

UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

LÍGIA MARIA CARVALHO SOUSA CORDEIRO

STATISTICAL MODELLING AND PRECODER DESIGN
FOR COORDINATED MIMO WIRELESS COMMUNICATIONS SYSTEMS

FORTALEZA
2011



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

Statistical Modelling and Precoder Design for Coordinated MIMO Wireless Communications Systems

Doctor of Science Thesis

Lígia Maria Carvalho Sousa Cordeiro

Advisor

Prof. Dr. Charles Casimiro Cavalcante

FORTALEZA
2011

LÍGIA MARIA CARVALHO SOUSA CORDEIRO

STATISTICAL MODELLING AND PRECODER DESIGN
FOR COORDINATED MIMO WIRELESS COMMUNICATIONS SYSTEMS

Tese submetida à Coordenação do Programa de Pós-graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como parte dos requisitos para obtenção do grau de Doutor em Engenharia de Teleinformática.

Área de concentração: Sinais e Sistemas

Orientador: Prof. Dr. Charles Casimiro
Cavalcante

FORTALEZA

2011

C819s Cordeiro, Lígia Maria Carvalho Sousa
Statistical modelling and precoder design for coordinated MIMO
wireless communications systems / Lígia Maria Carvalho Sousa Cordeiro. -
Fortaleza, 2011.
83f. ; il.; enc.

Orientador: Prof. Dr. Charles Casimiro Cavalcante
Área de concentração: Sinais e sistemas
Tese (Doutorado) - Universidade Federal do Ceará, Centro de
Tecnologia, Departamento de Engenharia de Teleinformática, Fortaleza,
2011.

1. Sistemas de comunicação sem fio. 2. Telecomunicações 3.
Processamento de sinais. I. Cavalcante, Charles Casimiro. II. Universidade
Federal do Ceará – Programa de Pós-Graduação em Engenharia de
Teleinformática. III. Título.

CDD 621.38

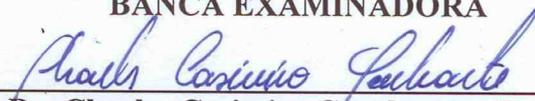
LÍGIA MARIA CARVALHO SOUSA

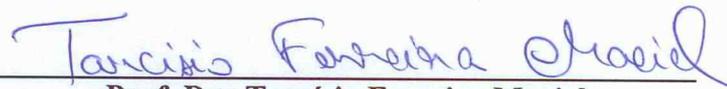
STATISTICAL MODELLING AND PRECODER DESIGN FOR COORDINATED
MIMO WIRELESS COMMUNICATION SYSTEMS

Tese submetida à Coordenação do Curso de Pós-Graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Doutor em Engenharia de Teleinformática. Área de concentração Sinais e Sistemas.

Aprovada em 06/05/2011.

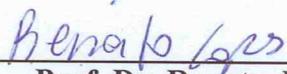
BANCA EXAMINADORA

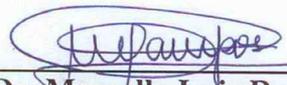

Prof. Dr. Charles Casimiro Cavalcante (Orientador)
Universidade Federal do Ceará - UFC


Prof. Dr. Tarcísio Ferreira Maciel
Universidade Federal do Ceará - UFC


Prof. Dr. Yuri Carvalho Barbosa Silva
Universidade Federal do Ceará - UFC


Prof. Dr. Walter da Cruz Freitas Júnior
Universidade Federal do Ceará - UFC


Prof. Dr. Renato da Rocha Lopes
Universidade Estadual de Campinas - Unicamp


Prof. Dr. Marcello Luiz Rodrigues de Campos
Universidade Federal do Rio de Janeiro - UFRJ

*To my parents João and Margarida for the unconditional love,
my brother Lívio for teaching and support and
my husband Eduardo for love, encouragement and companionship.*

Acknowledgements

First and foremost, I would like to thank the one who is above all of us, the omnipresent God, for answering my prayers and for being always by my side. I thank You God for showing me inner peace and for all my blessings. I was not able to complete this thesis on my own, but only with the strength of Jesus Christ in me, the Almighty.

This thesis would not have been possible without the guidance and the help of several people who in one way or another contributed and extended their assistance in the preparation and conclusion of this work.

I would like to thank my parents João and Margarida, who are the responsible for my existence and who I am today. I thank for their inseparable support and prayers. They always give me their caring and their endless love and I am really appreciative of all their dedication.

I am very grateful to my brother Lívio Sousa for his encouragement and support. He is a great example for me with his patience and his brilliant intelligence.

My lovely thanks to my husband Eduardo Cordeiro for his unconditional love, support and dedication. Although we do not share the passion for the academic area, he understood the value of this thesis for me and was very encouraging and comprehensive whenever I needed. Lovely, we complete ourselves and thanks for bringing so much sincere love and happiness to my life.

I would like to show my deepest gratitude to my advisor and friend Prof. Charles Cavalcante for his brilliant supervision and his continual guidance and support so far. Since even before the end of my master, he has advised me either on my research or on my career. Without his wide technical insight, understanding and encouragement, this thesis would have never been possible. I am indebted to him more than he knows. Thanks for the confidence that you have placed in me.

I wish to express my warm and sincere thanks to Prof. Tarcísio Maciel, who gave me important support during a crucial phase of this thesis. His large technical knowledge guided me during the research phase and his readiness was very important during both the simulations and writing phases. Thanks for all comments and discussions, even in very busy times.

I owe my sincere gratitude to Prof. João César Moura Mota for the great help in the master-doctorate transient period and for providing valuable pieces of advice on occasions I spoke with him. Thanks very much for your support and encouragement.

I would like to express my appreciation to Prof. Francisco Rodrigo Cavalcanti for giving me the opportunity to work at GTEL and get involved in diverse interesting projects. At GTEL I have had the opportunity to know and interact with many other researchers, which have enriched and made my experience there unique. Throughout these years there, I have

learned and grown so much. Extending my appreciation, I would like to thank my colleagues and friends at GTEL for the excellent and truly enjoyable ambiance. Special thanks to Darlan Moreira and Igor Guerreiro for their friendship and for helping me in both technical and non-technical issues. Another person from GTEL that I am very grateful is my colleague Alysson Guimarães, who gave me valuable insights into convex optimization tools and on how to apply them in this work. His mathematical support have always present whenever I needed and I appreciated so much his availability.

My personal friends also had an important role in my academic life. My warmest thanks to my dear friends Mônica Reis, Eliza Pinto, Elisiane Pinto, Rafael Barbosa (“Babinha”) and to my family in-law. Thanks for the fellowship and for your presence in my life.

Further thanks to the *Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico* (FUNCAP) for giving me an initial financial support through Doctor’s scholarship during the two first years.

Finally, I would like to thanks everybody who cheered for my success and sent always positive vibrations.

*“Tudo é do Pai
Toda honra e toda glória
É dEle a vitória
Alcançada em minha vida.”*

Resumo

Transmissão/recepção coordenada de múltiplos pontos (do inglês *Coordinated Multipoint Transmission/Reception-coordinated multipoint (CoMP)*) é uma técnica que visa aumentar a vazão na célula dos sistemas sem-fio das próximas gerações. O processamento conjunto (do inglês *Joint Processing-joint processing (JP)*) é um tipo de sistema CoMP que permite aumentar a performance dos sistemas empregando algoritmos de precodificação baseados em informação do canal no transmissor (do inglês *Channel State Information at the Transmitter-channel state information at the transmitter (CSIT)*).

Atualmente, muitas pesquisas têm focado na redução do *feedback* e na otimização de técnicas de precodificação com CSIT parcial. Nossa primeira proposta consiste em generalizar um modelo de canal estatístico para um sistema CoMP MIMO múltiusuário que leva em consideração a variação temporal do canal. Derivado de uma medição de canal desatualizada e das estatísticas do canal (média e covariância), o modelo de CSIT proposto consiste de uma estimativa do canal e da covariância do erro de estimação, os quais agem como média e covariância efetiva do canal, respectivamente. Ambos os parâmetros dependem do fator de correlação temporal do canal, o qual indica a qualidade do CSIT.

Em seguida, considerando que o transmissor tem a média do canal e a correlação espacial entre as antenas como forma de CSIT, nós propomos dois algoritmos que maximizam a aproximação de primeira e segunda-ordem da taxa soma média do sistema CoMP multiple-input multiple-output (MIMO) multiusuário, respectivamente. Os algoritmos propostos são computacionalmente simples, reduzem as informações de *feedback* e têm rápida convergência.

Os resultados de simulação mostram que os algoritmos propostos são quase-ótimos comparados com a técnica de *water-filling* iterativo (caso ótimo), e que apresentam perdas de taxa soma média moderadas para baixos valores de SNR e desprezíveis para altos valores de SNR.

Palavras-chave: Estações base coordenadas, sistemas MIMO multiusuário, Técnicas de precodificação, *Feedback* Limitado.

Abstract

Coordinated multipoint (CoMP) transmission/reception is a candidate technique for increasing cell-average and cell-edge throughputs in next-generation wireless systems. Joint processing (JP) is a branch of CoMP systems which can enhance the systems' performance, mainly by employing precoding algorithms based on channel state information at the transmitter (CSIT).

Many research efforts focus on reducing feedback and optimizing precoding with partial CSIT. Our first proposal is to generalize a statistical channel model for a multi-user multiple-input multiple-output (MU-MIMO) CoMP system, which takes into account the channel time-variation. Derived from a potentially outdated channel measurement and from the channel statistics (mean and covariance), this proposed CSIT model consists of a channel estimate and its error covariance, which act as the effective channel mean and covariance, respectively. Both parameters depend on a temporal correlation factor, indicating the CSIT quality.

Secondly, considering that the transmitter has the channel mean and the spatial correlation among the antennas as partial CSIT, we propose two algorithms to maximize the first and second-order approximations of the ergodic sum rate of a MU-MIMO CoMP system, respectively. The proposed algorithms are computationally simple, highly reduce feedback overheads, and have fast convergence.

Simulation results show that the proposed algorithms are near-optimal compared to the iterative water-filling (optimal) case and present only moderate and negligible sum rate losses for low and high SNR values, respectively.

Keywords: Coordinated base-station (BS)s, MU-MIMO systems, Precoding techniques, Limited feedback.

Contents

Acknowledgements	vii
Resumo	x
Abstract	xi
List of Figures	xiv
Nomenclature	xvi
1 Introduction	1
1.1 Background and Motivation	1
1.2 Thesis Contributions	2
1.2.1 Statistical Channel Model	2
1.2.2 Approximation of the multi-user MIMO (MU-MIMO) CoMP Ergodic Sum-Rate	2
1.2.3 Convex Optimization of the Approximated Ergodic Sum-Rate	3
1.3 Scientific Production and Contributions	3
1.4 Document Organization	4
2 Multi-User Multiple-Input Multiple-Output (MU-MIMO) Coordinated Multipoint (CoMP) Wireless Channel	5
2.1 The Wireless Channel	5
2.1.1 Path Loss	5
2.1.2 Shadowing	6
2.1.3 Fast Fading	6
2.1.4 Channel Selectivity	7
2.2 Multiple-input Multiple-output (MIMO) Channel	8
2.2.1 Spatial MIMO Channel Modeling	8
2.3 Multi-user (MU) MIMO Channel	10
2.3.1 Downlink Multi-user MIMO Channel	11
2.3.2 Uplink Multi-user MIMO Channel	11
2.3.3 Capacity Region of the Uplink MU-MIMO Channel	12
2.3.4 Capacity Region of the Downlink MU-MIMO Channel	14
2.3.5 Duality Between Downlink MU-MIMO Channel and Uplink MU-MIMO Channels	15
2.4 Coordinated Multipoint (CoMP) Multiuser MIMO Channels	19

2.5	Summary	22
3	Multi-user MIMO CoMP Channel Model	23
3.1	Proposed Statistical MU-MIMO CoMP Channel Model	24
3.1.1	Channel Auto-covariance	24
3.1.2	Channel Estimation at the Transmitter and Proposed Statistical CSIT	25
3.2	Simulations with the Proposed Statistical MU-MIMO CoMP Channel Model	30
3.2.1	Cell Scenario, Channel Model and Simulation Steps	32
3.2.2	Simulation Parameters and Performance Metrics	32
3.2.3	Simulation Results	34
3.3	Summary	35
4	Statistical Transmit Scheme for MU-MIMO CoMP Systems	39
4.1	Duality and Multiuser MIMO Downlink CoMP Capacity	39
4.2	Approximations of the MU-MIMO CoMP Ergodic Sum-Rate	42
4.2.1	Proposed Input Covariance Matrices using the First-Order Approximation of the Sum Rate - FOIC Approach	43
4.2.2	Proposed Input Covariance Matrices using the Second-Order Approximation of the Sum Rate - SOIC Approach	44
4.3	Input Covariance Matrix Maximizing the Second-Order Approximation of the Ergodic Sum Rate	46
4.3.1	Convergence Analysis	46
4.3.2	Convex Optimization of the Second-Order Approximation of the Ergodic Capacity	48
4.4	Simulations and Results	51
4.4.1	Scenario, Channel Model and Main Simulation Parameters	51
4.4.2	A Multi-user Analysis of the Proposed Transmit Scheme	51
4.4.3	A Single-user Analysis of the Proposed Transmit Scheme	52
4.4.4	A Single-user Analysis of the Proposed Transmit Scheme with the Proposed Channel Model	54
4.4.5	Analysis of the Phase Mismatch between the Uplink and Downlink Input Covariance Matrices	54
4.5	Summary	56
5	Conclusions and Perspectives	69
Appendix A Mathematical Manipulations and Proofs		71
A.1	Proof of Equation (4.7)	71
A.2	Proof of $\mathbb{E} \left\{ \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right\}$	72
A.3	Proof of $\mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{B}) \} = 0$ and $\mathbb{E} \{ \text{tr}(\mathbf{B}^H \mathbf{B}^H) \} = 0$	72
A.4	Proof of $\mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{C}) \} = 0$ and $\mathbb{E} \{ \text{tr}(\mathbf{B}^H \mathbf{C}) \} = 0$	73
A.5	Evaluation of $\mathbb{E} \{ \text{tr}(\mathbf{A}\mathbf{C}) \}$ and $\mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{B}^H) \}$	75
A.6	Evaluation of $\mathbb{E} \{ \text{tr}(\mathbf{C}^2) \}$	76
Bibliography		79

List of Figures

2.1	System model of the downlink MU-MIMO channel.	11
2.2	System model of the uplink MU-MIMO channel.	12
2.3	Capacity region for the uplink MU-MIMO channel when the receivers have a single antenna.	13
2.4	Capacity region for the downlink MU-MIMO channel considering two users. . . .	15
2.5	System model of the downlink MU-MIMO (left) and uplink MU-MIMO (right) channels.	16
2.6	MU-MIMO CoMP system model with $N_b = K = 3$	20
3.1	The transmitter in a system with precoder.	31
3.2	Modelling of the proposed channel model.	33
3.3	SE curves of linear precoding techniques using perfect and the proposed channel model with $N_S = 5$ and $N_t = N_r = 2$	36
3.4	SE curves of MMSE precoding technique using the proposed channel model and varying the number of symbols in each block N_S for $N_t = N_r = 2$	37
3.5	SE curves of MMSE precoding technique using the proposed channel model and varying the number of symbols in each block N_S for $N_t = N_r = 4$	38
4.1	System model of the downlink MU-MIMO CoMP (left) and uplink MU-MIMO CoMP (right) channels.	40
4.2	Flow chart of the proposed algorithm.	51
4.3	Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 2$ and $\kappa = 0$	57
4.4	Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 2$ and $\kappa = 3$	58
4.5	Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 4$ and $\kappa = 0$	59
4.6	Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 4$ and $\kappa = 3$	60
4.7	Analysis of the convergence of the SOIC approach. $N_t = N_r = 2$ and $\kappa = 3$	61
4.8	Ergodic rate curves of the single-user scenario considering the Kronecker channel model and $N_t = N_r = 2$	62
4.9	Ergodic rate curves of the single-user scenario considering the Kronecker channel model, $N_t = 2$ and $N_r = 4$	63
4.10	Comparison of the ergodic rate curves obtained by SOIC approach considering the Kronecker channel model and the proposed one. Scenario $N_t = N_r = 2$	64

4.11 Ergodic rate curves of the single-user scenario considering the proposed channel model, $N_t = N_r = 2$	65
4.12 Ergodic rate curves of the single-user scenario considering the proposed channel model, $N_t = 2$ and $N_r = 4$	66
4.13 Comparison of the ergodic sum rate curves obtained by SOIC approach considering the Kronecker channel model and a phase mismatch between uplink and downlink channels. Scenario $N_t = N_r = 2$	67
4.14 Comparison of the ergodic sum rate curves obtained by SOIC approach considering the proposed channel model and a phase mismatch between uplink and downlink channels. Scenario $N_t = N_r = 2$	68

Nomenclature

In this section, we summarize the conventional notation of this thesis. Firtly, we present a list of acronyms, followed by an overview of the notation of more general nature. We conclude with the specific notation for this thesis.

