Optimal Rank-Constrained Transmission in Block Diagonalized Cooperative Multi-Cell MIMO with Multiple Power Constraints

A Project Report

submitted by

SAI VIHARI CHATURVEDULA

in partial fulfilment of the requirements
for the award of the degree of

MASTER OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY MADRAS.

MAY 2019

THESIS CERTIFICATE

This is to certify that the thesis titled **Optimal Rank-Constrained Transmission in Block Diagonalized Cooperative Multi-Cell MIMO with Multiple Power Constraints**, submitted by **Sai Vihari Chaturvedula**, to the Indian Institute of Technology, Madras, for the award of the degree of **Master of Technology**, is a bona fide record of the research work done by him under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Srikrishna Bhashyam
Professor
Dept. Electrical Engineering
IIT-Madras, 600 036

Place: Chennai

Date: 10 May 2019

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude and respect to my advisor, Prof.Sri Krishna Bhashyam for his continuous support of my M.Tech study and research. It is his valuable guidance, patience, motivation, passion, timely responses and immense knowledge in the subject that has driven me in the right path during the project. Any words to describe his constructive approach towards problem solving would be less. I would like to thank you from the bottom of my heart, Sir for all the weekly feedback sessions on the project progress and introducing me to the interesting fields of Detection and Estimation, Adaptive Signal Processing.

I am ever grateful to a very special person, Prof.David Koilpillai who I believe has been very instrumental in building up both my academic and personal skill sets. He has been my inspiration in every walk of institute life and would always continue to be. I would also like to thank him for introducing me to the areas of Wireless Communications, 4G and Multirate DSP.

A special mention to Prof.Rachel for her innovative way of teaching convex optimization and her timely advices during the project. I am thankful to all the Professors who have taught me over 60 courses during my 5 years tenure in this great institute.

I am grateful to all my fellow students in the lab who spared their valuable time and help whenever required during this project. I could not have imagined my stay in IITM without my best friends, Kumar, Madhuri and Bhavya who have been a souce of perennial support.

Finally, this acknowledge would be incomplete without mentioning the affection, love and encouragement my parents and sweet little sister have bestowed upon me throughout the years. My mother and father are my biggest inspiration and I thank them for giving me the best life I could have ever asked for.

ABSTRACT

In this paper, we primarily focus on obtaining an optimal rank-constrained downlink transmission strategy in a cooperative multi-user multiple-input-multiple-output (MU-MIMO) system. More specifically, we look at the optimal precoder design that maximizes the weighted sum-rate of all the Mobile Stations (MSs) under a set of joint power constraints. This set includes (1) Sum power constraint over all the transmit antennas (SPC), (2) Per Base Station (BS) power constraints (PBPC) and (3) Per Antenna power constarints (PAPC). To eliminate the inter user interference, we apply a linear precoding technique called Block Diagonalisation (BD) to the downlink transmission. We assume perfect knowledge of the downlink channels and the transmit messages in designing signals from different BSs to all the MSs. It's important to study the BD rank-constrained transmission under joint power constraints set in the context of mmWave systems where the number of antennas could be large, but only a limited number of streams are allowed because of the low rank of the channel and also the implementation complexity of spatial multiplexing that comes with a large number of streams. Our main result in this paper discusses a very efficient algorithm, Projected Factored Gradient Descent (PFGD) to arrive at the optimal rank-constrained precoder matrix numerically. Achievable rates are plotted for some cooperative MU-MIMO systems under different low rank constraints. By using convex optimisation and linear algebra techniques, we also extend a popular work slightly to present another efficient optimal algorithm to solve the same problem without a rank constraint *viz.* the number of allowed data streams being the total number of receive antennas, and derive the closed form expression for the optimal BD precoding matrix.

Moreover, the proposed solution reduces to the optimal zero-forcing beamforming (ZF-BF) precoder design for the weighted sum-rate maximization in the multi-user multiple-input-single-output (MU-MISO) broadcast channel (BC) with *joint SPC and PAPC*. A sub optimal but low complexity BD precoding scheme is also presented and their achievable rates are compared against the rates achieved by both the optimal schemes in full possible rank case.

Index terms— **Multi User MIMO, coperative multi-cell system, multi-antenna broadcast channel, convex optimisation, linear algebra, block diagonalization, per antenna power constraints, sum power constraint, per base station power constraints, rank constraint, zero-forcing beamforming.**

TABLE OF CONTENTS

ACKNOWLEDGEMENTS **i**

ABSTRACT **ii**

LIST OF FIGURES **viii**

ABBREVIATIONS **ix**

NOTATIONS **x**

1 INTRODUCTION **1**

1.1 Single-Cell Setup and Block Diagonalisation 1

1.2 Multi-Cell Systems and Motivation behind the project 3

1.3 Project's Main Contributions 5

1.3.1 What is new in the project ? 5

1.3.2 Optimal Solution without rank constraints 5

1.3.3 Suboptimal Solution without rank constraints 6

1.3.4 PFGD Algorithm with rank constraints 6

1.3.5 Simulation Results 7

2 SYSTEM MODEL AND PROBLEM FORMULATION **8**

2.1 System Model 8

2.2 Relation between Transmit Covariance Matrices and Precoding Matrices 9

2.3 Multiple Power Constraints 10

2.3.1 Sum Power Constraints (SPC) 10

2.3.2 Per-Base Station Power Constraints (PBPC) 10

2.3.3 Per-Antenna Power Constraints (PAPC) 10

2.4 Quasi Static Fading Assumption 11

2.5 Block Diagonalisation Precoding Constraints 11

2.6 Convex Optimisation Problem to be solved in Chapter 3 - (P1) . . . 12
2.7 Non-Convex Optimisation Problem to be solved in Chapter 4 - (P2) 13

3 OPTIMAL PRECODING MATRIX SOLUTION WITHOUT RANK CONSTRAINTS 14

3.1 General MU-MIMO BC Case 14
3.1.1 Eliminate BD Precoding Constraints 14
3.1.2 Convex Optimisation Problem (P3) - A Modified Version of (P1) 15
3.1.3 Procedure to obtain Optimal Solution of (P3) 16
3.1.4 Relevant Power Constraints 17
3.1.5 Maximisation Problem (P4) and it's Optimal Solution . . . 18
3.1.6 Optimal Solution of (P1) 21
3.1.7 Optimal Algorithm (A1) to solve (P1) 21
3.1.8 Optimal Precoding matrix structure without rank constraints and it's miscellaneous properties 22
3.2 Special MU-MISO BC Case 24
3.2.1 Optimal Structure of the Transmit Covariance Matrices . . . 24
3.2.2 Optimal Precoding Matrix Structure 24
3.2.3 Interesting Properties of the above Optimal Precoding Matrix Structure 25

4 SUB OPTIMAL PRECODING MATRIX SOLUTION WITHOUT RANK CONSTRAINTS 27

4.1 Procedure to obtain Sub Optimal Precoding Matrix Structure 27
4.2 Formulation of Convex Optimisation Problem (P6) to be solved for obtaining Sub Optimal Solution 28
4.3 Procedure to obtain Optimal Solution of (P6) 30
4.4 Maximisation Problem (P7) and it's Optimal Solution 31
4.5 Sub Optimal Algorithm (A2) to solve (P1) 32
4.6 Reason behind the lower implementation complexity of (A2) compared to (A1) 33

5 RANK-CONSTRAINED MULTI-USER MIMO CAPACITY UNDER JOINT-SPC-PBPC-PAPC 34

5.1 Important reasons for choosing PFGD Approach 34

5.2 Formulation of two Non Convex Optimisation problems (P8) & (P9) equivalent to (P2) 35
5.3 Individual Constraint Sets - Notations 36
5.4 PFGD Algorithm (A3) to solve (P2) 37
5.5 Initialisation Step in the PFGD Algorithm 38
5.6 Procedure to describe the projection of a set of matrices onto the Total Constraints Set $\mathcal{C}^{Tot}$ 39
5.7 Projection Onto BD Precoding Constraints Set 40
5.8 Projection Onto Joint-SPC-PBPC Constraints Set 41
5.9 Projection Onto Joint-SPC-PBPC-PAPC Constraints Set 43
5.10 Dykstra's Projection Onto Total Constraints Set 44

6 NUMERICAL RESULTS 46

6.1 Convergence Behavior 46
6.1.1 Convergence Behavior of Dykstra's Projection Algorithm . 46
6.1.2 Convergence Behavior of PFGD Algorithm 46
6.1.3 Comments on Complexity of the PFGD Algorithm 47
6.2 MU-MISO BC with Joint-SPC-PAPC - Comparison of (A1) and (A2) performance 49
6.2.1 Case Study 1 49
6.2.2 Case Study 2 50
6.3 MU-MIMO BC with Joint-SPC-PBPC-PAPC - Comparison of (A1), (A2), PFGD (A3) and CVX performance 51
6.3.1 Case Study 1 51
6.3.2 Case Study 2 52
6.4 MU-MIMO BC with Joint-SPC-PBPC-PAPC - Comparison of PFGD (A3) performance under different rank constraints 53
6.4.1 Case Study 1 53
6.4.2 Case Study 2 54

7 SUMMARY 55

APPENDIX 56

A. Proof of Lemma 1 56
B. Proof of Lemma 2 57

C. Proof of Lemma 3 57
D. Proof of Theorem 1 58
E. Proof of Theorem 3 60
F. Proof of Theorem 4 61
G. Proof of Theorem 5 62

LIST OF FIGURES

2.1 Multi User - MIMO BC 9

6.1 $A = 4, M_B = 3, K = 4, N = 2$ & $\hat{P}_i = 1W$, $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ 47

6.2 $A = 4, M_B = 3, K = 4, N = 2$ & $\hat{P}_i = 1W$, $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ & $\eta = 2$ 48

6.3 $A = 2:1:10, M_B = 1, K = 2, N = 1$ & $\hat{P}_i = 1.25, P_{sum} = \frac{A M_B \hat{P}_i}{1.25}$ 49

6.4 $A = 4:1:12, M_B = 1, K = 4, N = 1$ & $\hat{P}_i = 1.25, P_{sum} = \frac{A M_B \hat{P}_i}{1.25}$ 50

6.5 $A = 4, M_B = 3, K = 4, N = 2$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ . . . 51

6.6 $A = 3, M_B = 3, K = 3, N = 2$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ . . . 52

6.7 $A = 4, M_B = 4, K = 4, N = 4$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ . . . 53

6.8 $A = 4, M_B = 5, K = 4, N = 4$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ . . . 54

ABBREVIATIONS

MU	Multi-User
MIMO	Multiple-Input-Multiple-Output
MISO	Multiple-Input-Single-Output
SU	Single-User
BS	Base Station
MS	Mobile Station
SPC	Sum Power Constraint
PBPC	Per-Base Station Power constraints
PAPC	Per-Antenna Power Constraints
BD	Block Diagonalisation
mmWave	Millimeter Wave
PFGD	Projected Factored Gradient
BC	Broadcast Channel
DPC	Dirty Paper Coding
MMSE	Minimum Mean Squared Error
ZF-BF	Zero Forcing - Beam Forming
SVD	Singular value Decomposition
CoMP	Coordinated Multi Point
DoF	Degrees of Freedom
SNR	Signal to Noise Ratio
CSCG	Circularly Symmetric Complex Gaussian

NOTATIONS

a	Scalar
$\mathbf{a}$	Vector
$\mathbf{A}$	Matrix
$\mathbf{I}$	Identity Matrix
$\mathbf{0}$	All-Zero Matrix
$\vert\mathbf{Q}\vert$	Determinant of a square matrix $\mathbf{Q}$
$\text{Tr}(\mathbf{Q})$	Trace of a square matrix $\mathbf{Q}$
$\mathbf{Q}^{-1}$	Inverse of a full-rank square matrix $\mathbf{Q}$
$\mathbf{Q}^{1/2}$	Square root of a square matrix $\mathbf{Q}$
$\mathbf{Q} \succeq \mathbf{0}$	$\mathbf{Q}$ is positive semi-definite
$\text{Diag}(\mathbf{x})$	A diagonal matrix with the main diagonal given by $\mathbf{x}$
$\mathbf{A}^H$	Conjugate Transpose of $\mathbf{A}$
$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$\text{Rank}(\mathbf{A})$	Rank of $\mathbf{A}$
$\mathbf{A}^\dagger$	Pseudo Inverse of $\mathbf{A}$
$\mathbb{E}[\mathbf{x}]$	Statistical expectation or mean of the random vector $\mathbf{x}$
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	The distribution of a CSCG random vector with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{C}$
$\sim$	Distributed as
$\mathbb{C}^{M\times N}$	Space of complex matrices of dimension $M \times N$
$\vert\vert\mathbf{x}\vert\vert$	Euclidean or L-2 norm of a complex vector $\mathbf{x}$
$\vert\vert\mathbf{U}\vert\vert_F$	Frobenius norm of the matrix $\mathbf{U}$
$\nabla_{\mathbf{Q}}$	Gradient w.r.t. $\mathbf{Q}$

CHAPTER 1

INTRODUCTION

1.1 Single-Cell Setup and Block Diagonalisation

An active research in the past few years has been carried out to obtain optimal downlink transmissison strategies in various kinds of cellular and wireless systems under different sets of power constraints. Conventionally, the earlier works' focus has been on the downlink beamforming in a single-cell set-up with a multiple-transmit antenna base station (BS) and multiple single-/multiple-output mobile stations (MSs). In this system, the transmission can be modeled by a multiple-input single-/multiple-output (MISO/MIMO) broadcast channel (BC). For the Gaussian MISO/MIMO BC, it is known that the dirty paper coding (DPC) technique achieves the capacity region which constitutes all the simultaneously achievable rates for all the MSs (Weingarten *et al.*, 2006). However owing to the nonlinear complicatedness and the difficulty of implementation in real-time systems, a massive drift of attention towards linear transmit and receive beamforming schemes for the Gaussian MISO/MIMO BC has been observed in the literature (Rashid-Farrokhi *et al.*, 1998; Schubert and Boche, 2004; Peel *et al.*, 2005; Wiesel *et al.*, 2006; Stojnic *et al.*, 2006).

One such simple linear precoding scheme is block diagonalisation (BD) discussed in (Spencer *et al.*, 2004; Wong *et al.*, 2003; Choi and Murch, 2004; Pan *et al.*, 2004). To eliminate the inter-user interferences and thereby allow each MS perceive an interference-free MIMO channel, the BD scheme restricts the precoding matrix that is multiplied by the transmitted signal from a BS to the designated MS to be orthogonal to the downlink channels corresponding to all other MSs. BD reduces down to well-known zero-forcing beamforming (ZF-BF) in the special case of a MISO channel (Peel *et al.*, 2005). Although BD is in general inferior as compared to the DPC based optimal nonlinear precoding scheme or the minimum-mean-squared-error (MMSE) based optimal linear precoding scheme in terms of achievable rate, it performs very well in the high signal-to-noise-ratio (SNR) regime and achieves the same degrees of freedom (DoF) for the

MISO-/MIMO-BC sum-rate as the optimal linear/nonlinear precoding schemes (Caire and Shamai, 2003). Moreover, (Caire and Shamai, 2003) shows how BD can be genearalized to incorporate nonlinear DPC processing, which leads to ZF-DPC precoding.

The optimal transmission strategy and capacity formulae for MIMO Gaussian channels under a sum-power constraint (SPC) have been given by Telatar in (Telatar, 1999). Gaussian signaling with a transmit covariance matrix determined using the singular value decomposition (SVD) of the channel matrix and a water-filling algorithm is optimal under the SPC. The SPC limits the total power than can be used by the transmitter and such a constraint is usually imposed by regulations and a need to control the total energy consumption. In recent years, there has been an increasing interest to derive the optimal schemes under many other transmission constraints. Per-Antenna power constraints (PAPC), Per-Group power constraints (PGPC) and Per-Base Station power constraints(PBPC) may arise due to the hardware limitations in sharing the total available power across all the transmit antennas. In distributed antenna systems, the transmit antennas are spread across multiple locations and are not driven by the same power amplifier. In such a setting, the total power cannot be arbitrarily allocated across the different geographically separated antennas. This kind of situation arises in cellular systems using coordinated multipoint transmission (CoMP) and in cell-free massive MIMO. If the multiple transmission points (BSs) in CoMP are single antenna transmitters, then we get PAPC. If the BSs are multi-antenna transmission points, we get PBPC. Therefore, it is important to study the optimal MIMO transmission schemes under multiple sets of power constraints such as SPC, PBPC, PAPC.

While Gaussian signaling is optimal even under multiple power constraints, there is no general analytical solution for the optimal transmit covariance matrix and the capacity as in the case of SPC. Exact analytical solutions are limited to the MISO and some full rank MIMO settings. MIMO and MISO BC under PAPC have been extensively studied in (Vu, 2011*b*; Pi, 2012; Tuninetti, 2014; Vu, 2011*a*). Closed form solutions for both capacity and optimal transmit covariance matrix have been obtained for the MISO case under PAPC in (Vu, 2011*b*). Under the two assumptions that channel matrix being full column rank and the optimal transmit covariance matrix being full rank, a closed form solution for MIMO capacity under PAPC is obtained in (Tuninetti, 2014). Capacity of MIMO Gaussian channels has also been studied in (Pi, 2012; Vu, 2011*a*). In (Pi, 2012), an iterative algorithm is proposed to compute the capacity for single-stream under PAPC and multi-stream MIMO with *per-stream* PAPC. Any kind of closed form solutions are not provided in (Pi, 2012). In (Vu, 2011a), the authors have proposed an algorithm to compute the MIMO capacity under PAPC for the special case when the channel matrix has full column rank or full row rank. Optimal transmit strategies for MIMO Gaussian channels and the MIMO capacity under Joint-SPC-PAPC have been obtained in (Le Cao and Oechtering, 2017) via an iterative algorithm proposal.

1.2 Multi-Cell Systems and Motivation behind the project

A recent work (Chaluvadi *et al.*, 2018), one of the two main inspirations for this project, discusses MIMO capacity under multiple simlultaneous power constraints - SPC, PGPC and PAPC. They have derived the analytical solutions under Joint-SPC-PGPC-PAPC for the MISO channel, full column rank MIMO channel assuming optimal transmit co-variance matrix to be full rank and a $2 \times N_r$ where N_r is the total number of receive antennas in the MIMO BC. There has been a significant growth of interest in MIMO systems in the context of mmWave communications (Torkildson *et al.*, 2011; Sun *et al.*, 2014; Raghavan *et al.*, 2016). In these systems, the number of antennas is expected to be large because of the high frequency of operation. However the rank of the transmit co-variance matrix *viz.* the number of spatial streams transmitted is likely to be limited by: (1) Sparsity in the channel or the rank of the channel itself (2) The complexity of implementing the spatial multiplexing on hardware and requirement of several RF chains. Similar constraints are expected to arise in the massive MIMO systems. Therefore, it is imporatnt to understand the capacity of the MIMO channel under the aforementioned rank constraint. The other main contribution of (Chaluvadi *et al.*, 2018) is a projected factored gradient descent algorithm in the general single user (SU) MIMO BC case, to find the optimal transmission strategy under Joint-SPC-PGPC-PAPC and rank constraint. In this proposed algorithm, the optimal precoding/beamforming matrix(square root of the transmit covariance matrix) is directly found instead of the optimal transmit covariance matrix.

Until now, the discussion has only been about a single-cell downlink transmission. A lot of work has also been done by shifting the design paradigm to a multi-cell cooperative downlink transmission in the last decade (Shamai and Zaidel, 2001; Zhang and

Dai, 2004; Karakayali *et al.*, 2006; Somekh *et al.*, 2007; Jing *et al.*, 2007; Kaviani and Krzymien, 2008). In these studies, it is assumed that BSs in a cellular network are connected via backhaul links to a central processing unit (e.g., a dedicated control station or a preassigned BS), which has the global knowledge of downlink channels from each BS to all the MSs transmit messages for all the MSs in the network. Thereby, the central processing unit is able to jointly design the downlink transmissions for all the BSs and provide them with appropriate transmit signals. As demonstrated in these works, the cooperative multi-cell downlink processing leads to enormous throughput gains by utilising the co-channel interference across different cells in a coherent fashion as compared to the conventional single-cell processing with the co-channel interference treated as noise. Moreover, distributed multi-cell downlink beamforming via the use of belief propagation and and message passing among BSs has also been proposed in (Ng *et al.*, 2008), without the need for a central controller.

Initially, the BD precoder design subject to Per-Base Station power constraints (PBPC) has been studied in (Boccardi and Huang, 2006; Liu *et al.*, 2009; Zhang *et al.*, 2008). In these woks, the BD precoders are designed essentially following the same principle as for the conventional sum-power constraint (SPC) i.e., the precoding vectors known for the SPC case are adopted, and then the power allocation is optimised to maximize the sum-rate under PBPC. However, it remains unclear whether the developed BD precoder solutions therein are indeed optimal for the weighted sum-rate maximization in a cooperative multi-cell system. Optimum zero-forcing beamforming (ZF-BF) with per-antenna power constraints is investigated in (Karakayali *et al.*, 2007) for a MISO BC case. The authors show that standard zero-forcing techniques, such as the Moore-Penrose pseudo inverse, which are optimal in the context of sum-power-constrained systems are actually suboptimal when there are per-antenna power constraints. They formulate convex optimization problems to find the optimum ZF-BF vectors.

Soon after the above works have been published, Rui Zhang has authored (Zhang, 2010) which is the other main inspirational work for this project. In this work, the author shows that the BD precoder designs following the above heuristic approach of *separating* the beamforming design and power allocation are indeed suboptimal for rate maximization, while the optimal BD precoder solution requires a new *joint* beamforming and power-allocation optimization approach. In more detail, (Zhang, 2010) formulates the MU-MIMO-BC transmit optimization problem with the BD precoding and equivalent Per-Base Station power constraints (PBPC) as a *convex optimization* problem and designs an efficient optimal algorithm to solve this problem. The author also derives the closed-form expression of the optimal BD precoding matrix to maximize the weighted sum-rate for the MU-MIMO-BC. It is also proved in this work that the optimal BD precoding(beamforming) vectors for each MS in the case of Per-BS power constraints are in general *non-orthogonal*, which differs from the conventional orthogonal BD precoder design for the sum-power constraint case. Consequently, it is proved that the orthogonal BD precoder designs proposed in prior works (Boccardi and Huang, 2006; Liu *et al.*, 2009; Zhang *et al.*, 2008) for the Per-BS power constraints are in general suboptimal(for weighted sum-rate maximization). The paper also presents a low-complexity, suboptimal scheme for the same problem.

1.3 Project's Main Contributions

In this paper, we focus our study on the BD-based downlink precoding for a fully cooperative multi-cell system equipped with a central processing unit, which is assumed to have the perfect knowledge of all the downlink channels and the transmit messages in the network.

This project's main contributions are summarized as follows :

1.3.1 What is new in the project ?

Both the Joint-SPC-PBPC-PAPC and the rank-constrained Joint-SPC-PBPC-PAPC problems in the context of a fully cooperative block diagonalised multi-cell(multi-user too) system have not been studied earlier and hence we have chosen these two topics to be presented in this work.

1.3.2 Optimal Solution without rank constraints

First, we formulate the cooperative multi-cell MIMO system in hand as a MU-MIMO-BC and design the transmit optimisation problem with the objective of maximizing the weighted sum-rate under BD precoding constraints and Joint-SPC-PBPC-PAPC power constraints as a *convex* optimization problem. Note that we still haven't introduced the rank constraints on any of the MSs as of now i.e. we assume the number of spatial streams allowed for each MS is the number of receive antennas on the MS (maximum allowable rank).

By applying linear algebra and convex optimization. techniques, we design an efficient optimal algorithm on the similar lines of the algorithm proposed in (Zhang, 2010) to solve this problem. We also derive the closed form expression of the optimal precoding matrix to maximize the weighted sum-rate for the MU-MIMO-BC in a similar fashion as in (Zhang, 2010).

More importantly, we conclude that the optimal BD precoding(beamforming) vectors for each MS in the case of Joint-SPC-PBPC-PAPC are in general *non-orthogonal*, which strengthens the argument in (Zhang, 2010). For the special case of single-antenna BSs and MSs, the proposed BD precoding design for the MU-MIMO-BC provides the optimal ZF-BF precoder solution to maximize the weighted sum-rate for the MU-MISO-BC with Joint-SPC-PAPC.

