Resource optimisation and interference management for multiuser communication systems in a cooperative environment

Barman Roy, Swagato

2017

Barman Roy, S. (2017). Resource optimisation and interference management for multiuser communication systems in a cooperative environment. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69826>

<https://doi.org/10.32657/10356/69826>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**Resource Optimisation and Interference
Management for Multiuser Communication Systems
in a Cooperative Environment**

Swagato Barman Roy
School of Computer Science & Engineering
2017

**Resource Optimisation and Interference
Management for Multiuser Communication Systems
in a Cooperative Environment**

Swagato Barman Roy

School of Computer Science & Engineering

A thesis submitted to the Nanyang Technological University in partial fulfilment
for the requirement of the degree of
Doctor of Philosophy

2017

Acknowledgment

My sincere gratitude goes to Dr. A. S. Madhukumar for giving me the opportunity to pursue research in the exciting field of multichannel communication and making me part of his research group. His continuous encouragement for my effort and reviews of my ideas in spite of his busy schedule are indispensable parts of this work. I look forward for his continued support as a teacher and mentor along my way.

In addition, I will thank my colleagues, research fellows and technical staffs of former CeMNet and current CNCL for supporting me at crucial times and the School of Computer Science & Engineering, NTU for giving me the opportunity to be part of its big family.

Summary

1	Introduction	1
2	Review of Rate Regions & Interference Management	10
3	Resource Optimisation in Single Cell Systems	46
4	Generalised Pareto Frontier & Relay Systems	68
5	Multicell Cooperation & Two Layer Decoding	86
6	Extension of Duality to Cooperative Multipoint Transmission	98
7	Conclusions and the Way Ahead	119
A	Review of Inner Product Space	124
B	Set Functions & Polymatroids	127
C	Duality with Linear Beamforming	129
D	Wirtinger Calculus Fundamentals	133
E	Constraints on Convex Problems	135
	Bibliography	137

Contents

List of Figures	viii
List of Tables	x
List of Abbreviations	xi
List of Mathematical Notations	xiii
Abstract	xiv
1 Introduction	1
1.1 Multi Antenna Communication Systems	2
1.2 Multi-Hop Communication Systems Aided by Relays	3
1.3 Multicell Communication Systems	4
1.4 Fifth Generation LTE Standards	5
1.5 Effects of Channel State Information	5
1.6 Motivation of the research	5
1.7 Contributions of this work	6
1.8 Organisation of the thesis	8
2 Review of Rate Regions & Interference Management	10
2.1 Mathematical Preliminaries	11
2.1.1 Degrees of Freedom of a Communication System	11
Degree of Freedom for Multiuser Networks	13
2.1.2 Convex Sets and Regions	13
Duality and Convex Optimisation	14
Convex Combination as Time Sharing	14
Achievable Rate from Signal to Noise Ratio	15
2.2 Review of Rate Regions for Single Cell Systems	16
2.2.1 Multiple Access Channel	16
2.2.2 Broadcast Channel	18

	Crosstalk Matrix	19
	Each User with Unique Steering Vectors	19
	Superposition Coding	21
2.2.3	Duality between System Models	23
	Uplink-Downlink Duality	23
	Minimax Duality	26
2.2.4	Duality in Interference Channel	27
2.2.5	Single Cell Optimisation Strategies	28
	Interference Function	28
	Sum Power Minimisation with Matrix Based Interference	30
	SINR Balancing with Spectral Radius Optimisation	31
2.3	Interference Management in Traditional Communication Systems	33
2.3.1	Interference Cancellation Decoding	33
2.3.2	Interference Ignorant Decoding	34
2.3.3	Interference Alignment Decoding	36
	Degree of Freedom Enhancement	36
	Linear Beamforming in Two User Interference Channel	36
	Interference Alignment in Multiuser System	37
2.4	Interference Management in Relay Channels	38
2.4.1	Amplify and Forward Relay Assisted Systems	39
2.4.2	Decode and Forward Relay Assisted Systems	39
2.5	Game Theoretic Interference Management	40
2.5.1	Basics of a Non-Cooperative Game	40
2.5.2	Interference Channel as a Non Cooperative Game	41
2.6	Interference Management in Multicell Cooperation	41
2.6.1	Coordinated Beamforming among Base Stations	42
2.6.2	Joint Processing with Perfect Data Sharing	42
2.6.3	Spatial Sensing	43
2.7	Summary and Shortcomings of Previous Work	43
2.7.1	Scope of this Thesis	45
2.8	Conclusion	45
3	Resource Optimisation in Single Cell Systems	46
3.1	The System Model	47

3.1.1	Sum Rate Maximisation	48
3.1.2	Pareto Frontier	48
3.2	Characterisation of Sum Rate Optimal Solution	48
3.2.1	Channel Norm Relationship for Two Users	51
3.2.2	Norm Relationship for Multiple Users	54
3.3	Proposed Norm Balancing Algorithms	55
3.3.1	Complexity of Norm Balancing Algorithms	56
3.4	Characterisation of the Pareto Frontier	57
3.4.1	Pareto Frontier for Two User System	57
3.4.2	Pareto Frontier for Multiple User System	59
3.5	Optimal Decoding Algorithms	60
3.5.1	Complexity of the Algorithms to Determine Optimal Decoding Order	62
3.6	Simulation Studies	62
3.6.1	Effectiveness of Decoding Order	64
3.7	Conclusion	66
4	Generalised Pareto Frontier & Relay Systems	68
4.1	The System Model	69
4.1.1	Communication Scenario	70
4.1.2	Problem Formulation for Sum Power Minimisation	70
4.2	Generalized Pareto Frontier for the Proposed System Models	71
4.2.1	Cost and Reward Functions	71
4.2.2	Linear Cost and Reward Functions	74
	For Non-Zero Power Allocation to All Users	77
	Special Case of Sum Power versus Sum Rate	77
4.3	Sum Power Minimisation Algorithms	77
4.3.1	Power Allocation in the Presence of MBS Interference	78
4.4	Polymatroid Intersection Problem	78
4.4.1	Brute Force Subset Comparison	79
4.4.2	Selective Subset Comparison	80
4.4.3	Fixed Subset Comparison	81
4.4.4	Suboptimal Power Allocation for Non-Joint Pareto Frontier	81
4.5	Numerical Simulations	81
4.5.1	The Complexity Count	82

4.6	Conclusion	85
5	Multicell Cooperation & Two Layer Decoding	86
5.1	The System Model	87
5.1.1	Duality between Uplink and Downlink	88
5.1.2	Optimal Beamforming Vectors for a Cooperating Cluster	88
5.1.3	Massive MIMO Systems for LTE-5G	88
5.2	Two Layer Decoding for Uplink	89
5.2.1	Two Layer with SIC	89
5.2.2	Antenna Partitioning for Two-Layer Decoding	91
5.3	Downlink Interference Cancellation	91
5.3.1	Centralised Interference Cancellation	91
5.3.2	Distributed Interference Cancellation	92
5.4	Simulation Studies	94
5.5	Conclusion	97
6	Extension of Duality to Cooperative Multipoint Transmission	98
6.1	System Model and Background	99
6.1.1	Coordinated Beamforming	99
6.1.2	Partial Data Cooperation	100
6.2	Duality between Uplink and Downlink with CBF	101
6.2.1	Zero-Forcing of Intercell Interference	101
6.2.2	Kernel of the Intercell Interference Matrix	102
6.2.3	CBF Uplink System Model	103
6.2.4	Dual MAC with Degree of Freedom Reduction	104
6.3	Algorithm for Zero-forcing	105
6.3.1	Complexity of the Distributed Zero Forcing Algorithm	106
6.4	Duality between Uplink and Downlink with PDC	106
6.4.1	Downlink System Model	107
6.4.2	Uplink System Model	108
6.4.3	Uplink to Downlink conversion	109
6.4.4	Downlink to Uplink conversion	110
6.5	Performance Optimisation with PDC	111
6.5.1	Suboptimal Power Allocation	112
	Bounds on the Suboptimal Performance	113

CONTENTS

6.6	Rate Region Boundary for Multicell Systems	114
6.6.1	Algorithm to Characterise Pareto Frontier for CBF Systems	114
6.6.2	Pareto Frontier for Multicell Systems Employing PDC	114
6.7	Numerical Simulation Results	116
6.8	Conclusion	118
7	Conclusions and the Way Ahead	119
7.1	Conclusions of the Present Research	119
7.1.1	Application of Proposed Algorithms in LTE-5G	121
7.2	Future Research Directions	121
A	Review of Inner Product Space	124
A.1	Fields & Vector Space	124
A.2	Inner Product Space & Norm	124
A.2.1	Norm & Inner Product with Respect to a Square Matrix	125
A.3	Cauchy Schwarz Inequality	126
B	Set Functions & Polymatroids	127
B.1	Submodular Set Function	127
B.2	Rank Function & Polymatroid	128
C	Duality with Linear Beamforming	129
C.1	MAC to BC Conversion	130
C.2	BC to MAC Conversion	131
C.3	Equivalence of Vector BC & Vector MAC	132
D	Wirtinger Calculus Fundamentals	133
D.1	Derivative from Linear Approximation	133
D.2	Optimisation Using Wirtinger Derivative	134
D.3	Generalisation to Matrix & Vector Arguments	134
E	Constraints on Convex Problems	135
	Bibliography	137

List of Figures

1.1	Growth of Data Rate with Time	2
1.2	Non Cooperating Users Interlaced with a Decode-Forward Relay	6
1.3	Organisation of the Optimisation Problems	9
2.1	Degrees of Freedom of Different Systems	12
2.2	Difference between convex and non-convex rate regions	15
2.3	Multiple Access Channel	17
2.4	Broadcast Channel	18
2.5	Two User Interference Channel	33
2.6	Two User Rate Region without Interference	34
2.7	Two User Relay Channel with Transmit and Receive Filters	37
3.1	Variation of Channel Norm when Sum Rate Passes through a Maximum	52
3.2	Variation of Channel Norm when Sum Rate is Monotonic	53
3.3	Two User Rate Region	57
3.4	Effect of Additional Power Constraint	60
3.5	Variation of Processor Time Requirement with Problem Size	63
3.6	Convergence of Norm Balancing Algorithm	63
3.7	Sum Rate Variation	64
3.8	Two User Pareto Frontiers with Different Decoding Orders	65
3.9	Three User Pareto Frontiers	66
4.1	MBS-FBS system model.	70
4.2	Intersection of Pareto Frontiers for Two user Rate Region	79
4.3	Variation of time requirement with problem size.	82
4.4	Sum rate variations with UE power.	82
4.5	Sum power variations with UE power.	83

LIST OF FIGURES

4.6	Simulation time requirement to test for intersection.	84
4.7	Power Requirement at FBS	85
5.1	Comp System Model	87
5.2	Block Diagram for Two Layer Decoding	90
5.3	Ordering for Two BSs for Intercellular noise cancellation	93
5.4	Comparison of Achievable Sum Rates.	95
5.5	Comparison of Interference Cancellation Techniques.	95
5.6	Comparison of BS Enumeration and Ordering.	96
5.7	Time Complexity of Interference Cancellation Techniques.	96
6.1	System Model for CoMP	99
6.2	BS power vs. single cell throughput	116
6.3	BS power vs. cluster throughput	117
6.4	Throughput variation with number of antenna	118
6.5	Pareto frontier for a two cell system	118
7.1	Broadcast Channel with Relay	122
7.2	Multiple Access Channel with Relay	122
7.3	Multiantenna Parallel Relay System	123

List of Tables

3.1	Rate Vectors under Different Conditions	58
4.1	The landscape for different problem formulations	72
5.1	Comparison of Centralised and Distributed Interference Management	94
6.1	Uplink and Downlink Systems with Different Cooperation Methods	100

List of Abbreviations

3GPP	Third Generation Partnership Project
AF	Amplify-and-Foward
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
BC-MAC	Broadcast and Multiple Access Channels
BS	Base Station
CAGR	Compound Annual Growth Rate
CBF	Coordinated Beamforming
CIC	Centralised Interference Cancellation
CoMP	Coordinated Multipoint Transmission
DIC	Distributed Interference Cancellation
DF	Decode-and-Foward
DL	Downlink
DoF	Deegree of Freedom
DMT	Diversity-Multiplexing Trade off
DPC	Dirty Paper Coding
EB	Exabyte
FBS	Femto Base Station
FD	Full-Duplex
FUE	Femto User Equipment
GPF	Generalised Pareto Frontier
HD	Half-Duplex
ICI	Intercell Interference
JP	Joint Processing
JPF	Joint Pareto Frontier
LSAS	Large Scale Antenna System
LTE-A	Long Term Evolution-Advanced
MAC	Multiple Access Channel
MBS	Macro Base Station
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MUE	Macro User Equipment
PDC	Partial Data Cooperation
QoS	Quality of Service
SDMA	Space Division Multiple Access
SIC	Successive Interference Cancellation

List of Abbreviations

SINR	Signal to Interference plus Noise Ratio
SLNR	Signal to Leakage plus Noise Ratio
SNR	Signal to Noise Ratio
SOBF	Successive Orthogonal Beamforming
SS	Spatial Sensing
SVD	Singular Value Decomposition
UE	User Equipment
UL	Uplink

List of Mathematical Notations

$\ \mathbf{v}\ $	Euclidean norm of a vector
$\ \mathbf{v}\ _1$	$\mathcal{L}_1$ norm of a vector
$\succcurlyeq$	Semidefinite partial order relation or tuple-wise inequality
$\langle \mathbf{u}, \mathbf{v} \rangle$	Inner product of vectors $\mathbf{u}$ and $\mathbf{v}$
$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{A}}$	Generalised inner product with respect to $\mathbf{A}$
$\ \mathbf{v}\ _{\mathbf{A}}$	Generalised norm with respect to $\mathbf{A}$
$v_1, v_2, \dots$	Scalar tuples of vector $\mathbf{v}$
$\mathbf{A}^H$	Hermitian transpose of $\mathbf{A}$
$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$ \mathbf{A} $	Determinant of square matrix $\mathbf{A}$
$\mathbb{R}$	Field of real numbers
$\mathbb{C}$	Field of complex numbers
$\mathbb{R}_+$	Set of non-negative real numbers
$\mathbb{R}_{++}$	Set of positive real numbers
$\mathbb{E}[X]$	Mean of a random variable X
$\mathcal{U}$	Set of all users in a multiuser system
$\mathcal{S}$	A subset of $\mathcal{U}$
$\mathcal{S}'$	The complementary set $\mathcal{U} - \mathcal{S}$
$\mathcal{O}$	Time complexity of algorithms
$!$	Factorial of a natural number

Abstract

As a result of the rapid exhaustion of available electromagnetic spectra and industry trend towards power limited devices, various cooperation schemes have been proposed as ways to achieve better performance in wireless systems. The currently deployed systems use a combination of time and frequency division multiplexing to grant each communication link an exclusive degree of freedom. While this limits the interference among users, the achievable performances are suboptimal under such circumstances. In contrast, when multiple users share the same time-frequency resource block, they suffer from channel noise as well as interference from each other. Mitigating the effect of interference in multiuser cooperative systems is a critical issue and the main focus of this work. Even though the best methods to deal with interference in a most general network is still unknown, two important classes, viz. the Broadcast (BC) and Multiple Access (MAC) channels have been well investigated. This thesis extends some of the known duality results to more general system models of multiuser relays and multicell environments. The existing duality principles state that same quality of service factors (error exponent, throughput or signal to interference and noise ratio) can be achieved via same total power in broadcast and multiple access channels. These duality principles allow the intractable problems in broadcast systems to be projected in multiple access systems for which efficient algorithms can be used to solve them.

The main contributions of this thesis are several optimisation algorithms based on power and rates in various system models. The rate region boundary of a broadcast channel is established first. Along with that, the thesis also proposes some novel power allocation algorithms and decoding order to achieve the efficient boundary (known as Pareto frontier) of a system. It is expected the efficiency of the proposed algorithms in terms of complexity and performance will ensure better service and computational ease for cellular service providers.

Furthermore, the existing duality principle for single-cell system has been a cornerstone of research on cellular information theory and optimisation. This result has prompted different optimisation approaches in literature and made the use of convex techniques possible in this context. The later part of this thesis extends the duality principle to a multicell system using two different cooperation paradigms, viz. Coordinated Beamforming and Partial Data Cooperation. While duality can be established for these two systems, optimal solution of the resource allocation problem proved to be infeasible for Partial data cooperation. However, it is expected that analysis of duality and the proposed suboptimal power allocation in this context will inform further research into basestation cooperation. The results have the potential to be useful for the standardisation of 5G standards based on significant Inter-cell cooperation to enhance user experience.

Keywords-Beamforming, Cooperative Communication, Duality, Pareto Optimality, Rate Region, Relay Channel, Resource Management.

Chapter 1

Introduction

WIRELESS communication involves transmission of information among entities wirelessly. When multiple entities communicate over a wireless medium, they have to be separated by some degree of freedom (spatial, time or frequency) or else their mutual communications will cause interference to each other. This interference, coupled with fading properties of the wireless channel, makes it challenging to recover the desired signal at the receiver end.

Two most important resources in a wireless system are power and bandwidth. Communication hardware is limited in terms of power consumption because of several physical reasons. In power limited systems, the rate of information transfer is also limited by the maximum mutual information between the receiver and transmitter. This limit (called channel capacity) was derived in terms of signal and noise powers present in a channel in [1]. Although the channel capacity (called Shannon's limit) is difficult to achieve in many practical systems, incorporation of efficient signal processing algorithms and innovative coding schemes helps the system designers to approach the capacity limit very close.

While Shannon's limit provided the performance benchmarks, the demand for wireless and wired services grew at an explosive rate around the world. With the birth of the Internet, the fields of computer networking and telecommunication became intricately merged together. The continuous rise in demand puts strain on the most fundamental resources in radio communication-viz. power and frequency band. Figure 1.1 shows the projected Compound Annual Growth Rate (CAGR) for half a decade. The years are plotted along the x-axis and the mobile traffic in Exabyte (EB= 10^{18} Byte) is plotted along the y-axis. The need to adapt innovative methods of resource utilisation is evident from the figure.

However, the demand for wireless data rate does not automatically translate into wider frequency bands over which transmissions can take place. Even though recent research has improved hardware, especially terahertz compatible systems [2] that provide more flexible frequency allocation schemes, the gain is not enough to meet the demand. Traditional systems dependent on a licensed spectrum regime allowed entities to be separated in frequency and time domain, eliminating the need to deal with interference arising out of multiuser communication. But the exhaustion of available spectra necessitated spectrum sharing among multiple users at the same time. This causes severe interference to each other and it is a challenge to recover the intended signal with reasonable accuracy from a host of users corrupted by channel noise. Among other works on interference management, [3] showed that it is possible to achieve sub-

Figure 1.1: Growth of Data Rate with Time

stantial throughput gain by exploiting two major paradigm shifts. The work proposed to exploit the spatial degrees of freedom at the transceiver nodes while sharing time-frequency resource blocks. In addition, it underlined the need of advanced channel acquisition techniques [4] to improve availability and predictability of channel information which results in better optimised transmit and receive strategies¹.

However, the shift towards a multiuser paradigm also necessitated more complicated user scheduling and selection strategies to minimise interference. Different competing approaches involving linear and non-linear encoding/decoding have been proposed in the literature which point out the trade off between performance and complexity which are reviewed in the next chapter. To investigate the effects of multiuser interference and methods to mitigate them in some of the advanced communication frameworks is the principal theme of this work.

1.1 Multi Antenna Communication Systems

To mitigate the effect of interference, and to improve data rate, Multiple Input Multiple Output (MIMO) systems had been proposed in earlier literature. The advantages of using

¹Image Courtesy of Figure 1.1: Visual Networking Index, Cisco Systems Inc. San Jose.

multiple antennas for transmission and reception are the presence of independent signal paths and an additional spatial domain to exploit the degrees of freedom. The available spatial degrees of freedom allow for transmission of independent data streams without interfering with each other [5]. Different varieties of systems models have been analysed in the literature with different assumptions on channel state distribution. Earlier literature have focussed on fundamental information theoretic aspects of an interference channel. The relevant analyses were carried out from a high level of abstraction to derive the mutual information upper bound (see [6–10] and the references therein). During the last decade or so, the attention has shifted more towards different space-time encoding procedures including power allocation and beamforming methods [11–13]. But still, the information theoretic results provide the ultimate limits and benchmarks to compare different systems. Additionally, the information theoretic expressions allow us to optimise the transmission-reception beamforming parameters which closely approximate the throughput achieved via practical coding methods [14].

When different transmitters communicate with different users (possibly with multiple transmit and receive antennas), each receiver will receive desired signal and interference from other users. Depending upon the system model, there may exist cooperation at the transmitter/receiver ends. A cooperation in this context implies that the network nodes are aware of the other nodes' data streams. Hence, the nodes can adapt appropriate strategies to mitigate the effect of interference amongst themselves. The classification of systems based on cooperation is given below.

- When there is cooperation at the transmitter end while the receivers are distributed, the channel is termed as a *Broadcast* channel.
- When there is a cooperation at the receiver end but no cooperation among the transmitters, it is a *Multiple Access* channel (MAC).
- When there is no cooperation at either end (which is the most difficult and unfavourable scenario), then it is a general interference channel.
- Where there is cooperation at both ends, then basically the system can be taken as a single user system (without interference) where all the users' data can be merged into a single data stream.

It is well known that having cooperation improves system performance and simplifies appropriate space-time code design [15]. In that regard, a broadcast or multiple access channel is more desirable than a general interference channel if it is possible to establish cooperation at the transmitter or receiver end. These are the kinds of channel we have in cellular communication systems and their analyses provide important insights into the interference management issue.

1.2 Multi-Hop Communication Systems Aided by Relays

Most traditional theoretical models dealt with a single hop communication. In such systems, the data transmitted by a source is received by the destination, perturbed by a channel and corrupted by noise. In contrast, recent research have focussed on more flexible systems, where the transmitter(s) and receiver(s) are coordinated by one or more relays. The relay node receives the source data and retransmits it to the receiver. This can introduce a form of cooperation and also possibly cascading channels using different network level protocols or modulation schemes. Depending on the strategy employed by the relay to retransmit received signal, two major categories have been proposed in literature [16].

- *Amplify and Forward Relay systems*: These relays have also been named as *non-regenerative* relays since they carry out linear processing of the received signal to minimise the effects of noise and interference. Although these relays suffer from noise propagation effect, they have been popular in research because of their operational simplicity. These relay systems provide a reasonable trade off between cost and performance.
- *Decode and Forward Relay Systems*: In this case, the relay fully decodes the received information before retransmitting. Hence, this category of relays have been termed as *regenerative* relays. They fully eliminate the effects of channel noise and fading before retransmission, hence preventing noise or error propagation. However, there is a trade off in terms of complexity and implementation.

1.3 Multicell Communication Systems

To further meet the traffic demand from bandwidth hungry applications, multicell cooperation has been proposed as a way forward by recent studies in 3rd generation partnership projects [17]. The objective is to facilitate cooperation among multiple reception and transmission points (i.e. cells). The paradigm has been termed as Cooperative Multipoint Transmission (CoMP), where multiple cells coordinate in such a manner that they reduce the interference incurred to each other. Depending upon the cooperation technique, part of the interference can even be exploited as useful signal. Many of the research in multicell cooperation has been driven by wider adaptation of small cell networks. In such systems, a larger number of adjacent base stations cover a unified coverage area, which makes it possible to exploit the cooperation [18–20]. Three major cooperation techniques have been proposed in prior research in this context.

- *Coordinated Beamforming (CBF)*: In this case, every BS transmits information to a selected number of users, while minimising interference to other users [21]. Since each user's downlink information is available only at one BS, the goal of BS cooperation is interference minimisation, rather than throughput maximisation.
- *Joint Processing (JP)*: All the BSs have access to downlink data streams for all users. Hence, the BSs can coordinate to maximise the throughput via a greater degree of cooperation compared to CBF [22]. This cooperation technique can improve the downlink throughput while it also increases the backhaul traffic. In the case of ideal cooperation, the transmit side is modelled by a distributed super base station which coordinate together to serve the users. The capacity of such a channel has been analysed in [23].
- *Spatial Sensing (SS)*: In this case, the users are served by a single BS. However, unlike CBF, there is more flexibility in assigning BSs to users via adaptive techniques depending on the channel strength [22]. Leveraging the flexibility, systems belonging to this category are able to exploit the spatial domain more efficiently using multiple antennas at the transceivers [24, 25].

In spite of extensive prior research into the area, complete characterisation of the rate region boundary of CoMP systems have been absent in the literature so far. Hence, the present research extends the relevant optimisation and interference management techniques to a multicell cooperative environment via a novel cooperation paradigm called Partial Data Cooperation (PDC).

1.4 Fifth Generation LTE Standards

The Fifth Generation (5G) cellular system is the upcoming standard resulting from the Long Term Evolution (LTE) of systems globally, and is slated to incorporate extensive cooperation and advanced multiplexing techniques to manage interference. According to [26] the standard is expected to mature by 2020, the effort being led by Group Special Mobile Association (GSMA) and its partners. Some of the salient features of the proposed standard are

- Throughputs of the order of 1-10 Gbps, a significant increase from earlier LTE standards
- A massive number of connected devices sharing resources to realise the vision of Internet of Things.
- Reduction of transmit power consumption leading to higher battery lives.

Based on these requirements, industrial and academic organisations came up with different visions of 5G standard, such as [27] and [28] for examples. In particular, Ericsson envisions fully backward compatible network architecture, ensuring a smooth transition from earlier standards and continued use of existing infrastructural resources [28].

Given the user equipment connectivity centric nature of the envisioned standard, it is expected that inter-device communication and cooperative resource management will play a significant role. Even though the system models relevant in 5G are yet to be specified, it is expected different cooperative and distributive algorithms proposed in this thesis will inform practical research in this direction.

1.5 Effects of Channel State Information

For multiple access and broadcast channels, it is assumed that accurate channel state information is available at the relevant transceiver nodes. This is a realistic assumption because of the recent progress with advanced channel acquisition and empirical measurement techniques. Among the recent papers, [29], for example, proposed a phase shift pilot training for channel acquisition in large scale antenna systems. [30] also explored the radio measurements of Over-The-Air (OTA) massive MIMO enabled systems to measure the channel strength and other fading effects.

However, under certain circumstances, the accuracy of the channel state information might be compromised. The algorithms proposed later in this thesis are dependent on channel values which are continuous in nature rather than discrete. Hence, a small error in measurement will result in a small error in the output of the algorithms, and in that sense, the algorithms are theoretically robust. As an example, robustness of resource allocations with imperfect channel state information was discussed in [31]. Extension of the algorithms proposed later for imperfect channel information and more rigorous analysis of robustness can be a future research avenue.

1.6 Motivation of the research

The present literature on multiple and broadcast channels have usually focussed on optimising an objective function with one or multiple linear power constraints. However, there have been two major unaddressed issues in the relevant literature.

- Prior works have treated the relevant optimisation problem as a special case of convex programming. Thus some special structure of the objective sum capacity function has not been exploited. The proposed norm balancing approach introduces certain distributed

Figure 1.2: Non Cooperating Users Interlaced with a Decode-Forward Relay

feature and improves the algorithmic time complexity for practical systems.

- Optimisation of a specific objective function does not always characterise the boundary of the rate region in a multiuser system. When the rate region is given by $\mathcal{C} \subset \mathbb{R}_+^K$, the boundary can be defined as a Pareto frontier. One of the motivations of this work has been a complete characterisation of the rate region boundary in a more generalised form than the objective functions. For non-linear beamforming, user ordering in the encoding/decoding procedure plays an important role to determine the achievable rate region [32, 33] and usually enhances the region over linear techniques. The optimal ordering depending upon transmitter parameter is an important issue which received relatively less attention so far.

The solutions of these problem give important intuitions into the optimisation problems of heterogeneous cellular systems and multicell cooperation. Extension of the concept of Pareto frontier to relay systems has been formally defined under the term Joint Pareto Frontier (JPF). In addition, the proposed norm balancing solution can also be applied to multicell systems in order to achieve higher throughput. Hence, establishing the connection between relevant optimisation problems in different systems with a goal to improve throughputs has been a principal motivation of this work.

The research was motivated by the system model shown in Figure 1.2 where a number of non-cooperating transceiver pairs communicate data symbol over the same time-frequency resource block. The objective is to identify the JPF in such a system under power constrained transmitters. The characterisation and resource allocations of individual BC and MAC systems branch out from this general problem, which are addressed later.

1.7 Contributions of this work

This work, along with a detailed review of the state of relevant literature, solves some relevant power and rate optimisation problems of relay and multicell channels. The problems of complete characterisations of rate region boundary for multiple access and broadcast channels branch out from the problem in relay channel.

An algorithm for sum rate maximisation in a multiple access channel has been developed exploiting the structure of the problem. In a usual multiple access channel, the transmitters are spread across a geographically wide coverage area and all of them transmit to a single node. The system model is applicable to a cellular system serving users with a base station or in case

of a wireless sensor network, where the sensors in a cluster transmit towards a data gathering node for further processing. In any case, it is assumed that there is no cooperation among the transmitters. Hence, each user can be supposed to have his own battery with individual transmit power constraint. Under the sum power constraint, the optimal rate region boundary can be shown to be where the sum rate is maximum and the only variable to control the user rates is the decoding orders. Different decoding orders of users give different data rates even though the sum rate remains the same. However, if the multiple access channel is being considered as a dual of a broadcast network, it is reasonable to assume that the transmitters draw power from a common source and hence they are constrained by a sum power. In that case it is shown that a specific power allocation maximises the sum rate. The sum rate maximisation is a convex optimisation problem whose solution can be found by standard optimisation methods like interior point method [34]. However, solving the problem by convex methods obscures some interesting properties of the system about how the different transmitters can cooperate in terms of power consumption to tolerate interference from each other. These properties can be exploited to formulate an efficient alternative algorithm which gives better insight into the problem at hand and also solves the problem with less time requirement for reasonable problem size. Further, the algorithm is self-contained in the sense that it provides a metric to test the suboptimality of any tentative feasible point and hence can be used to formulate a stopping criterion.

After formulating a novel algorithm for sum power maximisation, properties of the same problem have been used to characterise the rate region boundary of a dual multiple access channel in terms of the decoding order of users. It is shown that sum rate maximality is a sufficient condition for achieving the boundary, but not a necessary condition. In other words, every achievable point for which the sum rate of users is maximum, is obviously on the boundary, but there are points for which the sum rate is suboptimal. Those points have been partially characterised using some properties of the sum rate maximisation problem. It is shown that depending on the power allocation, different decoding orders give rate points on the boundary.

Further, the solution for single cell uplink-downlink systems have been extended to two different paradigms which have been proposed in recent literature as solutions to exponentially rising demand for wireless traffic. They are heterogeneous cellular systems modelled by a decode-forward relay and multicell systems with cooperating base stations to serve users in a unified coverage area.

For a heterogeneous cellular system modelled by a decode-forward relay, the novel concept of joint Pareto frontier has been introduced. To achieve the joint Pareto frontier, it is necessary to achieve an individual rate equality among the user pairs. Two user selection and testing algorithms have been proposed to achieve the joint Pareto frontier.

Further, combining the emerging architectures of massive MIMO systems with cooperative multipoint techniques can improve the data rate over existing systems. In this context, the duality principle is established for multicell systems to exploit a novel cooperation technique called Partial Data Cooperation (PDC).

In short, the major contributions of the present thesis can be summarised as below.

- *Norm Balancing Algorithm and Pareto Frontier for Single Cell Systems*: An efficient algorithm to solve the sum rate maximisation problem in MAC using a norm balancing approach has been proposed. Using the solution of sum rate maximisation, two algorithms

have been proposed to characterise the Pareto frontier of a single cell downlink channel with sum power constraint.

- *Joint Pareto Frontier for Heterogeneous Systems:* A novel analysis based on cost and reward functions to represent optimisation problems have been proposed. This aids in extension of the resource allocation problem to decode-forward relay systems. In the same context, the concept of joint Pareto frontier has been proposed together with the necessary algorithms to achieve it.
- *Two Layer Decoding with Successive Interference Cancellation:* The proposed concepts of Pareto Frontiers and resource allocation have been discussed for Multicell cooperative systems. To that end, the earlier two-layer decoding techniques applicable for massive MIMO have been enhanced via two new interference cancellation techniques, viz. Centralised Interference Cancellation (CIC) and Distributed Interference Cancellation (DIC). The complexity-performance trade off between these two algorithms is illustrated via simulation results.
- *Resource Allocation and Duality in Multicell Cooperative Network:* The duality principle has been extended to a multicell system. It is proved that BSs employing massive MIMO in a cooperating cluster can achieve zero forcing of Inter-Cell Interference (ICI) via a distributed beamforming method in Coordinated Beamforming (CBF). The co-ordination among BSs has been further enhanced to propose a novel paradigm called Partial Data Cooperation (PDC) along with the proposed duality transformation for such a system. Based on the proposed transformation, the sum rate maximisation problem is formulated. However, owing to the intractable structure of the problem, a suboptimal but efficient algorithm has been proposed to allocate the power. The results are compared with prior art.

1.8 Organisation of the thesis

The thesis is organised as follows.

Chapter 2 will give a review of literature on interference management and achievable rate regions under power constraints. It has been pointed out via relevant references that interference management is a relevant research area under the broad purview of wireless communications. After reviewing the state of the art on interference management, the effect of interference on the achievable rate regions is described.

The important contributions of the thesis following the literature review are shown in Figure 1.3 to illustrate their interdependencies.

An algorithm for sum rate maximisation in dual multiple access channel will be proposed in Chapter 3. Even though the problem has already been solved as a standard convex optimisation problem, the proposed algorithm gives a novel perspective on the problem and solves the problem more efficiently taking advantage of its special structure. Using some of the results from the optimisation problem, Chapter 3 will also characterise the entire rate region boundary (known as Pareto frontier) and propose a decoding algorithm to achieve the boundary. It is explained even though the decoding order affects individual rates of the users, it does not affect the sum rate of all users under individual power constraints. However, for a sum power constraint, certain decoding orders are proved to be suboptimal, and the condition for optimality of a decoding order is derived.

Figure 1.3: Organisation of the Optimisation Problems

Chapter 4 focusses on norm balancing techniques based on cost-reward formulations on multiple access channels and the results are extended to a Decode-Forward relay system. The proposed norm balancing techniques show that any power allocation can be Pareto efficient in a multiple access channel provided certain conditions on the performance metric and decoding order are satisfied. Based on the analysis, a novel method for efficient power allocation has been proposed to meet a target sum rate. In addition, two algorithms to test for the existence of the joint Pareto frontier have been proposed and compared. As Figure 1.3 shows, the results for Joint Pareto Frontier (JPF) in Chapter 4 follow from the earlier chapters.

Finally, the original problems formulations of MAC and BC sum rate optimisations are used to solve the resource allocation problems in Cooperative Multipoint transmission under three different paradigms, viz. two layer decoding, Coordinated Beamforming (CBF) and Partial Data Cooperation (PDC).

Chapter 5 proposes the two layer decoding algorithms, viz. Centralised Interference Cancellation (CIC) and Distributed Interference Cancellation (DIC) to mitigate interference in multicell systems. The algorithms are shown to improve the performance over earlier two layer decoding techniques applicable to base stations having large scale antenna systems such as massive MIMO.

Chapter 6 extends the well known uplink downlink duality principles to a multicell environment. It also proposes a novel cooperation technique called PDC, which offers significant performance improvement over cooperative beamforming techniques, at the cost of increasing backhaul traffic.

Chapter 7 concludes the thesis with a summary of the contributions and some pointers for future work where the results from the present thesis can be extended.

The Appendices A-E complement the discussion reviewing some mathematical details and interesting results which are used in the thesis.

Chapter 2

Optimal Rate Regions and Interference Management in Communication Systems: A Review

INTERFERENCE arises out of multiple user pairs communicating through the same physical medium and at the same time and frequency domain. Interference is a necessary evil in the sense that it disturbs the communication of each user while allowing multiple users to maximally exploit the important communication resources. [3] was one of the pioneering works to point out the evolution towards a multiuser paradigm and along with that, interference mitigation became a critical issue in wireless systems. Depending on the system model and some assumptions on the channel information, there are different competing strategies to deal with this interference. In the prevalent literature, qualitative distinction has been made among high interference regime, low interference regime and moderate interference regime. Different strategies have been used to achieve optimal performance under these paradigms. Under this research, the focus will be on how interference affects the transmission rates of communication systems and characterisation of the best achievable schemes.

The organisation of this chapter is as follows.

Section 2.1 discusses some mathematical preliminaries. Section 2.1.1 will discuss the concept of degree of freedom in wireless communication. Degree of freedom basically characterises how higher power utilisation leads to faster communication systems. While it is unambiguously defined for single user scenarios, multiple definitions are possible for multiple users and the quantity may not even exist. Nevertheless, it plays an important role in interference management. Additionally, convexity of rate region is an important concept in the context of rate region boundaries. Convexity determines whether better performance can be achieved by time sharing among achievable points. Section 2.1.2 introduces the concept of convexity of sets.

Section 2.2 will discuss the idea of rate region for single cell system models. In particular, Section 2.2.1 and Section 2.2.2 will discuss the rate regions for multiple access and broadcast channels, respectively. Duality principles are some of the important theoretical results concerning different wireless communication models. These principles assert that under various resource constraints, broadcast and multiple access systems are capable of giving similar performance.

This is introduced in Section 2.2.3 for non-linear encoding/decoding techniques. Particularly, Section 2.2.3 will discuss the duality between multiple access channels and broadcast channels in terms of the achievable rate regions and power requirements. Then another interesting class of duality will be introduced in terms of worst case noise correlation at the receivers. Finally, Section 2.2.4 will extend the concepts to a more general interference channel. Section 2.2.5 will discuss some of the prevalent optimisation strategies in current literature under various constraints.

Section 2.3 will discuss several approaches to mitigate the interference in a multiuser communication systems where different users share the same time-frequency resource blocks. Sections 2.3.1 to 2.3.3 will discuss the known results with three different approaches of interference mitigation— viz. interference cancellation, interference ignorant and interference alignment. Even though the interference cancellation schemes are computationally complex, they give better performance in terms of data rate, compared to interference ignorant schemes. Interference alignment is an emerging topic, which focusses on characterising interference via different subspaces and nullifying the effect orthogonalisation.

Section 2.4 will further discuss the issues of interference management in multi-hop systems, aided by relay nodes. The relays will be classified into several categories. In particular, Section 2.4.1 will discuss the optimisation of transceiver beamforming matrices in AF relays. However, in spite of the simpler implementation, the AF relay systems suffer from lower throughput compared to DF relays. The DF relay systems will be discussed in Section 2.4.2.

Section 2.6 introduces the idea of multicell cooperation, where a number of adjacent cells operate on the time-frequency block. A number of cooperation paradigms among the base stations have already been proposed in the literature, which can be classified into three broad categories. Section 2.6.1 covers the resource allocations in coordinated beamforming. Then Section 2.6.2 discusses joint processing among all the transmitters, which offers higher throughput at the cost of higher backhaul traffic. Then Section 2.6.3 discusses the idea of spatial sensing which has been more popular in cognitive spectrum sharing systems.

Section 2.7 points out the gap in the present knowledge and scope of the present work. Finally, Section 2.8 concludes the chapter.

2.1 Mathematical Preliminaries

2.1.1 Degrees of Freedom of a Communication System

It is intuitively clear that the data rate of a communication system grows monotonically with available power. This is quantitatively measured by degree of freedom (abbreviated as DoF henceforth).

For a single user system, let the SNR be γ . If the capacity is represented as [32]

$$C(\gamma) = n \log \gamma + o(\log \gamma) \quad (2.1)$$

for some $n > 0$, the capacity grows with $\log \gamma$ with slope n . Hence, n is the degree of freedom in this case. Here $o(\log \gamma)$ represents any function $f(\log \gamma)$ that grows slower than $\log \gamma$, i.e.

$$\lim_{\gamma \rightarrow \infty} \frac{f(\log \gamma)}{\log \gamma} = 0$$

Figure 2.1: Degrees of Freedom of Different Systems

Hence, from (2.1), the degree of freedom for a single user system can be formally defined as

$$\text{DoF} = \lim_{\gamma \rightarrow \infty} \frac{C(\gamma)}{\log \gamma} \quad (2.2)$$

As an example, Figure 2.1 shows the variation of capacity with power for various system configurations. The advantages of MIMO systems is clear from the slope of capacity curve [5].

The degree of freedom as defined in (2.2) may not exist in all cases. Still it has been used as a metric to compare many different systems under a common benchmark. In fact, for MIMO systems exploiting spatial diversity [35], the same quantity can also represent the multiplexing gain by virtue of multiple antenna and can have a range of values even for a given system configuration. In fact, multiple antennas systems allow the use of diversity schemes while reducing the degree of freedom. This variation represents a fundamental trade off between transmission speed and accuracy captured in the concept of diversity multiplexing trade off introduced in [36].

As a simple example, for a faded channel corrupted by Additive White Gaussian Noise (AWGN), the input output relationship is

$$y = hx + \eta \quad (2.3)$$

where y is the output, h is the scalar channel, x is the input and η is the AWGN. The capacity is given by,

$$C(\gamma) = \log(1 + h^2 \gamma) \quad (2.4)$$

where γ is the SNR as before. It can be easily shown that

$$\text{DoF} = \lim_{\gamma \rightarrow \infty} \frac{C(\gamma)}{\log \gamma} = 1 \quad (2.5)$$

For a set of K parallel non-interfering channels where the same variables are indexed by k , the

input output relationship is

$$y_k = h_k x_k + \eta_k \quad \text{where} \quad 1 \leq k \leq K \quad (2.6)$$

When the sum power is constrained by P , the total capacity is

$$C(P) = K \log P + o(\log(P)) \quad (2.7)$$

From (2.7), the DoF = K for the sum capacity.

DoF can be considered as one instance of a transmission over a channel respecting the power constraint. Depending on the bandwidth and time available, there can be multiple such instances (the relationship is given by Nyquist theorem [37]) and when there are multiple users present, the instances can be shared in time or frequency domains, reserving DoF slots exclusively for various users.

Degree of Freedom for Multiuser Networks

Complete characterisation and definition of DoF for multiple user systems is still an open issue and an active research topic (see [38] for a review). For a system of K users each with a power constraint P , the DoF should be defined as a region instead of a pure number. If the set of achievable rate tuples $[r_1, r_2, \dots, r_K]$ is denoted as $\mathcal{C}(P) \subset \mathbb{R}_+^K$, a widely accepted definition used in [39, 40] can be formulated as

$$\mathcal{D} \equiv \{\mathbf{d} \in \mathbb{R}_+^K : \quad \text{where} \quad [w_1, w_2, \dots, w_K] \in \mathbb{R}_+^K, \lim_{P \rightarrow \infty} \sup_{\mathbf{r} \in \mathcal{C}(P)} \sum_{k=1}^K \frac{w_k r_k}{\log P} \geq \sum_{k=1}^K w_k d_k\} \quad (2.8)$$

Since the capacity region may not be closed in general, supremum, denoted by $\sup(\cdot)$ in (2.8), is used instead of a maximum. However, the capacity region is not known for sure in many cases and even when the capacity region is known, characterising the complete DoF region of (2.8) is a difficult task. Hence, in many applications, a sum total is used as a single useful metric to judge a system.

$$d \equiv \max_{\mathbf{d} \in \mathcal{D}} \sum_{k=1}^K d_k$$

where $\mathcal{D}$ is defined by (2.8).

The DoF depends on many assumptions from channel state information to degree of cooperation among transmitters and receivers, giving rise to configurations like X channels [41] and genie aided channel [12].

2.1.2 Convex Sets and Regions

Let $\mathcal{V}$ be a vector space defined over $\mathbb{R}$ and $\mathcal{S} \subseteq \mathcal{V}$. According to literature on mathematical structure and topology [42], the following attributes are defined for a set.

Definition 2.1 $\mathcal{S}$ is a convex set if $\mathbf{u}, \mathbf{v} \in \mathcal{S}$, $\alpha \in [0, 1] \Rightarrow \alpha \mathbf{u} + (1 - \alpha) \mathbf{v} \in \mathcal{S}$.

Definition 2.2 $\mathcal{S}$ is an affine set if $\mathbf{u}, \mathbf{v} \in \mathcal{S}$, $\alpha \in \mathbb{R} \Rightarrow \alpha \mathbf{u} + (1 - \alpha) \mathbf{v} \in \mathcal{S}$.

Although the definitions sound similar at first look, the difference lies in the range of α . It

should be obvious from the definitions that convex sets are a special case of affine sets. As a way of interpretation, $\alpha \in [0, 1]$ implies the segment connecting points $\mathbf{u}$ and $\mathbf{v}$, $\alpha \in \mathbb{R}$ represents the straight line passing through $\mathbf{u}$ and $\mathbf{v}$. So for convex sets, the segment connecting any two points will lie completely in the set and for affine sets the entire straight line will lie within the set.

It can be proved that an affine set $\mathcal{S} \subset \mathbb{R}^K$ can be represented as

$$\mathcal{S} = \{\mathbf{x} | \mathbf{a}^T \mathbf{x} = b\}$$

for some $\mathbf{a} \in \mathbb{R}^K$ and $b \in \mathbb{R}$. Since basically this represents a hyperplane, affine sets are often called affine planes.

Several properties of affine and convex sets have been discussed in [43] and they will be used in subsequent discussions.

Duality and Convex Optimisation

Following the notion of convex set, the definition of a convex function will be given in this section.

Definition 2.3 *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex iff the epigraph of the function is convex.*

According to the convention of [43], typical description of a constrained optimisation problem is as follows.

$$\min_{\mathbf{x}} f_0(\mathbf{x}) \tag{2.9}$$

subject to

$$f_i(\mathbf{x}) \leq 0 \quad \text{where } 1 \leq i \leq m$$

and

$$h_i(\mathbf{x}) = 0 \quad \text{where } 1 \leq i \leq p$$

Here $f_i(\mathbf{x})$ are assumed to be convex functions and $h_i(\mathbf{x})$ are affine functions. According to the duality principle [44], the problem can be solved exploiting certain properties of the Lagrangian given below.

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \equiv f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^p \mu_i h_i(\mathbf{x}) \tag{2.10}$$

[44] showed that, if the solution of (2.9) is given by $\mathbf{x}^*$, then $\mathbf{x}^*$ satisfies the following conditions.

- $\mathbf{x}^*$ is a stationary point of $L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.
- $f_i(\mathbf{x}^*) \leq 0$ where $1 \leq i \leq m$ and $h_i(\mathbf{x}^*) = 0$ where $1 \leq i \leq p$.
- $\boldsymbol{\lambda} \succeq \mathbf{0}$
- $\lambda_i f_i(\mathbf{x}^*) = 0$ where $1 \leq i \leq m$.

Convex Combination as Time Sharing

If the system operates at an achievable rate vector $\mathbf{r}_1$ for α fraction of time and another different achievable vector $\mathbf{r}_2$ for $(1 - \alpha)$ fraction of time, then it is called a time sharing [45]

Figure 2.2: Difference between convex and non-convex rate regions

between $\mathbf{r}_1$ and $\mathbf{r}_2$. The effective rate is given by

$$\mathbf{r} = \alpha \mathbf{r}_1 + (1 - \alpha) \mathbf{r}_2 \quad \text{for } \alpha \in [0, 1]$$

From definition, $\mathbf{r}$ is a convex combination of $\mathbf{r}_1$ and $\mathbf{r}_2$. Hence, although the system throughput is not $\mathbf{r}$ at all instants, the average throughput over time is measured at $\mathbf{r}$. The time sharing can be extended among any number of achievable points.

The outer boundary (formally called Pareto frontiers) in two different two user systems are shown (with continuous lines) in Figures 2.2(a) and 2.2(b). The important difference is that area under the continuous curve in Figure 2.2(b) is a convex region while in Figure 2.2(a), it is not. In Figure 2.2(a), if we do time sharing between two end points of the continuous curve (where each user is enjoying single user capacity), it is possible to operate at points on the broken line segment. The points on the broken line segment lies outside the boundary of the previous rate region. Hence, by time sharing, it is possible to do better than the previous rate region. However, for the rate region of Figure 2.2(b), time sharing does not move the operating point outside the boundary because of the definition of convex sets. Notably the system of Figure 2.2(a) can not operate at points between the broken line (included) and the continuous curve. Those points can be achieved only via time sharing among other points and hence the region can be termed as *Time Sharing* region. Time sharing gives very different results for broadcast systems [46, 47] and multiple access channels [48] with respect to outer boundary of rate region. At this point, the results from the Figures 2.2(a) and 2.2(b) can be summarised as

- For convex rate regions, time sharing does not shift the operating point to outside the region.
- When the rate region is not convex, the full achievable rate region is described by convex hull of the region where the system can operate. Pareto frontier of the convex hull will be different from the previous region and there will be a time sharing region.

Achievable Rate from Signal to Noise Ratio

For a simple single user system, if the SNR is γ , the achievable maximum rate or capacity is given by the well known Shannon's theorem

$$r = \log(1 + \gamma) \tag{2.11}$$

Basically the results in this thesis follow from (2.11) directly or its variance in the presence of an interference being treated as noise. However, none of the known channel coding methods achieve the rate with finite block size, although low density parity check codes approach it very closely. Still, the maximum rate of r in (2.11) has been used as a benchmark to solve the problems of resource allocation from a high level of abstractions, without going to the details of modulation and coding scheme. In short, (2.11) gives the supremum performance with a given Signal to Noise Ratio (SNR).

However, even though the rate is not achievable, another justification for using the quantity $\log(1 + \gamma)$ can be described in terms of the SNR gap [14]. For a practical quadrature amplitude shifting scheme offering rate of r and $\text{SNR} = \gamma$, let

$$r = \log\left(1 + \frac{\gamma}{\Gamma_m}\right) \quad (2.12)$$

where Γ_m is a parameter that accounts for the modulation technique and desired error performance. Γ_m is roughly independent of the SNR. For a high SNR, (2.12) be approximated as [14],

$$r \approx \log(1 + \gamma) - \log \Gamma_m \quad (2.13)$$

Hence, a virtual rate can be defined from (2.13) as

$$\tilde{r} = r + \log \Gamma_m = \log(1 + \gamma) \quad (2.14)$$

which is the actual desired rate added with a fixed constant (taking into account the desired error performance).

From (2.14), it is possible to say that if the system achieves a rate $\tilde{r}$ based on information theoretic results, then the actual rate is given by r . If the desired rate target is $\mathbf{r} = [r_1, r_2, \dots, r_K]$ in a multiuser system, then information theoretic target rate is

$$\tilde{\mathbf{r}} = [\tilde{r}_1 + \log \Gamma_m, \tilde{r}_2 + \log \Gamma_m, \dots, \tilde{r}_K + \log \Gamma_m]$$

Based on this assertion, the results in the present work will focus on achievable rate tuples and it will be assumed that the desired error performance is implicitly accounted for in the rate.

2.2 Review of Rate Regions for Single Cell Systems

Rate regions of broadcast [49–52] and multiple access channels [48, 53, 54] have been analysed in literature. This work will be chiefly concerned with the Pareto frontiers of rate regions and the goal will be to characterise those regions for different system models.

2.2.1 Multiple Access Channel

A multiple access system of K users transmitting to a base station of M antennas is shown in Figure 2.3. The channel vector from user k to base station is given by column vectors $\mathbf{h}_k \in \mathbb{C}^M$. In an usual multiple access system, each transmitter is supposed to have its own power constraint, since a collective power constraint on non-cooperating transmitters across a coverage area is not possible to implement.

