Robust Optimization of Multiple Antenna Systems With Channel Uncertainty

vorgelegt von

Diplom-Ingenieur

Nikola Vucic

aus Belgrad

Von der Fakultät IV - Elektrotechnik und Informatik

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

Dr.-Ing.

genehmigte Dissertation

Promotionsausschuss :

Vorsitzender :	Prof. Dr.-Ing. Gerhard Mönich
Berichter :	Prof. Dr.-Ing. Dr. rer. nat. Holger Boche
Berichter :	Prof. Dr. Alex Gershman (TU Darmstadt)

Tag der wissenschaftlichen Aussprache : 14.7.2009

Berlin 2009

D 83

Uspomeni na mamu

Zusammenfassung

Die Anpassung an den Mobilfunk Kanal ist ein kritischer Faktor um die räumliche Diversität als zusätzlichen Freiheitsgrad in Mehrantennensystemen effektiv auszunutzen. Auf Grund fehleranfälliger Kanalschätzung, Quantisierung, schnell wechselnder Umgebung kombiniert mit strikten Forderungen an die Verzögerung und Hardware-Einschränkungen ist die Annahme perfekter Kanalkenntnis in der Praxis nicht realistisch. In Abhängigkeit der Hauptfehlerquellen existieren verschiedene mathematische Modelle um die Unsicherheit zu beschreiben. Hierbei sind die Quantisierungsfehler üblicherweise beschränkt. Im Gegensatz dazu werden die Fehler bei der Kanalschätzung als unbeschränkte, gaußverteilte Zufallsvariablen angenommen.

In modernen Mobifunksystemen werden bei der Entwicklung der Send- und Empfangseinheiten bestimmte Qualitätsanforderungen berücksichtigt. Wird die Fehlerbehaftung der Kanalkoeffizienten bei der Systementwicklung nicht mitberücksichtigt kann dies zu einer häufigen Verletzung der Anforderungen führen. Deswegen ist Robustheit bezüglich imperfekter Kanalkenntnis von besonderer, praktischer Bedeutung. Qualitätsanforderungen können strikt sein und haben damit eine besonders hohe Priorität. Für die anderen Anforderungen könnte (oder müßte) der Dienstanbieter eine bestimmte Prozentzahl der Ausfälle erlauben. Dies führt zu zwei unterschiedlichen Philosophien, der *worst-case* (schlimmster Fall) Optimierung und der probabilistischen Optimierung. Bei *worst-case* Methoden wird das strenge Einhalten der Performanzziele für alle Kanäle, die in einer bestimmten Unsicherheitsregion enthalten sind, gefordert. Demgegenüber zielt die probabilistische Optimierung darauf ab die Qualitätsanforderungen mit bestimmten Wahrscheinlichkeiten zu erfüllen.

In dieser Arbeit werden Transceiver Optimierungsprobleme robust gelöst. Hauptsächlich wird die Abwärtsstrecke (*downlink*) im Mobilfunk untersucht. Hierbei liegt der Fokus auf imperfekter Kanalkentnis an der Senderseite. Für strikte Qualitätsanforderungen werden die Transceiver Optimierungsprobleme als semidefinite Programme mit effizienten iterativen Lösungen umgeschrieben. Bei probabilistischen Anforderungen werden die speziellen Eigenschaften der vorhandenen Zufallsvariablen (wie zum Beispiel die Unimodalität) genutzt, um Algorithmen abzuleiten, die auf der Interferenzfunktionentheorie oder auf konvexer Optimierung basieren. Für Sonderfälle werden Lösungen in geschlossener Form entwickelt. Im Vergleich zu Methoden aus der Literatur erzielen die in dieser Arbeit entwickelten Algorithmen erhebliche Verbesserungen bezüglich der Performanz oder der Komplexität.

Abstract

Exploitation of spatial diversity, as an additional degree of freedom in multiple antenna wireless systems, depends crucially on ability to adapt to channel conditions. The assumption of having perfect knowledge of the channel is, however, often unrealistic in practice. Noise-prone channel estimation, quantization effects, fast varying environment combined with delay requirements, and hardware limitations are some of the most important factors that cause errors. Depending on the primary source of errors in channel state information (CSI), various mathematical models for the uncertainty can be adopted. For example, quantization errors are usually bounded, while estimation errors are often modeled as Gaussian random variables.

Modern wireless systems are supposed to include quality-of-service (QoS) based transceiver designs. However, a transceiver design that does not account for the errors in CSI can result in frequent violation of the promised QoS targets. Providing robustness to imperfect CSI is, therefore, a task of significant practical interest. Some QoS targets might be of particular importance and defined as strict. In other cases, the operator will be either willing or forced to allow a certain percentage of outages in the system. These observations naturally lead to two robust philosophies: the worst-case optimization and the probabilistically constrained optimization. In the worst-case approaches, some performance targets must be satisfied for all channels contained in the uncertainty regions. On the other hand, the probabilistically constrained optimization has a goal of satisfying the QoS constraints with certain probabilities.

The principal task of this thesis is solving transceiver optimization problems in a robust manner. The accent is put on downlink systems and the critical problem of imperfect CSI at the transmitter. As main tools, extensions of mathematical programming for supporting uncertain problem parameters are employed. The used iterative algorithms benefit from a major recent progress in optimization theory, which was particularly noticeable in the area of convex programming. Robust transceiver optimization problems with strict QoS targets are rewritten as semidefinite programming problems, which have efficient numerical solutions. In the case of probabilistic constraints, the exploitation of specific properties of random variables at hand, such as unimodality, yielded problems solvable by theory of interference functions, or conic quadratic programming. In some special setups, closed-form solutions are derived. The obtained algorithms for the worst-case and probabilistically constrained robust transceiver designs outperform relevant results from the literature in terms of performance or computational complexity.

Acknowledgments

I am deeply grateful to my advisor, Prof. Holger Boche, for his support and for introducing me to many interesting areas of communications and mathematics. I would also like to express my sincere gratitude to Prof. Alex Gershman for accepting to examine the thesis and serve in the thesis committee.

This work has benefited from valuable discussions with my colleagues at Heinrich Hertz Institute, Fraunhofer German-Sino Lab for Mobile Communications, and Technical University Berlin. I especially thank Martin Schubert for guiding my first steps as a doctoral student. I am grateful to Volker Pohl, Shuying Shi, and Siddharth Naik for collaborating and co-authoring papers with me. In addition, I would like to thank Ingmar Blau for sharing room and ideas, and for improving my German.

My most heartfelt gratitude goes to my family. I thank my parents, Nevenka and Desimir, for their endless love and support. Last, but far from least, I thank my wife, Jelena. Without her understanding and encouragement, this work would not have come to fruition.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Ever increasing data rate demands from wireless devices can be answered by introducing another type of diversity, besides already widely used temporal and frequency diversities. Spatial diversity is realized by employing antenna arrays (multiple antennas) at transmitters and/or receivers, and by a joint consideration of the whole multiuser network [DADSC04]. The gains promised by information theory for multiple antenna systems [Tel95, Fos96] have instigated work on practical signaling schemes which could approach the limit performance. The employment of antenna arrays has been adopted by current and emerging standards. Some of the examples are the Universal Mobile Telecommunications System (UMTS) [rGPP02], Worldwide Interoperability for Microwave Access (WiMAX) [IEE06], Wireless Local Area Networks (WLANs) [IEE], or the Long Term Evolution (LTE) [rGPP07]. However, in order to fully utilize the introduced spatial diversity, multi-antenna systems require innovative signal processing. Most notably, the transceivers should be adapted to channel conditions.

1.1.1 Transceiver Optimization With Perfect Knowledge of the Channel

There is a great amount of research work on transceiver optimization in multi-antenna systems, under the assumption that the transmitters and receivers are provided with perfect channel state information (CSI). The single-user, multiple-input multiple-output (MIMO) setup can be considered as relatively well-understood [PGNB04, WFVH04, PJ06]. The focus has recently gradually shifted to resource allocation and transceiver optimization in multiuser uplink and downlink scenarios. By uplink, we will assume a setup where more non-cooperating users transmit signals to a single base station (BS). On the other hand, in downlink, the BS simultaneously sends independent signals for non-cooperating users. Improving both uplink [Yat95, LDGW04, SB05a] and downlink [RFLT98, SB04, BO02, WES06, KTA06, MJHU06, SSB07, SSB08a, SSB08b] system performance is of significant interest. However, it should be remarked that optimization problems in downlink exhibit often a more involved mathematical structure.

Due to the intricacy of the problems, closed-form solutions appear relatively rare. The optimization of power allocation and/or equalization at the transmitters and receivers is usually performed using iterative algorithms. The most common fundamental theoretical bases on which these rely are the interference functions [Yat95, SB05b], mathematical programming and convex optimization [Ber99, BTN01, BV04, LY06], duality theory [VT03, SB04], and majorization theory [MO79, PJ06].

1.1.2 Imperfect Knowledge of Wireless Channels

The fast varying wireless environment, in combination with often very stringent delay constraints, makes the provision of perfect CSI extremely difficult. The receivers in wireless systems typically obtain the information about the channel using estimation algorithms for prearranged pilot sequences [HH03, BG06]. This process is inherently erroneous. At the transmitters, some additional obstacles appear. In time-division-duplex (TDD) systems, the transmitter might be able to exploit the channel reciprocity. This means using estimated channels from the phase where it operated as a receiver. Other option is the supply of CSI from the receiver using feedback channels. In both scenarios, the rapidly changing wireless environment can result in outdated estimates. Furthermore, the feedback links are typically of limited capacity [LHSH04], which increases the uncertainty in the CSI at the transmitter.

It is well-known that imperfect CSI can significantly degrade the system performance [HH03, ZG04, Jin06]. In other words, if one derives algorithms for transceiver design based on erroneous channel coefficients as if they were perfect, some promised quality-of-service (QoS) targets in the system might often be violated. Therefore, designing transmitters and receivers which are robust to imperfect CSI is a task of great practical interest.

1.2 Related Work on Robust Designs

There are various ways to introduce robustness to CSI errors in wireless systems. For example, the limited feedback channels, mentioned in the previous section, have attracted a lot of attention recently [LHL+08]. In this approach, the receiver uses estimates of the channel to inform the transmitter by quantizing either the channel coefficients themselves or the required properties of the transmitted signal [LHSH04, LH05, CMJH08]. The accent in the work on limited feedback is put either on vector quantization of the channel or on construction of precoding codebooks.

In this thesis, the focus will be on schemes where only an erroneous estimate of the channel is known at the transmitter. Furthermore, it will be assumed that the CSI errors have certain properties, either in terms of shapes of uncertainty regions, or statistics. The model is, therefore, quite general. Depending on the error properties, it can successfully represent the disturbance due to the quantization for the limited feedback channels, or the estimation errors. The main goal is the provision of some QoS targets despite the channel uncertainty. For this problem formulation, two methodologies might be of particular interest, namely, the worst-case and the probabilistic approaches.

One of the oldest, well-established approaches in robust designs is the worst-case philosophy [KP85]. In worst-case methods, one usually assumes that the errors in channel knowledge are bounded. Typically, no statistical assumption on the mismatch is needed, which indeed might not exist. The aim is to optimize the system to guarantee certain performance for all channels from the uncertainty regions. Recently, there has been a number of applications of this concept in wireless communications. Single-user beamforming with the steering vector mismatch was analyzed in [VGL03, LB05]. Single-user MIMO systems with imperfect CSI were studied in [GL05, PIPPNL06, GL06, PJ06]. Some interesting modifications of the worst-case idea, such as the worst-case regret and ratio optimizations, were proposed in the context of estimation theory in [EM04, EM05]. Most of the work in robust design for single-user MIMO focuses on flat-fading channels with straightforward extensions for the frequency selective case using orthogonal frequency division multiplexing (ODFM) [Gol05]. However, there are also methods which directly approach uncertain frequency selective MIMO channels in a single-carrier manner. In this case, typically the worst-case methodology [KP85, GL05, GL06, BEBT07], or the H^∞ approaches [EHK04, HEK06] are used. Robust, downlink, multiuser, multiple-input, single-output (MISO) scenarios, with a multi-antenna BS and single-antenna users, are studied in [BO02, BSG04, PPIL07, BD07a, MKB07, BD08b]. Robust worst-case optimization of multiuser MIMO systems, where the users generally have more than one antenna, is significantly less understood. Linear receivers with imperfect CSI for multiuser, uplink, space-time block coded MIMO systems are proposed in [RSG05].

Worst-case approaches present a convenient framework for modeling quantization errors, since these errors are normally bounded. They are also appropriate for handling slow fading channels, where no sufficient statistics for the averaging is available [WES07]. However, the requirement to satisfy some goals for all uncertain channels from a specified region might sometimes lead to designs which are overly conservative. Furthermore, the CSI errors are often unbounded. The most prominent example is the channel estimation process. The estimation errors are typically modeled as random variables with the Gaussian distribution. In the unbounded case, it is usually not possible to guarantee any QoS targets with an absolute certainty. The idea to exploit statistical properties of the CSI error, if they exist, and require the fulfillment of some performance targets with certain probabilities, leads to the concept of probabilistically constrained (also coined as chance constrained, or outage based) signal processing [Pré95, BL97]. The knowledge about the error might be more precise, as in the case when the exact distribution is available. Sometimes, however, only some parameters, like the error covariance matrix, could be provided. The first applications of these ideas in the contexts of estimation theory and wireless communications have been proposed recently in [VENG05, RVG06, CSCG07, Vor08, VCG08]. Most of the problems in this young area remain open though. There exists also a related group of stochastic (or Bayesian) robust designs which exploit the statistical information about the error and improve the average performance of the system [ZG02, JG04, WSM06, BD07b, BD08b, ZPO08]. It should be remarked though, that the outage based formulations exhibit often a very different mathematical structure from these problems.

1.3 Outline of the Thesis

The main problem considered in this thesis is the optimization of multiple antenna transceivers under the assumption of having imperfect CSI. Depending on a considered scenario, various problems can be defined. The principal contributions are divided in four chapters (Chapters 3, 4, 5, and 6), with respect to assumptions on system setup and channel uncertainty model.

In Chapter 3, downlink, flat-fading, multiuser MISO setups are analyzed. A robust counterpart of the standard beamforming problem is defined for these systems. The transceiver optimization is performed so that the total transmit power of the BS is minimized, subject to predefined quality-of-service (QoS) targets. The targets are satisfied for all channels that belong to the given ball uncertainty regions. The solutions are obtained using iterative algorithms, based on the semidefinite programming (SDP) methods [VB07b, VB08a, VB09a].

These results are extended in Chapter 4 in order to support downlink, flat-fading, multiuser MIMO systems, where the users also have multiple antennas. The problems become, however, more intricate. The reason is that matrix instead of scalar receivers must be optimized. An iterative, alternating framework, flexible enough to handle a number of optimization tasks in this setup, is presented [VBS07, VBS08, VBS09]. The studied problems include the power minimization subject to per-user or per-layer mean-square-error (MSE) constraints, minimization of the sum MSE under a total transmit power constraint, or minimization of the maximum MSE among the users. The iterations are shown to have equivalent SDP representations. Uncertain knowledge of the noise covariance, non-linear precoding, and various types of power constraints are covered, as well. In all cases, the MSEs are satisfied for any channel that might appear in a channel uncertainty region of ellipsoidal shape. This chapter also derives a general approach of calculating the actual worst-case channels in multi-antenna systems with ellipsoidal uncertainty sets, based on a strong duality result for trust region subproblems. The ability to determine the least favorable channels is important since it shows the possible decrease in transmission quality in practice due to imperfect CSI.

Frequency selective, single-user MIMO channels are the subject of Chapter 5. Here, we assume the less-understood, single-carrier, spatiotemporal precoding/equalization. The precoder and the receive equalizer are assumed to be practically realizable in the sense of being causal and having finite impulse responses (FIRs). The uncertainty is modeled by bounding the H_2 norm of the error matrix transfer function (MTF). Several optimization problems with MSEs as the performance criteria are analyzed using a framework based on the alternating transmitter/receiver optimization [VB07a, VB08c]. The iterations are, similarly to Chapter 4, resolved using SDP methods.

The analyzed uncertainty model is fundamentally changed in Chapter 6, where firstly the Gaussian CSI mismatch in a downlink multiuser MISO system is assumed. The task is to optimize the power control so that the QoS targets are satisfied with the specified probabilities. A conservative, deterministic approximation of this chance-constrained optimization problem is obtained by applying the Vysochanskii-Petunin inequality [VB08b, VB09b]. Various solutions for the resulting optimization problem are shown, based on theory of interference functions, or convex optimization. Some special cases regarding the channel uncertainty model, which admit closed-form solutions, are explained, as well. The power control problem is analyzed also in the scenario where only the covariance matrices of the CSI errors are known at the transmitter, and not the exact distribution. In this case, the posed problem is solved optimally in a closed form.

Finally, the thesis is concluded in Chapter 7. The main contributions are summarized in this chapter, and some open problems and possible directions for future research are discussed.

1.4 Notation

If not stated otherwise, the notation will be in accordance with the following table. Small and large bold fonts denote vectors and matrices, respectively. The dimensions will either be clear from the context, or they will be explicitely written.

$\boldsymbol{I}$	Identity matrix ($\boldsymbol{I}_M$ is the identity $M \times M$ matrix).
$\mathbf{0}$	Zero-matrix.
$\boldsymbol{A} \geq \boldsymbol{B}$	$\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite ("$>$" is used for the positive definiteness) [BV04].
$\boldsymbol{A} \otimes \boldsymbol{B}$	Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$ [HJ91].
$\mathrm{Tr}\{\cdot\}$	Trace of a matrix.
$\lvert(\cdot)\rvert$	Absolute value.
$\lVert(\cdot)\rVert_1$	l_1 (sum) norm [HJ85].
$\lVert(\cdot)\rVert_2$	Spectral norm of a matrix (the Euclidean norm of a vector) [HJ85], or H_2 norm of a matrix transfer function [BGFB94]. The meaning will be clear from the context.
$\lVert(\cdot)\rVert_\infty$	l_∞ (max) norm of a vector [HJ85].
$\lVert(\cdot)\rVert_F$	Frobenius norm [HJ85].
$\lambda_{\max}(\cdot)$	Maximum eigenvalue of a matrix.
$(\cdot)^T$	Matrix transpose.
$(\cdot)^*$	Conjugate (Hermitian) transpose.
$(\cdot)^\dagger$	Moore-Penrose pseudoinverse [Mey00].
$\boldsymbol{A}_{(k,m)}$	Element in position (k, m) of the matrix $\boldsymbol{A}$.
$\boldsymbol{A}_{(:,m)}$	mth column of $\boldsymbol{A}$.
$\boldsymbol{A}_{(k,:)}$	kth row of $\boldsymbol{A}$.
$\mathrm{vec}(\boldsymbol{A})$	Vector that contains the columns of the matrix $\boldsymbol{A}$, stacked below each other.
$\mathrm{diag}(a_1, \ldots, a_K)$	Diagonal $K \times K$ matrix having elements $a_1, \ldots, a_K$ on the main diagonal.
j	Imaginary unit.
$\Re\{\cdot\}$	Real part of the argument.
$\Im\{\cdot\}$	Imaginary part of the argument.
$\triangleq$	Defined as.
$\mathrm{E}\{\cdot\}$	Mean value of a random variable.
$\mathrm{Var}\{\cdot\}$	Variance of a random variable.
$\mathrm{Pr}\{\cdot\}$	Probability of an event.
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with mean value μ and variance σ^2.
$\mathbb{R}$	Set of real numbers.
$\mathbb{R}_+$	Set of non-negative real numbers.
$\mathbb{R}_{++}$	Set of strictly positive real numbers.
$\mathbb{C}$	Set of complex numbers.
$\mathcal{L}^K$	Second order (Lorentz) cone of dimension K [BV04].

Chapter 2

Mathematical Preliminaries

In this chapter, an overview of main mathematical results that will be used in the thesis is given. The principal tool of interest is convex optimization theory, together with its extensions for supporting uncertain problem parameters. Before explaining convex problems, some general lemmas from matrix analysis are reviewed, which will be frequently exploited in the sequel.

Lemma 1 (Cauchy-Schwarz Inequality)**.** For any $\boldsymbol{x} \in \mathbb{C}^N$, $\boldsymbol{y} \in \mathbb{C}^N$

$$|\boldsymbol{x}^* \boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \, \|\boldsymbol{y}\|_2 \,. \tag{2.1}$$

Proof. See, e.g., [Mey00], p. 272. □

Lemma 2. For any $\boldsymbol{x} \in \mathbb{C}^N$, $y \in \mathbb{R}$ the following relation is valid

$$\|\boldsymbol{x}\|_2 \leq y \quad \Leftrightarrow \quad \begin{bmatrix} y & \boldsymbol{x}^* \\ \boldsymbol{x} & y\boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0}. \tag{2.2}$$

Proof. The proof for the real-valued case is given in [BTN01], Lecture 4.2. The generalization for complex vectors is straightforward. □

Lemma 3. Let $\boldsymbol{X} \in \mathbb{C}^{N\times N}$ be a Hermitian matrix, and let $y \in \mathbb{R}$. It holds that

$$\lambda_{\max}(\boldsymbol{X}) \leq y \quad \Leftrightarrow \quad y\boldsymbol{I}_N - \boldsymbol{X} \succeq \boldsymbol{0} \tag{2.3}$$

where $\lambda_{\max}(\boldsymbol{X})$ is the maximum eigenvalue of $\boldsymbol{X}$.

Proof. See [BTN01], Lecture 4.2. □

Lemma 4 (Schur Complement Lemma)**.** Consider a Hermitian matrix

$$\boldsymbol{X} \triangleq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^* & \boldsymbol{C} \end{bmatrix} \tag{2.4}$$

with $\boldsymbol{A} \succ \boldsymbol{0}$. $\boldsymbol{X}$ is positive (semi-) definite if and only if $\boldsymbol{C} - \boldsymbol{B}^* \boldsymbol{A}^{-1} \boldsymbol{B}$ is positive (semi-) definite.

Proof. The proof can be found in [HJ85], Section 7.7.6, see also [BTN01], Lecture 4.2. □

2.1 Convex Optimization Theory

The brief overview of convex optimization starts with the basic definitions of convex sets, convex functions, and convex optimization problems.

Definition 2.1 (Convex Sets [BV04])**.** A set $\mathcal{S} \in \mathbb{R}^N$ is convex if for any $\boldsymbol{x} \in \mathcal{S}$ and $\boldsymbol{y} \in \mathcal{S}$, and any θ, $0 \leq \theta \leq 1$, it holds that $\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y} \in \mathcal{S}$.

The optimization in this work is often performed over convex cones. A set $\mathcal{C}$ is a convex cone if for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}$, and $\theta_1, \theta_2 \in \mathbb{R}_+$, $\theta_1\boldsymbol{x} + \theta_2\boldsymbol{y} \in \mathcal{C}$. A well-known example is the second order cone (SOC), also known as quadratic, Lorentz, or the ice-cream cone [BV04]

$$\mathcal{L}^{N+1} \triangleq \left\{(\boldsymbol{x}, y) : \ \boldsymbol{x} \in \mathbb{R}^N, \, y \in \mathbb{R}_+, \, ||\boldsymbol{x}||_2 \leq y\right\}. \tag{2.5}$$

It can be proved that the set of positive semidefinite matrices forms a convex cone, as well [BV04]. Closely related to this is the concept of linear matrix inequalities (LMIs). The LMIs are defined as

$$\mathcal{A}(\boldsymbol{x}) = \left\{\boldsymbol{x} \triangleq [x_1, \ldots, x_N]^T \in \mathbb{R}^N : \ \boldsymbol{A}_0 + x_1\boldsymbol{A}_1 + x_2\boldsymbol{A}_2 + \ldots + x_N\boldsymbol{A}_N \succeq \boldsymbol{0}\right\} \tag{2.6}$$

where $\boldsymbol{A}_0, \ldots, \boldsymbol{A}_N$ are symmetric matrices. The solution set of (2.6) is also convex [BTN01, BV04].

At this point, it is important to remark that, as common in the literature on convex optimization, the definitions in this chapter refer to sets in $\mathbb{R}^N$. On the other hand, communications theory prefers often complex-valued models. The transformation from the complex-valued into the real-valued model is, however, simple for the most problems of interest in this thesis. In the case of SOCs, for any $\boldsymbol{x} \in \mathbb{C}^N$ it clearly holds that

$$\|\boldsymbol{x}\|_2 \leq y \quad \Leftrightarrow \quad \left\| \begin{bmatrix} \Re\{\boldsymbol{x}\} \\ \Im\{\boldsymbol{x}\} \end{bmatrix} \right\|_2 \leq y. \tag{2.7}$$

Complex LMIs can be transformed into real-valued ones by noticing that for any Hermitian matrix $\boldsymbol{A}$

$$\boldsymbol{A} \geq \boldsymbol{0} \quad \Leftrightarrow \quad \begin{bmatrix} \Re\{\boldsymbol{A}\} & \Im\{\boldsymbol{A}\}^T \\ \Im\{\boldsymbol{A}\} & \Re\{\boldsymbol{A}\} \end{bmatrix} \geq \boldsymbol{0}. \tag{2.8}$$

Definition 2.2 (Convex Function [BV04])**.** A function $f : \mathbb{R}^N \to \mathbb{R}$ is convex if its domain is convex, and if for all $\boldsymbol{x}$ and $\boldsymbol{y}$ from the domain,

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}) \tag{2.9}$$

for any θ, $0 \leq \theta \leq 1$.

Definition 2.3 (Convex Optimization Problem [BV04])**.** A convex optimization problem has the form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & f_i(\boldsymbol{x}) \leq 0, \quad i = 1, \ldots, M \\ & \boldsymbol{a}_l^T \boldsymbol{x} = b_l, \quad l = 1, \ldots, P \end{aligned} \tag{2.10}$$

where $f_0, \ldots, f_M$ are convex functions, and $\boldsymbol{a}_l, b_l$, $l = 1, \ldots, P$, are fixed parameters.

The convexity is often considered as a criterion that separates efficiently solvable from difficult optimization problems. Almost all convex problems can be solved, either in a closed form or using iterative algorithms. Some classes of them have very efficient numerical solutions. Among the used algorithms, the interior point methods are considered to be one of the most promising techniques, applicable to a wide range of convex problems [NN94, VB96, BTN01, BV04, LY06]. In the rest of this section, linear, conic quadratic, and SDP problems are defined. General upper bounds for their solutions by interior point methods are presented, as well. It should be noted that these bounds are typically conservative [BTN01]. They also do not exploit special structure, which the problems at hand might exhibit.

In the sequel, the vector $\boldsymbol{d}$ will represent the data that determines the problem of interest. The size of this vector is s, and $\mathrm{Opt}(\boldsymbol{d})$ is the optimal value of the problem with the parameters in $\boldsymbol{d}$. The number of accuracy digits in ϵ-solution (the term ϵ-solution will be specified for particular problems) is defined as (see Lecture 5 in [BTN01])

$$\mathrm{dig}(\boldsymbol{d}, \epsilon) = \ln\left(\frac{s + \|\boldsymbol{d}\|_1 + \epsilon^2}{\epsilon}\right). \tag{2.11}$$

The Newton complexity means the maximum number of iterations for the termination of the interior point methods, either in finding ϵ-solution, or in determining that the problem is infeasible. The arithmetic complexity is the maximum number of arithmetic operations for these methods.

2.1.1 Linear Problems

A linear problem has the form

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^N} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{a}_i^T\boldsymbol{x} \le b_i, \quad i = 1, \dots, M \\ & \|\boldsymbol{x}\|_\infty \le R. \end{aligned} \tag{2.12}$$

The data vector and its size are given as

$$\boldsymbol{d} = \left[N, M, \boldsymbol{c}^T, \boldsymbol{a}_1^T, b_1, \cdots, \boldsymbol{a}_M^T, b_M, R\right]^T \tag{2.13}$$

$$s = (M+1)(N+1) + 2. \tag{2.14}$$

The ϵ-solution is defined as $\boldsymbol{x} \in \mathbb{R}^N$ such that

$$\begin{aligned} &\|\boldsymbol{x}\|_\infty \leq R \\ &\boldsymbol{a}_i^T \boldsymbol{x} \leq b_i + \epsilon, \quad i = 1, \ldots, M \\ &\boldsymbol{c}^T \boldsymbol{x} \leq \mathrm{Opt}(\boldsymbol{d}) + \epsilon. \end{aligned} \tag{2.15}$$

The Newton complexity of ϵ-solution (2.15) for linear problems (2.12) is

$$\mathrm{C}_{\mathrm{LP}}^{\mathrm{New}}(\boldsymbol{d}, \epsilon) = O(1)\sqrt{M+N}\,\mathrm{dig}(\boldsymbol{d}, \epsilon). \tag{2.16}$$

The arithmetic complexity of ϵ-solution is given as

$$\mathrm{C}_{\mathrm{LP}}^{\mathrm{Ari}}(\boldsymbol{d}, \epsilon) = O(1)(M+N)^{3/2}\,N^2\,\mathrm{dig}(\boldsymbol{d}, \epsilon). \tag{2.17}$$

2.1.2 Conic Quadratic Problems

Conic quadratic (or SOC) problems are convex problems of the following form

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^N} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} \quad & \|\boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i\|_2 \leq \boldsymbol{c}_i^T \boldsymbol{x} + d_i, \quad i = 1, \ldots, M \\ & \|\boldsymbol{x}\|_2 \leq R \end{aligned} \tag{2.18}$$

where $\boldsymbol{b}_i \in \mathbb{R}^{K_i}$. The data vector and its size are given in this case as

$$\begin{aligned} \boldsymbol{d} &= \left[N, M, K_1, \cdots, K_M, \boldsymbol{c}^T, \mathrm{vec}\,(\boldsymbol{A}_1)^T, \boldsymbol{b}_1^T, \boldsymbol{c}_1^T, d_1, \cdots, \mathrm{vec}\,(\boldsymbol{A}_M)^T, \boldsymbol{b}_M^T, \boldsymbol{c}_M^T, d_M, R\right]^T \\ s &= \left(M + \sum_{i=1}^{M} K_i\right)(N+1) + M + N + 3. \end{aligned} \tag{2.19}$$

For conic quadratic problems, ϵ-solution is defined as $\boldsymbol{x} \in \mathbb{R}^N$ such that

$$\begin{aligned} &\|\boldsymbol{x}\|_2 \leq R \\ &\|\boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i\|_2 \leq \boldsymbol{c}_i^T \boldsymbol{x} + d_i + \epsilon, \quad i = 1, \ldots, M \\ &\boldsymbol{c}^T \boldsymbol{x} \leq \mathrm{Opt}(\boldsymbol{d}) + \epsilon. \end{aligned} \tag{2.20}$$

The Newton complexity of ϵ-solution for the problem (2.18) is

$$\mathrm{C}_{\mathrm{CQP}}^{\mathrm{New}}(\boldsymbol{d}, \epsilon) = O(1)\sqrt{M+1}\,\mathrm{dig}(\boldsymbol{d}, \epsilon). \tag{2.21}$$

The arithmetic complexity for conic quadratic problems is given as

$$\mathrm{C}_{\mathrm{CQP}}^{\mathrm{Ari}}(\boldsymbol{d}, \epsilon) = O(1)\sqrt{M+1}\,N\left(N^2 + M + \sum_{i=1}^{M} K_i^2\right)\mathrm{dig}(\boldsymbol{d}, \epsilon). \tag{2.22}$$

2.1.3 SDP Problems

An SDP problem is defined as a minimization of a linear function subject to LMIs

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^N} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}_0 + \sum_{l=1}^{N} x_l \boldsymbol{A}_l \succeq \boldsymbol{0} \\ & \|\boldsymbol{x}\|_2 \le R \end{aligned} \tag{2.23}$$

where $\boldsymbol{A}_l$ are symmetric block-diagonal matrices with M diagonal blocks $\boldsymbol{A}_l^i$ of size $K_i \times K_i$, $i = 1, \ldots, M$ (notice that this is equivalent to having a set of M LMIs defined by the respective blocks). The data vector and its size are

$$\begin{aligned} \boldsymbol{d} &= \left[N, M, K_1, \cdots, K_M, \boldsymbol{c}^T, \mathrm{vec}\left(\boldsymbol{A}_0^1\right)^T, \cdots, \mathrm{vec}\left(\boldsymbol{A}_0^M\right)^T, \cdots, \mathrm{vec}\left(\boldsymbol{A}_N^1\right)^T, \cdots, \mathrm{vec}\left(\boldsymbol{A}_N^M\right)^T, R\right]^T \\ s &= \left(\sum_{i=1}^{M} \frac{K_i(K_i+1)}{2}\right)(N+1) + M + N + 3. \end{aligned} \tag{2.24}$$

The ϵ-solution is $\boldsymbol{x} \in \mathbb{R}^N$ that satisfies

$$\begin{aligned} & \|\boldsymbol{x}\|_2 \le R \\ & \boldsymbol{A}_0 + \sum_{l=1}^{N} x_l \boldsymbol{A}_l \succeq -\epsilon \boldsymbol{I} \\ & \boldsymbol{c}^T\boldsymbol{x} \le \mathrm{Opt}(\boldsymbol{d}) + \epsilon. \end{aligned} \tag{2.25}$$

The Newton complexity of ϵ-solution in SDP is

$$\mathrm{C}_{\mathrm{SDP}}^{\mathrm{New}}(\boldsymbol{d}, \epsilon) = O(1)\left(1 + \sum_{i=1}^{M} K_i\right)^{1/2} \mathrm{dig}(\boldsymbol{d}, \epsilon). \tag{2.26}$$

The arithmetic complexity for SDP problems is given as

$$\mathrm{C}_{\mathrm{SDP}}^{\mathrm{Ari}}(\boldsymbol{d}, \epsilon) = O(1)\left(1 + \sum_{i=1}^{M} K_i\right)^{1/2} N\left(N^2 + N\sum_{i=1}^{M} K_i^2 + \sum_{i=1}^{M} K_i^3\right)\mathrm{dig}(\boldsymbol{d}, \epsilon). \tag{2.27}$$

2.2 Robust Optimization

This section provides some lemmas that are of particular interest when modeling uncertain parameters in optimization problems. They have already found important applications in operations research, estimation, and control theory. In this thesis, they will be used for designing wireless communication systems, robust to imperfect knowledge of the channel.

Lemma 5. Let $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ be positive semidefinite matrices with $\boldsymbol{C} \succeq \boldsymbol{B}$. Then, $\mathrm{Tr}\{\boldsymbol{AB}\} \leq \mathrm{Tr}\{\boldsymbol{AC}\}$.

Proof. The proof can be found in [EM04]. □

One of the fundamental results in worst-case robust optimization theory is the $\mathcal{S}$-Lemma ($\mathcal{S}$-Procedure) [BTN01, BV04].

Lemma 6 ($\mathcal{S}$-Lemma)**.** Let $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ be symmetric $N \times N$ matrices, $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathbb{R}^N$, and $c_1, c_2 \in \mathbb{R}$. If there exists a vector $\bar{\boldsymbol{x}}$ satisfying

$$\bar{\boldsymbol{x}}^T \boldsymbol{A}_1 \bar{\boldsymbol{x}} + 2\boldsymbol{b}_1^T \bar{\boldsymbol{x}} + c_1 < 0 \tag{2.28}$$

the implication

$$\boldsymbol{x}^T \boldsymbol{A}_1 \boldsymbol{x} + 2\boldsymbol{b}_1^T \boldsymbol{x} + c_1 \leq 0 \quad \Rightarrow \quad \boldsymbol{x}^T \boldsymbol{A}_2 \boldsymbol{x} + 2\boldsymbol{b}_2^T \boldsymbol{x} + c_2 \leq 0 \tag{2.29}$$

is true if and only if

$$\exists\lambda \geq 0: \quad \begin{bmatrix} \lambda \boldsymbol{A}_1 - \boldsymbol{A}_2 & \lambda \boldsymbol{a}_1 - \boldsymbol{a}_2 \\ \lambda \boldsymbol{a}_1^T - \boldsymbol{a}_2^T & \lambda c_1 - c_2 \end{bmatrix} \geq \boldsymbol{0}. \tag{2.30}$$

Finally, the following proposition, whose proof, based on the $\mathcal{S}$-lemma, is given in [EM04], will be of great use in the following chapters.

Lemma 7. Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be given matrices, with $\boldsymbol{A} = \boldsymbol{A}^*$. The relation

$$\boldsymbol{A} \geq \boldsymbol{B}^* \boldsymbol{D} \boldsymbol{C} + \boldsymbol{C}^* \boldsymbol{D}^* \boldsymbol{B}, \quad \forall \boldsymbol{D}: \ \|\boldsymbol{D}\|_2 \leq \varepsilon \tag{2.31}$$

is valid, if and only if

$$\exists\lambda \geq 0: \quad \begin{bmatrix} \boldsymbol{A} - \lambda \boldsymbol{B}^* \boldsymbol{B} & -\varepsilon \boldsymbol{C}^* \\ -\varepsilon \boldsymbol{C} & \lambda \boldsymbol{I} \end{bmatrix} \geq \boldsymbol{0}. \tag{2.32}$$

Chapter 3

Worst-Case Optimization of Downlink Multiuser MISO Systems

Systems in which a multi-antenna BS serves single-antenna users present promising candidates for future cellular networks. The complex signal processing is performed at the BS, where the issue of computational power is less problematic, while the mobile units can be kept relatively simple and inexpensive.

In the downlink of such setups, which is a multiuser MISO scenario, precoding methods can be applied to boost the performance. The idea is to pre-equalize signals at the transmitter, and mitigate the channel-induced interference [Fis02]. These techniques naturally demand that CSI is supplied at the transmitter. However, as explained in Section 1.1.2, the provision of perfect CSI is often an involved task in wireless systems. The fact that the downlink precoding methods can be quite sensitive to imperfect knowledge of the channel [Jin06] necessitates robust designs.

In this chapter, we adhere to the worst-case philosophy. The CSI errors are supposed to belong to the given bounded uncertainty sets, and the system is optimized to satisfy certain requirements for all channels from the uncertainty regions.