Acronyms

The abbreviations and acronyms used throughout this thesis are listed here. The meaning of each abbreviation or acronym is indicated once, when it first appears in the text.

BC	broadcast channel
BS	base-station
CoMP	coordinated multipoint
CS/CB	coordinated scheduling/coordinated beamforming
CSI	channel state information
CSIT	channel state information at the transmitter
DPC	dirty-paper coding
FUNCAP	<i>Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico</i>
i.i.d.	independent identically distributed
ICI	inter-channel interference
JP	joint processing
LOS	line-of-sight
MAC	multiple access channel
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
ML	maximum likelihood
MMSE	minimum mean-square error
MU	multi-user
MU-MIMO	multi-user MIMO

NLOS	non-line-of-sight
PDF	probability density function
QoS	quality of service
SDMA	space-division multiple access
SIC	successive interference cancellation
SIMO	single-input multiple-output
SNR	signal-to-noise ratio
SVD	singular value decomposition
ZMCSG	zero-mean circularly symmetric complex Gaussian
ZF	zero-forcing

Notation

The following notation is used throughout this thesis. We use boldface letters to denote matrices and vectors and italic ones for scalars. Other notational conventions are summarized as follows:

$\mathbb{E}\{ \cdot \}$	- The expectation operator
\mathbb{C}	- The set of complex numbers
\mathbb{R}	- The set of real numbers
$\mathcal{CN}(\mathbf{x}, \mathbf{X})$	- The circularly symmetric complex Gaussian distribution with mean \mathbf{x} and covariance matrix \mathbf{X}
$ x $	- The absolute value of a scalar x
$ \mathbf{X} $	- The determinant of matrix \mathbf{X}
$\ \mathbf{X}\ $	- The Euclidean norm of matrix \mathbf{X}
$\ \mathbf{X}\ _F$	- The Frobenius norm of matrix \mathbf{X}
$(\cdot)^*$	- The complex conjugate operator
$(\cdot)^T$	- The transpose operator
$(\cdot)^H$	- The complex conjugate transpose (Hermitian) operator
$(\mathbf{X})^{-1}$	- The inverse of matrix \mathbf{X}
$\mathbf{I}_{X \times X}$	- The identity matrix of dimensions $X \times X$
$\text{tr}(\mathbf{X})$	- The trace of matrix \mathbf{X}
$\text{vec}(\mathbf{X})$	- The vector obtained by stacking the columns of \mathbf{X}
\otimes	- The Kronecker product
$\exp(\cdot)$	- The exponential function
$\log(\cdot)$	- The base 2 logarithm

Specific Notation of the Thesis

We summarize here the symbols and notation that are commonly used in this thesis.

κ	- Ricean Factor
f_d	- Doppler spread
P_b	- Maximum transmit power constraint of base-station (BS) b
K	- Number of users
N_b	- Number of BSs
N_t	- Number of transmit antennas in each BS
N_r	- Number of receive antennas in each user equipment
$\rho[\tau]$	- Temporal auto-correlation of the channel at time delay τ
T_S	- Symbol time interval
N_S	- Number of symbols inside each simulation block
η	- Additive zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise vector
η_k	- Additive ZMCSCG noise vector at the user k of dimensions $N_r \times 1$
σ_η^2	- Variance of the additive ZMCSCG noise vector η
$\mathbf{H}_{k,b}$	- Channel matrix from BS b to user k of dimensions $N_r \times N_t$
\mathbf{H}_k	- Channel matrix of user k
$\bar{\mathbf{H}}_k$	- Channel mean matrix to user k
$\hat{\mathbf{H}}_k$	- Channel estimation matrix to user k
\mathbf{H}_w	- Channel matrix represented by a spatially white (independent identically distributed (i.i.d.)) ZMCSCG matrix with unit variance
$\mathbf{H}_k[0]$	- Channel matrix at initial time 0
\mathbf{R}_{t_k}	- Transmit covariance matrix per each user k
\mathbf{R}_{r_k}	- Receive covariance matrix per each user k
$\mathbf{R}_k[0]$	- Covariance matrix of the channel per each user k
$\mathbf{R}_k[n]$	- Channel auto-covariance matrix per each user k at time n
$\tilde{\mathbf{T}}_{k,b}$	- Precoding matrix from BS b to user k without considering any power loading
$\mathbf{T}_{k,b}$	- Precoding matrix from BS b for user k
$\tilde{\mathbf{T}}_k$	- Precoding matrix for user k without considering any power loading
$\mathbf{x}_{k,b}$	- Transmitted signal from BS b to user k
\mathbf{x}_k	- Transmitted signal for user k
$\mathbf{\Omega}_k$	- Covariance matrix of transmitted signal for user k , \mathbf{x}_k , in the downlink model
\mathbf{Q}_k	- Covariance matrix of transmitted signal from user k , \mathbf{x}_k , in the uplink model
R_k	- The rate that can be achieved for the user k
$R_{k,DL}$	- The rate that can be achieved for the user k in the downlink
$R_{k,UL}$	- The rate that can be achieved for the user k in the uplink
\mathbf{E}_k	- Estimation error matrix of the channel of user k
\mathbf{R}_{e_k}	- Covariance matrix of the channel estimation error of user k
Θ	- Power loading matrix

It is valuable to state that an matrix with the index k (for example \mathbf{X}_k) can mean that this matrix is considered from all N_b BSs to user k in the multi-user (MU) CoMP case; or that is considered from BS b to user k for MU case, in which the index b was omitted for simplicity. The used representation will be explained properly when it appears in the text.

Chapter

1

Introduction

1.1 Background and Motivation

The wireless communication area has required an enormous demand for higher data rates and enhanced quality of service (QoS) during the last decade under various scenarios. New types of services, such as streaming multimedia, video telephony and Internet, demanding different QoS requirements are nowadays offered to customers. Therefore, future wireless communication systems should be flexible and adaptive to accommodate these various services. For this reason, the system should be robust to the influence of fading, interference and hardware impairments.

Since the transmit power and the bandwidth of the wireless system are limited, the required advances in these systems can be achieved through the use of multiple antennas at both the transmitter and the receiver. The multiple-input multiple-output (MIMO) technology has attracted much attention in wireless communications, because it offers significant increase in transmission rate which has a linear growth with the minimum of the numbers of antennas at each end [1, 2]. Moreover, MIMO also enhances the link reliability and improves coverage without additional bandwidth or transmit power [3]. All these features are achieved when the channel exhibits rich scattering and the channel state information (CSI) can be accurately known. These initial promises of incredible gains resulted in large research activities to extend the MIMO concepts to the multi-user (MU) systems.

While single-user MIMO only considers access to multiple antennas that are physically connected to a specific individual terminal, multi-user MIMO (MU-MIMO) is a set of advanced MIMO technologies that exploits the availability of multiple independent terminals in order to enhance the communication capabilities of each individual terminal. MU-MIMO can be seen as an extension of the concept of space-division multiple access (SDMA) which allows a terminal to transmit (or receive) signal to (or from) multiple users in the same band, simultaneously.

The coordinated multipoint (CoMP) transmission/reception is a tool to improve the coverage of high data rates, the cell-edge throughput and/or to increase system throughput. It is characterized by the dynamic coordination among multiple geographically separated transmission points [4].

The benefits of MU-MIMO CoMP systems are enhanced when the transmitter exploits CSI to process, in an intelligent way, the signal before transmission. This can be accomplished by precoding techniques, which often rely on the assumption that the transmitter knows perfectly the MIMO channel matrix [5, 6]. However, this may not be realistic in many

practical scenarios and considering partial availability of channel state information at the transmitter (CSIT) in MIMO systems becomes an important issue. This assumption might have a significant impact on the spectral efficiency that can be reliably obtained by the system.

Some limited feedback multiuser MIMO schemes let each user quantize a function of the channel coefficients and feed this information back to the transmitter [7, 8]. Problems occur, e.g., when user signals can not be perfectly orthogonalized by precoding due to channel quantization errors. In order to avoid this, some schemes have been proposed which select, at the receiver, a quantized precoder from a codebook. Then, only the precoder index is fed back to the transmitter [9, 10]. However, designing precoder codebooks is a difficult task, which must take into account the statistical characteristics of the channel. Other approaches focus on feeding back the mean [11] or the covariance matrix [12] of the channel, which convey important information about the slow fading and the mean spatial separability of the users, being a non-complex and effective form of CSIT. Moreover, in a statistical feedback scheme, better results are obtained if the employed precoder technique exploits the advantages of the statistical CSIT, i.e., if the used precoder is designed taking into account the availability of only the channel statistics.

This thesis is devoted to study the availability of statistical channel knowledge at the transmitter and to design a precoder technique that, using this statistical CSIT, maximizes some function of the ergodic sum rate in a MU-MIMO CoMP scenario.

1.2 Thesis Contributions

This section summarizes the contribution of this thesis, which is divided into 3 parts: proposing a statistical channel model; approximating the ergodic sum-rate of the considered scenario; optimizing the obtained ergodic sum-rate approximation and obtaining near-optimal input covariance matrices.

1.2.1 Statistical Channel Model

The time-variation of the channel degrades the accuracy of the channel information obtained at the transmitter due to the delay involved in the process of the channel estimation at the transmitter using reciprocity or feedback. Therefore, it is interesting to take into account this channel time-variation in the channel model. The work in [13] proposes to accomplish this objective. It relies on the stochastic processes and estimation theories. Derived from a potentially outdated channel measurement and the channel statistics (mean and covariance), this proposed CSIT model consists of a channel estimate and its error covariance, acting as the effective channel mean and covariance. Both parameters depend on a temporal correlation factor, indicating the CSIT quality. Depending on this quality, the model switches smoothly from perfect to statistical channel information. This proposed CSIT is applicable to all Gaussian random channels, however it was only proposed for the MIMO single-user system. Our first contribution is a generalization of the work presented in [13] for a more general multicell multiuser context, i.e., for the MU-MIMO CoMP channel model.

1.2.2 Approximation of the MU-MIMO CoMP Ergodic Sum-Rate

Some works in the literature propose a transmit scheme that maximizes the sum-rate of the MU-MIMO channel [14–16]. These works assume that the channel is perfectly known at the transmitter. This assumption can have a significant impact on the maximum ergodic sum rate that can be reliably communicated over the channel, but may not be realistic in many practical scenarios. Our second contribution consists of deriving a first- and second-order approximation of the ergodic sum rate for a MU-MIMO CoMP system considering that the

transmitter has access to statistical channel state information (CSI), while the receiver has access to instantaneous CSI. We use the duality theory [14, 15] to compute the downlink MU-MIMO CoMP ergodic sum-rate since it states that the achievable sum-rate of the downlink MU-MIMO channel is equal to the achievable sum rate of the uplink MU-MIMO channel.

1.2.3 Convex Optimization of the Approximated Ergodic Sum-Rate

In the third contribution, we use convex optimization tools to find covariance matrices of the transmitted signal (known as input covariance matrices) that maximizes the first- and second-order approximation of the MU-MIMO CoMP ergodic sum-rate. Thus, we propose an efficient and fast convergent algorithm for obtaining these input covariance matrices.

1.3 Scientific Production and Contributions

During the doctorate, a United States patent regular application related to the proposed algorithm that maximizes the approximation of the ergodic sum-rate has been filed, with the following information:

- ▶ P32276-US2 “*Statistical Precoder Design for Coordinated Wireless Systems*”, **L. M. C. Sousa**, C. C. Cavalcante, T. F. Maciel and A. A. Guimarães.

Some parts of the work presented in this thesis have been published in the following journal and conferences:

- ▶ **SOUSA, L. M. C.**; GUIMARAES, A. A. ; CAVALCANTE, C. C. ; MACIEL, T. F.. “*Ergodic Sum Rate Maximization with Statistical CSIT in a Cooperative MIMO System*”. Submitted to IEEE Transactions on Vehicular Technology.
- ▶ GUIMARAES, A. A. ; GUERREIRO, I. M. ; **SOUSA, L. M. C.** ; MOREIRA, D. ; MACIEL, T. F. ; CAVALCANTE, C. C. . “*A (Very) Brief Survey on Optimization Methods for Wireless Communication Systems*”. International Telecommunications Symposium, ITS, 2010, Manaus
- ▶ **SOUSA, L. M. C.** ; CAVALCANTE, C. C. ; MACIEL, T. F. . “*Multiuser CoMP Transmit Processing with Statistical Channel State Information at the Transmitter*”. The Seventh International Symposium on Wireless Communication Systems - ISWCS'2010, 2010, York, UK.
- ▶ **SOUSA, L. M. C.** ; CAVALCANTE, C. C. . “*Performance Analysis of Multicell Multiuser MIMO Precoding Using Partial Knowledge at the Transmitter*”. XXVII Simpósio Brasileiro de Telecomunicações 2009 - SBrT 2009, Blumenau - SC.

The core of this Doctor's thesis has been developed in the context of three research projects funded by Ericsson Research, where eight technical reports have been produced and one more is being written. The list follows below:

- ▶ CAVALCANTI, F. R. P., CAVALCANTE, C. C., CARDIERI, P., **SOUSA, L. M. C.**, MOREIRA, R. B., CORREIA, V. D. B. “*First Technical Report UFC21 - Precoding Strategies for Interference Management*”, May 2007.
- ▶ CAVALCANTI, F. R. P., CARDIERI, P., CAVALCANTE, C. C., **SOUSA, L. M. C.**, MOREIRA, R. B., CORREIA, V. D. B. “*Second Technical Report UFC21 - Precoding Strategies for Interference Management*”, November 2007.

- ▶ CAVALCANTI, F. R. P., CARDIERI, P., CAVALCANTE, C. C., **SOUSA, L. M. C.**, MOREIRA, R. B., CORREIA, V. D. B. “*Third Technical Report UFC21 - Impact of Outdated CSI on the Performance of Precoding Algorithms*”, May 2008.
- ▶ CAVALCANTI, F. R. P., CAVALCANTE, C. C., CARDIERI, P., **SOUSA, L. M. C.**, MOREIRA, R. B., CORREIA, V. D. B. “*Final Technical Report UFC21 - Multiuser Precoding for MIMO Distributed Wireless Systems with Partial Dynamic CSI*”, January 2009.
- ▶ SILVA, Y. C. B., CAVALCANTE, C. C., **SOUSA, L. M. C.**, LOPES, P. R., CORREIA, V. D. B. “*First Technical Report UFC25 - Performance Analysis of Multicell Multiuser MIMO Precoding using Partial Knowledge at the Transmitter*”, August 2009.
- ▶ SILVA, Y. C. B., CAVALCANTE, C. C., **SOUSA, L. M. C.**, LOPES, P. R., CORREIA, V. D. B. “*Second Technical Report UFC25 - Cooperative MIMO Precoding with Limited Feedback*”, January 2010.
- ▶ SILVA, Y. C. B., CAVALCANTE, C. C., **SOUSA, L. M. C.**, LOPES, P. R., CORREIA, V. D. B. “*Third Technical Report UFC25 - Cooperative MIMO Precoding with Limited Feedback*”, July 2010.
- ▶ **SOUSA, L. M. C.**, CAVALCANTE, C. C., CAVALCANTI, F. R. P., FERNANDES, C. E. R., GUERREIRO, I. M., SILVA, I. L. J. “*First Technical Report UFC31 - A Single-user Analysis of a Coordinated Transmit Processing with Statistical CSIT*”, December 2010.

1.4 Document Organization

We provide below an outline of the thesis and describe the contribution of each chapter.

Chapter 2 – Discusses the wireless channel characteristics; the MIMO parameters and its spatial structure; the MU-MIMO channel and its derivations; and the capacity region of each type of MU-MIMO channel. Moreover, in this chapter, we describe the MU-MIMO CoMP channel model, the considered input-output signal model and its advantages.

