

Analysis and Optimization of Wireless Cellular Multi-User Multiple-Input Multiple-Output Communication Systems

Von der Fakultät Informatik, Elektrotechnik und
Informationstechnik der Universität Stuttgart zur Erlangung der
Würde eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte
Abhandlung

Vorgelegt von
Venkata Krishna Chitti
Hyderabad (India)

Hauptberichter:	Prof. Dr.-Ing. Stephan ten Brink
Mitberichter:	Prof. Dr.-Ing. Andreas Kirstädter
Tag der mündlichen Prüfung:	19. November 2015

Institut für Nachrichtenübertragung der Universität Stuttgart
2015

Acknowledgements

I am ever indebted to my parents for the numerous sacrifices they made for me.

At the INÜ, I firstly thank Professor Joachim Spiedel for providing me a rare opportunity to carry out the research by providing necessary financial assistance and infrastructure. Next, I thank Professor Stephan ten Brink who took over at the last phase of my research and agreed to be the first supervisor of my work. All my INÜ colleagues have been more than friendly to me. This made my tenure at the institute enjoyable.

I thank everyone who has contributed in some way or the other to my journey since childhood. Above all, I must realize that an ever pervading source that keeps me connected from within has been guiding me throughout.

Contents

Abbreviations and Notations	v
Problem Definitions	xi
Abstract	xiii
Kurzfassung	xiv
1 Introduction	1
1.1 Basics of Wireless Communication	3
1.1.1 Fading	3
1.1.2 Frequency Reuse	4
1.1.3 Multiple Access	5
1.1.4 Multiple-Input Multiple-Output	6
1.2 Multi-User MIMO	7
1.2.1 Fairness	11
1.2.2 Transceiver design	12
2 Mathematical Techniques for Optimization	17
2.1 Constraints	17
2.2 Problem Formulation	19
2.3 Problem Solving	21
2.3.1 Base Station Association	21
2.3.2 Power Allocation and Beamforming	23

3	Uplink Resource Allocation under perfect-CSI	27
3.1	Algorithmic Framework	27
3.2	Rate Allocation	29
3.2.1	Iterative Power Allocation	30
3.2.2	Iterative Base Station Association	30
3.2.3	Single Stage Formulation	30
3.3	Power Control	33
3.3.1	Feasibility	34
4	Uplink Power Control under statistical-CSI and imperfect-CSI	41
4.1	Power Control under statistical-CSI	42
4.2	Value-at-Risk and Conditional Value-at-Risk	43
4.3	Extreme Value Theory	44
4.3.1	Generalized Extreme Value Distribution	44
4.3.2	Generalized Pareto Distribution for CVaR	48
4.4	Power Control under imperfect-CSI	50
5	Downlink Performance Analysis under perfect-CSI	55
5.1	Uplink-Downlink Duality	55
5.2	Problem Reformulation	57
5.3	Downlink Rate and Power Allocation	58
6	Conclusion	65
	Publications	67

Abbreviations and Notations

3GPP	3rd Generation Partnership Project
AAS	Active Antenna System
ACO	Ant Colony Optimization
AWGN	Additive White Gaussian Noise
BA	Bernstein Approximation
BB	Branch and Bound
BC	Broadcast Channel
BD	Bender Decomposition
BER	Bit Error Rate
BLAST	Bell-Labs Layered Space Time Architecture
BI	Bernstein-type Inequality
BF	Beamforming
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BS	Base Station
BSA	Base Station Association
cdf	Cumulative Distribution Function
CoMP	Coordinated Multi-point transmission reception
CCI	Cochannel Interference
CDMA	Code Division Multiple Access
CSB	Coordinated Scheduling and Beamforming
CSI	Channel State Information
CVaR	Conditional Value-at-Risk
DC	Difference in Convex
DoA	Domain-of-Attraction
DoF	Degrees-of-Freedom
DPC	Dirty Paper Coding
DL	Downlink
DWD	Danzig-Wolfe Decomposition
EVT	Extreme Value Theory
FDMA	Frequency Division Multiple Access
GA	Gaussian Approximation
GAP	Generalized Assignment Problem
GeA	Genetic Algorithm

GEVD	Generalized Extreme Value Distribution
GEVP	Generalized EigenValue Problem
GFDM	Generalized Frequency Division Multiplexing
GP	Geometric Programming
GPD	Generalized Pareto Distribution
HetNets	Heterogeneous Networks
i.i.d.	independent and identically distributed
IPM	Interior Point Method
IM	Interference Management
IP	Integer Programming
IMT-A	International Mobile Telecommunications Advanced
ITU-R	International Telecommunication Union Radiocommunication Sector
JP	Joint Processing
KKT	Karash-Kuhn-Tucker
KP	Knapsack Problem
LICQ	Linear Independent Constraint Qualification
LP	Linear Programming
LR	Lagrange Relaxation
LSE	Least Squared Error
LTE-A	Long Term Evolution Advanced
MAC	Multiple Access Channel
MAI	Multiple Access Interference
MC	Multi-carrier
MINLP	Mixed Integer Non-Linear Programming
MLE	Maximum Likelihood Estimator
MMKP	Multi-dimension Multi-Choice Knapsack Problem
MSE	Mean Squared Error
MLSE	Minimum Least Squared Error
MMSE	Minimum Mean Squared Error
MUD	Multi-User Diversity
MUI	Multi-User Interference
MU-MIMO	Multi-User Multiple-Input Multiple-Output
MVDR	dsdfsdfs
NDMA	Network assisted Diversity Multiple Access
NLP	Non-Linear Programming
NOMA	Non-orthogonal Multiple Access
NP	Non-deterministic Polynomial-time
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
pdf	Probability Density Function
PoT	Peaks over Threshold
PrCons	Probabilistic Constraints

PA	Power Allocation
PAPR	Peak to Average Power Ratio
PC	Power Control
PF	Proportional Fairness
PSO	Particle Swarm Optimization
QoS	Quality of Service
r.v.	Random Variable
RA	Resource Allocation
RR	Round Robin
SA	Simulated Annealing
SC-OFDMA	Single Carrier OFDMA
SDR	Semidefinite Relaxation
SDMA	Space Division Multiple Access
SDP	Semidefinite Programming
SGINR	Signal to Generated Interference Noise Ratio
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-Noise-Ratio
SISO	Single-Input Single-Output
SNR	Signal to Noise Ratio
SO	Successive Optimization
SOCP	Second Order Cone Programming
SON	Self Organizing Networks
SP	Signomial Programming
SQP	Sequential Quadratic Programming
STBC	Space Time Block Code
STTC	Space Time Trellis Code
SUS	Semi-orthogonal User Selection
TDMA	Time Division Multiple Access
THP	Tomlinson Harashima Precoding
TS	Tabu Search
TSD	Transmit Selection Diversity
UE	User Equipment
UL	Uplink
VaR	Value-at-Risk
VUL	Virtual-Uplink
ZF	Zero-Forcing
M	number of BSs
N	number of UEs
T	number of antenna elements at the BS
R	number of antenna elements at the UE, $R = 1$ assumed throughout
L	number of simultaneous data streams or spatial layers for a UE,

	$L = 1$ assumed throughout
α_{ij}	0 – 1 integer BSA variable, no macro-diversity throughout
$\boldsymbol{\alpha}_j$	BSA vector of UE $_j$
$\boldsymbol{\alpha}$	overall BSA matrix of all the UEs
$\underline{\boldsymbol{\alpha}}_i$	i^{th} row of $\boldsymbol{\alpha}$
p_j	UL transmit power variable of UE $_j$
q_{ij}	DL transmit power variable of UE $_j$ from BS $_i$
\mathbf{q}_j	DL power vector for UE $_j$ across all the BSs
\mathbf{Q}	overall DL power vector of all the UEs across all the BSs
\mathbf{q}	active DL power vector of all the UEs across all the BSs
$\underline{\mathbf{q}}_i$	i^{th} row of \mathbf{Q}
\mathbf{h}_{ij}	UL channel vector from UE $_j$ to BS $_i$
\mathbf{h}_{ij}^H	DL channel vector from BS $_i$ to UE $_j$
\mathbf{u}_{ij}	UL receive BF vectors of UE $_j$ at BS $_i$
\mathbf{u}_{ij}^H	DL transmit BF vectors at BS $_i$ for UE $_j$
\mathbf{U}	overall BF matrix of all the UEs across all the BSs
a_j	UL transmit symbol of UE $_j$
\hat{a}_j^{UL}	estimated UL transmit symbol of UE $_j$
a_{ij}	DL transmit symbol of UE $_j$ from BS $_i$
\hat{a}_{ij}^{DL}	estimated DL transmit symbol of UE $_j$
$\bar{\mathbf{z}}_j$	UL noise vector for UE $_j$ at a BS
$\bar{\sigma}_j^2$	UL AWGN noise variance per-antenna element for UE $_j$ at a BS
\underline{z}_j	DL noise variable at UE $_j$
$\underline{\sigma}_j^2$	DL AWGN noise variance per-antenna element at UE $_j$
\mathcal{P}	A primal problem
\mathcal{D}	Mathematical dual problem of \mathcal{P}
\mathcal{L}	Lagrangian of \mathcal{P}
N_2	total number of constraints across all the UEs in \mathcal{P}
\mathbf{z}_j	primal vector of UE $_j$
\mathbf{c}_j	constraint vector of UE $_j$
\mathbf{s}_j	slack vector for the \mathbf{c}_j
t	iteration index
δ	step length in IPM
Δ	search directions in IPM
\mathbf{F}	KKT conditions in vector form
\mathbf{J}	Jacobian of \mathbf{F} w.r.t. the primal and dual variables
μ	complementarity measure in perturbed KKT equation
τ	parameter for δ calculation
P_{\max}	maximum UL transmit power of each UE
\bar{P}_i	maximum DL transmit power at BS $_i$
CSI_p	perfect-CSI
$\bar{\gamma}_j$	UL SINR of UE $_j$ under CSI_p

\bar{r}_j	UL rate of UE _j under CSI _p
$\underline{\gamma}_j$	DL SINR at UE _j under CSI _p
\underline{r}_j	DL rate at UE _j under CSI _p
$\mathbf{1}_M$	$(M \times 1)$ vector with all elements equal to unity
r_{th}	QoS threshold for each UE on UL and DL
$\Re\{\cdot\}$	Real part of a complex number
$\Im\{\cdot\}$	Imaginary part of a complex number
(\circ)	Hadamard product
$(\cdot)^H$	Hermitian operator
$(\cdot)^T$	Transpose operator
$(\cdot)^{-1}$	Inverse operator
$(\cdot)^*$	complex conjugate operator
$\max(\cdot, \cdot)$	maximum operator
$\min(\cdot, \cdot)$	minimum operator
$\text{tr}(\cdot)$	trace operator
$\text{diag}(\cdot)$	diagonal operator
$\text{blkdiag}(\cdot)$	block diagonal operator
$\text{vec}(\cdot)$	vectorization operator
$\ \cdot\ _1$	ℓ_1 -norm operation
$\ \cdot\ _2$	ℓ_2 -norm operation
$\log_2(\cdot)$	logarithm base-2
$\ln(\cdot)$	natural logarithm
$\text{erf}^{-1}(\cdot)$	Inverse error function
$\exp(\cdot)$	Exponential function
$\text{arg-max}(\cdot)$	arg-max operator, chooses the argument that maximizes the function
$(\cdot, 0)_+$	non-negative operator, chooses 0 or non-negative argument
K	number of i.i.d. random variables in a block for EVT approach
B	number of blocks of i.i.d. random variables for EVT approach
η	confidence interval for the chance constraints
$\mathcal{N}(\cdot, \cdot)$	normal distribution
$\mathcal{CN}(\cdot, \cdot)$	complex normal distribution
ξ	shape parameter of GEVD and GPD
\bar{l}	location parameter of GEVD
\bar{s}	scale parameter of GEVD
\underline{l}	location parameter of GPD
\underline{s}	scale parameter of GPD
CSI _s	statistical-CSI
$\bar{\gamma}_j^{(s)}$	Statistical UL SINR of UE _j with no macro-diversity under CSI _s
$\bar{r}_j^{(s)}$	Statistical UL rate of UE _j with no macro-diversity under CSI _s
CSI _i	imperfect-CSI
$\bar{\gamma}_j^{(i)}$	Imperfect UL SINR of UE _j with no macro-diversity under CSI _i
$\bar{r}_j^{(i)}$	imperfect UL rate of UE _j with no macro-diversity under CSI _i

θ	A large positive constant in $P3_f$
\mathcal{F}	A single function that evaluates VaR and CVaR together
$Q(\cdot, \cdot)$	Marcum Q -function
$I_0(\cdot)$	0^{th} order modified Bessel's function of first kind
b	$2^{r_{th}} - 1$
g_{ijk}	effective channel gain in the MUI term from UE $_k$ at BS $_i$ caused to UE $_j$
\mathcal{I}_{ij}	MUI power for UE $_j$ at BS $_i$
\mathbf{I}	Identity matrix
χ_2^2	central chi-squared distributed r.v.
nc- χ_2^2	non-central chi-squared distributed r.v.
$F_Y(\cdot)$	cdf of r.v. Y
$F_Y^{-1}(\cdot)$	quantile function of r.v. Y or inverse CDF of r.v. Y
$\Psi_Y(\cdot)$	cdf of r.v. Y
$\phi(\cdot)$	A loss function
$\beta_j^{(s)}$	CVaR of UE $_j$
$\zeta_j^{(s)}$	VaR of UE $_j$
$\bar{\zeta}_j^{(s)}$	optimal $\zeta_j^{(s)}$
ν	ℓ_1 -norm parameter for feasibility handle in PC

Problem Definitions

P1	UL sum-rate maximization with QoS under CSI_p
$P1_b$	BSA sub-problem of P1
$P1_{bn}$	NLP formulation of $P1_b$
$P1_{bs}$	SDP formulation of $P1_b$
$P1_p$	PA sub-problem of P1
$P1_s$	Single stage NLP formulation of P1
P2	DL sum-rate maximization with QoS under CSI_p
$P2_b$	BSA sub-problem of P2
$P2_p$	PA sub-problem of P2
$P2_v$	NLP formulation of P2 with duality constraints
P3	UL sum-power minimization with QoS under CSI_p
$P3_b$	BSA sub-problem of P3
$P3_p$	PA sub-problem of P3
$P3_s$	Single stage NLP formulation of P3
$P3_f$	$P3_s$ with ℓ_1 -norm feasibility handle
$P3_c$	$P3_s$ under CSI_s
$P3_w$	Worst case performance problem of $P3_s$ under CSI_s
$P3_i$	$P3_s$ under CSI_i
P4	DL sum-power minimization with QoS under CSI_p
$P4_v$	VUL formulation of P4
$P4_{vb}$	BSA sub-problem of $P4_v$
$P4_{vp}$	PA sub-problem of $P4_v$
$P4_{vm}$	max-min sub-problem equivalent of $P4_{vp}$
P5	UL max-min problem under CSI_p

Abstract

Multi-User Multiple-Input Multiple-Output (MU-MIMO) communication systems are an attractive solution to the increasing demand in wireless services, limited availability of resources and increasing user density. MIMO achieves an increased spectral efficiency and an almost reliable link-quality without the need to expend other valuable network resources such as bandwidth and power. MIMO exploits the spatial separation among the network entities and the spatial richness in the environment to achieve this objective. The effect of shared nature of the wireless channel, i.e., multiple transmissions on the same resource unit at the same instance, is perceived as the cochannel interference. This is severe at the cell-edges. Hence the network performance is interference-limited and the need to optimally allocate resources while providing a reliable service to the network entities arises.

This thesis focuses on the development of iterative algorithms for the link-level physical layer abstraction for the task of Interference Management (IM) and Resource Allocation (RA) in conventional multi-cell multi-user wireless systems. For this, well understood mathematical methods with some modifications and new optimization principles are applied. Analysis is carried out separately on both the Uplink and the Downlink. With a required quality-of-service in a multi-cell MU-MIMO set-up, the problems of rate maximization and power control under different circumstances and assumptions are considered. These include analysis in the presence of perfect Channel State Information (CSI), statistical-CSI, imperfect-CSI and probabilistic constraints.

To achieve the task of RA, an equivalent mathematical optimization problem, usually a Mixed Integer Non-Linear Programming (MINLP) problem is formulated. Feasibility of the MINLP problem is in itself a requirement before well known mathematical tools can be applied to find the optimum (maxima/minima or global/local). Finding the optimum for a feasible MINLP problem is mathematically difficult due to non-convexity and highly coupled nature of the involved functions. The first step towards improving these results could be to identify a single mathematical framework to solve a general MINLP problem that performs atleast as good as the existing methods. This considerably reduces the need to migrate between different methods with a change in problem. The next step is to apply new mathematical methods and optimization principles to effectively solve the problem. This is important to improve the quality of the solution, since the mathematical problem may not be easy to handle due to various implicit reasons such as non-invertible functions and non-availability of analytic expressions. For the numerical results, a simplified link-level simulation model that is an abstraction to the actual system-level model is used. The setup is verified for a minimum working example and can easily be scaled to match the practical parameters. The setup assumes 2 Base Stations (BSs) each with 2 antennas serving a set of homogeneous User Equipments (UEs) each with a single antenna. Also each BS serves

a single cell and all the network entities operate in the same time-frequency slot. In addition to the thermal noise, the primary source of degradation to a UE signal arises from the intra-cell and the inter-cell interference. Simulation results at the link-level compare the performance of the proposed techniques against frequently used methods. The former outperforms the later in most of the cases and performs at least as good as the later in some cases. System performance is measured in terms of the sum-rate and sum-power objectives. Depending on the problem, an improvement could refer to an increase in the sum-rate, a decrease in sum-power, produce better bounds on the objective or better handling of the problem.

Kurzfassung

MIMO erreicht eine vergrößerte spektrale Effizienz und eine nahezu zuverlässige Verbindungsqualität ohne andere wertvolle Netzwerkressourcen wie Bandbreite und Leistung zu vergrößern. MIMO nutzt die räumliche Trennung der Netzwerkeinheiten und die räumliche Fülle der Umgebung aus, um dieses Ziel zu erreichen. Aufgrund der vorliegenden Natur des geteilten Mediums eines drahtlosen Kanals, d.h. mehrere Übertragungen bei Benutzung der gleichen Resource zur gleichen Zeit, wirkt sich dies als Co-Kanal-Interferenz aus. Dies ist besonders stark an den Zellrändern. Deswegen ist die Performance des Netzwerks durch die Interferenz begrenzt und das Bedürfnis einer optimalen Ressourcen-Allokation beim Bereitstellen sicherer Dienste für die Netzteilnehmer steigt.

Diese Ausarbeitung/Dissertation behandelt die Entwicklung von iterativen Algorithmen der abstrahierten Bitübertragungsschicht für das Interferenz-Management (IM) und der Resourcen-Allokation (RA) in konventionellen Mehr-Zellen Mehr-Nutzer Drahtlossystemen. Dazu werden ausgereifte mathematische Methoden mit ein paar Änderungen und neuartigen Optimierungen angewandt. Untersuchungen wurden getrennt für den Uplink sowie für den Downlink durchgeführt.

Aufgrund der erforderlichen Dienste-Qualität (quality-of-service) in Mehr-Zellen MU-MIMO Systemen wurde die Problematik der Maximierung der Rate und Leistungskontrolle bei verschiedenen Gegebenheiten berücksichtigt. Dies beinhaltet Untersuchungen der perfekten Kanal-Status-Information (CSI), statistischen CSI, mangelhafter-CSI und wahrscheinlichkeitstheoretischen Randbedingungen.

Um das Ziel der RA zu erreichen, wurde ein äquivalentes mathematisches Optimierungsproblem, gewöhnlich ein "Mixed Integer Non-Linear Programming"(MINLP) Problem formuliert. Die Durchführbarkeit des MINLP-Problems ist eine Voraussetzung um bekannte mathematische Werkzeuge anwenden zu können um das Optimum (Maximum/Minimum, global/lokal) zu finden. Das Herausfinden des Optimums eines durchführbaren MINLP-Problems ist aufgrund von Nicht-Konvexität und der hohen Verkopplung der einbezogenen Funktionen mathematisch anspruchsvoll.

Der erste Schritt um diese Ergebnisse zu verbessern, könnte durch ein einfaches mathematisches Framework erfolgen um ein allgemeines MINLP-Problem zu lösen, welches mindestens genau so gute Ergebnisse liefert wie die bereits existierenden Methoden. Dies vereinfacht drastisch die Notwendigkeit zwischen verschiedenen Methoden zu wählen wenn sich die Problemstellung ändert.

Der nächste Schritt ist eine neue mathematische Methode und Optimierungsprinzip anzuwenden, um dieses Problem effektiv zu lösen. Dies ist wichtig, um die Qualität der Lösungsfindung zu verbessern, da es nicht einfach ist mit dem mathematischen Problem zu handhaben aufgrund verschiedener Gründe wie nicht invertierbarer Funktionen und fehlender Existenz von analytischen Ausdrücken.

Für die numerischen Ergebnisse wird ein vereinfachtes Verbindungs-Modell benutzt anstatt dem aktuellen System-Modell. Diese Ergebnisse werden für eine minimale Anzahl an lauffähigen Beispielen verifiziert, welche für praktisch benutzten Parameter skaliert werden können. Der Aufbau enthält zwei Basisstationen (BSs) welche jeweils zwei Antennen besitzen, die eine homogene Menge von Benutzer-Geräten (UEs) mit jeweils einer Antenne bedient. Außerdem bedient jede BS eine Zelle und alle Netzwerkeinheiten operieren im gleichen Zeit-Frequenz-Schlitz. Die hauptsächliche Verschlechterung eines UE Signals kommt von den Intra-Zell- und Inter-Zell-Interferenzen neben dem thermischem Rauschen.

Simulationsergebnisse der Verbindungsschicht vergleicht die Performance der vorgeschlagenen Techniken gegenüber den meist genutzten Methoden. Die ehemaligen Methoden, übertreffen die letzten in den meisten Fällen und sind mindestens genau so gut wie die letzten in manchen Fällen. Die Systemperformance wird anhand von Gesamtrate und Gesamtleistung ermittelt. In Abhängigkeit des Problems könnte eine Verbesserung zur Erhöhung der Gesamtrate, eine Erniedrigung der Gesamtleistung, bessere Rahmenbedingungen oder bessere Handhabung des Problems führen.

Chapter 1

Introduction

Meeting the ever-increasing demand for new services and providing better Quality of Service (QoS) requires higher data rates on the link layer, which basically is limited by the availability of wireless resources such as bandwidth and power. The specifications of International Telecommunication Union Radio communication sector (ITU-R) in its IMT-Advanced are being standardized by the working group 3rd Generation Partnership Project (3GPP) [1] for the Radio Interface Technology as Long Term Evolution Advanced (LTE-A). With 4G technologies being gradually deployed globally, 5G technologies are already on the horizon. Several key issues were identified in the standardization, some requirements include enhancing cell-edge user throughput and cell spectral efficiency, reduction in latency and cost per data bit, improved coverage, provide a provision for flexible bandwidth, multi-antenna support and allow compatibility with existing technologies. Peak data rates as high as 1-Gb/s on the Downlink (DL) is suggested to meet the increasing wireless traffic, mainly arising from video. This requires evolution of both the wireless cellular technologies and the core networks. To address the coverage and capacity issues along with the task of Interference Management (IM) and Resource Allocation (RA) several candidate technologies have been identified. These include technologies such as Heterogeneous Networks (HetNets), Self Organizing Networks (SON), Coordinated Multi Point (CoMP) Transmission and Reception and 3D-MIMO. Affordable wireless devices, increasing demand for new wireless services, increasing User Equipment (UE) density and limited available resources are driving the wireless network to a small cell architecture. Small cells with cheaper base stations are being considered to increase the spectral efficiency. This leads to a heterogeneous architecture where several low power access points serving a small region, called femto cells, coexist with the macro base station (eNodeB in LTE) [2]. Inter-networking with WiFi is also under consideration. So the resource reuse pattern must be carefully addressed. With femto cells, the operating and maintenance costs are transferred to the end user with this architecture. SON [3] can be used to optimize the network dynamically. Load balancing, interference control, self-configuring and self-healing properties are defined to optimize the capacity. With frequency reuse factor one, multiple antennas are being deployed to exploit the space dimension. Massive-MIMO [4], employing

large-scale antenna systems with tens and hundreds of antenna elements is also under consideration. This improves the spectral efficiency and energy efficiency in the system. The increase in spectral efficiency is due to spatial-multiplexing. Energy efficiency increases due to the possibility of focusing energy more precisely into small space regions. 3D-MIMO [5] technology in the Active Antenna Systems (AAS) of LTE-A is an advanced technique to completely utilize the spatial domain. In addition to expanded coverage and increase in spectral efficiency, it enables 3D-beamforming where space domain beamforming in both vertical and horizontal directions is achieved. CoMP [6] in LTE-A is a technique where more than one distributed serving nodes cooperate to serve a geographical area. On the DL CoMP [7], if the UE data is available only at the serving cell then it is categorized as Coordinated Scheduling and Beamforming (CSB) and if the UE data is available at multiple points it is categorized as Joint Processing (JP). In JP, UE perceives a virtual antenna array, when more than one access points are simultaneously serving it. Such fully coordinated array is called the network-MIMO [8]. In CSB, even though the UE is served by a single Base Station (BS), multiple BSs coordinate by sharing cell specific and UE specific information to adjust the parameters in their respective cells to manage the interference levels. CoMP improves cell-edge spectral efficiency and reduces cost per bit. On the Uplink (UL), CoMP facilitates joint multi-cell scheduling and joint multi-cell signal processing.

The basic aim of the network operator is to design a cost effective network to maximize the revenue while providing a quality service to the UEs. There are several design and implementation issues at every point of a wireless link. The maximum possible theoretical transmission rate is the capacity. Identifying the capacity cells and coverage cells in a geographical area to carry out the link-budget analysis could be the beginning. The capacity cells focus on providing high data rates and enhanced services, while the coverage cells mainly focus on providing a minimum service. In a dense UE environment, capacity cells could be created by cell splitting and using small cell architecture. The region in which the system operates and the allocated bandwidth in this region impacts the design of equipment and technology. Wireless transmission is impaired by attenuation, distortion, interference and noise hence it cannot be error-free. Due to the shared nature of the wireless channel and existence of various wireless services, interference is the primary source for signal degradation. Cross-layer optimization [9] is a proposed technique to close the gap between the achieved throughput and the theoretical capacity. This cross-layered architecture simplifies functionality and inter-portability, but a lack of coordination among the layers limits the performance. Working only over the physical layer may not satisfy the network throughput requirement. A wireless network design based on cross-layer optimization coordinates the resources allocated to different layers to improve network performance e.g., reduce handoff latency, improve energy efficiency, exploit network diversity by Network assisted Diversity Multiple Access (NDMA) [10]. The functionality of each layer is still hidden from other layers but coordination, interaction and joint optimization across them is allowed. Cognitive Radio (CR) [11] is one of the techniques that aims at efficient utilization of the radio

spectrum. With CR spectral holes are identified and the spectrum is opportunistically used to serve secondary UEs without a drop in primary UE's QoS. CR senses the surroundings, adapts and reconfigures its parameters for dynamic spectrum utilization. Software Defined Radio is a practical implementation of a CR.

1.1 Basics of Wireless Communication

1.1.1 Fading

Wireless channel is an unguided transmission medium which is time-varying in nature. This variation could be due to the motion of transmitter or receiver, motion of the obstacles in the medium and atmospheric changes. These variations cause propagation loss in the received signal. To simplify the propagation analysis and model the effects of these variations, pathloss, shadowing and multipath fading are usually defined. Pathloss is the distance-dependent parameter that explains the long term or average decrease in the received power level with increasing distance between transmitter and receiver. A pathloss exponent determines how fast the received power decays with increasing distance. This depends on the propagation environment terrain e.g., buildings and vegetation. Hence the pathloss is different for urban, rural, indoor and outdoor environments. Free space pathloss is well modelled in the Friis transmission equation [12]. Due to the dependence of pathloss on various factors, empirical models [13] based on practical measurements were developed, such as, Hata and COST models [14]. Shadowing arises due to the obstacles in the medium. The attenuation of the signal due to shadowing is statistically described by a log-normally distributed random variable. Pathloss and shadowing are usually considered as long term fading effects. Transmitted signal arrives at the receiver from multiple reflected paths due to reflection, diffraction and scattering. This is a multipath phenomena. The multiple reflected rays usually have different amplitude, phase and angle of arrival. They combine either constructively or destructively, leading to severe short term fluctuations in the received signal. The fluctuations due to multipath is called multipath fading [15]. Frequently used statistical models for short term fading are the Rayleigh, Rician, Nakagami- m , Hoyt, Weibull fading models [16]. The fading model name suggests the distribution of the magnitude of the channel coefficient. In a rich scattering environment, there is a multitude of independent received signal components. In such a case, if there exists no line-of-sight communication path between the transmitter and the receiver, the fading can be statistically modelled by Rayleigh distribution. The mean square value of the received signal will represent the average long term received power. Rician model is often used when a line-of-sight exists. Nakagami- m is a more general model that can capture a variety of fading scenarios. Fading parameter m , captures the severity of fluctuations. To study the non-homogeneity and the non-linearity of the propagation medium, η - μ [17], κ - μ [18] and α - μ [19] distributions were proposed. They consider

a signal composed of clusters of multipath waves propagating in a non-homogeneous environment. Within any one cluster, the phases of the scattered waves are random and have similar delay times, with delay-time spreads of different clusters being relatively large. In κ - μ model there exists within each cluster, a dominant component with arbitrary power. Similar generalized model in homogeneous medium was suggested by the λ - μ distribution [20]. To further generalize fading distributions, α - η - μ and α - κ - μ distributions [21] were proposed.

