Multiuser Downlink Beamforming Techniques for Cognitive Radio Networks

Dem Fachbereich 18 der
Technischen Universität Darmstadt
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von

Dipl. Ing. **Dana Silvia Ciochina**
geboren am 10 Dez. 1984 in Bukarest, Rumänien

Referent:	Prof. Dr. Marius Pesavento
Korreferent:	Prof. Dr. Dirk T. M. Slock
Tag der Einreichung:	Sep. 22, 2015
Tag der mündlichen Prüfung:	Dec. 2, 2015

D17
Darmstädter Dissertationen
Darmstadt, 2016

Declaration

I, the undersigned, hereby declare that this dissertation is my original work and has not been submitted or accepted for the award of any other degree in any institution.

Erklärung laut §9 der Promotionsordnung

Ich versichere hiermit dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Signature/Unterschrift: . . . . . .

Place/Ort: . . . . . .

Date/Datum: . . . . . .

Acknowledgments

First and foremost, I would like to express my deep gratitude to my supervisor Prof Dr Marius Pesavento. His guidance, his availability for questions, his persistence and support, as well as his careful reviews and critical comments have played a strong role in the development of this work. I am indebted to late Prof Alex Gershman, who has given me the chance to start my PhD at the TU Darmstadt. Despite the very short time I have had the chance to know him, his enthusiasm for research and his sharp comments have been very motivating and made a strong impact.

I would like to thank Prof Dr Dirk Slock for the careful revision of the thesis and his kind and encouraging comments. I am grateful to Prof Max Wong, for having given me a chance to get acquainted to Riemannian geometry, and for the useful discussions on this topic.

Special thanks go to Imran Wajid and Christian Steffens for reviewing parts of this manuscript and for having been helpful and supportive in many occasions. Furthermore, I am grateful to all my former colleagues of the Communication Systems Group for the many comments and discussions and for having created an active and happy work environment.

Finally, but above all, I am thankful to my parents, for their love, for providing me with everything I have ever needed, for supporting my decisions and constantly encouraging me.

Abstract

Spectrum expansion and a significant network densification are key elements in meeting the ever increasing demands in data rates and traffic loads of future communication systems. In this context, cognitive radio (CR) techniques, which sense and opportunistically use spectrum resources, as well as beamforming methods, which increase spectral efficiency by exploiting spatial dimensions, are particularly promising. Thus, the scope of this thesis is to propose efficient downlink (DL) beamforming and power allocation schemes, in a CR framework. The methods developed here, can be further applied to various practical scenarios such as hierarchical multi-tier, heterogenous or dense networks. In this work, the particular CR underlay paradigm is considered, according to which, secondary users (SUs) opportunistically use the spectrum held by primary users (PUs), without disturbing the operation of the latter. Developing beamforming algorithms, in this scenario, requires that channel state information (CSI) from both SUs and PUs is required at the BS. Since in CR networks PUs have typically limited or no cooperation with the SUs, we particularly focus on designing beamforming schemes based on statistical CSI, which can be obtained with limited or no feedback. To further meet the energy efficiency requirements, the proposed beamforming designs aim to minimize the transmitted power at the BS, which serves SUs at their desired Quality-of-Service (QoS), in form of Signal-to-interference-plus-noise (SINR), while respecting the interference requirements of the primary network.

In the first stage, this problem is considered under the assumption of perfect CSI of both SUs and PUs. The difficulty of this problem consists on one hand, in its non-convexity and, on the other hand, in the fact that the beamformers are coupled in all constraints. State-of-the-art approaches are based on convex approximations, given by semidefinite relaxation (SDR) methods, and suffer from large computational complexity per iteration, as well as the

drawback that optimal beamformers cannot always be retrieved from the obtained solutions. The approach, proposed in this thesis, aims to overcome these limitations by exploiting the structure of the problem. We show that the original downlink problem can be equivalently represented in a so called 'virtual' uplink domain (VUL), where the beamformers and powers are allocated, such that uplink SINR constraints of the SUs are satisfied, while both SUs and PUs transmit to the BS. The resulting VUL problem has a simpler structure than the original formulation, as the beamformers are decoupled in the SINR constraints. This allows us to develop algorithms, which solve the original problem, with significantly less computational complexity than the state-of-the-art methods. The rigurous analysis of the Lagrange duality, performed next, exposes scenarios, in which the equivalence between VUL and DL problems can be theroretically proven and shows the relation between the obtained powers in the VUL domain and the optimal Lagrange multipliers, corresponding to the original problem.

We further use the duality results and the intuition of the VUL reformulation, in the extended problem of joint admission control and beamforming. The aim of this is to find a maximal set of SUs, which can be jointly served, as well as the corresponding beamforming and power allocation. Our approach uses Lagrange duality, to detect infeasible cases and the intuition of the VUL reformulation to decide upon the users, which have the largest contribution to the infeasibiity of the problem. With these elements, we construct a deflation based algorithm for the joint beamforming and admission control problem, which benefits from low complexity, yet close to optimal perfomance. To make the method also suitable for dense networks, with a large number of SUs and PUs, a cluster aided approach is further proposed and consists in grouping users, based on their long term spatial signatures. The information in the clusters serves as an initial indication of the SUs which cannot be simultaneously served and the PUs which pose similar interference constraints to the BS. Thus, the cluster information can be used to significantly reduce the dimension of the problem in scenarios with large number of SUs and PUs, and this fact is further validated by extensive simulations.

In the second part of this thesis, the practical case of imperfect covariance based CSI, available at the transmitter, is considered. To account for the uncertainty in the channel knowledge, a worst case approach is taken, in which the SINR and the interference

constraints are considered for all CSI mismatches in a predefined set One important factor, which influences the performance of the worst case beamforming approach is a proper choice of the the defined uncertainty set, to accurately model the possible uncertainties in the CSI. In this thesis, we show that recently derived Riemannian distances are better suited to measure the mismatches in the statistical CSI than the commonly used Frobenius norms, as they better capture the properties of the covariance matrices, than the latter. Therefore, we formulate a novel worst case robust beamforming problem, in which the uncertainty set is bounded based on these measures and for this, we derive a convex approximation, to which a solution can be efficiently found in polynomial time. Theoretical and numerical results confirm the significantly better performance of our proposed methods, as compared to the state-of-the-art methods, in which Frobenius norms are used to bound the mismatches. The consistently better results of the designs utilizing Riemannian distances also manifest in scenarios with large number of users, where admission control techniques must supplement the beamforming design with imperfect CSI. Both benchmark methods as well as low complexity techniques, developed in this thesis to solve this problem, show that designs based on Riemannian distance outperform their competitors, in both required transmit power as well as number of users, which can be simultaneously served.

Zusammenfassung

Die Erweiterung des genutzten Frequenzspektrums sowie eine erhebliche Verdichtung der Mobilfunknetze in Sinne kleiner Zellen sind Schlüsselemente um die ständig wachsenden Ansprüchen an Datenraten und Datenvolumen in zukünftigen Mobilfunknetzen zu erfüllen. Besonders vielversprechende Techniken, um den zuvor gennannten Ansprüchen gerecht zu werden, bieten, zum einen Cognitive Radios (CR, deut. kognitive Funksysteme), welche verfügbare Frequenzbänder detektieren und opportunistisch belegen, und zum anderen das Beamforming (deut. Strahlenformung), welches die spektrale Effizienz durch die Nutzung von räumlich selektiver Antenne-Abstrahlung erhöhen. Vor diesem Hintergrund werden in dieser Dissertation effiziente Verfahren der Beamformings und der Leistungszuteilung im Rahmen von CR vorgestellt. Die in dieser Arbeit entwickelten Verfahren können auf zahlreiche weitere, praktische Szenarien angewandt werden, wie etwa den gemischten Betrieb verschiedener Funknetzwerke, sowie heterogene oder dichte Funknetzwerke. Der Schwerpunkt der Betrachtung liegt dabei auf dem "Underlay"-Konzept für CR, demgemäß die für Primary Users (PUs, deut. primäre Nutzer) zugewiesenen Frequenzbänder zusätzlich von Secondary Usern (SUs, sekundäre Nutzer) verwendet werden, ohne dass dabei die Übertragungsqualität der Ersteren beeinträchtigt wird. Eine Voraussetzung für den Beamformer-Entwurf in diesen Szenarien ist die Verfügbarkeit der Kanalzustandsinformation der SUs und PUs an der Basisstation (BS). Aufgrund der begrentzten oder nicht vorhandenen Kooperation zwischen PUs und SUs, werden für die hier vorgestellten Verfahren statistische Kanalzustandsinformationen der PUs verwendet, die nur langsam mit der Zeit variieren und mit geringem oder gar ohne Feedback an der BS bestimmt werden können. Um weiterhin eine hohe Energieeffizienz zu gewährleisten, zielen die vorgestellten Verfahren auf eine Minimierung der

Sendeleistung auf Seiten der BS ab, welche die SUs mit Bezug auf eine gewünschte Dienstgüte bedient, während die an den PUs messbare Interferenz unterhalb eines vorgegebenen Schwellwerts liegt.

In einem ersten Ansatz wird das Problem des Beamforming in CR zunächst auf Basis idealer Kanalkenntnis für die PUs und SUs betrachtet. Die Schwierigkeiten dieses Problems bestehen zum einen in der nicht-konvexen Formulierung des zugrundeliegenden Optimierungsproblems und zum anderen in der Kopplung der Beamformer-Koeffizienten in allen Nebenbedingungen. Verfahren auf dem aktuellen Stand der Technik verwenden daher eine konvexe Approximationen in Form von Semidefiniter Relaxierung (SDR). Diese Methoden leiden jedoch an einer erhöhten Rechen-Komplexität pro Iteration sowie der Einschränkung, dass sich die optimalen Beamformer-Koeffizienten nicht immer aus der Lösung der SDR-Formulierung ableiten lassen. Der in dieser Dissertation vorgestellte neue Ansatz zielt darauf ab, diese Einschränkungen durch Ausnutzung der Problemstruktur zu umgehen. Es wird gezeigt, dass das ursprüngliche Problem für den Downlink (DL, deut. Abwärtsstrecke) in ein äquivalentes Problem in einem "virtuellen" Uplink (VUL, deut. virtuelle Aufwärtsstrecke) Bereich dargestellt werden kann, in welchem die Beamformer-Koeffizienten und Sendeleistungen so bestimmt werden, dass die Übertragungsqualität der SUs im Uplink gewährleistet ist, während sowohl PUs und SUs an die Basisstation senden. Das resultierende VUL-Problem besitzt eine vereinfachte Struktur gegenüber der ursprünglichen DL-Problemformulierung, in dem die Beamformer-Koeffizienten in allen Nebenbedingungen entkoppelt sind. Dies ermöglicht die Entwicklung von Algorithmen, welche das ursprüngliche DL-Problem, gegenüber Methoden auf aktuellem Stand der Technik, mit deutlich reduzierter Rechen-Komplexität lösen. Es folgt weiter eine ausführliche Analyse der Lagrangeschen Dualitätseigenschaften, in der die Äquivalenz des VUL- und des DL-Problems für bestimmte Szenarien theoretisch nachgewiesen wird und eine Beziehung zwischen der Leistungszuweisung im VUL-Bereich und den optimalen Lagrangeschen Multiplikatoren, welche der Leistungszuweisung im DL-Bereich entsprechen, aufgewiesen wird.

Ferner werden die Ergebnisse der Dualitäts-Analyse und der Intuition der VUL Formulierung auf das erweiterte Problem der gleichzeitigen Teilnehmerauswahl und Beamformer-Bestimmung angewendet. Ziel dieser Betrachtung ist es eine maximale Anzahl an ausgewälten SUs bei minimaler Sendeleistung an der BS und begrenzten Interferenz-Bedingungen

bezüglich der PUs zu bedienen. In dem vorgestellten Ansatz, lassen sich unzulässige Fälle, durch eine auf Lagrangescher Dualität basierende Methode, effizient identifizieren. Darüber hinaus, werden anhand der Intuition der VUL-Formulierung, die Nutzer identifiziert, welche den größten Effekt auf die Unzulässigkeit verursachen. Mit Hilfe dieser Komponenten wird ein auf Deflation basierenden Algorithmus für die gleichzeitige Teilnehmerauswahl und Beamformer-Bestimmung konstruiert, welcher ein geringe Rechen-Komplexität aufweist, gleichzeitig aber nahezu optimale Leistungsfähigkeit besitzt. Um das entwickelte Verfahren zusätzlich für dichte Netzwerke, mit einer Vielzahl von PUs und SUs, anwendbar zu machen, wird des Weiteren ein Ansatz zur Gruppierung von Nutzern anhand ihrer räumlichen Langzeitmerkmale vorgestellt. Die Informations innerhalb der einzelnen Gruppen dienen dabei zu einer ersten Feststellung welche SUs nicht gleichzeitig bedient werden können und welche PUs gleichwertige Interferenz-Bedingungen besitzen. In Szenarien mit einer Vielzahl an SUs und PUs können die Gruppeninformationen somit für eine deutliche Reduzierung der Problemgröße und der damit verbunden Rechen-Komplexität genutzt werden, was anhand umfangreicher Simulationen belegt wird.

Der praxisrelevante Fall von nicht idealen, statistischen Kanalkenntnissen auf der Senderseite wird im zweiten Teil der Dissertation betrachtet. Um Unsicherheiten in den Kanalinformationen zu berücksichtigen wird ein “worst case” (deut. Extremfall)-Ansatz betrachtet, in welchem die resultierende Dienstgüte und Interferenz für alle möglichen Kanalinformationsfehler innerhalb einer zuvor definierten Menge betrachtet werden. Ein gewichtiger Faktor, der die Leistungsfähigkeit von diesen Verfahren beeinflusst, ist, die geeignete Auswahl des Models für die präzise Beschreibung der vorhandenen Unsicherheiten in den Kanalzustandsinformationen. In dieser Dissertation wird nachgewiesen, dass die kürzlich hergeleiteten Riemannschen Distanzen besser zur Messung von Abweichungen der statistischen Kanalzustandsinformationen geeignet sind als die sonst übliche Frobenius-Norm, da sie, gegenüber den Letzteren, die Eigenschaften von Kovarianz-Matrizen besser darstellen. Daher wird in dieser Dissertation ein neuwertiges ”worst case”-robustes Beamforming Problem definiert, in welchem die Menge der Unsicherheiten mit Hilfe der Riemannschen Distanzen begrenzt wird, und darüber hinaus eine konvexe Approximation hergeleitet, welche eine Lösung in Polynomialzeit erlaubt. Theoretische und numerische Ergebnisse bestätigen die deutlich verbesserte Leistungsfähigkeit des vorgestellten Verfahrens gegenüber Verfahren

auf aktuellem Stand der Technik, in welchen die Frobenius-Norm zur Begrenzung der Unsicherheiten genutzt wird. Die durchgehend besseren Ergebnisse, des auf Riemannschen Distanzen basierenden Verfahrens, zeigen sich auch in Szenarien mit einer Vielzahl von Nutzern, in welchen der Beamformer-Entwurf mit nicht idealen Kanalzustandsinformationen durch eine Teilnehmerauswahl erweitert werden muss. Sowohl die rechenintensiven, optimalen Verfahren als auch die heuristischen Verfahren mit geringer Rechen-Komplexität, welche jeweils in dieser Dissertation vorgestellt werden, zeigen, dass die auf Riemannschen Distanzen basierenden Verfahren die Vergleichsverfahren sowohl bezüglich der benötigten Sendeleistung als auch bezüglich der Zahl der Nutzer, welche gleichzeitig bedient werden können, übertreffen.

Table of Contents

Mathematical Notation

Sets

$\mathbb{R}$	the set of real numbers
$\mathbb{R}_+^N$	the set of real non-negative vectors of length N
$\mathbb{C}$	the set of complex numbers
$\mathbb{C}^{N\times K}$	the set of complex matrices of N rows and K columns
$\lvert A\rvert$	cardinality of the set A
$A \cup B$	reunion of sets A and B
$A \cap B$	intersection of sets A and B
$A \subset B$	sets A is a subset of B
$x \in A$	element x belongs to the set A

Vectors and Matrices

range($\mathbf{A}$)	range of matrix $\mathbf{A}$
rank($\mathbf{A}$)	rank of matrix $\mathbf{A}$
Tr $\{\mathbf{A}\}$	trace of matrix $\mathbf{A}$
vec($\mathbf{A}$)	vectorization of matrix $\mathbf{A}$
$\lambda_{\min}(\cdot)$	smallest eigenvalue of a matrix
$\lambda_{\max}(\cdot)$	largest eigenvalue of a matrix
$\mathbf{0}_N$	all zeros column vector
$\mathbf{1}_N$	all ones column vector

$\mathbf{I}_N$	Identity matrix of dimension N
$\mathbf{A}^\dagger$	pseudo-inverse of matrix $\mathbf{A}$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker matrix product
$(\cdot)^T$	transpose
$(\cdot)^*$	complex conjugate
$(\cdot)^H$	transpose complex conjugate (Hermitian) operation
$\mathbf{a} \geq \mathbf{b}$	element-wise inequality between vectors $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{a} > \mathbf{b}$	element-wise strict inequality between vectors $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{A} \succeq 0$	matrix $\mathbf{A}$ is positive semidefinite
$\mathbf{A} \succ 0$	matrix $\mathbf{A}$ is positive definite

Norms and Distances

$\|\cdot\|_F$	Frobenius norm of a matrix
$\|\cdot\|$	second norm of a vector or a matrix
$d_F(\mathbf{A}, \mathbf{B})$	Euclidean distance between the matrices $\mathbf{A}$ and $\mathbf{B}$
$d_R(\mathbf{A}, \mathbf{B})$	Riemannian distance between the matrices $\mathbf{A}$ and $\mathbf{B}$

Statistical Operators

$\mathrm{E}\{\cdot\}$	statistical expectation operation
$\mathcal{W}_N(\mathbf{A}, N_s)$	Wishart distribution with scale matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ and N_s degrees of freedom
$\mathcal{N}(0, \mathbf{A})$	Gaussian distribution of zero mean and covariance matrix $\mathbf{A}$

Miscellaneous

$\mathrm{Re}(\cdot)$	real part of $(\cdot)$
$\mathrm{Im}(\cdot)$	imaginary part of $(\cdot)$

Acronyms

2G	second generation
4G	fourth generation
BS	base-station
cdf	cumulative distribution function
CoMP	cooperative multipoint
CR	cognitive radio
CSI	channel state information
DL	downlink
IA	interference alignment
iCSI	instantaneous CSI
IC	interference control
GSM	global system of mobile communications
KKT	Karush-Kuhn-Tucker
LP	linear program
LS	least-square
LTE	long term evolution
LTE-A	long term evolution advanced
MILP	mixed integer-linear program
MISDP	mixed integer semidefinite program

MISO	multiple-input-multiple-output
MISOCP	mixed integer second order cone program
MIMO	multiple-input multiple-output
ML	maximum likelihood
MMSE	minimum mean-square error
mmWave	milimiter Wave
MSE	mean-square error
QCQP	quadratically constrained quadratic program
QoS	Quality of Service
psd	positive semidefinite
PN	primary network
PU	primary user
RF	radio frequency
RMSE	root mean square error
sCSI	statistical CSI
SDP	semidefinite program
SDR	semidefinite relaxation
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
SOCP	second-order cone programming
SN	secondary network
SU	secondary user
SVD	singular value decomposition
UMTS	universal mobile telecommunications system
UL	uplink
VUL	virtual uplink

WI	worst case interference
WSINR	worst case signal-to-interference-plus-noise ratio

Chapter 1

Introduction

1.1 Background

The highly productive interplay between the semiconductor and communication research, together with the innovative development of mobile devices and their applications have led to an explosive demand in wireless services. Only last year, mobile data traffic has increased by 69% and the compound anual growth is expected to be around 57% from 2014 to 2019 [1]. To respond to the exponential increase in connected devices and data traffic, ambitious goals are envisioned for the future communication technologies, emerging under the keyword "5G" [3]. However, in their evolution of the last two decades, the wireless services have already undergone a paradigm shift from low rate low latency voice demands to high data rate streaming applications with an increase in supported peak data rates of several orders of magnitude. To sustain this, current technologies such as LTE and LTE-A are already designed to operate close to the theoretical limits in their particular bandwidths, due to e.g., advances in coding, modulation and link adaptation [4]. To understand potential directions in which this development is still possible, we first address some of the main challenges, which must be overcome.

An essential factor to increase data rates is the ability of the network to efficiently utilize larger bandwidths. Looking at the evolution of communication systems, the efforts, carried out in the past years to operate over constantly increasing bandwiths can be clearly

noted. Whereas 2G standards such as GSM were designed for chunks of 200 kHz of bandwidth, current 4G LTE-A can operate over 20MHz over contiguous bandwidths and even up to 100MHz, with carrier aggregation [5]. However, this dramatic increase in bandwidth requirements together with the fixed frequency allocation, currently imposed by regulators, has led to a situation where most spectrum below 3GHz, which has been regarded as most valuable to mobile communications, due to its favourable propagation properties, is already occupied. Therefore, future wireless systems must either find spectral opportunities in already licensed bandwidths or explore the frequency bands above 3GHz.

Additionally, it is acknowledged that to achieve a significant capacity increase, a large network densification is required. This can be accomplished by supplementing the traditional macrocell with small cell networks, which improve coverage and offload traffic. Heterogenous networks, consisting of femtocells and picocells, have already been well accepted by industry [10], [11]. To benefit, however, from a dense deployment of small networks, a rigurous interference management is needed. Techniques to achieve this, such as Interference alignment(IA) [12]-[15] and coordinate multipoint (CoMP) [17]-[21], are currently receiving significant attention in the research community. The goal of the former technique is to enable the simultaneous communication of multiple pairs of transmitters and receivers over a common channel. This is achieved by allowing transitters to jointly design their signals so that the interference at the receivers only occupies a portion of the signal space. This technique was shown to exhibit remarkable properties in theory [13], however, in practice, significant challenges must be overcome, in order to fully benefit from it. Among these are the low performance in low SNR regime, the stringent synchronization requirements of all transmitters as well as the need for full knowledge of all instantaneous channels [16]. Similarly, in CoMP schemes, multiple base-stations (BS) form clusters and cooperatively serve users [17]-[21], in order to achieve the desired performance metrics. Also, in the case of CoMP, significant research is directed towards making these techniques applicable in practice, e.g., by optimizing cooperation clusters [18], [21], limiting the feedback overhead [20], and the amount of backhaul signaling [19]. Thus, even though significant progress in improving these schemes is expected, it is unlikely that these techniques alone, will enable the large network densification, envisioned by future standards.

Last but not least, the increasing concerns on the carbon footprint are pushing the need

to actively consider energy efficiency aspects in the design of the future communication systems. Even though currently contributing to only around 2% of the global carbon emission, the exponentially increasing energy consumption, makes this problem impossible to be neglected [22]-[24].

A promising technology, which holds a potential solution to these challenges is cognitive radio (CR). The concept of CR has been introduced by Mitola in [6], where a 'flexible' radio, able to sense and adapt to the environment, was envisioned. This idea received immediate interest in the research community, particularly motivated by measurement studies which showed that, even though licensed, large amounts of spectrum are underutilized for large amounts of time and in wide geographical areas [2]. Consequently, significant efforts have been devoted to designing sensing algorithms, which are able to reliably discover available spectrum opportunities [25].

Initially considered as a solution to reutilize TV bands [8], the CR framework has evolved to encompass a variety of techniques, which enable an unlicensed secondary user (SU) or base-station (BS) to 'sense' transmit opportunities in the spectrum allocated to a licensed primary network (PN) and 'underlay', 'overlay' or 'interweave' its transmission with those of the incumbent primary users (PUs) [9]. Specifically, these transmission techniques define how the CR users utilize the spectrum of the PN, and differ in the amount of cooperation between the CR and the primary networks, as well as in the strictness of the interference protection the former must respect towards the latter. Thus, 'interweave' and 'underlay' schemes assume limited or no cooperation between the two networks, whereas in 'overlay', the CR transmitter assists the PN, by relaying messages to PUs and thus improving the primary transmission. Further, according to the interweave paradigm, a CR transmitter can only utilize frequency slots, which are not utilized by the PUs at the particular time instant, while the underlay scheme is less restrictive, allowing CRs to utilize the same frequency spectrum, as long as the interference leaked to the PUs is rigurosly controlled to lie below imposed interference thresholds.

Naturally, these CR transmission paradigms also apply to hierarchical heterogenous networks, in which, e.g., a femtocell serves users without disturbing the operation of a macrocell, by limiting the interference leaked to the users of the macro network, or by operating over a frequency slot not used by the macrocell at the particular time instant [11].

Furthermore, when multiple BSs are clustered to serve users, the underlay paradigm can also be useful, since, in order to be practical, only a small number of BSs can cooperate, while the users served by the remaining BSs must be protected from interference. Multi-antenna beamforming is considered as a significant enabler for CR underlay and overlay transmission schemes in multiple-input-multiple-output (MIMO) networks. Beamforming techniques consist in adaptively modifying the directivity patterns of the antennas, such that the transmission towards intended users is improved, whereas, interference is reduced and controlled [26]. Thus, the coexistance of multiple tiers, e.g., femtocells and macrocells can be achieved.

A particularly promising technique, which complements CR in the efforts to provide more available bandwidth and creates a relevant context for beamforming in underlay and overlay scenarios is milimiter-wave (mmWave) transmission. This technique advocates the use of the frequency bands above 3 GHz, more precisely between 3-300GHz [27]-[30], where large amounts of available spectrum can be found. The reason, why these bandwidths have been to date scarcely utilized is due to the concerns regarding the poor propagation conditions at very high frequency: increased pathloss, larger penetration losses and adverse meteorological conditions in this domain. Recent studies show that these concerns regarding large rain attenuation and penetration losses are indeed justified [27]. On the other hand, the statement that in free space, higher frequencies propagate less than lower ones, as given by the well known Friis equation, is valid when isotropic antennas are used at transmitter and receiver. If, however, directive antennas are employed, higher frequencies can even propagate farther than lower frequencies, for fixed aperture sizes [30]. The reason is that, higher frequencies translate into smaller wavelengths, and thus enable the use of more antenna elements, with which narrower beams, better gains and directivity characteristics can be obtained [27]. The research area of mmWave transmission for cellular communications is in its infancy, however feasibility studies [27], as well as measurement campaigns with prototypes designed to operate in 28 and 38 GHz have already shown promising results and spurred a great interest in the topic [28], [29]. Thus, beamforming is expected to be a key enabler for mmWave transmission, by means of which access to large amounts of available spectrum can be achieved. Furthermore the use of large numbers of antenna elements,

enabled by the higher frequencies, together with CR based interference management techniques are expected to provide essential solutions to the envisioned network densification.

1.2 Thesis Overview and Contributions

Our focus, in this thesis, lies on the design of efficient beamforming and power allocation schemes for multiuser CR underlay scenarios. Various aspects of this problem, such as the mathematical structure, the existence of feasible solutions, the control of the user admission and the robustness against imperfections in the channel knowledge, are considered and carefully analysed.

This dissertation is organized as follows. In Chapter 2, after a short introduction into the vast field of transmit beamforming techniques and the various design considerations that have been taken in literature, we formulate the core problem, on which we focus, in this thesis. Specifically, this consists in optimizing the beamformers allocating the power per user and minimizing the overall transmit power, in order to serve users with desired Quality-of-service (QoS) in terms of achieved Signal-to-Interference-plus-Noise-Ratio (SINR), while respecting the interference constraints, imposed by a PN. An overview of the techniques, available in literature, in both conventional scenarios, i.e., in the absence of PUs, as well as in CR networks, with PUs present, completes this chapter.

In Chapter 3, we consider the case, in which the problem with a given user assignment admits a feasible solution and perfect channel state information (CSI) is available at the BS. Under this assumption, we propose a virtual uplink (VUL) reformulation of the original problem. The advantage of this reformulation is that the beamformers are decoupled in the constraints, which allows us to devise iterative algorithms, with significantly reduced computational complexity as compared to the state-of-the-art interior point methods. Next, we analyze the duality properties of the original problem. For this purpose, we consider an alternative uplink reformulation, derived based on Lagrange duality, and show that, under certain assumptions, the original problem attains strong duality. Specifically, we show that this result applies if *i)* instantaneous CSI of SUs and covariance based CSI of PUs is considered, *ii)* covariance based CSI in the form of Toeplitz matrices is assumed for both SUs and PUs, or *iii)* a strictly feasible power allocation exists for any feasible

beamformer. A consequence of strong duality is that, when it holds, optimality of the solutions obtained by our proposed algorithms, as well as by the state-of-the-art methods is always guaranteed. Furthermore, we show through counterexamples, that extensions of the duality results obtained for conventional networks, i.e., in the absence of the PUs, do not necessarily hold for CR networks. Based on the uplink reformulations, two algorithms are proposed and their computational complexity and convergence properties are analysed. The chapter closes with a presentation of extensive numerical results, under both instantaneous and covariance based CSI for SUs and PUs.

In Chapter 4, we relax the feasibility assumption, and propose a method which is able to detect whether a beamforming and power allocation exists, such that the SINR and interference constraints can be simultaneously met. The method stems from the theorem of alternatives in duality theory and has as an advantage, that it can be naturally incorporated in the previously proposed technique, thus benefiting from its low computational complexity. For the cases, in which infeasibility is signaled, we design a deflation technique, which successively removes one SU per iteration until feasible set of users, that can be served simultaneously at there required QoS constraints is found. To decide upon the user to be removed, we first propose a heuristic measure that is suggested by the intuition, conferred by the VUL interpretation of the problem. This scheme performs reasonably well, especially in cases where there exists a user with a significantly large heuristic measure. A more elaborate approach is proposed next, in which a depth first search is taken to discern between SUs, with similar heuristic measures. The deflation approach considered so far, benefits from low complexity and good performance. However, it is impractical in dense networks, where a large number of SUs require access to resources or a large number of PUs imposes difficult interference limitations on the network. To make this method also applicable in these cases, we propose next a preprocessing clustering phase, based on the long term spatial signatures of the users. These statistical measures are used in order to allow the clustering scheme to be performed offline and require less frequent updates, as required in the case of instantaneous CSI. To measure the similarity between long term spatial signatures, we use a Riemannian distance, recently derived in [134]. Based on this, two clustering algorithms are proposed. Next, a method to retrieve meaningful cluster representatives is devised and, for these, the deflation based algorithm, previously proposed, is applied. Numerical results

show that this method provides good performance in point of served number of users and, most importantly, scales well with an increase in the number of users. The reduction in complexity, achieved by applying the preprocessing scheme, as compared to the successive deflation without clustering, can reach up to 80%, both in point of runtimes and number of total iterations, required to find an appropriate set of users and its corresponding optimal beamformers and powers.

In the remaining part of the thesis, we relax the assumptions on the perfect CSI, and turn our attention to designs, which account for mismaches in these values. Chapter 5 is a brief introduction into the topic of robust beamforming. Here, the worst case problem formulation with mismatches in the covariance based CSI is introduced and the related approaches taken in literature to solve this problem are shown, together with their drawbacks. In Chapter 6, we show our proposed worst case robust beamforming design. In our approach, rather than using the common Euclidean norms, which measure the distance between matrices as a cumulative difference between their individual elements, we consider a Riemannian distance measure, which is derived to intrinsically capture the positive semidefiniteness properties of the covariance matrices. We analyse the properties of the Riemannian distance, which impact on the robust beamforming design. We then proceed to derive a convex approximation to the newly formulated robust beamforming problem, where mismatches are bounded with the Riemannian distances. The final convex formulation can be implemented using interior point methods. Simulation reults show a significantly improved performance of our approach, with respect to the state-of-the-art methods, especially in difficult scenarios, in which the QoS requirements are large or the separation between users is small. Finally, we consider the general problem, in which both the assumptions on feasibility and perfect CSI are relaxed, and propose a solution to the problem of admission control and optimal beamforming. Our aim in this section is twofold. On one hand, we are interested to provide a solution to the admission control and beamforming problem with imperfect CSI. On the other hand, we complete the analysis of robust beamforming design using Riemannian distance by developing a benchmark method, based on which the impact of the mismatch set characterization can be shown, in a large number of users.

The results from this thesis have been presented in the following conference and journal papers.

- D. Ciochina and M. Pesavento, "A Clustering Approach for Admission Control and Optimal Beamforming in Cognitive Radio Networks", *IEEE GLOBECOM Workshop on Green Broadband Access (GLOBECOM-GBA'14),* Dec. 2014, Austin, Texas

- D. Ciochina, M. Pesavento, and K. M. Wong, "Worst Case Robust Beamforming on the Riemannian Manifold", *IEEE 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013),* May 2013, Darmstadt, Germany,

- L. Xu, K. M. Wong, J. K. Zhang, D. Ciochina, and M. Pesavento, "A Riemannian Distance for Robust Downlink Beamforming", *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC'13),* June 2013, Darmstadt, Germany

- I. Wajid, M. Pesavento, Y. C. Eldar, and D. Ciochina, "Robust downlink beamforming with partail channel state information for conventional and cognitive radio networks," *Proc IEEE Trans. on Signal Processing*, vol. 60, no. 2, Aug. 2013

- D. Ciochina and M. Pesavento, "Joint User Selection and Beamforming in Interference Limited Cognitive Radio Networks," *European Signal Processing Conference (EUSIPCO'12)*, Aug. 2012, Bucharest, Romania,

- M. Pesavento, D. Ciochina, and A.B. Gershman, "Iterative Dual Downlink Beamforming for Cognitive Radio Networks," *5th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2010)*, June 9-11, Cannes, France.

Chapter 2

Multiuser Downlink Beamforming - Problem Statement and Prior Work

2.1 Introduction

Transmit beamforming is a powerful technique, which makes it possible for multiple users to be simultaneously served, by efficiently exploiting the spatial dimensions. Therefore, this has been addressed in literature from a variety of view points, and techniques have been developed to respond to a wide range of design requirements.

To address the ever increasing demand in data rates, many approaches have focused on proposing beamforming designs which maximize the sum rate, [32]-[41], under further considerations such as total transmit power constraints, per antenna power constraints or, in the case of CR networks, maximum interference thresholds. Motivated by energy efficiency concerns, a formulation of most interest for telecom operators is the transmit power minimization, required to satisfy the QoS constraints of the users [58]-[66], [70]-[74]. A closely related design criterion is max-min fairness [64], [42]-[44], in which the beamformers are computed to maximize the smallest SINR of the users in the network. In [56] and [67], it was shown that these two problems, i.e., the transmit power minimization with SINR constraints and the maximin of SINR with power constraints are dual to each other, in the sense that, when the transmit power constraint in the latter is fixed to the optimum of the former problem the same optimum beamformers are obtained. The same result holds,

if the SINR targets in the power minimization problem are chosen as the optima of the max-min formulation. Other design criteria, of particular interest for CR networks are the minimization of the interference at the PUs, or proposing tradeoffs between intereference minimization and QoS satisfaction or finding thresholds, which ensure the operation of both [45].

A large amount of the beamforming designs have been developed, based on instantaneous CSI metrics[32]-[41], [61]. The channel information can be obtained at the receiver, but needs to be constantly fed back to the BS. This incurs significant overhead and is prone to errors, especially when channels are rapidly fluctuating. Thus to the first problem, limited feedback schemes have been developed [91], whereas, for the latter, possible solutions have been explored, in which statistical knowledge of the channel is used in the designs [46], [47], [58], [70], [71].

A common difficulty of multiuser downlink beamforming problems, for all the design considerations, mentioned above is the interference coupling. To handle this, two powerful tools have been developed, namely uplink downlink duality and the theory of interference functions. The former framework aims to find equivalence transformations, converting the downlink problems into uplink problems, which are not affected by interference coupling. Uplink downlink duality results were obtained either by complicated changes of variables [39], [40], by adopting Perron-Frobenius theory for non-negative matrices [58], [64] or by using Lagrange duality theory [41]. Alternatively, using the theory of interference functions, introduced in [59] and further developed in [48], [63] interesting results were obtained, regarding the convexity of the feasible regions under various QoS measures.

Here, we propose multiuser downlink beamforming techniques which aim to minimize the transmit power, while satisfying QoS and interference requirements. Our aim is to extend the uplink downlink duality results to CR scenarios, with arbitrary number of constraints, and under both instantaneous as well as covariance based CSI. We formulate in Section 3.2 the problem, which we are interested to solve. Further, in Section 3.3, we have a closer look at the state-of-the-art approaches, proposed in literature for the problem formulation, we deal with in this thesis, and discuss their limitations.

2.2 Signal Model and Problem Formulation

We consider a CR scenario, in which one BS with N antennas opportunistically uses the spectrum of a PN, in an underlay fashion, to serve SUs with a guaranteed QoS. We assume that K single antenna SUs, indexed as 1 to K require access to resources, and from them a set $S \in \{1, \ldots, K\}$ can be served at a particular time-slot. The signal transmitted by the BS at time instant n represents a linear combination of the precoded data symbols destined for all users and can be expressed as

$$\mathbf{x}(n) = \sum_{k \in S} \sqrt{p}_k \mathbf{u}_k s_k(n), \tag{2.1}$$

where p_k, $\mathbf{u}_k$ and $s_k(n)$ denote the power, unit norm beamformer and data symbol at time instant n, respectively, destined to the k-th user. Further, let $h_{k,i}$ be the channel from the ith antenna element to the kth SU, with which the channel vector from the BS to this respective user denoted as $\mathbf{h}_k = [h_{k,1}, \ldots, h_{k,N}]^T$ is formed. With this notation and the definition in Eq. (2.1), the received signal at the kth SU can be written as

$$y_k(n) = \mathbf{h}_k^H \mathbf{x}(n) + z_k, \tag{2.2}$$

where z_k represents the temporally and spatially white complex circularly Gaussian noise at the kth receiver, having zero mean and σ_k^2 variance. To ensure a required level of QoS at the BS, we consider as metric the SINR, which for the kth SU can be expressed as

$$\mathrm{SINR}_k^{\mathrm{DL}} \triangleq \frac{p_k \mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\sum_{\substack{i \neq k \\ i \in S}} p_i \mathbf{u}_i^H \mathbf{R}_k \mathbf{u}_i + \sigma_k^2}. \tag{2.3}$$

In Eq. (2.3), $\mathbf{R}_k$ denotes the channel covariance matrix of the k-th SU, which is computed as $\mathbf{R}_k = \mathrm{E}\left\{\mathbf{h}_k \mathbf{h}_k^H\right\}$.

In the same spectrum, L single antenna PUs coexist and we consider without loss of generality that they are indexed as K+1 to K+L. The set of the PU indices is then defined as $S_{PU} \triangleq \{K+1, \ldots, K+L\}$. Using a similar notation as in the SU case, we let $h_{K+l,i}$ be the channel coefficient between the ith BS antenna element and the lth PU. Correspondingly, we form the channel vector from the BS to the lth PU, i.e., $\mathbf{h}_{K+l} = [h_{K+l,1}, \ldots, h_{K+l,N}]$. With this notation and the definition in Eq. (2.1), the signal received at lth PU can be expressed as

$$y_{K+l}(n) = \mathbf{h}_{K+l}^H \mathbf{x}(n). \tag{2.4}$$

Consequently, the interference that the BS leaks to the lth PU can be written as

$$\mathrm{I}_l^{\mathrm{DL}} \triangleq \sum_{k\in S} p_k \mathbf{u}_k^H \mathbf{R}_{K+l} \mathbf{u}_k, \tag{2.5}$$

where, similarly to the SU case, we have denoted by $\mathbf{R}_{K+l}$ the channel covariance matrix of the lth PU, i.e., $\mathbf{R}_{K+l} = \mathrm{E}\left\{\mathbf{h}_{K+l}\mathbf{h}_{K+l}^H\right\}$. According to the underlay paradigm, to use the spectrum of the PN, the BS must ensure that the interference level at each PU is below imposed thresholds. This condition translates to

$$I_l \leq \gamma_{K+l}^{-1};\ l = 1, \ldots L, \tag{2.6}$$

where γ_{K+l} represents the interference threshold of the lth PU. Due to the form of the interference constraints in (2.6), we note that the absence of the noise components in Eq. (2.4) does not incur any loss of generality, as the contribution of these terms can be considered to be incorporated in the imposed interference thresholds.

The goal of the joint beamforming and power allocation problem is to compute optimal beamformers and powers, such that the QoS requirements for all the SUs, as well as the interference thresholds, imposed by all the PUs, are respected, while minimizing the total transmitted power. This can be compactly formulated as

$$\mathcal{PD}(S, S_{PU}): \min_{\{\mathbf{u}_k, p_k\}} \sum_{k\in S} p_k \tag{2.7a}$$

$$\text{s.to } \frac{p_k \mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\sum_{\substack{i\neq k \\ i\in S}} p_i \mathbf{u}_i^H \mathbf{R}_k \mathbf{u}_i + \sigma_k^2} \geq \gamma_k;\ k \in S \tag{2.7b}$$

$$\sum_{k\in S} p_k \mathbf{u}_k^H \mathbf{R}_l \mathbf{u}_k \leq \frac{1}{\gamma_l};\ l \in S_{PU} \tag{2.7c}$$

$$p_k \geq 0;\ ||\mathbf{u}_k|| = 1. \tag{2.7d}$$

In practice, an additional regulatory sum power constraint may be required. This can be straightforwardly considered in the formulation of the original problem (2.7) as an $L+1$-th PU, with a channel covariance matrix equal to the identity matrix. We will use this constraint explictly in the admission control problem.

It is easy to see that the problem in (2.7) can be equivalently considered in variables $\mathbf{w}_k = \sqrt{p_k}\mathbf{u}_k$, for $k = 1, \ldots, K$. With these, the Lagrange dual problem corresponding to

(2.7) can be formulated as

$$\max_{\{q_k\}} \min_{\{\mathbf{w}_k\}} \mathcal{L}(q_1, \ldots, q_{K+l}, \mathbf{w}_1, \ldots, \mathbf{w}_K). \tag{2.8}$$

In (2.8), $\mathcal{L}$ is the Lagrangian function defined as

$$\begin{aligned}\mathcal{L}(q_1, \ldots, q_{K+L}, \mathbf{w}_1, \ldots, \mathbf{w}_K) \triangleq & \sum_{k\in S} q_k \left(-\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k + \sum_{\substack{i\in S \\ i\neq k}} \mathbf{w}_i^H \gamma_k \mathbf{R}_k \mathbf{w}_i + \gamma_k \sigma_k^2 \right) \\ & + \sum_{l\in S_{PU}} q_l \left(\sum_{k=1}^{K} \mathbf{w}_k^H \gamma_l \mathbf{R}_l \mathbf{w}_k - 1 \right) \qquad (2.9a) \\ = & \sum_{k\in S} \mathbf{w}_k^H \left(\mathbf{I} - q_k \mathbf{R}_k + \sum_{\substack{i\in S \\ i\neq k}} q_i \gamma_i \mathbf{R}_i + \sum_{l\in S_{PU}} q_l \gamma_l \mathbf{R}_l \right) \mathbf{w}_k \\ & + \sum_{i=1}^{K} q_i \gamma_i \sigma_i^2 - \sum_{l\in S_{PU}} q_l. \qquad (2.9b)\end{aligned}$$

The equality in Eq. (2.9b) is simply a rearrangement of the terms in Eq. (2.9a), meant to emphasize that the Lagrangian function is a quadratic form, and that the matrices between parentheses must be positive semidefinite, in order for the inner minimization in (2.8) to be bounded below. Using the fact that the optimal weighting factors $\{\mathbf{w}_k\}_{k\in S}$ must respect the first order optimality condition, i.e., $\partial\mathcal{L}(q_1, \ldots, q_{K+L}, \mathbf{w}_1, \ldots, \mathbf{w}_K)/\partial\mathbf{w}_k = 0$, for all k, it follows that the problem in (2.8) can be equivalently expressed as

$$\min_{\{q_k\}} \sum_{k\in S} \gamma_k \sigma_k^2 q_k - \sum_{l\in S_{PU}} q_l \tag{2.10a}$$

$$\text{s.to } \mathbf{Z}_k \triangleq \mathbf{I} - q_k \mathbf{R}_k + \sum_{\substack{i\neq k \\ i\in S}} \gamma_i q_i \mathbf{R}_i + \sum_{l\in S_{PU}} \gamma_l q_l \mathbf{R}_l \succcurlyeq 0; \ k \in S \tag{2.10b}$$

$$q_k \geq 0; \ q_l \geq 0; \ l \in S_{PU}. \tag{2.10c}$$

Next, let us make several remarks regarding the CSI assumptions, considered in this thesis. In order to design the beamformers, CSI in the form of instantaneous or second order channel information of all users must be available at the transmitter. In this thesis, we consider the following cases:

i) Instantaneous CSI (iCSI) for the SUs, i.e., rank one covariance matrices $\mathbf{R}_k = \mathbf{h}_k\mathbf{h}_k^H$, for $k \in S$ and statistical CSI (sCSI) in the form of general rank channel covariance matrices for the PUs.

ii) Covariance based sCSI for both SUs and PUs.

Case i) is motivated by the fact that iCSI is a common assumption in literature and is practically motivated for the secondary system, as the channels may be acquired through training by the SUs and fed back to the BS. Estimation techniques and training procedures are available in literature and documented in standards. Considering this type of CSI is also mathematically attractive, because in this case the original problem (2.7) holds some interesting properties, which we reveal further in Chapter 3. On the other hand, assuming iCSI in the case of the PUs is generally unrealistic, as these users have little incentive to spend resources on feeding back their channel information to the secondary BS. Second order statistics of the channels are known to vary more slowly with time, which makes them a more appropriate choice in these cases.

On the other hand, sCSI is also of interest in the case of SUs. This is because sending frequent updates of the iCSI becomes impractical when, e.g., a large number of antennas are considered at the BS, a large number of users are present in the network, or when mobility is considered. This is the motivation behind the case ii).

2.3 Related Works and Contribution

The problem in (2.7) is a non-convex quadratically constrained quadratic program (QCQP). To be able to find optimal solutions in polynomial time, it is essential to prove that strong duality [49] holds for the problem. For this generic class of problems, however, strong duality results have so far only been proven, if the number of constraints is limited to two or three in [49] and [50], respectively. The proofs presented in these works are constructive and therefore extensions to larger number of constraints are not straightforward.