For a multiple access channel of K users (each having a single antenna) and the base station

Figure 2.3: Multiple Access Channel

with M antennas, the received signal is given by

$$\mathbf{y} = \sum_{i=1}^K \mathbf{h}_i x_i + \boldsymbol{\eta} \quad (2.15)$$

where x_i is the symbol transmitted by user i and $\boldsymbol{\eta}$ is the zero mean Gaussian additive noise vector with $\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^H] = \mathbf{I}_M$. Power constraints of users are captured by restrictions on the transmitted symbols, where $\mathbb{E}[|x_i|^2] \leq p_i$ is the power available for user i . Let $\mathcal{U}$ be the set of all users. The achievable rate region is given by [53]

$$\mathcal{C}_{\text{MAC}}(\mathbf{p}) = \{\mathbf{r} \in \mathbb{R}_+^K | 0 \leq \sum_{i \in \mathcal{S}} r_i \leq \log |\mathbf{I} + \sum_{i \in \mathcal{S}} p_i \mathbf{h}_i \mathbf{h}_i^H| \text{ where } \mathcal{S} \subseteq \mathcal{U}\} \subset \mathbb{R}_+^K \quad (2.16)$$

For a K user system, the above set definition gives $2^K - 1$ sum rate ceilings ($\mathcal{S} = \emptyset$ being considered a trivial case) and there exists K other obvious constraints given by $\mathbf{r} \succeq \mathbf{0}$. Since each constraint represents a half space, $\mathcal{C}_{\text{MAC}}$ is an intersection of $2^K + K - 1$ half spaces and a convex set. However, for the vertex representation of the polytope $\mathcal{C}_{\text{MAC}}(\mathbf{p})$, it can be seen that there are interesting properties of the vertices situated in the $\mathbb{R}_{++}^K$ cone. Application of this class of polytopes in multiple access interference channel was shown in [48]. See Appendix B for a review of the related concepts in set functions and computational geometry.

For ease of notation, define the set function

$$f(\mathcal{S}) \equiv \log |\mathbf{I} + \sum_{i \in \mathcal{S}} p_i \mathbf{h}_i \mathbf{h}_i^H| \quad (2.17)$$

It is easy to check that $f(\mathcal{S})$ satisfies the condition of a submodular function as in (B.1) and it is also normalised and monotonic. So $f(\mathcal{S})$ is a rank function and $\mathcal{C}_{\text{MAC}}(\mathbf{p})$ as defined in (2.16) is a polymatroid. The optimisation problems on the polymatroid can be solved by characterising its outer boundary.

Figure 2.4: Broadcast Channel

2.2.2 Broadcast Channel

In an information theoretic broadcast system, a transmitter sends different information symbols to a number of receivers. The receivers cannot cooperate amongst themselves, although the transmitter can carry out joint beamforming for the symbols. The system model is shown in Figure 2.4.

The base station (relay) has M antennas transmitting to K distributed users (each with a single antenna). The $1 \times M$ channel vector from relay to receiver i is denoted by $\mathbf{g}_i^H$. If the transmitter transmits $\mathbf{x} \in \mathbb{C}^M$, signal received by user i

$$y_i = \mathbf{g}_i^H \mathbf{x} + \eta \in \mathbb{C} \quad \text{where } i \in \{1, 2, \dots, K\} \quad (2.18)$$

where η is a zero mean circularly symmetric, unit variance Gaussian additive noise.

If we assume the data symbol intended for user i is $u_i \in \mathbb{C}$, it is necessary to multiply it with an $M \times 1$ steering vector to transmit it from M different antennas. Total power constraint on the base station is captured by the restriction

$$\mathbb{E} [\mathbf{x}^H \mathbf{x}] = \text{trace} [\mathbf{\Sigma}] = P$$

where $\text{trace}(\cdot)$ is the operation of summing diagonal elements and $\mathbf{\Sigma}$ is the transmit covariance matrix.

Since the transmitter is aware of the channel vectors and the symbol for each user, it has a non-causal knowledge of interference among the users, interference caused by user i to user j for example. This non causal knowledge can be used to reduce or eliminate the interference. [55] showed that if the interference is known to the transmitter, then it is possible to encode the signal in a way so as to completely eliminate the interference. The technique is popularly known as *Dirty Paper Coding* (DPC) scheme. The broadcast channel rate region with DPC has been already analysed [49]. [23] showed that DPC, although computationally complex, gives the best performance and better than zero forcing or interference ignorant schemes. In fact, [56] extended the concept to a more general class of non-ergodic channels. [54] proves that the DPC

scheme achieves the cooperative sum rate upper bound proposed by Sato [57].

Crosstalk Matrix

When there is possible interference among different users, a crosstalk matrix Φ is a succinct way of representing these relationships [58]. For the K users scenario, $\Phi \in \mathbb{B}^{K \times K}$ where $\mathbb{B} \equiv \{0, 1\}$ is the Boolean set. $\Phi_{ij} = 1$, if user i causes interference to user j and $\Phi_{ij} = 0$, otherwise. Obviously the diagonal elements $\Phi_{ii} = 0$ where $j = i$.

In a system with linear steering vector without DPC, every user will cause interference to every other user. So $\Phi_{ij} = 1$, where $i \neq j$. However, after choosing a codeword for user 2, as an example, when the transmitter forms the codeword for user 1, it can do so with full knowledge of interference from user 2 to 1 and hence precancel [59] the interference before transmission. In terms of the crosstalk matrix, a two user system can be written as follows.

$$\Phi = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Once the ordering has been done, the reverse (precanceling interference of user 1 caused to user 2) is not possible. In general, if an ordering of $K, K-1, K-2, \dots, 2, 1$ is used, the cross talk matrix will be an upper triangular matrix with

- All the elements above the major diagonal=1.
- Elements on or below the major diagonal=0.

It is obvious that fewer number of 1s on the crosstalk matrix is better so far as performance goes. DPC eliminates the ones below the main diagonal. Additionally, the numbering of users (and hence ordering) is arbitrary and the freedom can be exploited to get optimal performance in some sense. In conclusion, while *designing* the broadcast system with given total power constraint, channel vectors and pre-cancellation of known interference, we have three different freedoms and any optimisation strategy should make full use of all these freedoms.

- Choice of the permutation $\pi : \mathcal{U} \rightarrow \mathcal{U}$ where $\pi(1)$ is the user encoded first and $\pi(K)$ is the user encoded last.
- Power allocation to different users.
- Choosing the unit norm steering vectors.

Each User with Unique Steering Vectors

Let $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ be the steering vector for user k and (with a little abuse of notation) $p_1, p_2, \dots, p_K$ be the powers allocated to the users. To respect the power constraint $\sum_{k=1}^K p_k = P$, it is assumed that $\|\mathbf{w}_k\| = 1$.

So the transmitted vector is given by

$$\mathbf{x} = \sum_{k=1}^K \sqrt{p_k} \mathbf{w}_k u_k$$

where u_k is the unit norm symbol for user k after doing DPC.

The proper sequencing is done via choosing u_k for each user.

As already outlined, the freedoms are in choosing p_k , $\mathbf{w}_k$ and the permutation π . With given

values of those factors, the rate region $\mathcal{C}_{\text{BC}} \subset \mathbb{R}_+^K$ is given by [33]

$$\mathcal{C}_{\text{BC}} = \{\mathbf{r} \in \mathbb{R}_+^K | r_{\pi(k)} \leq \log \left(1 + \frac{|\mathbf{g}_{\pi(k)}^H \mathbf{w}_{\pi(k)}|^2 p_{\pi(k)}}{1 + \sum_{j>k} |\mathbf{g}_{\pi(k)}^H \mathbf{w}_{\pi(j)}|^2 p_{\pi(j)}} \right), \quad \text{where } 1 \leq k \leq K. \} \quad (2.19)$$

A point on the rate region boundary is achieved when all K inequalities in (2.19) are satisfied with equality. When the equality conditions are satisfied, the point is called a Pareto optimal point. For the sake of brevity, the steering vectors can be compactly collected in a single $M \times K$ matrix $\mathbf{W}$ as

$$\mathbf{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \cdots \mathbf{w}_K] \in \mathbb{C}^{M \times K}$$

where feasibility implies $(\mathbf{W}^H \mathbf{W})_{ii} = 1$, where $1 \leq i \leq K$.

In a similar system model, maximisation of weighted sum rate under linear power constraint was done in [33]. In this work, the objective will be to characterise the Pareto frontier instead of specific objective function. This, in turn will aid in achieving a given rate performance with minimal power consumption.

For any choice of steering matrix $\mathbf{W}$, the power allocation $\mathbf{p} = [p_1, p_2, \dots, p_K] \in \mathbb{R}_{++}^K$ and the user ordering π will give a single Pareto optimal point which will be denoted by $\mathbf{r}_p(\mathbf{W}, \mathbf{p}, \pi)$ (the suffix denoting Pareto frontier). Based on this, those triplets which give the maximum possible throughput based on a fixed sum power, will be called optimal triplet as follows.

Definition 2.4 *The triplet $(\mathbf{W}^*, \mathbf{p}^*, \pi^*)$ will be called an optimal triplet if and only if the corresponding Pareto optimal point $\mathbf{r}_p(\mathbf{W}^*, \mathbf{p}^*, \pi^*)$ does not satisfy*

$$\mathbf{r}_p(\mathbf{W}^*, \mathbf{p}^*, \pi^*) \preceq \mathbf{r}_p(\mathbf{W}, \mathbf{p}, \pi) \in \mathbb{R}_+$$

for any other feasible triplet $(\mathbf{W}, \mathbf{p}, \pi)$ with $\|\mathbf{p}\|_1 \leq \|\mathbf{p}^*\|_1$.

The optimal triplet is not a unique one and there can be an infinite number of optimal triplets with same total power. Complete characterisations of all the optimal triplets is a difficult task. However, the following theorem attempts to characterise the triplets partially.

Theorem 2.1 *If $(\mathbf{W}^*, \mathbf{p}^*, \pi^*)$ is an optimal triplet and $\mathbf{r}_p(\mathbf{W}^*, \mathbf{p}^*, \pi^*) = \mathbf{r} \in \mathbb{R}_+$, then minimum broadcast power to achieve $\mathbf{r}$ is $\|\mathbf{p}^*\|_1$.*

Proof: From the achievable rates given in (2.19), if all user powers are increased by the same proportion, the rates improve for all users. This is true if the steering matrix and user ordering remain fixed. In general, $\mathbf{r}_p(\mathbf{W}, \mathbf{p}, \pi) \succeq \mathbf{r}_p(\mathbf{W}, \alpha \mathbf{p}, \pi)$, where $\alpha > 1$. To prove the theorem by contradiction, assume another triplet $(\mathbf{W}, \mathbf{p}, \pi)$ with $\|\mathbf{p}\|_1 < \|\mathbf{p}^*\|_1$ achieves $\mathbf{r}$. Consider the triplet $(\mathbf{W}, \frac{\|\mathbf{p}^*\|_1 \mathbf{p}}{\|\mathbf{p}\|_1}, \pi)$. Then

$$\mathbf{r}_p \left(\mathbf{W}, \frac{\|\mathbf{p}^*\|_1 \mathbf{p}}{\|\mathbf{p}\|_1}, \pi \right) \succeq \mathbf{r}_p(\mathbf{W}, \mathbf{p}, \pi) \succeq \mathbf{r} = \mathbf{r}_p(\mathbf{W}^*, \mathbf{p}^*, \pi^*). \quad (2.20)$$

So contrary to the previous assumption, the triplet $(\mathbf{W}^*, \mathbf{p}^*, \pi^*)$ is not optimal, since $(\mathbf{W}, \frac{\|\mathbf{p}^*\|_1 \mathbf{p}}{\|\mathbf{p}\|_1}, \pi)$ achieves a better rate with same power. $\blacksquare$

It is expected that these results will help in characterising the rate region and the optimal triplets of Definition 2.4. In the end, some important characteristics of the broadcast channel can be summarised as follows. It is worth comparing it with the multiple access channel of Section 2.2.1.

- Pareto optimal region for a sum power constraint may not be convex in general. Hence, by time sharing between points achievable by specific triplets, it is possible to further improve the achievable rates, which give rise to the time sharing region.
- It is possible to move along the Pareto optimal boundary of a rate region by changing the power allocations to users (while keeping the total power fixed), while in a multiple access system, the only way to move along the boundary is via time sharing among vertices.
- User orderings have very different implications for broadcast and multiple access channels. For a given power allocation and beamforming matrix, one among $K!$ different user permutations can outperform all others. An example will be shown as applicable to superposition coding or scalar broadcast systems. In contrast, for multiple access channel one user ordering can never completely outperform any other different ordering. There will be a trade off.

Superposition Coding

For the most general steering matrix $\mathbf{W}$, where the only constraint is unit column norm for each of its columns, optimisation strategies may be difficult to achieve. However, it is possible with superposition coding [12, 60] to merge data streams of users in a single data stream at the transmitter and that, in effect, results in same steering vector $\mathbf{w} \in \mathbb{C}^M$ for all users [61]. In that case, it is possible to define a scalar called *Channel Energy* for each user as $E_k \equiv |\mathbf{w}^H \mathbf{g}_k|^2$, where $1 \leq k \leq K$ and the rate region defined in (2.19) boils down to

$$\mathcal{C}_{\text{BC}} = \{\mathbf{r} | r_{\pi(k)} \leq \log \left(1 + \frac{E_{\pi(k)} p_{\pi(k)}}{1 + E_{\pi(k)} \sum_{i>k} p_{\pi(i)}} \right) \quad \text{where } 1 \leq k \leq K\} \quad (2.21)$$

Since the user numbering is arbitrary, without loss of generality it is possible to assume that $0 < E_K \leq E_{K-1} \leq \dots \leq E_2 \leq E_1$. [61] proved that the optimal user ordering to achieve maximum throughput or minimise power consumption is given by $K \rightarrow K-1 \rightarrow K-2 \rightarrow \dots \rightarrow 1$. This means K is the first user to be encoded and 1 is the last, in each case precanceling the interference from previous users.

For a rate vector $\mathbf{r}$, let $f_k(\mathbf{r})$ be the sum power requirement users $1, 2, \dots, k$ for $1 \leq k \leq K$ with the optimal user ordering as stated above. Important points to be noted about $f_k(\mathbf{r})$ are

- Total power requirement $P = f_K(\mathbf{r})$.
- If a particular user i precancels the interference from user j , then rate and power of user j does not matter for i . So, with our optimal user ordering, $f_k(\mathbf{r})$ is independent of r_i where $k < i \leq K$.
- Power allocated to user k is given by $f_k(\mathbf{r}) - f_{k-1}(\mathbf{r})$.

Theorem 2.2 *For the optimal user ordering, the power requirement to achieve any rate vector $\mathbf{r}$ is convex in $\mathbf{r}$.*

Proof: From the set defined in (2.19) and using the encoding order $K, K-1, \dots, 2, 1$, the power allocated to user k is $f_k(\mathbf{r}) - f_{k-1}(\mathbf{r})$. The interfering power is given by $f_{k-1}(\mathbf{r})$. Hence,

from Shannon's formula, the rate for user k is given by

$$r_k = \log \left(1 + \frac{E_k (f_k(\mathbf{r}) - f_{k-1}(\mathbf{r}))}{1 + E_k f_{k-1}(\mathbf{r})} \right) \quad \text{where } 1 \leq k \leq K \quad (2.22)$$

From (2.22), and noting that user 1 has no interference, $f_k(\mathbf{r})$ can be defined recursively as

$$f_1(\mathbf{r}) = \frac{e^{r_1} - 1}{E_1} f_k(\mathbf{r}) = e^{r_k} \left(f_{k-1}(\mathbf{r}) + \frac{1}{E_k} \right) - \frac{1}{E_k} \quad \text{where } 1 < k \leq K \quad (2.23)$$

Interestingly, the encoding procedure encodes the highest priority user last, but while computing the power requirements we have to go the other way round, computing the power requirement of highest priority user first.

Two properties of a convex function [43] will be helpful here.

- Composition of a convex function and an affine function is convex. Particularly, if $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is convex, $\mathbf{x}, \mathbf{a} \in \mathbb{R}^{N \times 1}$, $b \in \mathbb{R}$, then $\phi(\mathbf{a}^T \mathbf{x} + b)$ is convex in $\mathbf{x}$.
- Positive combination of convex functions is convex¹.

Then from (2.23), expanding $f_k(\mathbf{r})$ and using induction, it can be shown that

$$f_k(\mathbf{r}) = \frac{1}{E_1} e^{\sum_{i=1}^k r_i} + \sum_{j=2}^k \left(\frac{1}{E_j} - \frac{1}{E_{j-1}} \right) e^{\sum_{i=j}^k r_i} - \frac{1}{E_k} \quad (2.24)$$

In (2.24), $\frac{1}{E_1} e^{\sum_{i=1}^k r_i}$: Convex in $\mathbf{r}$,

$\sum_{j=2}^k \left(\frac{1}{E_j} - \frac{1}{E_{j-1}} \right) e^{\sum_{i=j}^k r_i}$: Positive combination of individual convex functions, since $E_j \leq E_{j-1}$, where $1 < j \leq K$ and finally $\frac{1}{E_k}$ is a constant. So $f_k(\mathbf{r})$ is convex in $\mathbf{r}$, where $1 \leq k \leq K$. ■

Lemma 2.1 *Under optimal user ordering and given steering vector $\mathbf{w}$, the achievable rate region is convex.*

Proof: Let $\mathbf{r}_1$ and $\mathbf{r}_2$ be two different achievable vectors located on the Pareto frontier for some total power P and steering vector $\mathbf{w}$, i. e. $f_K(\mathbf{r}_1) = f_K(\mathbf{r}_2) = P$. As already proven in Theorem 2.2

$$f_K(\alpha \mathbf{r}_1 + (1 - \alpha) \mathbf{r}_2) \leq \alpha f_K(\mathbf{r}_1) + (1 - \alpha) f_K(\mathbf{r}_2) = P, \quad \text{where } \alpha \in [0, 1] \quad (2.25)$$

So the power required to achieve a time shared point $\alpha \mathbf{r}_1 + (1 - \alpha) \mathbf{r}_2$ is less than total available power. Hence, time sharing is suboptimal. ■

For given slow fading channels, we can consider the steering vector $\mathbf{w}$ as argument of a mapping $\phi : \mathbb{C}^M \rightarrow \mathbb{R}_+^K$ such that

$$\phi(\mathbf{w}) \equiv [|\mathbf{g}_1^H \mathbf{w}|^2, |\mathbf{g}_2^H \mathbf{w}|^2, \dots, |\mathbf{g}_K^H \mathbf{w}|^2] = \mathbf{E} \in \mathbb{R}_+^K \quad (2.26)$$

Just like the rate-space, it is possible to define a Pareto frontier on the $\mathbf{E}$ space in the same way. It is evident from equations (2.22) that the rate vector is a component-wise monotonically

¹Can be shown using the second derivative or Hessian matrix test [43].

increasing function of the channel energy vector. If $\mathcal{F} \equiv \{\mathbf{w} \mid \|\mathbf{w}\| \leq 1, \mathbf{w} \in \mathbb{C}^M\}$ is the set of all feasible steering vectors respecting the power constraint, it can be proved [61] that $\phi(\mathcal{F})$ is a convex set. So it is desired to have a steering vector $\mathbf{w}$ whose image is located on the Pareto frontier of the $\mathbf{E}$ space. The problem was solved with appropriate power allocation in [61] and the significance of this solution has been discussed therein. However, to the best of our knowledge, characterisation of the pareto frontier for a given rate vector $\mathbf{r}$ is not addressed in literature. The present thesis will address the issues in the next chapter.

2.2.3 Duality between System Models

For a single user multi-antenna channel, the duality was proposed in [5]. For the multiple user case, the duality between uplink and downlink channels is an important principle to connect the two systems. It appears in several forms under various Quality of Service (QoS) constraints like throughput, data rate, SINR or MMSE between the transmitted and received signals. One of the first works to propose the duality was [62]. It was extended for a scalar channel in [63] and for a vector channel in [32, 54]. The current section will give some overview of the duality principles.

Uplink-Downlink Duality

The difference between a broadcast and multiple access channel lies in the difference direction of signal transmission. In a broadcast channel, the transmitter sends independent information to the geographically distant receivers whereas for a multiple access channel, the non-cooperating transmitters send information to the receivers. The scenario is similar to the familiar system of cellular base station and users. In current systems communication occur in orthogonal time or frequencies to avoid interference, which is not the case for systems discussed in the present thesis. In this context, [45] showed the degraded property of a scalar Gaussian broadcast channel where the users can be arranged in an appropriate order of channel strength. Further, [64] showed that superposition coding achieves optimal performance in these systems from an information theoretic point of view. However, superposition coding also requires the stronger receiver to cancel out the signal for weaker receivers. This entails very high computational complexity. Hence, in both broadcast and multiple access channel, successive decoding of various users has been the norm.

[63] showed that a Gaussian multiple access channel has the same rate region as a Gaussian degraded scalar broadcast channel. However, that requires a sum power constraint on the multiple access system rather than individual power constraints. Let the powers of users in a multiple access channel be $P_1, P_2, \dots, P_K$. The power limitation of the M antenna base station in the broadcast channel is P . The user channels are given by $\mathbf{h} = [h_1, h_2, h_K]$. $\mathcal{C}_{\text{BC}}$ and $\mathcal{C}_{\text{MAC}}$ represent the rate regions for the respective channels. Then the following theorem introduces the duality principle for multiple access and broadcast channels. It is assumed that the base station broadcasts under a total power constraint P . In the multiple access system, the users are transmitting with powers $P_1, P_2, \dots, P_K$.

Theorem 2.3 *For a single antenna base-station and K single antenna users having channels $h_1, h_2, \dots, h_K \in \mathbb{C}$, the power constrained rate regions are the same, i.e.*

$$\mathcal{C}_{\text{BC}}(P, \mathbf{h}) = \bigcup_{\sum_{i=1}^K P_i = P} \mathcal{C}_{\text{MAC}}(P_1, P_2, \dots, P_K, \mathbf{h}) \quad (2.27)$$

Proof: For convex capacity regions $\mathcal{C}_{\text{BC}}$ and $\mathcal{C}_{\text{MAC}}$, it suffices to prove that every corner point of $\mathcal{C}_{\text{BC}}$ is a corner point of $\mathcal{C}_{\text{MAC}}$ and vice versa.

Consider a decoding order $\pi(1), \pi(2), \dots, \pi(K)$ for a permutation $\pi : \{1, 2, \dots, K\} \rightarrow \{1, 2, \dots, K\}$ and multiple access power allocations $P_1, P_2, \dots, P_K$. The rate of user $\pi(i)$ in this multiple access system is

$$r_{\pi(i)} = \log \left(1 + \frac{h_{\pi(i)} P_{\pi(i)}}{1 + \sum_{j=i+1}^K h_{\pi(j)} P_{\pi(j)}} \right) \quad (2.28)$$

If the opposite $\pi(K), \pi(K-1), \dots, \pi(1)$ encoding order is used for a broadcast system with powers $P'_1, P'_2, \dots, P'_K$, the rate of user i is

$$r'_{\pi(i)} = \log \left(1 + \frac{h_{\pi(i)} P'_{\pi(i)}}{1 + h_{\pi(i)} \sum_{j=1}^{i-1} P'_{\pi(j)}} \right) \quad (2.29)$$

Define

$$a_i \equiv 1 + h_{\pi(i)} \sum_{j=1}^{i-1} P'_{\pi(j)} \quad \text{and} \quad b_i \equiv 1 + \sum_{j=i+1}^K h_{\pi(j)} P_{\pi(j)} \quad (2.30)$$

which are basically the denominator terms in the multiple access and broadcast SINRs respectively.

It can be shown [63, Appendix A] that if the powers satisfy

$$\frac{P_{\pi(i)}}{a_i} = \frac{P'_{\pi(i)}}{b_i} \quad \text{where} \quad i \in \{1, 2, \dots, K\} \quad (2.31)$$

then

$$\sum_{i=1}^K P_i = \sum_{i=1}^K P'_i$$

Obviously (2.31) is the condition of rate equality for the broadcast and multiple access systems. Hence, it is necessary to show that for any given set of power constraints in a MAC, there exists a set of powers in a BC which satisfy the condition (2.31). The broadcast power allocation can be derived from the multiple access power allocation based on duality transformation [63] as follows.

$$\begin{aligned} P'_{\pi(1)} &= \frac{P_{\pi(1)}}{1 + \sum_{j=2}^K h_{\pi(j)} P_{\pi(j)}} \\ P'_{\pi(2)} &= \frac{P_{\pi(1)} \left(1 + h_{\pi(2)} P'_{\pi(1)} \right)}{1 + \sum_{j=3}^K h_{\pi(j)} P_{\pi(j)}} \\ P'_{\pi(K)} &= P_{\pi(K)} \left(1 + h_{\pi(K)} \sum_{j=1}^{K-1} P'_{\pi(j)} \right) \end{aligned} \quad (2.32)$$

Note that (2.32) gives a set of broadcast powers from the corresponding multiple access

powers. Similarly (2.33) gives the conversions the other way round.

$$\begin{aligned}
 P_{\pi(K)} &= \frac{P'_{\pi(K)}}{1 + h_{\pi(K)} \sum_{j=1}^{K-1} P'_{\pi(j)}} \\
 P_{\pi(K-1)} &= \frac{P'_{\pi(K-1)} (1 + h_{\pi(K)} P_{\pi(K)})}{1 + h_{\pi(K-1)} \sum_{j=1}^{K-2} P'_{\pi(j)}} \\
 P_{\pi(1)} &= P'_{\pi(K)} \left(1 + \sum_{j=2}^K h_{\pi(j)} P_{\pi(j)} \right)
 \end{aligned} \tag{2.33}$$

The conversions outlined in (2.33) and (2.32) characterise the duality property of the channels and give the methods to switch back and forth between the systems. ■

In a multiple access channel, each user's transmit power and channel coefficient occur only in product form. Hence, the concept of scaling the channel and power in inverse proportions is useful in this context. For some $\alpha > 0$, if the power P_i of user i is replaced by $\frac{P_i}{\alpha}$ and the channel strength is replaced by αh_i , the rates of all users remain unaltered. Based on this observation, Theorem 2.4 was proved [63] for a MAC system with power allocation $\mathbf{p} = [P_1, P_2, \dots, P_K]$ and a broadcast system with scaled powers².

Theorem 2.4

$$\bigcap_{\alpha > 0} \mathcal{C}_{\text{BC}}(\mathbf{1}^T \frac{\mathbf{p}}{\alpha}, \alpha \mathbf{h}) = \mathcal{C}_{\text{MAC}}(P_1, P_2, \dots, P_K, \mathbf{h})$$

However, this theorem assumes that the broadcast channel is a degraded one where it is possible to somehow order the users according to their channel strengths. More general characterisation of a MIMO broadcast channel was carried out in [49, 65]. Both of them use the basic premise of DPC. Hence, if the interference is known non-causally at the transmitter, its effect can be completely nullified. Lemma 2.2 proposed in [65] expresses the result rigorously.

Lemma 2.2 *For an $M \times K$ broadcast channel with single antenna users having channels $\mathbf{h}_k \in \mathbb{C}^M$, let the receive signals for user k be*

$$y_k = \mathbf{h}_k x_k + s_k + \eta_k$$

where s_k is a Gaussian interference. If s_k is known at the transmitter and is independent of η_k , the channel capacity is same as if $s_k = 0$, i.e. no interference.

As a way of interpretation, suppose the transmitter needs to send a specific codeword to user 1. While formulating the transmit signal for user 2, the transmitter already knows the interference from user 1. So it can precancel the interference. User 1, however, sees the signal of user 2 as an interference. This is how the user ordering affects the rates. One more issue is that for a broadcast channel with multiple transmit antennas, the transmit covariance matrix

²Vector division or multiplication indicate a new vector formed by element-wise divisions or multiplications.

Σ_k for each user affects the rate and the power constraint is captured by

$$\sum_{k=1}^K \text{trace} [\Sigma_k] = P$$

Theorem 2.5 due to [54] establishes the duality for vector multiple access and broadcast systems. In what follows, $\mathbf{H} \equiv [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]$ will denote the collection of all users' channels in the broadcast system and $\mathbf{H}^H$ will denote the corresponding Hermitian conjugate, which is the effective channel of the multiple access system.

Theorem 2.5

$$\bigcup_{\mathbf{1}, \mathbf{p}=P} \mathcal{C}_{\text{MAC}}(\mathbf{p}, \mathbf{H}) = \mathcal{C}_{\text{BC}}(P, \mathbf{H}^H)$$

The duality relation has been used to characterise the optimal power allocation and to characterise the Pareto frontier under sum power constraint.

Minimax Duality

The minimax duality is another important class of duality proposed in literature. [66] establishes duality as the saddle point as a maximisation with respect to one argument and simultaneous minimisation with respect to another. The work unifies several different types of duality under a common optimisation framework derived directly from the fundamental principles of information theory for broadcast channels [67].

Also, [66] introduced the concept of duality under the condition of least favourable noise in a broadcast channel. For the broadcast channel, the capacity region depends only on the marginal distributions of outputs conditioned on the input and not on the joint distribution. Additionally, when the receivers are cooperating with each other, it is always possible to get a better sum rate compared to non-cooperating receivers. Based on that observation in [57], the minimax sum capacity obeys the following bound proposed in [66].

$$C \leq \min_{p(\mathbf{z})} \max_{p(\mathbf{x})} I(\mathbf{x}; y_1, y_2, \dots, y_K) \quad (2.34)$$

For a Gaussian broadcast channel, the solutions for the above optimisation problem can be shown to be necessarily Gaussian and actually achievable. Assuming unit noise variance, the sum capacity can be actually expressed as the optimal solution of the following problem.

$$\max_{\mathbf{S}_x} \min_{\mathbf{S}_z} \log |\mathbf{H}^H \mathbf{S}_x \mathbf{H} + \mathbf{S}_z| - \log |\mathbf{S}_z| \quad (2.35)$$

with constraints

$$\begin{aligned} \mathbf{S}_z(k, k) &= 1 \quad \text{where} \quad 1 \leq k \leq K \\ \text{trace} [\mathbf{S}_x] &= P \\ \mathbf{S}_z, \mathbf{S}_x &\succeq \mathbf{0} \end{aligned}$$

The constraints on diagonal element comes from the fact that each user by itself perceives a unit variance noise and the actual achievable rate is less than the worst case of noise correlation.

Based on that minimax expression, a novel perspective on uplink-downlink duality was proposed in [66]. In this context, the throughput of a Gaussian BC is the saddle point of a minimax mutual information expression. Then a joint optimisation problem is solved using the following parameters.

- A maximisation over the set of feasible transmit covariance matrices which satisfy the power constraint.
- A minimisation over the noise covariance matrices which satisfy the diagonal constraint.

The sum throughput can be derived using the decision-feedback equaliser and also the uplink-downlink duality in BC and MAC systems. The important contribution of the work [66] is to establish the equivalence between uplink downlink duality and Lagrangian duality in minmax optimisation. The best case transmit covariance and worst case noise covariance matrices in the Gaussian channel can be expressed via the dual variables by the formulation. In addition, the Lagrangian formulation of uplink downlink duality allows for the generalisation of sum power constraints to arbitrary linear constraints. This was adapted in subsequent works such as [34]. However, the duality relations are critically dependent on linearity of the constraints and are found to be invalid for general convex constraints.

2.2.4 Duality in Interference Channel

When multiple user pairs communicate without cooperation at either end, the interference is added to the noise floor, which does not allow the multiuser decoding/encoding. A distributed approach was proposed within a game theoretic framework [68] where the resource allocation was modelled as a non-cooperative game among users. The same work also showed the convergence towards a Nash equilibrium. However, the game theoretic approach failed to approach the sum rate optimal and [39] proposed some interference alignment method to reach the capacity. Based on these works, [69] extended the BC-MAC duality to the interference channel under the assumption of perfect channel information and proposed a transformation. Theorem 2.6 states the duality result.

Theorem 2.6 *If each receiver treats the signal intended for other users as noise, then the forward and reverse links of a general interference channel have the same achievable rate region under the same sum power constraint.*

The work of [69] was further generalised to a case of interference channel. In that case, each transmitter may transmit data to multiple receivers and each receiver may be interested in the data from multiple transmitters. For example, [70] expressed the interference channel in terms of links, where each desired link is basically a pair of physical transmitter and physical receiver. For L transmitters and L receivers, the signal received by receiver l is

$$\mathbf{y}_l = \sum_{k=1}^L \mathbf{H}_{kl} x_k + \boldsymbol{\eta}_l \quad (2.36)$$

where $\mathbf{x}_k \in \mathbb{C}^{M_k}$ is symbol from transmitter k , $\mathbf{y}_l \in \mathbb{C}^{N_l}$, $\mathbf{H}_{kl} \in \mathbb{C}^{N_l \times M_k}$ is the channel from transmitter k to receiver l , M_k is the number of antenna in transmitter k , N_l is the number of antenna in receiver l . $\boldsymbol{\eta}_l$ is the noise vector at receiver l having an identity covariance of appropriate order. The system is assumed to have some desired data links and some interfering links [70, Fig. 1], all transmitting at the same time and frequency. The coupling or crosstalk

matrix Φ of Section 2.2.2 is used here with the same notation. In that case, the interference and noise matrix is expressed as

$$\Omega_l = \mathbf{I} + \sum_{k=1}^L \Phi_{lk} \mathbf{H}_{lk} \Sigma_k \mathbf{H}_{lk}^H. \quad (2.37)$$

Here $\Phi_{ll} = 0$ by definition and Σ_k is the transmit covariance matrix of user k . The power of user k is $\text{trace}[\Sigma_k]$. Given this, the capacity of link l is [5]

$$c_l = \log |\mathbf{I} + \mathbf{H}_{ll} \Sigma_l \mathbf{H}_{ll}^H \Omega_l^{-1}| \quad (2.38)$$

Under a sum power constraint P_t , the achievable rate region is given by

$$\mathcal{R} \equiv \bigcup_{\sum_{i=1}^L \text{trace}[\Sigma_i] \leq P_t} \{\mathbf{r} | r_l \leq c_l \text{ where } 1 \leq l \leq L\} \quad (2.39)$$

The sum power minimisation problem and rate balancing problems have been solved in [70] using this framework. The work proposed a general interference network architecture termed B-MAC, which is a combination of Broadcast systems and Multiple Access systems. The goal has been to ensure the QoS for the network via a proposed mapping between SINR and rate vectors.

2.2.5 Resource Optimisation Strategies on Single Cell Systems

There are several optimisation strategies available for resource allocation in communication systems. In this section, the focus will be on resource optimisation for multiple access channels with joint linear decoding strategies [71]. Earlier, the duality principle had been discussed for multiple access and broadcast channels with Successive Interference Cancellation (SIC) or DPC schemes. However, Appendix C shows that the duality still holds between linear decoding and beamforming in multiple access and broadcast channels even without DPC or SIC schemes. Many of the optimisation strategies available in present literature (examples are [71] and [72]) depend on a specific formulation of the interference function to model the intracellular interference as described below.

Interference Function

For the K user multiple access system, let the power allocations be $\mathbf{p} = [p_1, p_2, \dots, p_K] \in \mathbb{R}_+^K$ and the interference plus noise vector tolerated by users be $\mathbf{I}(\mathbf{p}) = [I_1(\mathbf{p}), I_2(\mathbf{p}), \dots, I_K(\mathbf{p})] \in \mathbb{R}_+^K$. Given this notation, the SINR for user γ_k is given by $\frac{p_k}{I_k(\mathbf{p})}$. The following properties hold [73] for the interference vector.

- *Monotonicity:* $\mathbf{I}(\mathbf{p}_1) \succeq \mathbf{I}(\mathbf{p}_2) \Leftrightarrow \mathbf{p}_1 \succeq \mathbf{p}_2$,
- *Scalability:* $\mathbf{I}(\alpha \mathbf{p}) \succeq \alpha \mathbf{I}(\mathbf{p})$ when $\alpha \geq 1$ and $\mathbf{p} \in \mathbb{R}_+^K$

The monotonicity property captures the fact that for larger power allocation, the interference is more. The scalability implies that for scaling of powers, the corresponding interference plus noise grows sublinearly. To see the implication of this property, consider

$$p_j \geq I_j(\mathbf{p}) \Rightarrow \alpha p_j \geq \alpha I_j(\mathbf{p}) \geq I_j(\alpha \mathbf{p})$$

So we have

$$\frac{\alpha p_j}{I_j(\alpha \mathbf{p})} \geq \frac{p_j}{I_j(\mathbf{p})} \quad \text{where } \alpha \geq 1$$

implying upscaling improves SINR.

Given the multiple access system, the following optimisation problems are formulated for resource allocation in cellular systems. Further, these algorithms will also provide some useful benchmarks to show the effectiveness of the results proposed in the present work.

- *Sum Power Minimisation*: Given certain SINR targets for all the users, to find the minimum power requirement to achieve the target. i.e. finding $\min \|\mathbf{p}\|_1$ such that $\text{SINR}_k \geq \gamma_k$.
- *SINR Balancing*: Given a sum power constraint, balancing the SINRs with a certain ratio, i.e. finding $\max_{\mathbf{p} > \mathbf{0}} \min_k \frac{\text{SINR}_k}{\gamma_k}$ under the constraint $\|\mathbf{p}\|_1 \leq P_t$.

Using the monotonicity and scalability properties of the interference function as well as available solutions of the sum power minimisation problem will be discussed. It is assumed that the transmitted signals x_i from user i undergo fading $\mathbf{h}_i \in \mathbb{C}^M$ before being received and added up at the receiver. The receiver employs different linear decoding strategies for each receiver [43], instead of successive decoding and interference cancellation. Thus the received signal is

$$\mathbf{y} = \sum_{i=1}^K \mathbf{h}_i x_i + \boldsymbol{\eta} \quad (2.40)$$

and the decoding statistic for user k is given by $\mathbf{f}_k^H \mathbf{y}$ where $\mathbf{f}_k \in \mathbb{C}^M$ is the filter coefficient for user k . Based on this system, the following theorem was introduced in [73] which paves the way for the sum power minimisation algorithm to meet certain SINR targets.

Theorem 2.7 *If any feasible power allocation exists which meets the SINR targets, at the optimal point, all SINR constraints will be satisfied with equality, i.e. at optimal power allocation, the following conditions hold, $\text{SINR}_k = \gamma_k$*

Proof: Let at a feasible power allocation $\mathbf{p}^*$, $\text{SINR}_i > \gamma_i$ for some $i \in [1, K]$. Then for user i , we can still reduce p_i keeping $\text{SINR}_i \geq \gamma_i$. (Note that other users will have even better SINRs because of reduced interference from user i). In this new scenario, $\|\mathbf{p}\|_1 < \|\mathbf{p}^*\|_1$. So $\mathbf{p}^*$ is not the optimal power allocation with the given SINR targets. Hence, proved by contradiction. ■

It is important to note that, when the objective is to achieve (possibly different) SINR targets for users, mathematically, it means pre-multiply the interference vector $\mathbf{I}(\mathbf{p})$ with a diagonal square matrix $\boldsymbol{\Gamma}$ given by $\text{diag}(\gamma_1, \gamma_2, \dots, \gamma_K)$. So the vector $\mathbf{I}(\mathbf{p})$ can be taken as the sum of actual interference and noise vectors pre-multiplied by $\boldsymbol{\Gamma}$ in this framework. At optimal power allocation, we can say that $\mathbf{p} = \mathbf{I}(\mathbf{p})$ i.e. the optimal point is a fixed point of the function $\mathbf{I}(\mathbf{p})$.

Theorem 2.8 from [73] helps to formulate the fixed point iteration algorithm for sum power minimisation.

Theorem 2.8 *If the function $\mathbf{I}(\mathbf{p})$ has a fixed point, then that point is unique. In other words, there exists only one fixed point or none.*

Proof: To prove by contradiction, suppose $\mathbf{p}$ and $\mathbf{p}'$ are two fixed points. Now, we must have $p_i > 0$ and $p'_i > 0$ when $i \in [1, K]$. Without any loss of generality, we can assume there

exists one or more i such that $p_i \leq p'_i$. Hence, there exists $\alpha > 1$ such that $\alpha \mathbf{p} \geq \mathbf{p}'$ and $\alpha p_i = p'_i$ for some i . The monotonicity and scalability properties imply

$$p'_i = I_i(\mathbf{p}') \leq I_i(\alpha \mathbf{p}) < \alpha I_i(\mathbf{p}) = \alpha p_i$$

Since $p_j = \alpha p'_j$, it is a contradiction, implying that the fixed point, if exists, is unique. $\blacksquare$

Next the following lemmas show that the fixed point can be approached by an iterative calculation of the interference function as proposed in [73].

Lemma 2.3 *If $\mathbf{p}$ is a feasible power vector then $\mathbf{I}^n(\mathbf{p})$ is a monotonically decreasing sequence of feasible power vectors that converges towards the fixed point $\mathbf{p}^*$*

Lemma 2.4 *If the target is feasible, then starting from the all zero power vector $\mathbf{0}$, the standard power control algorithm produces a monotonic increasing sequence $\mathbf{I}^n(\mathbf{0})$ converging towards the fixed point $\mathbf{p}^*$.*

However, as shown in the convergence analysis of [72], the fixed point iteration is rather inefficient in terms of resource requirement. The same work also proposes a better method to solve the sum power minimisation problem.

Sum Power Minimisation with Matrix Based Interference

For almost any practically relevant case, the interference crosstalk can be modelled by a $K \times K$ non-negative coupling matrix $\Psi(\mathbf{F})$ which depends on a parameter $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K)$ where $\mathbf{f}_k$ are chosen from a compact set $\mathfrak{F}_k$, the set of all possible receive strategies for user k . To keep the analysis as general as possible, no assumption is made about the nature of f , except that the k -th row of $\Psi(\mathbf{F})$ depends only upon $\mathbf{f}_k$. This means the interference of any user for a fixed power allocation is affected by the strategy of that user only and none else. As an example, for the present case, $\mathbf{f}_k$ could stand for specific realisation of the filter coefficients. The users are coupled by not only the interference, but also by power constraints. The power constraint affects the SINR only in the presence of noise. For k -th user, the noise is $\sigma^2 N_k(\mathbf{f}_k)$, where $N_k(\mathbf{f}_k)$ is the noise enhancement factor depending on the receive strategy $\mathbf{f}_k$. All K noise enhancement factors are collected in $\mathbf{N}(\mathbf{F}) = [N_1(\mathbf{f}_1), N_2(\mathbf{f}_2), \dots, N_K(\mathbf{f}_K)]^T$.

Combining the coefficients in an extended coupling matrix

$$\mathbf{G}(\mathbf{F}) = \Psi(\mathbf{F})|\mathbf{N}(\mathbf{F}) \quad (2.41)$$

The $K \times (K+1)$ coupling matrix determines overall coupling between the users. In particular, $[\Psi(\mathbf{F})\mathbf{p}^{n+1} + \mathbf{N}(\mathbf{F})]_k$ is the interference+noise power for k -th user. Without loss of generality, we are assuming the noise power is normalised to unity and any multiplication factor of the noise power is absorbed in defining the channel coefficient itself. Using this notation, the SINR constraints can be written as

$$\mathbf{p} \geq \mathbf{I}[\Psi(\mathbf{F})\mathbf{p} + \mathbf{N}(\mathbf{F})] \quad (2.42)$$

The minimum transmit power requirement can be found by joint optimisation over $\mathbf{p}$ and $\mathbf{F}$. Exploiting this further, the k -th row of the coupling matrix depends only upon $\mathbf{f}_k$. Thus, for a given power allocation, the optimal parameters $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K$ can be found independently of each

other. For each user, the optimal strategy is to choose the parameter $\mathbf{f}_k \in \mathfrak{F}_k$ that minimises the interference. This leads to the interference function

$$I_k(\mathbf{p}) = \min_{\mathbf{f}_k \in \mathfrak{F}_k} [\psi_k^T(\mathbf{f}_k)\mathbf{p} + N_k(\mathbf{f}_k)] \quad (2.43)$$

where $\psi_k^T(\mathbf{f}_k)$ is the k -th row of $\Psi(\mathbf{F})$. The optimal strategy as a solution of the minimisation problem in (2.43) can be found using the Wirtinger complex derivative (See Appendix D). The optimal filter coefficient can be shown to be

$$\hat{\mathbf{f}}_k = \mathbf{A}^{-1}\mathbf{h}_k \quad (2.44)$$

where $\mathbf{A} = \sum_{i=1, i \neq k}^K \mathbf{h}_i \mathbf{h}_i^H p_i + \mathbf{I}_K$.

The interference function as defined in (2.43) satisfies all the criterion monotonicity and scalability properties [73] discussed earlier. It is easily verified that the above definition of $\mathbf{I}(\mathbf{p})$ obeys the axioms of standard interference function. So, now the optimisation problem can be formulated as finding the $\min_{\mathbf{p} > \mathbf{0}} \|\mathbf{p}\|_1$ obeying $p_k \geq \gamma_k I_k(\mathbf{p})$. The only optimisation parameter here is $\mathbf{p}(> \mathbf{0})$. However, optimisation with respect to the receive strategy for individual users is implicit in the definition of $I_k(\mathbf{p})$. This is possible because the SINR of every user is independent of the receive strategy of all others. The following theorem due to [72] further characterises the optimal point and gives us a new algorithm to solve the optimisation problem.

Theorem 2.9 *A power allocation vector $\mathbf{p}^*$ is optimal if and only if there exists receive strategies $\hat{\mathbf{f}}_k \in \mathfrak{F}_k$ satisfying the following conditions.*

$$\hat{\mathbf{f}}_k = \operatorname{argmin}_{\mathbf{f}_k \in \mathfrak{F}_k} [\psi_k^T(\mathbf{f}_k)\mathbf{p} + N_k(\mathbf{f}_k)]$$

and

$$\mathbf{p}^* = \Gamma[\Psi(\hat{\mathbf{F}})\mathbf{p} + \mathbf{N}(\hat{\mathbf{F}})], \quad \text{where } \hat{\mathbf{F}} = (\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \dots, \hat{\mathbf{f}}_K) \text{ is the unified optimal receive strategy.}$$

Based on the results, Algorithm 2.1 was proposed in [72] which has a better convergence property compared to [73].

Algorithm 2.1: Matrix Iteration Algorithm for Power Minimisation

Require: SINR targets $\gamma_1, \gamma_2, \dots, \gamma_K$

- 1: $\mathbf{p}_0 \leftarrow$ Any feasible power allocation to meet the SINR targets
 - 2: **while** $\|\mathbf{p}_n - \mathbf{p}_{n+1}\| \geq \epsilon$ **do**
 - 3: $\mathbf{f}_k^n \leftarrow \operatorname{argmin}_{\mathbf{f}_k \in \mathfrak{F}_k} [\psi_k^T(\mathbf{f}_k)\mathbf{p}^n + N_k(\mathbf{f}_k)]$ {Optimise Receive Strategy}
 - 4: $\mathbf{p}^{n+1} \leftarrow \Gamma[\Psi(\mathbf{f}^n)\mathbf{p}^{n+1} + \mathbf{N}(\mathbf{f}^n)]$ {Optimise Power Allocation}
 - 5: **end while**
-

SINR Balancing with Spectral Radius Optimisation

From Section 2.2.5, finding a feasible power allocation satisfying the SINR constraints is essential for initialisation of the matrix based iterative algorithm. However, this in general is not a trivial task. When the number of antenna is larger than or same as the number of users, any SINR target is feasible. However, an efficient alternative method was proposed in [72]. To do so, it is necessary to define an extended power vector (assuming unit noise variance) as

$\bar{\mathbf{p}} = [p_1, p_2, \dots, p_K, 1]^T$ i.e. $\bar{\mathbf{p}}$ contains as its tuples all the transmitted powers and the noise variance. Then the interference function can be computed as

$$\mathbf{I}(\mathbf{p}) = \mathbf{G}(\mathbf{F})\bar{\mathbf{p}} \quad (2.45)$$

For the SINR balancing problem, it is seen that at the optimal point, all the inequalities will be satisfied with equality, i.e. with full utilisation of available power, where $|\mathbf{p}|_1 = P_t$, the relative SINRs, $\frac{\text{SINR}_k}{\gamma_k}$ will be balanced at same level for all users. For mathematical convenience, let the inverse of this relative SINR be denoted by C , i.e.

$$C = \min_{\mathbf{p} > 0} \max_k \frac{\gamma_k}{\text{SINR}_k} \quad (2.46)$$

with the sum power constraint.

Clearly, C is a function of $\mathbf{\Gamma}$ and total available power P_t . If $C(\mathbf{\Gamma}, P_t) \leq 1$, then the target $\mathbf{\Gamma}$ is feasible with P_t , otherwise not.

Hence, at optimal point of the SINR balancing problem

$$C \cdot p_k^* = \gamma_k I_k(\mathbf{p}^*)$$

where optimal receive strategies $\hat{\mathbf{f}}_k$ are employed by all users to maximise individual SINRs. Since all constraints are satisfied with equality, the optimisers of $\mathbf{f}_k$ and $\mathbf{p}$ are related as

$$C \cdot \mathbf{p}^* = \mathbf{\Gamma}[\Psi(\hat{\mathbf{F}})\mathbf{p}^* + \mathbf{N}(\hat{\mathbf{f}})] = \mathbf{\Gamma}\mathbf{G}(\hat{\mathbf{f}})\bar{\mathbf{p}}^* \quad (2.47)$$

where $\bar{\mathbf{p}}^*$ is the optimal power vector extended by noise variance. If the level C is known, then the above equation can be solved for $\mathbf{p}^*$. However, the balanced relative SINRs are not known a-priori in general. A method to compute the balanced inverse SINR factor C and corresponding power allocation $\mathbf{p}^*$ for given $\hat{\mathbf{f}}$ by Eigen value decomposition was proposed in [72]. Since at optimal point, total available power is utilised with $\mathbf{1}^T \mathbf{p}^* = P_t$, adding all the scalar equations of (2.47)

$$C = \frac{1}{P_t} \mathbf{1}^T \mathbf{\Gamma} \mathbf{G}(\hat{\mathbf{f}}) \bar{\mathbf{p}}^* \quad (2.48)$$

Here $\mathbf{1} \equiv [1, 1, \dots, 1]$ is the vector of all 1s which effectively adds the elements of other vectors. So the $K + 1$ equations of (2.47) and (2.48) are combined in an Eigen System

$$\phi(\hat{\mathbf{F}}, P_t) \bar{\mathbf{p}}^* = \frac{1}{C} \bar{\mathbf{p}}^* \quad (2.49)$$

where $\Phi(\mathbf{F}, P)$ is defined as $\begin{bmatrix} \mathbf{\Gamma} \mathbf{G}(\mathbf{F}) \\ \frac{1}{P} \mathbf{1}^T \mathbf{\Gamma} \mathbf{G}(\mathbf{F}) \end{bmatrix}$. So, it shows that the extended power vector $\bar{\mathbf{p}}^*$ is an Eigen Vector of the extended coupling matrix $\Phi(\mathbf{F}, P)$. Since $\Phi(\mathbf{F}, P)$ is a non-negative matrix, the Perron Frobenius Theorem (see [74]) says that there is only a single Eigen Vector with all positive components corresponding to a positive Eigen value, called Perron Root of the system. So the Perron root of extended coupling matrix gives the relative inverse SINR C at balanced condition and the corresponding Eigen Vector (with last component normalised to noise variance) gives the extended power allocation vector $\bar{\mathbf{p}}^*$. Based on these observations, Algorithm 2.2 has been proposed in [72].

Algorithm 2.2: SINR Balancing with Spectral Radius Optimisation

Require: Total Power P_t and SINR priorities $\gamma_1, \gamma_2, \dots, \gamma_K$
 1: Find any $\mathbf{p} \in \mathbb{R}_+^K$ obeying $\|\mathbf{p}_0\|_1 = P_t$ {Initialise with any power allocation}
 2: **while** $\|\mathbf{p}_n - \mathbf{p}_{n+1}\| \geq \epsilon$ **do**
 3: Find the optimal receive strategy from (2.44)
 4: Find the fixed point satisfying $\mathbf{p}^{n+1} = \Gamma[\Psi(\mathbf{f}^n)\mathbf{p}^{n+1} + \mathbf{N}(\mathbf{f}^n)]$
 5: **end while**

Figure 2.5: Two User Interference Channel

2.3 Interference Management in Traditional Communication Systems

When multiple nodes transmit information in the same time and frequency resource blocks, the signals cause interference to each other. Mitigating the effects of interference for optimal decoding of the interference has been a crucial challenge in multiuser communication systems. Although the specific coding and optimisation techniques depend on the cooperation scheme employed in a system, the methods employed are usually classified into three major paradigms.

- Interference cancellation decoding
- Interference ignorant decoding
- Interference alignment decoding

These are described in detail with a survey of the relevant literature in the following sections.

2.3.1 Interference Cancellation Decoding

For a strong interference in a two user system, it is better to decode and cancel the interference before decoding the desired signal. A simplistic version of the scheme was first proposed in [6] and can be used in many different networks.