Chapter 3 – From error estimation theory, this chapter proposes a statistical channel model which considers the channel temporal variation for the considered MU-MIMO CoMP scenario. The proposed channel model is compared to the case with perfect channel knowledge at the transmitter using the well-known zero-forcing (ZF) and minimum mean-square error (MMSE) precoding techniques. Performance results are also discussed.

Chapter 4 – Finds an approximation of the ergodic sum-rate for the MU-MIMO CoMP scenario. Next, optimizes the found approximation in order to obtain a near-optimal input covariance matrix per user using convex optimization tools. Simulation results are shown and discussed.

Chapter 5 The overall work presented herein is concluded analyzing the results achieved and pointing out future research directions.

Chapter 2

Multi-User Multiple-Input Multiple-Output (MU-MIMO) Coordinated Multipoint (CoMP) Wireless Channel

A good understanding of the multi-user MIMO (MU-MIMO) coordinated multipoint (CoMP) wireless channel, its key physical parameters and the modeling issues is the goal of this chapter. Firstly, some concepts related to the variations of the channel response over time and over frequency are discussed. After that, the multiple-input multiple-output (MIMO) channel is defined and its spatial characteristics are presented. Next, the MU-MIMO system is discussed and its downlink and uplink structures are defined. The capacity regions of these MU-MIMO models and the uplink-downlink duality are studied. Finally, the MU-MIMO CoMP system is discussed and the input/output channel model used in this thesis is derived, jointly with some important physical parameters.

2.1 The Wireless Channel

The wireless radio channel places fundamental limitations on the performance of communication systems. The transmission path between the transmitter and the receiver can vary from simple line-of-sight (LOS) to one that is severely obstructed by buildings, mountains and foliage. These various obstructions cause on the transmitted signal reflections on large surfaces, diffraction on edges and scattering on several objects. Therefore, the received signal is a superposition of multiple signals arriving from different directions at different time instants and with different phases and powers. In this section, we review the physical phenomena that modify the signal power. For a more detailed presentation of this subject, the reader is referred to [17].

2.1.1 Path Loss

Path loss is a range-dependent effect and is defined as the difference (in dB) between the effective transmitted power and the received power, which is due to the distance d between the transmitter and the receiver [3, 17]. In ideal free space, the received signal power is described by the Friis equation and follows an inverse square law power loss [18]. Several deterministic and empirical path loss models have been presented for various cellular environments

(microcells, macrocells, picocells, etc.), such as Okumura-Hata [19, 20], Walfisch-Ikegami and their COST-231 extensions [21] and plane-earth model [17]. A general path loss model can be given by:

$$PL = \xi d^{-\alpha}, \quad (2.1)$$

where α is the path loss exponent which indicates the rate at which the path loss increases with distance, and ξ is a scaling factor that accounts for the average channel attenuation and may or may not consider the antenna characteristics [17]. The path loss exponent varies typically from 2 to 6, depending on the propagation environment. For the free space case α is equal to 2, while for the case of full specular reflections from the ground α is 4. For buildings and indoor environments α can take values from 4 to 6 [3].

2.1.2 Shadowing

Shadowing is a phenomenon which results from large obstacles blocking the main signal path between the transmitter and receiver. It is known as *large-scale fading* and is determined by the local mean of the signal. The shadowing effects are random and influenced by antenna heights, operating frequency and features of the propagation environment. A usual shadowing model follows the log-normal distribution with probability density function (PDF), given by:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log(x) - \mu)^2}{2\sigma^2}\right) \quad x > 0 \quad (2.2)$$

where μ and σ are the mean and the standard deviation of the shadowing's logarithm, respectively [3].

2.1.3 Fast Fading

Fast fading describes the rapid fluctuation of the amplitude of a radio signal over a short period of time or travel distance. It is known as *small-scale fading* and is caused by the interference between two or more versions of the transmitted signal which arrive at the receiver at slightly different times. These waves, called *multipath waves*, are combined at the receive antenna to give a resultant signal that can vary widely in amplitude and phase depending on the distribution of the intensity and relative propagation time of the waves and the bandwidth of the transmitted signal. The statistical time-varying nature of the received amplitude is commonly described by the following two fading distributions:

Rayleigh Fading

The simplest probabilistic channel model is based on the assumption that there is a large number of independent scattered paths and there is no dominant propagation path, i.e., there is no LOS between the transmitter and receiver. Applying the central limit theorem, the channel impulse response can be considered as a complex-valued Gaussian process irrespective of the distribution of the individual paths. In this non-line-of-sight (NLOS) configuration, this random process is assumed to have zero mean and phase uniformly distributed between 0 and 2π radians. The magnitude of the received signal is a Rayleigh random variable with PDF given by:

$$p(x) = \frac{2x}{\Omega} \exp\left(-\frac{x^2}{\Omega}\right) \quad x > 0 \quad (2.3)$$

where $\Omega = \mathbb{E}\{x^2\}$ is the average received power [3].

The Rayleigh model is quite reasonable for scattering mechanisms where there are many small reflectors, but is adopted, primarily for its simplicity, even in typical cellular situations

with a relatively small number of reflectors.

Ricean Fading

The Ricean fading model is used when a direct, possibly a LOS, path exists and the assumption of a zero-mean fading process does no longer hold. This model is often defined in terms of the Ricean factor κ which denotes the ratio of the power in the mean component of the channel (direct path) to the power in the scattered paths, i.e. [3]:

$$\kappa = \frac{\|\bar{\mathbf{h}}\|_F^2}{\text{tr}(\mathbb{E}\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H\})} \quad (2.4)$$

where $\|\cdot\|_F$ is the Frobenius norm, $\text{tr}(\cdot)$ is the trace of a matrix, $\bar{\mathbf{h}}$ is the complex channel mean vector and $\tilde{\mathbf{h}}$ is the zero-mean circularly symmetric complex Gaussian (ZMCSCG) channel vector.

The magnitude of the received signal is a Ricean random variable with PDF given by [3]:

$$p(x) = \frac{2x(\kappa + 1)}{\Omega} \exp\left(-\kappa - \frac{(\kappa + 1)x^2}{\Omega}\right) J_0\left(2x\sqrt{\frac{\kappa(\kappa + 1)}{\Omega}}\right) \quad x > 0 \quad (2.5)$$

where $\Omega = \mathbb{E}\{x^2\}$ and J_0 is the zero-order modified Bessel function of the first kind defined as

$$J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(-x \cos \theta) d\theta \quad x > 0. \quad (2.6)$$

Although the Ricean PDF has a more complicated form, it is often a better model for fading than the Rayleigh model.

2.1.4 Channel Selectivity

The presence of reflecting objects and scatterers in the channel creates dissipating effects which result in the spreading of the signal in different dimensions, affecting significantly the received signal. These dimensions are time (Doppler spread), space (angle spread) and frequency (delay spread). Angle spread will not be discussed in this thesis, but the reference [3] is given for the interested reader.

Temporal Selectivity: Doppler Spread and Coherence Time

The motions of the transmitter, of the receiver or of the scatterers in the channel cause the temporal selectivity of the channel. These motions lead to a transmitted single tone to be spread in frequency over a finite spectral bandwidth, an effect known as *Doppler shift*. This effect can be captured in the power spectrum of the channel and the variation due to Doppler shifts are specific to each path and dependent on their angle with respect to the movement direction of the transmitter/receiver. Different Doppler shifts lead to the so-called *Doppler spread*, which is the maximum frequency spread among all Doppler shifts, and is given by:

$$f_d = \frac{v}{\lambda_c}, \quad (2.7)$$

where v is the mobile speed and λ_c is the carrier wavelength.

How fast the channel decorrelates with time is specified by the temporal autocorrelation function ρ . The most commonly used model for the autocorrelation function is the Clarke-Jakes' model, which assumes uniformly distributed scatterers on a circle around the

antenna, and is given by [22]:

$$\rho[\tau] = J_0(2\pi f_d \tau), \quad (2.8)$$

where τ is the sampling interval.

Higher mobility in a system commonly causes large Doppler spread and faster channel variation in time. Thus, a larger Doppler spread is associated with a higher temporal selectivity. A measure of the temporal selectivity is the channel coherence time, defined as the time interval over which the channel remains strongly correlated. The shorter the coherence time, the faster the channel changes with time. Since the coherence time is a statistically defined quantity, an approximated relation to the Doppler spread is:

$$T_c \approx \frac{1}{f_d}. \quad (2.9)$$

It is common to consider a constant such as 2, 4 or 8 in front of f_d in this relation; but there is no single agreed-number [17]. The important property is the inverse-proportionality between T_c and f_d .

Spectral Selectivity: Delay Spread and Coherence Bandwidth

Spectral selectivity is caused by the presence of multiple scaled versions of the transmitted signal arriving at different time instants at the receiver. An indicator of this selectivity is the *delay spread*, defined as time difference between the maximum multipath delay τ_{max} and the minimum path delay τ_{min} (typically the arrival time of the LOS component). Delay spread causes frequency selective fading and the range of frequencies over which the channel can be considered “flat” defines the coherence bandwidth B_c and depends on the form of the power delay spectrum. Similarly to equation (2.9), the coherence bandwidth is defined as $B_c = 1/\tau_{max}$. A channel is considered as flat or frequency non-selective if the signal bandwidth B is significantly small compared to the channel coherence bandwidth, i.e., $B \ll B_c$. In this thesis, only flat fading channels are considered.

2.2 Multiple-input Multiple-output (MIMO) Channel

Multiple-input multiple-output (MIMO) communication techniques have been an important research area for next-generation wireless systems due to their potential for high capacity, increased diversity and interference suppression [23].

The MIMO wireless channel is characterized by using multiple antennas at both the transmitter and the receiver. It generalizes the special cases of having a single antenna at only one side: multiple-input single-output (MISO) and single-input multiple-output (SIMO). Therefore, in addition to spanning the temporal and spectral dimensions, a MIMO channel exhibits a new spatial dimension across the antennas. The channel contains multiple elements and is often represented in an elegant, compact and unified way by a channel matrix.

2.2.1 Spatial MIMO Channel Modeling

In a system with N_t transmit antennas and N_r receive antennas, the frequency-flat MIMO channel at discrete time index n can be represented as a matrix $\mathbf{H}[n]$ of size $N_r \times N_t$

$$\mathbf{H}[n] = \begin{bmatrix} h_{11}[n] & h_{12}[n] & \dots & h_{1N_t}[n] \\ h_{21}[n] & h_{22}[n] & \dots & h_{2N_t}[n] \\ \vdots & \vdots & \dots & \vdots \\ h_{N_r 1}[n] & h_{N_r 2}[n] & \dots & h_{N_r N_t}[n] \end{bmatrix} \quad (2.10)$$

in which $h_{ij}[n]$ is the channel from transmit antenna j to receive antenna i .

Each element $h_{ij}[n]$ in a MIMO channel can be modeled as a complex random process. These elements can be correlated and can have different mean values. Hence, it is possible to decompose the channel given in the equation (2.10) into a fixed part and a variable part as [3]:

$$\mathbf{H} = \sqrt{\frac{\kappa}{1+\kappa}} \bar{\mathbf{H}} + \sqrt{\frac{1}{1+\kappa}} \tilde{\mathbf{H}} \quad (2.11)$$

where $\bar{\mathbf{H}}$ is the complex channel mean matrix, $\tilde{\mathbf{H}}$ is a ZMCSCG matrix and κ is the Ricean factor which represents the ratio between the energy of the mean channel part and the energy of the random one. The time index n has been omitted for simplicity of notation.

Channel Covariance and the Correlation-based Models

The channel covariance gives information about the spatial correlation among all the transmit and receive antennas. Thus, it is defined among all $N_r N_t$ channel elements as being the positive semi-definite Hermitian matrix given by:

$$\mathbf{R}[0] = \mathbb{E}\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H\} \quad (2.12)$$

where $\tilde{\mathbf{h}} = \text{vec}(\tilde{\mathbf{H}})$, the $\text{vec}(\cdot)$ operator vectorizes a matrix by stacking its columns and $\mathbb{E}\{\cdot\}$ is the expectation operator. The diagonal elements of the matrix $\mathbf{R}[0]$ represent the power gain of the $N_r N_t$ scalar channels, and its off-diagonal elements are the cross-coupling between these scalar channels.

A full-correlation model, disregarding the channel mean ($\bar{\mathbf{H}} = 0$), is described as [3]

$$\mathbf{H} = \text{unvec}(\mathbf{R}[0]^{1/2} \text{vec}(\mathbf{H}_{\mathbf{w}})) \quad (2.13)$$

where $\text{unvec}(\cdot)$ is the inverse operator of $\text{vec}(\cdot)$ and $\mathbf{H}_{\mathbf{w}}$ is an $N_r \times N_t$ independent identically distributed (i.i.d.) ZMCSCG matrix with unit variance. This full-correlation model is the most accurate model but it is not common due to its complexity. For simplicity, the covariance matrix $\mathbf{R}[0]$ is often assumed to have a less general, separable structure called Kronecker structure [3]. This model assumes that the covariance of the scalar channels seen from all N_t transmit antennas to a single receive antenna (a row of \mathbf{H}) is the same for any receive antenna (any row) and equal to the $N_t \times N_t$ matrix $\mathbf{R}_{\mathbf{t}}$

$$\mathbf{R}_{\mathbf{t}} = \mathbb{E}\{\tilde{\mathbf{h}}_i^T \tilde{\mathbf{h}}_i^H\} \quad \text{for } \forall i, \quad (2.14)$$

where $\tilde{\mathbf{h}}_i$ is the i -th row of the matrix $\tilde{\mathbf{H}}$ with $i = 1, \dots, N_r$.

Similarly, the covariance of the scalar channel seen from a single transmit antenna to all N_r receive antennas (a column of \mathbf{H}) is assumed to be the same for any transmit antenna (any column) and equal to the matrix $\mathbf{R}_{\mathbf{r}}$ of dimensions $N_r \times N_r$

$$\mathbf{R}_{\mathbf{r}} = \mathbb{E}\{\tilde{\mathbf{h}}_j \tilde{\mathbf{h}}_j^H\} \quad \text{for } \forall j, \quad (2.15)$$

where $\tilde{\mathbf{h}}_j$ is the j -th column of the matrix $\tilde{\mathbf{H}}$ with $j = 1, \dots, N_t$.

In this case, the channel covariance matrix can be decomposed by both complex Hermitian positive semidefinite matrices $\mathbf{R}_{\mathbf{t}}$ and $\mathbf{R}_{\mathbf{r}}$ as:

$$\mathbf{R}[0] = \mathbf{R}_{\mathbf{t}}^T \otimes \mathbf{R}_{\mathbf{r}}, \quad (2.16)$$

where the operator \otimes means the Kronecker product [24]. The Kronecker model is satisfied for small number of antennas or large antenna spacing [17] and it will be considered in the scenario of this thesis.

Channel Mean and the Factor κ

The channel mean is the fixed component of the channel, usually corresponding to a LOS propagation path or a cluster of strong paths. The mean of a MIMO channel is the complex matrix $\bar{\mathbf{H}}$ of size $N_r \times N_t$ obtained as [3]

$$\bar{\mathbf{H}} = \frac{\mathbb{E}\{\mathbf{H}\}}{\sqrt{\frac{\kappa}{1+\kappa}}}. \quad (2.17)$$

The elements of the mean can have different amplitudes and arbitrary phase, caused by the spatial selectivity. The strength of a channel mean can be quantified by the Rician κ factor. It is defined as the ratio of the power in the channel mean and the average power in the channel variable component. In Section 2.1 the Rician factor was defined in (2.4) for a channel vector, here we have the definition for a MIMO channel matrix [17]:

$$\kappa = \frac{\|\bar{\mathbf{H}}\|_F^2}{\text{tr}(\mathbf{R}[0])}. \quad (2.18)$$

The κ factor can take any real value between 0 and infinity. When $\kappa = 0$, the channel has the Rayleigh distribution. When $\kappa \rightarrow \infty$, the channel becomes deterministic. Measurements of fixed broadband channels have shown that the κ factor can take values from 0 up to 30dB in practice, and it tends to decrease with increasing distance between the transmitter and the receiver [25].