However, beamforming and power allocation problems, exhibit a favourable structure, based on which strong duality results have been obtained [58], [61] and efficient polynomial time algorithms have been designed [61], [62], [65], [66] for conventional DL scenarios i.e., $L = 0$. Instrumental in developing these algorithms is the fact that the original problem

can be equivalently solved by a reformulation, which exhibits an uplink structure, generally referred to as 'virtual' uplink (VUL) [61], [62]. This transformation is used directly in [61], whereas the approach in [62] proposes an extension to take into account different values of the noise variance terms. Once translated to the VUL domain, a fixed point iteration can be used, which was first proposed in [59] and extended in [60]. Alternatively, in [58], strong duality of the original problem is shown using the elegant Perron-Frobenius theory of nonnegative matrices [63].

The works we have mentioned so far [58]-[66], only consider the conventional communication scenario, in which no additional PUs are present and the constraints are solely posed by the SINR requirements. On the other hand, extending the strong duality results in [58] to accomodate additional constraints is not straightforward and scenario-dependant. This problem has been initially addressed in [68], where the authors considered one group of individual beam-shaping constraints in addition to the QoS constraints and proved, using uplink downlink duality, that strong duality holds in this case. In a CR context, the problem with interference constraints has been considered in [70]-[76].

The works of [70]-[74] approach the problem via its semidefinite relaxation (SDR) reformulation. Specifically, the SDR technique consists in writing the original problem as

$$\min_{\{\mathbf{W}_k\}} \sum_{k\in S} \mathrm{Tr}\left\{\mathbf{W}_\mathrm{k}\right\} \tag{2.11a}$$

$$\text{s.to} \quad \mathrm{Tr}\left\{\mathbf{W}_k\mathbf{R}_k\right\} - \sum_{\substack{i\neq k \\ i\in S}} \gamma_k \mathbf{W}_i\mathbf{R}_k \geq \gamma_k\sigma_k^2 \tag{2.11b}$$

$$\sum_{k\in S} \mathrm{Tr}\left\{\gamma_{K+l}\mathbf{R}_{K+l}\mathbf{W}_k\right\} \leq 1;\ l=1,\ldots,L \tag{2.11c}$$

$$\mathbf{W}_k \succeq 0;\ k\in S, \tag{2.11d}$$

where the change of variables $\mathbf{W}_k \triangleq \mathbf{w}_k\mathbf{w}_k^H$ is used. In this form, (2.11) belongs to the class of convex conic semidefinite programs (SDP), which are actively studied and for which algorithms have been proposed [77]-[79]. Note that, to write the original problem as a convex SDP in (2.11), the rank one constraints on the variables $\mathbf{W}_k$ have been ignored. Therefore, (2.11) is a relaxation of the original problem, which is exact if there exist optimal weighting matrices $\{\mathbf{W}_k^\star\}$, such that the rank of $\mathbf{W}_k^\star$ is one for all k.

Using the SDR technique is convenient, however it raises two main questions: "when

can it be expected that the relaxation of the rank constraints is indeed tight" and "how can a valid feasible beamformer be retrieved from the SDR solution, when this is not rank one". Interestingly, numerical results show that in many cases, solving the SDR reformulation indeed leads to rank one solutions. It is thus interesting to find analytical conditions under which this equivalence holds. In [58], it is shown that this is the case when $L = 0$, and the proof is based on the strong duality, which can be concluded by the equivalence between the original, the VUL and the dual problems. In [86], it is claimed that a rank one solution is always obtained and strong duality holds, however the proof relies on a result from quadratic programming, which shows strong duality in a constructive manner for three quadratic constraints. Therefore, the applicability of the result in [86] is limited to a CR network with $K + L = 3$. Stronger results were obtained in [69] for $L = 1$ and in [71] for $L \leq 2$. The importance of these works is that they give upper bounds for the rank of the solution obtained from the SDR, which depend on the number of constraints. Furthermore, based on rank decomposition techniques, optimal solutions to the original quadratic problem can be found, for a limited number of additional interference temperature constraints. Thus, the advantage of rank decomposition techniques is that, when applicable, they are able to 'purify' the solution obtained by the SDR to an exact solution for the original problem. The disadvantage, however, is that, when the number of constraints is larger than three, there is no guarantee that such a solution can be obtained. Furthermore, even though attractive due to their convex cone properties, the computational complexity per iteration of teh SDR techniques is still large.

Alternative approaches for the MISO CR scenarios have been proposed in [75] and [76]. In [75], the linear precoding technique in [67], for conventional scenarios, is extended into a subgradient method for CR setups. A limitation of the method, proposed in [75], is that it only considers the case of iCSI for both SUs and PUs.

Our aim, in this thesis, is to study the beamforming problem in a more general context, i.e., with arbitrary number of PU constraints and instantaneous as well as sCSI feedback. We show several duality properties in Chapter 3 and point out, through counterexamples, that some of the features of the beamforming problem in the conventional scenarios do not naturally extend to CR setups with $L > 0$. Furthermore, we expose the VUL structure of the original problem (2.7) in a generic CR scenario, for an arbitrary number of users and iCSI as

well as sCSI, available at the transmitter. Based on this, we propose algorithms to optimally devise beamforming and power allocation schemes for the considered CR scenario. These techniques, developed in Chapter 3, are further extended in Chapter 4 to handle feasibility and admission control.

Chapter 3

Proposed Downlink Beamforming Techniques for CR Scenarios

3.1 Introduction

In this chapter, we address the beamforming problem (2.7), under the assumption that a feasible solution exists, and we propose two iterative low complexity algorithms to find the optimal beamformer and power allocation, as well as an analysis of the conditions under which strong duality is attained for the considered problem. Specifically, in Section 3.1, we propose an initial reformulation of (2.7), which exposes the VUL structure of the downlink problem. This formulation offers a nice intuition of the problem and enables the derivation of an efficient fixed point algorithm, which we present in Section 3.3. However, the performance of the algorithm derived in this manner depends on whether strong duality for the original problem holds. Therefore, to understand, under which conditions it is possible to make strong duality statements, we take a minimax approach in Section 3.2 to derive an alternative uplink formulation. Based on this, we prove, in a first stage, the equivalence between the new reformulation, the SDR and the dual problems. To show the connection to the original problem, we separately consider the cases of iCSI and sCSI. We prove equivalence for the iCSI case, and show that in the case of sCSI several additional assumptions are necessary. Moreover, we point out differences between the CR and the

conventional scenarios, which add to the difficulty of the problem in the former in the former setup. Based on the proposed formulations, we construct, in Section 3.3, two iterative low complexity algorithms and analyse their convergence. Numerical results in Section 3.4 show the performance of the proposed approaches.

3.2 An Approach Based on Uplink Downlink Duality Theory

3.2.1 Uplink Downlink Reformulation

Here, we show that the original downlink problem can be equivalently solved via an uplink reformulation, given by

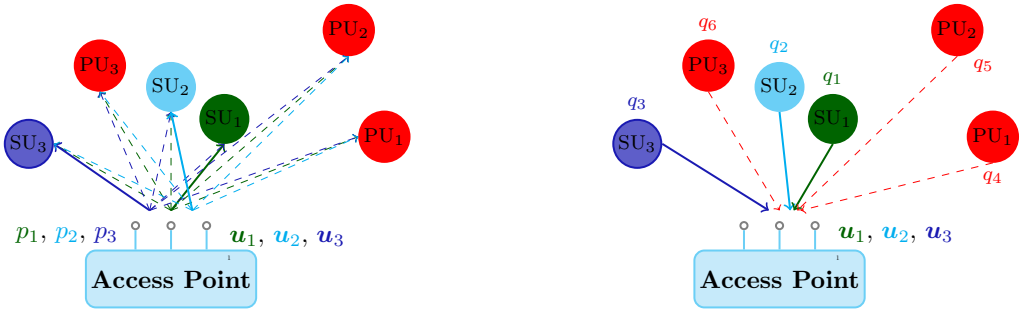
$$\mathcal{PU}_1(S, S_{PU}): \min_{\{\mathbf{u}_k, q_k\}, \{q_l\}} \sum_{k\in S} \gamma_k \sigma_k^2 q_k - \sum_{l\in S_{PU}} p_l\left(\mathbf{u}_1, \ldots, \mathbf{u}_K\right) q_l \tag{3.1a}$$

$$\text{s.to} \quad \text{SINR}_k^{\text{VUL}} = \frac{q_k \mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\mathbf{u}_k^H \left(\mathbf{I} + \sum_{\substack{i\neq k \\ i\in S}} q_i \gamma_i \mathbf{R}_i + \sum_{l\in S_{PU}} q_l \mathbf{R}_l \gamma_l\right) \mathbf{u}_k} \geq 1 \tag{3.1b}$$

$$0 \leq q_l;\ p_l(\mathbf{u}_1, \ldots, \mathbf{u}_k) \leq 1;\ l \in S_{PU}; \tag{3.1c}$$

$$||\mathbf{u}_k|| = 1;\ q_k \geq 0; \quad k \in S, \tag{3.1d}$$

where $\{q_k\}_{k\in S}$ and $\{q_l\}_{l\in S_{PU}}$ can be seen as VUL powers of the SUs and PUs, respectively. In (3.1), $\{p_l\}_{l\in S_{PU}}$ represent the interference levels experienced by the PUs, normed with respect to the corresponding imposed thresholds, and are functions of the unit norm beamformers. These interference levels are reflected in the VUL powers of the PUs. More precisely, we show here that an optimal solution for the variables $\{q_l\}_{l\in S_{PU}}$ is the one which is null when the PUs are not affected by interference, i.e., $p_l(\mathbf{u}_1, \ldots, \mathbf{u}_K) < 1$ and non-negative otherwise. Moreover, the QoS constraints are reformulated from the downlink SINR to the UL SINR in Eq. (3.1b), with the essential difference that in the latter, the beamformers are decoupled in the constraints, i.e., each constraint only contains a single beamformer. A depiction of the downlink problem and the corresponding VUL counterpart is shown in Figure 3.1. We now proceed to formally state and prove the equivalence between the original problem (2.7) and the VUL reformulation in (3.1).



(a) Downlink Scheme

(b) 'Virtual' Uplink Scheme

Figure 3.1: Illustration of the relation between the downlink and uplink scenarios represented formally by the original DL problem in Eq. (2.7) and the uplink reformulation Eq. (3.1)

Proposition 3.1 The original downlink problem (2.7) and the VUL reformulation (3.1) are equivalent, in the sense that they achieve the same optimal value, and have the same optimal beamformers. The downlink powers can be obtained from the uplink powers, via a linear transformation.

Before proceeding to the proof, we first introduce several definitions to ease the further exposition. Let

$$[\mathbf{D}]_{k,k} \triangleq \mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k \tag{3.2a}$$

$$[\mathbf{G}_1]_{i,j} \triangleq \begin{cases} 0, & i = j; \\ \gamma_i \mathbf{u}_j^H \mathbf{R}_i \mathbf{u}_j, & i \neq j \end{cases} \tag{3.2b}$$

$$[\mathbf{G}_2]_{l,j} \triangleq \gamma_l \mathbf{u}_j^H \mathbf{R}_l \mathbf{u}_j \tag{3.2c}$$

for all $l \in S_{PU}$ and $k, i, j \in S$, where $\mathbf{u}_k$ and γ_k are the kth unit norm beamformer and SINR target, respectively, as defined in Section 2.2. Note that, due to the positive semidefiniteness of the covariance matrices, all entries of $\mathbf{D}$, $\mathbf{G}_1$ and $\mathbf{G}_2$ are non-negative. Moreover, for all the beamformers, which are feasible for (2.7), $\mathbf{D}$ is strictly positive definite, as otherwise, if any element on the diagonal of $\mathbf{D}$ is null, the corresponding SINR is null and thus the constraints cannot be satisfied. We further gather the SU powers and the weighted targets

in vectors, respectively as:

$$\mathbf{p}_1 = [p_1, \ldots, p_{|S|}]^T \tag{3.3}$$

and

$$\boldsymbol{\eta} \triangleq [\gamma_1\sigma_1^2, \ldots, \gamma_{|S|}\sigma_{|S|}^2]^T. \tag{3.4}$$

Finally, we define the downlink PU powers as the interference levels experienced by each PU, and with these, we form the vector $\mathbf{p}_2 \triangleq [p_{K+1}, \ldots, p_{K+L}]$, where each element is given as $p_{K+l} = \sum_{k\in S} p_k \mathbf{u}_k^H \mathbf{R}_{K+l}\gamma_{K+l}\mathbf{u}_k$. With these definitions, (2.7) can be compactly expressed as

$$\min_{\{\mathbf{u}_k\},\mathbf{p}_1} \mathbf{1}^T\mathbf{p}_1 \tag{3.5a}$$

$$\text{s.to} \quad (\mathbf{D} - \mathbf{G}_1)\,\mathbf{p}_1 = \boldsymbol{\eta} \tag{3.5b}$$

$$\mathbf{p}_2 = \mathbf{G}_2\mathbf{p}_1 \tag{3.5c}$$

$$\mathbf{p}_1 > \mathbf{0};\ \mathbf{p}_2 \leq \mathbf{1}. \tag{3.5d}$$

Note that at the optimum of (3.5), as well as at that of (2.7), the SINR constraints in Eqs. (3.5b) and (2.7b), respectively, are satisfied with equality. This is evident from the fact that, if the SINR constraint of a particular user k is strictly larger than one, downscaling its power p_k leads to a smaller optimal value while satisfying all constraints, thus contradicting optimality. Therefore, the inequality sign in (3.5b) is replaced with equality, without loss of optimality.

Proof of Proposition 3.1 Since $\mathbf{D}$ is invertible, and $\mathbf{D}^{-1}\mathbf{G}_1$ has non-negative elements and is irreducible [1], it follows from Perron-Frobenius theory that $\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) \in (0,1)$, where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of $(\cdot)$. This implies that the matrix $(\mathbf{D} - \mathbf{G}_1)$ is invertible, and all the entries of its inverse are nonnegative [58]. Consequently, the power vector $\mathbf{p}_1 = (\mathbf{D} - \mathbf{G}_1)^{-1}\boldsymbol{\eta}$ has only nonnegative entries and thus, the power constraint $\mathbf{p}_1 \geq \mathbf{0}_{K\times 1}$ can be replaced by a constraint on the maximum eigenvalue of the matrix

[1] A matrix is said to be irreducible, if its connected graph is strongly connected [53]

$\mathbf{D}^{-1}\mathbf{G}_1$. In this way, Eq. (3.5) becomes

$$\min_{\{\mathbf{u}_k\},\mathbf{p}_1} \mathbf{1}^T\mathbf{p}_1 \tag{3.6a}$$

$$\text{s.to} \quad (\mathbf{D}-\mathbf{G}_1)\,\mathbf{p}_1 = \boldsymbol{\eta} \tag{3.6b}$$

$$\mathbf{p}_2 = \mathbf{G}_2\mathbf{p}_1;\ \mathbf{p}_2 \leq \mathbf{1} \tag{3.6c}$$

$$\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) < 1. \tag{3.6d}$$

Introducing the non-negative vector $\mathbf{q}_2 \triangleq [q_{K+1},\ldots,q_{K+L}]^T$ as slack variable, the cost function of the problem (3.6) can be expanded with the term $\mathbf{p}_2^T\mathbf{q}_2 - \boldsymbol{\eta}_1^T(\mathbf{D}-\mathbf{G}_1)^{-T}\mathbf{G}_2^T\mathbf{q}_2 = 0$, as

$$\min_{\{\mathbf{u}_k\},\mathbf{p}_1,\mathbf{q}_2} \mathbf{p}_1^T\mathbf{1}_{K\times 1} + \boldsymbol{\eta}^T(\mathbf{D}-\mathbf{G}_1)^{-T}\mathbf{G}_2^T\mathbf{q}_2 - \mathbf{p}_2^T\mathbf{q}_2 \tag{3.7a}$$

$$\text{s.to} \quad (\mathbf{D}-\mathbf{G}_1)\,\mathbf{p}_1 = \boldsymbol{\eta}_1 \tag{3.7b}$$

$$\mathbf{p}_2 = \mathbf{G}_2\mathbf{p}_1 \tag{3.7c}$$

$$\mathbf{p}_2 \leq \mathbf{1}_{L\times 1};\ \mathbf{q}_2 \geq \mathbf{0} \tag{3.7d}$$

$$\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) < 1;\ ||\mathbf{u}_k|| = 1; \quad k = 1,\ldots,K. \tag{3.7e}$$

Making explicit the power term from Eq. (3.7b) and introducing the vector

$$\mathbf{q}_1 \triangleq (\mathbf{D}-\mathbf{G}_1)^{-T}(\mathbf{1}+\mathbf{G}_2^T\mathbf{q}_2), \tag{3.8}$$

we can rewrite Eq. (3.7) as

$$\min_{\{\mathbf{u}_k\},\mathbf{q}_1,\mathbf{q}_2} \mathbf{q}_1^T\boldsymbol{\eta} - \mathbf{p}_2^T\mathbf{q}_2 \tag{3.9a}$$

$$\text{s.to} \quad (\mathbf{D}-\mathbf{G}_1)^T\,\mathbf{q}_1 - \mathbf{G}_2^T\mathbf{q}_2 = \mathbf{1} \tag{3.9b}$$

$$\mathbf{p}_2 \leq \mathbf{1}_{L\times 1}; \mathbf{q}_2 \geq \mathbf{0} \tag{3.9c}$$

$$\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) < 1;\ ||\mathbf{u}_k|| = 1; \quad k = 1,\ldots,K. \tag{3.9d}$$

Since $\mathbf{q}_2$ is non-negative, $1+\mathbf{G}_2^T\mathbf{q}_2$ is strictly positive. Then, with the same argument as before, the constraints on $\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) < 1$ can be equivalently replaced by $\mathbf{q}_1 > \mathbf{0}$ and

thus Eq. (3.9) is equivalent to

$$\min_{\{\mathbf{u}_k\},\mathbf{q}_1,\mathbf{q}_2} \quad \mathbf{q}_1^T\boldsymbol{\eta} - \mathbf{p}_2^T\mathbf{q}_2 \tag{3.10a}$$

$$\text{s.to} \quad (\mathbf{D}-\mathbf{G}_1)^T\mathbf{q}_1 - \mathbf{G}_2^T\mathbf{q}_2 = \mathbf{1} \tag{3.10b}$$

$$\mathbf{p}_2 \leq \mathbf{1}_{L\times 1};\ \mathbf{q}_2 \geq \mathbf{0} \tag{3.10c}$$

$$\mathbf{q}_1 > \mathbf{0};\ ||\mathbf{u}_k|| = 1; \quad k = 1,\ldots,K. \tag{3.10d}$$

Thus, the original problem (2.7) attains the same optimum as the reformulation in (3.10), which, can be equivalently expressed as shown in (3.1). □

Note that in problem (3.10), $\mathbf{p}_2$ cannot be considered as simple slack variables. Disregarding the dependency of $\mathbf{p}_2$ on the beamformers in Eq. (3.7c), can lead to a strictly smaller optimum of the problem in (3.10), due to the fact that the beamformers and DL SU powers are not guaranteed to satisfy the interference constraints, and thus correspond to a relaxed problem. Therefore, in this form the problem can only be seen as partially decoupled in the beamformers. However, the decoupling of the beamformers in the SINR constraints, together with optimality conditions which we show in the next section, enable the derivation of an efficient algorithm, based on the formulation in (3.10).

3.2.2 Discussion on Optimality and Strong Duality

In the conventional scenario, in which no PUs are present, it has been shown that the VUL powers can be chosen to correspond to the optimal solutions of the dual problem. We are interested to see if such a correspondence is also valid in the case of the more general CR scenarios.

First note that the existence of a global minimum of the original problem (2.7) is ensured, as in the conventional case, by the fact that the cost function is coercive and the constraint set is closed. These conditions imply that, when a feasible solution exists, a global minimum exists, as guaranteed by the Weierstrass Theorem [Theorem A.2, [51]].

Secondly, at any local optimum of the original problem (2.7), there exist Lagrange multipliers to satisfy the Karush-Kuhn-Tucker (KKT) conditions [49], which can be written

as

$$\left(\mathbf{I} - q_k\mathbf{R}_k + \sum_{\substack{i\in S\\ i\neq k}} q_i\mathbf{R}_i\gamma_i + \sum_{l\in S_{PU}} q_l\mathbf{R}_l\gamma_l\right)\mathbf{u}_k^\star = 0 \tag{3.11a}$$

$$q_k\left(\mathrm{SINR}_\mathrm{k}^\mathrm{DL} - 1\right) = 0;\ k \in S \tag{3.11b}$$

$$q_l\left(p_l\left(\mathbf{u}_1,\ldots,\mathbf{u}_K\right) - 1\right) = 0;\ l \in S_{PU}. \tag{3.11c}$$

The existence of Lagrange multipliers, to satisfy the KKT conditions formulated in (3.11), is ensured under mild assumptions, e.g., if the original problem (2.7) admits an optimal solution, which is regular. Specifically, this means that the gradients of the constraints at an optimum solution are linearly independent. Note that, in our case, this translates into a linearly independence condition of $K + L$ vectors of dimension NK. Since, in general, NK must be significantly larger than $K + L$ for the original beamforming problem to be feasible, we conclude that in most practical scenarios this condition holds. Additionally, when the regularity assumption holds, the optimum Lagrange multipliers are unique [51].

Returning to our problem, it can be easily seen that any set of optimal beamformers for the original problem, together with the corresponding Lagrange multiplers form a solution of the VUL reformulation in Eq. (3.1). Since the original problem is not convex, the KKT conditions are, however, only necessary and not sufficient for optimality. A sufficient condition for global optimality is that for each k, the dual matrix $\mathbf{Z}_k$ is psd [51], [55], i.e.,

$$\mathbf{Z}_k \triangleq \mathbf{I} - q_k\mathbf{R}_k + \sum_{\substack{i\in S\\ i\neq k}} q_i\mathbf{R}_i\gamma_i + \sum_{l\in S_{PU}} q_l\mathbf{R}_l\gamma_l \succeq 0. \tag{3.12}$$

In conventional scenarios, i.e., $L = 0$, this condition can be shown to hold at the optimum, as then, the solutions of the VUL problem (3.1) lie in the same set as the optima of the dual problem (2.10). The argument for this statement, in the particular case $L = 0$, is that, if positive semidefiniteness of some $\mathbf{Z}_k$ does not hold, then the corresponding beamformer $\mathbf{u}_k$ can be 'moved' in the direction of the negative eigenvalue. In this manner, the corresponding VUL power, q_k, can be reduced, while respecting all constraints, which contradicts optimality. In our case however, the argument holds only if it is possible to guarantee the existence of such a change in the beamformer, that does not violate the PU interference constraints, which are all functions of these variables. Even if, there may exist

beamformers for which this is satisfied, the condition does not hold in general. In fact, it is possible to construct a solution of the VUL reformulation, which attains the same optimum, but whose VUL powers violate the positive semidefiniteness of at least one of the matrices $\mathbf{Z}_k$. We show this in Appendix A. Such an example is interesting as it marks the differences between the conventional and the CR scenarios and the challenges associated with the latter.

In conclusion, a feasible solution $\left(\mathbf{q}_1^\star, \mathbf{q}_2^\star, \{\mathbf{u}_k\}_{k\in S}\right)$ of $\mathcal{PU}_1(S, S_{PU})$ is guaranteed to be optimal if the conditions in (3.11) and (3.12) are satisfied. Note that, the positive semidefiniteness can be enforced by computing the beamformers as generalized eigenvectors in the UL SINR expressions. This represents in fact the optimal beamformer strategy at a receiver in an uplink domain, where the goal is to minimize total uplink power, while satifying uplink SINR constraints. We use these ideas in Section 3.4, where we devise an algorithm to solve $\mathcal{PD}(S, S_{PU})$, by the alternative $\mathcal{PU}_1(S, S_{PU})$, while ensuring that the conditions (3.11) and (3.12) are respected.

3.3 An Approach Based on Minimax Theory

In this section, we show an alternative uplink reformulation of the original problem (2.7), which is constructed based on Lagrange duality. Specifically, we show that the original problem can be solved with an uplink reformulation, which writes as

$$\mathcal{PU}_2(S, S_{PU}) : \max_{\{q_l\}} \min_{\{\mathbf{u}_k, q_k\}} \sum_{k\in S} \gamma_k \sigma_k^2 q_k - \sum_{l\in S_{PU}} q_l \tag{3.13a}$$

$$\text{s.to} \qquad \mathrm{SINR}_{\mathrm{k}}^{\mathrm{VUL}} \triangleq \frac{\mathrm{q_k}\mathbf{u}_{\mathrm{k}}^{\mathrm{H}}\mathbf{R}_{\mathrm{k}}\mathbf{u}_{\mathrm{k}}}{\mathbf{u}_{\mathrm{k}}^{\mathrm{H}}\left(\sum_{\substack{j\neq k\\ j\in S}} \mathrm{q_j}\gamma_{\mathrm{j}}\mathbf{R}_{\mathrm{j}} + \sum_{\mathrm{l}\in\bar{\mathrm{S}}_{\mathrm{PU}}} \mathrm{q_l}\gamma_{\mathrm{l}}\mathbf{R}_{\mathrm{l}} + \mathbf{I}\right)\mathbf{u}_{\mathrm{k}}} = 1 \tag{3.13b}$$

$$q_k \geq 0;\ ||\mathbf{u}_k|| = 1;\ k \in S. \tag{3.13c}$$

As opposed to $\mathcal{PU}_1(S, S_{PU})$, in the mimimax formulation $\mathcal{PU}_2(S, S_{PU})$ the coupling between the beamformers is completely removed. We first show that this reformulation is equivalent to a partial Lagrange dual problem, corresponding to the original (2.7). Therefore, its optimum is smaller or equal than that of the original problem, and, at the same time, larger or equal than that of the total dual problem (2.10). Next, we close the gap

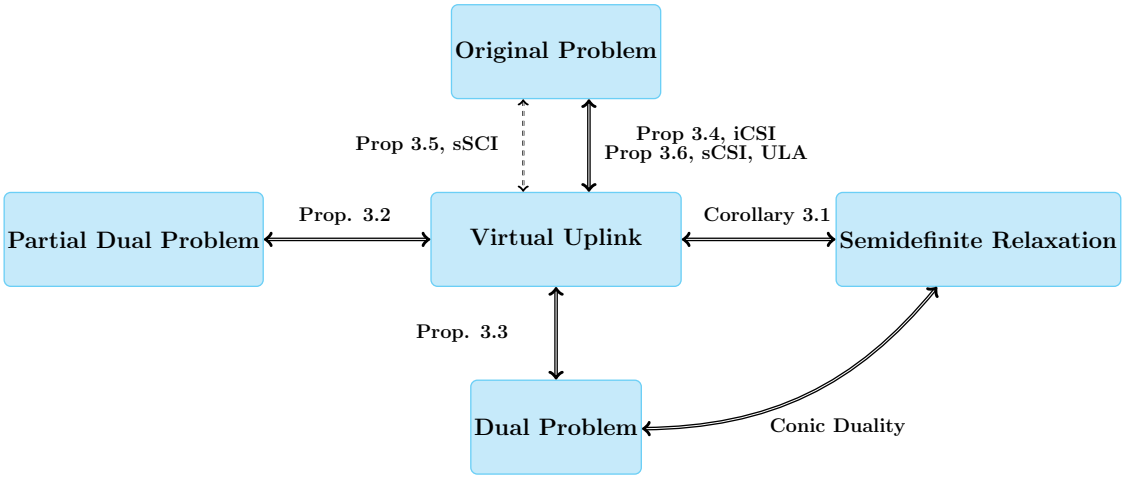


Figure 3.2: Correspondence between the problems considered in Section 3.2. From duality theory the original problem has a larger optimum than the partial dual and the total dual. The arrows with solid lines imply that the gap is tight between the connected blocks, whereas dashed line marks that additional assumptions are necessary.

between the uplink reformulation and the total dual problem, by showing that the two formulations achieve the same optimal solution. Further, to close the gap between the uplink formulation and the original problem, the separate cases of iCSI and sCSI of the SUs are considered. For the former case, duality is proven to always hold, while for the latter, we can only show this, under certain assumptions. In any case, the solution obtained by the uplink formulation in (3.13), is as tight as the state-of-the-art SDR solution, while benefiting from a significant computational reduction, as compared to this method, due to the decoupling of the beamformers. The relations mentioned so far, between the problems considered in this section are visually represented in Figure 3.2.

3.3.1 Minimax Problem Reformulation

We first construct the partial dual problem corresponding to (2.7), by relaxing the PU interference constraints. Specifically, the partial dual Lagrange function is defined as f_p :

$\mathbb{R}_+^L \to \mathbb{R}$ with

$$f_p\left(\mathbf{q}_2\right) = \min_{\{\mathbf{u}_k, p_k\}} \sum_{k\in S} p_k + \sum_{l\in S_{PU}} q_l\gamma_l \left(\sum_{k\in S} p_k \mathbf{u}_k^H \mathbf{R}_l \mathbf{u}_k - 1\right) \tag{3.14a}$$

$$\text{s.to} \quad \frac{p_k \mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\sum_{\substack{j\neq k \\ j\in S}} p_j \mathbf{u}_j^H \mathbf{R}_k \mathbf{u}_j + \sigma_k^2} \geq \gamma_k, \tag{3.14b}$$

$$p_k \geq 0; \ k \in S. \tag{3.14c}$$

The partial dual problem corresponding to (2.7) consists in the maximization of $f_p(\mathbf{q}_2)$, defined in (3.14), with respect to all non-negative $\mathbf{q}_2$, i.e.,

$$\max_{\mathbf{q}_2} f_p(\mathbf{q}_2). \tag{3.15}$$

Then, the equivalence between the partial dual problem (3.15) and the uplink formulation in (3.13), can be established, as formally stated and proved next.

Proposition 3.2 Let $f_d : \mathbb{R}_+^L \to \mathbb{R}$ be a function defined as

$$f_d\left(\mathbf{q}_2\right) \triangleq \min_{\{\mathbf{u}_k, q_k\}} \sum_{k\in S} \gamma_k \sigma_k^2 q_k - \sum_{l\in S_{PU}} q_l \tag{3.16a}$$

$$\text{s.to} \quad \mathrm{SINR}_{\mathrm{k}}^{\mathrm{VUL}} \triangleq \frac{\mathrm{q_k} \mathbf{u}_{\mathrm{k}}^{\mathrm{H}} \mathbf{R}_{\mathrm{k}} \mathbf{u}_{\mathrm{k}}}{\mathbf{u}_{\mathrm{k}}^{\mathrm{H}} \left(\sum_{\substack{j\neq k \\ j\in S}} \mathrm{q_j}\gamma_{\mathrm{j}} \mathbf{R}_{\mathrm{j}} + \sum_{\mathrm{l}\in \mathrm{S_{PU}}} \mathrm{q_l}\gamma_{\mathrm{l}} \mathbf{R}_{\mathrm{l}} + \mathbf{I}\right) \mathbf{u}_{\mathrm{k}}} = 1. \tag{3.16b}$$

Then, $\max_{\mathbf{q}_2 \geq \mathbf{0}} f_p = \max_{\mathbf{q}_2 \geq \mathbf{0}} f_d$ and the optimal beamformers which minimize the two functions are the same. Furthermore, the optimal downlink powers $\{p_k\}$ can be obtained from the variables $\{q_k\}$, using a linear transformation.

Proof Similar to the proof in Proposition 3.1, we can show that for any non-negative $\mathbf{q}_2$, $f_p(\mathbf{q}_2) = f_d(\mathbf{q}_2)$. With the definitions in Eq. (3.2), the partial dual function, f_p, can be written, for a fixed $\mathbf{q}_2$, as

$$f_p(\mathbf{q}_2) = \min_{\{\mathbf{u}_k\}_{k=1}^K, \mathbf{p}_1} \mathbf{1}_{K\times 1}^T \mathbf{p}_1 + \mathbf{q}_2^T \left(\mathbf{G}_2 \mathbf{p}_1 - \mathbf{1}_{L\times 1}\right) \tag{3.17a}$$

$$\text{s.to} \quad \left(\mathbf{D} - \mathbf{G}_1\right) \mathbf{p}_1 = \boldsymbol{\eta} \tag{3.17b}$$

$$\mathbf{p}_1 \geq \mathbf{0}. \tag{3.17c}$$

With the same argument, based on Perron-Frobenius theorem, as in the previous section, the power constraints in Eq. (3.17c) can be replaced with an inequality on the spectral radius of $\mathbf{D}^{-1}\mathbf{G}_1$, i.e., $\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) < 1$. This leads to

$$f_p(\mathbf{q}_2) = \min_{\{\mathbf{u}_k\}_{k=1}^K, \mathbf{p}_1} \mathbf{1}_{K\times 1}^T \mathbf{p}_1 + \mathbf{q}_2^T \left(\mathbf{G}_2\mathbf{p}_1 - \mathbf{1}_{L\times 1}\right) \tag{3.18a}$$

$$\text{s.to} \quad \left(\mathbf{D} - \mathbf{G}_1\right)\mathbf{p}_1 = \boldsymbol{\eta} \tag{3.18b}$$

$$\lambda_{\max}\left(\mathbf{D}^{-1}\mathbf{G}_1\right) < 1;\ k \in S. \tag{3.18c}$$

Further introducing the vector $\mathbf{q}_1$, with the definition in Eq. (3.8), we have that

$$f_d(\mathbf{q}_2) \triangleq \min_{\{\mathbf{u}_k\}_{k=1}^K, \mathbf{q}_1} \boldsymbol{\eta}^T \mathbf{q}_1 - \mathbf{1}_{L\times 1}^T \mathbf{q}_2 \tag{3.19a}$$

$$\text{s.to} \quad \left(\mathbf{D} - \mathbf{G}_1^T\right)\mathbf{q}_1 - \mathbf{G}_2^T\mathbf{q}_2 = \mathbf{1} \tag{3.19b}$$

$$\mathbf{q}_1 \geq \mathbf{0}_{K\times 1};\quad k \in S. \tag{3.19c}$$

Finally, since the equivalence of f_p and f_d holds for any $\mathbf{q}_2$, it also holds for the optimal $\mathbf{q}_2$, which maximizes f_p and f_d, respectively. Thus, $\mathcal{PU}_2(S, S_{PU})$ attains the same optimum as the partial dual problem in (3.15). □

Note that, by construction, the optimum value of the partial Lagrange dual problem (3.15) lies between the the optima of the original problem (2.7) and of the total dual (2.10). Consequently, by Proposition 3.2, the same statement holds for the VUL problem (3.13). In the following proposition, we show that the gap between the total dual and the VUL is however tight.

Proposition 3.3 The VUL reformulation (3.13) achieves the same optimal value as the dual problem (2.10).

Proof The proof implies showing that an optimal solution of (2.10) is in the feasible set of (3.13) and the optimal solution of (3.13) is in the feasible set of (2.10). Let $(\mathbf{q}_{V,1}^\star, \ldots, \mathbf{q}_{V,K+L}^\star)$ be an optimal solution for Eq. (3.13). Then, for any SU, it holds that

$$\mathbf{I} - q_{V,k}^\star \mathbf{R}_k + \sum_{\substack{j\neq k \\ j\in S}} q_{V,j}^\star \gamma_j \mathbf{R}_j \succeq 0;\ k \in S \tag{3.20}$$

This can be proven by contradiction, as follows. Note first, that the problem in Eq. (3.13) is equivalent to one, in which the virtual uplink SINR constraints are satisfied with inequality,

i.e., $\text{SINR}^{\text{UL}} \geq 1$. The argument for this is that, downscaling the virtual uplink powers leads to a contradiction of optimality if equality is not satisfied at the optimum.

Assume that the expression in (3.20) admits a negative eigenvalue. In this case, the corresponding beamformer $\mathbf{u}_k^\star$ can be 'moved' in the direction of the negative eigenvalue and thus $q_{V,k}^\star$ can be decreased, without violating the constraints. This, however, implies a smaller cost for the inner minimization, while satisfying all constraints and thus leading to a contradiction. In conclusion, any optimum of $\mathcal{PU}_2(S, S_{PU})$ lies in the feasible set of the dual problem (2.10). On the other hand, if $(\mathbf{q}_{D,1}^\star, \ldots, \mathbf{q}_{D,K+L}^\star)$ is optimal for the dual, then

$$\mathbf{I} - q_{D,k}^\star \mathbf{R}_k + \sum_{\substack{j \neq k \\ j \in S}} q_{D,j}^\star \gamma_j \mathbf{R}_j \succeq 0; \ k \in S \tag{3.21}$$

has at least one zero eigenvalue. Indeed, if for any k, the expression in Eq. (3.21) is positive definite, then $q_{D,k}^\star$ can be increased, while satisfying all constraints, which contradicts optimality. Thus, the two problems, (3.13) and (2.10), have the optima in the same set and both problems achieve the same objective function value. Then, since in Eq. (2.10) a maximization is performed with respect to the variables, the optimum of the dual must be larger than or equal to that of the VUL reformulation. On the other hand, since from Proposition 3.1, the VUL has the same optimum value as the partial dual problem, it follows from weak duality that the optimum of (3.13) is larger than or equal to that of (2.10). Therefore, equality between the two optima is obtained. □

Corollary 3.1 Assume that a strictly feasible solution exists for the original problem in (2.7). Then, the VUL formulation (3.13) attains the same optimum as the SDR formulation in (2.11).

Proof This statement results straightforwardly from conic duality. If a strictly feasible solution for the original problem exists, then a strictly feasible solution for the SDR in Eq. (2.11) exists, and, consequently, strong duality holds [Proposition 4, [52]]. Then, with Proposition 3.3, the conclusion of the proof follows.

3.3.2 Duality Results for Scenarios with Instantaneous CSI

In this subsection, we show that, when iCSI is available the VUL reformulation (3.13) is exact. This is due to the following result.

Proposition 3.4 If iCSI, i.e., $\mathbf{R}_k = \mathbf{h}_k\mathbf{h}_k^H$ is available for the SUs and a strictly feasible solution exists for $\mathcal{PD}(S, S_{PU})$ in (2.7), then the VUL reformulation (3.13) attains the same optimum as the original problem.

Proof From the existence of a strictly feasible solution for $\mathcal{PD}(S, S_{PU})$, it follows that strong duality holds for the SDR. Therefore, for any optimal solution of (2.11), $\mathbf{W}_k^\star$, and any optimal dual matrix $\mathbf{Z}_k^\star$, the slackness condition holds, i.e.,

$$\mathbf{W}_k^\star\mathbf{Z}_k^\star = \mathbf{0}. \tag{3.22}$$

From the definition of $\mathbf{Z}_k$ in (2.10b), it is easy to see that if $\mathbf{R}_k$ has rank one, $\mathbf{Z}_k$ has rank $N-1$, therefore from (3.22), it follows that $\text{rank}(\mathbf{W}_\text{k}) \leq 1$. Noting that $\mathbf{W}_k^\star \succ 0$, as otherwise the kth constraint is not feasible, completes the proof. □

Since the rank of each dual matrix is $N-1$, it follows that $\mathbf{Z}_k$ admits exactly one null eigenvalue for each k, whose corresponding eigenvector is unique up to a scaling. Furthermore, from cone duality, it follows that for every optimal $\mathbf{W}_k^\star$ and $\mathbf{Z}_k^\star$, the complementary slackness condition (3.22) holds. Therefore, normalizing this eigenvector, provides a feasible beamformer in the downlink domain. Although a phase ambiguity exists, it cancels out in the terms of the form $\mathbf{u}_i^H\mathbf{R}_j\mathbf{u}_i$, which further implies that the resulting downlink powers and interference levels are unique. A similar observation as the one in Proposition 3.4 has been made in [71]. Alternatively, the result of this proposition, can be obtained from the more general analysis, performed in the next section.

3.3.3 Duality Results for Scenarios with Covariance Based CSI

Let us now consider the general case, in which sCSI is available for both SUs and PUs. In the first statement we show the necessary and sufficient condition for the VUL formulation (3.13) and the total dual problem (2.10) to attain strong duality.

Proposition 3.5 Assume a strictly feasible solution for the original problem in Eq. (2.7) exists. Then, the VUL reformulation $\mathcal{PU}_2(S, S_{PU})$ in (3.13) attains the same optimal as the original problem, and implicitly strong duality holds, if and only if there exists a set of unit norm beamformers $\{\mathbf{u}_k\}_{k\in S}$, such that $\mathbf{G}_2\left(\mathbf{D}-\mathbf{G}_1\right)^{-1}\boldsymbol{\eta} \leq 1$.

Proof "⇒" We first show by contradiction that strong duality implies the existance of a set of beamformers, which satisfy the PU constraints. Assume that there exists no

such set of unit norm beamformers, which at the optimal of $\mathcal{PU}_2(S, S_{PU})$ satisfy the PU conditions. Let $\overline{\mathbf{q}}_1$ and $\overline{\mathbf{q}}_2$ be the optimal solution for the VUL problem (3.13), and $\overline{\mathbf{p}}_1$ and $\overline{\mathbf{p}}_2$ the corresponding power and interference vectors, respectively, obtained by minimizing the partial dual function $f_p(\overline{\mathbf{q}}_2)$. Further, let $\{\underline{\mathbf{u}}_k\}_{k\in S}$ be a set of beamformers, which are feasible for the downlink problem with $\underline{\mathbf{p}}_1$ and $\underline{\mathbf{p}}_2$, the corresponding SU power and PU interference vectors, respectively. Then, for the optimal PU VUL power vector, $\overline{\mathbf{q}}_2$, we have that

$$\overline{\mathbf{q}}_1^T\boldsymbol{\eta} - \overline{\mathbf{q}}_2^T\mathbf{1} = \overline{\mathbf{p}}_1^T\mathbf{1} + (\overline{\mathbf{p}}_2 - \mathbf{1})^T\overline{\mathbf{q}}_2 \tag{3.23a}$$

$$< \underline{\mathbf{p}}_1^T\mathbf{1} + \left(\underline{\mathbf{p}}_2 - \mathbf{1}\right)^T\overline{\mathbf{q}}_2 \tag{3.23b}$$

$$\leq \mathbf{p}_1^{\star T}\mathbf{1}. \tag{3.23c}$$

The inequality in (3.23b) is due to the fact that, for any $\mathbf{q}_2$, if $\{\underline{\mathbf{u}}_k\}$ is feasible for the corresponding DL problem, then it is feasible for the the system of VUL SINRs. Consequently, the cost function for the solution obtained with this set of beamformers must be larger than that obtained by the optimizer of VUL. Further, since strict inequality holds for all feasible solutions in (3.23b), then it also holds for the optimum of (2.7). This, however, contradicts the initial assumption, that the two problems attain the same optimum.

"$\Leftarrow$" Let $\{\overline{\mathbf{u}}_k\}_{k\in S}$ be a set of optimal beamformers for the VUL reformulation (3.13), which satisfies $\mathbf{G}_2\left(\mathbf{D} - \mathbf{G}_1\right)^{-1}\boldsymbol{\eta} \leq 1$. It can be easily seen that if this set of beamformers satisfies the constraints, then fixing $\mathbf{p} = (\mathbf{D} - \mathbf{G}_1)^{-1}\boldsymbol{\eta}$ and further defining $\mathbf{W}_k = p_k\overline{\mathbf{u}}_k\overline{\mathbf{u}}_k^H$, for all $k \in S$, we obtain a feasible solution for the SDR reformulation (2.11). Further, from Proposition 3.2 any solution of the VUL problem (3.13) is a solution of the dual (2.10), and whenerver strict feasibility is satisfied these two problems attain the same optimal value as the SDR formulation (2.11). Since $\{\mathbf{W}_k\}$ and $\{\mathbf{q}_k\}$ are feasible in the primal and dual domain, respectively, it follows from cone duality [Proposition 4, [52]], that the pair is optimal and strong duality holds. Thus, a rank one solution is optimal for the SDR formulation (2.11), which implies that strong duality holds for the original problem. □

In conclusion, the equivalence between the original problem $\mathcal{PD}(S, S_{PU})$ and $\mathcal{PU}_2(S, S_{PU})$ is guranteed if and only if there exists a vector in the null space of each $\mathbf{Z}_k$, for which the primary constraints are satisfied. For the case when iCSI of the SUs is available at the transmitter, it follows from the discussion in Section 3.3.2, that this condition must be

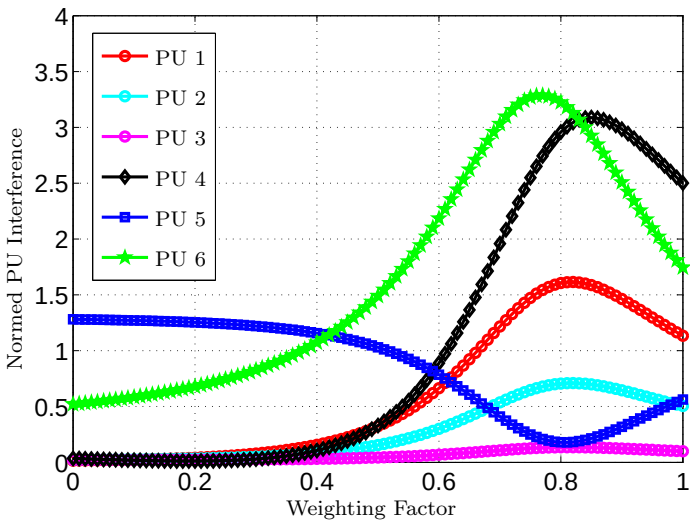


Figure 3.3: Example of violation of the PU constraints by random vectors in the null space of the dual matrices. Values larger than one imply that the obtained PU interferences are larger than the imposed interference thresholds.

satisfied, due to the complementarity slackness constraints. In the case of sCSI, this statement is however not straightforward. The following remark gives a hint regarding where the difficulties lie in evaluating the CR scenarios with sCSI.

Remark 1 (regarding optimal beamformers) Contrary to the conventional case with $L = 0$, when CR scenarios are assumed, and the channel covariance matrices have general rank, the statement, that any vector in the null space of the dual matrices $\{\mathbf{Z}_k\}$ is a feasible beamformer for the original problem (2.7), is not necessarily true. It is possible to expose cases, in which, vectors from the null space of $\mathbf{Z}_k$ do not lead to feasible solutions for the primary constraints, even for problem instances which do not exhibit a duality gap. Such an example is depicted in Figure 3.3, and corresponds to a scenario with $N = 12$ antennas, 6 PUs and 3 SUs, positioned at $[-65,\ -60,\ -55,\ -35,\ -25,\ -15]^\circ$ and, at $[-5,\ 10,\ 25]^\circ$ relative to the antenna broadside, respectively. A scattering angle of 2° is considered for the SUs, whereas this parameter is chosen as zero for the PUs. In this case, the null space of the dual matrix corresponding to the first two SUs has dimension one, while that of the third SU is two. Considering two orthogonal eigenvectors i.e., $\mathbf{u}_{3,1}$ and $\mathbf{u}_{3,2}$ in the null space of this user, we can construct a beamformer as a normed linear combination

$\mathbf{u}_3 \triangleq (\alpha\mathbf{u}_{3,1} + (1-\alpha)\mathbf{u}_{3,2}) / \|\alpha\mathbf{u}_{3,1} + (1-\alpha)\mathbf{u}_{3,2}\|$. With this newly obtained beamformer, downlink powers can be computed using Eq. (3.5b). However, as seen in Figure 3.3, for a real choice of the parameter $\alpha \in [0,1]$, no solution simultaneously satisfies all PU interference constraints, although for this problem strong duality holds. However, a beamformer exists in the null space of the dual matrices, and is feasible in the downlink domain.

