag{4.29}$$

which is of the form Weighted Sum MSE (WSMSE). Optimizing w.r.t. Rxs $\boldsymbol{\mathcal{F}}_k$ leads to the MMSE solution mentioned in Lemma 1. To optimizing w.r.t. the Txs $\boldsymbol{\mathcal{G}}_k$, we can follow the approach in [26], which is based on [78], but needs to be adapted to a weighted sum power constraint. We get a shorter derivation by following [20]. To that end, consider a reparameterization of the Tx filters to inherently satisfy the power constraint (see (4.21) where $p_i = \text{tr}\{\boldsymbol{\mathcal{G}}_i^{\text{H}}\boldsymbol{\mathcal{G}}_i\}$) :

$$\boldsymbol{\mathcal{G}}_k = \sqrt{\frac{\boldsymbol{\theta}^{\text{T}}\boldsymbol{p}_{\max}}{\sum_i \theta_{b_i}\text{tr}\{\mathbf{G}_i^{\text{H}}\mathbf{G}_i\}}}\,\mathbf{G}_k\ . \tag{4.30}$$

involving a unique system-wide scale factor (and note $\mathbf{G}_k \neq \boldsymbol{G}_k$). Introducing (4.30) directly into (4.29) does not lead to a quadratic criterion. However, reinterpreting the WSMSE (4.29) as a weighted sum rate via Lemma 2, we get

$$\text{WSR} = \sum_k \frac{\lambda_k}{\xi_k}\,\ln\det(\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{R}_k) \tag{4.31}$$

with $\boldsymbol{R}_{\overline{k}}$ as in (4.4) but with $\boldsymbol{\mathcal{G}}_i$ replaced by $\mathbf{G}_i$ and with the noise covariance term replaced by

$$\frac{\sum_i \theta_{b_i}\text{tr}\{\mathbf{G}_i^{\text{H}}\mathbf{G}_i\}}{\boldsymbol{\theta}^{\text{T}}\boldsymbol{p}_{\max}}\Sigma_k \tag{4.32}$$

where in fact $\Sigma_k = \sigma_n^2\boldsymbol{I}$. Using $\partial\ln\det(\boldsymbol{A}) = \text{tr}\{\boldsymbol{A}^{-1}\,\partial\boldsymbol{A}\}$ and e.g. $(\boldsymbol{R}_{\overline{i}}^{-1}\boldsymbol{R}_i)^{-1}\boldsymbol{R}_{\overline{i}}^{-1} = \boldsymbol{R}_i^{-1}$, we get

$$\begin{aligned}\tfrac{\partial\text{WSR}}{\partial\mathbf{G}_k^*} &= 0 = \tfrac{\lambda_k}{\xi_k}\boldsymbol{H}_{k,b_k}^{\text{H}}\boldsymbol{R}_k^{-1}\boldsymbol{H}_{k,b_k}\mathbf{G}_k \\ &-\big(\textstyle\sum_{i\neq k}\tfrac{\lambda_i}{\xi_i}\boldsymbol{H}_{i,b_k}^{\text{H}}\boldsymbol{R}_{\overline{i}}^{-1}\boldsymbol{H}_{i,b_k}\mathbf{G}_i\mathbf{G}_i^{\text{H}}\boldsymbol{H}_{i,b_k}^{\text{H}}\boldsymbol{R}_i^{-1}\boldsymbol{H}_{i,b_k}\big)\,\mathbf{G}_k \\ &-\big(\textstyle\sum_{i=1}^{K}\tfrac{\lambda_i}{\xi_i}\text{tr}\{\Sigma_i\boldsymbol{R}_{\overline{i}}^{-1}\boldsymbol{H}_{i,b_k}\mathbf{G}_i\mathbf{G}_i^{\text{H}}\boldsymbol{H}_{i,b_k}^{\text{H}}\boldsymbol{R}_i^{-1}\}\big)\,\theta_{b_k}\mathbf{G}_k\end{aligned} \tag{4.33}$$

Now if we note that $\boldsymbol{\mathcal{F}}_i = \boldsymbol{R}_i^{-1}\boldsymbol{H}_{i,b_i}\mathbf{G}_i = \boldsymbol{R}_{\overline{i}}^{-1}\boldsymbol{H}_{i,b_i}\mathbf{G}_i\boldsymbol{E}_i$, $\boldsymbol{W}_i = \boldsymbol{E}_i^{-1}$ and $\boldsymbol{R}_k^{-1}\boldsymbol{H}_{k,b_k}\mathbf{G}_k = \boldsymbol{\mathcal{F}}_k = \boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k\boldsymbol{E}_k = \boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k(\boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\text{H}}\boldsymbol{H}_{k,b_k}\mathbf{G}_k)$, then (4.33) leads to

$$\begin{aligned}\mathbf{G}_k &= \Big(\sum_{l=1}^{K}\boldsymbol{H}_{l,b_k}^{\text{H}}\boldsymbol{\mathcal{F}}_l\boldsymbol{W}_l'\boldsymbol{\mathcal{F}}_l^{\text{H}}\boldsymbol{H}_{l,b_k} + \sigma_n^2\theta_{b_k}\frac{\sum_l\text{tr}(\boldsymbol{W}_l'\boldsymbol{\mathcal{F}}_l^{\text{H}}\boldsymbol{\mathcal{F}}_l)}{\sum_c\theta_c P_{\max,c}}\boldsymbol{I}\Big)^{-1}\boldsymbol{H}_{k,b_k}^{\text{H}}\boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k' \\ \boldsymbol{\mathcal{G}}_k &= \sqrt{p_k}\,\boldsymbol{G}_k\,,\boldsymbol{G}_k = \frac{1}{\sqrt{\text{tr}\{\mathbf{G}_k^{\text{H}}\mathbf{G}_k\}}}\,\mathbf{G}_k\end{aligned} \tag{4.34}$$

where $\boldsymbol{W}_k' = \lambda_k/\xi_k\boldsymbol{W}_k$, and accounting for the fact that the user powers are actually optimized by the Perron-Frobenius theory. Note that this result for $\boldsymbol{G}_k$ would also be obtained by direct optimization of (4.28), but we needed the extra development above to get the expression for the Lagrange multiplier μ. The Perron-Frobenius theory also allows for the optimization of the Lagrange multipliers λ_k. With (4.23), we can reformulate (4.29) as

$$\Delta = \max_{\lambda:\sum_k\lambda_k=1}\min_{\boldsymbol{p}}\sum_k\lambda_k\frac{[\boldsymbol{\Lambda}\,\boldsymbol{p}]_k}{p_k} \tag{4.35}$$

which is the Donsker–Varadhan–Friedland formula [77, Chapter 8] for the Perron Frobenius eigenvalue of Λ. A related formula is the Rayleigh quotient

$$\Delta = \max_{\boldsymbol{q}}\min_{\boldsymbol{p}}\frac{\boldsymbol{q}^T\boldsymbol{\Lambda}\,\boldsymbol{p}}{\boldsymbol{q}^T\boldsymbol{p}} \tag{4.36}$$

where $\boldsymbol{p}$, $\boldsymbol{q}$ are the right and left Perron Frobenius eigenvectors. Comparing (4.36) to (4.35), then apart from normalization factors, we get $\lambda_k/p_k = q_k$ or hence $\lambda_k = p_k q_k$.

The proposed optimization framework is summarized in Table 4.1; Table 4.2 represents the power optimization algorithm ensuring the per cell power constraints. Superscripts refer to iteration numbers. The algorithm in Table 4.1 is based on a double loop. The inner loop solves the WMSE balancing problem in (4.10) whereas the outer loop iteratively transforms the WMSE balancing problem into the original rate balancing problem in (4.2).

4.4 Results

In this section, we numerically evaluate the performance of the proposed algorithm. The simulations are carried out over normalized flat fading channels, i.e., each element is i.i.d. and normally distributed : starting from i.i.d. $[\boldsymbol{H}_{k,j}]_{mn} \sim \mathcal{CN}(0,1)$. For all simulations, we take 500 channel realisations and $n_{\max} = 20$. The algorithm converges after 4-5 (or 13-15) iterations of m at SNR = 15dB (or 30dB).

In Figure 4.2 (only), the singular values are modified to a geometric series α^i to control the MIMO channel conditioning, in particular $\alpha = 0.3$. Fig. 4.2 plots the minimum achieved per user rate using *i)* our max-min user rate approach for equal priorities with total sum-power constraint and per cell power constraints, and *ii)* the user MSE balancing approach [75] w.r.t. the Signal to Noise Ratio (SNR). We observe that our approach outperforms significantly the unweighted MSE balancing optimization. Furthermore, the gap between the achieved balanced rate using per cell power constraints and the one obtained with total sum-power constraint over cells is very tiny.

In Figure 4.3, which illustrates the difference between the per cell power constraint $P_{\max,c}$ and the total power allocated per cell, i.e., $\boldsymbol{c}_{K_c,c}^{\mathrm{T}}\boldsymbol{p}$, for per cell power contraints and total sum-power constraint, we observe that using Table 4.2 we ensure the per cell power constraint with equality, unlike the total sum-power constraint which verifies the total power over cells.

In Figure 4.4, we illustrate how rate is distributed among users according to their priorities represented by the rate targets r_k°. We can see that, using the min-max weighted MSE approach, the rate is equally distributed between the users with equal user priorities, i.e., $r_k^\circ = r_1^\circ \ \forall k$, whereas with different user priorities, the rate differs from one user to another accordingly.

4.5 Closing Remarks

In this chapter, we addressed the multiple streams per user case (MIMO links) for which we considered user rate balancing, not stream rate balancing, in multi-cell downlink channel. Actually, we optimized the rate distribution over the streams of a user, within the rate

Table 4.1: User Rate Balancing via Lagrangian Duality Algorithm

1. initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0,0)} = \boldsymbol{q}_k^{(0,0)} = \frac{P_{\max,c}}{K_c}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$ and $r_k^{\circ(0)}$, initialize $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}_{d_k}$ and $\xi_k^{(0)} = d_k$
2. initialize $\boldsymbol{F}_k^{(0,0)}$ in $\boldsymbol{\mathcal{F}}_k^{(0,0)} = p_k^{(0,0)-1/2}\boldsymbol{F}_k$ from (4.8)
3. **repeat**

   3.1. $m \leftarrow m + 1$

   3.2. **repeat**

   $n \leftarrow n + 1$

   i update $\boldsymbol{G}_k$ in $\boldsymbol{\mathcal{G}}_k = p_k^{1/2}\boldsymbol{G}_k$ from (4.34)

   ii update $\boldsymbol{F}_k$ in $\boldsymbol{\mathcal{F}}_k = p_k^{-1/2}\boldsymbol{F}_k$ from (4.8)

   iii update $\boldsymbol{p}$ and $\boldsymbol{q}$ using Table 4.2

   3.3 **until** required accuracy is reached or $n \geq n_{\max}$

   3.4 compute $\boldsymbol{E}_k^{(m)}$ with (4.7) and update $\boldsymbol{W}_k^{(m)} = (\boldsymbol{E}_k^{(m)})^{-1}$

   3.5 compute $r_k^{(m)} = \ln\det(\boldsymbol{W}_k^{(m)})$ and determine $t = \min_k \frac{r_k^{(m)}}{r_{k_c}^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$,
   and $\xi_k^{(m)} = d_k + r_k^{(m)} - r_k^{\circ(m)}$

   3.6 set $n \leftarrow 0$ and set $(.)^{(n_{\max},m-1)} \rightarrow (.)^{(0,m)}$ in order to re-enter the inner loop

4. **until** required accuracy is reached or $m \geq m_{\max}$

Table 4.2: Power Distribution Optimization

1. for given $\boldsymbol{\theta}^{(0)}, \alpha = \frac{\alpha_0}{\sum_c P_{\max,c}}$, $\mathcal{C} = \{\boldsymbol{\theta} \in \mathbb{R}^{C}_{\geq 0} : \boldsymbol{\theta} \gneq 0,\ \mathbf{1}_C^{\mathrm{T}} \boldsymbol{\theta} = 1\}$
2. **repeat**

    2.1. compute $\boldsymbol{\Lambda}(\boldsymbol{\theta})$, update $\boldsymbol{p}$ and $\boldsymbol{q}$ as right and left Perron Frobenius eigen vectors of $\boldsymbol{\Lambda}$

    2.2. update $\boldsymbol{\theta}$ using the subgradient projection method, [79] :

    $$\boldsymbol{\theta}^{(i+1)} = \mathcal{P}_c\big\{\boldsymbol{\theta}^{(i)} - \alpha \hat{\boldsymbol{g}}(\boldsymbol{p}^{(i)})\big\},$$

    where $\hat{\boldsymbol{g}}(\boldsymbol{p}^{(i)}) = \boldsymbol{p}_{\max} - \boldsymbol{C}_C^{\mathrm{T}} \boldsymbol{p}^{(i)}$ and $\mathcal{P}_{\mathcal{C}}$ is the projection operator onto $\mathcal{C}$.

    2.3. $i \leftarrow i + 1$

3. **until** required accuracy is reached

balancing of the users under per cell power constraints. In this regard, we proposed an iterative algorithm to balance the rate between the users in a MIMO system. The latter was derived by transforming the max-min rate optimization problem into a min-max weighted MSE optimization problem which itself was shown to be related to a weighted sum MSE minimization via Langrangian duality. Moreover, we reformulated the multiple power constraints as a single weighted constraint satisfying with equality of all power contraints. We provided comparison between our weighted MSE balancing approach and the min-max unweighted MSE optimization. Simulation results showed that our solution maximizes the minimum rate.

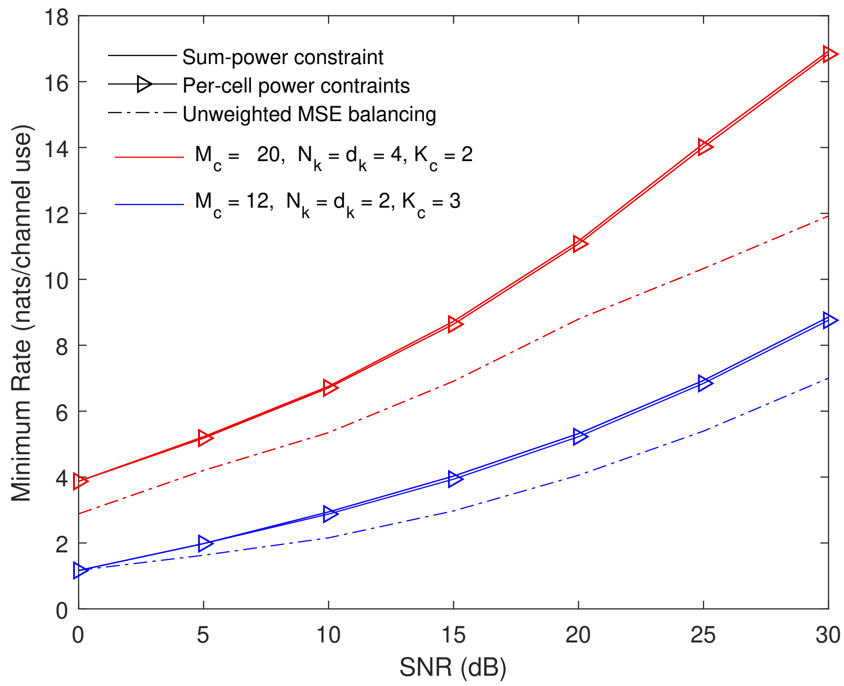

Figure 4.2: Minimum rate in the system vs. SNR, $C = 2$.

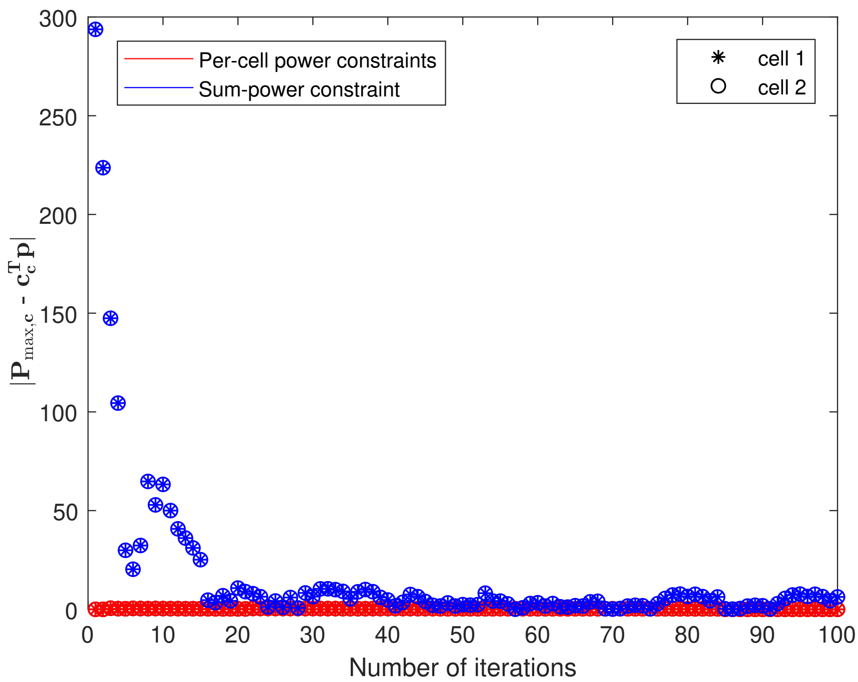

Figure 4.3: Difference between the per cell total power and its respective power constraint vs. number of iterations: SNR = 30 dB, $C = 2, K_c = 3, M_c = 20, N_k = d_k = 2, P_{\max,1} = P_{\max,2}$.

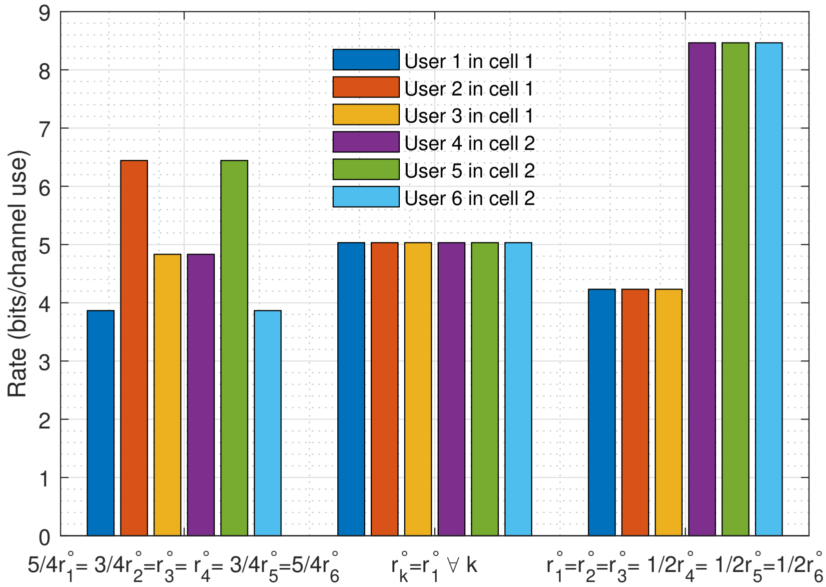

Figure 4.4: Rate distribution among users: SNR= 15 dB, $C = 2, K_c = 3, M_c = 12, N_k = d_k = 2$.

Chapter 5

From Rate Balancing to Power Minimization

5.1 Overview

In the previous chapters, we have focused on (weighted) user rate balancing. So far, we have considered the weighting factors as user priorities since the considered optimization aims to maximize (as large as possible) the minimum user rate in the system w.r.t. the users priorities. However, these weights can represent target rates, and every set of these targets also corresponds to a point on the boundary of the achievable rate region, which is defined as the set of all feasible rate points, when all users are active simultaneously under a total power constraint. Each point on the boundary corresponds to an optimal transmission strategy, depending on the chosen weights. When the achievable rates are larger than the target rates, a power efficient strategy w.r.t. that individual rate targets are jointly achieved with minimum total transmit power. In this Chapter, we study the power minimization problem w.r.t. per user rate constraints, via rate balancing optimization.

The rest of the chapter is organized as follows. In Section 5.2 we formulate the power minimization problem w.r.t. per user rate targets. Solution via rate balancing is provided in Section 5.3, using either UL/DL duality or Lagrangian duality. We present the numerical results for the solution in Section 5.4, and finally, conclude in Section 5.5.

5.2 Problem Formulation

Under the same system model assumptions as in Section 3.2, we consider downlink BC. We have so far considered the rate balancing problem, we refer to this problem as Max-Min Rate (MMR), namely

$$\text{MMR:} \quad \max \min_{1\le k\le K} r_k/r_k^o$$

$$\text{under total power constraints } P_{\max},$$

where r_k and r_k^o denote the individual user rate and the corresponding priority, respectively.

Here, we consider the total transmit power minimization (PM) problem. In fact, when having a set of balanced $\{r_k/r_k^o\}_{1\le k\le K}$, we can consider the following problem

$$\text{PM:} \quad \text{minimize the total transmission power } P$$

$$\text{while fulfilling } \min_k \tfrac{r_k}{r_k^o} = \tfrac{r_k}{r_k^o} = 1 \ \ \forall k.$$

where $P = ||\boldsymbol{p}||_1$ is the total transmit power.

This optimization is interesting from a network operator's perspective. In fact, it minimizes intercell interference and improves the power efficiency of the system. In order to make the PM problem more practical, let us consider the user priorities r_k^o, as a set of defined user target rates, i.e., $r_k^o = r_k^\triangle$. The latters are considered as feasible if and only if the optimum of MMR is greater than or equal to one, i.e.,

$$\frac{r_k}{r_k^o} \ge 1, \quad \forall k.$$

While optimizing the PM problem, we have to take into account that the predefined target rates may be unsupported along with the power minimization. Therefore, the design of the algorithmic solution for PM should be in a two-stage approach: First test for feasibility, then minimize the transmission power. In case the rate targets are infeasible, the user rates are *fairly* balanced between users according to their targets, without reaching them. In other words, users achieve reduced rates. If this drop in rates is important, resource management is needed to properly relax the initial conditions (e.g., by reducing the number of users).

We consider the following optimization problem

$$\begin{aligned} \min_{\boldsymbol{p},\boldsymbol{G}} \ & f(||\boldsymbol{p}||_1, t) \\ \text{s.t.} \ & r_k(\boldsymbol{p},\boldsymbol{G})/r_k^\triangle \ge t, \ \ \forall k \\ & ||\boldsymbol{p}||_1 \le P_{\max} \end{aligned} \tag{5.1}$$

where

$$f(||\boldsymbol{p}||_1, t) = u(t-1)(||\boldsymbol{p}||_1 + t) - t \tag{5.2}$$

$$\text{with } u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0. \end{cases} \tag{5.3}$$

The problem in (5.1) describes MMR problem when $t < 1$. When $t \geq 1$ (5.1) becomes as follows

$$\begin{aligned} P^{\text{opt}} = \min_{\boldsymbol{p},\boldsymbol{G}} \quad & ||\boldsymbol{p}||_1 \\ \text{s.t.} \quad & r_k(\boldsymbol{p}, \boldsymbol{G})/r_k^o \geq 1, \quad \forall k \\ & ||\boldsymbol{p}||_1 \leq P_{\max} \end{aligned} \tag{5.4}$$

5.3 Proposed Solution

Let us denote the function $\mathcal{R}(P_{\max}, \boldsymbol{G})$ as follows

$$\begin{aligned} \mathcal{R}(P_{\max}) = \max_{\boldsymbol{G}} \mathcal{R}(P_{\max}, \boldsymbol{G}) = \max_{\boldsymbol{p},\boldsymbol{G}} \min_{k} \quad & r_k(\boldsymbol{p}, \boldsymbol{G})/r_k^o \\ \text{s.t.} \quad & ||\boldsymbol{p}||_1 \leq P_{\max}, \quad \forall k \end{aligned} \tag{5.5}$$

such that, at iteration (i), we have

$$\min_k \frac{r_k(\boldsymbol{p}^{(i-1)}, \boldsymbol{G}^{(i)})}{r_k^o} \leq \mathcal{R}(P_{\max}, \boldsymbol{G}^{(i)}) \leq \max_k \frac{r_k(\boldsymbol{p}^{(i-1)}, \boldsymbol{G}^{(i)})}{r_k^o},$$

and at convergence

$$\min_k \frac{r_k(\boldsymbol{p}, \boldsymbol{G})}{r_k^o} = \mathcal{R}(P_{\max}, \boldsymbol{G}) = \max_k \frac{r_k(\boldsymbol{p}, \boldsymbol{G})}{r_k^o}.$$

Let us now assume that $\mathcal{R}(P_{\max}, \boldsymbol{G}) > 1$ holds, then $t > 1$. In other words, the rate targets are feasible; thus, we have additional degrees of freedom that can be used to minimize the total transmission power. In fact, the PM problem is closely related to the MMR problem. Both of them become equivalent if we set $P_{\max} = P^{\text{opt}}$ in (5.5). Therefore, a modified version of the algorithms solving (5.5) from Chapters 3 and 4 can be used to solve PM in (5.4).

The designed algorithm is summarized within two steps:

- First, we have to make sure that the predefined targets $r_1^o, \ldots, r_K^o$ are feasible. In other words, there exists at least one iteration n which verifies $t^{(n)} \geq 1$. For that, the algorithm iterates the same steps as for the rate balancing problem in Tables 3.1 and 4.1 until the condition is verified. In case the targets remain infeasible, i.e., $t^{(n)} < 1$ for $n \to \infty$, we must relax the initial conditions.

- The second step is taken into consideration only if the targets are feasible. Thus, the condition $t^{(n)} \geq 1$ here is fulfilled, and the power minimum (5.4) can be found **by changing the power allocation policy** for the subsequent iterations. In fact, we proceed to minimizing the total transmit power while constraining $r_k = r_k^o, \quad \forall k$, instead of maximizing the achievable rate margin under total power constraint.

The key idea of this design is to change the power control policy when the user rate targets are feasible. In fact we have the following

$$t = \min_k \frac{r_k}{r_k^o} = 1 \tag{5.6}$$

$$\Leftrightarrow \frac{\text{WMSE}_k}{\xi_k} = 1 \tag{5.7}$$

or,

$$r_k = r_k^o \Leftrightarrow \text{WMSE}_k = \xi_k \tag{5.8}$$

In the previous chapters, the transmit beamformers are optimized along with power allocation for MMR using either UL/DL duality or Lagrangian duality. Hereafter, we provide the power allocation which fulfills the design goal for both approaches.

5.3.1 UL-DL duality based approach

In the MMR optimization using UL/DL duality, the power allocation is optimized in the virtual UL. Therefore, we minimize the total transmit power in the UL channel, i.e., $||\boldsymbol{q}||_1$, and the DL total power will be minimized accordingly. Moreover, the constraints are considered under the matrix-weighted MSE formulation in the dual UL, namely

$$\text{WMSE}_k^{\text{UL}} = \xi_k, \quad \forall k.$$

Collecting the per user UL WMSE in a diagonal matrix $\boldsymbol{\epsilon}_w^{\text{UL}}$, we have (3.19)

$$\boldsymbol{\epsilon}_w^{\text{UL}} \boldsymbol{q} = \boldsymbol{A}\boldsymbol{q} + \sigma_n^2 \boldsymbol{C} \mathbf{1}_K$$

with $\mathbf{A}$ and $\boldsymbol{C}$ respectively defined in (3.21) and (3.20).