The channel mean is assumed to be generated in this thesis as:

$$\bar{\mathbf{H}} = \text{diag} \left(\begin{bmatrix} a_0(\theta_c) \\ a_1(\theta_c) \\ \vdots \\ a_{N_r}(\theta_c) \end{bmatrix} \right) \cdot \mathbf{S}_{N_r \times N_t} \cdot \text{diag} \left(\begin{bmatrix} a_0(\theta_s) \\ a_1(\theta_s) \\ \vdots \\ a_{N_t}(\theta_s) \end{bmatrix} \right) \quad (2.19)$$

where θ_s and θ_c are the angle between the BS and the user with respect to transmit and receive array broadside direction, respectively, and $a_m(\theta) = \exp(j\pi(m-1)\sin(\theta))$. The matrix \mathbf{S} is a matrix composed by 1's and -1's and with rank equal to $\min(N_t, N_r)$. Thus, the matrix \mathbf{S} has the role to condition the channel mean $\bar{\mathbf{H}}$, i.e., to make the matrix $\bar{\mathbf{H}}$ be a full rank one [3].

The channel given in equation (2.11) can then be written considering the Kronecker channel model and the mean channel matrix as:

$$\mathbf{H} = \sqrt{\frac{\kappa}{1+\kappa}} \bar{\mathbf{H}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{R}_r^{1/2} \mathbf{H}_w \mathbf{R}_t^{1/2}. \quad (2.20)$$

where $\mathbf{R}_t^{1/2}$ is the matrix square-root of \mathbf{R}_t , such that, $\mathbf{R}_t = \mathbf{R}_t^{1/2} \mathbf{R}_t^{1/2}$. Analogously, $\mathbf{R}_r = \mathbf{R}_r^{1/2} \mathbf{R}_r^{1/2}$.

2.3 Multi-user (MU) MIMO Channel

Multi-user MIMO (MU-MIMO) channels are a set of advanced MIMO channels that allow to exploit the availability of multiple independent terminals in order to enhance the communication capabilities of each individual terminal. There are two types of multi-user

(MU) channels: the uplink and the downlink channel. In the uplink MU channel, many transmitters (users) send signals to one receiver (base-station (BS)) in the same frequency band. It is also known as *multiple access channel (MAC)* or *reverse channel*. The downlink MU channel has one transmitter (BS) sending signals to many receivers (users) and it is also known as *broadcast channel (BC)* or *forward channel*. In the downlink, the transmit antennas can cooperate at the BS to mitigate the MU interference and, in the uplink, this processing is performed at the receiver (BS). Next, we analyze both MU-MIMO channels and their main features.

2.3.1 Downlink Multi-user MIMO Channel

Consider the communication between a BS b equipped with N_t transmit antennas and K active users, where each active user k is equipped with N_r receive antennas. The set of active users is defined by the users simultaneously downloading or uploading packets during one given scheduling window.

In the downlink MU-MIMO model, the BS b transmits simultaneously to K users as shown in Figure 2.1. Assuming a frequency-flat channel, the signal $\mathbf{y}_k \in \mathbb{C}^{N_r \times 1}$ received by each user k can be written as:

$$\mathbf{y}_k = \mathbf{H}_{k,b}\mathbf{x} + \boldsymbol{\eta}_k \quad \text{for } k = 1, \dots, K \quad (2.21)$$

where $\boldsymbol{\eta}_k \in \mathbb{C}^{N_r \times 1}$ is the additive ZMCSCG white noise vector at the receiver k with covariance matrix $\sigma_{\eta}^2 \mathbf{I}$, $\mathbf{H}_{k,b} \in \mathbb{C}^{N_r \times N_t}$ represents the frequency-flat channel matrix from BS b to user k and $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the signal vector transmitted by BS b . The signal \mathbf{x} is composed by signal desired by all users, i.e., $\mathbf{x} = \sum_{k=1}^K \mathbf{x}_k$ with \mathbf{x}_k being the signal desired by user k . The covariance matrix of the transmitted signal in the downlink model is also known as downlink input covariance matrix and it is given by $\boldsymbol{\Omega} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$. The input covariance matrix is subject to a power constraint $\text{tr}(\boldsymbol{\Omega}) \leq P_b$, where P_b is the maximum transmit power constraint of BS b .

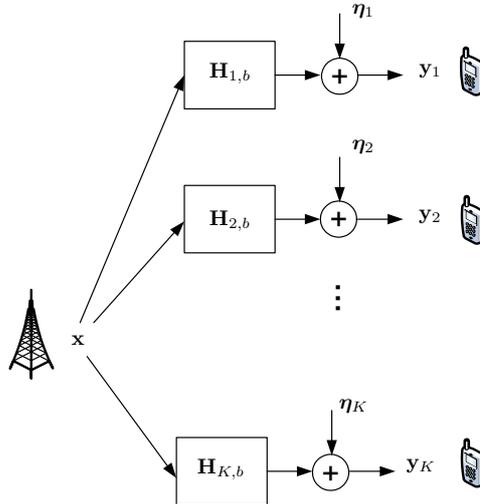


Figure 2.1: System model of the downlink MU-MIMO channel.

2.3.2 Uplink Multi-user MIMO Channel

In the uplink MU-MIMO model, the K users transmit simultaneously to BS b as shown in Figure 2.2. The received signal $\mathbf{y} \in \mathbb{C}^{N_t \times 1}$ at BS b is given by

$$\mathbf{y} = \sum_{k=1}^K \mathbf{H}_{k,b}^H \mathbf{x}_k + \boldsymbol{\eta}, \quad (2.22)$$

where $\mathbf{x}_k \in \mathbb{C}^{N_r \times 1}$ is the k -th user signal vector, $\mathbf{H}_{k,b} \in \mathbb{C}^{N_r \times N_t}$ represents the frequency-flat channel matrix from BS b to user k on the downlink model and $\boldsymbol{\eta} \in \mathbb{C}^{N_t \times 1}$ is the additive ZMCSCG white noise vector at the BS with variance σ_η^2 . We have considered that the uplink channel matrix is the dual corresponding downlink channel matrix, thus the channel matrix in the uplink model is the Hermitian of the channel matrix in the downlink model $\mathbf{H}_{k,b}^H$. This consideration will be taken along all thesis and its impact will be analyzed later. The covariance matrix of the transmitted signal by user k in the uplink model is also known as uplink input covariance matrix and it is given by $\mathbf{Q}_k = \mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\}$. These input covariance matrices are subject to an individual power constraint $\text{tr}(\mathbf{Q}_k) \leq P_k$, where P_k is the maximum transmit power of user k .

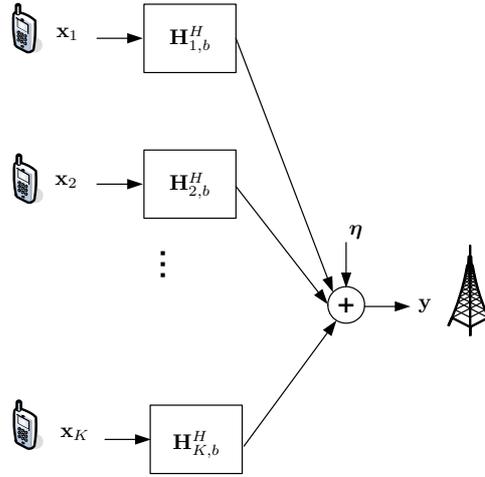


Figure 2.2: System model of the uplink MU-MIMO channel.

In the MU-MIMO channel with K users, the capacity is characterized by a K -dimensional rate region, where each point is a vector of achievable rates by all K users simultaneously [26]. This capacity region defines the limit of error-free communications given certain channel characteristics and it is used as the ultimate measure of channel capacity. In the following, we study the capacity region for the two MU-MIMO channels: downlink and uplink. For simplicity of notation, we consider in the next two subsections that the channel matrix from BS b to user k is only denoted by \mathbf{H}_k , since there is only one BS. Moreover, from this point forward, we consider that the noise power is equal to unity, i.e., $\sigma_\eta^2 = 1$.

2.3.3 Capacity Region of the Uplink MU-MIMO Channel

Firstly, we denote $R_{k,\text{UL}}$ as being the rate that can be reliably (error-free) maintained for the user k in the uplink channel model given in bits per second per Hertz (bps/Hz). We assume Gaussian signaling for each user and consider joint decoding of the users signals, i.e., decoding all signals simultaneously through maximum likelihood (ML) decoding. The uplink MU capacity region with joint decoding and individual power constraints P_1, \dots, P_K on each user has been shown to satisfy [15] [27]:

$$C_{\text{UL}}(\mathbf{P}; \mathbf{H}_1^H, \mathbf{H}_2^H, \dots, \mathbf{H}_K^H) = \sum_{k=1}^K R_{k,\text{UL}} \leq \max_{\{\mathbf{Q}_k\}_{k=1}^K; \mathbf{Q}_k \succeq 0; \text{tr}(\mathbf{Q}_k) \leq P_k} \log \left| \mathbf{I}_{N_t} + \sum_{k=1}^K \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right| \quad (2.23)$$

where $|\mathbf{X}|$ represents the determinant of the matrix \mathbf{X} , $\mathbf{P} = [P_1, \dots, P_K]$ is the set of powers and $C_{\text{UL}}(\mathbf{P}; \mathbf{H}_1^H, \mathbf{H}_2^H, \dots, \mathbf{H}_K^H)$ is the maximum sum rate of the uplink MU-MIMO channel. Each set

of covariance matrices determines a K -dimensional polytope and the uplink capacity region is equal to union of all such polytopes, in which the union is only performed over all covariance matrices satisfying the power constraints. For a case with 2 users with a single antenna, the capacity region is a pentagon as shown in Figure 2.3 [3]. Along the bold line the sum rate ($R_{1,UL} + R_{2,UL}$) is constant and is the maximum achievable sum rate. Every point along this line is achieved by each user transmitting at the maximum available power. The corner points of the pentagon can be achieved by successive decoding [3]. In the lower corner point A, user 1 transmits at full rate $R_{1,UL} = \log(1 + \text{tr}(\mathbf{Q}_1) \|\mathbf{h}_1\|_F^2)$, thus assuming no interference. User 2 transmits assuming that the signal from user 1 is additional noise. The same procedure can be performed to the upper corner B, user 1 transmits treating user 2 as additional noise and user 2 transmits with full rate equal to $R_{2,UL} = \log(1 + \text{tr}(\mathbf{Q}_2) \|\mathbf{h}_2\|_F^2)$. All other points along the bold line can be achieved by time-sharing between the two schemes [3].

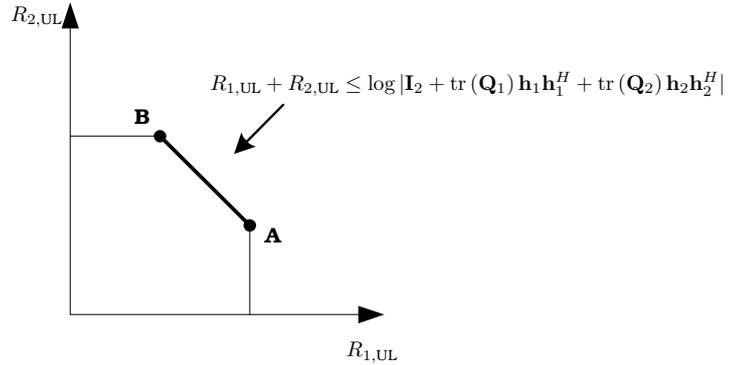


Figure 2.3: Capacity region for the uplink MU-MIMO channel when the receivers have a single antenna.

If the decoding uses minimum mean-square error (MMSE) with successive cancellation, as reported in [28], the capacity region of this channel now satisfies:

$$C_{UL}(\mathbf{P}; \mathbf{H}_1^H, \mathbf{H}_2^H, \dots, \mathbf{H}_K^H) = \sum_{k=1}^K R_{k,UL} \leq \max_{\{\mathbf{Q}_k\}_{k=1}^K; \mathbf{Q}_k \succeq 0; \text{tr}(\mathbf{Q}_k) \leq P_k} \sum_{k=1}^K \log \frac{|\mathbf{I} + \sum_{i=k}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i|}{|\mathbf{I} + \sum_{i=k+1}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i|} \quad (2.24)$$

considering that the user 1 is decoded first, then the user 2 is decoded and so on.

The objective function in (2.23) is a convex function of the uplink input covariance matrices \mathbf{Q}_k and the constraints are separable because there is an individual trace constraint on each correlation matrix \mathbf{Q}_k . In such situations, it is generally sufficient to optimize with respect to the first variable while holding all other variables constant, then optimize with respect to the second variable, and so on, in order to reach a globally optimum point [29].

In [16], the authors consider the MMSE receiver with successive interference cancellation (SIC) instead of the ML receiver. As previously mentioned, this receiver decodes the users' signals successively and the user to be decoded treats all the other users as interference. Moreover, it subtracts out the signal transmitted by the users already decoded from the received signal. In that work, the authors have proposed a block-coordinated ascent algorithm to reach the maximum sum rate of the uplink MU-MIMO channel. This algorithm consists of optimizing (2.24) with respect to the first variable while holding all other variables constant, then optimize with respect to the second variable, and so on, until reaching a globally optimum point. The work in [16] proposed to evaluate the sum rate maximizing the covariance matrix of any user through the single-user water-filling of its own channel with noise equal to the

actual noise and interference from the other $K - 1$ transmitters. Next, we study the capacity region of the downlink MU-MIMO channel.

2.3.4 Capacity Region of the Downlink MU-MIMO Channel

The sum rate capacity of a downlink MU-MIMO Gaussian channel C_{DL} for multiple users having multiple antennas each has been shown by Sato in [30] to satisfy:

$$C_{\text{DL}} \leq C_{\text{Sato}} = \min_{\sigma_\eta^2 \mathbf{I} > 0} \left(\max_{\text{tr}(\Omega) < P_b} \log \frac{|\sigma_\eta^2 \mathbf{I} + \mathbf{H}\Omega\mathbf{H}^H|}{|\sigma_\eta^2 \mathbf{I}|} \right) \quad (2.25)$$

where $\mathbf{H} = [\mathbf{H}_1^T \ \mathbf{H}_2^T \ \dots \ \mathbf{H}_K^T]^T$ is the matrix composed by channel matrices \mathbf{H}_k of all K users and Ω is the downlink input covariance matrix. This equation is known as Sato's upper bound on the capacity region of general broadcast channels which considers cooperation among users. As already mentioned, since in the MU-MIMO system there is only one BS, we omit the index b for simplicity of notation.

An achievable capacity region for downlink MU-MIMO channels has been derived in [31] using Costa's dirty-paper coding (DPC) [32], in which the authors demonstrate that the achievable rates meet the Sato upper bound. The basic idea of the DPC is that if the transmitter has perfect, non-causal channel state information (CSI) regarding an additive Gaussian interference, then the capacity of the channel is the same as if there was no additive interference. The main point is that the transmitter can subtract the known interference before the transmission of the desired signal, but in such way that the transmit power is not increased. Considering successive decoding and that the user 1 is encoded first, followed by user 2 and so on, then the rate $R_{k,\text{DL}}$ achieved by user k is given by:

$$R_{k,\text{DL}} = \log \frac{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j \geq k} \Omega_j \right) \mathbf{H}_k^H \right|}{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j > k} \Omega_j \right) \mathbf{H}_k^H \right|}, \quad \text{for } k = 1, \dots, K. \quad (2.26)$$

where $\Omega_k = \mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\}$ is the input covariance matrix of the signal desired by user k since $\mathbf{x} = \sum_{k=1}^K \mathbf{x}_k$ is the transmitted signal.

If the consideration now is that the user K is encoded first, followed by user $K - 1$ and so on, then the achievable rate $R_{k,\text{DL}}$ is now given by:

$$R_{k,\text{DL}} = \log \frac{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j \leq k} \Omega_j \right) \mathbf{H}_k^H \right|}{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j < k} \Omega_j \right) \mathbf{H}_k^H \right|}, \quad \text{for } k = 1, \dots, K. \quad (2.27)$$

Therefore, in [31], the authors state that the capacity region is the convex hull of the union of all such rates over all positive semi-definite correlation matrices satisfying the sum power constraint:

$$C_{\text{DL}}(\mathbf{H}_1, \dots, \mathbf{H}_K, P_b) = \max_{\{\Omega_k\}_{k=1}^K; \Omega_k \succeq 0, \sum_{k=1}^K \text{Tr}(\Omega_k) \leq P_b} \sum_k R_{k,\text{DL}}. \quad (2.28)$$

This optimization problem is hard to solve, since the sum rate equation is in general neither a concave nor a convex function of the input covariance matrices. Thus, a numerical solution for this problem requires a brute force search over the entire space of all input covariance

matrices that meets the sum power constraint. For a channel with two user and single antenna each, the capacity region is shown in Figure 2.4 [3].