1.3.3 Suboptimal Solution without rank constraints

We also present a low-complexity, sub-optimal scheme for the studied problem, which is obtained by computing the conventional BD precoder design for the SPC case with an optimal power allocation to meet the Joint-SPC-PBPC-PAPC. We also derive the closed-form expression of the suboptimal precoding matrix. This scheme can be considered as a direct extension of that given in (Zhang, 2010) for the MU-MIMO BC with the BD precoding constraints and Per-Base Station power constraints.

1.3.4 PFGD Algorithm with rank constraints

Next , we formulate the MU-MIMO-BC transmit optimisation problem for the given fully cooperative multi-cell system with the BD precoding constraints, Joint SPC-PBPC-PAPC power constraints and *rank constraints* (must be obviously less than the number of receive antennas on the MS) on all the MSs i.e. we are limiting the number of spatial streams through which data can be transmitted for each MS. Again by using linear algebra and convex optimization techniques, we propose an efficient numerical projected factored gradient descent (PFGD) algorithm to find the optimal transmisssion strategy which maximizes the weighted sum-rate under the aforementioned constraints. This algorithm can be thought of as an extension of the PFGD algorithm presented in (Chaluvadi *et al.*, 2018) for the general SU-MIMO-BC case in the scenario of a single-cell system set-up. In this method, instead of solving for the covariance matrix, the algorithm determines the precoding /beamforming matrix *viz.* square root of the transmit covariance matrix directly. Numerical results show that the solution from the PFGD algorithm matches with the solution proided by standard convex optimization package, CVX (Grant and Boyd, 2014). The PFGD algorithm can actually take advantage of the low rank structure for a reduced complexity solution.

1.3.5 Simulation Results

In both the MU-MIMO BC and MU-MISO BC cases for two cooperative multi-cell systems without a rank constraint(or assuming full possible rank), we compare the achievable rates of both the optimal and suboptimal schemes as discussed in the second and the fourth points respectively. We also plot the rate curves using PFGD algorithm for full-rank case and show that eventually PFGD algorithm matches the optimal algorithm that we have proposed initially. We also plot and compare the achievable rate curves obtained using PFGD algorithm for MU-MIMO BC under different rank constraints for two different multi-cell systems. We also provide a plot on how fast the the PFGD algorithm converges for a particular network realization.

CHAPTER 2

SYSTEM MODEL AND PROBLEM FORMULATION

2.1 System Model

We consider a multi-cell system consisting of A cells each of which has a single BS. For our convenience, let us assume there are $M_B \geq 1$ transmit antennas on each BS. Let us consider there are $K_a \geq 1$ users in every cell where "a" is the BS-/cell-index. Let us denote the total number of MSs (users) in the multi-cell system by $K = \sum_{a=1}^{A} K_a$, and also assume that each and every MS has the same number of receive antennas on it, denoted by $N \geq 1$. Denote the total number of transmit antennas across all the BSs in the system as $M = M_B A$. We can conveniently model the jointly designed downlink transmission for all the BSs in this fully cooperative multi-cell system as an auxiliary MU-MIMO BC with M transmit antennas and K users each with N receive antennas. In this auxiliary MU-MIMO BC, we assign the indices from $(a-1)M_B + 1$ to aM_B to the transmit antennas on the BS indexed with label $a = 1, \cdots, A$ for convenience of notation later on in the paper. Similarly, the indices of MSs in the auxiliary MU-MIMO BC are associated with the index of the cell they correspond to. In detail, cell "a" contains MSs with indices from $\sum_{i=1}^{a-1} K_i + 1$ to $\sum_{i=1}^{a} K_i$.

The discrete-time baseband signal of each auxiliary MIMO BC corresponding to K MSs is thus given by the linear model

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \sum_{j \neq k} \mathbf{H}_k \mathbf{x}_j + \mathbf{n}_k, \quad k = 1, \cdots, K \tag{2.1}$$

where $\mathbf{x}_k \in \mathbb{C}^{M \times 1}$ denotes the transmit signal vector designated for k-th MS from M transmit antennas, $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$ denotes the received signal vector at the k-th MS. $\mathbf{H}_k \in \mathbb{C}^{N \times M}$ denotes the downlink channel from all the M transmit antennas on all the BSs to the k-th MS. $\mathbf{n}_k \in \mathbb{C}^{N \times 1}$ denotes the receiver noise at the k-th MS. For our convenience, we assume that $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}), \forall k = 1, \cdots, K$.

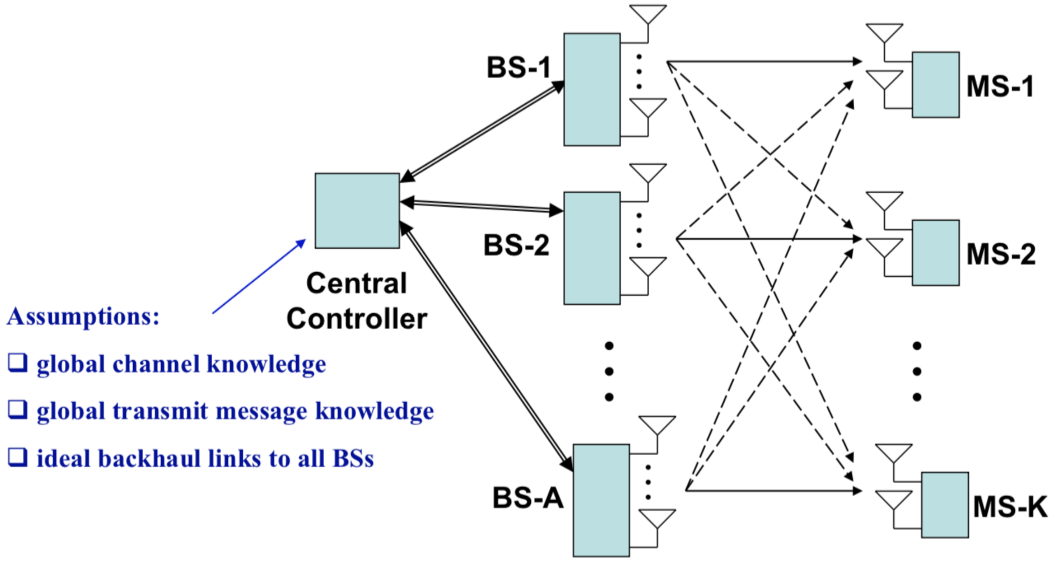

Figure 2.1: Multi User - MIMO BC

2.2 Relation between Transmit Covariance Matrices and Precoding Matrices

Without loss of generality, we can further express $\mathbf{x}_k$ as a product of the precoding matrix of the k-th MS, $\mathbf{U}_k \in \mathbb{C}^{M\times R_k}$ and the information-signal for the k-th MS, $\mathbf{s}_k \in \mathbb{C}^{R_k\times 1}$.

$$\mathbf{x}_k = \mathbf{U}_k\mathbf{s}_k, \quad k = 1, \cdots, K \tag{2.2}$$

We should note that the precoding matrices specify both the transmit beamforming vectors and allocated power values for different data streams. Here, R_k denotes the number of permitted data streams for the k-th MS due to spatial multiplexing. We note that $R_k \leq \min(M, N), \forall k = 1, \cdots, K$. Denote $\mathbf{Q}_k = \mathbb{E}[\mathbf{x}_k\mathbf{x}_k^H] \in \mathbb{C}^{M\times M}$ as the transmit covariance matrix corresponding to the k-th MS, with $\mathbf{Q}_k \succeq \mathbf{0}$. Now, we note that both the matrices $\mathbf{Q}_k$ and $\mathbf{U}_k$ have rank R_k. We assume the information-bearing signal vectors $\mathbf{s}_k$ are independent over k and we also assume a Gaussian codebook is used for each MS at the transmitter antennas and thus, $\mathbf{s}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}), \forall k = 1, \cdots, K$. Therefore, we can express the precoding matrix as a square root of the transmit covariance matrix i.e., $\mathbf{Q}_k = \mathbf{U}_k\mathbf{U}_k^H$. The overall downlink transmit covariance matrix for the M cooperating transmit antennas is thus given by $\mathbf{Q} = \sum_{k=1}^{K}\mathbf{Q}_k$.

2.3 Multiple Power Constraints

As discuused earlier, these transmit covariance matrices need to satisfy three different sets of power constraints briefly discussed below.

2.3.1 Sum Power Constraints (SPC)

In this case, the total average power across all the M transmitting antennas is limited to P_{sum}. Mathematically speaking, the transmit covariance matrices $\mathbf{Q}_k$ must satisfy the inequality given below.

$$\text{Tr}(\mathbf{Q}) \leq P_{sum} \ \text{ or } \ \sum_{k=1}^{K} \text{Tr}(\mathbf{Q}_k) \leq P_{sum} \tag{2.3}$$

2.3.2 Per-Base Station Power Constraints (PBPC)

Under PBPC, the BS “a" has an average sum transmit power constraint across all it's M_B transmit antennas given by $\tilde{P}_a \forall a$. Hence, the transmit covariance matrices corresponding to the K MSs must satisfy the following inequalities.

$$\text{Tr}(\mathbf{B}_a\mathbf{Q}) \leq \tilde{P}_a \ \text{ or } \ \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_a\mathbf{Q}_k) \leq \tilde{P}_a, \ \ a = 1, \cdots, A \tag{2.4}$$

where

$$\mathbf{B}_a \triangleq \text{Diag}(\underbrace{0, \cdots, 0}_{(a-1)M_B}, \underbrace{1, \cdots, 1}_{M_B}, \underbrace{0, \cdots, 0}_{(A-a)M_B}) \tag{2.5}$$

2.3.3 Per-Antenna Power Constraints (PAPC)

Under PAPC, the average transmit power of the i-th transmit antenna is constrained by $\hat{P}_i \forall i = 1, \cdots, M$. Therefore in mathematical terms, the transmit covariance matrices corresponding to the K MSs must satisfy the inequalities given below.

$$\text{Tr}(\mathbf{B}_i\mathbf{Q}) \leq \hat{P}_i \ \text{ or } \ \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_i\mathbf{Q}_k) \leq \hat{P}_i, \ \ i = 1, \cdots, M \tag{2.6}$$

where

$$\mathbf{B}_i \triangleq \text{Diag}(\underbrace{0, \cdots, 0}_{(i-1)}, 1, \underbrace{0, \cdots, 0}_{(M-i)}) \tag{2.7}$$

Let us denote this joint power constraints set as Joint-SPC-PBPC-PAPC. We note that in the special case of single-antenna BSs and MSs i.e., $M_B = N = 1$, the per-BS power constraints in (2.4) reduce to per-antenna power constraints in (2.6) for an equivalent MU-MISO BC.

2.4 Quasi Static Fading Assumption

We consider a scenario where the downlink channels corresponding to K MSs in the auxiliary MU-MIMO BC remain constant for a particular given downlink transmission frame i.e., we are basically dealing with a quasi-static fading environment.

2.5 Block Diagonalisation Precoding Constraints

Let us now consider the inter-user interference eliminating linear BD precoding scheme for each downlink transmission frame in the MU-MIMO BC, i.e., we should have for each given k, $\mathbf{H}_j \mathbf{x}_k = \mathbf{0}$ in (2.1) or more precisely $\mathbf{H}_j \mathbf{U}_k = \mathbf{0}, \forall j \neq k$. We can also present the aforementioned variant of ZF-BF constraints as below.

$$\mathbf{H}_j \mathbf{Q}_k \mathbf{H}_j^H = \mathbf{0}, \ \ \forall j \neq k \tag{2.8}$$

We assume that the row vectors in all downlink channels, $\mathbf{H}_k$'s are linearly independent (due to the fact that we have considered an independent fading scenario). From the BD precoding constraints in (2.8) $\forall k = 1, \cdots, K$, it follows that $M \geq NK$ should be a necessary condition in order to get feasible transmit covariance matrices, $\mathbf{Q}_k$'s with rank $R_k \leq \min(M, N) = N, \forall k$ (since $K \geq 1$).

Without loss of generality, we assume that all the MSs in the multi-cell system have the same number of permitted data streams i.e., $\mathbf{Q}_k$'s and $\mathbf{U}_k$'s have the same rank, $\forall k = 1, \cdots, K$. In Chapter 3, the rank constraint is made inactive i.e., $R_k = N, \forall k$ whereas in Chapter 4, we consider the rank constraint to be active but equal for all MSs as mentioned just now.

In practice, the total number of MSs in the system can be very large such that the above condition fails. In such scenarios, the central processing unit (CPU) schedules the transmission from BSs to MSs into different time-slots/frequency-bands. Now the CPU also makes sure that in each time-slot/frequency-band, the number of MSs scheduled for transmission satisfies the aforementioned condition. For further reference, the interested readers may go to (Yoo and Goldsmith, 2006; Shen *et al.*, 2005) for understanding the detailed design of downlink transmission scheduling in the MISO/MIMO BC with ZF-BF/BD precoding. For the rest of this paper, we assume $M \geq NK$.

2.6 Convex Optimisation Problem to be solved in Chapter 3 - (P1)

We are now all set to present the weighted sum-rate maximization problem for the downlink transmission in a fully cooperative multi-cell system modeled as an auxiliary MU-MIMO BC under Joint-SPC-PBPC-PAPC and BD precoding power constraints as follows.

$$
\text{(P1)}: \quad
\begin{aligned}
\max_{\mathbf{Q}_1,\cdots,\mathbf{Q}_K} \quad & \sum_{k=1}^{K} w_k \log|\mathbf{I} + \mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H| \\
\text{s.t.} \quad & \mathbf{H}_j\mathbf{Q}_k\mathbf{H}_j^H = 0, \ \forall j \neq k \\
& \sum_{k=1}^{K} \text{Tr}(\mathbf{Q}_k) \leq P_{sum} \\
& \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_a\mathbf{Q}_k) \leq \tilde{P}_a, \ \ a = 1,\cdots,A \\
& \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_i\mathbf{Q}_k) \leq \hat{P}_i, \ \ i = 1,\cdots,M \\
& \mathbf{Q}_k \succeq \mathbf{0}, \ \forall k
\end{aligned}
$$

where w_k is the given non-negative weight for the capacity corresponding to the k-th MS. In order to expose the capacities of all MSs, we let $w_k \geq 0, \forall k$.

A very important observation in (P1) is that we have posed the optimization problem with the transmit covariance matrices, $\mathbf{Q}_k, \forall k$ as the optimization variables instead of the precoding matrices $\mathbf{U}_k, \forall k$. We have purposefully done this beacuse it is easy to verify that (P1) is a convex optimization problem with $\mathbf{Q}_k$'s, since the objective function is concave over $\mathbf{Q}_k$'s and all the constraints specify a convex set over $\mathbf{Q}_k$'s.

However, we should note that we haven't introduced any rank constraints in (P1) as they don't specify a convex set over $\mathbf{Q}_k$'s. Thus, (P1) can be solved using standard optimization techinques, e.g., interior-point method (Boyd and Vandenberghe, 2004) or directly using CVX software package (Grant and Boyd, 2014) but such an approach would not reveal the optimal BD precoding matrix structure which is investigated in detail in Chapter 3.

2.7 Non-Convex Optimisation Problem to be solved in Chapter 4 - (P2)

If we choose $\mathbf{U}_k$'s as the design variables, both the objective and BD precoding constraints fail the requirements for a convex optimization problem formulation.

Therefore if we substitute $\mathbf{Q}_k = \mathbf{U}_k\mathbf{U}_k^H, \forall k$ in (P1), we have a non-convex optimization problem in hand. Let's utilise this opportunity and introduce the non-convex rank constraints also now to formulate a non-convex optimization problem (P2) as shown above. This problem (P2) would be solved numerically in Chapter 4.

$$
\text{(P2)}: \quad
\begin{aligned}
\max_{\mathbf{U}_1,\cdots,\mathbf{U}_K} \quad & \sum_{k=1}^{K} w_k \log |\mathbf{I} + (\mathbf{H}_k\mathbf{U}_k)(\mathbf{H}_k\mathbf{U}_k)^H| \\
\text{s.t.} \quad & \mathbf{H}_j\mathbf{U}_k = 0, \ \ \forall j \neq k \\
& \sum_{k=1}^{K} \text{Tr}(\mathbf{U}_k\mathbf{U}_k^H) \leq P_{sum} \\
& \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_a\mathbf{U}_k\mathbf{U}_k^H) \leq \tilde{P}_a, \ \ a = 1,\cdots,A \\
& \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_i\mathbf{U}_k\mathbf{U}_k^H) \leq \hat{P}_i, \ \ i = 1,\cdots,M \\
& \text{rank}(\mathbf{U}_k) = R_k(\leq N), \ \ k = 1,\cdots,K
\end{aligned}
$$

CHAPTER 3

OPTIMAL PRECODING MATRIX SOLUTION WITHOUT RANK CONSTRAINTS

In this chapter, we first derive an optimal structure for the precoding matrices and also present a very efficient algorithm to solve (P1) both of which are direct extensions of Rui Zhang's work in (Zhang, 2010) for the general case with arbitrary number of transmit and receive antennas at the BSs and the MSs respectively. Then, we extend this developed solution for the special case of single-antenna BSs/MSs i.e., auxiliary MISO BC.

3.1 General MU-MIMO BC Case

3.1.1 Eliminate BD Precoding Constraints

To solve (P1), we would first want to get rid of the BD precoding constraints given in (2.8), as follows:

Define $\mathbf{G}_k = [\mathbf{H}_1^T, \cdots, \mathbf{H}_{k-1}^T, \mathbf{H}_{k+1}^T, \cdots, \mathbf{H}_K^T]^T, \forall k = 1, \cdots, K$, where $\mathbf{G}_k \in \mathbb{C}^{L \times M}$ with $L = N(K-1)$. Let the normal singular value decomposition (SVD) of $\mathbf{G}_k$ be denoted as $\mathbf{G}_k = \mathbf{A}_k \mathbf{\Sigma_k} \mathbf{B}_k^H$ where $\mathbf{A}_k \in \mathbb{C}^{L \times L}, \mathbf{B}_k \in \mathbb{C}^{M \times M}$ are both unitary matrices i.e., $\mathbf{A}_k \mathbf{A}_k^H = \mathbf{A}_k^H \mathbf{A}_k = \mathbf{I}_L$ and $\mathbf{B}_k \mathbf{B}_k^H = \mathbf{B}_k^H \mathbf{B}_k = \mathbf{I}$ and $\mathbf{\Sigma}_k$ is a non-negative $L \times M$ matrix.

Anyways, $L \leq M$ (since $NK \leq M$) which means that the rank of the matrix $\mathbf{G}_k$ is given by rank($\mathbf{G}_k$) = min(L, M) = L. We can further express $\mathbf{\Sigma}_k, \mathbf{B}_k$ as $\mathbf{\Sigma}_k = [\mathbf{\Sigma}_{k_p} | \mathbf{0}]$ and $\mathbf{B}_k = [\mathbf{B}_{k_1} | \mathbf{B}_{k_0}]$ where $\mathbf{\Sigma}_{k_p}$ is a positive diagonal matrix of dimension $L \times L$ containing only positive singular values of the matrix $\mathbf{G}_k$, $\mathbf{B}_{k_1} \in \mathbb{C}^{M \times L}$ and $\mathbf{B}_{k_0} \in \mathbb{C}^{M \times (M-L)}$ are two parts of the right singular matrix, $\mathbf{B}_k$ that correspond to positive and zero singular values respectively.

It is easy to verify now that the normal SVD of $\mathbf{G}_k$ can be reduced to $\mathbf{G}_k = \mathbf{A}_k \mathbf{\Sigma}_{k_p} \mathbf{B}_{k_1}^H$. Based on the above build-up, we next propose the following lemma.

Lemma 1 *The optimal solution of (P1) takes the form*

$$\mathbf{Q}_k = \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H, \quad k = 1, \cdots, K \tag{3.1}$$

where $\mathbf{S}_k \in \mathbb{C}^{(M-L)\times(M-L)}$ *and* $\mathbf{S}_k \succeq \mathbf{0}$

Proof: Please refer to Appendix A.

Remark 1: In prior works (Spencer *et al.*, 2004; Wong *et al.*, 2003; Choi and Murch, 2004; Pan *et al.*, 2004) and (Zhang, 2010) on the design of BD precoder for the MIMO BC in the case of SPC and PBPC respectively, it has been proved that the columns(precoding vectors) in the BD precoding matrix for the k-th MS, $\mathbf{U}_k$ with $\mathbf{U}_k \mathbf{U}_k^H = \mathbf{Q}_k$, must be a linear combinations of the columns in $\mathbf{B}_{k_0}$ in order to satisfy the BD precoding constraints given in (P2) $\mathbf{H}_j \mathbf{U}_k = \mathbf{0}, \forall j \neq k$. Lemma 1 just extends this result to the case of Joint-SPC-PBPC-PAPC constraints.

3.1.2 Convex Optimisation Problem (P3) - A Modified Version of (P1)

So, the optimal structure for the transmit covariance matrix suggested in (3.1) satisfies all the Zero-forcing constraints given in (2.8) and hence they can be eliminated from (P1). Thus we can reduce (P1) to the following equivalent problem.

$$\text{(P3)}: \quad \begin{aligned} \max_{\mathbf{S}_1,\cdots,\mathbf{S}_K} \quad & \sum_{k=1}^{K} w_k \log |\mathbf{I} + \mathbf{H}_k \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H \mathbf{H}_k^H| \\ \text{s.t.} \quad & \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H) \leq P_{sum} \\ & \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_a \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H) \leq \tilde{P}_a, \quad a = 1, \cdots, A \\ & \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_i \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H) \leq \hat{P}_i, \quad i = 1, \cdots, M \\ & \mathbf{S}_k \succeq \mathbf{0}, \quad \forall k \end{aligned}$$

We can verify that (P3) is a convex optimization problem based on similar arguments that have been made earlier in proving (P1) as convex.

3.1.3 Procedure to obtain Optimal Solution of (P3)

We would like the solve the problem (P3) using Lagrange-Duality method. Let the dual variables associated with the SPC, PBPC and PAPC in (P3) be $\gamma, \{\lambda_a\}_{a=1}^{A}, \{\mu_i\}_{i=1}^{M}$ respectively, then we can write the Lagrangian function for the above problem (P3) as follows.

$$\begin{aligned} L(\{\mathbf{S}_k\}, \gamma, \{\lambda_a\}, \{\mu_i\}) = \sum_{k=1}^{K} w_k \log |\mathbf{I} + \mathbf{H}_k \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H \mathbf{H}_k^H| \\ +\gamma(P_{sum} - \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H)) \\ +\sum_{a=1}^{A} \lambda_a(\tilde{P}_a - \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_a \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H)) \\ +\sum_{i=1}^{M} \mu_i(\hat{P}_i - \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_i \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H)) \end{aligned} \tag{3.2}$$

where $\{\mathbf{S}_k\}, \{\lambda_a\}$ and $\{\mu_i\}$ denote the set of $\mathbf{S}_k$'s, λ_a's and μ_i's respectively. The Lagrange Dual function for (P3) is then defined as

$$g(\gamma, \{\lambda_a\}, \{\mu_i\}) = \max_{\mathbf{S}_k \succeq \mathbf{0}, \forall k} L(\{\mathbf{S}_k\}, \gamma, \{\lambda_a\}, \{\mu_i\}) \tag{3.3}$$

Also, the dual problem of (P3) is defined as

$$\text{(P3-D)}: \quad \begin{aligned} \min_{\gamma, \{\lambda_a\}, \{\mu_i\}} \quad & g(\gamma, \{\lambda_a\}, \{\mu_i\}) \\ \text{s.t.} \quad & \gamma \geq 0 \\ & \lambda_a \geq 0, \ \ \forall a = 1, \cdots, A \\ & \mu_i \geq 0, \ \ \forall i = 1, \cdots, M \end{aligned}$$

Since it is already known that (P3) is a convex optimization problem and satisfies the Slater's conditions (Boyd and Vandenberghe, 2004), the primal optimal objective value is equal to the dual optimal objective value i.e., the duality gap between the optimal objective values of (P3) and (P3-D) is zero. Therefore, we can equivalently solve (P3-

D) instead of solving (P3). Towards this end, we first need to solve for optimal $\{\mathbf{S}_k^*\}$ that maximizes the Lagrangian function in (3.2) in terms of the given dual variables. We should also note that $\{\mathbf{S}_k^*\}$ is the optimal solution for the maximization problem in (3.3) with the set of given $\gamma, \{\lambda_a\}, \{\mu_i\}$.