To achieve the targets $\boldsymbol{\xi} = \text{diag}([\xi_1 \ldots \xi_K])$ for fixed $\boldsymbol{G}_k, \boldsymbol{F}_k$ and β_k, the optimal power allocation is given by

$$\boldsymbol{q} = \sigma_n^2 (\boldsymbol{\xi} - \boldsymbol{A})^{-1} \boldsymbol{C} \mathbf{1}_K. \tag{5.9}$$

Thus, we set the new total power constraint $P_{\max}$ for the MMR problem as $P_{\max} = P$, with

$$P = \mathbf{1}_K^T \boldsymbol{q}. \tag{5.10}$$

The corresponding iterative algorithm summarized in Table 5.1.

Note1: In the designed algorithm, we drop the inner loop from MMR optimization. Thus, we optimize the PM problem under one loop, unlike in Table 3.1.

Note2: We have followed here the derivations from Chapter 3 which considers BC, with a total power constraint. However, the extension to IBC with total power constraint is straightforward.

Table 5.1: Power Minimization via UL/DL Duality Algorithm

1. for fixed r_k^o, initialize: $\boldsymbol{F}_k^{\mathrm{H}(0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})$, $\bar{\boldsymbol{Q}}^{(0)} = \frac{P_{\max}}{N_d}\boldsymbol{I}$, $P = P_{\max}$, $m = 0$ and $m_{\max}$
2. compute UL receive filter $\boldsymbol{G}^{(0)}$ and $\boldsymbol{\beta}^{(0)}$ with (3.11)
3. set $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}$, $\xi_k^{(0)} = d_k$ and $t^{(0)} \leftarrow 0$
4. find optimal user power allocation $\tilde{\boldsymbol{q}}^{(0)}$ by solving (3.24) and compute $\boldsymbol{Q}_k^{(0)} = \tilde{q}_k^{(0)}\bar{\boldsymbol{Q}}_k^{(0)}$
5. **repeat**

   5.1 $m \leftarrow m + 1$

   5.2 *UL channel:*

   - update $\boldsymbol{G}^{(m)}$ and $\boldsymbol{\beta}^{(tmp)}$ with (3.11)
   - compute the MSE values $\boldsymbol{\varepsilon}^{\mathrm{UL},(m)}$ with (3.8)

   5.3 *DL channel:*

   - compute $\boldsymbol{P}^{(m)}$ with (3.10)
   - update $\boldsymbol{F}^{(m)}$ and $\boldsymbol{\beta}^{(tmp)}$ with (3.12)
   - compute the MSE values $\boldsymbol{\varepsilon}^{\mathrm{DL}(m)}$ with (3.1)

   5.4 *UL channel:*

   - compute $\boldsymbol{Q}^{(tmp)}$ with (3.9) and $\bar{\boldsymbol{Q}}_k^{(m)} = \boldsymbol{Q}_k^{(tmp)}/\mathrm{tr}\big(\boldsymbol{Q}_k^{(tmp)}\big)$

   if $t^{(m-1)} < 1$
   find optimal user power allocation $\tilde{\boldsymbol{q}}^{(m)}$ by solving (3.24) with $P_{max} = P$, $\boldsymbol{Q}_k^{(m)} = \tilde{q}_k^{(m)}\bar{\boldsymbol{Q}}_k^{(m)}$
   else
   update $\boldsymbol{q}$ with (5.9) and set $P = ||\boldsymbol{q}||_1$
   end if

   5.5 update $\boldsymbol{W}_k^{(m)} = (\mathbf{E}_k^{\mathrm{UL}(m)})^{-1}$, $r_k^{(m)} = \ln\det(\boldsymbol{W}_k^{(m)})$, $t^{(m)} = \min_k \frac{r_k^{(m)}}{r_k^o}$
   if $t^{(m)} < 1$
   update $\xi_k^{(m)} = d_k + r_k^{(m)} - t^{(m)} r_k^o$
   else
   update $\xi_k^{(m)} = d_k + r_k^{(m)} - r_k^o$ (imposing rate targets are fulfilled, i.e., $t = 1$)
   end if
6. **until** required accuracy is reached or $m \geq m_{\max}$

5.3.2 Lagrangian duality based approach

Considering the rate balancing via Lagrangian duality as in Chapter 4, the power allocation is obtained directly in the DL channel. Since the MMR problem is formulated as max-min weighted-matrix MSE, the related power minimization problem is constrained by

$$\text{WMSE}_k = \xi_k, \quad \forall k.$$

Similarly to what has been considered for the PM based on UL/DL duality, we collect the per user WMSE in a diagonal matrix $\boldsymbol{\epsilon}_w$ to obtain (4.17) as follows

$$\boldsymbol{\epsilon}_w \boldsymbol{p} = (\boldsymbol{D} + \boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\sigma}.$$

where $\boldsymbol{D}, \boldsymbol{\Psi}$ and $\boldsymbol{\sigma}$, are defined in (4.12), (4.13) and (4.16), respectively.

The corresponding optimal power allocation to achieve the targets $\boldsymbol{\xi}$ is then

$$\boldsymbol{p} = (\boldsymbol{\xi} - \boldsymbol{D} - \boldsymbol{\Psi})^{-1}\boldsymbol{\sigma}. \tag{5.11}$$

Then, we set the new power constraint for MMR optimization as $P_{\max} = P$ with

$$P = \mathbf{1}_K^T \boldsymbol{p}. \tag{5.12}$$

which completes the optimization framework.

The proposed algorithm is summarized in Table 5.2. In Chapter 4, IBC with per cell power constraints is considered, and subgradient method is used to verify the latters. The corresponding iterative alogrithm turns inside the MMR algorithm when optimizing the power allocation. Here, we drop the subgradient loop and consider BC with total power constraint. The joint optimization of beamformers and power minimization is obtained with Table 5.2.

Note4: Extension to IBC with a total power constraint is straightforward.

5.4 Results

In this section, we numerically evaluate the performance of the proposed algorithm. The simulations are carried out over normalized flat fading channels, i.e., each element is i.i.d. and normally distributed : starting from i.i.d. $[\boldsymbol{H}_{k,j}]_{mn} \sim \mathcal{CN}(0, 1)$. For all simulations, we take 500 channel realisations and $m_{\max} = 100$.

Figure 5.1 illustrates an example of the PM algorithm execution. The upper subfigure shows the achieved per user rates vs. the number of iterations, while the lower subfigure shows the total transmit power vs. the number of iterations. As we can see, the user rates meet their targets in a very good approximation after convergence, and the total transmit power is minimized accordingly.

Table 5.2: Power Minimization via Lagrangian Duality Algorithm

1. for predefined targets r_k^o, initialize: $\boldsymbol{G}_k^{(0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0)} = \boldsymbol{q}_k^{(0)} = \frac{P_{\max}}{K}$, $m = 0$ and fix $m_{\max}$,
2. initialize $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}_{d_k}$, $\xi_k^{(0)} = d_k$ and do $t^{(0)} \leftarrow 0$
3. compute $\boldsymbol{F}_k^{(0)}$ in $\boldsymbol{\mathcal{F}}_k^{(0)} = p_k^{(0)-1/2}\boldsymbol{F}_k$ from (4.8)
4. **repeat**

   4.1 $m \leftarrow m + 1$

   4.2 update $\boldsymbol{G}_k$ in $\boldsymbol{\mathcal{G}}_k = p_k^{1/2}\boldsymbol{G}_k$ from (4.34)

   4.3 update $\boldsymbol{F}_k$ in $\boldsymbol{\mathcal{F}}_k = p_k^{-1/2}\boldsymbol{F}_k$ from (4.8)

   4.4 **if** $t^{(m-1)} < 1$

   update $\boldsymbol{p}$ and $\boldsymbol{q}$ as right and left Perron Frobenius eigen vectors of $\boldsymbol{\Lambda}$

   else

   update $\boldsymbol{p}$ using (5.11) and set $P = ||\boldsymbol{p}||_1$

   update $\boldsymbol{q}$ as left Perron Frobenius eigen vector of $\boldsymbol{\Lambda}(P_{\max} = P)$

   end if

   4.5 compute $\boldsymbol{E}_k^{(m)}$ with (4.7) and update $\boldsymbol{W}_k^{(m)} = (\boldsymbol{E}_k^{(m)})^{-1}$

   4.6 compute $r_k^{(m)} = \ln\det(\boldsymbol{W}_k^{(m)})$ and determine $t = \min_k \frac{r_k^{(m)}}{r_k^o}$,

   if $t^{(m)} < 1$

   update $\xi_k^{(m)} = d_k + r_k^{(m)} - t^{(m)} r_k^o$

   else

   update $\xi_k^{(m)} = d_k + r_k^{(m)} - r_k^o$ (imposing rate targets are fulfilled, i.e., $t = 1$)

   end if

5. **until** required accuracy is reached or $m \geq m_{\max}$

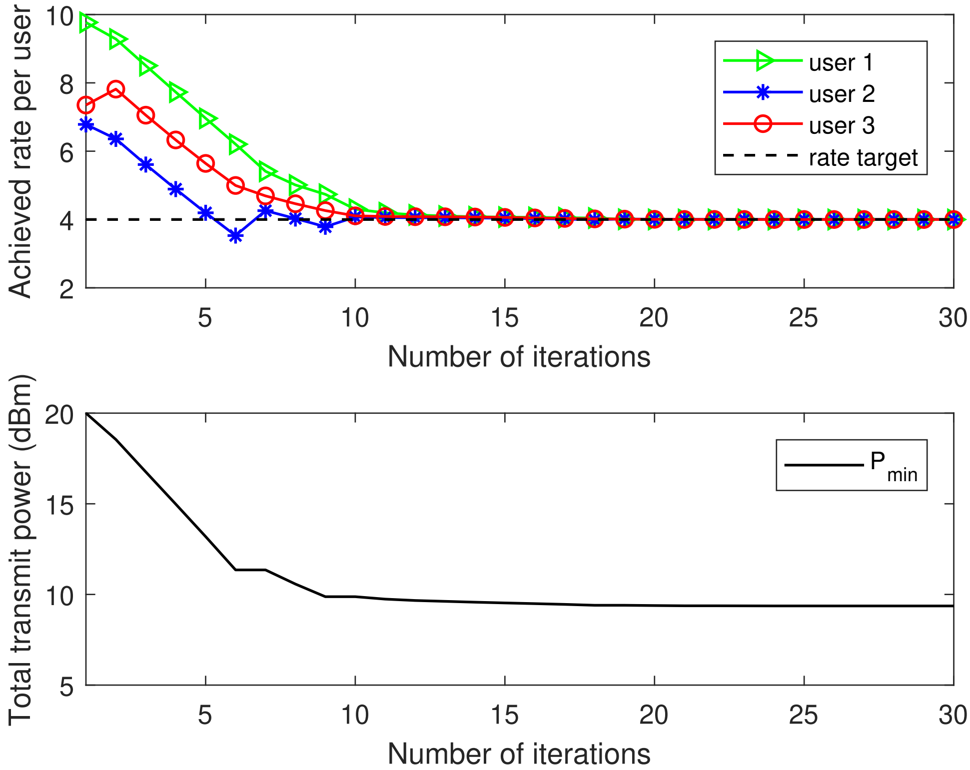

Figure 5.1: Example of total PM algorithm execution: $M = 12, K = 3, d_k = N_k = 2$.

In Table 5.3, we can see that for a fixed number of users, the more we increase the number of transmit antennas, the more we reduce the number of iterations; the average execution time however remains the same because of the increasing computational operations per iteration. On the other hand, for a fixed number of transmit antennas, the number of iterations increases with the increase of the number of active users as well as the average execution time.

In Figure 5.2, we plot the achieved minimum total power w.r.t. user targets for different configurations. In particual, we consider BC, limited by $P_{\max} = 40$dBm, and evaluate how performs the PM algorithm to guarantee the predefined target rates. The targets here are considered equal for all the active users in the system. We can see that

Table 5.3: Average Execution Time (AET) and number of iterations

antennas M	12	32	64
AET (s)	2.1983	2.2784	2.299
iterations	~ 15	~ 10	~ 8

(a) SNR = 30dB, $K = 3$, $N_k = d_k = 2$

users K	8	10	16	32
AET (s)	5.2029	6.3609	12.554	163.269
iterations	~ 13	~ 15	~ 16	~ 38

(b) SNR = 30dB, $M = 64$, $N_k = d_k = 2$

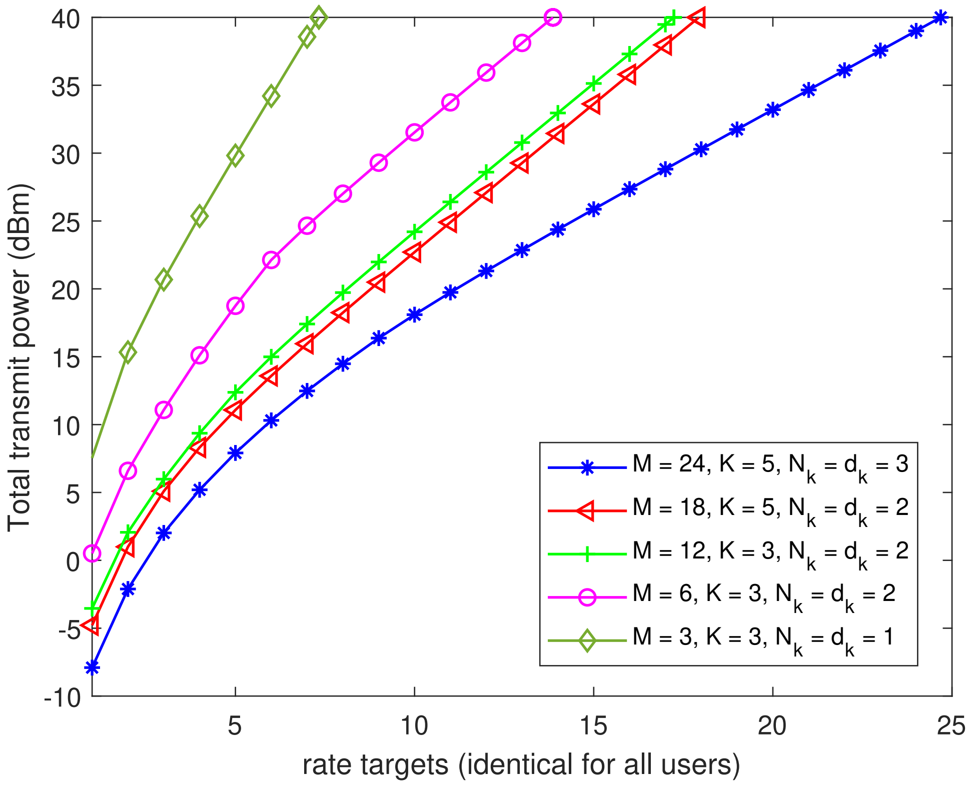

Figure 5.2: Minimum power vs. user rate targets.

the total transmit power is reduced considerably according to the number of users, their rate targets and the configuration of the system. Of course, the more we increase transmit antennas, the more we reduce the total transmit power for the same number of users/rate targets configuration.

Figure 5.3 represents the achieved rate and the total transmit power for both PM and MMR approaches. We observe that for equal user rate targets, under constrained $P_{\max} = 10$dBm, MMR guarantees the same achievable rate no matter the values of targets (but of course, for feasible targets). In contrast, the PM approach minimizes the total transmit power as long as $\mathcal{R} \geq 1$ till equality of the condition, and thus user rates equal the predefined user targets. See that when the target rates equal the achievable balanced rate with MMR, $P = P_{\max}$, i.e., no further minimization is supported. Therefore, the optimization problem of MMR provides performance measure that reflects the quality of the corporate multi-user channel and insights about the maximum feasible user rate targets.

5.5 Closing Remarks

In this chapter, we investigated the total power minimization problem subject to a set of user rate targets. An iterative strategy was derived via rate balancing to optimize

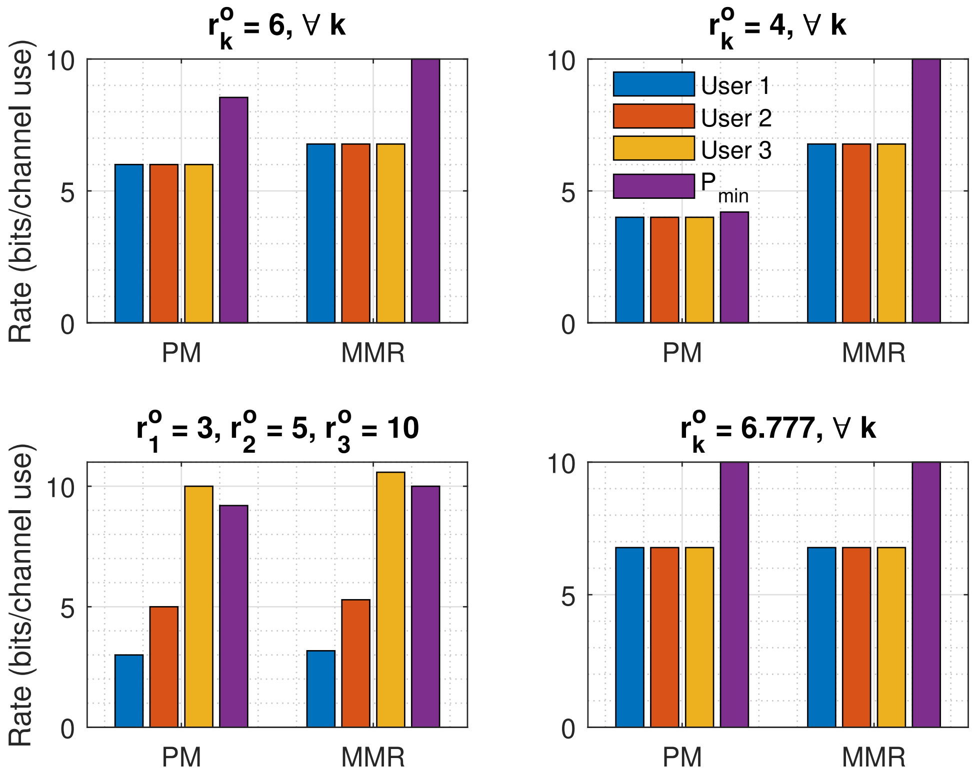

Figure 5.3: Comparison of the achieved per user rates with PM and MMR approaches: $C = 1$, $M_c = 12$, $N_k = d_k = 2$ and $K = 3$.

the problem. Simulation results showed that, for appropriate (feasible) user rate targets, arbitrary points within the achievable rate region can be achieved with minimal expense of transmission power.

Part III

Rate Balancing with Imperfect CSIT

Chapter 6

Channel Model with Partial CSIT

6.1 Overview

In the previous part, we have studied the joint beamforming design and power allocation w.r.t. rate balancing assuming perfect knowledge of the channel. Indeed, in order to fully exploit the spatial diversity gain in the MIMO broadcast channel, Channel State Information at the Transmitter (CSIT) is required to separate the spatial channels for different users. However, CSIT is difficult to obtain and is never perfect. In this part, we consider imperfect CSIT; namely, partial CSIT in terms of transmit covariance matrices and channel estimates.

In this chapter, we introduce the channel model with partial CSIT. In Section 6.2, we provide the motivation to considering imperfect CSIT and related works w.r.t. max-min balancing problems. In Section 6.3, we describe the considered channel model with the presence of only partial CSIT. Finally, we conclude in Section 6.4.

6.2 Why Partial CSI

In downlink communications, when a certain knowledge of the Channel State Information (CSI) at the transmitter is available, the system throughput can be maximized. In practical, obtaining CSI at the receiver is easily possible via training, whereas CSI at the transmitter acquires reciprocity or feedback from the receiver. Therefore, many works address the problem of optimizing the performance of MIMO systems with the presence of CSIT uncertainties, better known as partial CSIT. Among the different optimization criteria, we distinguish the transmit power minimization, and the max-min/min-max problems w.r.t. either SINR [8, 44–46, 79], MSE [47–49] or user rate. The latter is the focus of this work. In particular, we study Multi-Cell MIMO User Rate Balancing with Partial CSIT.

Perfect CSIT

In perfect CSIT case, [43] studies the balancing problem w.r.t. SINR for MISO system using uplink/downlink duality. In fact, most of max-min beamforming problems are transformed into the dual problem power minimization problem in the uplink. In [80], SINR balancing problem subject to multiple weighted-sum power constraints for MISO system is solved by exploiting Perron-Frobenius theory and uplink/downlink duality, and an iterative subgradient projection algorithm is used to satisfy the per-stream power constraints. Similarly, MSE duality, which states that the same MSE values are achievable in the downlink and the uplink with the same transmit power constraint, has been exploited to solve max-min beamforming problems w.r.t. MSE. In [49], three levels of MSE dualities are established between MIMO BC and MIMO MAC with the same transmit power constraint; these dualities are exploited to reduce the computational complexity of the sum-MSE and weighted sum-MSE minimization problems (with fixed weights) in a MIMO BC. On the other hand, we prove that user-wise rate balancing outperforms user-wise MSE balancing or streamwise rate (or MSE/SINR) balancing when the streams of any MIMO user are quite unbalanced in [76, 81, 82].

Partial CSIT

In contrast, due to the inevitability of channel estimation error, CSI can never be perfect. This motivates [83] to consider an MSE-based transceiver design problem where the channel knowledge is modeled in terms of channel mean and variance both at the transmitter and receivers. Then, an iterative algorithm is proposed to solve the expected MSE balancing problem by switching between the broadcast and the multiple access channels. Also, SINR balancing problem with imperfect CSIT is studied in [84] for multi-cell multi-user MISO system. Therein, the authors introduce an alternative biased SINR estimate to incorporate the knowledge of the channel estimation error, outperforming the unbiased maximum-likelihood estimate. In [85], CSI error matrix is represented as a bounded hyper-spherical region within some radius, leading to a robust max-min SINR problem for single-stream MIMO system. The latter is solved as semidefinite program problem, where robust transmit and receive beamformers are obtained using alternating optimization. Rate balancing problem is studied in [86], for broadcast MISO channel, where the case of erroneous CSI at the receiver is considered. The authors use duality w.r.t. SINR to solve the balancing problem: they transform the BC problem into dual MAC problem taking into consideration the erroneous receiver CSI. Actually, in the single stream per user case (e.g. in MISO systems), balancing w.r.t. SINR, MSE or user rate is equivalent (in the unweighted case). Another rate balancing work for MISO system is studied in [87], wherein the statistical properties of the channel are exploited and an algorithm for optimal downlink beamforming is derived using uplink/downlink duality.

6.3 Considered Model

6.3.1 Joint Mean and Covariance Gaussian CSIT

In this section we drop the user index k for simplicity. Assume that the channel has a (prior) Gaussian distribution with zero mean and separable correlation model

$$\boldsymbol{H} = \boldsymbol{C}_r^{1/2}\,\boldsymbol{H}^{'}\,\boldsymbol{C}_t^{1/2} \tag{6.1}$$

where $\boldsymbol{C}_r^{1/2}$, $\boldsymbol{C}_t^{1/2}$ are Hermitian square-roots of the Rx and Tx side covariance matrices

$$\begin{aligned} \mathbf{E}\,\boldsymbol{H}\,\boldsymbol{H}^H &= \mathrm{tr}\{\boldsymbol{C}_t\}\,\boldsymbol{C}_r \\ \mathbf{E}\,\boldsymbol{H}^H\boldsymbol{H} &= \mathrm{tr}\{\boldsymbol{C}_r\}\,\boldsymbol{C}_t \end{aligned} \tag{6.2}$$

and the elements of $\boldsymbol{H}^{'}$ are i.i.d. $\sim \mathcal{CN}(0,1)$. Now, the Tx dispose of a (deterministic) channel estimate

$$\widehat{\boldsymbol{H}}_d = \boldsymbol{H} + \boldsymbol{C}_r^{1/2}\,\widetilde{\boldsymbol{H}}_d^{'}\,\boldsymbol{C}_d^{1/2} \tag{6.3}$$

where again the elements of $\widetilde{\boldsymbol{H}}_d^{'}$ are i.i.d. $\sim \mathcal{CN}(0,1)$, and typically $\boldsymbol{C}_d = \sigma^2_{\widetilde{\boldsymbol{H}}} I_{N_t}$. The combination of the estimate with the prior information leads to the (posterior) Linear Minimum Mean Square Error (LMMSE) estimate

$$\begin{aligned} \widehat{\boldsymbol{H}} &= \widehat{\boldsymbol{H}}_d\,(\boldsymbol{C}_t + \boldsymbol{C}_d)^{-1}\boldsymbol{C}_t = \boldsymbol{H} + \boldsymbol{C}_r^{1/2}\,\widetilde{\boldsymbol{H}}_p^{'}\,\boldsymbol{C}_p^{1/2} \\ \boldsymbol{C}_p &= \boldsymbol{C}_d\,(\boldsymbol{C}_t + \boldsymbol{C}_d)^{-1}\boldsymbol{C}_t \end{aligned} \tag{6.4}$$

where $\widehat{\boldsymbol{H}}$ and $\widetilde{\boldsymbol{H}}_p^{'}$ are independent (or decorrelated if not Gaussian). Note that we get for the MMSE estimate of a quadratic quantity of the form

$$\mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d}\boldsymbol{H}^H\boldsymbol{H} = \widehat{\boldsymbol{H}}^H\widehat{\boldsymbol{H}} + \mathrm{tr}\{\boldsymbol{C}_r\}\boldsymbol{C}_p = \mathbf{S}\;. \tag{6.5}$$

Let us emphasize that this MMSE estimate implies $\mathbf{S} = \arg\min_{\boldsymbol{T}} \mathbf{E}_{\boldsymbol{H}|\hat{\boldsymbol{H}}_d}||\boldsymbol{H}^H\boldsymbol{H} - \boldsymbol{T}||^2$. It averages out to

$$\mathbf{E}_{\boldsymbol{H}_d}\mathbf{S} = \mathbf{E}_{\boldsymbol{H},\boldsymbol{H}_d}\boldsymbol{H}^H\boldsymbol{H} = \mathbf{E}_{\boldsymbol{H}}\boldsymbol{H}^H\boldsymbol{H} = \mathrm{tr}\{\boldsymbol{C}_r\}\boldsymbol{C}_t\;. \tag{6.6}$$