Remark 2 (regarding the VUL problem with fixed beamformers) One approach to prove uplink downlink duality is to show first, that the statement holds for fixed beamformers, and then extrapolate the result to the general problem. In our case, for fixed feasible beamformers, the VUL reformulation in (3.13) is indeed equivalent to the original problem (2.7). This can be shown as follows. As shown in Appendix A, when the beamformers are fixed and feasible, the original problem (2.7) is an LP, for which strong duality holds. Thus, for fixed beamformers, (2.7) achieves the same optimum as the LP dual problem

$$\max_{\mathbf{q}_1,\mathbf{q}_2} \ \mathbf{q}_1^T\boldsymbol{\eta} - \mathbf{q}_2^T\mathbf{1} \tag{3.24a}$$

$$\text{s.to } (\mathbf{D}-\mathbf{G}_1)^T\mathbf{q}_1 - \mathbf{G}_2^T\mathbf{q}_2 = \mathbf{1} \tag{3.24b}$$

$$\mathbf{q}_1 \geq \mathbf{0};\ \mathbf{q}_2 \geq \mathbf{0}. \tag{3.24c}$$

Furthermore, the spectral radius condition, i.e., $\lambda_{\max}(\mathbf{D}^{-1}\mathbf{G}_1) < 1$ is satisfied for any feasible set of beamformers. This implies that $\mathbf{q}_1$ is unique for any non-negative $\mathbf{q}_2$, and Eq. (3.24) is equivalent to

$$\max_{\mathbf{q}_2} \min_{\mathbf{q}_1} \ \mathbf{q}_1^T\boldsymbol{\eta} - \mathbf{q}_2^T\mathbf{1} \tag{3.25a}$$

$$\text{s.to} \quad \frac{q_k\mathbf{u}_k^H\mathbf{R}_k\mathbf{u}_k}{\mathbf{u}_k^H\left(\sum_{\substack{i\in S\\ i\neq k}} q_i\mathbf{R}_i\gamma_i + \sum_{l\in S_{PU}} q_l\mathbf{R}_l\gamma_l + 1\right)\mathbf{u}_k} \geq 1. \tag{3.25b}$$

In conclusion, when the transmit beamformers are fixed, the VUL formulation (3.13) achieves the same optimum as the original problem. The result, however, does not generalize directly to the case of variable beamformers, unless it is possible to show that the optimum beamformers for the original problems are also optimal for the inner minimization problem of the VUL (3.13).

3.3.4 A Particular Case of Strong Duality

When only sCSI is available for both SUs and PUs, a more general statement regarding the strong duality of the original problem is still an open problem. Under particular model assumptions it is however possible to prove that the SDR formulation (2.11) admits a rank one solution, therefore ensuring that the original problem (2.7) and its dual achieve the same optimal. More precisely, we show that strong duality holds for a channel model which simultaneously respects the following set of assumptions.

Assumptions A.1

i) the antenna configuration at the BS is uniform linear array (ULA)

ii) the distribution of rays around the nominal direction from which the signal from a user impinges on the array, is approximately normal with zero mean and σ_θ standard deviatiation

iii) the complex gains of the paths are assumed independent from snapshot to snapshot and from ray to ray.

Let θ_k be the angle of incidence of the signal from user k to the BS array. Under the above mentioned assumptions A.1, the covariance matrix can be written in closed form [80] as

$$\mathbf{R}_k(\theta_k) = \begin{bmatrix} \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} e^{j0g(\theta_k,\phi)} d\phi & \dots & \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} e^{-j(N-1)g(\theta_k,\phi)} d\phi \\ \dots & & \dots \\ \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} e^{j(N-1)g(\theta_k,\phi)} d\phi & \dots & \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} e^{j0g(\theta_k,\phi)} d\phi \end{bmatrix} \tag{3.26}$$

where

$$g(\theta_k, \phi) = 2\pi\Delta \sin(\theta_k + \phi). \tag{3.27}$$

Using a small spread angle approximation and explicitating the integrals, the expression in Eq. (3.26) is equivalent to the familiar model [80], which is commonly used in simulations.

To prove strong duality, we show that the SDR of the original problem (2.7) always admits a rank one solution. Furthermore, the proof bases on the Fejer-Riesz theorem which states the following.

Theorem (Fejer Riesz [82]) Let $P_c(e^{jt}) = \sum_{l=-N+1}^{N-1} c_l e^{jlt}$ be a trigonometric polynomial, which takes real and non-negative values for any real t. Then P_c can be factorized as $P_c(e^{jt}) = P_a^*(e^{jt})P_a(e^{jt})$, where $P_a(z) = \sum_{l=0}^{N} a_l z^l$ is a polynomial obtained as

$$P_a(z) = \sqrt{C}\prod_{l=1}^{N}(z-\alpha_l), \tag{3.28}$$

$\alpha_1, \ldots, \alpha_N$ represent the roots of $z^N P_c(z)$, which lie within the unit circle, and $(\cdot)^*$ denotes the complex conjugate of $(\cdot)$.

Sketch of Proof The proof of this statement stems from the observation that $P_c(z) = P_c^*(1/z^*)$, for any complex z, which implies that the roots of the poylnomial $z^N P_c(z)$ appear in pairs $\{\alpha_l,\ 1/\alpha_l^*\}$, with equal multiplicity. □

Proposition 3.6 Let $\widetilde{\mathbf{W}}$ be an N-by-N Hermitian psd matrix of rank r. Then, there exists $\widetilde{\mathbf{w}}$ such that $\mathrm{Tr}\{\widetilde{\mathbf{W}}\mathbf{R}(\theta_\mathrm{k})\} = \widetilde{\mathbf{w}}^\mathrm{H}\mathbf{R}(\theta_\mathrm{k})\widetilde{\mathbf{w}}$ for any θ_k.

Proof Since $\widetilde{\mathbf{W}}$ is Hermitian positive semidefinite, there exist vectors $\widetilde{\mathbf{w}}_1 \ldots \widetilde{\mathbf{w}}_r$ such that

$$\widetilde{\mathbf{W}} = \sum_{i=1}^{r} \widetilde{\mathbf{w}}_i \widetilde{\mathbf{w}}_i^H. \tag{3.29}$$

Then, we have

$$\mathrm{Tr}\{\widetilde{\mathbf{W}}\mathbf{R}_k(\theta_k)\} = \mathrm{Tr}\left\{\sum_{i=1}^{r} \widetilde{\mathbf{w}}_i \widetilde{\mathbf{w}}_i^H \mathbf{R}_k(\theta_k)\right\} \tag{3.30}$$

$$= \sum_{l=0}^{N-1}\sum_{m=0}^{N-1}\sum_{i=1}^{r} \widetilde{w}_{i,l}\widetilde{w}_{i,m}^* \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} e^{-j2\pi(l-m)g(\theta_k,\phi)} d\phi \tag{3.31}$$

$$= \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} \underbrace{\sum_{l=-N+1}^{N-1} f_l(\widetilde{\mathbf{w}}_1, \ldots, \widetilde{\mathbf{w}}_r) e^{jlg(\theta_k,\phi)}}_{P_f(e^{jg(\theta_k,\phi)})} d\phi \tag{3.32}$$

where the lth function $f_l(\widetilde{\mathbf{w}}_1, \ldots \widetilde{\mathbf{w}}_r)$ results from grouping the terms corresponding to the l-th exponentials and can be written as

$$f_l(\widetilde{\mathbf{w}}_1, \ldots \widetilde{\mathbf{w}}_r) \triangleq \begin{cases} \sum_{m=l}^{N-1}\sum_{i=1}^{r} \widetilde{w}_{i,m}\widetilde{w}_{i,m-l}^*, & l = 0, \ldots, N-1 \\ f_{-l}^*, & l = -(N-1), \ldots, -1 \end{cases} \tag{3.33}$$

We can easily notice that $P_f(e^{jg(\theta_k,\phi)}) \triangleq \sum_{l=-N+1}^{N-1} f_l(\widetilde{\mathbf{w}}_1, \ldots \widetilde{\mathbf{w}}_r) e^{jlg(\theta_k,\phi)}$ is a trigonometric polynomial, which is non-negative for any real θ_k and ϕ, as P_f results from a sum of

quadratic forms

$$P_f = e^{\frac{\phi^2}{2\sigma_\phi^2}} \sum_{i=1}^{r} \widetilde{\mathbf{w}}_i^H \mathbf{R}_k(\theta_k) \widetilde{\mathbf{w}}_i. \tag{3.34}$$

From the Fejer-Riesz theorem, it follows that P_f can be equivalently represented as $P_f(e^{jg(\theta_k,\phi)}) = P_w^*(e^{jg(\theta_k,\phi)}) P_w(e^{jg(\theta_k,\phi)})$, where $P_w(e^{jg(\theta_k,\phi)} = \sum_{i=0}^{N} \widetilde{w}_{o,i} e^{jg(\theta_k,\phi)}$, such that the coefficients $\widetilde{w}_{o,0}, \dots \widetilde{w}_{o,N}$ of P_w do not depend on θ_k but solely on $\mathbf{w}_1, \dots, \mathbf{w}_N$ through the relations given by Eq. (3.33). Denoting $\mathbf{b}_{\theta_k,\phi} \triangleq \left[1,\ e^{jg(\theta_k+\phi)}, \dots, e^{j(N-1)g(\theta_k+\phi)}\right]^T$ and $\widetilde{\mathbf{w}}_o = [\widetilde{w}_{o,0} \dots \widetilde{w}_{o,N-1}]$, (3.30) can now be written as

$$\mathrm{Tr}\{\widetilde{\mathbf{W}}\mathbf{R}_{\mathrm{k}}(\theta_{\mathrm{k}})\} = \int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} \widetilde{\mathbf{w}}_{o,i}^H \mathbf{b}_{\phi,\theta_k} \mathbf{b}_{\phi,\theta_k}^H \widetilde{\mathbf{w}}_{o,i} d\phi \tag{3.35}$$

$$= \widetilde{\mathbf{w}}_o^H \left(\int_{-\infty}^{\infty} e^{-\frac{\phi^2}{2\sigma_\phi^2}} \mathbf{b}_{\phi,\theta_k} \mathbf{b}_{\phi,\theta_k}^H d\phi \right) \widetilde{\mathbf{w}}_o \tag{3.36}$$

$$= \widetilde{\mathbf{w}}_o^H \mathbf{R}_k(\theta_k) \widetilde{\mathbf{w}}_o. \tag{3.37}$$

□

From Proposition 6.1, it follows that for any beamforming matrix $\mathbf{W}_k$, obtained as an optimal solution of the SDR formulation (2.11), a beamforming vector $\mathbf{w}_k$ can be constructed such that $\mathrm{Tr}\{\mathbf{W}_k \mathbf{R}_l\} = \mathbf{w}_k^H \mathbf{R}_l \mathbf{w}_k$, for all $l \in S \cup S_{PU}$, as long as the covariance matrices have the form in (3.26).

We remark, in the end, that a related result has been very recently published [72], in which strong duality is proven for a homogenous QCQP problem, in which all matrices in the quadratic constraints are Toeplitz Hermitian. However, the result in [72] does not directly apply to our problem, as bringing (2.7) to the form of the problem in [72], destroys the Hermitian Toeplitz structure of the matrices in the constraints. On the other hand, a careful inspection of the approach in [72] may provide additional insight into the problem structure and possibly new approaches, which are also applicable to the scenario, considered in this thesis.

3.4 Proposed Algorithms

In this section, we propose two algorithms to solve the original problem in (2.7), using the uplink reformulations proposed in Sections 3.4.1 and 3.4.2. The first algorithm, in

Section 3.4.1 is based on the reformulation in (3.1), together with the optimality conditions, mentioned in Section 3.1.2. On the other hand, the algorithm in Section 3.4.2 employs the partial dual structure of the reformulation in (3.13), and, based on this, a subgradient method is devised.

3.4.1 Fixed Point Algorithm

Our approach to solve the original problem in (2.7), through the VUL reformulation in (3.1), is summarized in Table 3.1.

We assume for now that, at initialization, a set of beamformers and powers is available, such that the SINR constraints of all SUs are simultaneously satisfied under a generic total sum power constraint and in the absence of the PUs. We term such a point SU feasible and note that it can be efficiently obtained as the solution of an SINR balancing problem, expressed as

$$\max_{\{p_k,\mathbf{u}_k\}} \min_{k\in S} \frac{\mathrm{SINR}_k^{\mathrm{DL}}}{\gamma_k} \tag{3.42a}$$

$$\text{s.to } \mathbf{p}_1^T\mathbf{1} \leq P_{\max}, \tag{3.42b}$$

when the optimum of this is larger than one. With the beamformers obtained from (3.42), the initial downlink powers, $\mathbf{p}_1(0)$, and interference terms, $\mathbf{p}_2(0)$, can be computed from Eq. (3.5b) and Eq. (3.5c), respectively. Finally, we initialize the PU VUL powers as $\mathbf{q}_2(0) = \mathbf{p}_2(0)$, and the SU VUL powers, $\mathbf{q}_1(0)$, as the solution of the linear system of equations in (3.10), for the chosen beamformers and VUL PU values. An algorithm for solving (3.42) has been proposed in [64] and shown to exhibit superlinear convergence. A discussion about infeasibility is deferred to the next chapter.

After the initilization procedure, we successively update the beamformers and VUL powers, as follows. The decoupled nature of the $\mathrm{SINR}^{\mathrm{VUL}}$ enables an efficient computation of the beamformers as generalized eigenvectors, as shown in Step 2 of Table 3.1. The interference levels, caused by the new beamformers to the PUs, are computed in Step 3, and, according to these, the VUL PU powers, $\{q_l\}_{l\in S_{\mathrm{PU}}}$, are updated. More precisely, the logic of the update is that, if some interference threshold $p_l(t)$ at the t-th iteration is violated, the corresponding PU VUL power q_l is increased, in order to put more effort on satisfying this constraint in the following iterations. If, on the contrary, $p_l(t)$ is strictly

Algorithm 3.1 Fixed Point Method

Step 1 Initialize $\{\mathbf{u}_k\}_{k\in S}$ such these are SU feasible and $\{q_k\}_{k\in S}$ and $\{q_l\}_{l\in S_{PU}}$ with the values corresponding to the SU feasible point

Step 2 Compute the uplink SU powers and beamformers

2.1 Define the function $\mathcal{E}_k$ as

$$\mathcal{E}_k(\mathbf{q}_1(t),\mathbf{q}_2(t),\mathbf{u}_k) \triangleq \frac{\mathbf{u}_k^H \left(\sum_{i\in S,\ i\neq k}\ q_i(t)\gamma_i\mathbf{R}_i + \sum_{l\in S_{PU}} q_l(t)\gamma_l\mathbf{R}_l + 1\right)\mathbf{u}_k}{\mathbf{u}_k^H\mathbf{R}_k\mathbf{u}_k} \tag{3.38}$$

2.2 Update the beamformers with the solutions of the minimization of $\mathcal{E}_k$ with respect to $\mathbf{u}_k$, for all $k \in S$ as

$$\mathbf{u}_k(t+1) = \text{argmin}_{\mathbf{u}_k}\mathcal{E}_k(\mathbf{u}_k,\mathbf{q}_1(t),\mathbf{q}_2(t)) \tag{3.39}$$

2.3 Reconstruct the matrices $\mathbf{D}(t+1)$, $\mathbf{G}_1(t+1)$ and $\mathbf{G}_2(t+1)$ as in Eq. (3.2)

Step 3 Update the downlink powers, $\mathbf{p}_1(t+1)$, and the interference levels, $\mathbf{p}_2(t+1)$, with Eq. (3.5b) and Eq. (3.5c), respectively

Step 4 Update the VUL PU powers

$$q_l(t+1) = q_l(t)p_l(t+1));\ l \in S_{PU} \tag{3.40}$$

Step 5 Update the VUL SU powers with the solution of the linear system of equations

$$q_k(t+1) = \mathcal{E}_k(\mathbf{q}_1(t+1),\mathbf{q}_2(t+1),\mathbf{u}_k(t+1));\ k \in S \tag{3.41}$$

Step 6 Check convergence

if $\sqrt{\sum_{i\in S\cup S_{PU}} |q_i(t+1) - q_i(t)|^2} > \epsilon$, go to Step 2, **else** exit

Table 3.1: The uplink downlink algorithm corresponding to the formulation in Section 3.1

smaller than one, then its corresponding PU power is decreased. Finally, the SU VUL powers are updated as the solution of the linear system of equations (3.41), in Step 5. The

procedure is repeated, until the norm of the difference between two consecutive updates is below a desired precision, denoted as ϵ.

Remark 3 (regarding positivity of DL powers and VUL feasibility) The initialization, together with the update in Step 5, ensures that, the intermediary downlink powers, obtained at each iteration, are positive. Consequently, at each step, the SU and PU VUL powers are positive and non-negative, respectively, and these form a feasible solution in the VUL domain, i.e., SINR^{VUL} constraints are satisfied. Indeed, if at initialization an SU feasible solution exists, then for the beamformers obtained at this step, the spectral radius condition $\lambda_{\max}(\mathbf{D}^{-1}\mathbf{G}_1^T) < 1$ holds. Therefore, feasibility is ensured for the VUL problem, and $\mathbf{q}_1$ exists, such that for each element k, it holds that $q_k \geq \mathcal{E}_k(\mathbf{q}_1, \mathbf{q}_2(0), \mathbf{u}_k(0))$, regardless of the choice of non-negative $\mathbf{q}_2(0)$. Initializing $\mathbf{q}_1(0)$ in this manner, we naturally have that the inequality also holds after the beamformer update in Step 2, i.e.,

$$q_k(0) \geq \min_{\mathbf{u}_k} \mathcal{E}_k(\mathbf{q}_1(0), \mathbf{q}_2(0), \mathbf{u}_k) \; k \in S. \tag{3.43}$$

To show that for the matrices $\mathbf{D}(1), \mathbf{G}_1(1)$, constructed with the newly updated beamformers, the spectral radius conditions $\lambda_{\max}(\mathbf{D}^{-1}\mathbf{G}_1) < 1$ holds, Collatz-Wielandt maximin characterization can be employed. Specifically, this can be proven as follows. Explicitating the terms $\mathcal{E}_k$, in Eq. (3.43), we have that, for any positive vector $\mathbf{v}$

$$\mathbf{v}^T\mathbf{q}_1(0) \geq \mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_1^T(1)\mathbf{q}_1(0) + \mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_2^T(1)\mathbf{q}_2(0) + \mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{1} \tag{3.44}$$

Since $\mathbf{q}_1(0) > 0$, it follows that $\mathbf{v}^T\mathbf{q}_1(0) > 0$ and subsequently that

$$1 \geq \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_1^T(1)\mathbf{q}_1(0)}{\mathbf{v}^T\mathbf{q}_1(0)} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_2^T(1)\mathbf{q}_2(0)}{\mathbf{v}^T\mathbf{q}_1^T(0)} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{1}}{\mathbf{v}^T\mathbf{q}_1^T(0)} \tag{3.45}$$

$$\geq \sup_{\mathbf{v}} \left\{ \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_1^T(1)\mathbf{q}_1(0)}{\mathbf{v}^T\mathbf{q}_1(0)} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_2^T(1)\mathbf{q}_2(0)}{\mathbf{v}^T\mathbf{q}_1^T(1)} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{1}}{\mathbf{v}^T\mathbf{q}_1^T(0)} \right\} \tag{3.46}$$

$$\geq \sup_{\mathbf{v}} \inf_{\mathbf{q}_1 > \mathbf{0}} \left\{ \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_1^T(1)\mathbf{q}_1}{\mathbf{v}^T\mathbf{q}_1} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_2^T(1)\mathbf{q}_2}{\mathbf{v}^T\mathbf{q}_1^T} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{1}}{\mathbf{v}^T\mathbf{q}_1^T}. \right\} \tag{3.47}$$

Since $\sup_{\mathbf{v}} \inf_{\mathbf{q}_1 > \mathbf{0}} \left\{ \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_2^T(1)\mathbf{q}_2}{\mathbf{v}^T\mathbf{q}_1^T(1)} + \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{1}}{\mathbf{v}^T\mathbf{q}_1^T(1)} \right\} > 0$, it follows that

$$\sup_{\mathbf{v}} \inf_{\mathbf{q}_1 > \mathbf{0}} \left\{ \frac{\mathbf{v}^T\mathbf{D}^{-1}(1)\mathbf{G}_1^T(1)\mathbf{q}_1}{\mathbf{v}^T\mathbf{q}_1} \right\} < 1, \tag{3.48}$$

which implies $\lambda_{\max}(\mathbf{D}^{-1}(1)\mathbf{G}_1(1)) < 1$, due to the Collatz-Wielandt max min characterization of the spectral radius [53]. The same reasoning can be applied, to show that VUL

feasibility, and implicitly the spectral radius condition is satisfied at all steps of the algorithm. Obtaining meaningful values in the DL domain, i.e., positive DL powers and non-negative interference levels, proves mostly useful, when feasibility control and user selection procedures are required, as we will show in Chapter 4.

Remark 4 (regarding alternative update rules for $\mathbf{q}_1$) The SU VUL powers can also be updated with a fixed point iteration of the form $q_k(t+1) = \mathcal{E}_k(\mathbf{q}_1(t), \mathbf{q}_2(t+1), \mathbf{u}_k(t+1))$, for all k, instead of computing the solution of the linear equations in Step 5. Simulations show that, also in this case, the spectral radius condition and the VUL feasibility are satisfied. A formal proof is however complicated, in this case, due to the fact that PU interference levels can either increase or decrease with the VUL PU updates. If, at an iteration, the VUL powers are feasible, and the PU interference levels obtained after the beamformer updates are below one, then, VUL feasibility is respected after the SU and PU power updates, which can be shown with a similar argument as before. On the other hand, if some of the PU interference conditions are not respected, and some entry of $\mathbf{p}_2$ is larger than one, then it is not sure whether VUL feasibility continues to be satisfied. Thus it cannot be guaranteed that the $\lambda_{max}(\mathbf{D}^{-1}\mathbf{G}_1) < 1$, and implicitly that the corresponding DL powers are positive. Furthermore, other initializations than the one based on SU feasibility, may also lead to negative intermediary powers, and therefore, the PU updates must be modified to accomodate these cases, as we showed in [76].

Remark 5 (regarding global optimality) At convergence, the resulting downlink powers represent the global optimal solution of Problem (2.7). This is due to the fact that the resulting $\mathbf{q}_1^\star$, $\mathbf{q}_2^\star$ and $\{\mathbf{u}_k^\star\}$, satisfy the necessary KKT conditions, as well as the sufficient conditions for global optimality, formulated in Section 3.1.2. Specifically, the positive semidefiniteness of the dual matrices $\mathbf{Z}_k^\star$, formed with $\mathbf{q}_1^\star$ and $\mathbf{q}_2^\star$ is ensured by the generalized eigenvalue problem. Further, since $\mathbf{Z}_k^\star$ is psd, $\mathbf{u}_k^{\star H}\mathbf{Z}_k^\star\mathbf{u}_k^\star = 0$, implies $\mathbf{Z}_k^\star\mathbf{u}_k^\star = 0$, thus the first order optimality condition in Eq. (3.11a) is satisfied. Finally, the PU update enforces that the complementarity slackness in Eq. (3.11c) holds. The algorithm converges, only if strong duality holds, as otherwise, there exists no pair of VUL powers, $\mathbf{q}_1$ and $\mathbf{q}_2$ such that the KKT conditions and the positive semidefiniteness of all $\mathbf{Z}_k$ are simultaneously satisfied.

Remark 6 (regarding complexity) The fixed point algorithm, presented in Table

3.1, mainly consists of K maximum eigenvalue/eigenvector problems to obtain the optimum beamformers, and two systems of linear equations of dimension K, to obtain the intermediary DL and VUL powers, respectively, for the SUs. Thus, the complexity is of the order $\mathcal{O}\left(KN^2 + K^2\right)$. In comparison, the SDP, which is the benchmark and state-of-the-art solution, has a complexity of $\mathcal{O}\left(\max\left(K+L, NK\right)^4\right)$ or $\mathcal{O}\left((KN)^3\right)$, depending on the implementation [79]. The number of iterations, which the SDR requires to converge within a precision ϵ is of order $\mathcal{O}\left((KN)^{1/2}\log 1/\epsilon\right)$. As discussed in Section 2.3, it is possible that the solution of the SDR has a rank higher than one, in which case, randomization procedures [56], stochastic optimization [57], or rank reduction techniques [71] must be performed, which further add to the complexity.

3.4.2 Subgradient Algorithm

The reformulation $\mathcal{PU}_2(S, S_{PU})$ in Section 3.2 suggests an alternative implementation of the originl problem (2.7), based on subgradient methods, which we show here. We remind that the subgradient of a function $f : \mathbb{R}^N \to \mathbb{R}$ at a point $\mathbf{a}$ is a function $g : \mathbb{R}^N \to \mathbb{R}$, which satisfies $f(\mathbf{b}) \geq f(\mathbf{a}) + g^T(\mathbf{a})(\mathbf{b}-\mathbf{a})$ for all $\mathbf{b} \in \mathbb{R}^N$. Similarly, g is a supragradient at $\mathbf{a}$, if the inequality is reversed, i.e., $f(\mathbf{b}) \leq f(\mathbf{a}) + g^T(\mathbf{a})(\mathbf{b}-\mathbf{a})$, for all $\mathbf{b}$. It is easy to note that a supragradient of $f_p(\mathbf{q}_2)$, at a non-negative $\overline{\mathbf{q}}_2$, is $\overline{\mathbf{p}}_2 - \mathbf{1}$, where $\overline{\mathbf{p}}_2$ is the vector of the interference levels to the PUs, caused by the beamformers which minimize $f_d(\overline{\mathbf{q}}_2)$. This results from the following sequence of inequalities. For an arbitrary non-negative $\tilde{\mathbf{q}}_2$,

$$
\begin{aligned}
f_p(\tilde{\mathbf{q}}_2) = & \min_{\{\mathbf{u}_k\},\mathbf{p}_1,\mathbf{p}_2} \mathbf{p}_1^T\mathbf{1} + \tilde{\mathbf{q}}_2^T(\mathbf{p}_2 - \mathbf{1}) && \text{(3.49a)}\\
& \quad \text{s.to} \quad (\mathbf{D} - \mathbf{G}_1)\mathbf{p}_1 = \boldsymbol{\eta} \\
\leq & \ \overline{\mathbf{p}}_1^T\mathbf{1} + \tilde{\mathbf{q}}_2^T(\overline{\mathbf{p}}_2 - \mathbf{1}) && \text{(3.49b)}\\
= & \ \overline{\mathbf{p}}_1^T\mathbf{1} + \overline{\mathbf{q}}_2^T(\overline{\mathbf{p}}_2 - \mathbf{1}) + (\tilde{\mathbf{q}}_2 - \overline{\mathbf{q}}_2)^T(\overline{\mathbf{p}}_2 - \mathbf{1}) && \text{(3.49c)}\\
= & \ f_p(\overline{\mathbf{q}}_2) + (\overline{\mathbf{p}}_2 - \mathbf{1})^T(\tilde{\mathbf{q}}_2 - \overline{\mathbf{q}}_2). && \text{(3.49d)}
\end{aligned}
$$

Since $\overline{\mathbf{p}}_1$ is optimal for $f_p(\overline{\mathbf{q}}_2)$, then it is in the feasible set of the minimization problem contained in $f_p(\tilde{\mathbf{q}}_2)$, for any non-negative $\tilde{\mathbf{q}}_2$. However, these may not be optimal for $f_p(\tilde{\mathbf{q}}_2)$, thus the inequality in (3.49b) results.

With these observations, we are ready to present the algorithm, which is summarized

in Table 3.2. After the initialization stage, which is similar to the one employed in the previous algorithm, we proceed to solve the inner minimization problem in Eq. (3.13). Due to the decoupled nature of the VUL SINRs, this can be performed by alternatively updating the beamformers as generalized eigenvectors, in Step 2.1 and updating the VUL SU powers, in Step 2.2, until convergence. In order to clearly differentiate the operations of the inner and outer optimization problems, we have introduced the variables t_{in} and t_{out}, to mark the iteration numbers of the inner, and respectively outer loop. To update the PU VUL powers, the equivalence between $f_p(\mathbf{q}_2)$ and $f_d(\mathbf{q}_2)$, shown in Proposition 3.2, is used. Thus, the elements of $\mathbf{q}_2$ are updated based on the supragradient of $f_p(\mathbf{q}_2)$. To compute the supragradient, the powers and PU interference levels must be evaluated in the downlink domain, as shown in Step 3.1 and 3.2, respectively. Finally, the PU powers are updated with the classical supragradient rule, as shown in Step 3.3, where α_l denotes the stepsize, for the l-th PU constraint. The algorithm is repeated iteratively, until the convergence of the PU VUL powers, to a desired precision ϵ.

Remark 7 (regarding the PU updates) The general form of the VUL PU powers updates, if projective subgradient methods are employed is the one given by Eq. (3.55) in the Step 3.3 of Algorithm 3.2, where $[\cdot]_+$, represents the projection onto the non-negative orthant. It is easy to note that, choosing the step size as the PU VUL power obtained at the previous iterations, i.e., $\alpha_l = q_{2,l}(t_{out})$, the update rule is the same as in the algorithm in Section 3.4.1. Thus, the iterative algorithm in Section 3.4.1 can be regarded as an approximate subgradient method, which only takes one step in the direction of the optimal downlink powers and beamformers, for each $\mathbf{q}_2$, rather than computing them thoroughly.

Remark 8 (regarding complexity and inner loop) The inner loop follows the same principle as the algorithm proposed in [65], for conventional networks. The updates in the inner loop can also be done by means of fixed point iteration, thus with lower complexity but also with lower convergence rate. A comparison between the convergence rates of the two methods has been done in [65], where it was shown that the former algorithm achieves a superlinear convergence rate, whereas, the latter only a quadratic convergence rate. These results can be straightforwardly proven to hold for the inner loop.

Remark 9 (regarding differentiability) When iCSI is available for the SUs, $f_p(\mathbf{q}_2)$

Algorithm 3.2 A Subgradient Based Power and Beamformer Allocation

Step 1 Initialize $\{q_k\}_{k\in S}$ and $\{q_l\}_{l\in S_{PU}}$ and the corresponding beamforming matrix, such that we have an SU-feasible point

Until convergence of $\{q_l\}_{l\in S_{PU}}$:

Step 2 Solve $\mathcal{PU}_2(S, S_{PU})$ for fixed $\{q_l\}_{l\in S_{PU}}$ until convergence of $\{q_k\}_{k\in S}$, as follows

2.1 Update the SU VUL power with

$$\mathbf{q}_1(t_{in}+1) = \left(\mathbf{D}\left(t_{in}\right) - \mathbf{G}_1^T\left(t_{in}\right)\right)^{-1}\left(\mathbf{1}_{K\times 1} + \mathbf{G}_2^T\left(t_{in}\right)\mathbf{q}_2\left(t_{out}\right)\right) \tag{3.50}$$

2.2 Update the unit norm beamformers, for fixed powers

$$\mathbf{u}_k(t_{in}+1) = \text{argmin}_{\mathbf{u}_k} \frac{\mathbf{u}_k^H\left(\mathbf{I} + \sum_{\substack{i\neq k\\ i\in S}}\mathbf{R}_i\gamma_i q_i\left(t_{in}\right) + \sum_{l\in S_{PU}}\mathbf{R}_l\gamma_l q_l\left(t_{out}\right)\right)\mathbf{u}_k}{\mathbf{u}_k^H\mathbf{R}_k\mathbf{u}_k} \tag{3.51}$$

2.3 Reconstruct the matrices $\mathbf{D}(t_{in}+1),\ \mathbf{G}_1(t_{in}+1),\ \mathbf{G}_2(t_{in}+1)$ with Eq. (3.2)

2.4 Check convergence

$$\|\mathbf{q}_1\left(t_{in}+1\right) - \mathbf{q}_1\left(t_{in}\right)\| \leq \epsilon \tag{3.52}$$

Step 3 Update the PU VUL powers

3.1 Compute the downlink powers

$$\mathbf{p}_1\left(t_{out}+1\right) = \left(\mathbf{D}\left(t_{in}\right) - \mathbf{G}_1\left(t_{in}\right)\right)^{-1}\boldsymbol{\eta} \tag{3.53}$$

3.2 Compute the PU interference terms

$$\mathbf{p}_2(t_{out}+1) = \mathbf{G}_2(t_{in})\mathbf{p}_1(t_{out}+1), \tag{3.54}$$

3.3 Update the PU VUL powers

$$q_l(t_{out}+1) = \left[q_l(t_{out}) + \alpha_l\left(p_l(t_{out}+1) - 1\right)\right]_+;\ l\in S_{PU} \tag{3.55}$$

Step 4. Check convergence of the PU VUL powers

if $\|\mathbf{q}_2(t_{out}+1) - \mathbf{q}_2(t_{out})\| \geq \epsilon$, go to Step 2, **else exit**

Table 3.2: Subgradient based algorithm to solve the original problem (2.7) by means of $\mathcal{PU}_2(S, S_{PU})$

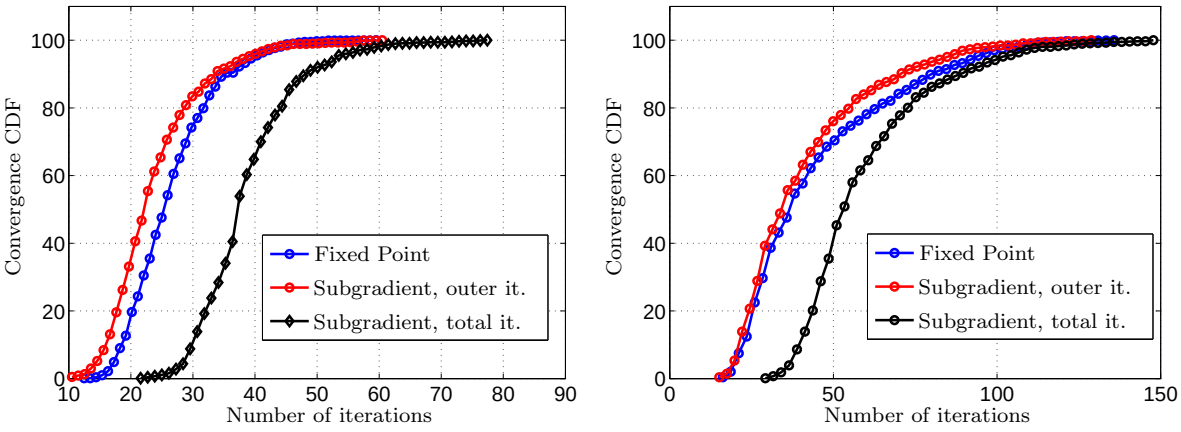
is differentiable and the algorithm in Table 3.2, is in fact a dual ascent method. Differentiability results from the fact that, in this case, the supragradient $\mathbf{p}_2 - \mathbf{1}$ is unique, as discussed in Section 3.2.1. Thus, since $f_p(\mathbf{q}_2)$ is concave, due to its structure as optimum objective of a piecewise minimization problem, it follows from [Proposition 6.1.1, [51]] that f_p is differentiable at any $\mathbf{q}_2$.

3.5 Simulation Results

In this section, we show the performance of the fixed point and subgradient algorithms, under gradually relaxed requirements in point of the CSI, available at the BS, i.e., from iCSI of all users, to iCSI of SUs and sCSI of PUs, and finally, to sCSI of all SUs and PUs. For comparison, we consider the benchmark SDR solution, which is implemented using the CVX sofware package, with the particular solver chosen as SeDuMi. We consider one BS with $N = 7$ antennas, 3 SUs and 3 PUs. The QoS thresholds are chosen randomly, at each run, from a uniform distribution in the interval $[-3, 8]$ dB. The interference threshold levels for the PUs, as well as the noise variance are set to -10 dBs. We compare the fixed point algorithm in Table 3.1 to the subgradient method in Table 3.2, where the stepsize is chosen equal to the VUL PU power at the current iterations, as discussed in Remark 7.

In the first scenario, the channels are generated as random complex Gaussian variables, to take Rayleigh fading effects into account, and we assume that perfect iCSI of both SUs and PUs is available at the BS. We show in Figures 3.4a and 3.4b, the cdf of the number of iterations, required to converge with a precision 10^{-5} and 10^{-6}, respectively. For the subgradient method, we consider noth the number of outer iterations, as well as the total number of iterations. The simulations are performed over 1000 independent Monte Carlo runs. As shown in Figure 3.4, the algorithms converge relatively fast, in this scenario. The number of outer iterations required by the subgradient method is similar to that of the the fixed point method, as expected. On the other hand, if all inner and outer iterations are taken into account, a clear difference can be seen between the two methods, even when the fastest procedure for the inner minimization in the subgradient method is used, i.e., the one using matrix inversion as shown in Table 3.2.

Furthermore, in Figure 3.5, we show the root-mean-square-error (RMSE) of the sum of



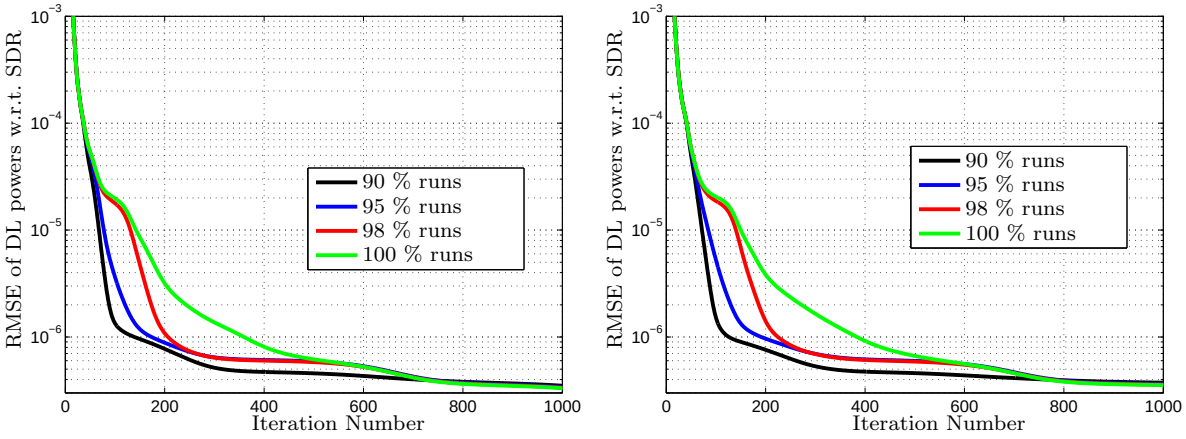
(a) precision 10^{-5} (b) precision 10^{-6}

Figure 3.4: CDF of the number of iterations in which the algorithms have converged, for two different choices of the precision factor ϵ

.

intermediary DL powers, obtained at each iteration, and the solution of the SDR reformulation. Specifically, this was computed as $\sqrt{1/N_M \sum_{i=1}^{N_M} \sum_{k \in S} \left(p_k(i) - p_k^{SDR}\right)^2}$, where N_M is the number of Monte Carlo runs, p_k^{SDR} is the optimal power of the kth user, as obtained by the SDR reformulation and i represents the iteration number. We consider the cdf as computed for all Monte Carlo runs, as well as for the best 90%, 95% and 98% of the cases. It can be seen that, whereas the RMSE decreases slowly with the iterations, when all feasible Monte Carlo runs are considered, significant improvement in the average convergence can be obtained if a small outage is acceptable.

We repeat the simulations for the cases of iCSI at the SUs and sCSI at the PUs, and show the results in Figures 3.6 and 3.7. In this scenario, the channel covariance matrices of both SUs and PUs are generated using the model in [80]. The positions of the users with respect to the antenna broadside are drawn at random from a uniform distribution. To model scattering effects, a spread angle of 3° is considered. The instantaneous channel realization at the k-th SU is further generated as $\mathbf{R}_k^{1/2}\mathbf{e}_k$, where $\mathbf{e}_k$ is a unit norm Gaussian variable, modelling Rayleigh fading effects. As shown by Figure 3.6, the number of iterations required to converge within a precision of 10^{-6} is significantly larger, compared to the



(a) RMSE for Fixed Point Method

(b) RMSE for Subgradient Method

Figure 3.5: RMSE of intermediary DL powers w.r.t. the optimal SDR value, for a scenario, in which, iCSI of both SUs and PUs is available

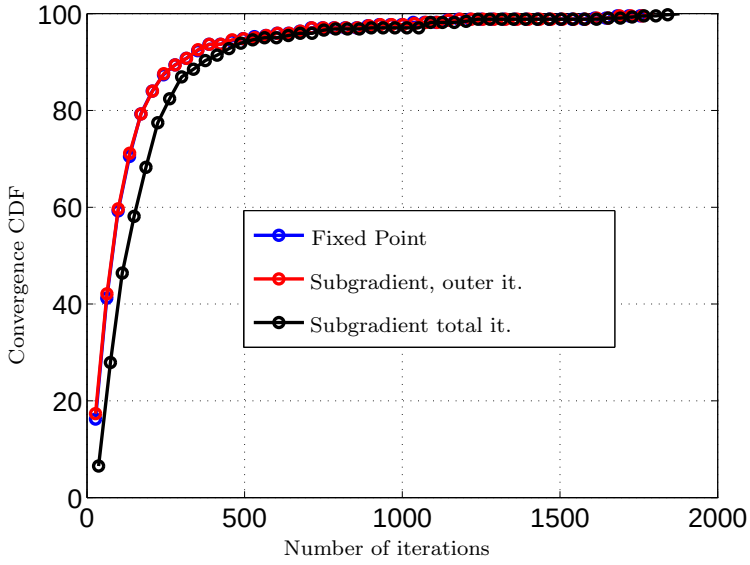
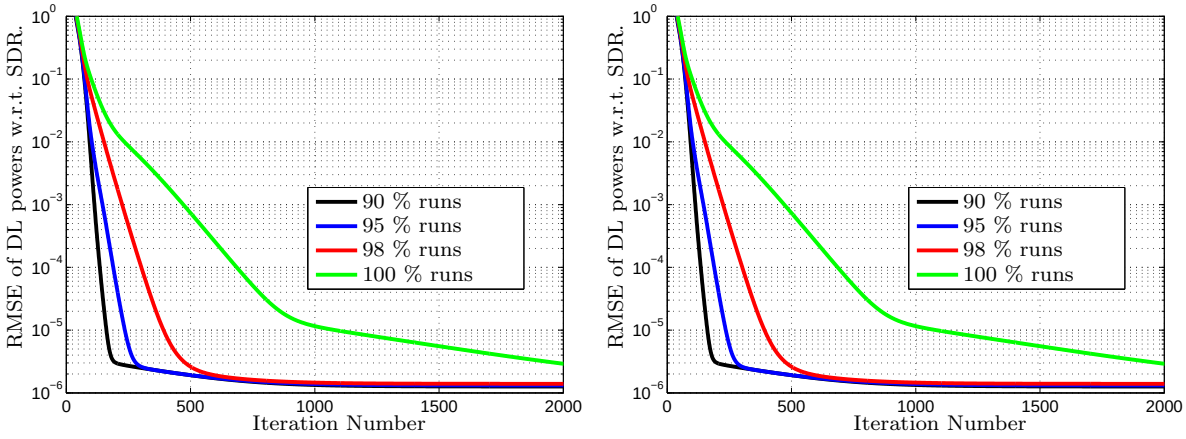


Figure 3.6: CDF of the number of iterations in which the algorithms have converged

previous case. On the other hand, from the RMSE graphic in Figure 3.7, it can be seen that the algorithms converge relatively fast up to a certain precision, e.g., of around 10^{-5}, in this scenario. However, from that point onwards, attaining higher precisions can only be achieved in a much larger number of iterations. We remark that the interior point algorithm,



(a) RMSE for Fixed Point Method

(b) RMSE for Subgradient Method

Figure 3.7: RMSE of intermediary DL powers w.r.t. the optimal SDR value, for a scenario, in which, iCSI of SUs and sCSI of PU is available

has comparatively a much faster covergence rate. On the other hand, the large difference between the complexity of the iterations, for the two algorithms, makes our method more efficient, for reasonable choices of the precision thresholds. Moreover, the updates of the beamformers, which represent the most computationally complex part of our algorithms, can be performed in parallel, which may further speed up the performance in practice.

Finally, we show in Figures 3.8 and 3.9, the results for a scenario in which the available CSI of both SUs and PUs is in the form of channel covariance matrices.

The RMSE for cases, in which convergence was achieved, is shown in Figure 3.9. The percentage of cases, in which convergence was not reached are 2-3%. Methods to force feasibility, e.g., using augumented Lagrangian or barrier functions can be developed, to address such cases. However, given the rare number of cases in which convergence problems occur, it is questionable whether the increase in complexity implied by treating these cases is indeed justifiable.