Together the long term and short term fading are represented by a multiplicative expression, i.e., the short term fading is superimposed onto the large scale fading. Short term fading manifests into frequency-flat fading, frequency-selective fading, slow fading and fast fading. To study and model the time-varying nature of the channel impulse response, observation-time (t -domain), time-delay (τ -domain), frequency (f -domain), Doppler-shift (λ -domain) are introduced. To categorize various mechanisms into their respective domains, the parameters coherence bandwidth B_c , symbol rate R_s , channel fade-rate R_c , coherence time T_c , delay spread T_d and symbol time T_s are defined. The time-spreading mechanism due to multipath studied in the τ -domain and f -domain results is frequency-flat fading and frequency-selective fading [15]. In the τ -domain, $T_d > T_s$ results in frequency-selective fading, while $T_d < T_s$ results in frequency-flat fading. In the f -domain, frequency-selective fading results when $B_c < R_s$ and frequency-flat fading when $B_c > R_s$. The time-variant mechanism due to motion, studied in the t -domain and λ -domain, results in fast fading and slow fading [15]. In the t -domain, fast fading occurs when $T_s > T_c$, while $T_s < T_c$ results in slow fading. In the λ -domain, $R_c > R_s$ results in fast fading, while slow fading occurs when $R_c < R_s$. Frequency-selective fading causes inter-symbol interference which increases the Bit-Error Rate (BER) if not proper estimation and compensation is not performed. Fast fading increases the Doppler-spread and also increases the BER.

1.1.2 Frequency Reuse

Radio spectrum is the fundamental requirement for wireless radio networks and reuse of resource elements such as time, frequency and space is an important concept. This directly relates to the network planning and link-budget. In a fractional frequency reuse scheme [22] each cell is allotted a specific frequency band and adjacent cells are allotted different bands. The same frequency pattern is repeated over the coverage area. Each cell gets to use only a fraction of the total spectrum. A minimum distance called reuse distance must be maintained by the BSs, such that sufficient Signal to Interference plus Noise Ratio (SINR) is achieved at each UE. Radio links suffer from cochannel interference (CCI) from cells that share the same frequency. Large frequency reuse factor reduces the CCI but the bandwidth is not fully utilized, resulting in a reduced system capacity. Hence a trade-off between the CCI and system capacity. In a soft frequency reuse [23] scheme the cell-center and cell-edge UEs are first identified.

This process varies with the actual traffic distribution. It can be expected that the received signal at the cell-center UE is stronger with a lower inter-cell interference, so at the cell-center a frequency reuse factor one, that utilizes complete bandwidth is used and fractional frequency reuse is adopted for cell-edge UEs. Designing a spatial reuse pattern could be a marginal solution to optimize network performance. Soft frequency reuse may also refer to the pattern in which all the BSs operate with reuse factor one [24]. The CCI could be either intra-cell or inter-cell with full frequency reuse. Nevertheless by adjusting the power levels of each UE, the reuse pattern and reuse factor can be dynamically decided with cooperation among various BSs. With a transmit power limitation on the UL and frequency reuse one, the cell size is reduced, thereby making the system interference-limited. Inter-cell interference cancellation techniques [25] become essential.

1.1.3 Multiple Access

To allow multiple UEs to access and share the system resources simultaneously, radio access technologies are characterized by multiple access schemes. Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA) are well known multiple-access techniques. In FDMA, the total spectrum is divided into a number of channels and the available channels are allotted for communication. A guard band and avoiding allocation of adjacent bands in the same location are usually adopted to prevent power leakage to adjacent bands. UEs communicate at the same time on their assigned frequency bands. By limiting the transmit power level, it forms the basis for the frequency reuse in the cellular communication. In TDMA, each UE communicates in the allotted time slot at the same frequency. This gives access to an increased bandwidth to each UE. In addition, a proper frame structure, time synchronization between transmitter and receiver, provision for guard interval must be considered. With CDMA, each UE gets access to the complete spectrum while communicating at the same time. This is possible due to a unique spreading code assigned to each UE in the system. Each code must satisfy some correlation properties and must be orthogonal to the other spreading codes in the system. In Space Division Multiple Access (SDMA), the angular separation or spatial separation between the UEs is exploited to simultaneously provide the service on the same time-frequency unit. It relies on the beamforming technique.

Multi-Carrier (MC) modulation [26] is an attractive solution to achieve high data rates. Orthogonal Frequency Division Multiplexing (OFDM) [27] is one such scheme that converts the incoming serial data stream into several lower data rate parallel sub-streams and transmits the sub-streams over orthogonal sub-carriers, one per sub-stream. As a result a frequency-selective channel is converted to several flat-fading channels. It is possible to have MC multiple-access schemes such as MC-TDMA and MC-FDMA with OFDM symbols [28]. Orthogonal Frequency Division Multiple Access (OFDMA) is an

extension of OFDM, where OFDM is combined with either TDMA or FDMA to assign a set of sub-carriers to UEs as required. Nevertheless, OFDMA in general refers to OFDM-FDMA [29]. OFDMA can be treated as MC-FDMA, where a set of OFDM sub-carriers are exclusively allocated to a UE. In OFDM-TDMA, the complete bandwidth is provided to a UE for a number of assigned OFDM symbols. An adaptive technique [30] which dynamically allocates OFDM sub-carriers and power to UEs could improve the multiple-access technique. High Peak-to-Average-Power Ratio (PAPR) is one of the drawbacks of OFDM, hence on the LTE-UL, Single-Carrier OFDMA (SC-OFDMA) is adopted. This is attractive since there could be some sub-carriers that are in deep fade and not suitable to carry data, while some sub-carriers are more suitable to certain UEs. A new multiplexing scheme Non-Orthogonal Multiple Access (NOMA) [31] utilizes a new power-domain. Non-orthogonality is intentionally introduced via power-domain UE multiplexing. A significant throughput gain over OFDMA is observed. Generalized Frequency Division Multiplexing (GFDM) [32] proposed to deal with the heavily fragmented spectrum is another alternative. It is based on the multi-branch, MC transmission system. When compared to the OFDM, the PAPR is reduced and each sub-carrier can be individually modulated, also the GFDM sub-carriers are not orthogonal.

1.1.4 Multiple-Input Multiple-Output

Shannon's information capacity theorem [33] relates the channel bandwidth and transmission power. For a Single-Input Single-Output (SISO), where there is only one antenna element at either end of the communication link, the capacity scales logarithmically with the Signal to Noise Ratio (SNR) and scales linearly with the bandwidth. However there is an upper limit to this capacity per unit bandwidth as the $\text{SNR} \rightarrow \infty$. To meet the ever increasing data rate requirements, another Degree of Freedom (DoF) must be introduced into the system. Multiple antenna elements at both the transmitter and the receiver, called Multiple-Input Multiple-Output (MIMO) system, introduces a new space dimension into the system. The capacity for a MIMO system increases linearly with the number of antenna elements [34]. This increase in the capacity, termed as multiplexing-gain is at the expense of increase in hardware and signal processing complexity. The effect is similar to an increase in the bandwidth. If the channel is diagonalized, a number of spatial streams or eigenmodes between the transmitter and the receiver are created by this new spatial DoF. Data can be parallelly sent over them and is differentiated at the receiver based on the spatial-signature. The capacity of a MIMO system can be further increased by spatial waterfilling [35] when the transmitter has full Channel State Information (CSI), i.e., a closed loop setup where eigenmodes are accessible. Antenna elements spaced sufficiently wide in a rich-scattering environment at the receiver may result in each element having independent copy of the transmitted signal. This is termed as spatial diversity on receive and makes the communication link more reliable. Assuming no correlation among the antenna elements, it could be

expected that at least one antenna element will receive a strong signal component. This reduces the probability of the received signal being in deep fade when compared to a SISO link. Receive diversity techniques, such as maximal ratio combining can be employed to increase the received SNR [36].

Single-UE MIMO is a point-to-point communication where the full spatial dimension is dedicated to a UE. The capacity scales linearly with $\min(R, T)$, where T is the number of antenna elements at the BS, R is the number of antenna elements at the UE and $\min(\cdot, \cdot)$ is the minimum operator. For a MIMO fading channel, the ergodic capacity is the Shannon capacity averaged over the fading process. Other capacity metrics such as delay limited capacity, which is the transmission rate possible for any given fading state of the channel, and outage capacity, which is the rate beyond which the channel is in outage, are also defined to study the performance. Diversity on transmit is achieved by spreading the symbols over space and time. Under the open-loop configuration where there is no feedback from the receiver to the transmitter, space time codes are employed [37]. They are categorized as Space Time Block Codes (STBC) and Space Time Trellis Codes (STTC). STTC operate on serial data for encoding and require memory elements. A multi-dimensional Viterbi decoder is required at the receiver. The exponential increase in decoding complexity is a drawback of STTC. STBC works on the orthogonality in the space and time domain for the codewords. This simplifies the receiver structure by facilitating linear post-processing. Linearity, full-diversity or orthogonality may have to be compromised for achieving higher data rates. Transmit Selection Diversity (TSD) is the form of transmit diversity under closed loop setup. TSD basically selects only a subset of transmit antennas thereby reducing the number of spatial chains on the link. Reduction in complexity and spatial interference are the advantages here. Even with reduced transmit antennas, the diversity order remains the same. A Layered Space Time (BLAST) architecture [38] was designed to provide spatial multiplexing over MIMO. Data is spatially demultiplexed into sub-streams and then mapped onto the transmit antennas. Each sub-stream is independently temporal encoded, interleaved and mapped. The receiver can independently decode the sub-streams thus avoiding joint decoding.

1.2 Multi-User MIMO

In a single cell, multiple MIMO links can exist, that share the space dimension. The shared channel is the MU-MIMO channel. Antennas across all the UEs must be considered for simultaneous scheduling. For understanding, Figure 1.1 shows the system-level diagram of a multi-cell MU-MIMO setup on the DL with 2 cells. Each cell is served by a BS, and the UEs are already associated to their respective BSs. All the network entities operate at a frequency reuse factor one. Each UE has a single antenna, the DL MIMO channel vector is \mathbf{h}^H . UE₃ and UE₄ are almost equidistant from either BSs, i.e., they are located at the cell-edge, hence they experience severe Multi-User Interference

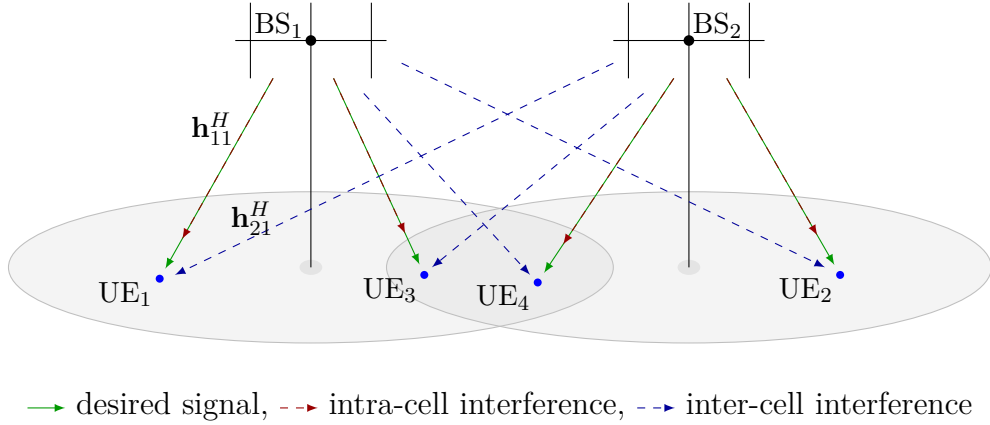


Fig. 1.1: Multi-Cell Multi-User MIMO Downlink.

(MUI). As seen, UE₁ and UE₃ are served by BS₁. For UE₁, both the desired signal and the intra-cell interference signal arrive over the channel \mathbf{h}_{11}^H . The intra-cell interference signal is the spatially multiplexed data of UE₃ with UE₁ at BS₁. The spatially multiplexed data of UE₂ and UE₄ at BS₂, arrives over \mathbf{h}_{21}^H as inter-cell interference signal. The contribution to the interference term of UE₁ includes the signals of all the UEs arriving at UE₁ over the DL channels between all the BSs and UE₁. Similarly, other UEs receive their signal in the presence of MUI. Thermal noise is added at each UE receiver.

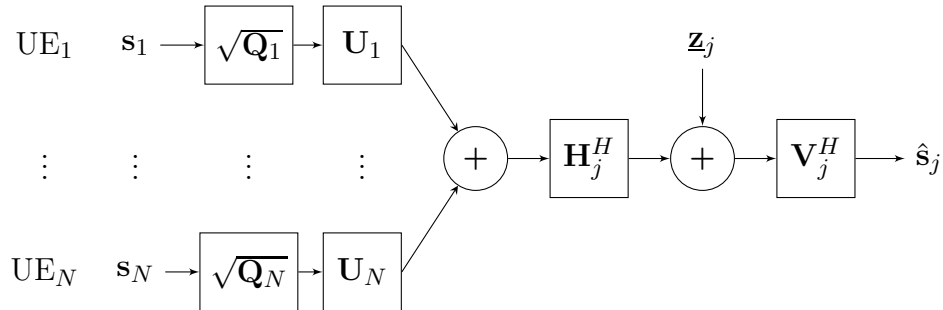
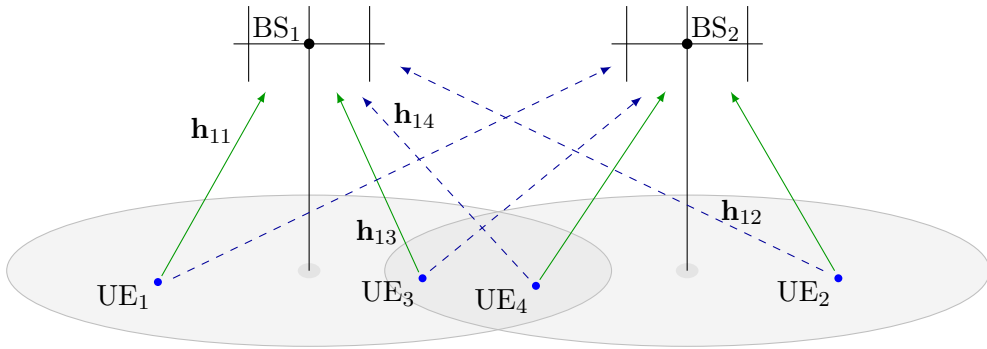


Fig. 1.2: Single-Cell, single BS, MU-MIMO Downlink (BC) transmission for UE_j.

A Broadcast Channel (BC) in Figure 1.2 is an abstraction of the DL transmission for N UEs, i.e., a single-point to multi-point transmission [39]. However, it considers only a single-cell case of Figure 1.1. With each UE equipped with R antennas, it is possible to schedule $L \leq R$ simultaneous data streams to each UE on the DL. Per-layer processing is considered and in addition to the inter-user interference, the CCI will have an intra-layer term. The vector $\mathbf{s}_j = [s_{j1}, \dots, s_{jL}]^T$ is the $(L \times 1)$ symbol vector of UE_j, $\{j = 1, \dots, N\}$. The transmitted symbol for UE_j on layer l , $\{l = 1, \dots, L\}$ is s_{jl} . It is normalized such that $\mathbb{E}[s_{jl}s_{jl}^*] = 1, \forall j, \forall l$. $\mathbb{E}[\cdot]$ is the expectation operator, $(\cdot)^T$ is the transpose operator and $(\cdot)^*$ is the complex conjugate operator. The transmitted

symbols are independent, i.e., $E[s_{jl}s_{jm}^*]=0, \forall j, \forall m \neq j$ and $E[s_{jl}s_{kn}^*]=0, \forall j, \forall k$. The matrix $\mathbf{Q}_j = \text{diag}(\mathbf{q}_j)$ is the $(L \times L)$ diagonal power loading matrix for UE $_j$ at the BS. The vector $\mathbf{q}_j = [q_{j1}, \dots, q_{jL}]^T$ is the $(L \times 1)$ power vector of UE $_j$. The power allocated to layer l of UE $_j$ is q_{jl} . For a vector input, the $\text{diag}(\cdot)$ operator stacks the vector as the principal diagonal elements of a diagonal matrix and for a matrix input $\text{diag}(\cdot)$ operator extracts the diagonal elements of a square matrix. To effectively manipulate interference among various entities when transmission occurs on the same time-frequency unit, a spatial pre-processing called precoding at the transmitter and a spatial post-processing called decoding at the receiver are employed. This precoding is not to be confused with the source coding and channel coding, it can be viewed as changing the effective channel. The suboptimal strategy when the transmitter employs linear precoding is called Beamforming (BF). It is possible by taking the advantage of SDMA to schedule more UEs and control the MUI, if the BF vectors are carefully chosen. The design of precoders and decoders depend on the system objective and link direction. $\mathbf{U}_j = [\mathbf{u}_{j1}, \dots, \mathbf{u}_{jL}]$ is the $(T \times L)$ transmit BF matrix of UE $_j$, \mathbf{u}_{jl} is the $(T \times 1)$ transmit vector of layer l for UE $_j$. It is required that $\|\mathbf{u}_{jl}\|_2 = 1, \forall j, \forall l$, to satisfy the power constraint at the BS, where $\|\cdot\|_2$ is the ℓ_2 -norm. Beamformer $\mathbf{U}_j, \forall j$, spatially multiplexes the data of UE $_j$. \mathbf{H}_j^H is the $(R \times T)$ DL channel matrix from the BS to UE $_j$. $(\cdot)^H$ is the Hermitian operator. The input to the channel is the composite signal containing spatially multiplexed data of all the UEs. At the receiver, Additive White Gaussian Noise (AWGN) gets added to the composite signal, which is further subjected to a decoding process. \mathbf{z}_j is the $(R \times 1)$ AWGN vector at UE $_j$. The noise variance per receive antenna element at UE $_j$ is σ^2 , such that, $E[\mathbf{z}_j\mathbf{z}_j^H] = \sigma^2\mathbf{I}_R$. \mathbf{I}_R is the $(R \times R)$ Identity matrix. $\mathbf{V}_j^H = [\mathbf{v}_{j1}, \dots, \mathbf{v}_{jL}]^H$ is the $(L \times R)$ receive BF matrix at UE $_j$. \mathbf{v}_{jl} is the $(R \times 1)$ receive BF vector of layer l for UE $_j$ and $\|\mathbf{v}_{jl}\|_2 = 1, \forall j, \forall l$. $\hat{\mathbf{s}}_j = [\hat{s}_{j1}, \dots, \hat{s}_{jL}]^T$ is the $(L \times 1)$ post-processed received symbol vector at UE $_j$ providing an estimate for \mathbf{s}_j .



→ desired signal, intra-cell interference for cochannel UEs, - - → inter-cell interference

Fig. 1.3: Multi-Cell Multi-User MIMO Uplink.

Similar to Figure 1.1, Figure 1.3 shows the system-level diagram of a multi-cell MU-

MIMO setup on the UL. For UE₁, the desired signal is communicated over \mathbf{h}_{11} . The desired signal of UE₃ will be the intra-cell interference component of UE₁ via channel \mathbf{h}_{13} . Due to reuse factor one, the signals from UE₂ and UE₄ also contribute to the inter-cell interference via channels \mathbf{h}_{12} and \mathbf{h}_{14} respectively. The contribution to the interference term of UE₁ includes the signals from all the UEs arriving at BS₁ over different UL channels. A Multiple Access Channel (MAC) in Figure 1.4 models a multi-point to single-point UL transmission, i.e., a single-cell link-level case of Figure 1.3. $\mathbf{P}_j = \text{diag}(\mathbf{p}_j)$ is the $(L \times L)$ diagonal power allocation matrix at UE_j. $\mathbf{p}_j = [p_{j1}, \dots, p_{jL}]$ is the $(L \times 1)$ power vector of UE_j. The power allocated to layer l of UE_j is p_{jl} . $\bar{\mathbf{z}}_j$ is the AWGN vector for UE_j at the BS. The noise variance per receive antenna element at the BS for UE_j is $\bar{\sigma}^2$.

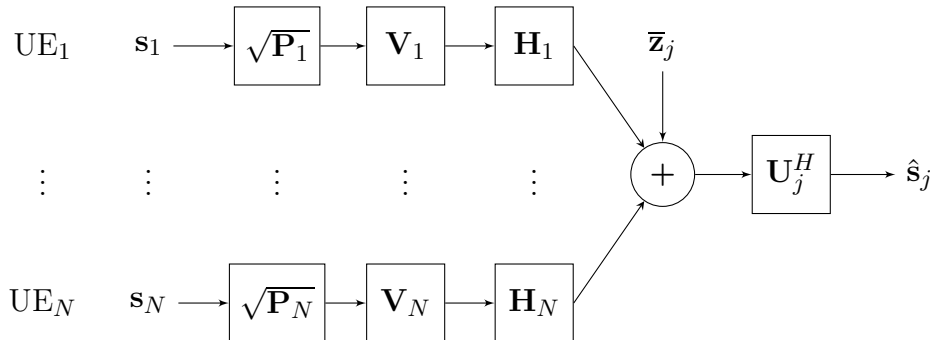


Fig. 1.4: Single-Cell MU-MIMO Uplink (MAC) transmission for UE_j.

In general, the channel matrices, transmit and receive beamformers are different on both the DL and UL. From mathematical point of view and structure of various matrices, for each UE, the UL channel is the transposed version of the DL channel, the transposed DL transmit matrix is viewed as the UL receive matrix and the transposed DL receive matrix is the UL transmit matrix. A duality between the BC and MAC has been explored and extensively applied for various transceiver optimization problems. This duality is different from the mathematical duality in optimization theory. Capacity of a MU-MIMO channel is given by an N -dimensional rate vector where N is the number of active UEs in the system. This vector is the set of simultaneously achieved rates of all active UEs. The set of all achievable rate vectors is the capacity region. Similar SINR and MSE region can also be defined. Performance evaluation by N -dimensional performance metric vector comparison may not be a suitable criteria. So a scalar quantity for the system-wide objective such as sum of all UE rates or UE powers or UE MSEs is usually considered. The objective is optimized w.r.t. various system and design constraints. By conventional BC-MAC duality [40] under the single total transmit sum-power constraint, the SINR region obtained for the BC and the MAC is the same with conjugate transposed channel and BF matrices. Also, the noise covariance matrices are identity matrices. Duality based on Mean Squared Error (MSE) metric [41] is also defined similar to the SINR duality. Under a sum-power con-

straint by minimax duality [40], every point on the BC capacity region can be obtained by solving a MAC minimax optimization problem with the channel transposed but the MAC noise covariance matrix is an optimization variable.

It is common that multiple BSs serve a geographical area, i.e., a multi-point to multi-point transmission. With increasing UE density and wireless traffic the need for coordination among the BSs arises. The level of coordination, i.e., the amount of CSI and UE-data at each BS, determines the scheduling and processing techniques. The processing could either be centralized or distributed based on how the BSs are connected and the level of side information, i.e., global CSI or local CSI availability. With multiple BSs, the problem of Base Station Association (BSA) arises. This is crucial, since the processing node and coordinating nodes are defined at this stage before the service is provided to a UE. BSA is another DoF introduced into the system that needs to be optimized. Each UE has a performance metric such as SINR, MSE or rate that needs to be optimized. Optimizing individual UE performance metric may not result in optimizing the system-wide metric. So a utility function that is a measure of satisfaction is assigned to each UE by the operator for RA of certain type.

1.2.1 Fairness

Fading statistics of the UEs are not always similar, so the fairness among the UEs must also be considered for design. Fairness in RA is important since the UEs are competing for the finite network resources. In this case, a Round Robin (RR) scheduler is the simplest way to serve each UE at a time in an ordered manner. It provides equal opportunity to all the UEs but does not provide any multiplexing gain. A weighted scheduling can capture certain degree of fairness among UEs e.g., weighted sum-rate where each UE rate in the sum is multiplied by a weight that reflects certain fairness criteria. When all the weights are equal to unity, then the obtained scalar objective is a simple sum-rate. For a simple sum-rate, the BS, in order to improve the system throughput can select a subset of UEs with favourable channel state. This form of selection diversity among the UEs is called Multi-User Diversity (MUD) and it may deprive service to some UEs. The improvement in the performance is the MUD gain. The sum-capacity of the fading channel is greater than that of a non-fading channel due to this MUD [42]. The Proportional Fairness (PF) [43] is a weighted scheduling scheme based on average UE throughput. If the channel statistics of the UEs are similar then the PF-rates will converge to the same value [42]. While maintaining fairness among the UEs, it also maintains MUD. A max-min fairness criteria [44] that maximizes the minimum allocation among the active UEs in the system can be considered in the interest of the most weak UE in the system. Such a criteria is fair from homogenous UEs point of view where all the UEs are treated with equal priority. α -fairness [45] captures a variety of fairness in the system. It uses a generalized utility function for each UE. $\alpha \rightarrow 1$ is the PF while $\alpha \rightarrow \infty$ is the max-min fairness. $\alpha = 2$ corresponds to

minimum potential delay fairness which seeks to minimize the total transfer time at a given rate. Another criteria that focuses on the extent of improvement in the efficiency and the extent of compromise on fairness is the (α, β) -fairness [46].

1.2.2 Transceiver design

The task of RA is to obtain the desired performance goals with limited resources by manipulating various DoF in the system. It is challenging since there are many DoF that need to be simultaneously optimized in a situation where each UE's performance is coupled to every other UE's performance. Also the order in which the updates must be carried out is also not certain. RA is solved by an equivalent constrained mathematical optimization problem with various system and design constraints. An optimal point in certain sense is obtained from the problem. Optimal may refer to the quality of solution, i.e., local or global optimum or one of the many solutions satisfying all the constraints in a reasonable way. Joint transmit and receive filter design finds a pair of optimal precoder and decoder. The signal processing and design techniques for the detection of UE signals in the presence of Multiple Access Interference (MAI) is called multi-user detection [47]. Joint optimum multi-user detection reduces the overall BER and has no near-far problem. One advantage of joint optimization is that a series of relations can be established under some assumptions among various performance metrics, such as under Linear Minimum Mean Squared Error (LMMSE) transceiver, maximizing SINR is equivalent to the minimization of normalized MSE [48], minimizing the product of MSE is equivalent to maximizing the sum-rate [49]. RA in a system is called Pareto efficient, if any UE's allocation cannot be improved without reducing any other UE's allocation. From Figures 1.2 and 1.4 various optimization problems such as Power Allocation (PA), that finds power loading matrices \mathbf{P}_j and \mathbf{Q}_j , BF that finds \mathbf{U}_j and \mathbf{V}_j , UE selection that finds the required active set of UEs, BSA that finds UE-BS association and so on arise. For each UE on the DL, $\sqrt{\mathbf{Q}_j} \cdot \mathbf{U}_j$ can be treated as a single transmit BF matrix such that ℓ_2 -norm for each column in the resultant matrix is equal to q_{jl} . Similarly on the UL, $\sqrt{\mathbf{P}_j} \cdot \mathbf{V}_j$ can be combined. Depending on the link direction, parameters and constraints change.

The capacity region of a MIMO BC is achieved by a non-linear precoding scheme called Dirty Paper Coding (DPC) [50] at the transmitter and Successive Interference Cancellation (SIC) at the receivers. DPC is an encoding strategy that involves an ordering of UEs and pre-subtraction of interference. As $N > T$, a linear increase in the capacity in terms of T is achieved, irrespective of R . DPC may be impractical since non-causal knowledge of interference is required and the involved expressions are non-convex. The capacity region of a MIMO MAC is achieved by performing SIC at the receiver. The basic aim of precoding at the BS is to simplify the receiver structure and also achieve the necessary diversity and multiplexing gain. Under the zero-MUI criteria and $N \cdot R \leq T$ as a primary requirement, Block Diagonalization (BD) [51] and

Zero-forcing (ZF) pre-processing are the simplest form of precoding on the DL. More often the objective in this case is the weighted sum-rate. The basic idea is that the precoder of a UE lies in the null space of all other UE channels. BD has capacity loss due to the nulling of overlapping subspaces of the different UEs [52]. ZF is a special case of BD when $R = 1, \forall j$, i.e., each UE has only one spatial layer. With BD, the MU-MIMO channel is decomposed into several independent parallel channels. To the interference-free streams, waterfilling can be used to optimally allocate system-wide power. For $N \cdot L > T$, it is not feasible to find null vectors for all the UEs, so a subset of UEs must be selected for zero-MUI criteria. In general, for $N \cdot L > T$, certain amount of interference is allowed for each UE while providing service as long as the objective is achieved and the constraints are satisfied. It is possible to have a ZF precoding in combination with a RR or PF criteria [53]. The scheduler must first select a UE subset for this. When only one UE with the largest channel gain is selected for transmission, the sum-rate achieved is called the TDMA rate. The precoders are obtained by channel inversion and waterfilling is used for optimal power allocation. This cannot achieve linear increase in the sum-capacity. This TDMA is not to be confused with the multiple access technique TDMA in section 1.1.3. If an ordering of the UEs is considered and partial MUI is allowed for the UEs then Successive Optimization (SO) [51] precoding arises. Instead of having null space requirement of all the UEs, only the null space requirement of the previous UEs in the order is considered. Random BF [54] is a BF technique where orthogonal vectors are evaluated and sent to each UE. Each UE then evaluates the SNR for all the vectors and reports it to the BS scheduler. A UE that has reported the maximum SNR is provided with a BF vector.

Opportunistic BF [42] is a BF technique that intentionally introduces fluctuations in the system to exploit the MUD while transmitting to a single UE. It is based on the fact that the MUD gain is higher with increase in channel randomization. The introduced random coefficients are closely matched to the channel coefficients to achieve higher SNR. The LMMSE [55] decoding is the Wiener filter that minimizes the overall MSE. At the optimal solution with the LMMSE beamformers, the minimum individual normalized MSEs and the maximum SINRs are simultaneously achieved [56], hence the LMMSE maximizes the SINR. LMMSE transceiver is optimal for total MSE minimization [57] under individual power constraint on the UL and maximizing individual SINR on the DL [58]. Other linear BF designs based on Signal to Generated Interference Noise Ratio (SGINR) [59] and Minimum Variance Distortionless Response (MVDR) [60] are also considered in certain cases. SGINR considers the locally available CSI and self generated CCI for distributed processing while MVDR considers and minimizes only the CCI without the desired signal power for BF. For the MMSE beamformers, it is required to know the signal power which may be precisely unavailable. Under such cases, a Least Squared Error (LSE) based Minimum Least Squared Error (MLSE) beamformer [61] can be a good alternative. It minimizes the LSE of the observation and no knowledge of signal power is needed. Thomlinson Harashima Precoding (THP) [62], a non-linear precoding technique can be used to avoid error propagation. THP

can be combined with MMSE and ZF precoders to increase diversity and reduce MUI [63]. To reduce latency and feedback overhead, a predefined codebook based precoding can be performed [64].