Hence, if we want the best estimate for $\boldsymbol{H}^H\boldsymbol{H}$ (which appears in the signal or interference powers), it is not sufficient to replace $\boldsymbol{H}$ by $\widehat{\boldsymbol{H}}$ but also the (estimation error) covariance information should be exploited. Other useful expressions are

$$\mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d}\boldsymbol{H}^H\boldsymbol{Q}\boldsymbol{H} = \widehat{\boldsymbol{H}}^H\boldsymbol{Q}\widehat{\boldsymbol{H}} + \mathrm{tr}\{\boldsymbol{C}_r\boldsymbol{Q}\}\boldsymbol{C}_p \tag{6.7}$$

and

$$\mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d}\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^H = \widehat{\boldsymbol{H}}\boldsymbol{P}\widehat{\boldsymbol{H}}^H + \mathrm{tr}\{\boldsymbol{C}_p\boldsymbol{P}\}\boldsymbol{C}_r\ . \tag{6.8}$$

Note that $\rho_P = \frac{\mathbf{E}\,\mathrm{tr}\{\widehat{\boldsymbol{H}}^H\widehat{\boldsymbol{H}}\}}{\mathrm{tr}\{\boldsymbol{C}_r\}\,\mathrm{tr}\{\boldsymbol{C}_p\}}$ is a form of Ricean factor that represents the posterior channel estimation quality. Perhaps more instructive is the deterministic channel estimation quality. From (6.3)

$$\begin{aligned}\rho_D &= \frac{\mathbf{E}\,\mathrm{tr}\{\boldsymbol{H}^H\boldsymbol{H}\}}{\mathbf{E}\,\mathrm{tr}\{(\boldsymbol{C}_r^{1/2}\,\widetilde{\boldsymbol{H}}'_d\,\boldsymbol{C}_d^{1/2})^H\boldsymbol{C}_r^{1/2}\,\widetilde{\boldsymbol{H}}'_d\,\boldsymbol{C}_d^{1/2}\}}\\ &= \frac{\mathbf{E}\,\mathrm{tr}\{\boldsymbol{C}_t^{1/2}\,\boldsymbol{H}'^H\,\boldsymbol{C}_r\,\boldsymbol{H}'\,\boldsymbol{C}_t^{1/2}\}}{\mathbf{E}\,\mathrm{tr}\{\boldsymbol{C}_d^{1/2}\,\widetilde{\boldsymbol{H}}'^H_d\,\boldsymbol{C}_r\,\widetilde{\boldsymbol{H}}'_d\,\boldsymbol{C}_d^{1/2}\}}\\ &= \frac{\mathrm{tr}\{\boldsymbol{C}_r\}\,\mathrm{tr}\{\boldsymbol{C}_t\}}{\mathrm{tr}\{\boldsymbol{C}_r\}\,\mathrm{tr}\{\boldsymbol{C}_d\}} = \frac{\mathrm{tr}\{\boldsymbol{C}_t\}}{\mathrm{tr}\{\boldsymbol{C}_d\}} = \frac{\frac{1}{M}\mathrm{tr}\{\boldsymbol{C}_t\}}{\sigma^2_{\widetilde{\boldsymbol{H}}}} = \frac{1}{\sigma^2_{\widetilde{\boldsymbol{H}}}}\end{aligned} \tag{6.9}$$

where we used $\mathrm{tr}\{\boldsymbol{C}_t\} = M$ from Section 6.3.3. On the other hand

$$\rho_P = \frac{\mathrm{tr}\{\boldsymbol{C}_t - \boldsymbol{C}_p\}}{\mathrm{tr}\{\boldsymbol{C}_p\}} = \frac{\mathrm{tr}\{\boldsymbol{C}_t\,(\boldsymbol{C}_t+\boldsymbol{C}_d)^{-1}\boldsymbol{C}_t\}}{\mathrm{tr}\{\boldsymbol{C}_d\,(\boldsymbol{C}_t+\boldsymbol{C}_d)^{-1}\boldsymbol{C}_t\}}\,. \tag{6.10}$$

6.3.2 Further Explanations

If $\boldsymbol{C}_r^{1/2}$, $\boldsymbol{C}_t^{1/2}$ are not Hermitian square-roots, which means e.g. $\boldsymbol{C}_t = \boldsymbol{C}_t^{1/2}\boldsymbol{C}_t^{H/2}$, then the more general formula is

$$\boldsymbol{H} = \boldsymbol{C}_r^{1/2}\,\boldsymbol{H}'\,\boldsymbol{C}_t^{H/2} \tag{6.11}$$

where the elements of $\widetilde{\boldsymbol{H}}$ are i.i.d. $\sim \mathcal{CN}(0,1)$. Note that the matrices $\boldsymbol{C}_r$, $\boldsymbol{C}_t$ are unique up to scale factor that leaves their (Kronecker) product unchanged. One can check that we still have

$$\begin{aligned}\mathbf{E}\,\boldsymbol{H}\,\boldsymbol{H}^H &= \mathrm{tr}\{\boldsymbol{C}_t\}\;\boldsymbol{C}_r\\ \mathbf{E}\,\boldsymbol{H}^H\boldsymbol{H} &= \mathrm{tr}\{\boldsymbol{C}_r\}\;\boldsymbol{C}_t\ .\end{aligned} \tag{6.12}$$

Now let $\boldsymbol{h} = \mathrm{vec}(\boldsymbol{H})$, which also has zero mean. From (6.11), we get, using $\mathrm{vec}(ABC) = (C^T \otimes A)\,\mathrm{vec}(B)$,

$$\boldsymbol{h} = (\boldsymbol{C}_t^{*/2} \otimes \boldsymbol{C}_r^{1/2})\,\boldsymbol{h}' \tag{6.13}$$

where $\boldsymbol{h}' = \mathrm{vec}(\boldsymbol{H}')$. With $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$ and $\boldsymbol{C}_{\boldsymbol{h}'\boldsymbol{h}'} = \boldsymbol{I}_{N_rN_t}$, this leads to the covariance matrix

$$\boldsymbol{C}_{\boldsymbol{hh}} = \boldsymbol{C}_t^T \otimes \boldsymbol{C}_r \tag{6.14}$$

exploiting $\boldsymbol{C}_t = \boldsymbol{C}_t^H$ and hence $\boldsymbol{C}_t^* = \boldsymbol{C}_t^T$. This leads to the term "Kronecker model". In a first instance, consider a deterministic channel estimate

$$\hat{\boldsymbol{h}}_d = \boldsymbol{h} - \widetilde{\boldsymbol{h}}_d = \boldsymbol{h} + \boldsymbol{C}^{1/2}_{\widetilde{\boldsymbol{h}}_d\widetilde{\boldsymbol{h}}_d}\,\widetilde{\boldsymbol{h}}'_d \tag{6.15}$$

where $\boldsymbol{C}_{\widetilde{\boldsymbol{h}}'_d\widetilde{\boldsymbol{h}}'_d} = \boldsymbol{I}_{N_rN_t}$. The combination of the estimate with the prior information leads to the (posterior) LMMSE estimate

$$\hat{\boldsymbol{h}} = \boldsymbol{C}_{\boldsymbol{hh}}(\boldsymbol{C}_{\boldsymbol{hh}} + \boldsymbol{C}_{\widetilde{\boldsymbol{h}}_d\widetilde{\boldsymbol{h}}_d})^{-1}\hat{\boldsymbol{h}}_d = \boldsymbol{h} - \widetilde{\boldsymbol{h}}_p \tag{6.16}$$

with error covariance matrix

$$\boldsymbol{C}_{\widetilde{\boldsymbol{h}}_p\widetilde{\boldsymbol{h}}_p} = (\boldsymbol{C}^{-1}_{\boldsymbol{hh}} + \boldsymbol{C}^{-1}_{\widetilde{\boldsymbol{h}}_d\widetilde{\boldsymbol{h}}_d})^{-1} \tag{6.17}$$

assuming the last two matrices are invertible. Now consider the Kronecker model in (6.14). Under what conditions can we have a Kronecker model for $\boldsymbol{C}_{\widetilde{\boldsymbol{h}}_p\widetilde{\boldsymbol{h}}_p}$? Clearly it should be helpful to consider a Kronecker model for

$$\boldsymbol{C}_{\widetilde{\boldsymbol{h}}_d\widetilde{\boldsymbol{h}}_d} = \boldsymbol{C}^T_{t,d} \otimes \boldsymbol{C}_{r,d}\ . \tag{6.18}$$

So the question becomes, when can a sum of two Kronecker products be written as a Kronecker product? We can scale one of the terms in the sum to become identity (by multiplying with its inverse), and we can use the eigen decomposition of the second term to reduce it to a diagonal form. The question then is, when can we write

$$\boldsymbol{I} \otimes \boldsymbol{I} + \boldsymbol{C} \otimes \boldsymbol{D} = \mathbf{A} \otimes \mathbf{B} \tag{6.19}$$

where $\boldsymbol{C}$, $\boldsymbol{D}$ are diagonal, which implies $\mathbf{A}$, $\mathbf{B}$ to be diagonal. Note that for diagonal matrices, (6.19) could be written in terms of the vectors of the diagonal elements, leading to Khatri-Rao products, but we shall stay with the diagonal matrices. We shall denote the diagonal elements of $\mathbf{A}$, $\mathbf{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ as a_i, b_i, c_i, d_i. Due to the scale ambiguity in a product, we can fix $a_1 = 1$. Equation diagonal block i in (6.19), we get

$$a_i\,\mathbf{B} = \boldsymbol{I} + c_i\,\boldsymbol{D} \tag{6.20}$$

which for $i = 1$ becomes $\mathbf{B} = \boldsymbol{I} + c_1\,\boldsymbol{D}$. Subtracting (6.20) for $i > 1$ from (6.20) for $i = 1$, we get

$$(a_i - 1)\,\mathbf{B} = (c_i - c_1)\,\boldsymbol{D} \tag{6.21}$$

where we assume that for at least one i, $a_i \neq 1$ and $c_i \neq c_1$ (note that the case "or" would lead to the implication of either $\mathbf{B}$ or $\boldsymbol{D}$ being zero which is not possible for positive (semi) definite matrices). So, we assume that $\mathbf{A}$ and $\boldsymbol{C}$ are not a multiple of identity. Then (6.21) together with (6.20) for $i = 1$ imply that $\mathbf{B}$ and $\boldsymbol{D}$ are multiples of $\boldsymbol{I}$. In summary, either $\mathbf{A}$ and $\boldsymbol{C}$ or $\mathbf{B}$ and $\boldsymbol{D}$ are multiples of $\boldsymbol{I}$. Going back to (6.17), for $\boldsymbol{C}_{\widetilde{\boldsymbol{h}}_p\widetilde{\boldsymbol{h}}_p}$ to have a Kronecker structure requires that either $\boldsymbol{C}_{t,d}$ is a multiple of $\boldsymbol{C}_t$, or $\boldsymbol{C}_{r,d}$ is a multiple of $\boldsymbol{C}_r$. Since the Tx side is too important to have such constraint, we shall consider that $\boldsymbol{C}_r$ is a multiple of $\boldsymbol{C}_{r,d}$.

Now, using orthogonal pilots and white Rx noise (or i.i.d. pilots) will lead to $\boldsymbol{C}_{\widetilde{\boldsymbol{h}}_d \widetilde{\boldsymbol{h}}_d} = \sigma^2_{\widetilde{h}} \boldsymbol{I}$ or

$$\boldsymbol{C}_{t,d} = \sigma^2_{\widetilde{h}} \boldsymbol{I} = \boldsymbol{C}_d \,, \; \boldsymbol{C}_{r,d} = \boldsymbol{C}_r = \boldsymbol{I} \tag{6.22}$$

and

$$\boldsymbol{C}_{\widetilde{\boldsymbol{h}}_p \widetilde{\boldsymbol{h}}_p} = \boldsymbol{C}_p^T \otimes \boldsymbol{I} \,. \tag{6.23}$$

We can now write the (posterior) LMMSE estimate (6.16) as

$$\hat{\boldsymbol{h}} = ((\boldsymbol{C}_t^T (\boldsymbol{C}_t^T + \boldsymbol{C}_d^T)^{-1}) \otimes \boldsymbol{I}) \hat{\boldsymbol{h}}_d = \boldsymbol{h} - \widetilde{\boldsymbol{h}}_p \tag{6.24}$$

which can be written as

$$\widehat{\boldsymbol{H}} = \widehat{\boldsymbol{H}}_d (\boldsymbol{C}_t + \boldsymbol{C}_d)^{-1} \boldsymbol{C}_t = \boldsymbol{H} - \widetilde{\boldsymbol{H}}_p \tag{6.25}$$

with (Tx side) error covariance matrix

$$\boldsymbol{C}_p = (\boldsymbol{C}_t^{-1} + \boldsymbol{C}_d^{-1})^{-1} = \boldsymbol{C}_d \, (\boldsymbol{C}_t + \boldsymbol{C}_d)^{-1} \boldsymbol{C}_t \,. \tag{6.26}$$

6.3.3 Normalizations and Rx SNR

Consider the Rx signal model

$$\boldsymbol{y} = \boldsymbol{H}\, \boldsymbol{x} + \boldsymbol{v} \tag{6.27}$$

with white noise $\boldsymbol{C}_{\boldsymbol{vv}} = \sigma_v^2 \boldsymbol{I}_N$. For computation of the Rx SNR, we consider omnidirectional transmission $\boldsymbol{C}_{\boldsymbol{xx}} = \frac{P}{M} \boldsymbol{I}_M$ where $P = \mathrm{tr}\{\boldsymbol{C}_{\boldsymbol{xx}}\}$ is the BS Tx power constraint. Given the white noise, for the Rx SNR we need to compute the Rx signal power as

$$\begin{aligned} \mathbf{E}\|\boldsymbol{H}\, \boldsymbol{x}\|^2 &= \tfrac{P}{M}\, \mathbf{E}\, \mathrm{tr}\{\boldsymbol{H}^H \boldsymbol{H}\} = \tfrac{P}{M}\, \mathbf{E}\|\boldsymbol{h}\|^2 \\ &= \tfrac{P}{M}\, \mathrm{tr}\{\boldsymbol{C}_{\boldsymbol{hh}}\} = \tfrac{P}{M}\, \mathrm{tr}\{\boldsymbol{C}_t\}\, \mathrm{tr}\{\boldsymbol{C}_r\} \end{aligned} \tag{6.28}$$

where we used (6.14). Consider normalizing the channel

$$\mathrm{tr}\{\boldsymbol{C}_{\boldsymbol{hh}}\} = \mathrm{tr}\{\boldsymbol{C}_t\}\, \mathrm{tr}\{\boldsymbol{C}_r\} = N\, \mathrm{tr}\{\boldsymbol{C}_t\} = N\, M \tag{6.29}$$

where N is the number of Rx antennas, i.e. each MIMO channel element has variance 1. So, this implies $\mathrm{tr}\{\boldsymbol{C}_t\} = M$. Then we get

$$\rho_R = \mathrm{SNR}_{\mathrm{Rx}} = \mathbf{E}\|\boldsymbol{H}\, \boldsymbol{x}\|^2 / \sigma_v^2 = N\, \frac{P}{\sigma_v^2} = N \mathrm{SNR}_{\mathrm{Tx}} = N\, \rho_T. \tag{6.30}$$

The channel estimation SNR is ρ_D.

6.3.4 Concrete Channel Model for Simulations

So we'll use $\boldsymbol{C}_r = \boldsymbol{I}_N$. Essentially what remains to be specified is the model for $\boldsymbol{C}_t$. We can consider a multipath channel model

$$\boldsymbol{C}_t = \sum_{n=1}^{N_p} \frac{\alpha_i}{\boldsymbol{v}_i^H \boldsymbol{v}_i} \boldsymbol{v}_i \boldsymbol{v}_i^H \tag{6.31}$$

which leads to $\mathrm{tr}\{\boldsymbol{C}_t\} = \sum_{n=1}^{N_p} \alpha_i = M$. Here the number of (equivalent) paths N_P leads to the rank of $\boldsymbol{C}_t$. So the "Power Delay Profile (PDP)" $\{\alpha_i\}$ should be normalized to have total energy equal to M. Particular choices for the PDP is a uniform PDP, in which $\alpha_i = M/N_p$. Another choice would be exponentially decreasing $\alpha_i = c^{i-1}\alpha_1$ induced by a particular choice for $\alpha_{N_P}/\alpha_1 = c^{N_P-1}$.

The antenna array responses $\boldsymbol{v}_i$ could be chosen as i.i.d. vectors of M i.i.d. elements $\mathcal{CN}(0,1)$. Note that in (8.56) the $\boldsymbol{v}_i$ get equivalently normalized to unit norm. Another choice would be to choose the Vandermonde vectors of a Uniform Linear Array (ULA) at $\lambda/2$ spacing with a random i.i.d. distribution of Direction of Arrival (DoA) angles according to a certain pdf (e.g. uniform over a certain sub interval of $(-\frac{\pi}{2}, \frac{\pi}{2})$).

Finally another important choice is the variation of channel estimation SNR. One choice is for it to be proportional to the SNR: $\rho_D = \beta\, \rho_R$ for some β (of the order of 1). Another choice is to take ρ_D to be constant.

6.4 Closing Remarks

In this chapter, we provided an overview of the channel knowledge assumptions in the state of the art considering balancing problem. Also, we described in details the considered model for partial CSIT part.

Chapter 7

WEMSE Balancing

7.1 Overview

In chapter 4, we have considered the user rate balancing problem for perfect CSIT. In this chapter, we consider a multi-cell multi-user MIMO system with partial CSIT, which combines both channel estimates and channel (error) covariance information. In fact, we focus on ergodic user rate balancing, which corresponds to maximizing the minimum (weighted) per user expected rate in the network. In particular, we introduce a novel extension of chapter 4 to partial CSIT, maximizing an expected rate lower bound in terms of expected MSE. Furthermore, we introduce analytical expression for the per cell power constraints by solving the problem via Lagrangian duality.

The rest of chapter is organized as follows. The considered system model is described in Section 7.2 along with the problem formulation for partial CSI. The proposed solution for the dual Lagrangian problem and the corresponding precoders are derived in Section 7.3, including the analytical expression of the Lagrangian parameter for per cell power constraints. Numerical results are carried out in Section 7.4. We finally conclude in Section 7.5.

7.2 System Model

For the same multi-cell multi-user MIMO system model described in Chapter 4, we consider the channel model detailed in Section 6.3 to solve the ergodic rate balancing problem formulated below.

7.2.1 Expected Rate Balancing Problem

Here, we aim to solve the weighted user-rate max-min optimization problem under per cell total transmit power constraint, i.e., the user rate balancing problem expressed as follows

$$\begin{aligned} &\max_{\boldsymbol{G},p} \min_{k} \; r_k/r_k^{\circ} \\ \text{s.t.} \; &\sum_{k:b_k=c} p_k \leq P_{\max,c}, 1 \leq c \leq C \end{aligned} \tag{7.1}$$

where r_k is the kth user-rate

$$r_k = \ln\det\Big(\boldsymbol{I}+\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\rm H}\boldsymbol{H}_{k,b_k}^{\rm H}\Big) = \ln\det\Big(\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{R}_k\Big)\,, \tag{7.2}$$

$$\boldsymbol{R}_{\overline{k}} = \sigma_n^2\boldsymbol{I}+\sum_{l\neq k}\boldsymbol{H}_{k,b_l}\boldsymbol{\mathcal{G}}_l\boldsymbol{\mathcal{G}}_l^{\rm H}\boldsymbol{H}_{k,b_l}^{\rm H}\,, \tag{7.3}$$

$$\boldsymbol{R}_k = \boldsymbol{R}_{\overline{k}}+\boldsymbol{H}_{k,b_k}\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\rm H}\boldsymbol{H}_{k,b_k}^{\rm H}\,, \tag{7.4}$$

$\boldsymbol{R}_{\overline{k}}$ and $\boldsymbol{R}_k$ are the interference plus noise and total received signal covariances, and r_k° is the rate priority (weight) for user k. Actually, in the presence of partial CSIT, we shall be interested in balancing the expected (or ergodic) rates

$$\begin{aligned} &\max_{\boldsymbol{G},p} \min_{k} \; \overline{r}_k/r_k^{\circ} \\ \text{s.t.} \; &\sum_{k:b_k=c} p_k \leq P_{\max,c},\; c = 1,\ldots,C \end{aligned} \tag{7.5}$$

where $\overline{r}_k = \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\, r_k$. We shall need

$$\overline{\mathbf{S}}_{k,i} = \widehat{\boldsymbol{H}}_{k,b_i}\boldsymbol{G}_i\boldsymbol{G}_i^{\rm H}\widehat{\boldsymbol{H}}_{k,b_i}^{\rm H} + {\rm tr}\{\boldsymbol{G}_i^{\rm H}\boldsymbol{C}_{k,b_i}\boldsymbol{G}_i\}\boldsymbol{I},\; \overline{\mathbf{S}}_k = \overline{\mathbf{S}}_{k,k} \tag{7.6}$$

$$\overline{\boldsymbol{R}}_{\overline{k}} = \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\boldsymbol{R}_{\overline{k}} = \sigma_n^2\boldsymbol{I}+\sum_{i\neq k} p_i\overline{\mathbf{S}}_{k,i}\,,\;\; \overline{\boldsymbol{R}}_k = \overline{\boldsymbol{R}}_{\overline{k}} + p_k\overline{\mathbf{S}}_k \tag{7.7}$$

However, the problem presented in (7.5) is complex and can not be solved directly.

Lemma 3. *The rate of user k in (7.2) is lower bounded as [88]*

$$\overline{r}_k = \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k} \left[\ln\det\left(\boldsymbol{W}_k\right) - \mathrm{tr}\left(\boldsymbol{W}_k\mathbf{E}_k\right) + d_k\right] \tag{7.8}$$

$$\geq \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k} \overline{r}_k^l\,,\ \overline{r}_k^l = \ln\det\left(\boldsymbol{W}_k\right) - \mathrm{tr}\left(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\right) + d_k \tag{7.9}$$

where
$$\begin{aligned}\overline{\boldsymbol{E}}_k &= \mathbb{E}\Big[(\hat{\mathbf{s}}_k - \mathbf{s}_k)(\hat{\mathbf{s}}_k - \mathbf{s}_k)^{\mathrm{H}}\Big] \\ &= \boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{\mathcal{G}}_k - \boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k + \sigma_n^2\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k \\ &+ \sum_{l=1}^{K}\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\big(\widehat{\boldsymbol{H}}_{k,b_l}\boldsymbol{\mathcal{G}}_l\boldsymbol{\mathcal{G}}_l^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_l}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{\mathcal{G}}_l^{\mathrm{H}}\boldsymbol{C}_{k,b_l}\boldsymbol{\mathcal{G}}_l\}\boldsymbol{I}\big)\boldsymbol{\mathcal{F}}_k\end{aligned} \tag{7.10}$$

is the kth-user downlink Expected Mean Square Error (EMSE) matrix between the decision variable $\hat{\mathbf{s}}_k$ and the transmit signal $\mathbf{s}_k$, and $\{\boldsymbol{W}_k\}_{1\leq k\leq K}$ are auxiliary weight matrix variables with optimal solution $\boldsymbol{W}_k^{\mathrm{opt}} = \overline{\boldsymbol{E}}_k^{-1}$ and the optimal receivers are

$$\boldsymbol{\mathcal{F}}_k = \overline{\boldsymbol{R}}_k^{-1}\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{\mathcal{G}}_k. \tag{7.11}$$

Note that $\overline{r}_k^l$ is a lower bound for any $\boldsymbol{W}_k, \boldsymbol{\mathcal{F}}_k$ and so is $\max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k}\overline{r}_k^l$. Now consider both (7.5) and (7.9), and let us introduce $\xi_k = \ln\det\left(\boldsymbol{W}_k\right) + d_k - r_k^{\triangle}$, the matrix-Weighted Expected Mean Square Error (WEMSE) requirement, with target rate $r_k^{\triangle}$. Assume that we shall be able to concoct an optimization algorithm that ensures that at all times and for all users the WEMSE satisfies $\epsilon_{w,k} = \mathrm{tr}\left(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\right) \leq d_k$ and $\ln\det\left(\boldsymbol{W}_k\right) \geq r_k^{\triangle}$ or hence $\xi_k \geq d_k$. This leads $\forall k$ to

$$\begin{aligned}\frac{\epsilon_{w,k}}{\xi_k} \leq 1 \iff \ln\det\left(\boldsymbol{W}_k\right) + d_k - \mathrm{tr}\left(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\right) \geq r_k^{\triangle} \\ \overset{(a)}{\Longrightarrow} \overline{r}_k^l/r_k^{\triangle} \geq 1\end{aligned} \tag{7.12}$$

where (a) follows from (7.9). To get to (7.12), what we can exploit in (7.5) is a scale factor t that can be chosen freely in the rate weights r_k° in (7.5). We shall take $t = \min_k \overline{r}_k^l/r_k^{\circ}$, which allows to transform the rate weights r_k° into target rates $r_k^{\triangle} = tr_k^{\circ}$, and at the same time allows to interpret the WEMSE weights ξ_k as target WEMSE values.

Doing so, the initial rate balancing optimization problem (7.5) can be transformed into a WEMSE balancing problem expressed as follows

$$\begin{aligned}&\min_{\boldsymbol{G},p,\boldsymbol{\mathcal{F}}}\max_{k}\ \epsilon_{w,k}/\xi_k \\ &\text{s.t.} \sum_{k:b_k=c} p_k \leq P_{\max,c}, 1 \leq c \leq C\end{aligned} \tag{7.13}$$

which needs to be complemented with an outer loop in which $\boldsymbol{W}_k = \overline{\boldsymbol{E}}_k^{-1}$, $t = \min_k \overline{r}_k^l/r_k^{\circ}$, $r_k^{\triangle} = tr_k^{\circ}$ and $\xi_k = d_k + \overline{r}_k^l - r_k^{\triangle}$ get updated. The problem in (7.13) is still difficult to be handled directly.