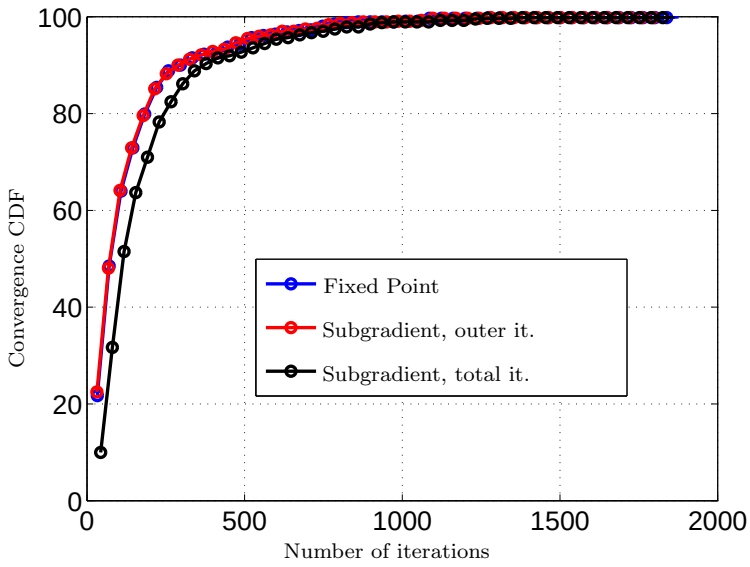
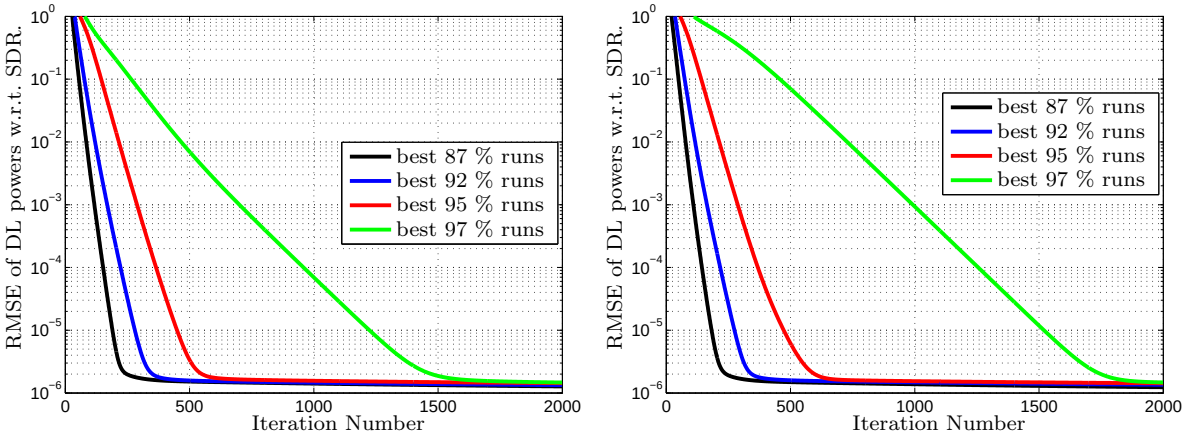


Figure 3.8: CDF of the number of iterations in which the algorithms have converged.



(a) RMSE for Fixed Point Method

(b) RMSE for Subgradient Method

Figure 3.9: RMSE of intermediary DL powers w.r.t. the optimal SDR value, for a scenario, in which sCSI of SUs and PUs is available.

3.6 Summary

In this chapter, we have exposed the VUL structure of downlink beamforming and power allocation problems, in CR scenarios. This resulted in the decoupling of the beamformers in

the SINR constraints, thus enabling efficient implementations, which achieve a significant reduction in complexity per iteration with respect to existing state-of the art techniques. Further, we have analysed the strong duality properties of the original DL problem in CR MISO scenario with arbitrary number of constraints. Even though a general duality statement for this is still an open problem, we have exposed several particular cases in which strong duality is indeed achieved. Furthermore, we have shown, through counterexamples, that some results from conventional networks, cannot be straightforwardly extended to CR scenarios, as might be presumed.

Chapter 4

Joint Downlink Beamforming and User Selection

4.1 Introduction

In the previous chapter, we have solved the beamforming and power allocation problem, under the assumption that a feasible solution exists. However, in practice, there may be no beamforming design, which can simultaneously satisfy all QoS and interference constraints imposed by the SUs and PUs, respectively, due to, e.g., excessively high SINR demands, stringent interference constraints or unfavourable channels. Therefore, in these cases, admission control at the base station becomes necessary. Selecting an appropriate set of users to be served is however a problem of combinatorial nature, and consequently, impossible to solve optimally, in polynomial time [83]. Therefore, two interesting objectives arise in this context. The first one, which has been approached in [83]-[86], concerns the design of low complexity suboptimal solutions with reasonable performance. The second one consists in devising convenient benchmark solutions, by customizing general purpose solvers with problem specific information, in order to reduce the search space, without compromising on optimality [88], [89]. Our focus, in this thesis, is on the first of type of these objectives. For this, a low complexity suboptimal scheme, to simultaneously select users to be served and their corresponding beamformers, has been proposed in [85]. The authors devised a

method, which consists in admitting SUs based on an orthogonality measure between their channel vectors and the channel vectors corresponding to the PUs. This makes this approach only applicable when perfect instantaneous CSI of both SUs and PUs is available at the BS. An alternative solution has been proposed in [83] for a setup with $L = 0$ and was later extended to a joint admission and power control scheme for cognitive radios in [84]. In [86], the more general admission control, with both power and beamformer design, was approached by extending the ideas in [83]. The methods in [83]-[86] consist of deflation and, respectively, branch and bound techniques, which, in each step, solve an SDR formulation, with a complexity similar to that of the problem in Eq. (2.11). This makes these techniques computationally demanding.

Our method, on the other hand, is designed based on the low complexity iterative scheme presented in Section 3. More precisely, we introduce a low complexity infeasibility detection technique, based on duality theory and the theorem of alternatives [49], which naturally integrates in the algorithm proposed in the previous chapter. With this, we further devise a deflation based user selection procedure, which successively removes SUs as long as infeasibility is detected. Heuristic measures, developed based on the intuition given by the VUL formulation, provide simple yet efficient measures to decide upon the users to be removed. Finally, for the feasible group of users obtained in this way, the algorithm provides the optimal beamformers and power allocation.

This method, which we show in detail in Section 4.2, performs well in small scenarios, however, it may become impractical when dense deployments are considered. Therefore, in Section 4.3, we consider a preprocessing phase in which a clustering of PUs and SUs, based on their long term spatial signatures, is performed. This information is further used in a computationally efficient admission and resource allocation scheme. Numerical results in Section 4.4. show the significant reduction in computational complexity achieved by this technique, when a large number of SUs and PUs are considered.

4.2 Proposed Approach for Small Scenarios

The problem we are interested in is to find the largest feasible set of users, for which there exists a beamforming and power allocation solution, satisfying all SINR and interference

constraints. More precisely, our aim is to find a set of SUs, S_o, as:

$$S_o = \text{argmax}_{S\subseteq\{1\dots K\}} \ |S| \ \text{ s.to } \ (2.7) \text{ is feasible for } \mathcal{S} \tag{4.1}$$

and to solve the resource allocation problem for the resulting set. In Eq. (4.1), we have denoted by $|\cdot|$ the cardinality of the set.

The approach, we take in this section, is based on the theorems of alternatives from duality theory [49]. The idea behind these theorems is as follows. In order to prove that a set of constraints is infeasible, it is sufficient to show that a consequence of simultaneously satisfying the constraints under consideration leads to a contradiction. This, in turn, can be shown by assessing whether a properly constructed 'complementary' set to the original constraint set is empty or not. If this set is empty, there exists a solution which simultaneously satisfies the original group of constraints. Duality theory provides elegant tools to easily construct such complementary sets and consequently expose the feasibility of a group of constraints.

We show next how we can use this principle to devise an infeasibility detection scheme. First, several remarks and definitions are necessary. Thus, we explicitly consider a total sum power constraint, in addition to the interference temperature constraints. This is practically motivated by regulatory requirements and represents a potential source of infeasibility. As discussed in Section 2.2, this can be incorporated as a $(K+L+1)$-th PU constraint. For ease of notation, we define the extended PU set as $\overline{S}_{PU} = S_{PU} \cup \{K+L+1\}$, to account for the additional constraint. Next, we define the primal feasibility set as

$$\mathcal{F} \triangleq \left\{ \{p_k, \mathbf{u}_k\}_{k\in S} \left| \text{SINR}_k^{\text{DL}} \geq \gamma_k;\ \text{I}_l^{DL} \leq \gamma_l^{-1};\ l \in \overline{S}_{PU};\ k \in S \right. \right\} \tag{4.2}$$

and the sets

$$\mathcal{D}_1 \triangleq \left\{ \left(\{q_k\}_{k\in S}, \{q_l\}_{l\in S_{PU}} \right) \left| -q_j\mathbf{R}_j + \sum_{\substack{k\neq j \\ k\in S}} q_k\gamma_k\mathbf{R}_k + \sum_{l\in\overline{S}_{PU}} q_l\gamma_l\mathbf{R}_l \succcurlyeq 0,\ j\in S \right. \right\} \tag{4.3}$$

$$\mathcal{D}_2 \triangleq \left\{ \left(\{q_k\}_{k\in S}, \{q_l\}_{l\in S_{PU}} \right) \left| \sum_{k\in S} q_k\gamma_k\sigma_k^2 - \sum_{l\in\overline{S}_{PU}} q_l > 0 \right. \right\}. \tag{4.4}$$

Further let $\mathcal{D} \triangleq \mathcal{D}_1 \cup \mathcal{D}_2$. Then, a feasibility certificate can be obtained using the theorem of alternatives [49], as follows.

Proposition 4.1 Assume the original problem (2.7) is strictly SU feasible. If there exist VUL powers such that $(q_1, \ldots, q_{K+L+1}) \in \mathcal{D}$, then no beamformer and power allocation can simultaneously satisfy the SINR and interference constraints in (2.7).

Proof The feasibility problem can be written as

$$\min_{\{\mathbf{w}_k\},s} \quad s \tag{4.5a}$$

$$\text{s.to} \quad \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k - \sum_{\substack{i \neq k \\ i \in S}} \mathbf{w}_i^H \mathbf{R}_k \mathbf{w}_i \gamma_k \geq \gamma_k \sigma_k^2; \ k \in S \tag{4.5b}$$

$$\sum_{k \in S} \mathbf{w}_k^H \mathbf{R}_l \mathbf{w}_k \gamma_l - s \leq 1/\gamma_l; \ l \in \overline{S}_{PU}. \tag{4.5c}$$

The dual problem corresponding to (4.5) can be easily derived as

$$g(\mathbf{q}_1, \mathbf{q}_2) = \max_{\mathbf{q}_1, \mathbf{q}_2} \quad \mathbf{q}_1^T \boldsymbol{\eta} - \mathbf{q}_2^T \mathbf{1} \tag{4.6a}$$

$$\text{s.to} \quad -q_j \mathbf{R}_j + \sum_{\substack{k \neq j \\ k \in S}} q_k \gamma_k \mathbf{R}_k + \sum_{l \in \overline{S}_{PU}} q_l \gamma_l \mathbf{R}_l \succcurlyeq 0 \tag{4.6b}$$

$$\mathbf{q}_2^T \mathbf{1} = 1. \tag{4.6c}$$

It can be seen that if a solution of the problem in (4.6) exists and is strictly larger than zero, then from weak duality, the optimum value of (4.5) is strictly positive, which implies that the original problem in (2.7) is infeasible. Thus if $\mathcal{D}$ is non-empty, the set $\mathcal{F}$ is empty. □

From Proposition 4.1, if the set $\mathcal{D}$ is non-empty, then the original problem is infeasible. However, this does not guarantee that, when the original problem is infeasible, $\mathcal{D}$ is not also empty. If both $\mathcal{F}$ and $\mathcal{D}$ are empty for some scenario, then infeasibility cannot be detected, and some different criteria must be found. However, we can show that under strict SU feasibility, this can only happen if the SDR reformulation (2.11) admits a solution, whereas the original problem (2.7) does not. Therefore, in scenarios in which iCSI or sCSI in Toeplitz form is available for the SUs, infeasibility can always be detected by using the result of Proposition 4.1.

To show that under strict SU feasibility, either $\mathcal{D}$ or the feasible set of the SDR (2.11) is empty, we can proceed as follows. The feasibility problem corresponding to the SDR

reformulation (2.11) can be written as

$$\min_{\{\mathbf{W}_k\},\, s} \quad s \tag{4.7a}$$

$$\text{s.to} \quad \mathrm{Tr}\left\{\mathbf{W}_k\mathbf{R}_k\right\} - \sum_{\substack{i\neq k \\ i\in S}} \mathrm{Tr}\left\{\gamma_{\mathrm{k}}\mathbf{W}_{\mathrm{i}}\mathbf{R}_{\mathrm{k}}\right\} \geq \gamma_{\mathrm{k}}\sigma_{\mathrm{k}}^2;\ \mathrm{k}\in\mathrm{S} \tag{4.7b}$$

$$\sum_{k\in S} \mathrm{Tr}\left\{\gamma_l\mathbf{W}_k\mathbf{R}_l\right\} - s \leq 1/\gamma_l;\ l\in\overline{S}_{PU}. \tag{4.7c}$$

It can be straightforwardly shown that the dual problem corresponding to (4.7) has the same expression as the dual of the original problem, (4.6). If the strict SU feasibility holds for the original problem (2.7), then a beamforming matrix exists, for which the inequalities in (4.7b) are strictly satisfied. Then, by choosing s such that $s > \max_{l\in\overline{S}_{PU}} \mathrm{Tr}\left\{\gamma_l\mathbf{W}_k\mathbf{R}_l\right\} - 1/\gamma_l$, we obtain a strictly feasible solution for the problem in (4.7). Therefore, Slater's condition is satisified [49], implying that strong duality holds for the feasibility problem (4.7), associated to the SDR formulation (2.11). Consequently, if there exists no feasible solution of the SDR problem (4.7), it follows that $s > 0$ and thus the set $\mathcal{D}$ is non-empty. Since the SDR in (2.11) is a relaxation of the original problem (2.7), its infeasibility implies that the latter problem is also infeasible. Thus, the sets $\mathcal{F}$ and $\mathcal{D}$ can only be simultaneously empty, if the SDR reformulation, (2.11), admits a solution, and the original problem does not.

4.2.1 Joint User Selection, Beamforming and Power Allocation Algorithm

According to the result in Proposition 4.1, if at an arbitrary step, the VUL powers are in the set $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$, then, the problem is infeasible. However, it can be easily shown that, the PU power updates based on complementary slackness, in Section 3.1, enforce the renewed VUL powers to be in the set $\mathcal{D}_2$. This holds for both the fixed point and subgradient based algorithm, presented in Section 3.

In the case of the subgradient based algorithm in Table 3.2, this can be shown as follows. Let us consider $\mathbf{q}_1(t_{out}+1)$ and $\{\mathbf{u}_k(t_{out}+1)\}$ as the optimum VUL power vector and beamformers respectively, obtained after the convergence of the inner loop for $q_2(t_{out})$.

We then have:

$$\boldsymbol{\eta}^T \mathbf{q}_1(t_{out}+1) = \boldsymbol{\eta}^T \left(\mathbf{D}\left(t_{out}+1\right) - \mathbf{G}_1\left(t_{out}+1\right)\right)^{-T} \left(\mathbf{1} + \mathbf{G}_2^T\left(t_{out}+1\right)\mathbf{q}_2\left(t_{out}\right)\right) \quad (4.8a)$$

$$= \mathbf{p}_1^T\left(t_{out}+1\right)\mathbf{1} + \mathbf{p}_2^T\left(t_{out}+1\right)\mathbf{q}_2(t_{out}) \quad (4.8b)$$

$$= \mathbf{p}_1^T\left(t_{out}+1\right)\mathbf{1} + \mathbf{q}_2^T\left(t_{out}+1\right)\mathbf{1} \quad (4.8c)$$

$$> \mathbf{q}_2^T\left(t_{out}+1\right)\mathbf{1}. \quad (4.8d)$$

Thus, $\{\mathbf{q}_1(t_{out}+1), \mathbf{q}_2(t_{out}+1)\}$ are in $\mathcal{D}_2$. A similar result holds for the fixed point algorithm in Table 3.1, if, instead of the updated SU VUL powers, the optima of the eigenvalue problems $\mathcal{E}_k\left(\mathbf{q}_1\left(t\right), \mathbf{q}_2\left(t\right), \mathbf{u}_k\left(t+1\right)\right)$ are employed, where t denotes the iteration number. This results as follows. Letting $\overline{q}_k \triangleq \mathcal{E}_k(\mathbf{q}_1(t), \mathbf{q}_2(t), \mathbf{u}_k(t+1))$, and stacking these elements into the vector $\overline{\mathbf{q}}_1$, we have

$$\boldsymbol{\eta}^T \overline{\mathbf{q}}_1 \geq \boldsymbol{\eta}^T \left(\mathbf{D}\left(t+1\right) - \mathbf{G}_1\left(t+1\right)\right)^{-T} \left(\mathbf{1} + \mathbf{G}_2^T\left(t+1\right)\mathbf{q}_2\left(t\right)\right). \quad (4.9)$$

The inequality in (4.9) holds due to VUL feasibility, as discussed in the third remark of Section 3.2. From (4.9), similar inequalities hold as in (4.8b)-(4.8d), thus we can conclude that $\boldsymbol{\eta}^T\overline{\mathbf{q}}_1 - \mathbf{q}_2(t+1)^T\mathbf{1} \geq 0$. Note that optimal values of $\mathcal{E}_k\left(\mathbf{q}_1\left(t\right), \mathbf{q}_2\left(t\right), \mathbf{u}_k\left(t+1\right)\right)$ are already available after solving the eigenvalue problems in Step 2, and are therefore obtained at virtually no cost. Furthermore, the result of the k-th eigenvalue problem can be viewed as an intermediary VUL power of the k-th SU.

In conclusion, to test infeasibility, it is sufficient to test if

$$\left(\left\{q_k\left(t_{out}+1\right)\right\}_{k\in S}, \left\{q_l\left(t_{out}+1\right)\right\}_{l\in S_{PU}}\right) \in \mathcal{D}_1, \quad (4.10)$$

in the case of the subgradient based algorithm in Table 3.2, and similarly that

$$\left(\left\{\mathcal{E}_k\left(\mathbf{q}_1\left(t\right), \mathbf{q}_2\left(t\right), \mathbf{u}_k\left(t+1\right)\right)\right\}_{k\in S}, \left\{q_l\left(t+1\right)\right\}_{l\in S_{PU}}\right) \in \mathcal{D}_1, \quad (4.11)$$

in the case of the fixed point algorithm in Table 3.1. If infeasibility is detected, an SU is removed, based on heuristics, which we derive in Section 4.2.2. Moreover, the feasibility test procedure can be terminated once a point in the set $\mathcal{F}$ has been found. This check comes at virtually no cost, as it only involves the verification of whether the PU interference constraints after the updates are respected. For completeness, we show in Table 4.1 the beamforming and admission control procedure, based on the subgradient method in Table 3.2. Naturally, the procedure can be adapted for the fixed point algorithm, where the infeasibility test is based on Eq. (4.11).

Algorithm 4.1 Heuristic Deflation Algorithm

Step 1. Initialize $\{q_k\}_{k\in S}$ and $\{q_l\}_{l\in S_{PU}}$ and the corresponding beamforming matrix such that we have an SU-feasible point

Step 2. Until convergence of $\{q_l\}_{l\in S_{PU}}$, iterate the following steps:

2.1 Update the SU uplink powers for fixed beamformers Eq. (3.50)

2.2 Update the unit norm beamformers for fixed powers Eq. (3.51)

Step 3. Update the PU virtual uplink powers with Eq. (3.54)

Step 4. Perform the infeasibility test: $[q_1, \ldots, q_{K+L+1}]\,(t_{out}+1) \in \mathcal{D}_1$.

If infeasible, remove user based on an appropriate heuristic, as described in Section 4.2.2 and go to Step 2.

Step 5. Perform convergence check

Table 4.1: Joint deflation based user selection and beamforming control, based on subgradient method.

4.2.2 Heuristic Selection

The proposed procedure assumes as initialization the largest SU feasible set. Then, the iterative beamforming algorithm with feasibility control is started and, whenever infeasibility is detected, one user is removed at a time, based on appropriately chosen heuristics. The structure of the problem and its interpretation in the VUL domain, makes it possible to choose powerful heuristics. Recalling that infeasibility is detected when the VUL powers are in the set $\mathcal{D}_1$, it is reasonable to consider the removal of the user, whose elimination 'moves' the remaining set of VUL powers the furthest away from $\mathcal{D}_1$. A potential candidate, to satisfy this, is the one with the largest weighted VUL power, i.e., $\gamma_k q_k$. The advantages of using this heuristic, in a deflation based scheme, lie in its simplicity and generally good performance, as shown in simulations. We refer to this scheme as 'Fast Removal', through the rest of the chapter.

This scheme performs well when there exists an SU with a significantly larger weighted VUL power. There may however occur cases when several SUs have similar such coefficients,

Algorithm 4.2 Look Ahead

Step 1. **for** j=1,2

1.1 Solve $\mathcal{PU}(S^j(t), S_{PU})$ until $\{q_k\}_{k\in S^j(t)\cup\overline{S}_{PU}}$ converges or infeasibility is detected.

1.2 With the obtained $\{q_k\}_{k\in S^j(t)\cup\overline{S}_{PU}}$ compute

$$m_j = \begin{cases} \sum_{k\in S^j(t)} \gamma_k\sigma_k^2 q_k - \sum_{l=1}^{L+1} q_{K+l} & \text{if } \mathcal{PU}(S^j(t),S_{PU}) \text{ feasible;} \\ |T| \text{ where } T=\{k\in S^j(t) \mid \gamma_k q_k > \gamma_j q_j\} & \text{otherwise.} \end{cases}$$

Return $(m_j, \text{infeasibility status})$

Step 2. **if** both $\mathcal{PU}\left(S^j(t), S_{PU}\right),\ j=1,2$ have the same infeasibility status eliminate user j with the smallest m_j

else eliminate user j for which $\mathcal{PU}(S^j(t), S_{PU})$ is feasible

Table 4.2: User removal based on depth one branch search. $\mathcal{PU}$ denotes generically the VUL problem and can be either $\mathcal{PU}_1(S, S_{PU})$ in (3.1) or $\mathcal{PU}_2(S, S_{PU})$ in (3.13)

e.g., when two SUs interfere with each other or create interference to the same PU. To overcome this problem we consider an alternative user selection scheme, in which a depth first tree search is performed. More specifically, when there is no user with a distinctively high weighted VUL power, we consider the two users, which attain the highest such values, as candidates for removal. Then, temporarily eliminating each one, we perform the iterative update and assess the evolution of the two candidates among the remaining users. Finally, we remove the SU, which at the next branch level has the poorest performance with respect to a chosen metric, e.g., weighted uplink power or interference created to the PUs. This removal approach, which we term 'Look-Ahead' is shown in Table 4.2, where for simplicity we consider only two SUs in the candidate set. For convenience, we use the notation $S^j(t)$ to represent the set of users S, considered at iteration t, from each user j has been removed, i.e., $S^j = S(t) \setminus \{j\}$.

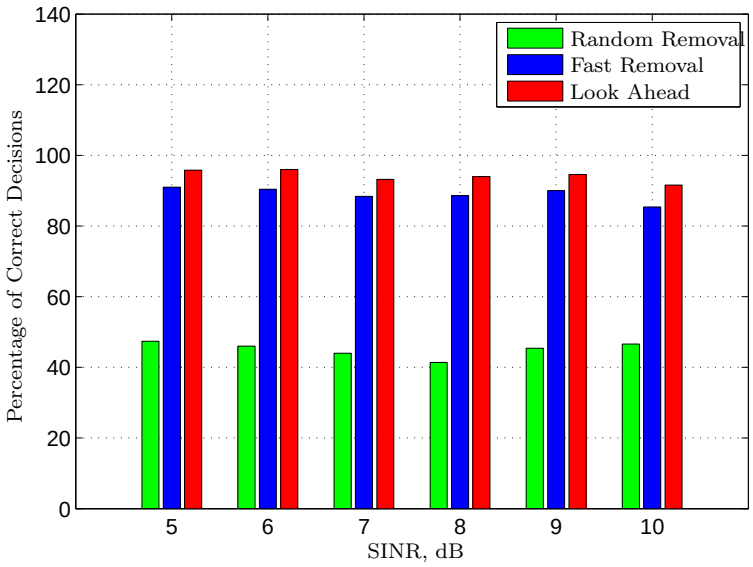


Figure 4.1: Percentage of correct decisions, in point of served number of users when the number of PUs is fixed to 6 and SINR target is increasing

4.2.3 Simulation Results

In this subsection, we evaluate by numerical simulations the performance of the algorithms in interference limited scenarios, in which infeasibility is due to a large number of coexisting PUs and increasing QoS demands, which must be respected without violating the imposed interference temperature constraints. We evaluate the performance in terms of served number of users and power consumption, as well as the number of iterations required to decide infeasibility. For comparison we consider the optimal solution, which can be obtained using a full search method and a solution, given by randomly removing one user at each step, for which infeasibility was detected. The simulation scenarios consist of one BS with $N = 7$ antennas and 6 single antenna SUs. The channels are assumed to be affected by Rayleigh fading. The interference thresholds for the PUs, as well as the noise levels at the SUs are set to -10dBs.

In the first scenario we consider that infeasibility caused by increasing QoS demands and show in Figure 4.1 the percentage of correct decisions in point of served number of users, when the number of PUs is 6. Specifically, we show the percentage of Monte Carlo runs, in which the number of served users, using the three deflation methods, i.e., our two proposed

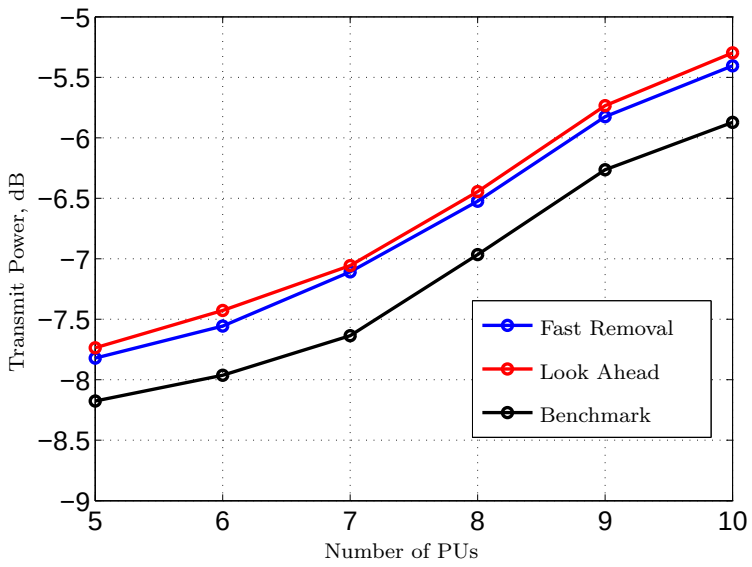


Figure 4.2: Average transmit power for the feasible user combinations, chosen by the three schemes

heuristic methods and the random removal scheme, coincide with the optimal number of users, given by the benchmark solution. We notice that the proposed low complexity algorithms serve the correct number of users in more than 85 % of the cases. As expected the 'Look-Ahead' algorithm which performs the tree search has a better performance, taking the correct decision in more than 90% of the cases. The performance of both proposed algorithms is significantly better than a random user selection scheme, thus certifying that the VUL powers at the iteration where infeasibility is detected, as well as our proposed heuristic measures, are indeed meaningful. The average transmitted power of the user combinations, chosen by our algorithms compared to the user combination, which achieves the optimal power, is depicted in Figure 4.2. For a fair comparison, only the Monte Carlo runs were considered, in which all three techniques served the same number of users. We note that even though our low complexity algorithms do not always choose the user combination with the smallest transmit power, the difference to the optimal scheme is below 0.5dB on average, which is acceptable once the significant difference in complexity between our methods and the exhaustive search is taken into account.

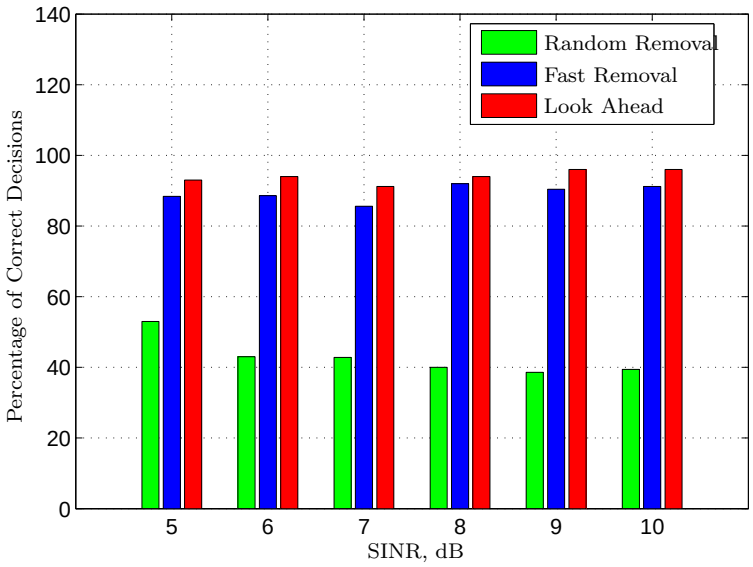


Figure 4.3: Percentage of correct decisions in point of served number of users when the SINR target is set to 8dB

In the second scenario, we evaluate the performance of our algorithms, when the infeasibility is due to the interference limitations, imposed by an increasing number of PUs. We show, in Figure 4.3, the percentage of the correct decisions taken by our algorithms and the random removal algorithm in point of served number of users. Similar to the previous case, the deflation procedures perform well, with the simple one step removal reaching the correct decisions in 85 to 90% of the cases, while being outperformed by the Look Ahead method. The random removal method only reaches the correct decisions in less than 40% of the cases. Finally, we show in Figure 4.4, the cdf of the number of iterations required to reach an infeasibility decision. More precisely, the number of outer iterations for the subgradient algorithm is considered in this figure, and is similar to the number of iterations for the fixed point algorithm. We note that, generally, the decision is taken relatively fast for these scenarios, with more than 90% of cases exhibiting infeasibility in less than 10 iterations.

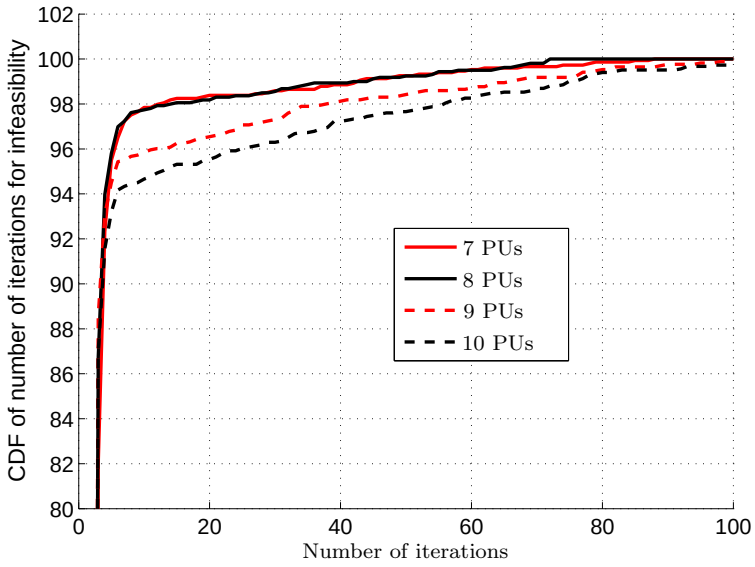


Figure 4.4: The cdf of number of iterations to decide infeasibility, for an SINR target of 8dBs and increasing number of PUs

4.3 Proposed Approach for Dense Networks

In this section we consider the case where the number of SUs competing for resources, as well as the number of PUs which require interference protection is large. Applying the deflation techniques from the previous section to these scenarios is clearly impractical. To reduce the complexity of the user selection scheme for such cases, we propose a preprocessing phase in which users are clustered based on the similarity between their long term spatial signatures. More precisely our motivation is as follows. Consider a scenario as the one depicted in Figure 4.5. Here, SUs with similar spatial signatures are prone to creating large levels of mutual interference, therefore are unlikely to be served jointly. Furthermore, PUs with similar spatial signatures pose similar constraints to the secondary system and, therefore, the use of PU clusters promises a more efficient interference management.

Clustering algorithms have been extensively studied in unsupervised machine learning and a significant number of algorithms has been developed to solve this problem on vector spaces. For our purposes, however, we are interested in grouping the users based on the second order statistics of their channels. In this case, the vector space prerequisite is

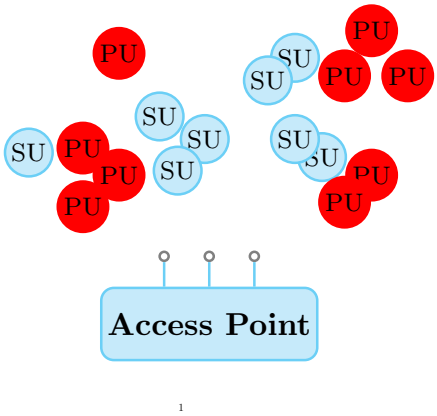


Figure 4.5: Clustering in dense CR networks

not met, due to the fact that the covariance matrices are positive semidefinite. Several techniques have been developed to cluster points based on covariance matrices, however these generally employ an exponential distance measure, which cannot be used when the covariance matrices are rank deficient. Therefore, in this thesis, we employ the distance measure derived in [134] and, based on this, we develop two clustering schemes. From each of these clusters, appropriate representatives can be chosen and fed to the uplink downlink algorithm, considered in Section 4.2.2. The use of long term channel statistics, which are known to vary slowly in time, makes the scheme practically attractive as the clustering schemes can be performed offline and do not require a frequent updates.

4.3.1 Covariance Based User Clustering

Let $\mathcal{G}$ denote a set of points, which, in our case, are psd channel covariance matrices. The purpose of a clustering scheme is to find non-overlapping sets $\{\mathcal{C}_i\}$ and corresponding cluster

centers $\overline{\mathbf{R}}_i$ such that

$$\min_{N_C,\{\mathcal{C}_i,\overline{\mathbf{R}}_i\}} \sum_i \sum_{j=1}^{|\mathcal{C}_i|} d^2\left(\mathbf{R}_{\mathcal{C}_i(j)}, \overline{\mathbf{R}}_i\right) \tag{4.12a}$$

$$\text{s.to} \qquad \bigcup_i \mathcal{C}_i = \mathcal{G}; \quad \bigcap_i \mathcal{C}_i = \emptyset, \tag{4.12b}$$

where $\mathcal{C}_i(j)$ represents the jth element in the ith cluster, N_C is the number of clusters and $d(\cdot,\cdot)$ is a similarity measure between two psd matrices which are its arguments.

To find appropriate clusters, the first important requirement, as pointed out by the problem formulated in (4.12), is to have a proper distance function to measure the similarity between channel covariance matrices. It has been acknowledged that Euclidean distances are generally not well suited to assess the proximity between covariance matrices. As shown in Appendix B, the set of covariance matrices can, however, be endowed with a Riemannian manifold structure, for which appropriate distance measures can be defined, to measure the similarity between points, i.e., psd matrices, while taking into account their inherent positive semidefiniteness properties. Thus, to measure similarity between the matrices on the Riemannian manifold, we use the distance derived in [134] as

$$d_R(\mathbf{R}_1, \mathbf{R}_2) = \sqrt{\mathrm{Tr}\left\{\left(\mathbf{R}_1^{1/2} - \mathbf{R}_2^{1/2}\right)^2\right\}}, \tag{4.13}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are two arbitrary psd matrices.

To group users around centers, which are not known apriori, it is necessary to establish the concept of the mean of points on the manifold. A definition of the mean on general Riemannian manifolds has been proposed by Karcher as a natural generalization of the arithmetic mean on Euclidean spaces. More precisely, the mean of P points on the manifold is defined as the point which minimizes the sum of square distances to all P points, i.e., psd matrices, and can be written as

$$\overline{\mathbf{R}} \triangleq \mathrm{argmin}_{\mathbf{R}\succ 0} \sum_{\mathrm{i}=1}^{\mathrm{P}} \mathrm{d}_{\mathrm{R}}^2(\mathbf{R}, \mathbf{R}_{\mathrm{i}}). \tag{4.14}$$

Inserting Eq. (4.13) in the definition (4.14), we obtain that the Karcher mean for the Riemannian manifold of psd matrices characterised with the distance measure (4.13), must satisfy

$$\min_{\mathbf{R}} P\mathrm{Tr}\left\{\mathbf{R}\right\} - 2\sum_{\mathrm{i}=1}^{\mathrm{P}} \mathrm{Tr}\left\{\mathbf{R}_{\mathrm{i}}^{1/2}\mathbf{R}^{1/2}\right\} + \sum_{\mathrm{i}=1}^{\mathrm{P}} \mathrm{Tr}\left\{\mathbf{R}_{\mathrm{i}}\right\}. \tag{4.15}$$

Naturally, a minimizer of (4.15) is

$$\overline{\mathbf{R}} = \left[\frac{1}{P} \sum_{i=1}^{P} \mathbf{R}_i^{1/2} \right]^2 . \tag{4.16}$$

Finally, one more aspect, which must be considered when approaching the problem in (4.12), is the knowledge of the number of clusters N_C. Many familiar clustering algorithms in literature, e.g., K-means, are designed under the assumption that N_C is known a-priori. In the case of SU clustering, such an assumption can be justified as, generally, the number of users for which a feasible beamforming design under SINR constraints exists, is limited by the number of antennas available at the BS. An upper bound for the number of SUs, which can be satisfied simultaneously, may also be derived in closed form as in [67]. Thus, for the case when the number of SUs is known, we can adapt a K-means algorithm to the distance measure in (4.13) and the corresponding center, defined by Eq. (4.16), as we show next.

SU Clustering

Choosing N_{SUC} as an upper bound for the number of users which can be simultaneously served, we perform a K-means clustering for the SUs. This algorithm is summarized in Table 4.3, where we denote the i_C-th SU cluster by $\mathcal{C}_{SU,i_c}$. The idea of the approach, is to iteratively update the members of the clusters, by evaluating the distance between each user and the estimated cluster center, at each iteration. Each user is then included in the cluster, whose center is the closest with respect to the chosen distance, (4.13), in Step 2. Finally, in Step 3, the mean of the newly formed groups are recomputed. The procedure is repeated iteratively, until the centers of the clusters stabilize.

PU Clustering

Contrary to the SU case, finding a number of PU clusters is more demanding, as in this case no condition can be found, e.g., w.r.t. number of antennas. Furthermore, performing the grouping for an erroneous number of clusters can lead to artificial results. Thus, to avoid these drawbacks, and an additional stage of estimating the number of clusters, we propose here a greedy scheme to group the PUs. Before we proceed, we introduce several definitions.

Algorithm 4.3 SU Clustering

Step 1. Initialize the cluster centers with N_{SUC} SUs and the center with the corresponding covariance matrix

$$\mathcal{C}_{SU,i_C} \leftarrow \pi(i_C); \quad i_C = 1, \dots N_{SUC} \tag{4.17}$$

$$\overline{R}_{SU,i_C}(1) \leftarrow \mathbf{R}_{\pi(i_C)}, \tag{4.18}$$

where $\pi(i_C)$ is the index of the SU which has been selected for cluster i_C

Repeat until convergence:

Step 2. For each $k = 1, \dots, K$, assign kth SU to a cluster i_C such that:

$$i_C = \mathrm{argmin}_{\mathrm{i=1,\dots,N_{SUC}}} \mathrm{d}^2_{\mathrm{R}}(\overline{\mathbf{R}}_{\mathrm{SU,i}}(\mathrm{t}), \mathbf{R}_{\mathrm{k}}) \tag{4.19}$$

$$\mathcal{C}_{SU,i_C} \leftarrow \mathcal{C}_{SU,i_C} \cup \{k\}\,. \tag{4.20}$$

Step 3. For $i_C = 1, \dots, N_{SUC}$, update the mean of the cluster i_C with Eq. (4.16)

$$\overline{R}_{SU,i_C}(t+1) = \left[\frac{1}{|\mathcal{C}_{SU,i_C}|} \sum_{k \in \mathcal{C}_{SU,i_C}} \mathbf{R}_k^{1/2} \right]^2. \tag{4.21}$$

Table 4.3: Clustering algorithm for known number of clusters.

Let D_{PU} be a set containing the distances between any two PU channels covariance matrices which are smaller than a threshold Υ as:

$$D_{PU} = \{d_R(\mathbf{R}_i, \mathbf{R}_j) < \Upsilon; \quad i, j = K{+}1, \dots, K{+}L\} \tag{4.24}$$

and assume the elements are sorted in ascending order. Further, consider the index set defined as

$$\mathcal{I}_{PU} \triangleq \{(i,j) | i < j, i, j = K+1, \dots, K+L\} \tag{4.25}$$

and a mapping $\pi : \{1, \dots, |D_{PU}|\} \rightarrow \mathcal{I}_{PU}$ such that $\pi(k)$ retrieves the indexes (i, j) of the k-th smallest distance in D_{PU}.

Algorithm 4.4 PU Clustering

Step 1. Initialisation: Form the set D_{PU} as in Eq. (4.24). Construct the first cluster and compute its mean:

$$(i,j) = \pi(1),$$
$$N_{PUC} \leftarrow 1,$$
$$\mathcal{C}_{PU,1} \leftarrow \{i\} \cup \{j\},$$
$$\overline{\mathbf{R}}_{PU,1} = \frac{1}{4}(\mathbf{R}_i^{1/2} + \mathbf{R}_j^{1/2})^2.$$

Step 2. **for** $t = 2, \ldots, |D_{PU}|$

$$(i,j) \leftarrow \pi(k)$$
$$i_{C,i} = \left[\mathrm{argmin}_{\mathrm{i_C}=1,\ldots,\mathrm{N_{PUC}}} \mathrm{d_R^2}(\mathbf{R}_\mathrm{i}, \overline{\mathbf{R}}_{\mathrm{PU,i_C}}) \leq \Upsilon/2\right]_+ \quad (4.22)$$
$$i_{C,j} = \left[\mathrm{argmin}_{\mathrm{i_C}=1,\ldots,\mathrm{N_{PUC}}} \mathrm{d_R^2}(\mathbf{R}_\mathrm{j}, \overline{\mathbf{R}}_{\mathrm{PU,i_C}}) \leq \Upsilon/2\right]_+ \quad (4.23)$$

if $i_{C,i} = i_{C,j} = 0$, create new cluster

$$N_{PUC} = N_{PUC} + 1$$
$$\mathcal{C}_{PU,N_{PUC}} \leftarrow \{i\} \cup \{j\}$$

else assign to the closest clusters

$$\mathcal{C}_{PU,i_{C,i}} \leftarrow \mathcal{C}_{PU,i_{C,i}} \cup \{i\}$$
$$\mathcal{C}_{PU,i_{C,j}} \leftarrow \mathcal{C}_{PU,i_{C,j}} \cup \{j\}$$

if all PUs assigned to a cluster, **exit**

Table 4.4: Clustering for unknown number of clusters

The idea of the algorithm, summarized in Table 4.4, is to first create one cluster with the two PUs, whose distance is the smallest, and compute its center, as shown in Step 1. The procedure is then successively performed, in ascending order of the distances between user pairs. At each iteration t, the similarity between the PUs, corresponding to the t-th smallest distance in $\mathcal{D}_{PU}$, and the centers of the existing clusters is evaluated. The PUs, corresponding to the t-th pair, which do not already belong to a cluster, are assigned to the group, whose center is the closest, in Riemannian distance. At the same time, the distance towards the closest center must not exceed Υ. This is evaluated by Eqs. (4.22) and (4.23), where the notation

$$[\text{argmin f} \leq \alpha]_+ = \begin{cases} f^\star, \text{if the minimum f}^\star \text{ satisfies } f^\star \leq \alpha \\ 0, \quad \text{otherwise.} \end{cases}$$

is used. If the smallest distance exceeds the largest threshold Υ, the users are not assigned to an existing cluster, but a new cluster containing these users is formed.

4.3.2 Proposed Algorithm for Cluster Aided User Selection and Optimal Beamforming

Our proposed algorithm for dense networks is depicted in Table 4.5. As mentioned in the previous section, the first stage consists of a preprocessing phase, in which PUs and SUs are clustered and which may be performed offline due to the slow varying nature of the second order channel statistics. If latency is a concern, the two clustering schemes may be performed in parallel. Otherwise, information about the PU clusters can be utilized, when deciding upon the initial centers of the SU clusters. In this manner, the SU clustering scheme can be designed to put more focus on the protection of the primary network from interference.

Once the clusters are formed, we determine one representative of each cluster and form the sets $\mathcal{R}_{SU}$ and $\mathcal{R}_{PU}$ of SU and PU representatives, respectively. The sets $\mathcal{R}_{SU}$ and $\mathcal{R}_{PU}$ are then used to solve the initial joint beamforming and user selection problem $\mathcal{PD}(\mathcal{R}_{SU}, \mathcal{R}_{PU})$ through the equivalent uplink formulation $\mathcal{PU}(\mathcal{R}_{SU}, \mathcal{R}_{PU})$. In Table 4.5, the subgradient method from Table 3.2 is used to achieve this. Naturally, the fixed point method can be alternatively employed, with the difference in the infeasibility test, as

Algorithm 4.5 Joint User Clustering, Beamforming and Power Allocation

Offline Preprocessing Stage:

Step 1. Perform PU clustering as in Algorithm 4.4.

Step 2. Perform SU clustering as in Algorithm 4.3.

Step 3. Find set of representatives, $\mathcal{R}_{PU}$, for each PU cluster

Online Stage:

Step 1. Find set of representatives, $\mathcal{R}_{SU}$, for each SU cluster and solve (3.42). If not strictly SU feasibile, remove the SU with the largest virtual power, update $\mathcal{R}_{SU}$ and repeat Step 1.

Until convergence of the VUL powers $\{q_l\}_{l\in\mathcal{R}_{PU}}$:

Step 2. Solve $\mathcal{PU}\left(\mathcal{R}_{SU},\mathcal{R}_{PU}\right)$ for fixed $\{q_l\}_{l\in\mathcal{R}_{PU}}$ by iterating the steps:

2.1 Update the SU uplink powers $\{q_k\}_{k\in\mathcal{R}_{SU}}$ with the solution of the linear system of equations in (3.13a) for fixed beamformers.

2.2 Update the beamformers $\{\mathbf{u}_k\}$ as generalized eigenvectors of the expressions defined in (3.13a) for fixed q_k.

Step 3. Update the PU VUL powers

3.1 Compute the DL powers from Eq. (2.7b).

3.2 Compute the PU interference levels $I_l^{\mathrm{DL}}(t+1)$ from Eq. (2.7c), for the fixed beamformers.

3.3 Perform the PU VUL power updates with

$$q_l(t+1) = q_l(t)\gamma_l I_l^{\mathrm{DL}}(t+1),\ l\in\mathcal{R}_{\mathcal{PU}}. \tag{4.26}$$

Step 4. **if** $\{q_k\}\in\mathcal{D}_1$ remove the SU with the largest weighted VUL power and go to Step 2.

Table 4.5: Clustering aided user selection, with a subgradient based algorithm.

mentioned in Section 4.1.1. We next describe the procedures for choosing the PU and SU representatives.