In a multi-cell setup, the cell-edge UEs perceive severe CCI since the signal power from more than one BS has almost the same strength. To further increase the system capacity the BSA must also be optimized. In a system with M BSs and a small set of UEs an exhaustive search with M^N combinations could be possible but when N becomes large this is prohibitive. So finding the optimal BS-UE combination becomes challenging. There could be per-BS constraints such as available power and load balancing that need to be considered. BSA can be viewed as a handoff mechanism where UE power levels at various BSs are compared and the BS resulting in the least power is chosen. Selection and dropping of UEs, called admission control [65] is also a design criteria since not every BS-UE combination is feasible or even if feasible, may have inferior performance. UE selection is often jointly optimized with other DoF, such as BF while optimizing an objective, such as sum-power or sum-rate. Selecting a UE in a cell that maximizes the capacity of that cell may not be an optimal choice on the system scale. Dominant eigenvalue based selection or norm based selection is the simplest form of UE selection. For ZF-BF, the selected UEs must have a high gain channel and must be nearly orthogonal. So a measure of orthogonality must be introduced for the UEs. As $N \rightarrow \infty$, ZF-BF with a semi-orthogonal UE selection (SUS) strategy can asymptotically achieve the capacity as DPC due to the MUD [53]. Construction of the semi-orthogonal group for SUS is based on the Gram-Schmidt orthogonalization. In Cognitive Radio, a protection margin that increases the minimum QoS to the operating links can be provided while admitting secondary UEs. Receive antenna selection [66] could be another DoF to be optimized. A subset of receive antennas are adaptively activated while retaining the diversity benefits. Power Control (PC) [67] has been studied to control the CCI and minimize the total transmit power while the required QoS for each UE is achieved. If the minimum required QoS, called minimum protection ratio [68], is the same for all UEs, then the PC technique achieves a balanced QoS across all the active UEs, i.e., at the optimal all the active UEs will have the same QoS value. Such a RA is fair w.r.t. the homogeneous UEs. Also this power minimization problem is coupled to the max-min problem where the minimum achievable QoS among the UEs is maximized. The optimal solution for both these problems is the same.

This thesis investigates conventional cochannel multi-cell MU-MIMO communication systems and optimizes various network DoF. In particular, the problems of Base Station Association and Power Allocation are solved w.r.t. the objective of sum-rate maximization and sum-power minimization. Frequency reuse factor one is applied, such that the main source of interference to a UE arises from the cochannel UEs. System-level setup of Figures 1.1 and 1.3 are analysed via the link-level abstractions of Figures 1.2 and 1.4 respectively. Link-level simulations are carried out for the CSB-CoMP based setup for Rayleigh flat fading channel. For numerical results in each chapter, a minimum working example is considered. It includes some assumptions, such as $T = 2$, $M = 2$,

$R = 1$, unit noise variance at BSs and UEs. Also certain variable values such as, transmit power and number of active UEs, are chosen to just demonstrate the effectiveness of the applied mathematical methods. These illustrations can be scaled to include the practical parameter values and analyse the network performance.

Chapter 2 provides the overview of various constraints, mathematical formulations and mathematical tools considered in various considered optimization problems. Basics of optimization, problem formulation and problem solving techniques are outlined. Additional literature results are also mentioned.

Chapter 3 outlines an Interior Point Method (IPM) framework based on numerical methods for Non-Linear Programming (NLP) techniques. It is applied to solve the NLP optimization problems in this thesis. A variety of problems can be solved with the same framework with a little or no modification, thus removing the need to apply problem specific mathematical tools. The IPM is applied to solve the UL sum-rate maximization and sum-power minimization problems under perfect-CSI. Though the results in this chapter appear trivial, they primarily serve as a benchmark for comparison with existing methods. All the considered Mixed Integer Non-Linear Programming (MINLP) problems under various assumptions in further chapters are reformulated as NLP problems to obtain an optimum. In addition, the feasibility issue is also addressed in the problem formulation.

Chapter 4 considers the optimization problems on the UL considered in chapter 3, but under different assumptions such as probabilistic constraints, statistical-CSI and imperfect-CSI. Under such assumptions, it is difficult to obtain good bounds on the objective. Concepts from Extreme Value Theory (EVT) and its application to Chance Constrained optimization are discussed in this chapter. New optimization metrics such as Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR) from the field of finance are considered to obtain tighter bounds. With these concepts, the optimization problem is basically reformulated into an NLP problem with no approximations. The generated bounds by this approach are tighter when compared to the existing approximation techniques. Also, these results match with the available analytic expression, which is not the case with the existing methods. For this, two cases for the UL sum-power minimization are examined.

Chapter 5 considers the Downlink (DL) setup under perfect-CSI. The expressions on the DL are highly coupled. This makes the problem difficult to handle, thus the UL-DL duality is exploited to find the optimum. An UL called virtual-UL (VUL) is formulated, which decouples the DL expressions and solves for the solution by alternately switching between the DL and VUL. Under a multi-cell setup, another concept of virtual-noise is also included. This deviates from the conventional single cell VUL-DL duality. Numerical results based on the implicit Lagrange-duality of the considered IPM show that, if the VUL-DL single cell duality is included into the constraint set of the optimization problem then, the IPM is capable of solving the problem without this deviation and also without the alternate switching.

Chapter 2

Mathematical Techniques for Optimization

There is no single method to solve the variety of optimization problems, each with its own set of objectives, requirements and constraints. Resource Allocation (RA) is usually a constrained Mixed Integer Non-linear Programming (MINLP) optimization problem. Non-convexity, highly non-linear functions or a combination of higher order functions make the problem intractable and closed form solutions may not exist. A possible local optimum based on numerical methods is obtained and further refined. Iterative greedy mechanism where an initial random point evolves in multiple stages hoping to converge to an optimum is usually adopted. Several sub-problems are successively solved where each sub-problem solves for a single variable while other variables are fixed. A rearrangement and manipulation of the involved functions and equations is required to simplify the sub-problems. This helps to identify the problem's structure and an appropriate mathematical tool can be applied to obtain the solution. Depending on the structure of the objective and the constraints, the involved sub-problems can be identified as Non-Linear Programming (NLP) [69], Integer Programming (IP), Geometric Programming (GP) [70], Semi-Definite Programming (SDP) [71], Second-Order Cone Programming (SOCP), Linear Programming (LP) and so on. To solve the optimization problem, Matlab based modelling language YALMIP [72] is used to present the problem to a solver. Free solvers based on Interior Point Methods (IPM) such as SDPT3 [73] an SDP solver, GPPOSY [74] a GP solver, SEDUMI [75] an SOCP solver, CPLEX [76, academic version] an IP solver can be used in combination with YALMIP.

2.1 Constraints

For a multi-cell setup, M BSs $\{i = 1, \dots, M\}$, each serving a cell are considered. Without loss of generality, $R = 1$ and $L = 1$ is assumed throughout. For a UE association

with a BS, 0–1 binary integer variable α_{ij} is used (different from α in section 1.2.1). If the UE $_j$ is associated with BS $_i$, then $\alpha_{ij} = 1$, else $\alpha_{ij} = 0$, i.e.,

$$\alpha_{ij} = \{0, 1\}, \forall i, \forall j. \quad (2.1)$$

The $(M \times 1)$ BSA vector of UE $_j$ is $\boldsymbol{\alpha}_j = [\alpha_{1j}, \dots, \alpha_{Mj}]^T$. The $(M \times N)$ overall BSA matrix is $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N]$. On the UL, every BS receives the signal from all UEs. Throughout this work, coordinated processing is assumed where the signal of a UE is decoded only at one assigned BS. Hence there is no macro-diversity and the assignment constraint is mathematically expressed as

$$\mathbf{1}_M^T \boldsymbol{\alpha}_j = 1, \forall j, \quad (2.2)$$

where $\mathbf{1}_M$ is an $(M \times 1)$ vector with all elements equal to unity. For independent and uncorrelated symbols and noise, equation (2.2) allows the received signal of UE $_j$ to be mathematically expressed as

$$\hat{a}_j^{UL} = \sum_{i=1}^M \alpha_{ij} \mathbf{u}_{ij}^H \mathbf{h}_{ij} \sqrt{p_j} a_j + \sum_{i=1}^M \sum_{k \neq j, k=1}^N \alpha_{ik} \mathbf{u}_{ij}^H \mathbf{h}_{ik} \sqrt{p_k} a_k + \sum_{i=1}^M \alpha_{ij} \mathbf{u}_{ij}^H \mathbf{z}_{ij}, \forall j. \quad (2.3)$$

On the right hand side of the equality, the first term in the desired signal of UE $_j$, the second term in the MUI, the third term in the thermal noise component. In the MUI term, the intra-cell and the inter-cell interference terms can be identified after a given BSA. Each UE is assumed to have a maximum available power P_{\max} . The UL transmit power is bounded as

$$p_j \in [0, P_{\max}], \forall j. \quad (2.4)$$

Equation (2.4) is a short term power constraint. For a long term constraint, the transmit power may exceed the total power for some channel states as long as it is compensated for other channel states [48]. The UL-SINR $\bar{\gamma}_j$ is expressed as

$$\bar{\gamma}_j = \frac{p_j \sum_{i=1}^M \alpha_{ij} \|\mathbf{u}_{ij}^H \mathbf{h}_{ij}\|_2^2}{\sum_{i=1}^M \sum_{k \neq j, k=1}^N p_k \alpha_{ik} \|\mathbf{u}_{ij}^H \mathbf{h}_{ik}\|_2^2 + \bar{\sigma}_j^2}, \forall j. \quad (2.5)$$

The UL-rate for UE $_j$ is $\bar{r}_j = \log_2(1 + \bar{\gamma}_j)$, where $\log_2(\cdot)$ is the logarithm to base-2. This expression of UE rate is derived by treating the interfering signals as additive Gaussian noise [77]. For a given $\boldsymbol{\alpha}$, $\bar{\gamma}_j$ and \bar{r}_j are written as $\bar{\gamma}_{ij}$ and \bar{r}_{ij} respectively. Similarly the DL received signal at UE $_j$ is

$$\hat{a}_j^{DL} = \sum_{i=1}^M \alpha_{ij} \mathbf{h}_{ij}^H \mathbf{u}_{ij} \sqrt{q_{ij}} a_{ij} + \sum_{m=1}^M \sum_{k \neq j, k=1}^N \alpha_{mk} \mathbf{h}_{mj}^H \mathbf{u}_{mk} \sqrt{q_{mk}} a_{mk} + \underline{z}_j, \forall j. \quad (2.6)$$

With maximum available transmit power \bar{P}_i at BS $_i$, the DL transmit power q_{ij} of UE $_j$ at BS $_i$ is bounded as

$$q_{ij} \in [0, \bar{P}_i], \forall i, \forall j. \quad (2.7)$$

The $(M \times 1)$ DL power vector for UE $_j$ across all the BSs is $\mathbf{q}_j = [q_{1j}, \dots, q_{Mj}]^T$. For a given $\boldsymbol{\alpha}$, not every element in \mathbf{q}_j of UE $_j$ will be active, i.e., only one element in \mathbf{q}_j corresponding to the associated BS will be equal to 1 due to equation (2.2). The active DL power vector of all the UEs is $\mathbf{q} = [\mathbf{1}_M^T \mathbf{q}_1, \dots, \mathbf{1}_M^T \mathbf{q}_N]$. The $(M \times N)$ overall DL power matrix is $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_N]$. The sum of transmit power on the DL allocated to all the UEs assigned to BS $_i$ is constrained by \bar{P}_i , i.e.,

$$\underline{\boldsymbol{\alpha}}_i^T \underline{\mathbf{q}}_i \leq \bar{P}_i, \forall i, \quad (2.8)$$

where $(N \times 1)$ vectors $\underline{\boldsymbol{\alpha}}_i$ and $\underline{\mathbf{q}}_i$ are the i^{th} rows of $\boldsymbol{\alpha}$ and \mathbf{Q} respectively. The DL-SINR $\underline{\gamma}_j$ of UE $_j$ to be expressed as

$$\underline{\gamma}_j = \frac{\sum_{i=1}^M q_{ij} \alpha_{ij} \|\mathbf{h}_{ij}^H \mathbf{u}_{ij}\|_2^2}{\sum_{m=1}^M \sum_{k=1, k \neq j}^N q_{mk} \alpha_{mk} \|\mathbf{h}_{mj}^H \mathbf{u}_{mk}\|_2^2 + \underline{\sigma}_j^2}, \forall j. \quad (2.9)$$

The DL-rate for UE $_j$ from BS $_i$ is $\underline{r}_j = \log_2(1 + \underline{\gamma}_j)$. For a given $\boldsymbol{\alpha}$, $\underline{\gamma}_j$ and \underline{r}_j are written as $\underline{\gamma}_{ij}$ and \underline{r}_{ij} respectively. The UL QoS requirement for each UE is the rate metric given as

$$\bar{r}_j \geq r_{th}, \forall j, \quad (2.10)$$

where r_{th} is the minimum threshold rate. Similarly, the DL QoS requirement is

$$\underline{r}_j \geq r_{th}, \forall j. \quad (2.11)$$

Due to the per unit bandwidth analysis, the rate metric is also the spectral efficiency which is measured in [bits/sec/Hz]. In equations (2.3), (2.5), (2.6) and (2.9) the BSA in variable $\boldsymbol{\alpha}$ is still not resolved. Also, the BF matrices $\mathbf{V}_j, \forall j$, shown in Figures (1.2) and (1.4) do not appear since $R = 1$. The MUI term in the denominator of SINR expressions (2.5) and (2.9) includes both the intra-cell and inter-cell interference. Also an implicit ordering of the UEs is present.

2.2 Problem Formulation

Branch and Bound (BB) methods [78] are frequently used to solve MINLP problems. BB involves the process of relaxation which makes discrete variables continuous, branching which solves two sub-problems, one each for a lower bound and an upper bound on the objective, fathoming that rejects or accepts a solution. In the process a binary tree evolves and pivoting around various solutions occurs. In the worst case this enumeration could lead to an exponential computation complexity. So to reduce the complexity or size of the problem, the integer and the non-linear parts are evaluated separately. With this separable approach, applying well known IP and NLP

techniques becomes easier. The involved problems and sub-problems solved in this thesis are outlined next.

The mathematical formulation of a multi-cell multi-user sum-rate maximization with required QoS under perfect-CSI (CSI_p) on the UL is given as

$$\text{P1: } \underset{\boldsymbol{\alpha}, \mathbf{p}, \mathbf{U}}{\text{maximize}} \quad \sum_{j=1}^N \bar{r}_j \quad (2.12)$$

$$\text{subject to: } \quad (2.1), (2.2), (2.4), (2.10). \quad (2.13)$$

P1 for a given (\mathbf{p}, \mathbf{U}) results in a linear IP sub-problem P1_b in $\boldsymbol{\alpha}$ and for a given $\boldsymbol{\alpha}$ results in an NLP sub-problem P1_p in (\mathbf{p}, \mathbf{U}) .

$$\text{P1}_b: \quad \max_{\boldsymbol{\alpha}} \quad \sum_{j=1}^N \bar{r}_j \quad (2.14)$$

$$\text{s.t: } \quad (2.1), (2.2), \quad (2.15) \\ (2.10).$$

$$\text{P1}_p: \quad \max_{\mathbf{p}, \mathbf{U}} \quad \sum_{j=1}^N \bar{r}_j \quad (2.16)$$

$$\text{s.t: } \quad (2.4), (2.10). \quad (2.17)$$

The same rate maximization problem on the DL is given as

$$\text{P2: } \underset{\boldsymbol{\alpha}, \mathbf{Q}, \mathbf{U}}{\text{maximize}} \quad \sum_{j=1}^N r_j \quad (2.18)$$

$$\text{subject to: } \quad (2.1), (2.2), (2.7), (2.8), (2.11). \quad (2.19)$$

Similarly P2 has two sub-problems, a linear IP sub-problem P2_b for a given (\mathbf{Q}, \mathbf{U}) and an NLP sub-problem P2_p for a given $\boldsymbol{\alpha}$.

$$\text{P2}_b: \quad \max_{\boldsymbol{\alpha}} \quad \sum_{j=1}^N r_j \quad (2.20)$$

$$\text{s.t: } \quad (2.1), (2.2), \quad (2.21) \\ (2.8), (2.11).$$

$$\text{P2}_p: \quad \max_{\mathbf{Q}, \mathbf{U}} \quad \sum_{j=1}^N r_j \quad (2.22)$$

$$\text{s.t: } \quad (2.7), (2.8), \quad (2.23) \\ (2.11).$$

The problem of simple sum-rate is solved in P1 and P2 for $r_{th} = 0$, where in the interest of the system, some UEs may be dropped or may have a very low assigned rate. The sum-power minimization problem on the UL with required QoS under CSI_p is given as

$$\text{P3: } \underset{\boldsymbol{\alpha}, \mathbf{p}, \mathbf{U}}{\text{minimize}} \quad \sum_{j=1}^N p_j \quad (2.24)$$

$$\text{subject to: } \quad (2.1), (2.2), (2.4), (2.10). \quad (2.25)$$

Similar sub-problems P3_b and P3_p are formulated for P3.

$$\text{P3}_b: \quad \min_{\boldsymbol{\alpha}} \quad \sum_{j=1}^N p_j \text{ (constant)} \quad (2.26)$$

$$\text{s.t: } \quad (2.1), (2.2), \quad (2.27) \\ (2.10).$$

$$\text{P3}_p: \quad \min_{\mathbf{p}, \mathbf{U}} \quad \sum_{j=1}^N p_j \quad (2.28)$$

$$\text{s.t: } \quad (2.4), (2.10). \quad (2.29)$$

Objective for $P3_b$ is a constant where a solution satisfying constraints (2.27) is the only requirement. The sum-power minimization problem when applied on the DL is given as

$$P4: \quad \underset{\alpha, \mathbf{Q}, \mathbf{U}}{\text{minimize}} \quad \sum_{i=1}^M \sum_{j=1}^N \alpha_{ij} q_{ij} \quad (2.30)$$

$$\text{subject to:} \quad (2.1), (2.2), (2.7), (2.8), (2.11). \quad (2.31)$$

It can be observed that both P1 and P3 have the same constraint set, i.e., equations (2.13) and (2.25) are equal, only their objective changes. Same is the case for P2 and P4. Another problem of interest is the UL max-min problem P5 that is related to P3.

$$P5: \quad \underset{\alpha, \mathbf{p}, \mathbf{U}}{\text{maximize}} \quad \underset{j}{\text{minimize}} \quad \bar{r}_j \quad (2.32)$$

$$\text{subject to:} \quad (2.1), (2.2), (2.4). \quad (2.33)$$

In P3, for homogeneous UEs each with the same QoS requirement r_{th} , the constraint (2.10) will be active at the optimum, i.e., each UE will achieve the same QoS. At the optimum, P5 also produces a balanced QoS across all the homogeneous UEs, i.e., the power vector \mathbf{p} is the same for P5 and P3.

2.3 Problem Solving

2.3.1 Base Station Association

BSA sub-problems $P1_b$, $P2_b$ and $P3_b$ are linear IP problems which are Non-deterministic Polynomial-time (NP) hard. To improve the bounds on a linear IP, Lagrangian Relaxation (LR) [79] and decomposition methods [80] are usually preferred. Primal decomposition based on Bender's Decomposition (BD), dual decomposition based on Dantzig-Wolfe Decomposition (DWD) and cross decomposition [81] are the frequently used decomposition methods. Decomposition works on the separability of variables where a master-problem and a slave-problem are successively solved. $P2_b$ has a structure similar to a Generalized Assignment Problem (GAP) [82] which in itself is a special case of Multi-dimensional Multiple-choice Knapsack Problem (MMKP) [83]. MMKP and variants are used in the RA problems where several heuristic algorithms exist to find the solution. Upper bounds are obtained based on the dual LR problems where either constraints (2.2) or (2.11) are added to the objective in a weighted manner. These weights are the Lagrange multipliers. Relaxing assignment constraints (2.2) decomposes $P2_b$ into M simple 0 – 1 KP problems [84]. Non-differential subgradient based methods [85] or BB methods are employed to further solve this relaxed problem. If the capacity constraints (2.11) are relaxed, then $P2_b$ will possess an integrality property where a simple linear relaxation as in equation (2.34) for equation (2.1) is sufficient

to produce a 0 – 1 binary solution. This relaxation produces weaker bounds when compared to the former relaxation.

To solve P1_b and P3_b as an NLP problem, integer constraints (2.1) are relaxed as

$$\alpha_{ij} \in [0, 1], \forall i, \forall j, \quad (2.34)$$

such that, α_{ij} is a continuous variable. Other IP to NLP conversion techniques also exist [86]. In addition to the assignment constraints (2.2), further restrictions on α_{ij} such as

$$\alpha_{ij}(\alpha_{ij} - 1) = 0, \forall i, \forall j, \quad (2.35)$$

$$\alpha_{mj} \cdot \alpha_{nj} = 0, \forall m \neq n, \forall j, \quad (2.36)$$

must be applied to ensure it converges to either a 0 or a 1. The NLP problem is given as P1_{bn}. Equation (2.35) is an SOCP constraint, hence P1_{bn} can be solved as an SOCP problem. For the sake of mathematical formulation, an SDP formulation P1_{bs} for P1_{bn} is also possible. In the interval (2.34), constraints (2.35) always satisfy

$$\alpha_{ij}(\alpha_{ij} - 1) \leq 0, \forall i, \forall j. \quad (2.37)$$

By Schur complement [87], a positive semi-definite criteria is obtained as

$$\begin{bmatrix} \alpha_{ij} & \alpha_{ij} \\ \alpha_{ij} & 1 \end{bmatrix} \geq 0, \forall i, \forall j. \quad (2.38)$$

$$\begin{array}{ll} \text{P1}_{bn}: & \max_{\alpha} \sum_{j=1}^N \bar{r}_j \quad (2.39) \\ & \text{s.t:} \quad (2.2), (2.10), (2.40) \\ & \quad (2.35) \text{ or } (2.36). \end{array} \quad \begin{array}{ll} \text{P1}_{bs}: & \max_{\alpha} \sum_{j=1}^N \bar{r}_j \quad (2.41) \\ & \text{s.t:} \quad (2.2), (2.10), (2.42) \\ & \quad (2.38). \end{array}$$

Relaxed constraints (2.34) are not required in P1_{bn} and P1_{bs}. Either P1_{bn} or P1_{bs} can be used to solve P1_b. With only a change in the objective, the same constraint formulation also applies for P3_b.

Neighbourhood search and memory based Tabu Search (TS) algorithm [88] can also be used to solve IP problems when the search space is very large. Some random solutions, each an M -dimensional 0 – 1 binary vector are initialized in the search space. A set of solution vectors called neighbours is generated for each obtained solution at every iteration. A fitness or a cost function is evaluated at each neighbour, i.e., an optimization problem is solved. For e.g., a utility function based on SINR or rate can serve as a fitness function and a utility function based on power consumption or BER can be a cost function. For each solution and for a predefined number of iterations called tabu-period, a short-term memory called tabu-list maintains the previously visited solution vectors. Tabu-list solutions must be avoided to prevent cycling, i.e., revisiting the solution. A best non-tabu move is made from the current state to the neighbour. To

escape from getting stuck at a local optima, a move to a lower quality neighbour is also accepted. At each iteration the process is repeated and the best current solution from various states is stored globally. A rule called aspiration criteria can override the tabu-list for diversified solutions. The best available solution after maximum iterations or any other exit criteria is declared the output. Similar evolution based algorithm called Ant Colony Optimization (ACO) [89] can also be used to solve the GAP problem. The mathematical techniques outlined are also applicable to the sub-carrier allocation in multi-carrier systems if the 0 – 1 variable α_{ij} is identified as a sub-carrier association variable instead of a BSA variable.

2.3.2 Power Allocation and Beamforming

A beamforming sub-problem for a given \mathbf{p} or \mathbf{Q} and the power allocation sub-problem for a given \mathbf{U} are further obtained in the NLP sub-problems P1_p, P2_p and P3_p. Depending on the problem formulation several iterative methods exist to solve the NLP sub-problem, few of them are outlined next. With $b = (2^{rth} - 1)$ and $g_{ijk} = \|\mathbf{u}_{ij}^H \mathbf{h}_{ik}\|_2^2$, the non-linear UL QoS constraints (2.10) can be rearranged to get a set of linear equations for LP problem as

$$\frac{p_j \cdot g_{ijj}}{b} = \bar{\sigma}_j^2 + \sum_{k \neq j, k=1}^N p_k \cdot g_{ijk}, \forall j. \quad (2.43)$$

Constraints (2.43) can be iteratively solved via Fixed-Point Theory [90]. The total interference term $\mathcal{I}_{ij} = (\bar{\sigma}_j^2 + \sum_{k \neq j, k=1}^N p_k \cdot g_{ijk})$ in equation (2.43) is known as a Standard-Function [91] and has a Fixed-Point property that allows the iterations to converge to a solution called Fixed-Point. The same set of equations can also be arranged as a posynomial (not polynomial) [70] for GP constraints [92] as

$$\frac{b \cdot \bar{\sigma}_j^2}{g_{ijj}} p_j^{-1} + \sum_{k \neq j, k=1}^N \frac{b \cdot g_{ijk}}{g_{ijj}} p_j^{-1} p_k \leq 1, \forall j. \quad (2.44)$$

A Signomial Programming (SP) [93] objective (2.45) which is the ratio of two posynomials can be obtained for the objective in P1 with constraints (2.44). To solve the SP objective, an approach called a complementary GP with a technique called a condensation technique [94] must be employed. Starting from an initial evaluation point $\bar{p}_j, \forall j$, an objective (2.46), which is a ratio of a posynomial and a monomial is obtained.

$$\min_{\mathbf{p}} \prod_{j=1}^N \frac{\mathcal{I}_{ij}}{\mathcal{S}_{ij}}, \quad (2.45) \quad \min_{\mathbf{p}} \prod_{j=1}^N \frac{\mathcal{I}_{ij}}{\bar{\mathcal{S}}_{ij}^2 / \bar{\mathcal{S}}_{ij} \prod_{j=1}^N (c_{ij} p_j)^{c_{ij}}}. \quad (2.46)$$

$\mathcal{S}_{ij} = (\mathcal{I}_{ij} + p_j g_{ijj})$, \mathcal{S}_{ij} for a given $\bar{p}_j, \forall j$, is $\bar{\mathcal{S}}_{ij}$ and $c_{ij} = \bar{\mathcal{S}}_{ij} / (g_{ijj} \bar{p}_j)$. Both these methods are used for comparison in the numerical results of chapter 3.

With Difference in Convex (DC) programming [95], constraints (2.10) can be expressed as a difference of convex functions. It is expressed as

$$\log_2(p_j g_{ijj} + \mathcal{I}_{ij}) - \log_2(\mathcal{I}_{ij}) \geq r_{th}, \forall j. \quad (2.47)$$

A first order Taylor series is obtained for $\log_2(\mathcal{I}_{ij})$ and the resultant constraints are iteratively solved by well known convex techniques. The objective in P1 must also be changed as equation (2.47) in DC programming.

Population based search methods such as Particle Swam Optimization [96] (PSO), Genetic Algorithms [97] (GeA) find the optimum for NLP by simultaneously solving the fitness function in different search directions. The initial population must be diverse to cover most of the search space while iterating. PSO is based on the social behaviour of a flock of birds searching for food. In PSO, several random solutions called particles each with a fitness value, velocity and position metric are initialized. The velocity and position influence the search direction. At each iteration each particle's and the populations' best fitness is evaluated and the corresponding velocity and position are also updated. The populations' best fitness and its location in the search space are declared at convergence. GeA is based on the evolution of biological organisms. In GeA, several random solutions called chromosomes, each with a fitness value are initialized. Chromosomes with best fitness called parents are selected and a crossover is performed to obtain new offsprings each corresponding to a solution. Several selections and crossovers are performed to obtain diverse solutions. Each iteration is called a generation which ends at convergence. PSO and GeA can also be used to obtain the BSA solutions where each particle or chromosome is an M -dimensional 0 – 1 binary vector.

Simulated Annealing [98] (SA) based on the heating and gradual cooling in metallurgical process can also be used to obtain the solution. In SA, a cost function and a solution acceptance criteria are defined. A temperature parameter is initialized and gradually reduced to simulate the cooling process. Trial solutions are generated and the cost function is evaluated, i.e., an optimization problem is solved. A solution may be accepted, dropped or may be accepted with a probability. An exponential probability function is chosen that depends on the temperature and the cost variation in successive iterations. With every iteration the probability of rejecting bad solutions is increased. PSO, GeA and SA are evolutionary algorithms where at each iteration several optimization problems are solved to obtain the overall cost or fitness value.

Sequential Quadratic Programming (SQP) [99] is also well known to handle NLP problems. SQP replaces the objective with a quadratic approximation and each constraint with a linear approximation. At each iteration, a quadratic optimization problem is solved around the trial point obtained from previous iteration along with a search direction and required step size. Positive semi-definiteness of the matrix in the quadratic term of the objective must be maintained to solve the problem as a convex optimization problem such as SDP.

IPM [100] is the most famous technique to handle NLP problems. Several formulations of IPM exist, such as potential-reduction methods and central-path methods [101]. In either case a log-barrier function exists which restricts the initial solution and its iterates to the interior of feasible region. A constrained optimization problem in the given form is called a primal-problem for which a dual-problem can be formulated. Depending on the number of constraints, the primal-problem may be easier to handle in the dual domain. Iterates generated by IPM reduce the gap between the primal and dual solutions. Another form of IPM called infeasible-IPM, that has a greater DoF while problem solving also exists, where the initialized solution and subsequent iterates are not in the primal feasible region. At each iterate a search direction and a step size is required. More than one way exists to choose these parameters. A primal-dual based IPM framework used to solve the NLP problems throughout this thesis is explained in chapter 3. Another mathematical tool recently being used to solve problems in energy efficient networks is the Fractional Programming [102].