7.2.2 The Weighted User EMSE Optimization

In this section, the problem (7.13) with respect to the matrix weighted user EMSE is studied. The per user matrix WEMSE can be expressed as follows

$$\begin{aligned}\epsilon_{w,k} &= \mathrm{tr}\big(\boldsymbol{W}_k \overline{\boldsymbol{E}}_k\big) \\ &= \mathrm{tr}\big(\boldsymbol{W}_k\big) - 2\,\mathrm{tr}\big(\boldsymbol{W}_k \boldsymbol{G}_k^{\mathrm{H}} \widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}} \boldsymbol{F}_k\big) + \sigma_n^2 p_k^{-1} \mathrm{tr}\big(\boldsymbol{W}_k \boldsymbol{F}_k^{\mathrm{H}} \boldsymbol{F}_k\big) \\ &+ p_k^{-1} \sum_{l=1}^{K} p_l \mathrm{tr}\big(\boldsymbol{W}_k \boldsymbol{F}_k^{\mathrm{H}} \big(\widehat{\boldsymbol{H}}_{k,b_l} \boldsymbol{G}_l \boldsymbol{G}_l^{\mathrm{H}} \widehat{\boldsymbol{H}}_{k,b_l}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_l^{\mathrm{H}} \boldsymbol{C}_{k,b_l} \boldsymbol{G}_l\} \boldsymbol{I}\big) \boldsymbol{F}_k\big)\end{aligned} \tag{7.14}$$

Define the diagonal matrix $\boldsymbol{D}$ of signal WEMSE contributions

$$[\boldsymbol{D}]_{ii} = \mathrm{tr}\big(\boldsymbol{W}_i\big) - 2\,\mathrm{tr}\big(\boldsymbol{W}_i \boldsymbol{G}_i^{\mathrm{H}} \widehat{\boldsymbol{H}}_{i,b_i}^{\mathrm{H}} \boldsymbol{F}_i\big) + \mathrm{tr}\big(\boldsymbol{W}_i \boldsymbol{F}_i^{\mathrm{H}} \big(\widehat{\boldsymbol{H}}_{i,b_i} \boldsymbol{G}_i \boldsymbol{G}_i^{\mathrm{H}} \widehat{\boldsymbol{H}}_{i,b_i}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_i^{\mathrm{H}} \boldsymbol{C}_{i,b_i} \boldsymbol{G}_i\} \boldsymbol{I}\big) \boldsymbol{F}_i\big),$$

and the matrix of weighted interference powers

$$[\boldsymbol{\Psi}]_{ij} = \begin{cases} \mathrm{tr}\{\boldsymbol{W}_i \boldsymbol{F}_i^{\mathrm{H}} \big(\widehat{\boldsymbol{H}}_{i,b_j} \boldsymbol{G}_j \boldsymbol{G}_j^{\mathrm{H}} \widehat{\boldsymbol{H}}_{i,b_j}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_j^{\mathrm{H}} \boldsymbol{C}_{i,b_j} \boldsymbol{G}_j\} \boldsymbol{I}\big) \boldsymbol{F}_i\}, & i \neq j \\ 0, & i = j. \end{cases}$$

We can rewrite (7.14) as, with $\boldsymbol{p} = [p_1 \cdots p_K]^{\mathrm{T}}$

$$\epsilon_{w,i} = [\boldsymbol{D}]_{ii} + p_i^{-1} [\boldsymbol{\Psi} \boldsymbol{p}]_i + \sigma_n^2 p_i^{-1} \mathrm{tr}\big(\boldsymbol{W}_i \boldsymbol{F}_i^{\mathrm{H}} \boldsymbol{F}_i\big) \tag{7.15}$$

Collecting all user WEMSEs in a vector $\boldsymbol{\epsilon}_w = \mathrm{diag}(\epsilon_{w,1}, \ldots, \epsilon_{w,K})$, we get

$$\boldsymbol{\epsilon}_w \mathbf{1}_K = \mathrm{diag}(\boldsymbol{p})^{-1} \left[(\boldsymbol{D} + \boldsymbol{\Psi}) \mathrm{diag}(\boldsymbol{p}) \mathbf{1}_K + \boldsymbol{\sigma} \right] \tag{7.16}$$

where the $K \times 1$ vector $\boldsymbol{\sigma}$ is defined as

$$\boldsymbol{\sigma}_i = \sigma_n^2\, \mathrm{tr}\big(\boldsymbol{W}_i \boldsymbol{F}_i^{\mathrm{H}} \boldsymbol{F}_i\big).$$

By multiplying both sides of (7.16) with $\mathrm{diag}(\boldsymbol{p})$, we get

$$\boldsymbol{\epsilon}_w \boldsymbol{p} = (\boldsymbol{D} + \boldsymbol{\Psi}) \boldsymbol{p} + \boldsymbol{\sigma}\,. \tag{7.17}$$

Let $\boldsymbol{\xi} = \mathrm{diag}(\xi_1, \ldots, \xi_K)$, then

$$\boldsymbol{\xi}^{-1} \boldsymbol{\epsilon}_w\, \boldsymbol{p} = \boldsymbol{\xi}^{-1} (\boldsymbol{D} + \boldsymbol{\Psi}) \boldsymbol{p} + \boldsymbol{\xi}^{-1} \boldsymbol{\sigma}\,. \tag{7.18}$$

Actually, problem (7.13) always has a global minimizer $\boldsymbol{p}$ characterized by the equality $\boldsymbol{\xi}^{-1} \boldsymbol{\epsilon}_w(\boldsymbol{p}) = \Delta \boldsymbol{I}$, i.e.,

$$\Delta\, \boldsymbol{p} = \boldsymbol{\xi}^{-1} (\boldsymbol{D} + \boldsymbol{\Psi}) \boldsymbol{p} + \boldsymbol{\xi}^{-1} \boldsymbol{\sigma}\,. \tag{7.19}$$

Now, consider the following problem

$$\begin{aligned} &\max_{\boldsymbol{G}, p, \boldsymbol{\mathcal{F}}} \min_k \quad \bar{r}_k / r_k^{\circ} \\ &\text{s.t.} \quad \sum_{c=1}^{C} \theta_c \boldsymbol{c}_c^{\mathrm{T}} \boldsymbol{p} \leq \sum_{c=1}^{C} \theta_c P_{\max,c} \end{aligned} \tag{7.20}$$

where $\boldsymbol{c}_c$ is a column vector with $\boldsymbol{c}_c(j) = 1$ for $K_{1:c-1} + 1 \leq j \leq K_{1:c}$, and 0 elsewhere. This problem formulation is a relaxation of (7.5), and $\boldsymbol{\theta} = [\theta_1 \cdots \theta_C]^{\mathrm{T}}$can be interpreted as the weights on the individual power constraints in the relaxed problem. The power constraint in (7.20) can be interpreted as a single weighted power constraint

$$(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}})\,\boldsymbol{p} \leq \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max} \tag{7.21}$$

with $\boldsymbol{C}_C = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_C] \in \mathbb{R}_+^{K_{1:C}\times C}$ and $\boldsymbol{p}_{\max} = [P_{\max,1} \cdots P_{\max,C}]^{\mathrm{T}}$. Reparameterize $\boldsymbol{p} = \frac{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}}\boldsymbol{p}'}\boldsymbol{p}'$ where now $\boldsymbol{p}'$ is unconstrained, which allows us to write (7.19) as follows (rewriting $\boldsymbol{p}'$ as $\boldsymbol{p}$)

$$\Delta\,\boldsymbol{p} = \boldsymbol{\Lambda}\,\boldsymbol{p} \text{ with } \boldsymbol{\Lambda} = \boldsymbol{\xi}^{-1}(\boldsymbol{D} + \boldsymbol{\Psi}) + \frac{1}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}\boldsymbol{\xi}^{-1}\boldsymbol{\sigma}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}}. \tag{7.22}$$

Now with (7.22), the WEMSE balancing problem of (7.13) becomes

$$\min_{\boldsymbol{p}} \max_k \frac{\epsilon_{w,k}}{\xi_k} = \min_{\boldsymbol{p}} \max_k \frac{[\boldsymbol{\Lambda}\,\boldsymbol{p}]_k}{p_k} \tag{7.23}$$

According to the Collatz–Wielandt formula [77, Chapter 8], the above expression corresponds to the Perron-Frobenius (maximal) eigenvalue Δ of $\boldsymbol{\Lambda}$ and the optimal $\boldsymbol{p}$ is the corresponding Perron-Frobenius (right) eigenvector

$$\boldsymbol{\Lambda}\boldsymbol{p} = \Delta\,\boldsymbol{p}. \tag{7.24}$$

Note that this implies the equality $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w = \Delta\,\boldsymbol{I}$ as announced in (7.19).

7.3 Proposed Solution

The max-min weighted user rate optimization problem (7.5) can be reformulated as

$$\begin{aligned} &\min_{t,\boldsymbol{G},\boldsymbol{p}} -t \\ &\text{s.t. } \; t\,r_k^{\circ} - \overline{r}_k^{l} \leq 0,\; \boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c} \leq 0\,, \forall k, c. \end{aligned} \tag{7.25}$$

Introducing Lagrange multipliers to augment the cost function with the constraints leads to the Lagrangian

$$\begin{aligned} &\max_{\lambda',\mu} \min_{t,\boldsymbol{G},\boldsymbol{p}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k'(t\,r_k^{\circ} - \overline{r}_k^{l}) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \tag{7.26}$$

Integrating the result (7.8), we get a modified Lagrangian

$$\begin{aligned} &\max_{\lambda',\mu} \min_{t,\boldsymbol{G},\boldsymbol{p},\boldsymbol{\mathcal{F}},\boldsymbol{W}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k'(\mathrm{tr}(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k) - \xi_k) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \tag{7.27}$$

We get $\mu_c = \mu_o \theta_c$ where μ_o is the Lagrange multiplier associated with the constraint in (7.20). Introducing $\lambda_k = \lambda_k^{'} \xi_k$, we can rewrite (with some abuse of notation since actually $\min_{\boldsymbol{W}}$ continues to apply to $\text{tr}(\boldsymbol{W}_k \overline{\boldsymbol{E}}_k) - \xi_k(\boldsymbol{W}_k)$)

$$\max_{\lambda,\mu} \min_{t,\boldsymbol{G},\boldsymbol{p},\boldsymbol{F},\boldsymbol{W}} \mathcal{L}$$
$$\mathcal{L} = -t + \sum_k \lambda_k (\frac{\text{tr}(\boldsymbol{W}_k \overline{\boldsymbol{E}}_k)}{\xi_k} - 1) + \mu_o \sum_c \theta_c (\boldsymbol{c}_c^{\text{T}} \boldsymbol{p} - P_{\text{max},c}) \tag{7.28}$$

We shall solve this saddlepoint condition for $\mathcal{L}$ by alternating optimization. As far as the dependence on $\lambda, \mu, \boldsymbol{G}, \boldsymbol{p}, \boldsymbol{\mathcal{F}}$ is concerned, we have

$$\max_{\lambda} \min_{\boldsymbol{G},\boldsymbol{p},\boldsymbol{\mathcal{F}}} \sum_k \frac{\lambda_k}{\xi_k} \text{tr}(\boldsymbol{W}_k \overline{\boldsymbol{E}}_k) + \sum_c \mu_c \big(\sum_{i:b_i=c} \text{tr}\{\boldsymbol{\mathcal{G}}_i^{\text{H}} \boldsymbol{\mathcal{G}}_i\} - P_{max,c}) \tag{7.29}$$

which is of the form Weighted Sum EMSE (WSEMSE). Optimizing w.r.t. Txs $\boldsymbol{\mathcal{G}}_k$:

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\mathcal{G}}_k^*} = 0 = -\frac{\lambda_k}{\xi_k} \widehat{\boldsymbol{H}}_{k,b_k}^{\text{H}} \boldsymbol{\mathcal{F}}_k \boldsymbol{W}_k + \mu_{b_k} \boldsymbol{\mathcal{G}}_k + \big(\sum_i \frac{\lambda_i}{\xi_i} \big(\widehat{\boldsymbol{H}}_{i,b_k}^{\text{H}} \boldsymbol{\mathcal{F}}_i \boldsymbol{W}_i \boldsymbol{\mathcal{F}}_i^{\text{H}} \widehat{\boldsymbol{H}}_{i,b_k} + \text{tr}\{\boldsymbol{\mathcal{F}}_i \boldsymbol{W}_i \boldsymbol{\mathcal{F}}_i^{\text{H}}\} \boldsymbol{C}_{i,b_k}\big)\big) \boldsymbol{\mathcal{G}}_k \tag{7.30}$$

This leads to

$$\boldsymbol{\mathcal{G}}_k' = \big(\sum_{l=1}^{K} \big(\widehat{\boldsymbol{H}}_{l,b_k}^{\text{H}} \boldsymbol{\mathcal{F}}_l \boldsymbol{W}_l' \boldsymbol{\mathcal{F}}_l^{\text{H}} \widehat{\boldsymbol{H}}_{l,b_k} + \text{tr}\{\boldsymbol{\mathcal{F}}_i \boldsymbol{W}_i' \boldsymbol{\mathcal{F}}_i^{\text{H}}\} \boldsymbol{C}_{i,b_k}\big) + \mu_{b_k} \boldsymbol{I}\big)^{-1} \widehat{\boldsymbol{H}}_{k,b_k}^{\text{H}} \boldsymbol{\mathcal{F}}_k \boldsymbol{W}_k',$$
$$\boldsymbol{\mathcal{G}}_k = \sqrt{p_k}\, \boldsymbol{G}_k \,, \; \boldsymbol{G}_k = \frac{1}{\sqrt{\text{tr}\{\boldsymbol{\mathcal{G}}_k'^{\text{H}} \boldsymbol{\mathcal{G}}_k'\}}} \boldsymbol{\mathcal{G}}_k' \tag{7.31}$$

where $\boldsymbol{W}_k' = \lambda_k / \xi_k \boldsymbol{W}_k$, and accounting for the fact that the user powers are actually optimized by the Perron-Frobenius theory. Note that we can solve for μ_c by multiplying (7.30) from the left by $\boldsymbol{\mathcal{G}}_k^{\text{H}}$ and summing over the users in cell c:

$$\mu_c = \frac{1}{P_{max,c}} \lfloor \text{tr}\{ \sum_{k:b_k=c} \big[\frac{\lambda_k}{\xi_k} \boldsymbol{\mathcal{G}}_k^{\text{H}} \widehat{\boldsymbol{H}}_{k,b_k}^{\text{H}} \boldsymbol{\mathcal{F}}_k \boldsymbol{W}_k - \boldsymbol{\mathcal{G}}_k^{\text{H}} \big(\sum_i \big(\widehat{\boldsymbol{H}}_{i,b_k}^{\text{H}} \boldsymbol{\mathcal{F}}_i \boldsymbol{W}_i' \boldsymbol{\mathcal{F}}_i^{\text{H}} \widehat{\boldsymbol{H}}_{i,b_k} + \text{tr}\{\boldsymbol{\mathcal{F}}_i \boldsymbol{W}_i' \boldsymbol{\mathcal{F}}_i^{\text{H}}\} \boldsymbol{C}_{i,b_k}\big)\big) \boldsymbol{\mathcal{G}}_k \big] \} \rfloor_+ \tag{7.32}$$

where we noted that $\boldsymbol{\mathcal{F}}_k = \boldsymbol{\mathcal{F}}_k \boldsymbol{W}_k \overline{\boldsymbol{E}}_k = \boldsymbol{\mathcal{F}}_k \boldsymbol{W}_k (\boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\text{H}} \widehat{\boldsymbol{H}}_{k,b_k} \boldsymbol{\mathcal{G}}_k)$ and $\lfloor x \rfloor_+ = x$ if $x \geq 0$ and is zero otherwise. This nonnegativity constraint on μ_c stems from the fact that $\mu_c = -\frac{\partial \mathcal{L}}{\partial P_{max,c}} \geq 0$ since indeed the WSMSE can only get smaller if we allow a larger power budget. We then get $\theta_c = \mu_c / \sum_{c'} \mu_{c'}$.

The Perron-Frobenius theory also allows for the optimization of the Lagrange multipliers λ_k. With (7.23), we can reformulate (7.29) as

$$\Delta = \max_{\lambda:\sum_k \lambda_k=1} \min_{\boldsymbol{p}} \sum_k \lambda_k \frac{[\boldsymbol{\Lambda}\,\boldsymbol{p}]_k}{p_k} \tag{7.33}$$

which is the Donsker–Varadhan–Friedland formula [77, Chapter 8] for the Perron Frobenius eigenvalue of Λ. A related formula is the Rayleigh quotient

$$\Delta = \max_{\boldsymbol{q}} \min_{\boldsymbol{p}} \frac{\boldsymbol{q}^T\boldsymbol{\Lambda}\,\boldsymbol{p}}{\boldsymbol{q}^T\boldsymbol{p}} \tag{7.34}$$

where $\boldsymbol{p}$, $\boldsymbol{q}$ are the right and left Perron Frobenius eigenvectors. Comparing (7.34) to (7.33), then apart from normalization factors, we get $\lambda_k/p_k = q_k$ or hence $\lambda_k = p_k q_k$.

The proposed optimization framework is summarized in Table 7.1. Superscripts refer to iteration numbers. The algorithm in Table 7.1 is based on a double loop. The inner loop solves the WEMSE balancing problem in (7.13) whereas the outer loop iteratively transforms the WEMSE balancing problem into the original rate balancing problem in (7.5). The proof of convergence of this transformation is similar to the one in [81].

7.4 Results

In this section, we numerically evaluate the performance of the proposed algorithm. Considering the channel model from Section 6.3, we use for the multipath channel model,

$$\boldsymbol{C}_t = \sum_{n=1}^{N_p} \frac{\alpha_i}{\boldsymbol{v}_i^H\boldsymbol{v}_i}\,\boldsymbol{v}_i\boldsymbol{v}_i^H \tag{7.35}$$

with $\mathrm{tr}\{\boldsymbol{C}_t\} = \sum_{n=1}^{N_p}\alpha_i = M_c$, $\alpha_i = c^{i-1}\alpha_1$ and the $\boldsymbol{v}_i$ are i.i.d. vectors of M_c i.i.d. elements $\mathcal{CN}(0,1)$. We take $N_p = M_c/K$. For all simulations, we take $n_{\max} = 20$, though typically 2-3 inner loop iterations suffice. The algorithm converges after 4-5 (or 13-15) (outer) iterations of m at SNR = 15dB (or 40dB). For all (partial CSIT) algorithms, we evaluate the actual expected rate $\overline{r}_k = \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\,r_k$ by Monte Carlo averaging over 500 channel realizations. The partial CSIT algorithms evaluated are the proposed WEMSE and also Naive Partial CSIT which corresponds to perfect CSIT by assuming the channel estimates to be the true channels. Perfect CSIT algorithms are obtained from WEMSE setting $\rho_D = \infty$. We also evaluate Lower Bound (LB) WEMSE, which considers $\ln(\det(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k))$.

We compare the average rate obtained from WEMSE-based method with the one obtained from balancing the EMSE between user in Figure 7.1, for BC MU-MIMO. Per user EMSE balancing approach is presented in [83], which consists on minimizing the maximum per user EMSE for a DL communicating cell. The latter is a direct extension of [47] under partial CSIT. We observe that solving the max-min expected rate, formulated as WEMSE, results in a better average rate as compared to the one obtained from EMSE approach.

Table 7.1: WEMSE based User Rate Balancing Algorithm

1. initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0,0)} = \boldsymbol{q}_k^{(0,0)} = \frac{P_{\max,c}}{K_c}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$ and $r_k^{\circ(0)}$, initialize $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}_{d_k}$ and $\xi_k^{(0)} = d_k$
2. initialize $\boldsymbol{F}_k^{(0,0)}$ in $\boldsymbol{\mathcal{F}}_k^{(0,0)} = p_k^{(0,0)-1/2}\boldsymbol{F}_k$ from (7.11)
3. **repeat**

    3.1. $m \leftarrow m + 1$

    3.2. **repeat**

    $n \leftarrow n + 1$

    i update $\boldsymbol{G}_k$, $\boldsymbol{\mathcal{G}}_k$, μ_c from (7.31),(7.32)

    ii update $\boldsymbol{F}_k = p_k^{1/2}\boldsymbol{\mathcal{F}}_k$ from (7.11)

    iii update $\boldsymbol{p}$ and $\boldsymbol{q}$ using (7.34)

    3.3 **until** required accuracy is reached or $n \geq n_{\max}$

    3.4 compute $\overline{\boldsymbol{E}}_k^{(m)}$ and update $\boldsymbol{W}_k^{(m)} = (\overline{\boldsymbol{E}}_k^{(m)})^{-1}$

    3.5 determine $t = \min_k \frac{\overline{r}_k^{l\,(m)}}{r_{k_c}^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$, and $\xi_k^{(m)} = d_k + \overline{r}_k^{l\,(m)} - r_k^{\circ(m)}$

    3.6 set $n \leftarrow 0$ and set $(.)^{(n_{\max},m-1)} \rightarrow (.)^{(0,m)}$ in order to re-enter the inner loop

4. **until** required accuracy is reached or $m \geq m_{\max}$

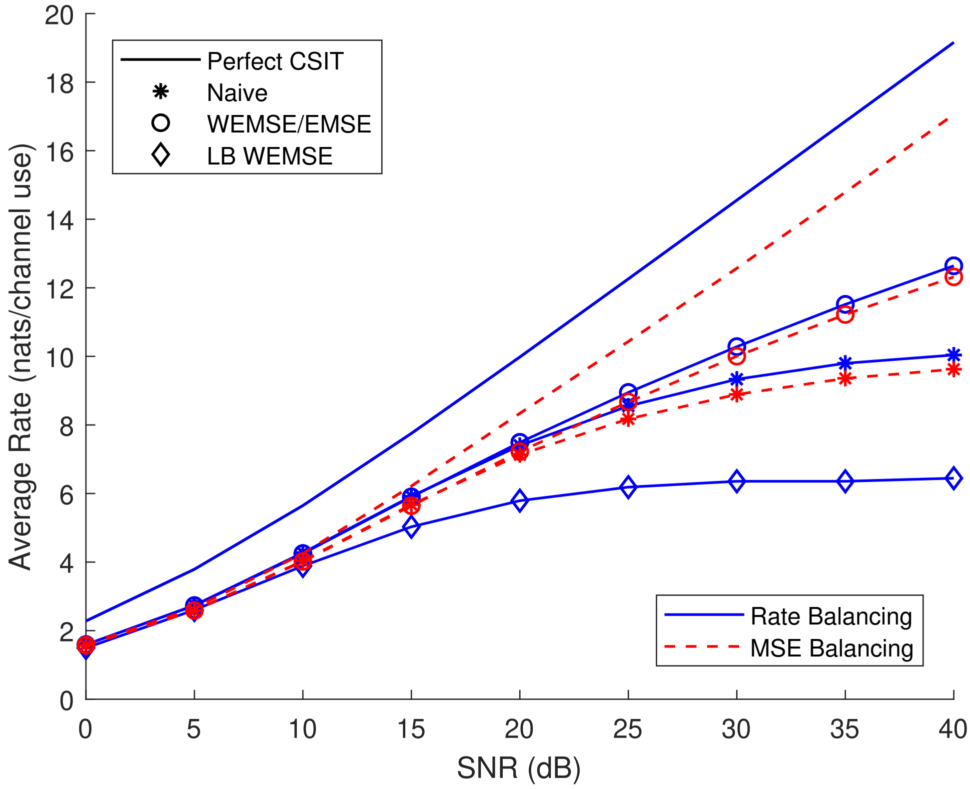

Figure 7.1: Average Rate with Partial CSIT vs. SNR for MU-MIMO BC Channel, $C = 1, K = 3, M_c = 12, N_k = d_k = 2, \rho_D = 10$.

Figure 7.2 considers a MU-MISO cell in order to compare the performances with the approach proposed in [86]. The latter solves the rate balancing problem for partial CSIT by optimizing the BC vector along with the constructed dual MAC vector achieving both, using duality, the same values of the mutual information lower bound. We can see that the resulting curve (refered to as Vect. Opt. in the Figure) coincides with the one from WEMSE approach along with the one from EMSE approach.

7.5 Closing Remarks

In this chapter, we addressed the multiple streams per MIMO user case for which we considered user Erate (Expected rate) balancing, in a multi-cell downlink channel. In particular, we provided an extension of the proposed solution in Chapter 4 to partial CSIT. We transformed the maxmin Erate optimization problem into a min-max weighted EMSE optimization problem which itself was shown to be related to a weighted sum EMSE minimization via Lagrangian duality, involving linearizing the EMSE balancing problem by transforming to EMSE constraints. The associated Lagrange multipliers and user powers get found as left and right eigen vectors of a weighted interference matrix in the Perron-Frobenius theory. Numerical results confirmed that our solution maximizes

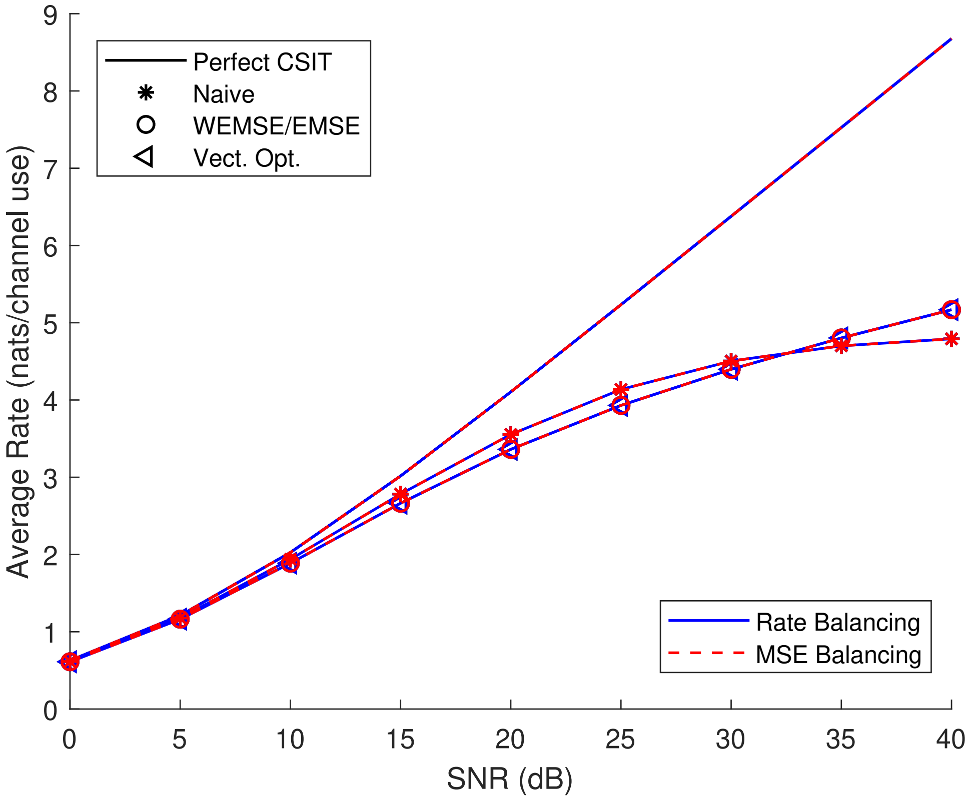

Figure 7.2: Average Rate with Partial CSIT vs. SNR for MU-MISO BC Channel, $C = 1, K = 4, M_c = 8, N_k = d_k = 1, \rho_D = 10$.

the average Erate as compared to the unweighted EMSE optimization for multistream MU MIMO scenario.

Chapter 8

ESIP method for rate balancing

8.1 Overview

In this chapter, we introduce a new algorithm by exploiting a better approximation of the expected rate as the Expected Signal and Interference Power (ESIP) rate. Whereas the ESIP approach have been considered in previous sum utility optimization work, the algorithm here is based on an original minorizer for every individual rate term, different from existing DC programming approaches in sum utility optimization. Furthermore, we investigate the ESIPrate approach within two approximations: *i)* Received signal level ESIP (R-ESIP) and *ii)* Stream level ESIP (S-ESIP). Actually, the use of the expectation operator makes the optimization a daunting task. Therefore, we aim to study which approximation is more simpler to optimize. In [89], a refined analysis of the gap between expected Weighted Sum Rate (WSR) and RESIP-WSR appears, where the actual gap disappears in case of only covariance CSIT. Here, we study study how different R/S-ESIP based approaches are.