Moreover, the dual problems are always convex and hence (P3-D) is also convex. Finally, we can obtain optimal dual variables $\gamma^*, \{\lambda_a^*\}, \{\mu_i^*\}$ iteratively using the subgradient method e.g., the ellipsoid method (Bland *et al.*, 1981; Boyd and Barratt, 2008), given the fact that the subgradients (partial derivatives) of function $g(\gamma, \{\lambda_a\}, \{\mu_i\})$ at a set of already fixed γ, λ_a's and μ_i's are $P_{sum} - \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_{k_0}\mathbf{S}_k^*\mathbf{B}_{k_0}^H)$ for γ, $\tilde{P}_a - \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_a\mathbf{B}_{k_0}\mathbf{S}_k^*\mathbf{B}_{k_0}^H)$ for $\lambda_a, a = 1, \cdots, A$ and $\hat{P}_i - \sum_{k=1}^{K} \text{Tr}(\mathbf{B}_i\mathbf{B}_{k_0}\mathbf{S}_k\mathbf{B}_{k_0}^H)$ for $\mu_i, i = 1, \cdots, M$ in that order.

Now we can substitute these optimal dual variables in the optimal structure obtained for $\mathbf{S}_k^*$ above and later obtain an optimal BD precoding matrix/transmit covariance matrix structure. This is validated by the fact that *strong duality holds* for the convex optimization problem (P3).

3.1.4 Relevant Power Constraints

Before we proceed further with our arguments on how to solve for optimal BD precoding matrix structure, it is important for us to take a step back and note down the following which are useful later on in the paper.

It is sufficient to consider the case where

$$P_{sum} \leq \sum_{a=1}^{A} \tilde{P}_a \quad \& \quad \tilde{P}_a \leq \sum_{i=(a-1)M_B+1}^{aM_B} \hat{P}_i \tag{3.4}$$

SPC is redundant if $P_{sum} > \sum_{a=1}^{A} \tilde{P}_a$ and in that case we can achieve the capacity without any loss by setting $P_{sum} = \sum_{a=1}^{A} \tilde{P}_a$.

Similarly for any $a \in \{1, \cdots, A\}$, if $\tilde{P}_a > \sum_{i=(a-1)M_B+1}^{aM_B} \hat{P}_i$ then the Per-Base Station power constraint corresponding to the BS "a" is redundant. In this case, we can still achieve the same capacity by making $\tilde{P}_a = \sum_{i=(a-1)M_B+1}^{aM_B} \hat{P}_i$ and obviously save some power.

For the rest of this paper hereafter, we work with power constraints that satisfy the sufficient conditions in (3.4). Next, we present an idea that has already been proposed in (Chaluvadi *et al.*, 2018).

Lemma 2 *In the case of Joint-SPC-PBPC-PAPC, with the sufficient conditions being enforced viz.* $P_{sum} \leq \sum_{a=1}^{A} \tilde{P}_a \; \& \; \tilde{P}_a \leq \sum_{i=(a-1)M_B+1}^{aM_B} \hat{P}_i, \forall a = 1, \cdots, A$*, the optimal transmission strategy* $\{\mathbf{Q}_k^*\}$ *uses full sum power i.e.,* $\sum_{k=1}^{K} Tr(\mathbf{Q}_k^*) = P_{sum}$*.*

Proof: Please refer to Appendix B

3.1.5 Maximisation Problem (P4) and it's Optimal Solution

Next, we continue forward to obtain the optimal $\{\mathbf{S}_k^*\}$ with a set of given dual variables. From (3.2), it is evident that the maximization problem in (3.3) can be split into K independent maximization subproblems each involving only one $\mathbf{S}_k$ as the primal variable. After discarding the irrelevant terms, we can express the corresponding subproblem, for each given k as below.

$$(\text{P4}): \quad \max_{\mathbf{S}_k \succeq \mathbf{0}} \; w_k \log |\mathbf{I} + \mathbf{H}_k \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H \mathbf{H}_k^H| \\ - \text{Tr}(\mathbf{B}_{\gamma\lambda\mu} \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H)$$

where $\mathbf{B}_{\gamma\lambda\mu} \triangleq \gamma \mathbf{I} + \sum_{a=1}^{A} \lambda_a \mathbf{B}_a + \sum_{i=1}^{M} \mu_i \mathbf{B}_i \triangleq \mathbf{B}_\gamma + \mathbf{B}_\lambda + \mathbf{B}_\mu$. The notations $\mathbf{B}_\gamma, \mathbf{B}_\lambda, \mathbf{B}_\mu$ are self explanatory. We must observe that $\mathbf{B}_{\gamma\lambda\mu}$ is a diagonal matrix with the diagonal elements given by $\gamma + \lambda_a + \mu_i$ in the order of $a = 1, \cdots, A$ and $i = 1, \cdots, M$.

We have to discuss about the rank of the above matrix i.e, the number of positive entries in $\mathbf{B}_{\gamma\lambda\mu}$. By applying one of the Karash-Kuhn-Tucker conditions (Boyd and Vandenberghe, 2004), the complementary slackness theorem, it is evident that if the dual variable corresponding to the SPC, $\gamma > 0$, then it leads to having a tight sum power constraint with the optimal solution for $\{\mathbf{S}_k\}$. Therefore, if $\gamma > 0$, then the diagonal matrix $\mathbf{B}_{\gamma\lambda\mu}$ is full rank as it has all positive entries in it's diagonal.

Now, we would want to find a lower bound on the rank($\mathbf{B}_{\gamma\lambda\mu}$) to deal with cases where $\gamma = 0$. This lower bound is presented in the following lemma, very crucial for our arguments next. The proof of this lemma goes on similar lines of the proof of Lemma 2 presented in (Zhang, 2010) where the author derives a lower bound on the number of active Per-Base Station power constraints in a block diagonalised multi-cell system with only PBPC.

Lemma 3 *For (P4) to have a bounded objective value, it holds that rank*$(\mathbf{B}_{\gamma\lambda\mu}) = R_{\mathbf{B}_{\gamma\lambda\mu}} \geq M - N(K-1)$.

Proof: Please refer to Appendix C.

Remark 2: The fact that the sum power constraint is tight with the optimal transmit covariance matrices set $\{\mathbf{Q}_k^*\}$ under sufficient conditions provided in (3.4) does not imply that the optimal dual variable corresponding to the sum power constraint γ^* is positive, it's value can be zero too. The complementary slackness theorem under the assumption of strong duality states that atleast one of the two following conditions have to be met. (1) The optimal dual variable corresponding to the constraint is zero, (2) The constraint is tight with the optimal primal variables. But, if we have some optimal dual variables $\lambda_a^* > 0$ for some $a \in \{1, \cdots, A\}/\mu_i^* > 0$ for some $i \in \{1, \cdots, M\}$, then we can straightaway conclude that the corresponding Per-Base Station /Per-Antenna power constraint is tight with optimal $\mathbf{Q}_k^*$.

Let's consider Tr$(\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0}\mathbf{S}_k\mathbf{B}_{k_0}^H)$ and perform some algebraic manipulations on this term to aid us later on in the paper as follows.

$$\begin{aligned} & \text{Tr}(\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0}\mathbf{S}_k\mathbf{B}_{k_0}^H) \\ & = \text{Tr}(\mathbf{S}_k\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0}) && (3.5) \\ & = \text{Tr}((\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{1/2}\mathbf{S}_k(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{1/2}) && (3.6) \\ & = \text{Tr}(\tilde{\mathbf{S}}_k) && (3.7) \end{aligned}$$

where (3.5), (3.6) follow from the fact that Tr$(\mathbf{XY}) =$ Tr$(\mathbf{YX})$. Also in (3.6), the matrix $(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{1/2}$ is the square root of the positive semi-definite matrix $\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0}$ (since $\mathbf{B}_{\gamma\lambda\mu} \succeq \mathbf{0}$). We should note that $\tilde{\mathbf{S}}_k \succeq \mathbf{0}$ if $\mathbf{S}_k \succeq \mathbf{0}$. With Lemma 3 and $L = N(K-1)$, we can assume without loss of generality that rank$(\mathbf{B}_{\gamma\lambda\mu}) \geq (M - L)$ since we are only interested in the scenario where the objective value of the problems, (P1) and (P4) are both bounded. Accordingly, we have rank$(\mathbf{B}_{k_0}\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0}^H) =$

$\min(\text{rank}(\mathbf{B}_{\gamma\lambda\mu}), M-L) = M-L$. Thus, $\mathbf{B}_{k_0}\mathbf{S}_k\mathbf{B}_{k_0}^H \in \mathbb{C}^{(M-L)\times(M-L)}$ is a full-rank matrix and hence invertible. Therefore, from (3.7), we can write

$$\mathbf{S}_k = (\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2}\tilde{\mathbf{S}}_k(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2} \tag{3.8}$$

and we can re-formulate (P4) to get the following maximization problem (P5).

$$\text{(P5)} : \quad \begin{aligned} \max_{\tilde{\mathbf{S}}_k \succeq \mathbf{0}} \; & w_k \log|\mathbf{I} + \mathbf{H}_k\mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2} \times \\ & \tilde{\mathbf{S}}_k(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2}\mathbf{B}_{k_0}^H\mathbf{H}_k^H| \\ & - \text{Tr}(\tilde{\mathbf{S}}_k) \end{aligned}$$

Note that $\text{rank}(\mathbf{H}_k\mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2}) = \min(N, M-L) = N$. Thus, we can obtain it's reduced SVD as follows

$$\mathbf{H}_k\mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2} = \hat{\mathbf{A}}_k\hat{\mathbf{\Sigma}}_{k_p}\hat{\mathbf{B}}_{k_1}^H \tag{3.9}$$

where $\hat{\mathbf{A}}_k \in \mathbb{C}^{N\times N}$, $\hat{\mathbf{B}}_{k_1} \in \mathbb{C}^{(M-L)\times N}$ and $\hat{\mathbf{\Sigma}}_{k_p} = \text{Diag}(\hat{\sigma}_{k_1}, \cdots, \hat{\sigma}_{k_N})$ with $\{\hat{\sigma}_{k_i}\}_{i=1}^N$ being all the positive singular values of the matrix $\mathbf{H}_k\mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2}$. We next substitute the above SVD in the objective function of (P5) and solve it using KKT conditions (Boyd and Vandenberghe, 2004) and present the optimal solution for (P5) in the following theorem.

Theorem 1 *The optimal solution of (P5) for a given set of dual varibles,* $\gamma, \{\lambda_a\}_{a=1}^A, \{\mu_i\}_{i=1}^M$ *is given as*

$$\tilde{\mathbf{S}}_k^* = \hat{\mathbf{B}}_{k_1}\mathbf{\Omega_k}\hat{\mathbf{B}}_{k_1}^H \tag{3.10}$$

where $\mathbf{\Omega_k} = \textit{Diag}(\omega_{k_1}, \cdots, \omega_{k_N})$ *with* ω_{k_i} *given by*

$$\omega_{k_i} = \left(w_k - \frac{1}{\hat{\sigma}_{k_i}^2}\right)^+ \;\; \forall i = 1, \cdots, N. \tag{3.11}$$

where $(y)^+ \triangleq \textit{max}(0, y)$.

Proof: Please refer to Appendix D.

To summarize from (3.10), the optimal solution of (P4) for a given set of dual variables, $\gamma, \{\lambda_a\}_{a=1}^A, \{\mu_i\}_{i=1}^M$ can be expressed as

$$\mathbf{S}_k^* = (\mathbf{B}_{k0}^H \mathbf{B}_{\gamma\lambda\mu} \mathbf{B}_{k_0})^{-1/2} \hat{\mathbf{B}}_{k_1} \mathbf{\Omega_k} \hat{\mathbf{B}}_{k_1}^H (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\lambda\mu} \mathbf{B}_{k_0})^{-1/2} \tag{3.12}$$

3.1.6 Optimal Solution of (P1)

As mentioned earlier, we next solve for the optimal dual variables $\gamma^*, \{\lambda_a^*\}_{a=1}^A, \{\mu_i^*\}_{i=1}^M$ in (P3-D) using a subgradient-based method such as the ellipsoid method (Boyd and Barratt, 2008; Bland *et al.*, 1981). Substituting these optimal dual variables in (3.12) gives us the optimal solution for the problem (P3). By combining this result with Lemma 1, we design the following theorem

Theorem 2 *The optimal solution of (P1) is given by*

$$\begin{aligned}\mathbf{Q}_k^* = \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\lambda\mu}^* \mathbf{B}_{k_0})^{-1/2} \hat{\mathbf{B}}_{k_1} \mathbf{\Omega_k} \hat{\mathbf{B}}_{k_1}^H \\ \times (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\lambda\mu}^* \mathbf{B}_{k_0})^{-1/2} \mathbf{B}_{k_0}^H\end{aligned} \tag{3.13}$$

$\forall k = 1, \cdots, K$, *where* $\mathbf{B}_{\gamma\lambda\mu}^* = \gamma^* \mathbf{I} + \sum_{a=1}^A \lambda_a^* \mathbf{B}_a + \sum_{i=1}^M \mu_i^* \mathbf{B}_i$ *with* $\gamma^*, \{\lambda_a^*\}_{a=1}^A, \{\mu_i^*\}_{i=1}^M$ *being the optimal dual solutions of (P3).*

3.1.7 Optimal Algorithm (A1) to solve (P1)

We next summarize the optimal algorithm(A1) for solving (P1) as follows.

1. **Initialize** the dual variables $\gamma \geq 0, \lambda_a \geq 0, \forall a = 1, \cdots, A$ and $\mu_i \geq 0, \forall i = 1, \cdots, M$.

2. **Repeat**

- Solve for $\mathbf{S}_k^*, k = 1, \cdots, K$ using the equation (3.12) with the given dual variables above.
- Compute the sub-gradients of the Lagrange Dual function $g(\gamma, \{\lambda_a\}, \{\mu_i\})$ as
  - $P_{sum} - \sum_{k=1}^K \text{Tr}(\mathbf{B}_{k_0} \mathbf{S}_k^* \mathbf{B}_{k_0}^H)$ w.r.t. γ.
  - $\tilde{P}_a - \sum_{k=1}^K \text{Tr}(\mathbf{B}_a \mathbf{B}_{k_0} \mathbf{S}_k^* \mathbf{B}_{k_0}^H)$ w.r.t. $\lambda_a, a = 1, \cdots, A$.
  - $\hat{P}_i - \sum_{k=1}^K \text{Tr}(\mathbf{B}_i \mathbf{B}_{k_0} \mathbf{S}_k \mathbf{B}_{k_0}^H)$ w.r.t. $\mu_i, i = 1, \cdots, M$.
- Update γ, $\{\lambda_a\}_{a=1}^A$ and $\{\mu_i\}_{i=1}^M$ based on the ellipsoid method (Bland *et al.*, 1981; Boyd and Barratt, 2008).

3. **Until** all the dual variables γ, $\{\lambda_a\}_{a=1}^{A}$ and $\{\mu_i\}_{i=1}^{M}$ converge to a prescribed accuracy.

4. **Set** $\mathbf{Q}_k^* = \mathbf{B}_{k_0}\mathbf{S}_k^*\mathbf{B}_{k_0}^H, k = 1, \cdots, K.$

3.1.8 Optimal Precoding matrix structure without rank constraints and it's miscellaneous properties

The following corollary is obtained from Theorem 2 and the fact that $\mathbf{Q}_k = \mathbf{U}_k\mathbf{U}_k^H, \forall k.$

Corollary 1. The optimal BD precoding matrices that maximize the weighted sum-rate in a multi-cell system posed as a MU-MIMO-BC subject to multiple power constraints e.g., SPC in (2.3), PBPC in (2.4) and PAPC in (2.6) are given by

$$\mathbf{U}_k^* = \mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}^*\mathbf{B}_{k_0})^{-1/2}\hat{\mathbf{B}}_{k_1}\mathbf{\Omega_k}^{1/2} \tag{3.14}$$

for $k = 1, \cdots, K$. Here, (3.14) follows from the fact that $(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}^*\mathbf{B}_{k_0})^{-1/2}$ is a hermitian matrix.

The following remarks have been made on similar lines of remarks 3.3 and 3.4 from (Zhang, 2010). We discuss some interesting properties of the optimal BD precoding matrix structure in (3.14) in these remarks.

Remark 3(*Channel Diagonalization*): One of the desirable properties of linear precoding techinques for a point-to-point MIMO channel is that the precoding matrix, when jointly deployed with a unitary decoding matrix at the receiver, must be able to diagonalize the MIMO channel into parallel scalar sub-channels, over which we can apply independent encoding and decoding to simplify the transceiver design. Here, we verify that the optimal BD precoding matrix, $\mathbf{U}_k^*$ in (3.14) satisfies this "*channel diagonalization*" property, as follows:

$$\hat{\mathbf{A}}_k^H\mathbf{H}_k\mathbf{U}_k^* = \hat{\mathbf{A}}_k^H\mathbf{H}_k\mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}^*\mathbf{B}_{k_0})^{-1/2}\hat{\mathbf{B}}_{k_1}\mathbf{\Omega_k}^{1/2} \tag{3.15}$$

$$= \hat{\mathbf{A}}_k^H\hat{\mathbf{A}}_k\hat{\mathbf{\Sigma}}_{k_p}\hat{\mathbf{B}}_{k_1}^H\hat{\mathbf{B}}_{k_1}\mathbf{\Omega_k}^{1/2} \tag{3.16}$$

$$= \hat{\mathbf{\Sigma}}_{k_p}\mathbf{\Omega_k}^{1/2} \tag{3.17}$$

where (3.15) follows from (3.14), (3.16) is due to (3.9) and (3.17) follows from the two facts that $\hat{\mathbf{A}}_k$ is a unitary matrix and $\hat{\mathbf{B}}_{k_1}^H \hat{\mathbf{B}}_{k_1} = \mathbf{I}$. Therefore, when we apply a unitary decoding matrix $\hat{\mathbf{A}}_k^H$ at the k-th MS receiver, the MIMO channel corresponding to the k-th MS with BD precoding is diagonalized into N scalar sub-channels with channel gains given by the main diagonal of the $N \times N$ diagonal matrix $\hat{\mathbf{\Sigma}}_{k_p} \mathbf{\Omega_k}^{1/2}$.

Remark 4*(Comparison with Sum Power Constraint)*: If we have only SPC as the power constraint instead of Joint-SPC-PBPC-PAPC in (P1), then the resulting problem corresponds to the conventional BD precoder design for the MIMO BC with a sum power constraint as studied earlier in (Spencer *et al.*, 2004; Wong *et al.*, 2003; Choi and Murch, 2004; Pan *et al.*, 2004). We can apply the developed solution for (P1) in this case, with the corresponding matrix $\mathbf{B}^*_{\gamma\lambda\mu}$ being modified as $\mathbf{B}^*_{\gamma} = \gamma^* \mathbf{I}$. From (3.13), it follows that the optimal transmit covariance matrix solution for this modified version of (P1) is given by

$$\mathbf{Q}_k^{**} = \frac{1}{\gamma^*} \mathbf{B}_{k_0} \hat{\mathbf{B}}_{k_1} \mathbf{\Omega_k} \hat{\mathbf{B}}_{k_1}^H \mathbf{B}_{k_0}^H, \;\; k = 1, \cdots, K. \tag{3.18}$$

Moreover, from (3.9) with $\mathbf{B}^*_{\gamma} = \gamma^* \mathbf{I}$, it follows that the matrix $\hat{\mathbf{B}}_{k_1}$ is obtained from the SVD:

$$\frac{1}{\sqrt{\gamma^*}} \mathbf{H}_k \mathbf{B}_{k_0} = \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tag{3.19}$$

and is thus independent of γ^* but $\mathbf{\Omega_k}$ in (3.18) will be dependent on γ^*. Accordingly, the optimal precoding matrix in the sum power constraint case is given by

$$\mathbf{U}_k^{**} = \frac{1}{\sqrt{\gamma^*}} \mathbf{B}_{k_0} \hat{\mathbf{B}}_{k_1} \mathbf{\Omega_k}^{1/2}, \;\; k = 1, \cdots, K. \tag{3.20}$$

Comparing $\mathbf{U}_k^{**}$ in (3.20) for the SPC case with $\mathbf{U}_k^{*}$ in (3.14) for the Joint SPC-PBPC-PAPC case, we see that $\mathbf{U}_k^{**}$ consists of *orthogonal columns* (orthogonal beamforming vectors) since $\hat{\mathbf{B}}_{k_1}^H \mathbf{B}_{k_0}^H \mathbf{B}_{k_0} \hat{\mathbf{B}}_{k_1} = \mathbf{I}$, while $\mathbf{U}_k^{*}$ in general consists of *non-orthogonal columns* (non-orthogonal beamforming vectors) if $\mathbf{B}^*_{\gamma\lambda\mu}$ is a non-identity diagonal matrix. This is the actual reason for the BD precoder designs in early works such as (Boccardi and Huang, 2006; Liu *et al.*, 2009; Zhang *et al.*, 2008) based on the orthogonal precoder structure $\mathbf{U}_k^{**}$ to be suboptimal in general for the Joint-SPC-PBPC-PAPC case or only-PBPC case as discussed in (Zhang, 2010).

3.2 Special MU-MISO BC Case

3.2.1 Optimal Structure of the Transmit Covariance Matrices

Here, we investigate the developed solution for (P1) for the special case of MISO BC i.e., $M_B = N = 1$. The auxiliary MU-MIMO BC with Joint-SPC-PBPC-PAPC reduces to an equivalent MU-MISO BC with the corresponding Joint-SPC-PAPC, and the BD precoding constraints reduce to the ZF-BF precoding constraints. When $M_B = 1$, the per-base station power constraints and the per-antenna power constraints are the same. Hence, we can write without loss of generality that $\mathbf{B}_{\gamma\lambda\mu} = \mathbf{B}_{\gamma\mu}$. With $N = 1$, we denote $\mathbf{H}_k = \mathbf{h}_k^H$ where $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ is a column-vector for $k = 1, \cdots, K$. Accordingly, the regular SVD in (3.9) can be re-written as

$$\mathbf{h}_k^H \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu} \mathbf{B}_{k_0})^{-1/2} = \hat{\sigma}_{k_p} \hat{\mathbf{b}}_{k_1}^H \tag{3.21}$$

where $\hat{\sigma}_{k_p} > 0$ and $\hat{\mathbf{b}}_{k_1} \in \mathbb{C}^{(M-L)\times 1}$. From (3.11), (3.13) and (3.21), it follows that the optimal downlink transmit covariance matrix for the k-th MS, $\mathbf{Q}_k^*$, in the case of $M_B = N = 1$ is given by

$$\begin{aligned} \mathbf{Q}_k^* = \omega_k \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu}^* \mathbf{B}_{k_0})^{-1/2} \hat{\mathbf{b}}_{k_1} \hat{\mathbf{b}}_{k_1}^H \\ \times (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu}^* \mathbf{B}_{k_0})^{-1/2} \mathbf{B}_{k_0}^H \end{aligned} \tag{3.22}$$

where $\omega_k = (w_k - 1/\hat{\sigma}_{k_p}^2)^+$. Now substitute the SVD given in (3.21) into the above equation (3.22) and obtain the following optimal structure for $\mathbf{Q}_k^*$.

$$\begin{aligned} \mathbf{Q}_k^* = \frac{\omega_k}{\hat{\sigma}_{k_p}^2} \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu}^* \mathbf{B}_{k_0})^{-1} \mathbf{B}_{k_0}^H \mathbf{h}_k \\ \times \mathbf{h}_k^H \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu}^* \mathbf{B}_{k_0})^{-1} \mathbf{B}_{k_0}^H \end{aligned} \tag{3.23}$$

3.2.2 Optimal Precoding Matrix Structure

From (3.23), we can easily observe that the rank($\mathbf{Q}_k^*$) ≤ 1 for the MISO BC case. Thus, we can express $\mathbf{Q}_k^*$ in this scenario as $\mathbf{Q}_k^* = \mathbf{u}_k^* (\mathbf{u}_k^*)^H$ where $\mathbf{u}_k^* \in \mathbb{C}^{M\times 1}$ is the optimal beamforming vector (precoding matrix in MIMO BC case reduced to a beamforming vector in MISO BC case). The expression for $\mathbf{u}_k^*$ is given as

$$\mathbf{u}_k^* = \frac{\omega_k^{1/2}}{\hat{\sigma}_{k_p}} \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu}^* \mathbf{B}_{k_0})^{-1} \mathbf{B}_{k_0}^H \mathbf{h}_k \tag{3.24}$$

A very important observation is that (3.24) holds regardless of the number of transmit antennas on each BS M_B. Furthermore, the optimal beamforming vector for the k-th MS in the conventional sum power constraint case (with $\mathbf{B}_{\gamma\mu}^* = \mathbf{B}_\gamma^* = \gamma^* \mathbf{I}$) is obtained from (3.24) as follows

$$\mathbf{u}_k^{**} = \frac{\omega_k^{1/2}}{\hat{\sigma}_{k_p} \gamma^*} \mathbf{B}_{k_0} \mathbf{B}_{k_0}^H \mathbf{h}_k \tag{3.25}$$

(3.25) follows from the fact that $\mathbf{B}_{k_0}^H \mathbf{B}_{k_0} = \mathbf{I}$.