Figure 2.5 shows a couple of transmitter-receiver pairs and each receiver is interested in the information of its corresponding transmitter. Suppose h_{11} and h_{22} are the direct channels and h_{12} and h_{21} are the interfering channels. The input output relationships are given by

$$y_i = h_{1i}x_1 + h_{2i}x_2 + \eta_i \quad \text{for } i = 1, 2 \quad (2.50)$$

with normalised Gaussian noise $\mathbb{E}(|\eta_i|^2) = 1$. In the absence of the cross-channels, the achievable

Figure 2.6: Two User Rate Region without Interference
rate region is a rectangle shown in Figure 2.6 given by (2.51).

$$r_i \leq \log(1 + P_i) \quad \text{for } i = 1, 2. \quad (2.51)$$

$$|h_{12}|^2 \geq 1 + P_2 \quad |h_{21}|^2 \geq 1 + P_1 \quad (2.52)$$

[6] showed that if the cross channels satisfy the conditions (2.52), the rectangular region given by (2.51) is achievable. The advantage of single user capacity in a two user system can not be shown to exist for more than two users, which is still an open problem. In the expanded work of [7], a distinction was made between interference network and interference channel. Interference channel was classified as a special case of the interference network. [6] characterised a K user interference network by K principal links³ and $K \times (K - 1)$ interfering links. In the absence of interference, let the principal link capacities be C_k for $1 \leq k \leq K$. With this assumption, Theorem 2.10 from [7] gives a pair of outer and inner bounds for the achievable rate regions.

Theorem 2.10 *The capacity of the interference channel $\mathcal{C}$ is bounded by $\mathcal{B}_1 \subset \mathcal{C} \subset \mathcal{B}_0$ where $\mathcal{B}_1 \equiv$ Region bounded by the $K - 1$ dimensional hyperplane containing rate vectors $[C_1, 0, 0, \dots, 0], [0, C_2, 0, \dots, 0], [0, 0, \dots, C_K]$ $\mathcal{B}_0 \equiv$ Parallelepiped containing all rate vectors tuple-wise dominated by $[C_1, C_2, \dots, C_K]$.*

Further refinement to the rate region of an interference channel was done in [60]. The same work, using the concept of jointly typical sequence, developed a coding technique which improves over [6]. [8] extended the concepts to moderately strong interference instead of weak interference.

2.3.2 Interference Ignorant Decoding

When the interference is weak enough compared to the signal strength and the noise, it is completely ignored in the decoding process. This in effect, adds the interference power to the noise floor (increasing the symbol error rate). From the practical perspective, this is the most feasible approach for dealing with interference and has been in deployment in many different systems, including cellular communication. Even the optimisation strategies described in Section 2.2.5 make use of this approach. Information theoretic validation of this was provided by the works of [12, 75, 76]. One important contribution of [76] was to characterise the two user

³The link between each transmitter-receiver pair is called a principal link. When a transmitter sends only interfering signals to a user, it is an interfering link.

capacity region with the help of the following input-output relations.

$$y_1 = x_1 + \sqrt{a}x_2 + z_1 \quad y_2 = \sqrt{b}x_1 + x_2 + z_2 \quad (2.53)$$

where $\sqrt{a}$ and $\sqrt{b}$ are the cross channel coefficients and users have power constraints P_1 and P_2 . [76] showed that the bounds (2.54)–(2.56) are satisfied by the achievable rate tuple $[r_1, r_2]$ when $\mu > 0$, $\frac{1+bP_1}{b(1+P_1)} \leq \eta_1 \leq \frac{1}{b}$ and $a \leq \eta_2 \leq \frac{a(1+P_2)}{1+aP_2}$.

$$\begin{aligned} r_1 + \mu r_2 \leq & \min_{0 \leq \rho_i \leq 1, (\sigma_1^2, \sigma_2^2) \in \Sigma} \frac{1}{2} \log(1 + \frac{P_1^*}{\sigma_1^2}) - \frac{1}{2} \log(aP_2^* + 1 - \rho_1^2) + \frac{1}{2} \log \left(1 + P_1 + aP_2 - \frac{(P_1 + \rho_1 \sigma_1)^2}{P_1 + \sigma_1^2} \right) \\ & + \frac{\mu}{2} \log \left(1 + \frac{P_2^*}{\sigma_2^2} \right) - \frac{\mu}{2} \log(1 + bP_1^* - \rho_2^2) + \frac{\mu}{2} \log \left(1 + P_2 + bP_1 - \frac{(P_2 + \rho_2 \sigma_2)^2}{P_2 + \sigma_2^2} \right) \end{aligned} \quad (2.54)$$

$$r_1 + \eta_1 r_2 \leq \frac{1}{2} \log \left(1 + \frac{b\eta_1 - 1}{b(1 - \eta_1)} \right) - \frac{\eta_1}{2} \log \left(1 + \frac{b\eta_1 - 1}{1 - \eta_1} \right) + \frac{\eta_1}{2} \log(1 + bP_1 + P_2) \quad (2.55)$$

$$r_1 + \eta_2 r_2 \leq \frac{1}{2} \log(1 + P_1 + aP_2) - \frac{1}{2} \log \left(1 + \frac{a - \eta_2}{\eta_2 - 1} \right) + \frac{\eta_2}{2} \log \left(1 + \frac{a - \eta_2}{a(\eta_2 - 1)} \right) \quad (2.56)$$

where

$$\Sigma = \begin{cases} \{(\sigma_1^2, \sigma_2^2) \mid \sigma_1^2 > 0, 0 < \sigma_2^2 \leq \frac{1-\rho_1^2}{a}\} & \text{if } \mu \geq 1 \\ \{(\sigma_1^2, \sigma_2^2) \mid 0 < \sigma_1^2 \leq \frac{1-\rho_2^2}{b}, \sigma_2^2 > 0\} & \text{otherwise} \end{cases} \quad (2.57)$$

Further, if $\mu \geq 1$

$$P_1^* = \begin{cases} P_1, & 0 < \sigma_1^2, \quad \text{if } 0 < \sigma_1^2 \leq \left(\left(\frac{1}{\mu} - 1 \right) P_1 + \frac{1-\rho_2^2}{b\mu} \right)^+ \\ \frac{1-\rho_2^2 - b\mu\sigma_1^2}{b(\mu-1)}, & \text{if } \left(\left(\frac{1}{\mu} - 1 \right) P_1 + \frac{1-\rho_2^2}{b\mu} \right)^+ < \sigma_1^2 \leq \frac{1-\rho_2^2}{b\mu} \\ 0 & \text{otherwise} \end{cases} \quad (2.58)$$

$$P_2^* = P_2, \quad 0 < \sigma_2^2 < \frac{1-\rho_1^2}{a} \quad (2.59)$$

where $(x)^+ \equiv \max\{x, 0\}$ and if $\mu \in (0, 1)$

$$P_1^* = P_1 \quad 0 < \sigma_1^2 \leq \frac{1-\rho_2^2}{b} \quad (2.60)$$

$$P_2^* = \begin{cases} P_2 & \text{if } 0 < \sigma_2^2 \leq \left((\mu - 1) P_2 + \mu \left(\frac{1-\rho_1^2}{a} \right) \right)^+ \\ \frac{\mu(1-\rho_1^2) - a\sigma_2^2}{a(1-\mu)} & \text{if } \left((\mu - 1) P_2 + \mu \left(\frac{1-\rho_1^2}{a} \right) \right)^+ < \sigma_2^2 \leq \mu \left(\frac{1-\rho_1^2}{a} \right) \\ 0 & \text{otherwise} \end{cases} \quad (2.61)$$

The different bounds depend on different models of a genie aided channels [45, Section 6.4] and in general, none of the bounds dominate the others. These bounds were obtained based on the extremal inequalities [77] and the entropy power inequality [78].

Further study of the interference ignorant decoding was done in [79] which improved the

state of the art over [12]. With an assumption of symmetric interference channels, the work [79] showed that if the interference falls below a certain threshold, completely ignoring the interference will achieve sum rate optimality. The analysis can be extended to more than two users, finding the threshold interference regime.

2.3.3 Interference Alignment Decoding

In contrast to interference cancellation or interference as noise, the concept of interference alignment introduces a completely new paradigm of treating interference as belonging to vectors in a space [80]. It can be said that the interference alignment approach uses orthogonalisation among users. This approach has been termed as *cake cutting* of degrees of freedom. However, the orthogonalisation takes place in a spatial domain. In other words, the approach attempts to orthogonalise spatial signatures of different receivers transmitting with multiple antennas to cancel the interference. The principle of orthogonalising interference can be applied in different forms in different contexts. One of the pioneering works was [81] which applied the concept to satellite channels. The basic concept involves processing the sum total of signals being distorted by vector channels and being added in the air. Suppose x_i is the scalar transmitted by user i for $1 \leq i \leq K$. After distortion by the vector channel from user i to the base station, it becomes $\mathbf{h}_i x_i$ and the base station receives the sum total of all signals.

$$\mathbf{y} = \sum_{i=1}^K \mathbf{h}_i x_i + \boldsymbol{\eta} \quad (2.62)$$

From the multiple access system described by (2.62), suppose the receiver is interested to recover only x_1 . In that case, $\sum_{i=2}^K \mathbf{h}_i x_i$ constitute interference. If it is possible to find a vector $\mathbf{g}$ which is orthogonal to the space spanned⁴ by $\{\mathbf{h}_2, \mathbf{h}_3, \dots, \mathbf{h}_K\}$, then pre-multiplying the received signal $\mathbf{y}$ with $\mathbf{g}^H$ eliminates all interference. So the basic concept of interference alignment is to confine the interfering vectors to a low dimensional subspace so that an orthogonal subspace can be obtained.

Degree of Freedom Enhancement via Interference Alignment

Application of orthogonalising interference in a multiuser communication system was introduced in [39]. It proposed the interference alignment for a two user channel (as in Figure 2.5) and gave an alternate method of tackling interference than other computationally complex algorithms. In addition, [39] also showed the result of Theorem 2.11.

Theorem 2.11 *The number of degree of freedom for a K user interference channel is $\frac{1}{2}$ per user, irrespective of K .*

The theorem says even when there are not enough degrees of freedom offered by the spatial dimensions, each user gets half of the total degrees of freedom. The detailed proof was given in [39].

Linear Beamforming in Two User Interference Channel

The principles of interference alignment were further explored [41, 82] with a system model as shown in Figure 2.7. The important difference with Figure 2.5, is that both of the receivers are interested to recover signals from both of the transmitters. So it is not possible to classify

⁴See Appendix A for a review of necessary algebraic concepts.

Figure 2.7: Two User Relay Channel with Transmit and Receive Filters

the channels as principal channels and interfering channels, as in Section 2.3.2. The given system model can be viewed as two broadcast systems from the transmitters' viewpoint or as two multiple access systems from the receivers' viewpoint. It is assumed that μ_{ij} independent data streams must be carried from transmitter i to receiver j for $i, j = 1, 2$. Also, transmitter i is assumed to have M_i antennas and receiver j has N_j antenna. Such a system is denoted as an (M_1, M_2, N_1, N_2) system. It is assumed that the transmitter uses a linear prefiltering before sending out the data (which may involve DPC) and multiplies the transmit vector by a matrix $\mathbf{Q}_i$, where $i = 1, 2$. Similarly, the receivers use linear receive filter to recover the data. The linear receive filter for receiver j to recover data from user i is given by Ψ_{ji} .

To *design* the system, [41] has used the auxiliary variables M'_i, N'_j and μ'_{ij} . Among the variables, M'_i and N'_j can be computed as functions of the transmit-receive antennas. Then μ'_{ij} can be calculated to satisfy the constraints given below.

$$\max\{\mu'_{11} + \mu'_{12} + \mu'_{22}, \mu'_{11} + \mu'_{12} + \mu'_{21}\} \leq N'_1 \quad (2.63)$$

$$\max\{\mu'_{11} + \mu'_{21} + \mu'_{22}, \mu'_{22} + \mu'_{12} + \mu'_{21}\} \leq N'_2 \quad (2.64)$$

$$\mu'_{11} + \mu'_{21} \leq M'_1 \quad (2.65)$$

$$\mu'_{22} + \mu'_{12} \leq M'_1 \quad (2.66)$$

To complete the design procedure, [41] proposed the following method.

- Algorithm to determine the auxiliary variables M'_i, N'_j .
- Compute μ_{ij} based on auxiliary variables μ'_{ij} .
- Compute the transmit prefilters $\mathbf{Q}_1, \mathbf{Q}_2$.

This analysis leads to the conclusion that the present system consists of two non-interfering broadcast channels exploiting the spatial dimensions and using the principles of interference alignment. The orthogonality of the channels ensure that interferences are nullified at the receiver end.

Interference Alignment in Multiuser System

Following the works of [41], linear interference alignment was generalised to a multiuser scenario [83]. The system consisted of K transmitter receiver pairs. Transmitter k (having M_k antennas) attempts to transmit μ_k independent data streams to receiver k (having N_k

antennas). Such a system has been denoted as $(M_1, N_1, \mu_1), (M_2, N_2, \mu_2), \dots, (M_K, N_K, \mu_K)$. When M_k, N_k, μ_k are all independent of k , that means every transmitter receiver pair has the same number of antennas and data streams. Hence, the symmetric system is denoted by $(M, N, \mu)^K$. For example, in Figure 2.7, if $\mu_{ij} = 0$ for $i \neq j$, it can be described as an $(M_1, N_1, \mu_1), (M_2, N_2, \mu_2)$ system. The feasibility of interference alignment has been studied in terms of the precoding matrices $\mathbf{Q}_i \in \mathbb{C}^{M_k \times \mu_k}$ and receive matrices $\Psi_j \in \mathbb{C}^{\mu_k \times N_k}$. [83] showed that if precoding and receive matrices satisfying the following conditions for all $i \neq j$, the linear interference alignment problem is feasible and the system can be constructed as a collection of non-interfering channels.

$$\text{Rank}(\mathbf{Q}_k) = \text{Rank}(\Psi_k) = \text{Rank}(\Psi_k^H \mathbf{H}_{kk} \mathbf{Q}_k) = \mu_k \quad (2.67)$$

$$\Psi_j^H \mathbf{H}_{jk} \mathbf{Q}_k = \mathbf{0} \quad (2.68)$$

Both of these conditions have their interesting interpretations. It can be shown that (2.68) implies there exists an interference free subspace along which desired transmission can take place. Further, (2.67) implies that this subspace has sufficient dimensions to transmit the desired data streams. Since taking Hermitian conjugates of both sides in (2.67) and (2.68) will satisfy the conditions, it can be said that switching roles of transmitters and receivers will be feasible as well. This is the reciprocity principle [5], which is a form of duality principle. The general interference channel of [83] is proper (in the sense that alignment is a feasible problem) if

$$d \leq \frac{M + N}{K + 1}$$

In summary, as pointed out, there has been an array of recent literature on interference alignment and exploiting spatial degrees of freedom (more detailed reviews can be found in [80, 84]). Still there are interesting open problems with reference to different system models within the context of the present research.

2.4 Interference Management in Heterogeneous and Relay-Assisted Systems

Most traditional theoretical models dealt with a single hop communication where the data transmitted by source is received by the destination, usually perturbed by a channel and corrupted by noise. Due to the exponential growth of wireless communication systems and devices, the recent research focus shifted to a more flexible system where the transmitter(s) and receiver(s) are being coordinated by one or more relays. The relay node receives the source data and retransmits to them to the receiver. This can introduce a form of cooperation and also possibly cascaded channels using different network level protocols or modulation schemes. In particular, MIMO relays have received significant attention in recent literature because of improved link performance and coverage expansion on traditional cellular networks [85]. In terms of processing of incoming and outgoing information, the relays have been classified into two broad categories

- Amplify and Forward Relay systems,
- Decode and Forward Relay Systems.

Further, based on the mode of operation, the relays have been classified into two categories

- Half Duplex relays,
- Full Duplex relays.

Different types of relay systems call for novel methods of interference mitigation and resource allocation. The following sections cover the state of the art research landscapes for relay assisted interference management.

2.4.1 Amplify and Forward Relay Assisted Systems

AF relays are categorised as *non-regenerative* relays. They process the incoming signal via receive and transmit combining matrices before a retransmission to the destination. The goal is to mitigate the effect of noise and interference. Although these relays do not decode the signal and hence provide suboptimal performance, they offer simplified implementation and lower computational complexity [86,87]. These relays offer flexible design schemes, which make it possible to achieve a continuum of complexity-information rate trade off.

As an example, when the receiver operates on maximum likelihood principle, the maximum information criterion has been used to optimise the relay transceiver matrix [88,89]. On the other hand, when the goal is to reduce the decoding complexity, it is possible to use a linear zero forcing or minimum mean square error equaliser at the destination. Then the relay transceiver can be optimised to improve the performance, as measured by bit error rate or Shannon's capacity [90,91]. In contrast, [92,93] focused on analytical studies with a goal to predict the performance based on a preset operating point.

Many of the other works focused on analysing the Diversity-Multiplexing Trade off (DMT) for relay channels. This trade off, proposed in [36], characterises the conflicting objectives of achieving high data rate with low bit error rates. Many works have contributed to the analysis of DMT in the context of MIMO relays schemes [94,95]. However, DMT, in itself, is not enough to characterise the diversity in high spectral efficiency and SINR regimes, since it cannot differentiate between different spectral efficiencies for the same multiplexing gain. Further, it was also shown that DMT fails to predict the performance in point to point MIMO systems when the channel capacity falls below a certain threshold. This was first observed in [96] and analysed comprehensively in [97].

2.4.2 Decode and Forward Relay Assisted Systems

In DF relay systems, the relays decode the received message fully or partially and forwards it to appropriate recipients. Although this makes the relaying operation more complex from a computational point of view, it improves the achievable data rate considerably. These relays have been studied extensively as part of network coding in the physical layer [98,99]. The idea of network coding has also been applied in Two Way Relay Networks to improve spectral efficiency and throughput. In [100] and [101], a system model was considered where two users are aided by a single relay node to exchange independent messages.

The power allocation strategies for OFDM transmission using a DF relay was discussed in [102]. In this case, a message transmitted by the source on one subcarrier is transmitted to the destination only if the relay decodes the information successfully. Most of the recent literature assumed the availability of perfect channel state information at the source. For example, [103] considered a general downlink relay system where each relay and the source all has its individual power constraint. However, the assumption is the destination decodes messages either from a relay (after multiple attempts) or from the source (at the first attempt).

It did not consider a case where the destination combines parts of the message received from multiple entities at different time-frequency slots. There is another additional difficulty as each relay uses all the available carriers to forward the messages to the destinations. There is a significant degree of suboptimality in the performance when the relay fails to adapt the available frequencies with the channel fading. Some of the issues were addressed in [104, 105] which proposed a fixed source and novel relay subcarrier pairing approach. [105] considered the power allocation problem with two different sets of constraints. First it considered different constraints on the source and the relay each. Then a total power constraint was considered with the assumption that these two nodes can share power with each other.

Further, in [106, 107] both power allocation and subcarrier pairing were considered with an alternative approach. [106], for example, sorted the subcarriers for source-relay link and relay-destination link. The source relay subcarriers and relay destination subcarriers with the same rank after sorting are paired together for efficient transmission. It became known as the sorted channel pairing which was shown to be optimal in [107]. However, these works usually used approximations to relax the constraints and make the optimisation solvable so that it is possible to arrive at a closed form solution. [66] investigated the same issue without the approximation. To be noted, traditional AF and DF relays have used either time or frequency division duplexing to transmit and receive information, a paradigm which is known as Half Duplex (HD) relays. However, for better utilisation of the degrees of freedom available, Full Duplex (FD) relays have been proposed in literature as an alternative. Many advances in FD relay systems have multiplied the achievable throughput gain and diversity gain to overcome the diversity-multiplexing trade off [36] for HD systems. The principal reason for the gain is the near doubling of channel capacity compared to HD channels. In addition, FD relays can employ a larger buffer size which aids in the reduction of packet loss ratios, which was shown in a comprehensive comparative study of FD versus HD in [108], although it was achieved with a trade off for the complexity. Traditional systems have preferred HD systems over FD for two principal reasons.

- The theoretical rise of capacity to twofold its original value is significantly impaired in the presence of self interference [109].
- HD systems are more attractive for low complexity implementations which do not allow for sophisticated processing of incoming signals [110].

2.5 Relay Interference Management via Game Theoretic Techniques

In many of the recent literature, Game theoretic analyses played an important role in frequency reuse and resource allocation. An efficient deployment of relays or base stations require addressing technical challenges for the uplink and downlink directions. Before introducing the interference management aspects, we will describe the basics of a non-cooperative game.

2.5.1 Basics of a Non-Cooperative Game

From a high level of abstraction, a game $\mathcal{G}$ is a triplet $[\mathcal{K}, \mathcal{S}_k, \{u_k\}]$ where

- $\mathcal{K} \equiv$ set of players,
- $\mathcal{S}_k \equiv$ set of possible strategies for player k and
- $u_k \equiv$ Utility function of player k .

To be noted, the utility gained by a player depends on the strategies of all players and hence, changes in strategies adapted by any player can start a dynamic chain which changes the outcome of every player. Based on this consideration, the Best Response Dynamics (BRD) for a game $\mathcal{G}$ can be defined as

$$\max_{s_k \in S_k} u_k(s_k, \mathbf{s}_{-k}) \quad \text{where } 1 \leq k \leq K. \quad (2.69)$$

Here $\mathbf{s}_{-k}$ denotes the collection of all strategies except player k . Based on BRD, the Nash Equilibrium (NE) is defined as

Definition 2.5 *Let*

$$(s_1, s_2, \dots, s_K) \in S_1 \times S_2 \times \dots \times S_K \quad (2.70)$$

represent a tuple of K strategies by the players. The point $(s_1, s_2, \dots, s_K)$ is a Nash Equilibrium iff

$$u_k(s_k, \mathbf{s}_{-k}) \geq u_k(s_k^*, \mathbf{s}_{-k})$$

when $s_k^ \neq s_k$ and $1 \leq k \leq K$.*

Hence, from the definition, an NE is a fixed point of a BRD, even though it does not guarantee the convergence of a game towards an NE.

2.5.2 Interference Channel as a Non Cooperative Game

[111] proposed modelling a multi-user interference channel by a non-cooperative game where every user seeks to maximise its own throughput (utility). It assumed K transceiver pairs are coordinated by an Amplify-forward relay. As a matter of practical hardware constraint, the relay normalises each received symbol by the square root of its power. Based on this system model, the energy efficiency (EE) was defined as

$$EE = \frac{\text{Throughput as a function of SINR}}{\text{Circuit power consumption} + \text{Transmit power}} \quad (2.71)$$

This quantity captures the tradeoff of achieving a very high data rate at the cost of minimum power, also analysed in [112]. In the solution proposed by [111], the game theoretic approach has been useful because there is a trade-off in the achievable rates of the users. If we consider each user as an autonomous agent motivated to get maximum throughput, the scenario is perfectly modelled by a non-cooperative game. To conclude, although most of the recent literature on game theoretic interference management, the principal theme has been the formulation of the utility function, which can be either throughput, or SINR, energy efficiency or bit error rate.

2.6 Interference Management in Multicell Cooperation

Traditional cellular systems depend on partitioning a broad coverage area into different cells so that every user equipment is served by only one base station. However, to meet the challenges of achieving higher data rate with low power devices, multicell cooperation is supposed to play a major role in next generation cellular networks [18–20]. In particular, the research community's interest in multicell networks is also led by wider adaptation of small cell networks where a number of small cell base stations are located in close proximity with each other. Hence, an efficient cooperation scheme among the base stations is capable of providing the full advantage

of small cells in terms of power consumption and achievable throughput. Upcoming cellular systems such as LTE are designed such that the macro cells can offload some of the users onto small cells. Therefore, the cellular standards such as 3GPP-LTE are supposed to incorporate different elements under the common umbrella term Coordinated Multipoint (CoMP), which is an active area of investigation [17].

When the transmitters cooperate jointly to mitigate the effect of Inter-cell Interference (ICI), it calls for an efficient backhaul network to provide the necessary overhead for the cooperation. In addition to the improved data rate, the coordination among base stations enables more accurate tracking and measurement of network utility performance. Although many CoMP systems can be characterised by a general MIMO system, complete characterisation of performance boundary for such a channel is still an open problem. Based on the different cooperation paradigms proposed in literature, the works on downlink multicell cooperation or CoMP can be classified into three broad categories.

- *Coordinated Beamforming (CBF)*: The data intended for a user equipment is available to only one fixed base station. However, the base stations cooperate to minimise the interference with the limited information available [113].
- *Joint Processing (JP)*: Each user can receive data from multiple base stations, where the base stations cooperate to transmit to all the users at the maximum data rate achievable [22].
- *Spatial Sensing (SS)*: Here each user is served by a single base station, similar to CBF. However, unlike CBF, the user-base station association adapts continuously to variable channel conditions to achieve optimal performance [22].

The following sections discuss these strategies in more detail.

2.6.1 Coordinated Beamforming among Base Stations

In this case, each base station is aware of the downlink channel information and the data to be sent to users inside the same cell. Although the multicell MIMO channel is analogous to an interference channel, the channel information sharing among the receivers makes it possible to reduce the intercell interference. A number of optimisation algorithms in CBF have been proposed in literature. Examples are [21, 114, 115], where the interference alignment approach dominates the work.

For multi-antenna receivers, transceiver algorithms at the cooperating nodes call for joint optimisation to optimise the sum capacity. [21] proposed a non-iterative interference aware algorithm applicable to a two cell system. It can be proved through an asymptotic closed form expression of the sum rate. It was shown that with increasing number of antennas, the sum rate approaches point-to-point upper bound.

2.6.2 Joint Processing with Perfect Data Sharing

In this case, each user receives transmitted data from multiple base stations. So the downlink information stream as well as the channel states must be available to all base stations via inter-transmitter cooperation. In the case of ideal cooperation among the base stations, the transmit side can be modelled by a distributed super base station which coordinate together to serve the users. The capacity of such a channel has been analysed in [23].

The idea of eliminating ICI is similar to spatial division multiple access, performed by

an optimal beamforming matrix at the multiple antenna transmitters. If the receiver has a single antenna, it can perform a simple decoding using the maximum a-posteriori principle. Otherwise, it can employ receive beamforming to further optimise the performance. Although linear precoding (as discussed in the context of broadcast channels) is one of the most popular solutions in this context, non-linear methods such as vector perturbation [116] or Tomlinson Harashima precoding [117] have been applied as well.

2.6.3 Spatial Sensing

Spatial sensing as a cooperation paradigm has been investigated particularly in cognitive radio frameworks. They make more efficient spectrum utilisation by allowing a secondary network to coexist with a primary network [118]. Thus cognitive networks are able to exploit the spatial domain more efficiently using multiple antennas at the cognitive transceivers [24, 25]. The same technique can be applied in multicellular communication where some base stations *sense* the spectrum utilisations and transmit patterns of other base stations to optimise their own transmit strategy.

[24] proposed a convex optimisation technique to maximise the throughput of the secondary link subject to constraints on interference to primary link. The same work also introduced a computationally simple low complexity algorithm to achieve the same. [25] proposed a multi-cast beamforming method for Multiple Input Single Output (MISO) cognitive radio networks. However, most of these users ignored the interference from primary or high priority base stations or treated the respective signals as noise. [114] considered the effect of primary network interference on secondary networks more thoroughly. This motivated further research into joint transceiver designs with two objectives [116] such as

- Prevent primary networks from experiencing any interference from cognitive networks,
- Cancel the primary users' interference to cognitive networks via innovative beamforming techniques.

Also, [119] and [120] proposed methods of how the secondary networks can overcome the effects of primary networks. In particular, [120] studied the effects of primary network interference on cognitive receivers and proposed a filter based approach in cognitive transceivers to cancel the interference. [119] proposed an alternative linear transceiver design method where the primary users have a single antenna each and the cognitive users have two antennas each.

2.7 Summary and Shortcomings of Previous Work

The present review discusses the topic of interference management and how interference affects the achievable rate vectors in relay assisted communication channels and multicell cooperative communication systems. The relays can enhance the data rates among non-cooperating transmitter-receiver pairs by introducing cooperation. Successful interference cancellation methods such as SIC and DPC are possible through a fruitful cooperation of transmitter and/or receiver sides.

Among the major interference management paradigms—viz. interference cancellation, interference ignorance and interference alignment, it is proven that the interference cancellation gives best possible performance at the cost of computational complexity. Interference alignment scheme nullifies the effect of interference with a subspace decoding approach while it enhances the noise. The noise enhancement is tolerable for high SNR scenario, whereas the noise power

can be negligible compared to the signal and interference power.

Subsequently, the trade off for performance in multiuser rate regions is introduced along with the effect of interference on it. For a MAC with individual users' power constraints, the outer boundary is a simplex of points obtained by different successive decoding orders. This does not apply to a BC, for which an additional degree of freedom exists on deciding the power allocation. The present research exploits the inherent duality principles to correlate the BC and MAC systems in terms of resource usage and rate region.

In a multiple access channel, the transmitters are spread across a geographically wide coverage area and all of them transmit to a single node. This information theoretic model can be applicable to two major practical deployments.

- A cellular base station which receives datastreams from users spread across a coverage area.
- A group of sensors transmitting information to a centralised data gathering node for further processing [121].

In both cases, it is assumed that there is no cooperation among the transmitters. Hence, each user can be supposed to have his own battery with individual transmit power constraint. Under the sum power constraint, the optimal rate region boundary can be where the sum rate is maximum and the only variable to control the user rates is the decoding orders. Different decoding orders of user give different data rates even though the sum rate remains the same. However, if the multiple access channel is being considered as a dual of a broadcast network, it is reasonable to assume that the transmitters draw power from a common source and hence they are constrained by a sum power. In that case, it is shown that a specific power allocation maximises the sum rate. The sum rate maximisation is a convex optimisation problem whose solution can be found by standard optimisation methods like interior point method [34]. However, solving the problem by convex methods hides some interesting properties of the system about how the different transmitters can cooperate in terms of power consumption to tolerate interference from each other. These properties can be exploited to formulate an efficient alternative algorithm which gives better insight into the problem at hand and also solves the problem with less time requirement for reasonable problem size. The simulation results support the analyses. Further, the algorithm is self-contained in the sense that it provides a metric to test the suboptimality of any tentative feasible point and hence can be used to formulate a stopping criterion.

Finally, the interference management techniques are discussed in the context of different categories of relay channel and cooperative multicell environments. The ideas of heterogeneous networks and multicell cooperation evolved to meet the ever increasing demand for higher data rate in present day communication systems. Some of the shortcomings of recent works which motivated this research are presented as follows.

- Some special structure of the sum rate capacity (which is a logarithmic sum function) has not been exploited in earlier literature, which treated it as a special case of convex optimisation. One goal of the thesis has been to propose novel methods exploiting the problem feature.
- Optimisation of a specific objective function is not sufficient to characterise the Pareto boundary of a system under resource constraint. The boundary is affected by user encoding/decoding and different types of constraints. A unified coverage of these system models have been absent in literature.

- While the duality principle has been an influential result to motivate several optimisation techniques in single cell system, the same principle has been absent from multicell systems so far. It is expected that a generalised formulation applicable to multicell systems can simply more complicated cooperative resource allocation problems.

2.7.1 Scope of this Thesis

Based on these observations the important goals of this work are twofold.

- Complement the existing works on different optimisation strategies for single cell uplink-downlink systems to enhance the algorithmic efficiency and establish some novel methods to achieve the rate region boundaries.
- Based on the novel formulations of rate region boundaries and sum rate capacities, the results are extended to more generalised category of heterogeneous and cooperative networks. In particular, a heterogeneous network has been modelled by a DF relay system, which helps us to see the interference mitigation and resource allocation problems from a new perspective. Based on the works of single cell systems, a new cooperation paradigm is introduced for multicell systems. The novel technique called Partial Data Cooperation achieves an optimal trade off between Coordinated Beamforming and Joint Processing from an operational point of view. The new multicell system also obeys the duality principle hitherto known only for single cell systems. This offers significant advantage in formulation and solution of the relevant optimisation problems.

2.8 Conclusion

The chapter serves as a review on interference management and the relevant optimisation techniques employed in different communication networks. The concepts of rate region boundaries in different systems are discussed initially. This is followed by a survey of literature on mitigating the effects of interference in multiuser communication systems. Effective interference management techniques are crucial to improve throughput gain and meet the demands for wireless service. The wireless communication landscape has been shifting to a greater degree of transceiver cooperation by exploiting the greater connectivity aspects due to exponential increase in wireless devices. The applications of earlier known techniques to the cooperative framework come with some unique optimisation challenges. Hence, proposing novel optimal solutions for single cell systems and their eventual extension to heterogeneous and multicell systems within a cooperative framework will form the most important aspects of the present research.

Chapter 3

Throughput Maximisation and Rate Region in Single Cell Systems

DUALITY relations among achievable QoS parameters help us transform the problems in the Broadcast Channel (BC) to problems in Multiple Access Channel (MAC). The QoS optimisation problems in a BC frequently turn out to be non-convex problems which are difficult to solve using interior point methods [34]. The same problems in a MAC are convex problems for which efficient algorithms exist [43] to solve the problems.

It was discussed in the previous chapter that the MAC and BC rate regions are equivalent, as predicted by the duality principle. However, when a MAC is considered as a dual of a BC, it is necessary to assume a virtual cooperation among users in terms of power consumption. Even though in reality, the users transmit data independently, it is assumed that they draw power from a common source, as in the case of a BC. In other words, power can be transferred from one user to another in order to achieve a different operating point. The reason to make this assumption is, the resource allocation problems are difficult to address directly in BC because they are typically not convex. Hence, for the dual MAC, it is necessary to consider a sum power constraint in the form of $\sum p_i \leq P$, where P is the total base station power.

To be noted, when a single cell MAC is considered with individual power constraints on the users, the relevant resource allocation problems have been explored in [48, 122]. It was shown in these two papers that the achievable rate region is a polymatroid, along with developing some original algorithms. Some properties of the polymatroid formulations will be reviewed later. Also, the algorithms proposed in this chapter and later ones make use of these previous results.

This chapter considers the Shannon's capacity as a measure of the throughput and proposes an algorithm for optimising the sum capacity (also referred to as sum rate) of a MAC under sum power constraint, where the users are coupled by individual power usage as well as interference. The optimal resource allocation has a distributed feature which can be exploited to formulate the algorithm. If the users are partitioned into different sets and each set maximises its own rate, treating the signal for the other set as noise, the sum rate of all users is maximised. Additionally, the sum rate of users has an interesting relationship with the generalised channel norms. Power allocation problem for any number of users is solved by dealing with only two users at a time. The algorithm gives useful insights about how the decoding priority affects achievable rates of different users.

Also, it is shown that the sum rate optimisation problem has a mathematical relationship with the formal characterisation of the rate region boundary in a multi user system, formally called as the Pareto frontier. Achieving the rate region Pareto frontier depends on the encoding/decoding order among users. Hence, the chapter proposes two algorithms to find the optimal decoding order(s) among users to achieve the Pareto frontier.

The chapter is organised as follows. Section 3.1 discusses the system model and the problem formulations. Section 3.2 discusses some properties of the optimal power allocation problem and its relation to decoding order among the users, which plays an important role to find the achievable rate region. Section 3.3 proposes the algorithms to find the optimal power allocation to achieve the maximum sum rate. Section 3.4 characterises the broader rate region boundary in terms of Pareto frontier for which the sum rate maximality is a special case. Based on this characterisation, Section 3.5 proposes two algorithms to achieve the Pareto frontier with any given power allocation. Section 3.6 gives some numerical results to compare the proposed norm-balancing algorithm with standard convex optimisation algorithms. Finally, Section 3.7 concludes the chapter.

3.1 The System Model

The multiple access system of K single antenna users transmitting to a base station of M antennas is already given in Figure 2.3. The channel from user k to base station is given by column vectors $\mathbf{h}_k \in \mathbb{C}^M$, where $1 \leq k \leq K$. A flat and slow fading channel is assumed where $\mathbf{h}_k$ represents channel coefficients accurately tracked by the receiver via pilot sequences. In a multiple access system, each transmitter is supposed to have its own power constraint because a collective power constraint on non-cooperating transmitters across a coverage area is not important or possible to implement. However, when the multiple access system is considered as dual of a broadcast system, the sum power constraint is more important compared to individual power constraints. For the current system model, the received signal is given by

$$\mathbf{y} = \sum_{i=1}^K \mathbf{h}_i x_i + \boldsymbol{\eta} \quad (3.1)$$

where x_i is the symbol transmitted by user i and $\boldsymbol{\eta}$ is the zero mean Gaussian additive noise vector with $\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^H] = \mathbf{I}_M$. Power constraints of users imply that $\mathbb{E}[|x_i|^2] = p_i$ for user i . If $\mathcal{U}$ is the set of all users, the achievable rate region with successive interference cancellation [123]

$$\mathcal{C}_{\text{MAC}}(\mathbf{p}) = \{\mathbf{r} | 0 \leq \sum_{i \in \mathcal{S}} r_i \leq \log |\mathbf{I} + \sum_{i \in \mathcal{S}} p_i \mathbf{h}_i \mathbf{h}_i^H| \text{ where } \mathcal{S} \subseteq \mathcal{U}\} \subset \mathbb{R}_+^K \quad (3.2)$$

For a BC with base station power constraint P and channels $\mathbf{h}_k^H$, from the duality principle [34], the capacity region can be obtained as

$$\mathcal{C}_{\text{BC}}(P) = \bigcup_{\sum p_i = P} \mathcal{C}_{\text{MAC}}(\mathbf{p}) \quad (3.3)$$

3.1.1 Sum Rate Maximisation

From the capacity region given in (3.2), it is always best to use the system at its maximum sum rate (see Theorem 3.1). However, when the power constraint is on sum power rather than individual user powers, the sum rate is a function of $\mathbf{p}$. This chapter is concerned with maximising the sum rate for a given sum power, viz.

$$\max_{\mathbf{p}} \log |\mathbf{I} + \sum_{i \in \mathcal{U}} p_i \mathbf{h}_i \mathbf{h}_i^H| \quad (3.4)$$

subject to the convex constraints

$$\mathbf{p} \succeq \mathbf{0} \quad \text{and} \quad \sum_{i=1}^K p_i = P \quad (3.5)$$

3.1.2 Pareto Frontier

Maximising the sum rate as described in (3.4) does not completely characterise the entire rate region boundary, formally called as the Pareto frontier as defined below.

Definition 3.1 *For an achievable rate vector $\mathbf{r} \in \mathbb{R}_+^K$, with given sum power, if there exists no other achievable rate vector $\mathbf{r}' \in \mathbb{R}_+^K$ satisfying $\mathbf{r} \preceq \mathbf{r}'$, then $\mathbf{r}$ is a Pareto optimal point¹ and the set of all such points is the Pareto frontier.*

For any rate vector $\mathbf{r} \in \mathbb{R}_+^K$ achievable with a total power P , if $\sum r_i$ is at its maximum, obviously there cannot be any other point $\mathbf{r}' \succeq \mathbf{r}$, since it will violate the sum rate optimality constraint by $\sum r'_i > \sum r_i$. However, there can be other rate vectors for which the sum rate is less than $\sum r_i$ which are still located on the Pareto frontier in the sense that it is not possible to improve on any user.

A novel sum rate maximisation algorithm and a characterisation of the Pareto frontier based on its relationship with the sum rate optimisation problem are the important contributions of this chapter. The significance of decoding order on the achievable rate is considered to arrive at two optimal decoding algorithms. It is important to mention that similar issue of Pareto optimality was considered from a game theoretic viewpoint in [124, 125]. This work considers a BC with interference pre-subtraction/cancellation scheme.

3.2 Characterisation of Sum Rate Optimal Solution

Since the sum rate of a set of users is given by (3.2), let us denote the sum rate matrix of all users by $(\mathbf{I} + \sum_{i \in \mathcal{U}} p_i \mathbf{h}_i \mathbf{h}_i^H)$, because the determinant of this matrix gives us the sum rate of all users. It can be shown that the matrix is an affine function of $\mathbf{p}$ and because of concavity of logarithm, the problem of (3.4) is a concave maximisation over a convex feasible region. This problem has been solved as special case of more general linear constrained problems in [34]. However, in this chapter the problem will be solved using its convenient structure and thus partially characterising the Pareto frontier for sum power constraint. The major advantages of the proposed approach are as follows.

¹ $\preceq$ will denote tuple-wise inequality among vectors or positive definite property of a matrix depending on the context.

- The norm balancing technique performs better than standard convex optimisation techniques when the number of user in each cell is lower. This has been quantified later via simulation results. However, the advent of heterogenous systems [126] is expected to accelerate the trend towards smaller coverage areas, which underlines the relevance of our proposed norm balancing approach.
- The analyses of the algorithms can be used to judge how close a specific power allocation is to the optimal distribution, simply via calculating the relevant channel norms, as shown later.
- The algorithm also aids in relating the problem of determining Pareto frontier with the sum rate problem and helps characterise the rate region boundary. This will be exploited to formulate the algorithms to determine the Pareto frontier in Section 3.5.

To exploit the structure of the optimisation problem and find the Pareto frontier, the concept of set function is useful which is defined below.

Definition 3.2 For a set of users $\mathcal{S} \subseteq \mathcal{U}$ and power allocation $\mathbf{p}$, define the set function

$$f(\mathcal{S}, \mathbf{p}) \equiv \log |\mathbf{I} + \sum_{i \in \mathcal{S}} p_i \mathbf{h}_i \mathbf{h}_i^H| \quad (3.6)$$

From the rate region given in (3.2), it is easy to check that $f(\mathcal{S}, \mathbf{p})$ gives the sum rate upper bound of a set of users and it depends only on $\{p_i\}_{i \in \mathcal{S}}$. Specifically, if all the users in $\mathcal{S}'$ have been decoded before decoding any of the users in $\mathcal{S}$, then the interference from $\mathcal{S}'$ is already known and subtracted from received signal. Hence, for any $i \in \mathcal{S}$, the achievable rate is independent of $\{p_i\}_{i \in \mathcal{S}'}$.

The function $f(\mathcal{S}, \mathbf{p})$ is an example of submodular set functions [127] and the corresponding region is a polymatroid [48]². Theorem 3.1 and Lemma 3.1 will characterise the Pareto frontier for a specific power allocation vector $\mathbf{p}$.

Theorem 3.1 For any given power allocation vector, the points where the sum rate constraint is satisfied with equality are Pareto optimal points and those are the only Pareto optimal points, i.e. at all points on the Pareto frontier, $\sum_{i \in \mathcal{U}} r_i = f(\mathcal{U}, \mathbf{p})$.

Proof: Since $\mathbf{p}$ remains constant throughout the context of this theorem, that argument of the submodular function will be dropped for brevity.

Proving the direct part is trivial. If for an achievable rate vector $\mathbf{r}$, the sum rate constraint for all user is satisfied with equality, i.e. $\sum_{i=1}^K r_i = f(\mathcal{U})$, then obviously $\mathbf{r}$ is Pareto optimal.

To prove the converse, let $\mathcal{S}$ and $\mathcal{T}$ be two distinct non-empty sets of users. So $\mathcal{U} \supset \mathcal{S} \neq \emptyset$ and $\mathcal{U} \supset \mathcal{T} \neq \emptyset$. Also assume that at an achievable rate vector $\mathbf{r}$ the sum rate constraints of these two sets are satisfied with equality. Hence,

$$\sum_{i \in \mathcal{S}} r_i = f(\mathcal{S}) \quad \text{and} \quad \sum_{i \in \mathcal{T}} r_i = f(\mathcal{T}) \quad (3.7)$$

²For basics of these concepts, see Appendix B.

We have

$$\begin{aligned} \sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i &= \sum_{i \in \mathcal{S}} r_i + \sum_{i \in \mathcal{T}} r_i - \sum_{i \in \mathcal{S} \cap \mathcal{T}} r_i \\ &= f(\mathcal{S}) + f(\mathcal{T}) - \sum_{i \in \mathcal{S} \cap \mathcal{T}} r_i \end{aligned} \quad (3.8)$$

$$\geq f(\mathcal{S} \cup \mathcal{T}) + \left(f(\mathcal{S} \cap \mathcal{T}) - \sum_{i \in \mathcal{S} \cap \mathcal{T}} r_i \right) \quad (3.9)$$

The submodularity constraint dictates that

$$\sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i \leq f(\mathcal{S} \cup \mathcal{T}) \quad (3.10)$$

Inequalities (3.9) and (3.10) can be satisfied together only if

$$\sum_{i \in \mathcal{S} \cup \mathcal{T}} r_i = f(\mathcal{S} \cup \mathcal{T}) \quad (3.11)$$

and

$$\sum_{i \in \mathcal{S} \cap \mathcal{T}} r_i = f(\mathcal{S} \cap \mathcal{T}) \quad (3.12)$$

Hence, if the maximum sum rate is reached for two different non-empty subsets of users, the sum rate constraints are satisfied with equality for the intersection and for union of those two sets. It is trivial to extend the result to more than two different sets.

Suppose for a rate vector $\mathbf{r}$, none of the sum rate constraints are satisfied with equality. In that case obviously $\mathbf{r}$ is not a Pareto optimal point since it is possible to improve any of the rates marginally without violating any constraint.

To the contrary, suppose for the achievable vector $\mathbf{r}$, the equality constraints are satisfied for N sets $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_N \subseteq \mathcal{U}$.

If $\bigcup_{i=1}^N \mathcal{S}_i \subset \mathcal{U}$, then it is possible to increase data rates of the users $\mathcal{U} - \bigcup_{i=1}^N \mathcal{S}_i$ and the point $\mathbf{r}$ is not Pareto optimal.

If $\bigcup_{i=1}^N \mathcal{S}_i = \mathcal{U}$, the sum rate bound is satisfied for $\mathcal{U}$. ■

Lemma 3.1 will give a lower bound on the sum rates as a necessary and sufficient condition for the Pareto frontier of the polymatroid.

Lemma 3.1 *For any non-empty set of users $\emptyset \neq \mathcal{S} \subseteq \mathcal{U}$, the following condition is satisfied at all points on the Pareto frontier.*

$$f(\mathcal{U}, \mathbf{p}) - f(\mathcal{S}', \mathbf{p}) \leq \sum_{i \in \mathcal{S}} r_i \leq f(\mathcal{S}, \mathbf{p}) \quad (3.13)$$

Conversely, if the condition is satisfied for all $\mathcal{S} \subseteq \mathcal{U}$, the point is Pareto optimal.

Proof: Proof of the converse is trivial. If the condition is satisfied for all sets of users, it is satisfied for $\mathcal{S} = \mathcal{U}$ and hence from Theorem 3.1, the point is Pareto optimal.

The upper bounds for every set $\mathcal{S}$ follow directly from the rate region defined in (3.2).

To prove the direct statement, consider the case where

$$\begin{aligned}
 \sum_{i \in \mathcal{S}} r_i &< f(\mathcal{U}, \mathbf{p}) - f(\mathcal{S}', \mathbf{p}) \\
 \Rightarrow f(\mathcal{S}', \mathbf{p}) + \sum_{i \in \mathcal{S}} r_i &< f(\mathcal{U}, \mathbf{p}) \\
 \Rightarrow \sum_{i \in \mathcal{U}} r_i = \sum_{i \in \mathcal{S}'} r_i + \sum_{i \in \mathcal{S}} r_i &\leq f(\mathcal{S}', \mathbf{p}) + \sum_{i \in \mathcal{S}} r_i < f(\mathcal{U}, \mathbf{p})
 \end{aligned} \tag{3.14}$$

From Theorem 3.1, the point is not on Pareto frontier. ■

To be noted, upper bounds in Lemma 3.1 do not depend on the powers $\{p_i\}_{i \notin \mathcal{S}}$. This is consistent with the decoding priority rule set out in [32]. If successive interference cancellation is used, the set $\mathcal{S}$ achieves the upper bound for decoding order³ $\mathcal{S}' \rightarrow \mathcal{S}$ and lower bound for decoding order $\mathcal{S} \rightarrow \mathcal{S}'$. For different power allocations $\mathbf{p}$, the Pareto frontiers are convex segments on parallel hyper planes in $\mathbb{R}_+^K$ satisfying $\sum_{i=1}^K r_i = c > 0$.

Hence, the problem (3.4) can be formulated as

$$\max_{\mathbf{p} \succeq \mathbf{0}, \|\mathbf{p}\|_1 = P} f(\mathcal{U}, \mathbf{p}) \tag{3.15}$$

It can be geometrically interpreted as finding the hyperplane at maximal separation from the origin. Since it is a convex problem, a local optimum will be the global optimal solution. The results in the next sections will give some properties of the optimal solution $\mathbf{p}^* \equiv [p_1^*, p_2^*, \dots, p_K^*] \in \mathbb{R}_+^K$.

3.2.1 Channel Norm Relationship for Two Users

To characterise the sum rate optimal power allocation for a general scenario of K users, it is necessary to characterise the solution for two users. The characterisation for a pair of users will aid in the more general case of multiple users.

Theorem 3.2 *For two users with a sum power constraint P and possible interference from other users, having same generalised channel norm is a sufficient condition for sum rate optimality.*

Proof: Let the interference plus noise matrix be denoted by $\mathbf{S}$ and power allocated to user one be x . Then the sum rate is

$$r_1 + r_2 = \log |\mathbf{S} + x \mathbf{h}_1 \mathbf{h}_1^H + (P - x) \mathbf{h}_2 \mathbf{h}_2^H| - \log |\mathbf{S}| \quad \text{where } x \in [0, P] \tag{3.16}$$

Taking derivative,

$$\frac{d(r_1 + r_2)}{dx} = \mathbf{h}_1^H \mathbf{A} \mathbf{h}_1 - \mathbf{h}_2^H \mathbf{A} \mathbf{h}_2 \tag{3.17}$$

³This means all the elements of $\mathcal{S}'$ are decoded before the elements of $\mathcal{S}$

Figure 3.1: Variation of Channel Norm when Sum Rate Passes through a Maximum

where the inverse sum rate matrix⁴

$$\mathbf{A} \equiv (\mathbf{S} + x\mathbf{h}_1\mathbf{h}_1^H + (P - x)\mathbf{h}_2\mathbf{h}_2^H)^{-1} \succ 0 \quad (3.18)$$

Now define the generalised inner product in $\mathbb{C}^M$ as

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{A}} \equiv \mathbf{u}^H \mathbf{A} \mathbf{v} \quad \text{where } \mathbf{u}, \mathbf{v} \in \mathbb{C}^M \quad (3.19)$$

which satisfies the sesquilinear and positive definite properties⁵ for $\mathbf{A} \succ 0$. The corresponding norm defined on the same vector space is given by

$$\|\mathbf{v}\|_{\mathbf{A}} \equiv \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{A}}} > 0 \quad \text{where } \mathbf{v} \neq \mathbf{0} \quad (3.20)$$

Since it is a one dimensional concave function, stationarity condition implies the global maximum. Then equating (3.17) to zero, it can be put succinctly as a sufficient condition $\|\mathbf{h}_1\|_{\mathbf{A}} = \|\mathbf{h}_2\|_{\mathbf{A}}$. ■

It is illustrated in Figure 3.1 where the horizontal axis shows the fraction of total power for user one, the vertical axis shows sum rate and channel norms of users (numerical values have been suitably scaled). The plot is generated using Rayleigh fading channels for both users. That means each element is taken as Rayleigh fading coefficient. So $\mathbf{H}_{ij} \sim \mathcal{CN}(0, 1)$. A BS of 5 antennas and total power of 30dBm is considered. The sum rate reaches its peak when the curves denoting channel norms intersect with each other. Although the exact values of the sum rate and norms will depend on the channel variables, the shapes of the curves correspond with the analyses earlier. Note that the above theorem can be generalised to multiple users when the conditions are satisfied pairwise for any two users. This will be shown later in Theorem 3.3.

Lemma 3.2 *The difference between squares of norms for two user scenario is a monotonic function of either user's power allocation.*

⁴It is trivial to prove the matrix is positive definite for $x \in [0, P]$.

⁵See Appendix A for a review of the mathematical preliminaries.