The rest of this chapter is organized as follows. The problem formulation with the corresponding derivations for R/S-ESIP approches are presented in Section 8.2. The details of the proposed solution via Lagrangian duality is given in Sections 8.3. In Section 8.4, we consider the total transmit power minimization via ESIPrate approach. Discussions of numerical results are carried out in Section 8.5. Finally, Section 8.6 draws some conclusions.

8.2 ESIP Approach

Considering the same rate balancing problem as in Chapter 7, with imperfect CSIT, we follow here another approximation of the expected rate expression. The following approach will use a rate minorizer for every r_k, similar but not identical to what is used as in the DC programming approach which for the optimization of $\boldsymbol{G}_k$ keeps r_k and linearizes the $r_{\overline{k}}$. The approach does not require the introduction of Rxs. We consider again the (ergodic) rate balancing problem (7.5) where $\overline{r}_k = \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\, r_k$ is now approximated by the Expected Signal and Interference Power (ESIP) rate

$$\begin{aligned}
\overline{r}_k &= \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \ln\det\Big(\boldsymbol{I}+p_k\, \boldsymbol{G}_k^{\rm H}\, \boldsymbol{H}_{k,b_k}^{\rm H}\boldsymbol{R}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\Big)\\
&\approx\ \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \ln\det\Big(\boldsymbol{I}+p_k\, \boldsymbol{G}_k^{\rm H}\, \boldsymbol{H}_{k,b_k}^{\rm H}(\mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\boldsymbol{R}_{\overline{k}})^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\Big)\\
&=\ \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \ln\det\Big(\boldsymbol{I}+p_k\, \boldsymbol{G}_k^{\rm H}\, \boldsymbol{H}_{k,b_k}^{\rm H}\overline{\boldsymbol{R}}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\Big)\\
&\leq \ln\det\Big(\boldsymbol{I}+p_k\, \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\boldsymbol{G}_k^{\rm H}\boldsymbol{H}_{k,b_k}^{\rm H}\overline{\boldsymbol{R}}_{\overline{k}}^{-1}\, \boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\Big) && (8.1)\\
&= \overline{r}_k^{s,{\rm S}} = f_k^{s,{\rm S}}(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) =\ \ln\det\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\rm H}\,\overline{\boldsymbol{B}}_k^{\rm S}(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})\,\boldsymbol{G}_k\Big)\,, && (8.2)
\end{aligned}$$

$$\overline{\boldsymbol{B}}_k^{\rm S}(\overline{\boldsymbol{T}}_k) = \widehat{\boldsymbol{H}}_{k,b_k}^{\rm H}\overline{\boldsymbol{T}}_k^{-1}\,\widehat{\boldsymbol{H}}_{k,b_k} + {\rm tr}\{\overline{\boldsymbol{T}}_k^{-1}\}\,\boldsymbol{C}_{k,b_k} \qquad (8.3)$$

where the $\overline{r}_k$ approximation $\overline{r}_k^s$ in (8.2) in general is neither an upper nor lower bound but in the Massive MIMO limit becomes a tight upper bound.

Let us now consider the following

$$\begin{aligned}
\overline{r}_k &\approx\ \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \ln\det\Big(\boldsymbol{I}+p_k\, \boldsymbol{G}_k^{\rm H}\, \boldsymbol{H}_{k,b_k}^{\rm H}\overline{\boldsymbol{R}}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\Big)\\
&=\ \mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \ln\det\Big(\boldsymbol{I}+p_k\overline{\boldsymbol{R}}_{\overline{k}}^{-1}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\,\boldsymbol{G}_k^{\rm H}\,\boldsymbol{H}_{k,b_k}^{\rm H}\Big)\\
&\leq \ln\det\Big(\boldsymbol{I}+p_k\,\overline{\boldsymbol{R}}_{\overline{k}}^{-1}\,\mathbf{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\boldsymbol{H}_{k,b_k}\boldsymbol{G}_k\boldsymbol{G}_k^{\rm H}\boldsymbol{H}_{k,b_k}^{\rm H}\Big) && (8.4)\\
&= \overline{r}_k^{s,{\rm R}} = f_k^{s,{\rm R}}(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) =\ \ln\det\Big(\boldsymbol{I}+p_k\overline{\boldsymbol{R}}_{\overline{k}}^{-1}\overline{\mathbf{S}}_k\Big)\,. && (8.5)
\end{aligned}$$

This ergodic rate differs from the ergodic rate in (8.2), because in (8.5) the average is taken over the expected received signal covariance matrix of user k, $\overline{\mathbf{S}}_k$ of dimension $(N_k \times N_k)$. Whereas in (8.2), the average is taken over the expected stream level signal covariance matrix of dimension $(d_k \times d_k)$.

Lemma 4. *The approximates* $\overline{r}_k$, $\overline{r}_k^{s,\cdot}$, *can be obtained as* $f_k^{s,\cdot}(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) = \min_{\overline{\boldsymbol{T}}_k} \underline{f}_k^{s,\cdot}(\overline{\boldsymbol{T}}_k, \frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})$ *with*

$$\underline{f}_k^{s,{\rm S}} = \ln\det\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\rm H}\overline{\boldsymbol{B}}_k^{\rm S}(\overline{\boldsymbol{T}}_k)\,\boldsymbol{G}_k\Big)+{\rm tr}\{\breve{\boldsymbol{W}}_k^{\rm S}(\overline{\boldsymbol{T}}_k-\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})\} \qquad (8.6)$$

and

$$\underline{f}_k^{s,{\rm R}} = \ln\det\Big(\boldsymbol{I}+\overline{\boldsymbol{T}}_k^{-1}\,\overline{\mathbf{S}}_k\Big)+{\rm tr}\{\breve{\boldsymbol{W}}_k^{\rm R}(\overline{\boldsymbol{T}}_k-\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})\} \qquad (8.7)$$

where

$$\breve{\boldsymbol{W}}_k^{\mathrm{S}} = \overline{\boldsymbol{T}}_k^{-1}\big(\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{X}_k\,\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{X}_k\,\boldsymbol{C}_{k,b_k}\}\boldsymbol{I}\big)\overline{\boldsymbol{T}}_k^{-1} \tag{8.8}$$

$$\textit{with } \boldsymbol{X}_k = \boldsymbol{G}_k\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k^{\mathrm{S}}(\overline{\boldsymbol{T}}_k)\,\boldsymbol{G}_k\Big)^{-1}\boldsymbol{G}_k^{\mathrm{H}} \tag{8.9}$$

and

$$\breve{\boldsymbol{W}}_k^{\mathrm{R}} = \overline{\boldsymbol{T}}_k^{-1} - \big(\overline{\boldsymbol{T}}_k + \overline{\mathbf{S}}_k\big)^{-1} \tag{8.10}$$

The optimizer is $\overline{\boldsymbol{T}}_k = \frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}$. *Also,* $\underline{f}_k^s$ *is a minorizer for* $f_k^s(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})$ *as a function of* $\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}$.

Indeed, since $f_k^s(.)$ is a convex function, it gets minorized by its tangent at any point:

$$f_k^s(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) \geq \underline{f}_k^s = f_k^s(\overline{\boldsymbol{T}}_k)+\mathrm{tr}\{\frac{\partial f_k^s(\overline{\boldsymbol{T}}_k)}{\partial \overline{\boldsymbol{T}}_k}\,(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}-\overline{\boldsymbol{T}}_k)\} \tag{8.11}$$

and $\breve{\boldsymbol{W}}_k = -\frac{\partial f_k^s(\overline{\boldsymbol{T}}_k)}{\partial \overline{\boldsymbol{T}}_k}$. Note that for the Perron-Frobenius theory, we need a function that is linear in $\frac{\boldsymbol{p}_{\overline{k}}}{p_k}$, hence we need to work with $\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}$ instead of $\overline{\boldsymbol{R}}_{\overline{k}}$.

8.3 Proposed Solution

The modifications in the Lagrangian formulation in Section 4.3.1 are now

$$\mathcal{L}^{\mathrm{ESIP}} = -t + \sum_k \breve{\lambda}_k'(t\,r_k^o - \underline{f}_k^s) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}), \tag{8.12}$$

where the Lagrangian for stream level ESIP is

$$\begin{aligned}\mathcal{L}^{\mathrm{SESIP}} &= -\,t - \sum_k \breve{\lambda}_k'\Big(\,\mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k^{\mathrm{S}}\boldsymbol{G}_k\Big) - \frac{1}{p_k}\,\mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}}\overline{\boldsymbol{R}}_{\overline{k}}\}+\mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}}\overline{\boldsymbol{T}}_k\}-t\,r_k^o\Big) \\ &\quad + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \qquad (8.13) \\ &= -\,t + \sum_k \breve{\lambda}_k(\frac{1}{p_k\,\breve{\xi}_k^{\mathrm{S}}}\,\mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}}\overline{\boldsymbol{R}}_{\overline{k}}\}-1) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \qquad (8.14) \\ &\text{with } \breve{\xi}_k^{\mathrm{S}} = \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}}\overline{\boldsymbol{T}}_k\} + \mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k^{\mathrm{S}}\boldsymbol{G}_k\Big) - t\,r_k^o, \qquad (8.15)\end{aligned}$$

$$\breve{\lambda}_k' = \breve{\lambda}_k/\breve{\xi}_k^{\mathrm{S}}, \overline{\boldsymbol{B}}_k^{\mathrm{S}} = \overline{\boldsymbol{B}}_k^{\mathrm{S}}(\overline{\boldsymbol{T}}_k),$$

and the Lagrangian for received signal level is

$$\begin{aligned}\mathcal{L}^{\mathrm{RESIP}} =& -t - \sum_k \breve{\lambda}'_k \Big(\mathrm{lndet}\Big(\boldsymbol{I} + \overline{\boldsymbol{T}}_k^{-1}\,\overline{\mathbf{S}}_k\Big) - \frac{1}{p_k}\,\mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{R}}\overline{\boldsymbol{R}}_{\overline{k}}\} + \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{R}}\overline{\boldsymbol{T}}_k\} - t\,r_k^o\Big) \\ & + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \quad (8.16)\\ =& -t + \sum_k \breve{\lambda}_k \Big(\frac{1}{p_k\,\breve{\xi}_k^{\mathrm{R}}}\,\mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{R}}\overline{\boldsymbol{R}}_{\overline{k}}\} - 1\Big) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \quad (8.17)\\ & \text{with } \breve{\xi}_k^{\mathrm{R}} = \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{R}}\overline{\boldsymbol{T}}_k\} + \mathrm{lndet}\Big(\boldsymbol{I} + \overline{\boldsymbol{T}}_k^{-1}\,\overline{\mathbf{S}}_k\Big) - t\,r_k^o, \quad (8.18)\end{aligned}$$

The balancing of the rates in (7.5) or equivalently the weighted interference plus noise powers in (8.14) or (8.17), i.e.,

$$\max_{\breve{\lambda}} \min_{\boldsymbol{G},\boldsymbol{p}} \sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k} \frac{\mathrm{tr}(\breve{\boldsymbol{W}}_k \overline{\boldsymbol{R}}_{\overline{k}})}{p_k} + \sum_c \mu_c \Big(\sum_{i:b_i=c} \mathrm{tr}\{\boldsymbol{\mathcal{G}}_i^{\mathrm{H}}\boldsymbol{\mathcal{G}}_i\} - P_{max,c}\Big) \quad (8.19)$$

leads to the same problem formulation as in (7.23) with this time

$$\breve{\boldsymbol{\Lambda}} = \breve{\boldsymbol{\xi}}^{-1}\breve{\boldsymbol{\Psi}} + \frac{1}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}\breve{\boldsymbol{\xi}}^{-1}\breve{\boldsymbol{\sigma}}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}} \text{ with} \quad (8.20)$$

$$[\breve{\boldsymbol{\Psi}}]_{ij} = \begin{cases} \mathrm{tr}\{\breve{\boldsymbol{W}}_i(\widehat{\boldsymbol{H}}_{i,b_j}\boldsymbol{G}_j\boldsymbol{G}_j^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_j}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_j^{\mathrm{H}}\boldsymbol{C}_{i,b_j}\boldsymbol{G}_j\}\boldsymbol{I})\}, & i \neq j \\ 0, & i = j \end{cases} \quad (8.21)$$

$$\boldsymbol{\sigma}_i = \sigma_n^2\,\mathrm{tr}\{\breve{\boldsymbol{W}}_i\}, \breve{\boldsymbol{\xi}} = \mathrm{diag}(\breve{\xi}_1, \ldots, \breve{\xi}_K)\,. \quad (8.22)$$

The Tx BF and stream power optimization will be based on $\sum_i \frac{\breve{\lambda}_i}{\breve{\xi}_i} \underline{f}_i^s$, for both SESIP and RESIP approximates, which becomes (apart from noise terms) as described in the following.

SESIP Tx BF

We have from (8.13) (see Appendix B.1)

$$\sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k} \underline{f}_k^{s,\mathrm{S}} = \sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k^{\mathrm{S}}}\,\mathrm{lndet}\Big(\boldsymbol{I} + \boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k^{\mathrm{S}}\,\boldsymbol{G}_k\Big) - \sum_k \mathrm{tr}\{p_k \boldsymbol{G}_k^H \overline{\boldsymbol{A}}_k^{\mathrm{S}} \boldsymbol{G}_k\} \quad (8.23)$$

$$\text{with } \overline{\boldsymbol{A}}_k^{\mathrm{S}} = \sum_{i \neq k} \frac{\breve{\lambda}_i}{p_i\,\breve{\xi}_i} \Big(\widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}} \breve{\boldsymbol{W}}_i^{\mathrm{S}} \widehat{\boldsymbol{H}}_{i,b_k} + \mathrm{tr}\{\breve{\boldsymbol{W}}_i^{\mathrm{S}}\}\boldsymbol{C}_{i,b_k}\Big)\,. \quad (8.24)$$

For the optimal Tx BF $\boldsymbol{G}_k$, the gradient of $\mathcal{L}^{\mathrm{SESIP}}$ yields

$$\frac{\partial \mathcal{L}^{\mathrm{SESIP}}}{\partial \boldsymbol{G}^*} = 0 \Leftrightarrow \frac{\breve{\lambda}_k}{\breve{\xi}_k^{\mathrm{S}}} \overline{\boldsymbol{B}}_k^{\mathrm{S}} \boldsymbol{G}_k\,(\boldsymbol{I} + \boldsymbol{G}_k^H \overline{\boldsymbol{B}}_k^{\mathrm{S}} \boldsymbol{G}_k)^{-1} - p_k(\overline{\boldsymbol{A}}_k^{\mathrm{S}} + \mu_{b_k}\boldsymbol{I})\,\boldsymbol{G}_k = 0\,. \quad (8.25)$$

The solution is the d_k maximal generalized eigen vectors

$$\boldsymbol{G}'_k = V_{1:d_k}(\overline{\boldsymbol{B}}_k^{\rm S}, \overline{\boldsymbol{A}}_k^{\rm S} + \mu_{b_k}\boldsymbol{I}), \boldsymbol{G}_k = \boldsymbol{G}'_k \overline{\boldsymbol{P}}_k^{1/2}, \boldsymbol{\mathcal{G}}_k = \boldsymbol{G}_k\sqrt{p_k} \tag{8.26}$$

where the $\overline{\boldsymbol{P}}_k = \text{diag}(p_{k,1}, \ldots, p_{k,d_k})$, $\text{tr}\{\overline{\boldsymbol{P}}_k\} = 1$, are the relative stream powers. Indeed, (8.25) represents the definition of generalized eigen vectors. Consider

$$\Sigma_k^{(1)} = \boldsymbol{G}_k'^{\rm H}\overline{\boldsymbol{B}}_k^{\rm S}\boldsymbol{G}'_k,\ \Sigma_k^{(2)} = \boldsymbol{G}_k'^{\rm H}\overline{\boldsymbol{A}}_k^{\rm S}\boldsymbol{G}'_k \tag{8.27}$$

then the generalized eigen vectors $\boldsymbol{G}'_k$ of $\overline{\boldsymbol{B}}_k, \overline{\boldsymbol{A}}_k^{\rm S} + \mu_{b_k}\boldsymbol{I}$ lead to diagonal matrices $\Sigma_k^{(1)}$, $\Sigma_k^{(2)} + \mu_{b_k}\boldsymbol{G}_k'^{\rm H}\boldsymbol{G}'_k$. Note that the normalized $\boldsymbol{G}'_k$ are not orthogonal. Then (8.25) represents the generalized eigen vector condition with associated generalized eigen values in the diagonal matrix $\frac{p_k\breve{\xi}_k}{\breve{\lambda}_k}(\boldsymbol{I} + \Sigma_k^{(1)}\overline{\boldsymbol{P}}_k)$. Also, plugging in generalized eigen vectors into (8.23) reveals that one should choose the eigen vectors associated to d_k maximal eigen values to maximize (8.23). Now, premultiplying both sides of (8.25) by $p_k\boldsymbol{G}_k^{\rm H}$, summing over all users $k : b_k = c$, taking trace and identifying the last term with $\sum_{k:b_k=c} p_k\text{tr}\{\boldsymbol{G}_k^{\rm H}\boldsymbol{G}_k\} = P_{max,c}$ allows to solve for

$$\mu_c = \frac{1}{P_{max,c}}\left\lfloor \sum_{k:b_k=c}\text{tr}\{\frac{\breve{\lambda}_k}{\breve{\xi}_k}\Sigma_k^{(1)}\overline{\boldsymbol{P}}_k(\boldsymbol{I}+\Sigma_k^{(1)}\overline{\boldsymbol{P}}_k)^{-1} - p_k\Sigma_k^{(2)}\overline{\boldsymbol{P}}_k\}\right\rfloor_+ . \tag{8.28}$$

The $\overline{\boldsymbol{P}}_k$ are themselves found from an interference leakage aware water filling (ILAWF) operation. Substituting $\boldsymbol{G}'_k$ into term k of (8.23), dividing by p_k, and accounting for the constraint $\text{tr}\{\overline{\boldsymbol{P}}_k\} = 1$ by Lagrange multiplier ν_k, we get the Lagrangian

$$\begin{aligned} &\frac{\breve{\lambda}_k}{p_k\,\breve{\xi}_k}\ln\det\big(\boldsymbol{I}+\Sigma_k^{(1)}\overline{\boldsymbol{P}}_k\big) - \text{tr}\{(\Sigma_k^{(2)} + \nu_k\boldsymbol{I})\overline{\boldsymbol{P}}_k\} = \\ &\frac{\breve{\lambda}_k}{p_k\,\breve{\xi}_k}\ln\det\big(\boldsymbol{I}+\Sigma_k^{(1)}\overline{\boldsymbol{P}}_k\big) - \text{tr}\{(\text{diag}(\Sigma_k^{(2)})+\nu_k\boldsymbol{I})\overline{\boldsymbol{P}}_k\}. \end{aligned} \tag{8.29}$$

Maximizing w.r.t. $\overline{\boldsymbol{P}}_k$ leads to the ILAWF

$$\overline{\boldsymbol{P}}_k = \left\lfloor \frac{\breve{\lambda}_k}{p_k\,\breve{\xi}_k}\,(\text{diag}(\Sigma_k^{(2)}) + \nu_k\boldsymbol{I})^{-1} - \Sigma_k^{-(1)}\right\rfloor_+ \tag{8.30}$$

where the Lagrange multiplier ν_k is adjusted (e.g. by bisection) to satisfy $\text{tr}\{\overline{\boldsymbol{P}}_k\} = 1$. Elements in $\overline{\boldsymbol{P}}_k$ corresponding to zeros in $\Sigma_k^{(1)}$ should also be zero.

RESIP Tx BF

We have from (8.16)

$$\sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k}\underline{f}_k^{s,\rm R} = \sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k^{\rm R}}\ln\det\Big(\boldsymbol{I}+\overline{\boldsymbol{T}}_k^{-1}\,\overline{\mathbf{S}}_k\Big) - \sum_k \text{tr}\{p_k\boldsymbol{G}_k^H\overline{\boldsymbol{A}}_k^{\rm R}\boldsymbol{G}_k\}, \tag{8.31}$$

$$\text{with } \overline{\boldsymbol{A}}_k^{\rm R} = \sum_{i\neq k}\frac{\breve{\lambda}_i}{p_i\,\breve{\xi}_i}\big(\widehat{\boldsymbol{H}}_{i,b_k}^{\rm H}\breve{\boldsymbol{W}}_i^{\rm R}\widehat{\boldsymbol{H}}_{i,b_k} + \text{tr}\{\breve{\boldsymbol{W}}_i^{\rm R}\}\boldsymbol{C}_{i,b_k}\big)\,. \tag{8.32}$$

For the optimal Tx BF $\boldsymbol{G}_k$, the gradient of $\mathcal{L}^{\text{RESIP}}$ yields

$$\frac{\partial \mathcal{L}^{\text{RESIP}}}{\partial \boldsymbol{G}^*} = 0 \Leftrightarrow \frac{\breve{\lambda}_k}{\breve{\xi}_k^{\text{R}}} \overline{\boldsymbol{B}}_k^{\text{R}} \boldsymbol{G}_k - p_k (\overline{\boldsymbol{A}}_k^{\text{R}} + \mu_{b_k} \boldsymbol{I}) \, \boldsymbol{G}_k = 0 \,, \tag{8.33}$$

$$\text{with } \overline{\boldsymbol{B}}_k^{\text{R}} = \widehat{\boldsymbol{H}}_{k,b_k}^{\text{H}} (\boldsymbol{I} + \overline{\boldsymbol{T}}_k^{-1} \overline{\text{S}}_k)^{-1} \overline{\boldsymbol{T}}_k^{-1} \widehat{\boldsymbol{H}}_{k,b_k} + \text{tr}\{(\boldsymbol{I} + \overline{\boldsymbol{T}}_k^{-1} \overline{\text{S}}_k)^{-1} \overline{\boldsymbol{T}}_k^{-1}\} \boldsymbol{C}_{k,b_k} \,. \tag{8.34}$$

The solution is the d_k maximal generalized eigen vectors

$$\boldsymbol{G}_k^{'} = V_{1:d_k} (\overline{\boldsymbol{B}}_k^{\text{R}}, \overline{\boldsymbol{A}}_k^{\text{R}} + \mu_{b_k} \boldsymbol{I}), \boldsymbol{G}_k = \boldsymbol{G}_k^{'} \overline{\boldsymbol{P}}_k^{1/2}, \boldsymbol{\mathcal{G}}_k = \boldsymbol{G}_k \sqrt{p_k}. \tag{8.35}$$

Then, we can solve for μ_c by multiplying (8.33) from the left by $\boldsymbol{G}_k^{\text{H}}$ and summing over the users in cell c, i.e.,

$$\mu_c = 1/P_{\text{max},c} \sum_{k:b_k=c} \Big[\frac{\breve{\lambda}_k}{\breve{\xi}_k} \boldsymbol{G}_k^H \overline{\boldsymbol{B}}_k^{\text{R}} \boldsymbol{G}_k - p_k \boldsymbol{G}_k^H \overline{\boldsymbol{A}}_k^{\text{R}} \boldsymbol{G}_k \Big] \,. \tag{8.36}$$

Now, we have to find the corresponding $\bar{\boldsymbol{P}}_k$. Substituting $\bar{\boldsymbol{P}}_k$ in (8.31), we can write the following

$$\text{lndet}\Big(\boldsymbol{I} + \sum_{i=1}^{d_k} \overline{p}_{k,i} [\boldsymbol{J}_k]_{i,i} \Big) - \sum_{i=1}^{d_k} \overline{p}_{k,i} a_{k,i} \tag{8.37}$$

where $\bar{\boldsymbol{P}}_k = \text{diag}(\bar{\boldsymbol{p}}_k), \bar{\boldsymbol{p}}_k = [\overline{p}_1 \ldots \overline{p}_{d_k}]$, $a_{k,i} = \frac{p_k \breve{\xi}_k^{\text{R}}}{\breve{\lambda}_k} [\boldsymbol{G}_k^{'}]_{:,i}^{\text{H}} \overline{\boldsymbol{A}}_k^{\text{R}} [\boldsymbol{G}_k^{'}]_{:,i}$, and

$$[\boldsymbol{J}_k]_{i,i} = \overline{\boldsymbol{T}}_k^{-1} (\widehat{\boldsymbol{H}}_{k,b_k} [\boldsymbol{G}_k^{'}]_{:,i} [\boldsymbol{G}_k^{'}]_{:,i}^{\text{H}} \widehat{\boldsymbol{H}}_{k,b_k}^{\text{H}} + [\boldsymbol{G}_k^{'}]_{:,i}^{\text{H}} \boldsymbol{C}_{k,b_k} [\boldsymbol{G}_k^{'}]_{:,i} \boldsymbol{I}).$$

Let $\boldsymbol{U}_k(\bar{\boldsymbol{p}}_k) = \sum_{i=1}^{d_k} \overline{p}_{k,i} [\boldsymbol{J}_k]_{i,i}$ and $\boldsymbol{a}_k = [a_1 \ldots a_{d_k}]$, we can rewrite (8.37) as

$$\text{lndet}\Big(\boldsymbol{I} + \boldsymbol{U}_k(\bar{\boldsymbol{p}}_k) \Big) - \boldsymbol{a}_k \bar{\boldsymbol{p}}_k^T - \nu_k \boldsymbol{1}_{d_k} \bar{\boldsymbol{p}}_k^T. \tag{8.38}$$

with ν_k being the Lagrangian multiplier for the constraint $||\bar{\boldsymbol{p}}_k||_1 = 1$ and $\boldsymbol{1}_{d_k}$ is a line vector of ones, of length d_k.