The beamformers that maximize the received SINR are the linear MMSE beamformers [103]. For a given α they are given as

$$\mathbf{u}_{ij} = \left(\bar{\sigma}_j^2 \mathbf{I}_T + \sum_{k=1, k \neq j}^N p_k \mathbf{h}_{ik} \mathbf{h}_{ik}^H \right)^{-1} \mathbf{h}_{ij}, \forall i, \forall j. \quad (2.48)$$

The MMSE decoder solution is the Generalized EigenValue Problem (GEVP) solution [104], i.e., finding the generalized eigenvector of $(\mathbf{h}_{ij} \mathbf{h}_{ij}^H p_j, \sum_{k=1, k \neq j}^N \mathbf{h}_{ik} \mathbf{h}_{ik}^H p_k + \bar{\sigma}_j^2 \mathbf{I}_T)$ results in the MMSE vectors (2.48). The order of matrices must be maintained while finding the eigenvector. For a joint transmit and receive BF with MMSE precoders and decoders, a Fixed-Point iteration [57] is also possible, but due to the multi-modal nature of the MSE function, convergence to a global optimum cannot be guaranteed. So an iterative technique can be applied to find the beamformers. The PA and BF sub-problems can be solved in a single step by an SDP approach, if $\sqrt{p_k} \mathbf{u}_{ij}$ is considered a single variable and the substitutions $\mathbf{B}_{ijk} = p_k \mathbf{u}_{ij} \mathbf{u}_{ij}^H$ and $\mathbf{R}_{ik} = \mathbf{h}_{ik} \mathbf{h}_{ik}^H$ are made. Further, the non-convex rank-1 constraints are relaxed to $\mathbf{B}_{ijk} \geq p_k \mathbf{u}_{ij} \mathbf{u}_{ij}^H$ to obtain a Semi-Definite Relaxation (SDR) problem. By Schur complement these relaxed constraints can be expressed as equation (2.50). Additional techniques based on randomization techniques [105] must be used to extract the rank-1 solution from the relaxed solution. The overall transceiver constraints are given as

$$\text{tr}(\mathbf{R}_{ij} \mathbf{B}_{ijj}) \geq \sum_{k=1, k \neq j}^N b \cdot \text{tr}(\mathbf{R}_{ik} \mathbf{B}_{ijk}) + b \cdot \bar{\sigma}_j^2, \forall j, \quad (2.49)$$

$$\begin{bmatrix} \mathbf{B}_{ijk} & \sqrt{p_k} \mathbf{u}_{ij} \\ \sqrt{p_k} \mathbf{u}_{ij} & \mathbf{I}_T \end{bmatrix} \geq 0, \forall j. \quad (2.50)$$

$\text{tr}(\cdot)$ is the trace operator. As mentioned before, several other BF techniques based on SGINR, MVDR, MLSE exist.

Chapter 3

Uplink Resource Allocation under perfect-CSI

In this chapter an Interior Point Method (IPM) framework that is used to solve a general Non-Linear Programming (NLP) problem is outlined with its application to sum-rate maximization problem P1 and sum-power minimization problem P3.

3.1 Algorithmic Framework

Consider a constrained optimization problem [106] \mathcal{P} in N_1 non-negative primal variables \mathbf{z} , and with a scalar objective $f_0(\mathbf{z})$. Constraints for UE $_j$ are arranged in a vector as $\mathbf{c}_j(\mathbf{z})$ such that $\mathbf{c}_j(\mathbf{z}) \geq 0$. For each $\mathbf{c}_j(\mathbf{z})$ a slack vector \mathbf{s}_j is included such that $\mathbf{s}_j \geq 0$. The overall constraint vector and slack vector of N UEs are $\mathbf{c}(\mathbf{z}) = [\mathbf{c}_1^T(\mathbf{z}), \mathbf{c}_2^T(\mathbf{z}), \dots, \mathbf{c}_N^T(\mathbf{z})]^T$ and $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_N^T]^T$, respectively. $\mathbf{c}(\mathbf{z})$ may have both equality constraints and inequality constraints.

$$\mathcal{P}: \quad \underset{\mathbf{z}}{\text{minimize}} \quad f_0(\mathbf{z}) \quad (3.1)$$

$$\text{subject to:} \quad \mathbf{c}(\mathbf{z}) - \mathbf{s} = 0. \quad (3.2)$$

For a maximization problem, $-f_0(\mathbf{z})$ is used in the objective of \mathcal{P} . Let N_2 be the total number of constraints in equation (3.2). The Lagrangian for \mathcal{P} is

$$\mathcal{L}(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}) = f_0(\mathbf{z}) - \boldsymbol{\lambda}^T(\mathbf{c}(\mathbf{z}) - \mathbf{s}). \quad (3.3)$$

$\boldsymbol{\lambda}$ is the non-negative vector of Lagrangian variables corresponding to the dual feasibility. Entities in $\boldsymbol{\lambda}$ for the corresponding equality constraints in equation (3.2) are unrestricted. The slack variables converge to values such that the non-negativity is ensured. For \mathcal{P} , a dual problem in dual decision vector $\boldsymbol{\lambda}$ is defined as

$$\mathcal{D}: \quad \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad \underset{\mathbf{z}, \mathbf{s}}{\text{infimum}} \quad \mathcal{L}(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}). \quad (3.4)$$

If \mathcal{P} is infeasible, then \mathcal{D} is unbounded and vice versa. \mathcal{D} provides a lower bound to \mathcal{P} and the difference in the objective values is the duality gap. For a zero duality gap, a strong duality is said to exist, i.e., \mathcal{P} and \mathcal{D} converge to the same solution. Let $c_n(\mathbf{z}) - s_n$ represent an individual constraint from equation (3.2) and λ_n be its corresponding dual multiplier. At the optimum $(\mathbf{z}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*)$, either the primal constraint $c_n(\mathbf{z}^*)$ or its dual multiplier λ_n^* is zero, i.e., $s_n^* \cdot \lambda_n^* = 0, \forall n$. This is called complementarity slackness. If the set $\nabla_{\mathbf{z}} c_n(\mathbf{z}^*)$ for the active constraints are independent then the condition is called Linear Independent Constraint Qualification (LICQ). The gradient operator w.r.t. \mathbf{z} is given as $\nabla_{\mathbf{z}}(\cdot)$. The implication of LICQ is that it ensures feasibility and the existence of unique dual multipliers. Also, $\nabla_{(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda})} \mathcal{L}(\mathbf{z}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*) = 0$. Together these necessary conditions called Karash-Kuhn-Tucker (KKT) conditions in vector form are

$$\mathbf{F} = \begin{bmatrix} \nabla_{(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda})} \mathcal{L}(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}) \\ \mathbf{c}(\mathbf{z}) - \mathbf{s} \\ \mathbf{D}_{\boldsymbol{\lambda}} \mathbf{D}_{\mathbf{s}} \mathbf{1}_{N_2} \end{bmatrix} = 0. \quad (3.5)$$

The first equation set in \mathbf{F} is the first-order derivative condition w.r.t. $[\mathbf{z}^T, \mathbf{s}^T, \boldsymbol{\lambda}^T]^T$. The second equation set is the primal feasibility condition and the third equation set is the complementarity slackness condition. $\mathbf{D}_{\mathbf{s}}$ and $\mathbf{D}_{\boldsymbol{\lambda}}$ are the diagonal matrices with \mathbf{s} and $\boldsymbol{\lambda}$ in their principal diagonal, respectively. For sufficiency condition, the Hessian $\nabla_{\mathbf{z}\mathbf{z}}^2 \mathcal{L}(\mathbf{z}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*)$ is verified and a minima occurs for a positive definite Hessian. For a convex constraint set the KKT conditions are also the sufficient conditions and the obtained optimum is also the global optimum.

A gradient based iterative algorithm [107] can be derived by perturbing the KKT vector (3.5). As a result, a feasible instance of \mathcal{P} gets solved as a set of linear equations only. With t as the iteration index and given initial values of \mathbf{z} , \mathbf{s} and $\boldsymbol{\lambda}$ the following steps constitute the algorithm. Search directions $[\Delta_{\mathbf{z}}^T, \Delta_{\mathbf{s}}^T, \Delta_{\boldsymbol{\lambda}}^T]^T$ called the Newton directions for the next iteration are evaluated as

$$\mathbf{J}_t [\Delta_{\mathbf{z}_{t+1}}^T, \Delta_{\mathbf{s}_{t+1}}^T, \Delta_{\boldsymbol{\lambda}_{t+1}}^T]^T = \mathbf{F}_t + [\mathbf{0}_{N_1}^T, -\mu_t \mathbf{1}_{N_2}^T, \mathbf{0}_{N_2}^T]^T, \quad (3.6)$$

where \mathbf{J} is the Jacobian of \mathbf{F} w.r.t. $[\mathbf{z}^T, \mathbf{s}^T, \boldsymbol{\lambda}^T]^T$, $\mu_t = \frac{\lambda_t^T \mathbf{s}_t}{N_2}$ is a complementarity measure. From a finite value, $\mu \rightarrow 0$ during the iterations. The system of equations in (3.6) form a perturbed KKT system. A certain set of continuous rows and columns in \mathbf{J} represent the Hessian $\nabla_{\mathbf{z}\mathbf{z}}^2 \mathcal{L}$. To calculate the direction, the steepest descent method uses the gradient, while the Newton method uses the Hessian. If the Hessian is unavailable, then quasi-Newton methods such as Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [69] can be used, where an approximation of the Hessian is used at each t . The variables are updated as

$$[\mathbf{s}_{t+1}^T, \mathbf{z}_{t+1}^T]^T = [\mathbf{s}_t^T, \mathbf{z}_t^T]^T + \delta_{\mathbf{s}_{t+1}} [\Delta_{\mathbf{s}_{t+1}}^T, \Delta_{\mathbf{z}_{t+1}}^T]^T, \quad (3.7)$$

$$\boldsymbol{\lambda}_{t+1} = \boldsymbol{\lambda}_t + \delta_{\boldsymbol{\lambda}_{t+1}} \Delta_{\boldsymbol{\lambda}_{t+1}}^T, \quad (3.8)$$

where $\delta_{\mathbf{s}_{t+1}}$ and $\delta_{\boldsymbol{\lambda}_{t+1}}$ are the step sizes that are a solution to

$$\max \left\{ \delta_{\mathbf{s}_{t+1}} \mid \delta_{\mathbf{s}_{t+1}} \in (0, 1], \mathbf{s}_{t+1} \geq (1 - \tau)\mathbf{s}_t \right\}, \quad (3.9)$$

$$\max \left\{ \delta_{\boldsymbol{\lambda}_{t+1}} \mid \delta_{\boldsymbol{\lambda}_{t+1}} \in (0, 1], \boldsymbol{\lambda}_{t+1} \geq (1 - \tau)\boldsymbol{\lambda}_t \right\}, \quad (3.10)$$

and the parameter $\tau \in (0, 1)$. Several ways to choose $\delta_{\mathbf{s}_{t+1}}$ and $\delta_{\boldsymbol{\lambda}_{t+1}}$ based on line search, such as Armijo condition, Goldstein condition, Wolfe condition [100], Zjoutendijk condition [108] exist. They are applied for unconstrained optimization problem and can be extended to \mathcal{P} while optimizing the unconstrained \mathcal{L} . The described sequence of steps are given in Algorithm 1. These iterations called Newton iterations are repeated till

Algorithm 1 : General iteration steps of IPM based on Newton directions.

- 1: **Initialize:** $t = 0, \mathbf{z}_t, \tau, (\boldsymbol{\lambda}_t, \mathbf{s}_t) > 0$
 - 2: **repeat**
 - 3: evaluate $\mathcal{L}(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}), \nabla_{(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda})} \mathcal{L}, \mu_t, \mathbf{F}_t, \mathbf{J}_t$
 - 4: $t = t + 1$
 - 5: solve (3.6) to obtain new search directions
 - 6: find step size from (3.9) and (3.10)
 - 7: evaluate (3.7) and (3.8) to update variables
 - 8: **until** convergence
 - 9: **Output:** $\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda}$.
-

the required convergence is met. The problem \mathcal{P} is assumed feasible and the involved functions are continuous and continuously differentiable. At $(\mathbf{z}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*)$, the conditions of strict complementarity and LICQ hold. The Lagrangian bound $\mathcal{L}(\mathbf{z}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*)$ is equal to the primal optimum $f(\mathbf{z}^*)$ as $\mu \rightarrow 0$ and $\nabla_{(\mathbf{z}, \mathbf{s}, \boldsymbol{\lambda})} \mathcal{L} \rightarrow 0$. Hence no duality gap between \mathcal{P} and \mathcal{D} . The problem is deemed infeasible if a predefined maximum number of iterations is exceeded, or when $\mathcal{L}(\mathbf{z}^*, \mathbf{s}^*, \boldsymbol{\lambda}^*) \rightarrow \infty$. These conditions can be used as the convergence and exit criteria for the algorithm. The problem is primal-infeasible till final convergence. But the iterates are interior to the non-negative orthant $(\mathbf{s}, \boldsymbol{\lambda})$. So the IPM is called a primal-dual infeasible interior-point method. It can be applied to a broad class of utility functions with little or no modification. More often choosing the initial solution may in itself be as hard as solving the original problem. With the IPM, the need for such a particular initialization is also eliminated, i.e., it eliminates the requirement of an initial primal feasible point to begin the algorithm

3.2 Rate Allocation

The explained IPM is applied to solve the sum-rate maximization problem P1 in this section. Problems P1_b and P1_p are alternately solved to obtain an optimum. Linear MMSE beamformers in equation (2.48) are used for beamforming.

3.2.1 Iterative Power Allocation

To solve sub-problem $P1_p$ by Algorithm 1, the primal variable \mathbf{z} is replaced by \mathbf{p} and the constraint vector for UE_j is

$$\mathbf{c}_j(\mathbf{p}) = \begin{bmatrix} P_{\max} - p_j \\ \bar{r}_j - r_{th} \end{bmatrix}. \quad (3.11)$$

The first constraint arises due to the upper bound on variable p_j . Optimization w.r.t. variable \mathbf{U} is implicit to $P1_p$, since \mathbf{U} is updated by the beamformers whenever \mathbf{p} changes.

3.2.2 Iterative Base Station Association

To solve $P1_b$ iteratively, i.e., $P1_{bn}$, primal variable \mathbf{z} is replaced by $\boldsymbol{\alpha}$. The constraint vector for UE_j is given as

$$\mathbf{c}_j(\boldsymbol{\alpha}) = \begin{bmatrix} -\boldsymbol{\alpha}_j \circ (\boldsymbol{\alpha}_j - 1) \\ 1 - \mathbf{1}_M^T \boldsymbol{\alpha}_j \\ \bar{r}_j - r_{th} \end{bmatrix}. \quad (3.12)$$

The constraint $(1 - \alpha_{ij})$ due to the upper bound on α_{ij} is not added to $\mathbf{c}_j(\boldsymbol{\alpha})$ since the inequality (2.37) is true only in this unit interval. The symbol \circ is the Hadamard product. The iterative steps given in Algorithm 2 solve the two stage problem, i.e., BSA and PA are solved alternately.

Algorithm 2 : Two stage formulation of P1

- 1: **Initialize:** \mathbf{p}
 - 2: **repeat**
 - 3: solve for $\boldsymbol{\alpha}$ by $P1_{bn}$ or $P1_{bs}$
 - 4: break on convergence
 - 5: solve for \mathbf{p} by $P1_p$
 - 6: break on convergence
 - 7: **until** final convergence
 - 8: **Output:** $\mathbf{p}, \boldsymbol{\alpha}$.
-

3.2.3 Single Stage Formulation

Algorithm 1 can be directly applied to solve P1 in a single stage, i.e., to solve $P1_b$ and $P1_p$ simultaneously for BSA and PA variables. It is given as

$$P1_s: \quad \underset{\boldsymbol{\alpha}, \mathbf{p}, \mathbf{U}}{\text{maximize}} \quad \sum_{j=1}^N \bar{r}_j \quad (3.13)$$

$$\text{subject to:} \quad (2.2), (2.4), (2.10), (2.34), (2.35). \quad (3.14)$$

The primal variables in this case are \mathbf{p} and $\boldsymbol{\alpha}$. Constraint vector $\mathbf{c}_j(\boldsymbol{\alpha}, \mathbf{p})$ is obtained by appending constraint $[P_{\max} - p_j]$ to constraint vector (3.12). Since the problem size is big when compared to the two stage formulation, both time and space complexity increase for each iteration.

Figures 3.1 to 3.4 show the convergence of various parameters for a single feasible instance of $P1_s$ under CSI_p , i.e., the intermediate Newton-iterations are shown. For illustration, the results are restricted to a minimum working example. For the link-level simulations, $T = 2$, $\tau = 0.9$, $P_{\max} = 0.1, \forall j$, $\bar{\sigma}_j^2 = 1$, $M = 2$, $N = 5$, $r_{th} = 0.2$ is set. The maximum number of iterations is set to 30. $M > 2$ can also be set, but it is easy to visualize a 2 dimensional 0–1 binary convergence plot as in Figure 3.3. With an increase in M or T , the number of active UEs can be increased for a given r_{th} , or for a given N , the threshold r_{th} can be increased. In either case the convergence trend remains the same as shown when $M = 2$, $T = 2$ is set. On the dB scale, $P_{\max} = 0.1$ is 20 [dBm]. This is close to the actual maximum available UL power of a UE, which is 23 [dBm]. But 20 [dBm] is good enough for the considered minimum working example. The channel vector $\mathbf{h}_{ij} \sim \mathcal{CN}(0, \mathbf{I}_T), \forall i, \forall j$, i.e., it follows a complex normal distribution with zero-mean and unit-variance. Elements of $\mathbf{h}_{ij}, \forall i, \forall j$, are independent and identically distributed (i.i.d). The absolute value $|\mathbf{h}_{ij}|, \forall i, \forall j$, is Rayleigh distributed which is the modelling of the assumed wireless fading channel. Vectors \mathbf{s} and $\boldsymbol{\lambda}$ are initialized to unit vectors with all elements equal to 1. Initial values $p_j = 0.2, \forall j$, $\alpha_{ij} = 1.5, \forall i, \forall j$ are set, i.e., these variables p_j and α_{ij} are initialized to the values outside their respective intervals (2.4) and (2.34).

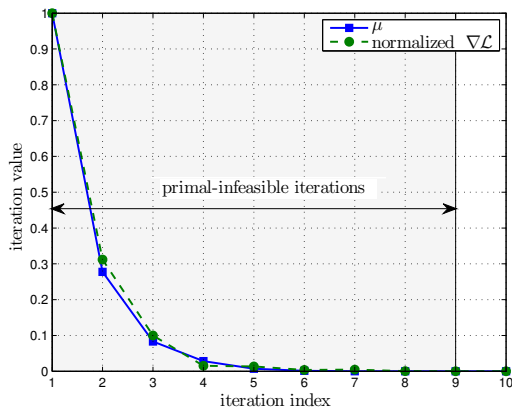


Fig. 3.1: Convergence of μ and $\nabla \mathcal{L}$.

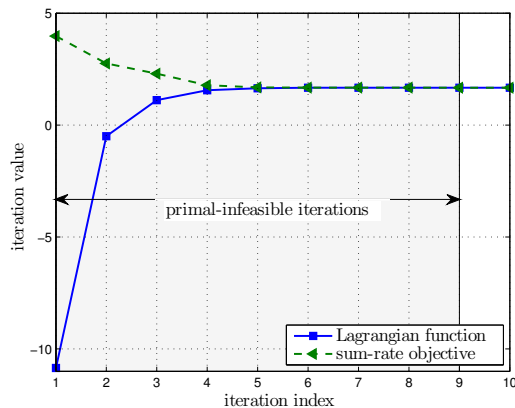


Fig. 3.2: Convergence of \mathcal{L} and sum-rate objective.

Figure 3.1 plots the convergence of the complementarity measure μ and gradient of the Lagrangian $\nabla \mathcal{L}$. Their convergence to the value 0 shows that KKT conditions are satisfied at the optimal. The convergence of either variable need not be monotonic. Figure 3.2 plots the convergence of the Lagrangian \mathcal{L} and the objective of $P1_s$. Convergence of both plots to the same value shows a zero duality gap between \mathcal{P} and \mathcal{D}

in the problem $P1_s$. Figure 3.3 plots the convergence of the BSA variable α . Since

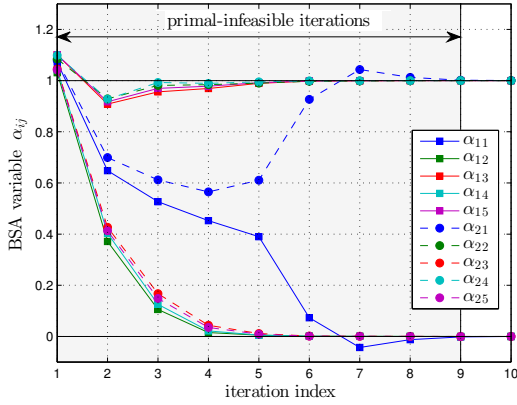


Fig. 3.3: Convergence of $\alpha_j, \forall j$.

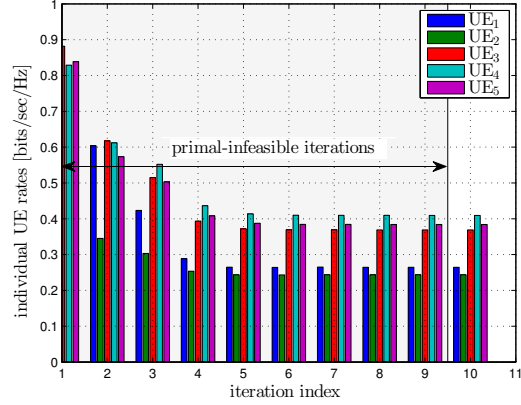


Fig. 3.4: Convergence of $\bar{r}_j, \forall j$.

$M = 2$ one element in each α_j converges to the value 1 corresponding to the associated BS, while the other element in it converges to a 0 thereby satisfying the 0 – 1 integer constraint (2.1) and the assignment constraint (2.2). So additional recovery of the 0 – 1 solution for the BSA variables is not required. Figure 3.3 plots the convergence of UE rates $\bar{r}_j, \forall j$. Each UE achieves a rate greater than the required r_{th} . This satisfies the QoS constraint (2.10). So formulation $P1_s$ solves the NP-hard MINLP problem $P1$ as an NLP problem and obtains a solution to the discrete decision matrix α and the continuous decision vector \mathbf{p} simultaneously. It can be observed that the intermediate iterates of $P1_s$ are all primal infeasible. Hence $P1_s$ is primal-infeasible till the final convergence. Only the non-negativity of \mathbf{s} and λ must be ensured throughout the Newton-iterations.

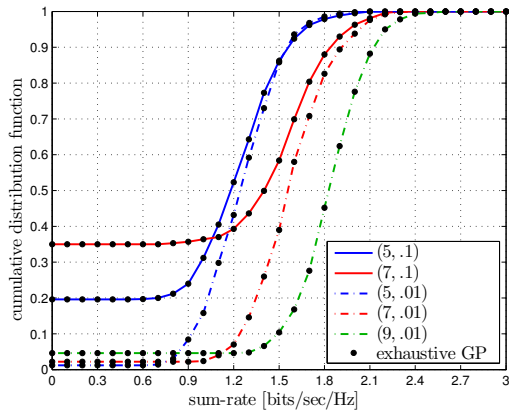


Fig. 3.5: Comparison of $P1_{bs}$ - $P1_p$ and exhaustive-GP.

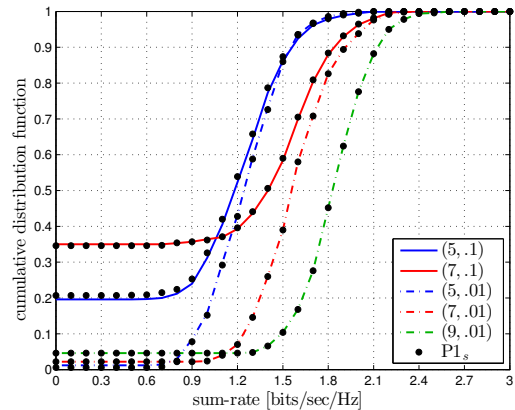


Fig. 3.6: Comparison of $P1_{bs}$ - $P1_p$ and $P1_s$.

Figure 3.5 compares the two stage formulation $P1_{bs}$ - $P1_p$ with an exhaustive method for $M = 2$ and various (N, r_{th}) by plotting the Cumulative Distribution Function (CDF).

Other parameters remain the same as before. PA in the $P1_{bs}$ - $P1_p$ combination uses the IPM. In the compared GP formulation [109], the PA uses equations (2.44) and (2.46), i.e., with single condensation method [110] for all the M^N possible combinations of BSA. Not every BSA-UE combination is feasible and more than one feasible combination is also possible in the exhaustive search. The feasible combination with the maximum sum-rate objective value is taken as the reference for comparison. It can be seen that the curves match for both the methods, i.e., same objective value at the same BSA. Since there is a possibility to achieve the same sum-rate value at more than one BS, $P1_{bs}$ is also verified by IP solver CPLEX. The obtained α in both the cases is the same. To reduce the number of outer iterations, implementing $P1_{bs}$ is better than $P1_{bn}$. For a fixed N , with increase in r_{th} , the CDF curves shift to the left showing a drop in the objective. The increased QoS requirement of each UE increases the MUI at every other UE. As a result, individual power levels are increased to overcome the MUI and attain the QoS objective. For a fixed r_{th} , with increase in N , the CDF curves move to upward. This is due to the increase in number of infeasible problem instances, i.e., increase in the system outage. Figure 3.6 compares the two stage formulation $P1_{bs}$ - $P1_p$ with its single stage formulation $P1_s$ for $M = 2$ and various (N, r_{th}) . It can be seen that either formulation leads to the same objective value. As expected, a right shift in the curves with increasing N at a fixed rate r_{th} and an upward shift in the curves with increase in r_{th} for a fixed N is observed. Both these show the competing nature of the UEs and also the increased interference with QoS and N . There is a slight variation in the $P1_s$ curve for the case (0.5,0.1). This variation is due to the chosen initial values that resulted in a convergence beyond the predefined maximum number of iterations. But under the assumption, the algorithm has to terminate if this number is exceeded. To solve P1 as an NLP problem, a single stage formulation $P1_s$, or two stage formulations $P1_{bs}$ - $P1_p$, or $P1_{bn}$ - $P1_p$ can be chosen.

3.3 Power Control

Similar to $P1_s$, sum-power minimization problem P3 can be solved with only a change in the objective, it is given as

$$P3_s: \quad \underset{\alpha, \mathbf{p}, \mathbf{U}}{\text{minimize}} \quad \sum_{j=1}^N p_j \quad (3.15)$$

$$\text{subject to:} \quad (2.2), (2.4), (2.10), (2.34), (2.35). \quad (3.16)$$

Since the constraint set for $P1_s$ and $P3_s$ is the same, other aspects applied for solving $P1_s$ are retained, i.e., $\mathbf{c}_j(\alpha, \mathbf{p})$ fed to the IPM is the same. The problem of sum-power minimization under SINR constraints is coupled to the problem of maximizing the minimum SINR, i.e., the max-min problem where the objective is to achieve the same SINR level across all the UEs. The feasibility of the later leads to formulation of the former [111]. Solution to the max-min problem is equivalent to solving the PC problem.

Some properties of the interference constraints \mathcal{I}_{ij} are identified [112] based on Fixed-Point Theory. This establishes the existence of iterative algorithms to solve the PC problem. The asynchronous iterative method under constrained power [113], under sum-power constraint [111] is well explained for the PC problem. However maximum possible threshold must be evaluated before applying the framework. In this approach, \mathbf{p} is initialized to a random point preferably within its bounds. Then the constraint (2.43) is evaluated for each UE as $p_j = \mathcal{I}_{ij}/g_{ijj}, \forall j$, to get one power variable at a time. These iterations are repeated and will converge if the feasibility is met as defined by Perron-Frobenius Theory. The matrix form representation of the LP constraints (2.43) for a given α is given as

$$(\mathbf{I}_N - \mathbf{A})\mathbf{p} \geq \mathbf{v}. \quad (3.17)$$

\mathbf{A} is an $(N \times N)$ non-negative, irreducible matrix. The diagonal elements $\mathbf{A}_{jj} = 0$ and other elements $\mathbf{A}_{jk} = b \cdot g_{ijk}/g_{ijj}$. Elements in the $(N \times 1)$ vector \mathbf{v} are given as $v_j = b \cdot \bar{\sigma}_j^2/g_{ijj}$. The PC sub-problem $P3_p$ is feasible only if the eigenvalue of \mathbf{A} is real, positive and less than 1. With these conditions, $(\mathbf{I}_N - \mathbf{A})$ is invertible and a positive power vector exists. The transmit power increases without bound as each UE approaches the maximum attainable threshold. However the upper bound on each p_j must also be taken into consideration to evaluate the feasible QoS. The structure of the coupling matrix \mathbf{A} is different when the sum-power constraint is considered. So either LP problem (3.17) or the asynchronous iterations based on Fixed-Point Theory can be used to solve $P3_p$. To extend these to a multi-cell case, they must be solved for every BSA [114]. As the number of network entities increase, this approach is practically not viable. PC with macro-diversity [115] is also an interesting problem, however the level of coordination among the BSs increases to perform joint processing. Other extensions based on PC are also possible [116].