Figure 3.2: Variation of Channel Norm when Sum Rate is Monotonic

Proof: For $p_1 = x$ and $p_2 = P - x$, with some algebraic manipulation it can be shown that

$$\begin{aligned} \frac{d \left(\|\mathbf{h}_2\|_{\mathbf{A}}^2 - \|\mathbf{h}_1\|_{\mathbf{A}}^2 \right)}{dx} &= \|\mathbf{h}_1\|_{\mathbf{A}}^4 + \|\mathbf{h}_2\|_{\mathbf{A}}^4 - 2|\langle \mathbf{h}_1, \mathbf{h}_2 \rangle_{\mathbf{A}}|^2 \\ &\geq \|\mathbf{h}_1\|_{\mathbf{A}}^4 + \|\mathbf{h}_2\|_{\mathbf{A}}^4 - 2\|\mathbf{h}_2\|_{\mathbf{A}}^2 \|\mathbf{h}_1\|_{\mathbf{A}}^2 \\ &= \left(\|\mathbf{h}_1\|_{\mathbf{A}}^2 - \|\mathbf{h}_2\|_{\mathbf{A}}^2 \right)^2 \geq 0 \end{aligned} \quad (3.21)$$

using Cauchy Schwarz inequality (proved in Appendix A). ■

Lemma 3.2 shows that (if any) there can be only one zero crossing of the norm difference, shown again in Figure 3.1 for a Rayleigh faded channel. However, the condition of Theorem 3.2 is not necessary when there is no zero crossing. In such cases, however, one among the two users will be completely deprived to attain the sum rate optimality. This is shown in Lemma 3.3.

Lemma 3.3 *If the channel norms are not equal at the optimal distribution, that implies the available power is allocated to the user with greater channel norm.*

Proof: Absence of zero-crossing implies that $\|\mathbf{h}_1\|_{\mathbf{A}}^2 - \|\mathbf{h}_2\|_{\mathbf{A}}^2$ does not change sign. Without loss of generality, we can assume that

$$\|\mathbf{h}_1\|_{\mathbf{A}}^2 > \|\mathbf{h}_2\|_{\mathbf{A}}^2 \quad \text{where } x \in [0, P] \quad (3.22)$$

From (3.17), $r_1 + r_2$ becomes a monotonically increasing function of x and hence maximum is attained at $x = P$. ■

This scenario has been illustrated in Figure 3.2 where the sum rate changes monotonically with power allocation of any user and for maximum sumrate, one of the users will be deprived. It is seen that the curves denoting the channel norms do not intersect and zero crossing never occurs.

3.2.2 Norm Relationship for Multiple Users

Based on the norm equality for two users encountered in Section 3.2.1, it is possible to generalise it to multiple users. Hence, Theorem 3.3 will be proven to pave the way for the optimisation algorithms. A similar result was shown for a multiuser case in [59, Theorem 1]. However, it was shown there for optimisation of co-variance matrices with individual power constraints in a BC MIMO channel. Theorem 3.3 proves it with a sum power constraint for single antenna users.

Theorem 3.3 *At the optimal power allocation $\mathbf{p} = \mathbf{p}^*$ for every non-empty $\mathcal{S} \subseteq \mathcal{U}$, its own power is optimally distributed to maximise the lower bound of Lemma 3.1.*

Conversely, if for all non-empty proper subset $\mathcal{S} \subset \mathcal{U}$, the power distribution is optimal (to maximise the lower bound) given the values of $\{p_i\}_{i \in \mathcal{S}'}$, then the power distribution for $\mathcal{U}$ is optimal under the sum power constraint.

Proof: To prove the direct part, use the lower bound from Lemma 3.1 to get

$$\sum_{i \in \mathcal{S}} r_i \geq f(\mathcal{U}, \mathbf{p}^*) - f(\mathcal{S}', \mathbf{p}^*). \quad (3.23)$$

$f(\mathcal{S}', \mathbf{p}^*)$ is independent of $\{p_i\}_{i \in \mathcal{S}}$. Hence, for given $p_i = p_i^*$ where $i \in \mathcal{S}'$, the choice of $\{p_i\}_{i \in \mathcal{S}}$ to maximise the lower bound of (3.23) is $p_i = p_i^*$ where $i \in \mathcal{S}$.

To prove the converse, for the specific power allocation $\mathbf{p}^* \in \mathbb{R}_+^K$ satisfying the condition, define the sum rate matrix

$$\mathbf{A} \equiv \left(\mathbf{I} + \sum_{i=1}^K p_i^* \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \succ 0 \quad (3.24)$$

Define the set

$$\mathcal{V} \equiv \{i | p_i > 0\} \subseteq \{1, 2, \dots, K\}. \quad (3.25)$$

Thus $\mathcal{V}$ is the set of users being served and $\mathcal{V}' \equiv \mathcal{U} - \mathcal{V}$ is the set of users to whom service is denied⁶.

For any pair of users $1 \leq i, j \leq K$, the power between them is optimally balanced. From Theorem 3.2

$$C \equiv \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i > 0 \quad \text{where } i \in \mathcal{V} \quad (3.26)$$

From Lemma 3.3,

$$\mathbf{h}_i^H \mathbf{A} \mathbf{h}_i > \mathbf{h}_j^H \mathbf{A} \mathbf{h}_j \quad \text{where } i \in \mathcal{V} \quad \text{and } j \in \mathcal{V}' \quad (3.27)$$

For the optimisation problem of (3.4), consider the Lagrangian

$$L(\mathbf{p}, \lambda, \boldsymbol{\mu}) = \log |\mathbf{I} + \sum_{i \in \mathcal{U}} p_i \mathbf{h}_i \mathbf{h}_i^H| + \lambda \left(\sum_{i=1}^K p_i - P \right) + \sum_{i=1}^K \mu_i p_i \quad (3.28)$$

where $\lambda \in \mathbb{R}$ and $\boldsymbol{\mu} \in \mathbb{R}_+^K$. Now if we set $\mathbf{p} = \mathbf{p}^*$, $\lambda = -C$ and $\mu_i = C - \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i$, it is easy to check that $(\mathbf{p}^*, -C, \boldsymbol{\mu})$ is a stationary point of (3.28) and satisfies the following Karush-Kuhn-

⁶As Lemma 3.3 proves, such users may exist.

Tucker conditions laid out in [44].

$$\mathbf{h}_i^H \mathbf{A} \mathbf{h}_i - C + \mu_i = 0 \quad \text{Stationarity from (3.26),} \quad (3.29)$$

$$\mathbf{p}^* \succeq \mathbf{0} \quad \text{and} \quad \sum_{i=1}^K p_i^* = P \quad \text{Primal feasibility,} \quad (3.30)$$

$$\mu_i = C - \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i \geq 0 \quad \text{Dual feasibility from (3.26) and (3.28)} \quad (3.31)$$

and

$$\mu_i p_i^* = 0 \quad \text{where } i \quad \text{Complementary Slackness.} \quad (3.32)$$

Hence, $\mathbf{p}^*$ solves the optimisation problem (3.4). $\blacksquare$

Lemma 3.4 *For the $\mathbf{p}^*$ to be sum rate optimal for any number of users, the norm equality of (3.26) among all non-deprived users is a necessary and sufficient condition.*

Proof: Follows directly from (3.26). $\blacksquare$

Lemma 3.4, which is an extension of Theorem 3.2, gives a simple test to check whether a power allocation $\mathbf{p} \in \mathbb{R}_+^K$ is optimal and even whether there is any set of users $\mathcal{S} \subseteq \mathcal{U}$ for which the power is optimally distributed (against interference from others). If for any number of users the channel norms are equal, then the power distribution among them is optimal. Additionally, Lemma 3.4 and Theorem 3.3 together convert the power allocation problem of Section 3.1.1 to a norm balancing problem.

Define the quantity

$$\Delta(\mathbf{p}) \equiv \max_{i \in \mathcal{V}} \|\mathbf{h}_i\|_{\mathbf{A}} - \min_{i \in \mathcal{V}} \|\mathbf{h}_i\|_{\mathbf{A}} \geq 0 \quad (3.33)$$

where $\mathcal{V}$ is the set of users with non-zero power as defined in (3.25). Clearly, $\Delta = 0$ only at the optimal power allocation $\mathbf{p}^*$ and it gives a measure of *imbalance* among the channel norms. So Δ can be used as a metric of how close a certain power allocation is to the optimal point (without knowledge of the optimal point itself). Moreover, finding the maximally imbalanced pair helps in determining iterative steps as shown next.

3.3 Proposed Norm Balancing Algorithms

Any noise plus interference covariance can be represented by a Hermitian symmetric positive definite matrix. For two users with an arbitrary noise (positive definite) matrix, according to monotonicity property of Lemma 3.2 there can be only a single point satisfying the norm criterion of Lemma 3.4. However, before asserting that, it is necessary to check whether the sum rate increases monotonically. If that is not the case, the zero crossing point can be easily found by Bisection method [128]. Algorithm 3.1 gives the steps in proper order. Once it is possible to solve the power allocation problem for two users, the method can be extended to any number of users by dealing with two *maximally imbalanced* users at a time as proposed in Algorithm 3.2.

Here $t > 0$ is a threshold which gives a stopping criterion. For smaller value of t , it is possible to approach closer to the optimal point but with more number of iterations.

Lemma 3.5 *Algorithm 3.2 produces a sequence of feasible power allocations (satisfying the sum*

Algorithm 3.1: Optimal Power Allocation Between Two Users with Noise and Interference

Require: $P > 0$, $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^M$ and $\mathbf{S}_z \succ 0 \in \mathbb{C}^{M \times M}$

- 1: **if** $\mathbf{h}_1^H (\mathbf{S}_z + P\mathbf{h}_1\mathbf{h}_1^H)^{-1} \mathbf{h}_1 > \mathbf{h}_2^H (\mathbf{S}_z + P\mathbf{h}_1\mathbf{h}_1^H)^{-1} \mathbf{h}_2$ **then**
- 2: $p_1 \leftarrow P, p_2 \leftarrow 0$
- 3: **else if**

$$\mathbf{h}_1^H (\mathbf{S}_z + P\mathbf{h}_2\mathbf{h}_2^H)^{-1} \mathbf{h}_1 < \mathbf{h}_2^H (\mathbf{S}_z + P\mathbf{h}_2\mathbf{h}_2^H)^{-1} \mathbf{h}_2$$
- then**
- 4: $p_2 \leftarrow P, p_1 \leftarrow 0$
- 5: **else**
- 6: Use Bisection method to find zero crossing.
- 7: **end if**

Algorithm 3.2: To Calculate Optimal Power Allocation Among Multiple Users

Require: $P > 0$ and $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K \in \mathbb{C}^M$

- 1: $p_i \leftarrow \frac{P}{K}$ where $i \in \{1, \dots, K\}$ {Initialise with equal power allocation}
- 2: **while** $\Delta > t$ **do**
- 3: $\mathcal{V} \leftarrow \{i | p_i > 0\}$, $\mathbf{A} \leftarrow (\mathbf{I} + \sum_{i \in \mathcal{V}} p_i \mathbf{h}_i \mathbf{h}_i^H)^{-1}$
- 4: $k \leftarrow \arg\max_{i \in \mathcal{V}} \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i$, $l \leftarrow \arg\min_{i \in \mathcal{V}} \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i$
- 5: Optimise power allocation between users k and l with sum power constraint $p_k + p_l$ using Algorithm 3.1.
- 6: Update $p_k, p_l, \mathbf{A}, \mathcal{V}$
- 7: $\Delta \leftarrow \max_{i \in \mathcal{V}} \|\mathbf{h}_i\|_{\mathbf{A}} - \min_{i \in \mathcal{V}} \|\mathbf{h}_i\|_{\mathbf{A}}$
- 8: **end while**

power constraint) with monotonically increasing sum rate with each iteration.

Proof: The algorithm begins with $p_i = P/K$ and with each iteration, it redistributes the powers of two users keeping the other powers constant. Hence, after n number of iterations $\sum_{i=1}^K p_i^{(n)} = P$.

If the set of users $\mathcal{S}$ is decoded after $\mathcal{S}'$ to achieve the lower bound of Lemma 3.1, the sum rate is

$$\begin{aligned} \sum_{i \in \mathcal{S}} r_i &= \log |\mathbf{I} + \left(\mathbf{I} + \sum_{i \in \mathcal{S}'} p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \sum_{i \in \mathcal{S}} p_i \mathbf{h}_i \mathbf{h}_i^H| \\ &= \log |\mathbf{I} + \sum_{i \in \mathcal{S}'} p_i \mathbf{h}_i \mathbf{h}_i^H + \sum_{i \in \mathcal{S}} p_i \mathbf{h}_i \mathbf{h}_i^H| - \log |\mathbf{I} + \sum_{i \in \mathcal{S}'} p_i \mathbf{h}_i \mathbf{h}_i^H| \end{aligned} \quad (3.34)$$

Since $\{p_i\}_{i \in \mathcal{S}'}$ is held fixed, optimising $\sum_{i \in \mathcal{S}} r_i$ results in increasing sum rate for all users. ■

3.3.1 Complexity of Norm Balancing Algorithms

Algorithm 3.1 inverts $M \times M$ matrices to solve the two user power allocation problem. This has complexity order of M^3 . Then, Algorithm 3.2 iterates the same procedure on a set of K users taking two at a time. It can be shown that the algorithm converges rapidly enough to declare a halt when all possible two user combinations have been tested. Hence, the number of iterations is bounded by $\frac{K(K-1)}{2}$. The complexity in each iteration is $\mathcal{O}(M^3)$. Hence, the

Figure 3.3: Two User Rate Region

complexity order of Algorithm 3.2 is $\mathcal{O}(K^2M^3)$. Even though the complexity is more than that of standard interior point method (given by $\mathcal{O}(K^2)$ in [129]), within the range of interest (user number up to 150) the proposed algorithm outperforms standard convex methods which will be shown later.

3.4 Characterisation of the Pareto Frontier

For a specific power allocation $\mathbf{p}$, the rate region has a polymatroid structure and sum rate maximum of the rate vector with the specific power allocation is a necessary and sufficient condition for Pareto optimality (see Theorem 3.1). However, when the constraint is on sum power rather than individual powers, the sum rate maximum point is a subset of the entire Pareto frontier.

It is illustrated for two users in Figure 3.3. The horizontal axis corresponds to the rate of first user r_1 and the vertical axis corresponds to that of the second user r_2 . The line segment corresponds to the maximum sum rate obtained for a specific power allocation whereas the curved segments correspond to other points on the Pareto frontier. The figure is a union of two user pentagonal rate regions having same sum power [123]. It is known that for a specific power allocation, the Pareto frontier is a convex polytope having $K!$ vertices, each corresponding to a decoding order of the users [48]. In the presence of SIC, the lower priority users are invisible to the higher priority users [32]. The SIC technique the multiple access receiver is dual of the interference pre-subtraction in a broadcast transmitter [23]. It was shown that a specific decoding order in a MAC is equivalent to the opposite encoding order in a BC [54]. So even though this work considers a MAC, the decoding order can be transposed to a BC. Further, it is known that the time sharing among achievable points and convexity of the rate region plays an important role in QoS optimisation problem. This work will consider time sharing among different decoding orders with a given power allocation, but not time sharing among different power allocations.

3.4.1 Pareto Frontier for Two User System

For a two user system obviously there exist only two distinct decoding orders. Theorem 3.4 will give the correct decoding order depending on the power allocation.

Theorem 3.4 *For a two user system constrained by sum power P and any power allocation which is not sum rate optimal, exactly one decoding order will give a Pareto optimal point.*

Proof: Let the powers of two users be denoted by $[q, P - q]$ obeying the feasibility condition and the solution of the optimisation problem (3.35) be $[q^*, P - q^*]$. Hence,

$$q^* = \operatorname{argmax}_{0 \leq q \leq P} \log |\mathbf{I} + q\mathbf{h}_1\mathbf{h}_1^H + (P - q)\mathbf{h}_2\mathbf{h}_2^H| \quad (3.35)$$

With the different power allocations and two different decoding orders, let the rates be denoted as in Table 3.1. In this context, a decoding order of $1 \rightarrow 2$ means user 1 is decoded first. After cancelling the interference, user 2 is decoded. It can be shown that [48]

Power allocation	$1 \rightarrow 2$ Decoding	$2 \rightarrow 1$ Decoding
$[q, P - q]$	r_{11}, r_{21}	r_{21}, r_{22}
$[q^*, P - q^*]$	r_{11}^*, r_{21}^*	r_{21}^*, r_{22}^*

Table 3.1: Rate Vectors under Different Conditions

$$r_{11} + r_{21} = r_{21} + r_{22} \leq r_{11}^* + r_{21}^* = r_{21}^* + r_{22}^*$$

It is trivial to prove that if $q = q^*$, any decoding order and their time sharing will give a sum rate maximal point, which is Pareto optimal by definition. Now it will be proved that if $q < q^*$, then $1 \rightarrow 2$ decoding order will give the Pareto optimal rate vector $[r_{11}, r_{21}]$. Also, conversely it will be shown the other decoding order $2 \rightarrow 1$ will give the suboptimal vector $[r_{21}, r_{22}]$.

For the direct part of the theorem, if q is increased, that means user 2 power will decrease and hence r_{21} will have to decrease. So it is not possible to get a better rate for the second user by increasing power of first user. On the other hand, if q is decreased, as $q < q^*$ already, the power allocation will move away from the optimal value and the sum rate will decrease monotonically. So it is not possible to do better for both users. This proves that for $q \leq q^*$, the rate vector $[r_{11}, r_{21}]$ is on the Pareto frontier.

To prove the converse, assume with $q < q^*$, $2 \rightarrow 1$ decoding order is used. This gives the rate vector $[r_{21}, r_{22}]$. It will be shown that we can do better for both users by some other power allocation. To that end, consider another different power allocation $[q', P - q']$ which gives the rate vector $[r_{21}, r'_{22}]$ with decoding order $1 \rightarrow 2$. Obviously since the first user priority has reduced in the new allocation, to keep his rate same, $q' > q$. Now consider two different cases.

$q' \leq q^*$ When $q < q' \leq q^*$, the power allocation $[q', P - q']$ is nearer to sum rate optimal compared to $[q, P - q]$. So $r_{21} + r'_{22} > r_{21} + r_{22}$ and $r'_{22} > r_{22}$. So $[r_{21}, r'_{22}]$ is not Pareto optimal.

$q' > q^*$ In this case $q < q^* < q'$. So $r_{11}^* < r_{21} < r_{21}^*$. For some $\alpha \in (0, 1)$, $r_{21} = \alpha r_{11}^* + (1 - \alpha)r_{21}^*$. Hence, a time sharing between between the points $[r_{11}^*, r_{21}^*]$ and $[r_{21}^*, r_{22}^*]$ (with weights α and $(1 - \alpha)$, respectively) will give a point $[r_{11}, \alpha r_{21}^* + (1 - \alpha)r_{22}^*]$, where rate of the first user is same as $[r_{21}, r_{22}]$ but the sum rate increases. So $[r_{21}, r_{22}]$ is not Pareto optimal

■

The variation of sum rate with power allocation among two users is shown in Figure 3.1 as an example. It can be shown that the sum rate is maximum when the two users have equal channel norms with respect to the inverse data rate matrix $\mathbf{A} \equiv (\mathbf{I} + p_1\mathbf{h}_1\mathbf{h}_1^H + p_2\mathbf{h}_2\mathbf{h}_2^H)^{-1}$. According

to Theorem 3.4, when the power allocation is to the right of where the norm curves intersect, $2 \rightarrow 1$ decoding order gives a Pareto optimal point.

3.4.2 Pareto Frontier for Multiple User System

Based on the discussion of two user decoding in the previous section, Theorem 3.5 will generalise the two user case to multiple user scenario.

Theorem 3.5 *Let for the set of all users $\mathcal{U}$, the sum power constraint be P . A feasible power allocation $\mathbf{p} = [p_1, p_2, \dots, p_K] \in \mathcal{P}$ is Pareto optimal if the following conditions are satisfied.*

- *There exists a set of users $\mathcal{S} \subseteq \mathcal{U}$ which operates at its own Pareto frontier and the total power of the set is more than what is required for that set for sum power optimality over $\mathcal{U}$.*
- *The complementary set $\mathcal{S}'$ is also operating at its own Pareto frontier tolerating interference from $\mathcal{S}$.*

If these conditions are satisfied, a decoding order $\mathcal{S}' \rightarrow \mathcal{S}$ will give a Pareto optimal point.

Proof: The sum rate maximum over the set $\mathcal{U}$ is obtained by $\mathbf{p}^* = [p_1^*, p_2^*, \dots, p_K^*] \in \mathcal{P}$. According to the conditions, for a given feasible allocation $\mathbf{p}$,

$$\sum_{i \in \mathcal{S}} p_i > \sum_{i \in \mathcal{S}} p_i^* \quad \text{and} \quad \sum_{i \in \mathcal{S}'} p_i < \sum_{i \in \mathcal{S}'} p_i^* \quad (3.36)$$

Now consider the decoding order $\mathcal{S}' \rightarrow \mathcal{S}$ where each subset is operating at its own Pareto frontier⁷. If it is possible to do better by some other power allocation, the only possibility is by a net power transfer between $\mathcal{S}$ and $\mathcal{S}'$.

$\mathcal{S}$ to $\mathcal{S}'$: Since $\mathcal{S}$ was already being decoded without interference from $\mathcal{S}'$, it cannot have a more favourable decoding order and hence, any reduction of its power is bound to reduce the available data rate for at least one user $i \in \mathcal{S}$.

$\mathcal{S}'$ to $\mathcal{S}$: In this case, along with a change in decoding order, it is not possible to predict the effects separately on $\mathcal{S}$ and $\mathcal{S}'$. However, if the sum power available to $\mathcal{S} \subset \mathcal{U}$ is increased beyond $\sum_{i \in \mathcal{S}} p_i^*$, that is equivalent to imposing an additional affine constraint on $\mathbf{p}$ in the form of $\sum_{i \in \mathcal{S}} p_i = P' \leq P$ for the optimisation problem (3.4). From Theorem E.1 of Appendix E, the resulting maximum value will be a concave function of P' and if P' is increased beyond $\sum_{i \in \mathcal{S}} p_i^*$, sum rate for all users $\sum_{i=1}^K r_i$ will decrease. So betterment for all users is not possible.

Hence, for the given power allocation $\mathbf{p}$ satisfying the two conditions for some $\mathcal{S} \subseteq \mathcal{U}$, there is no way to do better for all users. ■

An illustration of the additional power constraint is given in Figure 3.4. Consider Problem (3.4) with an additional constraint for some set of users $\mathcal{S} \subseteq \mathcal{U}$, as presented in (3.37).

$$\max_{\mathbf{p}} \quad \log |\mathbf{I} + \sum_{i \in \mathcal{U}} p_i \mathbf{h}_i \mathbf{h}_i^H| \quad \text{subject to} \quad \mathbf{p} \succeq \mathbf{0}, \quad \sum_{i=1}^K p_i = P \quad \text{and} \quad \sum_{i \in \mathcal{S}} p_i \leq P' \quad (3.37)$$

⁷This means all the elements of $\mathcal{S}'$ are decoded before the elements of $\mathcal{S}$

Figure 3.4: Effect of Additional Power Constraint

Clearly the additional constraint is feasible only if $P' \in [0, P]$. In Figure 3.4, the dotted line shows the optimum $\sum_{i \in \mathcal{S}} r_i$, i.e. the maximum sum rate of set $\mathcal{S}$. The continuous line shows the sum rate of all users $\sum_{i \in \mathcal{U}} r_i$. Both $\sum_{i \in \mathcal{S}} r_i$ and $\sum_{i \in \mathcal{U}} r_i$ are shown as functions of $\frac{P'}{P}$ with P' varying from 0 to P . If $P' = 0$, then Problem (3.37) is the same as maximising sum rate among users $\mathcal{S}'$. Also, if $P' = P$, then Problem (3.37) is the same as maximising sum rate among users $\mathcal{S}$. Hence, as expected, the sum rate of users in $\mathcal{S}$ increases monotonically with P' in the figure. Also, to be noted, the two curves in Figure 3.4 intersect when $P' = P$. This is because the sum power and rates of users $\mathcal{S}'$ is zero. Hence, $\sum_{i \in \mathcal{S}} r_i = \sum_{i \in \mathcal{U}} r_i$, where r_i is the rate of user i .

The conditions pointed out in Theorem 3.5 are sufficient for a point to be on Pareto frontier. However, they are not necessary conditions.

3.5 Optimal Decoding Algorithms

For any power allocation $\mathbf{p}$, if it is possible to find the set $\mathcal{S}$, the allocation is Pareto optimal according to Theorem 3.5. In fact, all feasible power allocations can achieve the Pareto frontier with correct decoding order and Algorithm 3.3 provides a constructive proof.

To keep the notations concise, the symbol for optimal power allocation vector is overloaded. $\mathbf{p}^*(\mathcal{S}, P) \in \mathbb{R}_+^{|\mathcal{S}|}$ denotes the optimum power allocation among a set of users $\mathcal{S}$ with sum power constraint P , i.e. the solution of optimisation problem (3.4) for the set $\mathcal{S}$, which can be found using the interior point or norm-balancing method (see Algorithm 3.2).

To see how the algorithm works, it is to verify that once user i is eliminated from $\mathcal{S}$, the entire set operates at Pareto frontier, if $\mathcal{S}$ operates at Pareto frontier. By successive elimination of the users, the flow arrives at a single user $|\mathcal{S}| = 1$, which is trivially at the Pareto frontier with its own power constraint. More than one users can satisfy the criterion for step 8 of the algorithm and there can be multiple optimal decoding orders. A special case is when the condition of step 4 of the algorithm is satisfied, which means the given power allocation maximises $\sum_{i \in \mathcal{S}} r_i$. Since sum rate is independent of decoding order, any decoding order of the set gives a Pareto optimal point.

Algorithm 3.3: Optimal Decoding Order to Achieve Pareto Frontier

Require: $\mathbf{p} \in \mathbb{R}_+^K$ and $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K \in \mathbb{C}^M$

- 1: $\mathcal{S} \leftarrow \{1, 2, \dots, K\}$ {Initialise with all K Users}
- 2: **while** $|\mathcal{S}| > 1$ **do**
- 3: $P \leftarrow \sum_{i \in \mathcal{S}} p_i$
- 4: **if** $p_i = p_i^*(\mathcal{S}, P)$ where $i \in \mathcal{S}$ **then**
- 5: Decode $\mathcal{S}$ in any order.
- 6: **Break**
- 7: **end if**
- 8: Find $i \in \mathcal{S}$ such that $p_i < p_i^*(\mathcal{S}, P)$.
- 9: $\mathcal{S} \leftarrow \mathcal{S} - \{i\}$
- 10: Decode user i treating signals for $\mathcal{S}$ as interference.
- 11: Subtract the known signal for user i from the received signal.
- 12: **end while**

Algorithm 3.4: Alternative Decoding Order to Achieve Pareto Frontier

Require: $\mathbf{p} \in \mathbb{R}_+^K$ and $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K \in \mathbb{C}^M$

- 1: $\mathcal{S} \leftarrow \{1, 2, \dots, K\}$ {Initialise with all K Users}
- 2: $\mathcal{T} \leftarrow \emptyset$
- 3: $k \leftarrow K$
- 4: **while** $|\mathcal{T}| < K$ **do**
- 5: $P \leftarrow \sum_{i \in \mathcal{S}} p_i$
- 6: **if** $p_i = p_i^*(\mathcal{S}, P, \mathcal{T})$ where $i \in \mathcal{S}$ **then**
- 7: Assign elements of $\mathcal{S}$ from $\pi(1)$ to $\pi(k)$ in any order.
- 8: **Break**
- 9: **end if**
- 10: Find $i \in \mathcal{S}$ such that $p_i > p_i^*(\mathcal{S}, P, \mathcal{T})$.
- 11: $\mathcal{S} \leftarrow \mathcal{S} - \{i\}$
- 12: $\mathcal{T} \leftarrow \mathcal{T} \cup \{i\}$
- 13: $\pi(k) \leftarrow i$
- 14: $k \leftarrow k - 1$
- 15: **end while**
- 16: Decode in the order $\pi(1) \rightarrow \pi(2) \rightarrow \dots \rightarrow \pi(K)$ with successive interference cancellation.

Algorithm 3.3 was devised to find the optimal decoding order from the lower end. Another algorithm, dual in nature can be formulated to tackle the decoding order problem from the higher end. To that end, the solution of Problem 3.4 will be further overloaded. $\mathbf{p}^*(\mathcal{S}, P, \mathcal{T}) \in \mathbb{R}_+^{|\mathcal{S}|}$ will denote the optimal power allocation for sum rate where all the users are tolerating interference from the set of users $\mathcal{T}$ with $\mathcal{T} \cap \mathcal{S} = \emptyset$. In other words, $\mathcal{T}$ determines the noise and interference matrix and the notation $\mathbf{p}^*(\mathcal{S}, P)$ used for Algorithm 3.3 can be thought as a special case with $\mathcal{T} = \emptyset$.

Algorithm 3.4 gives the decoding order. To see how it works, let user i be selected at step 10. If the rest of $\mathcal{S}$ can be made to work at Pareto frontier while tolerating interference from $\mathcal{T} \cup \{i\}$, then the entire set $\mathcal{T} \cup \{i\} \cup \mathcal{S}$ is at Pareto frontier from Theorem 3.5. So both Algorithms 3.3 and 3.4 give two alternate ways to arrive at an optimal decoding order.

As a comparison, Algorithm 3.3 appears conceptually simpler and more straightforward. Still the alternative algorithm is mentioned as an alternative, since it may be useful to prove the necessity of the decoding orders to achieve Pareto frontier. As a matter, it is also not proven that both of these algorithms will give the same set of optimal decoding orders among the $K!$ possible permutations. This can be an interesting open problem for further research, leading to a full characterisation of all optimal decoding orders.

3.5.1 Complexity of the Algorithms to Determine Optimal Decoding Order

As shown in Algorithms 3.3 and 3.4, the sum-rate optimisation problem (from Algorithm 3.2) has to be solved K times for K users, once after eliminating each user. Hence, the complexity is given by $K * \mathcal{O}(K^2 M^3) = \mathcal{O}(K^3 M^3)$.

3.6 Simulation Studies

Convex optimisation methods put the problems of least square programming, linear programming and quadratic programming in a more general framework. As already pointed out, the problem (3.4) is a maximisation of a concave function (equivalent to concave minimisation) over a convex region of feasible powers. It can be expressed as a special class of convex problems called *Disciplined* convex programming outlined in [130]. This concept of disciplined programming does not limit the generality of addressable problems, but aids in putting in a convenient structure. The same work [130] also proposes a framework to automate the solution procedure. For the purpose of this work, CVX solver module⁸ has been used as a benchmark against which the proposed algorithm is considered. The processor time requirements are plotted against the problem size (number of users). Figure 3.5 shows the time requirements when the algorithms are executed on a 64 bit Linux architecture. Even though the numerical values are platform specific, still it is seen that the proposed approach outperforms the standard convex optimisation approach for smaller and reasonable problem size. For convex optimisation method, the time requirement is relatively insensitive to the problem size. However, for norm balancing method bigger problem means more number of checks among the users and hence, more iterations. So the norm balancing method takes advantage of the problem size.

According to the trends evident from LTE, there is an increasing shift towards cooperative small cell systems. Even with growing user density, the miniaturisation of BSs and coverage areas (often concentrated within a single commercial block) is likely to lead to fewer users in

⁸Developed by CVX Research Inc. Austin (TX), US

Figure 3.5: Variation of Processor Time Requirement with Problem Size

each cell. Since the algorithm out-performs the comparable techniques with fewer users, it is expected that a real hardware implementation will lead to reduction of computational cost.

Figure 3.6: Convergence of Algorithm 3.2 with Number of Iteration

To demonstrate the convergence behaviour of Algorithm 3.2, an $M = 5$ antenna, $K = 15$ user system is considered, where the collection of channel vectors is $\mathbf{H} \in \mathbb{C}^{M \times K}$ and each element is taken as Rayleigh fading coefficient. So $\mathbf{H}_{ij} \sim \mathcal{CN}(0, 1)$. Two of the relevant quantities, viz. the norm difference Δ and the sum rate obtained for a tentative feasible power allocation (with each iteration) are plotted against the number of iteration in Figure 3.6. As can be

seen easily, the norm difference rapidly (but not monotonically) converges to zero (implying all non-deprived users have the same channel norm) and the sum rate curve tends to become horizontal. With a relatively low number of iterations the power allocation becomes nearly optimal. Exact number of iterations will certainly depend upon the number of users being considered. In system level implementations Δ serves as a metric of suboptimality.

Figure 3.7: Variation of Sum Rate Obtained by Algorithm 3.2 with Number of Users

Figure 3.7 shows the variation of sum rate (achievable by Algorithm 3.2) with the number of users. The curves start with power of 50dBm at the BS. Then the number of users is varied from 50 to 200. In the broken curve, the total power is kept constant at 50dBm throughout. In the continuous curve, the power is varied in proportion to the number of users. As it is seen, the sum-rate grows fast with the number of users when power for each user stays approximately same, but the growth is slow when the total power of users is same.

3.6.1 Effectiveness of Decoding Order

For a general K user system, there exists $K!$ different decoding orders and the corresponding rate regions belong to a $\mathbb{R}_+^K$ cone. Fig. 3.8 gives an illustration for a two user case without time sharing.

The figures shows the achievable rate regions for two users transmitting to a five antenna base station through a slow flat fading channel. The channel coefficients are taken from Rayleigh distribution. The users are have a total power constraint of 1 watt. Different points on the same curve represent different allocations of 1 watt between the users. The innermost curve represent the performance obtained when interference is treated as noise, for which the relevant optimisation strategies have been discussed in [71, 72]. The two outer curves represent the rate regions of two different decoding orders. As seen, the optimal decoding order to operate at a Pareto frontier depends on the power allocation. The figure basically results from taking the union of Pentagons in Fig. 2.3 for different power allocations.

To put it more rigorously, consider the rate regions with total power $P = p_1 + p_2$ where $[p_1, p_2]$ keeps changing.

- $\mathcal{C}_n(P)$: The achievable rate region by treating interference as noise, i.e. the region under

Figure 3.8: Two User Pareto Frontiers with Different Decoding Orders
the green curve. For this region, when the power distribution is $[p_1, p_2]$, the rates are

$$r_1 = \log |\mathbf{I} + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H| - \log (1 + p_2 \|\mathbf{h}_2\|^2)$$

$$r_2 = \log |\mathbf{I} + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H| - \log (1 + p_1 \|\mathbf{h}_1\|^2)$$

- $\mathcal{C}_{12}(P)$: The achievable rate region by decoding order $1 \rightarrow 2$, i.e. the region under the blue curve. For a power distribution $[p_1, p_2]$, the rates are given by

$$r_1 = \log |\mathbf{I} + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H| - \log (1 + p_2 \|\mathbf{h}_2\|^2)$$

$$r_2 = \log (1 + p_1 \|\mathbf{h}_1\|^2)$$

- $\mathcal{C}_{21}(P)$: The achievable rate region by decoding order $2 \rightarrow 1$, i.e. the region under the red curve. For a power distribution $[p_1, p_2]$, the rates are given by

$$r_1 = \log (1 + p_1 \|\mathbf{h}_1\|^2)$$

$$r_2 = \log |\mathbf{I} + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H| - \log (1 + p_1 \|\mathbf{h}_1\|^2)$$

The achievable rate region without any constraint on the decoding orders and allowing for time sharing is thus given by $\text{Co}(\mathcal{C}_{21}(P) \cup \mathcal{C}_{12}(P))$. Obviously both of the interference cancellation procedures consistently outperform interference ignorance, implying $\mathcal{C}_{21}(P) \cap \mathcal{C}_{12}(P) \supset \mathcal{C}_n(P)$.

When the constraint is on the total power P , then depending on the allocation, one of the decoding orders gives optimal performance according to Theorem 3.4.

Similarly, the surfaces in Fig. 3.9 show the outer boundaries for three users pulling power from a common source and transmitting to a base station. The rates of the users are plotted

Figure 3.9: Three User Pareto Frontiers

along the three axes. For three users, there are six different decoding orders out of which two have been chosen for demonstration.

- For the decoding order $3 \rightarrow 2 \rightarrow 1$ and power allocation $[p_1, p_2, p_3] \in \mathbb{R}_+^3$, the rates are given by

$$r_1 = \log \left(1 + p_1 \|\mathbf{h}_1\|^2 \right) \quad (3.38)$$

$$r_2 = \log |\mathbf{I} + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H| - \log \left(1 + p_1 \|\mathbf{h}_1\|^2 \right) \quad (3.39)$$

$$r_3 = \log |\mathbf{I} + (\mathbf{I} + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H)^{-1} p_3 \mathbf{h}_3 \mathbf{h}_3^H| \quad (3.40)$$

- For the decoding order $1 \rightarrow 2 \rightarrow 3$ and power allocation $[p_1, p_2, p_3] \in \mathbb{R}_+^3$, the rates are given by

$$r'_1 = \log |\mathbf{I} + (\mathbf{I} + p_3 \mathbf{h}_3 \mathbf{h}_3^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H)^{-1} p_1 \mathbf{h}_1 \mathbf{h}_1^H| \quad (3.41)$$

$$r'_2 = \log |\mathbf{I} + p_2 \mathbf{h}_2 \mathbf{h}_2^H + p_3 \mathbf{h}_3 \mathbf{h}_3^H| - \log \left(1 + p_3 \|\mathbf{h}_3\|^2 \right) \quad (3.42)$$

$$r'_3 = \log \left(1 + p_3 \|\mathbf{h}_3\|^2 \right) \quad (3.43)$$

It is seen that depending upon the power allocation, different decoding orders outperform one another. For example towards the point where the rate of user 1 reaches a its maximum, the curves intersect each other, implying a change in the optimal decoding order. The algorithms proposed in the literature serve to determine the optimal decoding orders.

3.7 Conclusion

The analyses done in this chapter allow us to see the optimal power allocation problem from a different point of view, viz. the channel norm balancing with respect to the inverse rate matrix. It is shown that by optimising the power allocation of two users at a time, eventually the entire set of users can be served with maximum sum rate.

The power allocation has an interesting relationship with generalised channel norms. For only two users, either the channel norms must be equal or the user with lesser channel norm has to be denied service to maximise the sum rate of users. This property resembles the classical water filling problem of resource allocation [43]. This channel norm property helps to establish the optimality of power allocation for different subsets of users as shown in Theorem 3.3. The characterisation of optimal power allocation makes use of some properties of an inner product space to test an allocation for optimality.

Numerical results illustrate convergence behaviour of the norm balancing algorithm. The relationship between optimal power allocation and norms helps us to readily identify if a given power allocation is optimal or if there is any set $\mathcal{S} \subseteq \mathcal{U}$ with optimal allocation among its own users (while tolerating interference from others). This approach can be used to further characterise the outer boundary of rate region for multiantenna receivers.

The chapter also discussed about achieving the Pareto frontier on a sum power constrained MAC. The sum power constraint is actually important only in BC. Because of the established duality principle, the multiple access results can be directly projected to a BC.

It is known that to attend the same rate point with sum power constraint, the dirty paper encoding order in a broadcast transmitter (which is basically interference pre-cancellation) must be the opposite of successive decoding at a multiple access receiver. For a MAC with a power allocation $\mathbf{p} \in \mathbb{R}_{++}^K$, Algorithm 3.3 gives the proper decoding order to achieve a Pareto optimal point under sum power $\sum p_i$. However, to implement the concept in a BC, it is necessary to carry out the MAC-BC transformation [54, Section IV.B] and reverse the decoding order.

Chapter 4

Generalised Pareto Frontier and Extension to Relay Systems

MULTI-WAY relays combine spectral efficiency and Space Division Multiple Access (SDMA) to meet the growing demand for wireless data. Prior works discussed system models where a pair of users exchanged information via an Amplify-and-Foward (AF) relay. For example, a zero forcing beamforming strategy and block diagonalization have been proposed in [131] to nullify inter-user interference. Similarly, [132] proposed a beamforming scheme based on null space where each user's signal lies in the interfering users' null space. Subsequently, [133] proposed a linear precoding strategy, where an AF relay coordinates transmission between multiple user pairs. The goal was to maximise the minimum throughput, which has been termed as max-min fairness criteria in literature. Both of these two schemes (from [132] and [133]) extract each signal pair from a one dimensional space. Further, the outage performance of dual-hop relays was analysed in [134]. Also, the multiuser scenario has been extended to incorporate cellular systems with two way relay in [135] along with small cell applications in [126]. A zero forcing transmitter was designed at the relay station to differentiate all the signal flows from the base station and users in [136]. However, in contrast to these prior works, this chapter focusses on the resource allocation of MAC and DF relay channels using some known results of MAC and system level duality.

Many optimisation challenges for MAC and Broadcast (BC) systems have been analysed with the help of duality principles¹. For example, the weighed sum rate maximisation problem with given power allocation takes a mathematically elegant form. This was solved using properties of polymatroids [127]. maximisation of sum rate with given sum power was solved in the previous chapters. In addition, the weighted sum power minimisation and sum rate maximisation problems follow the convex structures which can be solved using well-known techniques. For sum power minimisation with individual rate target, the optimal method is unknown, but a suboptimal method has been proposed in [14].

The motivation for this chapter is to unify the existing framework of optimisation problems previously discussed in the context of dual MAC [14, 48, 138] and subsequently to show how the general framework can be applied to DF relay channels. To be noted, some other recent

¹In this chapter, *broadcast* means each user receiving his own private information from a common source, following prior literature on PHY layer and information theoretic analyses [23, 63, 137].

papers discussed resource optimisation in multi-antenna relay systems. [139] discussed full-duplex systems using an AF relay. Subsequently, [140] extended the results to a massive MIMO system. [141] studied physical layer network coding in two-way relays from a principal-angle perspective. [142] studied secured beamforming of relay systems with an untrusted relay. [143] proposed a dual-hop approach in a system model using a source node and a destination node, aided by multiple relays. However, a practical heterogeneous cellular system is modelled more accurately by a DF relay system between multiple user pairs, whose analysis is the most important focus of this chapter. The major contributions are

- It proposes a novel analysis based on cost and reward functions to represent optimisation problems for BC and MAC in a unified and systematic context.
- Based on the cost and reward formulation, two novel norm-balancing algorithms are proposed which can be used to minimise sum power for a given throughput or to maximise throughput for a given sum power.
- The analyses in MAC have been extended to heterogeneous cellular relay systems that employ the proposed resource allocation methods. The goal is to find the minimum transmit power at the Macro Base Station (MBS) to achieve a Joint Pareto frontier (JPF) depending on the transmit powers of the Femto User Equipments (FUE). To verify the existence of JPF, three methods have been proposed and compared with analyses on algorithmic complexity.

This chapter is organised as follows. Section 4.1 presents the heterogeneous cellular system model and the problem formulation. The concept of Generalised Pareto Frontier (GPF) has been introduced in Section 4.2. The sum power minimisation algorithm for single cell systems is presented in Section 4.3. Then the solution is extended to a heterogeneous cellular system via the Polymatroid intersection problem in Section 4.4. Simulation studies are shown in Section 4.5, where the results are compared with existing literature. Finally, Section 4.6 concludes the paper.

4.1 The System Model

The system model discussed in this chapter is a heterogeneous network which is modelled as a relay channel. However, earlier works focussed on AF relay systems whereas the focus of this work is a DF relay. Figure 4.1 gives a practical diagram where a low power smaller cell (which we will refer to as a femtocell henceforth) is operating under a macro cell of wider coverage area. The macro base station (MBS) and the femto base station (FBS) are connected by a fronthaul network. It is assumed that the macro user equipments (MUE) which are located near the MBS communicate pairwise with femto user equipments (FUE) located on the cell edge, supported by the FBS. For tractable analyses, we make some assumptions.

- The users do not have direct links as it will not give the advantage of cooperation.
- Since the users are basically mobile devices with low transmit power, the signals from the MUE do not reach FUE or the FBS. Likewise, the signals from the FUE do not reach the MUE or the MBS.
- The two base stations function as the same node.

In the context of heterogeneous cloud radio access networks, the inter base-station links are known as *fronthaul*, established by copper coaxial or fibre optic cable depending on the commercial budget, quality requirement and local regulations. The recent industry trend points

Figure 4.1: MBS-FBS system model.

towards a wider adaptation of fibre optic networks [144]. Hence, for the purpose of this work, the two base stations together are considered parts of the same node with high processing capability and acting as a relay [145, 146]. The system does not consider multiple macro cells and their cooperation, which can be incorporated in future works. Also, in the given system model, since the FBS is a low power terminal located at the macro cell edge, typically in an indoor environment [145], its interference with the MBS is assumed negligible. Any residue interference can be reduced by the Enhanced Intra-Cell Interference Cancellation technique [147]. However, the information theoretic analyses adopted in this work can be used to model many different practical systems apart from this one.

4.1.1 Communication Scenario

Consider the MUE are transmitting to the FUE as shown in Figure 4.1. Let there be K definite pairs of users, namely, K MUEs and K FUEs, whose information passes via the MBS and the FBS. The phase from the MUE to the MBS (denoted by MUE→MBS) is an example of MAC. Further, it is assumed the MBS completely decodes the information for all K users before transmitting them to FBS. Hence, the FBS to FUE phase (denoted by FBS→FUE) can be modelled by a BC. From the duality principle (which first appeared in [63] followed by some other works), the rate region for a BC and MAC are equivalent to each other. Hence, the analyses carried out for MAC will be extended to incorporate the co-existing MAC and BC systems.

4.1.2 Problem Formulation for Sum Power Minimisation

Consider power constraints of the MUEs as $[p'_1, p'_2, \dots, p'_K] \in \mathbb{R}_{++}^K$, i.e., all positive powers² and slow flat fading channels in both phases, namely, the MUE→MBS and FBS→FUE. If there is a user whose transmit power is upper bounded by 0, then the user is taken as non-existent for this work. Also assume perfect channel state information at the relevant nodes. The objective is to find the minimum power required at the FBS, such that both the MAC subsystem (MUE→ MBS) and the BC subsystem (FBS→ FUE) operate at their own Pareto frontiers [148, Definition 1].

² $\mathbb{R}_{++} \equiv \{x|x > 0, \quad x \in \mathbb{R}\}$ and $\mathbb{R}_+ \equiv \mathbb{R}_{++} \cup \{0\}$.

To present it in mathematical terms, assume that the Pareto frontier of the MUE→MBS phase is given by $\mathcal{P}_{\text{MAC}}(p'_1, p'_2, \dots, p'_K)$. Similarly, the Pareto frontier of the FBS→FUE is given by $\mathcal{P}_{\text{BC}}(P)$. The problem is to determine

$$\min P \quad \text{subject to} \quad \mathcal{P}_{\text{BC}}(P) \cap \mathcal{P}_{\text{MAC}}(p'_1, p'_2, \dots, p'_K) \neq \emptyset \quad (4.1)$$

To solve the problem, the duality principle between BC and MAC systems have been used. For given power vector at the MUEs $(p'_1, p'_2, \dots, p'_K)$, the sum rate is independent of the decoding order, i.e.

$$\sum_{i=1}^K r_i = \log |\mathbf{I} + \sum_{i=1}^K p'_i \mathbf{h}_i \mathbf{h}_i^H| \quad (4.2)$$

where $\mathbf{h}_i$ is the channel between MUE i and the MBS.

Since the FBS→FUE phase is a BC, it is possible to address the optimisation problem via a sum power minimisation problem to achieve the target sum rate of (4.2). If the solution to the power minimisation problem is denoted by $\mathbf{p}^*$, then the achievable sum rate for the FBS→FUE can be found by expressions similar to (4.2). Further, via the duality principle [54], this power can be transposed to a BC. However, the sum rate equality does not necessarily guarantee equality of individual rate vectors. In that case, $\mathcal{P}_{\text{BC}}(P) \cap \mathcal{P}_{\text{MAC}}(p'_1, p'_2, \dots, p'_K) = \emptyset$. Hence, it is necessary to test for the existence of intersection and also to ensure that an intersection exists for a power allocation. After addressing the sum power minimisation problem in Section 4.3, the polymatroid intersection issue has been discussed in Section 4.4.

4.2 Generalized Pareto Frontier for the Proposed System Models

Transmission power is one of the most important resources in wireless systems. The concept of Pareto frontier under constrained sum power and successive interference cancellation in MAC was discussed in [148]. This work extends the formulation of Pareto frontier to a more generalized class of constraints and objectives. To that end, we define the notions of cost and reward functions.

4.2.1 Cost and Reward Functions

In our system, power is considered as the input representing a cost to operate the system, while the achievable rates are considered as an output that represents the reward. Formally, the cost and reward functions for a K user system will be defined as follows.

Definition 4.1 A function $f(\mathbf{p})$ is defined as a cost function of the power allocation $\mathbf{p} = [p_1, p_2, \dots, p_K] \in \mathbb{R}_+^K$, if it satisfies the following properties.

- P1: $f: \mathbb{R}_+^K \rightarrow \mathbb{R}_+$.
- P2: $f(\mathbf{0}) = 0$ which means the cost is zero in the absence of any transmit power.
- P3: $f(\mathbf{p})$ is monotonically increasing with each of $[p_1, p_2, \dots, p_K]$ while keeping the others constant.
- P4: $f(\mathbf{p})$ is convex in its domain.

P3 ensures that higher power implies higher cost while P4 captures higher marginal cost for each additional unit of power, which is a practical assumption. To see why that is the case,

Table 4.1: The landscape for different problem formulations

Description	Constraints	Objective	Remarks
Sum rate maximisation with given sum power	$\sum_{i=1}^K p_i$	$\sum_{i=1}^K r_i$	Solved [138]
Sum power minimisation with given sum rate	$\sum_{i=1}^K r_i$	$\sum_{i=1}^K p_i$	Solved in this work.
Weighted rate maximisation with individual power constraints	$p_i \leq p'_i$ for $1 \leq i \leq K$	$\sum_{i=1}^K w_i r_i$	Solved [48]
Sum power minimisation with individual rate constraints	$r_i \leq r'_i$ for $1 \leq i \leq K$	$\sum_{i=1}^K p_i$	Suboptimal solution proposed [14]
Proportional fair rate optimisation with sum power constraint	$\sum_{i=1}^K p_i$	$\sum_{i=1}^K \log r_i$	Joint beamforming optimisation proposed via SINR balancing [71].
Weighted sum power minimisation with individual rate constraints	$r_i \leq r'_i$ for $1 \leq i \leq K$	$\sum_{i=1}^K \lambda_i p_i$	Not solved
Weighted sum power minimisation with weighted rate constraint	$\sum_{i=1}^K w_i r_i$	$\sum_{i=1}^K \lambda_i p_i$	Standard convex problem [43]
Weighted sum rate maximisation with target weighted power constraint	$\sum_{i=1}^K \lambda_i p_i$	$\sum_{i=1}^K w_i r_i$	

consider $\alpha \in (0, 1)$. Now from P1 and P4, $f(\alpha \mathbf{p}) \leq \alpha f(\mathbf{p})$. This implies that the cost grows superlinearly with power.

Definition 4.2 A function $g(\mathbf{r})$ of the rate vector $\mathbf{r} = [r_1, r_2, \dots, r_K]$ is defined as a reward function, if it satisfies the following properties.

- P5: $g : \mathbb{R}_+^K \rightarrow \mathbb{R}_+$.
- P6: $g(\mathbf{0}) = 0$ which means zero reward is obtained when the rate is zero.
- P7: $g(\mathbf{r})$ is monotonically increasing with each of $[r_1, r_2, \dots, r_K]$ while keeping the others constant.
- P8: $g(\mathbf{r})$ is concave in its domain.

P7 implies higher utility is derived from higher rate. Likewise, P8 captures the diminishing marginal utility. It follows from similar reasoning (P5 and P8) that for $\alpha \in (0, 1)$, $g(\alpha \mathbf{r}) \geq \alpha g(\mathbf{r})$. This implies that the reward grows sub-linearly with rate. All the properties P1–P8 are formulated based on observations from practical systems which demonstrate increasing marginal resource cost and diminishing marginal return.

Using cost and reward functions, we quantify profit as $g(\mathbf{r}) - f(\mathbf{p})$ and formulate different classes of optimisation problems.