The main problem of interest is the transceiver optimization with the goal of minimizing the total transmit power subject to predefined users' QoS targets, in a flat-fading environment. The transmitter will be provided with imperfect channel estimates and with the bounds on the uncertainty regions, which surely contain exact channel values. Differently from [BO02, BSG04],

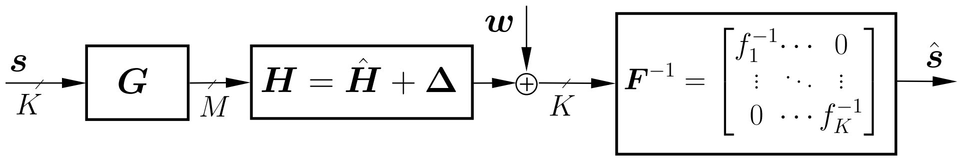

Figure 3.1: Block-scheme of the studied downlink multiuser MISO system.

our focus is on erroneous actual channels and not on disturbed channel covariance matrices, although these models can be considered as comparable [BD07a]. In the cases where the receivers' equalizers should be designed, it is also supposed that the users either know their equivalent channels perfectly, or that there exists an error-free mechanism in the system, which delivers scaling coefficients to the users. This formulation presents one way of defining a robust counterpart of the standard downlink beamforming problem, whose solution in the perfect CSI case is known from [BO02, SB04, SS05, WES06]. The considered robust scenario is considerably more involved comparing with the perfect CSI case. The reason is that the QoS requirements must be supported for an infinite number of possible channels, contained in the uncertainty regions, with a single set of filters.

In the rest of this chapter, firstly the flat-fading, downlink MISO system model is defined in Section 3.1. Sections 3.2 and 3.3 provide solutions for the two most common types of QoS measures, namely the MSEs and the signal-to-interference-plus-noise ratios (SINRs). Numerical examples for the obtained performance and comparisons with some related works are presented in Section 3.4.

3.1 Downlink Multiuser MISO System Model

The considered multiuser MISO system, with K single-antenna users (receivers) and an M-antenna BS (transmitter), is illustrated in Fig. 3.1. At one time instant, the BS transmits K independent symbols s_1, s_2, …, s_K, where the symbol s_k is intended for the kth user.* To simplify the expressions, the transmit symbols are grouped into a vector $\boldsymbol{s}$, $\boldsymbol{s} = [s_1, s_2, \ldots, s_K]^T$, with $\mathrm{E}(\boldsymbol{s}\boldsymbol{s}^*) = \boldsymbol{I}$ assumed, w.l.o.g.

The complete downlink, flat-fading channel is represented with the matrix $\boldsymbol{H}$, $\boldsymbol{H} \in \mathbb{C}^{K\times M}$. If

*Note that the problems of broadcasting and multicasting, where common information is transmitted to certain groups of users, are significantly different [SDL06].

multicarrier techniques, like orthogonal frequency division multiplexing (OFDM) [Gol05], are employed to combat the intersymbol interference in frequency selective channels, the model would correspond to the performance analysis for one subcarrier. The channel of the kth user is given as $\boldsymbol{H}_{(k,:)}$. It is assumed that the BS knows only erroneous channel estimates $\hat{\boldsymbol{H}}_{(k,:)}$, with

$$\boldsymbol{H}_{(k,:)} = \hat{\boldsymbol{H}}_{(k,:)} + \boldsymbol{\Delta}_{(k,:)}, \quad k = 1, \ldots, K \tag{3.1}$$

where $\boldsymbol{\Delta}_{(k,:)}$ is the error in CSI. The BS is also supposed to know the structure of the uncertainty regions $\mathcal{D}_k$, which surely contain the disturbances $\boldsymbol{\Delta}_{(k,:)}$. The requirement on the CSI at the receivers will be commented in more detail in the following sections, depending on the problem of interest.

The transmit filter (precoder) of the BS is denoted by $\boldsymbol{G}$, $\boldsymbol{G} \in \mathbb{C}^{M\times K}$. The eventual equalization of the non-cooperating users, is defined by a diagonal matrix $\boldsymbol{F}$, $\boldsymbol{F} = \mathrm{diag}(f_1, f_2, \ldots, f_K)$, $f_k \in \mathbb{C} \setminus \{0\}$, and the inverses are used for convenience. Finally, the system equations for the K users can be written as

$$\hat{s}_k = \frac{1}{f_k}\left(\boldsymbol{H}_{(k,:)}\boldsymbol{G}\boldsymbol{s} + w_k\right), \quad k = 1, \ldots, K \tag{3.2}$$

where w_k is the additive noise at the receivers, with $\boldsymbol{w} = [w_1, w_2, \ldots, w_K]^T$, and $\hat{s}_k$ is the estimate of s_k. The transmit signals and the receive noise are uncorrelated, and $\mathrm{E}(\boldsymbol{w}\boldsymbol{w}^*) = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_K^2)$ holds.

3.2 MSE-Constrained Optimization

In this section, the MSEs will be adopted as the QoS measure, with

$$\mathrm{MSE}_k = \mathrm{E}\left\{|s_k - \hat{s}_k|^2\right\}, \quad k = 1, \ldots, K. \tag{3.3}$$

The robust MSE-optimization problem assumes the minimization of the total transmit power P,

$$P = \mathrm{E}\left\{\mathrm{Tr}(\boldsymbol{G}\boldsymbol{s}\boldsymbol{s}^*\boldsymbol{G}^*)\right\} = ||\boldsymbol{G}||_F^2 \tag{3.4}$$

subject to predefined MSE targets μ_k, which must be satisfied under uncertainty:

$$\begin{aligned} \min_{\boldsymbol{G}, f_1, \ldots, f_K} \quad & P \\ \text{subject to} \quad & \mathrm{MSE}_k \leq \mu_k, \quad \forall \boldsymbol{\Delta}_{(k,:)} : \boldsymbol{\Delta}_{(k,:)} \in \mathcal{D}_k, \quad k = 1, \ldots, K. \end{aligned} \tag{3.5}$$

The problem (3.5) is meant to be solved by the BS, having in mind that it possesses only imperfect CSI. As far as the users are concerned, in this section, one of the following two options is assumed (see also the discussion in Section 4.1):

- There exists an error-free control mechanism in the system that delivers the equalization coefficients from the BS to the users.

- The users can estimate the equivalent channels $\boldsymbol{H}^e_{(k,:)} = \boldsymbol{H}_{(k,:)}\boldsymbol{G}$, which is, comparing to the availability of the CSI at the transmitter, generally considered as less critical in practice.

In the latter case, the users can adopt the standard minimum MSE (MMSE) solutions for the equalization coefficients f_k, with

$$f_k^{-1} = \boldsymbol{G}^*_{(:,k)} \boldsymbol{H}^*_{(k,:)} \left(\boldsymbol{H}^e_{(k,:)} \boldsymbol{H}^e_{(k,:)}{}^* + \sigma_k^2 \right)^{-1}. \tag{3.6}$$

This approach will result in the obtained MSEs that surely also comply to the uncertain constraints in (3.5), but without reducing the transmit power, obtained as the solution of (3.5). The key idea in optimizing the scaling factors in (3.5) is that the BS determines its beamformer $\boldsymbol{G}$ by having some (partial) knowledge of what could be done at the receivers. This joint optimization yields significant gains, as it will be shown in Section 3.4. Finally, it should be remarked that the incorporation of the MMSE solution (3.6) immediately into the problem formulation (3.5) would yield problems of a significantly more complex structure. This issue is discussed again at the end of Section 3.3.

The problem (3.5), with the assumption of having perfect CSI at both the transmitter and receivers, can be solved using the methodology derived in [BO02, SB04, SS05]. The uncertain problem (3.5) will be approached using numerical methods from convex optimization theory.

First, it can be noticed that the MSE expression of the kth user can be written as

$$\mathrm{MSE}_k = \frac{1}{|f_k|^2}\left(\left(\boldsymbol{H}_{(k,:)}\boldsymbol{G} - f_k\boldsymbol{e}_k^*\right)\left(\boldsymbol{H}_{(k,:)}\boldsymbol{G} - f_k\boldsymbol{e}_k^*\right)^* + \sigma_k^2\right) \tag{3.7}$$

where $\boldsymbol{e}_k$ are standard basis vectors of $\mathbb{R}^K$ [HJ85]. From (3.7), it can be concluded that assuming $f_k \in \mathbb{R}_{++}$ would not change the solution for the minimal transmit power of the optimization problem (3.5), due to the possibility of multiplying the columns of $\boldsymbol{G}$ with complex numbers of unit magnitude, without changing the objective function in (3.5). Therefore, one can proceed with an equivalent representation of the kth user's constraint

$$\mathrm{MSE}_k \le \mu_k \quad \Leftrightarrow \quad \left\|\begin{bmatrix} \boldsymbol{H}_{(k,:)}\boldsymbol{G} - f_k\boldsymbol{e}_k^* & \sigma_k \end{bmatrix}\right\|_2 \le f_k\sqrt{\mu_k}. \tag{3.8}$$

The convexity of the problem (3.5) can be proved now for any uncertainty region $\mathcal{D}_k$.

Theorem 1. The problem (3.5) is convex, irrespectively of the shape of the uncertainty regions $\mathcal{D}_k$.

Proof. Being the squared Frobenius norm of the transmit filter matrix $\boldsymbol{G}$, the objective function in (3.5) is clearly convex [BV04]. Consider now the kth user's constraint (3.8) under uncertainty. For any $\boldsymbol{H}_{(k,:)}$ from the uncertainty regions defined by $\hat{\boldsymbol{H}}_{(k,:)}$ and $\mathcal{D}_k$, (3.8) is a conic quadratic constraint, which yields a convex feasible set for the unknown transceiver parameters [BV04]. The complete domain of (3.8) is the intersection of all such convex sets, corresponding to each channel from the uncertainty region $\hat{\boldsymbol{H}}_{(k,:)} + \boldsymbol{\Delta}_{(k,:)}$, $\boldsymbol{\Delta}_{(k,:)} \in \mathcal{D}_k$. Since the convexity is preserved under intersection [BV04], it can be concluded that the domain of (3.8) is also convex. Finally, the complete problem domain in (3.5) is again convex as the intersection of the domains related to the constraints (3.8), with $k = 1, \ldots, K$. □

Although the problem of interest is convex, it is still not clear if it can be efficiently solved, because of the uncertainty. In the sequel, efficient numerical solutions will be derived for the often used, ball uncertainty model

$$\left\|\boldsymbol{\Delta}_{(k,:)}\right\|_2 \le \varepsilon_k, \quad k = 1, \ldots, K. \tag{3.9}$$

The model (3.9) naturally captures well bounded disturbances that result from quantization (see also Section 4.5), but it is also convenient for representing the CSI errors with Gaussian distribution and optimizing the outage performance [PIPPNL06, PPIL07].

Since the objective function in (3.5) is relatively simple, the focus in the sequel is on the transformation of the uncertain constraints in (3.5) into a tractable form. By applying the relation (2.2) on (3.8), the following equivalent representation of the users constraints is obtained

$$\boldsymbol{\Phi}_k + \begin{bmatrix} 0 & \boldsymbol{\Delta}_{(k,:)}\boldsymbol{G} & 0 \\ \boldsymbol{G}^*\boldsymbol{\Delta}^*_{(k,:)} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & 0 \end{bmatrix} \succeq \mathbf{0}, \quad \forall \boldsymbol{\Delta}_{(k,:)} : \left\|\boldsymbol{\Delta}_{(k,:)}\right\|_2 \leq \varepsilon_k, \quad k = 1, \ldots, K \tag{3.10}$$

where

$$\boldsymbol{\Phi}_k = \begin{bmatrix} f_k\sqrt{\mu_k} & \hat{\boldsymbol{H}}_{(k,:)}\boldsymbol{G} - f_k\boldsymbol{e}_k^* & \sigma_k \\ \boldsymbol{G}^*\hat{\boldsymbol{H}}^*_{(k,:)} - f_k\boldsymbol{e}_k & f_k\sqrt{\mu_k}\boldsymbol{I} & \mathbf{0} \\ \sigma_k & \mathbf{0} & f_k\sqrt{\mu_k} \end{bmatrix}. \tag{3.11}$$

With Lemma 7 recalled, all necessary tools are now provided for rewriting the MSE constraints into a form that is convenient for the numerical analysis.

Theorem 2. The uncertain MSE constraints in (3.5), with the uncertainty regions $\mathcal{D}_k$ defined by (3.9), are equivalent to a set of LMIs in the unknown transceiver coefficients

$$\begin{bmatrix} f_k\sqrt{\mu_k} - \lambda_k & \hat{\boldsymbol{H}}_{(k,:)}\boldsymbol{G} - f_k\boldsymbol{e}_k^* & \sigma_k & \mathbf{0} \\ \boldsymbol{G}^*\hat{\boldsymbol{H}}^*_{(k,:)} - f_k\boldsymbol{e}_k & f_k\sqrt{\mu_k}\boldsymbol{I} & \mathbf{0} & -\varepsilon_k\boldsymbol{G}^* \\ \sigma_k & \mathbf{0} & f_k\sqrt{\mu_k} & \mathbf{0} \\ \mathbf{0} & -\varepsilon_k\boldsymbol{G} & \mathbf{0} & \lambda_k\boldsymbol{I} \end{bmatrix} \succeq \mathbf{0}, \quad k = 1, \ldots, K \tag{3.12}$$

where λ_k are slack variables.

Proof. The LMIs in (3.12) are obtained by a direct application of Lemma 7 on the derived, equivalent representations (3.10) of the users' uncertain MSE constraints, with $\boldsymbol{A} = \boldsymbol{\Phi}_k$, $\boldsymbol{D} = \boldsymbol{\Delta}_{(k,:)}$, $\varepsilon = \varepsilon_k$ and

$$\boldsymbol{B} = -\begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \mathbf{0} & \boldsymbol{G} & \mathbf{0} \end{bmatrix}. \tag{3.13}$$

□

It can be noticed that the objective function in (3.5) can be immediately resolved by introducing one slack variable t that should be minimized and a rotated SOC constraint $\|\text{vec}(\boldsymbol{G})\|_2^2 \leq t$ [LY06]. Therefore, with the uncertainty model (3.9), the problem (3.5) is equivalent to an SDP problem (more rigorously, this is a linear conic optimization problem with second order and semidefinite cones). As explained in Section 2.1, this problem can be solved with interior point methods in a very efficient manner.

3.3 SINR as Performance Measure

Consider the system model described in Section 3.1. The SINR of the kth user can be calculated as

$$\text{SINR}_k = \frac{|\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}|^2}{\sum_{l=1,\, l\neq k}^{K}|\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,l)}|^2 + \sigma_k^2}. \tag{3.14}$$

Let the uncertainty model be defined as in (3.9). We can now formulate a robust counterpart of the standard SINR-constrained downlink beamforming problem

$$\begin{aligned} \min_{\boldsymbol{G}} \quad & P \\ \text{subject to} \quad & \text{SINR}_k \geq \gamma_k, \quad \forall \boldsymbol{\Delta}_{(k,:)} : \left\|\boldsymbol{\Delta}_{(k,:)}\right\|_2 \leq \varepsilon_k, \quad k = 1, \ldots, K \end{aligned} \tag{3.15}$$

where γ_k are the required QoS targets, expressed in terms of minimum tolerable SINRs.

The problem (3.15) is an involved, non-convex problem. This can be concluded by rewriting its constraints into a form

$$\left\|\begin{bmatrix} \boldsymbol{H}_{(k,:)}\boldsymbol{G} & \sigma_k \end{bmatrix}\right\|_2 \leq \sqrt{1+\gamma_k^{-1}}\,\left|\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}\right|, \quad \forall \boldsymbol{\Delta}_{(k,:)} : \left\|\boldsymbol{\Delta}_{(k,:)}\right\|_2 \leq \varepsilon_k \tag{3.16}$$

which appears to have no tractable representation due to the existence of the uncertainty on both sides and the absolute value. Therefore, the following problem of interest is formulated

$$\begin{aligned} \min_{\tilde{\boldsymbol{G}}} \quad & \frac{1}{2}\left\|\tilde{\boldsymbol{G}}\right\|_F^2 \\ \text{subject to} \quad & \left\|\begin{bmatrix} \tilde{\boldsymbol{H}}_{(k,:)}\tilde{\boldsymbol{G}} & \sigma_k \end{bmatrix}\right\|_2 \leq \sqrt{1+\gamma_k^{-1}}\,\tilde{\boldsymbol{H}}_{(k,:)}\tilde{\boldsymbol{G}}_{(:,k)}, \\ & \forall \tilde{\boldsymbol{\Delta}}_{(k,:)} : \left\|\tilde{\boldsymbol{\Delta}}_{(k,:)}\right\|_2 \leq \varepsilon_k, \quad k = 1, \ldots, K \end{aligned} \tag{3.17}$$

where

$$\tilde{\boldsymbol{H}}_{(k,:)} = \begin{bmatrix} \Re\{\boldsymbol{H}_{(k,:)}\} & \Im\{\boldsymbol{H}_{(k,:)}\} \end{bmatrix}, \quad \tilde{\boldsymbol{\Delta}}_{(k,:)} = \begin{bmatrix} \Re\{\boldsymbol{\Delta}_{(k,:)}\} & \Im\{\boldsymbol{\Delta}_{(k,:)}\} \end{bmatrix}, \quad \tilde{\boldsymbol{G}} = \begin{bmatrix} \Re\{\boldsymbol{G}\} & \Im\{\boldsymbol{G}\} \\ -\Im\{\boldsymbol{G}\} & \Re\{\boldsymbol{G}\} \end{bmatrix}. \tag{3.18}$$

It is proved in [BO02, WES06] that non-robust variations of (3.15) and (3.17) are equivalent. The idea was in showing that the term $\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}$ in the numerator of (3.14) can be real and positive, w.l.o.g.. This removes the absolute value in the right hand side of (3.16), yields SOC constraints, and, consequently, a conic quadratic optimization problem, whose real-valued version is exactly (3.17) with $\varepsilon_k = 0$. Unfortunately, the same principle is not applicable to the robust designs, due to the fact that an infinite number of channels must be supported with a single filter. However, the fulfillment of uncertain constraints in (3.17) clearly implies that the conditions (3.16) are also valid. Therefore, (3.17) can be viewed as a conservative approximation of (3.15). In other words, a filter obtained by solving (3.17) will surely be a feasible solution for (3.15), as well. However, the feasible set of (3.17) is smaller than in (3.15), so it is more likely to have infeasible scenarios. Consequently, the resulting transmit power from solving (3.17) is an upper bound for the optimal solution of (3.15). Finally, it should be noted that in the SINR-constrained problem formulations, no particular requirement on the CSI at the receiver is necessary, since the users' equalization plays no role in (3.14).

The robust problem (3.17) was posed originally in [BD07a], where three further restrictions of (3.17) are optimally solved using tractable SDP reformulations. Notice that due to the existence of uncertainty on both sides in the constraints in (3.17), rewriting these constraints as LMIs would not yield a suitable structure needed for the application of Lemma 7, as observed also in [BD07a, BTBG⁺]. In the sequel, it will be shown that the problem (3.17) can actually be solved optimally, by applying the ellipsoid method from convex optimization theory. Due to the fact that the ellipsoid method exhibits high (though still polynomial [BTN01]) complexity, an efficient SDP-based suboptimal solution for the same problem will also be derived. The latter method exhibits better properties in terms of both performance and complexity, comparing with the related results from the literature. Furthermore, numerical simulations will reveal that this SDP-based conservative solution seems to perform as well as the theoretically optimal benchmark provided by the ellipsoid method.

3.3.1 Solution by Ellipsoid Method

The ellipsoid method presents a multidimensional extension of the bisection approach. The main idea behind the algorithm can be formulated as follows [BTN01]. At a given iteration one is equipped with an ellipsoid which surely contains an optimal solution. Depending on whether the center of the ellipsoid is in the problem domain or not, there are two possibilities:

- If the center of the ellipsoid is in the problem domain, a subgradient of the objective function in the center of the ellipsoid should be calculated. Zero value of the subgradient means that an optimal solution is found, otherwise, the subgradient defines a hyperplane and a half-ellipsoid which contains the optimal solution. In the latter case, the new search ellipsoid is the minimal volume ellipsoidal container of the above-described half-ellipsoid.

- If the center of the ellipsoid is not in the problem domain, the hyperplane that separates the center from the problem domain can be constructed. In this way, a new half-ellipsoid is obtained, whose minimal volume ellipsoidal container is used for the subsequent search.

The procedure yields a sequence of shrinking ellipsoids which converges to an optimal solution.

The ellipsoid method is recalled in Algorithm 1, as described in [BTN01], omitting technical details regarding the initialization and the formalization of the convergence criterion. The unknown vector to be calculated, which uniquely determines the transmit filter $\boldsymbol{G}$, is denoted by

$$\boldsymbol{g} = \begin{bmatrix} \tilde{\boldsymbol{G}}^T_{(:,1)} & \cdots & \tilde{\boldsymbol{G}}^T_{(:,K)} \end{bmatrix}^T \in \mathbb{R}^d \tag{3.19}$$

where $d = 2MK$. Notice that in Algorithm 1, $\bar{\boldsymbol{g}}_t$ is the sequence of the search ellipsoids' centers, while in each iteration the vector $\boldsymbol{a}_t$ defines the hyperplane that divides the current search ellipsoid.

The main conclusion of this section is formulated as a theorem:

Algorithm 1 Algorithmic solution for the robust, SINR-constrained precoder.

Initialize the first search ellipsoid $\{\bar{\boldsymbol{g}}_0 + \boldsymbol{B}_0\boldsymbol{u}, \|\boldsymbol{u}\|_2 \leq 1\}$. Set $t = 0$ (step number).
repeat
  $t \leftarrow t + 1$.
  if $\bar{\boldsymbol{g}}_{t-1} \in \mathcal{S}$, where $\mathcal{S}$ is the domain of (3.17): Calculate the vector $\boldsymbol{a}_t$ as the subgradient of the objective function in $\bar{\boldsymbol{g}}_{t-1}$. If $\boldsymbol{a}_t = \boldsymbol{0}$ the optimal solution is $\bar{\boldsymbol{g}}_{t-1}$. Skip the following step.
  if $\bar{\boldsymbol{g}}_{t-1} \notin \mathcal{S}$: Calculate the separation vector $\boldsymbol{a}_t$ that satisfies

$$\boldsymbol{a}_t^T \bar{\boldsymbol{g}}_{t-1} \geq \sup_{\boldsymbol{g}\in\mathcal{S}} \boldsymbol{a}_t^T \boldsymbol{g}, \quad \boldsymbol{a}_t \neq \boldsymbol{0}. \tag{3.20}$$

  Update the search ellipsoid:

$$\boldsymbol{p}_t = \frac{\boldsymbol{B}_{t-1}^T \boldsymbol{a}_t}{\sqrt{\boldsymbol{a}_t^T \boldsymbol{B}_{t-1} \boldsymbol{B}_{t-1}^T \boldsymbol{a}_t}}, \quad \bar{\boldsymbol{g}}_t = \bar{\boldsymbol{g}}_{t-1} - \frac{1}{d+1} \boldsymbol{B}_{t-1} \boldsymbol{p}_t \tag{3.21}$$

$$\boldsymbol{B}_t = \frac{d}{\sqrt{d^2-1}} \boldsymbol{B}_{t-1} + \left(\frac{d}{d+1} - \frac{d}{\sqrt{d^2-1}} \right) \boldsymbol{B}_{t-1} \boldsymbol{p}_t \boldsymbol{p}_t^T \tag{3.22}$$

until : convergence is reached

Theorem 3. The sequence $\bar{\boldsymbol{g}}_t$ from Algorithm 1 converges to the optimal solution of the problem (3.17). The feasibility check for $\bar{\boldsymbol{g}}_{t-1}$ and the construction of $\boldsymbol{a}_t$ can be efficiently performed, and the complexity of Algorithm 1 is polynomial.

Proof. The proof is given in Section 3.5. □

3.3.2 SDP-Based Solution

In this section, the problem (3.17) is approached indirectly. Define a virtual MSE-optimization problem (3.5), with the same channels and the uncertainty regions as in (3.15), and with QoS targets

$$\mu_k = \frac{1}{1+\gamma_k}, \quad k = 1, \ldots, K. \tag{3.23}$$

It is known that if perfect CSI is provided, such MSE-constrained problem is equivalent to (3.17) and (3.15), as far as the transmit filter is concerned [SS05]. The following theorem shows that in the robust case, the virtual MSE-optimization problem can serve as a conservative, tractable approximation of (3.17) and consequently (3.15).

Theorem 4. Let the MSE problem (3.5) be feasible with the targets $\mu_k \in (0,1)$, $k = 1,\ldots,K$, and let $\boldsymbol{G}_{\mathrm{opt}}$ be the resulting optimal transmit filter. The SINR constraints in (3.17) are satisfied then for γ_k obtained from (3.23), with the same transmit filter $\boldsymbol{G}_{\mathrm{opt}}$.

Proof. It can be seen that the condition $\mathrm{MSE}_k \le \mu_k$ is equivalent to

$$\sum_{l=1,\, l\neq k}^{K} \left| \frac{1}{f_k} \boldsymbol{H}_{(k,:)} \boldsymbol{G}_{(:,l)} \right|^2 + \left| \frac{1}{f_k} \boldsymbol{H}_{(k,:)} \boldsymbol{G}_{(:,k)} - 1 \right|^2 + \frac{\sigma_k^2}{f_k^2} \le \mu_k. \tag{3.24}$$

Since $f_k > 0$ and $\mu_k \in (0,1)$, it follows that $\Re\{\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}\} \ge 0$, because otherwise the absolute value of $\Re\left\{\frac{1}{f_k}\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)} - 1\right\}$ would be larger than 1. Therefore, it is sufficient to prove that $\mathrm{MSE}_k \le \mu_k$ implies

$$\left\| \begin{bmatrix} \boldsymbol{H}_{(k,:)}\boldsymbol{G} & \sigma_k \end{bmatrix} \right\|_2^2 \le \frac{1}{1-\mu_k} \Big(\Re\{\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}\} \Big)^2 \tag{3.25}$$

since $1 + \gamma_k^{-1} = (1-\mu_k)^{-1}$, $\Re\{\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}\} = \tilde{\boldsymbol{H}}_{(k,:)}\tilde{\boldsymbol{G}}_{(:,k)}$, and the Euclidean norms of vectors $[\boldsymbol{H}_{(k,:)}\boldsymbol{G},\, \sigma_k]$ and $\left[\tilde{\boldsymbol{H}}_{(k,:)}\tilde{\boldsymbol{G}},\, \sigma_k\right]$ are equal, so the SINR constraints in (3.17) would be fulfilled. The condition (3.24) can be rewritten into an equivalent form

$$\left\| \begin{bmatrix} \boldsymbol{H}_{(k,:)}\boldsymbol{G} & \sigma_k \end{bmatrix} \right\|_2^2 \le f_k^2 \mu_k + \left| \boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)} \right|^2 - \left| \boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)} - f_k \right|^2. \tag{3.26}$$

The theorem is proved by noticing that

$$f_k^2 \mu_k + \left| \boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)} \right|^2 - \left| \boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)} - f_k \right|^2 \le \frac{1}{1-\mu_k} \Big(\Re\{\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}\} \Big)^2 \tag{3.27}$$

is always true, because it is equivalent to

$$\Big(\Re\{\boldsymbol{H}_{(k,:)}\boldsymbol{G}_{(:,k)}\} - f_k(1-\mu_k) \Big)^2 \ge 0. \tag{3.28}$$

□

The problem (3.17) can now be approached in the following way: For a fixed set of targets $\mathrm{SINR}_k \ge \gamma_k$, the virtual MSE constraints $\mathrm{MSE} \le \mu_k$ are defined using (3.23). The problem (3.5) is solved then by applying the SDP method from Section 3.2, and the resulting transmit filter surely complies to the specified SINR requirements γ_k. The excellent quality of this approach with respect to the optimal solution from Section 3.3.1 will be seen from simulation results in Section 3.4.

While it has been shown how an optimal solution of (3.17) can be computed by employing the ellipsoid method, an optimal solution of the SINR-constrained problem (3.15) remains as an open topic for the future work. It should be added that if the users can estimate the equivalent channels perfectly, it can be easily proved that the application of the standard MMSE solution (3.6) for the receiver's coefficients in the MSE-constrained problem (3.5), with the substitution $\mu_k = 1/(1 + \gamma_k)$, would lead to the same non-convex problem (3.15).

3.4 Numerical Examples

Except for the ellipsoid method from Section 3.3.1 which is implemented directly, numerical simulations in this section are obtained obtained using SeDuMi [Stu99, Löf04].

First, a simulation result, that compares the robust with the non-robust solution of the MSE-optimization problem (3.5) in a 3-user system, is given. The noise variance, the MSE targets and the bounds on the uncertainty (3.9) are assumed to be the same for all users: $\sigma_k^2 = 0.1$, $\mu_k = 0.15$ and $\varepsilon_k = 0.1$. The BS is equipped with $M = 4$ antennas. The distributions of the third user's MSE, if the non-robust solution from [SS05] and the robust solution from Section 3.2 are implemented, are plotted in Fig. 3.2 for 1000 randomly chosen channel realizations which yielded feasible problems for the specified targets. The real and imaginary parts of each channel coefficient had normal distribution with zero mean and variance 0.5, and the errors were uniformly chosen from the uncertainty regions (3.9). It can be concluded that if the transceiver optimization is performed disregarding the disturbances, the system does not satisfy the QoS constraint for more than 50% of the channel realizations. Contrary to this, the robust algorithm proposed in Section 3.2 guarantees the targets all the time, and the system never experiences an outage.

For introducing robustness in the system, clearly, a certain price had to be paid. In the problem formulation (3.5), this would correspond to the increased transmit power with respect to the situation when the obtained channel knowledge was perfect. In Fig. 3.3, the performance of the solution for the MSE-optimization problem from Section 3.2 is compared with three

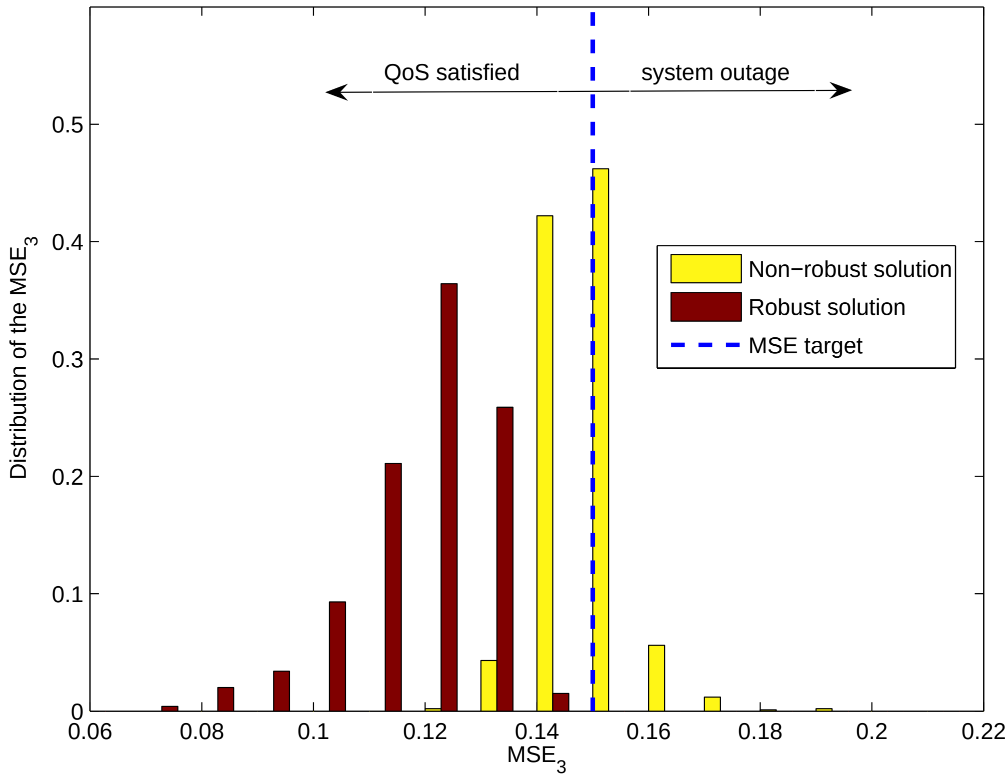

Figure 3.2: Comparison of robust (Section 3.2) and non-robust [SS05] solutions in an MSE-constrained multiuser MISO system. MSE_3 is in linear scale.

strategies. The first one is the robust power allocation proposed in [PPIL07]. The result of optimization from Section 3.2, if the receivers are fixed with $f_k = 1$, $k \in \{1, 2, 3\}$, is plotted, as well. Finally, the performance of a hypothetical system, where the obtained CSI is correct, is also given. The minimum transmit power is plotted for various MSE targets μ_3 of the third user, with other simulation parameters being the same as in the previous example. The results present an average over 1000 channels, chosen so that the problem was feasible for the observed range of QoS constraints. It can be seen that the proposed SDP-based solution from Section 3.2 outperforms in performance other robust strategies. The SDP-based robustness, which enabled the fulfillment of the targets, required roughly 25-35% more power comparing with the case when the obtained channel knowledge was exact. It should be remarked that the performance gaps between the investigated strategies vary depending on the system size, the uncertainty level and the signal-to-noise ratio (noise power) in the system.

Due to the fact that the uncertain MSE constraints have the equivalent representations with

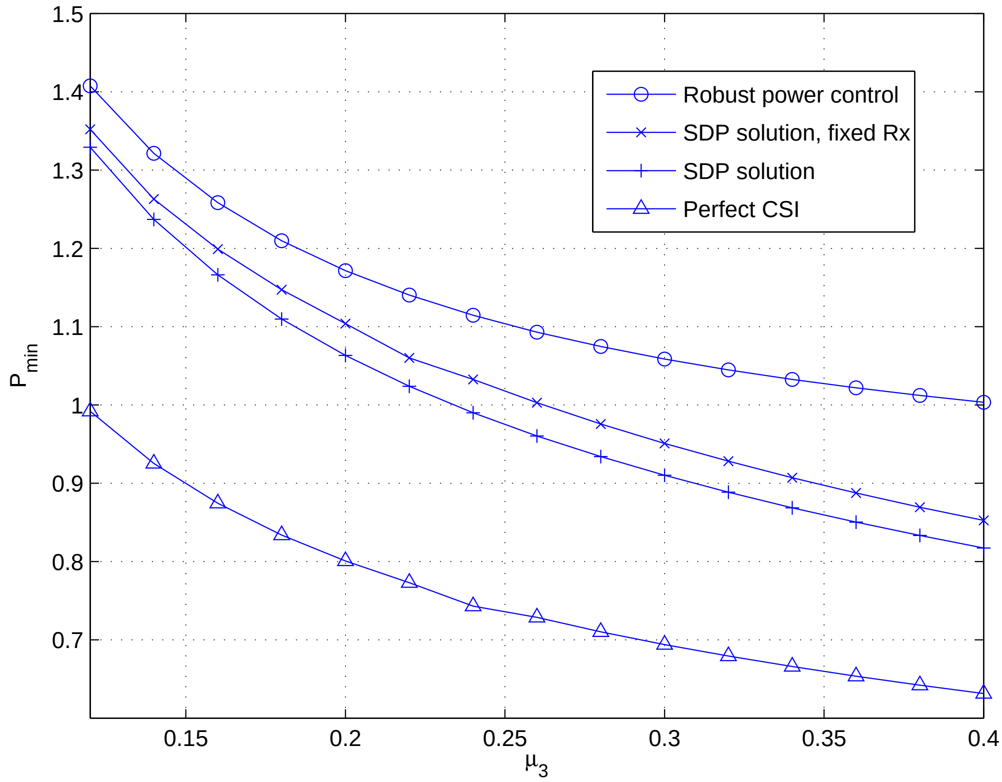

Figure 3.3: Minimal transmit power versus the QoS (MSE) target of the third user (linear scales), $\sigma^2_{1,2,3} = 0.1$, $\mu_{1,2} = 0.15$.

σ_k^2	Perfect CSI	SDP (Sec. 3.2)	Power Control [PPIL07]	SDP, $f_k = 1$
0.01	0.0011	0.0083	0.0112	0.0202
0.1	0.0056	0.0169	0.0208	0.1213

Table 3.1: Minimum feasible MSE target μ_3 in a 3-user MISO system with $\mu_{1,2} = 0.15$.

LMIs, the feasibility range of individual MSE constraints μ_k can be calculated by the bisection method [BV04]. Table 3.1 shows the average feasibility ranges of the MSE target of the third user μ_3, with $\mu_{1,2} = 0.15$ fixed, for two noise variances. Other system parameters are the same as in the first simulation example. Additionally, a total power constraint $P \leq 10$ is imposed, which is introduced into the optimization framework in a straightforward manner.