In the following, we omit the user indices k for simplicity. Consider the Taylor series expansion for matrices $\boldsymbol{X}, \boldsymbol{Y}$ of dimension N_k,

$$\ln\det(\boldsymbol{X} + \boldsymbol{Y}) \approx \ln\det(\boldsymbol{X}) + \text{tr}\{\boldsymbol{X}^{-1}\boldsymbol{Y}\} - \frac{1}{2}\text{tr}\{\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{Y}\}. \tag{8.39}$$

Let $\bar{\boldsymbol{p}} = \hat{\bar{\boldsymbol{p}}} + \tilde{\bar{\boldsymbol{p}}}$ and choose $\boldsymbol{X} = \boldsymbol{I} + \boldsymbol{U}(\hat{\bar{\boldsymbol{p}}})$ and $\boldsymbol{Y} = \boldsymbol{U}(\tilde{\bar{\boldsymbol{p}}})$, we obtain the following Lagrangian

$$\tilde{\bar{\boldsymbol{p}}} \mathbf{v}^T - \frac{1}{2} \tilde{\bar{\boldsymbol{p}}} \boldsymbol{Z} \tilde{\bar{\boldsymbol{p}}}^T - \boldsymbol{a} \tilde{\bar{\boldsymbol{p}}}^T - \nu \boldsymbol{1} \tilde{\bar{\boldsymbol{p}}}^T \tag{8.40}$$

where $\mathbf{v} = [\text{v}_1 \ldots \text{v}_{d_k}]$ with $\text{v}_i = \text{tr}\{(\boldsymbol{I} + \boldsymbol{U}(\hat{\bar{\boldsymbol{p}}}))^{-1} [\boldsymbol{J}]_{i,i}\}$, and $\boldsymbol{Z}$ is a matrix with the elements $[\boldsymbol{Z}]_{i,j} = \text{tr}\{(\boldsymbol{I} + \boldsymbol{U}(\hat{\bar{\boldsymbol{p}}}))^{-1} [\boldsymbol{J}]_{i,i} (\boldsymbol{I} + \boldsymbol{U}(\hat{\bar{\boldsymbol{p}}}))^{-1} [\boldsymbol{J}]_{j,j}\}$. Taking the gradient of (8.40) w.r.t. $\tilde{\bar{\boldsymbol{p}}}$, we get

$$\mathbf{v}^T - \boldsymbol{Z} \tilde{\bar{\boldsymbol{p}}}^T - \boldsymbol{a}^T - \nu \boldsymbol{1}^T = 0 \tag{8.41}$$

$$\Leftrightarrow \tilde{\bar{\boldsymbol{p}}}^T = \boldsymbol{Z}^{-1} (\mathbf{v}^T - \boldsymbol{a}^T - \nu \boldsymbol{1}^T) \tag{8.42}$$

thus

$$\bar{\boldsymbol{p}}^T = [\hat{\bar{\boldsymbol{p}}}^T + \boldsymbol{Z}^{-1}(\mathbf{v}^T - \boldsymbol{a}^T - \nu\mathbf{1}^T)]_+ \tag{8.43}$$

and ν gets determined by $\bar{\boldsymbol{p}}\mathbf{1}^T = 1$. Having $\hat{\bar{\boldsymbol{p}}}$ being $\bar{\boldsymbol{p}}$ at the current iteration i, we can write

$$\bar{\boldsymbol{p}}^{T(i+1)} = [\bar{\boldsymbol{p}}^{T(i)} + \boldsymbol{Z}^{-1}(\mathbf{v}^T - \boldsymbol{a}^T - \nu\mathbf{1}^T)]_+ \tag{8.44}$$

and thus

$$\overline{\boldsymbol{P}} = \mathrm{diag}(\bar{\boldsymbol{p}}). \tag{8.45}$$

8.3.1 Discussion

The approaches SESIP and RESIP lead to two precoder designs resulting from taking the expectation of d_k dimensional vs. N_k dimensional matrices, respectively. Actually, the smaller the dimension, the more averaging occurs of the fixed amount of random entries, and so bringing in the expectation $\mathbb{E}$ inside $\log\det()$ should be a tighter upper bound for the smaller dimension case. Nevertheless, when $d_k = N_k$, this reasoning stops applying and either one can be larger or smaller. Simulation results, however, have shown that the difference is negligible when $d_k = N_k$, S/R-ESIP becoming then equivalent.

8.3.2 Algorithm

Table 8.1 details the Algorithms 7.1 and 7.2 for SESIP and RESIP, respectively.

8.3.3 Covergence

From the problem formulation in (8.19) (omitting the power constraints), let us define $\Delta^{\mathrm{ESIP}} = \frac{1}{\breve{\xi}_k}\frac{\mathrm{tr}(\breve{\boldsymbol{W}}_k\overline{\boldsymbol{R}}_{\bar{k}})}{p_k}$, we have

$$\Delta^{\mathrm{ESIP}(m)} = \max_k \frac{1}{\breve{\xi}_k^{(m-1)}}\frac{\mathrm{tr}(\breve{\boldsymbol{W}}_k^{(m-1)}\overline{\boldsymbol{R}}_{\bar{k}}^{(m)})}{p_k^{(m)}} \overset{(a)}{\leq} \max_k \frac{1}{\breve{\xi}_k^{(m-1)}}\frac{\mathrm{tr}(\breve{\boldsymbol{W}}_k^{(m-1)}\overline{\boldsymbol{R}}_{\bar{k}}^{(m-1)})}{p_k^{(m-1)}} \overset{(b)}{\leq} 1 \tag{8.46}$$

where (a) holds due the minimization step at the mth iteration, and (b) is satisfied due to the definition of $\breve{\xi}_k^{(m-1)} = \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{(m-1)}\overline{\boldsymbol{T}}_k^{(m-1)}\} + \overline{r}_k^{s(m-1)} - t\, r_k^{o(m-1)} \geq \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{(m-1)}\overline{\boldsymbol{T}}_k^{(m-1)}\} = \frac{\mathrm{tr}(\breve{\boldsymbol{W}}_k^{(m-1)}\overline{\boldsymbol{R}}_{\bar{k}}^{(m-1)})}{p_k^{(m-1)}}$. Hence, $t \geq 1$. Of course, during the convergence $t > 1$. The increasing rate targets $\{r_k^{\circ(m)}\}$ constantly catch up with the increasing rates $\{r_k^{(m)}\}$. Now, the rates are upper bounded by the single user MIMO rates (using all power), and hence the rates will converge and the sequence t will converge to 1. That means that for at least one user k, $r_k^{(\infty)} = r_k^{\circ(\infty)}$. The question is whether this will be the case for all users, as is required for rate balancing. Now, the weighted interference plus noise powers balancing leads at every outer iteration m to $\frac{1}{\breve{\xi}_k^{(m-1)}}\frac{\mathrm{tr}(\breve{\boldsymbol{W}}_k^{(m-1)}\overline{\boldsymbol{R}}_{\bar{k}}^{(m)})}{p_k^{(m)}} = \Delta^{\mathrm{ESIP}(m)}, \forall k_c$. At convergence, this becomes $\frac{\text{tr}\{\breve{\boldsymbol{W}}_k^{(\infty)}\overline{\boldsymbol{T}}_k^{(\infty)}\}}{\breve{\xi}_{k_c}^{(\infty)}} = \Delta^{(\infty)}$ where $\breve{\xi}_k^{(\infty)} = \text{tr}\{\breve{\boldsymbol{W}}_k^{(\infty)}\overline{\boldsymbol{T}}_k^{(\infty)}\} + \overline{r}_k^{s(\infty)} - r_k^{\circ(\infty)}$. Hence, if we have convergence because for one user k_∞ we arrive at $r_{k_\infty}^{(\infty)} = r_{k_\infty}^{\circ(\infty)}$, then this implies $\Delta^{(\infty)} = 1$ which implies $r_k^{(\infty)} = r_k^{\circ(\infty)}, \forall k$. Hence, the rates will be maximized and balanced.

8.4 To Power Minimization

8.4.1 Total power minimization

We consider following optimization problem

$$\begin{aligned} \min_{\boldsymbol{p},\boldsymbol{G}} \ & f(||\boldsymbol{p}||_1, t) \\ \text{s.t.} \ & r_k(\boldsymbol{p},\boldsymbol{G})/r_k^{\triangle} \geq t, \ \ \forall k \\ & ||\boldsymbol{p}||_1 \leq P_{\max} \end{aligned} \tag{8.47}$$

where

$$f(||\boldsymbol{p}||_1, t) = u(t-1)(||\boldsymbol{p}||_1 + t) - t \tag{8.48}$$

$$\text{with } u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0. \end{cases} \tag{8.49}$$

The problem in (8.47) describes MMR problem when $t < 1$. When $t \geq 1$ (8.47) becomes a total transmit power minimization problem subject to per user rate targets (BC case), as

$$\begin{aligned} \min_{\boldsymbol{p},\boldsymbol{G}} \ & ||\boldsymbol{p}||_1 \\ \text{s.t.} \ & r_k(\boldsymbol{p},\boldsymbol{G})/r_k^{o} \geq 1, \ \ \forall k \\ & ||\boldsymbol{p}|| \leq P_{\max} \end{aligned} \tag{8.50}$$

r_k^o here are the rate targets, i.e., $r_k^o = r_k^{\triangle}$, thus, $t = 1$.

Similar to Chapter 5, total power minimization is held in a two-stage process:

1) When $t < 1$, we proceed to rate balancing as we do not have enough power to guarantee all rate targets.

2) When $t \geq 1$, we change the power allocation strategy considering the constraints are fulfilled which minimizes the total transmit power.

Since the ESIP-based MMR problem is formulated as max-min weighted interference plus noise powers, the related power minimization problem is constrained by

$$r_k = r_k^o \Leftrightarrow 1/p_k \text{tr}\{\breve{\boldsymbol{W}}_k \overline{\boldsymbol{R}}_{\overline{k}}\} = \breve{\xi}_k, \ \ \forall k.$$

Table 8.1: ESIPrate based User Rate Balancing Algorithm

1. initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0,0)} = \boldsymbol{q}_k^{(0,0)} = \frac{P_{\max,c}}{K}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$, $r_k^{\circ(0)}$, $\breve{\xi}_k^{(0)}$, and $\breve{\boldsymbol{W}}_k^{(0)} = \boldsymbol{I}$

2. **repeat**

   2.1. $m \leftarrow m + 1$

   2.2. **repeat**

   $n \leftarrow n + 1$

   i update $\overline{\boldsymbol{A}}_k$ from (8.24) **Algorithm 7.1**] or (8.32) [**Algorithm 7.2**]

   ii update μ_c and $\boldsymbol{G}_k^{'}$ from (8.26),(8.28) [**Algorithm 7.1**] or (8.35),(8.36) [**Algorithm 7.2**]

   iii update $\overline{\boldsymbol{P}}_k$ from (8.30) **Algorithm 7.1**] or (8.45) [**Algorithm 7.2**]

   iv update $\boldsymbol{p}$ and $\boldsymbol{q}$ as maximal eigen vectors of $\breve{\boldsymbol{\Lambda}}$ in (8.20)

   2.3 **until** required accuracy is reached or $n \geq n_{\max}$

   2.4 compute $\overline{\boldsymbol{B}}_k^{\mathrm{S}}(\overline{\boldsymbol{T}}_k)$ and update $\breve{\boldsymbol{W}}_k^{\mathrm{S}}$ from (8.8)[**Algorithm 7.1**] or compute $\overline{\boldsymbol{B}}_k^{\mathrm{R}}(\overline{\boldsymbol{T}}_k)$ and update $\breve{\boldsymbol{W}}_k^{\mathrm{R}}$ from (8.10)[**Algorithm 7.2**]

   2.5 compute $\overline{r}_k^{s\,(m)} = \mathrm{lndet}\Big(\boldsymbol{I} + \boldsymbol{G}_k^{\mathrm{H}}\,\overline{\boldsymbol{B}}_k(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})\ \boldsymbol{G}_k\Big)$ [**Algorithm 7.1**] or $\overline{r}_k^{s\,(m)} = \mathrm{lndet}\Big(\boldsymbol{I} + \overline{\boldsymbol{T}}_k^{-1}\ \mathbf{S}_k\Big)$ [**Algorithm 7.2**] and determine $t = \min_k \frac{\overline{r}_k^{s\,(m)}}{r_{k_c}^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$

   2.6 update $\breve{\xi}_k^{\mathrm{S}}$ from (8.15) [**Algorithm 7.1**] or $\breve{\xi}_k^{\mathrm{R}}$ from (8.18) [**Algorithm 7.2**]

3. **until** required accuracy is reached or $m \geq m_{\max}$

Similarly to what has been considered for the PM in Chapter 5, we collect the per user weighted interference plus noise powers in a diagonal matrix $\breve{\boldsymbol{\epsilon}}_w$ as follows

$$[\breve{\boldsymbol{\epsilon}}_w]_{k,k} = 1/p_k \mathrm{tr}\{\breve{\boldsymbol{W}}_k \overline{\boldsymbol{R}}_{\overline{k}}\} \tag{8.51}$$

$$\breve{\boldsymbol{\epsilon}}_w \mathbf{1}_K = \mathrm{diag}(\boldsymbol{p})^{-1} \left[\breve{\boldsymbol{\Psi}}\boldsymbol{p} + \boldsymbol{\sigma}\right] \tag{8.52}$$

The corresponding optimal power allocation to achieve the targets $\breve{\boldsymbol{\xi}}$ is then

$$\boldsymbol{p} = (\breve{\boldsymbol{\xi}} - \breve{\boldsymbol{\Psi}})^{-1}\boldsymbol{\sigma},. \tag{8.53}$$

Then, we set the new power constraint for MMR optimization as $P_{\max} = P$ with

$$P = ||\boldsymbol{p}||_1. \tag{8.54}$$

which completes the optimization framework. The proposed algorithm is summarized in Table 8.2.

Total power minimization for IBC

For IBC case, one can proceed as follows:

1) when $t < 1$, optimize the rate balancing problem while fulfilling the per cell power constraints by the means of Lagrangian multipliers μ_c.

2) when $t \geq 1$, change the power allocation strategy $\boldsymbol{p}$ to meet the targets with equality, which minimizes the total transmit power $P = ||\boldsymbol{p}||_1$. In this case, the corresponding MMR problem is constrained only by this new total transmit power, thus, the Lagrangian depends on $\mu_o = \sum_c \mu_c$.

8.4.2 Per cell power minimization/balancing

Now, consider the following power minimization problem

$$\begin{aligned}
\min_{\boldsymbol{p},\boldsymbol{G}} \ & P^s \\
\text{s.t.} \ & r_k(\boldsymbol{p},\boldsymbol{G})/r_k^o \geq 1, \ \ \forall k \\
& \sum_{k:b_k=c} \mathrm{tr}\{\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{\mathcal{G}}_k\} \leq P^s, \ \ \forall c \\
& P_c \leq P_{\max,c}, \ \ \forall c
\end{aligned} \tag{8.55}$$

where $P^s = \max_c P_c$ and $P_c = \boldsymbol{c}_c^T\boldsymbol{p} = \sum_{k:b_k=c} \mathrm{tr}\{\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{\mathcal{G}}_k\}$.

Similar to the total power minimization case, the optimal power allocation to achieve the targets $\breve{\boldsymbol{\xi}}$ is again

$$\boldsymbol{p} = (\breve{\boldsymbol{\xi}} - \breve{\boldsymbol{\Psi}})^{-1}\boldsymbol{\sigma}.$$

Table 8.2: SESIPrate based Total Power Minimization Algorithm

1. For predefined r_k^o, initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0,0)} = \boldsymbol{q}_k^{(0,0)} = \frac{P_{\max,c}}{K}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$, $\breve{\xi}_k^{(0)}$, and $\breve{\boldsymbol{W}}_k^{(0)} = \boldsymbol{I}$, $t^{(0)} = 0$
2. **repeat**
   - 2.1 $m \leftarrow m + 1$
   - 2.2 update $\overline{\boldsymbol{A}}_k$ from (8.24)
   - 2.3 update $\boldsymbol{G}_k^{'}$ from (8.26)
   - 2.4 update $\overline{\boldsymbol{P}}_k$ from (8.30)

     if $t^{(m-1)} < 1$

     update $\boldsymbol{p}$ and $\boldsymbol{q}$ as maximal eigen vectors of $\breve{\boldsymbol{\Lambda}}$ in (8.20)

     else

     update $\boldsymbol{p}$ with (8.53) and do $P = ||\boldsymbol{p}||_1$ (8.54)

     update $\boldsymbol{q}$ as maximal left eigen vector of $\breve{\boldsymbol{\Lambda}}(P_{\max} = P)$

     end if
   - 2.5 compute $\overline{\boldsymbol{B}}_k^{\mathrm{S}}(\overline{\boldsymbol{T}}_k)$ and update $\breve{\boldsymbol{W}}_k^{\mathrm{S}}$ from (8.8)
   - 2.6 compute $\overline{r}_k^{s\,(m)} = \mathrm{lndet}\Big(\boldsymbol{I} + \boldsymbol{G}_k^{\mathrm{H}}\, \overline{\boldsymbol{B}}_k^{\mathrm{S}}(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\bar{k}})\;\; \boldsymbol{G}_k\Big)$ and determine $t^{(m)} = \min_k \frac{\overline{r}_k^{s\,(m)}}{r_k^o}$
   - 2.7 **if** $t^{(m)} < 1$

     update $\breve{\xi}_k^{\mathrm{S},(m)} = \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S},(m)}\overline{\boldsymbol{T}}_k^{(m)}\} + \overline{r}_k^{s(m)} - t^{(m)} r_k^o$

     else

     update $\breve{\xi}_k^{\mathrm{S},(m)} = \mathrm{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S},(m)}\overline{\boldsymbol{T}}_k^{(m)}\} + \overline{r}_k^{s(m)} - r_k^o$

     end if
3. **until** required accuracy is reached or $m \geq m_{\max}$

The per cell transmit powers P_c are obtained as follows

$$P_c = \boldsymbol{c}_c^T \boldsymbol{p}.$$

The defined optimization problem aims to minimize the maximum transmit power among BSs, namely $P^s = \max_c P_c$. Therefore, we set the new power constraints for MMR optimization as $P_{\max,c} = P^s$, identical $\forall\ c$. Doing so, $\mu_c(P_{\max,c} = P^s)$ will make sure that all transmit powers $\boldsymbol{c}_c^T \boldsymbol{p}$ do not exceed the minimized maximum P^s, i.e., $P_c = \boldsymbol{c}_c^T \boldsymbol{p} \leq P^s$. Of course, at convergence, we have $P_c = P^s\ \forall\ c$.

8.5 Results

In this section, we numerically evaluate the performance of the proposed algorithms. We use the channel model from Section 6.3, as in Chapter 7, we consider for the multipath channel model,

$$\boldsymbol{C}_t = \sum_{n=1}^{N_p} \frac{\alpha_i}{\boldsymbol{v}_i^H \boldsymbol{v}_i} \boldsymbol{v}_i \boldsymbol{v}_i^H \tag{8.56}$$

with $\text{tr}\{\boldsymbol{C}_t\} = \sum_{n=1}^{N_p} \alpha_i = M_c$, $\alpha_i = c^{i-1}\alpha_1$ and the $\boldsymbol{v}_i$ are i.i.d. vectors of M_c i.i.d. elements $\mathcal{CN}(0,1)$. We take $N_p = M_c/K$.

Figure 8.1 shows the difference between the approximates received signal level and stream level ESIP, by considering $N_k \neq d_k$. We can see that for $N_k = d_k$, both R- and S-ESIP are equivalent, whereas for $N_k \geq d_k$, S-ESIPrate outperforms R-ESIPrate, and the more N_k/d_k increases, the more we have gap, especially at intermediate values for SNR.

In Figure 8.2, we evaluate the average rate w.r.t. SNR, in broadcast channel (only one communicating cell), for varying levels of channel estimation error $\sigma^2_{\widetilde{\boldsymbol{H}}}$. It is clear that the gap between SESIP and RESIP increases when $\rho_D = 1/\sigma^2_{\widetilde{\boldsymbol{H}}}$ decreases. In the following, we take $N_k = d_k$ and refer to the (S/R)-ESIP approches by ESIP as they are both equivalent.

Figure 8.3 represents the average attained rate using the proposed algorithms for different configurations of the system (single and multi-cell). We can see that ESIPrate outperforms WEMSE and suffers little loss compared to perfect CSIT, and that the UB ESIPrate provides a tight upper bound. Note also that for fixed ρ_D as considered here, Naive saturates at high SNR, whereas WEMSE appears to suffer Degree-of-Freedom (DoF) (slope) loss.

In Figure 8.4, we consider varying levels of channel estimation error $\sigma^2_{\widetilde{\boldsymbol{H}}}$. It is clear that when $\rho_D = 1/\sigma^2_{\widetilde{\boldsymbol{H}}}$ is proportional to SNR, all algorithms (only) suffer from varying SNR offset, but ESIPrate still outperforms WEMSE as the signal link channel error covariance is accounted for in the interference power instead of in the signal power.

Figure 8.5 which illustrates the convergence of the user rates w.r.t. the number of iterations for one channel realization, with SNR = 20dB and $\rho_D = 10$. We observe that

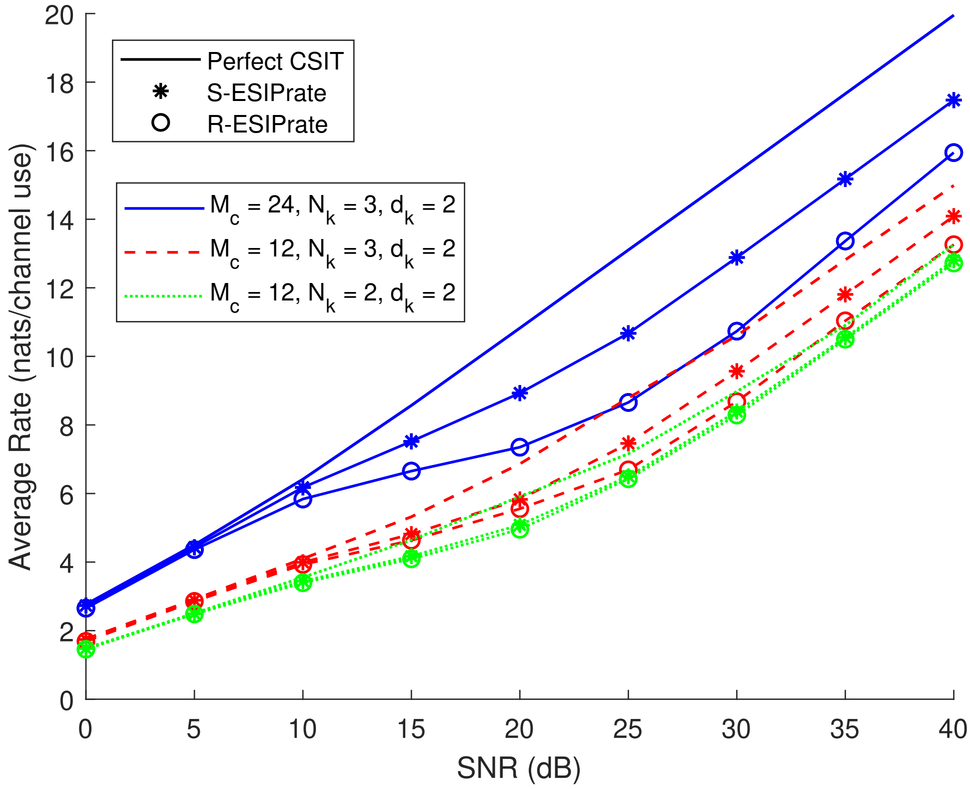

Figure 8.1: Average Rate with Partial CSIT w.r.t. SNR: R-ESIP vs. S-ESIP, $C = 2, K_c = 3$, and $\rho_D = 10$.

the rates obtained using ESIP (either with (8.2) or (8.5)) and $\ln(\det(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k))$ are balanced. Both approaches converge after about 5 iterations of the outer loop, while the inner loop converges after 2-3 iterations. Of course, due to the CSIT imperfections, the actual rates exhibit some randomness.

In Figure 8.6, we plot the achieved average rate and total transmit power using Table 8.2, for BC. We set identical user targets $r_k^o = 4$, $\forall k$ and $P_{\max} = 10^{\mathrm{SNR}\sigma_n^2/10}$. We can see that, when the rate targets are feasible, (i.e., $r_k(P_{\max})/r_k^o \geq 1$ with MMR optimization), the user rates using perfect CSIT and ESIPrate UB meet the targets with equality and the total transmit power is minimized accordingly. Also, the same total minimized power P is achieved $\forall$ $P_{\max}$. In this figure, this case corresponds to SNR$>$ 10dB. When the targets are infeasible, Table 8.2 acts as MMR algorithm since the power minimization is not possible.

Figure 8.7 illustrates the minimized total transmit power for both perfect and imperfect CSIT vs. number of transmit antennas. We observe that we need more power in partial CSIT case to reach the user targets, due to the uncertainties. However, when we increase the number of Tx antennas for a fixed number of users, the transmit power gap between perfect and partial CSIT gets smaller.

Figure 8.8 plots the achieved rate using per cell power minimization for IBC scenario,

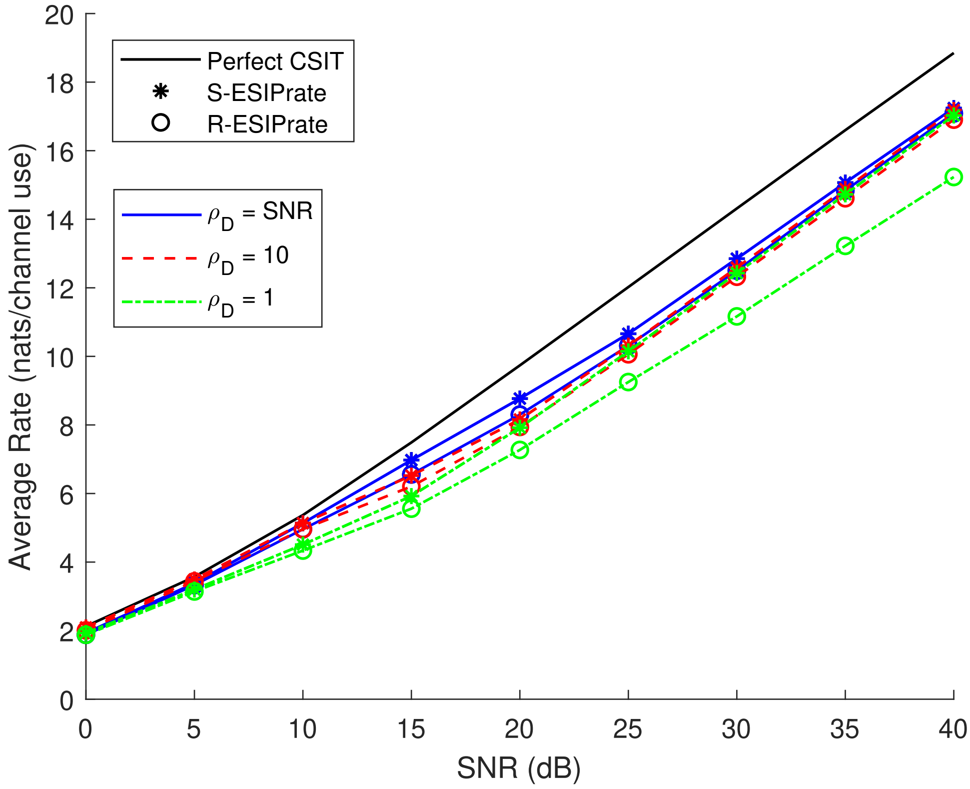

Figure 8.2: Average Rate with Partial CSIT w.r.t. SNR: R-ESIP vs. S-ESIP, $C = 1, K = 3, N_k = 3$, and $d_k = 2$.

and Figure 8.9 illustrates the corresponding transmitted power per cell. We can that, when the rate targets are feasible, the transmit power is minimized within each cell with equality while fulfilling the rate targets.