- Cost minimisation under reward constraint

$$\min_{\mathbf{p} \succcurlyeq \mathbf{0}} f(\mathbf{p}), \quad \text{subject to} \quad g(\mathbf{r}) \geq R; \quad (4.3)$$

- Reward maximisation under cost constraint

$$\max_{\mathbf{p} \succcurlyeq \mathbf{0}} g(\mathbf{r}), \quad \text{subject to} \quad f(\mathbf{p}) \leq P; \quad (4.4)$$

- Profit maximisation under feasible (non-negative) power constraint

$$\max_{\mathbf{p} \succcurlyeq \mathbf{0}} (g(\mathbf{r}) - f(\mathbf{p})); \quad (4.5)$$

where $\succcurlyeq$ denotes the component-wise inequality of vectors. Table 4.1 presents the current landscape of some special cases of these problems as applied to MAC. In the table, i indicates a general user index, p_i and r_i are, respectively, the variable power and rate of user i . p'_i and r'_i are, respectively, the constant constraints.

For instance, the first row in Table 4.1 is a special case of problem 4.4. Also, the second row is a special case of problem 4.3. It is seen that the sum power minimisation problem constitutes a gap in the existing body of literatures and its solution is one of the contributions of this work. In this context, this chapter proposes a power allocation $\mathbf{p}$ to be at the *Generalized Pareto frontier* (GPF), if it is a solution of problems (4.3)–(4.5) for some $f(\mathbf{p})$ and $g(\mathbf{r})$ satisfying

Definitions 4.1 and 4.2.

Lemma 4.1 *If $\mathbf{p}^* \in \mathbb{R}_+^K$ is a solution of problem (4.3) for some f and g , then $\mathbf{p}^*$ is a solution of problem (4.4) as well, and vice versa.*

Proof: The result follows from the dual nature of the problems (4.3) and (4.4). The detailed derivation is straight-forward. ■

Further, problem (4.5) is also related to the other problems as shown in Lemma 4.2.

Lemma 4.2 *For appropriately defined cost and reward functions, if a power vector $\mathbf{p}^* \in \mathbb{R}_+^K$ is a solution of problem (4.5), then it is also a solution of problems (4.3) and (4.4) for the same cost and reward.*

On the other hand, if $\mathbf{p}^$ is a solution of problems (4.3) and (4.4), then it is possible to construct a pair of rate reward functions such that the same $\mathbf{p}^*$ is an optimal solution of problem (4.5).*

Proof: Suppose $\mathbf{p}^* \equiv \operatorname{argmax}_{\mathbf{p} \succeq \mathbf{0}} (g(\mathbf{r}) - f(\mathbf{p}))$. Then it is obvious that,

$$\mathbf{p}^* = \operatorname{argmax}_{\mathbf{p} \succeq \mathbf{0}} g(\mathbf{r})$$

subject to $f(\mathbf{p}) \leq f(\mathbf{p}^*)$. If any other $\mathbf{p}$ could improve the reward $g(\mathbf{r})$ while satisfying the same cost constraint on $f(\mathbf{p})$, then that $\mathbf{p}$ would also be the solution of problem (4.4). To prove the converse, take the Lagrangian for (4.4) expressed as $g(\mathbf{r}) - \mu(f(\mathbf{p}) - P)$. If $\mathbf{p}^* \in \mathbb{R}_+^K$ is optimal, from Karush-Kuhn-Tucker conditions [44], $(\mathbf{p}^*, \mu)$ is a stationary argument for some $\mu \in \mathbb{R}_+$. Hence,

$$\frac{\partial g(\mathbf{r})}{\partial p_i} \Big|_{p_i=p_i^*} - \mu \frac{\partial f(\mathbf{p})}{\partial p_i} \Big|_{p_i=p_i^*} = 0, \quad 1 \leq i \leq K. \quad (4.6)$$

Now construct a new cost function³ $f'(\mathbf{p}) = \mu f(\mathbf{p})$ and imagine dual variables $\lambda_i = 0$, $1 \leq i \leq K$. From (4.6),

$$\frac{\partial g(\mathbf{r})}{\partial p_i} \Big|_{p_i=p_i^*} - \frac{\partial f'(\mathbf{p})}{\partial p_i} \Big|_{p_i=p_i^*} + \lambda_i = 0, \quad 1 \leq i \leq K. \quad (4.7)$$

It can be verified that (4.7) is the stationarity condition for the Lagrangian of problem (4.5) where the dual variables corresponding to the constraints $p_i \geq 0$ are given by λ_i and $f(\mathbf{p})$ is replaced by $f'(\mathbf{p})$. Hence, any $\mathbf{p}^* \in \mathbb{R}_+^K$ is a solution of problem (4.5) under appropriate cost function. ■

It is important to consider the possibility of time sharing between multiple operating points, using different power allocations and rate rewards. The following lemma shows that time sharing is not necessary to arrive at a solution of problems (4.3) or (4.4) for appropriately defined cost and rewards.

Lemma 4.3 *For any cost function f and reward function g satisfying Definitions 4.1 and 4.2 respectively, the problems (4.3)–(4.5) will have a single power allocation as a solution, as opposed to time sharing between multiple power allocations and decoding orders.*

³P1–P4 from Definition 4.1 can be easily verified.

Proof: Take N operating points as $(\mathbf{p}_1, \mathbf{r}_1), (\mathbf{p}_2, \mathbf{r}_2), \dots, (\mathbf{p}_N, \mathbf{r}_N)$ and weights given by $\alpha_1, \alpha_2, \dots, \alpha_N \geq 0$ and $\sum_{i=1}^N \alpha_i = 1$. The coefficient α_i represents the weight given to the operating point $(\mathbf{p}_i, \mathbf{r}_i)$ in a time sharing scheme. The net power vector is $\sum \alpha_i \mathbf{p}_i$ and the net rate vector is $\sum \alpha_i \mathbf{r}_i$. To find the bounds for the cost and reward, use P4 and P8 (in Section 4.2.1) respectively,

$$f\left(\sum_{i=1}^N \alpha_i \mathbf{p}_i\right) \leq \sum_{i=1}^N \alpha_i f(\mathbf{p}_i), \quad g\left(\sum_{i=1}^N \alpha_i \mathbf{r}_i\right) \geq \sum_{i=1}^N \alpha_i g(\mathbf{r}_i). \quad (4.8)$$

So time sharing increases the cost for the same average power consumption while reduces the reward for the same average rate. ■

4.2.2 Linear Cost and Reward Functions

From the general forms of cost and reward functions described earlier, it is possible to find the special cases of linear cost and reward functions obeying the definition. It will be shown that any power allocation $\mathbf{p}$ can be on a generalized Pareto frontier (in the sense of problems (4.3) and (4.4)). To that end, the importance of linear cost and reward functions will be pointed out. Then a constructive method to find an appropriate cost function given a reward will be proposed.

At this point, it is useful to introduce the concept of generalized inner product and norms. Let $\mathbf{u} \in \mathbb{F}^K$ and $\mathbf{v} \in \mathbb{F}^K$ be two vectors over the field $\mathbb{F}$, where $\mathbb{F}$ is $\mathbb{C}$ or $\mathbb{R}$. Also, let $\mathbf{A} \succ \mathbf{0}$ be an $\mathbb{F}^{K \times K}$ positive definite matrix over the same field. Then the inner product of $\mathbf{u}$ and $\mathbf{v}$ with respect to $\mathbf{A}$ is denoted and defined by $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{A}} \equiv \mathbf{u}^H \mathbf{A} \mathbf{v}$. The corresponding norm is $\|\mathbf{u}\|_{\mathbf{A}} \equiv \sqrt{\mathbf{u}^H \mathbf{A} \mathbf{u}}$. It can be verified that these definitions satisfy the required properties of inner product and norm. We will use the notations in the context of optimisation of sum power for multiple users. Also, $\langle \mathbf{u}, \mathbf{v} \rangle$ without the matrix subscript denotes the well-known dot product where $\mathbf{A} = \mathbf{I}$.

It has already been pointed out that the decoding order plays an important role in determining the operating point, given a power allocation. Under a fixed power vector, the achievable rate region has the shape of a polymatroid [48]. If the reward function is a weighted combination of rates given by $\sum_{i=1}^K w_i r_i$, where $w_i \geq 0$ are the weights, then the optimal decoding order to maximise the reward is the increasing order of user weights [48]. That means if $w_i \geq w_j$, then user j must be decoded before user i . Since in this section we are concerned with the weighted sum rate reward functions and the user numbering is arbitrary, let us assume that $w_1 \leq w_2 \leq \dots \leq w_K$. So the decoding order to maximise the reward is the increasing order of user index, i.e., $1 \rightarrow 2 \rightarrow \dots \rightarrow K$. Given a linear reward function, we can readily construct a linear cost function for any power allocation which solves problem (4.3). The method will be presented in the form of a constructive proof of Theorem 4.1. However, it is necessary to prove that the weighted sum rate function is indeed concave in power allocation, which is shown in the following lemma.

Lemma 4.4 *For any power allocation among the users and using optimal decoding order, the weighted sum rate is concave in power allocation.*

Proof: Consider the power allocation, rate and weight (corresponding to rate) vectors denoted respectively as $\mathbf{p}, \mathbf{r}, \mathbf{w} \in \mathbb{R}_{++}^K$ where $w_1 \leq w_2 \leq \dots \leq w_K$. Then $\langle \mathbf{w}, \mathbf{r} \rangle$ will be

maximised by the optimal decoding order $1 \rightarrow 2 \rightarrow \dots \rightarrow K$, which was shown in [48].

Using interference cancellation strategy [54, Section III-B],

$$\sum_{i=k}^K r_i = \log|\mathbf{I} + \sum_{i=k}^K p_i \mathbf{h}_i \mathbf{h}_i^H| \quad 1 \leq k \leq K. \quad (4.9)$$

Now the objective function can be written as

$$\langle \mathbf{w}, \mathbf{r} \rangle = w_1 \sum_{i=1}^K r_i + \sum_{k=2}^K \left((w_k - w_{k-1}) \sum_{i=k}^K r_i \right). \quad (4.10)$$

From properties of the log determinant function, it is easily shown that (4.9) is a concave functions of $\mathbf{p}$. Hence, each term having the same form in (4.10) is concave. Also, from our earlier assumption $w_1 \leq w_2 \leq \dots \leq w_K$, the coefficients $(w_k - w_{k-1})$ are positive. Hence, (4.10) is a conic combination [43, Chapter 2] of concave functions. ■

Now it is possible to formulate the relevant optimisation problem whose solution is given by a known power allocation. This will subsequently pave the way to formulate the sum power minimisation algorithm.

Theorem 4.1 *For the reward function $g(\mathbf{r}) = \langle \mathbf{w}, \mathbf{r} \rangle$ obeying $0 \leq w_1 \leq w_2 \leq \dots \leq w_K > 0$ and arbitrary power allocation $\mathbf{p}^* \in \mathbb{R}_+^K$, we can find a cost function f that yields $\mathbf{p}^*$ as a solution of problems (4.3) and (4.4).*

Proof: As noted already, with the given reward function, the optimal decoding order is $\pi = 1 \rightarrow 2 \rightarrow \dots \rightarrow K$. In that case, the rates of users (with power allocation $p_k \in \mathbb{R}_+$ and channel vectors $\mathbf{h}_k \in \mathbb{C}^M$) are

$$r_k = \log|\mathbf{I} + \sum_{i=k}^K p_i \mathbf{h}_i \mathbf{h}_i^H| - \log|\mathbf{I} + \sum_{i=k+1}^K p_i \mathbf{h}_i \mathbf{h}_i^H|. \quad (4.11)$$

For power allocation $\mathbf{p}^*$, denote the rate vector as $\mathbf{r}^* = [r_1^*, r_2^*, \dots, r_K^*]$ which is a fixed quantity. Let us take a weighted combination of powers as the cost function $f(\mathbf{p}) = \langle \boldsymbol{\lambda}, \mathbf{p} \rangle$ (which obviously satisfies Definition 4.1). It will be shown that for any $\mathbf{p}^* \in \mathbb{R}_+^K$, it is possible to find $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_K] \in \mathbb{R}_+^K$ such that $\mathbf{p}^*$ is a solution of (4.3).

With the linear cost and reward functions, the problem is

$$\min_{\mathbf{p} \succeq \mathbf{0}} \langle \boldsymbol{\lambda}, \mathbf{p} \rangle \quad \text{subject to} \quad \langle \mathbf{w}, \mathbf{r} \rangle \geq \langle \mathbf{w}, \mathbf{r}^* \rangle. \quad (4.12)$$

and the Lagrangian is

$$L(\mathbf{p}, \boldsymbol{\mu}, \zeta) = \langle \boldsymbol{\lambda}, \mathbf{p} \rangle - \langle \boldsymbol{\mu}, \mathbf{p} \rangle - \zeta (\langle \mathbf{w}, \mathbf{r} \rangle - \langle \mathbf{w}, \mathbf{r}^* \rangle). \quad (4.13)$$

where $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_K] \in \mathbb{R}_+^K$ and $\zeta \in \mathbb{R}$ are dual variables. From (4.11), we get the following partial derivatives.

$$\frac{\partial r_k}{\partial p_j} \quad (4.14)$$

$$\begin{cases} = 0, & \text{for } 1 \leq j < k, \\ = \mathbf{h}_k^H \left(\mathbf{I} + \sum_{i=k}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \mathbf{h}_k & \text{for } j = k, \\ = \mathbf{h}_j^H \left(\left(\mathbf{I} + \sum_{i=k}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} - \left(\mathbf{I} + \sum_{i=k+1}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \right) \mathbf{h}_j & \text{otherwise.} \end{cases}$$

So the relevant KKT optimality conditions can be expressed as $\mu_j p_j = 0$, $\zeta \langle \mathbf{w}, (\mathbf{r} - \mathbf{r}^*) \rangle = 0$ and

$$\lambda_j = \mu_j + \zeta \sum_{i=1}^j w_i \frac{\partial r_i}{\partial p_j}. \quad (4.15)$$

To satisfy the KKT optimality condition for power allocation $\mathbf{p}^*$, choose the parameters by Algorithm 4.1. If ζ , μ and λ are thus chosen, it is seen that $L(\mathbf{p}^*, \boldsymbol{\mu}, \zeta)$ will satisfy the KKT

Algorithm 4.1: Choice of KKT Parameters to optimise (4.13)

Require: $\mathbf{p}^*, \mathbf{w} \in \mathbb{R}_+^K$, $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K \in \mathbb{C}^M$.

- 1: Assign any $\zeta > 0$
 - 2: **for** $j = 1 : K$ **if** $p_j > 0$ $\mu_j \leftarrow 0$
 - 3: **else** Assign any $\mu_j \geq 0$ **end if**
 - 4: Assign $\lambda_j \leftarrow \mu_j + \zeta \sum_{i=1}^j w_i \frac{\partial r_i}{\partial p_j}$ to satisfy (4.15). **end for**
-

conditions. But to claim that $\mathbf{p}^*$ really minimises a cost function, it is necessary to prove $\langle \boldsymbol{\lambda}, \mathbf{p} \rangle$ is really a cost, which means showing $\boldsymbol{\lambda} \in \mathbb{R}_+^K$.

Since $\mu_j \geq 0$ and $\zeta > 0$ in (4.15), it suffices to show

$$\sum_{i=1}^j w_i \frac{\partial r_i}{\partial p_j} \geq 0 \quad 1 \leq j \leq K$$

To that end, note that

$$\begin{aligned} \sum_{i=1}^j w_i \frac{\partial r_i}{\partial p_j} &= \mathbf{h}_j^H \left(w_1 \left(\mathbf{I} + \sum_{i=1}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \right. \\ &\quad \left. + \sum_{i=1}^{j-1} (w_{i+1} - w_i) \left(\mathbf{I} + \sum_{l=i+1}^K p_l \mathbf{h}_l \mathbf{h}_l^H \right)^{-1} \right) \mathbf{h}_j. \quad 1 \leq j \leq K \end{aligned} \quad (4.16)$$

Now, all the matrices having the form $\mathbf{I} + \sum_i p_i \mathbf{h}_i \mathbf{h}_i^H$ are positive definite. So from (4.16), consider the matrix

$$w_1 \left(\mathbf{I} + \sum_{i=1}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} + \sum_{i=1}^{j-1} (w_{i+1} - w_i) \left(\mathbf{I} + \sum_{l=i+1}^K p_l \mathbf{h}_l \mathbf{h}_l^H \right)^{-1} \quad (4.17)$$

If at least one of the weights $w_i > 0$, as assumed in the premise, this matrix is positive definite. Hence, (4.16) is a norm of the vector $\mathbf{h}_j \in \mathbb{R}^K$. Hence, $\lambda_j \geq 0$ for $1 \leq j \leq K$. ■

For Non-Zero Power Allocation to All Users

For a candidate power allocation $\mathbf{p}^* \in \mathbb{R}_{++}^K$, the dual variable $\zeta > 0$ is a free parameter and $\mu_j = 0$, for $1 \leq j \leq K$. Then from (4.15), all λ_j become proportional to the free parameter. That is expected since proportional change in the weight of the objective function does not affect the optimisation problem (4.12).

Special Case of Sum Power versus Sum Rate : An interesting special case of the generalized Pareto frontier under given rate weight is when $w_j = 1$, $1 \leq j \leq K$. Let $\zeta = 1$ and

$\boldsymbol{\mu} = \mathbf{0}$. Combining (4.15) and (4.16), we get $\lambda_j = \mathbf{h}_j^H \left(\mathbf{I} + \sum_{i=1}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \mathbf{h}_j$. Denoting $\mathbf{A} \equiv \left(\mathbf{I} + \sum_{i=1}^K p_i \mathbf{h}_i \mathbf{h}_i^H \right)^{-1}$ we can express $\lambda_j = \|\mathbf{h}_j\|_{\mathbf{A}}^2$. Also, if $\|\mathbf{h}_j\|_{\mathbf{A}}$ is the same for all j then all weights become equal for the power allocation. Hence, the problem reduces to a sum power minimisation under given sum rate.

Based on this norm equality, the next section will propose two algorithms to achieve a sum rate with minimum sum power.

4.3 Sum Power Minimisation Algorithms

From the system model of Figure 4.1, it is clear that if the data rate through the fronthaul link is R , then the FBS must support the same data rate via its BC links to the FUE. Minimising the transmit power and finding the allocation among users (using the well-known duality principle [148, Section 3.2]) are important subproblems which can be written as in problem (4.18).

$$\min \sum_{i=1}^K p_i \quad \text{subject to} \quad \log |\mathbf{I} + \sum_{i=1}^K p_i \mathbf{h}_i \mathbf{h}_i^H| \geq R \quad (4.18)$$

Obviously problem (4.18) is a special case of weighted sum power minimisation to achieve a weighted sum rate target when all the weights (corresponding to powers and rates) become equal to one. This has been solved in Algorithms 4.2 and 4.3. Algorithm 4.2 solves the optimisation problem for two users at a time treating the signal for other users as part of the noise. It was shown in earlier works on MAC [138, Theorem 2] that for two users either the channel norms will have to be equal (implying $\|\mathbf{h}_1\| = \|\mathbf{h}_2\|$) or one user has to be deprived of any power, implying $p = 0$ for that user. Hence, it is possible to allocate optimal power by solving $\|\mathbf{h}_1\| - \|\mathbf{h}_2\| = 0$ by numerical methods.

Also, since optimal power allocation (to minimise total consumption or maximise sum rate) among different user pairs necessarily and sufficiently results in overall optimal allocation [138, Theorem 3], it is possible to accomplish the task pairwise. In this context, a channel norm imbalance implies suboptimal power allocation. Hence, the task is to equate the norms for maximally imbalanced pairs repeatedly. While choosing the pair to minimise the sum power, the goal is to minimise the number of iteration. Hence, as part of a heuristic reasoning, the user pair with maximally imbalanced norm is chosen. The steps are presented in Algorithm 4.3 for which Algorithm 4.2 serves as an intermediate step.

Algorithm 4.2: Power minimisation between two users to achieve a sum rate

Require: $[p_1, p_2] \in \mathbb{R}_+^2$, $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^M$ and $\mathbf{S}_z \succ 0 \in \mathbb{C}^{M \times M}$

- 1: $R \leftarrow |\mathbf{S}_z + p_1 \mathbf{h}_1 \mathbf{h}_1^H + p_2 \mathbf{h}_2 \mathbf{h}_2^H|$
- 2: $p'_1 \leftarrow \frac{e^R - 1}{\mathbf{h}_1^H \mathbf{S}_z \mathbf{h}_1}, p'_2 \leftarrow 0$
- 3: **if** $\mathbf{h}_1^H (\mathbf{S}_z + p'_1 \mathbf{h}_1 \mathbf{h}_1^H)^{-1} \mathbf{h}_1 > \mathbf{h}_2^H (\mathbf{S}_z + p'_1 \mathbf{h}_1 \mathbf{h}_1^H)^{-1} \mathbf{h}_2$ **return** p'_1 and p'_2
- 4: $p'_2 \leftarrow \frac{e^R - 1}{\mathbf{h}_2^H \mathbf{S}_z \mathbf{h}_2}, p'_1 \leftarrow 0$
- 5: **if** $\mathbf{h}_1^H (\mathbf{S}_z + p'_2 \mathbf{h}_2 \mathbf{h}_2^H)^{-1} \mathbf{h}_1 > \mathbf{h}_2^H (\mathbf{S}_z + p'_2 \mathbf{h}_2 \mathbf{h}_2^H)^{-1} \mathbf{h}_2$ **return** p'_1 and p'_2
- 6: Introduce dummy variables x and y .
 $A = \mathbf{h}_1^H (\mathbf{S}_z + x \mathbf{h}_1 \mathbf{h}_1^H + y \mathbf{h}_2 \mathbf{h}_2^H)^{-1} \mathbf{h}_1 - \mathbf{h}_2^H (\mathbf{S}_z + x \mathbf{h}_1 \mathbf{h}_1^H + y \mathbf{h}_2 \mathbf{h}_2^H)^{-1} \mathbf{h}_2$
- 7: Solve $A = 0$ and $|\mathbf{S}_z + x \mathbf{h}_1 \mathbf{h}_1^H + y \mathbf{h}_2 \mathbf{h}_2^H| = R$ using Levenberg–Marquardt algorithm.
- 8: **return** x and y

Algorithm 4.3: Power minimisation among multiple users to achieve a sum rate

Require: $\mathbf{p} \in \mathbb{R}_+^K$ and $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K \in \mathbb{C}^M$

- 1: **while** $\Delta > t$ $\mathcal{V} \leftarrow \{i | p_i > 0\}$, $\mathbf{A} \leftarrow (\mathbf{I} + \sum_{i \in \mathcal{V}} p_i \mathbf{h}_i \mathbf{h}_i^H)^{-1}$
- 2: $k \leftarrow \operatorname{argmax}_{i \in \mathcal{V}} \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i$, $l \leftarrow \operatorname{argmin}_{i \in \mathcal{V}} \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i$
- 3: minimise the sum power of k and l using Algorithm 4.2, Update $p_k, p_l, \mathbf{A}, \mathcal{V}$
- 4: $\Delta \leftarrow \max_{i \in \mathcal{V}} \|\mathbf{h}_i\|_{\mathbf{A}} - \min_{i \in \mathcal{V}} \|\mathbf{h}_i\|_{\mathbf{A}}$ **end while**

4.3.1 Power Allocation in the Presence of MBS Interference

As mentioned already, the MBS can cause interference to the different FUE by its own transmission. This adds an additional interference apart from the intra-cell interference inside the Femto cell. Hence, the transmission MUE $\rightarrow$ MBS $\rightarrow$ FBS $\rightarrow$ FUE suffers additional interference from the MBS $\rightarrow$ FUE links. So the channel vectors $\mathbf{h}_i^H$ from FBS to FUE i is modified by $\frac{\mathbf{h}_i}{\sqrt{\sigma_i^2 + 1}}$, where σ_i is the standard deviation of interference from MBS to FUE i . Algorithm 4.3 can be used to optimise the power allocation after this modification⁴.

4.4 Polymatroid Intersection Problem

The algorithms in Section 4.3 focussed on minimising the sum power in a MAC (thought as the dual of a BC). But as seen from Figure 4.1 we have a cascaded system of a MAC and a BC, both active in two directions. For example, if we consider the transmission from MUEs to the FUEs, then we have to equate the rate vectors in the two stages for pairwise communications. If the rate vector from the MUEs to the MBS is denoted by $\mathbf{r} \in \mathbb{R}_+^K$ and the rate vector from the FBS to the FUEs is $\mathbf{r}' \in \mathbb{R}_+^K$, then it is imperative to design the system in a way to have $\mathbf{r} = \mathbf{r}'$.

For example, if the channel vectors from the MUE i to the MBS (see Figure 4.1) is given by $\mathbf{g}_i \in \mathbb{C}^{M \times 1}$, the power allocation vector of the MUEs (in uplink) is $\mathbf{p} \in \mathbb{R}_+^K$ then the sum rate necessary for the FBS to the FUE BC must be $\sum_{i=1}^K r_i = \log |\mathbf{I} + \sum_{i=1}^K p_i \mathbf{g}_i \mathbf{g}_i^H|$.

Using Algorithm 4.3 (replacing $\mathbf{h}_i$ by $\mathbf{g}_i$) it is possible to find a power allocation $\mathbf{p}' \in \mathbb{R}_+^K$ for the dual MAC system for the FUE to the FBS. However, the algorithm does not guarantee

⁴MBS optimisation with simultaneous goal to minimise FUE interference can constitute extension of this work.

Figure 4.2: Intersection of Pareto Frontiers for Two user Rate Region

the same rate for each user since

$$\sum_{i=1}^K r_i = \sum_{i=1}^K r'_i \nRightarrow \mathbf{r} = \mathbf{r}'.$$

The scenario is illustrated for two user rate regions in Figures 4.2, where the horizontal and vertical axes show the rates of users 1 and 2 respectively. The two polymatroids [48] are upper bounded by the same sum rate at their Pareto frontiers, denoted by $\mathcal{P}'_1$ and $\mathcal{P}'_2$ respectively. In Figure 4.2(a), they intersect at their maximal sum rate points, i.e. the Pareto frontier. However, in Figure 4.2(b), there is no intersection. This section is concerned with the following general question. Construct two MAC systems $\mathcal{M}_1$ and $\mathcal{M}_2$ each having K users and same sum rate with the following parameters.

- $\mathcal{M}_1$ has power allocation $\mathbf{p}$ and channel vectors $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_K$.
- $\mathcal{M}_2$ has power allocation $\mathbf{p}'$ and channel vectors $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K$.

Is there any rate vector $\mathbf{r} \in \mathbb{R}_+^K$ that lies on the Pareto frontier of both of these channels under individual power constraints? The Pareto frontier of a polymatroid is the face where the sum rate is maximised [148, Theorem 1]. Hence, in the following sections, three different testing algorithms will be proposed to find the intersection of the Pareto frontiers.

4.4.1 Brute Force Subset Comparison

Let the polymatroid for the $\mathcal{M}_1$ be denoted by $\mathcal{P}_1 \subset \mathbb{R}_+^K$ and the same for $\mathcal{M}_2$ be denoted by $\mathcal{P}_2 \subset \mathbb{R}_+^K$. In addition, let us introduce the following set functions for two different polymatroids as $f_i : 2^{\mathcal{U}} \rightarrow \mathbb{R}_+$ for $i = 1$ or 2 such that⁵ $f_1(\mathcal{S}) = \log|\mathbf{I} + \sum_{i \in \mathcal{S}} p_i \mathbf{g}_i \mathbf{g}_i^H|$ and $f_2(\mathcal{S}) = \log|\mathbf{I} + \sum_{i \in \mathcal{S}} p'_i \mathbf{h}_i \mathbf{h}_i^H|$, where $\mathcal{S} \subseteq \mathcal{U}$. It can be shown that [54, Section III-B], $f_1(\mathcal{S})$ gives sum rate upper bound of $\mathcal{S} \subset \mathcal{U}$ in $\mathcal{M}_1$ and $f_1(\mathcal{S})$ for $\mathcal{M}_2$. From our sum rate matching assumption, $f_1(\mathcal{U}) = f_2(\mathcal{U})$. Now using [148, Lemma 1], it is possible to describe the sum power Pareto frontier of $\mathcal{P}_i$ in the $\mathbb{R}^K$ space (denoted by $\mathcal{P}'_i$) as $\mathcal{P}'_i = \{\mathbf{r} | f(\mathcal{U}) - f(\mathcal{S}') \leq \sum_{i \in \mathcal{S}} r_i \leq f(\mathcal{S})\}$ for $i = 1$ and 2 where $\mathcal{S}' \equiv \mathcal{U} - \mathcal{S}$. We can discard $\emptyset$ as a trivial subset and $\mathcal{U}$ itself since the sum rate condition is satisfied on the Pareto frontier. So we have $(2^K - 2)$ proper non-trivial subsets to compare. Hence, for $\mathbf{r} \in \mathcal{P}'_i$, it has to satisfy the upper bound constraint and lower bound constraints for all $\mathcal{S} \subset \mathcal{U}$ where the number of possible $\mathcal{S}$ is $2^K - 2$. So if there exists some

⁵ $\mathcal{U}$ denotes the set of users and $2^{\mathcal{U}}$ denotes its power set.

$\mathcal{P}'_1 \cap \mathcal{P}'_2$, all $\mathbf{r} \in \mathcal{P}'_1 \cap \mathcal{P}'_2$ has to satisfy the constraints corresponding to both $f_1(\mathcal{S})$ and $f_2(\mathcal{S})$. Hence, $\max\{f_1(\mathcal{U}) - f_1(\mathcal{S}'), f_2(\mathcal{U}) - f_2(\mathcal{S}')\} \leq \sum_{i \in \mathcal{S}} r_i \leq \min\{f_1(\mathcal{S}), f_2(\mathcal{S})\}$. Obviously the constraint presented is feasible iff

$$\max\{f_1(\mathcal{U}) - f_1(\mathcal{S}'), f_2(\mathcal{U}) - f_2(\mathcal{S}')\} \leq \min\{f_1(\mathcal{S}), f_2(\mathcal{S})\}. \quad (4.19)$$

Hence, it is possible to run the brute force test on all $\mathcal{S} \subset \mathcal{U}$ to see whether they satisfy (4.19). If the result is positive on all of them, then we can declare that that $\mathcal{P}'_1 \cap \mathcal{P}'_2 \neq \emptyset$.

However, to compare all the subsets, it is necessary to systematically enumerate them one by one. For the simulation purpose it is done by assigning a binary bit to every user. The bit is set if the user is selected in an $\mathcal{S} \subset \mathcal{U}$ and reset otherwise. Iteration through the binary equivalents of 1 to $2^K - 2$ will give us all the combinations. The Algorithm to enumerate all nontrivial proper subsets of $\mathcal{U} = \{1, 2, \dots, K\}$ is presented in Algorithm 4.4.

Algorithm 4.4: Nontrivial proper subsets of $\mathcal{U} = \{1, 2, \dots, K\}$.

Require: $K \in \mathbb{Z}_{++}$

- 1: $r \leftarrow 1, \mathcal{S} \leftarrow \emptyset$
 - 2: **while** $r < 2^K - 2$ $i \leftarrow 1, \mathcal{T} \leftarrow \emptyset$,
 - 3: Convert r to K digit binary format $\mathbf{v}$.
 - 4: **while** $i \leq K$ **if** $v_i == 1$ $\mathcal{T} \leftarrow \mathcal{T} \cup \{i\}$ **end if**
 - 5: $i = i + 1$ **end while**
 - 6: $\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathcal{T}\}, r = r + 1$ **end while**
-

Algorithm 4.5: $\lfloor \frac{K}{2} \rfloor$ combinations of subsets from $\mathcal{U}$.

Require: $K \in \mathbb{Z}_{++}$

- 1: Initialize $\mathcal{U}_1 \leftarrow \{\{1\}, \{2\}, \dots, \{K\}\}$ and $r \leftarrow 1$
 - 2: **while** $r \leq \lfloor \frac{K}{2} \rfloor$ $n \leftarrow 1, \mathcal{S} = \emptyset$
 - 3: **while** $n \leq K$ Find all $\mathcal{T} \in \mathcal{U}_r$ such that $1, 2, \dots, n \notin \mathcal{T}$
 - 4: $\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathcal{T} \cup \{n\}\}, n \leftarrow n + 1$ **end while**
 - 5: **if** $\min\{f_1(\mathcal{S}'), f_2(\mathcal{S}')\} + \min\{f_1(\mathcal{S}), f_2(\mathcal{S})\} < f_1(\mathcal{U})$ **end**
 - 6: $\mathcal{U}_{r+1} \leftarrow \mathcal{S}, r \leftarrow r + 1$ **end while**
-

4.4.2 Selective Subset Comparison

Notice that the left hand side of (4.19) can be written as $f_1(\mathcal{U}) - \min\{f_1(\mathcal{S}'), f_2(\mathcal{S}')\}$, since $f_1(\mathcal{U}) = f_2(\mathcal{U})$. Hence, the equivalent condition for (4.19) is

$$f_1(\mathcal{U}) \leq \min\{f_1(\mathcal{S}'), f_2(\mathcal{S}')\} + \min\{f_1(\mathcal{S}), f_2(\mathcal{S})\} \quad (4.20)$$

which is symmetric with respect to $\mathcal{S}$ and $\mathcal{S}'$. So if the condition (4.20) is satisfied for $\mathcal{S}$, then it is satisfied for the complementary set as well. This reduces the required number of comparisons by half. To take advantage of this, it is necessary to enumerate those subsets $\mathcal{S}$ of $\mathcal{U}$, which, together with their complements $\mathcal{S}'$ complete the set of non-trivial proper subsets of $\mathcal{U}$. To develop a method for that, observe that for any $\mathcal{S}$ with $|\mathcal{S}| = k < K$, we have $|\mathcal{S}'| = K - k$. Hence, if we enumerate all subsets $\mathcal{S}$ such that $|\mathcal{S}| \leq \lfloor \frac{K}{2} \rfloor$, then the corresponding larger complement sets will be taken care of. Algorithm 4.5 proposes a bottom up method for building the larger subsets (containing r elements, denoted by $\mathcal{U}_r$) from smaller subsets containing $(r - 1)$ elements starting from $r = 1$ singleton sets, where $\lfloor \bullet \rfloor$ denotes the floor

function. As argued earlier, it is enough to run the algorithm till $r = \lfloor \frac{K}{2} \rfloor$. If the condition is satisfied at step 5 for any $\mathcal{S} \subset \mathcal{U}$, that means there is no intersection.

4.4.3 Fixed Subset Comparison

Algorithm 4.5 still requires $2^{K-1} - 1$ comparisons for a K user system, which is still an exponential growth. To reduce this complexity further, it is possible to employ a suboptimal method, which requires comparison of each subset $\mathcal{S} \subset \mathcal{U}$ where $|\mathcal{S}| = r \leq \lfloor \frac{K}{2} \rfloor$. Greater the number of r , nearer the resulting rate point is to the Pareto frontier. However, the trade off is the increased algorithmic complexity given by K^r . For simulation purpose, $r = 2$ have been used to demonstrate the time requirement. In that case, the number of comparison is $\frac{K(K+1)}{2}$.

4.4.4 Suboptimal Power Allocation for Non-Joint Pareto Frontier

If any subset generated using Algorithm 4.5 violates (4.20), that implies that the channels $\mathcal{M}_1$ and $\mathcal{M}_2$ are severely imbalanced in spite of having the same sum rate. So it is impossible to operate them at a common point on the Pareto frontier. For illustration using a two user scenario, consider Figure 4.2 again. If the segments corresponding to the sum rate upper bound do not have a common point, the maximum rate of user 2 in $\mathcal{M}_1$ (corresponding to $\mathcal{P}'_1$) is below the minimum rate of user 2 in $\mathcal{M}_2$ (corresponding to $\mathcal{P}'_2$). In a practical environment, the imbalance arises if the same user pair has a weak (deep faded) channel at one side and a strong channel in the other side. In this case, even though the power allocations obtained by Algorithms 4.2 and 4.3 will achieve the sum rate, they may fail to achieve the rate vector. To address that, a suboptimal strategy to minimise power requirement is proposed following the approach of [14].

Let the Pareto frontier of the polymatroid rate region from the MUEs to the MBS be denoted by $\mathcal{P}'_1 \subset \mathbb{R}_{++}^K$. Now, at the FBS we need a BC strategy to achieve any $\mathbf{r} \in \mathcal{P}'_1$ with minimum sum power. To that end, consider the dual MAC at the femtocell, denoting the channels by column vectors $\mathbf{h}_i \in \mathbb{C}^M$. Now define a permutation $\pi : \mathcal{U} \rightarrow \mathcal{U}$ such that $\|\mathbf{h}_{\pi(i)}\| \leq \|\mathbf{h}_{\pi(i+1)}\|$, $1 \leq i < K$, arranging the users in the order of channel strengths. Then the proposed approach to minimise power consumption is

- Find the vertex of $\mathcal{P}'_1$ corresponding to the decoding order $\pi(1) \rightarrow \pi(2) \rightarrow \dots \rightarrow \pi(K)$. Let this rate vector be denoted by $\mathbf{r} \in \mathcal{P}'_1$.
- Find the minimum sum power to achieve $\mathbf{r}$ [14, Section IV-B].

The strategy exploits the diminishing marginal return of power against rate and assigns the highest decoding order priority to the strongest user according to the permutation. The strongest user $\pi(K)$ is the last one to be decoded and hence enjoys single user capacity without any interference.

Extension to Two-Way System

Communication may take place in the opposite direction in the same time slot if we have full-duplex transceivers and alternating time slots for half-duplex transceivers. Resource allocation in that case can be optimised via same methods adapted here.

4.5 Numerical Simulations

The simulation results of this section are based on the performance achieved by the sum power minimisation and polymatroid intersection algorithms proposed earlier. The plots are

Figure 4.3: Variation of time requirement with problem size.

Figure 4.4: Sum rate variations with UE power.

drawn by simulation on a five antenna base station and 15 single-antenna users, where the channel values are realized conforming to the Rayleigh distribution. The simulations were performed on a Linux machine running MATLAB[®] R2013, on a 64 bit processor.

In Figure 4.3 we compare three methods in terms of time requirement to solve the problem of sum power minimisation. To test against the convex optimisation techniques, we used the software CVX⁶ as a benchmark.

4.5.1 The Complexity Count

It is important to understand the complexity structure of the proposed optimisation algorithms. Algorithm 4.2 deals with the inversion of $M \times M$ matrices and then solution of equations involving the matrices of same dimensions, each of which has complexity order of M^3 . Then,

⁶Developed by CVX Inc., Texas, US.

Figure 4.5: Sum power variations with UE power.

Algorithm 4.3 iterates on a set of K users taking two at a time. It can be shown that the algorithm converges rapidly enough to declare a halt when all possible two user combinations have been tested. Hence, the number of iterations is bounded by $\frac{K(K-1)}{2}$. Since the complexity in each iteration is $\mathcal{O}(M^3)$, the complexity order of Algorithm 4.3 is $\mathcal{O}(K^2M^3)$. Even though the complexity is more than that of standard interior point method [43], within the range of interest (user number up to 150) the proposed algorithm outperforms standard convex methods as seen in Figure 4.3. A variation of the bisection method applicable to a MAC was proposed in earlier works [138], which is included to make the comparison complete.

Figs. 4.4 and 4.5 show the comparisons of the macro-cell and femto-cell in terms of base station powers and coverage areas. For the purpose of simulations, the parameters have been chosen to match the practical systems as described in [145]. The channel between any pair of antenna is given by $d^{-\alpha}h \in \mathbb{C}$, where α is the path-loss exponent, d is the distance and $h \in \mathbb{C}$ is a circularly symmetric standard Gaussian random variable. Also, the MBS–FUE distance is a kilometer, which means the corresponding interference is negligible compared to the intra femto-cell interference. The sum power of user equipments in Figs. 4.4 and 4.5 are assumed to vary starting from 30dBm to 2.2 times 30dBm=33.4 dBm. A value of x on the X-axis in Figures 4.4 or 4.5 imply a power of $30+10\log x$ dBm. x is varied from 1 to 2.2.

Although the values depend on the simulation parameters, the major goal of the simulation is to compare the achievable sum rates in the MUE→MBS→FBS→FUE direction as opposed to FUE→FBS→MBS→MUE direction. The important factors for this comparison are

- The FUE–FBS distance is taken to be 30 meters and the MBS–MUE distance is 100 meters. Although the macro cell radius is typically in the order of kilometers, the present assumption is to consider the users as being located sufficiently near the cell center.
- The path loss exponent α is assumed to be 2 for the FBS–FUE subsystem and $\alpha = 4$ for the MBS–MUE subsystem. This is because typically FBSs operate in indoor environments, often with line of sight advantage.
- The FUE and MUE are assumed to operate under same power constraints $\mathbf{p} \in \mathbb{R}^K$.

The dashed curve in Figure 4.4 shows the sum rate when the FUE has the power constraint

Figure 4.6: Simulation time requirement to test for intersection.

and the MBS operates in BC mode. The solid curve shows the opposite scenario, i.e. the MUE have the same power constraint and the FBS operates in BC mode.

Figure 4.5 shows the power requirements at the other base station to achieve the same data rate in the BC end, as determined by our proposed methods. For the set up in our simulation, it is evident that the MUE→MBS→FBS→FUE operates under less rate compared to the FUE→FBS→MBS→MUE side. This can be accounted for by the relative strength of the FBS–FUE channel compared to the MBS–MUE channel. In the broadcast modes, the MBS has more power to overcome the disadvantage, supporting higher data rate.

In Figure 4.6, we compare the proposed methods for the polymatroid intersection problem discussed in Section 4.4. Generally, the number of iterations will be variable. If once there is a mismatch for any $\mathcal{S} \subset \mathcal{U}$, the algorithm ends there. Hence, the comparison is done considering the worst case scenario. The curves are plotted for three different algorithms.

- Brute force comparison of Section 4.4.1 compares the sum rate for every subset of $\mathcal{U}$ according to Algorithm 4.4.
- Selective comparison performs the check only on sets containing less than $\frac{K}{2}$ elements according to Algorithm 4.5.
- The suboptimal method of Sec 4.4.2 checks the rate bound for every user, and then every user pair, resulting in a complexity of $\mathcal{O}(K^2)$. This corresponds to suboptimal two user comparison.

It is seen that the continuous line, even though higher, maintains a consistent difference with the broken line. This is because a proportional difference (of twice) in linear scale is translated to a fixed difference $\log 2$ in the logarithm scale. It is also observed that the broken line rises much slower than the other two, indicating the efficiency of the algorithm.

Further, Fig. 4.7 shows the effectiveness of the proposed method to achieve a sumrate when Algorithms 4.4 and 4.5 declare null intersection of the polymatroids. The simulations like the ones before, were undertaken with five antenna BS and 15 user pairs. The power of each MUE is increased from 10dBm to 10.3 dBm. The FBS power requirement is seen to vary from 23 dBm at the bottom of the curve to four times that value, approximately 23.6 dBm. The channel

Figure 4.7: Power Requirement at FBS to Achieve Joint Pareto Frontier

co-efficients are taken as Rayleigh faded channel like the earlier cases.

The method proposed in the Section 4.4.4 to find a vertex for target rate has been compared with the least sum power requirement among other vertices. It is seen that the proposed method (the continuous line) consumes less power to achieve the same target rate compared to other encoding orders [48].

4.6 Conclusion

The chapter uses the cost and reward functions to propose two efficient algorithms for sum power minimisation and the theoretical results are supported by simulations. The proposed algorithms have significant computational advantage over existing convex optimisation techniques. This is particularly relevant for relay channels and/or small cell scenarios with lower number of users, where the proposed algorithm outperforms the interior point methods. Then three testing algorithms to match rates at both ends of the relay are proposed along with complexity comparisons. These testing methods together with a suboptimal power allocation technique serve to extend the previous algorithms to DF relay channel. The theoretical and intuitive understanding of the Pareto frontier and Polymatroid structure from this work will guide future works on resource allocation problems in heterogeneous cellular systems.

Chapter 5

Throughput Maximisation through Cooperative Multipoint Transmission and Two Layer Decoding

COOPERATIVE multi-point transmission (CoMP) techniques have been proposed to blur the cell boundary and explore more efficient forms of intercell cooperation [19]. In the existing systems, each user equipment (UE) exchanges data with a predetermined single Base Station (BS). The downlink (broadcast) transmissions at adjacent cells go on independently of each other, usually separated in time/frequency domains so as to avoid interference. One of the major drawbacks is the non-uniform service quality, particularly, the poor performance suffered by users near the cell edge which was studied in [149]. Cooperative beamforming (CBF) is a special case CoMP, where the transmitting BSs still serve their own users. However, unlike existing systems in deployment, CBF entails minimisation of interference caused to other users as a secondary goal. Different interference management algorithms have been presented within cooperative frameworks [21], [19]. For example, [114] assumes that all the BSs have perfect cooperation and access to data streams of all users. Hence, the system model becomes equivalent to a multiuser single cell system with a sum power constraint. Further, a multicellular system was also considered in [32].

Also, a completely different paradigm of large scale antenna system (LSAS) or massive MIMO with a very high number of antennas has been proposed in literature [150]. Presence of large number of antennas in the BS supports transmission of multiple data streams along independent (interference free) spatial dimensions. To exploit this, zero forcing beamforming strategies have been proposed in [151], [152]. However, optimal utilisation of a big number of antennas faces serious challenges in terms of hardware and algorithmic complexity [153], [154]. Also, most MIMO architectures as deployed in existing wireless standards such as LTE and IEEE 802.11 etc. typically cater to single user systems.

The motivation behind this work is to enhance system throughput by incorporating the suboptimal two layer decoding [155] for massive MIMO in cooperative beamforming and present two novel interference cancellation methods, i.e, Centralised Interference Cancellation (CIC)

Figure 5.1: Comp System Model

and Distributed Interference Cancellation (DIC). Those two methods are compared between themselves along with existing techniques to identify the complexity-performance trade off in the design of next generation cellular systems.

The organisation of this chapter is as follows. Section 5.1 describes the system model of multicell transmission with a review of duality in Section 5.1.1. Then the optimal beamforming criteria is discussed in Section 5.1.2. The proposed two layer method for uplink management is discussed in Section 5.2. Section 5.2.1 discusses how to incorporate successive interference cancellation in a two layer massive MIMO system. Section 5.3 discusses the downlink beamforming via a two-layer approach with two alternative schemes, CIC and DIC. The proposed schemes are compared between themselves and prior beamforming schemes in Section 5.4. Finally, Section 5.5 concludes the chapter.

5.1 The System Model

The system model under consideration is a multi-cellular system where several BSs are serving multiple users scattered throughout the coverage area. It is shown in Figure 5.1. A transmission environment with three cooperating BSs is shown in Figure 5.1(a). It is assumed that a cluster of cooperating BSs use the same time-frequency resource block (unlike the currently deployed systems) to transmit downlink data. The other BSs adjacent to the cluster are assumed to use different time-frequency blocks and hence present zero interference. The frequency reuse happens across different clusters of cooperative cells and because of geographical distance, we can ignore the effect of inter-cluster interference. As we can see in Figure 5.1(a), the coverage areas can overlap with each other and each user is attached to only one BS, in line with the definition of CBF [22]. The channel carrying desired information for a user is shown by continuous line whereas the interfering channels from other BSs is shown as a broken line. Because of the backhaul cooperation (not shown in Figure 5.1(a)), each BS can access the data streams and channel state information among other BS-user pairs, which is important for the proposed beamforming procedure.

Figure 5.1(b) gives a schematic representation, where each user is attached to a BS and each BS in the cooperating cluster is controlled by central server. Note that every user has an existing link with every other BS, although some of them are omitted from the figure for

clarity. The continuous lines in the figures indicate signal path and the broken lines indicate the Intercell Interference (ICI) coming from other BSs in the cluster. From the implementation point of view, this is similar to the two layer schematic diagram for LSAS uplink (multiple-access) channel discussed in [155]. In Section 5.2, we will improve this idea of two layer decoding via Successive Interference Cancellation (SIC) and extend it to a broadcast channel.

5.1.1 Duality between Uplink and Downlink

It is well known that uplink channels render themselves to simpler analyses and optimisation procedure than broadcast channels. Fortunately, because of the duality principle [138, Section II], it is possible to solve any optimisation problem in a downlink channel using known solutions from an uplink channel. Suppose for a single cell uplink channel, the user equipment power constraints are $\mathbf{p} \in \mathbb{R}_+^K$ and the achievable rate region is given by $\mathcal{C}_{\text{MAC}}(\mathbf{p})$. Similarly, for a downlink channel, the BS power constraint is $P \in \mathbb{R}$ and rate region is given by $\mathcal{C}_{\text{BC}}(P)$. Then by duality principle [138, Section II]

$$\bigcup_{\sum p_i = P} \mathcal{C}_{\text{MAC}}(\mathbf{p}) = \mathcal{C}_{\text{BC}}(P)$$

Also, two algorithms have been proposed in [54] to translate any power allocation in the uplink channel to a power allocation in downlink channel and vice versa, so as to achieve the same rate region. Henceforth, we will attempt to solve the beamforming problems in the uplink direction and will assume that it is possible to transform the solutions to downlink.

5.1.2 Optimal Beamforming Vectors for a Cooperating Cluster

Following are the assumptions of each cooperating cluster. Each cluster has,

- L BSs each with N antennas and power constraints $P_l \in \mathbb{R}_{++}$, where $1 \leq l \leq L$.
- K single antenna users each attached to a BS and intends to receive a single data stream $x_k \in \mathbb{C}$ where $1 \leq k \leq K$.
- A central server with a high speed fibre optic connection with each BS facilitating instant communication of channel state and data signal.

The objective is to find the optimal beamforming vectors at the BSs required to transmit the data stream while obeying the power constraints. The optimal solution for single cell scenarios has been discussed in [138], [148]. However, for tractable analyses, this chapter will propose some suboptimal solutions and will compare them graphically with some existing methods in terms of performance and complexity.

5.1.3 Massive MIMO Systems for LTE-5G

Massive MIMO systems gained prominence in recent research to meet the challenges exponentially increasing data rate. According to [156], it is set to become a defining features of 5G networks coming up in 2020 and beyond. Some of the major advantages offered by Massive MIMO systems are as follows.

- Multiplexing gain offered by multiple independent signal paths.
- Large antenna arrays are capable of optimising transmit powers in both uplink and downlink communications. This is particularly helpful in exploiting the higher antenna aperture [157] and capture the directional waves, thus enhancing the SINR.

- The large number of antennas offer better robustness and reliability from antenna failure as well as channel outage. This results in better diversity gain.
- When the number of antenna at the BS is larger than the user antennas, linear precoding and receive processing are optimal [158].

However, practical implementations of massive MIMO post some significant research challenges. Particularly, the problems of channel acquisition, antenna calibration, manipulation of large matrices and the interference get exponentially more difficult with higher number of antennas, both from implementation and theoretical perspectives. Among the recent literature, [30] explored the implementation and radio measurements of Over-The-Air (OTA) massive MIMO enabled systems, as applicable to 5G systems. It also proposed some new requirements of such systems in light of the existing challenges.

In view of these, the goal of this chapter is to show the importance of massive MIMO via some suboptimal optimisation algorithms, which can outperform some other comparable systems in terms of performance. The work in this chapter does not achieve the information theoretic supremum. However, there are certain research challenges which can emerge from the techniques in this thesis and which can move us closer to the Shannon's limit. These are highlighted later in the chapter.

5.2 Two Layer Decoding Methods for Uplink Massive MIMO Systems

Before we discuss the cooperative beamforming methods, it is important to discuss the proposed two layer decoding method in massive MIMO uplink systems. Traditional zero forcing or MMSE receivers [54] need to find an inverse or pseudo inverse of a channel matrix for optimal reception. However, the computational complexity of finding an inverse varies with the cube of the matrix dimension, which makes it infeasible for large scale systems. In two layer systems, the data streams received at a large number of antennas are divided into different blocks. In the first layer, each individual block is processed using zero forcing or MMSE technique, in the second layer the results from the first block are processed using selection diversity.

5.2.1 Two Layer with SIC

In order to formulate the beamforming methods for downlink channel, a two layer decoding method for the uplink channel is proposed. A single BS is assumed, where a large number of antennas are partitioned into blocks. The system is similar to Figure 5.1(b), where each base-station is replaced by a block in the same BS. They cooperate with each other via the server to provide optimal performance.