3.2.3 Interesting Properties of the above Optimal Precoding Matrix Structure

Next, we discuss an interesting observation on the optimal ZF-BF precoding design in (3.24) as compared with a prior result published in (Peel *et al.*, 2005).

Remark 5: We first denote $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_K] \in \mathbb{C}^{M \times K}$ as the total precoding matrix for a MU-MISO BC with M transmitting antennas and K users/MSs each with a single-antenna. Then, for the sum power constraint case with $\mathbf{u}_k = \mathbf{u}_k^{**}$ shown in (3.25), the corresponding total optimal precoding matrix $\mathbf{U}^{**}$ becomes the conventional ZF-BF design for the MISO BC based on the channel pseudo inverse (Peel *et al.*, 2005), i.e., we can express $\mathbf{U}^{**}$ as $\mathbf{U}^{**} = \mathbf{H}^\dagger \hat{\mathbf{\Omega}}$ where $\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_K]^H \in \mathbb{C}^{K \times M}$ is the total downlink channel matrix and $\hat{\mathbf{\Omega}} = \text{Diag}(\hat{\omega}_1, \cdots, \hat{\omega}_K)$, where $\hat{\omega}_k = \omega_k^{1/2} \hat{\sigma}_{k_p}, k = 1, \cdots, K$.

However, it has been observed that this ZF-BF design based on the Moore-Penrose inverse of the channel matrix is in general suboptimal for the MISO BC with the per-antenna power constraints in (Karakayali *et al.*, 2007), per-BS power constraints in (Zhang, 2010). Hence, we conclude that the conventional pseudo-inverse-based method is suboptimal for the MISO BC under Joint-SPC-PAPC constraints also where the total optimal precoding matrix $\mathbf{U}^*$ is obtained by letting $\mathbf{u}_k = \mathbf{u}_k^*$ as given in (3.24).

We should note that $\mathbf{u}_k^*$ becomes collinear with $\mathbf{u}_k^{**}$ regardless of γ^*, $\{\mu_i^*\}_{i=1}^M$ when $M = K$ because of the fact that $(\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\mu}^* \mathbf{B}_{k_0})^{-1}$ in (3.24) becomes a scalar quantity as $(M - L) = M - N(K - 1) = M - K - 1 = 1$. In this particular case, $\mathbf{B}_{k_0}$ becomes a vector, $\mathbf{b}_{k_0} \in \mathbb{C}^{M \times 1}$, and $\mathbf{u}_k^*$ & $\mathbf{u}_k^{**}$ can both be expressed in the form $p_k \mathbf{b}_{k_0}$, with $p_k \geq 0$. Furthermore, we can show that this result holds regardless of the value M_B takes, provided that $N = 1$ and $M = M_B A = K$.

CHAPTER 4

SUB OPTIMAL PRECODING MATRIX SOLUTION WITHOUT RANK CONSTRAINTS

4.1 Procedure to obtain Sub Optimal Precoding Matrix Structure

In this section, we propose a suboptimal solution for (P1) which requires lesser implementation complexity than the optimal solution proposed in the previous chapter. A summary of steps in this procedure is as follows.

1. First, we consider the optimal structure of the transmit covariance matrix for the conventional BD precoding design problem under a single sum power constraint which has already been derived in (3.18). This structure is surely suboptimal for the Joint-SPC-PBPC-PAPC case. We can see from (3.18) that this matrix is unitarily diagonalizable into a diagonal matrix called the *power allocation* matrix.

2. Next, we substitute the above suboptimal transmit covariance structure in the objective function and the Joint-SPC-PBPC-PAPC power constraints and solve the resulting convex optimization problem for the best possible power allocation. Note that we have already gotten rid of the BD precoding constraints above by considering the solution of the conventional BD precoding design problem under SPC.

3. Finally, we discuss the algortihm(A2) used to obtain these suboptimal transmit covariance matrices and show why algorithm(A2) has lesser implementation complexity than the optimal algorithm(A1).

4.2 Formulation of Convex Optimisation Problem (P6) to be solved for obtaining Sub Optimal Solution

First, we define the projected channel of $\mathbf{H}_k$ associated with the projection matrix $\mathbf{P}_k$ as $\mathbf{H}_k^{\perp} = \mathbf{H}_k\mathbf{P}_k = \mathbf{H}_k\mathbf{B}_{k_0}\mathbf{B}_{k_0}^H, k = 1, \cdots, K$, where $\mathbf{H}_k^{\perp} \in \mathbb{C}^{N\times M}$ and rank($\mathbf{H}_k^{\perp}$) = min($N, M - L$) = N. Next, we define the reduced SVD of $\mathbf{H}_k^{\perp}$ as given below.

$$\mathbf{H}_k^{\perp} = \mathbf{A}_k^{\perp}\mathbf{\Sigma}_{k_p}^{\perp}(\mathbf{B}_{k_1}^{\perp})^H \tag{4.1}$$

where $\mathbf{A}_k^{\perp} \in \mathbb{C}^{N\times N}, \mathbf{B}_{k_1}^{\perp} \in \mathbb{C}^{M\times N}$ and $\mathbf{\Sigma}_{k_p} = \text{Diag}(\sigma_{k_1}^{\perp}, \cdots, \sigma_{k_N}^{\perp})$. Now taking a careful look at the SVD of $\mathbf{H}_k\mathbf{B}_{k_0}$ in (3.19) from Remark 4, we can easily deduce that

$$\mathbf{H}_k\mathbf{B}_{k_0}\mathbf{B}_{k_0}^H = \mathbf{H}_k^{\perp} = \hat{\mathbf{A}}_k(\sqrt{\gamma^*}\hat{\mathbf{\Sigma}}_{k_p})(\mathbf{B}_{k_0}\hat{\mathbf{B}}_{k_1})^H \tag{4.2}$$

where γ^* and $\mathbf{\Omega}_k$ from (3.18) gives the optimal power allocation for the k-th MS in the conventional BD precoder design under a single sum power constraint. Also, $(\mathbf{B}_{k_0}\hat{\mathbf{B}}_{k_1})^H(\mathbf{B}_{k_0}\hat{\mathbf{B}}_{k_1}) = \mathbf{I}$. Therefore, we can conclude from (4.1) and (4.2) that

$$\mathbf{B}_{k_1}^{\perp} = \mathbf{B}_{k_0}\hat{\mathbf{B}}_{k_1} \tag{4.3}$$

From (4.3) and (3.18), we can finally present a structure for the proposed suboptimal solution of (P1) as follows.

$$\bar{\mathbf{Q}}_k = \mathbf{B}_{k_1}^{\perp}\bar{\mathbf{\Omega}}_k(\mathbf{B}_{k_1}^{\perp})^H, \;\; k = 1, \cdots, K. \tag{4.4}$$

where $\bar{\mathbf{\Omega}}_k = \text{Diag}(\bar{\omega}_{k_1}, \cdots, \bar{\omega}_{k_N})$ denotes the power allocation for the k-th MS. With $\{\bar{\mathbf{Q}}_k\}_{k=1}^K$ given in (4.4), it can be shown that the ZF constraints in (P1) are satisfied and thus can be removed. Next, we substitute the above $\{\bar{\mathbf{Q}}_k\}$'s in the objective function and Joint-SPC-PBPC-PAPC of (P1) to solve for the optimal power allocation matrix $\bar{\mathbf{\Omega}}_k$.

Furthermore, in the objective function of (P1), the following equalities hold:

$$
\begin{aligned}
&\log|\mathbf{I}+\mathbf{H}_k\bar{\mathbf{Q}}_k\mathbf{H}_k^H| \\
=&\log|\mathbf{I}+\mathbf{H}_k(\mathbf{B}_{k_0}\mathbf{B}_{k_0}^H+\mathbf{B}_{k_1}\mathbf{B}_{k_1}^H)\bar{\mathbf{Q}}_k \\
&\times(\mathbf{H}_k(\mathbf{B}_{k_0}\mathbf{B}_{k_0}^H+\mathbf{B}_{k_1}\mathbf{B}_{k_1}^H))^H| \qquad (4.5)\\
=&\log|\mathbf{I}+(\mathbf{H}_k^\perp+\mathbf{H}_k\mathbf{B}_{k_1}\mathbf{B}_{k_1}^H)\bar{\mathbf{Q}}_k \\
&\times(\mathbf{H}_k^\perp+\mathbf{H}_k\mathbf{B}_{k_1}\mathbf{B}_{k_1}^H)^H| \qquad (4.6)\\
=&\log|\mathbf{I}+\mathbf{H}_k^\perp\bar{\mathbf{Q}}_k(\mathbf{H}_k^\perp)^H| \qquad (4.7)\\
=&\log|\mathbf{A}_k^\perp(\mathbf{A}_k^\perp)^H+\mathbf{A}_k^\perp\mathbf{\Sigma}_{k_p}^\perp\bar{\mathbf{\Omega}}_k\mathbf{\Sigma}_{k_p}^\perp(\mathbf{A}_k^\perp)^H| \qquad (4.8)\\
=&\log|\mathbf{I}+(\mathbf{\Sigma}_{k_p}^\perp)^2\bar{\mathbf{\Omega}}_k| \qquad (4.9)
\end{aligned}
$$

where (4.5) follows from the fact that $\mathbf{B}_{k_0}\mathbf{B}_{k_0}^H+\mathbf{B}_{k_1}\mathbf{B}_{k_1}^H=\mathbf{I}$; (4.6) follows from the definition of $\mathbf{H}_k^\perp$; (4.7) is due to the fact that $\mathbf{B}_{k_1}^H\mathbf{B}_{k_0}=\mathbf{0}$; (4.8) is because $\mathbf{A}_k^\perp(\mathbf{A}_k^\perp)^H=(\mathbf{B}_{k_1}^\perp)^H\mathbf{B}_{k_1}^\perp=\mathbf{I}$ and (4.9) follows from the fact that $\log|\mathbf{I}+\mathbf{XY}|=\log|\mathbf{I}+\mathbf{YX}|$. From (4.9), we can observe that the MIMO channel for the k-th MS has been diagonalized into N scalar sub-channels with channel gains given by $\bar{\omega}_{k_i}, i=1,\cdots,N$. Accordingly, (P1) is reduced to the following problem (P6).

$$
\text{(P6)}:\quad
\begin{aligned}
\max_{\{\bar{\omega}_{k_i}\}}\quad & \sum_{k=1}^{K}w_k\sum_{i=1}^{N}\log(1+(\sigma_{k_i}^\perp)^2\bar{\omega}_{k_i}) \\
\text{s.t.}\quad & \sum_{k=1}^{K}\sum_{i=1}^{N}\|\mathbf{b}_{k_1}^\perp[:,i]\|^2\bar{\omega}_{k_i}\le P_{sum} \\
& \sum_{k=1}^{K}\sum_{i=1}^{N}\|\mathbf{b}_{k_1}^\perp[a,i]\|^2\bar{\omega}_{k_i}\le\tilde{P}_a,\ \ \forall a=1,\cdots,A \\
& \sum_{k=1}^{K}\sum_{i=1}^{N}|\mathbf{b}_{k_1}^\perp(j,i)|^2\bar{\omega}_{k_i}\le\hat{P}_j,\ \ \forall j=1,\cdots,M \\
& \bar{\omega}_{k_i}\ge 0,\ \ \forall k,i
\end{aligned}
$$

where $\{\bar{\omega}_{k_i}\}$ denotes the set of $\bar{\omega}_{k_i}$'s, for $k=1,\cdots,K$ and $n=1,\cdots,N$. Note that $\mathbf{b}_{k_1}^\perp[:,i]$ denotes the i-th column, while $\mathbf{b}_{k_1}^\perp(j,i)$ represents the (j,i)-th element and $\mathbf{b}_{k_1}^\perp[a,i]$ denotes the vector consisting of the elements from the i-th column and the $((a-1)M_B+1)$-th to (aM_B)-th rows in the matrix $\mathbf{B}_{k_1}^\perp$, $a=1,\cdots,A$ and $i=1,\cdots,N$.

4.3 Procedure to obtain Optimal Solution of (P6)

We can verify that (P6) is a convex optimisation problem since the objective function is concave over $\{\bar{\omega}_{k_i}\}$, the primal optimization variables and all the three power constraints generate an affine set (which is also convex) over $\bar{\omega}_{k_i}$'s.

Thus, similar to (P3), we can apply Lagrange-Duality method to solve (P6) by introducing a set of dual variables, γ, $\{\lambda_a\}_{a=1}^{A}$, $\{\mu_j\}_{j=1}^{M}$ associated with the sum power constraint, per-BS power constraints and per-antenna power constraints respectively in (P6).

The Lagrangian function for the above problem (P6) is written as

$$\begin{aligned} L(\{\bar{\omega}_{k_i}\}, \gamma, \{\lambda_a\}, \{\mu_j\}) &= \sum_{k=1}^{K} w_k \sum_{i=1}^{N} \log(1 + (\sigma_{k_i}^{\perp})^2 \bar{\omega}_{k_i}) \\ &+ \gamma(P_{sum} - \sum_{k=1}^{K}\sum_{i=1}^{N} \|\mathbf{b}_{k_1}^{\perp}[:, i]\|^2 \bar{\omega}_{k_i}) \\ &+ \sum_{a=1}^{A} \lambda_a(\tilde{P}_a - \sum_{k=1}^{K}\sum_{i=1}^{N} \|\mathbf{b}_{k_1}^{\perp}[a, i]\|^2 \bar{\omega}_{k_i}) \\ &+ \sum_{j=1}^{M} \mu_j(\hat{P}_j - \sum_{k=1}^{K}\sum_{i=1}^{N} |\mathbf{b}_{k_1}^{\perp}(j, i)|^2 \bar{\omega}_{k_i}) \end{aligned} \tag{4.10}$$

where $\{\bar{\omega}_{k_i}\}$, $\{\lambda_a\}$ and $\{\mu_j\}$ denote the set of $\bar{\omega}_{k_i}$'s, λ_a's and μ_j's respectively.

The Lagrange Dual function for (P6) is then defined as

$$g(\gamma, \{\lambda_a\}, \{\mu_j\}) = \max_{\bar{\omega}_{k_i} \geq 0, \forall k, i} L(\{\bar{\omega}_{k_i}\}, \gamma, \{\lambda_a\}, \{\mu_j\}) \tag{4.11}$$

Again, the dual problem of (P6) is defined very similar to (P3-D) as follows.

$$\text{(P6-D)}: \quad \begin{aligned} \min_{\gamma, \{\lambda_a\}, \{\mu_j\}} \quad & g(\gamma, \{\lambda_a\}, \{\mu_j\}) \\ \text{s.t.} \quad & \gamma \geq 0 \\ & \lambda_a \geq 0, \;\; \forall a = 1, \cdots, A \\ & \mu_j \geq 0, \;\; \forall j = 1, \cdots, M \end{aligned}$$

It is already known that (P6) is a convex optimization problem and satisfies the

Slater's conditions (Boyd and Vandenberghe, 2004), so the duality gap between the optimal objective values of (P6) and (P6-D) is zero. Therefore, we can equivalently solve (P6-D) instead of solving (P6). Towards this end, we first need to solve for optimal $\{\bar{\omega}_{k_i}^*\}$ set that maximizes the Lagrangian function in (4.10) in terms of the given dual variables.

Moreover, the dual problems are always convex and hence (P6-D) is also convex. Finally, we can obtain optimal dual variables $\gamma^*, \{\lambda_a^*\}, \{\mu_j^*\}$ iteratively using the subgradient method e.g., the ellipsoid method (Bland *et al.*, 1981; Boyd and Barratt, 2008), given the fact that the subgradients (partial derivatives) of function $g(\gamma, \{\lambda_a\}, \{\mu_j\})$ at a set of already fixed γ, λ_a's and μ_j's are $P_{sum} - \sum_{k=1}^{K}\sum_{i=1}^{N} \|\mathbf{b}_{k_1}^{\perp}[:,i]\|^2\bar{\omega}_{k_i}^*$ for γ, $\tilde{P}_a - \sum_{k=1}^{K}\sum_{i=1}^{N} \|\mathbf{b}_{k_1}^{\perp}[a,i]\|^2\bar{\omega}_{k_i}^*$ for $\lambda_a, a = 1, \cdots, A$ and $\hat{P}_j - \sum_{k=1}^{K}\sum_{i=1}^{N} |\mathbf{b}_{k_1}^{\perp}(j,i)|^2\bar{\omega}_{k_i}^*$ for $\mu_j, j = 1, \cdots, M$ in that order.

Now we can substitute these optimal dual variables in the optimal structure obtained for $\bar{\omega}_{k_i}^*$ above and later obtain a suboptimal BD precoding matrix/transmit covariance matrix structure. This is validated by the fact that *strong duality holds* for the convex optimization problem (P6).

4.4 Maximisation Problem (P7) and it's Optimal Solution

Now, we continue forward to solve for the optimal $\{\bar{\omega}_{k_i}^*\}$ with a given set of dual variables. From (4.10), it is evident that the maximization problem in (4.10) can be separated into NK independent maximization subproblems each involving only one $\bar{\omega}_{k_i}$ as the primal variable. After discarding the irrelevant terms, we can express the corresponding subproblem in (P7), for each pair of $\{k, i\}$ as follows.

$$\text{(P7)}: \quad \begin{aligned} &\max_{\bar{\omega}_{k_i}\geq 0} w_k \log(1 + (\sigma_{k_i}^{\perp})^2\bar{\omega}_{k_i}) - \bar{\omega}_{k_i}\Big(\gamma\|\mathbf{b}_{k_1}^{\perp}[:,i]\|^2 \\ &\qquad + \sum_{a=1}^{A}\lambda_a\|\mathbf{b}_{k_1}^{\perp}[a,i]\|^2 + \sum_{j=1}^{M}\mu_j|\mathbf{b}_{k_1}^{\perp}(j,i)|^2\Big) \end{aligned}$$

Let us denote the scalar quantity multiplying $\bar{\omega}_{k_i}$ in the second term of the objective function above by f_{k_i}. This will be helpful for an easy notation later on. So,

$$f_{k_i} = \gamma \|\mathbf{b}_{k_1}^{\perp}[:,i]\|^2 + \sum_{a=1}^{A} \lambda_a \|\mathbf{b}_{k_1}^{\perp}[a,i]\|^2 + \sum_{j=1}^{M} \mu_j |\mathbf{b}_{k_1}^{\perp}(j,i)|^2 \tag{4.12}$$

The following theorem presents the optimal solution for (P7).

Theorem 3 *The optimal solution of (P7) for a given set of dual variables* $\gamma, \{\lambda_a\}_{a=1}^{A}, \{\mu_j\}_{j=1}^{M}$ *is given as*

$$\bar{\omega}_{k_i}^* = \left(\frac{w_k}{f_{k_i}} - \frac{1}{(\sigma_{k_i}^{\perp})^2}\right)^+ \tag{4.13}$$

where $(x)^+ \triangleq max(0,x)$.

Proof: Please refer to appendix E.

4.5 Sub Optimal Algorithm (A2) to solve (P1)

Similar to (A1), the following algorithm (A2) can be used to obtain the proposed sub-optimal solution for (P1).

1. **Initialize** the dual variables $\gamma \geq 0, \lambda_a \geq 0, \forall a = 1, \cdots, A$ and $\mu_j \geq 0, \forall j = 1, \cdots, M$.

2. **Compute** the SVDs: $\mathbf{H}_k \mathbf{B}_{k_0} \mathbf{B}_{k_0}^H = \mathbf{A}_k^{\perp} \mathbf{\Sigma}_{k_p}^{\perp} (\mathbf{B}_{k_1}^{\perp})^H, k = 1, \cdots, K$.

3. **Repeat**

- Solve for $\{\bar{\omega}_{k_i}^*\}$ using the equation (4.13) with the given dual variables above.
- Compute the sub-gradients of the Lagrange Dual function in (4.11), $g(\gamma, \{\lambda_a\}, \{\mu_j\})$ as
  - $P_{sum} - \sum_{k=1}^{K} \sum_{i=1}^{N} ||\mathbf{b}_{k_1}^{\perp}[:,i]||^2 \bar{\omega}_{k_i}^*$ for γ.
  - $\tilde{P}_a - \sum_{k=1}^{K} \sum_{i=1}^{N} ||\mathbf{b}_{k_1}^{\perp}[a,i]||^2 \bar{\omega}_{k_i}^*$ for $\lambda_a, a = 1, \cdots, A$.
  - $\hat{P}_j - \sum_{k=1}^{K} \sum_{i=1}^{N} |\mathbf{b}_{k_1}^{\perp}(j,i)|^2 \bar{\omega}_{k_i}^*$ for $\mu_j, j = 1, \cdots, M$.
- Update γ, $\{\lambda_a\}_{a=1}^{A}$ and $\{\mu_j\}_{j=1}^{M}$ based on the ellipsoid method.
- Update γ, $\{\lambda_a\}_{a=1}^{A}$ and $\{\mu_j\}_{j=1}^{M}$ based on the ellipsoid method (Bland *et al.*, 1981; Boyd and Barratt, 2008).

4. **Until** all the dual variables γ, $\{\lambda_a\}_{a=1}^{A}$ and $\{\mu_j\}_{j=1}^{M}$ converge to a prescribed accuracy.

5. **Set** $\bar{\mathbf{Q}}_k = \mathbf{B}_{k_1}^{\perp}\bar{\mathbf{\Omega}}_k^{*}(\mathbf{B}_{k_1}^{\perp})^H$ for $k = 1, \cdots, K$, where $\bar{\mathbf{\Omega}}_k^{*} = \text{Diag}(\bar{\omega}_{k_1}^{*}, \cdots, \bar{\omega}_{k_N}^{*})$.

4.6 Reason behind the lower implementation complexity of (A2) compared to (A1)

The following remark is inspired from Remark 4.1 in (Zhang, 2010). It discusses the reason behind the lower implementation complexity of the suboptimal algorithm (A2) when compared to that of the optimal algorithm (A1) and also a special case in which (A2) provides optimal solution for (P1).

Remark 6: For each loop in the “Repeat" section in (A2), only the power allocation computation in (4.13) is implemented, instead of whole matrix computation given in (3.12). This is the prime reason for (A2) having lower implementation complexity than (A1). Due to the suboptimal structure of the downlink transmit covariance matrix in (4.4) for (A2) as compared to the optimal one in (3.13) for (A1), (A2) in general leads to a suboptimal solution and a lower weighted sum-rate for (P1) than (A1) at reasonably high SNRs (power constraints). However, in the special case of $N = 1$ and $M = K$ where the suboptimal transmit covariance matrix structure in (4.4) is known to be collinear with the optimal structure (discussed already in Remark 5). In this special case, (A2) can be used as an alternative algorithm to (A1) to obtain the optimal solution for (P1). We can conclude saying that the algorithm (A2) will be indeed suboptimal for (P1) at reasonably high SNRs if $M > K$.