Similarly to the previous example, it can be concluded that the maximum size of the uncertainty region ε_k is also readily obtained by the bisection method. This parameter is of importance in determining the roughness of quantization of CSI, e.g., when returning it from the receivers to the BS using the feedback channels. The averaged results with the same system parameters as for the example in Fig. 3.3, with the additional power constraint $P \leq 10$, and

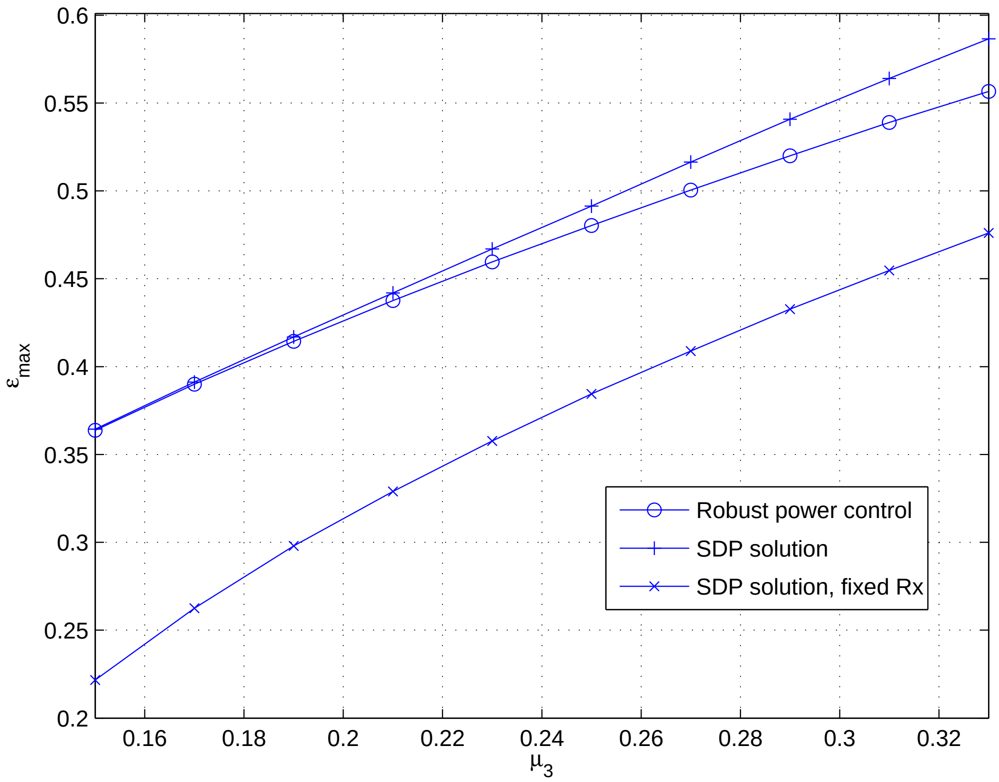

Figure 3.4: Maximum size of the uncertainty region ε_k, assumed to be the same for all users, for the MSE-constrained problem.

with the bounds ε_k assumed to be the same for all users, are given in Fig. 3.4. The results from Table 3.1 and Fig. 3.4 confirm that the proposed SDP-based approach exhibits better properties comparing with the other two robust methods also in terms of the feasibility regions and the sustained level of uncertainty.

In Fig. 3.5, the Monte Carlo simulations for the minimum transmit power against the SINR targets are plotted, with the targets assumed to be the same for all users. The system parameters are $M = K = 3$, $\sigma^2_{1,2,3} = 0.1$, and $\varepsilon_{1,2,3} = 0.1$. The results present an average over 2000 channel realizations that yielded feasible problems for the considered range of SINR targets. It can be seen that the solution based on the ellipsoid method and the SDP-based method from Section 3.3 perform the same. They both outperform the structured RSDP approach from [BD07a]. Furthermore, the complexity of the proposed SDP-based method appears to be smaller than the complexity of the approach in [BD07a], which employs larger LMIs and more unknown variables. In Section 3.4.1, a brief theoretical analysis of the complexity of all investigated

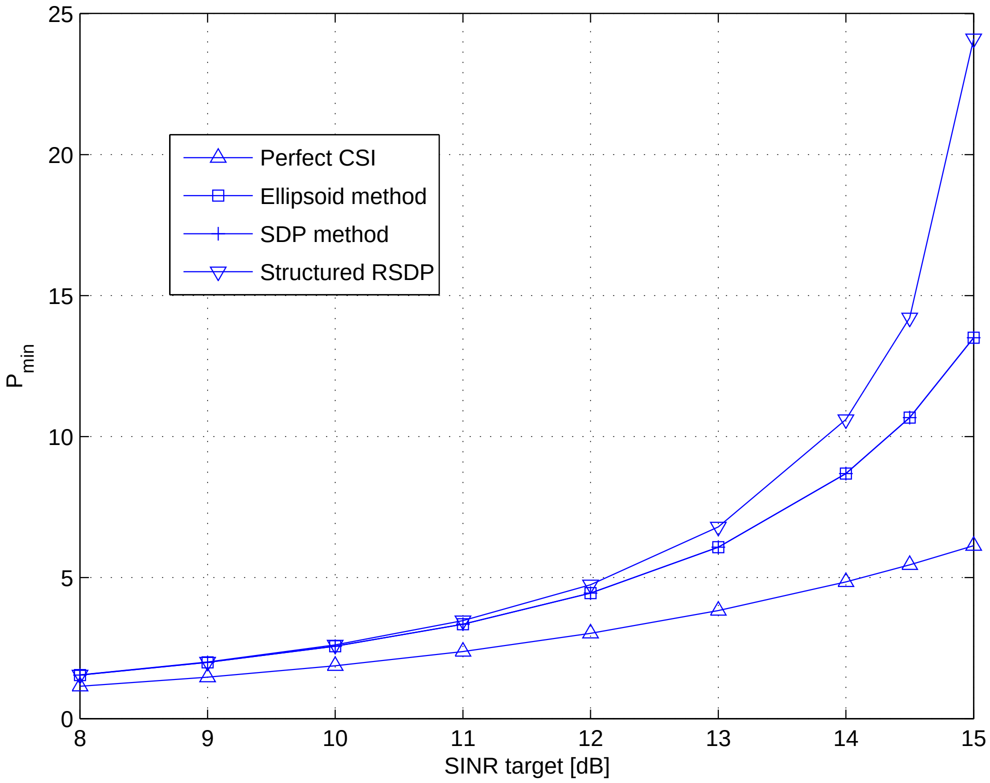

Figure 3.5: Minimal transmit power in linear scale versus the SINR target in dB (assumed to be the same for all users) for a system with $\sigma^2_{1,2,3} = 0.1$ and $\varepsilon_{1,2,3} = 0.1$.

methods is presented.

The analysis of the SINR-constrained problem is finished by plotting in Fig. 3.6 the percentage of feasible channel realizations in a system with the same parameters as in the previous simulation. The SINR target of 10 dB is assumed to be the same for all three users. 2000 randomly generated multiuser channels, with the complex Gaussian distribution of the channel coefficients remaining the same as for the first simulation result in this section, were inspected. The conclusions regarding the performance ordering are similar as in the power minimization case.

3.4.1 Notes on Computational Complexity

Except for the ellipsoid method from Section 3.3.1, that exhibits a significantly larger computational burden [BTN01], all studied methods in this paper and in the relevant results from the literature applied the SDP techniques, or they can be easily transformed into them for the sake

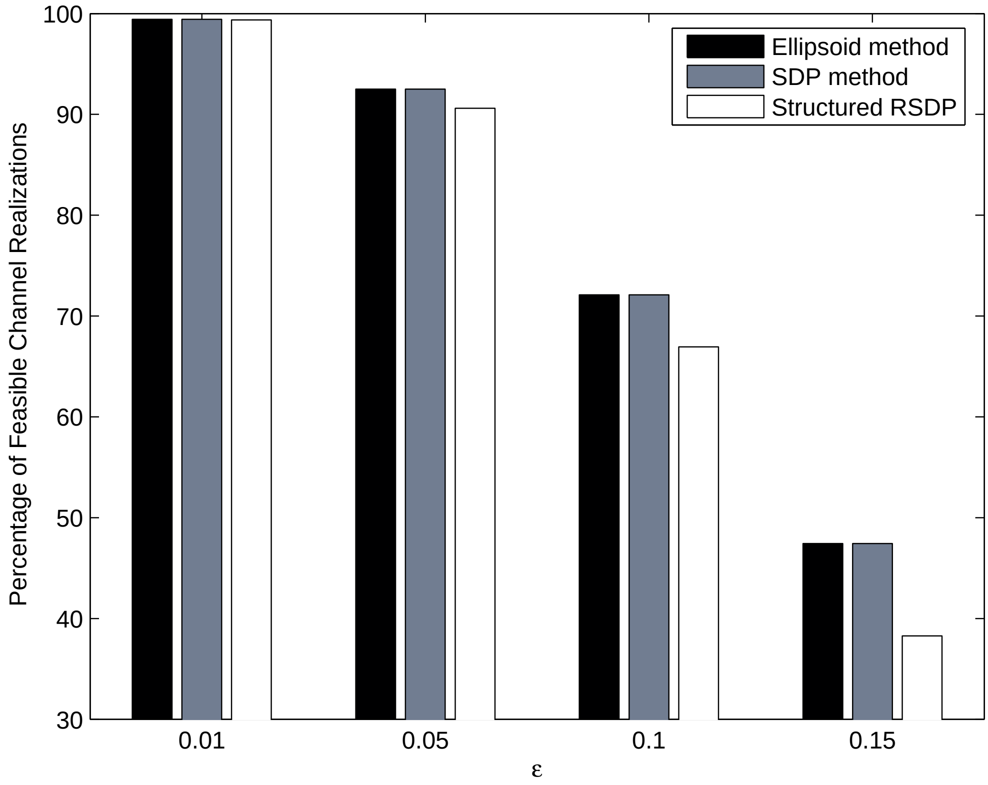

Figure 3.6: Percentage of feasible channel realizations in a system with $M = K = 3$, $\sigma^2_{1,2,3} = 0.1$, and $\gamma_{1,2,3} = 10\,\text{dB}$.

of a rough comparison. A good insight for the complexity can be gained already from studying the general bounds, given in Section 2.1.3. As seen from (2.27), the key influence on the arithmetic complexity is made by the size of the vector that contains the unknown variables N, and by the sizes of the LMIs K_l. These parameters are used to compare the examined strategies. The standard transformation of the objective function into a minimization of a linear function with slack variables and an SOC/LMI constraint, will be neglected, since this has a minor influence on the performance.

The SDP algorithm derived in Section 3.2 has $2MK + K$ original unknown variables (real and imaginary parts of the transceiver coefficients) and K additional slack variables λ_k. There are K LMIs in (3.12) corresponding to the users' constraints, with sizes $2(K + M + 2)$ in the real-valued SDP representation (2.8). This also determines the complexity of the conservative SINR-optimization from Section 3.3.2, with the users' equalization coefficients treated as slack variables.

The solution from [PPIL07] is simpler firstly in the sense that only power allocation, i.e., only K real-valued coefficients $q_1, \ldots, q_K$ are optimized. These coefficients are tightly connected with the equalization scaling factors $f_1, \ldots, f_K$ from Fig. 3.1. The constraints in [PPIL07] are originally given in the form

$$\lambda_{\max}\left(\hat{\boldsymbol{H}}^{\dagger}\mathrm{diag}(q_1, \ldots, q_K)\hat{\boldsymbol{H}}^{\dagger^*}\right) \leq c_{1,k}q_k + c_{2,k} \tag{3.29}$$

where $c_{1,k}$ and $c_{2,k}$ are constants depending on the MSE targets, the noise variances and the sizes of the uncertainty regions. Using Lemma 3, the condition (3.29) can be rewritten as $(c_{1,k}q_k + c_{2,k})\boldsymbol{I}_M - \hat{\boldsymbol{H}}^{\dagger}\mathrm{diag}(q_1, \ldots, q_K)\hat{\boldsymbol{H}}^{\dagger^*} \geq \boldsymbol{0}$, which is an LMI in $q_1, \ldots, q_K$.[†] Therefore, the size of the LMIs in the real-valued representation is $2M$ which is smaller comparing to the proposed solution in Section 3.2. However, as seen from Fig. 3.3, Fig. 3.4, and Table 3.1, the gains introduced by the proposed full optimization can be significant.

As far as the SINR-constrained downlink transmission is concerned, in [BD07a] the size of the LMI and the number of additional slack variables corresponding to the kth user constraint are $2(K + 1)(2M + 1)$ and $(K + 1)(2K + 3)$, respectively, which is larger comparing to the same parameters in the solution based on the virtual MSE-optimization problem, that is also seen to exhibit better performance.

3.5 Appendix: Proof of Theorem 3

In this section, it will be shown that all ingredients for Algorithm 1 are indeed available.

The ability to construct the subgradient of the objective function is not problematic. Clearly, $P = \|\boldsymbol{g}\|_2^2$ is valid, with $\boldsymbol{g}$ defined by (3.19). This function is differentiable, so its subgradient is equal to the usual gradient: $\partial P = 2\boldsymbol{g}$.

The problems of determining whether a particular vector $\boldsymbol{g} = \bar{\boldsymbol{g}}$ is feasible for (3.17) and the construction of the separation oracle, defined by (3.20), are considerably more involved. Notice that the domain $\mathcal{S}$ of (3.17) is an intersection of sets $\mathcal{S}_k$, $k = 1, \ldots, K$, where $\mathcal{S}_k$ is the feasible

[†] It should be remarked that in special cases of the problems studied in [PPIL07], there are possibilities for employing more efficient iterative algorithms than SDP.

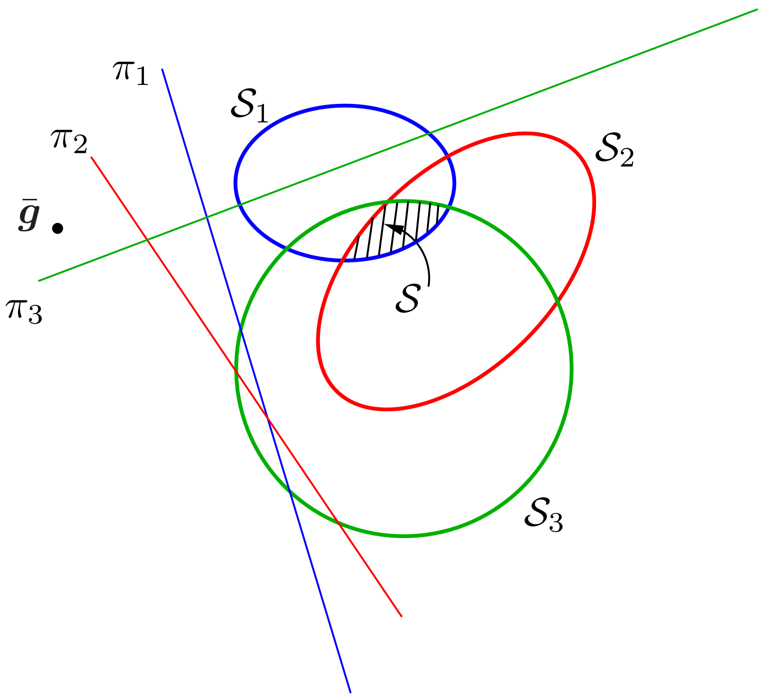

Figure 3.7: The separation hyperplanes π_k, defined by (3.20), and the domains $\mathcal{S}_k$, that correspond to users' constraints in (3.17).

region corresponding to the kth user's target in (3.17). By using similar reasoning as in the proof of Theorem 1, we conclude that $\mathcal{S}_k$ and consequently $\mathcal{S}$ are convex sets. Let us construct the procedure for determining the feasibility of a particular vector $\boldsymbol{g} = \bar{\boldsymbol{g}}$ and the separation hyperplane π_k according to (3.20), with respect to the kth user constraint and the domain $\mathcal{S}_k$ only. Clearly, as illustrated in Fig. 3.7 for a 3-user system, the global feasibility checking routine would consist of examining each of the user constraints. If, e.g., the kth constraint is declared infeasible, the hyperplane π_k separates $\bar{\boldsymbol{g}}$ and $\mathcal{S}$, as well. $\bar{\boldsymbol{g}}$ is feasible, if and only if each of the K constraints is fulfilled.

It can be noticed that the kth user's constraint in (3.17) can be rewritten as

$$\left\|\boldsymbol{\Psi}_k(\boldsymbol{g})\boldsymbol{\delta}_k + \boldsymbol{\psi}_k(\boldsymbol{g})\right\|_2 \leq \boldsymbol{\alpha}_k^T(\boldsymbol{g})\boldsymbol{\delta}_k + \beta_k(\boldsymbol{g}), \quad \forall \boldsymbol{\delta}_k : \|\boldsymbol{\delta}_k\|_2 \leq 1 \tag{3.30}$$

where $\boldsymbol{\delta}_k = (1/\varepsilon_k)\tilde{\boldsymbol{\Delta}}^T_{(k,:)}$, and $\boldsymbol{\Psi}_k$, $\boldsymbol{\psi}_k$, $\boldsymbol{\alpha}_k$ and β_k are affine in the unknown transmit filter coefficients grouped in the vector $\boldsymbol{g}$. For the constraints of type (3.30), robust optimization theory, developed in [BTBG$^+$], can be applied, in order to examine the feasibility and determine the separation hyperplane. In the rest of this section, the main ideas of the procedure are presented.

The user index k and the explicit dependence of the terms $\boldsymbol{\Psi}_k$, $\boldsymbol{\psi}_k$, $\boldsymbol{\alpha}_k$ and β_k on $\boldsymbol{g}$ will be occasionally omitted in rest of this section to make the formulas less cumbersome.

The feasibility problem is resolved by the following lemma:

Lemma 8. The uncertain constraint (3.30) is valid if and only if the following two inequalities are fulfilled

$$\begin{aligned} \|\boldsymbol{\alpha}\|_2 &\leq \beta && (3.31) \\ \boldsymbol{X}(\lambda) &\triangleq \begin{bmatrix} \lambda \boldsymbol{I} + \boldsymbol{\alpha}\boldsymbol{\alpha}^T - \boldsymbol{\Psi}^T\boldsymbol{\Psi} & \beta\boldsymbol{\alpha} - \boldsymbol{\Psi}^T\boldsymbol{\psi} \\ \beta\boldsymbol{\alpha}^T - \boldsymbol{\psi}^T\boldsymbol{\Psi} & \beta^2 - \boldsymbol{\psi}^T\boldsymbol{\psi} - \lambda \end{bmatrix} \succeq \boldsymbol{0} && (3.32) \end{aligned}$$

for some $\lambda \geq 0$. The feasibility of (3.31)-(3.32) for a fixed $\boldsymbol{g}$ can be determined in an efficient manner. In the case when (3.32) is infeasible, a finite set of vectors $\boldsymbol{z}_l$ with the corresponding affine functions $\phi_l(\lambda) = \boldsymbol{z}_l^T \boldsymbol{X}(\lambda)\boldsymbol{z}_l$, $l = 0, \ldots, L$ so that

$$\phi(\lambda) = \min\{\phi_0(\lambda), \ldots, \phi_L(\lambda)\} < 0, \quad \forall \lambda \geq 0 \tag{3.33}$$

where min means the pointwise minimum, is readily found.

Proof. It can be noticed that (3.30) requires

$$\boldsymbol{\alpha}^T(\boldsymbol{g})\boldsymbol{\delta} + \beta(\boldsymbol{g}) \geq 0, \quad \forall \boldsymbol{\delta} : \|\boldsymbol{\delta}\|_2 \leq 1. \tag{3.34}$$

By applying the Cauchy-Schwarz inequality (2.1) on (3.34), (3.31) is obtained. The inequalities in (3.30) can be squared now, and a direct use of the $\mathcal{S}$-Lemma (Lemma 6) renders the equivalent form (3.32).

It is straightforward to examine the validity of (3.31) for a given vector $\boldsymbol{g} = \bar{\boldsymbol{g}}$, so in the sequel, the accent is on analyzing (3.32). For a fixed $\boldsymbol{g}$, $\boldsymbol{X}(\lambda)$ is an LMI in λ. Notice that $\lambda_- \leq \lambda \leq \lambda_+$, where $\lambda_- = 0$, $\lambda_+ = \beta^2 - \boldsymbol{\psi}^T\boldsymbol{\psi}$, must be true, for (3.32) to hold. In other words, $[\lambda_-, \lambda_+]$ can be considered as a starting search interval for λ, so that the LMI (3.32) is valid. Therefore, for $\lambda_+ < 0$, it can be immediately concluded that (3.32) and consequently (3.30) are not feasible. If $\lambda_+ > 0$, check whether $\boldsymbol{X}(\lambda_l) \succeq \boldsymbol{0}$, with $\lambda_l = 0.5(\lambda_- + \lambda_+)$. If $\boldsymbol{X}(\lambda_l)$ is positive semidefinite, the required λ is λ_l, and the feasibility of (3.32) can be claimed. If $\boldsymbol{X}(\lambda_l)$ is not positive semidefinite, by examining the eigenvalues of $\boldsymbol{X}(\lambda_l)$, a vector $\boldsymbol{z}_l$ can be found so that the affine function $\phi_l(\lambda) = \boldsymbol{z}_l^T \boldsymbol{X}(\lambda)\boldsymbol{z}_l$ is negative for $\lambda = \lambda_l$. A new interval $[\lambda_-, \lambda_+]$ to search for λ satisfying $\boldsymbol{X}(\lambda) \geq \boldsymbol{0}$ is obtained as a set of values for which $\phi_l(\lambda) \geq 0$, and the procedure can be continued in an iterative manner. The new search interval is at least two times smaller than the original search space, which guarantees a rapid convergence. Finally, if no $\lambda \geq 0$ is found so that $\boldsymbol{X}(\lambda) \geq \boldsymbol{0}$, (3.32) is declared infeasible, with a set of functions $\phi_l(\lambda)$, that are affine in λ and satisfy (3.33), obtained. □

The infeasibility certificate of (3.30) is going to be built now. More precisely, when (3.30) is infeasible, one vector $\bar{\boldsymbol{\delta}}$, with $\left\|\bar{\boldsymbol{\delta}}\right\|_2 \leq 1$, can be found so that the first inequality in (3.30) is not satisfied for $\boldsymbol{\delta} = \bar{\boldsymbol{\delta}}$. Using the infeasibility certificate vector, the separating hyperplane (3.20) will be constructed.

In the case when (3.31) is infeasible, the required vector $\bar{\boldsymbol{\delta}}$ is easily obtained as $\bar{\boldsymbol{\delta}} = -\boldsymbol{\alpha}/\|\boldsymbol{\alpha}\|_2$. Therefore, the focus is put on the case when the infeasibility of (3.31) and (3.32) is due to (3.32).

It can be noticed that the obtained $\phi(\lambda)$ in (3.33) is a piecewise linear concave function for $\lambda \geq 0$. Observe its maximum, and conclude that $x \in (0, 1)$ and l_1, l_2, with $0 \leq l_1, l_2 \leq L$, can be found so that (the procedure should be slightly modified if one affine function is the minimizer of ϕ for all $\lambda \geq 0$)

$$\varphi(\lambda) = x\phi_{l_1}(\lambda) + (1 - x)\phi_{l_2}(\lambda) < 0, \quad \forall\lambda \geq 0. \tag{3.35}$$

Set $\boldsymbol{\mu} = \sqrt{x}\,\boldsymbol{z}_{l_1}$, $\boldsymbol{\zeta} = \sqrt{1-x}\,\boldsymbol{z}_{l_2}$, and $\boldsymbol{Z} = \boldsymbol{\mu}\boldsymbol{\mu}^T + \boldsymbol{\zeta}\boldsymbol{\zeta}^T$. It follows that

$$\varphi(\lambda) = \mathrm{Tr}\{\boldsymbol{X}(\lambda)\boldsymbol{Z}\} < 0, \quad \forall\lambda \geq 0. \tag{3.36}$$

By using basic linear algebra and theory of quadratic functions, it can be shown that a transformation $\boldsymbol{Z} = \boldsymbol{\eta}\boldsymbol{\eta}^T + \boldsymbol{\xi}\boldsymbol{\xi}^T$, where $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are elements of the second-order cone $\mathcal{L}^{2M+1}$, can be calculated [BTBG+]. From (3.36), one concludes that $\mathrm{Tr}\{\boldsymbol{Z}\boldsymbol{X}(0)\} < 0$. It follows that either $\mathrm{Tr}\{\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{X}(0)\} < 0$, or $\mathrm{Tr}\{\boldsymbol{\xi}\boldsymbol{\xi}^T\boldsymbol{X}(0)\} < 0$, or both. If the first inequality is true, the required infeasibility certificate $\bar{\boldsymbol{\delta}}$ is obtained by scaling $\boldsymbol{\eta}$ with its last element:

$$\begin{bmatrix} \bar{\boldsymbol{\delta}} \\ 1 \end{bmatrix} = \frac{1}{\eta_{2M+1}} \boldsymbol{\eta}. \tag{3.37}$$

The proof follows from the facts that

$$\begin{bmatrix} \bar{\boldsymbol{\delta}}^T & 1 \end{bmatrix} \boldsymbol{X}(0) \begin{bmatrix} \bar{\boldsymbol{\delta}}^T & 1 \end{bmatrix}^T < 0 \tag{3.38}$$

implies that the first inequality in (3.30) is not valid, and that $\boldsymbol{\eta} \in \mathcal{L}^{2M+1}$ yields $\|\bar{\boldsymbol{\delta}}\|_2 \leq 1$. The analog procedure can be performed when $\mathrm{Tr}\{\boldsymbol{\xi}\boldsymbol{\xi}^T \boldsymbol{X}(0)\} < 0$.

Let

$$\bar{\boldsymbol{y}} = \begin{bmatrix} \left(\boldsymbol{\Psi}(\bar{\boldsymbol{g}})\bar{\boldsymbol{\delta}} + \boldsymbol{\psi}(\bar{\boldsymbol{g}})\right)^T & \boldsymbol{\alpha}^T(\bar{\boldsymbol{g}})\bar{\boldsymbol{\delta}} + \beta(\bar{\boldsymbol{g}}) \end{bmatrix}^T \notin \mathcal{L}^{2K+2} \tag{3.39}$$

and set

$$\boldsymbol{v} = \begin{bmatrix} \left(\boldsymbol{\Psi}(\bar{\boldsymbol{g}})\bar{\boldsymbol{\delta}} + \boldsymbol{\psi}(\bar{\boldsymbol{g}})\right)^T \left\|\boldsymbol{\Psi}(\bar{\boldsymbol{g}})\bar{\boldsymbol{\delta}} + \boldsymbol{\psi}(\bar{\boldsymbol{g}})\right\|_2^{-1} & -1 \end{bmatrix}^T. \tag{3.40}$$

It can be noticed that

$$\boldsymbol{v}^T \bar{\boldsymbol{y}} \geq \boldsymbol{v}^T \boldsymbol{y}, \quad \forall \boldsymbol{y} \in \mathcal{L}^{2K+2}. \tag{3.41}$$

The homogenous part of

$$\boldsymbol{v}^T \begin{bmatrix} (\boldsymbol{\Psi}(\boldsymbol{g})\bar{\boldsymbol{\delta}} + \boldsymbol{\psi}(\boldsymbol{g}))^T & \boldsymbol{\alpha}^T(\boldsymbol{g})\bar{\boldsymbol{\delta}} + \beta(\boldsymbol{g}) \end{bmatrix}^T \tag{3.42}$$

which is immediately computed because of the affine nature of the respective terms in $\boldsymbol{g}$, yields the required vector $\boldsymbol{a}_t$.

The fact that the described ellipsoid method converges to an optimal solution with a polynomial complexity follows directly from Lecture 5 in [BTN01].

Chapter 4

Worst-Case Optimization of Downlink Multiuser MIMO Systems

The transceiver design in a multiuser setup, with both the BS and users having multiple antennas, attracted a lot of attention recently, especially in the downlink transmission [WES06, MJHU06, KTA06, SSB07, SSB08a]. Contrary to the MISO setting, where optimal solutions for transceivers exist under the assumption of knowing the channel perfectly, most of the QoS-constrained transceiver optimization problems are in the multiuser MIMO setup involved even with perfect CSI. The main problem in this chapter is the derivation of a framework, that is able to directly support the following illustrative robust optimization problems:

- Minimization of the total transmit power subject to strict per-user (or per-stream) MSE constraints that must be satisfied despite the uncertainty.
- Minimization of the worst-case sum MSE of all users with respect to the uncertainty, under a total transmit power constraint.
- Minimization of the maximum per-user (or per-stream) MSE taking into account the uncertainty, under a total transmit power constraint.

The solutions are obtained using iterative algorithms, where each iteration consists of two steps. In the first step, the transmit precoder is optimized with fixed receivers using SDP reformulations that will depend on the problem at hand. The second part of an iteration updates the receivers in a minimax fashion using the previously obtained solution for the transmitter. This step can also be rewritten as a semidefinite program, and the complete alternating algorithm converges (although not necessarily to a global optimum). There exist also numerous possibilities for extending the proposed framework. Uncertain noise covariance at the transmitter, nonlinear precoders, per-antenna power constraints (PAPCs), bounded rectangular and unbounded Gaussian CSI error models, and coordinated multi-cell (network) MIMO systems [KFV06] are readily incorporated.

The rest of this chapter is organized as follows. The multiuser MIMO system model and the main problems of interest are explained in Section 4.1. In Section 4.2, firstly a generic iterative framework for handling the stated problems is presented. Each of the problems is then separately discussed with respect to providing efficient numerical solutions. The algorithm for calculating the actual worst-case channels for fixed transceivers is derived in Section 4.3. Section 4.4 explains the above-mentioned possibilities for extending the proposed framework. Simulation results and a brief analysis of the computational complexity are presented in Section 4.5.

4.1 Downlink Multiuser MIMO System Model

The block-scheme of the considered flat-fading, single-cell, multiuser MIMO system with K users is given in Fig. 4.1. The antenna arrays at the BS and user k, $k = 1, \ldots, K$, are of sizes M and L_k, respectively. The BS is supposed to perform spatial multiplexing by transmitting an $N \times 1$ vector $\boldsymbol{s} = \left[\boldsymbol{s}_1^*, \boldsymbol{s}_2^*, \cdots, \boldsymbol{s}_K^*\right]^*$ of independent symbols, where $\boldsymbol{s}_k = \left[s_k^{(1)}, s_k^{(2)}, \cdots, s_k^{(N_k)}\right]^T \in \mathbb{C}^{N_k}$ is the sub-vector for the user k containing N_k data streams (layers), and $N = \sum_{k=1}^{K} N_k$. It is assumed, w.l.o.g., that $\mathrm{E}\{\boldsymbol{s}\boldsymbol{s}^*\} = \boldsymbol{I}$. The linear spatial precoder at the BS is denoted by $\boldsymbol{G} \in \mathbb{C}^{M \times N}$. The BS possesses only imperfect estimates $\hat{\boldsymbol{H}}_k$ of the users' flat-fading channels $\boldsymbol{H}_k \in \mathbb{C}^{L_k \times M}$, $k = 1, \ldots, K$. The expression for the actual channel $\boldsymbol{H}_k$ is

$$\boldsymbol{H}_k = \hat{\boldsymbol{H}}_k + \boldsymbol{\Delta}_k, \quad k = 1, \ldots, K \tag{4.1}$$

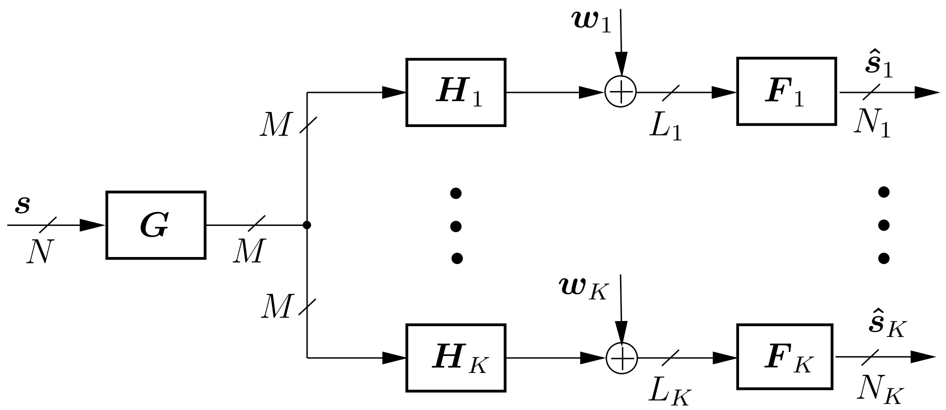

Figure 4.1: Schematic representation of the studied downlink multiuser MIMO system.

where $\mathbf{\Delta}_k$ is the error unknown to the transmitter. In the main problem formulations, it will be assumed that the Frobenius norm of $\mathbf{\Delta}_k$ is bounded

$$\left\|\mathbf{\Delta}_k\right\|_F \leq \varepsilon_k, \quad k = 1, \ldots, K. \tag{4.2}$$

The kth user's linear equalizer is denoted by $\boldsymbol{F}_k \in \mathbb{C}^{N_k \times L_k}$. At the receivers, zero-mean, additive noise $\boldsymbol{w}_k$ is present. The noise is independent from the input signals. To simplify the exposition, it is initially assumed that $\mathrm{E}\{\boldsymbol{w}_k \boldsymbol{w}_k^*\} = \sigma_k^2 \boldsymbol{I}$, and that σ_k, $k = 1, \ldots, K$, are known at the transmitter.

The received signal of the user k (after the equalization) is

$$\hat{\mathbf{s}}_k \triangleq \begin{bmatrix} \hat{s}_k^{(1)} & \hat{s}_k^{(2)} & \cdots & \hat{s}_k^{(N_k)} \end{bmatrix}^T = \boldsymbol{F}_k(\boldsymbol{H}_k \boldsymbol{G} \boldsymbol{s} + \boldsymbol{w}_k). \tag{4.3}$$

The total transmit power of the BS (the average for one channel use) can be calculated similarly to (3.4): $P = \|\boldsymbol{G}\|_F^2$. The performance measures are either the MSEs between the sent and the received vectors (per-user MSEs)

$$\mathrm{MSE}_k = \mathrm{E}\{\|\hat{\mathbf{s}}_k - \boldsymbol{s}_k\|_2^2\} \tag{4.4}$$

or the MSEs of individual streams (per-stream MSEs)

$$\mathrm{MSE}_k^{(n)} = \mathrm{E}\left\{\left|\hat{s}_k^{(n)} - s_k^{(n)}\right|_2^2\right\}, \qquad k = 1, \ldots, K, \quad n = 1, \ldots, N_k. \tag{4.5}$$

Three main transceiver design problems are formulated. They present robust counterparts of the MSE optimization problems studied in [MJHU06, KTA06, SSB07, SSB08a]:

- Minimize the total transmit power subject to predefined per-user MSE targets under the uncertainty

$$\min_{\boldsymbol{G},\boldsymbol{F}_1,\ldots,\boldsymbol{F}_K} \quad P \qquad \text{subject to} \quad \mathrm{MSE}_k \le \mu_k, \quad \forall \boldsymbol{\Delta}_k : \|\boldsymbol{\Delta}_k\|_F \le \varepsilon_k, \quad k = 1, \ldots, K. \tag{4.6}$$

  The variation of (4.6) in the case of constrained per-stream MSEs is

$$\begin{aligned} &\min_{\boldsymbol{G},\boldsymbol{F}_1,\ldots,\boldsymbol{F}_K} \quad P \qquad \text{subject to} \quad \mathrm{MSE}_k^{(n)} \le \mu_{k,n} \quad \forall \boldsymbol{\Delta}_k : \|\boldsymbol{\Delta}_k\|_F \le \varepsilon_k, \\ &k = 1, \ldots, K, \quad n = 1, \ldots, N_k. \end{aligned} \tag{4.7}$$

- Minimize the maximum sum MSE under the uncertainty, subject to a maximum transmit power $P_{\max}$

$$\min_{\boldsymbol{G},\boldsymbol{F}_1,\ldots,\boldsymbol{F}_K} \max_{\|\boldsymbol{\Delta}_k\|_F \le \varepsilon_k} \sum_{k=1}^{K} \mathrm{MSE}_k \qquad \text{subject to} \quad P \le P_{\max}. \tag{4.8}$$

- Minimize the maximum per-user MSE under the uncertainty, subject to a transmit power constraint

$$\min_{\boldsymbol{G},\boldsymbol{F}_1,\ldots,\boldsymbol{F}_K} \max_{\substack{\|\boldsymbol{\Delta}_k\|_F \le \varepsilon_k \\ 1 \le k \le K}} \mathrm{MSE}_k \qquad \text{subject to} \quad P \le P_{\max}. \tag{4.9}$$

  Similarly to (4.9), the min-max fairness problem with respect to the streams is also defined

$$\min_{\boldsymbol{G},\boldsymbol{F}_1,\ldots,\boldsymbol{F}_K} \max_{\substack{\|\boldsymbol{\Delta}_k\|_F \le \varepsilon_k \\ 1 \le k \le K,\, 1 \le n \le N_k}} \mathrm{MSE}_k^{(n)} \qquad \text{subject to} \quad P \le P_{\max}. \tag{4.10}$$

The problem (4.6) corresponds to a practical situation where each user defines its desired QoS target in terms of the maximum allowable MSE μ_k. The transmit power minimization is important then, among other reasons, for reducing the interference to the neighboring cells.

The problem (4.8) aims at improving the sum MSE as the global QoS measure for the complete system. In this case, the outcome might be that the differences in users' performances are large and that some users end up in a very high MSE regime. The formulation (4.9), also known as the min-max fairness problem, ensures that this does not happen. The variations (4.7) and (4.10) cover the cases when individual streams are of interest.

At this point, some remarks regarding the equalization and CSI at the receivers, that hold for all considered problems, should be given. The formulations (4.6)-(4.10) envisage that the BS calculates all unknown transceiver coefficients under uncertainty. The user k could be supplied with the data for its filter $\boldsymbol{F}_k$ from the BS, in an initial phase of communication over one channel, using control links, with an assumption that these links are error-free (see also [CMH06, CIHM08] for some results on limited feedforward schemes). If the users have the ability to estimate the equivalent channels $\boldsymbol{H}_k\boldsymbol{G}$, $k = 1, \ldots, K$, they can also self-adjust their equalizers using the MMSE solution

$$\boldsymbol{F}_k = (\boldsymbol{H}_k\boldsymbol{G}_k)^* \left(\boldsymbol{H}_k\boldsymbol{G}(\boldsymbol{H}_k\boldsymbol{G})^* + \sigma_k^2\boldsymbol{I}\right)^{-1}, \quad k = 1, \ldots, K \tag{4.11}$$

where $\boldsymbol{G}_k$ is obtained by taking the columns $1 + \sum_{l=1}^{k-1} N_l, \ldots, \sum_{l=1}^{k} N_l$ of $\boldsymbol{G}$. While the latter approach does not reduce the transmit power, it can decrease the users' MSEs. The advantage in obtaining the solution for $\boldsymbol{F}_k$ from the BS directly is in eliminating the computational burden at the receiver in calculating (4.11). Furthermore, some developing LTE standards do not allow training for calculation of the equivalent channels, but only $\boldsymbol{H}_k$ can be estimated with common pilots at the receiver [CMH06, CIHM08]. An interesting remaining question is which strategy would be more spectrally efficient in terms of overhead (see [KCJ08] for some results regarding these problems in a related scenario). The application of solution (4.11) is possible also in the case when the users obtain the coefficients of $\boldsymbol{G}$ using the control links, if they have already successfully estimated the actual channels $\boldsymbol{H}_k$. However, although in this strategy the coefficients of $\boldsymbol{G}$ can be broadcast simultaneously to all users, the overhead might still be considerably larger comparing to the delivery of all $\boldsymbol{F}_k$, $k = 1, \ldots, K$, depending on the number of transmit/receive antennas.