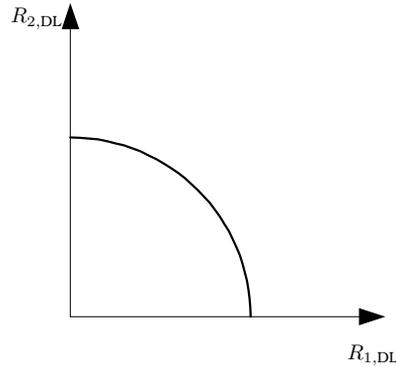


Figure 2.4: Capacity region for the downlink MU-MIMO channel considering two users.

In [15] and [33], the authors have shown the existence of duality between the uplink and the downlink. They have established that the downlink MU-MIMO capacity region is the same as the capacity region of the corresponding uplink MU-MIMO channel with the transmit power constraint of the downlink channel translated to the sum of powers in the uplink. The authors have named their result a *duality* connection. This is also equivalent to saying that the sum rate of the downlink MU-MIMO channel is equal to the sum rate of the dual uplink MU-MIMO channel subject to the sum power constraint P , i.e.:

$$C_{DL}(\mathbf{H}_1, \dots, \mathbf{H}_K, P_b) = C_{UL}(\mathbf{H}_1^H, \dots, \mathbf{H}_K^H, P_b), \quad (2.29)$$

where P_b , for the uplink channel model, is the sum of the power constraints of each user k , i.e., $P_b = \sum_{k=1}^K P_k$.

This equivalence between the uplink and downlink channel has been analyzed in many different situations. In [2] the authors have shown that the capacity of a single-user channel is unchanged when the role of transmitters and receivers is interchanged. In the case of the downlink of a multiple antenna system employing simple linear beamforming followed by single-user receivers, the works in [34] and [35] show that the optimal choice of transmit beamvectors is closely related to a virtual uplink problem. Other works in [33] and [36] show that the capacity region of degraded downlink Gaussian channel is the same as the capacity region of the corresponding uplink channel with the transmit power constraint of the downlink channel translated to the sum of powers in the uplink. In the sequel, the duality between the uplink and downlink channel is detailed.

2.3.5 Duality Between Downlink MU-MIMO Channel and Uplink MU-MIMO Channels

Figure 2.5 shows the downlink MU-MIMO channel with its dual uplink MU-MIMO channel, which is formed by reversing the roles of the transmitters and receivers. Thus, the dual uplink MU-MIMO channel is a K -user MU-MIMO uplink channel where each of the dual uplink channels is the conjugate transpose of the corresponding downlink channel.

From Figure 2.5, we can state the main differences between the uplink and the downlink channel:

- i. In the downlink there is a single power constraint associated with the BS (transmitter), whereas in the uplink a power constraint is associated with each user (transmitter).

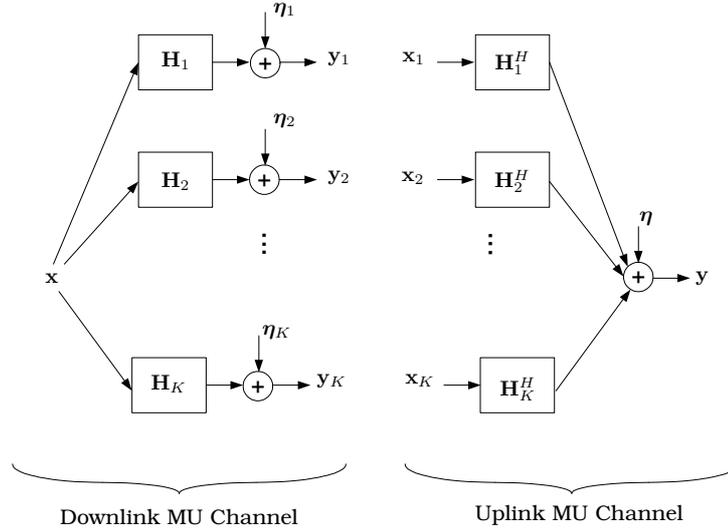


Figure 2.5: System model of the downlink MU-MIMO (left) and uplink MU-MIMO (right) channels.

- ii. The additive noise in the downlink is associated with each receiver (one noise vector for each receiver) and in the uplink the additive noise is only one vector since we have one receiver.
- iii. In the downlink, both signal and interference per each user travel through the same channel, whereas in the uplink these signals travel through different channels

Since the downlink and uplink channels look like mirror images of each other, the duality between these channels can be considered and the capacity region of either channel can be obtained from the capacity region of the other.

Next, we obtain the transformation that maps the uplink input covariance matrices into the downlink input covariance matrices [14]. With this transformation, we can use the fact that the achievable sum rate of the downlink MU-MIMO channel can be obtained from the achievable sum rate of the uplink MU-MIMO channel.

Assume that in the downlink the user K is encoded first, then the user $K - 1$ is encoded and so on; we assume that in the uplink the users are decoded in the reverse order, i.e., the user 1 is decoded first, next the user 2 is decoded, and so on. Then, from equation (2.24), we can write the rate $R_{k,\text{UL}}$ achieved by user k in the uplink as [15]:

$$\begin{aligned}
 R_{k,\text{UL}} &= \log \frac{|\mathbf{I}_{N_t} + \sum_{i=k}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i|}{|\mathbf{I}_{N_t} + \sum_{i=k+1}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i|}, \quad \text{for } k = 1, \dots, K. \\
 &= \log \frac{|\mathbf{I}_{N_t} + \sum_{i=k+1}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i + \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k|}{|\mathbf{I}_{N_t} + \sum_{i=k+1}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i|}, \quad (2.30)
 \end{aligned}$$

where \mathbf{I}_M is the identity matrix of dimensions $M \times M$.

Considering that $\log \frac{|\mathbf{A}+\mathbf{B}|}{|\mathbf{A}|} = \log |\mathbf{I} + \mathbf{A}^{-1}\mathbf{B}|$ and using this result in (2.30) we have:

$$R_{k,\text{UL}} = \log \left| \mathbf{I}_{N_t} + \left(\mathbf{I}_{N_t} + \sum_{j=k+1}^K \mathbf{H}_j^H \mathbf{Q}_j \mathbf{H}_j \right)^{-1} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right|, \quad \text{for } k = 1, \dots, K. \quad (2.31)$$

For the downlink channel and from equation (2.27) we can write the rate $R_{k,\text{DL}}$ achieved by

user k as:

$$\begin{aligned}
R_{k,DL} &= \log \frac{\left| \mathbf{I}_{N_r} + \mathbf{H}_k \left(\sum_{j=1}^k \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}{\left| \mathbf{I}_{N_r} + \mathbf{H}_k \left(\sum_{j=1}^{k-1} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}}, \quad \text{for } k = 1, \dots, K. \\
&= \log \frac{\left| \mathbf{I}_{N_r} + \mathbf{H}_k \left(\sum_{j=1}^{k-1} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H + \mathbf{H}_k \boldsymbol{\Omega}_k \mathbf{H}_k^H \right|}{\left| \mathbf{I}_{N_r} + \mathbf{H}_k \left(\sum_{j=1}^{k-1} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}}, \\
&= \log \left| \mathbf{I}_{N_r} + \left(\mathbf{I}_{N_r} + \mathbf{H}_k \left(\sum_{j=1}^{k-1} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right)^{-1} \mathbf{H}_k \boldsymbol{\Omega}_k \mathbf{H}_k^H \right|. \tag{2.32}
\end{aligned}$$

Let us define the following auxiliary matrices:

$$\mathbf{A}_k = \mathbf{I}_{N_r} + \mathbf{H}_k \left(\sum_{j=1}^{k-1} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H, \tag{2.33}$$

and

$$\mathbf{B}_k = \mathbf{I}_{N_t} + \sum_{j=k+1}^K \mathbf{H}_j^H \mathbf{Q}_j \mathbf{H}_j. \tag{2.34}$$

Substituting (2.33) and (2.34) into (2.32) and (2.31), we can write:

$$\begin{aligned}
R_{k,UL} &= \log \left| \mathbf{I}_{N_t} + \mathbf{B}_k^{-1} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right| \\
&= \log \left| \mathbf{I}_{N_t} + \mathbf{B}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{B}_k^{-1/2} \right| \\
&= \log \left| \mathbf{I}_{N_t} + \mathbf{B}_k^{-1/2} \mathbf{H}_k^H \mathbf{A}_k^{-1/2} \mathbf{A}_k^{1/2} \mathbf{Q}_k \mathbf{A}_k^{1/2} \mathbf{A}_k^{-1/2} \mathbf{H}_k \mathbf{B}_k^{-1/2} \right| \tag{2.35}
\end{aligned}$$

and

$$\begin{aligned}
R_{k,DL} &= \log \left| \mathbf{I}_{N_r} + \mathbf{A}_k^{-1} \mathbf{H}_k \boldsymbol{\Omega}_k \mathbf{H}_k^H \right| \\
&= \log \left| \mathbf{I}_{N_r} + \mathbf{A}_k^{-1/2} \mathbf{H}_k \boldsymbol{\Omega}_k \mathbf{H}_k^H \mathbf{A}_k^{-1/2} \right| \\
&= \log \left| \mathbf{I}_{N_r} + \mathbf{A}_k^{-1/2} \mathbf{H}_k \mathbf{B}_k^{-1/2} \mathbf{B}_k^{1/2} \boldsymbol{\Omega}_k \mathbf{B}_k^{1/2} \mathbf{B}_k^{-1/2} \mathbf{H}_k^H \mathbf{A}_k^{-1/2} \right| \tag{2.36}
\end{aligned}$$

In equations (2.35) and (2.36), the matrices \mathbf{B}_k and \mathbf{A}_k represent the interference experienced by user k in the uplink and downlink MU-MIMO channels, respectively.

We define the channel matrix taking into account the interference effect as the effective channel matrix. From the equation (2.35), we can treat $\mathbf{B}_k^{-1/2} \mathbf{H}_k^H \mathbf{A}_k^{-1/2}$ as the effective channel of the uplink system and $\mathbf{A}_k^{1/2} \mathbf{Q}_k \mathbf{A}_k^{1/2}$ as the input covariance matrix of this effective uplink channel. We can note from (2.36) that $\mathbf{A}_k^{-1/2} \mathbf{H}_k \mathbf{B}_k^{-1/2}$ is the effective downlink channel and it is equal to the Hermitian of the effective uplink channel. Thus, we can use the result shown in literature that, in a point-to-point system, the rate in the uplink and downlink are the same [2]:

$$R_{k,DL} = R_{k,UL}. \tag{2.37}$$

Moreover, we can use the transformation of the correlation matrices for the point-to-point

system in order to transform the uplink MU-MIMO input correlation matrices \mathbf{Q}_k into downlink MU-MIMO input correlation matrices $\mathbf{\Omega}_k$. Firstly, we need to define the concept of *flipped channel*. In [15], the authors state that the capacity of a system \mathbf{Y}_1 with effective channel \mathbf{H}_{ef} is equal to the capacity of another system \mathbf{Y}_2 with effective channel \mathbf{H}_{ef}^H , in which this channel \mathbf{H}_{ef}^H is called the *flipped channel*. Applying this result in the rate of the uplink MU-MIMO channel (equation (2.35)), we have that the flipped channel is $\mathbf{A}_k^{-1/2}\mathbf{H}_k\mathbf{B}_k^{-1/2}$ and the input covariance matrix of the flipped case is defined as the matrix of dimension $N_t \times N_t$ $\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2}$. Thus, we can write:

$$\mathbf{R}_{k,\text{UL}}(\mathbf{H}_j) = \mathbf{R}_{k,\text{UL}}(\overline{\mathbf{H}}_j) = \log \left| \mathbf{I}_{N_r} + \mathbf{A}_k^{-1/2}\mathbf{H}_k\mathbf{B}_k^{-1/2}\overline{\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2}}\mathbf{B}_k^{-1/2}\mathbf{H}_k^H\mathbf{A}_k^{-1/2} \right|. \quad (2.38)$$

Substituting (2.36) and (2.38) into (2.37), we have:

$$\begin{aligned} \mathbf{R}_{k,\text{UL}} &= \mathbf{R}_{k,\text{DL}} \\ \log \left| \mathbf{I}_{N_r} + \mathbf{A}_k^{-1/2}\mathbf{H}_k\mathbf{B}_k^{-1/2}\overline{\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2}}\mathbf{B}_k^{-1/2}\mathbf{H}_k^H\mathbf{A}_k^{-1/2} \right| &= \log \left| \mathbf{I}_{N_r} + \mathbf{A}_k^{-1/2}\mathbf{H}_k\mathbf{\Omega}_k\mathbf{H}_k^H\mathbf{A}_k^{-1/2} \right| \end{aligned}$$

It is possible to notice that if we chose the downlink input covariance matrix of user k as being:

$$\mathbf{\Omega}_k = \mathbf{B}_k^{-1/2}\overline{\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2}}\mathbf{B}_k^{-1/2} \quad (2.39)$$

we can guarantee that equality of the rates $\mathbf{R}_{k,\text{UL}} = \mathbf{R}_{k,\text{DL}}$.

Defining the singular value decomposition (SVD) of the effective uplink channel as $\mathbf{B}_k^{-1/2}\mathbf{H}_k^H\mathbf{A}_k^{-1/2} = \mathbf{U}_{\text{ef}_k}\mathbf{\Sigma}_{\text{ef}_k}\mathbf{V}_{\text{ef}_k}^H$, the input covariance matrix of the flipped channel $\mathbf{A}_k^{-1/2}\mathbf{H}_k\mathbf{B}_k^{-1/2}$ is equal to $\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2} = \mathbf{U}_{\text{ef}_k}\mathbf{V}_{\text{ef}_k}^H\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2}\mathbf{V}_{\text{ef}_k}\mathbf{U}_{\text{ef}_k}^H$ [15]. Substituting this result in equation (2.39), we have that the downlink input covariance matrix can be written as a function of the uplink input covariance matrix as [14]:

$$\mathbf{\Omega}_k = \mathbf{B}_k^{-1/2}\mathbf{U}_{\text{ef}_k}\mathbf{V}_{\text{ef}_k}^H\mathbf{A}_k^{1/2}\mathbf{Q}_k\mathbf{A}_k^{1/2}\mathbf{V}_{\text{ef}_k}\mathbf{U}_{\text{ef}_k}^H\mathbf{B}_k^{-1/2}, \quad (2.40)$$

where these input covariance matrices achieve the same rates under the same sum power constraint.

Since we have the transformation that maps the uplink input covariance matrices obtained into the downlink input covariance matrices, we can use the duality theorem to evaluate the achievable sum rate of the downlink MU-MIMO channel, i.e.:

$$\begin{aligned} \mathbf{C}_{\text{DL}} &= \mathbf{C}_{\text{UL}} \\ \mathbf{C}_{\text{DL}} &= \max_{\{\mathbf{Q}_k\}_{k=1}^K; \mathbf{Q}_k \succeq 0, \sum_k \text{tr}(\mathbf{Q}_k) \leq P_b} \log \left| \mathbf{I}_{N_t} + \sum_{k=1}^K \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right|, \end{aligned} \quad (2.41)$$

where the optimization is performed over the uplink input covariance matrices \mathbf{Q}_k , $k = 1, \dots, K$ subject to the same sum power constraint P_b . This function is a convex function of the uplink input covariance matrices \mathbf{Q}_k [37]. Thus, we can obtain the uplink input covariance matrices through convex optimization tools, and after that, perform the transformation in (2.40) which maps the obtained uplink input covariance matrices into the desired downlink input covariance matrices [14].

As already commented, the work in [16] proposed a iterative waterfilling algorithm to compute the input correlation matrices of the uplink MU-MIMO channel with individual

power constraints. The proposed solution consists in optimizing the sum rate of the uplink MU-MIMO channel with individual power constraints with respect to the first variable while holding all other variables constant, then optimizing with respect to the second variable and so on, until reaching a globally optimum point. This is referred to as the block-coordinate ascent algorithm [38].

However, in order to compute the sum rate of an uplink MU-MIMO channel subject to a sum power constraint, i.e., computing (2.41), the water-filling procedure of the individual power constraint problem can not be used. In the individual power constraint problem, the water level of each user is determined individually and can differ from one user to another. However, in the sum power constraint problem, the water level of all users must be equal and we must update all covariances simultaneously to maintain a constant water-level. Therefore, the algorithm must update all K covariance matrices simultaneously during each step based on the covariance matrices from the previous step. The work in [14] finds an iterative waterfilling algorithm to solve the optimization problem of the uplink MU-MIMO channel with sum power constraint. The algorithm is complex and requires numerous calculations of SVDs. Moreover, they consider a perfect channel state information at the transmitter (CSIT).