3.3.1 Feasibility

If the problem instance $P3$ or $P3_s$ is infeasible, then, to satisfy the eigenvalue criteria of \mathbf{A} , either some UEs are dropped or the QoS requirement is reduced. Branch and Bound techniques to select the QoS-satisfying UEs or dropping QoS-violating UEs exist with a reasonable performance [117]. It can be viewed as a sub-problem of admission control in $P3$. The admission control will maximize the cardinality of the active UE set and these UEs will be balanced due to homogeneity. Another way to define feasibility is that if an instance of $P3$ is infeasible then it is reasonable to serve as many UEs as possible that satisfy the QoS while still serving the QoS-violating UEs with the best possible rate capable of being achieved by them instead of dropping them. Again the selection and the dropping of UEs, which is an IP problem, needs to be evaluated. This is an admission control problem. To handle the infeasibility of $P3$, a formulation based on the ℓ_1 -norm heuristic [118] is applied. This is a single stage NLP formulation similar to $P1_s$ which also explicitly defines the QoS-satisfying and the QoS-violating

UE subsets among the UEs. A new vector $\boldsymbol{\nu} = [\nu_1, \dots, \nu_N]^T$ is introduced, such that $\nu_j \geq 0, \forall j$. Further, QoS constraint (2.10) is modified as

$$\nu_j + \bar{r}_j \geq r_{th}, \forall j. \quad (3.18)$$

The reformulated NLP problem for P3 that handles infeasibility is

$$\text{P3}_f: \quad \underset{\boldsymbol{\alpha}, \mathbf{p}, \mathbf{U}, \boldsymbol{\nu}}{\text{minimize}} \quad \sum_{j=1}^N p_j + \theta \|\boldsymbol{\nu}\|_1 \quad (3.19)$$

$$\text{subject to:} \quad (2.2), (2.4), (2.34), (2.35), (3.18). \quad (3.20)$$

The parameter θ is a large positive constant, and $\|\cdot\|_1$ is the ℓ_1 -norm. The constraint vector for the IPM is

$$\mathbf{c}_j(\boldsymbol{\alpha}, \mathbf{p}, \boldsymbol{\nu}) = \begin{bmatrix} -\boldsymbol{\alpha}_j \circ (\boldsymbol{\alpha}_j - 1) \\ 1 - \mathbf{1}_M^T \boldsymbol{\alpha}_j \\ \nu_j + \bar{r}_j - r_{th} \\ P_{\max} - p_j \\ p_j \end{bmatrix}. \quad (3.21)$$

Unlike the other variables, the lower bound on p_j has to be included in (3.21). For a feasible instance, P3_f and P3_s give the same result with $\nu_j = 0, \forall j$, while for an infeasible instance the number of QoS violations are minimized due to the sparsity requirement of the ℓ_1 -norm. P1_f is always feasible and the solution w.r.t. the QoS-satisfying UEs is balanced w.r.t. \bar{r}_j .

Figures 3.7 to 3.10 plot the convergence of various UE parameters for a feasible channel instance of P3_s as a special case of P3_f. For the minimum working example, $M = 2$, $N = 4$, $r_{th} = 0.2$, $P_{\max} = 20$ [dBm], $\theta = 1e2$ is set. Other parameters remain the same as in P1_s. Figure 3.7 plots the convergence of the BSA variable $\boldsymbol{\alpha}_j, \forall j$. As before only

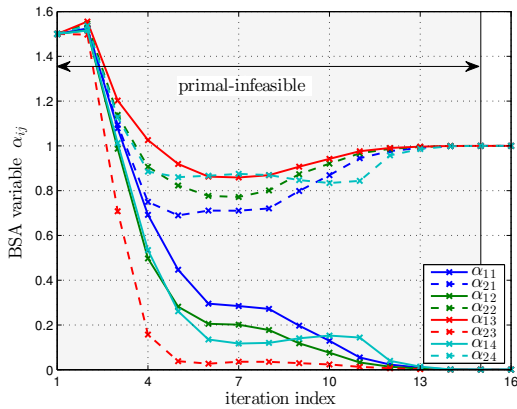


Fig. 3.7: Convergence of $\boldsymbol{\alpha}_j, \forall j$, in P3_f for a feasible P3_s.

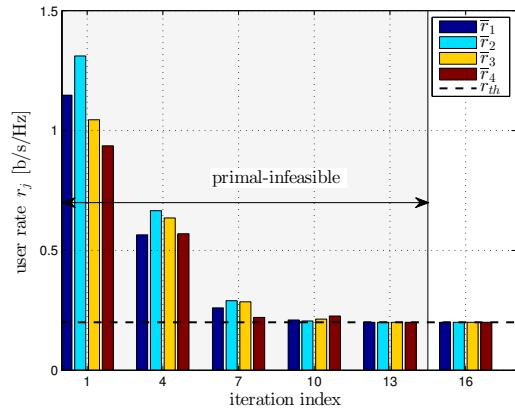


Fig. 3.8: Convergence of $\bar{r}_j, \forall j$, in P3_f for a feasible P3_s.

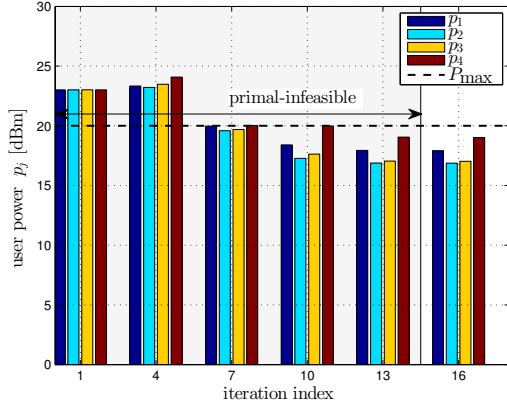


Fig. 3.9: Convergence of $p_j, \forall j$, in $P3_f$ for a feasible $P3_s$.

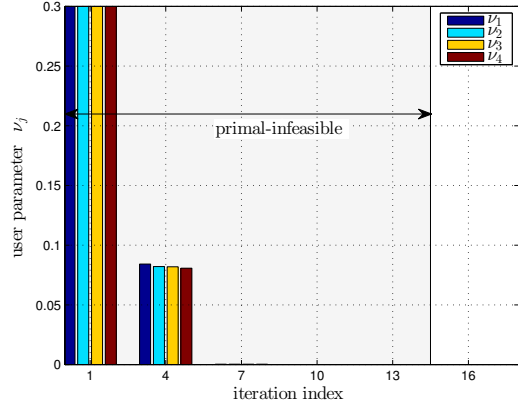


Fig. 3.10: Convergence of $\nu_j, \forall j$, in $P3_f$ for a feasible $P3_s$.

one element in each α_j converges to a 1, while other element converge to a 0. Figure 3.8 plots the convergence of the UE rates $\bar{r}_j, \forall j$. A balanced QoS is achieved across all the homogeneous UEs, i.e., $\bar{r}_j = r_{th}, \forall j$. Figure 3.9 plots the convergence of the PA variable $p_j, \forall j$. Each UE is transmitting at a power level lower than its corresponding P_{\max} while satisfying equation (2.10). Hence the objective of minimization of total transmit power in the system is achieved. Figure 3.10 plots the convergence of the ℓ_1 -norm parameter $\nu_j, \forall j$. Since the problem instance is feasible, $\nu_j, \forall j$, converges to zero, i.e., the ℓ_1 -norm parameter does not affect the objective $P3_f$. Solving $P3_f$ is the same as solving $P3_s$ for a feasible case.

Figures 3.11 to 3.14 plot the convergence of various UE parameters for the same setup as before. The only change is that the threshold is set to $r_{th} = 0.3$. An increase in the QoS makes $P3_s$ infeasible, however $P3_f$ is still feasible based on the new definition of feasibility. Figure 3.11 shows the convergence of the BSA variable $\alpha_j, \forall j$. Once

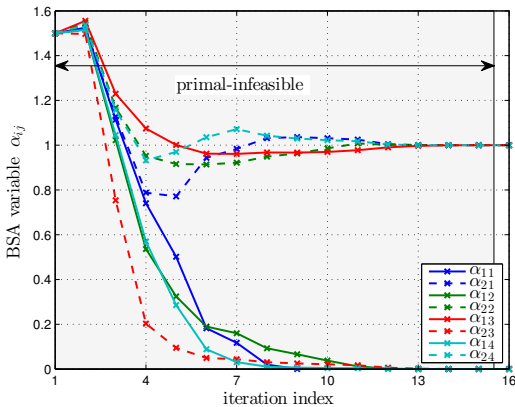


Fig. 3.11: Convergence of $\alpha_j, \forall j$, in $P3_f$ for an infeasible $P3_s$.

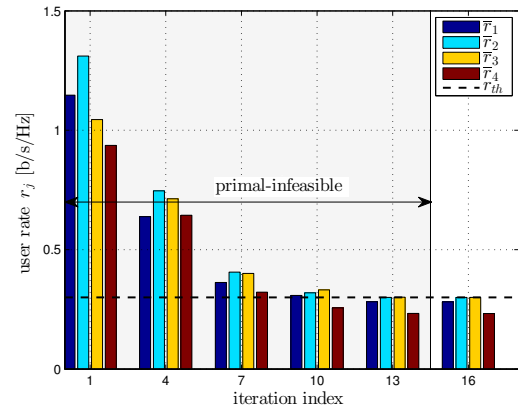


Fig. 3.12: Convergence of $\bar{r}_j, \forall j$, in $P3_f$ for an infeasible $P3_s$.

again only one element in each α_j , will converge to the value 1 while the other element

converge to a 0. In Figure 3.12, it can be observed that at convergence, UE₂ and UE₃ achieve the required QoS while UE₁ and UE₄ instead of being dropped receive a rate value lower than the QoS. So only two UEs satisfy the QoS constraint and the QoS-violating UEs achieve the best possible rate under infeasibility. It is evident that formulation P3_f, without dropping QoS-violating UEs or reducing the QoS value, handles an infeasible P3_s effectively. The initial UE selection sub-problem is eliminated in P3_f. The QoS-satisfying UEs are homogeneous, hence a balanced QoS is achieved among them. From Figure 3.13 it can be seen that the QoS-violating UEs are trans-

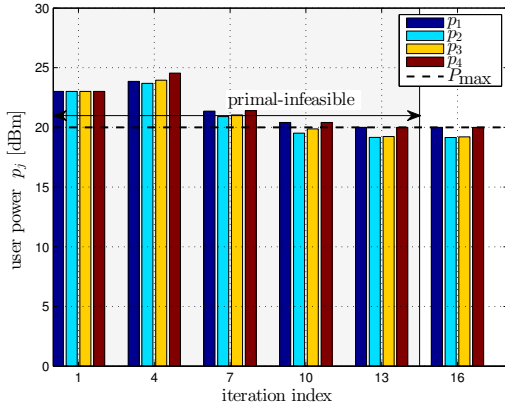


Fig. 3.13: Convergence of $p_j, \forall j$, in P3_f for an infeasible P3_s.

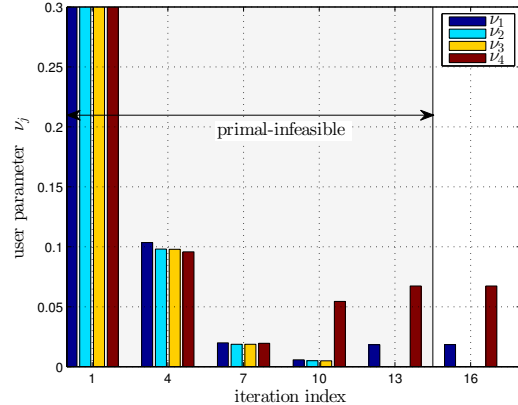


Fig. 3.14: Convergence of $\nu_j, \forall j$, in P3_f for an infeasible P3_s.

mitting at maximum power P_{\max} , i.e., $p_1 = p_4 = P_{\max}$. This increases the overall power level in the system when compared to serving only UE₂ and UE₃ while turning off UE₁ and UE₄. The QoS-achieving UEs are transmitting at a power level lower than the maximum value. However it could be possible that for achieving the QoS, a feasible UE may require P_{\max} . So focusing only on p_j cannot determine the feasibility. Figure 3.14 plots the convergence of the ℓ_1 -norm vector $\boldsymbol{\nu}$. It has a finite value for the QoS-violating UEs. The parameter ν_j for UE₂ and UE₃ is 0 as in the feasible case whereas for the infeasible UE₁ and UE₄ the value is finite. The larger the ν_j value, the lower the corresponding received rate value \bar{r}_j . So the set of feasible UEs can be determined by verifying the values in $\boldsymbol{\nu}$ in addition to the corresponding \bar{r}_j . Due to equation (3.18), P3_f achieves a balanced result w.r.t. the sum of \bar{r}_j and ν_j . Depending on the \bar{r}_j , the corresponding ν_j can take any positive value. The actual upper bound on ν_j is r_{th} and it is the case when the corresponding $\bar{r}_j = 0$. Hence an explicit upper limit on ν_j is not included in the constraint set. So $\nu_j, \forall j$, not only identifies the QoS-violating UEs but also specifies the difference between the achieved QoS and r_{th} , i.e., the amount by which the infeasibility is caused in the QoS constraint.

Figure 3.15 shows that the increasing average infeasibility with increasing N and also with r_{th} in P3_s. For the same channel realizations, Figure 3.16 compares the average number of UEs served for P3_s and P3_f. P3_f clearly outperforms P3_s by serving more

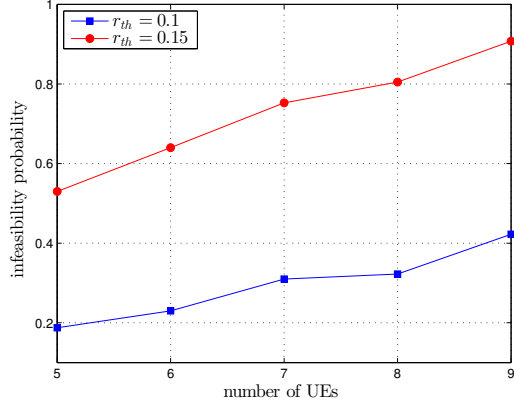


Fig. 3.15: Infeasibility increase with N and r_{th} for $P3_s$.

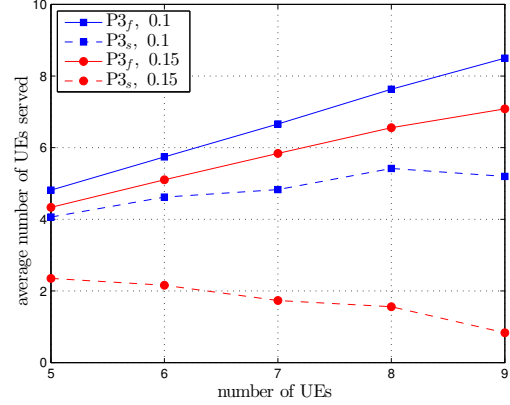


Fig. 3.16: Average number of UEs served for $P3_s$ and $P3_f$.

QoS-satisfying UEs on an average. A significant performance improvement is observed for the $r_{th} = 0.15$ curve.

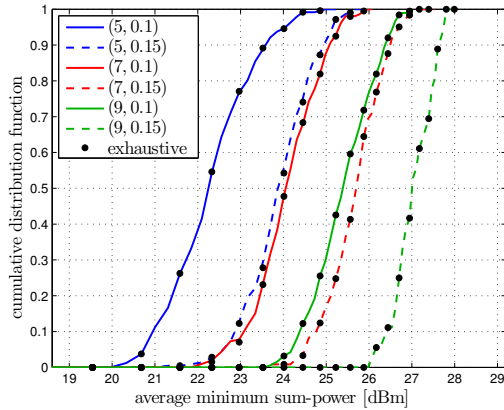


Fig. 3.17: CDF comparison of $P3_s$ and iterative exhaustive case.

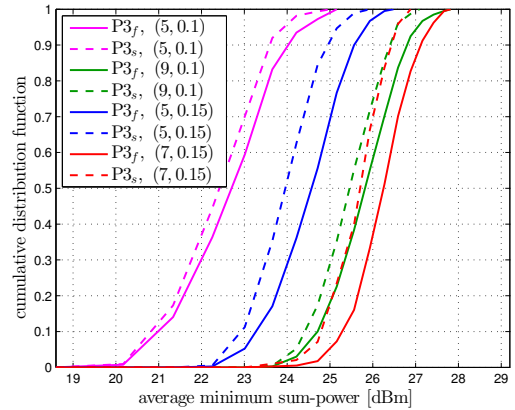


Fig. 3.18: Comparison of $P3_s$ and $P3_f$.

Figure 3.17 compares the CDF of $P3_s$ and an exhaustive method based on Fixed-Point Theory for various (N, r_{th}) , $M = 2$, $P_{\max} = 23$ [dBm]. For numerical verification, other arbitrary values of P_{\max} can also be used. The exhaustive search considers all possible M^N BSA-UE combinations in $P3_b$ and solution to $P3_p$ is obtained by the asynchronous iterative approach outlined before. More than one BSA-UE combination may be feasible and the one with the least sum-power objective in each feasible instance is taken as the reference to compare the performance of $P3_s$. It can be seen that the CDF of $P3_s$ matches with the CDF obtained by exhaustive search. Also the obtained α in both cases is the same. With increase in r_{th} , each UE increases its power level to maintain the required QoS. The increase in the right shift of the curve is reduced with increase in N or r_{th} due to the increase in the number of infeasible realizations. Figure 3.18 compares the CDF of $P3_s$ and $P3_f$ for various (N, r_{th}) and same set of

parameters as before. Only the actual transmit sum-power in the $P3_f$ objective must be considered rather than the complete $P3_f$ objective. Since the non-binding UEs are transmitting at full power, and the number of served UEs is more in $P3_f$, an increase in the objective value of $P3_f$ over $P3_s$ is observed. This is reflected by a right shift in the $P3_f$ CDF curve corresponding to the $P3_s$ curve.

Chapter 4

Uplink Power Control under statistical-CSI and imperfect-CSI

When perfect Channel State Information (CSI_p) is unavailable or when a small outage is permitted in the service, the deterministic constraints can be replaced by Probabilistic Constraints (PrCons) as a compromise. PrCons may at times be simplified to get a closed form expression [119], however an approximation function is required for this expression to further solve the problem. Obtaining such function may not be feasible all the time. Intractability and unavailability of closed-form expressions make PrCons mathematically difficult to handle. In this chapter, Uplink (UL) power control problems with PrCons under statistical-CSI (CSI_s) and imperfect-CSI (CSI_i) are solved for a multi-cell multi-user scenario. Existing bounds are further improved and the problem is well handled. To handle the PrCons and analyse the system performance, financial risk management measures Value-at-Risk and Conditional Value-at-Risk are applied. The resulting expressions may involve functions that are non-invertible and a combination of higher order functions, so obtaining closed form solutions may not be possible. To find the optimal value of the optimization problem Extreme Value Theory is applied. Worst case performance for a given probability is also evaluated in the process.

The problem of Power Control (PC) under CSI_p and deterministic constraints of the form $[\phi(y) \leq \zeta]$ in the decision variable y is well understood [120]. $\phi(\cdot)$ is usually a loss function in SINR and ζ is a constant. Implementing the deterministic constraints may be too costly or impossible at times. Further, CSI is partially available or is imperfect due to limited feedback or improper channel estimation. This affects the system's outage performance. In this case, PrCons of the form $\Pr[\phi(y, g) \leq \zeta] \geq \eta$, called chance constraints in uncertain variable g , may be implemented as a compromise [121]. $\Pr[\cdot]$ is the probability function and η is the confidence interval. Computationally tractable and conservative approximations are required to solve the chance constrained optimization problems. This approximation results in a solvable deterministic optimization problem with the feasible set contained in the original chance constrained problem. After obtaining the deterministic equivalent, standard deterministic optimization techniques

based on subgradients [122], Fixed-Point Theory based iterative methods [123] or SDR [124] can be applied for finding the decision vector. In this chapter it is assumed that for a UE $_j$ and a given $\boldsymbol{\alpha}$, the serving channel $\mathbf{h}_{ij}, \forall j$, is perfectly known.

4.1 Power Control under statistical-CSI

Under CSI $_s$, for a UE $_j$ and given $\boldsymbol{\alpha}$, the interfering channels $\mathbf{h}_{ik}, \forall k \neq j$, are only statistically known. Under these assumptions, the BS will have CSI $_p$ of the interfering channels for the UEs with the same BSA as UE $_j$, i.e., $\{\text{UE}_n \mid n \neq j, \alpha_{in} = \alpha_{ij}\}$. But to generalize the case, it is assumed that it is still statistically known. The uncertain variable for a given $\boldsymbol{\alpha}$ are the effective channel gains $g_{ijk} = \|\mathbf{u}_{ij}^H \mathbf{h}_{ik}\|_2^2, \forall k \neq j$ in the MUI component. So only the distribution of g_{ijk} in the MUI term is available. The statistical-SINR $\bar{\gamma}_j^{(s)}$ for the deterministic parameter in equation (2.5) is given as

$$\bar{\gamma}_j^{(s)} = \frac{p_j \sum_{i=1}^M \alpha_{ij} \|\mathbf{u}_{ij}^H \mathbf{h}_{ij}\|_2^2}{\sum_{i=1}^M \sum_{k \neq j, k=1}^N p_k \alpha_{ik} g_{ijk} + \bar{\sigma}_j^2}, \forall j. \quad (4.1)$$

For the deterministic QoS constraints (2.10), the PrCons under CSI $_s$ are given by

$$\Pr \left[\bar{r}_j^{(s)} > r_{th} \right] \geq \eta, \forall j, \quad (4.2)$$

where $\bar{r}_j^{(s)} = \log_2(1 + \bar{\gamma}_j^{(s)})$ is the statistical-rate for UE $_j$. With the substitutions

$$y_j^{(s)} = \sum_{i=1}^M \sum_{k \neq j, k=1}^N p_k \alpha_{ik} g_{ijk}, \quad \forall j, \quad (4.3)$$

$$\zeta_j^{(s)} = \frac{p_j}{b} \sum_{i=1}^M \alpha_{ij} g_{ijj} - \bar{\sigma}_j^2, \quad \forall j, \quad (4.4)$$

equation (4.2) can be rearranged as constraint (4.5) or (4.6).

$$F_{Y_j^{(s)}}(\zeta_j^{(s)}) \geq \eta, \forall j, \quad (4.5) \quad \zeta_j^{(s)} \geq F_{Y_j^{(s)}}^{-1}(\eta), \forall j. \quad (4.6)$$

$F_{Y_j^{(s)}}$ is the Cumulative Distribution Function (CDF) and $F_{Y_j^{(s)}}^{-1}(\eta)$ is the quantile function of the random variable (r.v.) $Y_j^{(s)}$. The considered multi-cell multi-user chance constrained MINLP optimization problem under CSI $_s$ is given by

$$\text{P3}_c: \underset{\boldsymbol{\alpha}, \mathbf{p}, \mathbf{U}}{\text{minimize}} \quad \sum_{j=1}^N p_j \quad (4.7)$$

$$\text{s.t:} \quad (2.1), (2.2), (2.4), (4.6). \quad (4.8)$$

The optimal value $\zeta_j^{(s)} = \bar{\zeta}_j^{(s)}$ is obtained when equation (4.6) is satisfied with equality, i.e., $\bar{\zeta}_j^{(s)}$ is η -quantile of r.v. $Y_j^{(s)}$. The analytic expression of the CDF of the r.v. must exist and be invertible to efficiently solve P3 $_c$. Due to the Pr[\cdot] function in constraint (4.2), the decision vectors $\boldsymbol{\alpha}$ and \mathbf{p} must be evaluated simultaneously.

4.2 Value-at-Risk and Conditional Value-at-Risk

In the field of finance, risk management measures are required for loss estimation. Such commonly used measures are Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) [125]. These metrics can be applied to the PC problem, if for each UE, the loss function $\phi(\cdot)$ is mapped to the MUI. Under CSI_s and a given η , VaR satisfies the chance constraint (4.6) and CVaR gives the worst case system performance. Worst case implies the possible performance loss beyond the evaluated optimal VaR. To evaluate VaR, the CDF of $\phi(y, g)$ is required. It may not be invertible. Further, to evaluate CVaR, the Probability Density Function (PDF) of $\phi(y, g)$ is needed. A closed form expression of CVaR may be unavailable if the PDF has higher order functions, such as a combination of Bessel and exponential functions. Approximations only give a bound which may not be tight enough to predict the exact performance.

For a UE, the MUI or more specifically $y_j^{(s)}$ is identified as the loss function. Then by the definition of VaR, a UE will experience a loss greater than VaR with probability $(1 - \eta)$. Or the MUI will not exceed VaR with probability η . Mathematically it is written as

$$\text{VaR}(y_j^{(s)}, \eta) = \inf\{y_j^{(s)} \mid \Pr[y_j^{(s)} \leq \zeta_j^{(s)}] \geq \eta\}, \forall j. \quad (4.9)$$

Equation (4.5) can be mapped to this VaR definition. Since there still exists a finite probability of exceedance beyond the maximum evaluated loss, it is further possible that for a given η , the loss $y_j^{(s)}$ exceeds $\bar{\zeta}_j^{(s)}$. So a worst case performance loss can be estimated. However the mathematical formulation of the worst case problem considers the mean value of the excess loss. The mean value of the exceedance of $y_j^{(s)}$ over $\bar{\zeta}_j^{(s)}$ is called the CVaR. It is given by the conditional expectation

$$\beta_j^{(s)} = \text{E} \left[y_j^{(s)} \mid y_j^{(s)} \geq \bar{\zeta}_j^{(s)} \right], \forall j. \quad (4.10)$$

The PDF of r.v. $Y_j^{(s)}$ is required to evaluate equation (4.10). Mathematically this worst case MINLP optimization problem is given as

$$\text{P3}_w: \underset{\alpha, \mathbf{p}, \mathbf{U}}{\text{minimize}} \quad \sum_{j=1}^N p_j \quad (4.11)$$

$$\text{s.t:} \quad (2.1), (2.2), (2.4), (4.10). \quad (4.12)$$

It can be observed that P3_w is not a chance constrained optimization problem. The outcome of P3_c is used as the input parameter to P3_w via equation (4.10). It implies, $\beta_j^{(s)} \geq \bar{\zeta}_j^{(s)}, \forall j$, i.e., CVaR can never be less than VaR. So in terms of loss at a given η , CVaR can be viewed as a loss value exceeding the already evaluated maximum loss, which is the VaR. Hence CVaR is a worst case loss value in system performance. It is mathematically possible to calculate $\bar{\zeta}_j^{(s)}$ and $\beta_j^{(s)}$ simultaneously [125] from a single function

$$\mathcal{F} = \zeta_j^{(s)} + \frac{1}{(1-\eta)} \int_{y_j^{(s)} > \zeta_j^{(s)}} (y_j^{(s)} - \zeta_j^{(s)}) \Psi_{Y_j^{(s)}}(y_j^{(s)}) dy_j^{(s)}. \quad (4.13)$$

$\Psi_{Y_j^{(s)}}(\cdot)$ is the PDF of r.v. $Y_j^{(s)}$. The stationary point of \mathcal{F} obtained from its derivative w.r.t. $\zeta_j^{(s)}$ is the VaR and \mathcal{F} evaluated at the stationary point gives the CVaR. For stationary point evaluation, equation (4.13) involves differentiation of an integral with a function of $\zeta_j^{(s)}$ in the limits, so the fundamental theorem of calculus [126] must be applied to obtain the result. It could be possible that analytic expressions may not exist for (4.9) and (4.10) or may be a combination of higher order functions. So numerical methods based on Extreme Value Theory (EVT) are applied.