Since each of the L blocks have N antennas, let the channel from user k to block l be denoted by $\mathbf{h}_{kl} \in \mathbb{C}^N$, where $1 \leq k \leq K$ and $1 \leq l \leq L$. The signal transmitted by user k is $x_k \in \mathbb{C}$ and the signal received at block l is $\mathbf{y}_l$. If we assume individual user power constraints as $p_1, p_2, \dots, p_K$, that means $\mathbf{E}(|\mathbf{x}_k|^2) = p_k$.

$$\mathbf{y}_l = \sum_{k=1}^K \mathbf{h}_{kl} x_k + \boldsymbol{\eta} \in \mathbb{C}^N, \quad \text{where } 1 \leq l \leq L \quad (5.1)$$

where $\boldsymbol{\eta}$ is centralised independent Gaussian noise with $\mathbb{E}(\boldsymbol{\eta}\boldsymbol{\eta}^H) = \mathbf{I}$.

Based on this system model, the novel two layer decoding method with SIC is shown in the form of a block diagram in Figure 5.2. Earlier works on uplink channels assumed full cooperation among the large number of antennas in the BS. However, this work assumes the antennas are partitioned into multiple blocks. Although the cooperation is limited, the presence of multiple blocks provide a diversity gain in the first layer of decoding. Exploiting this diversity gain and successively cancelling the interference are two most important parts of this work. After decoding each user, we subtract that user's signal from the received signals $\mathbf{y}_l$ so as to avoid propagation of the interference. As an example, if the decoding order is $1 \rightarrow 2 \rightarrow \dots \rightarrow K-1 \rightarrow K$, that means user k does not face interference from users 1 to $k-1$. Hence, while decoding user k , the interferences come only from users $k+1$ to K . Hence, adding the interferences for users $k+1$ to K in (5.1), the covariance of the interferences is $\sum_{i=k+1}^K p_i \mathbf{h}_{il} \mathbf{h}_{il}^H$. So the covariance matrix of noise and interference as seen in block l is given by

$$\mathbf{S}_{kl} = \mathbf{I} + \sum_{i=k+1}^K p_i \mathbf{h}_{il} \mathbf{h}_{il}^H. \quad (5.2)$$

With the given covariance matrix, if MMSE reception [159] is used, it can be shown that at block l , the SINR is given by $p_k \mathbf{h}_{kl}^H \mathbf{S}_{kl}^{-1} \mathbf{h}_{kl}$. After the first layer processing by MMSE, it is possible to employ selection diversity at the second layer to choose the block with highest SINR and after decoding. In the next step, the covariance matrices $\mathbf{S}_{kl}$ are updated for all l by subtracting the $p_k \mathbf{h}_{kl} \mathbf{h}_{kl}^H$ term.

Figure 5.2: Block Diagram for Two Layer Decoding

The steps are presented mathematically in Algorithm 5.1. In Algorithm 5.1, the subscript k is omitted from the covariance matrices $\mathbf{S}_l$ for the sake of brevity, since it is updated after each decoding.

Upon using this decoding method, the rate obtained for user k (after cancellation of interference by previous users) is given by

$$r_k = \log \left(1 + p_k \max_l \mathbf{h}_{kl} \left(\mathbf{I} + \sum_{i=k+1}^K p_i \mathbf{h}_{il} \mathbf{h}_{il}^H \right)^{-1} \mathbf{h}_{kl}^H \right) \quad (5.3)$$

Algorithm 5.1: Two Layer Decoding with SIC

Require: $\mathbf{p} \in \mathbb{R}_{++}^K$, $\mathbf{h}_{11}, \dots, \mathbf{h}_{kl}, \dots, \mathbf{h}_{KL} \in \mathbb{C}^N$ and $\mathbf{y}_1, \mathbf{y}_2 \dots \mathbf{y}_L$
 1: **for** $1 \leq l \leq L$
 2: $\mathbf{S}_l \leftarrow \mathbf{I} + \sum_{i=2}^K p_i \mathbf{h}_{il} \mathbf{h}_{il}^H$ **end**
 3: **for** $1 \leq k \leq K$
 4: $l^* \leftarrow \operatorname{argmax}_{1 \leq l \leq L} \mathbf{h}_{kl} \mathbf{S}_l^{-1} \mathbf{h}_{kl}^H$
 5: Decode user k from antenna cluster l^*
 6: **for** $1 \leq l \leq L$
 7: $\mathbf{S}_l \leftarrow \mathbf{S}_l - p_k \mathbf{h}_{kl} \mathbf{h}_{kl}^H$
 8: **end end**

5.2.2 Antenna Partitioning for Two-Layer Decoding

In this chapter, we have partitioned the antennas in a massive MIMO system according to their physical proximity or hardware implementation. Hence, the goal is to show the performance improvement via applying SIC and achieving an efficient trade-off. However, the partitioning based on hardware proximity is not optimal from an information theoretic point of view. While finding the optimal partition in this context has not been addressed in literature, some works proposed sub-optimal partitioning schemes in related system models. As an example, [160] proposed a two-stage constant-envelope precoding to reduce the cost of the RF chain. Also, [161] proposed a randomised scheme and analyses the computational complexity.

5.3 Downlink Interference Cancellation via Multicell Cooperation

The two layer method for uplink reception will be extended to downlink beamforming. While discussing downlink channels for single cells in this section, it is assumed that the noise at each user has the same power. If the powers are different, then the corresponding channel vector can be normalised, so that the noise variable having unit power.

Keeping that in mind, to maximise the sum rates constrained by individual BS powers $P_1, P_2, \dots, P_L$, it is necessary to optimise the following parameters:

- User Ordering,
- BS ordering,
- User association with BS,
- Power allocation and beamforming.

Clearly, for K users and L BSs, there can be $K!$ and $L!$ different orders respectively. In this work, for the sake of mathematical tractability, a specific ordering is assumed. Based on the given user and BS orderings, two interference cancellation methods will be proposed. The details are discussed below.

5.3.1 Centralised Interference Cancellation

In this scenario, a specific user ordering is decided by the central server out of $L!$ possibilities. After that, each user is associated with the BS which has the strongest channel (as determined by the $\mathcal{L}_2$ norm). Let the channel vector between user k and BS l be given by $\mathbf{h}_{kl}$. Without loss of generality, let us assume the user ordering is given by $1 \rightarrow 2 \rightarrow \dots \rightarrow K$. Consider an

association function ϕ , where

$$\phi : \{1, 2, \dots, K\} \rightarrow \{1, 2, \dots, L\}, \quad (5.4)$$

and $\phi(k) \equiv \operatorname{argmax}_l |\mathbf{h}_{kl}|^2$ represents the BS attached to user k . Also consider the inverse association function

$$\phi^{-1} : \{1, 2, \dots, L\} \rightarrow 2^{\{1, 2, \dots, K\}}, \quad (5.5)$$

which gives the set of users attached to any BS. Obviously in the case of CIC, there is no ordering of BSs since any two consecutive users can belong to two different BSs.

To find the required beamforming vectors, let us assume the vector transmitted by BS l is given by $\mathbf{x}_l \in \mathbb{C}^N$ where $1 \leq l \leq L$. Hence, the received signal at user k is

$$y_k = \sum_{l=1}^L \mathbf{h}_{kl}^H \mathbf{x}_l + \eta_k \quad (5.6)$$

where $\mathbb{E}(|\eta_k|^2) = 1$. Now, among all possible $\mathbf{x}_l$, only $\mathbf{x}_{\phi(k)}$ contains the useful data signal for k . The rest constitutes interference. However, because the BSs are cooperating with the central server, BS $\phi(k)$ is already aware of the interferences caused by signals for users $1, 2, \dots, k-1$, coming from BSs $\phi(1), \phi(2), \dots, \phi(k-1)$ respectively. So these interferences can be pre-subtracted before transmission by dirty paper coding [55]. Hence, for user k , the interferences will be coming from users $(k+1), \dots, K$, which must be treated as noise. Let the variance of this additional Gaussian interference be denoted by σ_k^2 . In that case, the total variance of noise and interference for user k is $(\sigma_k^2 + 1)$. The effective channel can be constructed as $\frac{\mathbf{h}_{k\phi(k)}}{\sqrt{\sigma_k^2 + 1}}$.

Hence, for BS l , the problem is to find the power allocation and beamforming vectors for users $\phi^{-1}(l)$. Since each user has different noise covariance, the channels must be normalised. Hence, it becomes a power allocation problem for single antenna users in a single cell downlink channel, performed independently at each BS according to power constraints P_l . It has been solved using duality principle and a norm balancing method as given in [138].

In summary, the steps for power allocation in CIC are

- All users are ordered according to pre-defined criteria by the server, out of $K!$ possibilities.
- Each user is associated with the BS which offers the strongest channel.
- Normalise the user channels with the noise and interference levels.
- Each BS will allocate power to maximise the sum rate according to the given noise levels using duality [54] and the algorithm proposed in [138].

Note that the last step in the procedure automatically cancels both the ICI (among users under the same BS) and intra-cell interference (from other BSs).

5.3.2 Distributed Interference Cancellation

In this scenario, the central server performs an ordering of the BS out of $L!$ possibilities. The users are associated with the BS in the same way as CIC and we will continue to use the functions ϕ and ϕ^{-1} to denote the user association.

However, the key difference with CIC is that the autonomous decisions are taken at the BSs to order their users and cancel the intercell interference. The ICI cancellation can be achieved

Figure 5.3: Ordering for Two BSs for Inter-cellular noise cancellation

using dirty paper coding among the BSs. Similar to what was done for the users in Chapter 3, the BS ordering concept can be applied among multiple cooperating cells. If the L cells are ordered according to index as $1, 2, \dots, l, \dots, L$, then BS i will tolerate interference from BS j for all $1 \leq i < j \leq L$. If the BS j has the downlink transmit symbol of BS i , then it can precancel the interference in the opposite direction.

This is shown for a two cell cluster in Figure 5.3. In the figure, BS l serves the group of users $\phi^{-1}(l)$ for $l = 1$ and 2 . However, BS 2 has higher priority than BS 1, which allows BS 2 to pre-subtract the interference from BS 1 to $\phi^{-1}(2)$, as illustrated by the absence of connecting lines. But this does not stop the interference from BS 2 to $\phi^{-1}(1)$ which is indicated by the dotted lines. For an L cell cluster, when the BS ordering is given by $1 \rightarrow 2 \rightarrow \dots \rightarrow L$, then for BS l , the interferences come from BSs $l+1$ to L . Hence, for users under a BS, the received signal can be written as

$$y_k = \sum_{i=l}^L \mathbf{h}_{ki}^H \mathbf{x}_i + \eta_k \quad \text{where} \quad k \in \phi^{-1}(l) = \mathbf{h}_{kl}^H \mathbf{x}_l + \sum_{i=l+1}^L \mathbf{h}_{ki}^H \mathbf{x}_i + \eta_k \quad (5.7)$$

where the second term represents ICI from other BSs with higher priorities.

The power allocation and beamforming problem can be solved as the sum rate maximisation problem in a single cell downlink channel with added intercell interference [138]. However, in the dual uplink channel, the sum rate is independent of the decoding order. So the BS l can use the appropriate decoding order not to maximise its own sum rate, but to minimise interference to users $\phi^{-1}(1), \dots, \phi^{-1}(l-1)$. To illustrate, let us assume the BS l transmits a vector $\mathbf{x}_l$. In that case, the ICI signals from BS l to other users is given by

$$\mathbf{h}_{kl}^H \mathbf{x}_l \quad \text{where} \quad k \in \bigcup_{i=1}^{l-1} \phi^{-1}(i)$$

Table 5.1: Comparison of Centralised and Distributed Interference Management

Scheme	Power Allocation	BS Ordering	User Ordering
CIC	Performed by BS	Results automatically from user ordering	Performed by server from $K!$ possibilities
DIC		Done by BS from $L!$ possibilities	Performed by each BS

In this work, the goal is to minimise the total interference power, which is given by

$$\mathcal{P}_l = \sum_{k \in \bigcup_{i=1}^{l-1} \phi^{-1}(i)} |\mathbf{h}_{kl}^H \mathbf{x}_l|^2 = \mathbf{x}_l^H \mathbf{A}_l \mathbf{x}_l \quad (5.8)$$

where

$$\mathbf{A}_l \equiv \sum_{k \in \bigcup_{i=1}^{l-1} \phi^{-1}(i)} \mathbf{h}_{kl} \mathbf{h}_{kl}^H$$

Since the positive definite matrix $\mathbf{A}_l$ depends solely on the channels and is independent of any beamforming strategy, the BS l has to find

$$\operatorname{argmin}_{\mathbf{x}_l} \mathbf{x}_l^H \mathbf{A}_l \mathbf{x}_l.$$

Now the optimisation parameter $\mathbf{x}_l$ can be taken only from a finite set of arguments, which we can find using the uplink-downlink transformation [54, Section IV–B]. After that it is possible to use brute force method to evaluate (5.8) for each resulting $\mathbf{x}_l$ and find the minimum. However, this step can be performed only if the available computational resources at each BS satisfy it.

In summary, steps for power allocation in DIC

- BSs are ordered according to pre-defined criteria by the server, out of $L!$ possibilities.
- Each user is associated with the BS which offers the strongest channel.
- Each BS will construct a dual multiple access channel [54] according to the given noise levels.
- Allocate the power to maximise the sum rate according to the BS power constraint P_l and using norm balance algorithm [138].
- [Optional: Depending on complexity] The user ordering is done by the BSs to minimise the determinant of interference covariance matrix according to duality.

Note that the power allocation to maximise sum rate is done exclusively to cancel the known intra-cell interference among the users. However, the optional last step seeks to minimise ICI with trade off on the algorithmic front.

Both CIC and DIC can be conceptualised as extensions of the two layer method in uplink to a cooperative downlink channel, where the roles of the antenna blocks [155] are replaced by individual BSs. The key difference between CIC and DIC have been summarised in Table 5.1.

5.4 Simulation Studies

To compare the proposed algorithms graphically with some existing results in the literature, the simulations were carried out using Matlab and Simulink on a Linux platform.

Figure 5.4 shows the performance of the two layer decoding system enabled by massive MIMO and enhanced by the SIC technique outlined in Section 5.2. The rates obtained from SIC are plotted according to (5.3). Two comparison benchmarks are the earlier two layer processing (proposed in [155, 162]) and the data rate achievable by an ordinary BS without the

Figure 5.4: Comparison of Achievable Sum Rates.

Figure 5.5: Comparison of Interference Cancellation Techniques.

aid of massive MIMO. For simulation purposes, a system consisting of 100 users is considered, along with a BS containing 1000 antenna divided into five blocks of 200 each. (For such a high number of antennas with uncorrelated channels, typically a millimetre range system is required [115].) The parameters have been chosen to match the practical systems as described in [145]. The channel between any pair of antenna is given by $d^{-\alpha}h \in \mathbb{C}$, where α is the path-loss exponent, d is the distance and $h \in \mathbb{C}$ is a circularly symmetric standard Gaussian random variable. Since the achievable data rate increases monotonically with power (with a diminishing marginal effect), the comparison is done for an appropriate range for illustration purpose.

The advantage of SIC along with two layer decoding is clear. Also, the last line indicates the data rate achievable for a single BS with 200 antenna, which does not allow the selection diversity employed at the second layer.

Further, in Figure 5.5, we compare the best possible outcomes of CIC and DIC. The performance of the interference cancellation techniques depend on the ordering of users in CIC and order of BSs in DIC. Since it is mathematically intractable to find the optimal ordering for maximum throughput performance, the algorithms have been run using brute force techniques to go through all possible orders and select the best possible one. For algorithmic efficiency,

Figure 5.6: Comparison of BS Enumeration and Ordering.

Figure 5.7: Time Complexity of Interference Cancellation Techniques.

the minimisation of (5.8) in DIC has been omitted and each BS has ordered the user according to the user index. It is clear that for the best possible scenarios, CIC outperforms DIC. However, the search space for CIC consists of $K!$ elements whereas for DIC, the search space has $L!$ elements, as pointed out in Table 4.1. The comparison benchmarks also include the conventional resource allocation strategy in downlink virtual MIMO, as pointed out in [72]. According to the framework proposed in [72], each BS allocates power without cooperation (to cancel intercell interference) and without dirty paper coding (to cancel intra-cell interference). The advantage of the proposed CIC and DIC is evident from the achievable throughputs.

Finally, Figure 5.6 shows the effect of enumeration (ordering among the BSs). If a cluster has $L = 5$ BSs, there are $5! = 120$ enumerations possible, giving different performance. In the figure, the data rates are plotted for the best and worst possible enumerations (determined by brute force) to illustrate the difference in throughput performance.

The computational trade off has been illustrated in Figure 5.7, where the time requirements

to execute the algorithms (in a 64 bit linux system) have been plotted against the number of users (problem size). Although the exact values are system dependent, the trend is evident from the variation of time requirements of the algorithms. Hence, DIC gives us a suboptimal result, but with superior algorithmic efficiency.

5.5 Conclusion

The important objective of the chapter is to identify the performance-complexity trade off by proposing several algorithms to allocate resources. Because of assumptions of specific ordering of users and BSs, the proposed techniques are suboptimal from the performance point of view. However, the assumptions are necessary to maintain analytical tractability and reduce the complexity. The proposed algorithms have been compared with several earlier methods to show the performance improvement where applicable. The work can be extended in the directions of further optimisation of the suboptimally allocated parameters. Also, it is possible to incorporate further cooperation among the BSs using backhauls which will enhance the data rate. This constitutes the discussion in the next chapter.

Chapter 6

Extension of Duality to Cooperative Multipoint Transmission

COOPERATIVE Multipoint systems coordinate transmissions and receptions across multiple BSs to enhance the data rate. As discussed earlier, the duality principle available for single cell systems plays a major role in solution of resource allocation/optimisation problems to achieve maximum throughput. It had been explored earlier in [54, 163] and generalised to linear covariance constraints in [34]. However, to the best of our knowledge, similar formulations for multicell cooperation have been absent in prior literature so far. It is in this context the present chapter explores the existence of duality principles in a multicell environment and proposes a novel formulation for the power allocation problem based on duality principle. To establish the duality in CoMP paradigm, two different cooperation scenarios are considered based on back-haul capacity. The major contributions of this chapter are

- It is shown that BSs employing massive MIMO systems [150] can achieve zero forcing of Inter-cell Interference (ICI) via a distributed beamforming method. The existence of UL-DL duality is shown to simplify the optimisation problem.
- A novel power allocation algorithm has been proposed to maximise the sum capacity in the CBF DL systems.
- The co-ordination among BSs have been further enhanced through a novel paradigm called Partial Data Cooperation (PDC) along with the proposed duality transformation.
- A sum rate maximisation problem is formulated which exploits the proposed duality transformation. Owing to the intractable structure of the problem, a suboptimal but efficient algorithm has been proposed to allocate the power. The results are compared with prior art.

The chapter is organised as follows. Section 6.1 introduces system model and two different cooperation paradigms studied in this chapter. Section 6.2 discusses how zero interference can be achieved with coordinated beamforming in both uplink and downlink channels. Section 6.3 presents the algorithm to achieve maximum sum rate via CBF. Then Section 6.4 discusses a novel cooperation paradigm called Partial Data Cooperation (PDC). The duality in PDC and the performance optimisation algorithm are presented in Section 6.5. Then Section 6.6 proposes two algorithms to characterise the Pareto frontiers for CBF and PDC based systems. Section 6.7 compares the performances of different paradigms numerically with earlier

Figure 6.1: Schematic of a CoMP System

proposed works. Section 6.8 concludes the chapter along with some pointers for future work.

6.1 System Model and Background

Consider a multicell system where L adjacent BSs communicate with users spread throughout a coverage area, as shown in Figure 6.1 for two BSs. The communication can happen in both UL and DL directions via time or frequency division duplexing¹. In the DL mode, a cluster of cooperating BSs use the same time-frequency resource block to transmit data to the users. Other BSs adjacent to the cluster are assumed to use different time-frequency blocks and hence present zero interference. The effect of inter-cluster interference is ignored because of geographical separation. Conversely, in the UL mode, the UEs transmit data to the BSs. Each UE operates under its own power constraint and without any cooperation with other UEs. However, the backhaul network can be utilised to have cooperation among the BSs. This chapter considers two different levels of cooperation among the BSs, viz. CBF and PDC. It is shown that the achievable performances obey the duality principle, which was known so far only for single cell systems [54]. Extension of duality principle to the more general cooperative setups with the respective algorithms for performance optimisation are the principal contributions of this chapter.

6.1.1 Coordinated Beamforming

According to the existing CBF paradigm [22], each BS is aware of its channels to all users. But each user is communicating with only a single BS. Hence, every BS interferes with the communication of other BS-user pairs. The goal is to coordinate among the BSs to minimise or nullify the ICI. The system model is shown in Figure 6.1 with $L = 2$ BSs and $K = 2$ users under each BS. The continuous lines indicate useful signal and the broken lines indicate ICI coming from other BSs. Hence, the downlink data for each user is available at only one BS. However, BSs in the cluster usually coordinate amongst each other on user scheduling and beamforming decisions. The scheduler targets to maximise the SINR or minimise leakage interference to other BSs while optimising the desired signal strength at the appropriate receiver. It has been called a Max-SLNR (Signal to Leakage and Noise Ratio) optimiser, which delivers better sum throughput [22] than schedulers based on SINR maximisation. In addition, scheduling allows

¹Analyses of full duplex communication requires modelling of the self-interference [109], which is not within the scope of this work.

Table 6.1: Uplink and Downlink Systems with Different Cooperation Methods

Method	Downlink	Uplink	Information needed to calculate DL Beamforming vector
CBF	Each BS transmits to UEs under its own coverage area.	Each UE transmits to only one BS.	Each BS needs to know its own channel to all the users.
PDC	Each BS transmits independent data streams to all UEs.	Each UE transmits independent data streams to the BSs.	Each BS needs to calculate $\mathbf{w}_{l,k}$ from (6.22).
JP	The BSs have all the data to be transmitted to UEs.	No uplink counterpart.	Each BS needs to know all the channels and data symbols.

a subset of receivers to be selected based on an orthogonal channel selection algorithm [164] so that the transmit and/or receive beamforming vectors can reduce the interference to/from other neighbouring users, while maximising the desired signal strength. CBF has proven particularly helpful in improving the signal strength and mitigating interference for cell-edge users, resulting in a more uniform throughput profile across a wide coverage area. Additionally, when the receivers are equipped with multiple receive antennas, transceiver algorithms at the cooperating transmitters and receivers should be jointly optimised to improve the throughput.

A major advantage of CBF is the lower overhead of backhaul traffic since the downlink data for any user will be stored only in a single BS. However, this also prevents leveraging the backhaul cooperation to its maximum extent which ultimately limits performance. This issue is addressed in the proposed PDC paradigm.

6.1.2 Partial Data Cooperation

The novel PDC paradigm introduces further cooperation using the backhaul network. In this case, each BS has access to part of the user's data to be transmitted (in DL) or will only decode part thereof (in the UL) when the users transmit to various base stations. The salient features of the proposed novel cooperation scheme are its practical and theoretical advantages, and its privacy and security features. These are explained below.

- *Practical Advantage:* It is a trade off between the earlier CBF and Joint Processing (JP) schemes in terms of two conflicting goals. PDC has less backhaul overhead traffic compared to JP and superior throughput performance compared to CBF.
- *Theoretical Advantage:* The cooperation scheme helps us to formulate earlier works on single cell UL–DL duality for a multicell scenario. This offers analytical advantages to solve the resource allocation problems of DL with efficient algorithms.
- *Data Privacy:* Since the BSs have access to parts of UEs' data, it is possible to support more privacy and security features at higher OSI layers such as presentation and transport. It is also possible to have more in-built redundancy among the BSs for encryption and decryption security.

Comparative descriptions of these two different cooperation paradigms in both UL and DL are summarised in Table 6.1, along with the existing JP with perfect data sharing [22], which will be used as a benchmark. The relative advantages in terms of necessary backhaul cooperation is presented in the last column of the table.

6.2 Duality between Uplink and Downlink with CBF

The aim of this section is to establish the uplink-downlink duality employing CBF paradigm and pave the way for the sum rate maximisation with zero forcing algorithm. To analyse the downlink systems, a group of cooperating BSs will be called a cluster. Assume that each cluster has

- L BSs each with M antennas and a power constraint $P_l \in \mathbb{R}_{++}$, where $1 \leq l \leq L$. Also, the BSs transmit in the same time-frequency block thus causing potential interference.
- K single antenna users being served by each BS, each user intending to receive a single data stream $x_k \in \mathbb{C}$, where $1 \leq k \leq KL$. Denote the number of users in other cells by $Q \equiv K(L - 1)$ and assume $M \geq Q$, i.e. the number of antennas in each BS is more than the number of intercell users. This assumption is consistent with heterogeneous systems with small coverage areas and the advent of large scale antenna systems [150].
- A central server with a high speed fibre optic connection with each BS facilitating instant communication of channel states.

The schematic form of a cluster is shown in Figure 6.1. Since each BS serves multiple users and each user also receives signals from multiple BSs, there are two components in the interference experienced by any user.

- ICI originating from other BSs in the cluster.
- Intracell interference because each BS serves multiple users at the same time-frequency resource block with finite power. This is akin to the interference in standard broadcast channels.

The present analysis uses the distributed zero-forcing approach to nullify the ICI, and the DPC to cancel the intra-cell interference present in the system. The advantages of combining distributed zero-forcing and DPC are

- Large scale antenna systems or massive MIMO can fabricate hundreds or thousands of antennas on relatively small terminals. Zero-forcing beamforming can make use of the spatial dimensions made available by large number of antennas and fewer number of users in small cell networks. It was shown that at high SINR regime, zero forcing approaches the optimal performance from an information theoretic perspective [49].
- The distributed zero forcing allows *decoupling* of the resource allocation problems of the BSs to obtain performance bounds drawing from the earlier results on single cell system. The decoupling allows us to ensure zero ICI even without having dedicated infrastructure for cooperation. This also makes possible to construct the dual MAC system from the zero forcing constraint so that the systems are equivalent from throughput perspective.

We assume the data vector of BS l is $\mathbf{x}_l \in \mathbb{C}^K$ and the corresponding beamforming matrices are $\mathbf{W}_l \in \mathbb{C}^{M \times K}$. Hence, the transmit vectors of the BSs are given by $\mathbf{W}_l \mathbf{x}_l \in \mathbb{C}^M$ where $1 \leq l \leq L$. Also assume the channel vector between BS l and user k is $\mathbf{h}_{l,k} \in \mathbb{C}^M$. Subsequently, we will drop the suffix l for brevity when we are dealing with a single cell system.

6.2.1 Zero-Forcing of Intercell Interference

In CBF, the goal is to ensure the zero forcing property of the beamforming vector of a single BS. To analyse zero forcing property mathematically, let us assume the users under BS 1 are indexed $1, 2, \dots, K$, those under BS 2 are indexed $K + 1, \dots, 2K$ etc. The signal received at

user k is given by

$$y_k = \sum_{l=1}^L y_{l,k} + \eta \quad \in \mathbb{C}, \quad \text{where } 1 \leq k \leq KL. \quad (6.1)$$

In (6.1), $y_{l,k} = \mathbf{h}_{l,k}^H \mathbf{W}_l x_l$ is the signal from BS l and η is circularly symmetric standard Gaussian noise satisfying $\mathbb{E}(|\eta|^2) = 1$. Since the BS indexing is arbitrary, consider the BS L which serves K users in its coverage area, and at the same time causes ICI to Q users under BSs $1, 2, \dots, (L-1)$. To zero-force the interference, it is necessary to ensure that

$$y_{L,k} = 0 \quad \text{where } 1 \leq k \leq Q \quad (6.2)$$

By augmenting all the channel vectors to users in other cells row wise, it is possible to obtain the Intercell Interference (ICI) matrix of cell L as follows.

$$\mathbf{H}_L \equiv \begin{bmatrix} \mathbf{h}_{1,L}^H \\ \mathbf{h}_{2,L}^H \\ \vdots \\ \mathbf{h}_{Q,L-1}^H \mathbf{h}_{Q,L}^H \end{bmatrix} \in \mathbb{C}^{Q \times M} \quad (6.3)$$

From the point of view of BS L , to ensure zero-forcing to all intercell users, the necessary condition is

$$\mathbf{H}_L \mathbf{W}_L = \mathbf{0}^{Q \times K}. \quad (6.4)$$

If the beamforming matrix is represented with the individual column vectors as

$$\mathbf{W}_L = [\mathbf{w}_{L,1}, \mathbf{w}_{L,2}, \dots, \mathbf{w}_{L,K}],$$

then the vectors $\mathbf{w}_{L,1}, \mathbf{w}_{L,2}, \dots, \mathbf{w}_{L,K}$ must be in the kernel of ICI matrix $\mathbf{H}_L$.

6.2.2 Kernel of the Intercell Interference Matrix

For non-correlated fading channels, the rank of $\mathbf{H}_L$ can be shown [165] to be Q . Let us assume that $Q + N = M$. This means N is the difference between number of antennas in each BS and number of intercell users. Hence, from the rank-nullity theorem of linear transformation, the kernel of $\mathbf{H}_L \in \mathbb{C}^{Q \times M}$ is an N dimensional subspace of $\mathbb{C}^M$. The goal of the proposed zero forcing strategy is to ensure that the beamforming matrices of every base station lie in the this subspace. Since the subspace is different for each ICI matrix, to characterise it, consider the Singular Value Decomposition (SVD) $\mathbf{H}_L = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^H$, where $\mathbf{U} \in \mathbb{C}^{Q \times Q}$, $\mathbf{\Lambda} \in \mathbb{R}_+^{Q \times M}$ and $\mathbf{V} \in \mathbb{C}^{M \times M}$. The matrices $\mathbf{U}$ and $\mathbf{V}$ contain the right Eigen vectors and left Eigen vectors of the ICI matrix respectively. Now, let us partition the matrix $\mathbf{V}$ as $[\mathbf{Q} \quad \mathbf{E}_L]$, where $\mathbf{Q} \in \mathbb{C}^{M \times Q}$, $\mathbf{E}_L \in \mathbb{C}^{M \times N}$. From orthonormal properties,

$$\mathbf{Q}^H \mathbf{Q} = \mathbf{I}_Q, \quad \mathbf{E}_L^H \mathbf{E}_L = \mathbf{I}_N \quad \text{and} \quad \mathbf{Q}^H \mathbf{E}_L = \mathbf{0}^{Q \times N} \quad (6.5)$$

Theorem 6.1 will characterise the kernel from the singular value decomposition and partitioning described above.

Theorem 6.1 *The columns of $\mathbf{E}_L$ form a set of orthonormal bases for the kernel of $\mathbf{H}_L$.*

Proof: Let us introduce another partition as $\Lambda = [\Delta \quad \mathbf{0}^{Q \times N}]$, where $\Delta \in \mathbb{R}_+^{Q \times Q}$ is the diagonal matrix of singular values. Consider the product

$$\begin{aligned} \Lambda \mathbf{V}^H \mathbf{E}_L &= \Lambda \begin{bmatrix} \mathbf{Q}^H \\ \mathbf{E}_L^H \end{bmatrix} \mathbf{E}_L \\ &= [\Delta \quad \mathbf{0}^{Q \times N}] \begin{bmatrix} \mathbf{0}^{Q \times N} \\ \mathbf{I}_N \end{bmatrix} \\ &= \mathbf{0}^{Q \times N} \end{aligned}$$

using (6.5). Hence, $\mathbf{H}_L \mathbf{E}_L = \mathbf{U} \Lambda \mathbf{V}^H \mathbf{E}_L = \mathbf{0}^{Q \times N}$ ■

Hence, for each cell l , it is possible to find a matrix $\mathbf{E}_l \in \mathbb{C}^{M \times N}$ whose columns are in the kernel of $\mathbf{H}_l$. Now consider a beamforming matrix $\mathbf{W}_l = \mathbf{E}_l \Upsilon_l$ for some arbitrary $\Upsilon_l \in \mathbb{C}^{N \times K}$. It can be shown that $\mathbf{H}_l \mathbf{W}_l = \mathbf{0}^{Q \times K}$, thus satisfying the zero forcing condition. In addition, since the number of columns of $\mathbf{E}_l$ is the same as the nullity of $\mathbf{H}_l$, the columns must span the entire kernel. So any vector in the kernel of $\mathbf{H}_l$ can be represented as a linear combination of the columns. Based on these observations Lemma 6.1 is proposed.

Lemma 6.1 *If the beamforming vectors of the users $\mathbf{w}_{l,1}, \dots, \mathbf{w}_{l,K}$ are in the subspace spanned by columns of $\mathbf{E}_l$, that is a necessary and sufficient condition for interference nullifications for all intercell users.*

Note that for a beamforming matrix $\mathbf{W}_l$ satisfying the zero-forcing condition, the transmit power is given by

$$\mathbb{E}[(\mathbf{W}_l \mathbf{x}_l)^H \mathbf{W}_l \mathbf{x}_l] = \text{Tr}[\mathbf{W}_l^H \mathbf{W}_l] = \text{Tr}[\Upsilon_l^H \Upsilon_l] \quad (6.6)$$

Since each BS has its own version of interference matrix $\mathbf{H}_l$, we have shown that it is possible to decouple the beamforming problems so long as the beamforming vectors lie in the appropriate subspaces. Hence, the challenge is to determine the beamforming vectors obeying the power constraint (6.6) and to achieve maximum sum rate or the Pareto frontier [148]. However, it is difficult to solve the relevant optimisation problems directly. This chapter will show the subspace constraints can be incorporated in an alternative MAC (as opposed to a BC discussed so far) with the duality principle.

6.2.3 CBF Uplink System Model

In the corresponding CBF uplink system (described in Table 6.1), each of the L BSs are receiving K data streams from UEs under its coverage. Each BS treats signals from intercell UEs (as in Figure 6.1) as noise. The goal of the receiver is to zero-force all intercell signals. As before, the UEs indexed by $(l-1)K+1$ to lK are being served by BS l , and consider the BS L without loss of generality. The received vector is given by

$$\mathbf{y}_L = \underbrace{\sum_{k=1}^Q \mathbf{h}_{k,L} x_k}_{\text{Intercell Interference}} + \underbrace{\sum_{k=Q+1}^K \mathbf{h}_{k,L} x_k}_{\text{Intracell Signal}} + \boldsymbol{\eta} \in \mathbb{C}^M \quad (6.7)$$

where $\boldsymbol{\eta}$ is the AWGN satisfying $\mathbb{E}(\boldsymbol{\eta} \boldsymbol{\eta}^H) = \mathbf{I}$.

The BSs employ zero-forcing strategies to cancel the interference. This means the statistic to decode x_k is given by $\mathbf{f}_{k,L}^H \mathbf{y}_L$ for some $\mathbf{f}_{k,L} \in \mathbb{C}^M$. Since the BSs are not aware of transmit powers of intercell UEs, the goal here is to nullify the ICI. Hence, from (6.7), the necessary condition is

$$\mathbf{f}_{i,L}^H \mathbf{h}_{k,L} = 0 \quad \text{where} \quad 1 \leq k \leq Q, \quad Q < i \leq KL \quad (6.8)$$

Condition (6.8) is the dual counterpart of (6.4). Hence, a single-cell dual MAC with reduced degrees of freedom can be constructed from this relation as shown below.

6.2.4 Dual MAC with Degree of Freedom Reduction

The duality principle allows us to characterise the performance parameters (power allocation and receive strategy) of a MAC and construct a dual BC with the same rate region and total power. The single cell duality between receive strategies and beamforming was discussed in earlier works [54, 163]. The goal of this section is to show how the subspace constraints in (6.4) apply to the dual CBF uplink system.

Represent a MAC of single antenna users by the tuples $(\mathbf{H}, \mathbf{p})$, where $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2 \dots \mathbf{h}_K] \in \mathbb{C}^{M \times K}$ and $\mathbf{p} \in \mathbb{R}_{++}^K$. This description contains the number of users K , number of antenna at the BS N , the channel vectors of the users $\mathbf{H}$ and the power constraints $\mathbf{p}$. The system can use the optimal MMSE vectors [159]. Further, to incorporate the subspace constraint of (6.4), introduce another MAC described by $(\mathbf{H}, \mathbf{p}, \mathbf{E})$, where $\mathbf{E} \in \mathbb{C}^{M \times N}$ is a matrix of orthonormal columns. This matrix represents a constraint on every receive beamforming vector $\mathbf{f}_k$, which has to lie in the column space of $\mathbf{E}$. To analyse and optimise the performance of MAC with subspace constraints, the following theorem is proposed. It establishes the information theoretic equivalence between a subspace constrained MAC and reduction of number of antennas.

Theorem 6.2 *The MAC system $\mathcal{M}_1 \equiv (\mathbf{H}, \mathbf{p}, \mathbf{E})$ is information theoretically equivalent to the system $\mathcal{M}_2 \equiv (\mathbf{E}^H \mathbf{H}, \mathbf{p})$ with both interference ignorant decoding and SIC schemes.*

Proof: Without loss of generality, let us assume a decoding order of $1 \rightarrow 2 \rightarrow \dots \rightarrow K$ employing Successive Interference Cancellation (SIC). It will be shown that users in $\mathcal{M}_1$ and $\mathcal{M}_2$ can achieve the same SINRs (and hence the same capacity) under this decoding order. Consider the vector decision statistic for user k in $\mathcal{M}_1$, obtained by cancelling the interferences from users 1 to $k-1$.

$$\begin{aligned} \mathbf{y}_k &= \sum_{i=k}^K \mathbf{h}_i x_i + \boldsymbol{\eta}_M \\ &= \underbrace{\mathbf{h}_k x_k}_{\text{Signal}} + \underbrace{\sum_{i=k+1}^K \mathbf{h}_i x_i}_{\text{Interference}} + \underbrace{\boldsymbol{\eta}_M}_{\text{Noise}} \in \mathbb{C}^M \end{aligned} \quad (6.9)$$

Let the strategy vector for this user be given by $\mathbf{E} \boldsymbol{\lambda}_k$ for some $\boldsymbol{\lambda}_k \in \mathbb{C}^N$. Clearly, it satisfies the subspace constraint and any strategy vector satisfying the constraint can be expressed in

this form. So the scalar decision statistic is

$$\lambda_k^H \mathbf{E}^H \mathbf{y}_k = \underbrace{\lambda_k^H \mathbf{E}^H \mathbf{h}_k x_k}_{\text{Signal}} + \underbrace{\lambda_k^H \mathbf{E}^H \sum_{i=k+1}^K \mathbf{h}_i x_i}_{\text{Interference}} + \underbrace{\lambda_k^H \mathbf{E}^H \boldsymbol{\eta}_M}_{\text{Noise}} \quad (6.10)$$

Similarly consider the scalar decision statistic of user k in $\mathcal{M}_2$ for a strategy vector $\boldsymbol{\lambda}_k \in \mathbb{C}^N$, as

$$\lambda_k^H \mathbf{y}'_k = \underbrace{\lambda_k^H \mathbf{E}^H \mathbf{h}_k x_k}_{\text{Signal}} + \underbrace{\lambda_k^H \sum_{i=k+1}^K \mathbf{E}^H \mathbf{h}_i x_i}_{\text{Interference}} + \underbrace{\lambda_k^H \boldsymbol{\eta}_N}_{\text{Noise}} \quad (6.11)$$

Comparing right hand sides of (6.10) and (6.11), it is easy to see the signal and interferences components are same, for appropriate strategies. The noise components are different, but considering the power of the noise terms, it is possible to show that

$$\mathbb{E}[|\lambda_k^H \mathbf{E}^H \boldsymbol{\eta}_M|^2] = \|\boldsymbol{\lambda}_k\|^2 = \mathbb{E}[|\lambda_k^H \boldsymbol{\eta}_N|^2]$$

Hence, it is possible to achieve the same SINRs for all k users in $\mathcal{M}_1$ and $\mathcal{M}_2$. $\blacksquare$

A conventional MAC with N BS antennas and channels given by $\mathbf{H}$ has N degrees of freedom in designing the strategy vectors. We have shown in Theorem 6.2 that the subspace constraint (imposed by the presence of intercell users and zero interference condition) is accounted by the transformation of the channels by $\mathbf{E}^H \mathbf{H}$. This makes the problems of power allocation, appropriate encoding-decoding orders etc. (addressed in earlier works including [71] and references therein) simpler.

6.3 Algorithm for Zero-forcing

Building on the results from the previous sections, here we present the steps in distributed beamforming based on zero-forcing (for ICI) in the form of Algorithm 6.1. There are several steps involved to achieve the maximum sum rate of each cell while obeying the zero ICI criterion. It was shown in Section 6.2.1 that the beamforming matrix must be at the kernel of the ICI matrix. It is also shown that this constraint can be incorporated in the dual uplink multiple access channel by finding a suitable transformation according to Theorem 6.2. Hence, the process can be repeated L number of times for different base stations. Then standard power allocation algorithms [138] can be applied to the transformed multiple access channels to achieve a maximum sum rate or weighted sum rate. The steps are summarised in Algorithm 6.1. Note that the power allocation problem for each of the L cells is solved independently. No joint optimisation strategy encompassing all cells is necessary in this instance. Hence, if we exploit advances in channel acquisition techniques (including the recently proposed phase shift pilot training [29] among others), it is reasonable to assume that in future systems each BS may have sufficient information of the transmission environment. This can be used for further enhancement in performance. Since $\mathbf{E}_l$ from Theorem 6.1 is the only information needed by each BS, this allows the BSs to act completely autonomously. This observation proves to be a major advantage of the proposed approach from the infrastructural perspective, since it eliminates the need of any dedicated backhaul for cooperation.

Algorithm 6.1: Distributed Zero Forcing

- Require:** $\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \dots, \mathbf{h}_{KL,L} \in \mathbb{C}^M$ and $P_1, P_2, \dots, P_L \in \mathbb{R}_{++}$.
- 1: **for** $1 \leq l \leq L$
 - 2: Construct the interference matrix $\mathbf{H}_l$ according to (6.3).
 - 3: Find the singular value decomposition [166] of $\mathbf{H}_l$.
 - 4: Determine the orthonormal bases for the kernel of $\mathbf{H}_l$ from the right unitary matrix, as outlined in Section 6.2.2. Denote the resulting matrix by $\mathbf{E}_l$
 - 5: Consider the dual MAC system for each BC cell and replace the channels $\mathbf{h}_{l,k}$ by $\mathbf{E}_l^H \mathbf{h}_{l,k}$ according to the transform in Theorem 6.2.
 - 6: Solve the power allocation problems (sum rate maximisation or weighted sum rate maximisation, sum power minimisation etc.) with known algorithms [71, 138].
 - 7: Transform back the power allocation to BC channel according to duality principle [54, Section IV-B]
 - 8: **end**
-

6.3.1 Complexity of the Distributed Zero Forcing Algorithm

The time complexity of Algorithm 6.1 are determined by three parameters of the problem size defined in Section 4.1—viz., L , K and M , which, respectively, denote the number of BSs, number of users served by each BS and number of antennas in each BS. The goal is to calculate the approximate complexity of the significant steps involved. The two most important steps in the iteration of Algorithm 6.1 are finding the bases of the kernel and then sum rate or sum power optimisation.

To find the orthonormal bases, note that $\mathbf{H}_k$ is a $Q \times N$ matrix. A combinatorial and numerical method to find the kernel via rank revealing factorisation was proposed in [167]. According to proposed algorithm, the complexity is given by $\frac{2Q^3}{3} + 4Q^2N$.

The next computationally intense step is to solve the optimisation problem. By the transformation of Theorem 6.2, the effective number of antennas at each BS is $(M - K(L - 1))$. Based on these considerations, the complexity of the sum power minimisation or sum rate maximisation by norm balancing was shown [168] to be $K^2(M - K(L - 1))^3$. Adding these two quantities, solving the problem for each BS in the cluster requires a complexity of

$$K^2(M - KL)^3 + \frac{2Q^3}{3} + 4Q^2M \quad (6.12)$$

Since the computation will be repeated for each BS, multiplying (6.12) with L , the final complexity is $\mathcal{O}(M^3K^5L^4)$.

6.4 Duality between Uplink and Downlink with PDC

In the case of PDC, each pair of UE and BS can communicate in UL or DL directions. Assume in each cluster there are

- L cooperating BSs, each with M antennas,
- K UEs² each with one antenna,
- LK data streams in each direction (UL and DL). The UL and DL communications can take place via time or frequency division duplexing.

²For notational convenience, in PDC schemes we denote the total number of UE by K , unlike the CBF case.

The transmission environment is same as in Figure 6.1. However, the only difference is in the case of PDC, every path carries a useful signal. Hence, the broken lines should be replaced by the continuous lines. The reason is, every BS-UE pair is communicating in both directions. Earlier literature on interference channel [79] did not consider cooperation at either the transmitter or the receiver end. However, in the present system model, since the BSs can have cooperation in terms of channel information, this can be exploited in both UL and DL directions. The details are described below.

6.4.1 Downlink System Model

Each of the L BSs are transmitting K independent data streams to the users. Hence, from each BS's perspective, it is a BC. The BSs are employing a DPC technique to precancel the interference among the transmit streams. For mathematical tractability, let us assume that each BS follows the same user ordering given by $1 \rightarrow 2 \rightarrow \dots \rightarrow (K-1) \rightarrow K$. Thus, each BS encodes the information for UE 1, then encodes UE 2 after presubtracting the interference from UE 1 and so on.

Conversely, since from each user's perspective, it is a MAC, they can employ SIC [53] to decode the incoming signals. It is assumed that each UE decodes the signals in the order given by $1 \rightarrow 2 \rightarrow \dots \rightarrow (L-1) \rightarrow L$. After decoding a BS's signal, it can be cancelled to decode the next BS.

The goal of this section is to establish the duality principle between uplink and downlink. Hence, to analyse the Shannon's capacity of the downlink system, let us introduce the following notations.

- $s_{l,k}^{\text{DL}} \in \mathbb{C}$ is the Gaussian symbol to be transmitted from BS l to UE k , with the constraint $\mathbb{E}(|s_{l,k}^{\text{DL}}|^2) = 1$,
- $\mathbf{w}_{l,k} \in \mathbb{C}^M$ is the beamforming vector for the above symbol,
- $r_{l,k}^{\text{DL}} \equiv$ Data rate from BS l to UE k .

With the given scheme, which uses DPC at the transmitter end and SIC at the receiver end, not all pairs of data symbols interfere with every other pair. The data streams $r_{l,k}^{\text{DL}}$ can be arranged in total order relation of interference. To denote the set of all interfering data streams, let us define $\mathcal{S}^{\text{DL}}(l, k)$ as follows.

$$\begin{aligned} \mathcal{S}^{\text{DL}}(l, k) &\equiv \{(j, i) | s_{ji}^{\text{DL}} \text{ interferes } r_{l,k}^{\text{DL}}\} \\ &\subset \{1, \dots, L\} \times \{1, \dots, K\} \end{aligned}$$

From the given encoding order of DPC and decoding order of SIC, it is seen that

$$(j, i) \in \mathcal{S}^{\text{DL}}(l, k) \Leftrightarrow j \leq l, i > k \quad \text{OR} \quad j > l, i \geq k \quad (6.13)$$

From (6.13), the following recursive properties of the set $\mathcal{S}^{\text{DL}}(l, k)$ can be proved which will be

useful to formulate the duality transformation³.

$$\mathcal{S}^{\text{DL}}(L, K) = \emptyset, \quad (6.14a)$$

$$\mathcal{S}^{\text{DL}}(l-1, k) = \mathcal{S}^{\text{DL}}(l, k) \cup \{(l, k)\} \quad \text{where } 1 < l \leq L \quad (6.14b)$$

$$\mathcal{S}^{\text{DL}}(L, k-1) = \mathcal{S}^{\text{DL}}(1, k) \cup \{(1, k)\} \quad \text{where } 1 < k \leq K \quad (6.14c)$$

So the signal power for $r_{l,k}^{\text{DL}}$ is given by $= |\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2$. If we assume each user has a circularly symmetric AWGN of unit variance, the sum of interference and noise power is

$$a_{l,k} \equiv 1 + \sum_{(j,i) \in \mathcal{S}^{\text{DL}}(l,k)} |\mathbf{h}_{k,j}^H \mathbf{w}_{j,i}|^2 \quad (6.15)$$

Hence, from Shannon's capacity rule

$$r_{l,k}^{\text{DL}} = \log(a_{l,k} + |\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2) - \log a_{l,k} \quad (6.16)$$

Remark: From (6.13), it follows that

$$(j, i) \in \mathcal{S}^{\text{DL}}(l, k) \Leftrightarrow (l, k) \notin \mathcal{S}^{\text{DL}}(j, i) \quad (6.17)$$

$$\text{where } 1 \leq j, l \leq L \quad \text{and} \quad 1 \leq i, k \leq K.$$

It emphasises the total order property [169] of the tuples. The ordering will determine the sequence of transformations to establish the duality principle.

6.4.2 Uplink System Model

Each of the K UEs are transmitting L independent data streams to the BSs. Since each UE transmits multiple data streams, from the UEs' perspective, it is a BC. The UEs are employing a DPC technique to precancel the interference among the transmit streams. As before, for mathematical tractability, let us assume that each UE follows the same BS ordering given by $L \rightarrow (L-1) \rightarrow \dots \rightarrow 2 \rightarrow 1$. Thus, each UE encodes the information for BS L , then encodes BS $(L-1)$ after presubtracting the interference from BS L and so on.

Conversely, since from each BS's perspective, it is a MAC, they can employ SIC to decode the incoming signals. For mathematical tractability, it is assumed that each BS decodes the signals in the order given by $K \rightarrow (K-1) \rightarrow \dots \rightarrow 2 \rightarrow 1$. So UE K 's signal is decoded first, then after subtracting the interference, UE $(K-1)$'s signal is decoded and so on.

Let us introduce the following notations for this UL system to analyse the performance.

- $s_{kl}^{\text{UL}} \in \mathbb{C}$ is the Gaussian symbol to be transmitted from UE k to BS l , with the constraint $\mathbb{E}(|s_{kl}^{\text{UL}}|^2) = 1$,
- $\sqrt{p_{k,l}} \in \mathbb{R}^{++}$ is the beamforming scaler for the above symbol,
- $r_{k,l}^{\text{UL}} \equiv$ Data rate from UE k to BS l .

With the given DPC-SIC, scheme, it is possible to arrive at another total order relation among

³(6.14b) can be proved by comparing $\mathcal{S}^{\text{DL}}(l-1, k)$ and $\mathcal{S}^{\text{DL}}(l, k)$ according to the definition from (6.13). Similarly, (6.14c) can be shown by comparing $\mathcal{S}^{\text{DL}}(L, k-1)$ and $\mathcal{S}^{\text{DL}}(1, k)$.

the data streams as follows.

$$\begin{aligned} \mathcal{S}^{\text{UL}}(k, l) &\equiv \{(i, j) | s_{ij}^{\text{UL}} \text{ interferes } r_{kl}^{\text{UL}}\} \\ &\subset \{1, \dots, K\} \times \{1, \dots, L\} \end{aligned}$$

From the given encoding order of DPC and decoding order of SIC, it is seen that

$$(i, j) \in \mathcal{S}^{\text{UL}}(k, l) \Leftrightarrow i \geq k, j < l \text{ OR } i < k, j \leq l \quad (6.18)$$

As in the DL case, the following recursive properties are useful to formulate the duality transformation.

$$\mathcal{S}^{\text{UL}}(1, 1) = \emptyset, \quad (6.19a)$$

$$\mathcal{S}^{\text{UL}}(k+1, l) = \mathcal{S}^{\text{UL}}(k, l) \cup \{(k, l)\} \text{ where } 1 \leq k < K \quad (6.19b)$$

$$\mathcal{S}^{\text{UL}}(1, l+1) = \mathcal{S}^{\text{UL}}(K, l) \cup \{(K, l)\} \text{ where } 1 \leq l < L \quad (6.19c)$$

Assume that each M antenna base station has an AWGN, whose covariance matrix is $\mathbf{I}$. Then the sum of noise and interference covariance matrix for $r_{k,l}^{\text{UL}}$ can be denoted as

$$\mathbf{B}_{k,l} = \mathbf{I} + \sum_{(i,j) \in \mathcal{S}^{\text{UL}}(k,l)} p_{i,j} \mathbf{h}_{i,l} \mathbf{h}_{i,l}^H \quad (6.20)$$

So, by Shannon's capacity rule,

$$r_{k,l}^{\text{UL}} = \log |\mathbf{B}_{k,l} + p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H| - \log |\mathbf{B}_{k,l}| \quad (6.21)$$

It is seen from (6.16) and (6.21) that the $\mathbf{w}_{l,k}$ and $p_{k,l}$ are the design parameters which affect the achievable rates and also the interferences ($a_{l,k}$ and $\mathbf{B}_{k,l}$) under given DPC–SIC ordering. Hence, the next sections will show how to arrive at a dual UL system from DL (and vice versa) so as to achieve the same data rate with sum power constraint.