Selecting PU representatives

In order to form the set $\mathcal{R}_{PU}$ of the PU representatives in Step 3 of the proposed algorithm in Table 4.5, we approach the problem as follows. Instead of directly selecting one of the PUs in the clusters, we construct 'virtual' users, for which the covariance matrices satisfy all the interference constraints, required by the PUs in the considered cluster. More precisely, a virtual user for cluster i_C is chosen, such that the satisfaction of the interference temperature constraints, w.r.t. its channel covariance matrix, $\hat{\mathbf{R}}_{i_C}$, implies that all interference thresholds of the PUs in the corresponding cluster are respected. Specifically, denoting by $\hat{\mathrm{I}}_{i_C}^{\mathrm{DL}}$, the interference experienced by an user with channel covariance matrix $\hat{\mathbf{R}}_{i_C}$, then the condition

$$\hat{\mathrm{I}}_{i_C}^{\mathrm{DL}} = \sum_{k \in S} p_k \mathbf{u}_k^H \hat{\mathbf{R}}_{i_C} \mathbf{u}_k \leq \frac{1}{\gamma_{i_C}} \tag{4.27}$$

implies that $\sum_{k \in \mathcal{R}_{SU}} p_k \mathbf{u}_k^H \hat{\mathbf{R}}_l \mathbf{u}_k \leq \gamma_l^{-1}$ for all $l \in \mathcal{C}_{PU,i_C}$. In Eq. (4.27), $\gamma_{i_C}^{-1}$ represents the interference threshold. For ease of exposition, we consider that all users in a cluster have the same interference thresholds i.e., $\gamma_{i_C} = \gamma_l, l \in \mathcal{C}_{PU,i_C}$. The case, in which the interference thresholds are different, can be easily treated by appropriately scaling the channel covariance matrices in (4.27). A virtual user, which satisfies (4.27), can be constructed as a solution to the following SDP problem

$$\begin{aligned} \hat{\mathbf{R}}_{i_C} = \mathrm{argmin}_{\mathbf{R} \succeq 0} & \mathrm{Tr}\{\mathbf{R}\} \\ \text{s.to} \quad & \mathbf{R} - \hat{\mathbf{R}}_i \succeq 0, \ i \in \mathcal{C}_{PU,i_C}. \end{aligned} \tag{4.28}$$

Selecting SU Representatives

If instantaneous CSI is available, a possible choice for cluster i_C is the SU with the strongest channel i.e., the SU such that

$$\mathrm{argmax}_{i \in \mathcal{C}_{SU,i_C}} \|\mathbf{h}_i\|^2. \tag{4.29}$$

In this way, multiuser diversity is exploited, by considering the fading nature of the wireless channels. Naturally, such a scheme must be performed online to take into account the constant channel feedbacks. This is the rule, which we considered in our simulations. However,

it is expected that, in practical scenarios, the decision, regarding the SU to be removed, will be taken by, or together with the scheduler on upper layers, in order to ensure, e.g., fairness and service quality aspects.

4.3.3 Simulation Results

We validate the performance of our algorithms in a scenario consisting of one BS with 8 antennas, 20 SUs and a number of PUs increasing from 20 to 40. The minimum SINR requirements for the SUs are set to 2dB, whereas, the interference thresholds imposed by the primary network are -10dB. The covariance matrices are generated using a ring scattering model [80], with 30 random sets of angles. We consider iCSI at the transmitter and generate the instantaneous channels from the covariance matrices as $\mathbf{h}_k = \mathbf{R}_k^{1/2}\mathbf{e}_k$. In this expression, $\mathbf{e}_k$ is a Gaussian random vector with zero mean and unitary variance, which models the Rayleigh fading. For each set of angles, we generate 300 instantaneous channels for the Monte Carlo simulations.

We compare the performances of the algorithms with the proposed clustering schemes to the deflation scheme in Section 4.2. Moreover, we consider the SU and PU clustering schemes alone and in combination, in order to show the effect of each of the procedures. For the SU clustering scheme, we consider two variants, which differ on the initialization point. Thus, the first, which we term 'SU-friendly', performs several rounds of Algorithm 4.3 with random initialization points, and chooses the one which achieves the smallest objective, as given by Eq. (4.12). The second scheme, which we term 'PU-friendly', assumes an initialization which is the farthest to the PU cluster centers, in terms of Riemannian distance. The number of SU clusters in the simulations was set to 9.

We show in Figure 4.6 the total number of inner iterations required to find a feasible user combination and its corresponding optimal beamforming and power allocation. We notice that, performing clustering only on the SUs, results in a decrease of approximately 20% in the number of iterations. Further, applying this technique on both the set of SUs and PUs leads to a significant reduction of 75%, with respect to the scheme without preprocessing. This effect is even more prominent if runtimes are compared, as shown in Figure 4.7. The reason for this is that the clustering leads to a reduction in the size of the problem.

Regarding the number of served users for all techniques, we notice from Figure 4.8,

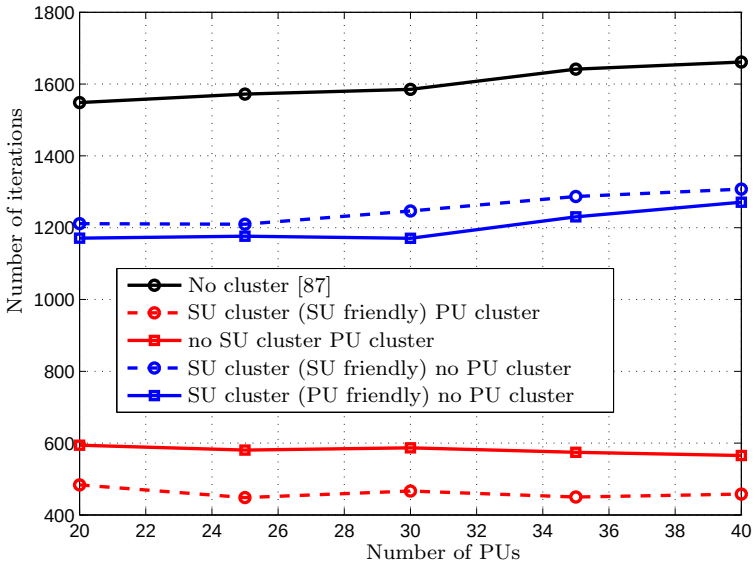


Figure 4.6: Total number of inner iterations to reach the feasible user combination and its corresponding optimal power and beamforming allocation

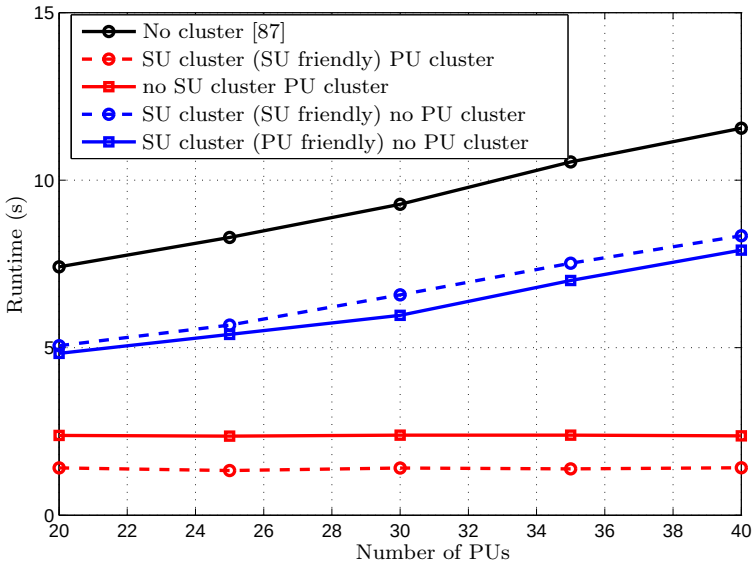


Figure 4.7: Runtime to reach the feasible user combination and its corresponding optimal power and beamforming allocation

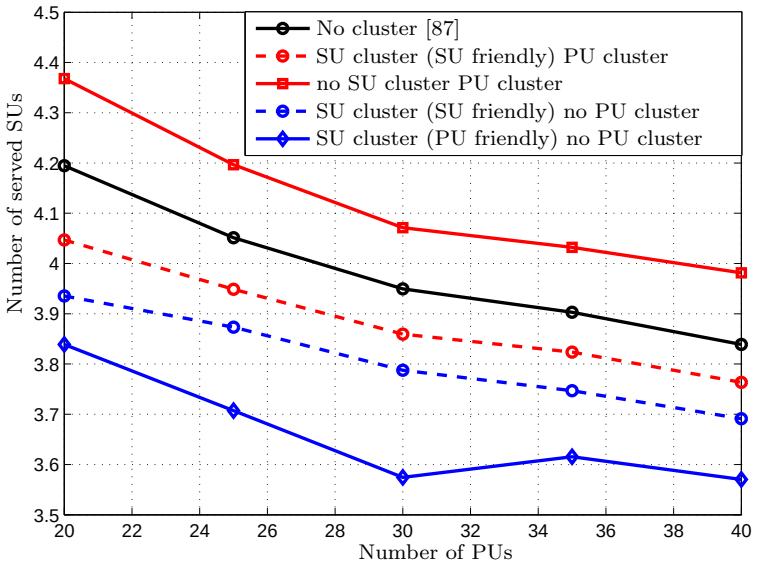


Figure 4.8: Number of secondary users which can be simultaneously served with required QoS

that in the clustering-based schemes less users are served than in the step by step deflation procedure. The difference, however, is around 0.1-0.25 users on average, in the case of SU clustering with random initializations and without PU clustering, and even smaller in the case of combined SU and PU clustering. Interestingly, clustering only the PUs results in a larger number of served users than the step by step deflation. A possible explanation for this can be that the effect of the PUs is emphasized more by the creation of the covering elipsoids serving as PU representatives, thus further contributing to the VUL powers of the SUs and reducing the ambiguities which were mentioned in Section 4.2.2.

Finally, in Figure 4.9, we show the transmitted power, required by all schemes to serve the set of users, they selected. We only considered for this comparison, the runs in which all techniques achieved the same number of users. We notice from Figure 4.9 that the differences between the SU clustering techniques are almost negligeable. On the other hand, as expected, the PU clustering induces an increase in the transmitted power. This power increase is however below 1.5dB in general.

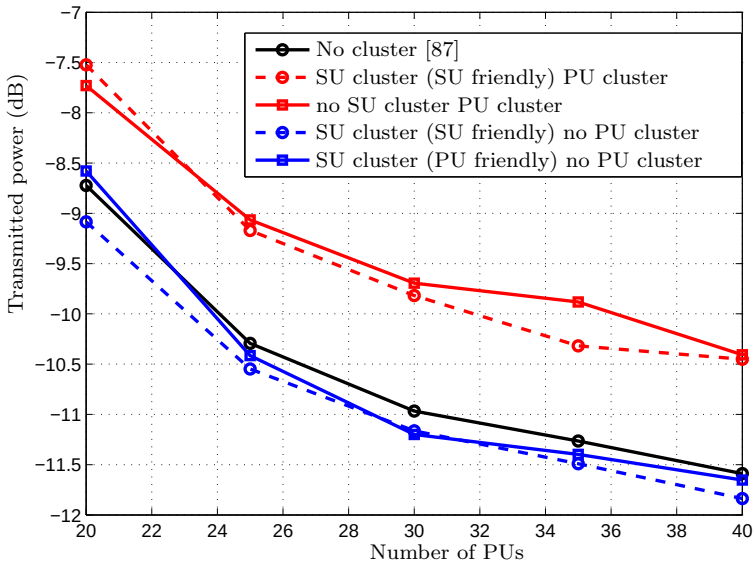


Figure 4.9: Transmitted power when all scenarios are feasible

4.4 Summary

We have shown, in this chapter, an approach to detect infeasibility, using Lagrange duality and the theorems of alternatives. Based on this and the iterative algorithms in the previous chapter, we have constructed a joint admission control and beamforming method, which selects a maximal set of SUs to be served, and for these, it compute the optimal beamformer and power allocation. Simulation results showed that the resulting methods perform close to optimal in point of served number of users. In achieving this, they require however a significantly reduced complexity as compared to the optimal search-based schemes. For the case of dense networks, we have proposed a cluster-aided admission and beamforming approach. We showed that, by forming clusters of SUs and PUs, based on their long term spatial signatures, a significant reduction in complexity can be achieved, as compared to the case when no grouping is performed. Furthermore, it was shown, through simulations, that using the clustering in both SUs and PUs, the resulting methods scale well with the increase in the number of PUs.

Chapter 5

Conventional Robust Downlink Beamforming

In the previous chapters, we have considered the beamforming design in which the BS has perfect knowledge of the CSI. This assumption is, however, rarely met in practice due to, e.g., imprecise estimation at the receiver [90], outdated or limited feedback [91] or quantization effects at transmitter and receiver. On the other hand, disregarding the errors in the channel information may lead to designs which violate the QoS or interference constraints for the true CSI. Therefore, our focus, in the rest of the thesis, is on robust beamforming methods, which are capable to reliably meet the imposed design requirements, even when the available sCSI is erroneous.

In this introductory chapter, we first give a short overview of the vast literature on the various instances of robust beamforming problems in Section 5.1. Thereafter, we introduce the signal model and the particular problem formulation, we are interested to solve in Section 5.2. The state-of-the-art methods for this scenario, together with their drawbacks, are presented in Section 5.3.

5.1 Background

Robust beamforming techniques were initially developed in the context of array processing, where mismatches in the steering vectors due, to e.g., imperfect calibration, look direction errors, manifold mismodeling and distorted antenna shapes cause significant degradation in the performance of the existing non-robust methods. The diagonal loading technique, which consists of adding small regularization parameters to the diagonal of the covariance matrices, has been recognized as a promising technique to overcome these drawbacks [92]. Attractive, due to its simplicity, the method suffers however from the shortcoming that the regularization parameters cannot be easily or optimally computed [92], [94]. A more general framework has been proposed in [93], namely worst case robust beamforming which consists in satisfying the constraints for all the errors lying in a sphere around the presumed signal steering vectors. In this work, the connection to the diagonal loading has been shown. The results have been further extended to more general ellipsoidal sets [95] and general rank constraints [96]. Less conservative approaches have been proposed in the form of probabilistic methods which make different assumptions on the statistics of the spatial signature. Considering a Gaussian distribution on the steering vector, the authors in [97] propose a robust design which ensures an outage probability of the QoS measure below a given threshold. Stochastic approaches, where the statistics of the uncertainties are chosen based on observed samples have been presented in [98].

The robust framework developed for array processing purposes has been extended and tailored to specific requirements of MIMO communication systems [99]-[120]. In the context of MIMO receiver techniques, in [99] and [100], linear equalizers have been derived that minimize the worst case mean square error. In [101], the framework of [96] has been extended and, based on this, a robust receive beamforming scheme for MIMO multiuser acces using orthogonal space time block codes has been developed.

Robust beamforming techniques at the transmitter side have been originally considered in a single user MIMO scenario. Worst case approaches have initially assumed predefined transmit directions and have been designed to optimize some specific QoS metric, e.g., SNR, only with respect to power while assuming the channel mismatches lie in arbitrary convex sets [102], [103]. Interestingly, it was shown in [106] that the transmit directions assumed in [103], i.e., the right singular eigenvectors of the nominal channels, are optimal

for uncertainty errors lying in an ellipsoid measured with weighted Frobenius norm, whereas a similar result was proven in [107] for mismatch sets measured by the spectral norm.

In the multiuser case, no such statements can be made about the optimal transmit directions and thus the optimization problem must be carried over both powers and beamformers. Multiuser beamforming techniques to minimize the transmit power under worst case QoS requirements were developed in [108]-[115], where SINR was chosen as QoS measure in [108], [112]-[115] and MSE in [104]. The robust SINR balancing problem was considered in [113] under both instantaneous and covariance based CSI.

The less conservative probabilistic approaches have been considered for robust downlink beamforming problems in [109], [119], [120], where Gaussian errors were assumed in the CSI. The approaches focused on finding appropriate convex approximations for the probabilistic constraints, e.g., ball approximations [120], Bernstein-type inequalities [120]. In this thesis we consider the general case of erroneous statistical CSI in a multiuser MISO scenario, for which the problem formulation is introduced in the next section.

5.2 Problem Formulation

Similar to the scenario considered in the previous chapters, we assume one BS with N antennas, which serves K SUs in the presence of L PUs. We remind that the transmitted signal at the BS is

$$\mathbf{x}(n) = \sum_{k \in S} \mathbf{w}_k s_k(n), \tag{5.1}$$

whereas the received signals at the kth SU and lth PU can be written as $\mathbf{y}_k = \mathbf{h}_k^H \mathbf{x} + n_k$ and $\mathbf{y}_l = \mathbf{h}_l^H \mathbf{x}$, respectively. At the transmitter, erroneous sCSI is available in the form of estimates of the channel covariance matrices, for both SUs and PUs, and can be expressed as

$$\hat{\mathbf{R}}_k = \mathbf{R}_k - \mathbf{\Delta}_k;\ k = 1, \ldots, K + L, \tag{5.2}$$

where $\mathbf{R}_k$ and $\mathbf{\Delta}_k$ denote the true channel covariance matrix and the error matrix, respectively, of the kth user.

A worst case approach for the robust downlink beamforming problem consists in finding a resource allocation, such that the QoS and interference constraints are satisfied for all

the channel mismatches, in a predefined uncertainty set. For a generic distance measure $d : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{R}^+$ and a positive theshold α_k, the uncertainty set at user k consists of all error matrices $\tilde{\mathbf{\Delta}}_k$ such that $d(\hat{\mathbf{R}}_k, \hat{\mathbf{R}}_k + \tilde{\mathbf{\Delta}}_k) \leq \alpha_k$. Consequently, the worst case SINR is defined as the smallest SINR, which is obtained with a mismatched channel covariance matrix from the predefined uncertainty set. Thus, the worst case SINR with respect to the distance function d, WSINR_d, is given by an optimization problem of the form

$$\mathrm{WSINR}_{d,k}\left(\mathbf{\Delta}_k^\star, \mathbf{w}_1, \ldots, \mathbf{w}_K\right) \triangleq \min_{\mathbf{\Delta}_k} \frac{\mathbf{w}_k^H \left(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k\right) \mathbf{w}_k}{\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{w}_i^H \left(\hat{\mathbf{R}}_k + \mathbf{\Delta}_k\right) \mathbf{w}_i + \sigma_k^2}, \tag{5.3a}$$

$$\text{s.to} \quad \mathrm{d}(\mathbf{R}_\mathrm{k}, \mathbf{R}_\mathrm{k} + \mathbf{\Delta}_\mathrm{k}) \leq \alpha_\mathrm{k}, \quad \hat{\mathbf{R}}_\mathrm{k} + \mathbf{\Delta}_\mathrm{k} \succ 0. \tag{5.3b}$$

In (5.3), $\mathbf{\Delta}_k^\star$ represents the mismatch, for which the optimum of the minimization problem (5.3) is attained. Similarly, the worst case interference term with respect to the distance d, WI_d, represents the largest interference experienced by the PUs for a mismatch covariance matrix in the uncertainty set defined with d and can be obtained from the optimization problem

$$\mathrm{WI}_{d,l}\left(\mathbf{\Delta}_{K+l}^\star, \mathbf{w}_1, \ldots, \mathbf{w}_K\right) \triangleq \max_{\mathbf{\Delta}_l} \sum_{i=1}^{K} \mathbf{w}_i^H \left(\hat{\mathbf{R}}_{K+l} + \mathbf{\Delta}_{K+l}\right) \mathbf{w}_i \tag{5.4a}$$

$$\text{s.to } \mathrm{d}(\mathbf{R}_{K+l}, \mathbf{R}_{K+l} + \mathbf{\Delta}_{K+l}) \leq \alpha_{K+l}. \tag{5.4b}$$

With the definitions in Eq. (5.3) and Eq. (5.4), the worst case robust beamforming problem can be written in the general form as

$$\min_{\{\mathbf{w}_i\}} \sum_{i=1}^{K} ||\mathbf{w}_i||^2 \tag{5.5a}$$

$$\text{s.to } \mathrm{WSINR}_{d,k}\left(\mathbf{\Delta}_k^\star, \mathbf{w}_1, \ldots, \mathbf{w}_K\right) \geq \gamma_k \tag{5.5b}$$

$$\mathrm{WI}_{d,l}\left(\mathbf{\Delta}_{K+l}^\star, \mathbf{w}_1, \ldots, \mathbf{w}_K\right) \leq \gamma_{K+l}^{-1} \tag{5.5c}$$

$$k = 1, \ldots, K, \; l = 1, \ldots, L. \tag{5.5d}$$

An alternative to the worst case formulation is solving the stochastic robust beamforming problem, in which it is assumed that the errors follow a given distribution and the SINR

and interference constraints are satisfied with a predefined outage

$$\min_{\{\mathbf{w}_i\}} \sum_{k=1}^{K} \mathbf{w}_k^H \mathbf{w}_k \tag{5.6a}$$

$$\text{s.to } \Pr_{\hat{\boldsymbol{R}}_k} \left\{ \frac{\mathbf{w}_k^H \hat{\boldsymbol{R}}_k \mathbf{w}_k}{\sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{w}_i^H \hat{\boldsymbol{R}}_k \mathbf{w}_i + \sigma_k^2} \geq \gamma_k \right\} \geq 1 - \epsilon_k, \ k = 1, \ldots, K \tag{5.6b}$$

$$\Pr_{\hat{\boldsymbol{R}}_{K+l}} \left\{ \sum_{i=1}^{K} \mathbf{w}_i^H \hat{\boldsymbol{R}}_{K+l} \mathbf{w}_i \leq \gamma_{K+l}^{-1} \right\} \geq 1 - \epsilon_{K+l}, \ l = 1, \ldots, L \tag{5.6c}$$

$$\tag{5.6d}$$

where $\Pr_{\mathcal{X}}\{\cdot\}$ and ϵ_k represent the probability operator, with respect to random variable $\mathcal{X}$ and the kth outage probability threshold, respectively.

5.3 Related Work

The worst case robust beamforming problem with sCSI, expressed in Eq. (5.5) has been previously approached in [58] and [112], for scenarios with $L = 0$ and in [113] - [118] for CR setups with $L \geq 0$. In all these works, the distance measure, to bound the uncertainty set, has been chosen to be the Euclidean norm. We refer to these methods as the conventional robust beamforming approaches to mark the difference to the method described in this thesis, which uses measures, developed on the Riemannian manifold.

A common idea of the existing approaches [58], [112]-[118] is to find a closed form solution or approximation to the inner problem posed by the worst case SINR and interference constraints, according to the chosen distance measure. With this, an adequate reformulation for the original problem is subsequently derived, to allow for an implementation with polynomial complexity.

To solve the inner problem the authors in [58] neglect the positive semidefinite constraints $\mathbf{R}_k + \boldsymbol{\Delta}_k$ and propose an approximation of the WSINRs as

$$\frac{\min_{\|\bar{\boldsymbol{\Delta}}_k\| \leq \alpha_k} \mathbf{w}_k^H (\hat{\boldsymbol{R}}_k + \bar{\boldsymbol{\Delta}}_k) \mathbf{w}_k}{\sum_{\substack{i=1 \\ i\neq k}}^{K} \max_{\|\bar{\boldsymbol{\Delta}}_k\| \leq \alpha_k} \mathbf{w}_i^H (\hat{\boldsymbol{R}}_k + \bar{\boldsymbol{\Delta}}_k) \mathbf{w}_i + \sigma_k^2} \geq \gamma_k. \tag{5.7}$$

A similar idea was used in [113], to deal with the WI constraints in a CR network. The relaxation of the positive semidefinite constraints in [58] and [113] is tight, as shown

in the simulations, however the inner approximations in Eq. (5.7) make the solution excessively conservative. To overcome this drawback, the authors in [112] employ Lagrange duality theory to propose exact closed form expressions for the inner maximization problems representing the SINR constraints.

This technique has been later extended to cognitive radio scenarios in [114], [115]. Interestingly, it was shown in [114], that using Lagrange duality to handle the inner optimization problems, leads to elegant expressions of the WSINR and WI, which resemble the non-robust counterparts but have additional penalty terms to account for robustness. Specifically, the WSINR case can be written as

$$\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k - \sum_{\substack{i=1 \\ i \neq k}} \mathbf{w}_i^H \gamma_k \mathbf{R}_k \mathbf{w}_i \geq \sigma_k^2 \gamma_k + \alpha_k \|\mathbf{A}_k\|, \tag{5.8}$$

where $\mathbf{A}_k$ is a function of the beamformers, as $\mathbf{A}_k \triangleq \gamma_k \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{w}_i \mathbf{w}_i^H - \mathbf{w}_k \mathbf{w}_k^H$. This observation has been used in [117] to develop an iterative robust technique, based on the uplink downlink algorithm presented in Section 3, thus benefiting from lower computational complexity as compared to the SDP-based counterparts in [114]. The inner optimization problems have been alternatively approached using S-lemma in [116], leading however to similar final formulations.

Even though the inner optimization problems are more rigurously treated using the techniques in [112] - [116], the robust beamforming designs proposed in these works can still be overly conservative, due to the use of the simple Euclidean distance. This is because the Frobenius norms do not take into account the properties of the mismatches and thus when errors are not spatially uniform, the methods tend to consume excessive transmit power to cover cases which in realistic situations rarely occur.

Alternatively, in [118], we showed an approach which is suitable for cases in which the errors in the channel covariance matrices are preponderent in certain known directions, an effect which may be caused, e.g., by pilot pollution or interference from neighbouring cells in the channel estimation phase. The idea of the approach is that, when the covariance matrices of the interferences from neighbouring cells are known, they can be appropriately incorporated in weight factors and the robust design can be carried out using weighted Forbenius norms, constructed in this manner.

Alternative approaches, which do not consider the Frobenius norms in multiuser setups

with sCSI, have been presented in [121] and [122]. In [121], a trace norm was proposed, which was shown to be appropriate for particular types of mismatches, e.g., due to finite sampling. Finally in [122], an approach based on the Euclidean distance between the square roots of the true and estimated channel covariance matrices has been proposed. Even though the uncertainty set with this distance resembles the one with the Riemannian measure, employed in this thesis, the motivation of the approach in [122] is merely the mathematical tractability and it is thus entirely different from ours. Furthermore, the approximations employed in the solution proposed in [122] make the approach relatively conservative as compared to the one presented in this thesis.

Chapter 6

Robust Downlink Beamforming Techniques on the Riemannian Manifold

6.1 Introduction

The techniques based on Frobenius distance tend to have a relatively poor performance when the errors are spatially not uniformly distributed [118]. This drawback can be overcome by measuring the mismatches with weighted Frobenius norms, in which the weighting matrices are designed to appropriately match the uncertainty sets. However, in order to appropriately design these weights, additional information is needed about the directions in which the errors preferentially lie, and for this purpose additional feedback is required.

Secondly, and most importantly, the techniques developed for robust beamforming with sCSI are treating matrices as elements of the Euclidean space, which is not a rigurous characterization from a mathematical point of view. This is because channel covariance matrices are positive semidefinite and, therefore, do not admit a vector space structure. However, they form a Riemannian manifold, which exhibits rich geometric properties. The difference between measuring with Frobenius and Riemannian distances can be intuitively perceived as

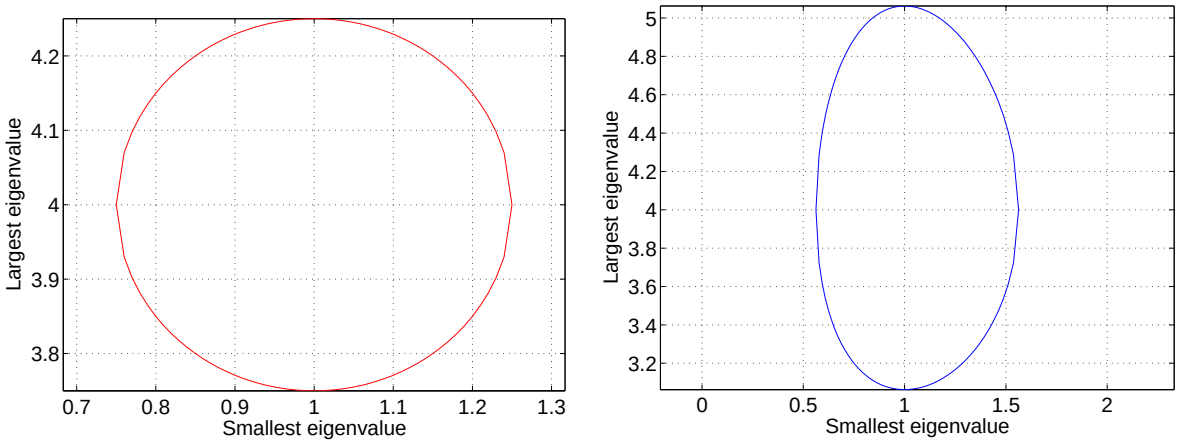
follows. While the former computes a sum of the distances between individual entries of the considered matrices, the latter measures the similarity between manifold components based on the intrinsic features which made them emerge from their original space, e.g., $\mathbb{C}^{N\times N}$. We therefore propose a robust beamforming design in which the mismatch set is bounded based on Riemannian distance derived recently in [134]. We study the implications of using this measure in Section 6.2, by pointing out the aspects which the Riemannian distance emphasizes as compared to the Frobenius norm and the impact of these differences on the robust beamforming design.

Unlike the previously developed exponential distance, e.g., [132], the measure in [134] allows for a closed form expression of the inner optimization problems describing the worst case SINR and interference constraints, which we expose in Section 6.3.1. Based on this, we derive a convex reformulation which can be solved by means of existing interior point methods. Simulations indicate that the properties of the Riemannian distances enable a more flexible robust design, which performs significantly better in point of feasibility and transmit power, when the SINR requirements are large and the system is prone to being interference-limited.

In Section 6.3.3, we show several statistical properties of the Riemannian distance, which can be utilized in establishing thresholds for the uncertainty sets, without requiring learning procedures. Furthermore, based on these properties, a probabilistic approach can be devised, as also shown in this section. Finally, in Section 6.4 we analyse the performance of the distance functions in a scenario with large number of users, and propose admission control schemes, which operate under imperfect sCSI.

6.2 Mismatch Characterization Based on the Riemannian Distance

In [134], two distance measures have been derived, by considering two possible mappings of the psd matrices. A detailed derivation of the distance metric along with intuitive interpretations is given in Appendix B. We remind that when the postive semidefinite matrices are assumed to emerge from $\mathbb{C}^{N\times N}$ by the mapping $\tilde{\mathbf{P}} \rightarrow \tilde{\mathbf{P}}\tilde{\mathbf{P}}^H$, the Riemannian



(a) Frobenius norm

(b) Riemannian distance

Figure 6.1: Isosurface of the eigenvalues of psd matrices which lie within a fixed distance to matrix with eigenvalues 4 and 1

distance is given as

$$d_{R,1}(\mathbf{P}_m,\mathbf{P}_n) = \min_{\substack{\mathbf{P}_m=\tilde{\mathbf{P}}_m\tilde{\mathbf{P}}_m^H \\ \mathbf{P}_n=\tilde{\mathbf{P}}_n^H\tilde{\mathbf{P}}_n}} \|\tilde{\mathbf{P}}_m-\tilde{\mathbf{P}}_n\| \tag{6.1}$$
$$= \sqrt{\mathrm{Tr}\left\{\mathbf{P}_\mathrm{m}\right\}+\mathrm{Tr}\left\{\mathbf{P}_\mathrm{n}\right\}-2\mathrm{Tr}\left\{\left(\mathbf{P}_\mathrm{m}^{1/2}\mathbf{P}_\mathrm{n}\mathbf{P}_\mathrm{m}^{1/2}\right)^{1/2}\right\}}.$$

On the other hand, when the mapping is such that the psd matrices are represented only with respect to their unique square root in $\mathbb{C}^{N\times N}$, the resulting distance measures can be derived as

$$d_{R,2}(\mathbf{P}_m,\mathbf{P}_n) = \sqrt{\mathrm{Tr}\left\{\mathbf{P}_\mathrm{m}\right\}+\mathrm{Tr}\left\{\mathbf{P}_\mathrm{n}\right\}-2\mathrm{Tr}\left\{\mathbf{P}_\mathrm{m}^{1/2}\mathbf{P}_\mathrm{n}^{1/2}\right\}}. \tag{6.2}$$

It is easy to note that the two distances are equal for matrices $\mathbf{P}_m$ and $\mathbf{P}_n$, which span the same space. Since, as shown in Appendix B, there exist different Riemannian measures which stem from very different derivations and underlying mappings. It would therefore be interesting to have a more thorough analysis, regarding the practical implications of these measures.

To this purpose, we visualize the shapes of the uncertainty sets bounded with respect to the Riemannian and Frobenius distance. Let $\mathbf{R}_\mathrm{ref}$ be a reference covariance matrix and the sets S_R and S_F comprise of all the covariance matrices, which are within a Riemannian

and Frobenius distance, respectively, from $\mathbf{R}_{\text{ref}}$, i.e., $S_R = \{\mathbf{R}|d_R(\mathbf{R}, \mathbf{R}_{\text{ref}}) < \alpha_R\}$ and $S_F = \{\mathbf{R}|d_F(\mathbf{R}, \mathbf{R}_{\text{ref}}) < \alpha_F\}$. The largest eigenvalue variations occur when the perturbed matrix $\mathbf{R}$ is in the same space as the original covariance matrix $\mathbf{R}_{\text{ref}}$. Let $\mathbf{R}_{\text{ref}} = \mathbf{U}^H\mathbf{\Lambda}_{\text{ref}}\mathbf{U}$, $\mathbf{R}_F = \mathbf{U}^H\mathbf{\Lambda}_F\mathbf{U}$ and $\mathbf{R}_R = \mathbf{U}^H\mathbf{\Lambda}_R\mathbf{U}$ be the eigenvalue decompositions of $\mathbf{R}_{\text{ref}}$ and of matrices $\mathbf{R}_F \in S_F$ and $\mathbf{R}_R \in S_R$, which span the same space as $\mathbf{R}_{\text{ref}}$. Then,

$$\text{Tr}\left\{(\mathbf{R}_{\text{F}} - \mathbf{R}_{\text{ref}})^{\text{H}}(\mathbf{R}_{\text{F}} - \mathbf{R}_{\text{ref}})\right\} = \text{Tr}\left\{\mathbf{\Lambda}_{\text{F}} - \mathbf{\Lambda}_{\text{ref}}\right\} = \sum(\lambda_{\text{F,i}} - \lambda_{\text{ref,i}})^2 = \alpha_{\text{F}}^2 \tag{6.3}$$

On the other hand, for the Riemannian distance we have

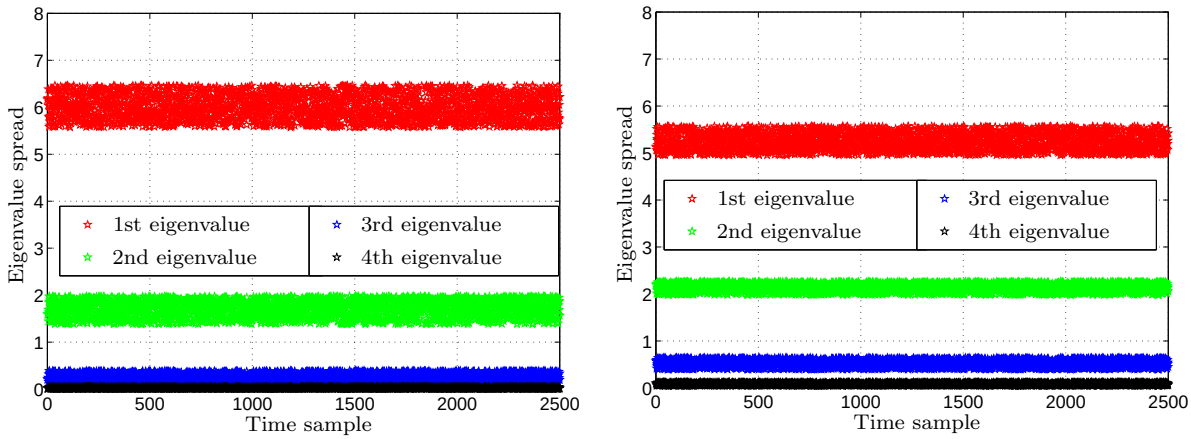
$$\text{Tr}\left\{(\mathbf{R}_{\text{R}}^{1/2} - \mathbf{R}_{\text{ref}}^{1/2})^{\text{H}}(\mathbf{R}_{\text{R}}^{1/2} - \mathbf{R}_{\text{ref}}^{1/2})\right\} = \text{Tr}\left\{\mathbf{\Lambda}_{\text{R}}^{1/2} - \mathbf{\Lambda}_{\text{ref}}^{1/2}\right\} = \sum(\lambda_{\text{R,i}}^{1/2} - \lambda_{\text{ref,i}}^{1/2})^2 = \alpha_{\text{R}}^2 \tag{6.4}$$

Note that (6.4) is valid for both $d_{R,1}$ and $d_{R,2}$, as the two distances are equivalent when their arguments span the same space.

It can then be easily seen from (6.3), that the eigenvalues of the matrices in S_F lie inside circles of radius α_F, centered around the eigenvalues of the reference matrix. On the other hand, for the Riemannian distance, the square roots of the eigenvalues of the matrices in S_R lie inside circles of radius α_R, around each of the eigenvalues of $\mathbf{R}_{\text{ref}}$. This implies that the interval in which each $\lambda_{R,i}$ varies depends on both the threshold α_R and the corresponding eigenvalue of the reference matrix. Therefore, the Riemannian distance allows larger variations of the large eigenvalues and smaller variations of the smaller eigenvalues, as shown in Figure 6.1.

To understand the implications of this observation on the robust beamforming problem we must first note that the channel covariance matrices, which frequently occur in practice, have generally large condition numbers for common antenna architectures, tending to be even low rank when the number of antennas increases [125], [126]. Furthermore, as shown in Figure 6.2, some common error models, e.g., the ones that assume errors due changes in the positions of the users or due to finite sampling, seem to favour the Riemannian model, in that the eigenvalues of the mismatched matrices have dissimilar spreads according to the value of the eigenvalues of the original channel matrices.

Secondly, even though a rigurous theoretical analysis is only available for the two users case [127], due to the complicated interference coupling, it has been observed that beamformers tend to align on the larger components of their corresponding channel covariance matrix and on the weaker components of the interference covariance matrices. Therefore,



(a) Erroneous Least Squares estimation

(b) Erroneous user positions

Figure 6.2: The variation of eigenvalues in erroneously estimated covariance matrices, under various error models

allowing larger variations around the strong components increases the 'price' for interference on the preferred transmit direction of each user, therefore 'encouraging' the beamformers of interfering users to align more on the low components. Thus, a larger separation between users is achieved, which makes it possible to find feasible solutions in more difficult scenarios. On the other hand, allowing larger variations of the small channel components, as permitted by the Frobenius distance, implies that an unnecessary large level of interference must be compensated, which results in an increase in transmitted power.

6.3 Worst Case Robust Downlink Beamforming on the Riemannian Manifold

6.3.1 Proposed algorithm

We first derive a closed form expression for the WSINR and WI expressions, choosing $d_{R,2}$ in Eq. (6.2) as distance function to bound the uncertainty set. Note that, since $d_{R,1}(\mathbf{P}_m, \mathbf{P}_n) \leq d_{R,2}(\mathbf{P}_m, \mathbf{P}_n)$ for all $\mathbf{P}_m$ and $\mathbf{P}_n$, the approach presented here also provides a suboptimal solution for the robust beamforming problem using the $d_{R,1}$ measure. Let us introduce the

variables

$$\mathbf{A}_k \triangleq \gamma_k \sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{w}_i \mathbf{w}_i^H - \mathbf{w}_k \mathbf{w}_k^H, \ k = 1, \dots K. \tag{6.5}$$

Then, the WSINR corresponding to the kth SU can be expressed as

$$\min_{\boldsymbol{\Delta}_k} \ -\mathrm{Tr}\left\{\boldsymbol{\Delta}_\mathrm{k}\mathbf{A}_\mathrm{k}\right\} - \mathrm{Tr}\left\{\hat{\mathbf{R}}_\mathrm{k}\mathbf{A}_\mathrm{k}\right\} - \sigma_\mathrm{k}^2\gamma_\mathrm{k} \geq 0 \tag{6.6a}$$

$$\text{s.to} \quad \mathrm{Tr}\left\{\left(\left(\hat{\mathbf{R}}_\mathrm{k}+\boldsymbol{\Delta}_\mathrm{k}\right)^{1/2} - \hat{\mathbf{R}}_\mathrm{k}^{1/2}\right)^2\right\} \leq \alpha_{\mathrm{R,k}}^2 \tag{6.6b}$$

$$\hat{\mathbf{R}}_k + \boldsymbol{\Delta}_k \succeq \mathbf{0}. \tag{6.6c}$$

A closed form solution to (6.6) can be found as follows.

Proposition 6.1 The kth WSINR resulting from the optimization problem in (6.6) is non-negative if and only if there exists a non-negative ξ_k such that

$$\mathbf{Z}_k(\xi_k, \alpha_{R,k}) \triangleq \begin{pmatrix} -\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2} & \xi_k \mathrm{vec}(\hat{\mathbf{R}}_\mathrm{k}^{1/2}) \\ \xi_k \mathrm{vec}^\mathrm{H}(\hat{\mathbf{R}}_\mathrm{k}^{1/2}) & \xi_k \mathrm{Tr}\{\hat{\mathbf{R}}_\mathrm{k}\} - \sigma_\mathrm{k}^2\gamma_\mathrm{k} - \xi_\mathrm{k}\alpha_{\mathrm{R,k}}^2 \end{pmatrix} \tag{6.7}$$

is positive semidefinite.

Proof The kth WSINR in (6.6) can be equivalently written as

$$\min_{\mathbf{Q}_k} \ -\mathrm{Tr}\left\{\mathbf{Q}_\mathrm{k}^\mathrm{H}\mathbf{A}_\mathrm{k}\mathbf{Q}_\mathrm{k}\right\} - \sigma_\mathrm{k}^2\gamma_\mathrm{k} \tag{6.8a}$$

$$\text{s.to} \quad \mathrm{Tr}\left\{\left(\mathbf{Q}_\mathrm{k} - \hat{\mathbf{R}}_\mathrm{k}^{1/2}\right)^\mathrm{H}\left(\mathbf{Q}_\mathrm{k} - \hat{\mathbf{R}}_\mathrm{k}^{1/2}\right)\right\} \leq \alpha_{\mathrm{R,k}}^2 \tag{6.8b}$$

We first show that omitting the constraints on the Hermitian structure of the variables $\mathbf{Q}_k$ in (6.8) has no effect on the reformulation as any solution of (6.8) is in fact Hermitian. To this purpose, we define the Lagrangian associated with (6.8) as

$$\mathcal{L}(\xi_k, \mathbf{Q}_k) = -\mathrm{Tr}\left\{\mathbf{Q}_\mathrm{k}^\mathrm{H}\mathbf{A}_\mathrm{k}\mathbf{Q}_\mathrm{k}\right\} - \sigma_\mathrm{k}^2\gamma_\mathrm{k} + \xi_\mathrm{k}\left(\mathrm{Tr}\left\{\left(\mathbf{Q}_\mathrm{k} - \hat{\mathbf{R}}_\mathrm{k}^{1/2}\right)^\mathrm{H}\left(\mathbf{Q}_\mathrm{k} - \hat{\mathbf{R}}_\mathrm{k}^{1/2}\right)\right\} - \alpha_{\mathrm{R,k}}^2\right) \tag{6.9}$$

and the corresponding Lagrangian function as

$$\min_{\mathbf{Q}_k} \mathcal{L}(\xi_k, \mathbf{Q}_k). \tag{6.10}$$

The first order optimality conditions can be subsequently derived as

$$\frac{\partial \mathcal{L}(\xi_k, \mathbf{Q}_k)}{\partial \mathbf{Q}_k^H} = (-\mathbf{A}_k + \xi_k \mathbf{I}_N)\mathbf{Q}_k - \hat{\boldsymbol{R}}_k^{1/2}. \tag{6.11}$$

Due to the symmetry of the problem it follows that, if $\mathbf{Q}_k^\star$ is a solution of the problem (6.8), then also its Hermitian $\mathbf{Q}_k^{\star H}$ is a solution of (6.8). Therefore both $\mathbf{Q}_k^\star$ and $\mathbf{Q}_k^{\star H}$ satisfy the first order optimality conditions in (6.11) and we consequently have

$$\mathbf{Q}_k^\star \left(-\mathbf{A}_k + \xi_k \mathbf{I}_N\right) = \left(-\mathbf{A}_k + \xi_k \mathbf{I}_N\right) \mathbf{Q}_k^\star = \hat{\mathbf{R}}_k^{1/2}. \tag{6.12}$$

From (6.12) and further using [53, Th 1.3.12] it follows that $\mathbf{Q}_k^\star = \mathbf{Q}_k^{\star H}$. The rest of the proof consists in deriving the dual problem corresponding to (6.8) and showing that strong duality holds. We then complete the proof using a reformulation based on Schur complement.