4.3 Extreme Value Theory

4.3.1 Generalized Extreme Value Distribution

EVT [127] can be applied to handle the mathematical formulations of VaR and CVaR. It is based on the properties of maximum order statistics of i.i.d. random variables. From the Central Limit Theorem it is known that the distribution of the sum of i.i.d. random variables tends to a Gaussian distribution as the number of r.v.s increase. This is called Sum-Stability. Similarly a Max-Stability concept exists that handles tail distributions. Consider a block of K i.i.d. random variables and B such blocks. Let ϱ_{mn} , $\{m = 1, \dots, B, n = 1, \dots, K\}$ be the n^{th} r.v. in m^{th} block. Let $\bar{\varrho}_m$ be the maximum order statistic of each block, i.e., the block-maxima given by $\bar{\varrho}_m = \max\{\varrho_{m1}, \dots, \varrho_{mK}\}$. $\max\{\dots\}$ is the maximum operator. The distribution of the set of maximum elements $\{\bar{\varrho}_m, m = 1, \dots, B\}$ converges to a non-degenerate limiting distribution $F_Y(\xi, \bar{s}, \bar{l})$ called Generalized Extreme Value Distribution (GEVD) and is given by

$$\Pr[Y \leq y] = \exp\left(-\left(1 + \xi \frac{(y - \bar{l})}{\bar{s}}\right)_+^{-\frac{1}{\xi}}\right), \quad (4.14)$$

where $(y)_+ = \max\{0, y\}$ and $\exp(\cdot)$ is the exponential function. This convergence to a limiting distribution is called Max-Stability. \bar{s} is the scale parameter, \bar{l} is the location parameter, ξ is the shape parameter. Depending on ξ values, GEVD has three types of limiting distributions. (1) The Gumbel family, when $\xi = 0$ and $-\infty < y < \infty$, (2) The Frechet family, when $\xi > 0$ and $y > -1/\xi$, (3) The Weibull family, when $\xi < 0$ and $y < -1/\xi$. To get VaR, GEVD from EVT is applied. The block-maxima approach based on GEVD solves for VaR without actually inverting the CDF. This is the primary requirement for P3_c. Optimal η -quantile [128] after parameter estimation is given by

$$\bar{\zeta}_j^{(s)} = \bar{l}_j - \bar{s}_j(1 - (-B \cdot \ln(\eta))^{-\xi_j})/\xi_j, \quad \forall j. \quad (4.15)$$

The natural logarithm is given by $\ln(\cdot)$. Algorithm 3 gives the sequence of steps to obtain η -quantile using GEVD based block-maxima approach. Step.14 estimates the

Algorithm 3 : GEVD approach for η -quantile

```

1: set  $\mathbf{p}$ ,  $\boldsymbol{\alpha}$ ,  $K$ ,  $B$ 
2: repeat
3:   for  $j = 1 : N$  do
4:      $\bar{\varrho}_m = [\cdot]$  {empty set}
5:     for  $n_1 = 1 : K$  do
6:        $\Omega = [\cdot]$  {empty set}
7:       for  $n_2 = 1 : B$  do
8:         generate  $\mathbf{h}_{ik}, \forall k \neq j, \forall i$ 
9:         for given  $\mathbf{h}_{ik}$ , evaluate loss  $y_j^{(s)}, \forall j$ 
10:         $\Omega = \Omega \cup y_j^{(s)}$ 
11:      end for
12:       $\bar{\varrho}_m = \bar{\varrho}_m \cup \max\{\Omega\}$ 
13:    end for
14:    estimate GEVD parameters  $\xi_j, \bar{l}_j, \bar{s}_j$ 
15:    evaluate optimal VaR  $\bar{\zeta}_j^{(s)}, \forall j$ , from equation (4.15)
16:  end for
17:  solve optimization problem to find  $\mathbf{p}$ ,  $\boldsymbol{\alpha}$  by Algorithm 1
18: until convergence
19: output :  $\bar{\zeta}_j^{(s)}, \forall j, \mathbf{p}, \boldsymbol{\alpha}$ 

```

GEVD parameters. A parametric approach based on the Maximum Likelihood Estimator (MLE) or a non-parametric approach based on Pickands estimator are usually used to estimate the GEVD parameters (ξ, \bar{s}, \bar{l}) . So an optimization in certain sense needs to be carried out at this step. Step.17 involves solving a deterministic optimization problem. Algorithm 1 used to simultaneously solve for $\boldsymbol{\alpha}$ and \mathbf{p} in chapter 3 is applied at this point. Under CSI_s two different cases are considered for simulation of a minimum working example. One where analytic expression for VaR exists and the other where it does not. If $\mathbf{h}_{ij} \sim \mathcal{N}(\underline{m}_{ij}, \rho_{ij}^2) + j\mathcal{N}(\underline{m}_{ij}, \rho_{ij}^2), \forall i, \forall j$, then $y_j^{(s)}$ for a given $\boldsymbol{\alpha}$ follows a non-central chi-squared distribution ($\text{nc-}\chi_2^2$) with 2 DoF. The CDF and PDF are respectively given by

$$F_{Y_j^{(s)}}(y) = 1 - Q\left(\frac{\sqrt{\bar{\lambda}}}{\bar{e}_{ij}}, \frac{\sqrt{y}}{\bar{e}_{ij}}\right), \quad (4.16)$$

$$f_{Y_j^{(s)}}(y) = \frac{1}{2 \cdot \bar{e}_{ij}^2} \cdot \exp\left(-\frac{(y + \bar{\lambda})}{2 \cdot \bar{e}_{ij}^2}\right) \cdot I_0\left(\frac{\sqrt{\bar{\lambda}y}}{\bar{e}_{ij}^2}\right). \quad (4.17)$$

The non-central parameter is given by $\bar{\lambda} = \bar{n}_{ij}^2 + \underline{n}_{ij}^2$, where $\bar{n}_{ij} = (\mathbf{1}_T^T \Re\{\mathbf{u}_{ij}\} + \mathbf{1}_T^T \Im\{\mathbf{u}_{ij}\}) \cdot \sum_{k \neq j, k=1}^N \sqrt{p_k} \underline{m}_{ij}$ and $\underline{n}_{ij} = (\mathbf{1}_T^T \Re\{\mathbf{u}_{ij}\} - \mathbf{1}_T^T \Im\{\mathbf{u}_{ij}\}) \cdot \sum_{k \neq j, k=1}^N \sqrt{p_k} \underline{m}_{ij}$, $\Re\{\cdot\}$ is the real part, $\Im\{\cdot\}$ is the imaginary part, $\bar{e}_{ij}^2 = \sum_{k \neq j} p_k \rho_{ik}^2$, $Q(\cdot, \cdot)$ is the generalized Marcum's

Q -function [129], $I_0(\cdot)$ is the 0^{th} order modified Bessel function of first kind. From equation (4.16) it can be observed that the analytic VaR expression does not exist. Also from equation (4.17) the evaluating a closed form expression for CVaR is not possible. If $\underline{m}_{ij} = 0, \forall i, \forall j$, then $y_j^{(s)}$ for a given $\boldsymbol{\alpha}$ follows a central chi-squared distribution (χ_2^2) with 2 DoF, i.e., an exponential distribution. An analytic expression for VaR exist for χ_2^2 -case and a closed form CVaR expression also exists. Equation (4.6) becomes

$$\zeta_j^{(s)} \geq -2 \cdot \ln(1 - \eta) \cdot \sum_{i=1}^M \sum_{k \neq j, k=1}^N \rho_{ik}^2 \cdot \alpha_{ij} p_k, \forall j. \quad (4.18)$$

The term to the right of inequality (4.18) is the VaR for χ_2^2 case [125].

For comparison, the deterministic equivalent of $P3_c$, i.e., $P3_s$ and frequently used methods to handle chance constraints under CSI_s such as Bernstein Approximation (BA) [121] and Gaussian Approximation (GA) [130] are considered. For GA, the CDF $F_{Y_j^{(s)}}(\cdot)$ is approximated by Gaussian CDF which is invertible. The GA to equation (4.2) is

$$\zeta_j^{(s)} \geq \bar{m}_j + \sqrt{2} \cdot \bar{\rho}_j \cdot \text{erf}^{-1}(2\eta - 1), \forall j, \quad (4.19)$$

where $\text{erf}^{-1}(\cdot)$ is the inverse error function [129]. Mean \bar{m}_j , variance $\bar{\rho}_j^2$ of the approximated Gaussian distribution must be estimated. GA may not produce tight bounds. For BA, equation (4.2) for a given $\boldsymbol{\alpha}$ is modified as

$$\bar{f}_0 + \sum_{k \neq j, k=1}^N \bar{\mu}_k \cdot \bar{f}_k + \bar{\delta}_k \cdot \sqrt{\sum_{k \neq j, k=1}^N \sigma_k^2 \cdot \bar{f}_k^2} \leq 0, \forall j, \quad (4.20)$$

where $\bar{f}_0 = \bar{\sigma}_j^2 - (p_j g_{ijj}/b) + \sum_{k \neq j} \bar{\varepsilon}_{ik} p_k$, $\bar{\varepsilon}_{ik} = (\bar{d}_{ik} + \bar{c}_{ik})/2$, $[\bar{c}_{ik}, \bar{d}_{ik}]$ is the support of g_{ijk} , $\bar{f}_k = p_k \bar{\varepsilon}_{ik}$, $\bar{\varepsilon}_{ik} = (\bar{d}_{ik} - \bar{c}_{ik})/2$, $\bar{\delta}_k = \sqrt{2 \cdot \ln(1 - \eta)^{-1}}$. Parameters $\bar{\mu}_k$ and σ_k are chosen as described in [121]. BA requires \bar{g}_{ijk} to be normalized such that $(g_{ijk} - \bar{\varepsilon}_{ik})/\bar{\varepsilon}_{ik}$ is supported on $[-1, 1]$, it also requires \bar{f}_k to be affine in the decision vector. But $y_j^{(s)}$ is bilinear in $\boldsymbol{\alpha}$ and \mathbf{p} , so for BA, two stage formulation for PA and BSA is used. Requirements such as choosing a support for unknown r.v. and formulating affine functions in decision variables make BA stringent. A support $[\bar{c}_{ik}, \bar{d}_{ik}]$ [122] for the exponential distribution must be set in BA. Since the χ_2^2 distribution is defined for the non-negative values of the r.v., $\bar{c}_{ik} = 0$ can be chosen. Finding the support may be difficult for most of the distributions and the obtained conservative approximation could be loose if it is not properly chosen.

$P3_c$ is assumed feasible. The considered beamformers are $\mathbf{u}_{ij} = \arg\text{-max}_{\mathbf{u}_{ij}} \|\mathbf{u}_{ij} \mathbf{h}_{ij}\|_2^2, \forall i, \forall j$ and $\|\mathbf{u}_{ij}\|_2^2 = 1, \forall i, \forall j$, i.e., the BF sub-problem is not included in PA. $\arg\text{-max}(\cdot)$ operator chooses the decision variable that maximizes the expression it is maximizing. For BA, an exhaustive search w.r.t. $\boldsymbol{\alpha}$ is performed. The chosen support for g_{ijk} is crucial for BA. No explicit estimation of EVT parameters is performed in Algorithm 3, MLE

based function `gevfit()` in Matlab is used. For simulation, $M = 2$, $N = 4$, $T = 2$, $\bar{\sigma}_j^2 = 1, \forall j$, $P_{\max} = 23$ [dBm], $\forall j$, $B = 1e3$, $K = 30$, $\underline{m}_{ij} = 0.6$, $\rho_{ij} = 0.5$ is set. These random parameter values are set as a reference to test the performance of Algorithm 3. They can be scaled according to the practical values. Both χ_2^2 and $\text{nc-}\chi_2^2$ are in the Domain of Attraction (DoA) of the Gumbel family where $\xi \rightarrow 0$.

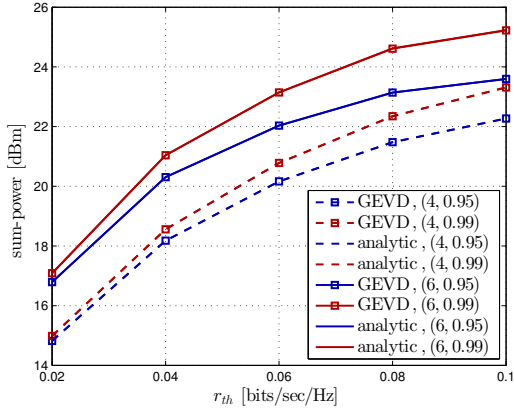


Fig. 4.1: VaR comparison with analytic values for χ_2^2 distribution of $y_j^{(s)}$.

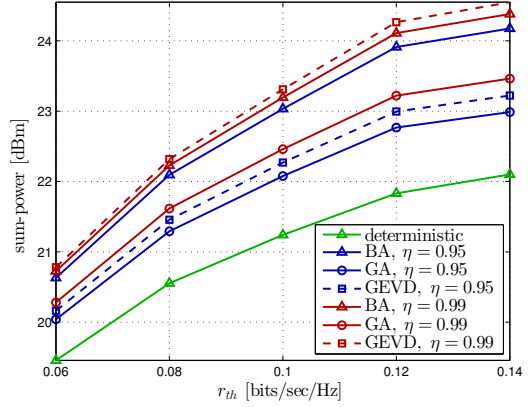


Fig. 4.2: VaR comparison with other methods for χ_2^2 distribution of $y_j^{(s)}$.

Figure 4.1 compares the numerical bounds obtained by solving $P3_c$ with GEVD approach and its analytic equivalent for the χ_2^2 distribution of $y_j^{(s)}$. Two confidence levels $\eta = 0.99$ and $\eta = 0.95$ are considered. The plots perfectly overlap showing that the GEVD approach indeed generates the optimal values. From $\eta = 0.95$ to $\eta = 0.99$ there is an outage requirement increase which results in higher objective value of $P3_c$. With GEVD, generating optimal values is possible without inverting the CDF. Figure 4.2 compares GEVD performance with other existing methods. For $\eta = 0.95$ the GA is underestimating the objective value while BA overestimates it. Neither plot matches the exact value. For $\eta = 0.99$ both GA and BA underestimate the objective, however the BA plot is very close to the optimal GEVD plot. GA produced weak bounds while bounds from BA were not tight for different η . An improvement in the bound with BA could be expected if the support for the unknown r.v. is chosen properly. Under CSI_p the objective value is the least as expected due to more available CSI at the BS. With the increased QoS, the MUI at each UE increases leading to an increased power level for each UE to achieve the QoS, hence the objective value increases. Figures 4.1 and 4.2 consider only instances which are feasible for all the methods simultaneously. Figure 4.3 compares the VaR for the non-invertible $\text{nc-}\chi_2^2$ CDF of $y_j^{(s)}$. No analytical values exist to verify the tightness of the generated numerical bound. The observed trend in the plots is very similar to the trend observed in Figure 4.2. GA and BA underestimate the GEVD bound for $\eta = 0.95$. For $\eta = 0.99$, BA is close to the GEVD bound while still underestimating it, GA also generates a weak underestimated bound. So the BA bounds are not consistent with the change in η while the GA bounds are loose.

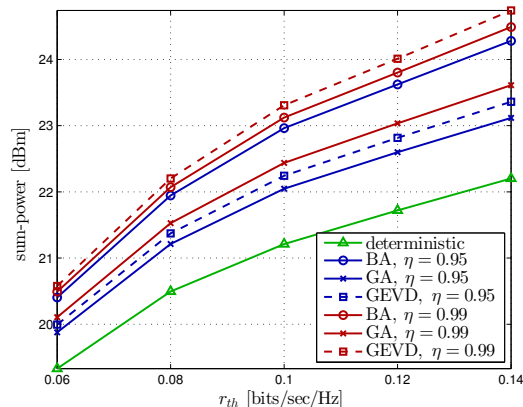


Fig. 4.3: VaR comparison for $nc-\chi_2^2$ distribution of $y_j^{(s)}$.

4.3.2 Generalized Pareto Distribution for CVaR

With $B = 1$, and by setting a threshold \bar{t} called return value, the exceedance for a block of i.i.d. random variables over \bar{t} is evaluated. This is called Peaks-over-Threshold (PoT). These excess values, i.e., the values exceeding \bar{t} , also converge to a limiting distribution $F_X(\xi, \underline{s}, \underline{l})$ called Generalized Pareto Distribution (GPD). It is given as

$$\Pr[X = Y - t | Y > \bar{t}] = 1 - \left(1 + \frac{\xi x}{\underline{s} + \xi(\bar{t} - \underline{l})}\right)_+^{-\frac{1}{\xi}}, \quad (4.21)$$

where \underline{s} is the scale parameter, \underline{l} is the location parameter, shape parameter ξ is analogous to ξ in GEVD. GPD is analogous to the Max-Stability. To get CVaR we apply GPD from EVT. CVaR can be evaluated if the threshold is $\bar{\zeta}_j^{(s)}$ which can be obtained from $P3_c$. CVaR after parameter estimation is given by

$$\beta_j^{(s)} = \bar{\zeta}_j^{(s)} + (\underline{s}_j + \xi(\bar{t} - \underline{l}_j))/(1 - \xi), \quad \forall j. \quad (4.22)$$

The second term on the right side is the mean excess over the chosen threshold $\bar{\zeta}_j^{(s)}$. Algorithm 4 gives the sequence of steps to obtain CVaR using GPD based PoT approach. Step.15 involves solving a deterministic optimization problem. Algorithm 1 is again used to find the optimal BSA and PA. The analytic expression for CVaR [128] for χ_2^2 distribution of the unknown r.v. under the worst case performance is

$$\zeta_j^{(s)} \geq 2 \cdot (1 - \ln(1 - \eta)) \cdot \sum_{i=1}^M \sum_{k \neq j, k=1}^N \rho_{ik}^2 \alpha_{ij} p_k, \quad \forall j. \quad (4.23)$$

The terms to the right of inequality (4.23) is the CVaR for χ_2^2 case. For the worst case performance using GA, equation (4.10) becomes

$$\zeta_j^{(s)} \geq \bar{m}_j + \bar{p}_j \cdot \exp(-(\text{erf}^{-1}(2\eta - 1))^2)/(\sqrt{2\pi}(1 - \eta)), \quad \forall j. \quad (4.24)$$

Algorithm 4 : GPD approach for CVaR

- 1: set \mathbf{p} , $\boldsymbol{\alpha}$, K .
 - 2: solve P3_c to get $\bar{\zeta}_j^{(s)}$, $\forall j$, from Algorithm 3.
 - 3: **for** $j = 1 : N$ **do**
 - 4: $\Omega = [\cdot]$ {empty set}
 - 5: **for** $n_1 = 1 : K$ **do**
 - 6: generate \mathbf{h}_{ik} , $\forall k \neq j, \forall i$.
 - 7: for given \mathbf{h}_{ik} , evaluate MUI $y_j^{(s)}$, $\forall j$.
 - 8: $\Omega = \Omega \cup y_j^{(s)}$.
 - 9: **end for**
 - 10: with threshold $\bar{t}_j = \bar{\zeta}_j^{(s)}$, $\forall j$.
 - 11: find PoT, entries greater than $\bar{\zeta}_j^{(s)}$ in Ω .
 - 12: estimate GPD parameters $\xi_j, \underline{s}_j, \underline{l}_j$.
 - 13: evaluate CVaR $\beta_j^{(s)}$, $\forall j$, from equation (4.22)
 - 14: **end for**
 - 15: solve optimization problem to find \mathbf{p} , $\boldsymbol{\alpha}$ by Algorithm 1.
 - 16: **output** : $\beta_j^{(s)}$, $\forall j$, \mathbf{p} , $\boldsymbol{\alpha}$
-

The terms to the right of inequality (4.24) is the CVaR [125] for GA. For simulation, same GEVD parameters values are assumed. The same feasible channel realizations are retained to further evaluate CVaR from the corresponding VaR values. GPD parameters $(\xi, \underline{s}, \underline{l})$ are estimated in Step.12 based on MLE based function `gpfitt()` in Matlab. Figure 4.4 compares the CVaR bounds $\bar{\zeta}_j$ with the analytic values and bounds from the

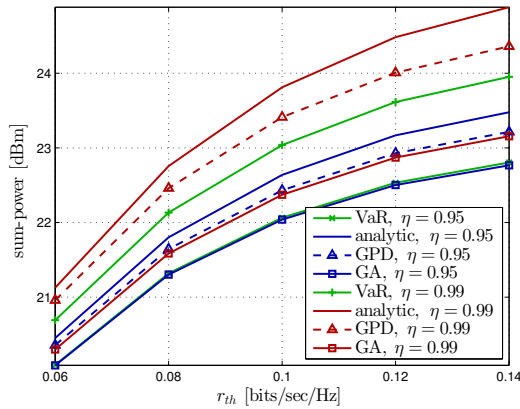


Fig. 4.4: CVaR comparison for χ_2^2 distribution of $y_j^{(s)}$.

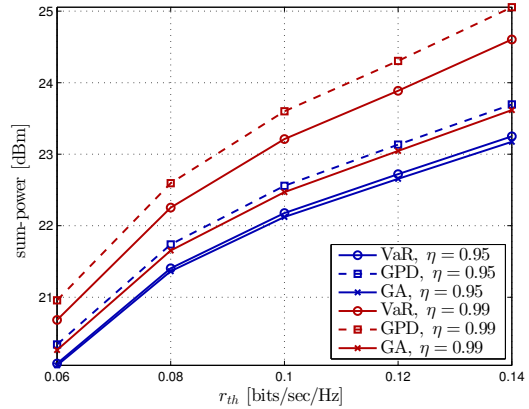


Fig. 4.5: CVaR comparison for $nc-\chi_2^2$ distribution of $y_j^{(s)}$.

GA method. Corresponding VaR bounds generated from GEVD are also plotted. As expected, $\beta_j^{(s)} \geq \bar{\zeta}_j^{(s)}$, $\forall j$. The GPD bounds are the closest to the analytic values. The gap between the plots is due to the GPD parameter estimation error. Increasing N may close this gap thus generating tighter bounds. This gap increases with η , i.e., when

compared to analytic values, $\eta = 0.95$ produced tighter bounds than the corresponding $\eta = 0.99$ curves. GA produced bounds that are very weak in either case. The values for $\eta = 0.95$ overlap with the deterministic curve, this is not a performance improvement. GA bounds are underestimating the analytic values. Figure 4.5 compares the CVaR curves for the $\text{nc}-\chi_2^2$ case. A trend similar to Figure 4.4 is observed, with the GPD bounds greater than the corresponding GEVD values. Clearly EVT outperforms other methods. The only requirement for EVT approach is that the underlying distribution is known.

4.4 Power Control under imperfect-CSI

Under CSI_i , the interfering channels are modelled by the CSI-error model. This uncertainty model is given as

$$\mathbf{h}_{ik} = \bar{\mathbf{h}}_{ik} + \mathbf{e}_{ik}, \forall i, \forall k \neq j, \quad (4.25)$$

where $\mathbf{e}_{ik} = \mathbf{E}_{ik}^{1/2} \bar{\mathbf{v}}_{ik}$, \mathbf{E}_{ik} is a known covariance matrix, $\bar{\mathbf{h}}_{ik}$ is also a known vector and $\bar{\mathbf{v}}_{ik} \sim \mathcal{CN}(0, \mathbf{I}_{N_r})$, i.e., unit-variance complex Gaussian vector. Hence $\mathbf{h}_{ik} \sim \mathcal{CN}(\bar{\mathbf{h}}_{ik}, \mathbf{E}_{ik}), \forall i, \forall k \neq j$. Under CSI_i the imperfect-SINR $\bar{\gamma}_j^{(i)}$ for the deterministic parameter in equation (2.5) is given as

$$\bar{\gamma}_j^{(i)} = \frac{p_j \sum_{i=1}^M \alpha_{ij} \|\mathbf{u}_{ij}^H \mathbf{h}_{ij}\|_2^2}{\sum_{i=1}^M \sum_{k \neq j, k=1}^N p_k \alpha_{ij} \|\mathbf{u}_{ij}^H (\bar{\mathbf{h}}_{ik} + \mathbf{e}_{ik})\|_2^2 + \bar{\sigma}_j^2}, \forall j. \quad (4.26)$$

For the deterministic QoS constraints (2.10), the PrCons under CSI_i are given by

$$\Pr \left[\bar{r}_j^{(i)} \geq r_{th} \right], \forall j, \quad (4.27)$$

where $\bar{r}_j^{(i)} = \log_2(1 + \bar{\gamma}_j^{(i)})$ is the imperfect-rate for UE_j . With the substitution

$$y_j^{(i)} = \sum_{i=1}^M \sum_{k \neq j, k=1}^N p_k \alpha_{ij} \mathbf{u}_{ij}^H (\bar{\mathbf{h}}_{ik} \mathbf{e}_{ik}^H + \mathbf{e}_{ik} \bar{\mathbf{h}}_{ik}^H + \mathbf{e}_{ik} \mathbf{e}_{ik}^H) \mathbf{u}_{ij}, \quad \forall j, \quad (4.28)$$

$$\zeta_j^{(i)} = \sum_{i=1}^M \frac{p_j \alpha_{ij}}{b} \|\mathbf{u}_{ij}^H \mathbf{h}_{ij}\|_2^2 - \sum_{i=1}^M \sum_{k \neq j, k=1}^N p_k \alpha_{ij} \|\mathbf{u}_{ij}^H \bar{\mathbf{h}}_{ik}\|_2^2 - \bar{\sigma}_j^2, \quad \forall j, \quad (4.29)$$

it is rearranged as equation (4.30) or (4.31)

$$F_{Y_j^{(i)}}(\zeta_j^{(i)}) \geq \eta, \forall j. \quad (4.30) \quad \zeta_j^{(i)} \geq F_{Y_j^{(i)}}^{-1}(\eta), \forall j. \quad (4.31)$$

$y_j^{(i)}$ is precisely not the loss for UE_j , since the MUI term is present in both $y_j^{(i)}$ and $\zeta_j^{(i)}$. $y_j^{(i)}$ is quadratic in \mathbf{e}_{ik} , hence obtaining an invertible expression is almost impossible.

Similar to P3_c, the chance constrained MINLP under CSI_i is given as

$$\text{P3}_i: \underset{\boldsymbol{\alpha}, \mathbf{p}, \mathbf{U}}{\text{minimize}} \quad \sum_{j=1}^N p_j \quad (4.32)$$

$$\text{s.t:} \quad (2.1), (2.2), (2.4), (4.31). \quad (4.33)$$

Under CSI_i, a Bernstein-type Inequality (BI) [131] is frequently used to obtain a conservative formulation based on SDP. It is based on identifying a quadratic form in zero-mean unit-variance Gaussian random variables of the PrCons. BI was also implemented for transmit BF [124] under CSI_i, secrecy-rate maximization in the presence of eavesdroppers [132]. For the BI formulation of PrCons, equation (4.27) must first be rearranged for a given $\boldsymbol{\alpha}$.

$$\begin{aligned} \mathbf{W}_{ijk} &= \mathbf{u}_{ij} \mathbf{u}_{ij}^H p_k, \\ \overline{\mathbf{W}}_{ij} &= -\text{blkdiag}(\mathbf{W}_{ij1}, \dots, \mathbf{W}_{ij(j-1)}, \mathbf{W}_{ij(j+1)}, \dots, \mathbf{W}_{ijN}), \\ \overline{\mathbf{E}}_{ij} &= \text{blkdiag}(\mathbf{E}_{i1}^{1/2}, \dots, \mathbf{E}_{i(j-1)}^{1/2}, \mathbf{E}_{i(j+1)}^{1/2}, \dots, \mathbf{E}_{iN}^{1/2}), \\ \overline{\mathbf{A}}_j &= \overline{\mathbf{E}}_{ij}^H \overline{\mathbf{W}}_{ij} \overline{\mathbf{E}}_{ij}, \\ \overline{\mathbf{g}}_{ij} &= [\overline{\mathbf{h}}_{i1}^H, \dots, \overline{\mathbf{h}}_{i(j-1)}^H, \overline{\mathbf{h}}_{i(j+1)}^H, \dots, \overline{\mathbf{h}}_{iN}^H]^H, \\ \underline{c}_j &= \overline{\sigma}_j^2 - \frac{\|\mathbf{h}_{ij}^H \mathbf{W}_{ijj}\|_2^2}{b} - \overline{\mathbf{g}}_{ij}^H \overline{\mathbf{W}}_{ij} \overline{\mathbf{g}}_{ij}, \\ \overline{\mathbf{b}}_j &= \overline{\mathbf{E}}_{ij}^H \overline{\mathbf{W}}_{ij} \overline{\mathbf{g}}_{ij}, \\ \overline{\mathbf{v}}_j &= [\overline{\mathbf{v}}_{i1}^H, \dots, \overline{\mathbf{v}}_{i(j-1)}^H, \overline{\mathbf{v}}_{i(j+1)}^H, \dots, \overline{\mathbf{v}}_{iN}^H]^H. \end{aligned} \quad (4.34)$$

is substituted for all UEs and equation (4.27) is rearranged into a known BI form as equation (4.35). The block diagonal operator that forms a new matrix by placing the given matrices on the main diagonal of the new matrix is given by $\text{blkdiag}(\cdot)$. From [131], for a zero-mean unit-variance Gaussian vector $\overline{\mathbf{v}}_j$ and a positive semidefinite matrix $\overline{\mathbf{A}}_j$, the constraint of the form

$$\Pr[\overline{\mathbf{v}}_j^H \overline{\mathbf{A}}_j \overline{\mathbf{v}}_j + 2 \cdot \Re\{\overline{\mathbf{v}}_j^H \overline{\mathbf{b}}_j\} \geq \underline{c}_j] \geq \eta, \forall j, \quad (4.35)$$

is approximated by a conservative deterministic form

$$\text{tr}(\overline{\mathbf{A}}_j) - x_j \cdot (-2 \cdot \ln(1 - \eta))^{1/2} + d_j \cdot \ln(1 - \eta) \geq \overline{c}_j, \forall j, \quad (4.36)$$

$$\overline{\mathbf{A}}_j + d_j \mathbf{I} \geq 0, \forall j, \quad (4.37)$$

$$\begin{bmatrix} x_j & \left[\text{vec}(\overline{\mathbf{A}}_j)^H, \overline{\mathbf{b}}_j^H \sqrt{2} \right] \\ \left[\text{vec}(\overline{\mathbf{A}}_j)^H, \overline{\mathbf{b}}_j^H \sqrt{2} \right]^H & x_j \cdot \mathbf{I} \end{bmatrix} \geq 0, \forall j, \quad (4.38)$$

$$d_j \geq 0, \forall j, \quad (4.39)$$

where $\text{tr}(\cdot)$ is the trace operator, the $\text{vec}(\cdot)$ operator stacks the columns of a matrix into a vector. The expression inside the $\Pr[\cdot]$ operator of equation (4.35) is quadratic in $\overline{\mathbf{v}}_j$.

Equations (4.37) and (4.38) are SDP constraints. Implicit constraints include $\overline{\mathbf{W}}_{ij} \geq 0, \forall j$. It can be observed that the problem of transmitter optimization is obtained by this reformulation, since for each UE, the power variable p_j and the beamforming vector \mathbf{u}_{ij} are considered as one matrix variable. The approximation of equation (4.27) by equations (4.36)-(4.39) makes $P3_i$ tractable, however it must be evaluated for all the BSA combinations. Also it might be difficult to expect CSI_i to always satisfy the quadratic form in Gaussian variables. Nevertheless, BI is one of the most frequently used methods in the presence of CSI_i and PrCons. BA, which is applicable to a family of distributions, is also often used.

For the link-level simulations of a minimum working example, $P3_i$ is assumed feasible. The same set of parameters as in $P3_c$ are set, i.e., $M = 2, \overline{\sigma}_j^2 = 1, \forall j, T = 2, P_{\max} = 23$ dBm. The procedure to obtain the η -quantile by Algorithm 3 is applied without any change. Step.6 generates the CSI-error model channel as in equation (4.25). For the MUI channels, $\overline{\mathbf{h}}_{ik} \sim \mathcal{CN}(0, \mathbf{I}_T), \forall i, \forall k \neq j, \mathbf{E}_{ik} = 0.2 \cdot \mathbf{I}_T, \forall i, \forall k \neq j$ is set, and for the deterministic channels, $\mathbf{h}_{ij} \sim \mathcal{CN}(0, \mathbf{I}_T), \forall i, \forall j$ is assumed. For a fair comparison of $P3_i$ with $P3_s$, $\mathbf{u}_{ij} = \arg\text{-max}_{\mathbf{u}_{ij}} \|\mathbf{u}_{ij} \mathbf{h}_{ij}\|_2^2, \forall i, \forall j$, is considered. It can be observed that the assumed CSI-error model in equation (4.25) gives a complex Gaussian distribution for \mathbf{h}_{ik} if it is treated as the sum of complex Gaussian r.v.s. The resulting complex normal distribution is under the DoA of the Gumbell limiting distribution, i.e., $\xi \rightarrow 0$. But for comparison, a non-zero mean Gaussian distribution is assumed and $\overline{\mathbf{h}}_{ik}$ represents the mean value. A given $\overline{\mathbf{h}}_{ik}$ and normally distributed \mathbf{v}_{ik} , gives a nc- χ_2^2 -distribution for $y_j^{(i)}$. For the BI formulation of $P3_i$, an exhaustive search w.r.t. $\boldsymbol{\alpha}$ is carried out, where the PA sub-problem is solved at all the M^N possible BSA combinations. The BSA combination which gives the least value of sum-power objective is taken as the optimal. EVT approach of $P3_i$ has no approximations. The CDF of the unknown r.v. is estimated using VaR and GEVD. Exact quantile function of the non-invertible CDF is evaluated and the PrCons are now deterministic.