CHAPTER 5

RANK-CONSTRAINED MULTI-USER MIMO CAPACITY UNDER JOINT-SPC-PBPC-PAPC

In this chapter, we propose a Projected Factored Gradient Descent (PFGD) algorithm to find the optimal transmission scheme that maximizes the weighted sum-rate of a multi-cell system modeled as an auxiliary MU-MIMO BC under BD precoding, joint SPC-PBPC-PAPC and additional rank constraints on transmit covariance matrices corresponding to K MSs as shown in (P2). In this paper, we consider a uniform rank constraint on all the precoding matrices i.e., we assume $R_k = \text{rank}(\mathbf{U}_k) = R(\leq N), \forall k = 1, \cdots, K$.

In the PFGD algorithm, the key idea is to project the updated precoding matrices corresponding to all the K MSs after each gradient descent step onto the constraint set consisting of Joint SPC-PBPC-PAPC (power constraints) and zero forcing constraints (BD precoding). We recall that $\mathbf{Q}_k \succeq \mathbf{0}$ iff $\exists$ a matrix $\mathbf{U}_k$ such that $\mathbf{Q}_k$ can be factored as $\mathbf{Q}_k = \mathbf{U}_k\mathbf{U}_k^H$. Note that $\mathbf{U}_k$ is also a square root of the matrix $\mathbf{Q}_k$. As shown in (P2), the whole formulation has been done in terms of $\{\mathbf{U}_k\}_{k=1}^K$.

5.1 Important reasons for choosing PFGD Approach

Inspired from (Chaluvadi *et al.*, 2018; Park *et al.*, 2016), we provide some important reasons for choosing PFGD approach as follows.

1. We actually desire to obtain the optimal precoder matrices set $\{\mathbf{U}_k\}_{k=1}^K$ that optimizes the transmission scheme in block diagonalized fully cooperative multi-cell system with multiple power constraints directly without finding the optimal downlink transmit covariance matrices $\{\mathbf{Q}_k\}_{k=1}^K$. Hence, we have formulated the non-convex optimization problem (P2) in Chapter 2 towards this end.

2. The constraint $\mathbf{Q}_k \succeq \mathbf{0}$ in (P1) is easily enforced by the factorization $\mathbf{Q}_k = \mathbf{U}_k\mathbf{U}_k^H, \forall k = 1, \cdots, K$.

3. The rank constraints rank$(\mathbf{Q}_k) = R$, for $k = 1, \cdots, K$ can be enforced simply by choosing the size of the precoding matrices $\{\mathbf{U}_k\}_{k=1}^{K}$ to be $M \times R$. It turns out that such rank constraints are very difficult to be enforced in an iterative algorithm while directly determining $\{\mathbf{Q}_k\}$.

4. Recently in (Chaluvadi *et al.*, 2018; Park *et al.*, 2016), it has been shown that PFGD algorithm can be implemented with a very low complexity when compared with the standard CVX package (Grant and Boyd, 2014).

5. Also, the optimal (A1) and the suboptimal (A2) algorithms presented in Chapter 3 for full-rank case have more SVD implementations than the PFGD algorithm at different stages. SVDs generally increase the complexity of algorithm by many folds. Hence, we can use PFGD as an alternative optimal rate-achieving algorithm in the full rank case too instead of (A1).

5.2 Formulation of two Non Convex Optimisation problems (P8) & (P9) equivalent to (P2)

We rewrite the power constraints in (P2) in two different ways and formulate two more problems equivalent to (P2). This is for the ease of notation later on in the chapter to help us solve the problem.

$$
\text{(P8)} : \quad
\begin{aligned}
\max_{\{\mathbf{U}_k\} \in \mathbb{C}^{M \times R}} \quad & \sum_{k=1}^{K} w_k \log |\mathbf{I} + (\mathbf{H}_k \mathbf{U}_k)(\mathbf{H}_k \mathbf{U}_k)^H| \\
\text{s.t.} \quad & \mathbf{H}_j \mathbf{U}_k = 0, \;\; \forall j \neq k \\
& \sum_{k=1}^{K} \sum_{l=1}^{M} (\mathbf{U}_k \mathbf{U}_k^H)_{ll} \leq P_{sum} \\
& \sum_{k=1}^{K} \sum_{l \in I(a)} (\mathbf{U}_k \mathbf{U}_k^H)_{ll} \leq \tilde{P}_a, \;\; a = 1, \cdots, A \\
& \sum_{k=1}^{K} (\mathbf{U}_k \mathbf{U}_k^H)_{ii} \leq \hat{P}_i, \;\; i = 1, \cdots, M
\end{aligned}
$$

where $I(a) = \{(a-1)M_B + 1, \cdots, aM_B\}$ and $(\mathbf{A})_{ll}$ denotes the (l, l)-th element in the matrix $\mathbf{A}$.

$$
\text{(P9)}: \quad \begin{aligned}
\max_{\{\mathbf{U}_k\}\in\mathbb{C}^{M\times R}} \quad & \sum_{k=1}^{K} w_k \log|\mathbf{I} + (\mathbf{H}_k\mathbf{U}_k)(\mathbf{H}_k\mathbf{U}_k)^H| \\
\text{s.t.} \quad & \mathbf{H}_j\mathbf{U}_k = 0, \ \forall j \neq k \\
& \sum_{k=1}^{K}\sum_{l=1}^{M} \|\mathbf{u}_{k_l}\|^2 \leq P_{sum} \\
& \sum_{k=1}^{K}\sum_{l\in I(a)} \|\mathbf{u}_{k_l}\|^2 \leq \tilde{P}_a, \ \ a = 1, \cdots, A \\
& \sum_{k=1}^{K} \|\mathbf{u}_{k_i}\|^2 \leq \hat{P}_i, \ \ i = 1, \cdots, M
\end{aligned}
$$

Another way of representing the power constraints is in terms of the rows of the precoding matrices $\{\mathbf{U}_k\}_{k=1}^{K}$ as shown above. Here, $\mathbf{u}_{k_l}$ denotes the l-th row of the precoding matrix $\mathbf{U}_k$. We again note that (P2) ≡ (P8) ≡ (P9).

5.3 Individual Constraint Sets - Notations

Let us denote the set consisting of BD precoding constraints in (P2) by $\mathcal{C}^{BD}$ which is given as

$$\mathcal{C}^{BD} = \left\{\{\mathbf{U}_k\}_{k=1}^{K} \in \mathbb{C}^{M\times R} : \mathbf{G}_k\mathbf{U}_k = \mathbf{0}, \forall k\right\}$$

where $\mathbf{G}_k = [\mathbf{H}_1^T, \cdots, \mathbf{H}_{k-1}^T, \mathbf{H}_{k+1}^T, \cdots, \mathbf{H}_K^T]^T, \forall k = 1, \cdots, K$.

The Joint-SPC-PBPC-PAPC power constraints set is denoted by $\mathcal{C}^{J-SBA}$ and it can be expressed as

$$
\begin{aligned}
\mathcal{C}^{J-SBA} = \Big\{ & \{\mathbf{U}_k\}_{k=1}^{K} \in \mathbb{C}^{M\times R} : \\
& \sum_{k=1}^{K}\sum_{l=1}^{M} (\mathbf{U}_k\mathbf{U}_k^H)_{ll} \leq P_{sum} \\
& \sum_{k=1}^{K}\sum_{l\in I(a)} (\mathbf{U}_k\mathbf{U}_k^H)_{ll} \leq \tilde{P}_a, \ \ a = 1, \cdots, A \\
& \sum_{k=1}^{K} (\mathbf{U}_k\mathbf{U}_k^H)_{ii} \leq \hat{P}_i, \ \ i = 1, \cdots, M \Big\}
\end{aligned}
$$

We will also need a separate notation for the Joint-SPC-PBPC constraints set later on in the projection step of the PFGD algorithm. Hence, we define Joint-SPC-PBPC power constraints set, $\mathcal{C}^{J-SB}$ as

$$\begin{aligned}\mathcal{C}^{J-SB} = \Big\{ & \{\mathbf{U}_k\}_{k=1}^{K} \in \mathbb{C}^{M\times R} : \\ & \sum_{k=1}^{K}\sum_{l=1}^{M}(\mathbf{U}_k\mathbf{U}_k^H)_{ll} \leq P_{sum} \\ & \sum_{k=1}^{K}\sum_{l\in I(a)}(\mathbf{U}_k\mathbf{U}_k^H)_{ll} \leq \tilde{P}_a,\ \ a = 1,\cdots,A\Big\}\end{aligned}$$

Now, the total constraint set is denoted as $\mathcal{C}^{Tot} = \mathcal{C}^{BD} \cap \mathcal{C}^{J-SBA}$. As mentioned earlier in the introduction, the ability to incorporate the rank constraint easily is useful when (1) Channel is sparse/Channel matrix has a low rank (2) Number of spatially multiplexed data streams is limited by the number of receive antennas on each MS.

5.4 PFGD Algorithm (A3) to solve (P2)

Let $f(\{\mathbf{Q}_k\}) = \sum_{k=1}^{K} w_k \log|\mathbf{I} + \mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H|$. We now present the PFGD algorithm (A3) below.

1. **Initialization**: Initialize the precoding matrices as $\{\mathbf{U}_{k_0}\}$. The initialization process will be discussed right after the algorithm.

2. **Projected Gradient Descent**: Compute the gradients of $f(\{\mathbf{U}_{k_p}\mathbf{U}_{k_p}^H\})$ w.r.t. the matrices $\{\mathbf{U}_{k_p}\}_{k=1}^{K}$, where "$p$" stands for the p-th iteration in the Gradient Descent method as follows.

$$\begin{aligned}&\nabla_{\mathbf{U}_{k_p}} f(\{\mathbf{U}_{k_p}\mathbf{U}_{k_p}^H\}) \\ =& 2w_k\mathbf{H}_k^H(\mathbf{I} + \mathbf{H}_k\mathbf{U}_{k_p}(\mathbf{H}_k\mathbf{U}_{k_p})^H)^{-1}\mathbf{H}_k\mathbf{U}_{k_p}\end{aligned} \tag{5.1}$$

for $k = 1,\cdots,K$. Then, the precoding matrices set for the next iteration, "$p+1$" is given by

$$\{\mathbf{U}_{k_{p+1}}\} = \Pi_{Tot}(\{\mathbf{U}_{k_p} + \eta\nabla_{\mathbf{U}_{k_p}} f(\{\mathbf{U}_{k_p}\mathbf{U}_{k_p}^H\})\}) \tag{5.2}$$

where $\Pi_{Tot}(\{\mathbf{V}_k\})$ in (5.2) is the projection of $\{\mathbf{V}_k\}_{k=1}^{K}$ onto the total constraints set

$\mathcal{C}^{Tot}$, and η is a randomly chosen fixed step size parameter for all MSs and iterations.

3. **Stopping Criterion**: Stop the algorithm when

$$\|\mathbf{U}_{k_{p+1}} - \mathbf{U}_{k_p}\|_F \le \epsilon, \forall k = 1, \cdots, K. \tag{5.3}$$

where $\|.\|_F$ in (5.3) denotes the frobenius norm of a matrix. Typically, we choose ϵ in the order of 10^{-3}.

4. **Optimal Solution**: The optimal downlink transmit covariance matrices are given by

$$\mathbf{Q}_k^* = \mathbf{U}_{k_{p+1}} \mathbf{U}_{k_{p+1}}^H, \ \ \forall k = 1, \cdots, K. \tag{5.4}$$

and the corresponding maximum weighted sum-rate is given by

$$C = f(\{\mathbf{Q}_k^*\}) = \sum_{k=1}^{K} w_k \log |\mathbf{I} + \mathbf{H}_k \mathbf{Q}_k^* \mathbf{H}_k^H| \tag{5.5}$$

Note that (5.1) follows from the fact in (Petersen *et al.*, 2008) that $\nabla_{\mathbf{U}} h(\mathbf{Q})|_{\mathbf{Q}=\mathbf{U}\mathbf{U}^H} = 2\,\nabla_{\mathbf{Q}} h(\mathbf{Q})\mathbf{U}|_{\mathbf{Q}=\mathbf{U}\mathbf{U}^H}$ and $\nabla_{\mathbf{Q}} h(\mathbf{Q}) = \mathbf{H}^H \big(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H\big)^{-1}\mathbf{H}$ when $h(\mathbf{Q}) = \log|\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H|$. (5.4),(5.5) follow directly from the basic definitions of precoding matrix and capacity(weighted sum-rate).

5.5 Initialisation Step in the PFGD Algorithm

The following remark discusses the initialization step in the PFGD algorithm.

Remark 7(*Initialization of Precoding Matrices*): Initialization of $\{\mathbf{U}_k\}$ is done as previously suggested in (Chaluvadi *et al.*, 2018; Park *et al.*, 2016). Let

$$\mathbf{X}_{k_0} = \frac{1}{\|\nabla_{\mathbf{Q}_k} f(\{\mathbf{0}\}) - \nabla_{\mathbf{Q}_k} f(\{\mathbf{e}_1 \mathbf{e}_1^H\})\|_F} \Pi_+(\nabla_{\mathbf{Q}_k} f(\{\mathbf{0}\})) \tag{5.6}$$

for $k = 1, \cdots, K$ where $\Pi_+(\mathbf{R})$ is the projection of $\mathbf{R}$ onto the positive semi-definite (PSD) matrices set.

In this scenario, $\Pi_+(\nabla_{\mathbf{Q}_k} f(\{\mathbf{0}\})) = w_k \mathbf{H}_k \mathbf{H}_k^H \succeq \mathbf{0}$. Note that $\mathbf{e}_1 = [1, 0, \cdots, 0]^T \in \mathbb{C}^{M\times 1}$.

Next, we need to compute a square root of $\mathbf{X}_{k_0}$ i.e., we have to find a $\tilde{\mathbf{U}}_{k_0}$ such that $\mathbf{X}_{k_0} = \tilde{\mathbf{U}}_{k_0}\tilde{\mathbf{U}}_{k_0}^H$ is satisfied for $k = 1, \cdots, K$.

Finally, we find the initial precoding matrices by projecting $\{\tilde{\mathbf{U}}_{k_0}\}$ onto the total constraint set $\mathcal{C}^{Tot}$, i.e., $\{\tilde{\mathbf{U}}_{k_0}\} = \Pi_{Tot}(\{\tilde{\mathbf{U}}_{k_0}\})$.

5.6 Procedure to describe the projection of a set of matrices onto the Total Constraints Set $\mathcal{C}^{Tot}$

We are now ready to describe the projection of a set of matrices $\{\mathbf{V}_k\}_{k=1}^K$ onto the total constraints set $\mathcal{C}^{Tot}$. The step by step procedure is given below.

1. First, we discuss the projection of $\{\mathbf{V}_k\}$ onto the BD precoding constraints set, $\mathcal{C}^{BD}$ using linear least squares technique in section 5.7.

2. Then, we discuss the projection of $\{\mathbf{V}_k\}$ onto the Joint-SPC-PBPC power constraints set, $\mathcal{C}^{J-SB}$ by minimizing the total linear least squares error in section 5.8.

3. Next, we use the projection of $\{\mathbf{V}_k\}$ onto Joint-SPC-PBPC constraints set as a vehicle in an algorithm to be proposed in section 5.9 that obtains the projection onto Joint-SPC-PBPC-PAPC power constraints set.

4. Now, we have the projections of $\{\mathbf{V}_k\}$ onto $\mathcal{C}^{BD}$ and $\mathcal{C}^{J-SBA}$ in our hand *separately*. Moreover, both $\mathcal{C}^{J-SBA}$ and $\mathcal{C}^{BD}$ are convex sets over the precoding matrices set $\{\mathbf{U}_k\}_{k=1}^K$.

5. However, we want the projection of $\{\mathbf{V}_k\}$ onto $\mathcal{C}^{Tot}$ which is the *intersection* of two convex sets, $\mathcal{C}^{BD}$ and $\mathcal{C}^{J-SBA}$. Hence, in section 5.10, we present *Dykstra's projection algorithm* (Bauschke and Borwein, 1994) which helps us find a projection in the intersection of two convex sets numerically thus giving us the projection of $\{\mathbf{V}_k\}$ onto $\mathcal{C}_{Tot}$.

5.7 Projection Onto BD Precoding Constraints Set

To find the projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto the BD precoding constraints set, we must find a set of matrices $\{\mathbf{U}_k^{BD}\}_{k=1}^K$ in $\mathcal{C}^{BD}$ that satisfies the following: $\mathbf{G}_k\mathbf{U}_k^{BD} = \mathbf{0}, \forall k = 1, \cdots, K$. We have already seen in Lemma 1 that the precoding matrix structure for the k-th MS obtained after eliminating the BD precoding constraints is given by $\mathbf{U}_k^{BD} = \mathbf{B}_{k_0}\mathbf{T}_k^{BD}$ where $\mathbf{T}_k^{BD} \in \mathbb{C}^{(M-L)\times R}$. We also note that rank($\mathbf{T}_k^{BD}$) = min($M - L, R$) = R since $M - L = M - N(K-1) > N \geq R$ (because $M \geq NK$).

Next, we find $\{\mathbf{T}_k^{BD}\}_{k=1}^K$ by posing the following least squares problems which ensure that $\{\mathbf{U}_k^{BD}\}$ is the closest to $\{\mathbf{V}_k\}$ in $\mathcal{C}^{BD}$.

$$\mathbf{T}_k^{BD} = \underset{\{\mathbf{t}_{k_j}\}}{\text{argmin}} \sum_{j=1}^{R} \|\mathbf{B}_{\mathbf{k_0}}\mathbf{t}_{k_j} - \mathbf{v}_{k_j}\|^2 \; \forall k = 1, \cdots, K. \quad (5.7)$$

where $\mathbf{t}_{k_j}$ and $\mathbf{v}_{k_j}$ are the j-th columns of the matrices $\mathbf{T}_k$ and $\mathbf{V}_k$ respectively.

The following theorem gives us $\{\mathbf{U}_k^{BD}\}_{k=1}^K$, the projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto the BD precoding constraints set, $\mathcal{C}^{BD}$.

Theorem 4 *The projection of a set of matrices $\{\mathbf{V}_k\}_{k=1}^K$ onto the BD precoding constraints set $\mathcal{C}^{BD}$ is given by*

$$\mathbf{U}_k^{BD} = \mathbf{B}_{k_0}\mathbf{B}_{k_0}^H\mathbf{V}_k, \;\; \forall k = 1, \cdots, K. \quad (5.8)$$

Proof: Please refer to Appendix F.

5.8 Projection Onto Joint-SPC-PBPC Constraints Set

Our goal is to find a set of matrices $\{\mathbf{U}_k^{J-SB}\}_{k=1}^{K}$ from $\mathcal{C}^{J-SB}$ that are closest to $\{\mathbf{V}_k\}$. Towards this end, we formulate the following problem.

$$\text{(P10)}: \quad \begin{aligned} \min_{\{\mathbf{u}_{k_l}\}} \quad & \sum_{k=1}^{K}\sum_{l=1}^{M} \|\mathbf{u}_{k_l} - \mathbf{v}_{k_l}\|^2 \\ \text{s.t.} \quad & \sum_{k=1}^{K}\sum_{l=1}^{M} \|\mathbf{u}_{k_l}\|^2 \leq P_{sum} \\ & \sum_{k=1}^{K}\sum_{l\in I(a)} \|\mathbf{u}_{k_l}\|^2 \leq \tilde{P}_a, \;\; a = 1, \cdots, A \end{aligned}$$

So, we conclude that the optimal primal variables obtained after solving the above problem (P10)are the required $\{\mathbf{U}_k^{J-SB}\}$. Also, (P10) is a convex optimization problem since the objective is a convex function over $\{\mathbf{u}_{k_l}\}$ and the Joint-SPC-PBPC power constraints form a convex set over $\{\mathbf{u}_{k_l}\}$. In fact, the SPC and PBPC in (P10) are norm balls over $\{\mathbf{u}_{k_l}\}$ and hence generate a convex set. By introducing a set of non-negative dual variables, $\gamma, \{\lambda_a\}_{a=1}^{A}$, associated with sum power constraint and per-BS power constraints respectively ,the Lagrangain function of (P10) can be written as

$$\begin{aligned} L(\mathbf{u}_{k_l}, \gamma, \{\lambda_a\}) = & \sum_{k=1}^{K}\sum_{l=1}^{M} \|\mathbf{u}_{k_l} - \mathbf{v}_{k_l}\|^2 \\ & + \gamma\Big(\sum_{k=1}^{K}\sum_{l=1}^{M} \|\mathbf{u}_{k_l}\|^2 - P_{sum}\Big) \\ & + \sum_{a=1}^{A} \lambda_a\Big(\sum_{k=1}^{K}\sum_{l\in I(a)} \|\mathbf{u}_{k_l}\|^2 - \tilde{P}_a\Big) \end{aligned} \tag{5.9}$$

In (P10), the objective function and the constraints are smooth, the slater's conditions are satisfied and strong duality holds i.e, the duality gap between the primal and dual objective functions of (P10) is zero. Therefore, all the assumptions required for KKT conditions to be valid are active. From the KKT conditions (Boyd and Vandenberghe, 2004), we know that the optimal primal variables $\{\mathbf{u}_{k_l}^*\}$ minimizes the Lagrangian function at optimal dual variables i.e.,

$$\nabla_{\mathbf{u}_{k_l}} L(\mathbf{u}_{k_l}, \gamma^*, \{\lambda_a^*\})\Big|_{\mathbf{u}_{k_l}^*} = \mathbf{0}, \;\; \forall k, l \tag{5.10}$$

Solving (5.10), we observe that each row $\mathbf{u}_{k_l}^*, \forall k = 1, \cdots, K$ and $\forall l = 1, \cdots, M$ is a scaled version of the original row $\mathbf{v}_{k_l}$ i.e., we get

$$\mathbf{u}_{k_l}^* = \frac{\mathbf{v}_{k_l}}{1 + \gamma^* + \lambda_a^*}, \quad l \in I(a).$$

For finding the projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto Joint-SPC-PBPC power constraints set $\mathcal{C}^{J-SB}$, we arrange the Base Stations (indexed by "a") in the ascending order of $\frac{\tilde{P}_a}{\sum_{k=1}^K \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2}$ and place them in an ordered set Δ where $\mathbf{v}_{k_l}$ denotes the l-th row of $\mathbf{V}_k$, $I(a) = \{(a-1)M_B + 1, \cdots, aM_B\}$ and $\Delta(j)$ for $j = 1, \cdots, A$ denotes the j-th element of Δ.

The following theorem gives us $\{\mathbf{U}_k^{J-SB}\}_{k=1}^K$, the projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto the Joint-SPC-PBPC constraints set, $\mathcal{C}^{J-SB}$.

Theorem 5 *Let $\{\mathbf{U}_k^{J-SB}\}_{k=1}^K$ be the projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto the Joint-SPC-PBPC constraints set, $\mathcal{C}^{J-SB}$. Then, each row $\mathbf{u}_{k_l}^{J-SB}$ of the matrix $\mathbf{U}_k^{J-SB}$ is a scaled version of the original row $\mathbf{v}_{k_l}$ of the matrix $\mathbf{V}_k$ i.e.,*

$$\mathbf{u}_{k_l}^{J-SB} = \xi_l \mathbf{v}_{k_l}, \quad \forall k, l \tag{5.11}$$

where

$$\xi_l = \sqrt{\frac{\tilde{P}_a}{\sum_{k=1}^K \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2}} \tag{5.12}$$

for $l \in I(a), a \in \{\Delta(1), \cdots, \Delta(m)\}$ and

$$\xi_l = \xi = \sqrt{\frac{P_{sum} - \sum_{a=\Delta(1)}^{\Delta(m)} \tilde{P}_a}{\sum_{k=1}^K \sum_{a=\Delta(m+1)}^{\Delta(A)} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2}} \tag{5.13}$$

for $l \in I(a), a \in \{\Delta(m+1), \cdots, \Delta(A)\}$ where m is the least element in $\{0, 1, 2, \cdots, A-1\}$ such that

$$\frac{\tilde{P}_{\Delta(m+1)}}{\sum_{k=1}^K \sum_{l \in I(\Delta(m+1))} \|\mathbf{v}_{k_l}\|^2} \geq \frac{P_{sum} - \sum_{a=\Delta(1)}^{\Delta(m)} \tilde{P}_a}{\sum_{k=1}^K \sum_{a=\Delta(m+1)}^{\Delta(A)} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2} \tag{5.14}$$

Note that the scaling factor ξ_l is independent of k always and independent of l also when $a \in \{\Delta(m+1), \cdots, \Delta(A)\}$.