Following, we present the CoMP structure which is the scenario considered in the present thesis.

2.4 Coordinated Multipoint (CoMP) Multiuser MIMO Channels

CoMP transmission/reception has been considered as a promising approach to improve coverage, cell-edge throughput, and/or system efficiency through suppressing the inter-channel interference (ICI) [4]. CoMP was originally proposed to extend communication coverage in environments hostile to radio propagation, such as in mines or tunnels. Saleh was the first to apply CoMP in the context of cellular mobile communications [39], and recent studies have shown that in addition to increasing coverage, CoMP can also reduce transmit power and hence co-channel interference and, as a result, increase system capacity [40] [41]. CoMP transmission can be categorized into two classes: 1) joint processing (JP), where data are simultaneously transmitted from multiple transmission points to a single user in order to improve the received signal quality and/or actively cancel interference from other users, and; 2) coordinated scheduling/coordinated beamforming (CS/CB), in which data are only available at the serving cell, but user scheduling/beamforming decisions are made with coordination among cells belonging to the CoMP group [4].

In JP CoMP systems, besides the serving cell, the cells in the coordination group are normally chosen as the ones that create the highest interference to the user. By this approach, the received signal levels will be improved and, at the same time, the ICI level will be decreased since part of it has been changed to a useful signal. A gain from JP CoMP is therefore obtained, not only by suppressing the interference sources but also by benefiting from it. Moreover, due to a higher number of transmit antennas involved in the joint transmission processing, better diversity gains can also be obtained.

Input-Output Signaling Model

In this thesis, the scenario considers the downlink of a MU-MIMO CoMP communication system composed by N_b base stations (BSs) and K cochannel users arbitrarily distributed within the system coverage area. Each BS b is equipped with N_t transmit antennas and each user k with N_r receive antennas. The BSs are connected through high capacity links (referred to as the backbone) to a central processing unit, which jointly processes the signals.

This characterizes a CoMP structure and, for simplicity, we will consider an ideal delayless, infinite-capacity backbone to connect all BSs to the central unit. Figure 2.6 shows this representation for a case with $N_b = K = 3$.

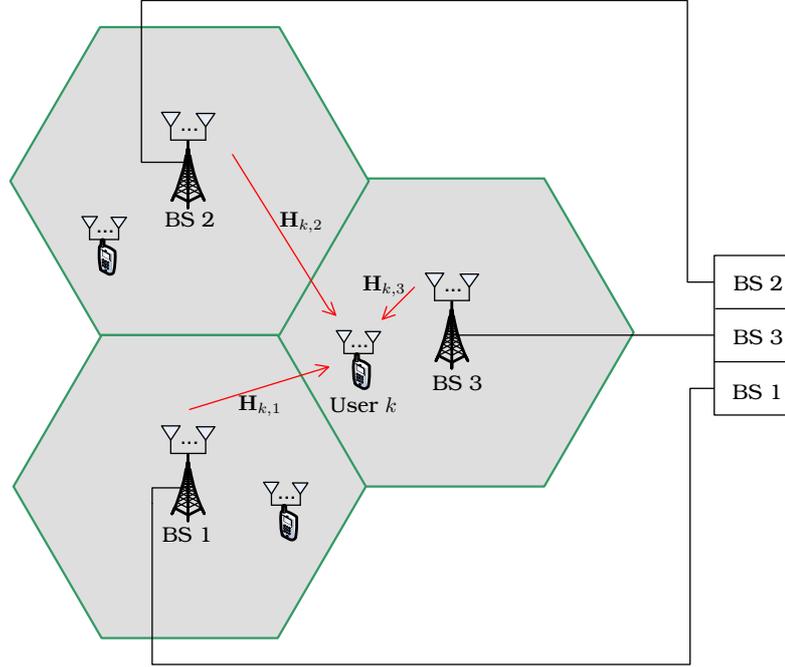


Figure 2.6: MU-MIMO CoMP system model with $N_b = K = 3$.

Moreover, the channel is considered frequency-flat and its spatial characteristics assume Kronecker-structured covariances with an $N_b N_t \times N_b N_t$ transmit covariance matrix \mathbf{R}_{t_k} and an $N_r \times N_r$ receive covariance matrix \mathbf{R}_{r_k} for each user k . Considering this, the channel matrix $\mathbf{H}_k[n]$ from all BSs to MS k at time n can be modeled as:

$$\mathbf{H}_k[n] = \sqrt{\frac{\kappa}{1+\kappa}} \bar{\mathbf{H}}_k + \sqrt{\frac{1}{1+\kappa}} \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{\mathbf{w}k}[n] \mathbf{R}_{t_k}^{1/2}, \quad (2.42)$$

where $\mathbf{H}_k = \begin{bmatrix} \mathbf{H}_{k,1} & \mathbf{H}_{k,2} & \dots & \mathbf{H}_{k,N_b} \end{bmatrix}_{N_r \times N_b N_t}$ is the joint channel matrix from all BSs to user k with $\mathbf{H}_{k,b}$ being the channel matrix from BS b to user k and $\mathbf{H}_{\mathbf{w}k}[n]$ is an $N_r \times N_t N_b$ small-scale fading channel matrix represented by an i.i.d. (spatially white) ZMCSCG matrix with unit variance.

As already mentioned, for JP CoMP systems, the transmit signal intended for each user k is spread over all N_b BSs. Omitting the time index n for simplicity of notation, the transmit signal for user k can be expressed as $\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_{k,1}^T & \mathbf{x}_{k,2}^T & \dots & \mathbf{x}_{k,N_b}^T \end{bmatrix}^T$, where $\mathbf{x}_{k,b}$ is the signal transmitted from BS b to user k . The signal \mathbf{y}_k received at user k is

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \sum_{j \neq k} \mathbf{H}_k \mathbf{x}_j + \boldsymbol{\eta}_k, \quad (2.43)$$

where $\boldsymbol{\eta}_k$ refers to a ZMCSCG noise vector with identity-covariance.

As we can observe in (2.43), the signal received by user k involves one term dependent on the desired signal \mathbf{x}_k and another dependent on the interfering signals \mathbf{x}_j . In this work, we assume that the transmitter knows the propagation delay for each BS-user pair. Thus, in (2.43), the signal \mathbf{x}_k transmitted from all BSs to user k can be pre-compensated for the different delays to user k associated with each BS. However, \mathbf{x}_j in the received signal by user

k can not be pre-compensated in the joint transmission, thus we have that \mathbf{x}_j is received by user k in an asynchronous way. Since the user k is not interested in correctly detecting the streams intended to user j (i.e. decoding correctly \mathbf{x}_j), we can simply view \mathbf{x}_j as the data of some virtual synchronous interfering users. Some recent works have considered that this interference signal is fundamentally asynchronous [40, 42]. These works analyze the impact of this asynchronism on existent precoding algorithms that ignore this behavior and suggest how to mitigate it. In this thesis we prefer to consider a synchronous interference and leave the study with the asynchronous one for future work. Moreover, this consideration is also often used in literature [43, 44].

Let L_k denote the number of data streams intended for user k . For each user k , an $N_t N_b \times L_k$ precoding matrix $\mathbf{T}_k = [\mathbf{T}_{k,1}^T \quad \mathbf{T}_{k,2}^T \quad \dots \quad \mathbf{T}_{k,N_b}^T]^T$ is designed based on the characteristics of \mathbf{H}_k , where $\mathbf{T}_{k,b}$ is the $N_t \times L_k$ matrix representing the precoder of BS b for user k . Thus, the transmitted signal from all BSs to user k is $\mathbf{x}_k = \mathbf{T}_k \mathbf{s}_k$, where $\mathbf{s}_k = [s_{k,1} \quad s_{k,2} \quad \dots \quad s_{k,L_k}]^T$ contains the data streams intended for user k . For simplicity, streams are assumed to have i.i.d. zero-mean unit-variance complex Gaussian entries, i.e., Gaussian signaling is assumed.

Since we consider a CoMP structure, each BS will have its own power constraint. Thus, we consider that:

$$\mathbb{E} \left\{ \sum_{k=1}^K \text{tr}(\mathbf{x}_{k,b} \mathbf{x}_{k,b}^H) \right\} = \text{tr} \left(\sum_{k=1}^K \mathbb{E}\{\mathbf{T}_{k,b} \mathbf{T}_{k,b}^H\} \right) = \sum_{k=1}^K \text{tr}(\boldsymbol{\Omega}_{k,b}) \leq P_b \quad b = 1, 2, \dots, N_b, \quad (2.44)$$

where P_b is the power constraint of BS b and $\boldsymbol{\Omega}_{k,b} = \mathbb{E}\{\mathbf{T}_{k,b} \mathbf{T}_{k,b}^H\}$ is the downlink input covariance matrix of user k and BS b .

This scenario is the considered one in this thesis and it is of great interest due to some advantages:

- ▶ Assuming JP in a MU-MIMO system, the overall transmit array is distributed among the cooperative BSs. In the resulting channel for user k , all subchannel matrices corresponding to the transmission from each BS to user k ($\mathbf{H}_{k,b}$ with $b = 1, \dots, N_b$) are independent of each other. Thus, the total number of independent links is given by $\sum_{b=1}^{N_b} \text{rank}(\mathbf{H}_{k,b})$ which is assured to be at least equal to N_b . Therefore, if $N_b \geq N_r$, the joint channel matrix \mathbf{H}_k of user k will always be full-rank. Moreover, even if local fading occurs at each BS, the channel conditioning will not be greatly degraded as the fading among different transmit antennas at different BSs is still uncorrelated.
- ▶ Shadowing is a position-dependent factor and thus transmit antennas placed at the same BS are generally subject to the same attenuation. For single-cell processing, strong shadowing conditions may degrade the capacity significantly. On the other hand, JP CoMP structures can offer macrodiversity protection for shadowing impairments as BSs are independent of each other and consequently there is a much lower probability that all $N_b N_t$ antennas be under deep fading compared to the case where the entire antenna array is co-located at the same BS.

Thus, the considered CoMP scenario is interesting due to the consideration of the heterogeneous nature of the networks to be deployed and due to the diversity of sources obtained with the JP scheme.

2.5 Summary

In this chapter the fundamental characteristics of the wireless channel were overviewed. In the sequel, we discussed about the multiple-input multiple-output (MIMO) parameters and its spatial characteristics, such as the channel mean and the channel covariance. After that, we detailed the multi-user MIMO (MU-MIMO) channel, its categories: uplink and downlink and the capacity region for each category of the MU-MIMO channel. Moreover, the duality theorem between the uplink MU-MIMO and the downlink MU-MIMO was explained. Finally, we described the MU-MIMO coordinated multipoint (CoMP) channel model, the input-output signal model that will be considered in this thesis and its advantages. In the next chapter, a channel model considering the channel statistics will be proposed for our considered scenario.

Chapter 3

Multi-user MIMO CoMP Channel Model

A major challenge in wireless communication is the time-variation of the channel. This time-variation creates difficulty in obtaining channel information, which is required for achieving the best performance. While the channel can be measured directly at the receiver with sufficient accuracy, the transmitter must obtain channel information indirectly, using either reciprocity or feedback. In a time-varying channel, the delay involved in such a process can degrade the information accuracy.

Most of transmit processing techniques rely on the fact that the transmitter knows perfectly the channel [5, 6, 45–48]. However, the random time-varying wireless medium makes it very difficult and often expensive to obtain perfect channel state information at the transmitter (CSIT). In closed-loop methods, CSIT is degraded by the limited feedback resources, associated feedback delays, and scheduling lags [49]. In open-loop methods, antenna calibration errors and turn-around time lags limit CSIT accuracy [50]. Therefore, the transmitter has often only partial channel information and schemes exploiting partial CSIT are both important and necessary.

The first use of partial CSIT has been introduced in [51], where the limits on system performance of a transmit array that uses partial side information to transmit to a single user has been quantified. In [51], the feedback information consists of an N -bit description of the channel parameters vector. The space of channel parameter vectors is quantized in 2^N regions and, for each region, the transmitter selects the transmission strategy that maximizes the expected SNR. However, this work does not consider that the antennas might be too closely spaced, making the channel parameters vector very correlated.

Some limited feedback multi-user MIMO (MU-MIMO) schemes let users quantize some function of the channel matrix and send this information to the base station [7] [8]. The problem is that, when the channel is quantized, the users' signals can not be perfectly orthogonalized due to inherent quantization errors. This leads to rate degradation as the SNR increases. In order to avoid this problem, schemes are proposed that directly select a quantized precoder from a codebook at the receiver and feedback the precoder index to the transmitter [52, 53]. However, it is too difficult to design the precoder codebook, which must take into account the channel distribution and the precoder design.

Other approaches focused on the mean matrix of the channel [54], on the covariance matrix of the channel as a form of feedback [55], or on both mean and covariance matrices [56].

This information reveals much about the slow fading and the mean separability of the users. In [54] optimum transmit strategies are proposed and the gains obtained are verified with the use of two imperfect channel feedback schemes: mean feedback, in which the channel state information resides in the mean of the channel distribution, with the channel covariance modeled as an white random matrix; and covariance feedback, in which the channel state information is assumed to be varying too rapidly to allow tracking its mean, so that the mean is set to zero and the information is captured by a nonwhite covariance matrix. It is interesting to comment that the considered system model in [54] was a single user one and that the optimum transmission strategies presented were based on random coding arguments, i.e., practical coding strategies for exploiting partial knowledge of the spatial channel were not considered. In [55], the author considers the narrowband point-to-point communication system employing multiple-element antenna arrays at both transmitter and receiver with channel covariance feedback. Under covariance feedback, the receiver is assumed to have perfect CSI while, at the transmitter, the channel matrix is modeled as consisting of zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables with known covariances. Most of the previous works have results for either mean information or the covariance information available to the transmitter, but not both.

As previously mentioned, the time-variation of channel degrades the accuracy of the channel information obtained at the transmitter. This occurs due to the delay involved in the process of the channel estimation at the transmitter when using reciprocity or feedback. The channel model proposed in [13] takes into account this channel time-variation. It relies on the stochastic processes and estimation theories. Derived from a potentially outdated channel measurement and from the channel statistics (mean and covariance), this proposed CSIT model consists of a channel estimate and its error covariance, which act as the effective channel mean and covariance, respectively. Both parameters depend on a temporal correlation factor, indicating the CSIT quality. Depending on this quality, the model switches smoothly from perfect to statistical channel information. This proposed CSIT model is applicable to all Gaussian random channels, however it was only proposed for multiple-input multiple-output (MIMO) single-user system. In this chapter, as one contribution of this thesis, we propose a generalization of the work presented in [13] for a more general multicell multiuser context, i.e., for the MU-MIMO coordinated multipoint (CoMP) channel model.

3.1 Proposed Statistical MU-MIMO CoMP Channel Model

Before we present our proposition, we discuss some concepts related to the temporal variation of the channel that will be taken into account in the proposed channel model.

3.1.1 Channel Auto-covariance

The channel auto-covariance characterizes how rapidly the channel decorrelates in time. Assuming stationarity, the auto-covariance between two channel samples $\mathbf{H}_k[m]$ and $\mathbf{H}_k[m+n]$ depends on the time difference but not on the absolute time, and is given by:

$$\mathbf{R}_k[n] = \mathbb{E} \left\{ \tilde{\mathbf{h}}_k[m] \tilde{\mathbf{h}}_k[m+n]^H \right\}, \quad (3.1)$$

where $\tilde{\mathbf{h}}_k = \text{vec}(\tilde{\mathbf{H}}_k)$ and $\tilde{\mathbf{H}}_k[m]$ is the random part of the channel $\mathbf{H}_k[m]$ from all BSs to user k represented by a joint ZMCSCG matrix and already defined in (2.11) on the page 9.

When $n = 0$, this auto-covariance coincides with the covariance of the channel from all transmit antennas to user k , i.e., $\mathbf{R}_k[0] = \mathbb{E} \left\{ \tilde{\mathbf{h}}_k \tilde{\mathbf{h}}_k^H \right\}$, which is given in equation (2.12) on the page 9; and when n becomes large it eventually decays to zero. For a MU-MIMO channel, the

channel covariance $\mathbf{R}_k[0]$ of user k captures the spatial correlation among transmit antennas and the receive antennas of user k , while the auto-covariance at a non-zero delay $\mathbf{R}_k[n]$ captures both channel spatial and temporal correlations.