6.4.3 Uplink to Downlink conversion

Consider the uplink channel with the given SIC–DPC orderings,

$$p_{k,l} \text{ where } (k, l) \in \{1, 2, \dots, K\} \times \{1, 2, \dots, L\}.$$

To design the dual downlink channel, the following transformation is proposed.

$$\mathbf{w}_{l,k} = \frac{\mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l} \sqrt{a_{l,k} p_{k,l}}}{\sqrt{\mathbf{h}_{k,l}^H \mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l}}} \quad (6.22)$$

By substituting this $\mathbf{w}_{l,k}$ in (6.16) and comparing with (6.21), it is easy to see that the above transformation satisfies $r_{k,l}^{\text{UL}} = r_{l,k}^{\text{DL}}$. However, according to (6.22), to evaluate the beamforming vector $\mathbf{w}_{l,k}$, it is necessary to know $a_{l,k}$ which depends on other beamforming vectors. Here it

is possible to exploit the recursive properties (6.14) to show that

$$a_{L,K} = 1, \quad (6.23a)$$

$$a_{l-1,k} = a_{l,k} + |\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2, \quad (6.23b)$$

$$a_{L,k-1} = 1 + \sum_{j=1}^L \sum_{i=k}^K |\mathbf{h}_{k-1,j}^H \mathbf{w}_{j,i}|^2 \quad (6.23c)$$

The above expressions allow us to calculate the beamforming vectors $\mathbf{w}_{l,k}$ and interferences $a_{l,k}$ starting from the boundary case of (L, K) and using backward induction. The correct order of (l, k) tuples to evaluate the beamforming vectors goes as

$$\begin{aligned} &L, K \rightarrow L-1, K \rightarrow \cdots 1, K \\ &\rightarrow L, K-1 \rightarrow L-1, K-1 \rightarrow \cdots 1, K-1 \\ &\quad \vdots \\ &\rightarrow L, 1 \rightarrow L-1, 1 \rightarrow \cdots 1, 1 \end{aligned}$$

It starts from the boundary value of $l = L, k = K$. It is necessary to reduce the BS index first and when the BS index reaches one, it is necessary to reduce the user index. Eventually, via the above sequence we get $l = 1, k = 1$. The steps are summarised in Algorithm 6.2.

Algorithm 6.2: Uplink to Downlink Conversion

Require: $L, K, p_{k,l}, \mathbf{h}_{k,l}$

- 1: $a_{L,K} \leftarrow 1$
- 2: **for** $k = K : 1$ **for** $l = L : 1$
- 3: $\mathbf{B}_{k,l} \leftarrow \mathbf{I} + \sum_{(i,j) \in \mathcal{S}^{\text{UL}}(k,l)} p_{i,j} \mathbf{h}_{i,l} \mathbf{h}_{i,l}^H$
- 4: $\mathbf{w}_{l,k} \leftarrow \frac{\mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l} \sqrt{a_{l,k} p_{k,l}}}{\sqrt{\mathbf{h}_{k,l}^H \mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l}}}$
- 5: $a_{l-1,k} \leftarrow a_{l,k} + |\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2$ **end**
- 6: $a_{L,k-1} \leftarrow 1 + \sum_{j=1}^L \sum_{i=k}^K |\mathbf{h}_{k-1,j}^H \mathbf{w}_{j,i}|^2$ **end**

6.4.4 Downlink to Uplink conversion

Now consider a given downlink channel with its beamforming vectors $\mathbf{w}_{l,k}$, the channel vectors and the DPC–SIC orderings. The required transformation to obtain the dual uplink channel can be derived by premultiplying (6.22) with $\mathbf{h}_{k,l}^H$ and evaluating $p_{k,l}$ as follows.

$$p_{k,l} = \frac{|\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2}{\mathbf{h}_{k,l}^H \mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l} a_{l,k}} \quad (6.24)$$

To find the correct order to determine $p_{k,l}$ sequentially, we make use of the recursive relations (6.19) to arrive at

$$\mathbf{B}_{1,1} = \mathbf{I}, \quad (6.25a)$$

$$\mathbf{B}_{k+1,l} = \mathbf{B}_{k,l} + p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H, \quad (6.25b)$$

$$\mathbf{B}_{1,l+1} = \mathbf{I} + \sum_{j=1}^l \sum_{i=1}^K p_{i,j} \mathbf{h}_{i,l+1} \mathbf{h}_{i,l+1}^H \quad (6.25c)$$

Hence, the correct sequence of (k, l) tuples to determine $p_{k,l}$ and $\mathbf{B}_{k,l}$ is given by

$$\begin{aligned} &1, 1 \rightarrow 2, 1 \rightarrow \cdots K, 1 \\ &\rightarrow 1, 2 \rightarrow 2, 2 \rightarrow \cdots K, 2 \end{aligned}$$

and continuing the same way,

$$\rightarrow 1, L \rightarrow 2, L \rightarrow \cdots K, L$$

It starts from the boundary values of $k = 1, l = 1$ and via the above sequence, eventually we have $k = K, l = L$. The steps are summarised in Algorithm 6.3. The transformation relations and Algorithms 6.2 and 6.3 give us the methods to convert a PDC uplink channel to downlink and vice versa while maintaining the same achievable performance.

Algorithm 6.3: Downlink to Uplink Conversion

Require: $L, K, \mathbf{w}_{k,l}, \mathbf{h}_{k,l}$

- 1: $\mathbf{B}_{1,1} \leftarrow \mathbf{I}$
 - 2: **for** $l = 1 : L$ **for** $k = 1 : K$
 - 3: $a_{l,k} \leftarrow 1 + \sum_{(j,i) \in \mathcal{S}^{\text{DL}}(l,k)} |\mathbf{h}_{k,j}^H \mathbf{w}_{j,i}|^2$
 - 4: $p_{k,l} \leftarrow \frac{|\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2}{\mathbf{h}_{k,l}^H \mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l} a_{l,k}}$
 - 5: $\mathbf{B}_{k+1,l} \leftarrow \mathbf{B}_{k,l} + p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H$ **end**
 - 6: $\mathbf{B}_{1,l+1} \leftarrow \mathbf{I} + \sum_{j=1}^l \sum_{i=1}^K p_{i,j} \mathbf{h}_{i,l+1} \mathbf{h}_{i,l+1}^H$ **end**
-

6.5 Performance Optimisation with PDC

Based on the duality principle described before, we will analyse the power consumption and performance of the system along with the corresponding resource allocation problem. From the transform relations, following reasoning similar to Appendix B of [54], it can be shown that for the downlink–uplink conversions of (6.24) and (6.22),

$$\sum_{k=1}^K p_{k,l} = \sum_{k=1}^K \mathbf{w}_{l,k}^H \mathbf{w}_{l,k} \quad \text{where } 1 \leq l \leq L \quad (6.26)$$

This means, under the transforms, power is conserved for each base station. Hence, any performance optimisation in the downlink system under BS power constraint can be undertaken by the corresponding optimisation in the uplink system. In this section, a power allocation algorithm will be proposed to maximise the sum rate of all K users in the uplink system.

To get the rate combine (6.21) and (6.25b). We get the rates as

$$r_{k,l}^{\text{UL}} = \log |\mathbf{B}_{k+1,l}| - \log |\mathbf{B}_{k,l}| \quad \text{where} \quad 1 \leq k < K \quad (6.27)$$

Summing over k gives the sum capacity of symbols to the BS l

$$\sum_{k=1}^K r_{k,l}^{\text{UL}} = \log |\mathbf{B}_{1,l} + \sum_{k=1}^K p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H| - \log |\mathbf{B}_{1,l}| \quad \text{where} \quad 1 \leq l \leq L. \quad (6.28)$$

Finally, the sum rate of all data streams can be expressed as

$$\sum_{l=1}^L \sum_{k=1}^K r_{k,l}^{\text{UL}} = \sum_{l=1}^L \left(\log |\mathbf{B}_{1,l} + \sum_{k=1}^K p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H| - \log |\mathbf{B}_{1,l}| \right) \quad (6.29)$$

where $\mathbf{B}_{1,l}$ can be found recursively from (6.19a) and (6.19c).

It is seen that (6.29) is a function of KL independent scalars $\{p_{k,l}\}_{k=1,l=1}^{K,L}$. Hence, the sum rate maximisation problem can be formulated as

$$\max_{p_{k,l} \geq 0} = \sum_{l=1}^L \left(\log |\mathbf{B}_{1,l} + \sum_{k=1}^K p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H| - \log |\mathbf{B}_{1,l}| \right) \quad (6.30)$$

subject to

$$\sum_{k=1}^K p_{k,l} = P_l \in \mathbb{R}_{++} \quad \text{where} \quad 1 \leq l \leq L.$$

The constraints represent the power limitations P_l of the BSs and the obvious physical requirements that the individual powers are non-negative.

In general, (6.29) is not concave with respect to $p_{k,l}$. So although the constraints are convex, it is impossible to use standard interior point optimisation techniques [43], unlike the optimisations discussed in previous MAC systems. Hence, we will propose a suboptimal power allocation scheme based on previous works on single cell MAC systems [168].

6.5.1 Suboptimal Power Allocation

Since global optimisation of (6.29) does not follow the rules of convex analysis, the goal of this section will be to develop a suboptimal power allocation algorithm which exploits the multicell duality transformations described earlier. The sum rate of all UL data streams is given by (6.28) for BS l . From known results of MAC [53], this sum rate can be considered as that of a transformed MAC where BS noise is given by the covariance matrix $\mathbf{B}_{1,l}$. Hence, the UL system can be considered as a L parallel MACs with the corresponding noise covariance matrices. Also note that (6.28) is independent of $p_{i,j}$ when $j > l$. Hence, the L MAC systems can be successively optimised for sum capacity by solving problem (6.31).

$$\max_{p_{k,l} \geq 0} \sum_{k=1}^K r_{k,l}^{\text{UL}} \quad \text{subject to} \quad \sum_{k=1}^K p_{k,l} = P_l. \quad (6.31)$$

Hence, to arrive at the optimal power allocation, the problem has to be solved for each cell sequentially⁴. So the steps are to first solve the power allocation problem (6.31) of BS L in uplink direction. Based on this power allocation, the interference plus noise covariance for the previous cell (indexed by $L - 1$) is determined. Then it is possible to transform the result to a dual downlink system. Hence, it allows us to compute the optimal power allocation for each base station successively. The steps are given in Algorithm 6.4.

Algorithm 6.4: To determine downlink beamforming vectors in PDC for Sum Rate Maximisation

Require: $\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \dots, \mathbf{h}_{K,L} \in \mathbb{C}^M$ and $P_1, P_2, \dots, P_L \in \mathbb{R}_{++}$.

- 1: $\mathbf{B}_{1,1} \leftarrow \mathbf{I}$
- 2: **for** $l = 1 : L$
- 3: $p_{k,l} \leftarrow \text{argmax} |\mathbf{B}_{1,l} + \sum_{k=1}^K p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H|$ (Solve via existing methods [138, 168].)
- 4: **if** ($l < L$) $\mathbf{B}_{1,l+1} \leftarrow \mathbf{I} + \sum_{j=1}^l \sum_{i=1}^K p_{i,j} \mathbf{h}_{i,l+1} \mathbf{h}_{i,l+1}^H$ **end end**
- 5: $a_{L,K} \leftarrow 1$
- 6: **for** $k = K : 1$
- 7: **for** $l = L : 1$
- 8: $\mathbf{w}_{l,k} \leftarrow \frac{\mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l} \sqrt{a_{l,k} p_{k,l}}}{\sqrt{\mathbf{h}_{k,l}^H \mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l}}}$
- 9: **if** ($l > 1$) $a_{l-1,k} \leftarrow a_{l,k} + |\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2$ **end end**
- 10: **if** ($k > 1$) $a_{L,k-1} \leftarrow 1 + \sum_{j=1}^L \sum_{i=k}^K |\mathbf{h}_{k-1,j}^H \mathbf{w}_{j,i}|^2$ **end end**

Bounds on the Suboptimal Performance

Algorithm 6.4 allocates power cell by cell, starting from BS 1 and continuing to BS L . A lower bound can be developed for the achievable sum rate by considering L parallel optimisation problems instead of sequentially optimising the cells as shown in Algorithm 6.4. To show the lower bound, a slightly modified version of Algorithm 6.4 is presented in Algorithm 6.5. In

Algorithm 6.5: To determine downlink beamforming vectors in PDC for Sum Rate Maximisation

Require: $\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \dots, \mathbf{h}_{K,L} \in \mathbb{C}^M$ and $P_1, P_2, \dots, P_L \in \mathbb{R}_{++}$.

- 1: **for** $l = 1 : L$ $\mathbf{B}_{1,l} \leftarrow \mathbf{I}$ **end**
- 2: **for** $l = 1 : L$
- 3: $p_{k,l} \leftarrow \text{argmax} |\mathbf{B}_{1,l} + \sum_{k=1}^K p_{k,l} \mathbf{h}_{k,l} \mathbf{h}_{k,l}^H|$ (Solve via existing methods [138, 168].) **end**
- 4: $a_{L,K} \leftarrow 1$
- 5: **for** $k = K : 1$
- 6: **for** $l = L : 1$
- 7: $\mathbf{w}_{l,k} \leftarrow \frac{\mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l} \sqrt{a_{l,k} p_{k,l}}}{\sqrt{\mathbf{h}_{k,l}^H \mathbf{B}_{k,l}^{-1} \mathbf{h}_{k,l}}}$
- 8: **if** ($l > 1$) $a_{l-1,k} \leftarrow a_{l,k} + |\mathbf{h}_{k,l}^H \mathbf{w}_{l,k}|^2$ **end end**
- 9: **if** ($k > 1$) $a_{L,k-1} \leftarrow 1 + \sum_{j=1}^L \sum_{i=k}^K |\mathbf{h}_{k-1,j}^H \mathbf{w}_{j,i}|^2$ **end end**

Algorithm 6.5, the optimisation at cell l does not take into account the noise and interference generated by cell $(l - 1)$. Hence, by the sequential optimisation of Algorithm 6.4, it is possible to get a better throughput.

⁴This is in contrast with Algorithm 6.1, where the problems are solved parallelly and solutions are independent.

6.6 Rate Region Boundary for Multicell Systems

Throughput optimisation with a given objective function are special cases of characterisation of the rate region boundaries for different communication systems. The objective of this section is to generalise the sum rate problem to broader categories of resource allocation and encoding/decoding orders to achieve the sum rate boundary.

6.6.1 Algorithm to Characterise Pareto Frontier for CBF Systems

To characterise the Pareto frontier of multicell systems while using the CBF cooperation paradigm, similar ideas of cost and reward can be applied. The result stated in Theorem 6.3 (due to [43]) is useful for full characterisation of the Pareto frontier.

Theorem 6.3 *Let $\mathcal{S} \subseteq \mathbb{R}^K$ be a non-empty convex region for $K \in \mathbb{N}$. Let $\mathbf{x} \in \mathbb{R}^K$ be on the boundary of $\mathcal{S}$. Then there exists at least one hyperplane passing through $\mathbf{x}$ and containing $\mathcal{S}$ at one of its halfspaces.*

Since any hyperplane can be represented as a linear combination of tuples, in the form of $a_1x_1 + a_2x_2 + \dots + a_Kx_K$, it is possible to characterise the outer boundary of rate region by varying the optimisation parameters of the objective function. Hence, for a CBF based MAC, the Pareto frontier can be defined by the following theorem.

Theorem 6.4 *A point $\mathbf{x}_0 \in \mathbb{R}^K$ is on the Pareto frontier of a convex rate region $\mathcal{R}$ if and only if*

$$\mathbf{x}_0 = \operatorname{argmax}_{\mathbf{x} \in \mathcal{R}} \mathbf{a}^H \mathbf{x}$$

for some $\mathbf{a} \succeq \mathbf{0}$.

Detailed proofs of Theorems 6.3 and 6.4 can be obtained via some results of convex analysis, as in [43]. To ensure that a specific power allocation among all users is optimal in the Pareto sense, it is sufficient to find the parameters $\mathbf{a}$. To that end, Algorithm 6.6 is proposed. It is assumed that a set of power allocations P_{ij} is supplied for $1 \leq i \leq L$ and $1 \leq j \leq K$. Hence, the individual MAC uplink systems are transformed via Theorem 6.2 to L reduced MACs. These MAC systems can operate parallelly without any ICI. Hence, the power vector is assumed to be the solution and exact coefficients are determined by the KKT method [44]. Algorithm 6.6 summarises the steps to determine the optimisation parameters λ_{lj} in a systematic manner. Note that, λ_{lj} in the algorithm is the same as a_i in Theorem 6.4, where the index l represents the base station. ζ is a dummy variable whose value does not affect the algorithm.

6.6.2 Pareto Frontier for Multicell Systems Employing PDC

For a multicell system employing PDC, it is necessary to consider the system as a cascade of multiple access channels as discussed in Section 6.5.1. Similar to the CBF case, let us assume that a give power allocation is P_{ij} , for $1 \leq i \leq L$ and $1 \leq j \leq K$. Although (6.28) gives the sum rate of all UL data streams towards BS l , that is a special case of the Pareto frontier as discussed earlier. But since (6.28) is independent of $p_{i,j}$ when $j > l$, the optimisation parameters for L MAC systems can be successively computed as in Algorithm 4.1 (in Chapter 4). Hence, the goal is to compute the KKT coefficients λ_{lj} of the following optimisation problem for each cell. As before, ζ is a dummy variable whose value does not affect the algorithm. The steps are given in Algorithm 6.7.

Algorithm 6.6: Pareto Frontier for Distributed Zero Forcing

Require: $\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \dots, \mathbf{h}_{K,L,L} \in \mathbb{C}^M$ and $P_{11}, P_{21}, \dots, P_{LK} \in \mathbb{R}_{++}$.

- 1: **for** $l = 1 : L$,
- 2: Construct the interference matrix $\mathbf{H}_l$ according to (6.3).
- 3: Find the singular value decomposition [166] of $\mathbf{H}_l$.
- 4: Determine the orthonormal bases for the kernel of $\mathbf{H}_l$ from the right unitary matrix, as outlined in Section 6.2.2. Denote the resulting matrix by $\mathbf{E}_l$
- 5: Consider the dual MAC system for each BC cell and replace the channels $\mathbf{h}_{l,k}$ by $\mathbf{E}_l^H \mathbf{h}_{l,k}$ according to the transform in Theorem 6.2.
- 6: Assign any $\zeta > 0$
- 7: **for** $j = 1 : K$ **if** $p_j > 0$ $\mu_j \leftarrow 0$
- 8: **else** Assign any $\mu_j \geq 0$ **end if**
- 9: Assign $\lambda_{lj} \leftarrow \mu_{lj} + \zeta \sum_{i=1}^j w_i \frac{\partial r_i}{\partial p_{lj}}$.
- 10: **end end**

Algorithm 6.7: Pareto Frontier for Partial Data Cooperation

Require: $\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \dots, \mathbf{h}_{K,L} \in \mathbb{C}^M$ and $P_{11}, P_{21}, \dots, P_{LK} \in \mathbb{R}_{++}$.

- 1: **for** $l = 1 : L$
- 2: Assign any $\zeta > 0$
- 3: **for** $j = 1 : K$ **if** $p_j > 0$ $\mu_j \leftarrow 0$
- 4: **else** Assign any $\mu_j \geq 0$ **end if**
- 5: Assign $\lambda_{lj} \leftarrow \mu_{lj} + \zeta \sum_{i=1}^j w_i \frac{\partial r_i}{\partial p_{lj}}$. **end**
- 6: **if** $\mathbf{B}_{1,l+1} \leftarrow \mathbf{I} + \sum_{j=1}^l \sum_{i=1}^K p_{i,j} \mathbf{h}_{i,l+1} \mathbf{h}_{i,l+1}^H$
- 7: **end end**

Figure 6.2: Variation of single cell throughput with BS power for different cooperation schemes.

6.7 Numerical Simulation Results

In this section, we show the efficacy of the proposed beamforming via simulation studies. For the simulations, a cluster of $L = 5$ cooperating BSs is used. Each BS is assumed to have $N = 100$ antennas and each cell has $K = 20$ single antenna users. The noise powers are assumed to be 1mW according to our mathematical formulations and the BS powers are varied from 30 dBm to five times this value, i.e. about 37 dBm. The channel value between each pair of antenna is chosen via $\mathcal{CN}(0, 1)$, i.e. according to Rayleigh distribution. In classical macro cell environments, typically each user will be sufficiently near only to one base station, and distant from most other base stations. However, in the future generation small cell systems, it is possible to have comparable channel strengths with multiple BSs in close proximity. Hence, we do not make any specific assumption about the relative strengths of the channels. The simulations have been performed for multiple realisations of the distributions and the average throughputs are plotted. The objective is to observe the trends in the throughput with varying power constraints.

Figure 6.2 shows the throughputs in a single cell (measured in bit per transmission) within the relevant power range. The distributed zero forcing case refers to a scenario where the power allocation in a single cell is optimised to achieve the highest throughput while causing zero ICI to other users outside the cell. The effectiveness of PDC is demonstrated via two different cases. In both cases, we plot the $\sum_{k=1}^K r_{k,l}^{\text{UL}}$ for a single value of l . In the best case scenario, the specific cell has the highest priority in the UL system, as described in Section 6.4.2. That means, according to the description of Section 6.4.2, $l = 1$, so that it is not affected by the ICI. Conversely, in the worst case scenario, this cell has the lowest priority, corresponding to $l = L$, so that its throughput is affected by ICI from all other cells.

Figure 6.3 shows the total throughput of all users across the cluster with variation of power in the relevant range. We have used four different beamforming techniques to arrive at the plots. The curves respectively correspond to the following techniques.

- Successive Orthogonal Beamforming (SOBF) [170] is used as a benchmark to compare

Figure 6.3: Variation of cluster throughput with BS Powers for different cooperation schemes.

the performance of the given system model.

- CBF via distributed zero forcing technique, as described in Algorithm 6.1.
- Weighted minimum mean square algorithm via DPC, which was proposed in the context of a multicell network under universal frequency reuse [171].
- PDC scheme, which uses the UL–DL duality of Section 6.4 and the Algorithm 6.4 to determine the power allocation.
- JP [22], which entails the maximum degree of cooperation where all the BSs share the full data and channel information to all the users. It effectively transforms the system to a BC with a per group antenna power constraint [172].

It is seen that the performances improve in the order of SOBF, Distributed Zero forcing, PDC and JP. However, PDC offers certain advantages in terms of lower backhaul traffic and also from the point of view of maintaining redundancy/data privacy when the information is divided among multiple BSs. The comparative advantages were discussed in more detail in Section 6.1.2.

Figure 6.4 shows the variation of sum rate (obtained using Algorithm 6.4) with power, when different number of antennas M per BS and all other parameters are constant. The degree of freedom of each BS in determining the zero–forcing beamformer is given by $N \equiv M - Q$. Hence, increasing the number of BS antenna results in corresponding increase in degree of freedom and improved performance. However, it is observed that increasing the number of antennas has a diminishing marginal return on the achievable data rate.

Figure 6.5 shows the Pareto frontiers in a two cell system, i.e. $L = 2$. The sum rate of the cell 1 is plotted along the horizontal axis and sum rate of cell 2 is plotted along the vertical axis. It is shown that as power is transferred between the cells, the achievable sum rates change according to the trade off. The two curves show the outer boundaries of rate regions for two different cooperation schemes, viz. CBF and PDC. The broken line shows achievable rates for PDC, while the continuous line shows achievable sum rate for CBF. It is seen that PDC outperforms CBF in terms of achievable rate even though it entails more overhead backhaul traffic.

Figure 6.4: Cluster throughput variation with BS power shown for different number of antennas.

Figure 6.5: Achievable Pareto Frontiers in a Two Cell System under CBF and PDC.

6.8 Conclusion

The chapter studied two important cooperation paradigms applicable to multicell environments. It showed the existence of UL–DL duality for CBF, as well as the novel paradigm of PDC. It is argued that CBF systems can achieve zero ICI without any explicit cooperating backhaul infrastructure, if each BS is aware of its own channels. However, in the presence of a higher capacity backhaul network, the novel PDC paradigm can be employed. Since the power allocation problem in the case of PDC is intractable, a suboptimal algorithm has been proposed based on earlier norm balancing approach and the results are compared with prior art. From the simulation results, JP is observed to outperform PDC. However, it is argued that PDC offers advantages in terms of the existence of duality principle and simpler backhaul infrastructure compared to JP. Further optimisation of performance can be attempted by a more comprehensive modelling of full-duplex self-interference [109] and optimising the encoding–decoding order based on channel strength.

Chapter 7

Conclusions and the Way Ahead

COOPERATIVE communication has been an important focus of research to meet the rising demand of throughput. As it offers exploitation of the spatial dimensions, several forms of cooperation have been established as possible candidates for future wireless communication systems. However, along with the advantages, the resource allocation and interference management in cooperative frameworks come with some unique challenges. Achieving the right trade off among power requirement, implementation complexity and throughput performance is a fundamental objective in this context. The goal of this thesis is to address the optimisation challenges and practical aspects of information transmission in cooperative communication systems.

7.1 Conclusions of the Present Research

The present research has been motivated by some fundamental trends evident in research and application landscapes in wireless communication. Radio spectrum is a most scarce and heavily regulated resource for wireless traffic and it should be efficiently utilised to meet the throughput demand. When multiple transmitters share the same time and frequency resource blocks to send data, the resulting interference critically affects the link throughputs. Although many interference suppression techniques have been proposed with different system models, the throughput gain so far cannot meet the projected demand of wireless traffic. Further, as portable devices become increasingly popular with lower cost, their potential to act as relays is yet to be exploited in practical systems. When a single relay is placed among multiple interfering users, the system can be considered as a cascade of broadcast and MACs. For each user, the end to end capacity is the minimum of two capacities achieved in two subsystems. Improving user throughputs by preventing the bottleneck in heterogeneous cellular systems is one principal objective of the thesis.

In a broadcast or downlink channel where a single centralised transmitter sends (possibly different) messages to disparate users, it is reasonable to assume that the users draw power from a common source, resulting in a collective power constraint instead of individual power constraints. This gives a greater degree of freedom while designing a broadcast system and should be considered to characterise the rate region boundary.

Duality relations among achievable QoS parameters help us to transform the problems in the BC to convex problems in MACs. However, a MAC being considered as the dual of a BC must

exploit the additional freedom of power allocation traditionally considered only in a BC. The duality principle offers the advantage of solving the typical intractable problems of broadcast system via convex methods available for multiple access or uplink systems. The sum capacity expression is a concave function of power allocation and the feasible power allocations have to obey linear constraints. Because of these properties, the sum rate maximisation problem is a standard convex optimisation problem which can be solved using disciplined convex programming methods. In this context, a novel algorithm is proposed for optimising the sum-capacity of a MAC under sum power constraint, where the users are coupled by individual power usage as well as interference. Optimal resource allocation has a certain distributed feature which can be exploited to formulate the algorithm. In the proposed algorithm, when each set of users has their total power distributed in a way so as to maximise the sum rate while treating other users' signal as noise, the sum rate of all users is maximised. This work compared the run time of disciplined convex programming method against the proposed norm balancing approach. The norm balancing approach is shown to have better performance for a reasonable problem size.

Then the solution of the sum capacity maximisation problem has been used to achieve the Pareto frontier of a BC. The Pareto frontier for a specific power allocation were known in literature [48]. However, the Pareto frontier for a sum power constraint is critically dependent on the decoding order of users and its importance in understanding the rate region is another novel aspect of work. An optimal user ordering procedure had been proposed for degraded BCs [45, 63], but this work extends the procedure to non-degraded channels. It is to be noted that for a degraded channel, maximising the sum rate implies using the total power on the strongest user [63, Appendix B] and the sum rate maximisation problem becomes a trivial problem. Additionally, for a degraded channel, only a single encoding order achieves a point on the boundary irrespective of power allocation. This is not the case for non-degraded MISO channels, where the power allocation plays an important role to determine the optimal decoding order.

Eventually, the concepts of sum capacity maximisation and Pareto frontier for a single cell system have extended for a heterogeneous system modelled by a DF relay. In this contexts, the concepts of cost and reward functions have been introduced which brings several classes of optimisation problem under a unified framework. It is shown that achieving a joint Pareto frontier is a necessary condition for optimal resource utilisation in a heterogeneous network. Two algorithms are proposed to solve the problem of joint Pareto frontier.

Traditional cellular systems depend on coverage areas partitioned into disjoint cells where each user is served by a single base station, without any cooperation. To deal with the resulting intercell interference, this work combined the advantages of Cooperative Multipoint Transmission along with detection and broadcasting techniques suitable for large scale antenna systems (massive MIMO). Two novel interference mitigation techniques, such as Centralised Interference Cancellation and Distributed Interference Cancellation are proposed to optimise resource allocation in a cooperative framework. The schemes illustrate a trade off between achievable performance and computational complexity. A large number of antennas entails mathematical manipulation of big matrices to arrive at optimal beamforming vectors. To overcome the difficulty, a suboptimal two layer decoding method is proposed as an improvement to earlier techniques. The proposed method combines selection diversity with least mean square filtering.

Finally, the work explores novel resource allocation algorithms with two different cooperation paradigms, viz. the coordinated beamforming and partial data cooperation. The well-known

duality principle, which has been used to solve uplink and downlink optimisation problems in previous literature, is also exist for multicell systems with both coordinated beamforming and the novel paradigm termed as partial data cooperation. Using the generalised version of duality, this thesis solves relevant resource allocation problems in both cases. The proposed power allocation problems are shown to outperform some of the existing optimisation techniques in terms of the throughput while having significant conceptual and theoretical advantage.

7.1.1 Application of Proposed Algorithms in LTE-5G

The challenges in the system level development of 5G communication systems and continuing the long term evolution stem from two major necessities.

- Demand for data rate
- Reduce the energy consumption compared to the 4G and pre 4G networks, which is important from practical power constraints as well as from an environmental perspective.

To be noted, the Information Communication Technology (ICT) has been collectively responsible for 2-3% of annual carbon emission in recent years [173], which is about the same as the aviation industry and a fourth of the automobile industry. Hence, innovative optimisations at various layers to reduce the carbon footprint while meeting the traffic demand is a primary concern. The importance of Green Communication network for sustainable development was recently also highlighted in [174].

In light of these, the goal of these thesis has been to underline the importance of inter-device and inter-BS cooperation in resource management, with the goal to achieve optimal performance from an information theoretic perspective. Given the device connectivity centric nature of the upcoming standard (envisioned in [27] and [28], as two examples) it is expected the cooperative algorithms and analyses presented throughout the thesis will inform future research in this area and similar techniques can be applied to solve optimisation problems in other system models as well.

7.2 Future Research Directions

This work focused mainly on the aspects of interference mitigation via interference pre-cancellation or via successive decoding. However, the use of interference alignment and spatial degrees of freedom (which was reviewed in Section 2.3.3) to a relay channel is not present in the literature as of now. A possible interference alignment approach can involve user selection algorithms [164] such that interferences on both sides of the relay lie on a restricted subspace. Multiuser diversity can be used to select the users with most appropriate channel matching the criteria.

A standard system model with relay was depicted earlier. It is also possible to consider a generalised broadcast system such as that proposed in [175], which can be conveniently merged to existing cellular setups as shown in Figure 7.1. In the proposed framework, two alternative modes of transmission are considered. One transmission mode is direct mode, which completely ignores the relay. Alternatively, the transmission can take place via the relay mode through the AF relay. The same work can be extended to the case of a DF relay. It is expected that if the DF relays replace the AF relays, the network will have a performance improvement, since similar comparison was made in [176] with a different architecture. The interference management and rate region characterization for this generalised channel is a challenging area

Figure 7.1: Broadcast Channel with Relay

Figure 7.2: Multiple Access Channel with Relay

of future research. The corresponding MAC with a relay in-between is shown in Figure 7.2. The duality between the systems in Figures 7.1 and 7.2 are not known till date. Existence of any correspondence between achievable rate regions of these two systems will greatly simplify the optimisation approaches, leading to more efficient algorithms. The problem formulation in the present research focuses on multiple users being coordinated by a single relay. However, [177] used a more general scenario as in Figure 7.3 where the coordination is done by multiple parallel relays (without any cooperation among themselves) in a two hop network. As before, there are K non-cooperating users aided by (in this instance) R non-cooperating relays, each containing M antennas¹. The channel from user i to relay j is denoted by $\mathbf{h}_{ij} \in \mathbb{C}^M$ and the channel from relay j to user i is $\mathbf{g}_{ji}^H \in \mathbb{C}^M$ for $1 \leq i \leq K$ and $1 \leq j \leq R$. Initially no assumption is made about relative values of K and R .

The system model can be taken as a cascaded system of interfering BCs or interfering MAC (following the approach of [41]). However, in [177] it was done with AF relays without $\mathbf{r} = \mathbf{r}'$ constraint. In a DF system using multiple relays, even the relay power allocation must be a vector instead of a scalar quantity, which presents a more challenging problem.

As the increasing bandwidth demand from real time applications strains the available re-

¹Initially the number of antennas in each relay is assumed to be same to keep the mathematical formulation simpler

Figure 7.3: Multiantenna Parallel Relay System

sources, it becomes imperative to accommodate more and more geographically spread users in any time, frequency dimensions. So far as base stations and other non-portable infrastructure based terminals are concerned, spatial dimension is the one in which the system can be scaled most conveniently, even though it does not come without its own challenges. To begin with, the massive MIMO paradigm is characterised by a massively large number of antennas attached to a base station and possibly, proportional number of users (see [153] for an overview). In each transmit period, a small subset of antennas focus their transmit energy in predetermined transmit directions (which may use some variations of Eigen mode beamforming or Zero forcing). This makes a better use of spatial dimensions by greatly enhancing the capacity and flexibility, while introducing more challenging system complexity. If multiple smaller relays are used to process signals transmitted to or arriving from such massive terminals, the variation of the corresponding rate region needs further investigation. For example, [178] proposed the capacity scaling analyses of relay networks which can be extended in the direction of the system models discussed earlier. [179] carried out the analyses from an algorithmic information theory perspective but complete characterisation of the achievable rate region in presence of multiple relays is still not available. [152] compared the available beamforming techniques, as they might apply to massive MIMO systems, which show significant performance deviation.

In summary, both multiple relay and multicell communications form two uncharted territories with great potentials and their application together will be an interesting area of research.

Appendix A

Review of Inner Product Space

Inner product and norm generalise the concepts of dot product and length, respectively to more abstract vector spaces. This appendix will review some of the concepts starting with basic definitions of some algebraic structures. The focus will be more on the application perspective of linear algebra as it applies to multiuser interference channels and some mathematical rigour will be compromised. For a more detailed discussion and basics of abstract algebra, see [180].

A.1 Fields & Vector Space

Definition A.1 *A field $\mathcal{F}$ is a set on which the commutative operations addition and multiplications are defined, there are unique inverses and identity elements with respect to both operations, multiplication is distributive over addition, and subtraction and division are defined.*

Some examples are $\mathbb{R}$ and $\mathbb{C}$. Note that for a set to be classified as field, the elements need not be ordered. Even though many other examples of finite and infinite elements fields are found in abstract algebra and coding theory, this work will be mainly concerned with the above two fields of real and complex numbers. In subsequent discussion, a field element will be referred to as scalar.

From the definition of field, a vector space $\mathcal{V}$ can be defined over a field $\mathcal{F}$ in which certain properties are defined. The most common example of a vector space over any field $\mathcal{F}$ is $\mathcal{F}^n$ for some natural number n . The properties to characterise a vector space can be found in [180].

Another example of vector spaces is the space of matrices of fixed dimension, consisting of elements from the field, i.e. $\mathcal{F}^{m \times n}$.

A.2 Inner Product Space & Norm

The concepts of inner product and norm define further structure for vector spaces defined over the fields of real and complex number. Let us denote the underlying field as $\mathcal{F}$ for sake of generality (which is $\mathbb{C}$ or $\mathbb{R}$) and the vector space as $\mathcal{V}$.

Definition A.2 *The mapping $\langle \bullet, \bullet \rangle : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{F}$ is an inner product if it satisfies the following conditions.*

Conjugate Symmetry : $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$.

Sesquilinear Property : Linear with respect to second argument and anti-linear with respect to the first¹.

Positive Definite Property : $\langle \mathbf{x}, \mathbf{x} \rangle > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$.

Once an inner product is defined for a vector space $\mathcal{V}$, it is called an inner product space. The dot product defined on the euclidean space $\mathbb{R}^3$ is an inner product.

Like the inner product, the concept of length can be generalised to norm.

Definition A.3 The mapping $\|\bullet\| : \mathcal{V} \rightarrow \mathbb{R}_+$ is a norm if the following conditions are satisfied.

Triangle Inequality : $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$

Positive Definite Property : $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

Homogeneity : $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathcal{V}, \alpha \in \mathcal{F}$

Any vector space for which a norm has been defined is called a normed vector space. Apart from the well known Euclidean norm, a more general definition is the $\mathcal{L}_p$ norm for a natural number p defined over $\mathcal{V} = \mathbb{R}^n$ as

$$\|\mathbf{x}\|_p \equiv \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

It is easy to check that the above definition of norm satisfies the necessary properties of the definition.

Any inner product space $\mathcal{V}$ can be taken as a normed vector space when the norm is defined in terms of the inner product. The definition is

$$\|\mathbf{x}\| \equiv \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \quad \forall \mathbf{x} \in \mathcal{V}$$

This definition satisfies the positive definite and homogeneity criteria of norm (since it is included in the definition of inner product itself). For the triangle inequality condition, it is satisfied because of the existence of Cauchy-Schwarz inequality (see Theorem A.1).

A.2.1 Norm & Inner Product with Respect to a Square Matrix

For the vector space $\mathcal{F}^{m \times n}$ defined over the field $\mathcal{F}$ (which is either $\mathbb{R}$ or $\mathbb{C}$), the following inner product can be defined with respect to a fixed Hermitian positive definite matrix $\mathbf{A} \in \mathcal{F}^{m \times m}$ as

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathbf{A}} \equiv \text{trace}(\mathbf{X}^H \mathbf{A} \mathbf{Y})$$

and the corresponding norm as $\|\mathbf{X}\|_{\mathbf{A}} \equiv \sqrt{\text{trace}(\mathbf{X}^H \mathbf{A} \mathbf{X})}$. These definitions satisfy the necessary criteria for $\mathbf{A} \succ \mathbf{0}$ and hence the Cauchy-Schwarz inequality holds. This definition has

¹The ordering of arguments is arbitrary. However, this is the convention used in this work.

been used for the proposed norm balancing algorithm to find optimal power allocation in dual multiple access channel (see Chapter 3). Note that for $n = 1$ and $\mathbf{A} = \mathbf{I}_m$, this definition reduces to the familiar dot product and Euclidean length in the corresponding vector space.

A.3 Cauchy Schwarz Inequality

The Cauchy Schwarz inequality imposes upper bound on the absolute value of inner product of two vectors, as defined in any space over the field of real or complex numbers. A basic proof is given here. For a detailed discussion on related concepts and implications, see [181]. To prove the inequality, the following lemma² will be used.

Lemma A.1 *For the normed vector space $\mathcal{V}$ over $\mathcal{F}$, if $\mathbf{x}$ and $\mathbf{y}$ are two orthogonal vectors, i.e. $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, then $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$*

Now it is possible to prove the following inequality.

Theorem A.1 (Cauchy Schwarz inequality) *For the normed vector space*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$$

and

$$|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \|\mathbf{y}\| \quad \Leftrightarrow \quad \mathbf{x} = \alpha \mathbf{y}$$

for some $\alpha \in \mathcal{F}$.

Proof: If either of the vectors is $\mathbf{0}$, equality holds trivially and the linear dependence condition is also satisfied with $\alpha = 0$. So assume that $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$.

Now consider the vector $\mathbf{z} \equiv \mathbf{y} - \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|^2} \mathbf{x}$

It is easy to check that $\mathbf{x}$ and $\mathbf{z}$ are mutually orthogonal. Hence $\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|^2} \mathbf{x}$ and $\mathbf{z}$ are orthogonal as well³.

$$\mathbf{y} = \mathbf{z} + \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|^2} \mathbf{x} \tag{A.1}$$

Taking norms of both sides in A.1 and using Lemma A.1,

$$\|\mathbf{y}\|^2 = \|\mathbf{z}\|^2 + \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\|\mathbf{x}\|^2} \geq \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\|\mathbf{x}\|^2} \tag{A.2}$$

The inequality in A.2 is an equality if and only if $\mathbf{z} = \mathbf{0} \Leftrightarrow \mathbf{x} = \frac{\|\mathbf{x}\|^2}{\langle \mathbf{x}, \mathbf{y} \rangle} \mathbf{y}$ ■

²This is an extension of Pythagoras's theorem to more abstract vector spaces and is easy to prove using the definitions.

³Scalar multiplication does not alter orthogonality.

Appendix B

Set Functions & Polymatroids

Polymatroids are an important class of algebraic structures relevant in the fields of combinatorics and computational geometry. It was introduced in [127] and the related concepts have interesting applications in the field of multiple access systems [48]. The current appendix will introduce the basics of these algebraic structure keeping the relevant applications in mind and avoiding the mathematical abstractions so far as possible. Before defining the geometrical structure, the basics of submodular property will be introduced.

A mapping whose domain is a family of sets and co-domain is the field of real numbers or some measure space is called a set function.

As an example, for any countable finite set $\mathcal{U}$, $f : 2^{\mathcal{U}} \rightarrow \mathbb{R}$ can be a set function. In this work, the focus will be on a set of distinct users.

B.1 Submodular Set Function

This special category of set function extends the concept of *diminishing marginal return* to the case when argument of the function is a set instead of a real number.

Definition B.1 *The set function $f : 2^{\mathcal{U}} \rightarrow \mathbb{R}$ is submodular if it satisfies the following property.*

$$f(\mathcal{S}) + f(\mathcal{T}) \geq f(\mathcal{T} \cap \mathcal{S}) + f(\mathcal{T} \cup \mathcal{S}) \quad \text{when } \mathcal{S}, \mathcal{T} \subseteq \mathcal{U} \quad (\text{B.1})$$

To see how the submodular property captures the diminishing marginal return, consider $\mathcal{S} \subset \mathcal{T} \subset \mathcal{U}$ and an element $e \in \mathcal{T}'$. From the submodular property,

$$\begin{aligned} f(\mathcal{S} \cup \{e\}) + f(\mathcal{T}) &\geq f(\mathcal{S}) + f(\mathcal{T} \cup \{e\}) \\ \Rightarrow f(\mathcal{S} \cup \{e\}) - f(\mathcal{S}) &\geq f(\mathcal{T} \cup \{e\}) - f(\mathcal{T}) \end{aligned} \quad (\text{B.2})$$

Effect of adding a new element on any subset is captured in the above proof and as B.2 shows, the bigger the initial set, smaller is the gain from adding an element. Moreover, B.2 can be used to define the submodularity and B.1 can be derived easily from the definition. It is explained in Section 2.2.1 that the submodular property is helpful to model the interference among different users in a multiple access system.

B.2 Rank Function & Polymatroid

Rank function is an important concept in polymatroid theory and a workable definition will be provided next.

Definition B.2 *The submodular function $f : 2^{\mathcal{U}} \rightarrow \mathbb{R}_+$ is called a rank function if it satisfies the following two properties.*

Normalised : $f(\emptyset) = 0$

Monotonic : $f(S) \leq f(T)$ when $S \subseteq T$

The definition is not absolutely rigorous since the objective here is to avoid excessive abstraction wherever possible. However, as seen in Section 2.2.1, the definition provides an elegant formulation of the achievable rate region in a multiple access channel and will suffice for our purpose. See [182] for a more detailed and rigorous treatment of associated concepts. The concept of optimisation with a matroid constraint has been discussed in [183] and adapted in [48] for application to multiple access interference systems.

Definition B.3 *For a rank function $f : 2^{\mathcal{U}} \rightarrow \mathbb{R}_+$ and $K = |\mathcal{U}|$, the region*

$$\mathcal{B}(f) \equiv \{\mathbf{x} | \mathbf{x} \in \mathbb{R}_+^K, \sum_{i \in \mathcal{S}} x_i \leq f(\mathcal{S}) \text{ where } \mathcal{S} \subseteq \mathcal{U}\} \subseteq \mathbb{R}_+^K \quad (\text{B.3})$$

is a polymatroid.

It should be obvious that a polymatroid is a convex polytope in a vector space. As proved in [127], a polymatroid in K dimension has $K!$ vertices in $\mathbb{R}_{++}^K$ and in the context of a K user multiple access channel, they correspond to different decoding orders of users.

Appendix C

Duality with Linear Beamforming

In the vector MAC, the zero-mean unit-variance data signal s_k for user k with $\mathbb{E}(s_k) = 0$ and $\mathbb{E}(|s_k|^2) = 1$ is weighted with the beamformer $\mathbf{w}_k \in \mathbb{C}^M$ to generate k -th additive component of the transmitted signal.

The total transmitted signal for K single antenna users and M antennas base station

$$\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k \in \mathbb{C}^M$$

which is transmitted over the k -th channel $\mathbf{h}_k^T \in \mathbb{C}^{1 \times M}$ and perturbed by the additive zero-mean noise $\eta_k \in \mathbb{C}$ $\mathbb{E}(\eta_k) = 0$ and $\mathbb{E}(|\eta_k|^2) = \sigma_{\eta_k}^2$ to form the received signal of user k

$$y_k = \mathbf{h}_k^T \mathbf{w}_k s_k + \mathbf{h}_k^T \sum_{i=1, i \neq k}^K \mathbf{w}_i s_i + \eta_k \in \mathbb{C} \quad (\text{C.1})$$

The signal to interference and noise ratio is given by

$$\text{SINR}_k^{\text{BC}} = \frac{\mathbb{E}[|\mathbf{h}_k^T \mathbf{w}_k s_k|^2]}{\mathbb{E}[|\mathbf{h}_k^T \sum_{i=1, i \neq k}^K \mathbf{w}_i s_i|^2] + \mathbb{E}[|\eta_k|^2]} = \frac{|\mathbf{h}_k^T \mathbf{w}_k|^2}{\sum_{i=1, i \neq k}^K |\mathbf{h}_k^T \mathbf{w}_i|^2 + \sigma_{\eta_k}^2} \quad (\text{C.2})$$

Here the standard assumptions are the data symbols of different users are uncorrelated, i.e. $\mathbb{E}(s_i s_j^*) = 0$ where $i \neq j$) and that the symbols and the noise are uncorrelated, i.e. $\mathbb{E}(s_i \eta_j^*) = 0$. Note that above SINR is independent of a non-zero weighting of the received signal x_k , since all signal portions are weighted by the same factor and consequently, also their powers.

The transmit power used in the vector broadcast channel can be written as

$$P^{\text{BC}} = \mathbb{E}[\|\mathbf{y}\|^2] = \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \quad (\text{C.3})$$

The vector multiple access channel has the same channel vectors $\mathbf{h}_k^T \in \mathbb{C}^{1 \times N}$ for user k , but

weighted with $\frac{1}{\sigma_{\eta_k}}$ to account for the different noise variances in the vector BC. If the vector MAC is constructed to solve problems originally formulated in the vector BC, covariance of the non-zero noise signals $\eta \in \mathbb{C}^N$ can be taken to be identity, i.e. $\mathbb{E}[\eta] = \mathbf{0}_N$ and $\mathbb{E}[\eta\eta^H] = \mathbf{I}_N$. The resulting SINR for user k in vector MAC (with receive filter coefficients $\mathbf{f}_k$)

$$\text{SINR}_k^{\text{MAC}} = \frac{\mathbb{E}[|\mathbf{f}_k^H \mathbf{h}_k^* \sigma_{\eta_k}^{-1} t_k s_k|^2]}{\mathbb{E}[|\mathbf{f}_k^H \sum_{i=1, i \neq k}^K \mathbf{h}_i^* \sigma_{\eta_i}^{-1} t_i s_i|^2] + \mathbb{E}[|\mathbf{f}_k^H \eta|^2]} = \frac{|\mathbf{f}_k^H \mathbf{h}_k^* \sigma_{\eta_k}^{-1} t_k|^2}{\sum_{i=1, i \neq k}^K |\mathbf{f}_k^H \mathbf{h}_i^* \sigma_{\eta_i}^{-1} t_i|^2 + \|\mathbf{f}_k\|^2} \quad (\text{C.4})$$

Since a non-zero scalar weighting of the vector MAC equalizer $\mathbf{g}_k$ does not change $\text{SINR}_k^{\text{MAC}}$, as all terms are scaled equally, it is possible impose the additional constraint $\|\mathbf{g}_k\|_2 = 1$ without loss of generality. The total transmit power in MAC is

$$P^{\text{MAC}} = \mathbb{E}[\|y_k\|^2] = \sum_{k=1}^K |t_k|^2 \quad (\text{C.5})$$

In a MAC, the transmitters are spread through a coverage area without a common power source. While a constraint of the total transmit power P^{BC} in the broadcast channel is plausible, a constraint of the total transmit power P^{MAC} in the multiple access channel is impractical. Hence, to derive the SINR duality, the total transmit powers P^{BC} and P^{MAC} will be compared. This illustrates that even if the sum power constrained multiple access channel has no physical meaning, it is constructed to solve SINR problems in the vector broadcast system.

C.1 MAC to BC Conversion

Starting with filters $\{t_k\}_{k=1}^K$ and $\{\mathbf{f}_k\}_{k=1}^K$ given in the vector MAC, we find the precoders in the vector BC with the transform

$$\mathbf{w}_k = \alpha_k \mathbf{f}_k \quad \text{where} \quad k \in \{1, 2, \dots, K\} \quad (\text{C.6})$$

So the precoders are defined besides the scalars $\alpha_k = \mathbb{R}_+$ where $k \in \{1, 2, \dots, K\}$ that are chosen such that the SINRs in the vector BC are the same as the SINRs in the vector MAC, i.e.

$$\text{SINR}_k^{\text{BC}} = \text{SINR}_k^{\text{MAC}} \quad \text{where} \quad k \in \{1, 2, \dots, K\}$$

The power used for user k in the vector BC simplifies to

$$\|\mathbf{w}_k\|_2^2 = \|\mathbf{f}_k\|_2^2 \alpha_k^2 = \alpha_k^2$$

using the normalisation carried out. Thus, the problem of finding the α_k -s that lead to the same SINRs in the vector BC as in the vector MAC is in fact a power allocation problem. Applying C.6 to C.2 and with C.4, after some algebraic manipulation

$$\left(\sum_{i=1, i \neq k}^K |\mathbf{h}_i^T \mathbf{f}_k|^2 \sigma_{\eta_i}^{-2} |t_i|^2 + 1 \right) \alpha_k^2 - \sum_{i=1, i \neq k}^K |\mathbf{h}_k^T \mathbf{f}_i|^2 \sigma_{\eta_k}^{-2} |t_k|^2 \alpha_i^2 = |t_k|^2 \quad \text{where} \quad k \in \{1, 2, \dots, K\} \quad (\text{C.7})$$

It is a square system of linear equations with K unknowns $\{\alpha_k^2\}_{k=1}^K$ which can be more concisely written as

$$\mathbf{\Gamma} \mathbf{a} = \boldsymbol{\tau} \quad (\text{C.8})$$

with $\mathbf{a} = [\alpha_1^2, \alpha_2^2, \dots, \alpha_K^2]^T \in \mathbb{R}_+^K$, $\boldsymbol{\tau} = [|t_1|^2, |t_2|^2, \dots, |t_K|^2]^T \in \mathbb{R}_+^K$ and

$$\gamma_{k,j} = \begin{cases} \sum_{i=1, i \neq k}^K |\mathbf{h}_i^T \mathbf{f}_k|^2 \sigma_{\eta_i}^{-2} |t_i|^2 + 1 & \text{where } j = k \\ -|\mathbf{h}_k^T \mathbf{f}_i|^2 \sigma_{\eta_k}^{-2} |t_k|^2 & \text{for } j \neq k \end{cases}$$

$\mathbf{\Gamma}$ is a strictly column diagonally dominant matrix with positive diagonal and non-positive off-diagonal entries, $\boldsymbol{\tau}$ is a positive real vector and the entries of the resulting $\mathbf{a}$ are positive. Consequently, the coefficients $\{\alpha_k\}_{k=1}^K$ in the transform of C.6 can always be chosen such that the resulting SINRs in the vector BC are the same as that in the vector MAC.