Ideally, under the assumption that the users can estimate perfectly the equivalent channels, the BS could use knowledge of the solution (4.11) for $\boldsymbol{F}_k$ (along with the uncertainty in it), and try to optimize only the precoder $\boldsymbol{G}$ in the worst-case sense. Notice, however, that this approach yields very involved problems even in the perfect CSI case.

4.2 Iterative Solutions

The expression (4.4) for the MSE of the user k can be rewritten as

$$\mathrm{MSE}_k = \mathrm{E}\left\{\left\|\boldsymbol{F}_k(\boldsymbol{H}_k\boldsymbol{G}\boldsymbol{s} + \boldsymbol{w}_k) - \boldsymbol{Q}_k\boldsymbol{s}\right\|_2^2\right\} = \left\|\boldsymbol{F}_k\boldsymbol{H}_k\boldsymbol{G} - \boldsymbol{Q}_k\right\|_F^2 + \sigma_k^2 \left\|\boldsymbol{F}_k\right\|_F^2 \tag{4.12}$$

where $\boldsymbol{Q}_k = \begin{bmatrix} \boldsymbol{0}_{N_k \times \sum_{l=1}^{k-1} N_l} & \boldsymbol{I}_{N_k} & \boldsymbol{0}_{N_k \times \sum_{l=k+1}^{K} N_l} \end{bmatrix}$. The per-stream MSEs (4.5) are similarly given as

$$\mathrm{MSE}_k^{(n)} = \left\|\boldsymbol{F}_{k(n,:)}\boldsymbol{H}_k\boldsymbol{G} - \boldsymbol{Q}_{k(n,:)}\right\|_F^2 + \sigma_k^2 \left\|\boldsymbol{F}_{k(n,:)}\right\|_F^2 . \tag{4.13}$$

It can be noticed that in all problem formulations, the per-user and per-stream MSEs are either simply upper-bounded or constitute the objective function in an affine manner. However, from (4.12) and (4.13) it follows that the expressions for MSE_k and $\mathrm{MSE}_k^{(n)}$ are not convex in all transceiver coefficients jointly. The convex reformulations of the problems (4.6)-(4.10) are not known even for the perfect CSI case. A partial exception is the sum MSE problem, in the case of transmit covariance matrix optimization and non-fixed number of data streams [SSB07]. Most of the work in the literature under the assumption of perfect CSI is focused on providing efficient suboptimal solutions with good performance in practice [MJHU06, KTA06, SSB07, SSB08a].

The robust problems (4.6)-(4.10) impose serious additional obstacles because of the uncertainty. Differently from the related work for the perfect CSI case, which mainly exploits uplink-downlink duality [SB04, SS05], the idea for an iterative algorithm in this work is based on the fact that the expressions for MSE_k and $\mathrm{MSE}_k^{(n)}$ are clearly convex in the perfect CSI case, if either $\boldsymbol{G}$ or $\boldsymbol{F}_k$ is fixed. Therefore, a non-robust optimization of the transmit filter and the receive equalizers might be possible in a way similar to block coordinate descent methods [Ber99]. In the robust case, a general claim about the convexity of the MSE expression, if one of the filters is fixed, is not obvious because of the uncertainty. Furthermore, contrary to [Ber99], the solutions of the subproblems that optimize the blocks of variables are in general not unique. However, it will be shown that the alternating principle can be adapted to efficiently handle the introduced error model (4.2). The generic iterative procedure is summarized in Algorithm 2.

Algorithm 2 Iterative solution for problems (4.6)-(4.10).

1: *initialization:* Set $t = 0$ (the iteration number), the maximum number of iterations $t_{\max}$, the desired accuracy $\epsilon > 0$, and initial receivers $\boldsymbol{F}_k = \tilde{\boldsymbol{F}}_k$, $k = 1, \ldots, K$. Select arbitrary $P_{\text{new}} \gg 0$ for problems (4.6) and (4.7), or $\text{MSE}_{\text{new}} \gg 0$ for problems (4.8), (4.9) and (4.10).

2: **repeat**

3: $t \leftarrow t + 1$. Fix the current values for $\boldsymbol{F}_k$, $k = 1, \ldots, K$.

4: If (4.6), $P_{\text{old}} \leftarrow P_{\text{new}}$. Solve the problem

$$P_{\text{new}} \leftarrow \min_{\boldsymbol{G}} \quad P \quad \text{subject to} \quad \text{MSE}_k \leq \mu_k, \quad \forall ||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k, \; k = 1, \ldots, K. \tag{4.14}$$

5: If (4.7), $P_{\text{old}} \leftarrow P_{\text{new}}$. Solve the problem

$$P_{\text{new}} \leftarrow \min_{\boldsymbol{G}} \quad P \quad \text{s. t.} \quad \text{MSE}_k^{(n)} \leq \mu_{k,n}, \quad \forall ||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k, \; k = 1, \ldots, K, \; n = 1, \ldots, N_k. \tag{4.15}$$

6: If (4.8), $\text{MSE}_{\text{old}} \leftarrow \text{MSE}_{\text{new}}$. Solve the problem

$$\text{MSE}_{\text{new}} \leftarrow \min_{\boldsymbol{G}} \max_{||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k} \sum_{k=1}^{K} \text{MSE}_k \quad \text{subject to} \quad P \leq P_{\max}. \tag{4.16}$$

7: If (4.9), $\text{MSE}_{\text{old}} \leftarrow \text{MSE}_{\text{new}}$. Solve the problem

$$\text{MSE}_{\text{new}} \leftarrow \min_{\boldsymbol{G}} \max_{\substack{||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k \\ 1 \leq k \leq K}} \text{MSE}_k \quad \text{subject to} \quad P \leq P_{\max}. \tag{4.17}$$

8: If (4.10), $\text{MSE}_{\text{old}} \leftarrow \text{MSE}_{\text{new}}$. Solve the problem

$$\text{MSE}_{\text{new}} \leftarrow \min_{\boldsymbol{G}} \max_{\substack{||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k \\ 1 \leq k \leq K,\, 1 \leq n \leq N_k}} \text{MSE}_k^{(n)} \quad \text{subject to} \quad P \leq P_{\max}. \tag{4.18}$$

9: Update of the receivers for per-user MSE problems, with the fixed precoder from (4.14), (4.16), or (4.17):

$$\boldsymbol{F}_k \leftarrow \arg\min_{\boldsymbol{F}_k} \max_{||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k} \text{MSE}_k, \quad k = 1, \ldots, K. \tag{4.19}$$

10: Update of the receivers for per-stream MSE problems, with the fixed precoder from (4.15) or (4.18):

$$\boldsymbol{F}_{k_{(n,:)}} \leftarrow \arg\min_{\boldsymbol{F}_{k_{(n,:)}}} \max_{||\boldsymbol{\Delta}_k||_F \leq \varepsilon_k} \text{MSE}_k^{(n)}, \quad k = 1, \ldots, K, \quad n = 1, \ldots, N_k. \tag{4.20}$$

11: **until** $t \geq t_{\max}$, or $P_{\text{old}} - P_{\text{new}} < \epsilon$ for (4.6) and (4.7), or $\text{MSE}_{\text{old}} - \text{MSE}_{\text{new}} < \epsilon$ for (4.8)-(4.10).

The rest of this section provides the necessary ingredients for solving the subproblems of

Algorithm 2. More precisely, it will be shown that (4.14)-(4.20) are SDP problems. The convergence of the iterative procedures, described by the algorithm, will be proved, as well.

4.2.1 Power Minimization Problems

Consider firstly the per-user MSE variation (4.6). By using the relation

$$\mathrm{vec}\,(\boldsymbol{ABC}) = \left(\boldsymbol{C}^T \otimes \boldsymbol{A}\right) \mathrm{vec}\,(\boldsymbol{B}) \tag{4.21}$$

valid for arbitrary matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ ([HJ91], Section 4.3), MSE_k given by (4.12) can be expressed as

$$\mathrm{MSE}_k = \left\|\boldsymbol{\psi}_k(\boldsymbol{G}, \boldsymbol{F}_k) + \boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)\boldsymbol{\delta}_k\right\|_2^2 + \sigma_k^2 \, ||\boldsymbol{F}_k||_F^2 \tag{4.22}$$

where

$$\boldsymbol{\psi}_k(\boldsymbol{G}, \boldsymbol{F}_k) = \mathrm{vec}(\boldsymbol{F}_k\hat{\boldsymbol{H}}_k\boldsymbol{G} - \boldsymbol{Q}_k), \quad \boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k) = \boldsymbol{G}^T \otimes \boldsymbol{F}_k, \quad \boldsymbol{\delta}_k = \mathrm{vec}(\boldsymbol{\Delta}_k). \tag{4.23}$$

Notice that the functions $\boldsymbol{\psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ and $\boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ are affine in $\boldsymbol{G}$, if $\boldsymbol{F}_k$ is kept fixed, and vice versa. The explicit dependence of $\boldsymbol{\psi}_k$ and $\boldsymbol{\Psi}_k$ on $\boldsymbol{G}$ and $\boldsymbol{F}_k$ will be occasionally omitted in the rest of this section to simplify the expressions.

The optimization of the transmit precoder can be resolved by a transformation into an SDP problem using Lemma 7.

Theorem 5. The problem (4.14) is equivalent to

$$\begin{aligned} &\min_{\boldsymbol{G},\lambda_1,\ldots,\lambda_K} \quad ||\boldsymbol{G}||_F^2 \\ &\text{subject to} \quad \begin{bmatrix} \sqrt{\mu_k - \sigma_k^2 \, ||\boldsymbol{F}_k||_F^2} - \lambda_k & \boldsymbol{\psi}_k^* & \boldsymbol{0} \\ \boldsymbol{\psi}_k & \sqrt{\mu_k - \sigma_k^2 \, ||\boldsymbol{F}_k||_F^2}\, \boldsymbol{I} & -\varepsilon_k \boldsymbol{\Psi}_k \\ \boldsymbol{0} & -\varepsilon_k \boldsymbol{\Psi}_k^* & \lambda_k \boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0}, \quad k = 1, \ldots, K \end{aligned} \tag{4.24}$$

where $\lambda_1, \ldots, \lambda_K$ are slack variables.

Proof. Using the relation (2.2), the constraint $\text{MSE}_k \leq \mu_k$ can be rewritten as

$$\begin{bmatrix} \sqrt{\mu_k - \sigma_k^2 \|\boldsymbol{F}_k\|_F^2} & \boldsymbol{\psi}_k^* + \boldsymbol{\delta}_k^* \boldsymbol{\Psi}_k^* \\ \boldsymbol{\psi}_k + \boldsymbol{\Psi}_k \boldsymbol{\delta}_k & \sqrt{\mu_k - \sigma_k^2 \|\boldsymbol{F}_k\|_F^2}\, \boldsymbol{I} \end{bmatrix} \geq \boldsymbol{0}. \tag{4.25}$$

Since $\|\boldsymbol{\Delta}_k\|_F \leq \varepsilon_k \Leftrightarrow \|\boldsymbol{\delta}_k\|_2 \leq \varepsilon_k \Leftrightarrow \left\|\boldsymbol{\delta}_k^*\right\|_2 \leq \varepsilon_k$, a direct application of Lemma 7 on (4.25) with

$$\begin{aligned} &\boldsymbol{A} = \begin{bmatrix} \sqrt{\mu_k - \sigma_k^2 \|\boldsymbol{F}_k\|_F^2} & \boldsymbol{\psi}_k^* \\ \boldsymbol{\psi}_k & \sqrt{\mu_k - \sigma_k^2 \|\boldsymbol{F}_k\|_F^2}\, \boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{B} = -\begin{bmatrix} 1 & \boldsymbol{0} \end{bmatrix} \\ &\boldsymbol{C} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Psi}_k^* \end{bmatrix}, \quad \boldsymbol{D} = \boldsymbol{\delta}_k^*, \quad \varepsilon = \varepsilon_k \end{aligned} \tag{4.26}$$

yields the problem (4.24). □

Due to the affinity of $\boldsymbol{\psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ and $\boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ in $\boldsymbol{G}$, it can be concluded that the constraints in (4.24) are linear matrix inequalities (LMIs) in the unknown precoder coefficients. Therefore, when the receivers are fixed, an SDP problem, known to be solvable with interior point methods in a very efficient manner, is obtained.

The equivalence of (4.19) to a semidefinite program can now be proved. Notice that the step (4.19) is the same for all per-user MSE problems of interest.

Theorem 6. The problem (4.19) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{F}_k, \lambda_k, \tau_k} \quad & \tau_k + \sigma_k^2 \|\boldsymbol{F}_k\|_F^2 \\ \text{subject to} \quad & \begin{bmatrix} \tau_k - \lambda_k & \boldsymbol{\psi}_k^* & \boldsymbol{0} \\ \boldsymbol{\psi}_k & \boldsymbol{I} & -\varepsilon_k \boldsymbol{\Psi}_k \\ \boldsymbol{0} & -\varepsilon_k \boldsymbol{\Psi}_k^* & \lambda_k \boldsymbol{I} \end{bmatrix} \geq \boldsymbol{0}, \quad k = 1, \ldots, K \end{aligned} \tag{4.27}$$

where λ_k and τ_k, $k = 1, \ldots, K$, are slack variables.

Proof. Consider (4.22) and introduce the slack variable τ_k so that $\|\boldsymbol{\psi}_k + \boldsymbol{\Psi}_k \boldsymbol{\delta}_k\|_2^2 \leq \tau_k$. Rewrite this uncertain SOC constraint using the Schur Complement Lemma (Lemma 4) as

$$\begin{bmatrix} \tau_k & \boldsymbol{\psi}_k^* \\ \boldsymbol{\psi}_k & \boldsymbol{I} \end{bmatrix} + \begin{bmatrix} 0 & \boldsymbol{\delta}_k^* \boldsymbol{\Psi}_k^* \\ \boldsymbol{\Psi}_k \boldsymbol{\delta}_k & \boldsymbol{0} \end{bmatrix} \geq \boldsymbol{0}. \tag{4.28}$$

The uncertainty in (4.28) is resolved by applying Lemma 7 with

$$\boldsymbol{A} = \begin{bmatrix} \tau_k & \boldsymbol{\psi}_k^* \\ \boldsymbol{\psi}_k & \boldsymbol{I} \end{bmatrix} \tag{4.29}$$

and with parameters $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, and ε being the same as in (4.26). The procedure can be independently performed for all equalizers $\boldsymbol{F}_k$, $k = 1, \dots, K$. For a fixed $\boldsymbol{G}$, a set of SDPs in the unknown receivers' coefficients is obtained. □

It should be remarked that using a strong duality result from [SNC02, GL05, GL06], one can construct another approach for obtaining an explicit convex program equivalent to (4.19). However, the existence of the fixed precoder $\boldsymbol{G}$ in (4.19) makes the optimization considerably more involved comparing to [GL06]. We will return to this approach also in Section 4.3.

Finally, it can be proved that the iterative procedure from Algorithm 2 converges for the problem (4.6).

Theorem 7. If the problem (4.14) is feasible for the initial receivers $\tilde{\boldsymbol{F}}_k$, the convergence of the sequence of powers generated by Algorithm 2 is guaranteed for the setting (4.6).

Proof. Let the problem (4.14) have a non-empty domain for initial receive filters, and let $P_{\text{new}}^{(1)}$ be the solution of (4.14) for the minimal transmit power. Let $\text{MSE}_k^{1a} \leq \mu_k$, $k = 1, \dots, K$, be the achieved maximum MSEs under the uncertainty after this step. It can be noticed that the minimax problem (4.19) is always feasible. After solving (4.19), a new set of guaranteed MSEs under the uncertainty, denote them by MSE_k^{1b}, is obtained, where $\text{MSE}_k^{1b} \leq \text{MSE}_k^{1a}$, $\forall k$, due to the definition of (4.19). In the second iteration, the precoder $\boldsymbol{G}$ is recalculated with the receivers $\boldsymbol{F}_k$ updated using (4.19). However, the solution of (4.14) in the second iteration can only give the total transmit power $P_{\text{new}}^{(2)}$ which is smaller than or equal to $P_{\text{new}}^{(1)}$. The reason for this is that the solution for $\boldsymbol{G}$ from the first iteration, which yielded $P_{\text{new}}^{(1)}$, is feasible in the second iteration, as well. Therefore, the proposed iterative procedure yields a monotonically decreasing sequence of transmit powers (and the corresponding transceiver coefficients) that is clearly bounded below (at least with zero), which concludes the proof of the theorem. □

The selection of initial receive equalizers $\tilde{\boldsymbol{F}}_k$ is of significant importance for the problem

(4.6), since their improper choice can result in an immediate infeasibility of this problem. For example, from (4.12), it follows that a necessary condition for the feasibility of the constraint $\text{MSE}_k \leq \mu_k$ in (4.6) is

$$\mu_k - \sigma_k^2 \|\boldsymbol{F}_k\|_F^2 \geq 0. \tag{4.30}$$

Consider, therefore, a slightly modified system model, where the receivers have the form $\frac{1}{p_k}\boldsymbol{F}_k$, $p_k \in \mathbb{R}_{++}$, $k = 1, \ldots, K$. In this case, the condition $\text{MSE}_k \leq \mu_k$ is equivalent to

$$\left\| \begin{bmatrix} \boldsymbol{\psi}_k^p(\boldsymbol{G}, \boldsymbol{F}_k, p_k)^* + \boldsymbol{\delta}_k^* \boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)^* & \sigma_k \|\boldsymbol{F}_k\|_F \end{bmatrix} \right\|_2 \leq p_k \sqrt{\mu_k} \tag{4.31}$$

where

$$\boldsymbol{\psi}_k^p(\boldsymbol{G}, \boldsymbol{F}_k, p_k) = \text{vec}\left(\boldsymbol{F}_k \hat{\boldsymbol{H}}_k \boldsymbol{G} - p_k \boldsymbol{Q}_k\right) \tag{4.32}$$

is affine in $\boldsymbol{G}$ and p_k jointly, if $\boldsymbol{F}_k$ is kept fixed, and vice versa. By applying the same methodology as in Section 3.2, equivalent LMIs with no uncertainty, that can replace the constraints in (4.24) to attempt a feasible start of Algorithm 2, are obtained

$$\begin{bmatrix} p_k \sqrt{\mu_k} - \lambda_k & \boldsymbol{\psi}_k^{p*} & \sigma_k \|\boldsymbol{F}_k\|_F & \boldsymbol{0} \\ \boldsymbol{\psi}_k^p & p_k \sqrt{\mu_k} \boldsymbol{I} & \boldsymbol{0} & -\varepsilon_k \boldsymbol{\Psi}_k \\ \sigma_k \|\boldsymbol{F}_k\|_F & \boldsymbol{0} & p_k \sqrt{\mu_k} & \boldsymbol{0} \\ \boldsymbol{0} & -\varepsilon_k \boldsymbol{\Psi}_k^* & \boldsymbol{0} & \lambda_k \boldsymbol{I} \end{bmatrix} \geq \boldsymbol{0}, \quad k = 1, \ldots, K. \tag{4.33}$$

Furthermore, the LMIs (4.33) can be used to optimize the BS precoder together with the scaling factors p_k, $k = 1, \ldots, K$, not only in the first, but also in the following iterations, when applying Algorithm 2 for the problem (4.6). The update of the complete receivers can be done with the obtained transmitter $\boldsymbol{G}$ as described by Theorem 6. In this update, the scaling factors can be neglected (set to one), and the obtained $\boldsymbol{F}_k$ can be used in the subsequent construction of (4.33). The convergence analysis assuming the feasible initial constraints (4.33) is the same as in the proof of Theorem 7.

Consider now the problem (4.7), which addresses the MSE constraints on individual streams. We focus immediately on the more complex receiver model $(1/p_{k,n})\boldsymbol{F}_{k_{(n,:)}}$, similarly to the discussion above. The condition $\text{MSE}_k^{(n)} \le \mu_{k,n}$ is equivalent to

$$\left\| \left[\boldsymbol{\psi}_{k,n}^{p}(\boldsymbol{G}, \boldsymbol{F}_{k(n,:)}, p_{k,n})^{*} + \boldsymbol{\delta}_k^{*}\boldsymbol{\Psi}_{k,n}(\boldsymbol{G}, \boldsymbol{F}_{k(n,:)})^{*} \quad \sigma_k \left\| \boldsymbol{F}_{k(n,:)} \right\|_F \right] \right\|_2 \le p_{k,n} \sqrt{\mu_{k,n}} \tag{4.34}$$

where

$$\begin{aligned} \boldsymbol{\psi}_{k,n}^{p}(\boldsymbol{G}, \boldsymbol{F}_{k(n,:)}, p_{k,n}) &= \left(\boldsymbol{F}_{k(n,:)} \hat{\boldsymbol{H}}_k \boldsymbol{G} - p_{k,n} \boldsymbol{Q}_{k(n,:)} \right)^T \\ \boldsymbol{\Psi}_{k,n}(\boldsymbol{G}, \boldsymbol{F}_{k(n,:)}) &= \boldsymbol{G}^T \otimes \boldsymbol{F}_{k(n,:)}. \end{aligned} \tag{4.35}$$

Using (4.34) and (4.35), one can construct equivalent, explicit SDP transmitter optimization problems analogously to (4.33). The update of the receivers, defined now by (4.20), simultaneously minimizes the worst-case MSEs of all streams with respect to the uncertainty. The convergence is ensured similarly as in the proof of Theorem 7.

4.2.2 Sum MSE Minimization Problem

Since the update of the receivers has already been resolved, we focus on the tractable reformulation of (4.16) and the convergence in the branch of Algorithm 2 for supporting the problem (4.8).

Theorem 8. The problem (4.16) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{G}, \lambda_1, \ldots, \lambda_K, \tau_1, \ldots, \tau_K} \quad & \sum_{k=1}^{K} \tau_k \\ \text{subject to} \quad & \begin{bmatrix} \tau_k - \lambda_k & \boldsymbol{\psi}_k^{*} & \boldsymbol{0} \\ \boldsymbol{\psi}_k & \boldsymbol{I} & -\varepsilon_k \boldsymbol{\Psi}_k \\ \boldsymbol{0} & -\varepsilon_k \boldsymbol{\Psi}_k^{*} & \lambda_k \boldsymbol{I} \end{bmatrix} \ge \boldsymbol{0}, \quad k = 1, \ldots, K \\ & \|\boldsymbol{G}\|_F \le \sqrt{P_{\max}} \end{aligned} \tag{4.36}$$

where $\tau_1, \ldots, \tau_K$ and $\lambda_1, \ldots, \lambda_K$ are slack variables.

Proof. Introduce the slack variables τ_k, $k = 1, \ldots, K$, in the same way as in the proof of Theorem 6, and notice that the problem (4.16) assumes fixed receivers. Therefore, the term $\sigma_k^2 \|\boldsymbol{F}_k\|_F^2$ in the expression for the MSE_k (4.22) has no influence on the optimal transmitter coefficients, and it is omitted in the objective function of (4.36). The equivalent SDP problem is now derived analogously to the proof of Theorem 6, except that the affinity of $\boldsymbol{\psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ and $\boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ in $\boldsymbol{G}$ is exploited. The total transmit power constraint is rewritten as an SOC constraint. □

Differently from the power minimization problem, the convergence of Algorithm 2 for the sum MSE minimization is always guaranteed.

Theorem 9. The MSE sequence, generated by Algorithm 2 for the problem (4.8), converges for any choice of the initial receivers.

Proof. Firstly, it should be noticed that differently from (4.14), the minimax problem (4.16), that presents the first step in one iteration of Algorithm 2 for the case (4.8), is always feasible. Let MSE^{1a} be the achieved minimum sum MSE under the uncertainty after the first application of (4.36). Due to its definition, the subsequent employment of (4.19) yields a new sum MSE, MSE^{1b}, with $\mathrm{MSE}^{1b} \leq \mathrm{MSE}^{1a}$. In the second iteration, the new application of (4.36) additionally reduces the MSE value, because the previous solution for the transmit filter is feasible for this problem, as well. Therefore, the iterative transceiver optimization creates a monotonically decreasing sequence of the achievable sum MSEs under uncertainty, which is clearly bounded below. □

4.2.3 Min-Max Fairness Problems

The equivalent SDP representation of the first part in one outer iteration for the fairness problem (4.9) is given by the following theorem.

Theorem 10. The problem (4.17) is equivalent to

$$\begin{aligned}
&\min_{\boldsymbol{G}, t, \lambda_1, \ldots, \lambda_K} \quad \tau \\
&\text{subject to} \quad \begin{bmatrix} \tau - \sigma_k^2 \|\boldsymbol{F}_k\|_F^2 - \lambda_k & \boldsymbol{\psi}_k^* & \boldsymbol{0} \\ \boldsymbol{\psi}_k & \boldsymbol{I} & -\varepsilon_k \boldsymbol{\Psi}_k \\ \boldsymbol{0} & -\varepsilon_k \boldsymbol{\Psi}_k^* & \lambda_k \boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0}, \quad k = 1, \ldots, K \\
&\qquad\qquad \|\boldsymbol{G}\|_F \leq \sqrt{P_{\max}}
\end{aligned} \tag{4.37}$$

where τ, $\lambda_1, \ldots, \lambda_K$ are slack variables.

Proof. Introduce the slack variable τ through a constraint

$$\mathrm{MSE}_k \leq \tau, \quad k = 1, \ldots, K. \tag{4.38}$$

Notice that (4.38) is equivalent to

$$\left\| \boldsymbol{\psi}_k + \boldsymbol{\Psi}_k \boldsymbol{\delta}_k \right\|_2^2 \leq \tau - \sigma_k^2 \left\| \boldsymbol{F}_k \right\|_F^2, \quad k = 1, \ldots, K. \tag{4.39}$$

The constraint (4.39) can be rewritten using the Schur Complement Lemma as

$$\begin{bmatrix} \tau - \sigma_k^2 \left\| \boldsymbol{F}_k \right\|_F^2 & \boldsymbol{\psi}_k^* + \boldsymbol{\delta}_k^* \boldsymbol{\Psi}_k^* \\ \boldsymbol{\psi}_k + \boldsymbol{\Psi}_k \boldsymbol{\delta}_k & \boldsymbol{I} \end{bmatrix} \succeq \mathbf{0}. \tag{4.40}$$

The equivalent LMIs with no uncertainty in (4.37) are obtained by an application of Lemma 7 on (4.40), with $\boldsymbol{A}$ being the matrix on the left hand side of (4.40) for $\boldsymbol{\delta}_k = \mathbf{0}$, and with parameters $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, and ε being the same as in (4.26). □

Analogously to the proof of Theorem 9, it can be shown that the convergence of Algorithm 2 for the case (4.9) is ensured for any initial receivers. Finally, the per-stream, min-max, fairness problem (4.10) can be supported similarly to the analysis presented at the end of Section 4.2.1, but without scaling factors in the case of multi-antenna receivers. In the case of single-antenna users, clearly, an optimal, non-alternating approach is possible using bisection with respect to τ, with one semidefinite program per step.

The convergence of the iterative procedure described by Algorithm 2 to global optima of the respective problems is unfortunately not guaranteed. However, the legitimacy of the used approach is supported by the fact that similar issues regarding the convergence to a global optimum appear if perfect CSI is assumed, as analyzed in [MJHU06, KTA06, SSB07, SSB08a]. Finally, it will be shown in Section 4.5, how the robust algorithms presented above approach in performance the methods from [MJHU06, KTA06, SSB07, SSB08a], if the bound on the disturbance ε_k is set to zero.

4.3 Calculation of Worst-Case Channels

Being able to determine the actual worst-case channel for some fixed transceiver is of significant interest, since it can show the potential deterioration of any adopted scheme. Mathematically, the calculation of the worst-case channels involves solving

$$\max_{\|\boldsymbol{\Delta}_k\|_F^2 \leq \varepsilon_k^2} \mathrm{MSE}_k(\boldsymbol{\Delta}_k), \quad k = 1, \ldots, K. \tag{4.41}$$

The MISO system studied in Chapter 3 presents just a special case on which the subsequent analysis is applicable, as well.

Let

$$L(\boldsymbol{\Delta}_k, \lambda_k) = \left\|\boldsymbol{F}_k(\hat{\boldsymbol{H}}_k + \boldsymbol{\Delta}_k)\boldsymbol{G} - \boldsymbol{Q}_k\right\|_F^2 + \sigma_k^2 \|\boldsymbol{F}_k\|_F^2 + \lambda_k \left(\varepsilon_k^2 - \|\boldsymbol{\Delta}_k\|_F^2\right) \tag{4.42}$$

be the Lagrangian function with the Lagrange multiplier λ_k for the problem (4.41). Since (4.41) maximizes a convex function over a convex set, it is obviously a non-convex problem. However, (4.41) belongs to a class of trust region subproblems. The equivalent conditions for its optimum can be formulated similarly as in [SW95] (see also [Ber99, SNC02, GL05, GL06])

$$\nabla_{\boldsymbol{\Delta}_k} L(\boldsymbol{\Delta}_k, \lambda_k) = \boldsymbol{0} \tag{4.43}$$

$$\nabla_{\lambda_k} L(\boldsymbol{\Delta}_k, \lambda_k) = 0 \tag{4.44}$$

$$\nabla^2_{\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k} L(\boldsymbol{\Delta}_k, \lambda_k) \preceq \boldsymbol{0}. \tag{4.45}$$

The condition (4.44) is equivalent to

$$\|\boldsymbol{\Delta}_k\|_F^2 = \varepsilon_k^2. \tag{4.46}$$

That (4.46) holds for the optimal, worst-case $\boldsymbol{\Delta}_k$ is clear also from the fact that the solution of (4.41) must lie on the boundary of the problem domain. Using the matrix differentiation [HG07], the equivalent representations of (4.43) and (4.45) are readily obtained as

$$\boldsymbol{F}_k^* \boldsymbol{F}_k \boldsymbol{\Delta}_k \boldsymbol{G}\boldsymbol{G}^* - \lambda_k \boldsymbol{\Delta}_k = \boldsymbol{F}_k^* \left(\boldsymbol{Q}_k - \boldsymbol{F}_k \hat{\boldsymbol{H}}_k \boldsymbol{G}\right) \boldsymbol{G}^* \tag{4.47}$$

$$\lambda_k \geq \|\boldsymbol{G}\|_2^2 \|\boldsymbol{F}_k\|_2^2 . \tag{4.48}$$

Let $\boldsymbol{\Delta}_k^{\mathrm{wc}}$ be an optimal point of (4.41). An optimal Lagrange multiplier λ_k^{opt} can be obtained from a convex problem [SW95, GL05]

$$\lambda_k^{\mathrm{opt}} = \arg \min_{\lambda_k \geq \|\boldsymbol{G}\|_2^2 \|\boldsymbol{F}_k\|_2^2} L(\boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k), \lambda_k). \tag{4.49}$$

In order to solve (4.49) a connection between λ_k and $\boldsymbol{\Delta}_k^{\mathrm{wc}}$ is needed. This is obtained from (4.47) using vectorization (4.21)

$$\mathrm{vec}(\boldsymbol{\Delta}_k^{\mathrm{wc}}) = \left((\boldsymbol{G}\boldsymbol{G}^*)^T \otimes (\boldsymbol{F}_k^*\boldsymbol{F}_k) - \lambda_k \boldsymbol{I}\right)^{\dagger} \mathrm{vec}\left(\boldsymbol{F}_k^*\left(\boldsymbol{Q}_k - \boldsymbol{F}_k\hat{\boldsymbol{H}}_k\boldsymbol{G}\right)\boldsymbol{G}^*\right). \tag{4.50}$$

Notice that $\boldsymbol{\Delta}_k^{\mathrm{wc}}$ from (4.50) is certainly a solution of (4.47) since (4.47) must be consistent (it is a necessary condition for an optimum, and the problem (4.41) is always feasible) [Mey00]. After (4.49) is solved, it remains to be checked whether $\boldsymbol{\Delta}_k^{\mathrm{wc}}$ obtained by (4.50) satisfies the condition (4.46) required for the global optimality. Two cases are of interest here:

- $\lambda_k^{\mathrm{opt}} > \|\boldsymbol{G}\|_2^2 \|\boldsymbol{F}_k\|_2^2$. Since (4.49) is a scalar convex problem, the derivative of $L(\boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k), \lambda_k)$ with respect to λ_k must be zero for λ_k^{opt}, as shown in Fig. 4.2. This derivative can be calculated as (see Appendix A.5 in [Ber99], Definition 2 in [HG07], and [GL05, GL06])

$$\frac{d}{d\lambda_k} L\left(\boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k), \lambda_k\right) = \left(\nabla_{\boldsymbol{\Delta}_k^{\mathrm{wc}}} L\left(\boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k), \lambda_k\right)\right)^* \nabla_{\lambda_k} \boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k) + \varepsilon_k^2 - \left\|\boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k)\right\|_F^2 \tag{4.51}$$

$$= \varepsilon_k^2 - \left\|\boldsymbol{\Delta}_k^{\mathrm{wc}}(\lambda_k)\right\|_F^2 \tag{4.52}$$

  since the first term in the product in (4.51) is zero due to (4.43). Therefore, the last condition for the optimality (4.46) is fulfilled, and $\boldsymbol{\Delta}_k^{\mathrm{wc}}$ is the actual worst-case channel.

- $\lambda_k^{\mathrm{opt}} = \|\boldsymbol{G}\|_2^2 \|\boldsymbol{F}_k\|_2^2$. In this case, (4.52) is just greater or equal to zero (see Fig. 4.2), so $\boldsymbol{\Delta}_k^{\mathrm{wc}}$ might not be the desired worst-case channel. Notice that for $\lambda_k = \|\boldsymbol{G}\|_2^2 \|\boldsymbol{F}_k\|_2^2$, $(\boldsymbol{G}\boldsymbol{G}^*)^T \otimes \boldsymbol{F}_k^*\boldsymbol{F}_k - \lambda_k \boldsymbol{I}$ is not invertible. The worst-case channel can be obtained by adding a vector $\alpha_k \boldsymbol{x}_k$ to $\mathrm{vec}(\boldsymbol{\Delta}_k^{\mathrm{wc}})$, where $\boldsymbol{x}_k$ is an arbitrary vector from the right null space of $(\boldsymbol{G}\boldsymbol{G}^*)^T \otimes \boldsymbol{F}_k^*\boldsymbol{F}_k - \lambda_k^{\mathrm{opt}} \boldsymbol{I}$, and $\alpha_k \in \mathbb{R}$ is determined so that (4.46) is satisfied. The latter

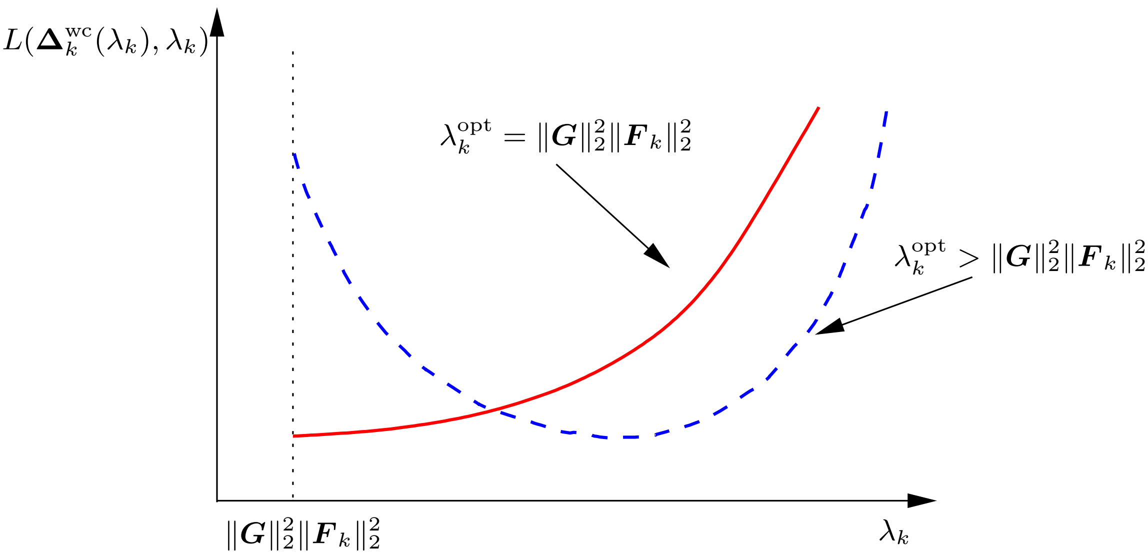

Figure 4.2: Two cases when solving (4.49) depending on whether λ_k^{opt} is at the boundary of the problem domain or not.

problem reduces to a simple, scalar, quadratic equation

$$\alpha_k^2 \|\mathbf{x}_k\|_2^2 + 2\alpha_k \Re\left\{(\mathrm{vec}(\mathbf{\Delta}_k^{\mathrm{wc}}))^* \mathbf{x}_k\right\} + \left\|\mathrm{vec}(\mathbf{\Delta}_k^{\mathrm{wc}})\right\|_2^2 - \epsilon_k^2 = 0. \tag{4.53}$$

The equation (4.53) has always real-valued solutions since $\varepsilon_k^2 - \left\|\mathbf{\Delta}_k^{\mathrm{wc}}\right\|_F^2 \geq 0$ in this case. The obtained solution for $\mathbf{\Delta}_k$ and λ_k clearly satisfies the necessary and sufficient conditions (4.43), (4.44), and (4.45).

4.4 Extensions and Applications

There are variations of the problems (4.6)-(4.10) which can be supported with minor modifications of the presented algorithms. One such example is the problem of transmit power minimization subject to a global sum MSE constraint under uncertainty. Weighted (scaled with fixed numbers) MSEs under uncertainty can be adopted as a performance measure in a straightforward manner. Furthermore, by using the connection between MSEs and SINRs of individual streams, explained in Section 3.3, robust problems that have uncertain SINRs of individual streams as the performance measure can also be approached.