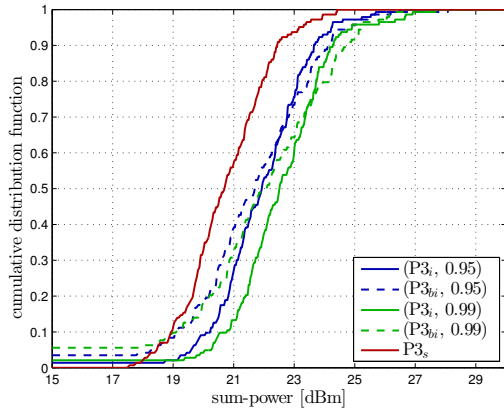


Fig. 4.6: Comparison of $P3_i, P3_{bi}, P3_s$ with $r_{th} = 0.08$.

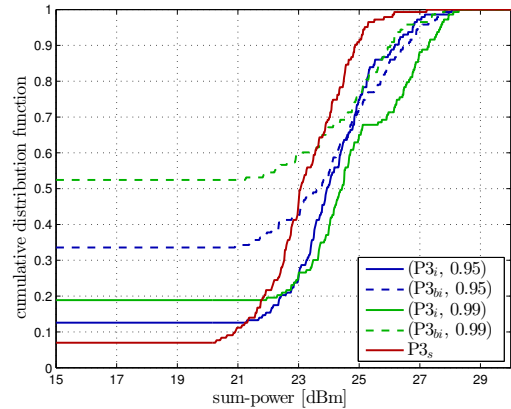


Fig. 4.7: Comparison of $P3_i, P3_{bi}, P3_s$ with $r_{th} = 0.14$.

Figure 4.6 plots the CDF curves of the sum-power objective obtained by $P3_s$, $P3_i$ and BI-formulation ($P3_{bi}$) with $\eta = 0.95$ and $\eta = 0.99$ at a low value of $r_{th} = 0.08$. As expected, the CDF plot for $P3_s$ is to the left most, giving the least objective value. The obtained objective at $\eta = 0.99$ is higher than the objective of the corresponding $\eta = 0.95$ for both $P3_i$ and $P3_{bi}$. This is due to the increased outage performance requirement. Though the performance of both $P3_i$ and $P3_{bi}$ is similar, $P3_{bi}$ performs slightly better than $P3_i$ at lower objective values, i.e., a left shift in the CDF curve. Figure 4.7 plots similar CDF curves when r_{th} is increased to 0.14. Though the CDF of $P3_i$ has a right shift to the corresponding $P3_{bi}$ for $\eta = 0.99$, it can be seen that the outage performance of $P3_i$ is better than $P3_{bi}$, i.e., an increase in the number of feasible instances for $P3_i$. For $\eta = 0.95$, the outage performance is better for $P3_i$ and the objective value is also lower at higher sum-power values. Figure 4.8 plots the average sum-power objective

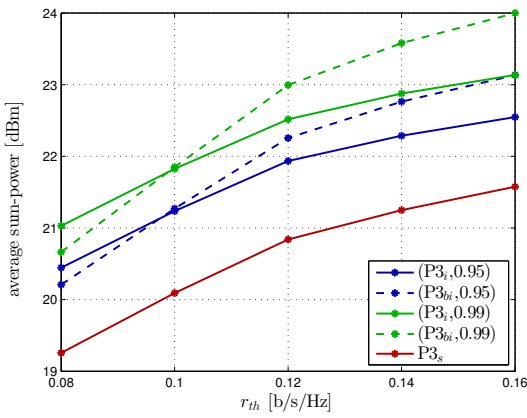


Fig. 4.8: average sum-power for $P3_i$ for $N = 3$.

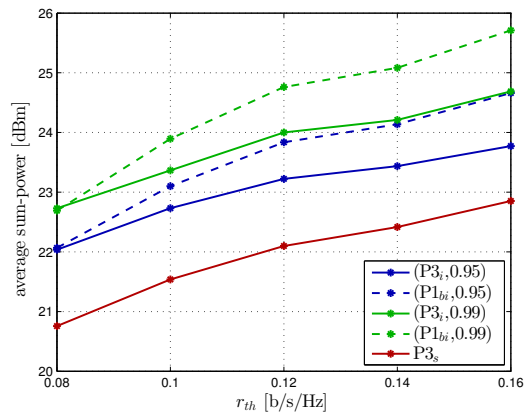


Fig. 4.9: average sum-power for $P3_i$ for $N = 4$.

vs the QoS for $N = 3$. As expected, $P3_s$ has the least objective value. The plot shows that for $r_{th} \leq 0.1$, $P3_{bi}$ performs slightly better than the $P3_i$. But beyond $r_{th} = 0.1$, $P3_i$ outperforms $P3_{bi}$ for both $\eta = 0.95$ and $\eta = 0.99$. Figure 4.9 plots the same set of curves for $N = 4$. Values of $P3_i$, $P3_{bi}$ and $P3_s$ for $N = 4$ are higher than the corresponding values in $N = 3$ i.e., the sum-power increases with increase in N . The plot shows that the average sum-power for $P3_i$ is lower than the value from $P3_{bi}$ for the considered range of QoS for both $\eta = 0.95$ and $\eta = 0.99$.

Chapter 5

Downlink Performance Analysis under perfect-CSI

From the Downlink (DL) QoS constraint (2.11), it can be observed that the expression is coupled in the beamformers and power variables. However the Uplink (UL) QoS constraint (2.10) exhibits a decoupled nature. This forms the basis to solve the DL problem via the UL. However this UL called a Virtual Uplink (VUL) exists only for mathematically solving the DL problem via a duality concept, i.e., there is no physical existence. It is very much the same as the actual UL in terms of the expressions and parameters. This duality is not the conventional mathematical optimization duality. It is a duality based on the DL-VUL reciprocity. Such a reciprocity technique is easy to understand and can be applied to a single-cell case where only a single total sum-power constraint exists. However it cannot be extended to the multi-cell case where a per-BS power constraint exists [133]. Even in a single-cell with per-antenna power constraint the simple DL-VUL duality fails since there is more than one sum-power constraint. In this case a duality different from the conventional single-cell duality with additional optimization variables must be formulated. In this chapter, problems P2 and P4 are solved via the implicit Lagrange-duality in Algorithm 1. This eliminates additional optimization sub-problems. Also the single-cell duality requirements are included into the Lagrange-duality.

5.1 Uplink-Downlink Duality

The concept of duality in single-cell is well understood [134]. For duality in a single-cell, under the same set of beamformers, transposed channels and equal total sum-power, the same SINR is achieved on both the DL and VUL. Also the noise variance remains the same in DL and VUL. The parameters and expressions for the actual-UL and VUL are the same. Understanding the sum-power constraint is important when dealing with the VUL, since on the DL it indicates the total available power at the BS and on the

VUL it is the sum of individual UE powers. So the same UL variables p_j , $\bar{\gamma}_j$, \bar{r}_j , $\bar{\sigma}_j^2$ are used as VUL transmit power, VUL-SINR, VUL-rate and VUL-noise variance respectively with no change. To apply the duality for a single-cell DL based on Perron Frobenius theory, a non-negative matrix

$$\mathbf{\Gamma} = \begin{bmatrix} \underline{\mathbf{D}} \cdot \underline{\mathbf{C}} & \underline{\mathbf{D}} \cdot \mathbf{1}_N \cdot \underline{\sigma}_j^2 \\ \mathbf{1}_N^T \cdot \underline{\mathbf{D}} \cdot \underline{\mathbf{C}} / \bar{P}_i & \mathbf{1}_N^T \cdot \underline{\mathbf{D}} \cdot \mathbf{1}_N \cdot \underline{\sigma}_j^2 / \bar{P}_i \end{bmatrix}, \quad (5.1)$$

is constructed. $\underline{\mathbf{D}}$ is an $(N \times N)$ diagonal matrix with $\underline{D}_{jj} = \underline{\gamma}_j / \|\mathbf{h}_{ij}^H \mathbf{u}_{ij}\|_2^2$, $\underline{\mathbf{C}}$ is an $(N \times N)$ matrix with $\underline{C}_{jj} = 0$ and $\underline{C}_{jk} = \|\mathbf{h}_{ij}^H \mathbf{u}_{ik}\|_2^2$. Even though $M = 1$, subscript i is retained to maintain consistency with the previous notations. For a feasible DL power vector to exist that attains the DL target SINR $\underline{\gamma}_j$ equal to the VUL SINR $\bar{\gamma}_j$, the largest eigenvalue of $\mathbf{\Gamma}$ must be positive, i.e., $\tilde{\lambda}_{\max}(\mathbf{\Gamma}) > 0$, where the maximum eigenvalue of a given matrix is given by $\tilde{\lambda}_{\max}(\cdot)$. Also, $\tilde{\lambda}_{\max}(\mathbf{\Gamma}) < 1$ [135] is needed. The achievable DL-SINR $\underline{\gamma}_j$ is $1/\tilde{\lambda}_{\max}(\mathbf{\Gamma})$. The eigenvector corresponding to this $\tilde{\lambda}_{\max}(\mathbf{\Gamma})$ is a non-negative $((N+1) \times 1)$ extended DL power vector $\tilde{\mathbf{q}} = [\mathbf{q}^T, 1]^T$. Such formulation encompasses the eigenvalue minimization problem [135], where the PA and BF are alternately optimized to reduce the $\tilde{\lambda}_{\max}(\mathbf{\Gamma})$ to less than 1. Once these conditions are satisfied, the power vector on the DL for a fixed beamformer set is given as

$$\mathbf{q} = (\mathbf{I}_N - \underline{\mathbf{D}} \cdot \underline{\mathbf{C}})^{-1} \cdot \underline{\mathbf{D}} \cdot (\underline{\sigma}^2 \mathbf{1}_N). \quad (5.2)$$

The DL noise variance for all the UEs is the same, i.e., $\underline{\sigma}_j^2 = \underline{\sigma}^2, \forall j$. The target SINR metric can also be replaced by the MSE metric [41], i.e., a duality based on achieving same MSE on the DL and the VUL can also be used. This can be extended to a joint transmitter-receiver optimization problem [136] with multiple spatial layers for each UE. In this case, per-layer UE metric must be considered, where intra-layer interference is also included into $\mathbf{\Gamma}$. With such a duality, a simple DL sum-rate maximization problem can be solved by equivalently solving the problem of minimizing the product of MSE matrices [49]. In the problem of weighted sum-power minimization with QoS constraints in SINR, the DL-VUL duality is equivalent to Lagrange-duality when the considered DL weight vector is the VUL-noise vector [137]. This Lagrange-dual includes the case where the DL-noise and VUL-noise have the same value and set to unity value. The coupled max-min problem [138] for transmit BF can also be solved via duality. The aim of MSE-duality, SINR-duality, max-min duality is to decouple the DL problem. Since $M = 1$, this single-cell DL-VUL duality has only one simple total-power constraint bounded by the available power at the BS.

This single-cell duality from equations (5.1) and (5.2) can be extended to the multi-cell case only if there is a single total power constraint, i.e., with $\sum_{i=1}^M \bar{P}_i$ as the only requirement. This may not provide a feasible solution since a possibility of $q_{ij} > \bar{P}_i$ can be mathematically obtained for some random UE $_j$ at its assigned BS, i.e., the DL transmit power is greater than the available BS power. A multi-cell BF problem by

max-min duality with per-BS power constraints [133] includes VUL-noise as an optimization variable to satisfy the per-BS constraints. Also instead of a single sum-power constraint, a sum-sum power constraint and a weighted sum of VUL-noise constraint are implemented to map the DL power constraints. This formulation for the multi-cell case uses a Lagrange-duality to satisfy the power constraints. In addition, the VUL-noise is also optimization variable and the value of DL-noise may be different from its corresponding VUL value. A multi-cell DL throughput maximization problem with per-BS power constraints by Lagrange-duality [139] also maps the weighted sum of the thermal noise to the per-BS power constraints. The optimization sub-problem in the VUL-noise variable is solved via the GP formulation. In most of the cases, either the BSA is known or only one UE per-BS is considered. The per-antenna power constraints in a single-cell [140] are similar to the per-BS power constraints. Lagrange-duality and uncertain VUL-noise is once again the solution approach for transmitter optimization. The DL-VUL duality in per-BS and per-antenna constrained problems is obtained by the primal DL problem's mathematical dual with uncertain VUL-noise. For single-cell duality the DL and VUL noise variance remains the same, also the total sum-power on the DL and VUL are equal. In the multi-cell case with uncertain noise and Lagrange-duality formulation, the noise variance on DL and VUL may not be the same, i.e., $\bar{\sigma}_j^2 \neq \underline{\sigma}_j^2$ and the total sum-power on the DL and VUL may not be satisfied with equality, though there is no mathematical violation. The implicit Lagrange-duality in Algorithm 1 satisfies both these conditions while considering the per-BS power constraints, i.e.,

$$\sum_{i=1}^M \sum_{j=1}^N \alpha_{ij} q_{ij} = p_j, \quad (5.3) \quad \bar{\sigma}_j^2 = \underline{\sigma}_j^2, \forall j. \quad (5.4)$$

and eliminates the need to treat VUL-noise as an optimization variable.

5.2 Problem Reformulation

P2_p is reformulated such that it includes the Lagrange-duality and single-cell duality. The reformulation begins with the single-cell duality requirements from equations (5.3) and (5.4). For a given α , the VUL sum-power is bounded by the total available power at the associated BS, i.e.,

$$\sum_{j=1}^N p_j \leq \sum_{i=1}^M \bar{P}_i. \quad (5.5)$$

Equations (5.3) and (5.5) imply that the total available power in the system is constant. Implicit constraints include

$$\sum_{i=1}^M \sum_{j=1}^N \alpha_{ij} q_{ij} \leq \sum_{i=1}^M \bar{P}_i, \quad (5.6) \quad p_j \in [0, \sum_{i=1}^M \bar{P}_i], \forall j. \quad (5.7)$$

It can be observed from equations (2.7) and (5.7) that each DL power variable q_{ij} is bounded by the per-BS power constraint while each VUL power variable p_j is bounded by the total sum-power across all the BSs. Hence p_j in general is not equal to its corresponding q_{ij} and can be greater than the maximum available per-BS power \bar{P}_i . In addition to the equality (5.4), the VUL-noise also remains constant, which is one of the objectives of the reformulation. For a given $\boldsymbol{\alpha}$, \mathbf{p} , \mathbf{U} and achievable $\bar{\gamma}_j, \forall j$, the required DL power to satisfy $\underline{\gamma}_j = \bar{\gamma}_j, \forall j$, is given by equation (5.2). To include the BSA variable $\boldsymbol{\alpha}$ into the reformulation, the elements of the diagonal desired signal matrix $\underline{\mathbf{D}}$ are modified as $\tilde{\mathbf{D}}_{jj} = \bar{\gamma}_j / \sum_{i=1}^N \|\alpha_{ij} \mathbf{u}_{ij}^H \mathbf{h}_{ij}\|_2^2$ and elements of crosstalk matrix $\underline{\mathbf{C}}$ are modified as $\tilde{\mathbf{C}}_{jk} = \sum_{m=1}^M \|\alpha_{mk} \mathbf{u}_{mk}^H \mathbf{h}_{mj}\|_2^2$ and $\tilde{\mathbf{C}}_{jj} = 0$. Under the feasibility conditions of Perron Frobenius theory equation (5.2) changes to

$$\mathbf{q} = (\mathbf{I} - \tilde{\mathbf{D}} \cdot \tilde{\mathbf{C}})^{-1} \cdot \tilde{\mathbf{D}} \cdot (\sigma^2 \mathbf{1}_N). \quad (5.8)$$

The concept of single-cell duality is included into the constraints of P2_p by equation (5.8). P2_b is also solved via VUL, i.e., an approach similar to solving P3_b .

5.3 Downlink Rate and Power Allocation

With these additional constraints, P2 is reformulated and the same without any change applies to P4 . To solve P2_b , the objective considers VUL-rate \bar{r}_j in equation (2.10) instead of DL-rate r_j in equation (2.11), i.e., the objective has a VUL-rate instead of DL-rate. It is given as

$$\text{P2}_v: \quad \underset{\boldsymbol{\alpha}, \mathbf{q}, \mathbf{P}, \mathbf{U}}{\text{maximize}} \quad \sum_{j=1}^N \bar{r}_j \quad (5.9)$$

$$\text{subject to :} \quad \begin{aligned} & (2.1), (2.2), (2.7), (2.8), (2.10), \\ & (5.3), (5.5), (5.6), (5.7), (5.8). \end{aligned} \quad (5.10)$$

It is not required to include all the constraints in equation (5.10), e.g., equation (5.6) can be dropped since satisfying equation (2.8) will satisfy equation (5.6). Similarly, equation (5.5) and the upper bound on p_j in equation (5.7) are dropped due to equation (5.3). For P2_v , two sub-problems P2_{vb} and P2_{vp} are formulated to solve for BSA and PA respectively. The VUL BSA sub-problem is

$$\text{P2}_{vb}: \quad \underset{\boldsymbol{\alpha}}{\text{maximize}} \quad \sum_{j=1}^N \bar{r}_j \quad (5.11)$$

$$\text{subject to:} \quad (2.1), (2.2), (2.8), (2.10). \quad (5.12)$$

The reformulated VUL PA sub-problem to solve P2_{vp} is

$$\text{P2}_{vp}: \underset{\mathbf{Q}, \mathbf{p}, \mathbf{U}}{\text{maximize}} \quad \sum_{j=1}^N \bar{r}_j \quad (5.13)$$

$$\text{subject to :} \quad \begin{aligned} & (2.7), (2.8), (2.10), (5.3), \\ & (5.5), (5.6), (5.7), (5.8). \end{aligned} \quad (5.14)$$

As before, \mathbf{U} is a part of PA sum-problem and is updated whenever there is a change in PA. Algorithm 5 gives the general sequence of steps to solve P2 via VUL. To solve

Algorithm 5 : General iteration steps to solve DL via VUL

- 1: **Initialize:** α , \mathbf{U} , IPM parameters
 - 2: **repeat**
 - 3: solve P2_{vp} by Algorithm-1, Output: \mathbf{Q} , \mathbf{p} , \mathbf{U}
 - 4: evaluate $\bar{r}_j, \forall j$
 - 5: solve P2_b, Output: α
 - 6: **until** convergence
 - 7: **Output:** BSA α , DL-power \mathbf{Q} , DL beamforming vectors \mathbf{U} and VUL-power \mathbf{p} .
-

P4, similar VUL reformulation is obtained. Only the objective in P2 is changed to solve P4. Due to equation (5.3), the VUL objective can be either equation (5.15) or (5.16).

$$\sum_{i=1}^M \sum_{j=1}^N \alpha_{ij} q_{ij}, \quad (5.15) \quad \sum_{j=1}^N p_j. \quad (5.16)$$

$$\text{P4}_v: \underset{\alpha, \mathbf{p}, \mathbf{Q}, \mathbf{U}}{\text{minimize}} \quad \sum_{j=1}^N p_j \quad (5.17)$$

$$\text{subject to :} \quad \begin{aligned} & (2.1), (2.2), (2.7), (2.8), (2.10), \\ & (5.3), (5.5), (5.6), (5.7), (5.8). \end{aligned} \quad (5.18)$$

Similar VUL sub-problems P4_{vb} for DL BSA and P4_{vp} for DL PA are formulated for P4. Constraints (5.12) and (5.14) are retained without any change to solve P4_{vb} and P4_{vp} respectively. With $\bar{P}_i = \bar{P}, \forall i$, the constraint vector for the IPM in Algorithm 1 to solve P2_{vp} and P4_{vp} has a system wide constraint vector $\mathbf{c}_s(\mathbf{p}, \mathbf{q}, \mathbf{U})$, a per-UE constraint vector $\mathbf{c}_j(\mathbf{q}, \mathbf{U}), \forall j$, a per-BS constraint vector $\mathbf{c}_i(\mathbf{q}), \forall i$. They are given as

$$\mathbf{c}_s(\mathbf{p}, \mathbf{q}, \mathbf{U}) = \begin{bmatrix} \bar{P} \cdot M - \sum_{i=1}^M \sum_{j=1}^N \alpha_{ij} q_{ij} \\ \bar{P} \cdot M - \sum_{j=1}^N p_j \\ \sum_{j=1}^N p_j - \sum_{i=1}^M \sum_{j=1}^N \alpha_{ij} q_{ij} \\ \mathbf{q} - (\mathbf{I} - \tilde{\mathbf{D}} \cdot \tilde{\mathbf{C}})^{-1} \cdot \tilde{\mathbf{D}} \cdot (\sigma^2 \mathbf{1}_N) \end{bmatrix}, \quad (5.19)$$

$$\mathbf{c}_j(\mathbf{q}, \mathbf{U}) = \begin{bmatrix} \bar{P} - \boldsymbol{\alpha}_j^T \mathbf{q}_j \\ \bar{r}_j - r_{th} \\ \bar{P} - \mathbf{q}_j \\ \mathbf{q}_j \end{bmatrix}, \forall j, \quad (5.20)$$

$$\mathbf{c}_i(\mathbf{q}) = \begin{bmatrix} \bar{P} - \boldsymbol{\alpha}_i^T \mathbf{q}_i \end{bmatrix}, \forall i. \quad (5.21)$$

For link-level simulation of a minimum working example, $\underline{\sigma}^2 = \bar{\sigma}^2 \mathbf{1}$, $\bar{P}_i = 33$ dBm, $\forall i$, $T = 2$ is set. The actual maximum available DL power at the macro cell is 46 [dBm]. So for the demonstration of the Algorithm via duality, a value higher than the previously assumed UL values is set for the transmit power. The problem is assumed feasible and the channel state is perfectly known. $h_{ij}, \forall i, \forall j$, is i.i.d. zero-mean complex Gaussian with variance 0.5 per dimension, i.e., $\mathbf{h}_{ij} = \mathcal{CN}(0, \mathbf{I}_T)$. MMSE beamforming vectors (2.48) that maximize the received SINR are assumed for \mathbf{u}_{ij} .

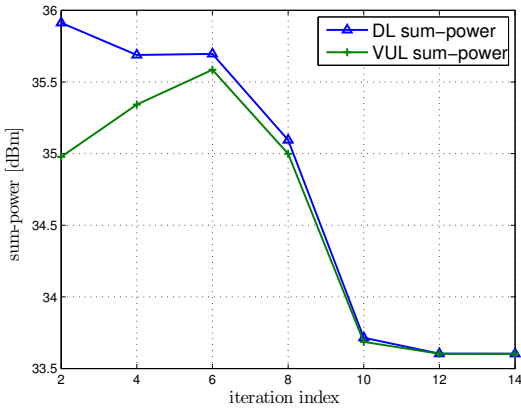


Fig. 5.1: Convergence of the objective in $P2_{vp}$ for a feasible realization.

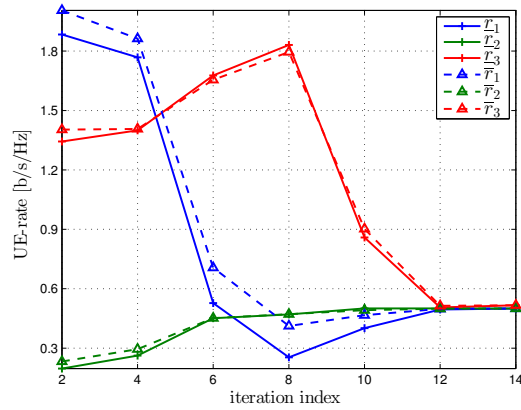


Fig. 5.2: Convergence of the DL and VUL UE rates in $P2_{vp}$ for a feasible realization.

Figure 5.1 shows the convergence of the objective for the $P2_{vp}$ sub-problem for a feasible channel instance with $N = 3$, $r_{th} = 0.5$, $M = 2$ and a given random $\boldsymbol{\alpha}$. Again these values are random to demonstrate the minimum working example. They can be scaled to match the actual requirement. It shows that both DL-power and VUL-power converge to the same value after starting from random initial points. As described in chapter 3, these initial points can be feasible or infeasible and the problem remains infeasible until the final convergence. At convergence the total sum-power in DL and VUL remains the same, which is one of the requirements of the reformulation. For a given BSA, only one element in \mathbf{q}_j is finite, this corresponds to the $\alpha_{ij} = 1$ in $\boldsymbol{\alpha}_j$. The remaining $(M - 1)$ elements in it correspond to the 0-elements. As mentioned before, in general for a UE $_j$, p_j is not equal to the corresponding non-zero q_{ij} . Figure 5.2 shows the convergence of the UE rates for the same feasible instance of Figure 5.1. The DL-rate \mathcal{L}_j and the corresponding VUL-rate \bar{r}_j in $P2_{vp}$ converge to the same value, i.e., satisfying the definition of DL-VUL duality. With this Lagrange-duality of

the IPM the single-cell duality becomes implicit in the formulation. The plots converge as required under the same total sum-power equality on the DL and VUL and also under the same VUL and DL noise variance.

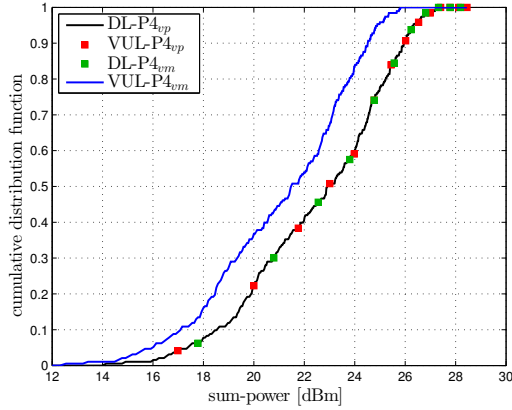


Fig. 5.3: CDF of VUL and DL sum-power for P_{4vp} , P_{4vm} .

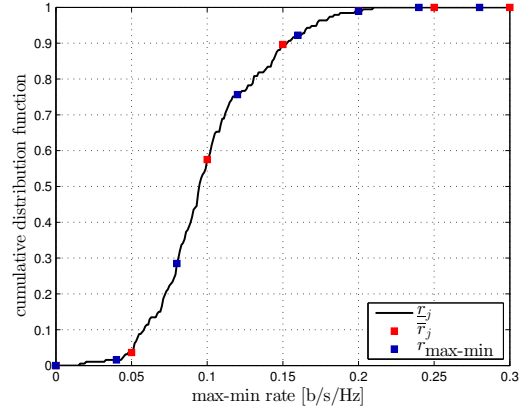


Fig. 5.4: CDF of max-min rate for P_{4vp} , P_{4vm} for a UE.

As explained in chapter 3, the sum-power minimization problem with QoS is coupled to the max-min problem. The feasibility of the later leads to the formulation of the former. A multi-cell max-min problem on the DL [141] can be formulated similar to P_5 on the UL. For a set of homogeneous UEs, the DL max-min problem achieves a balanced QoS across all the UEs after convergence. The max-min problem with VUL-noise as an optimization variable (P_{4vm}) [141] is one way to satisfy duality. For comparison, P_{4vm} is first solved for a randomly assigned BS and then the obtained balanced max-min rate $r_{\max\text{-min}}$ is used as the QoS for P_{4v} . Optimum α is not known for P_{4vm} until an exhaustive search is carried out. So to avoid an exhaustive BSA search a random feasible assignment is considered. Figure 5.3 plots the CDF of P_{4vp} and P_{4vm} for $M = N = 3$. It is observed that the DL sum-power and the VUL sum-power in P_{4vp} , the DL sum-power in P_{4vm} converge to the same value. However, the VUL sum-power of P_{4vm} is shifted the left, i.e., the sum-power on the VUL is lower than its corresponding sum-power on the DL. VUL-noise $\bar{\sigma}_j^2, \forall j$, is an optimization variable in P_{4vm} . It has been observed from the simulation results that $\bar{\sigma}_j^2, \forall j$, converges to a value less than its corresponding $\underline{\sigma}_j^2 = 1, \forall j$. So the thermal noise variance in the UE interference term is lower in the VUL and so is the VUL sum-power. Single-cell duality requirement in equation (5.3) is satisfied by P_{4vp} but not by P_{4vm} . Also for P_{4vp} , $\underline{\sigma}_j^2 = \bar{\sigma}_j^2 = 1, \forall j$, a constant and not an optimization variable. Figure 5.4 shows that the required $r_{\max\text{-min}}$, i.e., r_{th} is achieved by P_{4vp} on both DL and VUL. Since every UE has the same QoS value, only one UE rate is plotted. Also, P_{4vb} converges to the same chosen random BSA of P_{4vm} .

Figure 5.5 compares the CDF of P_{4v} with the VUL sum-power objective (5.15) and DL sum-power objective (5.16) for various (N, r_{th}) and $M = 2$. As mentioned before, both converge to the same optimal. An increase in N or r_{th} increases the objective, shown by

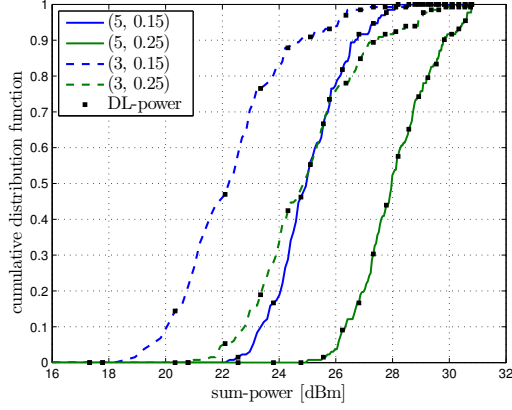


Fig. 5.5: CDF of $P4_v$ with VUL objective and DL objective.

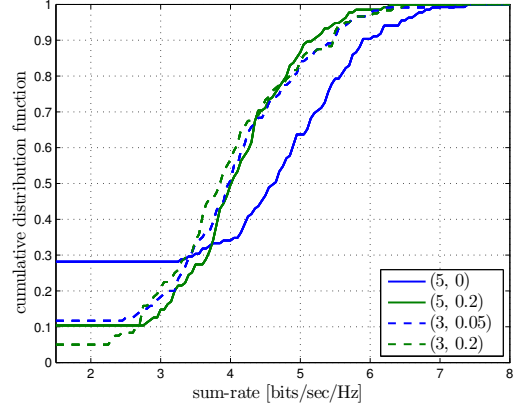


Fig. 5.6: CDF of sum-rate for $P2_v$.

the right shift in CDF curve. Figure 5.6 plots the CDF of the sum-rate objective in $P2_v$ for various (N, r_{th}) and $M = 2$. For a given N , the CDF curve shifts to left with increase in r_{th} , i.e., a drop in the objective. This is due to the increased MUI at each UE and also due to the competing nature of UEs to obtain resources. Maximum throughput is achieved when $r_{th} = 0$. This is the simple sum-rate maximization problem. This increase is due to the MUD, where in the interest of the system, UEs with good channel conditions are served with a higher resource value. UEs with unfavourable conditions may have the service deprived or may have a low resource value allocated. There is also an increase in the outage for this case.