Proof: Please refer to Appendix G.

Therefore, we can conclude that

$$\mathbf{U}_k^{J-SB} = \mathbf{\Xi^{J-SB}} \mathbf{V}_k, \ \forall k = 1, \cdots, K. \tag{5.15}$$

where $\Xi^{J-SB} = \text{Diag}(\xi_1, \cdots, \xi_M)$ in (5.15).

5.9 Projection Onto Joint-SPC-PBPC-PAPC Constraints Set

First and foremost, we note from Theorem 5 that each row $\mathbf{u}_{k_l}^{J-SB}$ of the matrix $\mathbf{U}_k^{J-SB}$ obtained after projection onto Joint-SPC-PBPC power constraints set is just a scaled version of the row $\mathbf{v}_{k_l}$ of the matrix $\mathbf{V}_k$ before projection. Therefore, the key idea here is to find the projection onto $\mathcal{C}^{J-SB}$ first and then check if the set of rows $\{\mathbf{u}_{k_i}^{J-SB}\}_{k=1}^K \forall i = 1, \cdots, M$ violate PAPC .

If the PAPC constraint corresponding to the antenna "i" is violated, then we must scale these rows $\{\mathbf{u}_{k_i}^{J-SB}\}_{k=1}^K$ appropriately such that the sum of their norms is equal to the PAPC constraint $\hat{P}_i$ and the scaled $\{\mathbf{u}_{k_i}^{J-SB}\}_{k=1}^K$ will be our new projection onto $\mathcal{C}^{J-SBA}$, $\{\mathbf{u}_{k_i}^{J-SBA}\}_{k=1}^K$. The justification for the above scaling is provided by the following lemma inspired from Lemma 2 in both (Le Cao and Oechtering, 2017; Chaluvadi *et al.*, 2018).

Lemma 4 *Denote* $Z = \{1, \cdots, M\}$. *Let* $C \subseteq Z$, *and* $S(C) := \Big\{ \{\mathbf{u}_{k_i}^{J-SB}\}_{k=1}^K$*:* $\sum_{k=1}^K \|\mathbf{u}_{k_i}^{J-SB}\|^2 \leq \hat{P}_i, \forall i \in C \Big\}$. *Let D :=* $\Big\{ i \in Z \backslash C : \sum_{k=1}^K \|\mathbf{u}_{k_i}^{J-SB}\|^2 > \hat{P}_i \Big\}$. *Let* $P_i^{S(Z)}$ *denote the optimal power allocation under the per-antenna power constraints set* $S(Z)$. *Note that* $Z = C \cup D$. *If* $D \neq \emptyset$, *then* $P_i^{S(Z)} = \hat{P}_i, \forall i \in D$.

Proof: Please refer to Lemma 2 in (Chaluvadi *et al.*, 2018).

Therefore, once the PAPC have been checked for violation by the Joint-SPC-PBPC projection, we know that the optimal power under Joint-SPC-PBPC-PAPC for atleast one antenna. Next, we can remove the transmit antennas whose optimal power have been determined and formulate a new Joint-SPC-PBPC problem by modifying $P_{sum}, \tilde{P}_a$ and $I(a), \forall a = 1, \cdots, A$ as described in the algortihm below. We determine the optimal Joint-SPC-PBPC-PAPC projection in atmost M steps repeating the above steps.

Denote $G = \{1, \cdots, A\}, I(a) = \{(a-1)M_B+1, \cdots, aM_B\}$ and $Z = \{1, \cdots, M\}$. The algorithm to obtain projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto $\mathcal{C}^{J-SBA}$ is explained in the following steps:

1. **Repeat**: Compute the per-antenna powers allocated for $\{\mathbf{u}_{k_i}^{J-SB}\}_{k=1}^K$ i.e., $P_i^{J-SB} = \sum_{k=1}^K \|\mathbf{u}_{k_i}^{J-SB}\|^2, \forall i \in Z.$

2. **Check for PAPC Violation**: Check if for any $i \in Z, P_i^{J-SB} > \hat{P}_i$.

- **If No**: Output the projection of $\{\mathbf{V}_k\}$ onto the Joint-SPC-PBPC-PAPC constraints set, $\mathcal{C}^{J-SBA}$ as

$$\mathbf{U}_k^{J-SBA} = \mathbf{\Xi^{J-SB}}\mathbf{V}_k, \ \ \forall k = 1, \cdots, K.$$

- **If Yes**:
  - Set $\mathcal{P}_a = \{j \in I(a) : P_j^{J-SB} > \hat{P}_j\}, \forall a \in G.$
  - Allocate optimal per-antenna powers for projection onto $\mathcal{C}^{J-SBA}$ as $P_i^{J-SBA} \leftarrow \hat{P}_i, \forall i \in \mathcal{P}_a, \forall a \in G.$
  - This means $\forall i \in \mathcal{P}_a, \forall a \in G$, we have to update the scaling factor as

$$\xi_i^{J-SBA} = \xi_i^{J-SB}\sqrt{\frac{\hat{P}_i}{P_i^{J-SB}}}$$

  - Formulate Joint-SPC-PBPC problem for the remaining antennas by updating as follows:
    i. $Z = Z \backslash \{\mathcal{P}_a, a \in G\}.$
    ii. $\mathbf{H}_k = [\mathbf{h}_{k_i}]_{i \in Z}, \forall k = 1, \cdots, K.$
    iii. $P_{sum} \leftarrow P_{sum} - \sum_{a \in G}\sum_{i \in \mathcal{P}_a} \hat{P}_i.$
    iv. $\tilde{P}_a \leftarrow \tilde{P}_a - \sum_{i \in \mathcal{P}_a} \hat{P}_i, \forall a \in G.$
    v. Go back to the step "**Repeat**" and solve the reduced size problem.

5.10 Dykstra's Projection Onto Total Constraints Set

We have the projections of $\{\mathbf{V}_k\}_{k=1}^K$ onto two convex sets $\mathcal{C}^{BD}$ and $\mathcal{C}^{J-SBA}$ separately in our hand now. Dykstra's projection algorithm finds for each set of matrices $\{\mathbf{V}_k\}_{k=1}^K$, the only $\{\mathbf{U}_k^{Tot}\}_{k=1}^K \in \mathcal{C}^{BD} \cap \mathcal{C}^{J-SBA} = \mathcal{C}^{Tot}$ numerically such that:

$$\sum_{k=1}^{K}\sum_{l=1}^{M} ||\mathbf{u}_{k_l}^{Tot} - \mathbf{v}_{k_l}||^2 \leq \sum_{k=1}^{K}\sum_{l=1}^{M} ||\mathbf{u}_{k_l} - \mathbf{v}_{k_l}||^2 \tag{5.16}$$

The above problem in (5.16) is equivalent to finding the projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto $\mathcal{C}^{BD} \cap \mathcal{C}^{J-SBA} = \mathcal{C}^{Tot}$, which we have already denoted by $\{\mathbf{U}_k^{Tot}\}_{k=1}^K$. The Dykstra's projection algorithm is described in the following steps:

1. **Initialize** $\mathbf{X}_{k_0} = \mathbf{V}_k, \mathbf{P}_{k_0} = \mathbf{Q}_{k_0} = \mathbf{0}, \forall k = 1, \cdots, K.$

2. **Update**: At the p-th iteration,

- $\{\mathbf{Y}_{k_p}\}_{k=1}^K = \Pi_{J-SBA}(\{\mathbf{X}_{k_p} + \mathbf{P}_{k_p}\}_{k=1}^K).$
- $\mathbf{P}_{k_{p+1}} = \mathbf{X}_{k_p} + \mathbf{P}_{k_p} - \mathbf{Y}_{k_p}, \forall k = 1, \cdots, K.$
- $\{\mathbf{X}_{k_{p+1}}\}_{k=1}^K = \Pi_{BD}(\{\mathbf{Y}_{k_p} + \mathbf{Q}_{k_p}\}_{k=1}^K).$
- $\mathbf{Q}_{k_{p+1}} = \mathbf{Y}_{k_p} + \mathbf{Q}_{k_p} - \mathbf{X}_{k_{p+1}}, \forall k = 1, \cdots, K.$

where $\Pi_{BD}(.)$ and Π_{J-SBA} denote the projections onto BD precoding constraints set and Joint-SPC-PBPC-PAPC power constraints set respectively.

3. **Stopping Criterion**: Stop the algorithm when

$$\|\mathbf{X}_{k_{p+1}} - \mathbf{X}_{k_p}\|_F \leq \epsilon, \forall k = 1, \cdots, K. \tag{5.17}$$

where $\|.\|_F$ in (5.17) denotes the frobenius norm of a matrix. Typically, we choose ϵ in the order of 10^{-3}.

4. **The Final Projection**: Therefore, the final projection of $\{\mathbf{V}_k\}_{k=1}^K$ onto $\mathcal{C}^{Tot} = \mathcal{C}^{BD} \cap \mathcal{C}^{J-SBA}$ is given by

$$\mathbf{U}_k^{Tot} = \mathbf{X}_{k_{p+1}}, \forall k = 1, \cdots, K.$$

CHAPTER 6

NUMERICAL RESULTS

In this chapter, we provide numerical examples to illustrate the results in the paper. For the purpose of exposition, we assume the downlink channels $\mathbf{H}_k$'s in (2.1) to be independent over k, and all the elements in each $\mathbf{H}_k$ to be independent CSCG random variables with zero mean and unit variance. Moreover, we consider the sum-rate maximisation for the fully cooperative multi-cell downlink transmission system, i.e., w_k's are all equal to one in (P1)-(P10). In the following sections, we present the obtained simulation results along with related discussions.

6.1 Convergence Behavior

6.1.1 Convergence Behavior of Dykstra's Projection Algorithm

In Fig. 6.1, we show the convergence behavior of Dykstra's projection algorithm for obtaining the projection onto $C^{Tot} = \mathcal{C}^{BD} \cap \mathcal{C}^{J-SBA}$. We assume that $A = 4, M_B = 3, K = 4, N = 2$. The transmit power constraint for each antenna is $\hat{P}_i = 1W$. The transmit power constraint for each base station is considered $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}$ and the sum power constraint is then taken as $P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ keeping in mind, the sufficient conditions on power constraints in (3.4). The error b/w two consecutive projections in the frobenius norm sense is shown against different iterations for all the four MSs in the system. As observed, the error converges to prescribed accuracy of 10^{-3} within almost 10 iterations.

6.1.2 Convergence Behavior of PFGD Algorithm

In Fig. 6.2, we show the convergence behavior of PFGD algorithm for obtaining the optimal precoding matrices that maximise the sum-rate under Joint-SPC-PBPC-PAPC. We assume that $A = 4, M_B = 3, K = 4, N = 2$, Rank = $R = 2$. We have chosen

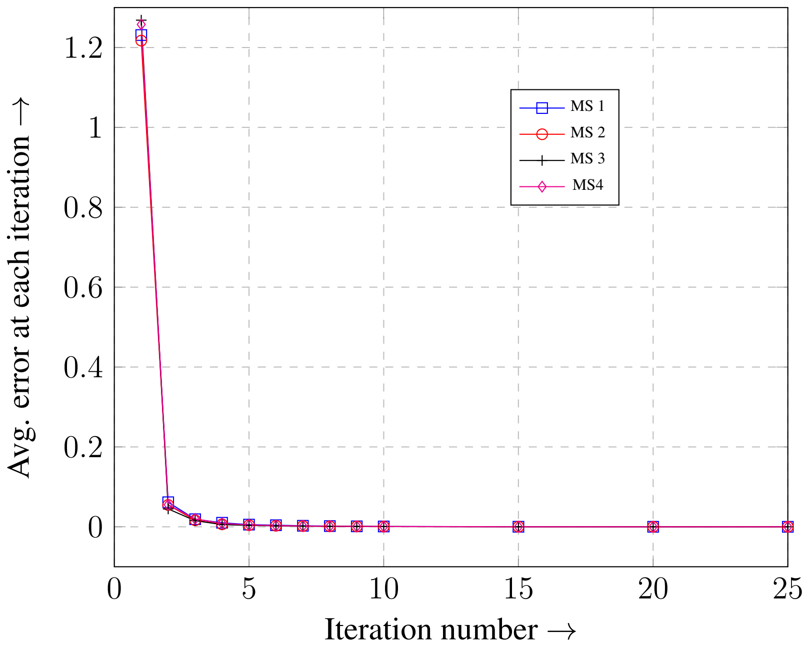

Figure 6.1: $A = 4, M_B = 3, K = 4, N = 2$ & $\hat{P}_i = 1W$, $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}$, $P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$

$\eta = 2$, the fixed step-size parameter in a trial and error method. The transmit power constraint for each antenna is $\hat{P}_i = 1W$. The transmit power constraint for each base station is considered $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}$ and the sum power constraint is then taken as $P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ keeping in mind, the sufficient conditions on power constraints in (3.4). The error b/w two consecutive precoding matrices in the frobenius norm sense is shown against different iterations for all the four MSs in the system. As observed, the error converges to prescribed accuracy of 10^{-3} within almost 7-8 iterations. Therefore, we observe that the convergence behavior is good even though there is no strong therotitical claim supporting this kind of convergence. In (Chaluvadi *et al.*, 2018), the authors have proved a local convergence gurantee for PFGD algorithm in SU-MIMO BC case under Joint-SPC-PBPC-PAPC and rank constraints (no BD precoding in this case).

6.1.3 Comments on Complexity of the PFGD Algorithm

In every iteration of the PFGD algorithm, we have to find "K" gradients and do a projection step. The dominant computations are the gradient computations and the projections onto BD precoding constraints set, $\mathcal{C}^{BD}$ since the projections onto $\mathcal{C}^{J-SBA}$ are mainly scaling operations. Gradient, $\nabla_{\mathbf{U}_{k_p}} f(\{\mathbf{U}_{k_p}\mathbf{U}_{k_p}^H\}) = 2w_k \mathbf{H}_k^H(\mathbf{I}_N + \mathbf{H}_k\mathbf{U}_{k_p}(\mathbf{H}_k\mathbf{U}_{k_p})^H)^{-1}\mathbf{H}_k\mathbf{U}_{k_p}$ can be written as $\nabla_{\mathbf{U}_{k_p}} f(\{\mathbf{U}_{k_p}\mathbf{U}_{k_p}^H\}) = 2w_k \mathbf{H}_k^H(\mathbf{H}_k\mathbf{U}_{k_p})(\mathbf{I}_R + (\mathbf{H}_k\mathbf{U}_{k_p})^H(\mathbf{H}_k\mathbf{U}_{k_p}))^{-1}$ using matrix inversion lemma to reduce complexity as $R \leq N$.

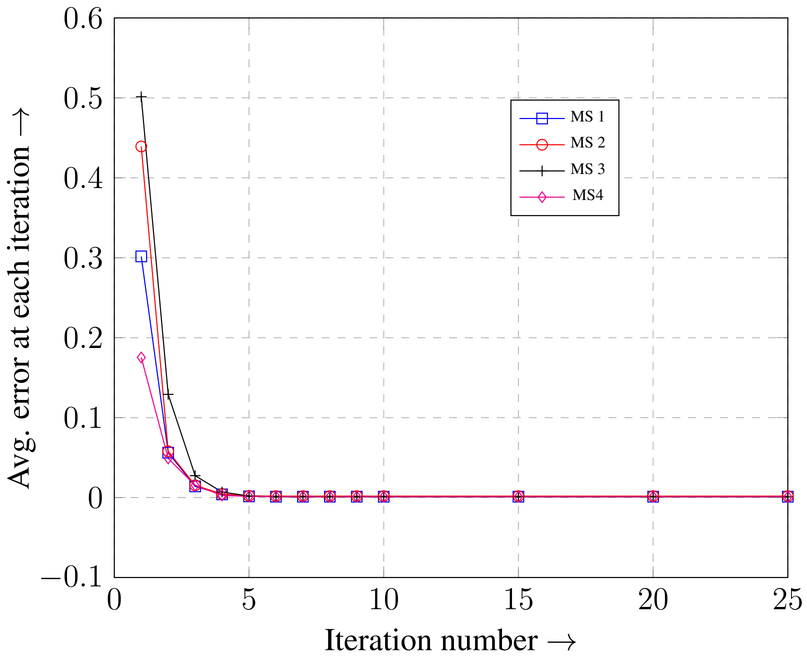

Figure 6.2: $A = 4, M_B = 3, K = 4, N = 2$ & $\hat{P}_i = 1W$, $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}$, $P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$ & $\eta = 2$

Finding the gradient involves finding inverse of a $R \times R$ matrix and multiplication of $M \times N, N \times M, M \times R, R \times R$ matrices and a scalar multiplication which can be ignored for now (as $w_k = 1$). Finding the projection onto $\mathcal{C}^{BD}$ involves finding the multiplication of $M \times (M-L), (M-L) \times M, M \times R$ matrices. Complexity of inverse operation for an $R \times R$ matrix is $6(2R^3)$ flops and multiplication of $M \times R$ matrix, $R \times L$ matrix is $6(2MRL)$ flops (Golub and Van Loan, 2012). Also, we count every complex operation as 6 real flops.

Therefore, complexity of PFGD algorithm is given by $6K(2R^3 + 2MR(R + 2N) + 2M^2(M - L + R))$ per iteration where $L = N(K-1)$. Note that the complexity of the optimal algorithm (A1) in full rank case requires more multiplications as SVD's are involved.

6.2 MU-MISO BC with Joint-SPC-PAPC - Comparison of (A1) and (A2) performance

6.2.1 Case Study 1

Next, we consider a special case of the fully cooperative multi-cell downlink transmission system with $M_B = N = 1$, which is equivalent to a MU-MISO BC with the corresponding Joint-SPC-PAPC power constraints. We assume that $K = 2$ and vary the number of transmitting antennas from 2 to 10. The transmit power constraint for each antenna is assumed to be $\hat{P}_i = 1.25W$. The sum power constraint is then taken as $P_{sum} = \frac{AM_B\hat{P}_i}{1.25}$ keeping in mind, the sufficient conditions on power constraints in (3.4). In Fig. 6.3, we compare the achievable sum-rate averaged over 1000 random network simulations with the optimal ZF-BF precoder obtained by (A1) against that with the suboptimal precoder obtained by (A2). It is observed that when $M = K = 2$, the achievable rates for both the optimal and suboptimal precoders are identical, which is in accordance with our discussion in Remark 6, Chapter 4. We can also observe that when $M > K$, the sum-rate gain of the optimal precoder solution over the suboptimal solution increases with M.

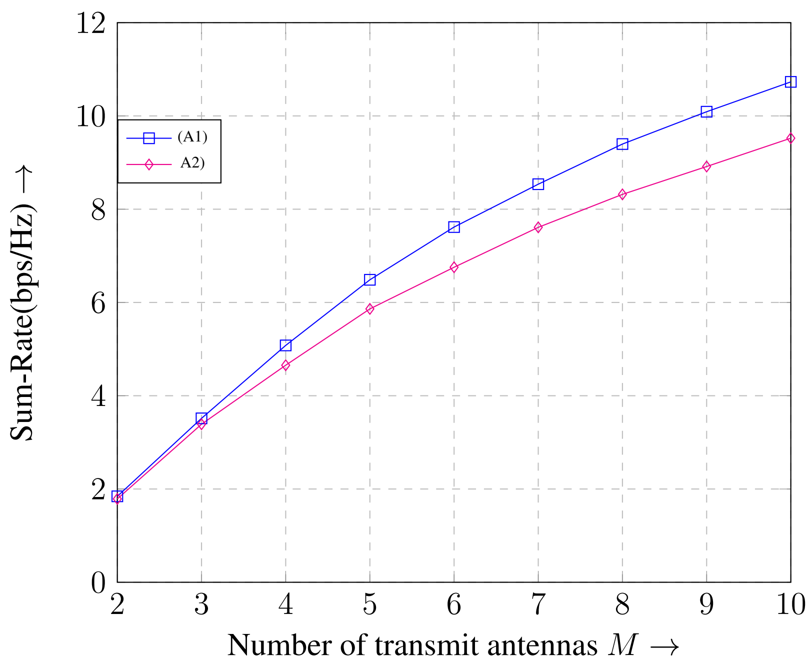

Figure 6.3: $A = 2 : 1 : 10, M_B = 1, K = 2, N = 1$ & $\hat{P}_i = 1.25, P_{sum} = \frac{AM_B\hat{P}_i}{1.25}$

6.2.2 Case Study 2

We also consider another special case of the fully cooperative multi-cell downlink transmisssion system with $M_B = N = 1$, which is equivalent to a MU-MISO BC with the corresponding Joint-SPC-PAPC power constraints. This time, we assume that $K = 4$ and vary the number of transmitting antennas from 4 to 12. The transmit power constraint for each antenna is assumed to be $\hat{P}_i = 1.25W$. The sum power constraint is then taken as $P_{sum} = \frac{AM_B\hat{P}_i}{1.25}$ keeping in mind, the sufficient conditions on power constraints in (3.4). In Fig. 6.4, we compare the achievable sum-rate averaged over 1000 random network simulations with the optimal ZF-BF precoder obtained by (A1) against that with the suboptimal precoder obtained by (A2). It is observed that when $M = K = 4$, the achievable rates for both the optimal and suboptimal precoders are identical, which is in accordance with our discussion in Remark 6, Chapter 4. We can also observe that when $M > K$, the sum-rate gain of the optimal precoder solution over the suboptimal solution increases with M.

Therefore, we can conclude that our observations in both the case studies have been consistent.

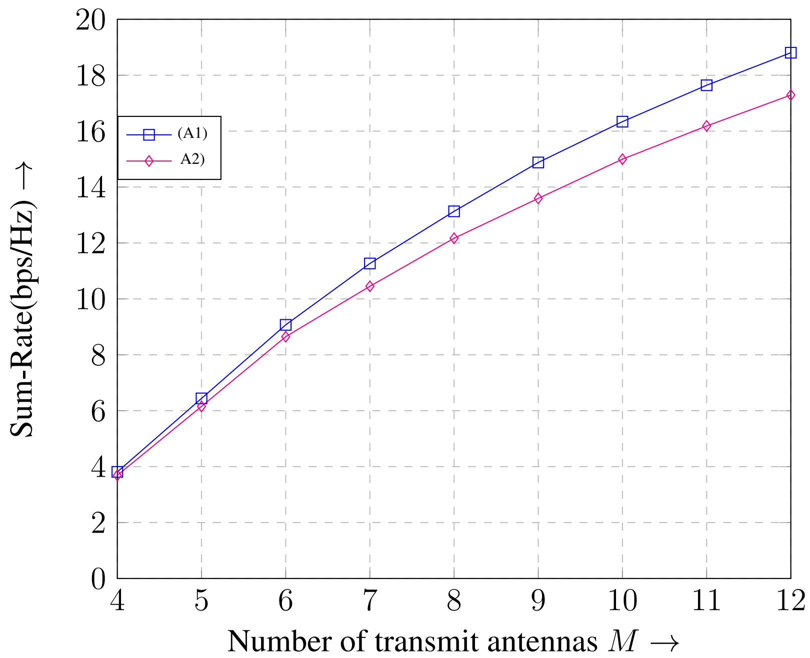

Figure 6.4: $A = 4 : 1 : 12, M_B = 1, K = 4, N = 1$ & $\hat{P}_i = 1.25, P_{sum} = \frac{AM_B\hat{P}_i}{1.25}$

6.3 MU-MIMO BC with Joint-SPC-PBPC-PAPC - Comparison of (A1), (A2), PFGD (A3) and CVX performance

6.3.1 Case Study 1

We consider the case of multi-antenna MS receivers. For the corresponding auxiliary BC, we assume that $A = 4, M_B = 3, K = 4, N = 2$. Also, we maintain $\hat{P}_i = \hat{P}, \forall i; \tilde{P}_a = \frac{M_B \hat{P}}{1.1}, \forall a; P_{sum} = \frac{A M_B \hat{P}}{(1.1)^2}$ so that the sufficient conditions for relevant power constraints in (3.4) are satisfied. In Fig. 6.5, we show the achievable sum-rates for all the optimal (A1), suboptimal (A2) BD precoders, PFGD (A3) algorithm for both Rank-2 and Rank-1 cases and Standard CVX package vs. the per-antenna transmit power constraint P. We vary P from 0 to 20 dB. It is observed that optimal BD precoder solution (A1), PFGD - Full Rank Case (Rank = 2) and Standard CVX package achieve the same sum-rates. We observe that although the optimal precoder solution still performs better than the suboptimal one, their rate gap is not large. Furthermore, the PFGD - Rank 1 case performs much worse than the remaining schemes as expected.