In the rest of this section, the accent is put on the noise covariance uncertainty at the transmitter, additional power constraints on each antenna separately or on groups of antennas, eventual multi-cellular setup, other channel uncertainty regions, and a non-linear transmitter structure. The analysis in the sequel will be applicable to all problems formulated in Section 4.1. Furthermore, the results hold also for the simpler MISO setup with a remark that the transmit/receive alternating optimization is not always necessary in the case of single-antenna users.

4.4.1 Robustness Against Uncertainty in the Noise Covariance

Consider the case when the noise covariance information at the transmitter also contains uncertainties. One might be interested then in handling the worst-case MSE with respect to the uncertain channel and noise covariance. In the case when the noise covariance matrices are $\sigma_k^2 \boldsymbol{I}$, and $\sigma_k \in [\sigma_k^{\min}, \sigma_k^{\max}]$, the values $\sigma_k^{\max}$ clearly correspond to the worst case and should be directly plugged in, instead of σ_k, in the algorithms in Section 4.2. If we consider a general case of the noise covariance matrices $\boldsymbol{C}_k$ that are not scaled identity matrices, various models for their uncertainty can be supported with simple modifications of the proposed framework. For example, let the user k feed back the square root of $\boldsymbol{C}_k$, $\boldsymbol{C}_k^{1/2}$ [GL96], and let $\boldsymbol{C}_k^{1/2} = \hat{\boldsymbol{C}}_k^{1/2} + \boldsymbol{\Delta}_k^w$, where $\hat{\boldsymbol{C}}_k^{1/2}$ is known to the transmitter and $\left\|\boldsymbol{\Delta}_k^w\right\|_F \leq \xi_k$. Notice that this assures the positive semidefiniteness of the noise covariance matrix at the BS. The noise term in the MSE expressions in all considered problems can be handled by introducing slack variables η_k with

$$\left\|\boldsymbol{F}_k\left(\hat{\boldsymbol{C}}_k^{1/2} + \boldsymbol{\Delta}_k^w\right)\right\|_F \leq \eta_k. \tag{4.54}$$

The condition (4.54) is equivalent to

$$\left\|\text{vec}\left(\boldsymbol{F}_k\hat{\boldsymbol{C}}_k^{1/2}\right) + (\boldsymbol{I} \otimes \boldsymbol{F}_k)\,\text{vec}\,(\boldsymbol{\Delta}_k^w)\right\|_2 \leq \eta_k. \tag{4.55}$$

Lemma 7 can be used now to obtain an equivalent, explicit LMI that covers the noise term of the kth user

$$\begin{bmatrix} \eta_k - \zeta_k & \text{vec}\left(\boldsymbol{F}_k\hat{\boldsymbol{C}}_k^{1/2}\right)^* & \boldsymbol{0} \\ \text{vec}\left(\boldsymbol{F}_k\hat{\boldsymbol{C}}_k^{1/2}\right) & \eta_k\boldsymbol{I} & -\xi_k\,(\boldsymbol{I} \otimes \boldsymbol{F}_k) \\ \boldsymbol{0} & -\xi_k\,(\boldsymbol{I} \otimes \boldsymbol{F}_k)^* & \zeta_k\boldsymbol{I} \end{bmatrix} \geq \boldsymbol{0} \tag{4.56}$$

where ζ_k is a slack variable. Consider as an illustration the condition $\text{MSE}_k \leq \mu_k$ in the power minimization problem from Section 4.2.1. This uncertain inequality can be rewritten as a combination of three inequalities $\left\|\boldsymbol{\psi}_k + \boldsymbol{\Psi}_k\boldsymbol{\delta}_k\right\|_2 \leq \omega_k$, (4.55), and $\omega_k^2 + \eta_k^2 \leq \mu_k$, where ω_k is an additional slack variable. From the analysis in Section 4.2.1 and (4.56), it can be concluded that the MSE constraints with the uncertain noise covariance model explained above, have SDP representations.

Another simple way for introducing robustness with respect to the noise is with a slightly different uncertainty model

$$\boldsymbol{C}_k = \hat{\boldsymbol{C}}_k + \boldsymbol{\Delta}_k^w \tag{4.57}$$

with a Hermitian matrix $\hat{\boldsymbol{C}}_k$ known to the transmitter, $\left\|\boldsymbol{\Delta}_k^w\right\|_2 \leq \xi_k$, and $\hat{\boldsymbol{C}}_k + \boldsymbol{\Delta}_k^w \geq \boldsymbol{0}$ for all $\left\|\boldsymbol{\Delta}_k^w\right\|_2 \leq \xi_k$. In this case, it holds that

$$\text{Tr}\{\boldsymbol{F}_k\boldsymbol{C}_k\boldsymbol{F}_k^*\} = \text{Tr}\{\boldsymbol{F}_k^*\boldsymbol{F}_k(\hat{\boldsymbol{C}}_k + \boldsymbol{\Delta}_k^w)\} \leq \text{Tr}\{\boldsymbol{F}_k^*\boldsymbol{F}_k(\hat{\boldsymbol{C}}_k + \xi_k\boldsymbol{I})\}. \tag{4.58}$$

For the last transformation, the relation

$$\left\|\boldsymbol{\Delta}_k^w\right\|_2 \leq \xi_k \quad \Leftrightarrow \quad \xi_k\boldsymbol{I} \geq \pm\boldsymbol{\Delta}_k^w \tag{4.59}$$

which follows from (2.3) (notice that in this model $\boldsymbol{\Delta}_k^w$ is a Hermitian matrix), and Lemma 5 are used. The explicit worst-case noise term that results from (4.58) can be directly inserted into the SDP framework.

4.4.2 Per-Antenna Power Constraints and Network MIMO Systems

In a practical system, constraints on transmit powers of individual antennas might be imposed. These conditions can be introduced in the framework described by Algorithm 2 in a straightforward manner since they are SOC constraints of the type

$$\left\|\boldsymbol{G}_{(m,:)}\right\|_2 \leq \sqrt{P_m} \tag{4.60}$$

where P_m is the power limit for antenna m. The analysis from Section 4.2 remains valid after including the constraints (4.60), because the updates of the transmitter will again always have the previous solution for $\boldsymbol{G}$, that satisfied PAPCs, as a feasible choice.

The power constraints can be imposed in a similar manner on groups of antennas by selecting appropriate rows of the transmit precoder $\boldsymbol{G}$. Let $m_l, \ldots, m_k$ be the antennas with a group power constraint P_g. This constraint has an equivalent SOC form

$$\left\| \begin{bmatrix} \boldsymbol{G}_{(m_l,:)} & \cdots & \boldsymbol{G}_{(m_k,:)} \end{bmatrix} \right\|_2 \leq \sqrt{P_g}. \tag{4.61}$$

While the group condition (4.61) has also a practical value in a single-cell system, it is of particular importance in a coordinated network MIMO system, where multi-antenna base stations perform fully cooperative downlink transmission [KFV06]. Clearly, in such multi-cell systems, the power constraints on groups of antennas can be used as per-BS power constraints.

4.4.3 Rectangular and Unbounded Stochastic CSI Uncertainty Models

The rectangular error model is defined as

$$|\Re\{\boldsymbol{\Delta}_{k_{(l,m)}}\}| \leq \varepsilon_k, \ |\Im\{\boldsymbol{\Delta}_{k_{(l,m)}}\}| \leq \varepsilon_k, \quad l = 1, \ldots, L_k, \quad m = 1, \ldots, M, \quad k = 1, \ldots, K. \tag{4.62}$$

The model (4.62) corresponds well to errors that are caused by scalar quantization. In this case, the uncertainty region for one channel $\boldsymbol{H}_k$ is actually the convex hull of the following set of fixed channel matrices

$$\hat{\boldsymbol{H}}_k + \varepsilon_k \begin{bmatrix} \pm 1 \pm j & \cdots & \pm 1 \pm j \\ \vdots & \ddots & \vdots \\ \pm 1 \pm j & \cdots & \pm 1 \pm j \end{bmatrix}_{L_k \times M} . \tag{4.63}$$

Denote the matrices from (4.63) with $\overline{\boldsymbol{H}}_k^i$, $i = 1, \ldots, 4^{L_k M}$, and, w.l.o.g., focus on the problem (4.6), with the disturbances described by (4.62), and with the receivers in the form $\frac{1}{p_k}\boldsymbol{F}_k$ to handle the initialization. The (suboptimal) solution can be obtained by a similar iterative procedure as in Algorithm 2. However, the two problems that should be addressed in each iteration are different:

- Instead of (4.14), the problem of minimizing the transmit power subject to the MSE constraints that must be satisfied for all *fixed* channel matrices $\bar{\boldsymbol{H}}_k^i$ and for fixed $\boldsymbol{F}_k$, $i = 1, \ldots, 4^{L_k M}$, $k = 1, \ldots, K$, should be solved. Using (4.12), it can be proved that this is a conic quadratic problem

$$\begin{aligned} \min_{\boldsymbol{G}, p_1, \ldots, p_K} \quad & \|\boldsymbol{G}\|_F^2 \\ \text{subject to} \quad & \left\| \left[\operatorname{vec}\left(\boldsymbol{F}_k \bar{\boldsymbol{H}}_k^i \boldsymbol{G} - p_k \boldsymbol{Q}_k\right)^* \quad \sigma_k \|\boldsymbol{F}_k\|_F \right] \right\|_2 \leq p_k \sqrt{\mu_k} \\ & i = 1, \ldots, 4^{L_k M}, \quad k = 1, \ldots, K. \end{aligned} \tag{4.64}$$

  The equivalence stems directly from the general convexity of SOC constraints and the representation with the convex hull (4.63).

- The update of the receivers is performed again in a minimax fashion. The uncertainty in (4.19) is resolved now similarly to (4.64), with a set of constraints corresponding to the channel matrices from (4.63). With a fixed precoder, it can be easily shown that a set of conic quadratic problems is obtained.

This principle can be applied for any CSI error modeling based on the convex hull of a fixed, finite set of matrices. However, although each iteration again consists of solving the problems that are in theory convex (SOC), it can be noticed that every channel matrix from the set (4.63) generates one additional SOC constraint in these problems. Consequently, for the rectangular model (4.62), the number of additional conditions grows exponentially with the system dimension. The computational complexity is therefore unacceptably high and cannot be considered as polynomial in time for one outer iteration anymore. Notice that in the multiuser MIMO case, it is not possible to apply approximations similar as in [BTNR02, BTBG+, VB09a], since the alternating transmit/receive optimization requires optimal solutions of the subproblems in each outer iteration.

In such cases, it is beneficial to robustly optimize the system with a certain degree of conservativeness. Notice that the ball of radius $\varepsilon_k \sqrt{2L_k M}$ with the center at the origin contains the kth user's uncertainty box defined by (4.62). Therefore, the uncertainty model (4.2) can be used in constructing a safe, tractable approximation for the problems with box-like uncertainty. The analysis for the ball uncertainty presented in Section 4.2 is easily extended for the ellipsoidal uncertainty sets $\|\boldsymbol{V}\boldsymbol{\delta}_k\|_2 \leq \varepsilon_k$, where $\boldsymbol{V}$ is a square, non-singular matrix. This can be done with a simple change of variables that describe the uncertainty. For this reason, the principle of conservative approximations has a broader application, because there are numerous methods for calculating the (unique) minimum volume ellipsoidal container [BTN01] for various shapes of uncertainty regions, that might in practice correspond to Voronoi sets in popular vector quantization methods.

If the actual CSI error is unbounded, certain QoS targets, naturally, cannot be strictly guaranteed anymore. However, the model (4.2) can be used to control the outage performance of the system, if the statistical information about the disturbance and channel is known. Consider, e.g., the case of channel estimation, where the channel and estimation error are typically assumed to be independent and jointly Gaussian. The radius ε_k of a ball that with a certain probability p contains the error $\boldsymbol{\Delta}_k$, can be calculated then by using the results from [PIPPNL06]. In this way, the performance achieved with the framework from Section 4.2 can be violated for a given erroneous channel realization with probability $1 - p$ at most. We will thoroughly review this concept also in Chapter 6.

4.4.4 Non-Linear Precoding

It is well-known that if CSI is provided at the transmitter, the employment of Tomlinson-Harashima precoding (THP) can additionally boost the performance, compared to linear precoding [Fis02]. It will be shown now, how the transceiver optimization in this case can be performed if only erroneous CSI is available.

Let the strictly lower triangular $N \times N$ matrix $\boldsymbol{R}$ describe the successive interference cancellation (SIC) process at the transmitter, with $\boldsymbol{G}$ being the linear spatial precoder in the cascade. The linear spatial equalizer of the user k remains to be denoted by $\boldsymbol{F}_k$. A detailed description of the system model for the THP methods can be found in [Fis02, WFVH04]. The MSE of the kth user can be approximately calculated for a system with quadrature amplitude modulation (QAM) as [BD08a]

$$\mathrm{MSE}_k^{\mathrm{THP}} = \|\boldsymbol{F}_k \boldsymbol{H}_k \boldsymbol{G} - [\boldsymbol{R} + \boldsymbol{I}]_k\|_F^2 + \sigma_k^2 \|\boldsymbol{F}_k\|_F^2 \tag{4.65}$$

where $[\boldsymbol{R}+\boldsymbol{I}]_k$ is obtained by taking the rows $1+\sum_{l=1}^{k-1} N_l, \ldots, \sum_{l=1}^{k} N_l$ of $\boldsymbol{R}+\boldsymbol{I}$. Let

$$\boldsymbol{\psi}_k^{\mathrm{THP}}(\boldsymbol{G}, \boldsymbol{F}_k, \boldsymbol{R}) = \operatorname{vec}\left(\boldsymbol{F}_k \hat{\boldsymbol{H}}_k \boldsymbol{G} - [\boldsymbol{R}+\boldsymbol{I}]_k\right). \tag{4.66}$$

The MSE expression (4.65) is equal to

$$\mathrm{MSE}_k^{\mathrm{THP}} = \left\|\boldsymbol{\psi}_k^{\mathrm{THP}}(\boldsymbol{G}, \boldsymbol{F}_k, \boldsymbol{R}) + \boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)\boldsymbol{\delta}_k\right\|_2^2 + \sigma_k^2 \|\boldsymbol{F}_k\|_F^2 \tag{4.67}$$

where $\boldsymbol{\Psi}_k(\boldsymbol{G}, \boldsymbol{F}_k)$ and $\boldsymbol{\delta}_k$ are defined as in (4.23). It is important to notice that $\boldsymbol{\psi}_k^{\mathrm{THP}}(\boldsymbol{G}, \boldsymbol{F}_k, \boldsymbol{R})$ depends affinely on all precoder coefficients, contained in $\boldsymbol{G}$ and $\boldsymbol{R}$, jointly, if $\boldsymbol{F}_k$ is fixed, and vice versa. With the precoding loss neglected, the transmit power of the BS is $P = \|\boldsymbol{G}\|_F^2$. Therefore, the idea from Algorithm 2 is applicable with minor modifications to this case, as well. In the first part of each iteration, the non-linear transmitter can be completely optimized. The update of the receivers can be performed in a minimax fashion, similarly to Theorem 6, and the reasoning behind all proofs remains the same as in Section 4.2.

4.5 Numerical Examples

First, the power minimization problem (4.6) for the ball and the rectangular uncertainty regions is analyzed. To illustrate the performance and show the range of the supported targets, in a 2-user system with parameters $M = 4$, $L_1 = L_2 = 2$, $N_1 = N_2 = 2$, and $\sigma_1^2 = \sigma_2^2 = -20$ dB one fixed set of erroneous users' channels

$$\hat{\boldsymbol{H}}_1 = \begin{bmatrix} -1.87 & 0.50 & 0.38 & 0.01 \\ -0.05 & 1.69 & -1.01 & -0.31 \end{bmatrix}, \quad \hat{\boldsymbol{H}}_2 = \begin{bmatrix} 0.62 & 0.21 & 1.08 & -0.63 \\ 0.79 & 0.23 & -0.13 & -0.55 \end{bmatrix} \tag{4.68}$$

is chosen randomly. In Fig. 4.3, the minimum transmit power versus the MSE target of the first user μ_1, while μ_2 is fixed to -10 dB, is shown. The ball uncertainty model (4.2), with $\varepsilon_1 = \varepsilon_2 = 0.01\sqrt{8}$ (ball 1) and $\varepsilon_1 = \varepsilon_2 = 0.02\sqrt{8}$ (ball 2), is considered, and the results are obtained using Algorithm 2 with the scaled receivers described in Section 4.2.1. These balls can also serve as approximations for the box-like uncertainty models, with $\varepsilon_1 = \varepsilon_2 = 0.01$

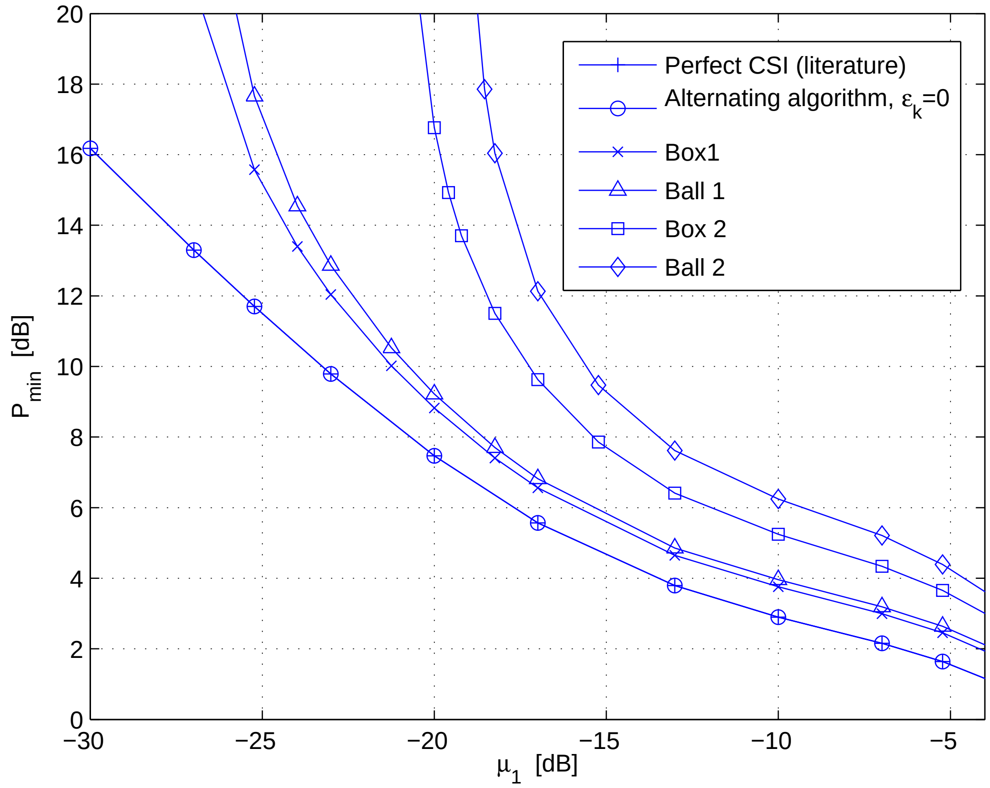

Figure 4.3: Minimal transmit power versus the MSE target μ_1 for the ball and rectangular uncertainty models in a 2-user system with $\mu_2 = -10$ dB.

(box 1) and $\varepsilon_1 = \varepsilon_2 = 0.02$ (box 2) in (4.62). We plot in Fig. 4.3 also the performance for the box models using the computationally involved approach from Section 4.4.3. It should be noticed that the loss in performance when using ball approximations is acceptable considering the drastically reduced complexity. While the performance naturally varies depending on a particular channel realization, qualitatively similar relations among the curves hold for other channel realizations, as well. The best convergence properties are obtained if the matrix parts of the initial receive equalizers, apart from the scaling factors p_k, $k = 1, \ldots, K$, are taken as solutions for the complete receivers with the assumption that $\hat{\boldsymbol{H}}_k$, $k = 1, \ldots, K$, are perfect (these solutions can be found, e.g., in [MJHU06, SSB08a]). In this case, usually just a couple of outer iterations is required for the termination of the algorithms. Finally, from Fig. 4.3 it can be seen that the alternating optimization procedure from Section 4.2 exhibits the same performance as the related schemes from the literature [MJHU06, SSB08a], if the obtained CSI (4.68) is perfect.

Consider now the sum MSE optimization problem (4.8). For the problems (4.8)-(4.10) the feasibility is not an issue, so in principle, any initial receivers $\tilde{\boldsymbol{F}}_k$, $k = 1, \ldots, K$, can be taken. In Fig. 4.4, the sum MSE curves are plotted versus the noise power, which is assumed to be equal for all users. The same 2-user setup is observed, but the results present now an average over 1000 erroneous channel realizations, where real and imaginary parts of each erroneous channel coefficient had normal distribution $\mathcal{N}(0, 1/2)$. A bound for the ball model (4.2), $\varepsilon_1 = \varepsilon_2 = 0.1$, is considered, and the guaranteed sum MSE performance as the outcome of Algorithm 2 is given. A total transmit power constraint $P_{\max} = 1$ (0 dB) is imposed. Besides the erroneous channels, a set of actual worst-case channels from the given uncertainty regions, for a determined transceiver, is calculated, as explained in Section 4.3. Fig. 4.4 shows the performance of the robust SDP approach from Section 4.2.2, both with the receivers that result from Algorithm 2 (notice the exact match of the actual and the guaranteed worst-case performance) and the optimal MMSE receivers (4.11). The results are compared with a non-robust method that takes the obtained, erroneous CSI as perfect [MJHU06, KTA06, SSB07]. In the latter approach, the difference is also made between the feedforward and optimal MMSE receivers (the precoder remains the same). It can be seen that the proposed robust strategies provide significantly better results comparing to the naive, non-robust designs. The sum MSE from Algorithm 2 is also plotted for $\varepsilon_1 = \varepsilon_2 = 0$. The same performance is obtained as for the approaches with perfect CSI from the literature [MJHU06, KTA06, SSB07].

In Fig. 4.5, the average number of outer iterations for the termination of Algorithm 2 for the sum MSE problem, described above, is given. The specified accuracy was $\epsilon = 10^{-4}$. The difference is made between naive initial equalizers $\tilde{\boldsymbol{F}}_1 = \tilde{\boldsymbol{F}}_2 = \boldsymbol{I}$, and heuristic $\tilde{\boldsymbol{F}}_1$ and $\tilde{\boldsymbol{F}}_2$ chosen as if the obtained CSI were perfect (the performance in terms of the obtained sum MSE is observed to be the same for both of these strategies). It can be seen that in the case of smaller disturbances $\varepsilon_1 = \varepsilon_2 = 0.1$ it makes sense to use heuristic initial receivers. However, this advantage is not present always if the disturbances become larger.

For the same scenario as in the sum MSE example, the min-max fairness problem with respect to the streams (4.10) is simulated, as well. This is a reasonable strategy if one aims at minimizing bit or symbol error rates (SERs). The results for the worst-case per-stream MSE and

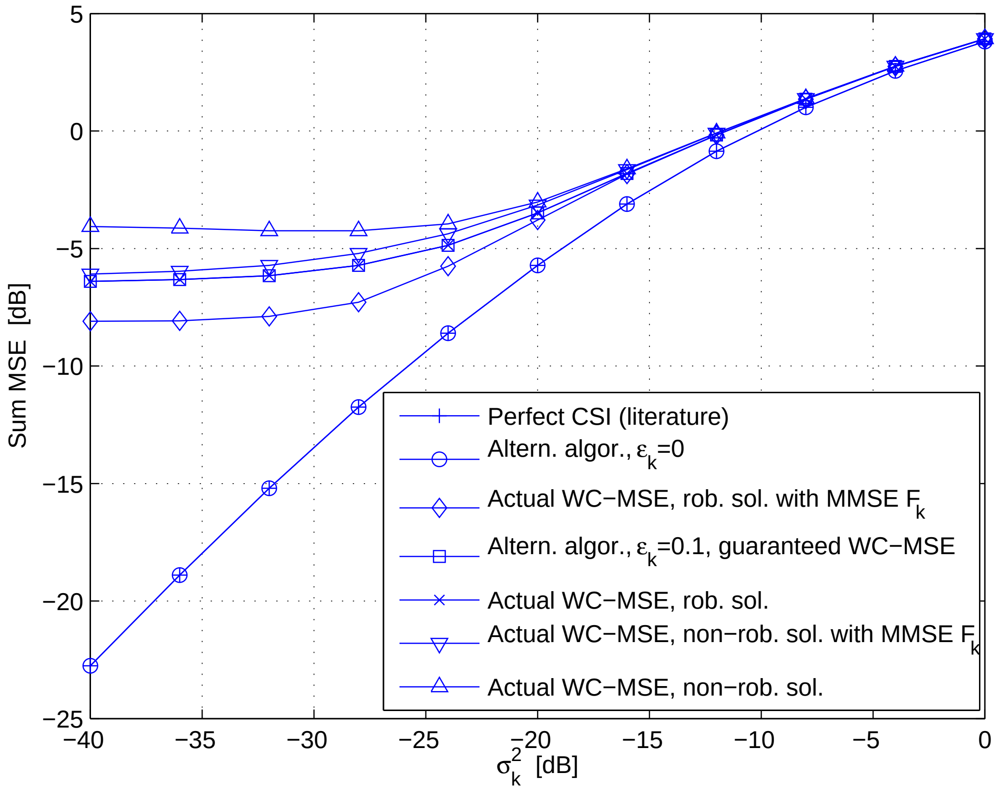

Figure 4.4: Minimal sum MSE versus the noise power for the ball uncertainty model with $\varepsilon_k = 0.1$. Guaranteed and actual sum MSEs for robust and non-robust methods with feedforward and MMSE receivers.

the SERs for the complete system, with the transceiver based on (4.10) and uniform 4-QAM input symbols, are shown in Fig. 4.6 and Fig. 4.7, respectively. Similarly to the sum MSE case, the robust strategy provides a considerably less degraded performance comparing to the non-robust scheme.

Finally, a brief remark concerning the computational complexity of the employed SDP methods is given, having in mind the results from Section 2.1.3. For example, the LMI in (4.24), that contains all K MSE constraints, has K blocks of the size $2(N_k N + ML_k + 1)$. The unknown vector to be determined is of size $2MN + K$, where the first term corresponds to the real and imaginary parts of $\boldsymbol{G}$ and the second represents the K slack variables. The size of the LMI in the update of the receivers (4.27) is also equal to $2(N_k N + ML_k + 1)$, and the size of the vector that contains the optimization variables is $2(L_k N_k + 1)$. The determination of all $\boldsymbol{F}_k$, $k = 1, \dots, K$, can be performed in parallel.

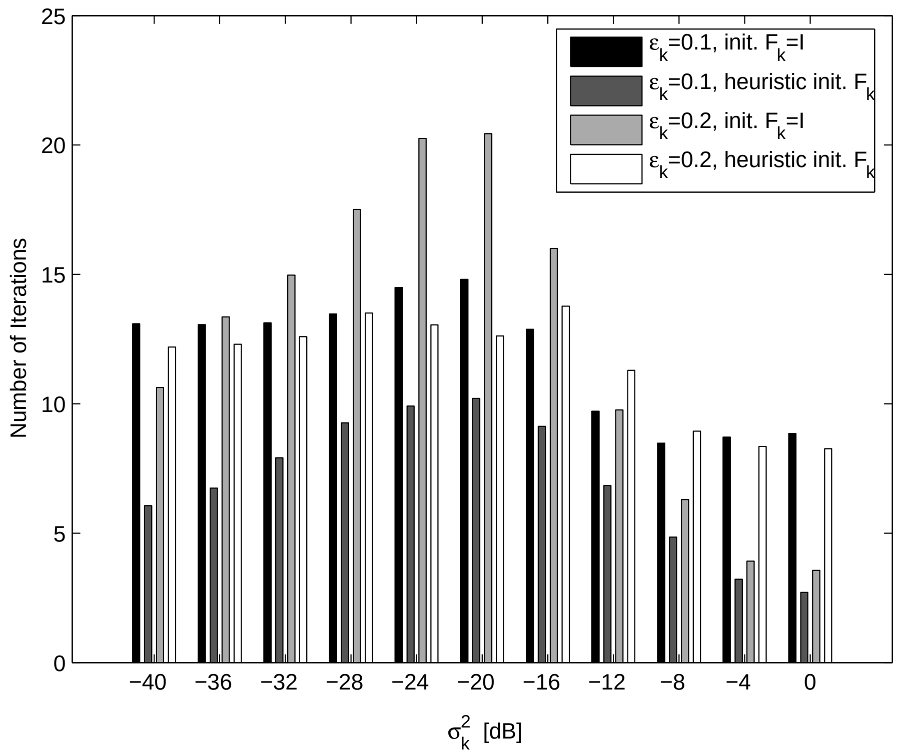

Figure 4.5: Number of outer iterations for the sum MSE problem with $\epsilon = 10^{-4}$. Comparison for $\tilde{\boldsymbol{F}}_k = \boldsymbol{I}$ and $\tilde{\boldsymbol{F}}_k$ taken as if the obtained CSI were perfect (heuristic $\tilde{\boldsymbol{F}}_k$).

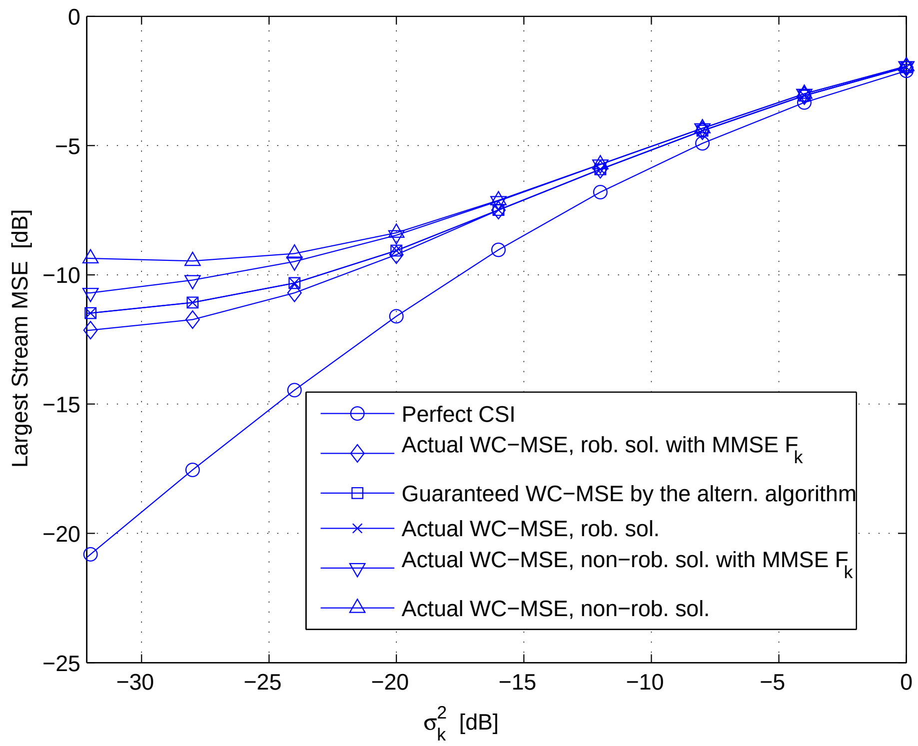

Figure 4.6: Largest per-stream MSE versus the noise power, with $\varepsilon_k = 0.1$. Comparison of robust and non-robust methods, both with feedforward and perfect MMSE receivers.

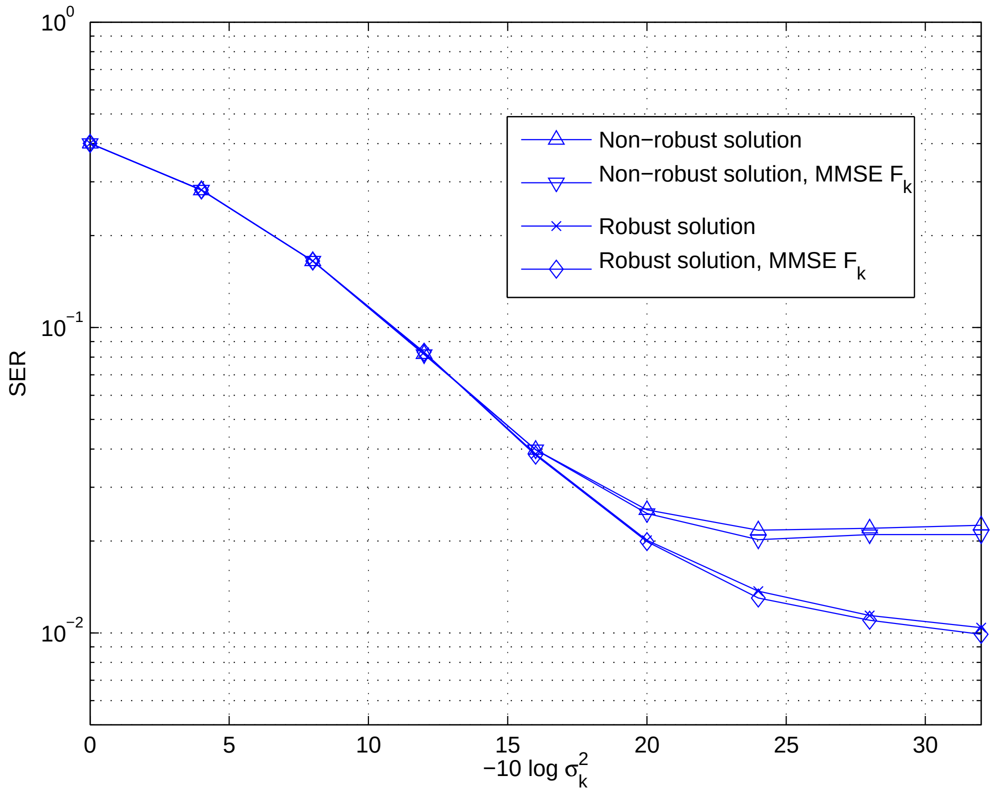

Figure 4.7: SERs in a system with the transceiver based on min-max, per-stream MSE design, 4-QAM input symbols, and $\varepsilon_k = 0.1$. Comparison of robust and non-robust methods, both with feedforward and perfect MMSE receivers.

Chapter 5

Worst-Case Optimization of Frequency Selective MIMO Systems

The task of combating the intersymbol interference (ISI) becomes more involved in multiple antenna systems compared to single-input, single-output (SISO) scenarios, since the spatial dimension of the system must be included, as well. The analysis in the previous two chapters assumed flat-fading channels. The extension for the frequency-selective case is straightforward, if multicarrier techniques (e.g., OFDM) are applied. In this chapter, another model for handling the MIMO-ISI channels in a robust manner will be considered, namely, a direct, spatiotemporal equalization. Such single-carrier schemes might be of interest in situations where OFDM is not a preferable choice, e.g., due to synchronization problems. Furthermore, by equalizing the signals directly, no guard intervals are needed, which increases the overall system throughput.

In MIMO-ISI scenarios, linear transmit filters (precoders) and receive equalizers can be designed to optimize the system, with the MSE taken as the performance measure. Most of the work in the literature assumed that perfect CSI is available at the end of the link where the optimization is performed. With an assumption of knowing perfectly the channel coefficients, if either the receiver or the transmitter is kept fixed, the solutions for the precoder and the receive equalizer, respectively, are readily obtained as Wiener filters. The joint transmit and receive filter optimization is considerably more involved [HCS92, YR94]. The solutions for several scenarios of interest in block-based ISI systems, where inserted additional non-information symbols eliminate the interblock interference, can be found in [SSB+02, LDGW04].

In this chapter, the system model consists of a precoder, MIMO-ISI channel and a receive equalizer. All three elements are causal and have FIRs. Robustness is defined in accordance with the worst-case philosophy, by imposing a bound on the H_2 norm of the error MTF. The main problem of interest will be the minimax design, where the largest MSE with respect to the uncertainty should be minimized. Some extensions, such as the problem of minimizing the total transmit power subject to a strict MSE constraint, per-stream MSEs as the performance measure, and PAPCs are explained, as well. Contrary to the block-based systems, the setup of interest is defined by standard convolutional equations, without any guard intervals.

In this particular setting, the problem of joint transmit and receive filters design is still open even with a perfect CSI assumption. An alternating algorithm, that provides a suboptimal solution in the perfect CSI case, is mentioned already in [HCS92]. In the sequel, a robust variation of the alternating principle, that uses SDP methods similarly to the methodology developed for flat-fading environments in Chapter 4, will be presented.

Finally, it should be remarked that related robust designs for MIMO-ISI channels are studied also in [GL05, GL06]. In these works, the optimization is performed only at one end of the MIMO link, and the filters had infinite impulse responses with no causality constraint imposed. In other words, the solutions are not realizable in practice and must be approximated by cutting and delaying the obtained impulse responses. Furthermore, the calculation of the singular value decomposition of the channel MTF for every frequency is necessary, which significantly increases the computational complexity. The robust optimization in [BEBT07] considered another ISI-suitable model, based on block-circulant matrices. This model, however, does not correspond to the above-explained convolutional setup.

The rest of this chapter is organized in the following way. Section 5.1 describes the MIMO-ISI system model. In Section 5.2, the illustrative minimax problem is defined and SDP-based solutions are provided. Modifications of the method to cover the above-mentioned, related problems are analyzed in Section 5.3. Simulation results that show the benefits obtained by the proposed algorithms are presented in Section 5.4.

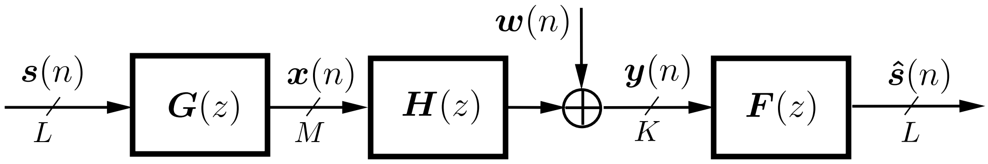

Figure 5.1: Block-scheme of the studied MIMO-ISI system.