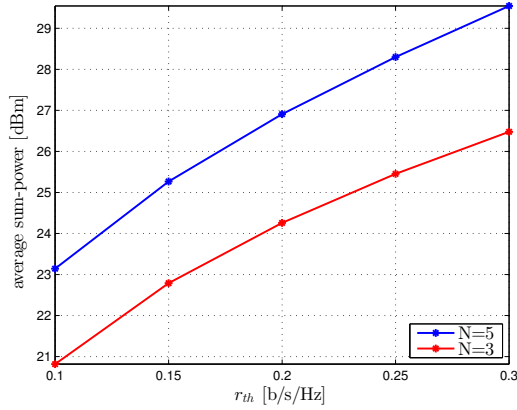


Fig. 5.7: Average sum-power vs QoS in $P4_v$.

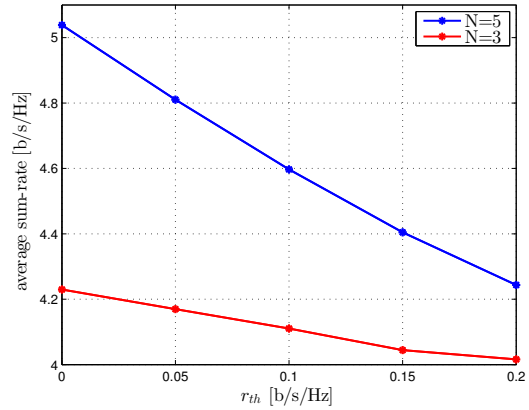


Fig. 5.8: Average sum-rate vs QoS in $P2_v$.

Figure 5.7 plots the average sum-power vs QoS for different N in $P4_v$. An increase in objective with increase in r_{th} and N is observed. This increment is due to the increase in MUI at each UE. On the VUL, an increase in the MUI prompts each UE to increase their respective power level to achieve the same QoS, thereby increasing the overall power in the system. Figure 5.8 plots the average sum-rate vs QoS for different N in

P2_v. For a given N , the maximum objective value is achieved for $r_{th} = 0$ due to MUD. Also, the sum-rate objective reduces with the increase in r_{th} . The upward shift in the curve from $N = 3$ to $N = 5$ shows an increase in the objective value. But the shift is reduced with increase in r_{th} . This shows the interference-limited nature of the shared channel and the competing nature of UEs.

Chapter 6

Conclusion

In this thesis various multi-cell multi-user resource allocation problems based on rate maximization and power allocation under perfect Channel State Information (CSI), statistical-CSI and imperfect-CSI were addressed on both the Uplink and Downlink. These optimization problems are in general Mixed Integer Nonlinear Programming problems. A mathematical framework based on Interior Point Method (IPM) was identified to solve a variety of non-linear optimization problems. This framework eliminates the need to choose problem specific mathematical tool. However, several other mathematical tools exist to obtain efficient distributed processing. The Integer Programming sub-problem of Base Station Association was also solved as a Nonlinear Programming sub-problem. This IPM facilitates simultaneous solving of both the Base Station Association and Power Allocation sub-problems. A new definition of feasibility via ℓ_1 -norm handle has been implemented for the Power Allocation. Under statistical Channel State Information, the concept of tail distributions in Extreme Value Theory (EVT) has been used to solve the problems involving Probabilistic Constraints. To handle the Chance Constrained problems, the definitions of risk measurement elements of Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) from the field of finance were applied. To find the optimal solution with the Probabilistic Constraints, the concept of VaR and a block-maxima based Generalized Extreme Value Distribution in EVT are together implemented. This reformulation has an advantage that the CDF of the unknown random variable need not be invertible and no approximations were required. It can be extended to most of the distributions where the quantile function does not exist. Performance loss based on the finite probability of exceeding the evaluated VaR is also considered. While evaluating this, CVaR and block-maxima based Generalized Pareto Distribution in EVT are combined to overcome the non-availability of closed form expressions. The VaR concept can also be extended to the resource allocation problem under imperfect-CSI after identifying the correct distribution of the imperfect channel. EVT based approach reformulates the problem into a Nonlinear Programming problem which can be solved by the IPM framework. The generated bounds clearly outperforms the existing results. On the highly coupled Downlink, the duality principle has been implemented based on the implicit Lagrange-duality of the

IPM. The IPM maintains consistency with the single-cell duality theory and makes no change to the noise components, which were excluded by the existing methods.

In this thesis, a small fraction of the mathematics as applied to mobile wireless communication has been addressed. The presented results have been numerically solved for a link-level setup under Rayleigh fading. To demonstrate the effectiveness of the considered mathematical tools, a minimum working example has been considered in each case. The analysis and the results contribute to a better understanding of a conventional wireless communication system. It is evident that this work needs a further investigation. As a part of future work, to realize the effectiveness of the obtained results, actual channel models and system-level parameters must be considered. Though in this thesis, a binary 0 – 1 problem has been addressed as a Base Station Association problem, with an additional power constraint inclusion, a multi-carrier resource allocation problem can be solved. Multi-carrier analysis is a more practical problem, since most of the deployed wireless networks are multi-carrier in nature. The current research trends demand the network architecture to be more ad-hoc, small-cell oriented and decentralized. Including the heterogeneous setup, where multiple small cells coexist in a macro cell is an interesting setup to further extend these results. The obtained results assume a centralized processing. Decentralized processing, the actual amount of CSI and processing capability of the network entities must also be addressed for more robust analysis. The derived results must also include cross layer architecture, since the functionalities at each layer are interdependent. To include all these high complexity design considerations, experimental data combined with efficient simulation environment is a must for the challenging task of Resource Allocation and Interference Management.

Publications

- [1] Chitti. K, Tan. X, ten Brink. S.: "Multi-User Downlink Resource Allocation with per-Base Station Power Constraints", 21st IEEE Symposium on Communications and Vehicular Technology (IEEE SCVT), Nov 2014, Delft, Netherlands.
- [2] Chitti. K, Tan. X, ten Brink. S.: "Multi-User Uplink Power Control under Imperfect-CSI and Probabilistic Constraints", 2014 International Conference on Wireless Communications and Signal Processing (WCSP), Oct 2014, Hefei, China.
- [3] Chitti, K.; Speidel, J.: "MU-MIMO Power Control under Statistical CSI and Probabilistic Constraints", 11th International Symposium on Wireless Communication Systems (ISWCS), Aug 2014, Barcelona, Spain.
- [4] Chitti. K, Speidel. J.: "Joint Base Station Association and Power Allocation for Uplink Sum-Power Minimization", 2013 IEEE 78th Vehicular Technology Conference: VTC 2013-Fall, September 2013, Las Vegas, USA.
- [5] Chitti, K; Kuang, Q; Speidel, J.: Joint Base Station Association and Power Allocation for Uplink Sum-Rate Maximization, 2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Darmstadt, Germany, Jun. 2013.

Bibliography

- [1] “Lte-a, requirements for further advancements for eutra . 3gpp tr 36.913,” 2008.
- [2] [Online]. Available: http://www.keysight.com/upload/cmc_upload/All/Agilent-LTE-Book-Chap8-Looking-Towards-4G-LTE-Advanced-2294001.pdf?&cc=DE&lc=ger
- [3] M. Marwangi, N. Fisal, S. Yusof, R. A. Rashid, A. S. Ghafar, F. A. Saparudin, and N. Katiran, “Challenges and practical implementation of self-organizing networks in lte/lte-advanced systems,” in *Information Technology and Multimedia (ICIM), 2011 International Conference on*. IEEE, 2011, pp. 1–5.
- [4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive mimo for next generation wireless systems,” *arXiv preprint arXiv:1304.6690*, 2013.
- [5] [Online]. Available: <http://www.huawei.com/en/static/AAS-129092-1-197969.pdf>
- [6] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: from theory to practice*. Cambridge University Press, 2011.
- [7] Z. Feng, W. Muqing, and L. Huixin, “Coordinated multi-point transmission and reception for lte-advanced,” in *Proceedings of the 5th International Conference on Wireless communications, networking and mobile computing*. IEEE Press, 2009, pp. 255–258.
- [8] H. Dahrouj and W. Yu, “Coordinated beamforming for the multicell multi-antenna wireless system,” *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1748–1759, 2010.
- [9] V. Srivastava and M. Motani, “Cross-layer design: a survey and the road ahead,” *Communications Magazine, IEEE*, vol. 43, no. 12, pp. 112–119, 2005.
- [10] R. Samano-Robles and A. Gameiro, “Stability properties of network diversity multiple access with multiple-antenna reception and imperfect collision multiplicity estimation,” *Journal of Computer Networks and Communications*, vol. 2013, 2013.

- [11] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201–220, 2005.
- [12] H. T. Friis, "A note on a simple transmission formula," *proc. IRE*, vol. 34, no. 5, pp. 254–256, 1946.
- [13] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [14] S. S. Haykin, M. Moher, and D. Koilpillai, *Modern wireless communications*. Pearson Education India, 2011.
- [15] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. i. characterization," *Communications Magazine, IEEE*, vol. 35, no. 7, pp. 90–100, 1997.
- [16] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels*. John Wiley & Sons, 2005, vol. 95.
- [17] M. D. Yacoub, "The η - μ distribution: a general fading distribution," in *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, vol. 2. IEEE, 2000, pp. 872–877.
- [18] —, "The κ - μ distribution: a general fading distribution," in *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, vol. 3. IEEE, 2001, pp. 1427–1431.
- [19] —, "The α - μ distribution: A general fading distribution," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2. Citeseer, 2002, pp. 629–633.
- [20] G. Fraidenraich and M. D. Yacoub, "The λ - μ general fading distribution," in *Microwave and Optoelectronics Conference, 2003. IMOC 2003. Proceedings of the 2003 SBMO/IEEE MTT-S International*, vol. 1. IEEE, 2003, pp. 49–54.
- [21] —, "The α - η - μ and α - κ - μ fading distributions," in *Spread Spectrum Techniques and Applications, 2006 IEEE Ninth International Symposium on*. IEEE, 2006, pp. 16–20.
- [22] T. Novlan, J. G. Andrews, I. Sohn, R. K. Ganti, and A. Ghosh, "Comparison of fractional frequency reuse approaches in the ofdma cellular downlink," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–5.
- [23] L. Chen and D. Yuan, "Soft frequency reuse in large networks with irregular cell pattern: How much gain to expect?" in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE, 2009, pp. 1467–1471.

- [24] R. Ghaffar and R. Knopp, "Fractional frequency reuse and interference suppression for ofdma networks," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*. IEEE, 2010, pp. 273–277.
- [25] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4g networks," *Communications Magazine, IEEE*, vol. 47, no. 4, pp. 74–81, 2009.
- [26] J. A. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *Communications Magazine, IEEE*, vol. 28, no. 5, pp. 5–14, 1990.
- [27] L. J. Cimini Jr and N. R. Sollenberger, "Ofdm with diversity and coding for high-bit-rate mobile data applications," in *Mobile Multimedia Communications*. Springer, 1997, pp. 247–254.
- [28] S. Kaiser, "mc-fdma and mc-tdma versus mc-cdma and ss-mc-ma: performance evaluation for fading channels," in *Spread Spectrum Techniques and Applications, 1998. Proceedings., 1998 IEEE 5th International Symposium on*, vol. 1. IEEE, 1998, pp. 200–204.
- [29] M. Morelli, C. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (ofdma): A tutorial review," *PROCEEDINGS-IEEE*, vol. 95, no. 7, p. 1394, 2007.
- [30] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser ofdm with adaptive subcarrier, bit, and power allocation," *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 10, pp. 1747–1758, 1999.
- [31] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (noma) for cellular future radio access," in *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*. IEEE, 2013, pp. 1–5.
- [32] G. Fettweis, M. Krondorf, and S. Bittner, "Gfdm-generalized frequency division multiplexing," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*. IEEE, 2009, pp. 1–4.
- [33] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.
- [34] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless personal communications*, vol. 6, no. 3, pp. 311–335, 1998.

- [35] M.-A. Khalighi, J. Brossier, G. Jourdain, and K. Raoof, "Water filling capacity of rayleigh mimo channels," in *Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on*, vol. 1. IEEE, 2001, pp. A-155.
- [36] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge university press, 2003.
- [37] H. Jafarkhani, *Space-time coding: theory and practice*. Cambridge university press, 2005.
- [38] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell labs technical journal*, vol. 1, no. 2, pp. 41-59, 1996.
- [39] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *Information Theory, IEEE Transactions on*, vol. 49, no. 10, pp. 2658-2668, 2003.
- [40] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "On gaussian mimo bc-mac duality with multiple transmit covariance constraints," *Information Theory, IEEE Transactions on*, vol. 58, no. 4, pp. 2064-2078, 2012.
- [41] S. Shi, M. Schubert, and H. Boche, "Downlink mmse transceiver optimization for multiuser mimo systems: Duality and sum-mse minimization," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5436-5446, 2007.
- [42] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1277-1294, 2002.
- [43] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of cdma-hdr a high efficiency-high data rate personal communication wireless system," in *Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo. 2000 IEEE 51st*, vol. 3. IEEE, 2000, pp. 1854-1858.
- [44] S. Shakkottai, S. G. Shakkottai, and R. Srikant, *Network optimization and control*. Now Publishers Inc, 2008.
- [45] E. Altman, K. Avrachenkov, and A. Garnaev, "Generalized alpha-fair resource allocation in wireless networks," in *Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico, 2008*, pp. 2414-2419.
- [46] M. Zukerman, M. Mammadov, L. Tan, I. Ouveysi, and L. L. Andrew, "To be fair or efficient or a bit of both," *Computers & Operations Research*, vol. 35, no. 12, pp. 3787-3806, 2008.

- [47] [Online]. Available: http://users.ece.utexas.edu/~jandrews/ee381k/EE381KTA/Chaper2_diss_Jeff.pdf
- [48] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint tx-rx beamforming design for multicarrier mimo channels: A unified framework for convex optimization," *Signal Processing, IEEE Transactions on*, vol. 51, no. 9, pp. 2381–2401, 2003.
- [49] A. J. Tenenbaum and R. S. Adve, "Linear processing and sum throughput in the multiuser mimo downlink," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2652–2661, 2009.
- [50] M. H. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, 1983.
- [51] Q. H. Spencer and M. Haardt, "Capacity and downlink transmission algorithms for a multi-user mimo channel," in *Signals, Systems and Computers, 2002. Conference Record of the Thirty-Sixth Asilomar Conference on*, vol. 2. IEEE, 2002, pp. 1384–1388.
- [52] V. Stankovic and M. Haardt, "Multi-user mimo downlink precoding for users with multiple antennas," in *Proc. of the 12-th Meeting of the Wireless World Research Forum (WWRF), Toronto, ON, Canada*, vol. 10, 2004, pp. 12–14.
- [53] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 3, pp. 528–541, 2006.
- [54] M. Sharif and B. Hassibi, "On the capacity of mimo broadcast channels with partial side information," *Information Theory, IEEE Transactions on*, vol. 51, no. 2, pp. 506–522, 2005.
- [55] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in mimo communications systems," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, pp. 2700–2712, 2005.
- [56] S. Shi and M. Schubert, "Mmse transmit optimization for multi-user multi-antenna systems," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings (ICASSP'05). IEEE International Conference on*, vol. 3. IEEE, 2005, pp. iii–409.
- [57] S. Serbetli and A. Yener, "Transceiver optimization for multiuser mimo systems," *Signal Processing, IEEE Transactions on*, vol. 52, no. 1, pp. 214–226, 2004.
- [58] M. Codreanu, A. Tolli, M. Juntti, and M. Latva-aho, "Mimo downlink weighted sum rate maximization with power constraints per antenna groups," in *IEEE 65th Vehicular Technology Conference, 2007, VTC2007-Spring*. IEEE, 2007, pp. 2048–2052.

- [59] B. O. Lee, H. Je, I. Sohn, O.-S. Shin, and K. B. Lee, "Interference-aware decentralized precoding for multicell mimo tdd systems," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*. IEEE, 2008, pp. 1–5.
- [60] J. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter and receiver beamforming for maximum capacity in spatial division multiaccess," in *PROCEEDINGS OF THE ANNUAL ALLERTON CONFERENCE ON COMMUNICATION CONTROL AND COMPUTING*, vol. 35. Citeseer, 1997, pp. 93–101.
- [61] Y. Rong, Y. C. Eldar, and A. B. Gershman, "Performance tradeoffs among beamforming approaches," in *Sensor Array and Multichannel Processing, 2006. Fourth IEEE Workshop on*. IEEE, 2006, pp. 26–30.
- [62] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Space-time transmission using tomlinson-harashima precoding," 2002. [Online]. Available: http://www.lit.lnt.de/papers/itg02_fi.pdf
- [63] M. Joham, J. Brehmer, and W. Utschick, "Mmse approaches to multiuser spatio-temporal tomlinson-harashima precoding," *ITG FACHBERICHT*, pp. 387–394, 2004.
- [64] E. Zeydan, D. Kivanc-Tureli, and U. Tureli, "Iterative beamforming and power control for mimo ad hoc networks," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–5.
- [65] E. Manskani, N. D. Sidiropoulos, Z.-q. Luo, and L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 7, pp. 2682–2693, 2008.
- [66] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for mimo spatial multiplexing: theory and algorithms," *Signal Processing, IEEE Transactions on*, vol. 51, no. 11, pp. 2796–2807, 2003.
- [67] S. A. Grandhi, R. VIJAVAN, D. J. Goodman, and J. Zander, "Centralized power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 466–468, 1993.
- [68] F. Rashid-Farrokhi, L. Tassiulas, and K. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *Communications, IEEE Transactions on*, vol. 46, no. 10, pp. 1313–1324, 1998.
- [69] D. M. Himmelblau, B. Clark, and M. Eichberg, *Applied nonlinear programming*. McGraw-Hill New York, 1972, vol. 111.
- [70] [Online]. Available: http://stanford.edu/~boyd/papers/pdf/gp_tutorial.pdf

- [71] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2009.
- [72] [Online]. Available: <http://users.isy.liu.se/johanl/yalmip/pmwiki.php?n=Main.HomePage>
- [73] [Online]. Available: <http://www.math.nus.edu.sg/~mattohkc/sdpt3.html>
- [74] [Online]. Available: <http://www.stanford.edu/~boyd/ggplab/>
- [75] [Online]. Available: <http://sedumi.ie.lehigh.edu/>
- [76] [Online]. Available: https://www-304.ibm.com/ibm/university/academic/pub/page/ban_ilog_programming
- [77] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. Slock, “On the mimo interference channel,” in *Information Theory and Applications Workshop (ITA), 2010*. IEEE, 2010, pp. 1–9.
- [78] S. Rao and S. Rao, *Engineering Optimization: Theory and Practice*. John Wiley & Sons, 2009. [Online]. Available: <http://books.google.de/books?id=YNt34dvnQLEC>
- [79] M. L. Fisher, “The lagrangian relaxation method for solving integer programming problems,” *Management science*, vol. 50, no. 12_supplement, pp. 1861–1871, 2004.
- [80] A. Conejo, E. Castillo, and R. Mínguez, *Decomposition techniques in mathematical programming*. Springer-Verlag Berlin Heidelberg, 2006.
- [81] T. J. Van Roy, “Cross decomposition for mixed integer programming,” *Mathematical programming*, vol. 25, no. 1, pp. 46–63, 1983.
- [82] K. Jörnsten and M. Näsberg, “A new lagrangian relaxation approach to the generalized assignment problem,” *European Journal of Operational Research*, vol. 27, no. 3, pp. 313–323, 1986.
- [83] B. Han, J. Leblet, and G. Simon, “Hard multidimensional multiple choice knapsack problems, an empirical study,” *Computers & operations research*, vol. 37, no. 1, pp. 172–181, 2010.
- [84] S. Martello and P. Toth, *Knapsack problems*. Wiley New York, 1990.
- [85] N. Shor, *Minimization methods for non-differentiable functions*, ser. Springer series in computational mathematics. Springer-Verlag, 1985. [Online]. Available: <http://books.google.de/books?id=Zi6rAAAIAAJ>

- [86] L. Han-Lin, “An approximate method for local optima for nonlinear mixed integer programming problems,” *Computers & operations research*, vol. 19, no. 5, pp. 435–444, 1992.
- [87] J. Gallier, “The schur complement and symmetric positive semidefinite (and definite) matrices,” *December*, vol. 44, pp. 1–12, 2010.
- [88] F. Glover and E. Taillard, “A user’s guide to tabu search,” *Annals of operations research*, vol. 41, no. 1, pp. 1–28, 1993.
- [89] M. Dorigo, M. Birattari, and T. Stutzle, “Ant colony optimization,” *Computational Intelligence Magazine, IEEE*, vol. 1, no. 4, pp. 28–39, Nov 2006.
- [90] S. Carl and S. Heikkilä, *Fixed Point Theory in Ordered Sets and Applications: From Differential and Integral Equations to Game Theory*. Springer, 2010.
- [91] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1341–1347, 1995.
- [92] M. Chiang, C. W. Tan, D. P. Palomar, D. O’Neill, and D. Julian, “Power control by geometric programming,” *Wireless Communications, IEEE Transactions on*, vol. 6, no. 7, pp. 2640–2651, 2007.
- [93] C. W. Tan, D. P. Palomar, and M. Chiang, “Solving nonconvex power control problems in wireless networks: Low sir regime and distributed algorithms,” in *Global Telecommunications Conference, 2005. GLOBECOM’05. IEEE*, vol. 6. IEEE, 2005, pp. 6–pp.
- [94] M. Avriel and A. Williams, “An extension of geometric programming with applications in engineering optimization,” *Journal of Engineering Mathematics*, vol. 5, no. 2, pp. 187–194, 1971.
- [95] N. Vucic, S. Shi, and M. Schubert, “Dc programming approach for resource allocation in wireless networks,” in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*. IEEE, 2010, pp. 380–386.
- [96] R. Eberhart and Y. Shi, “Particle swarm optimization: developments, applications and resources,” in *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on*, vol. 1, 2001, pp. 81–86 vol. 1.
- [97] T. Bäck and H.-P. Schwefel, “An overview of evolutionary algorithms for parameter optimization,” *Evolutionary computation*, vol. 1, no. 1, pp. 1–23, 1993.
- [98] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *et al.*, “Optimization by simulated annealing,” *science*, vol. 220, no. 4598, pp. 671–680, 1983.

- [99] [Online]. Available: <http://www.neos-guide.org/content/sequential-quadratic-programming>
- [100] S. Wright and J. Nocedal, *Numerical optimization*. Springer New York, 1999, vol. 2.
- [101] R. M. Freund and S. Mizuno, *Interior point methods: current status and future directions*. Springer, 2000.
- [102] S. Schaible, “Fractional programming: applications and algorithms,” *European Journal of Operational Research*, vol. 7, no. 2, pp. 111–120, 1981.
- [103] R. H. Louie, M. R. McKay, N. Jindal, and I. B. Collings, “Spatial multiplexing with mmse receivers: Single-stream optimality in ad hoc networks,” *arXiv preprint arXiv:1003.3056*, 2010.
- [104] A. Wiesel, Y. C. Eldar, and S. Shamai, “Linear precoding via conic optimization for fixed mimo receivers,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 1, pp. 161–176, 2006.
- [105] Z.-q. Luo, W.-k. Ma, A.-C. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” *Signal Processing Magazine, IEEE*, vol. 27, no. 3, pp. 20–34, 2010.
- [106] C. A. Floudas, *Nonlinear and mixed-integer optimization: fundamentals and applications*. Oxford University Press, 1995.
- [107] S. Wright, *Primal-dual interior-point methods*. Society for Industrial Mathematics, 1987, vol. 54.
- [108] H. Tomizuka and H. Yabe, “A hybrid conjugate gradient method for unconstrained optimization.”
- [109] C. W. Tan, D. Palomar, and M. Chiang, “Solving nonconvex power control problems in wireless networks: low sir regime and distributed algorithms,” in *Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE*, vol. 6, dec 2005, pp. 3445–3450.
- [110] M. Avriel and A. Williams, “An extension of geometric programming with applications in engineering optimization,” *Journal of Engineering Mathematics*, vol. 5, no. 2, pp. 187–194, 1971.
- [111] M. Schubert and H. Boche, “Solution of the multiuser downlink beamforming problem with individual sinr constraints,” *Vehicular Technology, IEEE Transactions on*, vol. 53, no. 1, pp. 18–28, 2004.

- [112] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1341–1347, 1995.
- [113] S. A. Grandhi and J. Zander, “Constrained power control in cellular radio systems,” in *Vehicular Technology Conference, 1994 IEEE 44th*. IEEE, 1994, pp. 824–828.
- [114] R. D. Yates and C.-Y. Huang, “Integrated power control and base station assignment,” *Vehicular Technology, IEEE Transactions on*, vol. 44, no. 3, pp. 638–644, 1995.
- [115] X. Lu, W. Li, A. Tölli, M. Juntti, E. Kunnari, and O. Piirainen, “Joint power control, receiver beamforming and adaptive multi base station coordination for uplink wireless communications,” in *Personal, Indoor and Mobile Radio Communications Workshops , 2010 IEEE 21st International Symposium on*. IEEE, 2010, pp. 446–450.
- [116] M. Chiang, P. Hande, and T. Lan, *Power control in wireless cellular networks*. Now Pub, 2008, vol. 2, no. 4.
- [117] D. Evangelinakis, N. Sidiropoulos, and A. Swami, “Joint admission and power control using branch & bound and gradual admissions,” in *Signal Processing Advances in Wireless Communications (SPAWC), 2010 IEEE Eleventh International Workshop on*. IEEE, 2010, pp. 1–5.
- [118] S. Boyd, “ l_1 -norm methods for convex-cardinality problems.” [Online]. Available: http://www.stanford.edu/class/ee364b/lectures/l1_slides.pdf
- [119] A. Grundinger, A. Butabayeva, M. Joham, and W. Utschick, “Chance constrained and ergodic robust qos power minimization in the satellite downlink,” in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*. IEEE, 2012, pp. 1147–1151.
- [120] K. Chitti and J. Speidel, “Joint base station association and power allocation for uplink sum-power minimization,” in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*. IEEE, 2013, pp. 1–5.
- [121] A. Nemirovski and A. Shapiro, “Convex approximations of chance constrained programs,” *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 969–996, 2006.
- [122] N. Y. Soltani, S. J. Kim, and G. B. Giannakis, “Chance constrained optimization of ofdma cognitive radio uplinks,” *IEEE transactions on wireless communications*, vol. 12, no. 3, pp. 1098–1107, 2013.

- [123] S. Parsaeefard, A. R. Sharafat, and M. Rasti, "Robust probabilistic distributed power allocation by chance constraint approach," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*. IEEE, 2010, pp. 2162–2167.
- [124] K.-Y. Wang, T.-H. Chang, W.-K. Ma, A.-C. So, and C.-Y. Chi, "Probabilistic sinr constrained robust transmit beamforming: A bernstein-type inequality based conservative approach," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 3080–3083.
- [125] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," 1999.
- [126] [Online]. Available: <http://www.mathmistakes.info/facts/CalculusFacts/learn/doi/doib.html>
- [127] R. Gençay and F. Selçuk, "Extreme value theory and value-at-risk: relative performance in emerging markets," *International Journal of Forecasting*, vol. 20, no. 2, pp. 287–303, 2004.
- [128] R. S. Tsay, "Extreme values and their applications in finance," 2007. [Online]. Available: <http://www.rmi.nus.edu.sg/events/files/paper/extreme%20values%20and%20their%20applications%20in%20finance.pdf>
- [129] J. G. Proakis, *Digital communications, 1995*. McGraw-Hill, New York.
- [130] Y. Yao and G. B. Giannakis, "Rate-maximizing power allocation in ofdm based on partial channel knowledge," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 3, pp. 1073–1083, 2005.
- [131] I. Bechar, "A bernstein-type inequality for stochastic processes of quadratic forms of gaussian variables," *arXiv:0909.3595*, 2009.
- [132] Q. Li, W.-K. Ma, and A.-C. So, "Safe convex approximation to outage-based miso secrecy rate optimization under imperfect csi and with artificial noise," in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, Nov 2011, pp. 207–211.
- [133] S. He, Y. Huang, H. Wang, A. Nallanathan, and L. Yang, "Coordinated multi-cell beamforming scheme using uplink-downlink max-min sinr duality," in *2012 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2012, pp. 3930–3934.
- [134] H. Boche and M. Schubert, "Duality theory for uplink and downlink multiuser beamforming," *Smart Antennas*, p. 545.

- [135] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual sinr constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, 2004.
- [136] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve, "Linear processing for the downlink in multiuser mimo systems with multiple data streams," in *IEEE International Conference on Communications, 2006. ICC'06*, vol. 9. IEEE, 2006, pp. 4113–4118.
- [137] B. Song, R. L. Cruz, and B. D. Rao, "Network duality and its application to multiuser mimo wireless networks with sinr constraints," in *2005 IEEE International Conference on Communications, 2005. ICC 2005.*, vol. 4. IEEE, 2005, pp. 2684–2689.
- [138] W. Yang and G. Xu, "The optimal power assignment for smart antenna downlink weighting vector design," in *48th IEEE Vehicular Technology Conference, 1998. VTC 98*, vol. 1, May 1998, pp. 485–488 vol.1.
- [139] J. Yang and D. K. Kim, "Multi-cell uplink-downlink beamforming throughput duality based on lagrangian duality with per-base station power constraints," *IEEE Communications Letters*, vol. 12, no. 4, pp. 277–279, 2008.
- [140] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, 2007.
- [141] Y.-H. Yang, S.-C. Lin, and H.-J. Su, "Multiuser mimo downlink beamforming design based on group maximum sinr filtering," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1746–1758, 2011.