8.6 Closing Remarks

In this chapter, we considered the ESIP Erate approximation, for which we introduced an original minorizer, judiciously chosen to be amenable to the Perron-Frobenius theory. We furthermore introduced original explicit power constraint Lagrange multiplier solutions, which can handle the case in which some cell power constraints are met with inequality, as can happen in a multi-cell scenario. The simulation results exhibit the different SNR behavior of the Erate lower bound vs. actual Erate, showed that the upper bound is a quite tight approximation, and that the ESIP partial CSIT approach with LMMSE channel estimation leads to very limited performance loss compared to perfect CSIT. In the multi-cell case, the proposed algorithms can handle scenarios in which the CSIT quality could be very different between intracell and intercell links.

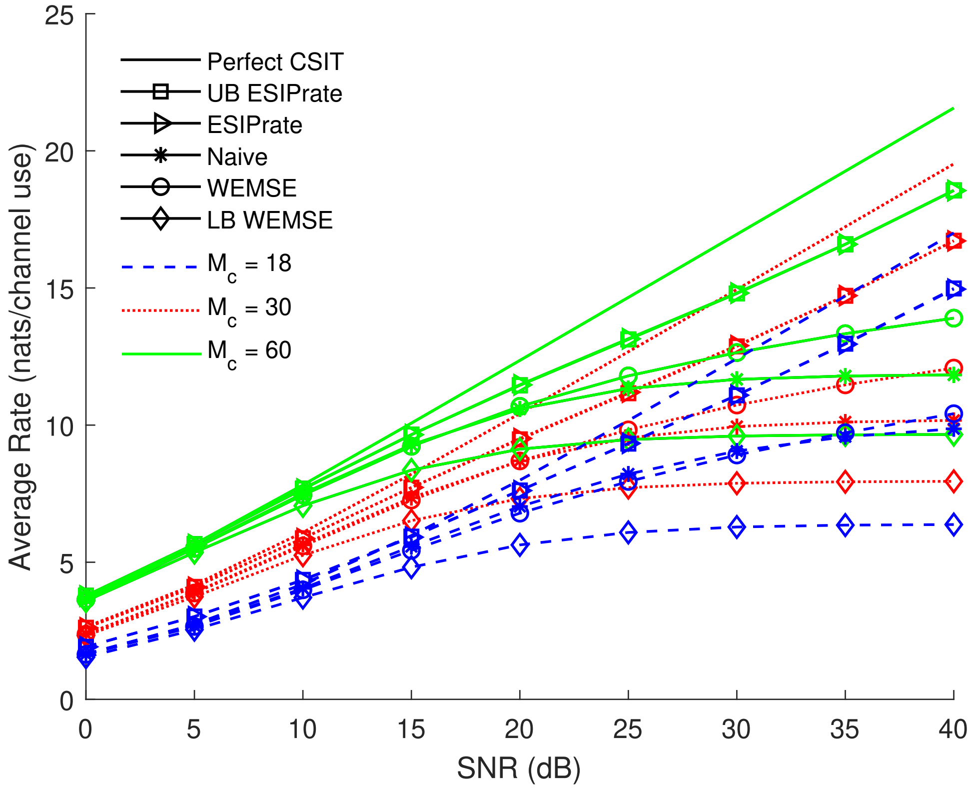

Figure 8.3: Average Rate with Partial CSIT vs. SNR, $\rho_D = 10$.

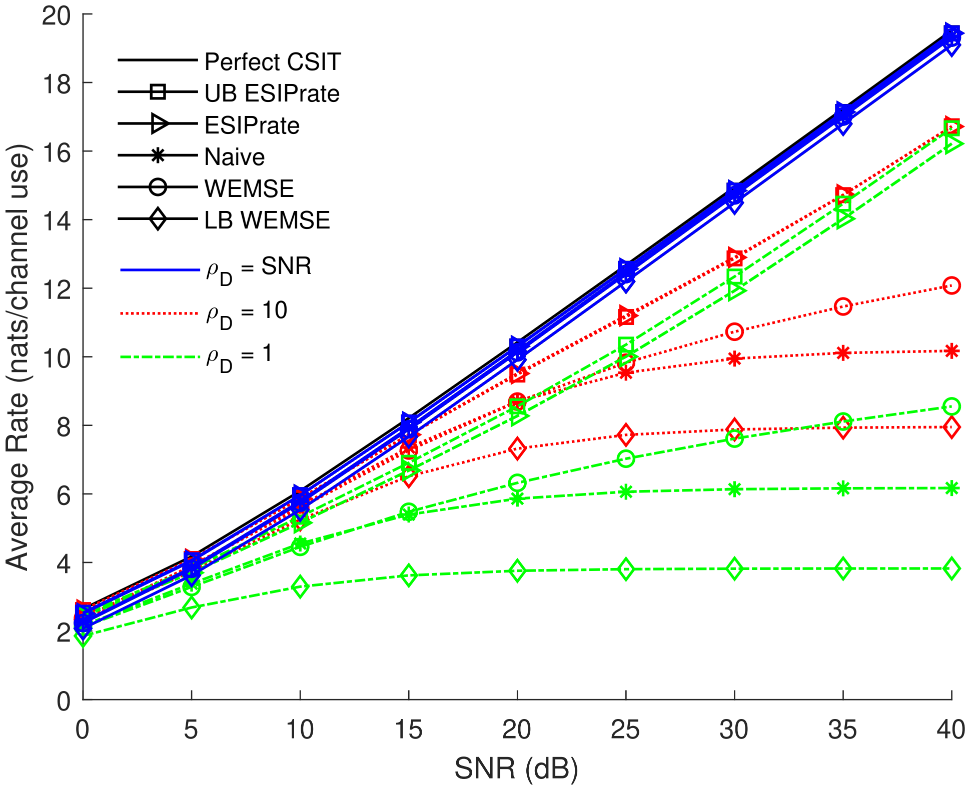

Figure 8.4: Average Rate with Partial CSIT vs. SNR for different instances of ρ_D, $C = 1, K = 3, M_c = 12, N_k = d_k = 2$.

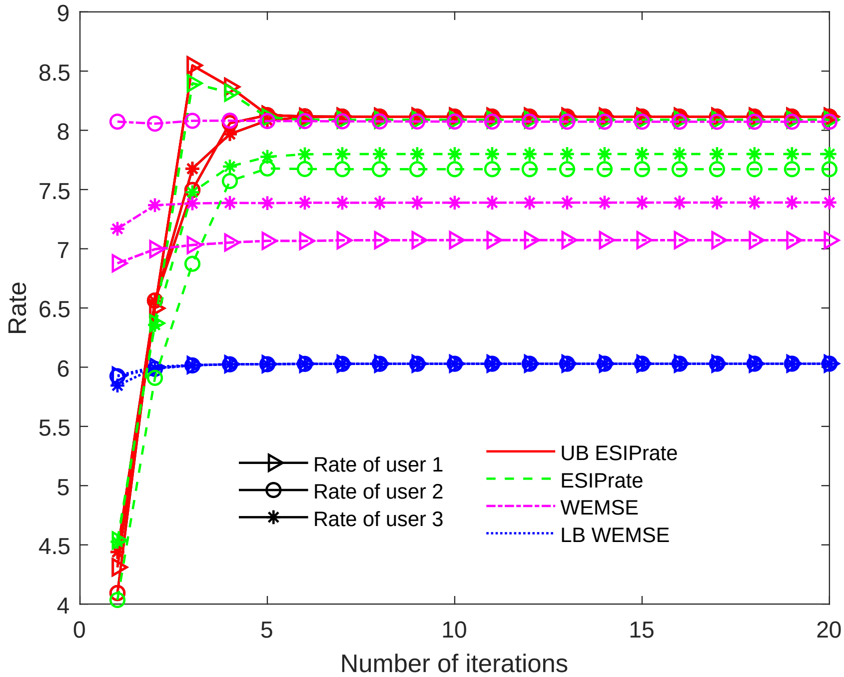

Figure 8.5: User Rates with Partial CSIT vs. number of outer iterations, $C = 1, K = 3, M_c = 12, N_k = d_k = 2, \rho_D = 10$.

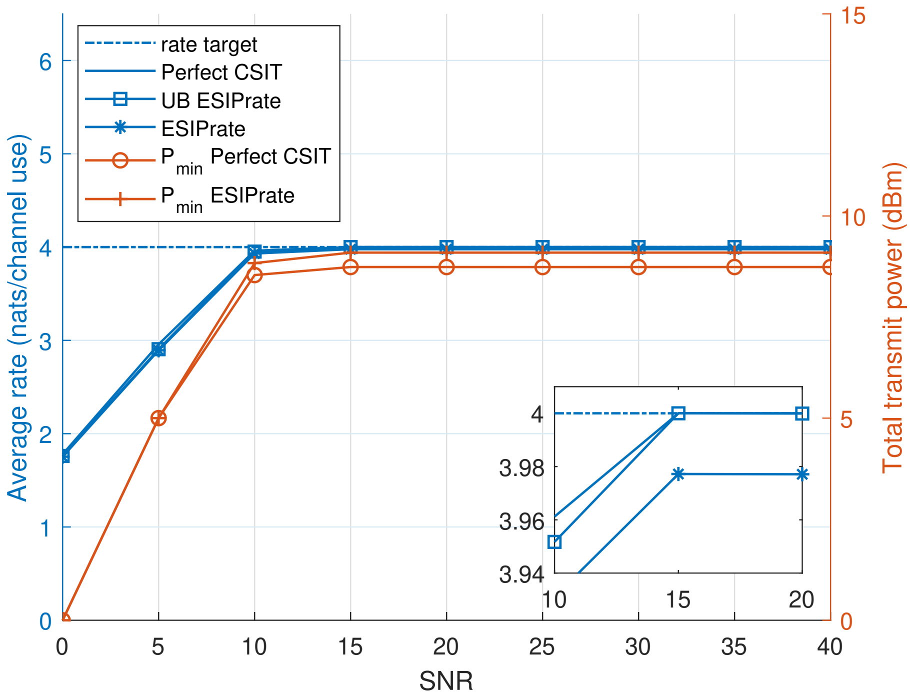

Figure 8.6: Total Power Minimization for BC via ESIP, $C = 1, K = 3, M_c = 12, N_k = d_k = 2, \rho_D = 10$.

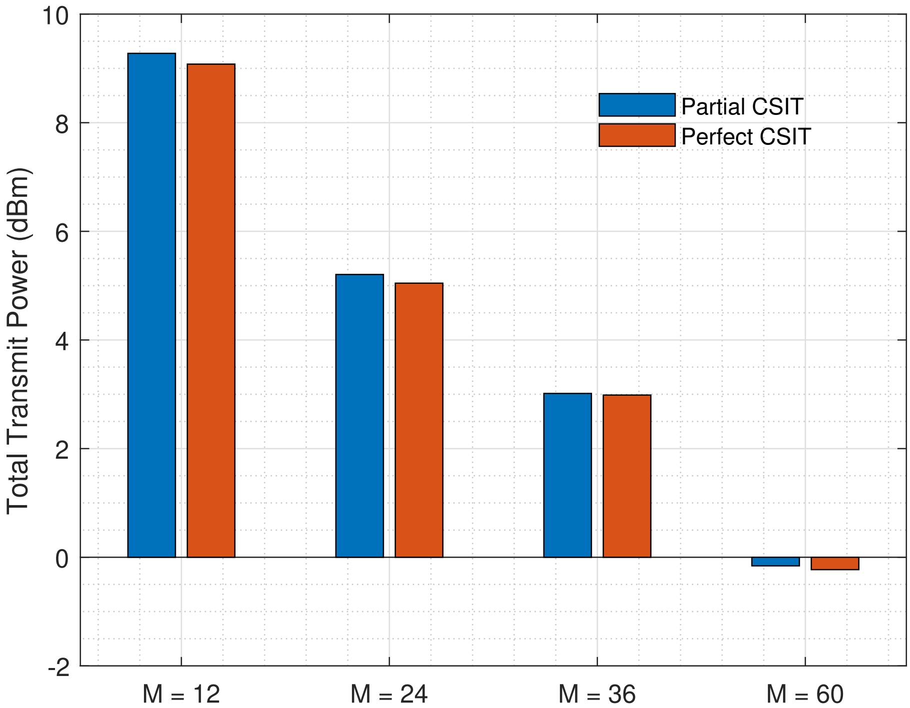

Figure 8.7: Total Power Minimization for BC via ESIP, $C = 1, K = 3, N_k = d_k = 2, r_k^o = 4, \rho_D = 10$.

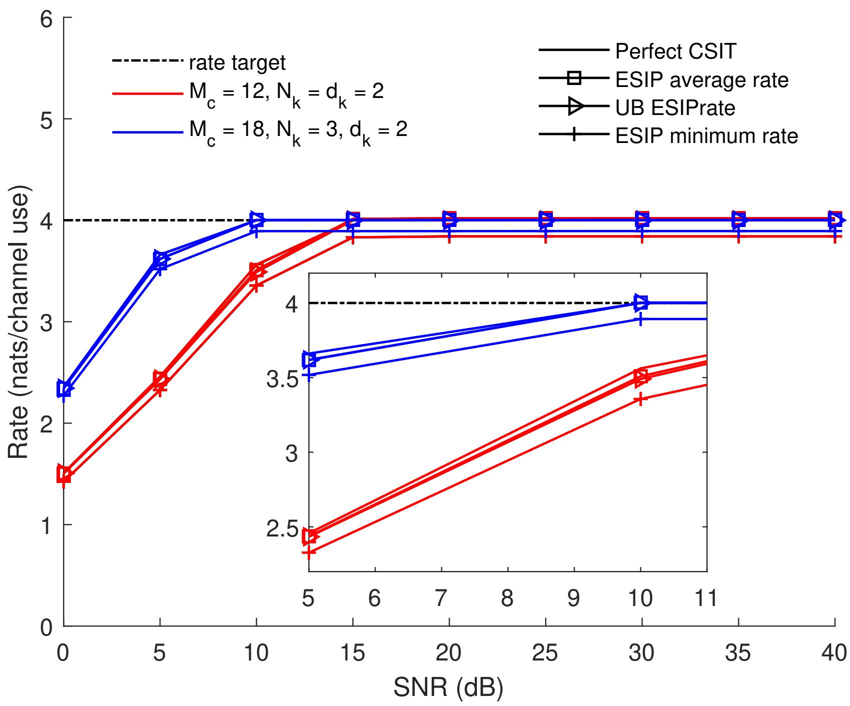

Figure 8.8: Achieved rate vs. SNR using per cell Power Minimization via ESIP: $C = 2, K_c = 3, \rho_D = 10$, and $r_k^o = 4 \forall k$.

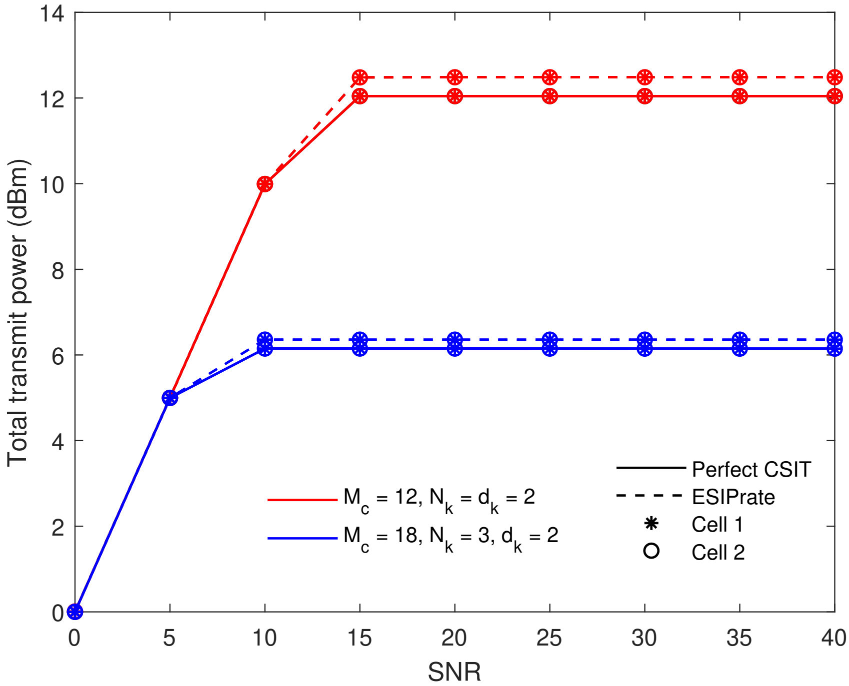

Figure 8.9: Per cell transmit power vs. SNR using per cell Power Minimization via ESIP: $C = 2, K_c = 3, \rho_D = 10$, and $r_k^o = 4 \forall k$.

Part IV

Conclusions, Outlook and Appendices

Chapter 9

Conclusions and Future Directions

In this thesis, we have investigated the weighted max-min fairness w.r.t. per user rates, namely, rate balancing problem subject to total transmit power constraint for multi-cell multi-user MIMO systems.

In the first part, we have addressed the problem with perfect knowledge of the channel at the transmitter, for broadcast and interfering broadcast channels. At a first stage, we have exploited the MSE - rate relation to enable uplink/downlink duality. In fact, the maximization of the minimum (weighted) rate has been reformulated as a minimization of the maximum matrix-weighted MSE. Then, MSE duality has been used, consisting on achieving the same per user weighted MSE at both uplink and downlink links. This has conducted to the solution in an alternating manner between both links. At a second stage, we have considered the Lagrangian duality to solve the rate balancing problem. Actually, the min max weighted matrix MSE balancing operation has been reformulated as constraints in the dual problem. The corresponding Lagrangian has been therefore introduced, leading to alternating optimization to reach the saddle point. Also, various aspects of Perron Frobenius theory have been exploited in the process. Moreover, we have considered per cell power constraints for which the Lagrangian multipliers could be formulated as a single weighted power constraint. The weighting could be then optimized leading to the satisfaction with equality of all power constraints. Subgradient projection method has been used to this end. The main conclusions of this part are the following

- the max-min rate optimization problem can be transformed into a min-max weighted MSE optimization problem which itself was shown to be related to a weighted sum MSE minimization via Langrangian duality;
- the optimization can be held in the uplink channel (via UL/DL duality), or directly in the downlink channel considering either diagonal or non-diagonal matrix weighted MSE;
- the min max matrix-weighted MSE optimization provides appreciable performance improvements as compared to optimizing the conventional unweighted per user MSE balancing problem;

- balancing the weighted user rate via matrix-weighted MSE approach distributed the rate equally between the users with equal when the weights are equal, when not, the rate differs from one user to another accordingly;
- reformulating the multiple power constraints as a single weighted constraint ensures the per cell power constraint with equality, unlike the total sum power constraint which verifies the total power over cells.

In the second part, we have considered only partial CSIT to solve the rate balancing problem, as the knowledge of the channel at the transmitter side is never perfect. In particular, we have focused on ergodic user rate balancing problem, which corresponds to maximizing the minimum (weighted) per user expected rate in the network. The partial CSIT that has been taken into account combines both channel estimates and channel (error) covariance information. Firstly, we have introduced the extension of the matrix-weighted MSE balancing formulation to partial CSIT, namely, maximizing an expected rate lower bound in terms of expected MSE. Secondly, we have provided a second algorithm by exploiting a better approximation of the expected rate as the expected signal and interference power rate, based on an original minorizer for every individual rate term. We have studied the latter within two approximations: *i)* Received signal level ESIP and *ii)* Stream level ESIP. Furthermore, we have introduced original explicit power constraint Lagrange multiplier solutions, which can handle the case in which some cell power constraints are met with inequality, as can happen in a multi-cell scenario. The main conclusions of this part are as follows

- optimizing the WEMSE provides a lower bound of the achieved average rate, both suffering from Degree-of-Freedom (slope) loss for fixed level of channel estimation error, as compared to perfect CSIT;
- ESIPrate is proved to outperform WEMSE and suffers little loss compared to perfect CSIT, moreover, ESIPrate based optimization provides a tight upper bound;
- the R/S-ESIP comparison via simulations motivates the use of SESIP rate approach when the number of streams is lower than the number of receive antennas, confirming the intuition: averaging a smaller dimensional matrix results in a tighter bound;
- for a number of streams equaling the number of receive antennas, as also in high or low SNR regimes, the difference between the two approximations (R/S-ESIP) is negligible.

On the other hand, we have considered a more practical scenario: the case when the per user weights/priorities represents individual user targets/requirements. In this context, maximizing the jointly achievable rate margin under total power constraint is closely related to minimizing the total transmission power while satisfying a set of user rate constraints. In particular, if we set the total power as the minimized power in the max-min optimization problem, both problems becomes equivalent. The main cgonclusions to this regard are in the following

- maximizing the minimum user rate to user rate target ratio provides a single performance measure that reflects the quality of the multi-user channels, such a measure is required by the upper layers (e.g., medium access control), to decide whether spatial multiplexing is meaningful or not;

- the transmission power minimization problem w.r.t. per user rate targets is a variation of user rate balancing problem subject to total power constraint;

- when the individual rate targets are feasible, i.e., if they can be jointly supported, the total transmission power can be minimized via rate balancing approaches and the achieved per user rates jointly meet their respective targets.

The extension of minimizing the *total* transmit power from BC to IBC is direct via rate balancing under total transmit power constraint instead of per cell power constraint. Doing so, the overall transmit power of the system is minimized accordingly when the rate targets are feasible; the distribution of this transmit power between cells is however not balanced. Therefore, *per cell* transmit power balancing problem has been considered, wherein, the optimization is achieved through the single objective optimization: minimizing the maximum transmit power among cells. The study has been handled with perfect and partial CSIT, the main obtained conclusions are the following

- minimizing the maximum transmit power among cells, subject to individual rate targets, minimizes the per cell transmit powers equally, when targets are feasible, otherwise MMR operates maintaining the per cell power constraints;

- when user rate targets are supported, the minimized per cell transmit powers with perfect CSIT are below the ones minimized considering partial CSIT, while both of them reach the target rates.

Despite the fact that the methods considered in this thesis present contributions in terms of user rates and power allocation, especially for MU MIMO scenarios, there still exist several challenges which need to be investigated in the future. In fact, we have considered the optimization problem in the physical layer, assuming full buffers (each user is continuously receiving data). It would be of great interest to consider the following directions.

With transmit power minimization via rate balancing, the algorithm optimizes the max-min (weighted) user rate problem if we are in some SNR regime that does not support the individual rate targets. In this case, *i)* the user rate to target rate ratios are equally balanced between users, *ii)* none of the users meet their targets, and *iii)* the obtained user rates to user rate target reduction is *fair* (equal) for all users.

- If this reduction can still ensure some QoS, (e.g., $r_k/r_k^o = 98\%, \quad \forall k$), we can still consider the allocation. Otherwise, we must relax the initial conditions, by reducing the number of users. Actually, MMR optimization balances the user rate ratios by allocating more power to users in bad conditions and less power to users with good channel quality, under a sum power equaling the total power constraint (or per cell power constraint in IBC). Given that, we can consider a scheduling policy which drops the user with the highest individual power, namely $k^* = \text{argmax}_k \, p_k$, with k^* denoting the user to be dropped, until feasibility of the targets for the remaining users.

- Balancing weighted user rates between the selected users ensures instantaneous fairness between the users that are scheduled together, not all active users. It would be of interest to study the average (in time) rate balancing problem, to ensure long term fairness between all active users. Indeed, there are several works of long term user fairness based on Proportional Fair (PF) scheduling [90–95]. Actually, PF represents a tradeoff between sum SE and max-min (unweighted) fairness. PF schemes exploit the weights in WSR optimization while scheduling. In particular, it considers the users' current average throughput in a period of time. Then, the instantaneous channel quality to average throughput ratio determines whether the user can be scheduled or not, taking account of long term fairness. In this regard, additional effort to exploit the user weights in MMR is of high interest, in order to have a more realistic view to the system.

- There are applications that require a quasi-constant Quality-of-Experience (QoE) such as multimedia transmissions. In such cases, buffering may be needed and it is important to maintain a minimum throughput by considering max-min rate and power minimization optimizations. However, an end-to-end study considering a well defined traffic model is needed to provide more realistic insights.

- Furthermore, one could extend our works for systems with large dimensions. Doing so, large system analysis can be used to provide approximations for Massive MIMO (ma-MIMO) to simplify beamforming. Also, a study of the gap between perfect and imperfect CSIT at high SNR regime can be handled for ESIPrate based optimizations. In fact, simulation results have shown that ESIPrate curves have a similar behavior to perfect CSIT curves, with a parallel gap at high SNR. It would be interesting to characterize analytically this gap in order to better serve user in partial CSIT.

- In case of unfeasibility of user rate targets, user selection is not the only way to relax the problem. One can also use antenna selection under ma-MIMO assumptions. In fact, the individual targets may be not supported when BSs are equipped with limited antennas. However, the QoS requirements may become feasible as the number of antennas increases. Thus, the total transmit power minimization problem can be considered as jointly solving the antenna selection and beamforming design. Paper [96] has solved the problem for fixed beamforming directions when the number of antennas is large. Therein, not only transmitter power, but also hardware-consumed power is considered which increases with the number of antennas. Thus, energy efficiency is improved by turning on and off antennas. Asymptotic results based analysis is handled to provide the number of antennas that are needed to ensure QoS constraints.

Bibliography

[1] A. Goldsmith, *Wireless Communications*.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, p. 311–335, 1998.

[3] R. K. Mueller and G. J. Foschini, "The capacity of linear channels with additive gaussian noise," *The Bell System Technical Journal*, vol. 49, no. 1, pp. 81–94, 1970.

[4] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, p. 585–595, 1999.

[5] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the pareto boundary for the miso interference channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5292–5296, 2008.

[6] L. Zheng, D. W. H. Cai, and C. W. Tan, "Max-min Fairness Rate Control in Wireless Networks: Optimality and Algorithms by Perron-Frobenius Theory," *IEEE Transactions on Mobile Computing*, Jan. 2018.

[7] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up mimo: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[8] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Vehic. Tech.*, Jun 2004.

[9] M. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions*, vol. 29, no. 3, p. 439–441, May 1983.

[10] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE TRANSACTIONS ON INFORMATION THEORY*, vol. 49, no. 7, July 2003.

[11] N. Jindal, W. Rhee, S. Viswanath, A. Jafar, and A. Goldsmith, "Sum Power Iterative Water-Filling for Multi-antenna Gaussian Broadcast Channels," *IEEE Trans. Inform Theory*, vol. 51, no. 4, pp. 1570–1580, Apr 2005.

[12] H. Wang, X. Xu, M. Zhao, W. Wu, and Y. Yao, "Robust transmission for multiuser mimo downlink systems with imperfect csit," *in Wireless Communications and Networking Conference, WCNC, IEEE*, p. 340–344, Apr 2008.

[13] L. Zhang, Y.Wu, S. Zhou, and J.Wang, "Joint linear transmitter and receiver design for the downlink of multiuser mimo systems," *Communications Letters, IEEE*, vol. 9, no. 11, Nov. 2005.

[14] M. Amara, Y. Yuan-Wu, and D. Slock, "Receiver and transmitter iterative optimization using maximum sum-rate criterion for multi-user mimo systems," *in ISCCSP 2010. IEEE International Symposium on Communications, Control and Signal Processing*, Mar. 2010.