5.1 Frequency Selective MIMO System Model

A schematic representation of the considered frequency selective MIMO setup is given in Fig. 5.1. The transmitter is supposed to send a sequence of vectors containing L spatial data streams $\boldsymbol{s}(n) = [s_1(n), \cdots, s_L(n)]^T$, where n denotes the time instant. It will be assumed that the transmit sequence is white, with

$$\mathrm{E}\{\boldsymbol{s}(n)\boldsymbol{s}^*(n-k)\} = \begin{cases} \sigma_s^2 \boldsymbol{I}_L, & k = 0 \\ \boldsymbol{0}_{L\times L}, & k \neq 0. \end{cases} \tag{5.1}$$

The transmitter and the receiver are equipped with M and K antennas, respectively. The linear, causal, FIR filters of the transmitter and the receiver are denoted by $\boldsymbol{G}(z)$ and $\boldsymbol{F}(z)$, and have orders N_G and N_F, respectively:

$$\boldsymbol{G}(z) = \sum_{k=0}^{N_G} \boldsymbol{G}_k z^{-k}, \quad \boldsymbol{F}(z) = \sum_{k=0}^{N_F} \boldsymbol{F}_k z^{-k}. \tag{5.2}$$

The $K \times M$ MIMO-ISI channel is of order N_H, with the MTF

$$\boldsymbol{H}(z) = \hat{\boldsymbol{H}}(z) + \boldsymbol{\Delta}(z). \tag{5.3}$$

The coefficients of

$$\hat{\boldsymbol{H}}(z) = \sum_{k=0}^{N_H} \hat{\boldsymbol{H}}_k z^{-k}, \quad \boldsymbol{\Delta}(z) = \sum_{k=0}^{N_H} \boldsymbol{\Delta}_k z^{-k} \tag{5.4}$$

determine the erroneous channel estimates and disturbances, respectively. It is assumed that the transmitter is provided only with $\hat{\boldsymbol{H}}(z)$ and a bound ε on the H_2 norm of the error MTF

$$||\boldsymbol{\Delta}||_2 \leq \varepsilon \tag{5.5}$$

with

$$\begin{aligned} ||\mathbf{\Delta}||_2 &\triangleq \sqrt{\frac{1}{2\pi}\int_0^{2\pi} \mathrm{Tr}\left\{\mathbf{\Delta}(e^{j\omega})\,\mathbf{\Delta}\,(e^{j\omega})^*\right\} d\omega} \\ &= \sqrt{\sum_{k=0}^{N_H} ||\mathbf{\Delta}_k||_F^2}. \end{aligned} \tag{5.6}$$

The receiver will be supposed to have perfect CSI, although this assumption can be somewhat relaxed, similarly to Chapters 3 and 4, as discussed in Section 5.2. The model (5.5) presents an ISI-suitable extension of the ellipsoidal uncertainty model that is often used for modeling errors in flat-fading channels. A similar error model for broadband channels is used also in [GL05, GL06]. In a practical application, the value ε would be determined by the roughness of channel quantization, or according to the desired outage probability in the system, if the CSI errors are unbounded.

At the receiver, additive white noise $\boldsymbol{w}(n)$ is present, with

$$\mathrm{E}\left\{\boldsymbol{w}(n)\boldsymbol{w}^*(n-k)\right\} = \begin{cases} \sigma_w^2 \boldsymbol{I}_K, & k = 0 \\ \mathbf{0}_{K\times K}, & k \neq 0. \end{cases} \tag{5.7}$$

The transmission of data in the system can be described with the following convolutional equations

$$\begin{aligned} \boldsymbol{x}(n) &= \sum_{k=0}^{N_G} \boldsymbol{G}_k \boldsymbol{s}(n-k) \\ \boldsymbol{y}(n) &= \sum_{k=0}^{N_H} \boldsymbol{H}_k \boldsymbol{x}(n-k) + \boldsymbol{w}(n) \\ \hat{\boldsymbol{s}}(n) &= \sum_{k=0}^{N_F} \boldsymbol{F}_k \boldsymbol{y}(n-k). \end{aligned} \tag{5.8}$$

Finally, as the performance measure, the MSE between the sent and the estimated signal vector

$$\mathrm{MSE} \triangleq \mathrm{E}\left\{||\boldsymbol{s}(n) - \hat{\boldsymbol{s}}(n)||_2^2\right\} \tag{5.9}$$

is initially used.

5.2 Minimax Transceiver Optimization

In this section, the problem of minimax robust design

$$\min_{\boldsymbol{G}(z),\, \boldsymbol{F}(z)} \ \max_{\|\boldsymbol{\Delta}\|_2 \leq \varepsilon} \text{MSE} \quad \text{subject to} \quad P \leq P_{\max} \tag{5.10}$$

is studied. MSE is given by (5.9), and P is the total transmit power that can be calculated as

$$P = \mathrm{E}\left\{\mathrm{Tr}\left\{\boldsymbol{x}(n)\boldsymbol{x}^*(n)\right\}\right\} = \sigma_s^2 \|\boldsymbol{G}\|_2^2 . \tag{5.11}$$

In order to understand the structure of the problem (5.10), the expression for the MSE (5.9) can be transformed in the following way

$$\text{MSE} = \sigma_s^2 \left\|\tilde{\boldsymbol{F}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{G}} - \tilde{\boldsymbol{I}}\right\|_F^2 + \sigma_w^2 \|\boldsymbol{f}\|_2^2 \tag{5.12}$$

where $\tilde{\boldsymbol{F}}$ and $\tilde{\boldsymbol{H}}$ are $(N_F + N_H + N_G + 1)L \times (N_H + N_G + 1)K$ and $(N_H + N_G + 1)K \times (N_G + 1)M$ matrices, respectively

$$\tilde{\boldsymbol{F}} = \begin{bmatrix} \boldsymbol{F}_0 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{F}_{N_F} & \ddots & \boldsymbol{F}_0 \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{F}_{N_F} \end{bmatrix}, \quad \tilde{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{H}_0 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{N_H} & \ddots & \boldsymbol{H}_0 \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{H}_{N_H} \end{bmatrix} \tag{5.13}$$

and where

$$\tilde{\boldsymbol{G}} = \begin{bmatrix} \boldsymbol{G}_0 \\ \vdots \\ \boldsymbol{G}_{N_G} \end{bmatrix}, \quad \tilde{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{I}_L \\ \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{f} = \mathrm{vec}\left(\begin{bmatrix} \boldsymbol{F}_0 \\ \vdots \\ \boldsymbol{F}_{N_F} \end{bmatrix}\right). \tag{5.14}$$

Using (4.21), (5.12) can be rewritten as

$$\begin{aligned} \text{MSE} &= \sigma_s^2 \left\|\mathrm{vec}\left(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{G}} - \tilde{\boldsymbol{I}}\right)\right\|_2^2 + \sigma_w^2 \|\boldsymbol{f}\|_2^2 \\ &= \sigma_s^2 \left\|\left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right)\mathrm{vec}\left(\tilde{\hat{\boldsymbol{H}}}\right) - \mathrm{vec}\left(\tilde{\boldsymbol{I}}\right) + \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right)\mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}\right)\right\|_2^2 + \sigma_w^2 \|\boldsymbol{f}\|_2^2 \end{aligned} \tag{5.15}$$

where $\tilde{\tilde{\boldsymbol{H}}}$ and $\tilde{\boldsymbol{\Delta}}$ are defined analogously to $\tilde{\boldsymbol{H}}$

$$\tilde{\tilde{\boldsymbol{H}}} = \begin{bmatrix} \hat{\boldsymbol{H}}_0 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \hat{\boldsymbol{H}}_{N_H} & \ddots & \hat{\boldsymbol{H}}_0 \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \hat{\boldsymbol{H}}_{N_H} \end{bmatrix}, \quad \tilde{\boldsymbol{\Delta}} = \begin{bmatrix} \boldsymbol{\Delta}_0 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Delta}_{N_H} & \ddots & \boldsymbol{\Delta}_0 \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \boldsymbol{\Delta}_{N_H} \end{bmatrix}. \tag{5.16}$$

Similarly to the last equation in (5.14), the unknown precoder coefficients are grouped into a vector $\boldsymbol{g}$

$$\boldsymbol{g} = \mathrm{vec}\left(\begin{bmatrix} \boldsymbol{G}_0^T & \cdots & \boldsymbol{G}_{N_G}^T \end{bmatrix}^T\right) \tag{5.17}$$

and, for notational convenience, a function $\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f})$ is defined

$$\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) = \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right) \mathrm{vec}\left(\tilde{\tilde{\boldsymbol{H}}}\right) - \mathrm{vec}\left(\tilde{\boldsymbol{I}}\right). \tag{5.18}$$

Finally, after some calculations one obtains the following equivalent expression for the MSE, that will be convenient for the robust analysis

$$\mathrm{MSE} = \sigma_s^2 \,\|\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) + \boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f})\boldsymbol{\delta}\|_2^2 + \sigma_w^2 \,\|\boldsymbol{f}\|_2^2 . \tag{5.19}$$

In (5.19)

$$\boldsymbol{\delta} = \mathrm{vec}\left(\begin{bmatrix} \boldsymbol{\Delta}_0^T & \cdots & \boldsymbol{\Delta}_{N_H}^T \end{bmatrix}^T\right) \tag{5.20}$$

contains the whole uncertainty, and

$$\boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f}) = \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right) \begin{bmatrix} \bar{\boldsymbol{I}} \\ \mathbf{0}_{(N_G+1)K\times(N_H+1)MK} \\ \bar{\boldsymbol{I}} \\ \mathbf{0}_{(N_G+1)K\times(N_H+1)MK} \\ \vdots \\ \bar{\boldsymbol{I}} \end{bmatrix} \tag{5.21}$$

with $((N_H+1)M+N_G(M-1))K \times (N_H+1)MK$ matrix $\bar{\boldsymbol{I}}$

$$\bar{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{I} & \mathbf{0}^{(1)} & \cdots & \mathbf{0}^{(1)} \\ \mathbf{0}^{(2)} & \mathbf{0}^{(2)} & \cdots & \mathbf{0}^{(2)} \\ \mathbf{0}^{(1)} & \boldsymbol{I} & \cdots & \mathbf{0}^{(1)} \\ \mathbf{0}^{(2)} & \mathbf{0}^{(2)} & \cdots & \mathbf{0}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}^{(1)} & \mathbf{0}^{(1)} & \cdots & \boldsymbol{I} \end{bmatrix} \tag{5.22}$$

where $\mathbf{0}^{(1)}$ and $\mathbf{0}^{(2)}$ have dimensions $(N_H+1)K \times (N_H+1)K$ and $N_G K \times (N_H+1)K$, respectively. Consider the formula (5.19) with no uncertainty, i.e., assuming $\boldsymbol{\delta} = \mathbf{0}$. By analyzing (5.18) and (5.21), it can be noticed that (5.19) is not convex in the unknown vectors $\boldsymbol{g}$ and $\boldsymbol{f}$ jointly because of the multiplication via the Kronecker product. However, due to the affinity of $\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f})$ and $\boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f})$ in $\boldsymbol{g}$ and $\boldsymbol{f}$ separately, it is easily seen that (5.19) is convex in $\boldsymbol{g}$, if $\boldsymbol{f}$ is kept fixed, and vice versa. This leads to an idea that the problem can be approached using an iterative procedure, if perfect CSI is provided. Algorithm 3 shows how the same principle is applicable in the robust case $\boldsymbol{\delta} \neq \mathbf{0}$, as well.

Algorithm 3 Iterative solution for the problem (5.10).

1: *initialization:* Set $t = 0$ (the iteration number), the maximal number of iterations $t_{\max}$, the desired accuracy $\epsilon > 0$, the initial $\mathrm{MSE}_{\mathrm{new}} \gg 0$, and the initial receive filter $\boldsymbol{F}(z)$.
2: **repeat**
3: $t \leftarrow t+1$, $\mathrm{MSE}_{\mathrm{old}} \leftarrow \mathrm{MSE}_{\mathrm{new}}$
4: Solve the optimization problem

$$\min_{\boldsymbol{G}(z)} \max_{\|\boldsymbol{\Delta}\|_2 \le \varepsilon} \mathrm{MSE} \qquad \text{subject to} \quad P \le P_{\max} \tag{5.23}$$

with the current values determining $\boldsymbol{F}(z)$ fixed.
5: Update the receive filter by solving the following problem with the fixed transmit filter $\boldsymbol{G}(z)$ obtained from (5.23):

$$\mathrm{MSE}_{\mathrm{new}} = \min_{\boldsymbol{F}(z)} \max_{\|\boldsymbol{\Delta}\|_2 \le \varepsilon} \mathrm{MSE}. \tag{5.24}$$

6: **until** : $t \ge t_{\max}$ or $\mathrm{MSE}_{\mathrm{old}} - \mathrm{MSE}_{\mathrm{new}} \le \epsilon$

In the sequel, it will be proved that the problems (5.23) and (5.24) are convex optimization problems with efficient numerical solutions, and that the algorithm converges. Notice that the bounded uncertainty set (5.5) has an equivalent representation

$$||\boldsymbol{\Delta}||_2 \leq \varepsilon \;\Leftrightarrow\; ||\boldsymbol{\delta}||_2 \leq \varepsilon. \tag{5.25}$$

Lemma 7 enables a transformation of the uncertain problems (5.23) and (5.24) into tractable forms.

Theorem 11. The uncertain problem (5.23) is equivalent to an SDP problem

$$\begin{aligned} \min_{\boldsymbol{g},\tau,\lambda} \quad & \tau + \sigma_w^2 \, ||\boldsymbol{f}||_2^2 \\ \text{subject to} \quad & \begin{bmatrix} \frac{\tau}{\sigma_s^2} - \lambda & \boldsymbol{\psi}^*(\boldsymbol{g},\boldsymbol{f}) & \boldsymbol{0} \\ \boldsymbol{\psi}(\boldsymbol{g},\boldsymbol{f}) & \boldsymbol{I} & -\varepsilon\boldsymbol{\Psi}(\boldsymbol{g},\boldsymbol{f}) \\ \boldsymbol{0} & -\varepsilon\boldsymbol{\Psi}^*(\boldsymbol{g},\boldsymbol{f}) & \lambda\boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0} \\ & \sigma_s \, ||\boldsymbol{g}||_2 \leq \sqrt{P_{\max}} \end{aligned} \tag{5.26}$$

where τ and λ are slack variables.

Proof. The objective function in (5.26) is obtained from (5.19), by introducing a slack variable τ, so that $\tau \geq \sigma_s^2 \, ||\boldsymbol{\psi}(\boldsymbol{g},\boldsymbol{f}) + \boldsymbol{\Psi}(\boldsymbol{g},\boldsymbol{f})\boldsymbol{\delta}||_2^2$. This constraint can be equivalently rewritten using the Schur Complement Lemma as

$$\begin{bmatrix} \frac{\tau}{\sigma_s^2} & \boldsymbol{\psi}^*(\boldsymbol{g},\boldsymbol{f}) \\ \boldsymbol{\psi}(\boldsymbol{g},\boldsymbol{f}) & \boldsymbol{I} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\Psi}(\boldsymbol{g},\boldsymbol{f}) \end{bmatrix} \boldsymbol{\delta} \begin{bmatrix} 1 & \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} 1 \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\delta}^* \begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Psi}^*(\boldsymbol{g},\boldsymbol{f}) \end{bmatrix} \succeq \boldsymbol{0}. \tag{5.27}$$

Now, Lemma 7 can be applied with

$$\boldsymbol{A} = \begin{bmatrix} \frac{\tau}{\sigma_s^2} & \boldsymbol{\psi}^*(\boldsymbol{g},\boldsymbol{f}) \\ \boldsymbol{\psi}(\boldsymbol{g},\boldsymbol{f}) & \boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{D} = \boldsymbol{\delta}^*, \quad \boldsymbol{B} = -\begin{bmatrix} 1 & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Psi}^*(\boldsymbol{g},\boldsymbol{f}) \end{bmatrix}. \tag{5.28}$$

Remember that in (5.23), the receiver, uniquely represented by the vector $\boldsymbol{f}$ is fixed. Since $\boldsymbol{\psi}(\boldsymbol{g},\boldsymbol{f})$ and $\boldsymbol{\Psi}(\boldsymbol{g},\boldsymbol{f})$ are affine in the unknown transmit filter coefficients, grouped in the vector $\boldsymbol{g}$, the first constraint in (5.26) is a convex constraint in the form of an LMI. The total transmit power constraint has an equivalent SOC representation since

$$\sigma_s^2 \|\boldsymbol{G}\|_2^2 \leq P_{\max} \quad \Leftrightarrow \quad \sigma_s \|\boldsymbol{g}\|_2 \leq \sqrt{P_{\max}} \,. \tag{5.29}$$

The objective function is linear, which concludes the proof of the theorem. □

The problem (5.24) can be solved using the same methodology. For an equivalent formulation of (5.24), (5.26) can be used with only two minor differences. Firstly, $\boldsymbol{g}$ will be fixed, and the optimization is over $\boldsymbol{f}$. Secondly, the total transmit power constraint can be omitted, since this condition will always remain satisfied after solving (5.24), because (5.24) updates the receiver only. These modifications clearly do not change the conclusions about the problem's convexity and solvability. Finally, the convergence of Algorithm 3 should be discussed.

Theorem 12. The convergence of Algorithm 3 is guaranteed for any choice of the initial receive equalizer.

Proof. The proof follows the same idea as in the proof of Theorem 9 adopted only to the single-user MIMO-ISI scenario of interest here. □

Although the convergence to the global optimum cannot be claimed, it will be shown in Section 5.4 that the described iterative procedure performs very well in practice and enables significant gains comparing to the optimization of the precoder only. Finally, it should be remarked that Algorithm 3 is meant to be implemented by the transmitter that possesses only imperfect CSI. The receiver either obtains the coefficients of $\boldsymbol{F}(z)$ from the transmitter in an initial phase of communication over one channel using the presumably error-free control channels. Other option is that, after the transmitter determines its precoder, the receiver estimates (perfectly) the equivalent channel $\boldsymbol{H}(z)\boldsymbol{G}(z)$ and adopts the Wiener (MMSE) solution for $\boldsymbol{F}(z)$. The latter approach can only additionally reduce the MSE, similarly to the corresponding discussions in Sections 3.2 and 4.1.

5.3 Related MSE-Optimization Problems

The problem of transmit power minimization subject to a fixed MSE target μ, that must be satisfied despite the uncertainty, is considered first

$$\min_{\boldsymbol{G}(z),\,\boldsymbol{F}(z)} P \qquad \text{subject to} \quad \text{MSE} \leq \mu, \quad \forall \boldsymbol{\Delta} : \ ||\boldsymbol{\Delta}||_2 \leq \varepsilon. \tag{5.30}$$

The problem (5.30) can be approached with a slight modification of the iterative algorithm proposed in Section 5.2.

Theorem 13. The problem (5.30) can be solved using Algorithm 3 with the following changes:

- Except in the problem (5.24), P_{old} and P_{new} should take places of MSE_{old} and MSE_{new}, respectively.
- Instead of (5.23), the optimization problem

$$\text{P}_{\text{new}} \leftarrow \min_{\boldsymbol{G}(z)} P \qquad \text{subject to} \quad \text{MSE} \leq \mu, \quad \forall \boldsymbol{\Delta} : \ ||\boldsymbol{\Delta}||_2 \leq \varepsilon \tag{5.31}$$

  should be solved with a fixed receive equalizer $\boldsymbol{F}(z)$.

The problem (5.31) has an efficiently solvable SDP representation. If the problem (5.31) is feasible in the first step of the first iteration, the convergence of the algorithm is guaranteed.

Proof. An equivalent SDP problem for (5.31)

$$\begin{aligned} \min_{\boldsymbol{g},\lambda} \quad & \sigma_s^2 \, ||\boldsymbol{g}||_2^2 \\ \text{subject to} \quad & \begin{bmatrix} \frac{\mu - \sigma_w^2 ||\boldsymbol{f}||_2^2}{\sigma_s^2} - \lambda & \boldsymbol{\psi}^*(\boldsymbol{g}, \boldsymbol{f}) & \boldsymbol{0} \\ \boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) & \boldsymbol{I} & -\varepsilon \boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f}) \\ \boldsymbol{0} & -\varepsilon \boldsymbol{\Psi}^*(\boldsymbol{g}, \boldsymbol{f}) & \lambda \boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0} \end{aligned} \tag{5.32}$$

is derived in a similar manner as the SDP in Theorem 11. Under a technical condition that the first optimization problem is feasible (otherwise the algorithm stops immediately), the convergence of the algorithm can be proved by noticing that it yields a monotonically decreasing sequence of transmit powers (and the corresponding transceiver coefficients), which is clearly bounded below. Notice that the update of the receivers in (5.24) does not change the transmit power, but it has beneficial influence since it reduces the MSE, having in mind the uncertainty. After one complete iteration, (5.31) operates with an updated receive filter, and the new resulting transmit power must be smaller or equal to the power obtained after the previous solution of (5.31). □

In the balancing problem, the spatial streams are treated individually. The maximal per-stream MSE, $\mathrm{MSE}_l = \mathrm{E}\left\{|s_l(n) - \hat{s}_l(n)|^2\right\}$, $l = 1, \ldots, L$, where $\hat{s}_l(n)$ is defined analogously to $s_l(n)$, is supposed to be minimized taking into account the uncertainty

$$\min_{\boldsymbol{F}(z), \boldsymbol{G}(z)} \max_{\substack{\|\Delta\|_2 \leq \varepsilon \\ 1 \leq l \leq L}} \mathrm{MSE}_l \qquad \text{subject to} \quad P \leq P_{\max}. \tag{5.33}$$

The problem (5.33) is also readily supported with small modifications of the presented framework, similarly to Chapter 4. Per-antenna power constraints are easily included in all problem formulations, since they are SOC constraints $\sigma_s \left\| \begin{bmatrix} \boldsymbol{G}_{0_{(m,:)}} & \cdots & \boldsymbol{G}_{N_{G(m,:)}} \end{bmatrix} \right\|_2 \leq \sqrt{P_m}$, where P_m is the power limit for antenna m, $m = 1, \ldots, M$.

5.4 Numerical Examples

System parameters $\sigma_s^2 = 1$, $L = 2$, $M = 3$, $K = 2$, $N_G = 3$, $N_H = 2$, $N_F = 1$, and $\epsilon = 10^{-3}$, are kept fixed in all simulations. The initial receive equalizer had always $\boldsymbol{F}_0 = \boldsymbol{I}$, and the other taps were zero. It should be remarked that any non-robustly designed transceiver can also be selected as a starting point. The issue of smart initialization is left as an interesting open topic for the future work.

In Fig. 5.2, the resulting minimax MSEs are plotted versus the bound on the uncertainty ε in (5.5), as a solution of Algorithm 3. The simulations present an average over 1000 channel realizations, where real and imaginary parts of each erroneous channel coefficient have distribution $\mathcal{N}(0, 1/2)$. Signal-to-noise ratio (SNR) scenarios 1 and 2, with $P_{\max} = 5$, $\sigma_w^2 = 0.1$ and $P_{\max} = 10$, $\sigma_w^2 = 0.01$, respectively, were considered. It can be seen that with the introduction of

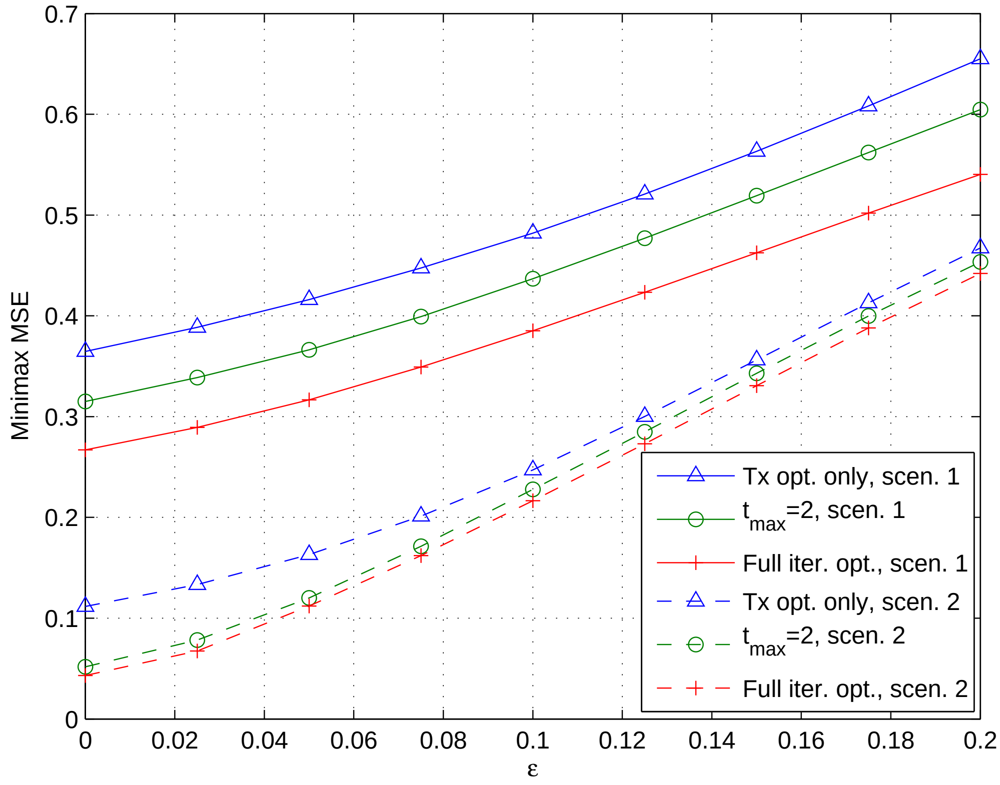

Figure 5.2: Minimax MSE (linear scale) versus the bound on the uncertainty ε. Two SNR scenarios considered for $t_{\max} = 1$ (precoder optimization only), $t_{\max} = 2$, and $t_{\max} = \infty$ in Algorithm 3.

only one additional tap at the receiver, the optimization of the whole system, renders significant gains comparing to the precoder optimization alone.

The necessary number of iterations for the termination of the algorithm ranged from 13.74 to 15.52 for scenario 1, and from 3.54 to 7.32 for scenario 2, depending on ε. From Fig. 5.2, it can be seen that already after the second iteration (or, in other words, with a fixed complexity), the performance improvement is conspicuous. The complexity of the complete iterative algorithm is practically of the same order as the optimization of the transmitter only (one application of (5.26)), however, with a multiplicative constant that corresponds to the number of outer iterations.

Finally, in Fig. 5.3 the benefits in reducing the minimum transmit power as a function of the MSE target μ for the problem (5.30) are plotted, with $\varepsilon = 0.1$ and $\sigma_w^2 = 0.01$. The averaging is performed over 1000 randomly-chosen channels which yielded feasible problems for the

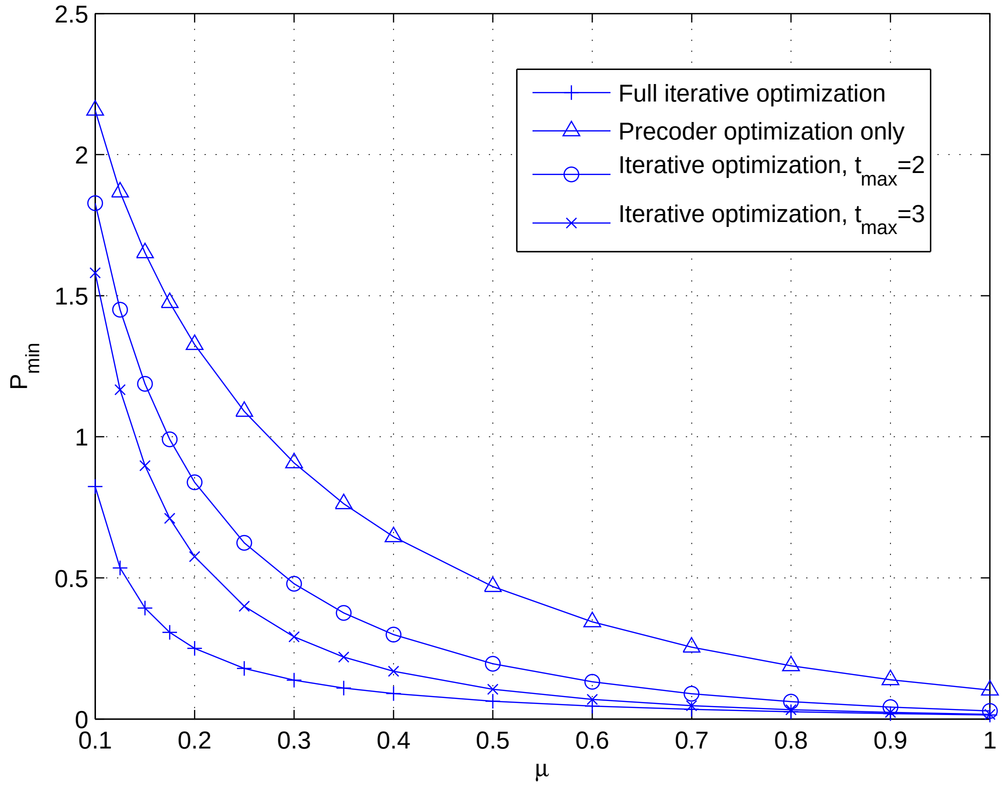

Figure 5.3: Minimal transmit power versus the MSE constraint μ (linear scales) for $t_{\max} \in \{1, 2, 3\}$, and $t_{\max} = \infty$ in Algorithm 3.

observed range of μ.

Chapter 6

Probabilistically Constrained Optimization of Multiuser MISO Systems

Probabilistically constrained methods suppose that CSI errors exhibit certain statistical properties, which is often the case in estimation and quantization procedures. Generally, the statistical information about the CSI mismatch can be used to optimize the mean or the outage performance of the system. As discussed in Section 1.2, for unbounded CSI errors, a desired QoS performance can often be promised only with a certain probability. The main feature of the worst-case optimization, to satisfy strictly some QoS targets for any channel in the uncertainty region, might sometimes present a drawback in practice. Namely, some channels from the uncertainty region, which have low probability of actual appearance, can require a lot of system resources (e.g., transmit power). This makes the stochastic optimization an attractive and natural approach in wireless communications.

The main contribution of this chapter are algorithms for robust power allocation under probabilistic QoS targets, in a downlink flat-fading MISO system. Provided with the error distribution, it is known that the size of the uncertainty region for the worst-case methods can be calculated for this system, so that the QoS targets are satisfied with the desired probability [PPIL07]. It will be shown that this worst-case based approach can be significantly outperformed for a wide range of scenarios, by exploiting the problem structure deeper.

In Section 6.1, a downlink MISO setup with a Gaussian CSI mismatch is introduced. Section

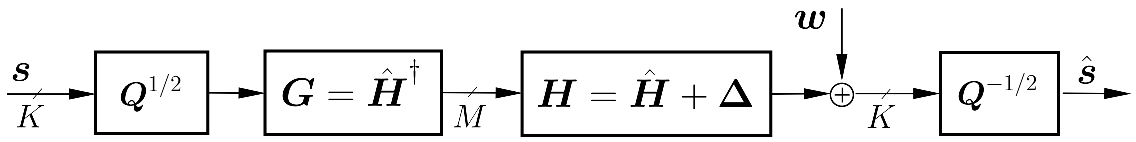

Figure 6.1: Downlink, power-controlled, multiuser MISO system.

6.2 shows how the stochastic uncertainty can be conservatively removed using the Vysochanskii-Petunin inequality in order to obtain a deterministic system optimization problem. The solutions for the latter problem are shown in Section 6.3 using interference functions theory and convex optimization. Finally, in Section 6.4, the uncertainty model is generalized in the sense that no exact distribution of the error is known, but only the covariance matrices of CSI error vectors are available at the transmitter.

6.1 CSI Errors With Gaussian Distribution

The considered multiuser, flat-fading MISO system is illustrated in Fig. 6.1. The BS is equipped with M antennas. It transmits a vector $\boldsymbol{s} = [s_1, \ldots, s_K]^T \in \mathbb{C}^K$, where the symbol s_k is intended for the kth single-antenna user, $k = 1, \ldots, K$. Similarly to Chapter 3, there exists no cooperation among users. The complete downlink channel is denoted with $\boldsymbol{H} \in \mathbb{C}^{K \times M}$. The BS is provided only with an estimate $\hat{\boldsymbol{H}}$ of $\boldsymbol{H}$, where $\hat{\boldsymbol{H}}$ has the full row rank. The CSI error matrix is defined as $\boldsymbol{\Delta} = \boldsymbol{H} - \hat{\boldsymbol{H}}$. It is assumed that

$$\boldsymbol{\Delta}_{(k,:)} = \bar{\boldsymbol{\Delta}}_{(k,:)} \boldsymbol{R}_k, \quad k = 1, \ldots, K \tag{6.1}$$

where $\bar{\boldsymbol{\Delta}}_{(k,m)}$, $m = 1, \ldots, M$, are i.i.d. complex Gaussian random variables (RVs) with independent real and imaginary parts that have $\mathcal{N}(0, 1/2)$-distribution, and $\boldsymbol{R}_k \in \mathbb{C}^{M \times M}$. An application example for the Gaussian error model is a TDD system, where the estimated channel coefficients at the BS from the uplink phase are used for the downlink precoding. Notice that differently from the work in [CSCG07], the focus here is on erroneous channels and not on erroneous channel covariance matrices.

The linear precoder of the BS is composed of two parts. The diagonal power control matrix

$\boldsymbol{Q}^{1/2} = \operatorname{diag}(\sqrt{q_1}, \ldots, \sqrt{q_K})$ is optimized. The transmission matrix $\boldsymbol{G} \in \mathbb{C}^{M \times K}$ is fixed, with $\boldsymbol{G} = \hat{\boldsymbol{H}}^{\dagger}$. In other words, power-controlled, Moore-Penrose pseudoinversion (zero-forcing) with respect to the available imperfect CSI is performed. The users equalize the received signals by a multiplication with $q_k^{-1/2}$, so the system equations are

$$s_k = q_k^{-1/2} \boldsymbol{H}_{(k,:)} \boldsymbol{G} \boldsymbol{Q}^{1/2} \boldsymbol{s} + w_k, \quad k = 1, \ldots, K. \tag{6.2}$$

A similar system model is analyzed with the goal of worst-case optimization in [PPIL07], as mentioned also in Section 3.4. As QoS measures, the MSEs between the sent and the signals after the equalization, defined by (3.3), are initially adopted. For the transmit vector $\boldsymbol{s}$ and the receive noise $\boldsymbol{w} = [w_1, \ldots, w_K]^T$, it is supposed that $\mathrm{E}\{\boldsymbol{s}\boldsymbol{s}^*\} = \boldsymbol{I}$, and $\mathrm{E}\{\boldsymbol{w}\boldsymbol{w}^*\} = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_K^2) \succ \boldsymbol{0}$. The total transmit power P is the sum of the scaled elements of $\boldsymbol{Q}$, and it can be calculated as $P = \operatorname{Tr}(\boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^*)$. The principal problem of interest is the minimization of the total transmit power, subject to fixed MSE targets $\mu_k \in (0, 1)$, that should be satisfied with probabilities $p_k \in (0, 1)$

$$\begin{aligned} \min_{\boldsymbol{Q}} \quad & P \\ \text{subject to} \quad & P \leq P_{\max} \\ & \Pr\{\mathrm{MSE}_k \leq \mu_k\} \geq p_k, \quad k = 1, \ldots, K. \end{aligned} \tag{6.3}$$

A block fading scenario is assumed, and the probability in (6.3) refers to the channel uncertainty. As usual in practical systems, the solution of (6.3) should also obey a total transmit power constraint $P \leq P_{\max}$.

6.2 Elimination of Stochastic Uncertainty

In this section, the accent is put on transforming the chance constraints in (6.3) into tractable deterministic problems. The fact that the QoS measures of interest are quadratic functions, makes the probabilistic optimization problem (6.3) intricate. For the zero-forcing transmission matrix $\boldsymbol{G}$ specified in the previous section, the MSE of the kth user can be calculated as

$$\text{MSE}_k = q_k^{-1}\boldsymbol{\Delta}_{(k,:)}\boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^*\boldsymbol{\Delta}^*_{(k,:)} + q_k^{-1}\sigma_k^2. \tag{6.4}$$

Therefore, the condition $\text{MSE}_k \leq \mu_k$ can be rewritten as

$$X_k \leq \mu_k q_k - \sigma_k^2, \quad X_k \triangleq \overline{\boldsymbol{\Delta}}_{(k,:)}\boldsymbol{G}_k\boldsymbol{Q}\boldsymbol{G}_k^*\overline{\boldsymbol{\Delta}}^*_{(k,:)} \tag{6.5}$$

where $\boldsymbol{G}_k = \boldsymbol{R}_k\boldsymbol{G}$. X_k is a quadratic form in normal variables, and the expression for its probability density function (PDF) is known to be involved [JK70]. Therefore, a conservative tractable approximation of (6.3) is going the be derived. In other words, deterministic constraints, that imply the stochastic constraints in (6.3), will be obtained.

First, it will be proved that X_k is unimodal. In other words, there exists at least one number $m \in \mathbb{R}$, so that the PDF of X_k is non-decreasing on $[0, m)$ and non-increasing on (m, ∞) [DJd88].

Lemma 9. The PDF of X_k is a unimodal function.