Based on the premise that the channel temporal statistics are the same for all antenna pairs of a same user k , it may be assumed that the channel gains between all the $N_t N_b$ transmit antennas and N_r receive antennas of that user k have the same temporal correlation function. Similar assumptions for MIMO temporal correlation have also been made in [57, 58]. Then, it is possible to separate the temporal correlation from the spatial correlation and the auto-covariance becomes their product as

$$\mathbf{R}_k[n] = \rho_k[n] \mathbf{R}_k[0] \quad (3.2)$$

where $\rho_k[\Delta n]$ is the channel temporal correlation and it is a function of the Doppler spread f_{d_k} of user k at time delay Δn . In Jakes' model, the channel temporal correlation is given by [3, 22]:

$$\rho_k[\Delta n] = J_0(2\pi\Delta n f_{d_k}), \quad (3.3)$$

where $J_0(\cdot)$ is the zero-th order Bessel function of the first kind.

3.1.2 Channel Estimation at the Transmitter and Proposed Statistical CSIT

Firstly, consider that $\hat{\mathbf{H}}_k[n]$ is the channel estimate of user k at time n and $\mathbf{E}_k[n]$ is the estimation error matrix with correlation $\mathbf{R}_{\mathbf{e}_k}[n]$. The CSIT model can be considered as:

$$\mathbf{H}_k[n] = \hat{\mathbf{H}}_k[n] + \mathbf{E}_k[n], \quad (3.4)$$

$$\mathbf{R}_{\mathbf{e}_k}[n] = \mathbb{E} \{ \mathbf{e}_k[n] \mathbf{e}_k[n]^H \} \quad (3.5)$$

where $\mathbf{e}_k[n] = \text{vec}(\mathbf{E}_k[n])$.

Moreover, we assume that the transmitter has an initial channel measurement at time $n = 0$ equal to $\mathbf{H}_k[0]$ and relevant channel statistics (the channel mean $\bar{\mathbf{H}}_k$ and the channel covariance matrix $\mathbf{R}_k[0]$) for each user k .

Considering that $\mathbf{h}_k[0] = \text{vec}(\mathbf{H}_k[0])$ and $\mathbf{h}_k[n] = \text{vec}(\mathbf{H}_k[n])$ are independent random variables, then there is little that can be said about the value assumed by one random variable ($\mathbf{h}_k[n]$) when the value assumed by the other ($\mathbf{h}_k[0]$) is known or measured [59]. However, $\mathbf{h}_k[0]$ and $\mathbf{h}_k[n]$ are dependent channel vectors, since the channel is correlated in the time domain. Thus, it is possible to say something about the value assumed by $\mathbf{h}_k[n]$ when the value assumed by $\mathbf{h}_k[0]$ is known or measured.

Therefore, an estimate of the vector $\mathbf{h}_k[n]$, say $\hat{\mathbf{h}}_k[n]$, can be described as a function of the value assumed by other dependent vector, i.e., $\mathbf{h}_k[0]$ [59]:

$$\hat{\mathbf{h}}_k[n] = f(\mathbf{h}_k[0]). \quad (3.6)$$

The channel vector known at the transmitter $\mathbf{h}_k[0]$ is considered valid in the estimation of the channel vector at time n $\mathbf{h}_k[n]$ during N_S symbol times. Thus, in each N_S symbol times, the channel vector $\mathbf{h}_k[0]$ need to be updated at the transmitter. This procedure decreases the amount of feedback needed to estimated the channel vector at time n $\mathbf{h}_k[n]$ compared to the case when the transmitter has the perfect channel, i.e., when $\mathbf{h}_k[n]$ is sent to transmitter in each symbol time through the feedback channel.

The challenge in such estimation procedure is to suitably choose the function $f(\cdot)$ to yield an estimate $\hat{\mathbf{h}}_k[n]$ that satisfies a desired optimality criterion. For signal processing, communications and control, one of the most important criteria is the least-mean-squares [59]. With the mean-square-error criterion, it can be shown that the optimum estimate is the conditional expectation of $\mathbf{h}_k[n]$ given $\mathbf{h}_k[0]$ represented by [59]:

$$\hat{\mathbf{h}}_k[n] = \mathbb{E} \{ \mathbf{h}_k[n] | \mathbf{h}_k[0] \}. \quad (3.7)$$

Calculating this expectation requires full knowledge of the joint probability density function of $\{\mathbf{h}_k[n], \mathbf{h}_k[0]\}$, which is often hard to obtain. If, however, we deliberately restrict the estimation function $f(\cdot)$ to be a linear function of the observations, then it turns out that all we shall need is knowledge of the first- and second-order statistics given by $\mathbb{E} \{ \mathbf{h}_k[n] \}$, $\mathbb{E} \{ \mathbf{h}_k[0] \}$, $\mathbb{E} \{ \mathbf{h}_k[n] \mathbf{h}_k[n]^H \}$, $\mathbb{E} \{ \mathbf{h}_k[0] \mathbf{h}_k[0]^H \}$ and $\mathbb{E} \{ \mathbf{h}_k[n] \mathbf{h}_k[0]^H \}$. This constraint is not so hard since when $\{\mathbf{h}_k[n], \mathbf{h}_k[0]\}$ are jointly Gaussian, an assumption that is often reasonable, the unconstrained least-mean-squares estimation function is indeed linear [59].

Channel with Zero-Mean Assumption

Firstly, we will consider that the vectors $\mathbf{h}_k[0]$ and $\mathbf{h}_k[n]$ are $N_r N_b N_t \times 1$ complex random variables with zero-mean. In a posterior analysis, we will take into account non-zero mean vectors. Our objective is to estimate the value assumed by variable $\mathbf{h}_k[n]$ given the value assumed by random variable $\mathbf{h}_k[0]$. As already commented, we assume that $\mathbf{h}_k[n]$ is estimated by a linear combination of the form

$$\hat{\mathbf{h}}_k[n] = \mathbf{F} \mathbf{h}_k[0], \quad (3.8)$$

where $\mathbf{F} \in \mathbb{C}^{N_r N_b N_t \times N_r N_b N_t}$ is a coefficient matrix we wish to determine so as to minimize the resulting error covariance matrix $\mathbf{R}_{\mathbf{e}_k}[n]$ for user k at time n ,

$$\mathbf{F}^O = \arg \min_{\mathbf{F}} \mathbf{R}_{\mathbf{e}_k}[n] = \arg \min_{\mathbf{F}} \mathbb{E} \left\{ \left(\mathbf{h}_k[n] - \hat{\mathbf{h}}_k[n] \right) \left(\mathbf{h}_k[n] - \hat{\mathbf{h}}_k[n] \right)^H \right\}. \quad (3.9)$$

Thus, it is desired to find \mathbf{F}^O such that for every $\mathbf{F} \in \mathbb{C}^{N_r N_b N_t \times N_r N_b N_t}$ we obtain

$$\mathbf{R}_{\mathbf{e}_k}[n] \triangleq \mathbb{E} \left\{ \left(\mathbf{h}_k[n] - \mathbf{F} \mathbf{h}_k[0] \right) \left(\mathbf{h}_k[n] - \mathbf{F} \mathbf{h}_k[0] \right)^H \right\} \succeq \mathbf{R}_{\mathbf{e}_k}^O[n] \quad (3.10)$$

where $\mathbf{R}_{\mathbf{e}_k}^O[n]$ is the minimum error covariance matrix and $\mathbf{X} \succeq \mathbf{Y}$ means that $\mathbf{X} - \mathbf{Y}$ is positive semi-definite.

This problem is equivalent to requiring that

$$\mathbf{a} \mathbf{R}_{\mathbf{e}_k}[n] \mathbf{a}^H \geq \mathbf{a} \mathbf{R}_{\mathbf{e}_k}^O[n] \mathbf{a}^H \quad (3.11)$$

for every \mathbf{F} and for every non null row vector \mathbf{a} [59].

In this problem, \mathbf{F}^O is a solution of the optimization problems (3.9) and (3.10) if, and only if, for all vectors \mathbf{a} , $\mathbf{a} \mathbf{F}^O$ is a minimum of $\mathbf{a} \mathbf{R}_{\mathbf{e}_k}[n] \mathbf{a}^H$ [59], where

$$\begin{aligned} \mathbf{a} \mathbf{R}_{\mathbf{e}_k}[n] \mathbf{a}^H &= \mathbf{a} \mathbb{E} \left\{ \left(\mathbf{h}_k[n] - \mathbf{F} \mathbf{h}_k[0] \right) \left(\mathbf{h}_k[n] - \mathbf{F} \mathbf{h}_k[0] \right)^H \right\} \mathbf{a}^H \\ &= \mathbf{a} \mathbb{E} \left\{ \mathbf{h}_k[n] \mathbf{h}_k^H[n] - \mathbf{h}_k[n] \mathbf{h}_k^H[0] \mathbf{F}^H - \mathbf{F} \mathbf{h}_k[0] \mathbf{h}_k^H[n] + \mathbf{F} \mathbf{h}_k[0] \mathbf{h}_k^H[0] \mathbf{F}^H \right\} \mathbf{a}^H \end{aligned} \quad (3.12)$$

From equations (2.12) and (3.1), we can rewrite (3.12) as:

$$\mathbf{a} \mathbf{R}_{\mathbf{e}_k}[n] \mathbf{a}^H = \mathbf{a} \left(\mathbf{R}_k[0] - \mathbf{R}_k[n]^H \mathbf{F}^H - \mathbf{F} \mathbf{R}_k[n] + \mathbf{F} \mathbf{R}_k[0] \mathbf{F}^H \right) \mathbf{a}^H \quad (3.13)$$

where $\mathbf{R}_k[0]$ and $\mathbf{R}_k[n]$ are the channel covariance and the channel auto-covariance for user k , respectively.

Note that $\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H$ is a scalar function of a complex-valued (row) vector quantity $\mathbf{a}\mathbf{F}$. In order to evaluate the differential of $\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H$ with respect to $\mathbf{a}\mathbf{F}$ and set the derivative equal to zero at $\mathbf{F} = \mathbf{F}^O$, we perform the gradient given by:

$$\left. \frac{\nabla (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\nabla \mathbf{a}\mathbf{F}} \right|_{\mathbf{F}=\mathbf{F}^O} = 0$$

From [60], the gradient of a scalar function of a complex-valued vector is given by:

$$\begin{aligned} \frac{\nabla (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\nabla \mathbf{a}\mathbf{F}} &= 2. \frac{\partial (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\partial (\mathbf{a}\mathbf{F})^*} \\ \frac{\nabla (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\nabla \mathbf{a}\mathbf{F}} &= 2. \frac{\partial (\mathbf{a} (\mathbf{R}_k[0] - \mathbf{R}_k[n]^H \mathbf{F}^H - \mathbf{F}\mathbf{R}_k[n] + \mathbf{F}\mathbf{R}_k[0]\mathbf{F}^H) \mathbf{a}^H)}{\partial (\mathbf{a}\mathbf{F})^*} \\ \frac{\nabla (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\nabla \mathbf{a}\mathbf{F}} &= 2. \left(\frac{\partial (\mathbf{a}\mathbf{R}_k[0]\mathbf{a}^H)}{\partial (\mathbf{a}\mathbf{F})^*} - \frac{\partial (\mathbf{a}\mathbf{R}_k[n]^H \mathbf{F}^H \mathbf{a}^H)}{\partial (\mathbf{a}\mathbf{F})^*} - \frac{\partial (\mathbf{a}\mathbf{F}\mathbf{R}_k[n]\mathbf{a}^H)}{\partial (\mathbf{a}\mathbf{F})^*} + \frac{\partial (\mathbf{a}\mathbf{F}\mathbf{R}_k[0]\mathbf{F}^H \mathbf{a}^H)}{\partial (\mathbf{a}\mathbf{F})^*} \right) \\ \frac{\nabla (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\nabla \mathbf{a}\mathbf{F}} &= 2. (0 - 0 - \mathbf{R}_k[n]\mathbf{a}^H + \mathbf{R}_k[0]\mathbf{F}^H \mathbf{a}^H) \\ \frac{\nabla (\mathbf{a}\mathbf{R}_{\mathbf{e}_k}[n]\mathbf{a}^H)}{\nabla \mathbf{a}\mathbf{F}} &= -2\mathbf{R}_k[n]\mathbf{a}^H + 2\mathbf{R}_k[0]\mathbf{F}^H \mathbf{a}^H \end{aligned} \quad (3.14)$$

Setting the result of the derivative obtained in (3.14) equal to zero at $\mathbf{F} = \mathbf{F}^O$ we have:

$$\begin{aligned} -2\mathbf{R}_k[n]\mathbf{a}^H + 2\mathbf{R}_k[0]\mathbf{F}^H \mathbf{a}^H &\Big|_{\mathbf{F}=\mathbf{F}^O} = 0 \\ \mathbf{R}_k[0]\mathbf{F}^{OH} \mathbf{a}^H &= \mathbf{R}_k[n]\mathbf{a}^H \\ \mathbf{R}_k[0]\mathbf{F}^{OH} \mathbf{a}^H &= \mathbf{R}_k[n]\mathbf{a}^H \\ \mathbf{F}^{OH} &= \mathbf{R}_k[0]^{-1}\mathbf{R}_k[n] \\ \mathbf{F}^O &= \mathbf{R}_k[n]^H \mathbf{R}_k[0]^{-H} \\ \mathbf{F}^O &= \mathbf{R}_k[n]^H \mathbf{R}_k[0]^{-1} \end{aligned} \quad (3.15)$$

The corresponding minimum error covariance matrix $\mathbf{R}_{\mathbf{e}_k}^O[n]$ can be written as

$$\begin{aligned} \mathbf{R}_{\mathbf{e}_k}^O[n] &= \mathbb{E} \left\{ \left(\mathbf{h}_k[n] - \hat{\mathbf{h}}_k[n] \right) \left(\mathbf{h}_k[n] - \hat{\mathbf{h}}_k[n] \right)^H \right\} \\ &= \mathbb{E} \left\{ \left(\mathbf{h}_k[n] - \mathbf{F}^O \mathbf{h}_k[0] \right) \left(\mathbf{h}_k[n] - \mathbf{F}^O \mathbf{h}_k[0] \right)^H \right\} \\ &= \mathbb{E} \left\{ \left(\mathbf{h}_k[n] - \mathbf{F}^O \mathbf{h}_k[0] \right) \mathbf{h}_k[n]^H - \left(\mathbf{h}_k[n] - \mathbf{F}^O \mathbf{h}_k[0] \right) \mathbf{h}_k[0]^H \mathbf{F}^{OH} \right\} \\ &= \mathbb{E} \left\{ \mathbf{h}_k[n]\mathbf{h}_k[n]^H - \mathbf{F}^O \mathbf{h}_k[0]\mathbf{h}_k[n]^H - \mathbf{h}_k[n]\mathbf{h}_k[0]^H \mathbf{F}^{OH} + \mathbf{F}^O \mathbf{h}_k[0]\mathbf{h}_k[0]^H \mathbf{F}^{OH} \right\} \\ &= \mathbf{R}_k[0] - \mathbf{F}^O \mathbf{R}_k[n] - \mathbf{R}_k[n]^H \mathbf{F}^{OH} + \mathbf{F}^O \mathbf{R}_k[0]\mathbf{F}^{OH}. \end{aligned} \quad (3.16)$$

Substituting the result obtained in (3.15) into (3.16), we obtain:

$$\begin{aligned} \mathbf{R}_{\mathbf{e}_k}^O[n] &= \mathbf{R}_k[0] - \mathbf{F}^O \mathbf{R}_k[n] - \mathbf{R}_k[n]^H \mathbf{F}^{OH} + \mathbf{R}_k[n]^H \mathbf{F}^{OH} \\ &= \mathbf{R}_k[0] - \mathbf{F}^O \mathbf{R}_k[n], \end{aligned} \quad (3.17)$$

and substituting (3.15) in (3.17) leads to

$$\mathbf{R}_{\mathbf{e}_k^O}[n] = \mathbf{R}_k[0] - \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} \mathbf{R}_k[n]. \quad (3.18)$$

Now, we have found the linear function which minimizes the mean-square-error (\mathbf{F}^O) and the minimum error covariance ($\mathbf{R}_{\mathbf{e}_k^O}[n]$). Using the equation (3.15) in (3.8), we have that the channel estimate $\hat{\mathbf{h}}_k[n]$ at time n for user k is given by:

$$\hat{\mathbf{h}}_k[n] = \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} \mathbf{h}_k[0]. \quad (3.19)$$

Using the homogeneous channel temporal correlation assumption given in (3.2), the channel estimate and its minimum error covariance become, respectively,

$$\begin{aligned} \hat{\mathbf{h}}_k[n] &= \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} \mathbf{h}_k[0] \\ &= (\rho_k[n] \mathbf{R}_k[0])^H (\mathbf{R}_k[0])^{-1} \mathbf{h}_k[0] \\ &= \rho_k[n] \mathbf{R}_k[0]^H (\mathbf{R}_k[0])^{-1} \mathbf{h}_k[0] \\ &= \rho_k[n] \mathbf{h}_k[0] \\ \hat{\mathbf{H}}_k[n] &= \rho_k[n] \mathbf{H}_k[0] \end{aligned} \quad (3.20)$$

and

$$\begin{aligned} \mathbf{R}_{\mathbf{e}_k^O}[n] &= \mathbf{R}_k[0] - \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} \mathbf{R}_k[n] \\ &= \mathbf{R}_k[0] - (\rho_k[n] \mathbf{R}_k[0])^H (\mathbf{R}_k[0])^{-1} (\rho_k[n] \mathbf{R}_k[0]) \\ &= \mathbf{R}_k[0] - \rho_k[n] \rho_k[n] \mathbf{R}_k[0] \\ &= (1 - \rho_k^2[n]) \mathbf{R}_k[0]. \end{aligned} \quad (3.21)$$

where the channel estimate has been taken in a matricial form, without loss of generality.