Let us define the vectors $\mathbf{v}_k$ by stacking the columns of $\mathbf{Q}_k$, i.e., $\mathbf{v}_k \triangleq \text{vec}(\mathbf{Q}_\text{k})$. Using the properties of the vectorization operator, namely that for any matrices $\mathbf{A}, \mathbf{B}, \mathbf{X}$ and $\mathbf{Y}$, $\text{Tr}\{\mathbf{XY}\} = \text{vec}^\text{H}(\mathbf{X}^\text{H})\text{vec}(\mathbf{Y})$ and $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^\text{T} \otimes \mathbf{A})\text{vec}(\mathbf{X})$, the problem in (6.8) can be equivalently written as

$$\begin{aligned} \min_{\mathbf{v}_k} \quad & -\mathbf{v}_k^H (\mathbf{I}_N \otimes \mathbf{A}_k)\mathbf{v}_k - \sigma_k^2 \gamma_k \\ \text{s.to} \quad & \mathbf{v}_k^H \mathbf{v}_k - 2\text{Re}\left\{\text{vec}^\text{H}\left(\hat{\mathbf{R}}_\text{k}^{1/2}\right)\mathbf{v}_\text{k}\right\} + \text{Tr}\{\hat{\mathbf{R}}_\text{k}\} - \alpha_\text{k}^2 \leq 0. \end{aligned} \tag{6.13}$$

The formulation in (6.13) belongs to the class of single constraint quadratic problems, for which it has been proven that strong duality holds if a strictly feasible solution exists [49]. Since the bound of the uncertainty set is assumed to be strictly positive, a strictly feasible solution to the problem defined by (6.13) is representend by the all zeros vectors, i.e., $\mathbf{v}_k = \mathbf{0}$, for all k. Therefore, in our case, strong duality holds. The Lagrange dual function corresponding to (6.13) is then

$$g_k(\xi_k) = \min_{\mathbf{v}_k} \mathcal{L}(\xi_k, \mathbf{v}_k), \tag{6.14}$$

where $\mathcal{L}(\xi_k, \mathbf{v}_k)$ represents the Lagrangian associated to (6.13), expressed as

$$\mathcal{L}(\xi_k, \mathbf{v}_k) = -\mathbf{v}_k^H (\mathbf{I}_N \otimes \mathbf{A}_k)\mathbf{v}_k - \sigma_k^2 \gamma_k + \xi_k \left(\mathbf{v}_k^H \mathbf{v}_k - 2\text{Re}\left\{\text{vec}^\text{H}\left(\hat{\mathbf{R}}_\text{k}^{1/2}\right)\mathbf{v}_\text{k}\right\} + \text{Tr}\left\{\hat{\mathbf{R}}_\text{k}\right\} - \alpha_\text{k}^2\right). \tag{6.15}$$

According to the KKT conditions, any optimum must satisfy

$$\frac{\partial \mathcal{L}(\xi_k, \mathbf{v}_k)}{\partial \mathbf{v}_k^H} = \left(-\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2}\right)\mathbf{v}_k - \xi_k \text{vec}\left(\hat{\mathbf{R}}_\text{k}^{1/2}\right) = \mathbf{0}. \tag{6.16}$$

Subsequently, we have from Eq. (6.16) that, for a critical point to exist, $\text{vec}\left(\hat{\mathbf{R}}_k^{1/2}\right)$ must be in the range of $-\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2}$. In this case, the optimum solution is

$$\mathbf{v}_k^\star = \left(-\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2}\right)^\dagger \xi_k \text{vec}\left(\hat{\mathbf{R}}_k^{1/2}\right). \tag{6.17}$$

In order for the minimization problem in Eq. (6.14) to be bounded, the matrix $-\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2}$ must be psd.

Using these observations and further introducing Eq. (6.17) in Eq. (6.15), we have that the Lagrange dual function is

$$g(\xi_k) = \begin{cases} c_k - \mathbf{b}_k(-\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2})^{\dagger}\mathbf{b}_k & \text{if } -\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2} \succeq 0, \\ & \mathbf{b}_k \in \text{range}\,(-\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2}) \\ -\infty & \text{otherwise} \end{cases} \tag{6.18}$$

where $c_k \triangleq \xi_k \text{Tr}\left\{\hat{\mathbf{R}}_{\text{k}}\right\} - \xi_{\text{k}}\alpha_{\text{k}}^2 - \sigma_{\text{k}}^2\gamma_{\text{k}}$ and $b_k \triangleq \xi_k \text{vec}\left(\hat{\mathbf{R}}_k^{1/2}\right)$. To satisfy the kth WSINR, we must have that $\max_{\xi_k \geq 0} g(\xi_k) \geq 0$. Consequently, in order to respect the kth SINR constraint, it is sufficient that there exists a $\xi_k \geq 0$, for which $g(\xi_k) \geq 0$. However, using the Schur complement [49], the condition that $g(\xi_k)$ is non-negative is equivalent to the condition that the matrix $\mathbf{Z}_k(\xi_k, \alpha_{R,k}) \triangleq \begin{pmatrix} -\mathbf{I}_N \otimes \mathbf{A}_k + \xi_k \mathbf{I}_{N^2} & \xi_k \text{vec}(\hat{\mathbf{R}}_{\text{k}}^{1/2}) \\ \xi_k \text{vec}^{\text{H}}(\hat{\mathbf{R}}_{\text{k}}^{1/2}) & \xi_k \text{Tr}\{\hat{\mathbf{R}}_{\text{k}}\} - \sigma_{\text{k}}^2\gamma_{\text{k}} - \xi_{\text{k}}\alpha_{\text{R,k}}^2 \end{pmatrix}$ is psd.

□

The worst case interference constraints can be handled in a similar way. We define

$$\mathbf{A}_{K+1} \triangleq \sum_{i=1}^{K} \mathbf{w}_i \mathbf{w}_i^H. \tag{6.19}$$

Thus, $\text{WI}_{d_R,l}$ can be expressed as

$$\max_{\mathbf{\Delta}_l} \ \text{Tr}\left\{\gamma_{K+l}\hat{\mathbf{R}}_{K+l}\mathbf{A}_{K+1}\right\} - 1 \geq 0 \tag{6.20a}$$

$$\text{s.to} \ \ \text{Tr}\left\{\left(\left(\hat{\mathbf{R}}_{K+l} + \mathbf{\Delta}_{K+l}\right)^{1/2} - \hat{\mathbf{R}}_{K+l}^{1/2}\right)^2\right\} \leq \alpha_{R,K+l}^2 \tag{6.20b}$$

$$\hat{\mathbf{R}}_{K+l} + \mathbf{\Delta}_{K+l} \succeq \mathbf{0}. \tag{6.20c}$$

The maximization can be transformed into a minimization, and the problem in (6.20) can be equivalently expressed as

$$\min_{\mathbf{\Delta}_l} \ -\text{Tr}\left\{\gamma_{K+l}\hat{\mathbf{R}}_{K+l}\mathbf{A}_{K+1}\right\} + 1 \geq 0 \tag{6.21}$$

$$\text{s.to} \quad (6.20b) \text{ and } (6.20\text{c}). \tag{6.22}$$

Therefore, using the reformulation in (6.21), a similar statement to that regarding the WSINR inner problems can be made. Specifically, the lth WI is non-positive if and only if

there exists a non-negative ξ_{K+l} such that

$$\mathbf{Z}_{K+l}(\xi_{K+l}, \alpha_{R,K+l}) \triangleq \begin{pmatrix} \mathbf{I}_N \otimes \gamma_{K+l}\mathbf{A}_{K+1} + \xi_{K+l}\mathbf{I}_{N^2} & \xi_{K+l}\mathrm{vec}(\hat{\mathbf{R}}_{\mathrm{K+l}}^{1/2}) \\ \xi_{K+l}\mathrm{vec}^H(\hat{\mathbf{R}}_{K+l}^{1/2}) & \mathrm{Tr}\{\hat{\mathbf{R}}_{K+l}\} - \xi_{K+l}\alpha_{R,K+l}^2 + 1 \end{pmatrix} \succeq 0. \tag{6.23}$$

This equivalence can be shown similarly as the result of Proposition 6.1, and therefore, we omit the proof for brevity. With the use of these two results, we can now reformulate the original worst case robust beamforming problem (5.5) as

$$\min_{\{\mathbf{W}_k\}} \quad \mathrm{Tr}\{\mathbf{W}_{\mathrm{k}}\} \tag{6.24a}$$

$$\text{s.to} \quad \mathbf{Z}_k(\xi_k, \alpha_{R,k}) \succeq 0, \ k = 1, \ldots K + L \tag{6.24b}$$

$$\mathbf{W}_k \succeq 0 \tag{6.24c}$$

In general, the formulation in (6.24) is a relaxation of the original problem in (5.5), due to the rank relaxation of the variables $\{\mathbf{W}_k\}$. However, simulations show that rank one solutions to (6.24) are obtained in the majority of the cases. Therefore, in most practical cases, the two are equivalent.

6.3.2 Simulation Results

In our simulations, we separately consider the cases of $L = 0$ i.e., no PUs are present, and the CR setup, in order to clearly show the impact of the proposed distance for bounding the uncertainty sets. We compare our proposed method to the robust beamforming scheme with the state-of-the-art Frobenius distance [118] and the non-robust scheme [58] for two error models, i.e., due to position changes of the users and due to a combination of erroneous LS estimation and finite sampling. To ensure a fair comparison between the techniques, the thresholds for the uncertainty sets are chosen to cover 95% of the mismatches, for the considered error models. Specifically, 2500 independent channel covariance matrices are generated, according to the considered error models, the Riemannian and Frobenius distances between the true and estimated covariance matrices are computed and the thresholds are chosen such that the desired percentage of matrices are within the uncertainty area. In all simulation setups, we have one BS, which is equipped with an ULA. The channel covariance matrices are chosen according to the model in [80].

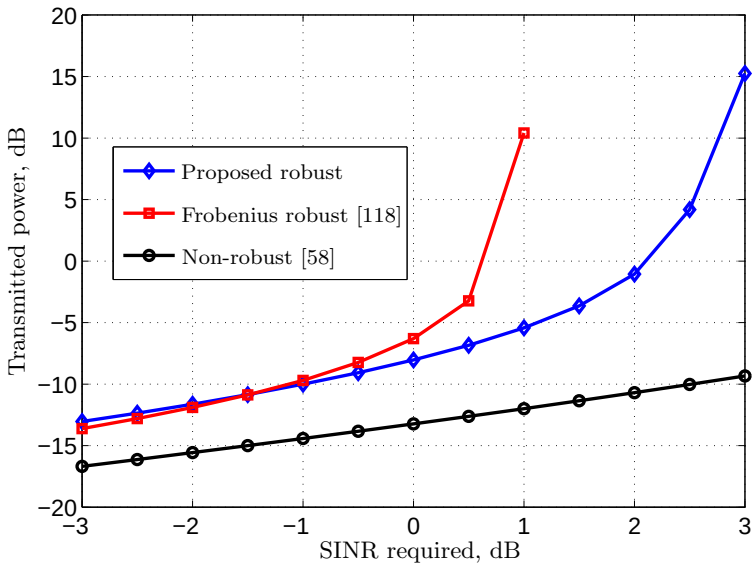


Figure 6.3: Transmitted power for increasing SINR, when error is due to changes in users positions.

In the first simulation setup, we assume that the mismatch is due to changes in the user positions, i.e., the angle of each SU is drawn from a uniform distribution $[0, 7]^\circ$, centered around the around the presumed values. These presumed positions of the SUs are at $[20\ 35\ 60]^\circ$, with respect to the antenna broadside. The required SINR values are between -3 and 3 dB. The noise variance is set to -10dB.

In Figure 6.3, we show the average transmit power to satisfy the increasing SINR demands. We first remark that, the robust technique using Frobenius norms is only feasible, when the imposed SINRs are below 1dB, while our proposed technique, based on Riemannian distance, is able to find feasible beamforming designs for required threshods of up to 3dB. Moreover, it can be noticed from Figure 6.3, that when the imposed SINR thresholds are low, the two robust techniques perform similar in point of transmit power, with a slight advantage in the case of the method using the Frobenius distance. However, as the SINR levels increase, our technique performs significantly better, achieving a reduction in transmit power of even more than 10dB, as compared to the Frobenius-norm counterpart, for imposed SINR levels of 1dB.

To understand the reason behind this large difference, we plot in Figures 6.4 and 6.5, the

contribution of the beamforming components to the interference and useful signal direction. Specifically, we show in Figure 6.4 how the optimum beamformers, obtained by the three techniques, align to components of the channel covariance matrix for each user. For a user k, this can be evaluated as $\mathbf{w}_k^{(t)H}\mathbf{R}_k\mathbf{w}_k^{(t)}/\|\mathbf{w}_k^{(t)}\|$, where $\mathbf{w}_k^{(t)} \in \left\{\mathbf{w}_k^{\star,F}, \mathbf{w}_k^{\star,R}, \mathbf{w}_k^{\star,NR}\right\}$ is the optimum beamformer obtained by the Frobenius, Riemannian and non-robust techniques, respectively. Similarly, we show in Figure 6.5, how the obtained beamfomers align to the components of user 2, which in this scenario is the one most affected by interference. Specifically, what we plot is $\mathbf{w}_k^{(t)H}\mathbf{R}_2\mathbf{w}_k^{(t)}/\|\mathbf{w}_k^{(t)}\|$, for $k = \{1, 3\}$ and $\mathbf{w}_k^{(t)} \in \left\{\mathbf{w}_k^{\star,F}, \mathbf{w}_k^{\star,R}, \mathbf{w}_k^{\star,NR}\right\}$. As can be seen from these two figures, the unit norm beamformers in the case of the Frobenius and non-robust techniques have similar contributions. This implies that the Frobenius method ensures robustness by merely adjusting the power. On the other hand, the Riemannian technique allows a much larger flexibility for the optimal beamformers, which leads to a decreased interference experienced by the users, mainly user 2 which is the one most affected in this scenario. Furthermore, this decrease in the interference terms towards large SINRs is even more significant with the Riemannian characterization, when the worst case mismatches are considered. This is shown in Figure 6.6, where $\mathbf{R}_{2,\text{worst}}$ represents the matrix in the uncertainty set, for which the largest interference term is obtained.

In the following simulation, we change the error model and assume that the channels are obtained using a least squares (LS) estimation method, in which a random Gaussian error with variance -20 dB occurs. The true channels, follow the Rayleigh model with spatial correlation, i.e., $\mathbf{h}_k = \mathbf{R}_k^{1/2}\tilde{\mathbf{n}}_k$, with $\mathbf{R}_k$, denoting the true covariance matrix generated as in [80], and $\tilde{\mathbf{n}}_k$ denoting a Gaussian random vector of zero mean and unit power. The estimated channel, corresponding to one realization is then $\hat{\mathbf{h}}_k = \mathbf{h}_k + \mathbf{T}^{\dagger}\mathbf{e}_k$, where $\mathbf{T}$ and $\mathbf{e}_k$ are the training matrix and the estimation error respectively. It was shown in [141], that any training matrix is optimal if it satisfies the condition $\mathbf{T}\mathbf{T}^H = \frac{P_T}{s_T}$ with P_T being the total training power and s_T the number of training samples. The estimated covariance matrix of the kth user is then constructed as $\hat{\mathbf{R}}_k = 1/N_s \sum_{i=1}^{N_s} \hat{\mathbf{h}}_{k,i}\hat{\mathbf{h}}_{k,i}^H$, where N_s denotes the number of samples, which, in our simultions is set to 512, and $\hat{\mathbf{h}}_{k,i}$ is the estimated channel at user k and time instant i.

We first consider the errors $\mathbf{e}_k$ to be white Gaussian with -20 dB variance and plot in Figures 6.7 and 6.8 the feasibility percentage and transmitted power, respectively, for a

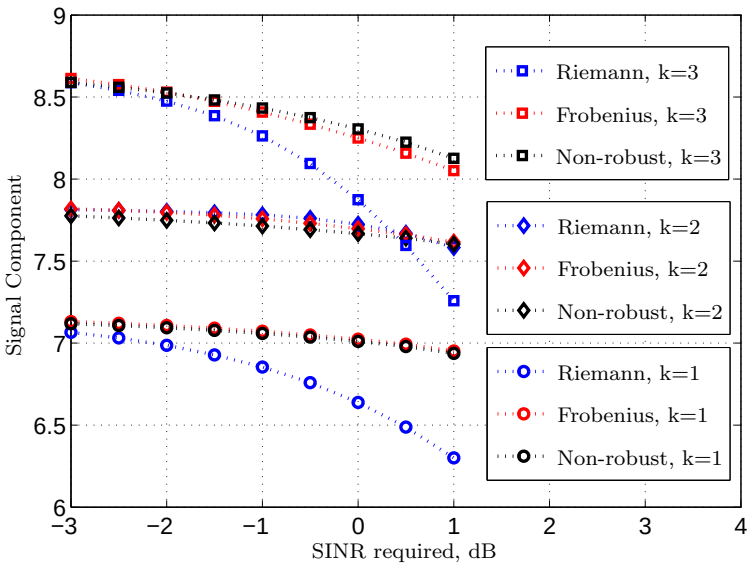


Figure 6.4: Average normed power on transmit direction of each user. For user k, this is assessed as $\mathbf{w}_k^{(t)H}\mathbf{R}_k\mathbf{w}_k^{(t)}/\|\mathbf{w}^(t)_k\|$

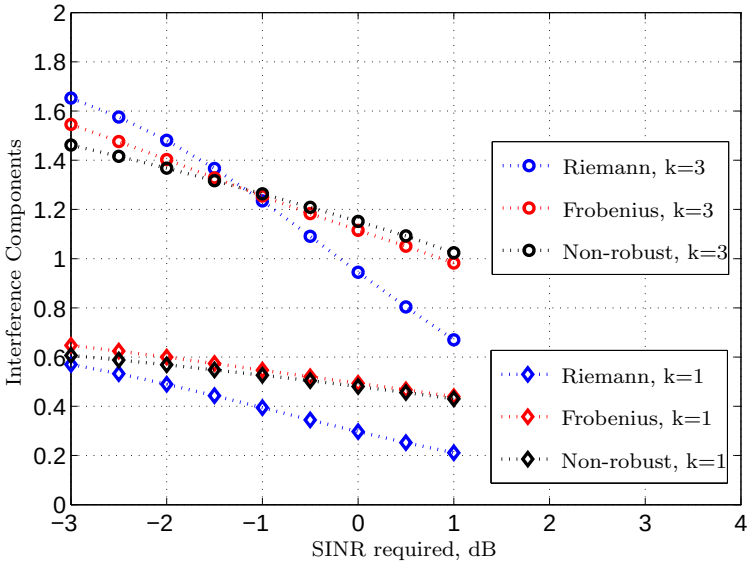


Figure 6.5: Average normed interference received by user 2. The normed interference from the kth user to the second user is assessed as $\mathbf{w}_k^{(t)H}\mathbf{R}_2\mathbf{w}_k^{(t)}/\|\mathbf{w}_k^{(t)}\|$

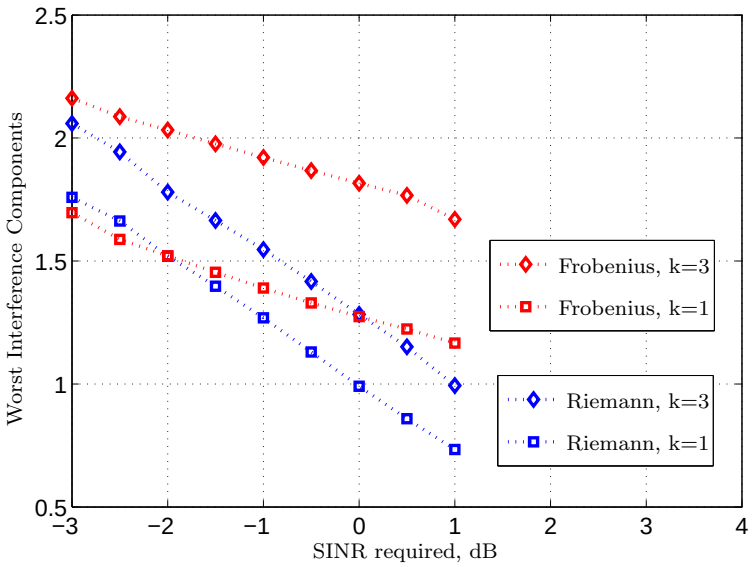


Figure 6.6: Average normed interference received by user 2, for worst mismatch in uncertainty set. The normed interference from the kth user to the second user is assessed as $\mathbf{w}_k^{(t)H}\mathbf{R}_{2,worst}\mathbf{w}_k^{(t)}/\|\mathbf{w}_k^{(t)}\|$

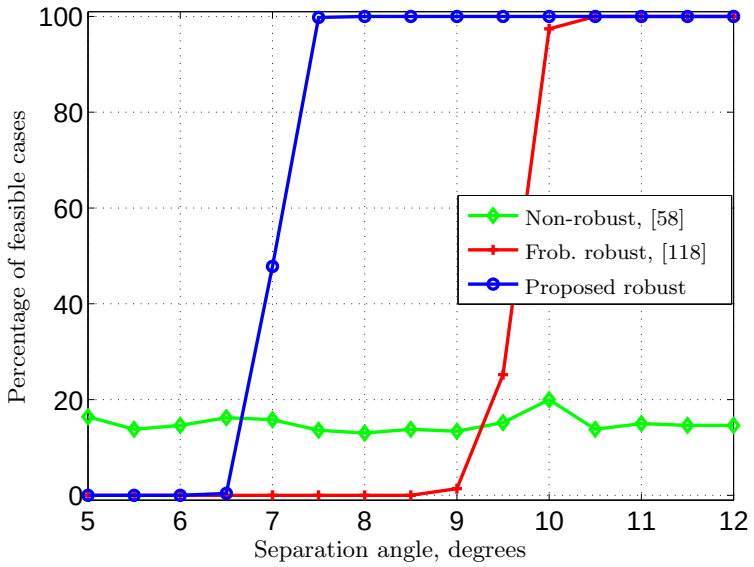


Figure 6.7: Feasibility percentage when error is due to imprecise LS estimation

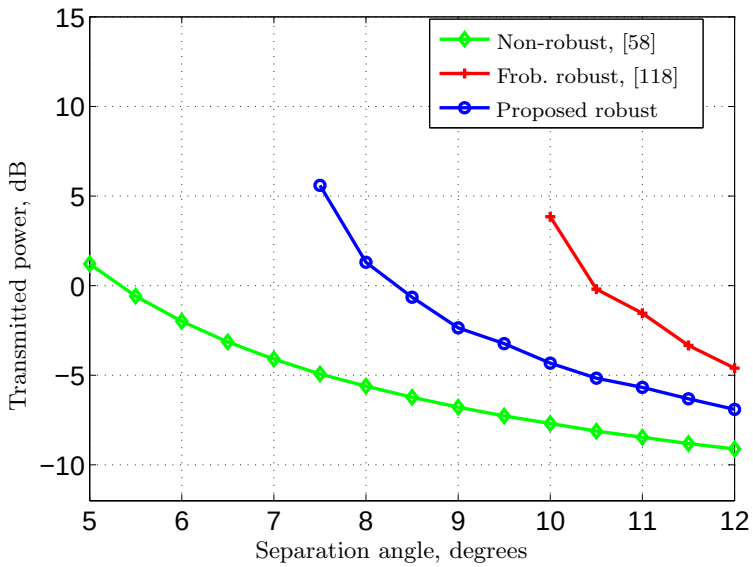


Figure 6.8: Transmitted power when error is due to imprecise LS estimation

separation angle between users varying from 5° to 12°. As in the previous case, we notice that the proposed approach can reliably solve the robust beamforming problem in more difficult scenarios, i.e., for smaller separation angles between the users, than the Frobenius norm counterpart. Moreover, when both approaches are feasible, the performance of our technique clearly outperforms the competing method in terms of transmitted power.

Further, in Figure 6.9, we show the histograms of the weighted SINR values for the three considered cases. The weighted SINR, defined as the ratio between the obtained SINR for the true channel matrix and the imposed threshold, is an indication of both the satisfaction and conservativeness of the approach. The goal is to satisfy these constraints, thus the obtained weighted SINR values must be larger than one. At the same time, however, the obtained SINRs must not be excessively over-satisfied, and, consequently, the weighted SINRs must be as close to one as possible as otherwise, the transmit power is unnecessarily large, without bringing additional benefits for the robustness problem. As we can see from Figure 6.9, even though both techniques lead to an oversatisfaction of these constraints, the characterization with the Riemannian distance, performs better in this respect. On the other hand, the non-robust technique is only able to satisfy about 50% of the constraints.

Finally, we assume that the errors on the LS estimation are coloured, with the covariance

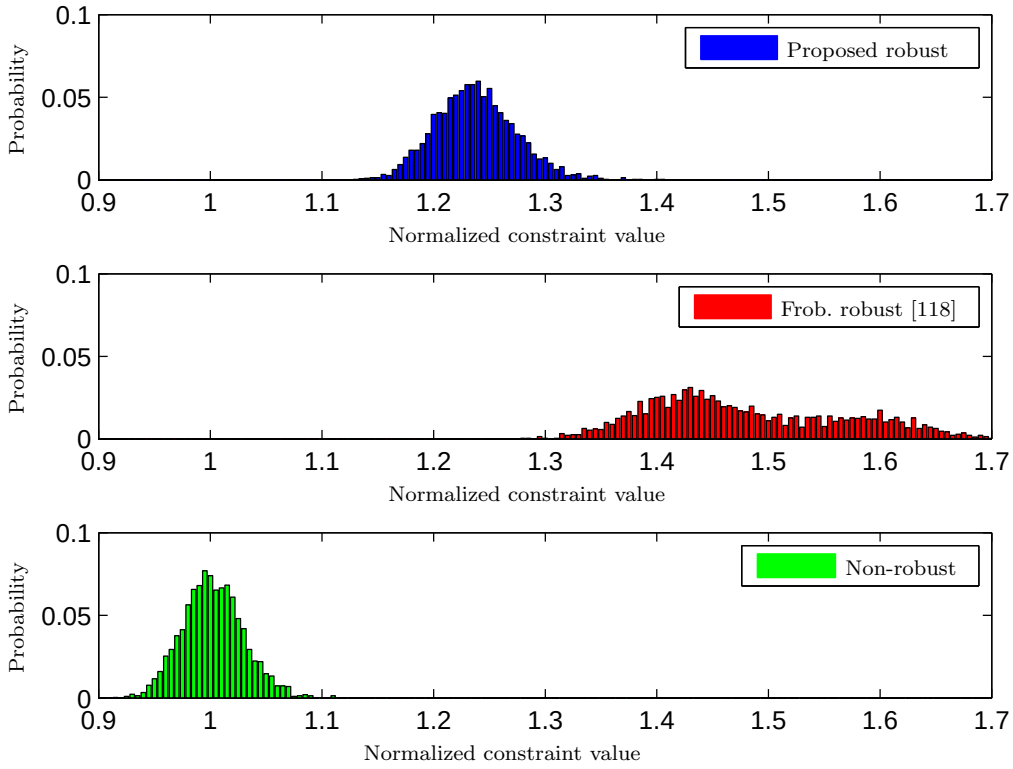


Figure 6.9: Histogram of normalized QoS constraints ($\gamma = -2$ dB).

matrix of the errors at user k, denoted by $\mathbf{R}_{I,k}$. This can be the case e.g., if interferences or reflections occur in the estimation process and results in errors that are preponderant in certain directions dictated by the space of $\mathbf{R}_{I,k}$. We consider both the cases when the error covariance matrices are known and unknown at the transmitter. When the error covariance matrix is known at the transmitter a weighted Frobenius norm can be used to model the uncertainty set and solve the robust beamforming problem [118]. The weight can be defined with respect to $\mathbf{R}_{I,k}$ to essentially whiten the estimation errors as follows. Under the coloured noise assumption we have that $\hat{\mathbf{h}}_k = \mathbf{h}_k + \mathbf{T}^{\dagger}\mathbf{R}_{I,K}^{1/2}\mathbf{e}_k$ and further constructing the sampled covariance matrices, we obtain that $\mathbf{R}_k = \hat{\mathbf{R}}_k + \mathbf{T}^{\dagger}\mathbf{R}_{I,k}^{1/2}\boldsymbol{\Delta}_{e,k}\hat{\mathbf{R}}_{I,k}^{1/2}\mathbf{T}^{\dagger H}$, where

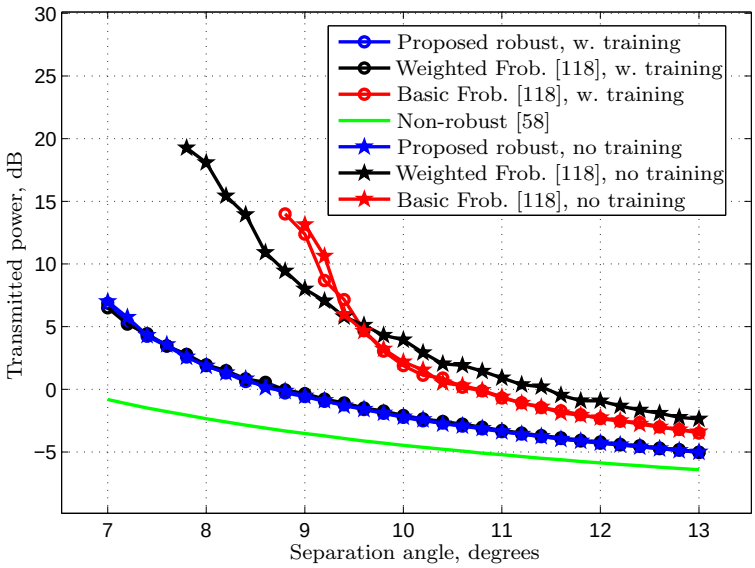


Figure 6.10: Transmitted power on colored noise LS

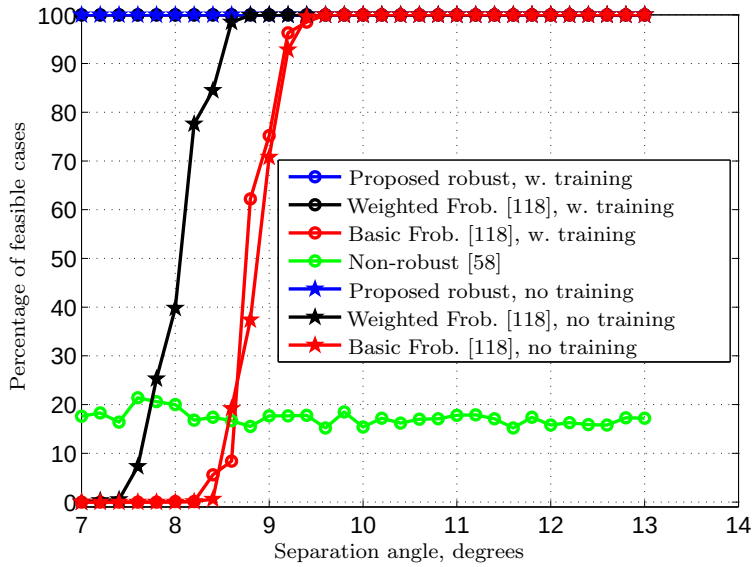


Figure 6.11: Feasibility percentage when colored error on LS

$\mathbf{\Delta}_{e,k} = \frac{1}{N_s}\sum_{i=1}^{N_s} \mathbf{e}_k \mathbf{e}_k^H$. The weighted Frobenius distance is defined as [118]

$$d_{F,\mathbf{\Xi}}(\hat{\mathbf{R}}_k, \mathbf{R}_k) = \sqrt{\mathrm{vec}^{\mathrm{H}}\left(\mathbf{R}_{\mathrm{k}} - \hat{\mathbf{R}}_{\mathrm{k}}\right)\mathbf{\Xi}_{\mathrm{k}}\mathrm{vec}\left(\mathbf{R}_{\mathrm{k}} - \hat{\mathbf{R}}_{\mathrm{k}}\right)}, \tag{6.25}$$

where $\mathbf{\Xi}_k$ denotes the positive definite weighting matrix and, in our case, it can be chosen as

$\mathbf{\Xi}_k = \frac{\left(\mathbf{T}^{\dagger H}\mathbf{R}_{I,k}^{1/2}\right)^T \otimes \left(\mathbf{T}^{\dagger H}\mathbf{R}_{I,k}^{1/2}\right)}{\|\mathbf{\Xi}_k\|}$ We derive in the Appendix C the form of the training matrices that are optimal in the sense that they minimize the power of the estimation error.

In Figures 6.10-6.11 we show the transmitted power necessary for the robust beamforming problem when the separation angle varies from 5 to 12° and consider that near each user an interfer at 10° affects the LS estimation. We consider the 3 cases of weighted, non-weighted and Riemannian distance with both the optimal training as resulting from the derivation in Appendix C and the basic training matrices, that we used for the white Gaussian LS estimation errors. The latter case is motivated by a scenario in which the transmitter is either uninformed of the interference covariance matrices or does not want to invest the additional effort in such a specialized training. Remarkably the 'uninformed' Riemannian distance has a close performance to the 'informed' weighted Frobenius distance. This may be intuitively explained by the fact that the Riemannian distance depends on the covariance matrix to which it is applied, therefore the effect that an error preponderant in some directions has on the shape of the estimated covariance matrix is intrinsically taken into account to a certain extent.

6.3.3 Relation to the Probabilistic Approach

Let us now consider a probabilistic model for CSI, instead of the deterministic one, assumed in the previous section. More precisely, we consider that each channel covariance matrix is a random variable drawn from a Wishart distribution with N_s degrees of freedom and a scale matrix corresponding to the true covariance matrix, i.e., $\hat{\mathbf{R}}_k \sim \mathcal{W}_N(\mathbf{R}_k, N_s)$. The practical case corresponding to this model is the one where the covariance matrices are constructed from N_s samples of estimated channels. If the channels are affected by Rayleigh fading, then we can assume each estimated channel $\hat{\mathbf{h}}_k$ to be independently drawn from a Gaussian N-variate distribution with zero mean and covariance $\mathbf{R}_k$ thus $\hat{\mathbf{R}}_k = \sum_{i=1}^{N_s} \hat{\mathbf{h}}_k(i)\hat{\mathbf{h}}_k^H(i) \sim \mathcal{W}_N(\mathbf{R}_k, N_s)$.

Then, interpreting the Riemannian distance as a statistical measure for random variables, we can gain more insight into the properties of the uncertainty set, considered so far. More precisely, we show in Proposition 6.2 that, for our particular model, the cdf of the Riemannian distance, considered in a statistical sense, can be computed in closed form.

This enables an efficient choice of the thresholds of the uncertainty region, without necessitating an a-priori learning phase, as shown in Section 6.3.2. The statistical properties of the Riemannian distance can be additionally used to derive an approximate solution to the probabilistic robust beamforming problem as further shown in this section.

Proposition 6.2 Let $\hat{\mathbf{R}}_{k,1}$ and $\hat{\mathbf{R}}_{k,2}$ two independent random matrices from a Wishart distribution with N_s degrees of freedom and scale matrix $\mathbf{R}_k$, $\mathcal{W}_N(\mathbf{R}_k, N_s)$. Then,

$$\Pr\left\{d_R(\hat{\mathbf{R}}_{k,1}, \hat{\mathbf{R}}_{k,2}) \leq \alpha_R^2\right\} = F(\alpha_R^2), \tag{6.26}$$

where F is defined as

$$F(t) = 1 - \prod_{j=1}^{r} \lambda_j^{N_s} \sum_{j=1}^{r} \sum_{l=1}^{N_s} \frac{\Psi_{j,l}(-\lambda_j) t^{N_s - l} e^{-\lambda_i t}}{(N_s - l)!(l-1)!}, \tag{6.27}$$

with λ_j denoting the jth eigenvalue of $\mathbf{R}_k$ and further,

$$\Psi_{j,l}(t) = (-1)^{l-1}(l-1)! \sum_{\Omega_j(l)} \prod_{m} \binom{i_m + N_s - 1}{i_m} (\lambda_m + t)^{-N_s - i_m} \tag{6.28}$$

and

$$\Omega_j(l) = \left\{ (i_1, \ldots, i_r) | \sum_{\substack{m=1 \\ m \neq j}}^{r} i_m = l - 1,\ i_m = 0, \ldots, N_s \right\}. \tag{6.29}$$

Proof The matrices $\hat{\mathbf{R}}_{k,1}, \hat{\mathbf{R}}_{k,2} \sim \mathcal{W}_N(\mathbf{R}_k, N_s)$ can be written as $\hat{\mathbf{R}}_{k,1} = \mathbf{R}_k^{1/2} \mathbf{G}_{k,1} \mathbf{G}_{k,1}^H \mathbf{R}_k^{1/2}$ and $\hat{\mathbf{R}}_{k,2} = \mathbf{R}_k^{1/2} \mathbf{G}_{k,2} \mathbf{G}_{k,2}^H \mathbf{R}_k^{1/2}$ where $\mathbf{G}_{k,1}, \mathbf{G}_{k,2}$ are complex $N \times N_s$ matrices whose columns are independent identically distributed complex circular Gaussian variables of zero mean and covariance $\mathbf{R}_k$.

It has been shown in B that the Riemannian distance $d_{R,1}$ between two psd matrices, e.g., $\mathbf{P}_1$ and $\mathbf{P}_2$ is essentially the smallest distance between the equivalence classes, comprised by all complex square matrices, which correspond to $\mathbf{P}_1, \mathbf{P}_2$, i.e., all $\tilde{\mathbf{P}}_1$ and $\tilde{\mathbf{P}}_2$ such that $\mathbf{P}_1 = \tilde{\mathbf{P}}_1 \tilde{\mathbf{P}}_1^H$ and $\mathbf{P}_2 = \tilde{\mathbf{P}}_2 \tilde{\mathbf{P}}_2^H$. It is however easy to see that the same result can be obtained, if the mapping is considered from the space of "fat" complex matrices, i.e., $\tilde{\mathbf{P}}_1$ and $\tilde{\mathbf{P}}_2$ are of dimensions $N \times N_s$ with $N_s > N$. We show this for completeness in Appendix B.3. Using this observation, we can write the Riemannian distance between the two random

matrices $\hat{\mathbf{R}}_{k,1}$ and $\hat{\mathbf{R}}_{k,2}$, by considering as fixed representatives in the complex N by N_s plane, the matrices $\mathbf{R}_{k,1}^{1/2}\mathbf{G}_{k,1}$ and $\mathbf{R}_{k,2}^{1/2}\mathbf{G}_{k,2}$. Thus, we have

$$\Pr_{\hat{\boldsymbol{R}}_1,\hat{\boldsymbol{R}}_2}\left\{\mathrm{d}_{\mathrm{R}}^2(\hat{\boldsymbol{R}}_1,\hat{\boldsymbol{R}}_2)\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.30}$$

$$=\Pr_{\mathbf{G}_{\mathrm{k},1},\mathbf{G}_{\mathrm{k},2}}\left\{\min_{\substack{\tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_2^H=\mathbf{I}_{N_s}\\ \tilde{\mathbf{U}}_2\tilde{\mathbf{U}}_2^H=\mathbf{I}_{N_s}}}\left\|\mathbf{R}_{\mathrm{k}}^{1/2}\mathbf{G}_{\mathrm{k},1}\tilde{\mathbf{U}}_1-\mathbf{R}_{\mathrm{k}}^{1/2}\mathbf{G}_{\mathrm{k},2}\tilde{\mathbf{U}}_2\right\|_{\mathrm{F}}^2\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.31}$$

$$=\Pr_{\mathbf{G}_{\mathrm{k},1},\mathbf{G}_{\mathrm{k},2}}\left\{\left\|\mathbf{R}_{\mathrm{k}}^{1/2}\mathbf{G}_{\mathrm{k},1}\tilde{\mathbf{U}}_1^{\star}-\mathbf{R}_{\mathrm{k}}^{1/2}\mathbf{G}_{\mathrm{k},2}\tilde{\mathbf{U}}_2^{\star}\right\|_{\mathrm{F}}^2\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.32}$$

$$=\Pr_{\underline{\mathbf{G}}_{\mathrm{k}}}\left\{\left\|\mathbf{R}_{\mathrm{k}}^{1/2}\underline{\mathbf{G}}_{\mathrm{k}}\right\|_{\mathrm{F}}^2\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.33}$$

In Eq. (6.32), we used the notations $\mathbf{U}_1^{\star}$ and $\mathbf{U}_2^{\star}$ for the matrices which attain the optimum of the minimization problem. Further, in Eq. (6.33), $\underline{\mathbf{G}}_k$ denotes the difference between the two matrices $\mathbf{G}_{k,1}\tilde{\mathbf{U}}_1^{\star}$ and $\mathbf{G}_{k,2}\tilde{\mathbf{U}}_2^{\star}$. Note that, since the Gaussian distribution is invariant to multiplication with unitary matrices, $\mathbf{G}_{k,1}\hat{\mathbf{U}}_1^{\star}$ and $\mathbf{G}_{k,2}\hat{\mathbf{U}}_2^{\star}$ are also Gaussian with the same properties as $\mathbf{G}_{k,1}$ and $\mathbf{G}_{k,2}$. Furthermore, the difference between two independent Gaussian matrices is also Gaussian thus the probability in (6.32) reduces to the form in Eq. (6.33) with $\underline{\mathbf{G}}_k$ Gaussian.

We now evaluate the probability in (6.33). To this purpose, Eq. (6.33) can be equivalently written as:

$$\Pr_{\underline{\mathbf{G}}_{\mathrm{k}}}\left\{\left\|\mathbf{R}_{\mathrm{k}}^{1/2}\underline{\mathbf{G}}_{\mathrm{k}}\right\|_{\mathrm{F}}^2\leq\alpha_{\mathrm{R}}^2\right\}=\Pr\left\{\mathrm{Tr}\left\{\left(\mathbf{R}_{\mathrm{k}}^{1/2}\underline{\mathbf{G}}_{\mathrm{k}}\right)^{\mathrm{H}}\left(\mathbf{R}_{\mathrm{k}}^{1/2}\underline{\mathbf{G}}_{\mathrm{k}}\right)\right\}\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.34}$$

$$=\Pr\left\{\mathrm{vec}^{\mathrm{H}}\left(\mathbf{R}_{\mathrm{k}}^{1/2}\underline{\mathbf{G}}_{\mathrm{k}}\right)\mathrm{vec}\left(\mathbf{R}_{\mathrm{k}}^{1/2}\underline{\mathbf{G}}_{\mathrm{k}}\right)\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.35}$$

$$=\Pr\left\{\mathrm{vec}^{\mathrm{H}}\underline{\mathbf{G}}_{\mathrm{k}}\left(\mathbf{I}_{\mathrm{N}}\otimes\mathbf{R}_{\mathrm{k}}\right)\mathrm{vec}^{\mathrm{H}}\underline{\mathbf{G}}_{\mathrm{k}}\leq\alpha_{\mathrm{R}}^2\right\} \tag{6.36}$$

$$=\Pr_{\underline{\mathbf{g}}_{\mathrm{k}}\sim\mathcal{N}(0,\mathbf{I}_{\mathrm{NN_s}})}\left\{\underline{\mathbf{g}}_{\mathrm{k}}^{\mathrm{H}}\overline{\boldsymbol{\Lambda}}_{\mathrm{k}}\underline{\mathbf{g}}_{\mathrm{k}}\leq\alpha_{\mathrm{R}}^2\right\}, \tag{6.37}$$

where $\underline{\mathbf{g}}_k\triangleq\mathrm{vec}\left(\underline{\mathbf{G}}_k\right)$. In (6.37), we further denoted by $\overline{\boldsymbol{\Lambda}}_k$ the diagonal matrix containing the eigenvalues of $\mathbf{I}_N\otimes\mathbf{R}_k$, where we considered, without loss of generality, that equal eigenvalues are grouped together, i.e.,

$$\overline{\lambda}_{k,N_s(j-1)+1}=\ldots=\overline{\lambda}_{k,N_s(j-1)+N_s}=\lambda_j;\ j=1,\ldots,r_k, \tag{6.38}$$

for $r_k=\mathrm{rank}(\mathbf{R}_k)$. We assume, for ease of exposition, that the positive eigenvalues of $\mathbf{R}_k$

are distinct. Then, the term in (6.37) can be compactly written as

$$\underline{\mathbf{g}}_k^H \overline{\mathbf{\Lambda}}_k \underline{\mathbf{g}}_k = \sum_{i=1}^{r} \sum_{j=1}^{N_s} \lambda_i |\underline{g}_{k,i+j}|^2, \tag{6.39}$$

where $g_{k,i+j}$ denotes the $i+j$-th element of $\underline{\mathbf{g}}_k$. The expression in (6.39) represents a sum of independent squared Gaussian variables with variances $\lambda_1, \ldots, \lambda_r$, each having the multiplicity N_s. The cumulative distribution function of an expression of this form has been derived in [139] and [140]. By writing the Laplace transformation of the pdf of a squared Gaussian variable as $f_k(s) = \lambda_k/(s+\lambda_k)$, the Laplace transformation of the pdf of the sum in Eq. (6.39) results as $f(s) = \prod_{i=1}^{r} \left(\frac{\lambda_i}{s+\lambda_i}\right)^{N_s}$. Thereafter, using the properties of the Laplace transformation, it follows that the cdf of the expression in (6.39) can be derived as the inverse Laplace transformation of $f(s)/s$. After further computations, this results in [139]:

$$F(t) = 1 - \prod_{j=1}^{r} \lambda_j^{N_S} \sum_{j=1}^{r} \sum_{l=1}^{N_S} \frac{\Psi_{j,l}(-\lambda_j) t^{N_S - l} e^{-\lambda_i t}}{(N_S - l)!(l-1)!}, \tag{6.40}$$

where

$$\Psi_{j,l}(t) = \frac{\partial^{l-1}}{\partial t^{l-1}} \prod_{\substack{m=0 \\ m \neq j}}^{r} (\lambda_m + t)^{-N_S}. \tag{6.41}$$

The partial derivatives in Eq. (6.41) have been further written in closed form in [139] and in our case lead to the expression in Eq. (6.28). □

The result in Proposition 6.2 can be used to derive an approximate solution to the robust beamforming problem with probabilistic constraints, as shown next. Using the notations in Eq. (6.5), the probabilistic robust beamforming problem can be expressed as

$$\min_{\{\mathbf{W}_i\}} \sum_{k=1}^{K} \mathrm{Tr}\{\mathbf{W}_\mathrm{k}\} \tag{6.42a}$$

$$\text{s.to } \Pr\left\{ \mathrm{Tr}\left\{ \left(\hat{\boldsymbol{R}}_k \right) \mathbf{A}_k \right\} \geq \sigma_k^2 \gamma_k \right\} \geq 1 - \epsilon_k;\ k = 1, \ldots, K \tag{6.42b}$$

$$\Pr\left\{ \mathrm{Tr}\left\{ \left(\hat{\boldsymbol{R}}_{K+l} \right) \mathbf{A}_{K+1} \right\} \leq \gamma_{K+l}^{-1} \right\} \geq 1 - \epsilon_{K+l};\ l = 1, \ldots, L \tag{6.42c}$$

The approach is based on the observation that the kth SINR constraint in Eq. (6.42b) is respected with an outage probability of $1 - \epsilon_k$, i.e,

$$\Pr\left\{ \mathrm{Tr}\left\{ \hat{\boldsymbol{R}}_\mathrm{k} \mathbf{A}_\mathrm{k} \right\} \geq \sigma_\mathrm{k}^2 \gamma_\mathrm{k} \right\} \geq 1 - \epsilon_k, \tag{6.43}$$

if, for some random matrix $\overline{\mathbf{R}}_k$, from the same Wishart distribution as $\hat{\boldsymbol{R}}_k$, it holds that

$$\mathrm{Tr}\left\{\hat{\boldsymbol{R}}_{\mathrm{k}}\mathbf{A}_{\mathrm{k}}\right\} \geq \sigma_{\mathrm{k}}^2\gamma_{\mathrm{k}} \text{ for all } \hat{\boldsymbol{R}}_{\mathrm{k}} \text{ such that } \mathrm{d}_{\mathrm{R},1}(\hat{\boldsymbol{R}}_{\mathrm{k}}, \overline{\mathbf{R}}_{\mathrm{k}}) \leq \mathrm{F}^{-1}(1-\epsilon_{\mathrm{k}}). \tag{6.44}$$

This can be proven by the following sequence of equalities and inequalities

$$\begin{aligned}
\Pr\left\{\mathrm{Tr}\left\{\hat{\boldsymbol{R}}_{\mathrm{k}}\mathbf{A}_{\mathrm{k}}\right\}\right\} &= \int_{\mathrm{Tr}\{\mathbf{R}_{\mathrm{k}}\mathbf{A}_{\mathrm{k}}\}\geq\sigma_{\mathrm{k}}^2\gamma_{\mathrm{k}}} p_f(\hat{\boldsymbol{R}}_k)d\hat{\boldsymbol{R}}_k & (6.45)\\
&\geq \int_{d_R^2\left(\hat{\boldsymbol{R}}_k,\overline{\mathbf{R}}_k\right)\leq F^{-1}(1-\epsilon_k)} p_f(\hat{\boldsymbol{R}}_k)d\hat{\boldsymbol{R}}_k & (6.46)\\
&= F(F^{-1}(1-\epsilon_k)) = 1-\epsilon_k, & (6.47)
\end{aligned}$$

where $p_f\left(\hat{\mathbf{R}}_k\right)$ denotes the probability distribution function of $\hat{\mathbf{R}}_k$ and the inequality in (6.46) holds due to (6.44). In this manner, it is posible to approximate the probabilistic constraints (6.42b) with WSINR constraints, of the form in (5.5b), where the boundary of the uncertainty region is chosen based on the statistical properties of the Riemannian distance. Naturally, a similar statement can be made for the PU probabilistic constraints in (6.42c).