Proof. Perform a unitary diagonalization $\boldsymbol{G}_k\boldsymbol{Q}\boldsymbol{G}_k^* = \boldsymbol{U}_k\boldsymbol{\Lambda}_k\boldsymbol{U}_k^*$, with unitary $\boldsymbol{U}_k \in \mathbb{C}^{M\times M}$ and $\boldsymbol{\Lambda}_k = \text{diag}(\lambda_{k,1}, \ldots, \lambda_{k,M})$, $\lambda_{k,1}, \ldots, \lambda_{k,M} \geq 0$. It can be seen that $\boldsymbol{d}_k \triangleq [D_{k,1}, \ldots, D_{k,M}]^T = \boldsymbol{U}_k^*\overline{\boldsymbol{\Delta}}^*_{(k,:)} \in \mathbb{C}^M$ is a vector of i.i.d. zero-mean complex Gaussian variables with independent real and imaginary parts. The standard quadratic form for X_k is then

$$X_k = \sum_{m=1}^{M} \lambda_{k,m}\left(\Re\{D_{k,m}\}^2 + \Im\{D_{k,m}\}^2\right). \tag{6.6}$$

The unimodality of the sum in (6.6) for any $\lambda_{k,1}, \ldots, \lambda_{k,M} \geq 0$ follows directly from Theorem 4 in [SB03]. □

X_k is real-valued, and it can be expressed using the real-valued system representation as a sum of correlated gamma variables

$$X_k = \sum_{l=1}^{2K} q_{\kappa(l)}\left(\tilde{\boldsymbol{G}}_{k_{(l,:)}}\tilde{\boldsymbol{\delta}}_k\right)^2 \tag{6.7}$$

where

$$\kappa(l) = \begin{cases} l, & l \in \{1, \ldots, K\} \\ l - K, & l \in \{K+1, \ldots, 2K\} \end{cases} \tag{6.8}$$

and

$$\tilde{\boldsymbol{G}}_k = \begin{bmatrix} \Re\{\boldsymbol{G}_k^*\} & -\Im\{\boldsymbol{G}_k^*\} \\ \Im\{\boldsymbol{G}_k^*\} & \Re\{\boldsymbol{G}_k^*\} \end{bmatrix}, \quad \tilde{\boldsymbol{\delta}}_k = \begin{bmatrix} \Re\{\bar{\boldsymbol{\Delta}}_{(k,:)}^*\} \\ \Im\{\bar{\boldsymbol{\Delta}}_{(k,:)}^*\} \end{bmatrix}. \tag{6.9}$$

The mean and variance of X_k can be calculated as

$$\mathrm{E}\{X_k\} = \sum_{l=1}^{K} q_l \left\|\boldsymbol{G}_{k_{(:,l)}}\right\|_2^2 \tag{6.10}$$

$$\mathrm{Var}\{X_k\} = \frac{1}{2} \sum_{i=1}^{2K} \sum_{j=1}^{2K} q_{\kappa(i)} q_{\kappa(j)} \left(\tilde{\boldsymbol{G}}_{k_{(i,:)}} \tilde{\boldsymbol{G}}_{k_{(j,:)}}^T\right)^2. \tag{6.11}$$

For the further analysis, a one-sided form of the Vysochanskii-Petunin inequality is derived.

Lemma 10. For a unimodal X and any $a \in \mathbb{R}_{++}$, it holds that

$$\Pr\{X - \mathrm{E}\{X\} \geq a\} \leq \max\left\{\frac{4}{9}\frac{\mathrm{Var}\{X\}}{\mathrm{Var}\{X\} + a^2}, \frac{4}{3}\frac{\mathrm{Var}\{X\}}{\mathrm{Var}\{X\} + a^2} - \frac{1}{3}\right\}. \tag{6.12}$$

Proof. Let $a \in \mathbb{R}_{++}$ and let $y \leq \mathrm{E}\{X\}$ be a deterministic parameter. It holds that

$$\Pr\{X - \mathrm{E}\{X\} \geq a\} \leq \Pr\{|X - y| \geq a + \mathrm{E}\{X\} - y\} \tag{6.13}$$

$$\leq \max\left\{\frac{4}{9}\frac{\mathrm{E}\left\{(X-y)^2\right\}}{(a + \mathrm{E}\{X\} - y)^2}, \frac{4}{3}\frac{\mathrm{E}\left\{(X-y)^2\right\}}{(a + \mathrm{E}\{X\} - y)^2} - \frac{1}{3}\right\} \tag{6.14}$$

where the last relation is the standard, two-sided Vysochanskii-Petunin inequality [VP80, DJd88, Puk94]. Since

$$\mathrm{E}\left\{(X-y)^2\right\} = \mathrm{Var}\{X\} + (\mathrm{E}\{X\} - y)^2 \tag{6.15}$$

and both terms in the maximum in (6.14) are minimized for $y = \mathrm{E}\{X\} - \mathrm{Var}\{X\}/a$, it can be concluded that (6.12) holds. □

The deterministic approximations of the constraints in (6.3) can now be formulated.

Theorem 14. The probabilistic constraints in (6.3) are implied by

$$f(p_k)\left(\mu_k q_k - \sigma_k^2 - \mathrm{E}\{X_k\}\right) \geq \sqrt{\mathrm{Var}\{X_k\}}, \quad k = 1, \ldots, K \tag{6.16}$$

where

$$f(p_k) = \begin{cases} \sqrt{(1-p_k)/(p_k - 5/9)}, & p_k \geq 5/6 \\ \sqrt{(4/3 - p_k)/p_k}, & p_k < 5/6. \end{cases} \tag{6.17}$$

Proof. Let $a_k = \mu_k q_k - \sigma_k^2 - \mathrm{E}\{X_k\}$. Notice that (6.16) implies $a_k > 0$. Using (6.12), the following relation is obtained

$$\begin{aligned} \Pr\{X_k \leq \mu_k q_k - \sigma_k^2\} &= 1 - \Pr\{X_k - \mathrm{E}\{X_k\} \geq a_k\} &(6.18)\\ &\geq 1 - \max\left\{\frac{4}{9}\frac{\mathrm{Var}\{X_k\}}{\mathrm{Var}\{X_k\} + a_k^2}, \frac{4}{3}\frac{\mathrm{Var}\{X_k\}}{\mathrm{Var}\{X_k\} + a_k^2} - \frac{1}{3}\right\} &(6.19) \end{aligned}$$

where in (6.18) the continuity of the PDF of X_k is used [Mos85, AAK01]. If (6.19) is greater or equal to p_k, a safe approximation is obtained in the form of two constraints that must be fulfilled

$$\begin{aligned} \frac{\mathrm{Var}\{X_k\}}{\mathrm{Var}\{X_k\} + a_k^2} &\leq \frac{9}{4}(1 - p_k) &(6.20)\\ \frac{\mathrm{Var}\{X_k\}}{\mathrm{Var}\{X_k\} + a_k^2} &\leq 1 - \frac{3}{4}p_k. &(6.21) \end{aligned}$$

For $p_k \geq 5/6$, the right hand side term in (6.20) is smaller than the right hand side term in (6.21). In this case,

$$a_k \sqrt{(1-p_k)/(p_k - 5/9)} \geq \sqrt{\mathrm{Var}\{X_k\}} \tag{6.22}$$

clearly implies both conditions (6.20) and (6.21), and, correspondingly, the probabilistic MSE constraint for the user k in (6.3). For $p_k < 5/6$, (6.21) is more critical. The probabilistic MSE constraint in (6.3) is conservatively approximated in this case by

$$a_k \sqrt{(4/3 - p_k)/p_k)} \geq \sqrt{\mathrm{Var}\{X_k\}}. \tag{6.23}$$

□

6.3 Iterative Solutions

Equipped with the conservative, deterministic approximation (6.16) of the chance constraints, the original problem (6.3) reduces to

$$\begin{aligned} \min_{\boldsymbol{Q}} \quad & \operatorname{Tr}(\boldsymbol{GQG}^*) \\ \text{subject to} \quad & \operatorname{Tr}(\boldsymbol{GQG}^*) \leq P_{\max} \\ & f(p_k)\left(\mu_k q_k - \sigma_k^2 - \operatorname{E}\{X_k\}\right) \geq \sqrt{\operatorname{Var}\{X_k\}}, \quad k = 1, \ldots, K \end{aligned} \tag{6.24}$$

with q_k being the diagonal elements of the diagonal matrix $\boldsymbol{Q}$, and with the mean and variance of X_k given as functions of $q_1, \ldots, q_K$ in (6.10) and (6.11), respectively. The problem (6.24) has a rich structure which is exploited in the following sections to derive various iterative algorithms that converge to an optimal solution.

6.3.1 Iterative Algorithm Based on Interference Functions Theory

The theory of interference functions enables a design of a simple, efficient, iterative algorithm for finding a global optimum of (6.24). Notice that the objective function in (6.3) can be rewritten as $\sum_{k=1}^{K} q_k \|\boldsymbol{G}_{(:,k)}\|_2^2$. Introduce for convenience new variables

$$\boldsymbol{t} \triangleq [t_1, \ldots, t_K]^T, \qquad t_k \triangleq q_k \left\|\boldsymbol{G}_{(:,k)}\right\|_2^2, \quad k = 1, \ldots, K. \tag{6.25}$$

The equivalent representation of (6.24) is then

$$\begin{aligned} \min_{\boldsymbol{t}} \quad & \sum_{l=1}^{K} t_l \\ \text{subject to} \quad & \sum_{l=1}^{K} t_l \leq P_{\max}, \quad \frac{t_k}{\mathcal{I}_k(\boldsymbol{t})} \geq \beta_k, \quad k = 1, \ldots, K \end{aligned} \tag{6.26}$$

where $\beta_k = \left\|\boldsymbol{G}_{(:,k)}\right\|_2^2 \sigma_k^2 / \mu_k$, and

$$\mathcal{I}_k(\boldsymbol{t}) = 1 + \frac{1}{\sigma_k^2} \sum_{l=1}^{K} t_l \psi_{k,l} + \frac{1}{\sigma_k^2 f(p_k)} \sqrt{\sum_{l=1}^{K} t_l^2 \psi_{k,l}^2 + \sum_{i=1}^{2K-1} \sum_{j=i+1}^{2K} t_{\kappa(i)} t_{\kappa(j)} \eta_{k,i,j}} \tag{6.27}$$

$$\psi_{k,l} = \frac{\left\|\boldsymbol{G}_{k_{(:,l)}}\right\|_2^2}{\left\|\boldsymbol{G}_{(:,l)}\right\|_2^2}, \quad \eta_{k,i,j} = \frac{\left(\tilde{\boldsymbol{G}}_{k_{(i,:)}}\tilde{\boldsymbol{G}}_{k_{(j,:)}}^T\right)^2}{\left\|\boldsymbol{G}_{(:,\kappa(i))}\right\|_2^2 \left\|\boldsymbol{G}_{(:,\kappa(j))}\right\|_2^2}. \tag{6.28}$$

The solution of (6.26) is given by the following theorem.

Theorem 15. Let $\boldsymbol{t}^{(0)} = \boldsymbol{0}$. The iterative procedure

$$\boldsymbol{t}^{(n+1)} = \operatorname{diag}(\beta_1, \ldots, \beta_K)\mathcal{I}\left(\boldsymbol{t}^{(n)}\right) \tag{6.29}$$

with

$$\mathcal{I}(\boldsymbol{t}) \triangleq \begin{bmatrix} \mathcal{I}_1(\boldsymbol{t}) & \cdots & \mathcal{I}_K(\boldsymbol{t}) \end{bmatrix}^T \tag{6.30}$$

yields a component-wise, monotonically increasing sequence $\boldsymbol{t}^{(n)}$. If this sequence converges to a point before exceeding the total transmit power constraint, the convergence point is an optimal point. Otherwise, the infeasibility can be immediately declared.

Proof. It can be easily checked that $\mathcal{I}(\boldsymbol{t})$ satisfies the following properties

$$\mathcal{I}(\boldsymbol{t}) > \boldsymbol{0} \tag{6.31}$$

$$\boldsymbol{t} \geq \boldsymbol{\tau} \quad \Rightarrow \quad \mathcal{I}(\boldsymbol{t}) \geq \mathcal{I}(\boldsymbol{\tau}) \tag{6.32}$$

$$\alpha\mathcal{I}(\boldsymbol{t}) > \mathcal{I}(\alpha\boldsymbol{t}), \quad \forall\alpha > 1 \tag{6.33}$$

where the inequalities between vectors are component-wise. Therefore, $\mathcal{I}_k(\boldsymbol{t})$, $k = 1, \ldots, K$, can be classified as standard interference functions. For such functions, the fixed point iteration (6.29) is known to converge to an optimal solution of the problem (6.26) for any initial vector $\boldsymbol{t}$, if (6.26) is feasible [Yat95, SB05b].

To handle the feasibility issue, notice that (6.31) implies that $\mathcal{I}(\boldsymbol{0}) > \boldsymbol{0}$. Furthermore, it holds that

$$\boldsymbol{t}^{(1)} = \operatorname{diag}(\beta_1, \ldots, \beta_K)\mathcal{I}(\boldsymbol{0}) > \boldsymbol{t}^{(0)} = \boldsymbol{0}. \tag{6.34}$$

Using (6.32), we obtain

$$\boldsymbol{t}^{(2)} = \operatorname{diag}(\beta_1, \ldots, \beta_K)\mathcal{I}\left(\boldsymbol{t}^{(1)}\right) \geq \operatorname{diag}(\beta_1, \ldots, \beta_K)\mathcal{I}\left(\boldsymbol{t}^{(0)}\right) = \boldsymbol{t}^{(1)}. \tag{6.35}$$

The procedure can be continued in the same manner, so the initialization $\boldsymbol{t}^{(0)} = \boldsymbol{0}$ in (6.29) yields $\boldsymbol{t}^{(n+1)} \geq \boldsymbol{t}^{(n)}$, $n \geq 1$.

Suppose now that the problem (6.26) is infeasible, and that there is a convergence point $\boldsymbol{t}^{\star}$ of the fixed-point iteration (6.29) with $\sum_{k=1}^{K} t_k^{\star} \leq P_{\max}$. If this is true, due to the continuity of the interference functions $\mathcal{I}_k(\boldsymbol{t})$ in (6.27), it holds that

$$t_k^{\star} = \lim_{n\to\infty} t_k^{(n)} = \lim_{n\to\infty} t_k^{(n+1)} = \lim_{n\to\infty} \beta_k \mathcal{I}_k\left(\boldsymbol{t}^{(n)}\right) = \beta_k \mathcal{I}_k\left(\boldsymbol{t}^{\star}\right) \tag{6.36}$$

for $k = 1, \ldots, K$. In other words, $\boldsymbol{t}^{\star}$ is a feasible solution, which is a contradiction. Therefore, it can be concluded that there can exist no (finite) convergence point if the problem domain is empty due to the constraints $t_k/\mathcal{I}_k(\boldsymbol{t}) \geq \beta_k$ in (6.26). □

6.3.2 Convexity

In this section, the problem (6.24) is approached from the perspective of convex optimization theory. The convexity of $\sqrt{\mathrm{Var}\{X_k\}}$ with respect to $\boldsymbol{q} \triangleq [q_1, \ldots, q_K]^T$ will be analyzed firstly. Using (6.7) and (6.10), one concludes that

$$X_k - \mathrm{E}\{X_k\} = \boldsymbol{q}^T \boldsymbol{z}_k \tag{6.37}$$

where

$$\begin{aligned} \boldsymbol{z}_k &= \begin{bmatrix} Z_{k,1} & Z_{k,2} & \cdots & Z_{k,K} \end{bmatrix}^T \\ Z_{k,l} &= \left(\tilde{\boldsymbol{G}}_{k(l,:)} \tilde{\boldsymbol{\delta}}_k\right)^2 + \left(\tilde{\boldsymbol{G}}_{k(l+K,:)} \tilde{\boldsymbol{\delta}}_k\right)^2 - \left\|\boldsymbol{G}_{k(:,l)}\right\|_2^2, \quad l = 1, \ldots, K. \end{aligned} \tag{6.38}$$

Similarly to (6.11), the covariance matrix $\mathrm{E}\{\boldsymbol{z}_k \boldsymbol{z}_k^T\}$ can be calculated, as well. This is a positive semidefinite matrix [Pap84], so it has the following representation [Mey00]

$$\mathrm{E}\{\boldsymbol{z}_k \boldsymbol{z}_k^T\} = \boldsymbol{R}_{z_k}^T \boldsymbol{R}_{z_k} \tag{6.39}$$

for some easily computable $\boldsymbol{R}_{z_k} \in \mathbb{R}^{K \times K}$ [GL96].

Therefore, the standard deviation of X_k

$$\begin{aligned}\sqrt{\mathrm{Var}\{X_k\}} &= \sqrt{\mathrm{E}\left\{(X_k - \mathrm{E}\{X_k\})^2\right\}} \\ &= \sqrt{\mathrm{E}\left\{\boldsymbol{q}^T \boldsymbol{z}_k \boldsymbol{z}_k^T \boldsymbol{q}\right\}} = \left\|\boldsymbol{R}_{z_k} \boldsymbol{q}\right\|_2\end{aligned} \tag{6.40}$$

is convex in $\boldsymbol{q}$ as a norm of an affine function [BV04].

Though it does have rather simple iterations, the algorithm (6.29) exhibits typically only linear convergence [BS08b].* The convexity of $\sqrt{\mathrm{Var}\{X_k\}}$ offers several possibilities to formulate alternative algorithms and possibly speed-up the iterative scheme (6.29):

- Since $\sqrt{\mathrm{Var}\{X_k\}}$ is convex in $\boldsymbol{q}$, it follows that the interference functions (6.27) are also convex. The fixed-point iteration (6.29) for convex interference functions can be replaced with the matrix-based approach, which exhibits superlinear convergence [BS08a].

- In the spirit of previous chapters, a conic quadratic problem

$$\begin{aligned}\min_{\boldsymbol{q} \geq \boldsymbol{0}} \quad & \sum_{k=1}^{K} q_k \|\boldsymbol{G}_{(:,k)}\|_2^2 \\ \text{subject to} \quad & f(p_k)\left(\mu_k q_k - \sigma_k^2 - \mathrm{E}\{X_k\}\right) \geq \left\|\boldsymbol{R}_{z_k} \boldsymbol{q}\right\|_2, \quad k = 1, \ldots, K\end{aligned} \tag{6.41}$$

  equivalent to (6.24), can be defined (for simplicity, the total power constraint is omitted in (6.41)). For solving (6.41), efficient interior point methods, with a guaranteed polynomial complexity explained in Section 2.1.2, can be used.

6.3.3 Special Cases and Extensions

In some special cases, there exist closed-form solutions for the problem (6.24). In the sequel, two scenarios which might be of practical interest are analyzed. First, the CSI error for all users will have the same Gaussian distribution. Second scenario covers the case when the covariance matrices of the CSI errors are arbitrarily scaled identity matrices.

Consider the case $\boldsymbol{R}_k = \boldsymbol{R}$, $k = 1, \ldots, K$, in (6.1), where $\boldsymbol{R} \in \mathbb{C}^{M \times M}$ is an arbitrary matrix. Clearly, the RVs X_k in (6.7) all have the same distribution, and their mean values (6.10) and

*See [Ber99] for definitions of convergence rates (linear, superlinear, quadratic) in terms of local analysis, which can be used as alternative measures of computational complexity to the bounds from Section 2.1.

variances (6.11) will also be the same. Similarly, it can be concluded that $\boldsymbol{R}_{z_k} = \boldsymbol{R}_z$ for all k, in the equivalent reformulation (6.41). From the discussion about the fixed-point iteration in Section 6.3.1, it can be concluded that in the optimum, all constraints in (6.41) are active (fulfilled with equality). Let

$$a = \mathrm{E}\{X_k\} = \sum_{l=1}^{K} q_l \|\boldsymbol{R}\boldsymbol{G}_{(:,l)}\|_2^2 \quad (6.42)$$

$$b = \sqrt{\mathrm{Var}\{X_k\}} = \|\boldsymbol{R}_z \boldsymbol{q}\|_2 . \quad (6.43)$$

The active constraints in (6.41) reduce to

$$f(p_k)\left(\mu_k q_k - \sigma_k^2 - a\right) = b, \quad k = 1, \ldots, K. \quad (6.44)$$

From (6.44), q_k can be expressed as a function of the introduced variables a and b

$$q_k = \frac{1}{\mu_k}\left(a + \sigma_k^2 + \frac{b}{f(p_k)}\right), \quad k = 1, \ldots, K. \quad (6.45)$$

By inserting q_k from (6.45) into (6.42), linear dependence of a on b is obtained

$$a = \left(1 - \sum_{l=1}^{K} \frac{\left\|\boldsymbol{R}\boldsymbol{G}_{(:,l)}\right\|_2^2}{\mu_l}\right)^{-1} \sum_{m=1}^{K} \frac{\sigma_m^2 + b/f(p_m)}{\mu_m} \left\|\boldsymbol{R}\boldsymbol{G}_{(:,m)}\right\|_2^2 . \quad (6.46)$$

Therefore, q_k is a linear function in b. If both sides in (6.43) are squared, only a scalar, quadratic equation with respect to b remains to be solved.

Identity CSI error covariance matrices can be modeled by taking $\boldsymbol{R}_k = \theta_k \boldsymbol{I}_M$ in (6.1). Let

$$u = \sum_{l=1}^{K} q_l \|\boldsymbol{G}_{(:,l)}\|_2^2 \quad (6.47)$$

$$v = \sqrt{\frac{1}{2}\sum_{i=1}^{2K}\sum_{j=1}^{2K} q_{\kappa(i)} q_{\kappa(j)} \left(\tilde{\boldsymbol{G}}_{(i,:)}\left(\tilde{\boldsymbol{G}}_{(j,:)}\right)^T\right)^2} \quad (6.48)$$

where

$$\tilde{\boldsymbol{G}} = \begin{bmatrix} \Re\{\boldsymbol{G}^*\} & -\Im\{\boldsymbol{G}^*\} \\ \Im\{\boldsymbol{G}^*\} & \Re\{\boldsymbol{G}^*\} \end{bmatrix}. \quad (6.49)$$

From (6.10) and (6.11), it holds that

$$\mathrm{E}\{X_k\} = u\theta_k^2, \qquad \sqrt{\mathrm{Var}\{X_k\}} = v\theta_k^2 \tag{6.50}$$

and a similar procedure as for the case of identical distributions can be performed.

Finally, the method presented in the previous sections can be easily modified to cover several interesting related problems, such as the QoS targets in terms of SINRs, or the min-max fairness problem.

Minimum tolerable SINR targets present often a more convenient performance measure than the MSEs. From Theorem 4, it can be concluded that guaranteeing certain MSE targets μ_k under uncertainty assures that the SINR targets $\gamma_k = \mu_k^{-1} - 1$ are fulfilled with the same transmit filter, as well (remember that for SINRs, the users' equalization plays no role). In this way, an SINR-constrained problem of minimizing the total transmit power subject to probabilistic constraints $\Pr\{\mathrm{SINR}_k \geq \gamma_k\} \geq p_k$ can be conservatively approached by defining a virtual MSE-constrained problem (6.3) with $\mu_k = (\gamma_k + 1)^{-1}$.

In the min-max fairness problem, we are interested in minimizing the "balanced" level of the users' MSEs μ, having in mind that a certain outage probability is allowed

$$\begin{aligned} \min_{\boldsymbol{Q},\mu} \quad & \mu \\ \text{subject to} \quad & \Pr\{\mathrm{MSE}_k \leq \mu\} \geq p_k, \quad k = 1, \ldots, K, \quad P \leq P_{\max}. \end{aligned} \tag{6.51}$$

This problem might be of particular practical importance, since the goal of the system design is often to minimize the error under some given power budget. The problem (6.51) can be solved using the bisection method with respect to the level μ. The starting search interval for μ is $[0, 1]$. In each step of the bisection procedure, for a given μ, it should be determined whether (6.51) is feasible. This is performed easily, either according to the interference functions based algorithm from Section 6.3.1, or by analyzing the constraints in the conic quadratic problem (6.41). Notice that, contrary to (6.3), the complete problem (6.51) can never be infeasible.

6.4 Probabilistic Constraints With Unknown Distribution of CSI Errors

In this section, a related problem of worst-case PDF optimization is going to be considered. The power minimization problem with probabilistic constraints (6.3) is studied for the same system model as in Section 6.1, except that the distribution of the CSI errors (6.1) is not known anymore. It is assumed that the following relations hold for $\boldsymbol{\Delta}_{(k,:)}$

$$\mathrm{E}\left\{\boldsymbol{\delta}_k\right\} = \mathbf{0}, \quad \mathrm{E}\left\{\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T\right\} = \tilde{\boldsymbol{C}}_k, \quad \boldsymbol{\delta}_k \triangleq \begin{bmatrix} \Re\{\boldsymbol{\Delta}^*_{(k,:)}\} \\ \Im\{\boldsymbol{\Delta}^*_{(k,:)}\} \end{bmatrix}, \quad k = 1, \ldots, K. \tag{6.52}$$

Clearly, this is a more general model comparing with Section 6.1. Its importance lies in the fact that in practice it might be much easier to estimate the error covariance matrix instead of the exact distribution. The probabilistic, worst-case PDF problem can now be formulated as

$$\begin{aligned} \min_{\boldsymbol{Q}} \quad & P \\ \text{subject to} \quad & \Pr\left\{\mathrm{MSE}_k(\boldsymbol{\delta}_k, \boldsymbol{Q}) < \mu_k\right\} \geq p_k \\ & \forall \boldsymbol{\delta}_k : \ \mathrm{E}\left\{\boldsymbol{\delta}_k\right\} = \mathbf{0}, \quad \mathrm{E}\left\{\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T\right\} = \tilde{\boldsymbol{C}}_k, \quad k = 1, \ldots, K. \end{aligned} \tag{6.53}$$

In other words, the probabilistic constraints are satisfied for any distribution of errors with the given covariance matrices. This explains also the term worst-case in the name of this approach.

The following lemma is a powerful result which can help in solving (6.53).

Lemma 11. Let $\mathcal{M}$ be a set in $\mathbb{R}^n$ defined as

$$\mathcal{M} = \left\{\boldsymbol{x} \in \mathbb{R}^n : \ \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} + 2\boldsymbol{b}_i^T\boldsymbol{x} + c_i < 0, \ i = 1, \ldots, m\right\} \tag{6.54}$$

where $\boldsymbol{A}_i = \boldsymbol{A}_i^T \in \mathbb{R}^{n\times n}$, $\boldsymbol{b}_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, and let

$$I(\mathcal{M}, \bar{\boldsymbol{x}}, \boldsymbol{C}) = \inf\left\{\Pr\{\boldsymbol{x} \in \mathcal{M}\} : \ \mathrm{E}\{\boldsymbol{x}\} = \bar{\boldsymbol{x}}, \ \mathrm{E}\left\{\boldsymbol{x}\boldsymbol{x}^T\right\} = \boldsymbol{C}\right\} \tag{6.55}$$

where the infimum is over all probability distributions on $\mathbb{R}^n$ that have the specified mean value and covariance matrix. Formulate the following SDPs

$$\begin{aligned} \min \quad & 1 - \sum_{i=1}^{m} \lambda_i \\ \text{subject to} \quad & \mathrm{Tr}\{\boldsymbol{A}_i \boldsymbol{Z}_i\} + 2\boldsymbol{b}_i^T \boldsymbol{z}_i + c_i \lambda_i \geq 0, \quad i = 1, \ldots, m \\ & \sum_{i=1}^{m} \begin{bmatrix} \boldsymbol{Z}_i & \boldsymbol{z}_i \\ \boldsymbol{z}_i^T & \lambda_i \end{bmatrix} \preceq \begin{bmatrix} \boldsymbol{C} & \bar{\boldsymbol{x}} \\ \bar{\boldsymbol{x}}^T & 1 \end{bmatrix} \\ & \begin{bmatrix} \boldsymbol{Z}_i & \boldsymbol{z}_i \\ \boldsymbol{z}_i^T & \lambda_i \end{bmatrix} \succeq \boldsymbol{0}, \quad i = 1, \ldots, m \end{aligned} \tag{6.56}$$

with the variables $\boldsymbol{Z}_i \in \mathbb{R}^{n \times n}$, $\boldsymbol{z}_i \in \mathbb{R}^n$, $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, m$, and

$$\begin{aligned} \max \quad & 1 - \mathrm{Tr}\{\boldsymbol{C}\boldsymbol{P}\} - 2\boldsymbol{y}^T \bar{\boldsymbol{x}} - r \\ \text{subject to} \quad & \begin{bmatrix} \boldsymbol{P} - \tau_i \boldsymbol{A}_i & \boldsymbol{y} - \tau_i \boldsymbol{b}_i \\ (\boldsymbol{y} - \tau_i \boldsymbol{b}_i)^T & r - 1 - \tau_i c_i \end{bmatrix} \succeq \boldsymbol{0}, \quad i = 1, \ldots, m \\ & \tau_i \geq 0, \quad i = 1, \ldots, m \\ & \begin{bmatrix} \boldsymbol{P} & \boldsymbol{y} \\ \boldsymbol{y}^T & r \end{bmatrix} \succeq \boldsymbol{0} \end{aligned} \tag{6.57}$$

with the variables $\boldsymbol{P} \in \mathbb{R}^{n \times n}$, $\boldsymbol{y} \in \mathbb{R}^n$, $r \in \mathbb{R}$, and $\tau_i \in \mathbb{R}$, $i = 1, \ldots, m$. The optimal values of both SDPs (6.56) and (6.57) are equal to $I(\mathcal{M}, \bar{\boldsymbol{x}}, \boldsymbol{C})$.

Proof. The proof can be found in [VBC07] (see also [VENG06, VEG06, VCG08] for some similar results). □

The problem (6.53) can now be rewritten as an equivalent linear program, which admits a direct, closed-form solution.

Theorem 16. The problem (6.53) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{q}\geq\boldsymbol{0}} \quad & \sum_{k=1}^{K} q_k\|\boldsymbol{G}_{(:,k)}\|_2^2 \\ \text{subject to} \quad & \sum_{l=1}^{K}\left(q_l\left(\tilde{\boldsymbol{G}}_{(l,:)}\tilde{\boldsymbol{C}}_k\tilde{\boldsymbol{G}}_{(l,:)}^T+\tilde{\boldsymbol{G}}_{(l+K,:)}\tilde{\boldsymbol{C}}_k\tilde{\boldsymbol{G}}_{(l+K,:)}^T\right)\right)+(1-p_k)(\sigma_k^2-\mu_k q_k)=0,\ k=1,\dots,K \end{aligned} \tag{6.58}$$

with $\tilde{\boldsymbol{G}}$ given by (6.49).

Proof. Notice that X_k in (6.5) can be rewritten as (see also (6.7))

$$X_k=\boldsymbol{\delta}_k^T\left(\sum_{l=1}^{K}q_l\left(\tilde{\boldsymbol{G}}_{(l,:)}^T\tilde{\boldsymbol{G}}_{(l,:)}+\tilde{\boldsymbol{G}}_{(l+K,:)}^T\tilde{\boldsymbol{G}}_{(l+K,:)}\right)\right)\boldsymbol{\delta}_k. \tag{6.59}$$

Apply now the SDP (6.57) from Lemma 11, on the kth user constraint in (6.53) with

$$\boldsymbol{x} = \boldsymbol{\delta}_k, \quad \overline{\boldsymbol{x}}=\boldsymbol{0} \tag{6.60}$$
$$\boldsymbol{A} = \sum_{l=1}^{K}q_l\left(\tilde{\boldsymbol{G}}_{(l,:)}^T\tilde{\boldsymbol{G}}_{(l,:)}+\tilde{\boldsymbol{G}}_{(l+K,:)}^T\tilde{\boldsymbol{G}}_{(l+K,:)}\right) \tag{6.61}$$
$$\boldsymbol{b} = \boldsymbol{0} \tag{6.62}$$
$$c = \sigma_k^2-\mu_k q_k \tag{6.63}$$
$$\boldsymbol{C} = \tilde{\boldsymbol{C}}_k \tag{6.64}$$

in (6.54) (the indexes are omitted since $m = 1$). Having in mind (6.62), for maximizing the objective function in (6.57), $r = 0$ can be taken (the last equation in (6.57) requires $r \geq 0$). From the last equation in (6.57), it follows also that $\boldsymbol{y} = \boldsymbol{0}$ and the second condition in (6.57) reduces to

$$\boldsymbol{P} \geq \tau\boldsymbol{A} \tag{6.65}$$
$$-1-\tau c \geq 0. \tag{6.66}$$

Using Lemma 5 combined with the fact that in problem (6.53) $c < 0$ must hold, it can be concluded that in the optimum of (6.57) the following relations are valid

$$\tau = -c^{-1}, \quad \boldsymbol{P} = \tau \boldsymbol{A}. \tag{6.67}$$

Finally, the solution for $\boldsymbol{P}$ from (6.67) can be inserted into the objective function of (6.57). The requirement that the optimal value of problem (6.57) is greater or equal to p_k yields a condition for user k

$$1 - \frac{1}{\mu_k q_k - \sigma_k^2} \mathrm{Tr}\left\{\tilde{\boldsymbol{C}}_k \boldsymbol{A}\right\} \geq p_k \tag{6.68}$$

or, equivalently

$$(1 - p_k)(\mu_k q_k - \sigma_k^2) \geq \sum_{l=1}^{K} q_l \left(\tilde{\boldsymbol{G}}_{(l,:)} \tilde{\boldsymbol{C}}_k \tilde{\boldsymbol{G}}_{(l,:)}^T + \tilde{\boldsymbol{G}}_{(l+K,:)} \tilde{\boldsymbol{C}}_k \tilde{\boldsymbol{G}}_{(l+K,:)}^T\right). \tag{6.69}$$

The activity of the constraints in (6.69) in the optimum, can be concluded, e.g., by noticing the same problem structure in the sense of interference functions as in Section 6.3.1. □

The problem (6.58) is clearly solvable directly by matrix inversion.

6.5 Numerical Examples

In the first experiment, the number of feasible channel realizations was counted for the problem (6.24) in systems with $M = K \in \{2, 3, 4\}$. The parameters $p_k = 0.95$, $\mu_k = -10\,\mathrm{dB}$, $\sigma_k^2 = 10^{-3}$, $\boldsymbol{R}_k = 2\sigma_k \boldsymbol{I}$, $k = 1, \ldots, K$, were fixed for all users. 1000 erroneous channel coefficients were generated as complex normal variables with zero mean and unit variance. Fig. 6.2 compares the feasibility of four schemes. The performance of the method from Section 6.3 is compared with the worst-case based method. In the latter approach, under the assumptions stated above, the radius of the ball uncertainty region, that guarantees the targets with the probability p_k, can be simply calculated as $\varepsilon_k = \sqrt{2}\sigma_k \sqrt{F^{-1}_{\chi^2,2M}(p_k)}$, with $F_{\chi^2,2M}$ being the cumulative density function of the central chi-square distribution with $2M$ degrees of freedom. The optimal solution for the worst-case power control, whose derivation was the subject of [PPIL07], can be applied then. The result for the worst-case PDF approach from Section 6.4, with the same CSI error

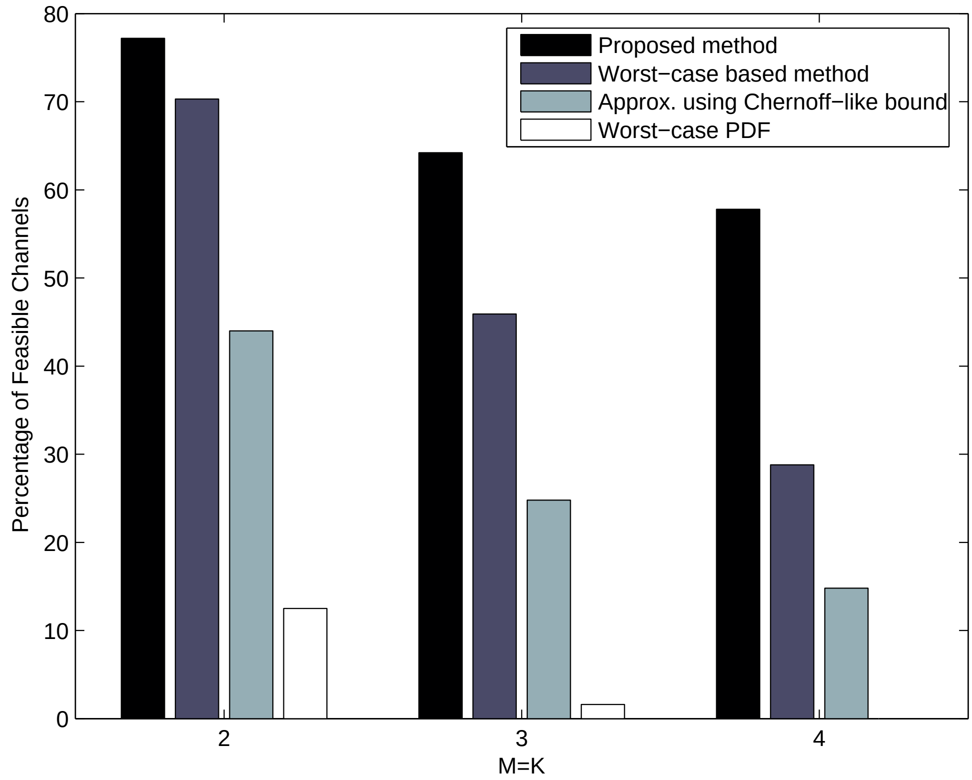

Figure 6.2: Percentage of feasible channel realizations. Probability of fulfillment of the MSE targets $\mu_k = -10\,\text{dB}$ set to 0.95 for all users.

covariance matrix, is shown, as well. Finally, the method based on an approximation using the Chernoff-like bound from [VEG06] can also be applied to conservatively solve the problem (6.3). It can be seen that the iterative algorithms from Section 6.3 significantly outperform other strategies. The worst-case PDF solution, although the only non-conservative one, suffers greatly from unknowing the error distribution. As far as the computational complexity is concerned, in the general case of arbitrary error covariance matrices, the worst-case based method [PPIL07] and the approach from [VEG06] are the most involved since they resort to SDP optimization.†

The proposed method from Section 6.3 and the most competitive, worst-case based method are compared in the sequel. A system with parameters $M = K = 3$, $\mu_1 = \mu_2 = -10\,\text{dB}$ is considered in Fig. 6.3. The average minimum transmit power against the target μ_3 is plotted

†It should be remarked that [VEG06] is a further approximation of the result in [VENG05]. The latter method is even more complex since it has additional non-linear convex constraints. Its performance is also significantly below the method from Section 6.3. It should be remarked, however, that the algorithms in [VENG05] and [VEG06] are derived for more general quadratic constraints than (6.5), and that the problems there do not exhibit the specific structure of (6.3).