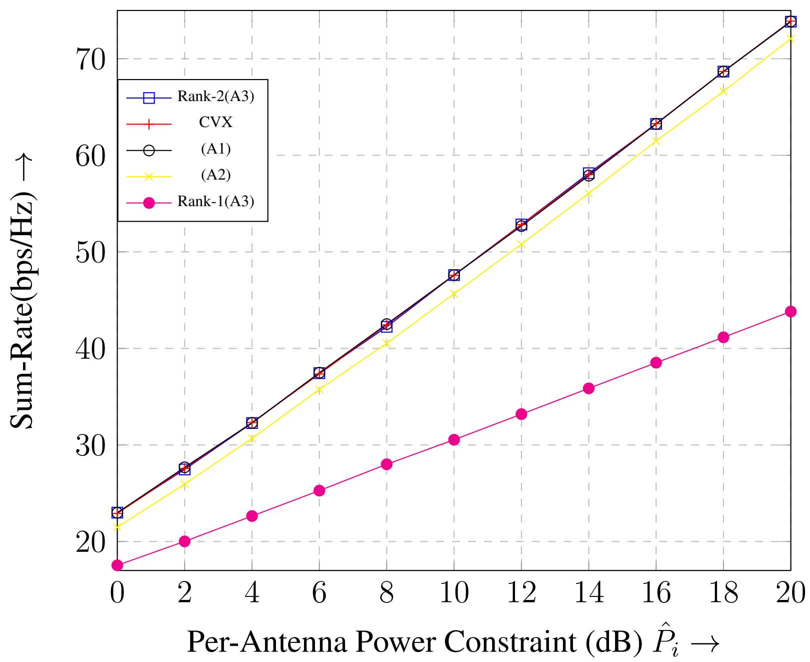

Figure 6.5: $A = 4, M_B = 3, K = 4, N = 2$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$

6.3.2 Case Study 2

Next, we consider another case of MU-MIMO BC. We assume that $A = 3, M_B = 3, K = 3, N = 2$. Also, we maintain $\hat{P}_i = \hat{P}, \forall i; \tilde{P}_a = \frac{M_B \hat{P}}{1.1}, \forall a; P_{sum} = \frac{AM_B\hat{P}}{(1.1)^2}$ so that the sufficient conditions for relevant power constraints in (3.4) are satisfied. In Fig. 6.6, we show the achievable sum-rates for all the optimal (A1), suboptimal (A2) BD precoders, PFGD (A3) algorithm for both Rank-2 and Rank-1 cases and Standard CVX package vs. the per-antenna transmit power constraint P. We vary P from 0 to 20 dB. It is again observed that optimal BD precoder solution (A1), PFGD - Full Rank Case (Rank = 2) and Standard CVX package achieve the same sum-rates. We also observe that although the optimal precoder solution still performs better than the suboptimal one, their rate gap is not large. Furthermore, the PFGD - Rank 1 case performs much worse than the remaining schemes as expected.

Therefore, the observations in both the case studies have been consistent.

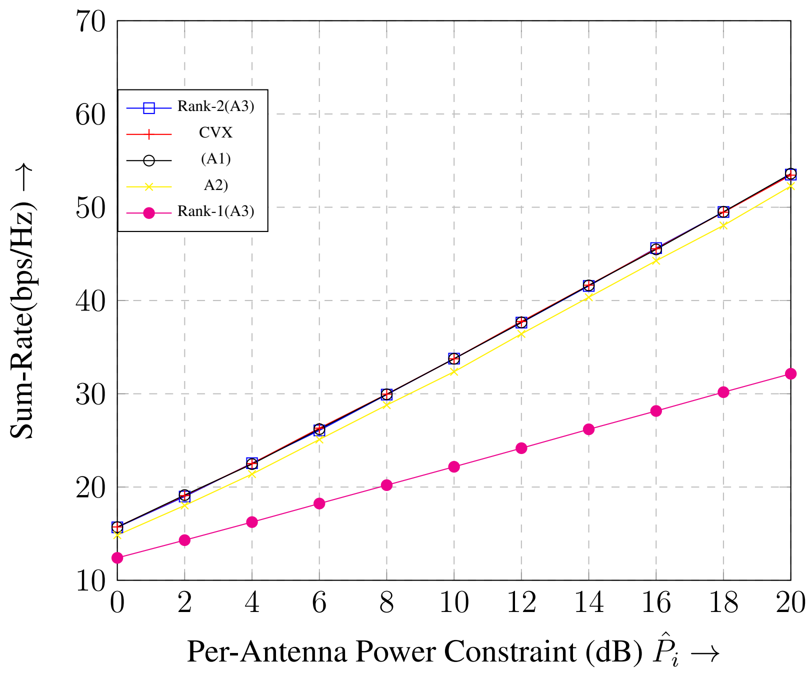

Figure 6.6: $A = 3, M_B = 3, K = 3, N = 2$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{AM_B\hat{P}_i}{(1.1)^2}$

6.4 MU-MIMO BC with Joint-SPC-PBPC-PAPC - Comparison of PFGD (A3) performance under different rank constraints

6.4.1 Case Study 1

We consider the case of an auxiliary MU-MIMO BC where we assume $A = 4, M_B = 4, K = 4, N = 4$. Since $N = 4$, we can vary the rank of the precoding matrices from 1 to 4. Also, we maintain $\hat{P}_i = \hat{P}, \forall i; \tilde{P}_a = \frac{M_B \hat{P}}{1.1}, \forall a; P_{sum} = \frac{A M_B \hat{P}}{(1.1)^2}$ so that the sufficient conditions for relevant power constraints in (3.4) are satisfied. In Fig. 6.7, we show the achievable sum-rates for the PFGD (A3) algorithm for all the four possible ranks vs. the per-antenna transmit power constraint P. We vary P from -10 to 30 dB. Furthermore, the PFGD (A3) performance for a higher rank is always superior to a lower rank in high SNR conditions (high values of P, i.e., P > -5). However, in low SNR conditions (low values of P, i.e., P < -5), the PFGD (A3) performance is almost the same for all the four possible ranks.

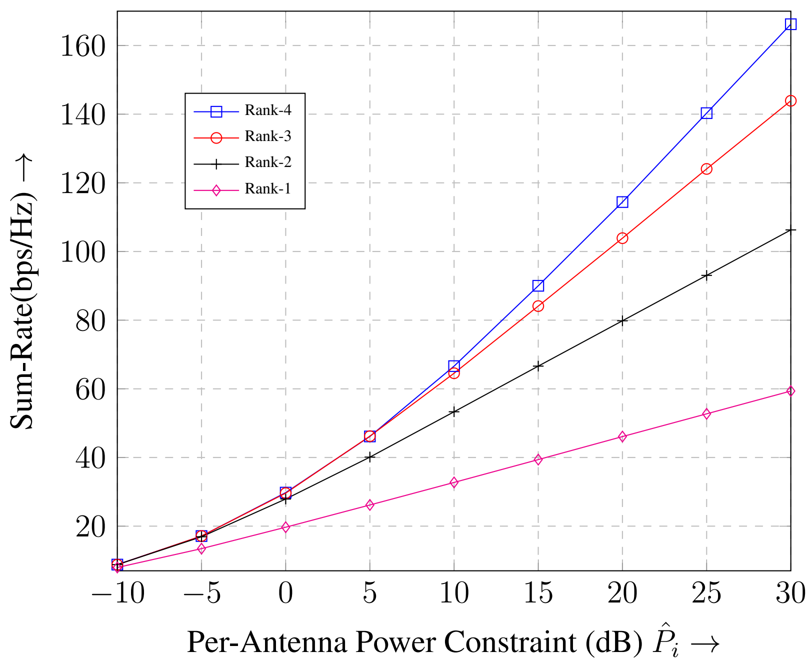

Figure 6.7: $A = 4, M_B = 4, K = 4, N = 4$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$

6.4.2 Case Study 2

Next, we consider another case of an auxiliary MU-MIMO BC where we assume $A = 4, M_B = 5, K = 4, N = 4$. Since $N = 4$, we can vary the rank of the precoding matrices from 2 to 4. Also, we maintain $\hat{P}_i = \hat{P}, \forall i; \tilde{P}_a = \frac{M_B \hat{P}}{1.1}, \forall a; P_{sum} = \frac{A M_B \hat{P}}{(1.1)^2}$ so that the sufficient conditions for relevant power constraints in (3.4) are satisfied. In Fig. 6.8, we show the achievable sum-rates for the PFGD (A3) algorithm for all the three chosen ranks ($R = 2, 3, 4$) vs. the per-antenna transmit power constraint P. We vary P from -10 to 30 dB. Furthermore, the PFGD (A3) performance for a higher rank is always superior to a lower rank in high SNR conditions (high values of P, i.e., P > -5). However, in low SNR conditions (low values of P, i.e., P < -5), the PFGD (A3) performance is almost the same for all the four possible ranks.

Therefore, we see that the observations have been consistent in both the case studies.

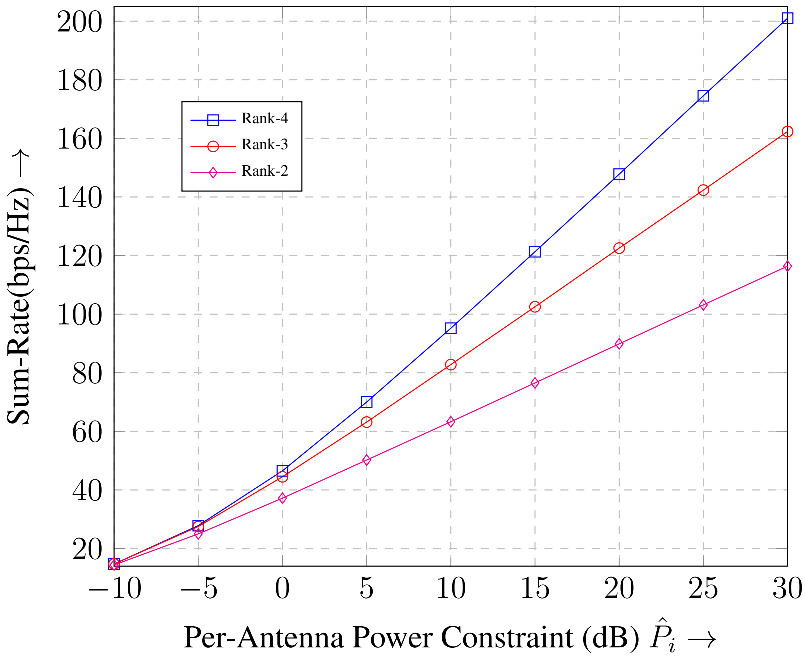

Figure 6.8: $A = 4, M_B = 5, K = 4, N = 4$ & $\tilde{P}_a = \frac{M_B \hat{P}_i}{1.1}, P_{sum} = \frac{A M_B \hat{P}_i}{(1.1)^2}$

CHAPTER 7

SUMMARY

This paper obtains the optimal transmission scheme for a fully cooperative multi-cell system (with multiple users) that can be modeled as an auxiliary MU-MIMO BC under Block Diagonalisation (BD) precoding constraints and multiple simulataneous power constraints such as sum, per-base station and per-antenna power constraints (Joint-SPC-PBPC-PAPC) with and without constraints on the number of spatial multiplexing data streams available for each MS (rank constraints). By applying linear algebra and convex optimization techniques, this paper derives the closed form expression for the optimal transmit covariance matrices corresponding to all the users to maximize the weighted sum-rate in the multi-cell system without rank constraints. In this case, the optimal BD precoding vectors for each user are shown to be *non-orthogonal* in general, which differs from the conventional orthogonal precoder design for the sum-power constraint case where the optimal zero-forcing beamforming vectors are orthogonal. A suboptimal heuristic method method is also proposed, which combines the conventional orthogonal BD precoder design with an optimised power allocation to meet the Joint-SPC-PBPC-PAPC. Furthermore, this paper shows that the proposed optimal BD precoder solution reduces down to the optimal zero-forcing beamforming (ZF-BF) solution for the special case of MU-MISO BC under Joint-SPC-PAPC. Since analytical solutions are not possible for the general MU-MIMO case, this paper proposes a Projected Factored Gradient Descent (PFGD) algorithm to find the optimal precoding matrices that maximize the weighted sum-rate of all the users in the multi-cell system under BD precoding and Joint-SPC-PBPC-PAPC along with rank constraints on transmit covariance matrices numerically. The proposed PFGD algorithm is extremely advantageous for finding optimal transmission strategies under low rank constraints and also preferred over the optimal algorithm proposed earlier in the full-rank case owing to it's very low implementation complexity.

APPENDIX

A. Proof of Lemma 1

We first argue that $\mathbf{B}_{k_0}$ forms an orthogonal basis for the null-space of $\mathbf{G}_k$. We already know that the rank of the matrix $\mathbf{G}_k \in \mathbb{C}^{L\times M}$ is min(L, M) = $L = N(K-1)$. Therefore, null($\mathbf{G}_k$) must have $(M-L)$ $M\times 1$ orthogonal basis vectors. From the fact that $\mathbf{B}_k$ is a unitary matrix, we can easily derive the following.

$$\mathbf{B}_k\mathbf{B}_k^H = \mathbf{I} \Longleftrightarrow \mathbf{B}_{k_1}\mathbf{B}_{k_1}^H + \mathbf{B}_{k_0}\mathbf{B}_{k_0}^H = \mathbf{I}$$

$$\begin{aligned} \mathbf{B}_k^H\mathbf{B}_k = \mathbf{I} \Longleftrightarrow \mathbf{B}_{k_1}^H\mathbf{B}_{k_1} &= \mathbf{B}_{k_0}^H\mathbf{B}_{k_0} = \mathbf{I} \\ \mathbf{B}_{k_1}^H\mathbf{B}_{k_0} &= \mathbf{B}_{k_0}^H\mathbf{B}_{k_1} = \mathbf{0} \end{aligned} \tag{7.1}$$

Recall from the BD precoding constraints in (P2) that $\mathbf{H}_j\mathbf{U}_k = \mathbf{0}, \forall j \neq k$ directly implies $\mathbf{G}_k\mathbf{U}_k = \mathbf{0}$. This means that $\mathbf{U}_k$ must lie in the null-space of $\mathbf{G}_k$. Now, consider the product $\mathbf{G}_k\mathbf{B}_{k_0}$ = $\mathbf{A}_k\mathbf{\Sigma}_{k_p}\mathbf{B}_{k_1}^H\mathbf{B}_{k_0} = \mathbf{0}$ from (7.1). This proves the argument that $\mathbf{B}_{k_0}$ = null($\mathbf{G}_k$). So, we can express the optimal solution for $\mathbf{U}_k$ in the following form.

$$\mathbf{U}_k = \mathbf{B}_{k_0}\mathbf{T}_k \tag{7.2}$$

where $\mathbf{T}_k \in \mathbb{C}^{(M-L)\times R_k}$ is an arbitrary matrix with rank($\mathbf{T}_k$) = min($M-L, R_k$) = R_k (since $M-L = M-N(K-1) \geq N \geq R_k$). From (7.2), it is very simple now to see that the optimal transmit covariance matrix $\mathbf{Q}_k$ takes the following form after the elimination of BD precoding constraints.

$$\mathbf{Q_k} = \mathbf{U}_k\mathbf{U}_k^H = \mathbf{B}_{k_0}\mathbf{T}_k\mathbf{T}_k^H\mathbf{B}_{k_0}^H = \mathbf{B}_{k_0}\mathbf{S}_k\mathbf{B}_{k_0}^H \tag{7.3}$$

where $\mathbf{S}_k \in \mathbb{C}^{(M-L)\times(M-L)}$ with rank($\mathbf{S}_k$) = R_k is evidently a positive semi-definite matrix from it's structure ($\mathbf{S}_k = \mathbf{T}_k\mathbf{T}_k^H$). Lemma 1 thus follows.

B. Proof of Lemma 2

We prove Lemma 2 by contradiction. We know that if $\mathbf{Q}_{k_1} - \mathbf{Q}_{k_2}$ is positive definite, then $\mathbf{Q}_{k_1} \succ \mathbf{Q}_{k_2}$. Let us define $f(\mathbf{Q}_k) = w_k \log|\mathbf{I} + \mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H|$ with $w_k > 0$ as previously mentioned in our discussion. Now, $f(\mathbf{Q}_k)$ is monotonic with respect to $\mathbf{Q}_k$ i.e., $f(\mathbf{Q}_{k_1}) > f(\mathbf{Q}_{k_2})$ if $\mathbf{Q}_{k_1} \succ \mathbf{Q}_{k_2}$. Also, the total capacity is given by $\sum_{k=1}^{K} f(\mathbf{Q}_k)$. Suppose all the available power is not used by the optimal transmit covariance matrices set $\{\mathbf{Q}_k^*\}$ i.e., $\sum_{k=1}^{K} \text{Tr}(\mathbf{Q}_k^*) < P_{sum}$ then we can find a new set of transmit covaraince matrices $\{\mathbf{Q}_k\}$ that uses full sum power i.e., $\sum_{k=1}^{K} \text{Tr}(\mathbf{Q}_k) = P_{sum}$ satisfying $\mathbf{Q}_k \succ \mathbf{Q}_k^*, \forall k = 1, \cdots, K$. This means $f(\mathbf{Q}_k) > f(\mathbf{Q}_k^*), \forall k$ and we can write $\sum_{k=1}^{K} f(\mathbf{Q}_k) > \sum_{k=1}^{K} f(\mathbf{Q}_k^*)$. Obviously, it leads to a contradiction that the total capacity achieved by the optimal set $\{\mathbf{Q}_k^*\}$ is less than the total capacity achieved by the set of transmit covariance matrices $\{\mathbf{Q}_k\}$. Hence, we conclude that the optimal transmission strategy makes the sum power constraint tight given the sufficient conditions in (3.4). Lemma 2 thus follows.

C. Proof of Lemma 3

We prove Lemma 3 by contradiction. Suppose $\gamma = 0$ and there exist a number of strictly positive λ_a's and μ_i's such that rank($\mathbf{B}_{\gamma\lambda\mu}$) $< M - N(K-1)$. Since $L = N(K-1)$, we can write rank($\mathbf{B}_{\gamma\lambda\mu}$) $< (M - L)$. This also validates the fact that $\gamma = 0$ because otherwise the rank($\mathbf{B}_{\gamma\lambda\mu}$) $= M \geq (M - L)$. Let S denote the set consisting of the indices corresponding to all the strictly positive λ_a's and μ_i's i.e., if $\lambda_a > 0$ for any $a \in \{1, \cdots, A\}$ and $\mu_i > 0$ for any $i \in \{1, \cdots, M\}$, then $\{(a-1)M_B + 1, \cdots, aM_B\}, i \in S$. Note that the cardinality of the set S is denoted by $|S| =$ rank($\mathbf{B}_{\gamma\lambda\mu}$). Let $\mathbf{E}_k(S)$ and $\mathbf{F}_k(S^c)$ denote the matrices consisting of the rows in $\mathbf{B}_{k_0} \in \mathbb{C}^{M\times(M-L)}$ with the row indices given by the elements in S and S^c, respectively. Here, S^c denotes the complement of set S. Note that $|S| + |S^c| = M$ and $|S^c| > 0$ since $|S| = R_{\mathbf{B}_{\gamma\lambda\mu}} < (M - L) < M$. From the facts that $\mathbf{E}_k(S) \in \mathbb{C}^{R_{\mathbf{B}_{\gamma\lambda\mu}}} \times (M - L)$ and $R_{\mathbf{B}_{\gamma\lambda\mu}} < (M - L)$, we can say that the null-space of $\mathbf{E}_k(S)$ is not empty and hence we could find a unit vector $\mathbf{s}_k \in \mathbb{C}^{(M-L)\times 1}$ such that $\mathbf{E}_k(S)\mathbf{s}_k = \mathbf{0}$ and at the same time satisfying $\mathbf{F}_k(S^c)\mathbf{s}_k \neq \mathbf{0}$. Since $\gamma = 0$, $\mathbf{B}_{\gamma\lambda\mu} = \mathbf{B}_{\lambda\mu} \triangleq \mathbf{B}_\lambda + \mathbf{B}_\mu$. Further, we can deduce the following.

$$\mathbf{E}_k(S)\mathbf{s}_k = \mathbf{0} \Longleftrightarrow \mathbf{B}_{\lambda\mu}\mathbf{B}_{k_0}\mathbf{s}_k = \mathbf{0}$$

$$\mathbf{F}_k(S^c)\mathbf{s}_k \neq \mathbf{0} \Longleftrightarrow \mathbf{B}_{k_0}\mathbf{s}_k \neq \mathbf{0}$$

Denote $\mathbf{z}_k = \mathbf{B}_{k_0}\mathbf{s}_k$ and $\mathbf{H}_k\mathbf{z}_k = \mathbf{r}_k$. Note that the indices of the non-zero elements in $\mathbf{z}_k$ belong to S^c. Suppose we take the solution of (P4) as $\mathbf{S}_k^* = p(\mathbf{s}_k\mathbf{s}_k^H)$ with $p \geq 0$. Substituting this solution for $\mathbf{S}_k^*$ into the objective function of (P4) gives us the following.

$$\begin{aligned} & w_k \log|\mathbf{I} + \mathbf{H}_k\mathbf{B}_{k_0}\mathbf{S}_k^*\mathbf{B}_{k_0}^H\mathbf{H}_k^H| - \mathrm{Tr}(\mathbf{B}_{\lambda\mu}\mathbf{B}_{k_0}\mathbf{S}_k^*\mathbf{B}_{k_0}^H) \\ &= w_k \log|\mathbf{I} + p\mathbf{H}_k\mathbf{z}_k\mathbf{z}_k^H\mathbf{H}_k^H| \\ &= w_k \log|\mathbf{I} + p\mathbf{r}_k\mathbf{r}_k^H| \end{aligned} \tag{7.4}$$

Provided that $\mathbf{r}_k\mathbf{r}_k^H \neq \mathbf{0}$, as $p \to \infty$, the value in (7.4) becomes unbounded (which holds with probability one with independent channel realizations). Therefore, we conclude that our presumption that rank($\mathbf{B}_{\gamma\lambda\mu}$) = $R_{\mathbf{B}_{\gamma\lambda\mu}} < M - N(K-1)$ is false in order to have a bounded objective value for (P4). Lemma 3 thus follows.