In practice, it may be of more interest to consider, instead of the random variables $\overline{\mathbf{R}}_k$, an observed sample of the estimated covariance matrix, say $\mathbf{R}'_k$. If this is the case, the computation of the cdf of $d_R(\hat{\boldsymbol{R}}_k, \mathbf{R}'_k)$ is more involved, since in the expressions in Eq. (6.39) the Gaussian variables have a mean corresponding to the fixed sample matrix. Thus the distribution is generalized non-central chi-square, for which several expressions for numerically computing the cdf have been proposed in e.g., [128], [129].

6.4 Joint User Selection and Beamforming with Imperfect Covariance Based CSI

6.4.1 Introduction

As we have seen in the previous chapter, the beamforming problems tend to become infeasible when a large number of users require access to resources or when the SINR constraints are too stringent. Compensating for imperfect sCSI is an additional cause for infeasibility and the ability of robust techniques to satisfy the constraints without being excessively

conservative is of high importance. Thus, the problem of joint users section and robust beamforming, that is considered in this section, is doubly motivated. Firstly, the need for admission control and beamforming under imperfect CSI is expected to be of great practical relevance, where beamforming designs to simultaneously serve all users may not always exist, and the acquisition of perfect CSI for all users is highly unlikely. Secondly, the problem posed here, completes the analysis of the impact of mismatch characterization on the feasibility and performance of the robust beamforming designs.

The rest of this chapter is organized as follows. In Section 6.4.2, we formulate the joint admission control and beamforming problem under the assumption of imperfect sCSI. For this, we derive in Section 6.4.3 a mixed integer SDP (MISDP) formulation, that can be implemented by Branch and Cut techniques, for which solvers started to appear, e.g, MOSEK and BARON. In this manner, the impact of the uncertainty sets on the robust beamforming problems can be evaluated, for various scenarios of interest. These methods, provide in essence a more intelligent search for optimal or close to optimal solutions, than basic exhaustive searches. However, in terms of complexity, these methods are still too computationally demanding to be used in real time implementations for admission control and beamforming. Therefore in Section 6.4.4, we devise a heuristic inflation algorithm to perform this task, with lower computation complexity. The heuristic measures, based on which the inflation is pursued are based on the evaluation of the Riemannian and Frobenius distances between users and their corresponding uncertainty sets. Simulations in Section 6.4.5 show the consistently and significantly better performance of the robust designs based on Riemannian distances w.r.t. the ones employing Frobenius norms. Furthermore, the heuristic algorithms for admission control and beamforming show a performance close to the optimal schemes, albeit at a significantly reduced complexity.

6.4.2 Problem Formulation and Proposed Approach

We consider the problem in the absense of the PUs, i.e., L=0, noting that the CR case can be straightforwardly extended from this analysis. To account for imperfections in the sCSI, we consider as in the previous section that the constraints must be satisfied for an uncertainty set defined around the presumed channel covariance matrix. Therefore, we build the admission control and beamforming problem on the robust formulation in (5.5).

To include admission control, we introduce the binary variables $\{s_k\}$ such that $s_k = 1$ when user k is admitted and $s_k = 0$ otherwise. Furthermore, the cost function must be modified to ensure that the largest number of users is admitted in the system. To accomodate all these aspects, the problem can be formulated as

$$\min_{\{\mathbf{w}_i\}} \quad \rho \sum_{i=1}^{K} ||\mathbf{w}_i||^2 - \sum_{i=1}^{K} s_i \tag{6.48a}$$

$$\text{s.to} \quad \min_{\substack{d(\hat{\mathbf{R}}_k, \hat{\mathbf{R}}_k+\Delta_k)\leq\alpha_k \\ \mathbf{R}_k+\Delta_k\succeq 0}} \frac{\mathbf{w}_k^H \left(\hat{\mathbf{R}}_k + \Delta_k\right) \mathbf{w}_k}{\sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{w}_i^H \left(\hat{\mathbf{R}}_k + \Delta_k\right) \mathbf{w}_i + \sigma_k^2} \geq s_k \gamma_k \tag{6.48b}$$

$$\sum_{i=1}^{K} \|\mathbf{w}_i\| \leq P_{\text{MAX}} \tag{6.48c}$$

$$s_k \in \{0, 1\}\,;\ k = 1, \ldots, K. \tag{6.48d}$$

In (6.48), the parameter ρ is chosen such that the part of the objective function corresponding to the minimum power objective does not favour a smaller number of admitted users than that could otherwise be simultaneously served. It can be shown that the choice of ρ such that $\rho \leq \frac{1}{1+P_{MAX}}$ ensures the optimality of the solution of (6.48) in point of admitted number of users. The proof of this can be done by contradiction, as follows. Assume that $(\mathbf{w}_k^\star, s_k^\star)_k$ is an optimal solution and there exists $(\overline{\mathbf{w}}_k, \overline{s}_k)_k$, which admits a larger number of users. Then $\sum_{k=1}^{K} \overline{s}_k > \sum_{k=1}^{K} s_k^\star$ or equivalently $-\sum_{k=1}^{K} \overline{s}_k \leq -\sum_{k=1}^{K} s_k^\star - 1$. Consequently,

$$\rho \sum_{k=1}^{K} \|\overline{\mathbf{w}}_k\|^2 - \sum_{k=1}^{K} \overline{s}_k \leq \rho P_{MAX} - \sum_{k=1}^{K} s_k^\star - 1. \tag{6.49}$$

Since ρ is chosen to satisfy $\rho P_{MAX} - 1 < 0$, it further follows that the right-hand-side of (6.49) is smaller that $\rho \sum_{k=1}^{K} \|\mathbf{w}_k^\star\|^2 - \sum_{k=1}^{K} s_k^\star$, which contradicts the optimality of $(\mathbf{w}_k^\star, s_k^\star)_k$.

In addition to the difficulties of the robust beamforming design posed by the non-convexity and inner optimization problems, as discussed in the previous section, the new formulation is further complicated by the presence of the binary variables. These give the problem a combinatorial nature, which results in exponential complexity. An exhaustive search procedure, which examines all possible combinations has a prohibitive computational complexity, even for relatively small problem setups with reduced number of users or antennas. A possible approach is to relax the integer variables. This method however rarely leads

to a meaningful solution, from which a set of users, with a feasible robust beamforming and power allocation can be retrieved. Nevertheless, relaxing some sets of the variables, while keeping the remaining variables fixed provides useful lower bounds, which can be used in Branch and Bound or Branch and Cut techniques. Specifically, these methods consist in constructing search trees, whose nodes represent subsets of the integer variables. Solving an optimization problem at a node and comparing to some appropriately chosen upper and lower bounds, decides whether the corresponding branch is worth pursuing or not. This is also the principle behind existing solvers, developed to handle mixed integer/continuous problem such as mixed integer linear programs (MILP), second order cone programs (MI-SOCP) or semidefinite programs (MISDP).

6.4.3 Benchmark Solution

In this subsection, we derive a MISDP reformulation of (6.48). Similar to the solution in Section 6.3, we are first interested in deriving a closed form expression for the inner optimization problems, represented by the WSINR constraints, which can be written as

$$\min_{\substack{d(\hat{\mathbf{R}}_k,\hat{\mathbf{R}}_k+\Delta_k)\leq\alpha_k \\ \hat{\mathbf{R}}_k+\Delta_k\succeq 0}} \mathbf{w}_k^H\left(\hat{\mathbf{R}}_k+\Delta_k\right)\mathbf{w}_k - s_k\gamma_k\sum_{\substack{i=1\\ i\neq k}}^{K}\mathbf{w}_i^H\left(\hat{\mathbf{R}}_k+\Delta_k\right)\mathbf{w}_i - s_k\gamma_k\sigma_k^2 \geq 0 \qquad (6.50)$$

Note that utilizing directly the result in Proposition 6.1 for $s_k\gamma_k$ has the undesired effect that the admission control and beamforming variables are coupled in the terms $\mathbf{A}_k$, and consequently in $\mathbf{Z}_k$, as can be seen from the definitions of these variables in (6.5) and (6.7), respectively. To overcome this, the 'big M' approach [123] can be taken. This consists in introducing variables $\{M_k\}$ such that the kth WSINR in (6.50) is automatically satisfied when the kth user is not scheduled. On the contrary, the contribution of the $\{M_k\}$ terms to the SINR constraints, corresponding to scheduled users, is null. In this manner, the kth WSINR constraint in (6.50) can be equivalently reformulated as

$$\min_{\substack{d(\hat{\mathbf{R}}_k,\hat{\mathbf{R}}_k+\Delta_k)\leq\alpha_k \\ \hat{\mathbf{R}}_k+\Delta_k\succeq 0}} \mathbf{w}_k^H\left(\hat{\mathbf{R}}_k+\Delta_k\right)\mathbf{w}_k - \gamma_k\sum_{\substack{i=1\\ i\neq k}}^{K}\mathbf{w}_i^H\left(\hat{\mathbf{R}}_k+\Delta_k\right)\mathbf{w}_i - \gamma_k\sigma_k^2 + M_k(1-s_k) \geq 0, \qquad (6.51)$$

where M_k is such that the optimum solution $\mathbf{w}_k^\star = 0$ when $s_k = 0$. It is easy to see that M_k has no effect on the WSINR in (6.51) when a user is admitted, i.e., $s_k = 1$. To find M_k, we

first make the following observation.

Lemma 6.1 Let d be the Riemannian distance $d_{R,2}$ and α_k be the threshold of the uncertainty set measured with this distance. If $M_k \geq P_{MAX}\gamma_k \mathrm{Tr}\left(\hat{\mathbf{R}}_\mathrm{k} + \Delta_\mathrm{k}\right) + \sigma_\mathrm{k}^2\gamma_\mathrm{k}$, then the SINR constraints are automatically satisfied for a non-admitted user, i.e., $s_k = 0$.

Proof Let $\overline{\boldsymbol{\Delta}}_k$ be an arbitrary mismatch matrix. Then,

$$\sum_{\substack{i=1\\i\neq k}}^{K} \mathbf{w}_i^H \gamma_k \left(\hat{\mathbf{R}}_k + \overline{\boldsymbol{\Delta}}_k\right)\mathbf{w}_i + \sigma_k^2\gamma_k = \sum_{\substack{i=1\\i\neq k}}^{K} \mathbf{w}_i^H \gamma_k \left\{\sum_{l=1}^{\mathrm{rank}\left(\hat{\mathbf{R}}_\mathrm{k}+\overline{\boldsymbol{\Delta}}_\mathrm{k}\right)} \lambda_l \mathbf{u}_l\mathbf{u}_l^H\right\}\mathbf{w}_i + \gamma_k\sigma_k^2 \quad (6.52a)$$

$$\leq \sum_{\substack{i=1\\i\neq k}}^{K} ||\mathbf{w}_i||^2\gamma_k \mathrm{Tr}\left(\hat{\mathbf{R}}_\mathrm{k} + \overline{\boldsymbol{\Delta}}_\mathrm{k}\right) + \sigma_\mathrm{k}^2\gamma_\mathrm{k} \quad (6.52b)$$

$$\leq P_{MAX}\gamma_k \mathrm{Tr}\left(\hat{\mathbf{R}}_\mathrm{k} + \overline{\boldsymbol{\Delta}}_\mathrm{k}\right) + \sigma_\mathrm{k}^2\gamma_\mathrm{k}, \quad (6.52c)$$

where $\mathbf{u}_l$ and λ_l in (6.52a) are, respectively, the eigenvectors and eigenvalues of the matrix $\hat{\mathbf{R}}_k + \overline{\boldsymbol{\Delta}}_k$. Note that (6.52b) results from the Cauchy-Schwartz inequality, applied to the inner terms of the sum in (6.52a). Finally, (6.52c) is a direct consequence of the power limitatition of the transmit power by P_{MAX}.

From the last inequality of (6.52) it follows that, if $M_k \geq P_{MAX}\gamma_k \mathrm{Tr}\left(\hat{\mathbf{R}}_\mathrm{k} + \Delta_\mathrm{k}\right) + \sigma_\mathrm{k}^2\gamma_\mathrm{k}$ for all $\boldsymbol{\Delta}_k$ in the uncertainty set, the WSINR constraint is respected for $s_k = 0$. This condition is equivalent to solving the optimization problem

$$\max_{\boldsymbol{\Delta}_k} \quad \mathrm{Tr}\left\{\hat{\boldsymbol{R}}_k + \boldsymbol{\Delta}_k\right\} \quad \text{s.to} \quad d_{R,2}^2(\hat{\boldsymbol{R}} + \boldsymbol{\Delta}_k, \hat{\boldsymbol{R}}_k) \leq \alpha_{R,k}^2. \quad (6.53)$$

Further, (6.53) can be equivalently reformulated as

$$\min_{\mathbf{Q}_k} \quad -\mathrm{Tr}\left\{\mathbf{Q}_k\mathbf{Q}_k^H\right\} \quad (6.54a)$$

$$\text{s.to} \quad \mathrm{Tr}\left\{\left(\mathbf{Q}_k - \hat{\boldsymbol{R}}_k^{1/2}\right)^H\left(\mathbf{Q}_k - \hat{\boldsymbol{R}}_k^{1/2}\right)\right\} \leq \alpha_{R,k}^2. \quad (6.54b)$$

Writing the Lagrangian dual function corresponding to (6.54) and deriving the first order optimality conditions, we have that each optimal solution $\mathbf{Q}_k^\star$ must satisfy

$$\mathbf{Q}_k^\star = \frac{\xi_k}{\xi_k - 1}\hat{\mathbf{R}}_k^{1/2}, \quad (6.55)$$

where ξ_k is the non-negative Lagrange multiplier. By further introducing (6.55) in (6.54b) and constraining that the latter must be satisfied with equality, we have that $\xi_k = 1 +$

$\frac{\sqrt{\mathrm{Tr}\{\hat{\boldsymbol{R}}_{\mathrm{k}}\}}}{\alpha_{R,k}}$ and, consequently, that the optimum of (6.53) is $\left(\alpha_{R,k}+\sqrt{\mathrm{Tr}\left\{\hat{\boldsymbol{R}}_{\mathrm{k}}\right\}}\right)^2$. Thus, the final expression for M_k results as

$$M_k = P_{MAX}\gamma_k\left(\alpha_{R,k}+\sqrt{\mathrm{Tr}\left\{\hat{\boldsymbol{R}}_{\mathrm{k}}\right\}}\right)^2+\sigma_k^2\gamma_k. \tag{6.56}$$

With M_k from Eq. (6.56), the kth WSINR constraint in (6.50) can be equivalently expressed as

$$\mathbf{Z}_k^{(a)} \triangleq \begin{pmatrix} -\mathbf{I}_N\otimes\mathbf{A}_k+\xi_k\mathbf{I}_{N^2} & \xi_k\mathrm{vec}(\hat{\mathbf{R}}_{\mathrm{k}}^{1/2}) \\ \xi_k\mathrm{vec}^{\mathrm{H}}(\hat{\mathbf{R}}_{\mathrm{k}}^{1/2}) & \widetilde{c_k} \end{pmatrix} \succeq \mathbf{0}, \tag{6.57}$$

where $\widetilde{c_k}=\sigma_k^2\gamma_k-M_k(1-s_k)\sigma_k^2\gamma_k+\xi_k\mathrm{Tr}\left\{\hat{\mathbf{R}}_{\mathrm{k}}\right\}-\xi_{\mathrm{k}}\alpha_{\mathrm{k}}^2$. Thus, when the distance measure is $d_{R,2}$ and the threshold $\alpha_{R,2}$, the joint admission control and beamforming problem can be reformulated as

$$\min_{\{\mathbf{W}_i\}} \quad \sum_{i=1}^{K}\rho\mathrm{Tr}\{\mathbf{W}_{\mathrm{i}}\}-\sum_{\mathrm{i}=1}^{\mathrm{K}}\mathrm{s_i} \tag{6.58a}$$

$$\text{s.to} \quad \mathbf{Z}_k^{(a)}\succeq 0 \tag{6.58b}$$

$$\sum_{i=1}^{K}\mathrm{Tr}\{\mathbf{W}_{\mathrm{i}}\}\le \mathrm{P_{MAX}} \tag{6.58c}$$

$$s_k\in\{0,1\};\ k=1,\ldots,K. \tag{6.58d}$$

Even though this class of problems is far less explored than the simpler MILP and MISOCPs, solvers for this have started to emerge, e.g., MOSEK and BARON. These operate based on the Branch and Bound and Branch and Cut techniques, to provide more efficient search methods. The basic idea of these techniques is the following. Starting from a node, representing a user, the continuous relaxation problem corresponding to (6.58)

$$\min_{\{\mathbf{W}_i\}} \sum_{i=1}^{K}\epsilon\mathrm{Tr}\{\mathbf{W}_{\mathrm{i}}\}-\sum_{\mathrm{i}=1}^{\mathrm{K}}\mathrm{s_i} \tag{6.59a}$$

$$\text{s.to } \mathbf{Z}_k^{(a)}\succeq 0 \tag{6.59b}$$

$$\sum_{i=1}^{K}\mathrm{Tr}\{\mathbf{W}_{\mathrm{i}}\}\le \mathrm{P_{MAX}} \tag{6.59c}$$

$$s_k\in(0,1);\ k=1,\ldots,K \tag{6.59d}$$

is solved. If the solution obtained is smaller than an upper bound, which can be chosen as the best integer solution obtained so far, the branch is followed. This implies solving (6.59), in which the s_k variables corresponding to the nodes of the branch, which have been visitied are fixed, while the ones corresponding to the remaining nodes are relaxed. Branches are pruned: *i*) if the continuous relaxation at a node is infeasible *ii*) if an integer solution is found, in which case the result is compared to the best obtained solution or *iii*) if the solution of the relaxation is larger than the cost of the best obtained solution, in which case, further pursuing the branch leads to suboptimal solutions. Additional constraints, commonly referred to as cuts, can be introduced in the formulation in (6.59) to tighten the continuous relaxation without affecting the optimality of the problem.

6.4.4 Heuristic Admission Control and Beamforming with Imperfect sCSI

In this subsection, we construct a low complexity inflation based admission and beamforming algorithm. The algorithm employs the Riemannian and Frobenius measures to assess the distance between users and between the mismatch sets between users.

Specifically, to define a measure, which asserts whether the uncertainty sets of users i and j intersect, we first introduce variables $m(i,j)$ as

$$m(i,j) = d_R(R_j, R_i) - \alpha_{R,i} - \alpha_{R,j}. \tag{6.60}$$

Naturally, if the uncertainty areas of the two users intersect, then $m(i,j) > 0$. We further rank each user, based on how often its uncertainty set intersects the mismatch regions of the other users, which have not yet been admitted at a particular time instant. Denoting this set with S_r, the ranking of user i can then be defined as $\overline{m}(i) = |\{j | m(i,j) > 0,\ j \in S_r\}|$.

The algorithm, summarized in Table 6.1 is then constructed as follows. We intialize the set of admitted users S_a with the user i, whose ranking $\overline{m}(i)$ is the largest. Then, users are added successively in decreasing order of their rankings, provided the distance towards already admitted users is larger than a certain threshold Υ_R. After each step, in which a user is added, the feasibility of the problem (5.5) is tested. If the problem is feasible the procedure is repeated and a new user is considered for admission. Otherwise the procedure is stopped and the set of users, for which the problem is feasible, is returned.

Algorithm 6.1 Heuristic Admission Control and Beamforming with Imperfect sCSI

Step 1. Construct set of users $\mathcal{S}_r$, in the descreasing order of their rankings $\overline{m}$

Step 2. Select the user with largest ranking $\overline{m}(i)$. Initialize the set of admitted users $\mathcal{S}_a$ and upldate the set of remaining users $\mathcal{S}_r$

$$\mathcal{S}_a \leftarrow \text{argmin } \overline{m}(i)$$

$$\mathcal{S}_r \leftarrow \{1, \ldots, K\} \setminus \{i\}$$

Step 3. Until $\mathcal{S}_r$ is empty repeat the following steps

3.1 Select user with the largest metric, from the users, which where not yet admitted $\mathcal{S}_r$

$$j \leftarrow \text{argmin}_{i \in \mathcal{S}_r} \ \overline{m}(i)$$

3.2 Check if the distance between the sCSI of user j and each of the selected users is below a threshold

if $\min_{i \in S_r} d_R(\hat{\mathbf{R}}_j, \hat{\mathbf{R}}_i) \leq \Upsilon_R$, remove $\{j\}$ from the set of users to be admitted at this time instant, $\mathcal{S}_r \leftarrow \mathcal{S}_r \setminus \{j\}$

else if, (5.5) is feasible, admit user $\mathcal{S}_a \leftarrow \mathcal{S}_a \cup \{j\}$

else remove from the set of remaining users $\mathcal{S}_r \leftarrow \mathcal{S}_r \setminus \{j\}$

Table 6.1: The heuristic inflation scheme for robust admission control and beamforming

6.4.5 Simulation results

In this subsection, we compare the optimal solutions of the admission control and beamforming problem with uncertainty sets measured with the Riemannian and Frobenius distances, and we evaluate the performance of the low complexity heuristic schemes.

In the first simulation scenario, we consider one BS with 6 antennas and 10 single antenna users. The positions of the users with respect to the antenna broadside are generated from a uniform distribution, between $[0, 180]^\circ$. Furthermore, we consider that the errors in the

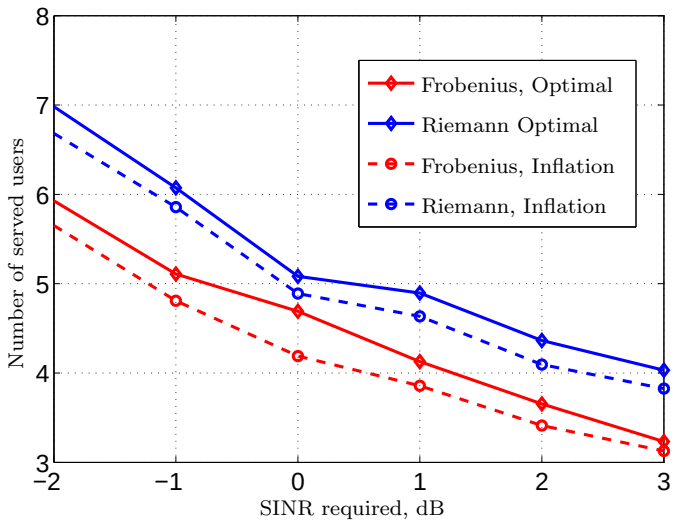


Figure 6.12: Number of users vs error variance

estimated channel covariance matrices are due to changes in the positions of the users. Specifically we consider that the angles based on which the channel covariance matrices are generated are $\{a_i \pm \epsilon_{p,i}\}$, where a_i is the correct position of user i and and $\epsilon_{p,i}$ is a random variable uniformly drawn from $[0, 7]^\circ$. We show in Figure 6.12, the number of served users, as obtained by the methods using the Frobenius and Riemannian distance to measure the uncertainty set, with both benchmark and heuristic solution. As can be observed, the Riemannian distance significantly outperforms the Frobenius distance, in both the optimal and heuristic algorithms, with the Riemannian based method consistently serving at least one more user than the Frobenius counterpart. This behaviour is also observed in regions where the required SINR targets are low, and thus in cases in which the analysis in small scenarios had suggested an almost imperceptible difference between the techniques. Furthermore, the heuristic techniques lose in average at most 0.22 users in the case of the robust admission control with Riemannian distance and 0.4 for the Frobenius one. Thus, they perform close to optimal, although at significantly reduced complexity.

Similar results are obtained for the second simulation scenario. In this simulation setup, we consider one BS with 5 antennas and a pool of $K = 10$ users, which demand access to resources. The SINR levels are fixed to 10dB, while the noise variance at each receiver is -10dB.

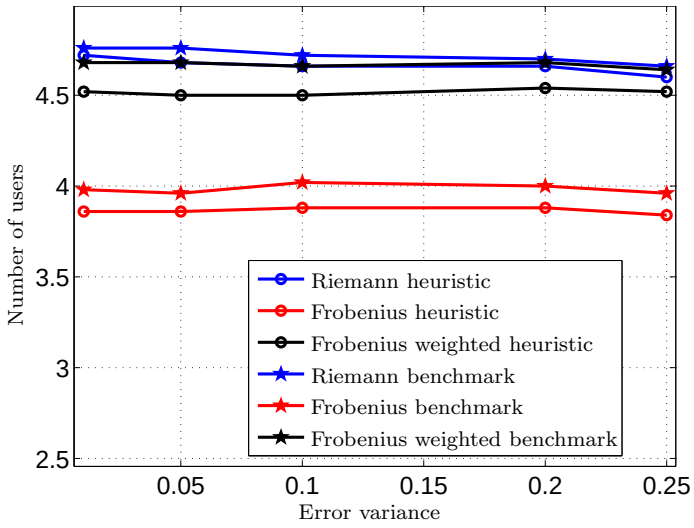


Figure 6.13: Number of users vs error variance

Similar to the previous scenario, the true channel covariance matrices are generated according to the model in [58]. The error in this scenario, occurs due to imprecise LS channel estimation and finite sampling in creating the covariance matrices. Specifically the true channels are generated as $\mathbf{R}_k^{1/2}\widetilde{\mathbf{n}}_k$, where $\mathbf{R}_k$ and $\widetilde{\mathbf{n}}_k$ are the true channel covariance matrix and a unit norm Gaussian vector, modelling the Rayleigh fading, respectively. Further, the estimated channel at user k is $\hat{\mathbf{h}}_k = \mathbf{h}_k + \mathbf{T}^{\dagger}\mathbf{R}_{I,k}^{1/2}\mathbf{e}_k$, with $\mathbf{R}_{I,k}$, $\mathbf{T}$ and $\mathbf{e}_k$ being the channel covariance matrix of interfering signals in the estimation phase, the training matrix and a Gaussian vector, modelling the error fluctuations, respectively. Similar to the simulation scenario in Section 6.3.4, the training matrix is chosen to minimize the error, given the colored noise and has the form derived in Appendix C. Finally, the estimated channel covariance matrix is $\hat{\mathbf{R}}_k = 1/N_s \sum_{i=1}^{N_s} \hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H$, where the numer of samples N_s is chosen $N_s = 512$. We compare the robust admission schemes with Frobenius, weighted Frobenius and Riemannian distance and show the results in point of number of served users, for an increasing variance of $\mathbf{e}_k$, for all k, in Figure 6.13. It can be observed from this figure that, similar to in the results obtained in the previous simulation scenario, the designs with the Riemannian distance are able to serve the largest number of users, significantly more than the scheme with the Frobenius norm and sligtly more than the weighted Frobenius distance. We remind however that the weighted Frobenius distance uses the information

of the error covariance $\mathbf{R}_{I,k}$ in defining the weighting matrices, whereas the Riemannian distance does not require this specific information. Furthermore, the heuristic scheme for the Riemannian distance is close to optimal, losing less than 0.1 users on average as compared to the benchmark scheme using Riemannian distances. Furthermore, with this scheme, almost the same number of users is served, as compared to the optimal admission scheme with weighted Frobenius distance and on average 0.5-0.6 users more than in the optimal Frobenius distance. Thus, bounding the uncertainty sets with Riemannian distance in beamforming designs with imperfect sCSI and which require admission control seems to bring significant improvements, with respect to the Frobenius counterpart.

6.5 Summary

In this chapter, we have proposed a new robust beamforming technique, in which the uncertainty set is measured using a Riemannian distance. We showed that this new measure has several interesting features, which which can be beneficially exploited in the the beamforming design. Particularly, we showed that the designs with Riemannian distances protect more the strong components of the channel covariance matrices, thus the ones of most interest for the beamforming problem. This leads to a significantly better performance of the designs with the Riemannian distance as compared to state-of-the-art methods employing Frobenius norms. Interesting results were obtained not only in small scenarios, but also in setups with larger number of users, where admission control was necessary. To evaluate this, we developed both a benchmark method and a low complexity heuristic method, suitable for real time implementations.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In this thesis, we have proposed DL beamforming algorithms for CR MISO scenarios, in which spectral resources are utilized in an underlay fashion. First, we have considered the problem under the assumption of perfect CSI, in the form of both instantaneous and covariance based CSI. For this, we have proposed efficient algorithms, based on uplink downlink and minimax duality, which exploit the underlying structure of the problem and consequently achieve a significantly reduced computational complexity, as compared to existing state-of-the-art techniques. We have then studied the duality properties of the original DL problem and exposed scenarios in which the original problem achieves strong duality, regardless of the number of PU constraints.

We then considered the joint admission control and beamforming problem under the assumption of perfect CSI at the SUs and PUs. We derived a method to detect infeasibility, based on the theorem of alternatives, from Lagrange duality theory and showed that this can be naturally incorporated in the iterative algorithm, without compromising on computational complexity. The nature of the variable updates in the resulting algorithm, not only has a nice intuitive explanation but also achieves two main goals. It ensures optimality of the solutions in feasible cases and it 'forces' the detection of infeasibility when no set of beamforming and power allocations exists to simultaneously satify all imposed constraints.

Furthermore, the heuristics constructed based on the updated SUs and PUs powers, at the infeasibility decisions, proved meaningful and strong, when used in our proposed deflation-based algorithm. Numerical results have further confirmed that the infeasibility decisions generally require a reduced number of iterations To make the approach also applicable to dense networks, we constructed a preprocessing phase based on unsupervised learning. In this, we clustered SUs and PUs, using their long term spatial signatures, in order to obtain an initial indication regarding the SUs which are unlikely to be served together and the PUs which are likely to pose similar interference constraints on the system. We showed that using this preprocessing phase together with an appropriate choice of representatives can significantly improve the deflation based technique, which we devised for small networks. Indeed, the cluster-aided deflation method was shown to scale well with an increase in the number of users and achieve reductions in runtimes and number of iterations of up to 80%, as compared to the simple deflation method.

In the second part of the thesis, we considered the problem under the more general assumptions on imperfect covariance based CSI. We first considered the problem in the worst case robust framework, with the aim to minimize transmit power while satisfying the SINR constraints for all mismatches in a predefined uncertainty set. We introduced a new characterization of the uncertainty set, in which the mismatches are measured based on a Riemannian distance. Compared to the Frobenius norm, this distance not only better captures the psd properties of the covariance matrices but also exhibits some interesting features, which prove useful in the beamforming design. Specifically we showed that these measures allow larger variations around large eigenvalues and smaller variations around small eigenvalues of the covariance matrices. Thus, when used in a robust beamforming context, the strong components of the channel covariance matrices, which are the ones of most interest for the problem, are protected more than the smaller components. This is particular useful in multiuser scenarios, in which the larger the small components of the covariance matrices are, the larger is the interference level experienced by a particular user. For the new robust beamforming formulation with uncertainty sets bounded based on the Riemannian distance, we have derived a convex approximation, which can be easily implemented with existing interior point solvers. Simulations showed a significantly better performance of the robust techniques using this measure than the state-of-the-art methods

based on Frobenius norms. Furthermore, we showed that, contrary to existing methods, the improvements in terms of feasibility and transmit power are achieved due to a larger flexibility in the beamformer design. This is essentially different than in the designs based on Frobenius distance, where the beamformers are similar to the ones obtained by the non-robust designs and the robustness is achieved mostly by power control.

We then showed some statistical properties of the Riemannian distance, based on which thresholds to bound the uncertainty sets can be analytically derived. Additionally they can employed in deriving probabilistic approaches, in which the SINR constraints are satisfied with a predefined outage and we showed a possible technique to achieve this.

Finally we considered the joint admission control and beamforming problem under the assumption of imperfect covariance based CSI. We have derived both a benchmark method to provide the optimal solutions in a computationally tractable manner and a heuristic low complexity inflation method, more suitable for real time implementation of robust admission control and beamforming problems. The analysis in a large number of users regime confirmed the consistently better performance of the Riemannian distances. Interestingly, simulations showed that the low complexity heuristic methods based on Riemannian distance can even outperform the optimal robust admission control and beamforming solutions based on Frobenius norms, in point of served number of users.

7.2 Future work

The problems addressed in this thesis can be further extended, as follows. First of all the techniques developped in this thesis, in the framework of CR underlay can be straightforwardly applied to heterogenous, multi-tier networks. Furthermore, the robust beamforming techniques with Riemannian distances, can be naturally applied to other scenarios of interest such as multicast, broadcast or relay networks. Besides these, non-trivial extensions can be explored, as discussed further.

First, regarding the duality analysis in a CR underlay framework with perfect CSI, we note that a proof of the string duality of (2.7) in the most general sense, is still an open problem. In this thesis, we have only shown that the original problem attains strong duality under several assumptions. The importance of a strong duality statement in the

most general case is two-fold. Firstly, it ensures that the downlink beamforming problem can always be solved by an equivalent uplink reformulation, with significantly reduced complexity. Secondly it ensures that infeasibility can always be detected. On the other hand, finding classes of problems, for which strong duality does not hold can also be beneficial. In these cases, it should be further analysed whether the scenarios for which strong duality does not hold are of practical interest, and if so, improved beamformer designs and feasibility problems have to be devised.

Regarding the use of Riemannian distances to measure the similarity between channel covariance matrices, several application can be further investigated. Based on this distance, codebooks can be created to appropriately reduce the feedback, required to transmit the channel covariance matrices. Thereafter, applying the robust beamforming problem in conjunction with these codebooks can yield efficient beamforming schemes, in which the feedback overhead is significantly reduced compared to existing techniques. This is achieved both by the use of the codebook itself, as well as by the need for less frequent feedback than in the case of iCSI, due to the slow varying nature of the channel covariance matrices. We note that such codebooks have been constructed for related Riemannian manifolds, i.e., the Grassmianian manifold of fixed rank matrices, and have already found application in codebook beamforming. There exists however to our knwoledge, no codebooks, based on distances derived on the Riemannian manifold of covariance matrices.

Regarding the user clustering problem, based on the Riemannian distance between the channel covariances of the users, one aspect can be further studied. One simple example is a decentralized clustering scheme, in which users exchange information and group themselves according to their channel information. In these ways the groups can act as entities, which choose one representative, which requires service at particular time-instances and feedbacks its channel to the BS. In this manner, the feedback overhead can be further reduced.

Finally, the techniques, developed in this thesis for the robust beamforming problem with imperfect sCSI are based on interior point methods which are still relatively complex. Methods to reduce complexity, by further analysing the structure of the problem would be useful.

Appendices

Appendix A

LP Duality and Implications

It is easy to see that whe the beamformers are fixed, the original problem (2.7) turns into a linear program (LP) in the power variables $\{p_k\}$. This can be expressed as

$$\min_{\mathbf{p}_1} \quad \mathbf{p}_1^T \mathbf{1} \tag{A.1a}$$

$$\text{s.to} \quad (\mathbf{D} - \mathbf{G}_1)\, \mathbf{p}_1 = \boldsymbol{\eta} \tag{A.1b}$$

$$\mathbf{G}_2 \mathbf{p}_1 \leq \mathbf{1}. \tag{A.1c}$$

The Lagrange dual problem, corresponding to (A.1) can then be easily derived as

$$\max_{\mathbf{q}_1, \mathbf{q}_2} \quad \mathbf{q}_1^T \boldsymbol{\eta} - \mathbf{q}_2^T \mathbf{1} \tag{A.2a}$$

$$\text{s.to} \quad (\mathbf{D} - \mathbf{G}_1)^T\, \mathbf{q}_1 - \mathbf{G}_2^T \mathbf{q}_2 = \mathbf{1} \tag{A.2b}$$

$$\mathbf{q}_1 \geq \mathbf{0};\ \mathbf{q}_2 \geq \mathbf{0}. \tag{A.2c}$$

When (A.1) is feasible and bounded, LP duality holds [52], thus, (A.1) and (A.2) attain the same optima. When any of the problems (A.1) or (A.2) is feasible and bounded, the other one is also feasible and bounded. If however, the original is not feasible, the dual LP (A.2) is unbounded above. Furthermore, it is known from LP duality that for any feasible solution for the primal problem (A.1), $\mathbf{p}_1^\star$, and for any feasible solution of the dual problem (A.2), if the complementary slackness condition is satisfied, the primal and dual solutions are optimal and the two problems attain the same solution. Based on this fact,

the construction of a solution, which contradicts positive semidefinitness can be done as follows.

A.1 An Example Which Contradicts Positive Semidefiniteness

Let us consider an instance of the downlink beamforming problem (2.7), for which the PU interference constraints are not implicit, i.e., the inequality between the optimum objective functions of the original problem with $L > 0$ and a problem obtained by removing the PU constraints is strict. Further, let $\{\mathbf{u}_k^\star\}$ be the set of optimal beamformers for this problem (2.7) with $L > 0$ and $\mathbf{q}_1^\star$, the solution of the linear system $(\mathbf{D} - \mathbf{G}_1)\mathbf{q}_1 = \mathbf{1}$, constructed with $\{\mathbf{u}_k^\star\}$.

First, we show that $(\{\mathbf{u}_k^\star\}, \mathbf{q}_1^\star, \mathbf{0}_L)$ is an optimal solution for $\mathcal{PU}_1(S, S_{PU})$. This results as follows. Since $\{\mathbf{u}_k^\star\}$ are feasible for the original problem (2.7), there exists $\mathbf{q}_1$ such that

$$(\mathbf{D} - \mathbf{G}_1)^T \mathbf{q}_1 = \mathbf{1}. \tag{A.3}$$

It is easy to see that choosing $\mathbf{q}_1^\star$ as the solution of (A.3) and $\mathbf{q}_2^\star = \mathbf{0}_L$ is feasible for the dual LP problem. Furthermore, this solution satisfies the complementary slackness conditions for the LP, therefore it is optimal for this. Therefore, we have from LP duality theory that $(\{\mathbf{u}_k\}, \mathbf{q}_1^\star, \mathbf{0}_L)$ is optimal for $\mathcal{PU}_1(S, S_{PU})$.

Next, we prove by contradiction that at least one matrix $\mathbf{Z}_k$ is not positve semidefinite. Indeed, assume that $\mathbf{I} - q_k^\star \mathbf{R}_k + \sum_i q_i^\star \mathbf{R}_i \gamma_i$ is positive semidefinite. Then $\{q_k^\star\}_{k \in S}$ is in the feasible set of the total dual problem, coresponding to a conventional beamforming problem, without PU constraints, expressed as

$$\min_{\{\mathbf{u}_k, p_k\}} \sum_{k=1}^{K} p_k \tag{A.4a}$$

$$\text{s.to } p_i \mathbf{u}_i^H \mathbf{R}_i \mathbf{u}_i - \sum_{k=1}^{K} p_k \mathbf{u}_k^H \mathbf{R}_i \gamma_i \mathbf{u}_k = \sigma_k^2 \gamma_k. \tag{A.4b}$$

Since, in our case $(\{\mathbf{u}_k\}, \mathbf{q}_1^\star, \mathbf{0}_L)$ attains a strictly smaller optimum than $\mathcal{PU}_1(S, S_{PU})$, thus a contradiction is found.

Note that in general (A.4) attains a lower optimum then the original problem (2.7), due to the relaxation of the PU constraints and examples where the difference between the two is strict often occur in simulations.

Appendix B

Deriving Geodesic Distances on the Riemannian Manifold

The aim of this section is to show how the Riemannian distances emerge and what is the intuition behind the expressions, used in this thesis. Interested users may refer to [137] for a rigurous, yet accessible introduction. The derivations of the distances, used in this thesis, have been proposed in [134].

B.1 Definitions

An Euclidean space is a vector space equipped with an inner product. In the case of Euclidean space of $N \times N$ complex matrices the inner product can be chosen as

$$< \mathbf{A}, \mathbf{B} >= \frac{1}{2}\mathrm{Tr}\left\{\mathbf{A}\mathbf{B}^{\mathrm{H}} + \mathbf{B}\mathbf{A}^{\mathrm{H}}\right\}. \tag{B.1}$$

A Riemannian manifold $\mathcal{M}$ is a differentiable manifold that can be locally approximated at each point by a finite dimensional vector space, called tangent space, as it comprises of all vectors tangent to the manifold at that specific point. A curve on a manifold $\mathcal{M}$ is defined as a smooth mapping $\gamma : \mathbb{R} \to \mathcal{M}$. At each point $\mathbf{P}_m$, the tangent space, which we denote by $\mathcal{T}_{\mathcal{M}}\mathbf{P}_m$, is the space of the vectors tangent to all curves passing through $\mathbf{P}_m$. Furthermore, at each point $\mathbf{P}_m$, $\mathcal{T}_{\mathcal{M}}\mathbf{P}_m$, can be endowed with a smooth inner product, called Riemannian

metric. Once a Riemannian metric is defined, the length of a curve $\mathbf{P}(\theta) : [\theta_1, \theta_2] \rightarrow \mathcal{M}$ between two points $\mathbf{P}_m = \mathbf{P}(\theta_1)$ and $\mathbf{P}_n = \mathbf{P}(\theta_2)$ on the manifold $\mathcal{M}$ can be computed as

$$l(\mathbf{P}(\theta)) = \int_{\theta_1}^{\theta_2} \sqrt{g_P\left(\frac{dP(\theta)}{d\theta}, \frac{dP(\theta)}{d\theta}\right)} d\theta, \tag{B.2}$$

where $g_{\mathbf{P}}$ is the inner product on the tangent space at $\mathbf{P}$. The distance between two points on the manifold can then be defined as the length of the shortest path between them, also called geodesic in differential geometry terminology and can be written as

$$d_R(\mathbf{P}_m, \mathbf{P}_n) = \min_{\mathbf{P}:[\theta_1,\theta_2]\rightarrow\mathcal{M}} l(\mathbf{P}(\theta)). \tag{B.3}$$

B.2 Riemannian Manifold of Positive Definite Matrices

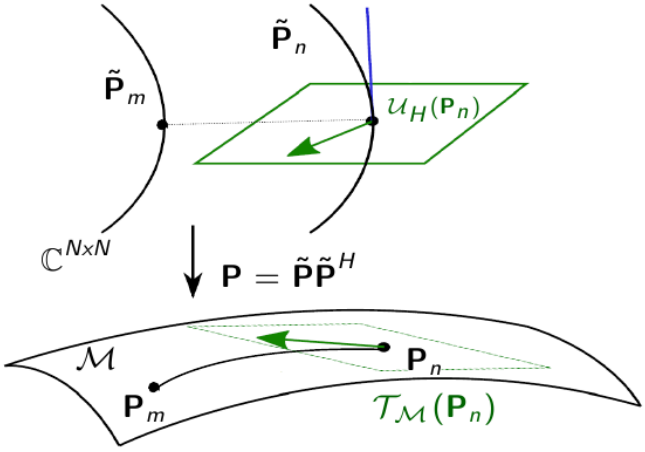
The set of the positive definite matrices can be represented as emerging from the Euclidean space of complex N-by-N matrices with the following mapping:

$$\pi_1 : \mathbb{C}^{N\times N} \rightarrow \mathcal{M}, \text{ such that } \pi_1(\mathbf{P}) = \tilde{\mathbf{P}}\tilde{\mathbf{P}}^{\mathrm{H}}. \tag{B.4}$$

Naturally, the preimage of the mapping in Eq. (B.4) is

$$\pi_1^{-1} : \mathcal{M} \rightarrow \mathbb{C}^{N\times N} \text{ such that } \pi_1^{-1}(\mathbf{P}_{\mathrm{m}}) = \left\{\tilde{\mathbf{P}}_{\mathrm{m}}\mathbf{U} \mid \ \mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I}_{\mathrm{N}}, \ \mathbf{U} \in \mathbb{C}^{\mathrm{N}\times\mathrm{N}}\right\}. \tag{B.5}$$

The matrix $\tilde{\mathbf{P}}_m$ in (B.5) is commonly referred to as representative of the equivalence class, formed by all matrices in the Euclidean plane, which after the mapping (**??**) correspond to $\mathbf{P}_m$. To compute the length of a curve $\mathbf{P}(\theta)$ more conveniently, it is useful to find a correspondence between the tangent spaces along $\mathbf{P}$ and the original Euclidean space. The difficulty of considering tangent vectors in $\mathcal{T}_{\mathcal{M}}\mathbf{P}_m$ with respect to elements of the original space $\mathcal{T}_{\mathbb{C}^{N\times N}}\tilde{\mathbf{P}}_m$ is that there are infintely many valid representations. The tangent space $\mathcal{T}_{\mathbb{C}^{N\times N}}\tilde{\mathbf{P}}_m$ can be however decomposed into two orthogonal subspaces namely the 'vertical' and 'horizontal' subspace. The advantage of this representation is that tangent vectors in the original space $\mathbb{C}$ can be uniquely 'lifted' in the horizontal space. Further, by defining a 'horizontal lift' of a curve $\mathbf{P}(\theta)$ on the manifold $\mathcal{M}$ as the curve in the original space $\mathbb{C}^{N\times N}$, for which the tangent vectors at each point lie in the horizontal space, it is possible to prove that the two curves have the same lengths, provided certain properties of the mapping hold. We show this in more detail, below.