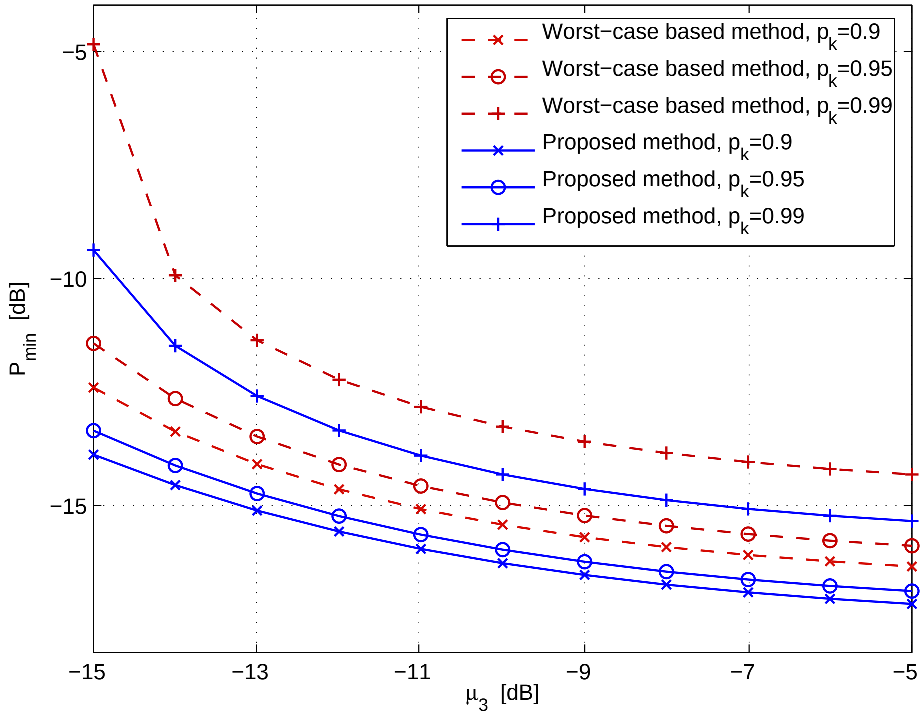

Figure 6.3: Minimal transmit power versus the MSE target μ_3, with $\mu_1 = \mu_2 = -10\,$dB.

for three scenarios regarding the probabilistic constraints: $p_k = 0.99$, $\forall k$, $p_k = 0.95$, $\forall k$, and $p_k = 0.9$, $\forall k$. Other system parameters are the same as in the previous example. The results present an average over 1000 feasible erroneous channels. It can be noticed that the worst-case based strategy performs substantially poorer than the algorithms from Section 6.3. As expected, for both strategies, the required transmit power decreases with increasing the outage probabilities $1 - p_k$ or the MSE target μ_3.

The computational complexity of all studied methods, except the fixed-point iteration (6.29), can be theoretically analyzed using the results from Section 2.1. Fig. 6.4 illustrates the complexity of (6.29) by showing the average number of iterations for the termination of algorithm when solving the power minimization problem (6.26). The relative difference of 10^{-3} for the norm of $\boldsymbol{t}$ was taken as the stopping criterion. Other system parameters are the same as in the previous example.

Finally, the gains over the worst-case based strategy for the min-max fairness problem (6.51) are shown in Table 6.1. The average value of the min-max level μ is given for a system with $M = K = 3$, $p_k = 0.95$, $\sigma_k^2 = 10^{-3}$, $\boldsymbol{R}_k = \sqrt{2}\sigma_\delta \boldsymbol{I}$, $k = 1, \ldots, K$.

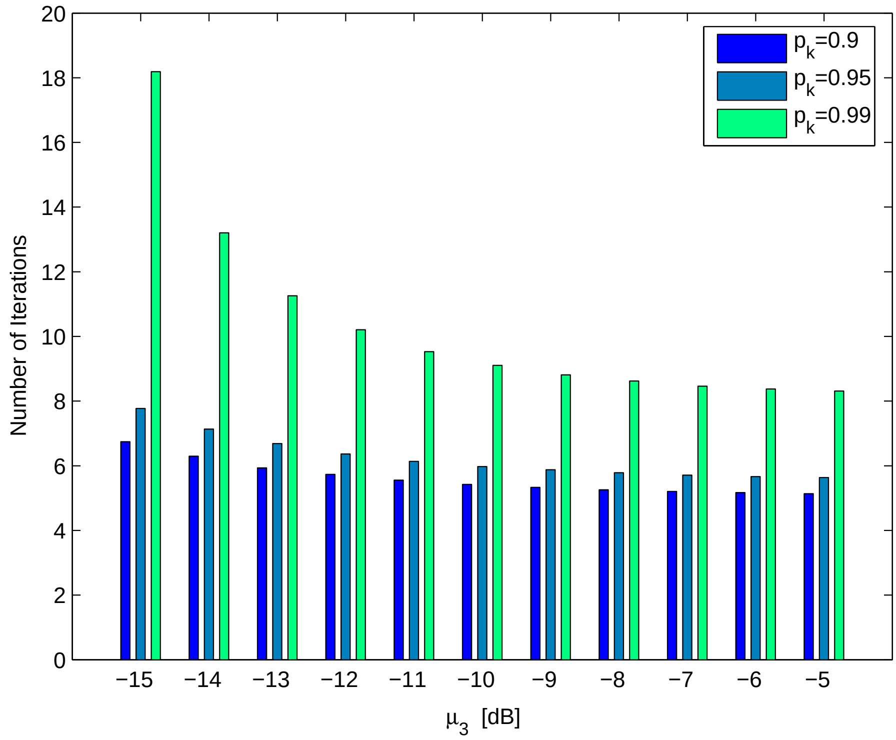

Figure 6.4: Number of iterations for the termination of algorithm (6.29) in a 3-user system with targets $\mu_1 = \mu_2 = -10\,\text{dB}$.

$\sigma_\delta^2/10^{-3}$	1	2	4	8
Worst-Case Based Method	0.0193	0.0325	0.0651	0.1301
Proposed Method	0.0106	0.0209	0.0417	0.0832

Table 6.1: Min-max MSE μ in a 3-user MISO system with $p_k = 0.95$.

It should be remarked that only in "small" systems of size $M \leq 3$, $1 < K \leq M$, with very high probabilities $p_k \rightarrow 1$ (these are often impractical because the feasibility region becomes considerably smaller), the worst-case approach exhibited a somewhat better performance. In all other scenarios, both regarding the system size and the probabilities p_k, the method based on the Vysochanskii-Petunin inequality yielded significant gains. An interesting possibility for the performance improvement even in the region of very high probabilities might be to use the Vysochanskii-Petunin inequality for higher moments ([DJd88], Section 1.5)

$$\Pr\{|X-b| \geq a\} \leq \max\left\{\frac{n\mathrm{E}\{|X-b|^r\} - a^r}{(n-1)a^r}, \frac{r^r\mathrm{E}\{|X-b|^r\}}{(r+1)^r a^r}\right\} \tag{6.70}$$

where $b \in \mathbb{R}$, $r > 0$, $a > 0$, and n is the unique number that satisfies the conditions

$$n > r + 1, \quad n(n - r - 1)^r = r^r. \tag{6.71}$$

However, it should be noticed that the employment of (6.70) results in an increased computational complexity.

Chapter 7

Conclusion

7.1 Summary of Contributions

In the thesis, algorithms for transceiver design in multiple antenna systems with channel uncertainty were considered. In the course of the research, the following results were obtained:

- The problem of complete linear transceiver optimization in the flat-fading, downlink MISO system, for total transmit power minimization with strict MSE targets and ellipsoidal channel uncertainty regions, was optimally solved using an equivalent SDP reformulation. Thereby is the performance of the methods from the literature, which optimized the system only partially, significantly outperformed. The proposed method exhibits a rather small increase in computational complexity over the existing approaches.

- The conservative, SINR-constrained, downlink MISO precoding problem, with imperfect CSI at the transmitter, was optimally solved using the ellipsoid method. The performance of the latter, academic approach was reached in practice using the derived virtual MSE-constrained problem and the corresponding, efficiently solvable, SDP representation. The obtained SDP solution performs substantially better and has a lower complexity in comparison with the state-of-the-art approaches from the literature for this problem.

- An alternating transmit/receive optimization scheme, with the iterations resolved using SDP methods, was proposed for a number of worst-case optimization tasks in downlink multiuser MIMO systems with imperfect CSI. The algorithm is able to support transmit power minimization with strict per-user or per-stream MSE targets, sum MSE minimization, and min-max per-user or per-stream MSE problems. Extensions for supporting uncertain knowledge of the noise covariance, non-linear precoding, and various types of transmit power constraints were derived, as well.

- An algorithm for calculating efficiently the actual worst-case channels in a general, multiuser MIMO scenario with ellipsoidal channel uncertainty sets, was derived using a strong duality result for trust region subproblems.

- The alternating principle was successfully applied in analyzing robust counterparts of several MSE-constrained, transceiver optimization problems in frequency selective MIMO systems with spatiotemporal (time domain) equalization.

- Downlink MISO systems with power controlled zero-forcing and a Gaussian CSI mismatch at the transmitter were also analyzed. A conservative solution for the chance constrained, power control problem, where the QoS constraints were satisfied with certain probabilities, was obtained. Iterative algorithms, based on the application of the Vysochanskii-Petunin inequality for unimodal variables in combination with theory of interference functions or convex optimization, were derived for the general case of arbitrary error covariance matrices. Closed-form solutions are shown to exist for several cases of practical interest. The proposed algorithms outperform the possible approaches from the literature, such as the worst-case based approach, in terms of performance and computational complexity.

- The chance constrained, power control problem in the downlink MISO system, where only the covariance matrices of channel errors were known at the transmitter (worst-case PDF problem), was optimally solved in a closed form.

7.2 Open Problems and Future Research

Interesting open problems for the future work are classified according to the two robust philosophies, employed in this thesis.

7.2.1 Worst-Case Designs

In the downlink MISO case, it would be interesting to establish an analytical connection between the virtual MSE-optimization problem, based on Theorem 4, and the conservative SINR-constrained problem (3.17). From the simulation results, it appears that these problems might be equivalent, as is the case if perfect CSI is provided.

As discussed in Section 3.2, the MSE-constrained problem was optimally solved in a centralized problem setting. To remind the reader, this means that the BS determines all transceiver parameters under uncertainty. However, in practice, CSI at the receiver is usually less critical, which opens several new directions for research. In an ideal situation at the receiver, i.e., equipped with perfect CSI, the users could adopt the exact MMSE equalizers (3.6). What remains is that the BS optimizes only its precoder, but having in mind the uncertainty in (3.6). As commented at the end of Section 3.3, this problem reduces to the non-conservative SINR-constrained problem (3.15). Mathematically, the uncertainty on the both sides of the inequalities

$$\left\| \left[\left(\hat{\boldsymbol{H}}_{(k,:)} + \boldsymbol{\Delta}_{(k,:)} \right) \boldsymbol{G} \quad \sigma_k \right] \right\|_2 \leq \sqrt{1 + \gamma_k^{-1}} \left| \left(\hat{\boldsymbol{H}}_{(k,:)} + \boldsymbol{\Delta}_{(k,:)} \right) \boldsymbol{G}_{(:,k)} \right| \tag{7.1}$$

should be resolved. This is a challenging open problem due to the absolute value on the right hand side.

For the downlink multiuser MIMO setup, besides the same issue of the centralized approach, the convergence of the alternating scheme to a global optimum remains as an interesting open issue. This is tightly connected with the problem of smart initialization, which is also present in the solutions for frequency selective MIMO scenarios. It should be repeated though, that most of the transceiver optimization problems for setups where the users also have multiple antennas are missing solutions in the perfect CSI case, as well. A further research based on strong duality results, along with the algorithms for calculating the actual worst-case channels, derived in Section 4.3, might lead to more efficient numerical solutions than the currently used SDP methods, and possibly to new methods for approaching global maxima of the problems.

It can be noticed that the research in this thesis was in the area of signal processing for wireless communications. However, one can consider also related problems of compound capacity optimization (see, e.g., [WES07]), which often have the form

$$\max_{\boldsymbol{v}} \min_{\boldsymbol{H}\in\mathcal{H}} I(\boldsymbol{v}, \boldsymbol{H}) \tag{7.2}$$

where I is the mutual information in a multiple antenna system described by the channel $\boldsymbol{H}$, $\boldsymbol{v}$ is a vector containing unknown transceiver coefficients, and $\mathcal{H}$ is an uncertainty region that describes imperfect CSI. Establishing connections between the worst-case robustness in signal processing and the currently very attractive, information theoretical concept of compound capacity is certainly an intriguing task.

Finally, it should be remarked that a tailored application of the used algorithms from convex optimization theory would reveal true potentials of the proposed methods in a real-world application.

7.2.2 Probabilistically Constrained Approaches

Probabilistically constrained optimization problems lack unfortunately the deep mathematical background which is partially present in the worst-case optimization. In the past, the contributors in optimization theory were mainly concerned with control theory problems. The system outages there are often considered as catastrophic and not reparable. How to allow a certain, controlled percentage of outages, which is in many situations a natural approach in communication systems, remained somewhat less understood.

The theory is relatively rich for linear models though. However, the QoS measures in communication systems are often quadratic functions in channel coefficients. This brings complex PDF expressions into the problems of interest, and requires innovative approaches. The future work should exploit special properties of the random variables at hand (the unimodality is one such example) to obtain at least tight approximations of the probabilistic constraints. Related to this, one should notice that the Gaussian CSI mismatch was analyzed in Chapter 6. It would be of significant practical interest to cover the case of uniform distribution over some bounded uncertainty region. This scenario clearly corresponds very well to errors caused by quantization.

In this thesis, only the optimization of the power control in a MISO system with probabilistic constraints has been tackled. The results are promising in the sense that one can significantly reduce the conservativeness comparing with the worst-case based methods. What remains to be done is an extension towards full transceiver optimization of MISO and MIMO systems. The optimization of the complete transceiver is supposed to introduce additional gains, especially in low SNR regimes. The application of bounds based on the Chernoff inequality in combination with the majorization theory, used already in the context of MIMO estimation theory [VENG05], is one interesting approach in this direction.

Abbreviations

3GPP	3rd Generation Partnership Project.
BS	Base Station.
CSI	Channel State Information.
FIR	Finite Impulse Response.
ISI	Intersymbol Interference.
LMI	Linear Matrix Inequality.
LTE	Long Term Evolution.
MIMO	Multiple Input Multiple Output.
MISO	Multiple Input Single Output.
MMSE	Minimum Mean Square Error.
MSE	Mean Square Error.
MTF	Matrix Transfer Function.
OFDM	Orthogonal Frequency Division Multiplexing.
PAPC	Per Antenna Power Constraint.
PDF	Probability Density Function.
QAM	Quadrature Amplitude Modulation.
QoS	Quality of Service.
RV	Random Variable.
SDP	Semidefinite Programming.
SER	Symbol Error Rate.
SIC	Successive Interference Cancellation.
SINR	Signal to Interference plus Noise Ratio.
SISO	Single Input Single Output.
SNR	Signal to Noise Ratio.
SOC	Second Order Cone.
TDD	Time Division Duplex.
THP	Tomlinson Harashima Precoding.
UMTS	Universal Mobile Telecommunications System.
WC	Worst Case.
WiMAX	Worldwide Interoperability for Microwave Access.
WLAN	Wireless Local Area Network.

Publication List

- N. Vucic, H. Boche, and S. Shi. Robust Transceiver Optimization in Downlink Multiuser MIMO Systems. *IEEE Transactions on Signal Processing*, accepted for publication, 2009.

- N. Vucic and H. Boche. A Tractable Method for Chance-Constrained Power Control in Downlink Multiuser MISO Systems With Channel Uncertainty. *IEEE Signal Processing Letters*, vol. 16, no. 5, pp. 346-349, May 2009.

- N. Vucic and H. Boche. Robust QoS-Constrained Optimization of Downlink Multiuser MISO Systems. *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 714-725, Feb. 2009.

- N. Vucic and H. Boche. Probabilistically Constrained Robust Power Allocation in Downlink Multiuser MISO Systems. *Proc. 42nd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Oct. 2008.

- N. Vucic and H. Boche. Robust Transceiver Optimization for Frequency Selective MIMO Channels. *Proc. 9th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2008)*, Recife, Brazil, July 2008.

- N. Vucic, H. Boche, and S. Shi. Robust Transceiver Optimization in Downlink Multiuser MIMO Systems With Channel Uncertainty. *Proc. IEEE International Conference on Communications (ICC 2008)*, Beijing, China, May 2008.

- S. Shi, M. Schubert, N. Vucic, and H. Boche. MMSE Optimization With Per-Base-Station Power Constraints for Network MIMO Systems. *Proc. IEEE International Conference on Communications (ICC 2008)*, Beijing, China, May 2008.

- N. Vucic and H. Boche. Downlink Precoding for Multiuser MISO Systems With Imperfect Channel Knowledge. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, Las Vegas, USA, Apr. 2008.

- N. Vucic and H. Boche. Robust Transceiver Optimization for Multiuser MISO Broadcast Systems with MSE Targets. *Proc. 2nd International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, St. Thomas, U.S. Virgin Islands, Dec. 2007.

- N. Vucic and H. Boche, Robust Minimax Equalization of Imperfectly Known Frequency Selective MIMO Channels, *Proc. 41st Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2007.

- N. Vucic, H. Boche, and S. Shi. Robust MSE-Constrained Downlink Precoding in Multiuser MIMO Systems With Channel Uncertainty. *Proc. 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Sept. 2007.

- H. Boche, M. Schubert, N. Vucic, and S. Naik. Non-Symmetric Nash Bargaining Solution for Resource Allocation in Wireless Networks and Connection to Interference Calculus. *Proc. 15th European Signal Processing Conference (EUSIPCO 2007)*, Poznan, Poland, Sept. 2007.

- N. Vucic and H. Boche. Equalization for MIMO ISI Systems using Channel Inversion under Causality, Stability and Robustness Constraints. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006)*, Toulouse, France, May 2006.

- V. Pohl, H. Boche, and N. Vucic. Relation between Robustness and Noise Enhancement of Causal Inverses for MIMO Systems. *Proc. Intern. ITG / IEEE Workshop on Smart Antennas (WSA), Schloss Reisensburg, Germany, Mar. 2006.*

- N. Vucic and H. Boche. Causal Channel Inversion for MIMO ISI Systems: Equalization Error Modeling and Robustness. *Proc. 5th International IMCAS Symposium on Mathematical Modelling (MATHMOD 2006)*, Vienna, Austria, Feb. 2006.

- N. Vucic and M. Schubert. Antenna Selection for BER Performance Improvement in Multi-Antenna Systems With MMSE-SIC Detection. *Proc. IEEE Vehicular Technology Conference (VTC 2005 - Spring)*, Stockholm, Sweden, May 2005.

Bibliography

[AAK01] M.-S. Alouini, A. Abdi, and M. Kaveh. Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels. *IEEE Trans. Veh. Technol.*, 50(6):1471–1480, November 2001.

[BD07a] M. Botros Shenouda and T. N. Davidson. Convex conic formulations of robust downlink precoder design with quality of service constraints. *IEEE J. Sel. Topics Signal Process.*, 1(4):714–724, December 2007.

[BD07b] M. Botros Shenouda and T. N. Davidson. Tomlinson-Harashima precoding for broadcast channels with uncertainty. *IEEE J. Sel. Areas Commun.*, 25:1380–1389, September 2007.

[BD08a] M. Botros Shenouda and T. N. Davidson. A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding. *IEEE J. Sel. Areas Commun.*, 26(2):401–411, February 2008.

[BD08b] M. Botros Shenouda and T. N. Davidson. On the design of linear transceivers for multiuser systems with channel uncertainty. *IEEE J. Sel. Areas Commun.*, 26(6):1015–1024, August 2008.

[BEBT07] A. Beck, Y. C. Eldar, and A. Ben-Tal. Mean-squared error estimation for linear systems with block circular uncertainty. *SIAM J. Matrix Anal. Appl.*, 29(3):712–730, 2007.

[Ber99] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, USA, 1999.

[BG06] M. Biguesh and A. B. Gershman. Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals. *IEEE Trans. Signal Process.*, 54(3):884–893, March 2006.

[BGFB94] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities* *in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, USA, June 1994.

[BL97] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer, New York, NY, USA, 1997.

[BO02] M. Bengtsson and B. Ottersten. Optimum and suboptimum transmit beamforming. In L. C. Godara, editor, *Handbook of Antennas in Wireless Communications*. CRC Press, Boca Raton, FL, USA, 2002.

[BS08a] H. Boche and M. Schubert. Concave and convex interference functions - General characterizations and applications. *IEEE Trans. Signal Process.*, 56(10):4951–4965, October 2008.

[BS08b] H. Boche and M. Schubert. A superlinearly and globally convergent algorithm for power control and resource allocation with general interference functions. *IEEE/ACM Trans. Networking*, 16(2):383–395, April 2008.

[BSG04] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman. Robust downlink power control in wireless cellular systems. *EURASIP J. Wireless Commun.*, 2:261–272, December 2004.

[BTBG+] A. Ben-Tal, D. Bertsimas, L. El Ghaoui, A. Nemirovski, and M. Sim. Robust Optimization. in preparation, available online: http://iew3.technion.ac.il/~nemirovs/index.html.

[BTN01] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization, Philadelphia, PA, USA, 2001.

[BTNR02] A. Ben-Tal, A. Nemirovski, and C. Roos. Robust solutions of uncertain quadratic and conic quadratic problems. *SIAM J. Optim.*, 13:535–560, 2002.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[CIHM08] C.-B. Chae, T. Inoue, R. W. Heath Jr., and D. Mazzarese. Non-iterative multiuser MIMO coordinated beamforming with limited feedforward. In *Proc. ICASSP 2008*, Las Vegas, NV, USA, March 2008.

[CMH06] C.-B. Chae, D. Mazzarese, and R. W. Heath Jr. Coordinated beamforming for multiuser MIMO systems with limited feedforward. In *Proc. Asilomar 2006*, pages 1511–1515, Pacific Grove, CA, USA, October 2006.

[CMJH08] C. B. Chae, D. Mazzarese, N. Jindal, and R. W. Heath Jr. Coordinated beamforming with limited feedback in the MIMO broadcast channel. *IEEE J. Sel. Areas Commun.*, 26(8):1505–1515, October 2008.

[CSCG07] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman. Robust downlink beamforming based on outage probability specifications. *IEEE Trans. Wireless Commun.*, 6(10):3498–3503, October 2007.

[DADSC04] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank. Great expectations: The value of spatial diversity in wireless networks. *Proc. IEEE*, 92(2):219–270, February 2004.

[DJd88] S. Dharmadhikari and K. Joag-dev. *Unimodality, Convexity, and Applications*. Academic Press, San Diego, CA, USA, 1988.

[EHK04] A. T. Erdogan, B. Hassibi, and T. Kailath. MIMO decision feedback equalization from an H^∞ perspective. *IEEE Trans. Signal Process.*, 52(3):734–745, March 2004.

[EM04] Y. C. Eldar and N. Merhav. A competitive minimax approach to robust estimation of random parameters. *IEEE Trans. Signal Process.*, 52(7):1931–1946, July 2004.

[EM05] Y. C. Eldar and N. Merhav. Minimax MSE-ratio estimation with signal covariance uncertainties. *IEEE Trans. Signal Process.*, 53(4):1335–1347, April 2005.

[Fis02] R. F. H. Fischer. *Precoding and Signal Shaping for Digital Transmission*. John Wiley and Sons, Inc., New York, NY, USA, 2002.

[Fos96] G. J. Foschini. Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, pages 41–59, Autumn 1996.

[GL96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, MD, USA, 1996.

[GL05] Y. Guo and B. C. Levy. Worst-case MSE precoder design for imperfectly known MIMO communication channels. *IEEE Trans. Signal Process.*, 53(8):2918–2930, August 2005.

[GL06] Y. Guo and B. C. Levy. Robust MSE equalizer design for MIMO communication systems in the presence of model uncertainties. *IEEE Trans. Signal Process.*, 54(5):1840–1852, May 2006.

[Gol05] A. Goldsmith. *Wireless Communications*. Cambridge University Press, New York, NY, USA, 2005.

[HCS92] M. L. Honig, P. Crespo, and K. Steiglitz. Suppression of near- and far-end crosstalk by linear pre- and post-filtering. *IEEE J. Sel. Areas Commun.*, 10(3):614–629, April 1992.

[HEK06] B. Hassibi, A. T. Erdogan, and T. Kailath. MIMO linear equalization with an H^∞ criterion. *IEEE Trans. Signal Process.*, 54(2):499–511, February 2006.

[HG07] A. Hjørungnes and D. Gesbert. Complex-valued matrix differentiation: Techniques and key results. *IEEE Trans. Signal Process.*, 55(6):2740–2746, June 2007.

[HH03] B. Hassibi and B. M. Hochwald. How much training is needed in multiple-antenna wireless links? *IEEE Trans. Inf. Theory*, 49(4):951–963, April 2003.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 1985.

[HJ91] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, New York, NY, USA, 1991.

[IEE] IEEE. *The Working Group for WLAN Standards, 802.11(n)*. http://grouper.ieee.org/groups/802/11/.

[IEE06] IEEE. *Air Interface for Fixed and Mobile Broadband Wireless Access Systems, Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, IEEE Std 802.16e-2005*, February 2006. http://standards.ieee.org/getieee802/802.16.html.

[JG04] S. A. Jafar and A. Goldsmith. Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback. *IEEE Trans. Wireless Commun.*, 3(4):1165–1175, July 2004.

[Jin06] N. Jindal. MIMO broadcast channels with finite-rate feedback. *IEEE Trans. Inf. Theory*, 52(11):5045–5060, November 2006.

[JK70] N. L. Johnson and S. Kotz. *Continuous Univariate Distributions - 2*. John Willey & Sons, New York, NY, USA, 1970.

[KCJ08] M. Kobayashi, G. Caire, and N. Jindal. How much training and feedback are needed in MIMO broadcast channels? In *Proc. ISIT 2008*, Toronto, Canada, July 2008.

[KFV06] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela. Network coordination for spectrally efficient communications in cellular systems. *IEEE Wireless Commun. Mag.*, 13(4):56–61, August 2006.

[KP85] S. A. Kassam and H. V. Poor. Robust techniques for signal processing: A survey. *Proc. IEEE*, 73(3):433–482, March 1985.

[KTA06] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve. Linear processing for the downlink in multiuser MIMO systems with multiple data streams. In *Proc. ICC 2006*, Istanbul, Turkey, June 2006.

[LB05] R. Lorenz and S. P. Boyd. Robust minimum variance beamforming. *IEEE Trans. Signal Process.*, 53(5):1684–1696, May 2005.

[LDGW04] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong. Transceiver optimization for block-based multiple access through ISI channels. *IEEE Trans. Signal Process.*, 52(4):1037–1052, April 2004.

[LH05] D. J. Love and R. W. Heath Jr. Limited feedback unitary precoding for spatial multiplexing systems. *IEEE Trans. Inf. Theory*, 51(8):2967–2976, August 2005.

[LHL+08] D. J. Love, R. W. Heath Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews. An overview of limited feedback in wireless communications. *IEEE J. Sel. Areas Commun.*, 26(8):1341–1365, October 2008.

[LHSH04] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig. What is the value of limited feedback for MIMO channels? *IEEE Comm. Mag.*, 42:54–59, October 2004.

[Löf04] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference, Taipei, Taiwan*, 2004. http://control.ee.ethz.ch/ joloef/wiki/pmwiki.php.

[LY06] Z.-Q. Luo and W. Yu. An introduction to convex optimization for communications and signal processing. *IEEE J. Sel. Areas Commun.*, 24(8):1426–1438, August 2006.

[Mey00] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, PA, USA, 2000.

[MJHU06] A. Mezghani, M. Joham, R. Hunger, and W. Utschick. Transceiver design for multi-user MIMO systems. In *Proc. ITG/IEEE WSA 2006*, Ulm, Germany, March 2006.

[MKB07] A. Mutapcic, S.-J. Kim, and S. Boyd. A tractable method for robust downlink beamforming in wireless communications. In *Proc. Asilomar 2007*, Pacific Grove, CA, USA, November 2007.

[MO79] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York, NY, USA, 1979.

[Mos85] P. G. Moschopoulos. The distribution of the sum of independent gamma random variables. *Ann. Inst. Statist. Math. (Part A)*, 37:541–544, 1985.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics, Philadelphia, PA, USA, 1994.

[Pap84] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, NY, USA, 1984.

[PGNB04] A. J. Paulraj, D. Gore, R. U. Nabar, and H. Bölcskei. An overview of MIMO communications - A key to gigabit wireless. *Proc. IEEE*, 92(2):198–218, February 2004.

[PIPPNL06] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas. A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization. *IEEE Trans. Signal Process.*, 54(1):346–360, January 2006.

[PJ06] D. P. Palomar and Y. Jiang. MIMO transceiver design via majorization theory. In S. Verdu, editor, *Foundations and Trends in Communications and Information Theory*. Now publishers, Hanover, NH, USA, 2006.

[PPIL07] M. Payaró, A. Pascual-Iserte, and M. Á. Lagunas. Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization. *IEEE J. Sel. Areas Commun.*, 25(7):1390–1401, September 2007.

[Pré95] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.

[Puk94] F. Pukelsheim. The three sigma rule. *The American Statistician*, 48:88–91, 1994.

[RFLT98] F. Rashid-Farrokhi, K. J. Ray Liu, and L. Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE J. Sel. Areas Commun.*, 16(8):1437–1450, October 1998.

[rGPP02] 3rd Generation Partnership Project. *Physical Channels and Mapping of Transport Channels onto Physical Channels*, *3GPP TS 25.211*, 2002. http://www.3gpp.org/ftp/Specs/archive.

[rGPP07] 3rd Generation Partnership Project. *LTE Physical Layer - General Description (Release 8)*, *3GPP TS 36.201*, March 2007. http://www.3gpp.org/ftp/Specs/archive.

[RSG05] Y. Rong, S. Shahbazpanahi, and A. B. Gershman. Robust linear receivers for space-time block coded multiaccess MIMO systems with imperfect channel state information. *IEEE Trans. Signal Process.*, 53(8):3081–3090, August 2005.

[RVG06] Y. Rong, S. A. Vorobyov, and A. B. Gershman. Robust linear receivers for multiaccess space-time block coded MIMO systems: A probabilistically constrained approach. *IEEE J. Sel. Areas Commun.*, 24(8):1560–1570, August 2006.

[SB03] G. J. Székely and N. K. Bakirov. Extremal probabilities for Gaussian quadratic forms. *Probab. Theory Relat. Fields*, 126:184–202, 2003.

[SB04] M. Schubert and H. Boche. Solution of the multi-user downlink beamforming problem with individual SINR constraints. *IEEE Trans. Veh. Technol.*, 53:18–28, January 2004.

[SB05a] M. Schubert and H. Boche. Iterative multiuser uplink and downlink beamforming under SINR constraints. *IEEE Trans. Signal Process.*, 53(7):2324–2334, July 2005.

[SB05b] M. Schubert and H. Boche. QoS based resource allocation and transceiver optimization. In S. Verdu, editor, *Foundations and Trends in Communications and Information Theory*. Now publishers, Hanover, NH, USA, 2005.

[SDL06] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo. Transmit beamforming for physical layer multicasting. *IEEE Trans. Signal Process.*, 54(6):2239–2251, June 2006.

[SNC02] A. H. Sayed, V. H. Nascimento, and F. A. M. Cipparrone. A regularized robust design criterion for uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 23(4):1120–1142, 2002.

[SS05] S. Shi and M. Schubert. MMSE transmit optimization for multi-user multi-antenna systems. In *Proc. ICASSP 2005*, Philadelphia, USA, March 2005.

[SSB+02] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath. Optimal designs for space-time linear precoders and decoders. *IEEE Trans. Signal Process.*, 50(5):1051–1064, May 2002.

[SSB07] S. Shi, M. Schubert, and H. Boche. Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization. *IEEE Trans. Signal Process.*, 55(11):5436–5446, November 2007.

[SSB08a] S. Shi, M. Schubert, and H. Boche. Downlink MMSE transceiver optimization for multiuser MIMO systems: MMSE balancing. *IEEE Trans. Signal Process.*, 56(8):3702–3712, August 2008.

[SSB08b] S. Shi, M. Schubert, and H. Boche. Rate optimization for multiuser MIMO systems with linear processing. *IEEE Trans. Signal Process.*, 56(8):4020–4030, August 2008.

[Stu99] J. F. Sturm. Using SeDuMi 1.02, a MATLAB tool for optimization over symmetric cones. In *Optimization Methods and Software*, volume 11-12, pages 625–653, 1999.

[SW95] R. J. Stern and H. Wolkowicz. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optimiz.*, 5(2):286–313, May 1995.

[Tel95] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *AT & T Bell Labs Tech. Rep.*, June 1995.

[VB96] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, March 1996.

[VB07a] N. Vucic and H. Boche. Robust minimax equalization of imperfectly known frequency selective MIMO channels. In *Proc. Asilomar 2007*, Pacific Grove, CA, USA, November 2007.

[VB07b] N. Vucic and H. Boche. Robust transceiver optimization for multiuser MISO broadcast systems with MSE targets. In *Proc. IEEE CAMSAP*, St. Thomas, U.S. Virgin Islands, December 2007.

[VB08a] N. Vucic and H. Boche. Downlink precoding for multiuser MISO systems with imperfect channel knowledge. In *Proc. ICASSP 2008*, Las Vegas, NV, USA, March 2008.

[VB08b] N. Vucic and H. Boche. Probabilistically constrained robust power allocation in downlink multiuser MISO systems. In *Proc. Asilomar 2008*, Pacific Grove, CA, USA, October 2008.

[VB08c] N. Vucic and H. Boche. Robust transceiver optimization for frequency selective MIMO channels. In *Proc. SPAWC 2008*, Recife, Brazil, July 2008.

[VB09a] N. Vucic and H. Boche. Robust QoS-constrained optimization of downlink multiuser MISO systems. *IEEE Trans. Signal Process.*, 57(2):714–725, February 2009.

[VB09b] N. Vucic and H. Boche. A tractable method for chance-constrained power control in downlink multiuser MISO systems with channel uncertainty. *IEEE Signal Process. Letters*, 16(5):346–349, May 2009.

[VBC07] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized Chebyshev bounds via semidefinite programming. *SIAM Review*, 49:52–64, 2007.

[VBS07] N. Vucic, H. Boche, and S. Shi. Robust MSE-constrained downlink precoding in multiuser MIMO systems with channel uncertainty. In *Proc. Allerton 2007*, Monticello, IL, USA, September 2007.

[VBS08] N. Vucic, H. Boche, and S. Shi. Robust transceiver optimization in downlink multiuser MIMO systems with channel uncertainty. In *Proc. ICC 2008*, Beijing, China, May 2008.

[VBS09] N. Vucic, H. Boche, and S. Shi. Robust transceiver optimization in downlink multiuser MIMO systems. *IEEE Trans. Signal Process.*, 2009. accepted for publication.

[VCG08] S. A. Vorobyov, H. Chen, and A. B. Gershman. On the relationship between robust minimum variance beamformers with probabilistic and worst-case distortionless response constraints. *IEEE Trans. Signal. Process.*, 56(11):5719–5724, November 2008.

[VEG06] S. A. Vorobyov, Y. C. Eldar, and A. B. Gershman. Parameter estimation in linear models based on outage probability minimization. In *Proc. Asilomar 2006*, pages 1415–1419, Pacific Grove, CA, USA, October 2006.

[VENG05] S. A. Vorobyov, Y. C. Eldar, A. Nemirovski, and A. B. Gershman. Probability-constrained approach to estimation of random Gaussian parameters. In *Proc. IEEE CAMSAP 2005*, pages 101–104, Puerto Vallarta, Mexico, December 2005.

[VENG06] S. A. Vorobyov, Y. C. Eldar, A. Nemirovski, and A. B. Gershman. Probability-constrained estimation of random parameters with unknown distribution. In *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Waltham, MA, USA, July 2006.

[VGL03] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo. Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem. *IEEE Trans. Signal Process.*, 51(2):313–324, February 2003.

[Vor08] S. A. Vorobyov. Robust CDMA multiuser detectors: Probability-constrained versus the worst-case based design. *IEEE Signal Process. Letters*, 15:273–276, 2008.

[VP80] D. F. Vysochanskii and Yu. I. Petunin. Justification of the 3σ rule for unimodal distributions. *Theory of Probability and Mathematical Statistics*, 21:25–36, 1980. (in Russian).

[VT03] P. Viswanath and D. Tse. Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. Inf. Theory*, 49(8):1912–1921, August 2003.

[WES06] A. Wiesel, Y. C. Eldar, and S. Shamai. Linear precoding via conic optimization for fixed MIMO receivers. *IEEE Trans. Signal Process.*, 54(1):161–176, January 2006.

[WES07] A. Wiesel, Y. C. Eldar, and S. Shamai. Optimization of the MIMO compound capacity. *IEEE Trans. Wireless Commun.*, 6(3):1094–1101, March 2007.

[WFVH04] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber. Precoding in multiantenna and multiuser communications. *IEEE Trans. Wireless Commun.*, 3(4):1305–1316, July 2004.

[WSM06] T. Weber, A. Sklavos, and M. Meurer. Imperfect channel-state information in MIMO transmission. *IEEE Trans. Commun.*, 54(3):543–552, March 2006.

[Yat95] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE J. Sel. Areas Commun.*, 13(7):1341–1347, September 1995.

[YR94] J. Yang and S. Roy. On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems. *IEEE Trans. Commun.*, 42(12):3221–3231, December 1994.

[ZG02] S. Zhou and G. B. Giannakis. Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback. *IEEE Trans. Signal Process.*, 50(10):2599–2613, October 2002.

[ZG04] S. Zhou and G. B. Giannakis. How accurate channel prediction needs to be for transmit beamforming with adaptive modulation over Rayleigh MIMO channels? *IEEE Trans. Wireless Commun.*, 3:1285–1294, July 2004.

[ZPO08] X. Zhang, D. P. Palomar, and B. Ottersten. Statistically robust design of linear MIMO transceivers. *IEEE Trans. Signal Process.*, 56(8):3678–3689, August 2008.