D. Proof of Theorem 1

We have to find an optimal $\tilde{\mathbf{S}}_k^*$ that maximizes the objective function in (P5) denoted by $L(\tilde{\mathbf{S}}_k, \gamma, \{\lambda_a\}, \{\mu_i\})$ for a given set of $\gamma, \{\lambda_a\}, \{\mu_i\}$. Hence, we have to solve the equation below

$$\nabla_{\tilde{\mathbf{S}}_k} L(\tilde{\mathbf{S}}_k, \gamma, \{\lambda_a\}, \{\mu_i\})\Big|_{\tilde{\mathbf{S}}_k^*} = \mathbf{0} \tag{7.5}$$

The results from (Petersen *et al.*, 2008) on derivatives of traces and determinants of standard matrix forms have been very helpful to derive the next few steps.

$$\nabla_{\tilde{\mathbf{S}}_k} w_k \log|\mathbf{I} + \mathbf{H}_k\mathbf{B}_{k_0}(\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2}\tilde{\mathbf{S}}_k \times (\mathbf{B}_{k_0}^H\mathbf{B}_{\gamma\lambda\mu}\mathbf{B}_{k_0})^{-1/2}\mathbf{B}_{k_0}^H\mathbf{H}_k^H|\big|_{\tilde{\mathbf{S}}_k^*}$$

$$= \nabla_{\tilde{\mathbf{S}}_k} w_k \log |\mathbf{I} + \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H| \Big|_{\tilde{\mathbf{S}}_k^*} \tag{7.6}$$

$$\begin{aligned} &= w_k \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H (\mathbf{I} + \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k^* \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H)^{-1} \\ &\quad \times \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \end{aligned} \tag{7.7}$$

where (7.6) is obtained by substituting the regular SVD of $\mathbf{H}_k \mathbf{B}_{k_0} (\mathbf{B}_{k_0}^H \mathbf{B}_{\gamma\lambda\mu} \mathbf{B}_{k_0})^{-1/2}$ in the objective function of (P5) and (7.7) follows from the result in (Petersen *et al.*, 2008), $\nabla_{\mathbf{Q}} \log |\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H| = \mathbf{H}^H (\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H)^{-1} \mathbf{H}$. Next, we have from (Petersen *et al.*, 2008) that

$$\nabla_{\tilde{\mathbf{S}}_k} \mathrm{Tr}(\tilde{\mathbf{S}}_k) \Big|_{\tilde{\mathbf{S}}_k^*} = \mathbf{I} \tag{7.8}$$

From the equations (7.5, 7.7, 7.8) above, it follows that

$$\begin{aligned} &w_k \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H (\mathbf{I} + \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k^* \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H)^{-1} \\ &\quad \times \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H = \mathbf{I} \end{aligned} \tag{7.9}$$

We should note that $\hat{\mathbf{B}}_{k_1}^H \hat{\mathbf{B}}_{k_1} = \mathbf{I}$ which has already been proven in Appendix A of this paper. $\hat{\mathbf{A}}_k$ is a unitary matrix and hence $\hat{\mathbf{A}}_k \hat{\mathbf{A}}_k^H = \mathbf{I}$. Also, $\hat{\mathbf{\Sigma}}_{k_p}$ is a diagonal matrix with only positive singular values as it's diagonal entries and hence is invertible. Applying all the above arguments on (7.9), we get

$$(\mathbf{I} + \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k^* \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H)^{-1} = \frac{1}{w_k} \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p}^{-2} \hat{\mathbf{A}}_k^H \tag{7.10}$$

Since $\hat{\mathbf{\Sigma}}_{k_p}^{-2} \succ \mathbf{0}$ (a diagonal matrix with only positive entries in the diagonal), $\hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p}^{-2} \hat{\mathbf{A}}_k^H \succ \mathbf{0}$ and hence we conclude that $\hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p}^{-2} \hat{\mathbf{A}}_k^H$ is invertible. Therefore, we can write from (7.10) as below

$$\begin{aligned} \mathbf{I} + \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k^* \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H &= w_k (\hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p}^{-2} \hat{\mathbf{A}}_k^H)^{-1} \\ &= w_k \hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p}^{2} \hat{\mathbf{A}}_k^H \end{aligned} \tag{7.11}$$

where (7.11) follows from the fact that a unitary matrix $\hat{\mathbf{A}}_k$ satisfies $\hat{\mathbf{A}}_k^H = \hat{\mathbf{A}}_k^{-1}$.

Further by substituting $\mathbf{I} = \hat{\mathbf{A}}_k \hat{\mathbf{A}}_k^H$ in the above equation, we get

$$\hat{\mathbf{A}}_k \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k^* \hat{\mathbf{B}}_{k_1} \hat{\mathbf{\Sigma}}_{k_p} \hat{\mathbf{A}}_k^H = \hat{\mathbf{A}}_k (w_k \hat{\mathbf{\Sigma}}_{k_p}^2 - \mathbf{I}) \hat{\mathbf{A}}_k^H \tag{7.12}$$

Using the fact that $\hat{\mathbf{A}}_k^H \hat{\mathbf{A}}_k = \mathbf{I}$ and $\hat{\mathbf{\Sigma}}_{k_p}$ is invertible, we can continue from (7.12) to write as below

$$\hat{\mathbf{B}}_{k_1}^H \tilde{\mathbf{S}}_k^* \hat{\mathbf{B}}_{k_1} = w_k \mathbf{I} - \hat{\mathbf{\Sigma}}_{k_p}^{-2} \tag{7.13}$$

Now, we use the fact that $\hat{\mathbf{B}}_{k_1}^H \hat{\mathbf{B}}_{k_1} = \mathbf{I}$ once again to conclude that the optimal stucture of $\tilde{\mathbf{S}}_k^*$ that satisfies (7.14) is given by

$$\tilde{\mathbf{S}}_k^* = \hat{\mathbf{B}}_{k_1} (w_k \mathbf{I} - \hat{\mathbf{\Sigma}}_{k_p}^{-2}) \hat{\mathbf{B}}_{k_1}^H \tag{7.14}$$

However, there is a small catch here. We are looking for an optimal $\tilde{\mathbf{S}}_k^*$ that is positive semi-definite. $\tilde{\mathbf{S}}_k^* \succeq \mathbf{0}$ iff $w_k \mathbf{I} \succeq \hat{\mathbf{\Sigma}}_{k_p}^{-2}$. We also know that $(w_k \mathbf{I} - \hat{\mathbf{\Sigma}}_{k_p}^{-2})$ is a diagonal matrix. Therefore, to make this diagonal matrix positive semi-definite, we should allow only non-negative values into it's diagonal. Let us define $\Omega_k = \text{Diag}(\omega_{k_1}, \cdots, \omega_{k_N})$ with ω_{k_i} given by

$$\omega_{k_i} = \left(w_k - \frac{1}{\hat{\sigma}_{k_i}^2} \right)^+ \ \forall i = 1, \cdots, N$$

where $(y)^+ \triangleq \max(0, y)$. Therefore, the optimal solution for problem (P5) is given by

$$\tilde{\mathbf{S}}_k^* = \hat{\mathbf{B}}_{k_1} \mathbf{\Omega_k} \hat{\mathbf{B}}_{k_1}^H$$

Theorem 1 thus follows.

E. Proof of Theorem 3

We have to find an optimal $\bar{\omega}_{k_i}^*$ that maximizes the objective function in (P7) denoted by $L(\bar{\omega}_{k_i}, \gamma, \{\lambda_a\}, \{\mu_j\})$ for a given set of $\gamma, \{\lambda_a\}, \{\mu_j\}$. Hence, we have to solve the equation below

$$\frac{\partial}{d\bar{\omega}_{k_i}} L(\bar{\omega}_{k_i}, \gamma, \{\lambda_a\}, \{\mu_j\}) \Big|_{\bar{\omega}_{k_i}^*} = 0 \tag{7.15}$$

Therefore, the objective function of (P7) and the equation (7.15) gives us the following.

$$\frac{w_k(\sigma_{k_i}^{\perp})^2}{1+(\sigma_{k_i}^{\perp})^2\bar{\omega}_{k_i}^*} = f_{k_i} \tag{7.16}$$

From (7.16) and the condition that $\bar{\omega}_{k_i}^*$ must be non-negative, we get

$$\bar{\omega}_{k_i}^* = \Big(\frac{w_k}{f_{k_i}} - \frac{1}{(\sigma_{k_i}^{\perp})^2}\Big)^+$$

where $(y)^+ \triangleq \max(0, y)$. Theorem 3 thus follows.

F. Proof of Theorem 4

From equation (5.7), it is evident that we have to solve the following equation.

$$\nabla_{\mathbf{t}_{k_j}} \sum_{j=1}^{R} \|\mathbf{B}_{k_0}\mathbf{t}_{k_j} - \mathbf{v}_{k_j}\|^2 \Big|_{\mathbf{t}_{k_j}^{BD}} = \mathbf{0} \tag{7.17}$$

We know that

$$\|\mathbf{B}_{k_0}\mathbf{t}_{k_j} - \mathbf{v}_{k_j}\|^2 = (\mathbf{B}_{k_0}\mathbf{t}_{k_j} - \mathbf{v}_{k_j})^H(\mathbf{B}_{k_0}\mathbf{t}_{k_j} - \mathbf{v}_{k_j}) \tag{7.18}$$

From (7.17) and (7.18), we can write

$$(\mathbf{t}_{k_j}^{BD})^H \mathbf{B}_{k_0}^H \mathbf{B}_{k_0} = \mathbf{v}_{k_j}^H \mathbf{B}_{k_0} \tag{7.19}$$

Since $\mathbf{B}_{k_0}^H\mathbf{B}_{k_0} = \mathbf{I}$, (7.19) reduces down to $\mathbf{t}_{k_j}^{BD} = \mathbf{B}_{k_0}^H\mathbf{v}_{k_j}$ and hence we can conclude that

$$\mathbf{T}_k^{BD} = \mathbf{B}_{k_0}^H \mathbf{V}_k \tag{7.20}$$

The fact that the projection is given by $\mathbf{U}_k^{BD} = \mathbf{B}_{k_0}\mathbf{T}_k^{BD}$ and (7.20) suggest that

$$\mathbf{U}_k^{BD} = \mathbf{B}_{k_0}\mathbf{B}_{k_0}^H\mathbf{V}_k, \;\; \forall k = 1, \cdots, K.$$

Theorem 4 thus follows.

G. Proof of Theorem 5

We build up the solution $\mathbf{u}_{k_l}^{J-SB}$ from the step below

$$\mathbf{u}_{k_l}^* = \frac{\mathbf{v}_{k_l}}{1+\gamma^*+\lambda_a^*}, \quad l \in I(a). \tag{7.21}$$

Now, we have to find $(1+\gamma^*+\lambda_a^*), a = 1, \cdots, A$ to complete the solution. We consider two cases.

(i) $\lambda_a^* > 0$ for Base Station "a": From complementary slackness theorem, one of the KKT conditions (Boyd and Vandenberghe, 2004), we have that

$$\lambda_a^* \Big(\sum_{k=1}^{K} \sum_{l \in I(a)} \|\mathbf{u}_{k_l}^*\|^2 - \tilde{P}_a \Big) = 0 \tag{7.22}$$

From the condition $\lambda_a^* > 0$, (7.21) and (7.22), it is clearly evident that

$$\begin{aligned} \frac{1}{1+\gamma^*+\lambda_a^*} &= \xi_l, \quad l \in I(a), \quad \forall a : \lambda_a^* > 0. \\ &= \sqrt{\frac{\tilde{P}_a}{\sum_{k=1}^{K} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2}} \end{aligned} \tag{7.23}$$

Also, we should note that since $\lambda_a^* > 0$,

$$\frac{\tilde{P}_a}{\sum_{k=1}^{K} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2} < \frac{1}{(1+\gamma^*)^2}, \quad \forall a : \lambda_a^* > 0 \tag{7.24}$$

(ii) $\lambda_a^* = 0$ for Base Station "a": In this case, firstly the optimal $\mathbf{u}_{k_l}^*$ is given as

$$\mathbf{u}_{k_l}^* = \frac{\mathbf{v}_{k_l}}{1+\gamma^*}, \quad l \in I(a) : \lambda_a^* = 0. \tag{7.25}$$

From the Per-Base Station power constraints, we have

$$\sum_{k=1}^{K}\sum_{l\in I(a)}\|\mathbf{u}_{k_l}^*\|^2 \le \tilde{P}_a,\ \ \forall a=1,\cdots,A. \tag{7.26}$$

From (7.25) and (7.26), we can conclude that

$$\frac{1}{(1+\gamma^*)^2} \le \frac{\tilde{P}_a}{\sum_{k=1}^{K}\sum_{l\in I(a)}\|\mathbf{v}_{k_l}\|^2},\ \ \forall a:\lambda_a^*=0. \tag{7.27}$$

Let "m" number of Base Stations satisfy $\lambda_a^*>0$ amomg a total of A BSs. The rest of the "$A-m$" Base Stations satisfy $\lambda_a^*=0$.

Given the arrangement of Base Stations in the ordered set Δ and a careful observation of the two equations (7.24) and (7.27), BSs corresponding to first m elements in Δ will satisfy $\lambda_a^*>0$ and the BSs corresponding to the last $A-m$ elements in Δ will satisfy $\lambda_a^*=0$.

Moreover, the sufficient conditions in (3.4) suggest that $P_{sum}\le\sum_{a=1}^{A}\tilde{P}_a$. We have also proved that the optimal $\{\mathbf{u}_{k_l}^*\}$ utilises the full available sum power in Lemma 2. Therefore, we have

$$\sum_{k=1}^{K}\sum_{a=\Delta(1)}^{\Delta(m)}\sum_{l\in I(a)}\|\mathbf{u}_{k_l}^*\|^2+\sum_{k=1}^{K}\sum_{a=\Delta(m+1)}^{\Delta(A)}\sum_{l\in I(a)}\|\mathbf{u}_{k_l}^*\|^2 = P_{sum} \tag{7.28}$$

Since $\lambda_a^*>0$ for $a\in\{\Delta(1),\cdots,\Delta(m)\}$ and due to the virtue of complementary slackness condition in (7.22), we can conclude that

$$\sum_{k=1}^{K}\sum_{a=\Delta(1)}^{\Delta(m)}\sum_{l\in I(a)}\|\mathbf{u}_{k_l}^*\|^2=\sum_{a=\Delta(1)}^{\Delta(m)}\tilde{P}_a \tag{7.29}$$

From (7.28) and (7.29), we get

$$\frac{1}{1+\gamma^*} = \xi = \xi_l, \;\; l \in I(a) \;\; \forall a : \lambda_a^* = 0.$$
$$= \sqrt{\frac{P_{sum} - \sum_{a=\Delta(1)}^{\Delta(m)} \tilde{P}_a}{\sum_{k=1}^{K} \sum_{a=\Delta(m+1)}^{\Delta(A)} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2}} \tag{7.30}$$

Now, it is only left to obtain "m". Because of our ordering in the set Δ, we can argue that "m" is the only solution that satisfies the following two inequalities.

$$\frac{\tilde{P}_{\Delta(m)}}{\sum_{k=1}^{K} \sum_{l \in I(\Delta(m))} \|\mathbf{v}_{k_l}\|^2} < \frac{P_{sum} - \sum_{a=\Delta(1)}^{\Delta(m)} \tilde{P}_a}{\sum_{k=1}^{K} \sum_{a=\Delta(m+1)}^{\Delta(A)} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2} \tag{7.31}$$

$$\frac{P_{sum} - \sum_{a=\Delta(1)}^{\Delta(m)} \tilde{P}_a}{\sum_{k=1}^{K} \sum_{a=\Delta(m+1)}^{\Delta(A)} \sum_{l \in I(a)} \|\mathbf{v}_{k_l}\|^2} \leq \frac{\tilde{P}_{\Delta(m+1)}}{\sum_{k=1}^{K} \sum_{l \in I(\Delta(m+1))} \|\mathbf{v}_{k_l}\|^2} \tag{7.32}$$

Because of the ordering of the BSs in Δ, $\frac{\tilde{P}_{\Delta(i)}}{\sum_{k=1}^{K} \sum_{l \in I(\Delta(i))}}$ is non-decreasing in i. Therefore, "m" is the least element i in $\{0, 1, 2, \cdots, A-1\}$ that satisfies (7.32). "m" is also the highest element i in $\{0, 1, 2, \cdots, A-1\}$ that satisfies (7.31).

We specifically note that $m = A - 1$ satisfies (7.32) since we have already imposed the condition, $P_{sum} \leq \sum_{a=1}^{A} \tilde{P}_a$. Therefore, we conclude that there exists atleast one solution for "m" in $\{0, 1, 2, \cdots, A-1\}$.

addcontentslinetocchapterREFERENCES

REFERENCES

1. **Bauschke, H. H.** and **J. M. Borwein** (1994). DykstraâĂš s alternating projection algorithm for two sets. *Journal of Approximation Theory*, **79**(3), 418–443.

2. **Bland, R. G.**, **D. Goldfarb**, and **M. J. Todd** (1981). The ellipsoid method: A survey. *Operations research*, **29**(6), 1039–1091.

3. **Boccardi, F.** and **H. Huang**, Optimum power allocation for the mimo-bc zero-forcing precoder with per-antenna power constraints. *In 2006 40th Annual Conference on Information Sciences and Systems*. IEEE, 2006.

4. **Boyd, S.** and **C. Barratt** (2008). Ellipsoid method. *Notes for EE364B, Stanford University*, **2008**.

5. **Boyd, S.** and **L. Vandenberghe**, *Convex optimization*. Cambridge university press, 2004.

6. **Caire, G.** and **S. Shamai** (2003). On the achievable throughput of a multiantenna gaussian broadcast channel. *IEEE Transactions on Information Theory*, **49**(7), 1691–1706.

7. **Chaluvadi, R.**, **S. S. Nair**, and **S. K. Bhashyam** (2018). Optimal multi-antenna transmission with multiple power constraints. *Submitted to IEEE Transactions on Wireless Communications*.

8. **Choi, L.-U.** and **R. D. Murch** (2004). A transmit preprocessing technique for multiuser mimo systems using a decomposition approach. *IEEE Transactions on Wireless Communications*, **3**(1), 20–24.

9. **Golub, G. H.** and **C. F. Van Loan**, *Matrix computations*, volume 3. JHU press, 2012.

10. **Grant, M.** and **S. Boyd** (2014). Cvx: Matlab software for disciplined convex programming, version 2.1.

11. **Jing, S.**, **N. David**, **J. B. Soriaga**, **J. Hou**, **J. E. Smee**, and **R. Padovani**, Downlink macro-diversity in cellular networks. *In 2007 IEEE International Symposium on Information Theory*. IEEE, 2007.

12. **Karakayali, K.**, **R. Yates**, **G. Foschini**, and **R. Valenzuela**, Optimum zero-forcing beamforming with per-antenna power constraints. *In 2007 IEEE International Symposium on Information Theory*. IEEE, 2007.

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

14. **Kaviani, S.** and **W. A. Krzymien**, Sum rate maximization of mimo broadcast channels with coordination of base stations. *In 2008 IEEE Wireless Communications and Networking Conference*. IEEE, 2008.

15. **Le Cao, P.** and **T. J. Oechtering**, Optimal transmit strategy for mimo channels with joint sum and per-antenna power constraints. *In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017.

16. **Liu, W.**, **S. X. Ng**, and **L. Hanzo**, Multicell cooperation based svd assisted multi-user mimo transmission. *In VTC Spring 2009-IEEE 69th Vehicular Technology Conference*. IEEE, 2009.

17. **Ng, B. L.**, **J. S. Evans**, **S. V. Hanly**, and **D. Aktas** (2008). Distributed downlink beamforming with cooperative base stations. *IEEE Transactions on Information Theory*, **54**(12), 5491–5499.

18. **Pan, Z.**, **K.-K. Wong**, and **T.-S. Ng** (2004). Generalized multiuser orthogonal space-division multiplexing. *IEEE Transactions on Wireless Communications*.

19. **Park, D.**, **A. Kyrillidis**, **S. Bhojanapalli**, **C. Caramanis**, and **S. Sanghavi** (2016). Provable burer-monteiro factorization for a class of norm-constrained matrix problems. *arXiv preprint arXiv:1606.01316*.

20. **Peel, C. B.**, **B. M. Hochwald**, and **A. L. Swindlehurst** (2005). A vector-perturbation technique for near-capacity multiantenna multiuser communication-part i: channel inversion and regularization. *IEEE Transactions on Communications*, **53**(1), 195–202.

21. **Petersen, K. B.**, **M. S. Pedersen**, *et al.* (2008). The matrix cookbook. *Technical University of Denmark*, **7**(15), 510.

22. **Pi, Z.**, Optimal transmitter beamforming with per-antenna power constraints. *In 2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012.

23. **Raghavan, V.**, **J. Cezanne**, **S. Subramanian**, **A. Sampath**, and **O. Koymen** (2016). Beamforming tradeoffs for initial ue discovery in millimeter-wave mimo systems. *IEEE Journal of Selected Topics in Signal Processing*, **10**(3), 543–559.

24. **Rashid-Farrokhi, F.**, **K. R. Liu**, and **L. Tassiulas** (1998). Transmit beamforming and power control for cellular wireless systems. *IEEE Journal on Selected Areas in Communications*, **16**(8), 1437–1450.

25. **Schubert, M.** and **H. Boche** (2004). Solution of the multiuser downlink beamforming problem with individual sinr constraints. *IEEE Transactions on Vehicular Technology*, **53**(1), 18–28.

26. **Shamai, S.** and **B. M. Zaidel**, Enhancing the cellular downlink capacity via co-processing at the transmitting end. *In IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No. 01CH37202)*, volume 3. IEEE, 2001.

27. **Shen, Z.**, **R. Chen**, **J. G. Andrews**, **R. W. Heath**, and **B. L. Evans**, Low complexity user selection algorithms for multiuser mimo systems with block diagonalization. *In Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005.*. IEEE, 2005.

28. **Somekh, O.**, **B. M. Zaidel**, and **S. Shamai** (2007). Sum rate characterization of joint multiple cell-site processing. *IEEE Transactions on Information Theory*, **53**(12), 4473–4497.

29. **Spencer, Q. H.**, **A. L. Swindlehurst**, and **M. Haardt** (2004). Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels. *IEEE transactions on signal processing*, **52**(2), 461–471.

30. **Stojnic, M.**, **H. Vikalo**, and **B. Hassibi** (2006). Rate maximization in multi-antenna broadcast channels with linear preprocessing. *IEEE Transactions on Wireless Communications*, **5**(9), 2338–2342.

31. **Sun, S.**, **T. S. Rappaport**, **R. W. Heath**, **A. Nix**, and **S. Rangan** (2014). Mimo for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both? *IEEE Communications Magazine*, **52**(12), 110–121.

32. **Telatar, E.** (1999). Capacity of multi-antenna gaussian channels. *European transactions on telecommunications*, **10**(6), 585–595.

33. **Torkildson, E.**, **U. Madhow**, and **M. Rodwell** (2011). Indoor millimeter wave mimo: Feasibility and performance. *IEEE Transactions on Wireless Communications*, **10**(12), 4150–4160.

34. **Tuninetti, D.**, On the capacity of the awgn mimo channel under per-antenna power constraints. *In 2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014.

35. **Vu, M.** (2011*a*). The capacity of mimo channels with per-antenna power constraint. *arXiv preprint arXiv:1106.5039*.

36. **Vu, M.** (2011*b*). Miso capacity with per-antenna power constraint. *Communications, IEEE Transactions on*, **59**, 1268 – 1274.

37. **Weingarten, H.**, **Y. Steinberg**, and **S. S. Shamai** (2006). The capacity region of the gaussian multiple-input multiple-output broadcast channel. *IEEE transactions on information theory*, **52**(9), 3936–3964.

38. **Wiesel, A.**, **Y. C. Eldar**, and **S. Shamai** (2006). Linear precoding via conic optimization for fixed mimo receivers. *IEEE Transactions on Signal Processing*, **54**(1), 161–176.

39. **Wong, K.-K.**, **R. D. Murch**, and **K. B. Letaief** (2003). A joint-channel diagonalization for multiuser mimo antenna systems. *IEEE Transactions on Wireless Communications*, **2**(4), 773–786.

40. **Yoo, T.** and **A. Goldsmith** (2006). On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. *IEEE Journal on selected areas in communications*, **24**(3), 528–541.

41. **Zhang, H.** and **H. Dai** (2004). Cochannel interference mitigation and cooperative processing in downlink multicell multiuser mimo networks. *EURASIP Journal on Wireless Communications and Networking*, **2004**(2), 222–235.

42. **Zhang, J.**, **R. Chen**, **J. G. Andrews**, **A. Ghosh**, and **R. W. Heath Jr** (2008). Networked mimo with clustered linear precoding. *arXiv preprint arXiv:0808.3971*.

43. **Zhang, R.** (2010). Cooperative multi-cell block diagonalization with per-base-station power constraints. *IEEE Journal on Selected Areas in Communications*, **28**(9), 1435–1445.