Summing up all scalar equations by multiplying with the all-ones vector from the left shows that equating the SINRs in vector BC and the vector MAC also leads to the same total transmit power in the vector BC as that in the vector MAC:

$$\mathbf{1}^T \mathbf{\Gamma} \mathbf{a} = \mathbf{1}^T \boldsymbol{\tau} \quad (\text{C.9})$$

giving

$$\sum_{k=1}^K \left[\left(\sum_{i=1, i \neq k}^K |\mathbf{h}_i^T \mathbf{f}_k|^2 \sigma_{\eta_i}^{-2} |t_i|^2 + 1 \right) \alpha_k^2 - \sum_{i=1, i \neq k}^K |\mathbf{h}_k^T \mathbf{f}_i|^2 \sigma_{\eta_k}^{-2} |t_k|^2 \alpha_i^2 \right] = \sum_{k=1}^K \alpha_k^2 = \sum_{k=1}^K |t_k|^2 \quad (\text{C.10})$$

The left-hand side is the total transmit power P^{BC} in the BC and the right-hand side is the total transmit power P^{MAC} in the MAC. C.10 shows that the total transmit power is conserved, when switching from the vector MAC to the vector BC with the same SINR.

C.2 BC to MAC Conversion

To fulfil the unit norm constraint, the transform to get the vector MAC filters from the vector BC filters is

$$\mathbf{g}_k = \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|_2} \quad \text{where } k \in \{1, 2, \dots, K\} \quad (\text{C.11})$$

The scalar weights $\{t_k\}_{k=1}^K$ applied at the transmitters in the vector MAC must be chosen such that

$$\text{SINR}_k^{\text{MAC}} = \text{SINR}_k^{\text{BC}} \quad \text{where } k \in \{1, 2, \dots, K\}$$

Consequently the following set of equations must be solved.

$$\boldsymbol{\Xi} \boldsymbol{\tau} = \boldsymbol{\pi} \quad (\text{C.12})$$

where $\boldsymbol{\pi} = [\|\mathbf{w}_1\|_2^2, \|\mathbf{w}_2\|_2^2, \dots, \|\mathbf{w}_K\|_2^2]^T \in \mathbb{R}_+^K$ and

$$\xi_{k,j} = \begin{cases} \sum_{i=1, i \neq k}^K |\mathbf{h}_i^T \mathbf{w}_j|^2 \sigma_{\eta_i}^{-2} + 1 & \text{when } j = k \\ -|\mathbf{h}_j^T \mathbf{w}_k|^2 \sigma_{\eta_j}^{-2} & \text{where } j \neq k \end{cases}$$

C.12 is basically a power allocation problem, since the entries of $\boldsymbol{\tau}$ are the powers given to the different users in the vector MAC. Due to the column-wise diagonal dominance of $\boldsymbol{\Xi}$, the positivity of the diagonal entries, and the non-positivity of the off-diagonal entries, the entries of $\boldsymbol{\Xi}^{-1}$ are non-negative. Together with the fact that the entries of $\boldsymbol{\pi}$ are positive, this leads to positive entries of $\boldsymbol{\tau} = \boldsymbol{\Xi}^{-1}\boldsymbol{\pi}$ i.e., values for the $|t_k|^2$ appropriate for equating the SINRs can always be found. Multiplying C.12 with the all 1 row vector from the left shows that the transform conserves the total transmit power, i.e. the total transmit power of the resulting vector MAC is the same as that of the vector BC.

C.3 Equivalence of Vector BC & Vector MAC

The results of the previous two subsections show that some SINRs are achievable in the vector BC under a total transmit power constraint, if and only if the same SINRs are achievable in the dual vector MAC under the same total transmit power constraint. The transform to switch from the vector MAC to the vector BC can be found in C.6 and the transform from the vector BC to the vector MAC in C.11. This result shows that any SINR problem under a total transmit power constraint in the vector BC can be solved in the dual vector MAC. The optimal result for the vector BC can be obtained by applying the transformation of C.6 to the optimal result in the dual vector MAC.

Appendix D

Wirtinger Calculus Fundamentals

For a complex function $f : \mathbb{C} \rightarrow \mathbb{C}$, the Holomorphic property is ascertained by the well known Cauchy-Riemann differentiability conditions. However, a function may not be differentiable in the analytic sense. So the Cauchy Riemann conditions do not help in constrained or non-constrained optimisation of those kind of functions. To address this issue, Wirtinger derivative was introduced [184] as partial derivatives with respect to imaginary and complex parts of a function. This appendix will introduce some basics of Wirtinger calculus which have been used in Chapter 2 in the context of linear decoding in multiple access channels.

D.1 Derivative from Linear Approximation

For complex argument $z = x + jy$, let $f : \mathbb{C} \rightarrow \mathbb{C}$ be represented as

$$f(z) = g(x, y) = \phi(x, y) + j\psi(x, y)$$

separating the real and imaginary parts. Using first order Taylor's series approximation

$$g(x, y) = g(x_0, y_0) + \alpha(x - x_0) + \beta(y - y_0) + h(x, y, x_0, y_0)\sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (\text{D.1})$$

The objective is to set the parameters α and β to ensure that $h(x, y, x_0, y_0) \rightarrow 0$ for $(x - x_0)^2 + (y - y_0)^2 \rightarrow 0$. As shown in [184],

$$\alpha = \frac{\partial \phi(x_0, y_0)}{\partial x} + j \frac{\partial \psi(x_0, y_0)}{\partial x} \quad \text{and} \quad \beta = \frac{\partial \phi(x_0, y_0)}{\partial y} + j \frac{\partial \psi(x_0, y_0)}{\partial y}$$

Since $x = \frac{z+z^*}{2}$ and $y = \frac{z-z^*}{2j}$, the linear approximation can be written as

$$f(z) = f(z_0) + \frac{(\alpha - j\beta)}{2}(z - z_0) + \frac{(\alpha + j\beta)}{2}(z^* - z_0^*) + r(z, z_0) \quad (\text{D.2})$$

where $r(z, z_0) = h(x, y, x_0, y_0)\sqrt{(x - x_0)^2 + (y - y_0)^2}$.

Then the Wirtinger derivative operators are defined as

$$\frac{\partial}{\partial z} \equiv \frac{1}{2} \left(\frac{\partial}{\partial x} - j \frac{\partial}{\partial y} \right) \quad \text{and} \quad \frac{\partial}{\partial z^*} \equiv \frac{1}{2} \left(\frac{\partial}{\partial x} + j \frac{\partial}{\partial y} \right)$$

such that [D.2](#) can be expressed as

$$f(z) = f(z_0) + \frac{\partial f(z_0)}{\partial z} (z - z_0) + \frac{\partial f(z_0)}{\partial z^*} (z^* - z_0^*) + r(z, z_0) \quad (\text{D.3})$$

D.2 Optimisation Using Wirtinger Derivative

The Wirtinger derivatives as defined above always exist if the real and imaginary parts of the function are differentiable. Additionally, the well known Cauchy Riemann conditions are fulfilled if $\frac{\partial f(z)}{\partial z^*} = 0$. Additionally, it can be shown [\[184\]](#) that the Cauchy Riemann derivative $f'(z)$, if exists, is equal to the Wirtinger derivative $\frac{\partial f(z)}{\partial z} = 0$.

Now to solve the optimisation problem, define the continuous function $f : \mathbb{C} \rightarrow \mathbb{R}$ which can be captured by the special case $\psi(x, y) = 0$. In that case, using the definition of Wirtinger derivative $\frac{\partial f(z)}{\partial z^*} = \left(\frac{\partial f(z)}{\partial z} \right)^*$. Hence for an unconstrained optimisation problem, $\frac{\partial f(z)}{\partial z} = 0$ is a necessary and sufficient condition for stationarity.

D.3 Generalisation to Matrix & Vector Arguments

Since the problem at hand in [Chapter 2](#) has a vector argument, it is necessary to extend the concept of Wirtinger derivative to vectors, which is actually straight forward. For $\mathbf{Z} \in \mathbb{C}^{m \times n}$ and $f : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}$, the Taylor series can be expressed as

$$f(\mathbf{Z}) = f(\mathbf{Z}_0) + \text{trace} \left(\frac{\partial f(\mathbf{Z}_0)}{\partial \mathbf{Z}} (\mathbf{Z} - \mathbf{Z}_0)^T \right) + \text{trace} \left(\frac{\partial f(\mathbf{Z}_0)}{\partial \mathbf{Z}^*} (\mathbf{Z} - \mathbf{Z}_0)^H \right) + r(\mathbf{Z}, \mathbf{Z}_0) \quad (\text{D.4})$$

For the linear approximation of [D.4](#), the Wirtinger derivative is defined as

$$\frac{\partial}{\partial \mathbf{Z}} \equiv \frac{1}{2} \left(\frac{\partial}{\partial \mathbf{X}} - j \frac{\partial}{\partial \mathbf{Y}} \right) \quad \text{and} \quad \frac{\partial}{\partial \mathbf{Z}^*} \equiv \frac{1}{2} \left(\frac{\partial}{\partial \mathbf{X}} + j \frac{\partial}{\partial \mathbf{Y}} \right)$$

Treating the complex conjugate as a constant while differentiating with respect to $\mathbf{Z}$, the following rules can be derived

- $\frac{\partial \text{trace} \mathbf{AZ}}{\partial \mathbf{Z}} = \mathbf{A}^T$
- $\frac{\partial \text{trace} \mathbf{AZ}}{\partial \mathbf{Z}^*} = 0$
- $\frac{\partial \mathcal{R}(\text{trace} \mathbf{AZ})}{\partial \mathbf{Z}} = \frac{1}{2} \mathbf{A}^T$
- $\frac{\partial \text{trace} \mathbf{Z}^H \mathbf{AZ}}{\partial \mathbf{Z}} = \mathbf{AZ}$

Appendix E

Constraints on Convex Problems

Let $f(\mathbf{x})$ be a concave function defined over the convex feasible region $\mathcal{F} \subseteq \mathbb{R}^n$. Now consider m additional affine constraints $\mathbf{Ax} = \mathbf{y}$ for some $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{y} \in \mathbb{R}^m$. Denote the optimal value as $\phi(\mathbf{y})$.

$$\phi(\mathbf{y}) \equiv \max_{\mathbf{x}} f(\mathbf{x})$$

subject to $\mathbf{x} \in \mathcal{F}$ and $\mathbf{Ax} = \mathbf{y}$. Let $\mathcal{Y} \subseteq \mathbb{R}^m$ be the set of $\mathbf{y}$ for which the problems (with additional constraints) are feasible.

Theorem E.1 $\phi : \mathcal{Y} \rightarrow \mathbb{R}$ is a concave function on its domain.

Proof: Obviously in presence of the additional constraints, the problem is a convex optimisation problem. Now for the additional constraint $\mathbf{Ax} = \mathbf{y}_1$, let the solution be $\mathbf{x}_1^*$ and similarly for $\mathbf{Ax} = \mathbf{y}_2$, let the solution be $\mathbf{x}_2^*$. Then

$$\phi(\mathbf{y}_1) = f(\mathbf{x}_1^*)$$

$$\phi(\mathbf{y}_2) = f(\mathbf{x}_2^*)$$

and from feasibility conditions $\mathbf{Ax}_1^* = \mathbf{y}_1$ and $\mathbf{Ax}_2^* = \mathbf{y}_2$. Now for any $\alpha \in [0, 1]$, consider the point

$$\alpha \mathbf{x}_1^* + (1 - \alpha) \mathbf{x}_2^* \in \mathcal{F}$$

which is a feasible point satisfying the constraints

$$\mathbf{Ax} = \alpha \mathbf{y}_1 + (1 - \alpha) \mathbf{y}_2$$

So $\mathcal{Y}$ is a convex region. Moreover

$$\begin{aligned} \phi(\alpha \mathbf{y}_1 + (1 - \alpha) \mathbf{y}_2) &\geq f(\alpha \mathbf{x}_1^* + (1 - \alpha) \mathbf{x}_2^*) \\ &\geq \alpha f(\mathbf{x}_1^*) + (1 - \alpha) f(\mathbf{x}_2^*) \quad \text{concave property of } f \\ &= \alpha \phi(\mathbf{y}_1) + (1 - \alpha) \phi(\mathbf{y}_2) \end{aligned}$$

■

Journal Publications

1. Barman Roy, S. & Madhukumar, A.S., 2015. Characterising the Pareto frontier of multiple access rate region. Springer Journal of Wireless Networks, 21(5), pp.1537–1548. Available at: <http://link.springer.com/article/10.1007/s11276-014-0863-2>
2. Barman Roy, S., Madhukumar, A. S., & Joung, J. 2016. On Joint Pareto Frontier in Multiple Access and Relay Rate Regions with Rayleigh Fading. IEEE Transactions on Vehicular Technology, PP(99). Available at: <http://doi.org/10.1109/TVT.2016.2598578>
3. Barman Roy, S., Madhukumar, A. S., & Chin, F. Resource allocation in multicell systems with coordinated beamforming and partial data cooperation. Under Review in IEEE IEEE Journal on Emerging Topics in Circuits and Systems.

Conference Publications

1. Barman Roy, S., Merchant, S.N. & Madhukumar, A.S., 2013. Random transmission in cognitive uplink network. In International Conference on Mobile Services, Resources, and Users. Lisbon, pp. 94–98. Available at: https://www.thinkmind.org/download.php?articleid=mobility_2013_4_40_40042
2. Barman Roy, S. & Madhukumar, A.S., 2014. Resource allocation strategy in multiple access interference channel. In IEEE Vehicular Technology Conference Fall. Vancouver, British Columbia, pp. 1–5. Available at: <http://ieeexplore.ieee.org/document/6966160/>
3. Barman Roy, S., Madhukumar, A.S. & Chin, F., 2016. Centralised and distributed interference management in coordinated downlink beam-forming. In IEEE Wireless Communications and Networking Conference. Doha, pp. 276–281. Available at: <http://ieeexplore.ieee.org/document/7552712/>
4. Barman Roy, S., Madhukumar, A.S. & Chin, F., 2017. Subspace Based Interference Cancellation in Cooperative Multicell Transmission. In IEEE Vehicular Technology Conference Spring. Sydney.

Bibliography

- [1] Claude Elwood Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(7):379–423, 1948.
- [2] Radoslaw Piesiewicz, Thomas Kleine-Ostmann, Norman Krumbholz, Daniel Mittleman, Martin Koch, Thomas Kürner, and Joerg Schoebel. Short-Range Ultra-Broadband Terahertz Communications: Concepts and Perspectives. *IEEE Antennas and Propagation Magazine*, 49(6):24–39, 2007.
- [3] David Gesbert, Marios Kountouris, Robert W. Heath, Chan Byoung Chae, and Thomas Sälzer. From Single User to Multiuser Communication: Shifting the MIMO Paradigm. *IEEE Signal Processing Magazine*, 24(5):36–46, 2007.
- [4] Haifan Yin, David Gesbert, Miltiades Filippou, and Yingzhuang Liu. A coordinated approach to channel estimation in large-scale multiple-antenna systems. *IEEE Journal on Selected Areas in Communications*, 31(2):264–273, apr 2013.
- [5] Emre Telatar. Capacity of multi-antenna Gaussian Channels. *European Transactions on Telecommunications*, 10(6):585–595, nov 1999.
- [6] Aydano Barreto Carleial. A Case Where Interference Does Not Reduce Capacity. *IEEE Transactions on Information Theory*, IT-21(5):569–570, 1975.
- [7] Aydano Barreto Carleial. Interference Channels. *IEEE Transactions on Information Theory*, IT-24(1):60–70, 1978.
- [8] Hiroshi Sato. The Capacity of the Gaussian Interference Channel Under Strong Interference. *IEEE Transactions on Information Theory*, IT-27(6):786–788, 1981.
- [9] Patrick P. Bergmans and Thomas M. Cover. Cooperative Broadcasting. *IEEE Transactions on Information Theory*, IT-20(3):317–324, 1974.
- [10] Patrick P. Bergmans. A Simple Converse for Broadcast Channels with Additive white Gaussian Noise. *IEEE Transactions on Information Theory*, IT-20(2):279–280, 1974.
- [11] Han Wu Chen and Kenjiro Yanagi. The Convex–Concave Characteristics of Gaussian Channel Capacity Functions. *IEEE Transactions on Information Theory*, 52(5):2167–2172, 2006.
- [12] Raúl H. Etkin, David N. C. Tse, and Hua Wang. Gaussian Interference Channel Capacity to Within One Bit. *IEEE Transactions on Information Theory*, 54(12):5534–5562, 2008.

- [13] An Liu, Ashutosh Sabharwal, Youjian Liu, Haige Xiang, and Wu Luo. Distributed MIMO Network Optimization Based on Local Message Passing and Duality. In *Allerton Conference on Communication, Control, and Computing*, pages 795–802, Monticello, 2009.
- [14] Chi-Hang Fred Fung, Wei Yu, and Teng Joon Lim. Precoding for the multiantenna downlink : Multiuser SNR gap and optimal user ordering. *IEEE Transactions on Communications*, 55(1):188–197, jan 2007.
- [15] J. Nicholas Laneman, David N. C. Tse, and Gregory W. Wornell. Cooperative Diversity in Wireless Networks : Efficient Protocols and Outage Behavior. *IEEE Transactions on Information Theory*, 50(12):3062–3080, 2004.
- [16] Luca Sanguinetti, Antonio Alberto D’Amico, and Yue Rong. A Tutorial on the Optimization of Amplify-and-Forward MIMO Relay Systems. *IEEE Journal on Selected Areas in Communications*, 30(8):1331–1346, sep 2012.
- [17] 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Coordinated multi-point operation for LTE Physical Layer Aspects. Technical Report 11, 3GPP, Valbonne, 2011.
- [18] Patrick Marsch and Gerhard P. Fettweis. *Coordinated Multi-Point in Mobile Communications From Theory to Practice*. Cambridge University Press, New York, 1 edition, 2011.
- [19] David Gesbert, Stephen Hanly, Howard Huang, Shlomo Shamai Shitz, Osvaldo Simeone, and Wei Yu. Multi-cell MIMO cooperative networks: A new look at interference. *IEEE Journal on Selected Areas in Communications*, 28(9):1380–1408, dec 2010.
- [20] Emil Bjornson, Randa Zakhour, David Gesbert, and Björn Ottersten. Cooperative multi-cell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI. *IEEE Transactions on Signal Processing*, 58(8):4298–4310, 2010.
- [21] Chan Byoung Chae, Insoo Hwang, Robert W. Heath, and Vahid Tarokh. Interference aware-coordinated beamforming in a multi-cell system. *IEEE Transactions on Wireless Communications*, 11(10):3692–3703, oct 2012.
- [22] Insoo Hwang, Chan Byoung Chae, Jungwon Lee, and Robert W. Heath. Multicell Cooperative Systems with Multiple Receive Antennas. *IEEE Wireless Communications*, 20(1):50–58, feb 2013.
- [23] Hanan Weingarten, Yossef Steinberg, and Shlomo Shamai Shitz. The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Transactions on Information Theory*, 52(9):3936–3964, sep 2006.
- [24] Rui Zhang and Ying Chang Liang. Exploiting Multi-Antennas for Opportunistic Spectrum Sharing in Cognitive Radio Networks. *IEEE Journal of Selected Topics in Signal Processing*, 2(1):88–102, 2008.
- [25] Khoa T. Phan, Sergiy A. Vorobyov, Nicholas D. Sidiropoulos, and Chintha Tellambura. Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming. *IEEE Transactions on Signal Processing*, 57(6):2323–2335, jun 2009.

- [26] Mamta Agiwal, Abhishek Roy, and Navrati Saxena. Next Generation 5G Wireless Networks : A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, 18(3):1617–1655, 2016.
- [27] DOCOMO 5G Radio Access: Requirements, Concept, and Technologies. *White Paper*, 2015.
- [28] Ericsson. 5G Radio Access What is 5G ? *White Paper*, (February):10, apr 2015.
- [29] Li You, Xiqi Gao, A. Lee Swindlehurst, and Wen Zhong. Channel Acquisition for Massive MIMO-OFDM with Adjustable Phase Shift Pilots. *IEEE Transactions on Signal Processing*, 64(6):1461–1476, mar 2016.
- [30] Wei Fan, Ines Carton, Pekka Kyösti, Anders Karstensen, Tommi Jämsä, Mattias Gustafsson, and Gert F. Pedersen. A step toward 5G in 2020: Low-cost OTA performance evaluation of massive MIMO base stations. *IEEE Antennas and Propagation Magazine*, feb 2017.
- [31] Peiran Wu, Robert Schober, and Vijay K. Bhargava. Robust Transceiver Design for SC-FDE Multi-hop Full-Duplex Decode-and-Forward Relaying Systems. *IEEE Transactions on Wireless Communications*, 15(2):1129–1145, 2016.
- [32] Syed Ali Jafar, Gerard Joseph Foschini, and Andrea J. Goldsmith. PhantomNet: Exploring optimal multicellular multiple antenna systems. *EURASIP Journal on Advances in Signal Processing*, (5):591–604, 2004.
- [33] Hoon Huh, Haralabos C. Papadopoulos, and Giuseppe Caire. MIMO broadcast channel optimization under general linear constraints. In *International Symposium on Information Theory*, pages 2664–2668, Seoul, 2009.
- [34] Lan Zhang, Rui Zhang, Ying Chang Liang, Yan Xin, and H. Vincent Poor. On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints. *IEEE Transactions on Information Theory*, 58(4):2064–2078, apr 2012.
- [35] David N. C. Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, New York, first edition, 2005.
- [36] Lizhong Zheng. *Diversity-Multiplexing Tradeoff: A Comprehensive View of Multiple Antenna Systems*. PhD thesis, University of California at Berkeley, 2002.
- [37] Harry Nyquist. Certain Topics in Telegraph Transmission Theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617 – 644, 1928.
- [38] Lei Ke. *Degrees of Freedom of Wireless Interference Network*. PhD thesis, Iowa State University, 2011.
- [39] Viveck R. Cadambe and Syed Ali Jafar. Interference Alignment and Degrees of Freedom of the K-User Interference Channel. *IEEE Transactions on Information Theory*, 54(8):3425–3441, 2008.
- [40] Syed Ali Jafar and Shlomo Shamai Shitz. Degrees of Freedom Region of the MIMO X Channel. *IEEE Transactions on Information Theory*, 54(1):151–170, 2008.

BIBLIOGRAPHY

- [41] Mohammad Ali Maddah-Ali, Abolfazl Seyed Motahari, and Amir Keyvan Khandani. Communication Over MIMO X Channels : Interference Alignment , Decomposition , and Performance Analysis. *IEEE Transactions on Information Theory*, 54(8):3457–3470, 2008.
- [42] Marcel L. J. van de Vel. *Theory of Convex Structures*. Elsevier B. V., North Holland, 1 edition, 1993.
- [43] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, New York, second edition, 2004.
- [44] Harold William Kuhn and Albert William Tucker. Nonlinear programming. In *Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492, 1951.
- [45] Abbas El Gamal and Young Han Kim. *Network information theory*. Cambridge University Press, New York, first edition, 2011.
- [46] Johannes Brehmer and Wolfgang Utschick. Utility maximization strategies in the multi-user MIMO downlink. In *International Workshop on Cross Layer Design*, pages 86–90, Jinan, 2007.
- [47] Johannes Brehmer, Qing Bai, and Wolfgang Utschick. Time sharing solutions in MIMO broadcast channel utility maximization. In *International ITG Workshop on Smart Antennas*, pages 153–156, Vienna, 2008.
- [48] David N. C. Tse and Stephen V. Hanly. Multiaccess fading channels-Part I: Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, nov 1998.
- [49] Giuseppe Caire and Shlomo Shamai Shitz. On the achievable throughput of a multi-antenna Gaussian broadcast channel. *IEEE Transactions on Information Theory*, 49(7):1691–1706, jul 2003.
- [50] Nihar Jindal and Andrea J. Goldsmith. Capacity and Optimal Power Allocation for Fading Broadcast Channels with Minimum Rates. *IEEE Transactions on Information Theory*, 49(11):2895–2909, 2003.
- [51] Nihar Jindal and Andrea J. Goldsmith. Dirty-paper Coding Versus TDMA for MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, 51(5):1783 – 1794, may 2005.
- [52] Le Nam Tran, Markku Juntti, Mats Bengtsson, and Björn Ottersten. Beamformer Designs for MISO Broadcast Channels with Zero-Forcing Dirty Paper Coding. *IEEE Transactions on Wireless Communications*, 12(3):1173–1185, 2013.
- [53] Roger Shu Cheng and Sergio Verdú. Gaussian multiaccess channels with ISI: Capacity region and multiuser water filling. *IEEE Transactions on Information Theory*, 39(3):773–785, may 1993.
- [54] Sriram Vishwanath, Nihar Jindal, and Andrea J. Goldsmith. Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *IEEE Transactions on Information Theory*, 49(10):2658–2668, oct 2003.

BIBLIOGRAPHY

- [55] Max H. M. Costa. Writing on dirty paper. *IEEE Transactions on Information Theory*, 29(3):439–441, 1983.
- [56] Wei Yu, Arak Sutivong, David Julian, Thomas M. Cover, and Mung Chiang. Writing on Colored Paper. In *International Symposium on Information Theory*, page 302, Washington DC, 2001.
- [57] Hiroshi Sato. An Outer Bound to the Capacity Region of Broadcast Channels. *IEEE Transactions on Information Theory*, 24(3):374–377, 1978.
- [58] Hieu T. Do and Sae-Young Chung. Linear Beamforming and Superposition Coding with Common Information for the Gaussian MIMO Broadcast Channel. *IEEE Transactions on Communications*, 57(8):2484–2494, 2009.
- [59] Wei Yu, Wonjong Rhee, Stephen Boyd, and John M. Cioffi. Iterative water-filling for vector multiple access channels. *IEEE Transactions on Information Theory*, 50(1):145–152, jan 2004.
- [60] Te Sun Han and Kingo Kobayashi. A New Achievable Rate Region for the Interference Channel. *IEEE Transactions on Information Theory*, 27(1):49–60, 1981.
- [61] Ignacio Santamaría, Javier Vía, Alfredo Nazábal, and Christian Lameiro. Capacity Region of the Multiantenna Gaussian Broadcast Channel with Analog Tx-Rx Beamforming. In *International ICST Conference on Communications and Networking*, pages 1–6, Beijing, China, 2010.
- [62] Pramod Viswanath and David N. C. Tse. Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink Downlink Duality. *IEEE Transactions on Information Theory*, 49(8):1912–1921, 2003.
- [63] Nihar Jindal, Sriram Vishwanath, and Andrea J. Goldsmith. On the duality of Gaussian multiple-access and broadcast channels. *IEEE Transactions on Information Theory*, 50(5):768–783, may 2004.
- [64] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley Series in Telecommunications. John Wiley & Sons, New Jersey, second edition, 2012.
- [65] Wei Yu and John M. Cioffi. Trellis Precoding for the Broadcast Channel. In *Global Telecommunications Conference*, volume 2, pages 1344–1348, San Antonio, 2001.
- [66] Wei Yu. Uplink-downlink duality via minimax duality. *IEEE Transactions on Information Theory*, 52(2):361–374, feb 2006.
- [67] Katalin Marton. A Coding Theorem for the Discrete Memoryless Broadcast Channel. *IEEE Transactions on Information Theory*, IT-s25(3):306–311, 1979.
- [68] Gesualdo Scutari, Daniel Pérez Palomar, and Sergio Barbarossa. Competitive Design of Multiuser MIMO Systems based on Game Theory : A Unified View. *IEEE Journal on Selected Areas in Communications*, 26(7):1089–1103, 2008.

BIBLIOGRAPHY

- [69] An Liu, Youjian Liu, Haige Xiang, and Wu Luo. On the Rate Duality of MIMO Interference Channel and Its Application to Sum Rate Maximization. In *Global Telecommunications Conference*, pages 1–6, Hawaii, 2009.
- [70] An Liu, Youjian Liu, Haige Xiang, and Wu Luo. MIMO B-MAC Interference Network Optimization Under Rate Constraints by Polite Water-Filling and Duality. *IEEE Transactions on Signal Processing*, 59(1):263–276, 2011.
- [71] Martin Schubert and Holger Boche. A generic approach to QoS-based transceiver optimization. *IEEE Transactions on Communications*, 55(8):1557–1566, aug 2007.
- [72] Holger Boche and Martin Schubert. A superlinearly and globally convergent algorithm for power control and resource allocation with general interference functions. *IEEE/ACM Transactions on Networking*, 16(2):383–395, apr 2008.
- [73] Roy D. Yates. A Framework for Uplink Power Control in Cellular Radio Systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, 1995.
- [74] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, second edition, 2001.
- [75] Abolfazl Seyed Motahari and Amir Keyvan Khandani. Capacity bounds for the Gaussian interference channel. *IEEE Transactions on Information Theory*, 55(2):620–643, 2009.
- [76] Xiaohu Shang, Gerhard Kramer, and Biao Chen. A New Outer Bound and the Noisy-Interference Sum-Rate Capacity for Gaussian Interference Channels. *IEEE Transactions on Information Theory*, 55(2):689–699, 2009.
- [77] Tie Liu and Pramod Viswanath. An Extremal Inequality Motivated by Multiterminal Information-Theoretic Problems. *IEEE Transactions on Information Theory*, 53(5):1839–1851, 2007.
- [78] Sergio Verdú and Dongning Guo. A Simple Proof of the Entropy-Power Inequality. *IEEE Transactions on Information Theory*, 52(5):2165–2166, 2006.
- [79] Venkata Sreekanth Annapureddy and Venugopal V. Veeravalli. Gaussian interference networks: sum capacity in the low interference regime and new outer bounds on the capacity region. *IEEE Transactions on Information Theory*, 55(7):3032–3050, 2009.
- [80] Syed Ali Jafar. *Interference Alignment: A New Look at Signal Dimensions in a Communication Network*, volume 7 of *Foundations and Trends in Communications and Information Theory*. now Publishers Inc, Delft, 2010.
- [81] Yitzhak Birk and Tomer Kol. Coding On Demand by an Informed Source (ISCOD) for Efficient Broadcast of Different Supplemental Data to Caching Clients. *IEEE Transactions on Information Theory*, 52(6):2825–2830, 2006.
- [82] Mohammad Ali Maddah-Ali, Amir Keyvan Khandani, and Abolfazl Seyed Motahari. Communication over X Channel: Signalling and Multiplexing Gain. Technical report, University of Waterloo, 2006.

- [83] Cenk M. Yetis, Tiangao Gou, Syed Ali Jafar, and Ahmet Hamdi Karyan. On Feasibility of Interference Alignment in MIMO Interference Networks. *IEEE Transactions on Signal Processing*, 58(9):4771–4782, 2010.
- [84] Sang Woon Jeon and Michael Gastpar. A Survey on Interference Networks: Interference Alignment and Neutralization. *Entropy*, 14(12):1842–1863, 2012.
- [85] Changick Song and Cong Ling. On the diversity of linear transceivers in MIMO AF relaying systems. *IEEE Transactions on Information Theory*, 62(1):272–289, 2016.
- [86] Yue Rong, Xiaojun Tang, and Yingbo Hua. A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems. *IEEE Transactions on Signal Processing*, 57(12):4837–4851, dec 2009.
- [87] Wenyi Zhang, Urbashi Mitra, and Mung Chiang. Optimization of amplify-and-forward multicarrier two-hop transmission. *IEEE Transactions on Communications*, 59(5):1434–1445, may 2011.
- [88] Xiaojun Tang and Yingbo Hua. Optimal design of non-regenerative MIMO wireless relays. *IEEE Transactions on Wireless Communications*, 6(4):1398–1406, apr 2007.
- [89] Olga Muñoz-Medina, Josep Vidal, and Adrián Agustín. Linear transceiver design in nonregenerative relays with channel state information. *IEEE Transactions on Signal Processing*, 55(6 I):2593–2604, jun 2007.
- [90] Chengwen Xing, Shaodan Ma, and Yik Chung Wu. Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems. *IEEE Transactions on Signal Processing*, 58(4):2273–2283, apr 2010.
- [91] Changick Song, Kyoung Jae Lee, and Inkyu Lee. MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems. *IEEE Transactions on Wireless Communications*, 9(7):2310–2319, jul 2010.
- [92] Raymond H. Y. Louie, Yonghui Li, Himal A. Suraweera, and Branka Vucetic. Performance analysis of beamforming in two hop amplify and forward relay networks with antenna correlation. *IEEE Transactions on Wireless Communications*, 8(6):3132–3141, jun 2009.
- [93] Changick Song, Kyoung Jae Lee, and Inkyu Lee. MMSE-based MIMO cooperative relaying systems. *IEEE Journal on Selected Areas in Communications*, 30(8):1390–1401, sep 2012.
- [94] Kambiz Azarian, Hesham El Gamal, and Philip Schniter. On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. *IEEE Transactions on Information Theory*, 51(12):4152–4172, dec 2005.
- [95] Sheng Yang and Jean Claude Belfiore. Optimal space-time codes for the MIMO amplify-and-forward cooperative channel. *IEEE Transactions on Information Theory*, 53(2):647–663, feb 2007.
- [96] Ahmadreza Hedayat and Aria Nosratinia. Outage and diversity of linear receivers in flat-fading MIMO channels. *IEEE Transactions on Signal Processing*, 55(12):5868–5873, dec 2007.

BIBLIOGRAPHY

- [97] A. H. Mehana and A. Nosratinia. Diversity of MMSE MIMO Receivers. *IEEE Transactions on Information Theory*, 58(11):6788–6805, nov 2012.
- [98] Shuo Yen Robert Li, Raymond Wai-Ho Yeung, and Ning Cai. Linear network coding. *IEEE Transactions on Information Theory*, 49(2):371–381, feb 2003.
- [99] Raymond Wai-Ho Yeung, Shuo Yen Robert Li, Ning Cai, and Zhen Zhang. *Network coding theory*. now Publishers Inc, Hanover, Massachusetts, 2006.
- [100] Sachin Katti, Hariharan Rahul, Wenjun Hu, Dina Katabi, Muriel Medard, and Jon Crowcroft. XORs in the air: Practical wireless network coding. *IEEE/ACM Transactions on Networking*, 16(3):497–510, jun 2008.
- [101] Boris Rankov and Armin Wittneben. Spectral efficient protocols for half-duplex fading relay channels. *IEEE Journal on Selected Areas in Communications*, 25(2):379–389, feb 2007.
- [102] Chih Ning Hsu, Hsuan Jung Su, and Pin Hsun Lin. Joint subcarrier pairing and power allocation for OFDM transmission with decode-and-forward relaying. *IEEE Transactions on Signal Processing*, 59(1):399–414, jan 2011.
- [103] Wooseok Nam, Woohyuk Chang, Sae Young Chung, and Yong H. Lee. Transmit optimization for relay-based cellular OFDMA systems. In *International Conference on Communications*, pages 5714–5719, Glasgow, 2007.
- [104] Luc Vandendorpe, Rodolfo Torrea Durán, Jérôme Louveaux, and Abdellatif Zaidi. Power allocation for OFDM transmission with DF relaying. In *International Conference on Communications*, pages 3795–3800, Beijing, 2008.
- [105] Luc Vandendorpe, Jérôme Louveaux, Onur Oguz, and Abdellatif Zaidi. Rate-optimized power allocation for DF-relayed OFDM transmission under sum and individual power constraints. *Eurasip Journal on Wireless Communications and Networking*, 2009.
- [106] Yong Li, Wenbo Wang, Jia Kong, Wei Hong, Xing Zhang, and Mugen Peng. Power allocation and subcarrier pairing in OFDM-based relaying networks. In *International Conference on Communications*, pages 2602–2606, Glasgow, 2008.
- [107] Yong Li, Wenbo Wang, Jia Kong, and Mugen Peng. Subcarrier pairing for amplify-and-forward and decode-and-forward OFDM relay links. *IEEE Communications Letters*, 13(4):209–211, apr 2009.
- [108] Taehoon Kwon, Youngju Kim, and Daesik Hong. Comparison of FDR and HDR under adaptive modulation with finite-length queues. *IEEE Transactions on Vehicular Technology*, 61(2):838–843, feb 2012.
- [109] Su Min Kim and Mats Bengtsson. Virtual full-duplex buffer-aided relaying in the presence of inter-relay interference. *IEEE Transactions on Wireless Communications*, 15(4):2966–2980, apr 2016.
- [110] Gerhard Kramer, Michael Gastpar, and Piyush Gupta. Cooperative Strategies and Capacity Theorems for Relay Networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, 2005.

- [111] Alessio Zappone, Stefano Buzzi, and Eduard Jorswieck. A Survey on Game-Theoretic Approaches to Energy-Efficient Relay-Assisted Communications. In *Tyrrhenian International Workshop on Digital Communications*, number 2, Genoa, sep 2013.
- [112] Sharon M. Betz and H. Vincent Poor. Energy Efficient Communications in CDMA Networks: A Game Theoretic Analysis Considering Operating Costs. *IEEE Transactions on Signal Processing*, 56(10):5181–5190, oct 2008.
- [113] Federico Boccardi, Robert W. Heath, Angel Lozano, Thomas L. Marzetta, and Petar Popovski. Five disruptive technology directions for 5G. *IEEE Communications Magazine*, 52(2):74–80, feb 2014.
- [114] Chan Byoung Chae, Sang Hyun Kim, and Robert W. Heath. Network coordinated beamforming for cell-boundary users: Linear and non-linear approaches. *IEEE Journal for Selected Topics in Signal Processing*, 3(6):1094–1105, 2009.
- [115] Omar El Ayach, Robert W. Heath, Sridhar Rajagopal, and Zhouyue Pi. Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays. In *Global Telecommunications Conference*, pages 3476–3480, Atlanta, 2013.
- [116] Daewon Lee, Hanbyul Seo, Bruno Clerckx, Eric Hardouin, David Mazzarese, Satoshi Nagata, and Krishna Sayana. Coordinated multipoint transmission and reception in LTE-advanced: Deployment scenarios and operational challenges. *IEEE Communications Magazine*, 50(2):148–155, 2012.
- [117] C. Huppert and J. G. Klotz. Required transmit power applying Tomlinson-Harashima-precoding in scalar and MIMO broadcast systems. *IEEE Transactions on Communications*, 58(10):3011–3020, oct 2010.
- [118] Simon Haykin. Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, feb 2005.
- [119] Luca Bixio, Giacomo Oliveri, Marina Ottonello, Mirco Raffetto, and Carlo S. Regazzoni. Cognitive radios with multiple antennas exploiting spatial opportunities. *IEEE Transactions on Signal Processing*, 58(8):4453–4459, aug 2010.
- [120] Xuemin Hong, Zengmao Chen, Cheng-Xiang Wang, Sergiy A. Vorobyov, and John S. Thompson. Cognitive radio networks. *IEEE Vehicular Technology Magazine*, 4(4):76–84, 2009.
- [121] Cong Wang, Ji Li, Fan Ye, and Yuanyuan Yang. A Mobile Data Gathering Framework for Wireless Rechargeable Sensor Networks with Vehicle Movement Costs and Capacity Constraints. *IEEE Transactions on Computers*, 65(8):2411–2427, 2016.
- [122] Stephen V. Hanly and David N. C. Tse. Multiaccess fading channels– Part II: Delay limited capacities. *IEEE Transactions on Information Theory*, 44(7):2816–2831, nov 1998.
- [123] Andrea J. Goldsmith, Syed Ali Jafar, Nihar Jindal, and Sriram Vishwanath. Capacity limits of MIMO Channels. *IEEE Journal on Selected Areas in Communications*, 21(5):684–702, jun 2003.

BIBLIOGRAPHY

- [124] Erik G. Larsson and Eduard A. Jorswieck. Competition Versus Cooperation on the MISO Interference Channel. *IEEE Journal on Selected Areas in Communications*, 26(7):1059–1069, 2008.
- [125] Johannes Lindblom, Eleftherios Karipidis, and Erik G. Larsson. Efficient Computation of Pareto Optimal Beamforming Vectors for the MISO Interference Channel with Successive Interference Cancellation. *IEEE Transactions on Signal Processing*, 61(19):4782–4795, 2013.
- [126] Vasileios K. Sakarellos, Dimitrios Skraparlis, and Athanasios D. Panagopoulos. Cooperation within the small cell: The indoor, correlated shadowing case. *Physical Communication*, 9:16–22, may 2013.
- [127] Jack Edmonds. Submodular functions, matroids, and certain polyhedra. In *International Conference on Combinatorial Structures and Applications*, pages 69–87, Calgary, 1969.
- [128] Mokhtar S. Bazaraa, Hanif D. Sherali, and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley-Interscience in Discrete Mathematics and Optimization, New Jersey, third edition, 2006.
- [129] Jacek Gondzio. Interior point methods 25 years later. *European Journal of Operational Research*, 218(3):587–601, may 2012.
- [130] Michael Charles Grant. *Disciplined Convex Programming*. PhD thesis, Stanford University, 2004.
- [131] Jington Joung and Ali H. Sayed. Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems. *IEEE Transactions on Signal Processing*, 58(3):1833–1846, mar 2010.
- [132] Chee Yen Leow, Zhiguo Ding, Kin K. Leung, and Dennis L. Goeckel. On the study of analogue network coding for multi-pair, bidirectional relay channels. *IEEE Transactions on Wireless Communications*, 10(2):670–681, feb 2011.
- [133] Meixia Tao and Rui Wang. Linear precoding for multi-pair two-way MIMO relay systems with max-min fairness. *IEEE Transactions on Signal Processing*, 60(10):5361–5370, oct 2012.
- [134] Vasileios K. Sakarellos, Dimitrios Skraparlis, Athanasios D. Panagopoulos, and John D. Kannellopoulos. Outage performance analysis of a dual-hop radio relay system operating at frequencies above 10GHz. *IEEE Transactions on Communications*, 58(11):3104–3109, nov 2010.
- [135] Can Sun, Chenyang Yang, Yonghui Li, and Branka Vucetic. Transceiver design for multi-user multi-antenna two-way relay cellular systems. *IEEE Transactions on Communications*, 60(10):2893–2903, oct 2012.
- [136] Celal Esli and Armin Wittneben. Multiuser MIMO two-way relaying for cellular communications. In *Proc. International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6, Cannes, 2008.

BIBLIOGRAPHY

- [137] Thomas M. Cover. Broadcast Channels. *IEEE Transactions on Information Theory*, IT-18(1):2–14, 1972.
- [138] Swagato Barman Roy and A. S. Madhukumar. Resource allocation strategy in multiple access interference channel. In *Proc. Vehicular Technology Conference Fall*, pages 1–5, Vancouver, British Columbia, 2014.
- [139] Hongyu Cui, Meng Ma, Lingyang Song, and Bingli Jiao. Relay selection for two-way full duplex relay networks with amplify-and-forward protocol. *IEEE Transactions on Wireless Communications*, 13(7):3768–3777, jul 2014.
- [140] Zhanzhan Zhang, Zhiyong Chen, Manyuan Shen, and Bin Xia. Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO. *IEEE Journal on Selected Areas in Communications*, 34(4):848–863, apr 2016.
- [141] Haiyang Xin, Xiaojun Yuan, and Soung-Chang Liew. Multi-pair MIMO two-way relaying : A principal-angle perspective. *IEEE Transactions on Signal Processing*, 64(2):380–394, jan 2015.
- [142] Jianhua Mo, Meixia Tao, Yuan Liu, and Rui Wang. Secure beamforming for MIMO two-way communications with an untrusted relay. *IEEE Transactions on Signal Processing*, 62(9):2185–2199, may 2014.
- [143] Athanasios Stavridis, Dushyantha Basnayaka, Sinan Sinanovic, Marco Di Renzo, and Harald Haas. A virtual MIMO dual-hop architecture based on hybrid spatial modulation. *IEEE Transactions on Communications*, 62(9):3161–3179, sep 2014.
- [144] Ulrich Kohn. Fronthaul networks—a key enabler for LTE-advanced. Technical report, ADVA Optical Networking SE, Martinsried, Bavaria, feb 2014.
- [145] Vikram Chandrasekhar, Jeffrey G. Andrews, Tarik Muharemovict, Zukang Shen, and Alan Gatherer. Power control in two-tier femtocell networks. *IEEE Transactions on Wireless Communications*, 8(8):4316–4328, aug 2009.
- [146] Shu Ping Yeh, Shilpa Talwar, Seong Choon Lee, and Heechang Kim. WiMAX femtocells: A perspective on network architecture, capacity, and coverage. *IEEE Communications Magazine*, 46(10):58–65, oct 2008.
- [147] The second phase of LTE Advance, LTE-B: 30-fold capacity boosting to LTE. Technical report, Huawei Technologies Co. Limited, Shenzhen, 2013.
- [148] Swagato Barman Roy and A. S. Madhukumar. Characterising the Pareto frontier of multiple access rate region. *Springer Journal of Wireless Networks*, 21(5):1537–1548, jun 2015.
- [149] Steve Perlman and Antonio Forenza. An Introduction to pCell. Technical report, Artemis Networks LLC, feb 2015.
- [150] Thomas L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, nov 2010.

BIBLIOGRAPHY

- [151] Hong Yang and Thomas L. Marzetta. Performance of conjugate and zero-forcing beamforming in large-scale antenna systems. *IEEE Journal on Selected Areas in Communications*, 31(2):172–179, 2013.
- [152] Jakob Hoydis, Stephan ten Brink, and Mérouane Debbah. Comparison of Linear Precoding Schemes for Downlink Massive MIMO. In *International Conference on Communications*, pages 2135–2139, Ottawa, 2012.
- [153] Fredrik Rusek, Daniel Persson, Buon Kiong Lau, Erik G. Larsson, Thomas L. Marzetta, Ove Edfors, and Fredrik Tufvesson. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1):40–60, 2013.
- [154] Erik G. Larsson. Massive MIMO for Next Generation Wireless Systems. *IEEE Communications Magazine*, pages 186–195, feb 2014.
- [155] Sébastien Roy. Two-layer linear processing for massive MIMO on the TitanMIMO platform. Technical report, Université de Sherbrooke, 2015.
- [156] Olakunle Elijah, Chee Yen Leow, Tharek Abdul Rahman, Solomon Nunoo, and Solomon Zakwui Iliya. A Comprehensive Survey of Pilot Contamination in Massive MIMO–5G System. *IEEE Communications Surveys and Tutorials*, 18(2):905–923, 2016.
- [157] Hien Quoc Ngo, Erik G. Larsson, and Thomas L. Marzetta. Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems. *IEEE Transactions on Communications*, 61(4):1436–1449, apr 2013.
- [158] Jakob Hoydis, Stephan Ten Brink, and Mérouane Debbah. Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE Journal on Selected Areas in Communications*, 31(2):160–171, feb 2013.
- [159] Alberto Zanella, Marco Chiani, and Moe Z. Win. MMSE reception and successive interference cancellation for MIMO systems with high spectral efficiency. *IEEE Transactions on Wireless Communications*, 4(3):1244–1253, may 2005.
- [160] An Liu and Vincent K. N. Lau. Two-stage constant-envelope precoding for low-cost massive MIMO systems. *IEEE Transactions on Signal Processing*, 64(2):485–494, jan 2016.
- [161] Dianjun Chen. A Randomized Space-Time Spreading Scheme for Massive MIMO Channels. *IEEE Transactions on Wireless Communications*, 15(5):3668–3678, may 2016.
- [162] Michael Brown and Martin Turgeon. TitanMIMO: A 100x100 Massive MIMO Testbed Based on xTCA Standards nutaq.com. Technical report, 2014.
- [163] Martin Schubert and Holger Boche. Solution of the multiuser downlink beamforming problem with individual SINR constraints. *IEEE Transactions on Vehicular Technology*, 53(1):18–28, jan 2004.
- [164] Karama Hamdi, Wei Zhang, and Khaled Ben Letaief. Opportunistic Spectrum Sharing in Cognitive MIMO Wireless Networks. *IEEE Transactions on Wireless Communications*, 8(8):4098–4109, 2009.

BIBLIOGRAPHY

- [165] Mario Kießling. *Statistical Analysis and Transmit Prefiltering for MIMO Wireless Systems in Correlated Fading Environments*. PhD thesis, Universität Stuttgart, 2004.
- [166] Michael W. Berry. Large scale singular value computations. *International Journal of Supercomputer Applications*, 6(1):13–49, jan 1992.
- [167] Thomas F. Colemanm and Alex Pothén. The null space problem II: Algorithms. *Journal on Algebraic Discrete Methods*, 8(4):544–563, 1987.
- [168] Swagato Barman Roy, A. S. Madhukumar, and Jingon Joung. On Joint Pareto Frontier in Multiple Access and Relay Rate Regions with Rayleigh Fading. *IEEE Transactions on Vehicular Technology*, PP(99), aug 2016.
- [169] Garrett Birkhoff. *Lattice Theory*. American Mathematical Society, New York, 1940.
- [170] Yingbo Li, Lin Bai, Chen Chen, Ye Jin, and Jinho Choi. Successive orthogonal beamforming for cooperative multi-point downlinks. *IET Communications*, 7(8):706–714, may 2013.
- [171] Duy H N Nguyen and Tho Le-Ngoc. Sum-rate maximization in the multicell MIMO multiple-access channel with interference coordination. *IEEE Transactions on Wireless Communications*, 13(1):36–48, 2014.
- [172] Hoon Huh, Haralabos C. Papadopoulos, and Giuseppe Caire. Multiuser MIMO transmitter optimization for inter-cell interference mitigation. *IEEE Transactions on Signal Processing*, 58(8):4272–4285, aug 2010.
- [173] Kemal Davaslioglu and Ender Ayanoglu. Quantifying Potential Energy Efficiency Gain in Green CellularWireless Networks. *IEEE Communications Surveys & Tutorials*, 16(4):2065–2091, 2014.
- [174] Wenchi Cheng, Xi Zhang, and Hailin Zhang. Statistical-QoS Driven Energy-Efficiency. *IEEE Journal on Selected Areas in Communications*, 34(12):3092–3107, dec 2016.
- [175] Rui Zhang, Chin Choy Chai, and Ying Chang Liang. Joint beamforming and power control for multiantenna relay broadcast channel With QoS constraints. *IEEE Transactions on Signal Processing*, 57(2):726–737, feb 2009.
- [176] Rui Cao and Liuqing Yang. What determines resource optimization in cooperative communications. In *IEEE Wireless Communications and Networking Conference*, Budapest, 2009.
- [177] Batu Krishna Chalise and Luc Vandendorpe. Optimization of MIMO Relays for Multipoint-to-Multipoint Communications: Nonrobust and Robust Designs. *IEEE Transactions on Signal Processing*, 58(12):6355–6368, 2010.
- [178] Helmut Bölcskei, Rohit U. Nabar, Özgür Oyman, and Arogyaswami J. Paulraj. Capacity Scaling Laws in MIMO Relay Networks. *IEEE Transactions on Information Theory*, 5(6):1433–1444, 2006.

BIBLIOGRAPHY

- [179] Jörg Wagner, Boris Rankov, and Armin Wittneben. Large n Analysis of Amplify-and-Forward MIMO Relay Channels with Correlated Rayleigh Fading. *IEEE Transactions on Information Theory*, 54(12):5735–5746, 2006.
- [180] Kenneth Hoffman and Ray Kunze. *Linear algebra*. Pearson Education Inc., New Jersey, second edition, 2001.
- [181] John Michael Steele. *The Cauchy-Schwarz master class: An introduction to the art of mathematical inequalities*. Mathematical Association of America Problem Books. Cambridge University Press, New York, first edition, 2004.
- [182] James G. Oxley. *Matroid Theory*. Oxford graduate texts in mathematics. Oxford University Press, Cary, second edition, 2011.
- [183] Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a Monotone Submodular Function Subject to a Matroid Constraint. *Society for Industrial and Applied Mathematics Journal on Computing*, 40(6):1740–1766, 2011.
- [184] Reinhold Remmert. *Theory of complex functions*. Springer-Verlag, Berlin, 1991.

..... THE END