(a) from the set of the complex matrices

Figure B.1: Illustration of the derivation of the Riemannian distances

The vertical space $\mathcal{V}_{\mathbb{C}^{N\times N}}\tilde{\mathbf{P}}_m$ can be considered as the tangent space of the pre-image $\pi^{-1}(\mathbf{P}_m)$. To derive its form, note first that the set $\mathcal{O}_N = \left\{\mathbf{U} \in \mathbb{C}^{N\times N} \mid \mathbf{U}\mathbf{U}^H = \mathbf{I}_N\right\}$ of square unitary matrices is the well known Stiefel manifold of N-by-N matrices. For this, the tangent space at an element $\overline{\mathbf{U}}$ can be written as [138]

$$\mathcal{T}_{\mathcal{O}_N}\overline{\mathbf{U}} = \left\{\overline{\mathbf{U}}\mathbf{S} \mid \mathbf{S} = -\mathbf{S}^H;\ \mathbf{S} \in \mathbb{C}^{N\times N}\right\}. \tag{B.6}$$

Since $\pi_1^{-1}(\mathbf{P}_m)$ is a right multiplication of a representative with all unitary N-by-N matrices, we have that the tangent space of π_1^{-1} and consequently the vertical space at $\tilde{\mathbf{P}}_m$ is

$$\mathcal{V}_{\mathbb{C}^{N\times N}}\tilde{\mathbf{P}}_m = \left\{\tilde{\mathbf{P}}_m\mathbf{S} \mid \mathbf{S}^H = -\mathbf{S}\ \mathbf{S} \in \mathbb{C}^{N\times N}\right\} \tag{B.7}$$

The horizontal space is formed by all matrices in the tangent space $\mathcal{T}_{\mathbb{C}^{N\times N}}\tilde{\mathbf{P}}_m$ which are orthogonal to $\mathcal{V}_{\mathbb{C}^{N\times N}}\mathbf{P}_m$ with respect to the inner product, which in our case is given in Eq. (B.1). Thus, an arbitrary tangent vector in $\mathcal{T}_{\mathbb{C}^{N\times N}}\tilde{\mathbf{P}}_m$, $\dot{\tilde{\mathbf{P}}}_m$, is in the horizontal space if for all for all skew symmetric matrices $\mathbf{S}$, it holds that

$$< \tilde{\mathbf{P}}_m\mathbf{S}, \dot{\tilde{\mathbf{P}}}_m >= \frac{1}{2}\mathrm{Tr}\left\{\left(\dot{\tilde{\mathbf{P}}}_{\mathrm{m}}^{\mathrm{H}}\tilde{\mathbf{P}}_{\mathrm{m}} - \tilde{\mathbf{P}}_{\mathrm{m}}^{\mathrm{H}}\dot{\tilde{\mathbf{P}}}_{\mathrm{m}}\right)\mathbf{S}\right\} = 0 \tag{B.8}$$

For (B.8) to hold for all S, we must have that $\left(\dot{\tilde{\mathbf{P}}}_m^H\tilde{\mathbf{P}}_m - \tilde{\mathbf{P}}_m^H\dot{\tilde{\mathbf{P}}}_m\right) = 0$. This is satisfied

$\dot{\tilde{\mathbf{P}}}_m = \mathbf{K}\tilde{\mathbf{P}}_m$ for all $\mathbf{K}$ Hermitian, thus the horizontal space writes as

$$\mathcal{H}_{\mathbb{C}^{N\times N}} = \left\{\mathbf{K}\tilde{\mathbf{P}}_m \mid \mathbf{K} = \mathbf{K}^H\right\} \tag{B.9}$$

Next, consider the Riemannian metric $g^{(1)}_{\mathbf{P}} : \mathcal{T}_{\mathcal{M}}\mathbf{P} \times \mathcal{T}_{\mathcal{M}}\mathbf{P} \to \mathbb{R}$ defined as

$$g^{(1)}_{\mathbf{P}}(\mathbf{T}_1, \mathbf{T}_2) = \frac{1}{2}\mathrm{Tr}\left\{\mathbf{T}_1\mathbf{K}\right\}, \tag{B.10}$$

where $\mathbf{K}$ is a square Hermitian matrix such that $\mathbf{KP} + \mathbf{PK} = \mathbf{T}_2$.

It was shown in [137] that for any $\mathbf{T}_1, \mathbf{T}_2 \in \mathcal{T}_{\mathcal{M}}(\mathbf{P})$, there exist $\tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_2 \in \mathcal{H}_{\mathbb{C}^{N\times N}}(\mathbf{P})$, such that $g^{(1)}_{\mathbf{P}}(\mathbf{T}_1, \mathbf{T}_2) =< \tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_2 >$ [136] When this is the case, the length of a geodesic in $\mathcal{M}$ is equal to the length of its horizontal lift [133]. Based on this, the following result was obtained.

Proposition 1 [136] If the mapping is chosen as (B.4) and the Riemannian metric as (B.10), then the Riemannian distance is derived as

$$\begin{aligned} d_{R,1}(\mathbf{P}_m, \mathbf{P}_n) &= \min_{\substack{\mathbf{P}_m = \tilde{\mathbf{P}}_m\tilde{\mathbf{P}}_m^H \\ \mathbf{P}_n = \tilde{\mathbf{P}}_n^H\tilde{\mathbf{P}}_n}} \|\tilde{\mathbf{P}}_m - \tilde{\mathbf{P}}_n\| \\ &= \sqrt{\mathrm{Tr}\left\{\mathbf{P}_{\mathrm{m}}\right\} + \mathrm{Tr}\left\{\mathbf{P}_{\mathrm{n}}\right\} - 2\mathrm{Tr}\left\{\left(\mathbf{P}_{\mathrm{m}}^{1/2}\mathbf{P}_{\mathrm{n}}\mathbf{P}_{\mathrm{m}}^{1/2}\right)^{1/2}\right\}} \end{aligned} \tag{B.11}$$

An alternative derivation of the Riemannian distance considers the manifold of positive definite matrices emerging from the set of Hermitian matrices as follows.

$$\pi_2 : \mathbb{C}_H^{N\times N} \to \mathcal{M},\ \mathbf{P} = \tilde{\mathbf{P}}^2,\ \mathbf{P} \in \mathcal{M}, \tilde{\mathbf{P}} \in \mathbb{C}^H. \tag{B.12}$$

Proposition 2 [136] Let $g^{(2)}_{\mathbf{P}}(\mathbf{T}_1, \mathbf{T}_2) : \mathcal{T}_{\mathcal{M}}\mathbf{P} \times \mathcal{T}_{\mathcal{M}}\mathbf{P} \to \mathbb{R}$ be a Riemannian metric defined as

$$g^{(2)}_{\mathbf{P}}(\mathbf{T}_1, \mathbf{T}_2) = \frac{1}{2}\mathrm{Tr}\left\{\mathbf{T}_1^{\mathrm{H}}\mathbf{K} + \mathbf{K}^{\mathrm{H}}\mathbf{T}_1\right\}, \tag{B.13}$$

where $\mathbf{K}$ is such that $\mathbf{PK} + \mathbf{KP} + 2\mathbf{P}^{1/2}\mathbf{KP}^{1/2} = \mathbf{T}_2$ and $\mathbf{T}_1, \mathbf{T}_2 \in \mathcal{T}_{\mathcal{M}}(\mathbf{P})$. Then the distance between $\mathbf{P}_m$ and $\mathbf{P}_n$ is given as

$$d_{R,2}(\mathbf{P}_m, \mathbf{P}_n) = \sqrt{\mathrm{Tr}\left\{\mathbf{P}_{\mathrm{m}}\right\} + \mathrm{Tr}\left\{\mathbf{P}_{\mathrm{n}}\right\} - 2\mathrm{Tr}\left\{\mathbf{P}_{\mathrm{m}}^{1/2}\mathbf{P}_{\mathrm{n}}^{1/2}\right\}} \tag{B.14}$$

Since in this case the pre-image consists of only one element, the introduction of horizontal and vertical spaces is no longer necessary and the proof that $g_{\mathbf{P}}(\mathbf{T}_1, \mathbf{T}_2) =< \tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_2 >$ was done in [136] directly using Lyapunov operators.

B.3 Riemannian Distance with Non-Square Representatives

In the previous section it was shown that the Riemannian distance $d_{R,1}$ between two psd matrices, e.g., $\mathbf{P}_1$ and $\mathbf{P}_2$ is essentially the smallest distance between the equivalence classes, formed by all matrices in the complex space of square matrices, which correspond to $\mathbf{P}_1, \mathbf{P}_2$, i.e., all $\tilde{\mathbf{P}}_1, \tilde{\mathbf{P}}_2$ such that $\mathbf{P}_1 = \tilde{\mathbf{P}}_1\tilde{\mathbf{P}}_1^H$ and $\mathbf{P}_2 = \tilde{\mathbf{P}}_2\tilde{\mathbf{P}}_2^H$. Here, we show that the same result holds, if the mapping is considered from the space of fat complex matrices, instead of that of square complex matrices.

For this, assume $\mathbf{P}_1$ and $\mathbf{P}_2$ be positive definite of dimension N and their corresponding representatives $\underline{\mathbf{P}}_i = \mathbf{P}_i^{1/2}\underline{\mathbf{U}}_i$, where $\underline{\mathbf{U}}_i$ is a fixed unitary matrix in $\mathbb{C}^{N\times N_s}$ and $N_s > N$. Naturally all $\mathbf{P}_i$ can be obtained as $\mathbf{P}_i = \tilde{\mathbf{P}}_i\tilde{\mathbf{P}}_i^H$, with $\tilde{\mathbf{P}}_i = \underline{\mathbf{P}}_i\tilde{\mathbf{U}}_i$ and $\tilde{\mathbf{U}}_i$ a unitary matrix in $\mathbb{C}^{N_s\times N_s}$.

In this case the distance writes as

$$\underline{d}_R(\mathbf{P}_1,\mathbf{P}_2) = \min_{\substack{\tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_1^H=\mathbf{I}_{N_s}\\ \tilde{\mathbf{U}}_2\tilde{\mathbf{U}}_2^H=\mathbf{I}_{N_s}}} \|\mathbf{P}_1^{1/2}\underline{\mathbf{U}}_1\tilde{\mathbf{U}}_i - \mathbf{P}_2^{1/2}\underline{\mathbf{U}}_2\tilde{\mathbf{U}}_2\|^2 \tag{B.15}$$

$$= \mathrm{Tr}\{\mathbf{P}_1\} + \mathrm{Tr}\{\mathbf{P}_2\} - \max_{\substack{\tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_1^H=\mathbf{I}_{N_s}\\ \tilde{\mathbf{U}}_2\tilde{\mathbf{U}}_2^H=\mathbf{I}_{N_s}}} \mathrm{Re}\left\{\tilde{\mathbf{U}}_1^{\mathrm{H}}\underline{\mathbf{U}}_1^{\mathrm{H}}\mathbf{P}_1^{1/2}\mathbf{P}_2^{1/2}\underline{\mathbf{U}}_2\tilde{\mathbf{U}}_2\right\} \tag{B.16}$$

$$= \mathrm{Tr}\{\mathbf{P}_1\} + \mathrm{Tr}\{\mathbf{P}_2\} - \sum_{\mathrm{j}} \underline{\sigma}_{\mathrm{j}}(\mathbf{P}_1^{1/2}\mathbf{P}_2^{1/2}) \tag{B.17}$$

$$= \mathrm{Tr}\{\mathbf{P}_1\} + \mathrm{Tr}\{\mathbf{P}_2\} - \max_{\substack{\overline{\mathbf{U}}_1\overline{\mathbf{U}}_1^H=\mathbf{I}_N\\ \overline{\mathbf{U}}_2\overline{\mathbf{U}}_2^H=\mathbf{I}_N}} \mathrm{Re}\left\{\overline{\mathbf{U}}_1^{\mathrm{H}}\mathbf{P}_1^{1/2}\mathbf{P}_2^{1/2}\overline{\mathbf{U}}_2\right\} \tag{B.18}$$

$$= d_R(\mathbf{P}_1,\mathbf{P}_2), \tag{B.19}$$

where in Eq. (B.17), we have used the notation $\underline{\sigma}_j(\mathbf{P}_1^{1/2}\mathbf{P}_2^{1/2})$ to represent the jth singular value of $\mathbf{P}_1^{1/2}\mathbf{P}_2^{1/2}$

Appendix C

Optimal Training for Coloured LS Estimation Noise

The optimal training matrices for the LS estimation are found by imposing that the power of the estimation noise is minimized [141]. Specifically, for user k, the training matrix can be found by solving the optimization problem:

$$\min_{\mathbf{T}_k} \quad E\left\{\mathrm{Tr}\left\{\mathbf{T}_{\mathrm{k}}^{+\mathrm{H}}\mathbf{E}_{\mathrm{k}}^{\mathrm{H}}\mathbf{E}_{\mathrm{k}}\mathbf{T}_{\mathrm{k}}^{+}\right\}\right\} \tag{C.1a}$$

$$\text{s.to} \quad \mathrm{Tr}\left\{\mathbf{T}_{\mathrm{k}}\mathbf{T}_{\mathrm{k}}^{\mathrm{H}}\right\} = \mathrm{P}_{\mathrm{T}}, \tag{C.1b}$$

where $\mathbf{E}_k$ is the matrix obtained by stacking all samples of the kth error vector, $\mathbf{e}_k$. Defining the error covariance matrix $\mathbf{R}_{\mathbf{E}_k} \triangleq E\left\{\mathbf{E}_k\mathbf{E}_k^H\right\}$, the optimization problem in (C.1) can be equivalently reformulated as

$$\min_{\mathbf{T}_k} \mathrm{Tr}\left\{\mathbf{R}_{\mathbf{E}_{\mathrm{k}}}\left(\mathbf{T}_{\mathrm{k}}^{\mathrm{H}}\mathbf{T}_{\mathrm{k}}\right)^{-1}\right\} \quad \text{s.to} \quad \mathrm{Tr}\left\{\mathbf{T}_{\mathrm{k}}\mathbf{T}_{\mathrm{k}}^{\mathrm{H}}\right\} = \mathrm{P}_{\mathrm{T}}. \tag{C.2}$$

The Lagrangian function, corresponding to (C.2) can be written as

$$\mathcal{L}(\mathbf{T}_k\mathbf{T}_k^H) = \mathrm{Tr}\left\{\mathbf{R}_{\mathbf{E}_{\mathrm{k}}}\left(\mathbf{T}_{\mathrm{k}}^{\mathrm{H}}\mathbf{T}_{\mathrm{k}}\right)^{-1}\right\} + \xi_{\mathrm{k}}\left(\mathrm{Tr}\left\{\mathbf{T}_{\mathrm{k}}\mathbf{T}_{\mathrm{k}}^{\mathrm{H}}\right\} - \mathrm{P}_{\mathrm{T}}\right) \tag{C.3}$$

With this, the first order optimality condition is

$$\frac{\partial\mathcal{L}\left(\mathbf{T}_k\mathbf{T}_k^H\right)}{\partial\mathbf{T}_k\mathbf{T}_k^H} = -\left(\mathbf{T}_k\mathbf{T}_k^H\right)^{-T}\mathbf{R}_{E_k}^T\left(\mathbf{T}_k\mathbf{T}_k^H\right)^{-T} + \xi_k\mathbf{I}_N = 0. \tag{C.4}$$

Naturally, from Eq. (C.4), for the optimum training matrices, it must hold that

$$\mathbf{T}_k \mathbf{T}_k^T = \frac{1}{\xi_k} \mathbf{R}_{\mathbf{E}_k}^{1/2}. \tag{C.5}$$

By further introducing this expression in Eq. (C.1b), we obtain that $\xi_k = \mathrm{Tr}\left\{\mathbf{R}_{\mathbf{E}_\mathrm{k}}^{1/2}\right\}/\xi_\mathrm{k}$ and, finally, that the optimum training matrices must satisfy

$$\mathbf{T}_k \mathbf{T}_k^T = \frac{P_T}{\mathrm{Tr}\left\{\mathbf{R}_{\mathbf{E}_\mathrm{k}}^{1/2}\right\}} \mathbf{R}_{E_k}^{1/2}. \tag{C.6}$$

Bibliography

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 20142019 White Paper," Feb. 2015, [Online]. Available: *http : //www.cisco.com/c/en/us/solutions/collateral/service − provider/visual − networking − index − vni/white_paper_c*11 − 520862.*pdf*

[2] "Spectrum policy task force," Federal Communications Commision, Tech. Rep. 02-135, Nov. 2002. [Online]. Available: http://hraunfoss.fcc.gov/edocs public/attachmatch/DOC-228542A1.pdf

[3] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong and J. C. Zhang, "What Will 5G Be?", *IEEE J. Sel. Areas of Commun.,* vol. 32, no. 6, pp. 1065-1082, Jun. 2014

[4] E. Dahlman, S. Parkvall, and J. Sköld, " 4G: LTE/LTE-Advanced for Mobile Broadband", *Elsevier*, May 2011

[5] 3GPP, "E-UTRA: Physical Channels And Modulation (Release 11)", 3GPP TS 36.211, V11.3.0, Jun. 2013

[6] J. Mitola, III; G. Q. Jr Maguire, "Cognitive Radio: Making Software Radios More Personal", *IEEE Pers. Commun.,* vol. 6, no. 4, Aug. 1999, pp. 13-18

[7] S. Haykin, "Cognitive Radio: Brain-empowered Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[8] C. Cordeiro, K. Challapali, and D. Birru, "IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive Radios", *Journal of Commun.,*, vol. 1, no. 1, Apr. 2006

[9] A. Goldsmith, S. A. Jafar, I. Maric, S. Srinivasa, "Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective" *IEEE Proc. of the IEEE,* vol. 97, no. 5, pp. 894 - 914

[10] D. Calin, H. Claussen, and H. Uzunalioglu, "On Femto Deployment Architectures and Macrocell Offloading Benefits in Joint Macro-Femto deployments," *IEEE Commun. Mag.*, vol. 48, no. 1, pp. 26-32, Jan. 2010.

[11] J. D. Hobby and H. Claussen, "Deployment Options for Femtocells and their Impact on Existing Macrocellular Networks ", *Bell Labs Technical Journal,* vol. 13 , no. 4, Winter 2009

[12] M. A. Maddah-Ali, A. S. Motahari and A. K. Khandani, "Communication Over MIMO X Channels: Interference Alignment, Decomposition, and Performance Analysis", *IEEE Trans. Inf. Theory*, vol. 54, no. 8, Aug. 2008

[13] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 34253441, Aug. 2008.

[14] F. Negro, S. P. Shenoy, I. Ghauri and D. T. M. Slock, "Interference alignment feasibility in constant coefficient MIMO interference channels," *IEEE Proc. of Signal Process. Advances in Wireless Commun. (SPAWC),* Jun. 2010

[15] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone and Wei Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference", *IEEE J. Sel. Areas in Commun.,*, vol. 28 , no. 9, pp. 1380-1408, Oct. 2010

[16] O. El Ayach, S. W. Peters and R. W. Heath Jr., "The Practical Challenges of Interference Alignment", *IEEE Trans. Wireless Commun.,*, vol. 20, no. 1, pp. 35-42, Feb. 2013

[17] M. K. Karakayali, G. J. Foschini and R. A. Valenzuela, "Network Coordination for Spectrally Efficient Communications in Cellular Systems", *IEEE Trans. Wireless Commun.,* vol. 13, no. 4, pp. 56-61, Aug. 2006

[18] M. Hong, R. Sun, H. Baligh and Zhi-Quan Luo, "Joint Base Station Clustering and Beamformer Design for Partial Coordinated Transmission in Heterogeneous Networks" *IEEE J. on Sel. Areas in Commun.,* vol. 31, no. 2, Feb. 2013

[19] Seung-Jun Kim, S. Jain, G. B. Giannakis, "Backhaul-constrained Multi-cell Cooperation Using Compressive Sensing and Spectral Clustering", *IEEE Proc. of 13th Sig. Process. Advances in Wireless Comm.(SPAWC'12)*, pp. 65-69, Jun. 2012

[20] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh and R. W. Heath, "Networked MIMO with Clustered Linear Precoding" *IEEE Trans. on Wireless Commun.,* vol. 8 , no. 4, pp. 1910 - 1921, Apr. 2009

[21] A. Papadogiannis, A., D. Gesbert, E. Hardouin, "A Dynamic Clustering Approach in Wireless Networks with Multi-Cell Cooperative Processing", *Proc. of IEEE Int. Conf. on Commun. (ICC'08)*, pp. 4033, May 2008

[22] A. J. Fehske, P. Marsch and G. P. Fettweis, "Bit per Joule efficiency of cooperating base stations in cellular networks", *IEEE Proc. of Global Telecommun. (GLOBECOM) Workshops*, Miami, Dec. 2010

[23] Z. Hasan, H. Boostanimehr and V. K. Bhargava, "Green Cellular Networks: A Survey, Some Research Issues and Challenges", *IEEE Commun. Surveys & Tutorials*, vol. 13, no. 4, fourth quarter 2011

[24] J. Hoydis, M. Kobayashi and M. Debbah, "Green Small-Cell Networks", *IEEE Mag. Vehicular Techn.,* vol. 6, no. 1, pp. 37-43, Mar. 2011

[25] Y. Tevfik and H. Arslan "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications", *IEEE Commun. Surveys and Tutorials*, vol. 11, no. 1, 2009

[26] A. B. Gershman, N. D. Sidiropoulos, S. ShahbazPanahi, M. Bengtsson and B. Ottersten, "Convex Optimization-Based Beamforming", *IEEE Signal Process. Mag.,* vol. 27, iss. 3, pp. 62-75, May 2010

[27] F. Khan and Z. Pi, "mmWave Mobile Broadband (MMB): Unleashing the 3-300GHz Spectrum", *IEEE 34th Sarnoff Symposium*, May 2011

[28] J. N. Murdock, E. Ben-Dor, Q. Yijun, J. I. Tamir and T. S. Rappaport, "A 38 GHz Cellular Outage Study for an Urban Outdoor Campus Environment", *IEEE Proc. on Wireless Commun. and Networking Conf. (WCNC'12)*, pp. 3085 - 3090, Apr. 2012

[29] Y. Azar, G. N. Wong, K. Wang, R. Mayzus, J. K. Schulz, H. Zhao, F. Gutierrez, D. Hwang, T. S. Rappaport, "28 GHz Propagation Measurements for Outdoor Cellular Communications Using Steerable Beam Antennas in New York City", *IEEE Proc. of Int. Conf. on Commun. (ICC'13)*, Jun. 2013

[30] S. Rajagopal, S. Abu-Surra, Z. Pi and F. Khan, "Antenna Array Design for Multi-Gbps mmWave Mobile Broadband Communication", *IEEE Proc. of Global Telecommun. Conf. (GLOBECOM 2011)*, pp. 1-6, Dec. 2011

[31] K. Hosseini, J. Hoydis, S. ten. Brink and M. Debbah, "Massive MIMO and Small Cells: How to densify Heterogenous Networks," *Proc. of the Int Conf. on Commun. (ICC'13)*, Jun 2013, Budapest, Hungary,

[32] K. Eriksson, S. Shi, N. Vucic, M. Schubert, and E. G. Larsson, "Globally Optimal Resource Allocation for Achieving Maximum Weighted Sum-rate", *IEEE Proc. of Global Telecommun. (GLOBECOM)*, Dec. 2010, pp. 16.

[33] L. Tran, M. Hanif, A. Tolli, and M. Juntti, "Fast Converging Algorithm for Weighted Sum Rate Maximization in Multicell MISO Downlink," *IEEE Signal Process. Lett.,* vol. 19, no. 12, pp. 872875, Dec. 2012.

[34] Q. Zhang, C. He, and L. Jiang, "Achieving Maximum Weighted Sum-Rate in Multicell Downlink MISO Systems," *IEEE Commun. Lett.,* vol. 16, no. 11, pp. 18081811, Nov. 2012.

[35] L. Gallo, F. Negro, I. Ghauri and D. T. M. Slock, "Weighted Sum Rate Maximization in the Underlay Cognitive MISO Interference Channel", *IEEE Int. Symposium on Pers. Indoor and Mobile Radio*, pp. 661 - 665, Sept. 2011

[36] F. Negro, S. P. Shenoy, I. Ghauri, D. T. M. Slock, "On the MIMO interference channel", *IEEE Inf. Theory and Applications Workshop (ITA)*, Feb. 2010

[37] H.-J. Choi, S.-H.-Park, S.-R. Lee, and I. Lee, "Distributed Beamforming Techniques for Weighted Sum-rate Maximization in MISO Interfering Broadcast Channels" *IEEE Trans. Wireless Commun.,* vol. 11, no. 4, pp. 13141320, Apr. 2012.

[38] Q. Shi, M. Razaviyayn, Z.-Q. Luo and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-utility Maximization for a MIMO Interfering Broadcast Channel", *IEEE Trans. on Signal Process.,* vol. 59, no. 9, pp. 4331-4340, 2011

[39] P. Viswanath, D. N. C. Tse,"Sum Capacity of the Multiple Antenna Gaussian Broadcast Channel And Uplink-Downlink Duality", *IEEE Trans. on Inf. Theory*, vol. 49 , no. 8, Aug. 2003

[40] P. Viswanath N. Jindal and A. Glodsmith, "Duality Achievable Rates and Sum-rate Capacity of Gaussian MIMO Broadcast Channels", *IEEE Trans. on Inf. Theory*, vol. 49, no. 8, pp. 2658-2668, Oct. 2003

[41] W. Yu, "Uplink Downlink Duality Via Minimax Duality", *IEEE Trans. on Inf. Theory*, vol. 52, no. 2, Feb. 2006

[42] D. Cai, T. Quek and C. W. Tan, "A Unified Analysis of Max-min Weighted SINR for MIMO Downlink System", *IEEE Trans. on Signal Process.,* vol. 59, no. 8, pp. 3850-3862, Aug. 2011

[43] H. Boche and M. Schubert, "Resource Allocation in Multiantenna Systems, Achieving Max-min Fairness by Optimizing a Sum of Inverse SIR", *IEEE Trans. on Signal. Process.,* vol. 54, no. 6, pp. 1990-1997, June 2006

[44] M. Razaviyayn , M. Hong, , Z.-Q. Luo, "Linear Transceiver Design for a MIMO Interfering Broadcast Channel Achieving Maxmin Fairness", *Elsevier Signal Process.,* vol. 92, no. 12, Dec. 2013, pp. 3327-2240

[45] K. T. Phan, N. D. Sidiropoulos and C. Tellambura, "Spectrum Sharing in Wireless Networks via QoS-Aware Secondary Multicast Beamforming", *IEEE Trans. on Signal Process.,* vol. 57, no. 6, pp. 2323-2334, June 2009

[46] D. Hammarwall, M. Bengtsson and B. Ottersten, "Acquiring Partial CSI for Spatially Selective Transmission by Instantaneous Channel Norm Feedback", *IEEE Trans on Signal Process.,*, vol. 56, no. 3,

[47] E. Bjornson, D. Hammarwall and B. Ottersten, "Exploiting Quantized Channel Norm Feedback Through Conditional Statistics in Arbitrarily Correlated MIMO Systems", *IEEE Trans. on Signal Process.,* vol. 57, no. 10, pp. 4027-4041, June 2009

[48] H. Boche and M. Schubert, "The Structure of General Interference Functions and Applications", *IEEE Trans. on Inf. Theory,* vol. 54, no. 11, pp. 4980-4990, Nov. 2008

[49] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[50] W. Ai, Y. Huang, and S. Zhang, "New Results on Hermitian Matrix Rank-one Decomposition," *Math. Program.: Series A*, vol. 128, no. 12, pp. 253-283, June 2011

[51] D. Bersekas, "Nonlinear Programming: 2nd Edition", Apr. 2004

[52] A. Ben-Tal, A. Nemirovski, "Lectures on Modern Convex Optimization", *http* : $//www2.isye.gatech.edu/\ nemirovs/Lect_{M}odConvOpt.pdf$

[53] R. A. Horn and C. Johnson, "Matrix Analysis", *Cambridge University Press*, 2007.

[54] A. Nemirovski, K. Roos, and T. Terlaky, "On Maximization of Quadratic Forms over Intersection of Ellipsoids with Common Center," *Math. Program., ser. A*, vol. 86, pp. 463-473, 1999.

[55] V. Jeyakumar, A. M. Rubinov, Z. Y. Wu, "Non-convex Quadratic Minimization Problems with Quadratic Constraints: Global Optimality Conditions" *Math. Program.,* vol. 110, no. 3, pp. 521-541, Sep. 2007

[56] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit Beamforming for Physical-layer Multicasting,", *IEEE Trans. on Signal Process.,* vol. 54, no. 6, pp. 22392251, 2006.

[57] Y. Huang and D. P. Palomar, "Randomized Algorithms for Optimal Solutions of Double-Sided QCQP With Applications in Signal Processing", *IEEE Trans. on Signal Process.*, vol. 62, no. 5, Feb. 2014, pp. 1093 - 1108

[58] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," *Handbook of Antennas in Wireless Commun.*, L. Godara, Editor, Boca Raton, CRC, FL, Aug. 2001.

[59] R. D. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems", *IEEE J. on Sel. Areas in Commun.,* vol. 13, no. 7, Sep. 1995

[60] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint Optimal Power Control and Beamforming in Wireless Networks Using Antenna Arrays," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1312-1324, Oct. 1998.

[61] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Transmit Beamforming and Power control for Cellular Wireless Systems," *IEEE J. on Sel. Areas in Commun.*, vol. 16, no. 8, pp. 1437-1450, Oct. 1998.

[62] E. Visotsky, U. Madhow, "Optimum Beamforming Using Transmit Antenna Arrays", *IEEE Proc. of the 49th Vehicular Tech. Conf. (VTC'99)*, Houston, Jul. 1999

[63] M. Schubert and H. Boche, "QoS-based Resource Allocation and Transceiver Optimization," *Foundations and Trends in Communications and Information Theory,* vol. 2, no. 6, 2006

[64] M. Schubert and H. Boche, "Solution of the Multiuser Downlink eamforming problem with individual SINR constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18-28, Jan. 2004.

[65] H. Boche, "A Superlinearly and Globally Convergent Algorithm for Power Control and Resource Allocation with General Interference Functions," *IEEE Trans. on Networking*, vol. 16, no. 2,Apr. 2008

[66] M. Schubert and H. Boche, "A Generic Approach to QoS based transceiver optimization", *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1557-1566, Aug. 2007

[67] A. Wiesel; Y. C. Eldar and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers", *IEEE Trans on Signal Processing*, Vol. 56, Iss. 1, Jan. 2006, pp. 161-176

[68] D. Hammarwall, M. Bengtsson and B. Ottersten, "On downlink Beamforming with Indefinite Shaping Constraints", *IEEE Trans. Signal Process.*, vol. 54, pp. 3566-3580, Sep. 2006

[69] G. Pataki, "On the Rank of Extreme Matrices in Semidefinite Programs and the Multiplicity of Optimal Eigenvalues," *Math. Operations Res.,* vol. 23, no. 2, pp. 339-358, 1998

[70] Y. Huang and D. Palomar, "A Dual Perspective on Separable Semidefinite Programming With Applications to Optimal Downlink Beamforming", *IEEE Trans. Signal Process.,* vol. 58, no. 8, pp. 4254-4271, Aug. 2010

[71] Y. Huang and D. Palomar, "Rank-constrained Separable semidefinite programming with applications to optimal beamforming", *IEEE Trans. Signal Process.,*, vol. 58, pp. 664-678, Feb. 2010

[72] A. Konar, N. D. Sidiropoulos, "Hidden Convexity in QCQP with Toeplitz-Hermitian Quadratics",*IEEE Signal Process. Lett.* vol. 22, no. 10, pp. 1623-1627, Apr. 2015

[73] K. Cumanan, R. Krishna, Z. Xiong, and S. Lambotharan, "SINR Balancing Technique and its Comparison to Semidefnite Programming based QoS Provision for Cognitive Radios," *IEEE Proc. Veh. Tech- Conf. (VTC)*, Barcelona, Spain, Apr. 2009.

[74] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR Balancing Technique for Downlink Beamforming in Cognitive Radio Networks," *IEEE Signal Proc. Lett.,* vol. 17, no. 2, pp. 133 - 136, Feb. 2010.

[75] X. Fu, J. Wang and S. Li, "Joint Power Management and Beamforming for Base Stations in Cognitive Radio systems", *IEEE Int. Symp. on Wire. Commun. Systems (ISWCS)*, pp. 403-407, Sept. 2009

[76] M. Pesavento, D. Ciochina, and A.B. Gershman, "Iterative Dual Downlink Beamforming for Cognitive Radio Networks," *Proc. of the 5th Int. Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2010)*, June 9-11, Cannes, France.

[77] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.,* vol. 38, pp. 49-95, 1996.

[78] J. Mattingley and S. Boyd, "Real-time Convex Optimization in Signal Processing," *IEEE Signal Process. Mag.,* vol. 27, no. 3, pp. 50-61, 2010

[79] Z. -Q. Luo, W. -K. Ma, A. M. -C. So, Y. Ye, and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems: From its Practical Deployments and Scope of Applicability to Key Theoretical Results," *IEEE Signal Process. Mag.,* vol. 27, no. 3, pp. 20-34, May 2010.

[80] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," *Elsevier, Signal Process.*, vol. 50, no. 1-2, pp. 57-69,Apr. 1996

[81] E. Karipidis,"Far-Field Multicast Beamforming for Uniform Linear Antenna Arrays," *IEEE Trans. on Signal Process.*, vol. 55, no. 10, pp. 4916-4927

[82] "$http://people.virginia.edu/\ jlr5m/Papers/FejerRiesz.pdf$"

[83] E. Matskani, N. Sidiropoulos, Z. Q. Luo and L Tassiulas "Convex Approxiamtion Techniques for Joint Multiuser Downlink Beamforming and Admission Control," *IEEE Trans. on Wireless Commun.*, vol. 7, no 7, pp. 2682-2693, Jul. 2008.

[84] I. Mitliagkas, N. D. Sidiropoulos, A. Swami "Convex Approximation-Based Joint Power and Admission Control for Cognitive Underlay Networks", *IEEE Wireless Commun. and Mobile Computing Conf., IWCMC '08* , Aug. 2008

[85] K. Hamdi, W. Zhang, K. W. Letaief, "Joint Beamforming and Scheduling in Cognitive Radio Networks ," *Proc. Global Telecommun. Conf. (GLOBECOM)*, pp. 2977-2981, Nov. 2007.

[86] K. Cumanan, R. Krishna, L. Musavian, and S. Lambotharan, "Joint Beamforming and User Maximization Techniques for Cognitive Radio Networks Based on Branch and Bound Method ," *IEEE Trans. on Wireless Commun.*, vol. 9, no 10, pp. 3082-3092, Oct. 2010.

[87] D. Ciochina, M. Pesavento, "Joint User Selection and Beamforming in Interference Limited Cognitive Radio Networks," *Proc. of the European Signal Process. Conf. (EUSIPCO'12)*, August 2012, Bucharest, Romania,

[88] Y. Cheng and M. Pesavento, "Joint Discrete Rate Adaptation and Downlink Beamforming Using Mixed Integer Conic Programming", *IEEE Trans. on Signal. Process.*, vol. 63, no. 7, Jan. 2015

[89] Y. Cheng and M. Pesavento, "Dynamic Rate Adaptation and Multiuser Downlink Beamforming using Mixed Integer Conic Programming", *IEEE Proc. of European Signal Process. Conf. (EUSIPCO'12),* Aug. 2012, Bucharest, Romania

[90] B. Hassibi and B. M. Hochwald, "How Much Training is Needed in Multiple Antenna Wireless Links?," *IEEE Trans. on Inf. Theory*, vol. 49, no. 4, pp. 951-963,April 2003.

[91] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the Value of Limited Feedback for MIMO Channels?," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54-59, Oct. 2004.

[92] H. Cox, R. M. Zeskind, and M. H. Owen, "Robust Adaptive Beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1365-1376, Oct. 1987

[93] 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.*, vol. 51, no. 2, pp. 313-324, Feb. 2003

[94] J. Li, P. Stoica and Z. Wang "On Robust Capon Beamforming and Diagonal Loading", *IEEE Trans. Sig. Process.,*, vol. 51, no. 7, pp. 1702 - 1715, Jul. 2003

[95] R. Lorenz and S. P. Boyd, "Robust Minimum Variance Beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684-1696, May 2005.

[96] S. Shahbazpanahi, A. B. Gershman, Z. -Q. Luo, and K. M. Wong, " Robust Adaptive Beamforming for General-rank Signal Models", *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257-2269, Sep. 2003

[97] S. 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.*, vol. 56, pp. 5719-5724, Nov. 2008.

[98] O. Besson and S. Bidon, "Bayesian robust beamforming based on random steering vector with Bingham prior distribution", *IEEE Proc of ICASSP*, May 2013

[99] Y. C. Eldar and N. Merhav, "A competitive minimax approach tp robust estimation of random parameters", *IEEE Trans. Signal Processing,* vol. 52, pp. 1931-1946, Jul. 2004

[100] Y.Guo and B. C. Levy, "Robust MSE equalizer design for MIMO communications systems in the presence of model uncertainties", *IEEE Trans. Singal Processing*, vol. 53, pp. 168-181, Jan. 2005

[101] 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. Siganl Processing*, vol. 53, pp. 3081-3090, Aug. 2005

[102] J. Wang, M Bengtsson, B. Ottersten, D. Palomar, " Robust maximin MIMO precoding for arbitrary convex uncertainty sets" *in Proc. IEEE ICASSP*, pp. 3045-3048, March 2012

[103] 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 Processing*, vol. 54, no. 1, pp. 346-360, Jan. 2006.

[104] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink Multiuser MISO systems," *IEEE Trans. in Signal Process.*, vol. 57, no. 2, Feb. 2009

[105] E. Song, Q. Shi, M. Sanjabi, R. Sun and Z-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: when is semidefinite programming relaxation tight?" *EURASIP Journal on Wireless Communications and Networking*, Aug. 2012

[106] J. Wang and D. Palomar, "Worst case robust MIMO Transmission with imperfect channel knwoledge", *IEEE Trans. Signal Processing*, vol. 57, no. 8, Aug. 2008

[107] J. Wang and M. Payaro, "Is Transmit Beamforming Robust?, " *in Proc. IEEE Int. Conf. on Commun. (ICC 2011)*, pp 1-5, Jun 2011

[108] M. B. Shenouda and T. N Davidson, "Conic convex formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics in Signal Process.*, vol. 1, pp. 714-724, Dec. 2007.

[109] M. B. Shenouda and T. N Davidson, "Nonlinear and Linear Broadcasting With QoS Requirements: Tractable Approaches for Bounded Channel Uncertainties", *IEEE Trans. Sig. Processing*, Vol. 57, Iss. 5, Pp. 1936 - 1947

[110] D. Ciochina, M. Pesavento, and K. M. Wong, "Worst Case Robust Beamforming on the Riemannian Manifold", *IEEE 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013)* , Vancouver, Canada, May 2013

[111] L. Xu, K. M. Wong, J. K. Zhang, D. Ciochina, and M. Pesavento, "A Riemannian Distance for Robust Downlink Beamforming", *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC'13)* , Darmstadt, Germany, June 2013

[112] I. Wajid, Y. C. Eldar, and A. B. Gershman, "Robust downlink beamforming using covariance channel state information,", *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'09.*, April 2009, pp. 2285-2288.

[113] K. Cumanan, R. Krishna, V. Sharma, and S. Lambotharan, "Robust interference control techniques for multiuser cognitive radios using worst-case performance optimization," in *Proc. 42nd Asilomar Conference on Signals, Systems and Computers*, pp. 378-382, Oct. 2008.

[114] I. Wajid, M. Pesavento, Y. C. Eldar, and A. B. Gershman, "Robust downlink beamforming for cognitive radio networks,", *IEEE Global Telecommunications Conference GLOBECOM'10*, Dec. 2010.

[115] E. A. Gharavol, Y. -C. Liang, and K. Mouthaan, "Robust downlink beamforming in multiuser MISO cognitive radio networks with imperfect channel-state information," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 2852-2860, July 2010

[116] G. Zheng, K. K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Sig. Processing*, pp. 4871-4881, Dec. 2009.

[117] I. Wajid, H. Nikolaeva and M. Pesavento, "Iterative robust downlink beamforming in cognitive radio networks," *Sixth International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications CROWNCOM'11*, June 2011, pp. 375-379.

[118] I. Wajid, M. Pesavento, Y. C. Eldar, and D. Ciochina, "Robust downlink beamforming with partail channel state information for conventional and cognitive radio networks," *Proc IEEE Trans. on Signal Processing*, vol. 60, Issue 2, Aug. 2013

[119] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman,"Robust downlink beamforming based on outage probability specifications," *IEEE Trans. Wireless Commun.,* vol. 6, no. 10, pp. 34983503, Oct. 2007

[120] K.-Yu Wang, A. Man-Cho So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage Constrained Robust Transmit Optimization for Multiuser MISO Downlinks: Tractable Approximations by Conic Optimization," *IEEE Trans. Sig. Processing*, Vol. 62, Iss. 21, pp. 5690 -5705, Sep. 2014

[121] K. Law, I Wajid and M. Pesavento, " Robust downlink beamforming in multi-group mulicasting using trace bounds on the covariance mismatches" *in Proc. IEEE ICASSP*, pp. 3045-3048 , April 2011

[122] V. Sharma, I. Wajid, A. Gershman, H. Chen and S. Lambotharan, 'Robust downlink beamforming using positive semidefinite covariance constraints," *IEEE Int. ITG Workshop on Smart Antennas, WSA*, Feb. 2008

[123] Y. Cheng, "Joint Downlink Beamforming and Discrete Resource Allocation Using Mixed-Integer Programming, " Dissertation, TU Darmstadt, Darmstadt, 2012

[124] H. Chen and A. B. Gershman, "Robust adaptive beamforming for general-rank signal models with positive semidefinite constraints," *IEEE Int. Conf. Acoustics, Speech and Signal Processing, (ICASSP'08)*, Apr. 2008, pp. 2341-2344

[125] A. Adhikary, J. Nam, J. -Y. Ahn, and G. Caire, "Joint spatial division and multiplexing," *arXiv preprint arXiv:1209.1402*, 2012

[126] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A Coordinated Approach to Channel Estimation in Large-scale Multiple-antenna Systems," *IEEE Journal on Selected Areas Commun., Special Issue on Large Scale Antenna Systems* , Feb. 2013

[127] V. Raghavan, S. Hanly and V. Veeravalli, " Statistical Beamforming on the Grassman Manifold for the two user broadcast channel", *IEEE Trans. Inf. Theory*, vol. 59, no- 10, Oct. 2013

[128] A. Castano-Martinez, F. Lopez-Blazquez, "Distribution of a sum of weighted non-central chi-square variables", *Springer Link, Test*, vol. 14, Iss. 2, pp. 397-415, , Dec. 2005

[129] A. Stuart "J. K. Kendall's Advanced Theory of Statistics, Vol. 2A: Classical Inference & the Linear Model", *Oxford University Press*, New York, 1999

[130] O. Tuzel, F. Porikli, P. Meer, "Pedestrian Detection via Classification on Riemannian Manifolds", *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, Vol. 30 , Iss. 10, 2008

[131] A. Barachant, S. Bonnet, M. Congedo and C. Jutten, "Riemannian geometry applied to BCI classification" Proc. of 9th International Conference Latent Variable Analysis and Signal Separation (LVA/ICA 2010), 2010

[132] S. T. Smith "Covariance, subspace, and intrinsic Cramer-Rao bounds", *IEEE Trans on Signal Processing*, Vol. 53, Iss. 5, May 2005

[133] S. Gallot, D. Hulin and J. Lafontaine, *Riemannian Geometry*, Springer 3rd Ed., 2004.

[134] Y. Li, K. M. Wong and H, de Bruin "Electroencephalogram signals classification for sleepstate decision - A Riemannian geometry approach," *IET Signal Processing*, vol. 6 , no. 4, pp 288-299

[135] Y. Li, K. M. Wong and H. de Bruin, "EEG signal classification based on a Riemannian distance measure", *Proc. IEEE Science and Technology for Humanity (TIC-STH)*, pp 268-273, Toronto 2009

[136] Y. Li and K. M. Wong, "Riemannian Distances for Signal Classification by Power Spectral Density", *IEEE Journal of Selected topics in Signal Processing*, vol. 7, no. 4, Aug. 2013

[137] P. A. Absil, R. Mahony and R. Sepulchre, "Optimization Algorithms on Matrix Manifolds' ', Princeton University Press, 2007

[138] J. H. Manton, " Optimization algorithms exploiting unitary constraints ", *IEEE Transactions on Signal Processing*, vol. 50, Iss. 3, pp. 635 - 650, Mar. 2002

[139] S. Amari, N. Mitra, "Closed-Form Expressions for Distribution of Sum of Exponential Random Variables", *IEEE Trans. Rel.*, vol 37, no. 1,pp 519-522, Dec 1997

[140] E. M. Scheuer, "Reliability of m-out-of-n System When Componenet Failure Induces Higher Failure Rates in Survivors", *IEEE Trans. Reliability.*, vol 37, no. 1,pp 73-74, Apr. 1988

[141] M. Biguesh and A. B. Gershman, "Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals", *IEEE Trans. in Signal Process.*, vol. 54, no. 3, May. 2006

Curriculum Vitae

Education

- since 03/2009, Ph.D. Candidate, Communication Systems Group, Darmstadt University of Technology, Darmstadt, Germany.
- 10/2003–08/2008, Engineering Degree, Faculty of Electronics, Telecommunications and Information Technology, Univ. "Politehinca" Bucuresti
  - specialization: Mobile Communications and Satellites,
  - final thesis: Multiband Excitation Model for the Parametric Representation of Speech Signals

Work Experience

- 03/2009-03/2015, Research Associate, Communication Systems Group, Darmstadt University of Technology, Darmstadt, Germany.
- 07/2008-03/2009, Access Technologies Engineer, Orange, Romania
- 10/2007-07/2008, Access Technologies Trainee, Orange, Romania
- 07/2007-08/2007, Industrial Attachment, Infineon Technologies, München, Germany

Teaching Experience

- Teaching Assistent for Information Technology I, WiSe 2012/2013 and WiSe 2013/2014
- Teaching Assisten for Smart Antenna Project Seminar, SS 2011 and SS 2014

Distinctions

- Co-author of the article, which received the Best Paper Award at the 5th International Conference on Cognitive Radio Oriented Networks and Communications (CrownCom'10)
- Scolarship from Orange, Romania
- Solarship from Infineon Technologies, Romania