[15] ——, "Optimal mu-mimo precoder with miso decomposition approach," *in SPAWC 2010. IEEE International Workshop on Signal Processing Advances for Wireless Communications*, June 2010.

[16] M. Ivrlac, R. Choi, R. Murch, and J. Nossek, "Effective use of long-term transmit channel state information in multi-user mimo communication systems," *in IEEE 58th Vehicular Technology Conference, 2003. VTC 2003-Fall*, vol. 1, p. 373–377, Oct. 2003.

[17] M. Lee and S. K. Oh, "A per-user successive mmse precoding technique in multiuser mimo systems," *in Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, p. 2374–2378, Apr. 2007.

[18] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *Signal Processing, IEEE Transactions on*, vol. 52, no. 2, p. 461–471, Apr. 2004.

[19] M. Amara, Y. Yuan-Wu, and D. Slock, "Optimized linear receivers and power allocation for two multi-user mimo downlink schemes with linear precoding," *in ISCCSP 2010. IEEE International Symposium on Communications, Control and Signal Processing*, Mar. 2010.

[20] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate Maximization in Multi-Antenna Broadcast Channels with Linear Preprocessing," *IEEE Trans. Wireless Comm.*, Sep. 2006.

[21] L. Sun and M. Mckay, "Eigen-based transceivers for the mimo broadcast channel with semi-orthogonal user selection," *Signal Processing, IEEE Transactions on*, 2010.

[22] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear Transceiver Design for a MIMO Interfering Broadcast Channel Achieving Max-min Fairness," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov 2011.

[23] L. P. Qian, Y. J. Zhang, and J. Huang, "Mapel: Achieving global optimality for a non-convex wireless power control problem," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1553–1563, 2009.

[24] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell miso systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2508–2523, 2012.

[25] P. C. Weeraddana, M. Codreanu, M. Latva-aho, A. Ephremides, and C. Fischione, 2012.

[26] 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 *Weighted Sum-Rate Maximization using Weighted MMSE for MIMO-BC Beamforming Design*, Dec 2008.

[27] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," in *IEEE Trans. Signal Processing*, Sept. 2011.

[28] P. Komulainen, A. Tölli, and M. Juntti, "Effective csi signaling and decentralized beam coordination in tdd multi-cell mimo systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2204–2218, 2013.

[29] D. H. N. Nguyen and T. Le-Ngoc, "Sum-rate maximization in the multicell mimo multiple-access channel with interference coordination," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 36–48, 2014.

[30] L. Tran, M. F. Hanif, A. Tolli, and M. Juntti, "Fast converging algorithm for weighted sum rate maximization in multicell miso downlink," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 872–875, 2012.

[31] D. P. Palomar and S. Verdu, "Gradient of mutual information in linear vector gaussian channels," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 141–154, 2006.

[32] A. J. Tenenbaum and R. S. Adve, "Minimizing sum-mse implies identical downlink and dual uplink power allocations," *IEEE Transactions on Communications*, vol. 59, no. 3, pp. 686–688, 2011.

[33] Zhi-Quan Luo, T. N. Davidson, G. B. Giannakis, and Kon Max Wong, "Transceiver optimization for block-based multiple access through isi channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1037–1052, 2004.

[34] M. Codreanu, A. Tolli, M. Juntti, and M. Latva-aho, "Joint design of tx-rx beamformers in mimo downlink channel," in *2007 IEEE International Conference on Communications*, 2007, pp. 4997–5002.

[35] S. Shi, M. Schubert, and H. Boche, "Rate optimization for multiuser mimo systems with linear processing," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4020–4030, 2008.

[36] J. Kaleva, A. Tölli, and M. Juntti, "Successive convex approximation for simultaneous linear tx/rx design in mimo bc," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, 2015, pp. 1227–1231.

[37] 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.

[38] Y. Liu, Y. Dai, and Z. Luo, "Coordinated beamforming for miso interference channel: Complexity analysis and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1142–1157, 2011.

[39] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear Precoding via Conic Optimization for Fixed MIMO Receivers," in *IEEE Transactions on Signal Processing*, vol. 54, no. 1, Jan 2006, pp. 161–176.

[40] M. Bengtsson and B. Ottersten, "Optimal and suboptimal Transmit Beamforming," *Handbook of Antennas in Wireless Communications*, 2001.

[41] D. Cai, T. Quek, and C. W. Tan, "A Unified Analysis of Max-min Weighted SINR for MIMO Downlink System," in *IEEE Transactions on Signal Processing*, vol. 59, no. 8, Aug 2011, p. 3850 –3862.

[42] G. Montalbano and D. Slock, "Matched Filter Bound Optimization For Multiuser Downlink Transmit Beamforming," in *Universal Personal Communications, 1998. ICUPC '98. IEEE 1998 International Conference*, vol. 1, Oct 1998, p. 677 –681.

[43] F. Negro, M. Cardone, I. Ghauri, and D. Slock, "SINR Balancing and Beamforming for the MISO Interference Channel," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep 2011.

[44] W. Yu and T. Lan, "Transmitter Optimization for the Multi-antenna Downlink with Per-antenna Power Constraints," in *IEEE Trans. Signal Processing*, June 2007.

[45] L. Zhang, R. Zhang, Y. C. Liang, Y. Xin, and H. V. Poor, "On Gaussian MIMO BC-MAC Duality with Multiple Transmit Covariance Constraints," *IEEE Trans. Inform. Theory*, Apr. 2012.

[46] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR Balancing Technique for Downlink Beamforming in Cognitive Radio Networks," *IEEE Signal Process. Lett.*, Feb. 2010.

[47] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and sum-MSE Minimization," in *IEEE Trans. Signal Process*, vol. 55, no. 11, Nov 2007.

[48] ——, "Capacity Balancing for Multiuser MIMO Systems," in *Proc. IEEE ICASSP*, Apr 2007.

[49] R. Hunger, M. Joham, and W. Utschick, "On the MSE-Duality of the Broadcast Channel and the Multiple Access Channel," in *IEEE Trans. Signal Processing*, vol. 57, no. 2, Feb 2009, p. 698–713.

[50] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 1, pp. 57–62, 1992.

[51] C. W. Tan, M. Chiang, and R. Srikant, "Maximizing sum rate and minimizing mse on multiuser downlink: Optimality, fast algorithms and equivalence via max-min sinr," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 6127–6143, 2011.

[52] C. Fischione and U. Jönsson, "Fast-lipschitz optimization with wireless sensor networks applications," in *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2011, pp. 378–389.

[53] C. Fischione and M. Jakobsson, "Optimality of radio power control via fast-lipschitz optimization," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2589–2601, 2016.

[54] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and user association optimization for massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6384–6399, 2016.

[55] H. V. Cheng, E. Björnson, and E. G. Larsson, "Optimal pilot and payload power control in single-cell massive mimo systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2363–2378, 2017.

[56] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint pilot design and uplink power allocation in multi-cell massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2000–2015, 2018.

[57] H. Yang and T. L. Marzetta, "Massive mimo with max-min power control in line-of-sight propagation environment," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4685–4693, 2017.

[58] Z. Xiang, M. Tao, and X. Wang, "Massive mimo multicasting in noncooperative cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1180–1193, 2014.

[59] S. Zarei, J. Aulin, W. Gerstacker, and R. Schober, "Max–min multicell-aware precoding and power allocation for downlink massive mimo systems," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1433–1437, 2017.

[60] N. Akbar, E. Björnson, N. Yang, and E. G. Larsson, "Max-min power control in downlink massive mimo with distributed antenna arrays," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 740–751, 2021.

[61] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink–downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.

[62] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1748–1759, 2010.

[63] D. Gesbert, S. G. Kiani, A. Gjendemsjo, and G. E. Oien, "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2393–2409, 2007.

[64] M. Ku and D. Kim, "Tx-rx beamforming with multiuser mimo channels in multiple-cell systems," in *2008 10th International Conference on Advanced Communication Technology*, vol. 3, 2008, pp. 1767–1771.

[65] Tianmin Ren and R. J. La, "Downlink beamforming algorithms with inter-cell interference in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2814–2823, 2006.

[66] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534–5562, 2008.

[67] Te Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, 1981.

[68] X. Shang, B. Chen, and H. V. Poor, "Multiuser miso interference channels with single-user detection: Optimality of beamforming and the achievable rate region," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4255–4273, 2011.

[69] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays.*

[70] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "On the MIMO Interference Channel," in *Proc. IEEE Inf. Theo. and Applic. Workshop (ITA)*, San Diego, CA, USA, 20110.

[71] F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel," in *Proc. IEEE Inf. Theo. and Applic. Workshop (ITA)*, San Diego, CA, USA, 2011.

[72] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, in *Transmit wiener filter for the downlink of TDD DS-CDMA systems*, vol. 01, Sept 2002.

[73] Seung-Jun Kim and G. B. Giannakis, "Optimal resource allocation for mimo ad hoc cognitive radio networks," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, 2008, pp. 39–45.

[74] H. Al-Shatri and T. Weber, "Achieving the maximum sum rate using DC programming in cellular networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1331–1341, 2012.

[75] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: MMSE Balancing," *IEEE Trans. Signal Processing*, vol. 56, no. 8, pp. 606–619, Aug 2008.

[76] I. Ghamnia, D. Slock, and Y. Yuan-Wu, "Rate balancing for multiuser multicell downlink MIMO Systems," in *27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain*, 2019.

[77] C. Meyer, *Matrix Analysis and Applied Linear Algebra.* SIAM, 2000.

[78] M. Joham, W. Utschick, and J. Nossek, "Linear Transmit Processing in MIMO Communications Systems," *IEEE Trans. Sig. Process.*, Aug. 2005.

[79] D. Cai, T. Quek, C. W. Tan, and S. H. Low, "Max-Min SINR Coordinated Multipoint Downlink Transmission—Duality and Algorithms," in *IEEE Transactions on Signal Processing*, vol. 60, no. 10, Oct 2012.

[80] D. W. H. Cai, T. Q. S. Quek, C. W. Tan, and S. H. Low, "Max-Min SINR Coordinated Multipoint Downlink Transmission—Duality and Algorithms," Oct. 2012.

[81] I. Ghamnia, D. Slock, and Y. Yuan-Wu, "Rate Balancing for Multiuser MIMO Systems," in *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cannes, France, Jul 2019.

[82] ——, "MIMO user rate balancing in multicell networks with per cell power constraints," in *IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium*, 2020.

[83] T. E. Bogale, B. K. Chalise, and L. Vandendorpe, "Robust transceiver optimization for downlink multiuser mimo systems," *IEEE Transactions on Signal Processing*, Jan. 2011.

[84] S. Malla and G. Abreu, "Transmission strategies under imperfect instantaneous csit," *IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA*, 2015.

[85] C. Li, C. He, L. Jiang, and F. Liu, "Robust beamforming design for max–min sinr in mimo interference channels," *IEEE Communications Letters*, April 2016.

[86] M. Joham, A. Gründinger, J. R. F. A. Pastore, and W. Utschick, "Rate balancing in the vector bc with erroneous csi at the receivers," *47th Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD*, 2013.

[87] D. Neumann, A. Gründinger, M. Joham, and W. Utschick, "Rate-balancing in massive mimo using statistical precoding," *IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm*, 2015.

[88] F. Negro, I. Ghauri, and D. T. Slock, "Sum Rate Maximization in the Noisy MIMO Interfering Broadcast Channel with Partial CSIT via the Expected Weighted MSE," in *Proc. Int'l Symp. on Wireless Communication Systems (ISWCS)*, Paris, France, 2012.

[89] K. Gopala and D. Slock, "A refined analysis of the gap between expected rate for partial csit and the massive mimo rate limit," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 3864–3868.

[90] H. J. Bang, T. Ekman, and D. Gesbert, "Channel predictive proportional fair scheduling," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 482–487, 2008.

[91] E. Björnson, J. Hoydis, and L. Sanguinetti, 2017.

[92] J. Ma, A. Aijaz, and M. Beach, "Recent results on proportional fair scheduling for mmwave-based industrial wireless networks," in *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, 2020, pp. 1–5.

[93] M. K. Müller, S. Schwarz, and M. Rupp, "Qos investigation of proportional fair scheduling in lte networks," in *2013 IFIP Wireless Days (WD)*, 2013, pp. 1–4.

[94] L. Liu, Y.-H. Nam, and J. Zhang, "Proportional fair scheduling for multi-cell multi-user mimo systems," in *2010 44th Annual Conference on Information Sciences and Systems (CISS)*, 2010, pp. 1–6.

[95] V. Valls and D. J. Leith, "Proportional fair mu-mimo in 802.11 wlans," *IEEE Wireless Communications Letters*, vol. 3, no. 2, pp. 221–224, 2014.

[96] K. Senel, E. Björnson, and E. G. Larsson, "Joint transmit and circuit power minimization in massive mimo with downlink sinr constraints: When to turn on massive mimo?" *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1834–1846, 2019.

[97] H. Boche and M. Schubert, "A General Duality Theory for Uplink and Downlink Beamforming," in *Vehicular Technology Conf*, Sept 2002.

[98] J. Zander and S. Kim, "Radio Resource Management for Wireless Networks," *Boston, MA: Artech House*, 2001.

Appendix A: Rate Balancing via MSE UL/DL Duality

A.1 Proof of Lemma 3.4

First, by checking the first order optimality condition of (3.4) with respect to $\boldsymbol{\mathcal{F}}_k$, we get

$$\begin{aligned} \boldsymbol{W}_k^H(\boldsymbol{R}_k\boldsymbol{\mathcal{F}}_k^* - \boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k) = 0 \\ \Rightarrow \boldsymbol{\mathcal{F}}_k^* = \boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k \end{aligned} \tag{A.1}$$

where $\boldsymbol{R}_k = \sigma_n^2\boldsymbol{I} + \sum_{j=1}^{K}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_j\boldsymbol{\mathcal{G}}_j^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}$ and $\boldsymbol{\mathcal{F}}^*$ is the optimal solution of (3.4). By plugging in the optimal value $\boldsymbol{\mathcal{F}}^*$ in (3.5), we obtain

$$\mathbf{E}_k^{\mathrm{DL},opt} = \boldsymbol{I} - \boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{R}_k^{-1}\boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k. \tag{A.2}$$

Hence plugging $\mathbf{E}_k^{\mathrm{DL},opt}$ in (3.4) yields

$$\begin{aligned} &\max_{\boldsymbol{\mathcal{G}}_k,\boldsymbol{W}_k} \log\det(\boldsymbol{W}_k) - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k\big) + d_k \\ &= \max_{\boldsymbol{W}_k} \log\det(\boldsymbol{W}_k) - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k^{\mathrm{DL},opt}\big) + d_k. \end{aligned} \tag{A.3}$$

The first order optimality condition of (A.3) with respect to $\boldsymbol{W}_k$ implies

$$\boldsymbol{W}_k^* = (\mathbf{E}_k^{\mathrm{DL},opt})^{-1} \tag{A.4}$$

By plugging in the optimal $\boldsymbol{W}_k^*$ in (A.3), we can write

$$\begin{aligned}
&\max_{\boldsymbol{\mathcal{G}}_k,\boldsymbol{W}_k} \log\det(\boldsymbol{W}_k) - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k^{\mathrm{DL}}\big) + d_k \\
&\quad = -\log\det(\mathbf{E}_k^{\mathrm{DL},opt}) \qquad (A.5)\\
&\quad = -\log\det(\boldsymbol{I} - \boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{R}_k^{-1})\\
&\quad = -\log\det((\boldsymbol{R}_k - \boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}})\boldsymbol{R}_k^{-1})\\
&\quad = \log\det\left(\boldsymbol{R}_k(\boldsymbol{R}_k - \boldsymbol{H}_k\boldsymbol{\mathcal{G}}_k\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}})^{-1}\right) = \log\det\left(\boldsymbol{R}_k\overline{\boldsymbol{R}}_{\bar{k}}^{-1}\right), \qquad (A.6)
\end{aligned}$$

which is the rate of user k in (3.3).

A.2 MSE Duality

Given $\boldsymbol{F}$, $\boldsymbol{G}$, and a total power limit $P_{\max}$, the same MSE values $\varepsilon_1 \dots \varepsilon_{N_d}$ can be achieved in the downlink channel (Figure 3.2) and uplink channel (Figure 3.3).

With matrices

$$[\boldsymbol{D}]_{ii} = \beta_i^2\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}\boldsymbol{H}\boldsymbol{g}_i - 2\beta_i\mathrm{Re}\{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{f}_i\} + 1$$

and

$$[\boldsymbol{\Psi}]_{ij} = \begin{cases}\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{f}_j\boldsymbol{f}_j^{\mathrm{H}}\boldsymbol{H}\boldsymbol{g}_i, & i \neq j\\ 0, & i = j.\end{cases}$$

we can rewrite the downlink MSE (3.1) and uplink MSE (3.8) as

$$\varepsilon_i^{\mathrm{DL}} = [\boldsymbol{D}]_{ii} + \beta_i^2/p_i[\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{p}]_i + \sigma_n^2\beta_i^2/p_i,\ \forall i, \qquad (A.7)$$

and

$$\varepsilon_i^{\mathrm{UL}} = [\boldsymbol{D}]_{ii} + \beta_i^2/q_i[\boldsymbol{\Psi}\boldsymbol{q}]_i + \sigma_n^2\beta_i^2/q_i,\ \forall i, \qquad (A.8)$$

respectively.

Collecting all layer MSEs in a diagonal matrix $\varepsilon^{\mathrm{UP/DL}} = \mathrm{diag}\big\{[\varepsilon_1^{\mathrm{UP/DL}} \dots \varepsilon_{N_d}^{\mathrm{UP/DL}}]\big\}$ we obtain

$$\boldsymbol{\varepsilon}^{\mathrm{UL}}\mathbf{1}_{N_d} = \boldsymbol{Q}^{-1}\left[(\boldsymbol{D} + \boldsymbol{\beta}^2\boldsymbol{\Psi})\boldsymbol{Q}\mathbf{1}_{N_d} + \sigma_n^2\boldsymbol{\beta}^2\mathbf{1}_{N_d}\right], \qquad (A.9)$$

and

$$\boldsymbol{\varepsilon}^{\mathrm{DL}}\mathbf{1}_{N_d} = \boldsymbol{P}^{-1}\left[(\boldsymbol{D} + \boldsymbol{\beta}^2\boldsymbol{\Psi}^T)\boldsymbol{P}\mathbf{1}_{N_d} + \sigma_n^2\boldsymbol{\beta}^2\mathbf{1}_{N_d}\right]. \qquad (A.10)$$

If $\lambda_{\max}((\boldsymbol{\varepsilon}-\boldsymbol{D})^{-1}\boldsymbol{\beta}^{\mathbf{2}}\boldsymbol{\Psi})<1$, then there exists a strictly positive power allocation

$$\boldsymbol{q}=\sigma_n^2(\boldsymbol{\varepsilon}-\boldsymbol{D}-\boldsymbol{\beta}^2\boldsymbol{\Psi})^{-1}\boldsymbol{\beta}^2\mathbf{1}_{N_d} \tag{A.11}$$

such that $\varepsilon_i^{\mathrm{UP}}=\varepsilon_i,\ \forall i$. This is an immediate consequence of the convergence properties of the Neumann series in which (A.11) can be decomposed (see [97], and the power control literature [98] for more details). Conversely, by a similar reasoning as shown in [97] in the context of SINR, we know that if there exists a $\boldsymbol{q}>0$ such that $\varepsilon_i^{\mathrm{UP}}=\varepsilon_i,\ \forall i$, then $\lambda_{\max}((\boldsymbol{\varepsilon}-\boldsymbol{D})^{-1}\boldsymbol{\beta}^{\mathbf{2}}\boldsymbol{\Psi})<1$ holds.

In the same way it can be shown that downlink targets $\varepsilon_i\ \forall i$, can be achieved if and only if $\lambda_{\max}((\boldsymbol{\varepsilon}-\boldsymbol{D})^{-1}\boldsymbol{\beta}^{\mathbf{2}}\boldsymbol{\Psi}^{\mathrm{T}})<1$. If the targets are feasible, then they can be achieved by a strictly positive power allocation

$$\boldsymbol{p}=\sigma_n^2(\boldsymbol{\varepsilon}-\boldsymbol{D}-\boldsymbol{\beta}^2\boldsymbol{\Psi}^T)^{-1}\boldsymbol{\beta}^2\mathbf{1}_{N_d}.$$

The same spectral radius of $(\boldsymbol{\varepsilon}-\boldsymbol{D})^{-1}\boldsymbol{\beta}^{\mathbf{2}}\boldsymbol{\Psi}$ and $(\boldsymbol{\varepsilon}-\boldsymbol{D})^{-1}\boldsymbol{\beta}^{\mathbf{2}}\boldsymbol{\Psi}^{\mathrm{T}}$, implies that $\boldsymbol{q}>1$ exists, if and only if $\boldsymbol{p}>1$ exists.

Both allocations have the same total power. This can be verified by

$$\begin{aligned}
\|\boldsymbol{q}\|_1 &= \mathbf{1}_{N_d}^T\boldsymbol{q}=\sigma_n^2\mathbf{1}_{N_d}^T(\boldsymbol{\varepsilon}-\boldsymbol{D}-\boldsymbol{\beta}^2\boldsymbol{\Psi})^{-1}\boldsymbol{\beta}^2\mathbf{1}_{N_d} && \text{(A.12)}\\
&= \sigma_n^2\mathbf{1}_{N_d}^T\left[\boldsymbol{\beta}^{-2}(\boldsymbol{\varepsilon}-\boldsymbol{D})-\boldsymbol{\Psi}\right]^{-1}\mathbf{1}_{N_d} && \text{(A.13)}\\
&= \sigma_n^2\mathbf{1}_{N_d}^T\left[\boldsymbol{\beta}^{-2}(\boldsymbol{\varepsilon}-\boldsymbol{D})-\boldsymbol{\Psi}^T\right]^{-1}\mathbf{1}_{N_d} && \text{(A.14)}\\
&= \sigma_n^2\mathbf{1}_{N_d}^T(\boldsymbol{\varepsilon}-\boldsymbol{D}-\boldsymbol{\beta}^2\boldsymbol{\Psi}^T)^{-1}\boldsymbol{\beta}^2\mathbf{1}_{N_d} && \text{(A.15)}\\
&= \mathbf{1}_{N_d}^T\boldsymbol{p}=\|\boldsymbol{p}\|_1\le P_{\max}, && \text{(A.16)}
\end{aligned}$$

i.e., the same feasible layer MSE values can be achieved in both links with the same total transmit power $\|\boldsymbol{q}\|_1=\|\boldsymbol{p}\|_1\le P_{\max}$.

Appendix B: ESIP Method for Rate Balancing

B.1 Minorizor Optimization Formulation

From (8.13), apart from the power constraints and constant terms, we focus the optimization on $\sum_k \breve{\lambda}'_k \operatorname{lndet}\Big(\boldsymbol{I} + \boldsymbol{G}_k^{\mathrm{H}} \overline{\boldsymbol{B}}_k^{\mathrm{S}} \boldsymbol{G}_k\Big) - \sum_k \frac{\breve{\lambda}'_k}{p_k} \operatorname{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}} \overline{\boldsymbol{R}}_{\bar{k}}\}$. The first term corresponds to the first term in (8.23) with $\breve{\lambda}'_k = \breve{\lambda}_k / \breve{\xi}_k^{\mathrm{S}}$. By developing the second term, we obtain the following

$$\sum_k \frac{\breve{\lambda}_k}{p_k \breve{\xi}_k^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}} \overline{\boldsymbol{R}}_{\bar{k}}\} = \sum_k \frac{\breve{\lambda}_k}{p_k \breve{\xi}_k^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}} (\sigma_n^2 \boldsymbol{I} + \sum_{i \neq k} p_i \overline{\mathbf{S}}_{k,i})\} \tag{B.1}$$

$$= \sigma_n^2 \sum_k \frac{\breve{\lambda}_k}{p_k \breve{\xi}_k^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}}\} + \sum_k \sum_{i \neq k} \frac{p_i \breve{\lambda}_k}{p_k \breve{\xi}_k^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}} \overline{\mathbf{S}}_{k,i}\}$$

$$= \sigma_n^2 \sum_k \frac{\breve{\lambda}_k}{p_k \breve{\xi}_k^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_k^{\mathrm{S}}\} + \sum_i \sum_{k \neq i} \frac{p_k \breve{\lambda}_i}{p_i \breve{\xi}_i^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_i^{\mathrm{S}} \overline{\mathbf{S}}_{i,k}\} \tag{B.2}$$

where (B.1) follows from (7.7) and we inverse between the indices k and i to get (B.2), using $\sum_k \sum_{i \neq k} = \sum_{k,i,i \neq k} = \sum_i \sum_{k \neq i}$. The first term in (B.2) is constant, we focus the optimization on the second term which becomes as follows using (7.6)

$$\sum_{k,i,k \neq i} \frac{p_k \breve{\lambda}_i}{p_i \breve{\xi}_i^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_i^{\mathrm{S}} \overline{\mathbf{S}}_{i,k}\} = \sum_{k,i,k \neq i} \frac{p_k \breve{\lambda}_i}{p_i \breve{\xi}_i^{\mathrm{S}}} \operatorname{tr}\{\breve{\boldsymbol{W}}_i^{\mathrm{S}} (\widehat{\boldsymbol{H}}_{i,b_k} \boldsymbol{G}_k \boldsymbol{G}_k^{\mathrm{H}} \widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}} + \operatorname{tr}\{\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{C}_{i,b_k} \boldsymbol{G}_k\} \boldsymbol{I})\}$$

$$= \sum_{k,i,k \neq i} \frac{p_k \breve{\lambda}_i}{p_i \breve{\xi}_i^{\mathrm{S}}} \operatorname{tr}\{\boldsymbol{G}_k^{\mathrm{H}} (\widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}} \breve{\boldsymbol{W}}_i^{\mathrm{S}} \widehat{\boldsymbol{H}}_{i,b_k} + \operatorname{tr}\{\breve{\boldsymbol{W}}_i^{\mathrm{S}}\} \boldsymbol{C}_{i,b_k}) \boldsymbol{G}_k\}$$

$$= \sum_k p_k \operatorname{tr}\{\boldsymbol{G}_k^{\mathrm{H}} \sum_{i \neq k} \frac{\breve{\lambda}_i}{p_i \breve{\xi}_i^{\mathrm{S}}} (\widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}} \breve{\boldsymbol{W}}_i^{\mathrm{S}} \widehat{\boldsymbol{H}}_{i,b_k} + \operatorname{tr}\{\breve{\boldsymbol{W}}_i^{\mathrm{S}}\} \boldsymbol{C}_{i,b_k}) \boldsymbol{G}_k\}$$

$$= \sum_k p_k \operatorname{tr}\{\boldsymbol{G}_k^{\mathrm{H}} \overline{\boldsymbol{A}}_k^{\mathrm{S}} \boldsymbol{G}_k\} \tag{B.3}$$

which corresponds to the second term of (8.23) and thus completes the formulation.