Thus, the CSIT at time n for user k is given by the channel estimate $\hat{\mathbf{H}}_k[n]$ (equation (3.20)) and by its minimum error covariance $\mathbf{R}_{\mathbf{e}_k^O}[n]$ (equation (3.21)). They effectively work as channel mean and channel covariance at a delay n for user k . Thus $\hat{\mathbf{H}}_k[n]$ and $\mathbf{R}_{\mathbf{e}_k^O}[n]$ are also referred to as the effective mean and effective covariance, respectively. Moreover, we can notice from (3.20) and (3.21) that the CSIT for each user k is simply characterized as a function of $\rho_k[n]$, the initial channel measurement $\mathbf{H}_k[0]$ of that user k and the covariance channel matrix $\mathbf{R}_k[0]$. The channel estimate $\hat{\mathbf{H}}_k[n]$ is a linear function of the initial channel measurement and the error covariance $\mathbf{R}_{\mathbf{e}_k^O}[n]$ is a linear function of the channel covariance.

Substituting in (3.4) the result obtained for the channel estimate, i.e., the equation (3.20), we have:

$$\mathbf{H}_k[n] = \rho_k[n] \mathbf{H}_k[0] + \mathbf{E}_k[n]. \quad (3.22)$$

From the covariance decomposition by Kronecker product (2.16), the error covariance matrix given in (3.21) can similarly be decomposed in effective antenna correlation matrices as

$$\begin{aligned} \mathbf{R}_{\mathbf{e}_k^O, \mathbf{t}_k}[n] &= \sqrt{1 - \rho_k^2[n]} \mathbf{R}_{\mathbf{t}_k}, \text{ and} \\ \mathbf{R}_{\mathbf{e}_k^O, \mathbf{r}_k}[n] &= \sqrt{1 - \rho_k^2[n]} \mathbf{R}_{\mathbf{r}_k}, \end{aligned} \quad (3.23)$$

which follow the Kronecker structure, where $\mathbf{R}_{\mathbf{t}_k}$ is the transmit covariance matrix and $\mathbf{R}_{\mathbf{r}_k}$ is the receive covariance matrix of user k , given by equations (2.14) and (2.15).

Thus, we can rewrite the equation (3.22) as:

$$\mathbf{H}_k[n] = \rho_k[n]\mathbf{H}_k[0] + \mathbf{R}_{\mathbf{e},\mathbf{r}_k^O}[n]^{1/2}\mathbf{H}_{\mathbf{w}k}[n]\mathbf{R}_{\mathbf{e},\mathbf{t}_k^O}[n]^{1/2}, \quad (3.24)$$

where $\mathbf{H}_{\mathbf{w}k}[n]$ is an $N_r \times N_b N_t$ channel matrix whose entries are independent identically distributed (i.i.d.) ZMCSCG with unit-variance.

Channel with Nonzero-Mean Assumption

So far in our channel estimate discussion, we have assumed that the channel vector $\mathbf{h}_k[n]$ is a zero-mean random one. However, in our channel model, there is a line-of-sight (LOS) component which determines a non-zero channel mean (see equation (2.11)). In order to consider the non-zero channel mean in the channel estimation, the simplest way is to proceed by centering the random variables $\mathbf{h}_k[0]$ and $\mathbf{h}_k[n]$. Let us define:

$$\mathbf{h}_k^{\text{cent}}[0] = \mathbf{h}_k[0] - \bar{\mathbf{h}}_k \quad \text{and} \quad \mathbf{h}_k^{\text{cent}}[n] = \mathbf{h}_k[n] - \bar{\mathbf{h}}_k \quad (3.25)$$

where $\mathbf{h}_k[0]$ and $\mathbf{h}_k[n]$ are the non-zero mean random vectors and $\bar{\mathbf{h}}_k$ is the channel mean vector for user k .

The transformation between $\{\mathbf{h}_k[0], \mathbf{h}_k[n]\}$ and $\{\mathbf{h}_k^{\text{cent}}[0], \mathbf{h}_k^{\text{cent}}[n]\}$ is reversible, so there is no loss of information in making such a transformation. Evaluating the covariance matrix of the $\mathbf{h}_k^{\text{cent}}[0]$ and the auto-covariance matrix between the new matrices $\mathbf{h}_k^{\text{cent}}[0]$ and $\mathbf{h}_k^{\text{cent}}[n]$, we have:

$$\mathbb{E}\{\mathbf{h}_k^{\text{cent}}[0]\mathbf{h}_k^{\text{cent}H}[0]\} = \mathbb{E}\{(\mathbf{h}_k[0] - \bar{\mathbf{h}}_k)(\mathbf{h}_k[0] - \bar{\mathbf{h}}_k)^H\} = \mathbb{E}\{\mathbf{h}_k[0]\mathbf{h}_k^H[0]\} - \bar{\mathbf{h}}_k\bar{\mathbf{h}}_k^H, \quad (3.26)$$

and

$$\mathbb{E}\{\mathbf{h}_k^{\text{cent}}[0]\mathbf{h}_k^{\text{cent}H}[n]\} = \mathbb{E}\{(\mathbf{h}_k[0] - \bar{\mathbf{h}}_k)(\mathbf{h}_k[n] - \bar{\mathbf{h}}_k)^H\} = \mathbb{E}\{\mathbf{h}_k[0]\mathbf{h}_k^H[n]\} - \bar{\mathbf{h}}_k\bar{\mathbf{h}}_k^H. \quad (3.27)$$

From equations (3.26) and (3.27), we can define:

$$\mathbb{E}\{\mathbf{h}_k^{\text{cent}}[0]\mathbf{h}_k^{\text{cent}H}[0]\} \triangleq \mathbf{R}_k[0], \text{ covariance matrix of } \mathbf{h}_k[0] \quad (3.28)$$

$$\mathbb{E}\{\mathbf{h}_k^{\text{cent}}[0]\mathbf{h}_k^{\text{cent}H}[n]\} \triangleq \mathbf{R}_k[n], \text{ auto-covariance matrix between } \mathbf{h}_k[0] \text{ and } \mathbf{h}_k[n]. \quad (3.29)$$

We can write the zero mean channel estimate $\hat{\mathbf{h}}_k^{\text{cent}}[n]$ from the equation obtained in (3.19) as:

$$\hat{\mathbf{h}}_k^{\text{cent}}[n] = (\mathbb{E}\{\mathbf{h}_k^{\text{cent}}[0]\mathbf{h}_k^{\text{cent}H}[n]\})^H (\mathbb{E}\{\mathbf{h}_k^{\text{cent}}[0]\mathbf{h}_k^{\text{cent}H}[0]\})^{-1} \mathbf{h}_k^{\text{cent}}[0]. \quad (3.30)$$

From equations (3.28) and (3.29), we can rewrite the equation (3.30) as:

$$\hat{\mathbf{h}}_k^{\text{cent}}[n] = \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} \mathbf{h}_k^{\text{cent}}[0]. \quad (3.31)$$

Substituting the equations that define the centered random variables $\mathbf{h}_k[0]$ and $\mathbf{h}_k[n]$ given

by (3.25) into (3.31),

$$\begin{aligned}\widehat{\mathbf{h}}_k[n] - \bar{\mathbf{h}}_k &= \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} (\mathbf{h}_k[0] - \bar{\mathbf{h}}_k) \\ \widehat{\mathbf{h}}_k[n] &= \bar{\mathbf{h}}_k + \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} (\mathbf{h}_k[0] - \bar{\mathbf{h}}_k) \\ \widehat{\mathbf{h}}_k[n] &= \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1} \mathbf{h}_k[0] + (1 - \mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1}) \bar{\mathbf{h}}_k.\end{aligned}\quad (3.32)$$

We have already shown that, using the homogeneous channel temporal correlation assumption, the term $\mathbf{R}_k[n]^H (\mathbf{R}_k[0])^{-1}$ is equal to the temporal correlation parameter for user k at time n , represented by $\rho_k[n]$. Hence, the channel estimate given by equation (3.32) can be rewritten as:

$$\begin{aligned}\widehat{\mathbf{h}}_k[n] &= \rho_k[n] \mathbf{h}_k[0] + (1 - \rho_k[n]) \bar{\mathbf{h}}_k \\ \widehat{\mathbf{H}}_k[n] &= \rho_k[n] \mathbf{H}_k[0] + (1 - \rho_k[n]) \bar{\mathbf{H}}_k.\end{aligned}\quad (3.33)$$

Therefore, the channel model for the non-zero mean case can be written as:

$$\mathbf{H}_k[n] = \rho_k[n] \mathbf{H}_k[0] + (1 - \rho_k[n]) \bar{\mathbf{H}}_k + \mathbf{R}_{e,r_k}[n]^{1/2} \mathbf{H}_{\mathbf{w}_k}[n] \mathbf{R}_{e,t_k}[n]^{1/2}.\quad (3.34)$$

We can note that in this CSIT model, $\rho_k[n]$ acts as a channel estimate quality measure that depends on the time delay n . When the delay is zero or a small value, the parameter $\rho_k[n]$ is close to 1 and the channel estimate is strongly determined by the initial channel measurement $\mathbf{H}_k[0]$. Moreover, the error covariance is small. As the delay increases, $\rho_k[n]$ decreases to 0 reducing the impact of the initial channel measurement and the error covariance grows towards the channel covariance $\mathbf{R}_k[0]$.

Since our channel model considers the Ricean factor κ (equation (2.42)), we need to insert this factor appropriately into (3.34). Substituting the effective antenna correlation equations (3.23) into (3.34) and making some mathematical manipulations, the channel model for a non-zero mean case considering the Ricean factor κ can be rewritten as:

$$\mathbf{H}_k[n] = \rho_k[n] \mathbf{H}_k[0] + (1 - \rho_k[n]) \sqrt{\frac{\kappa}{1 + \kappa}} \bar{\mathbf{H}}_k + \sqrt{1 - \rho_k^2[n]} \sqrt{\frac{1}{1 + \kappa}} \mathbf{R}_{r_k}[n]^{1/2} \mathbf{H}_{\mathbf{w}_k}[n] \mathbf{R}_{t_k}[n]^{1/2}.\quad (3.35)$$

We can note that the channel model works as a predictor, in which the channel matrix is evaluated from its initial matrix $\mathbf{H}_k[0]$, its mean matrix $\bar{\mathbf{H}}_k$ and the spatial covariance matrices.

3.2 Simulations with the Proposed Statistical MU-MIMO CoMP Channel Model

In order to verify the performance of the proposed statistical MU-MIMO CoMP channel model, we first need to present the transmitter processing technique used, which is commonly known as *precoding*. This technique exploits the channel state information at the transmitter (CSIT) by operating on the signal before transmission in order to manage or eliminate the interference. The transmitter in a system with precoding consists of an encoder and a precoder, as depicted in figure 3.1. The encoder intakes data bits and performs coding for error correction by adding redundancy, then maps the coded bits into vector symbols. The precoder processes these symbols before transmission by the antennas. At the other end, the receiver decodes the received signal corrupted by noise plus interference in order to recover the data bits, treating the combination of the precoder and the channel as an effective channel.

In the literature, there are various precoding techniques that can be divided into linear and

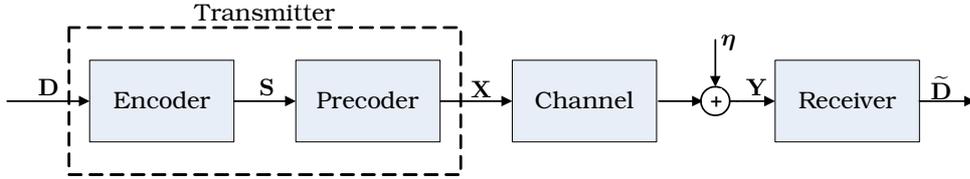


Figure 3.1: The transmitter in a system with precoder.

non-linear cases. The non-linear precoding techniques are based on the concept of coding technique proposed by Costa, known as dirty-paper coding (DPC) [32]. These techniques were developed having in mind interference cancellation and it was shown that the capacity of a channel in the case the transmitter knows the interfering signal is the same as if no interference is present. DPC techniques can achieve the maximum sum rate of the system and provide the maximum diversity order [61]. However, these techniques require the use of a complex sphere-decoder or an approximate closest-point solution, which makes them hard to implement in practice [62]. Moreover, non-linear MU-MIMO precoding techniques require the instantaneous knowledge of the channel matrix at the transmitter.

The linear precoding assumes that the transmitted signal is generated by a linear combination of input data symbols. Some examples of these techniques are zero-forcing (ZF) [3] and minimum mean-square error (MMSE) approach [63]. These linear techniques are less computationally demanding than DPC ones, and they can use either instantaneous channel knowledge or long-term statistics of the channel. Thus, they are more flexible and more favorable for practical implementation than non-linear techniques. In this section, we analyze the proposed statistical channel model using linear precoder techniques.

As already mentioned in Section 2.4, the MU-MIMO CoMP model has per-base power constraints. We define the joint precoder matrix \mathbf{T} as being the matrix composed by precoder matrices of all K users:

$$\mathbf{T} = [\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_K]_{N_t N_b \times (\sum_k L_k)}, \quad (3.36)$$

where L_k denotes the number of data streams intended for user k . A suboptimal way of obtaining this joint precoding matrix \mathbf{T} is to use the already known results for precoder techniques without considering their scale factors (which is the classical approach for global power constraints), and afterwards the per-base power constraints can be imposed by applying a power loading matrix [44].

Thus, the matrix \mathbf{T} can be seen as a product of two other matrices, given as:

$$\mathbf{T} = \mathbf{\Theta} \tilde{\mathbf{T}}, \quad (3.37)$$

where $\tilde{\mathbf{T}}$ is the joint precoder matrix without any power loading. In the same way as \mathbf{T} , the matrix $\tilde{\mathbf{T}}$ is one collection of submatrices $\tilde{\mathbf{T}}_k$, grouped side-by-side,

$$\tilde{\mathbf{T}} = [\tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_2, \dots, \tilde{\mathbf{T}}_K]_{N_t N_b \times (\sum_k L_k)}, \quad (3.38)$$

where $\tilde{\mathbf{T}}_k$ represents the precoding matrix for the user k without any power loading. The matrix $\mathbf{\Theta} = \mu \mathbf{I}$ is an $(\sum_k L_k) \times (\sum_k L_k)$ diagonal matrix with μ being the power allocated to the original data stream [44].

Let $\mathbf{P}_T = [P_1, P_2, \dots, P_{N_b}]^T$ be the per-base power constraint vector, where P_b is the power

constraint of BS b . Then the matrix Θ can be calculated as [44]:

$$\Theta = \mu \mathbf{I}, \quad \mu = \min_{b=1,2,\dots,N_B} \sqrt{\left(\frac{P_b}{\|\tilde{\mathbf{T}}^{[b]}\|_{\mathcal{F}}^2} \right)}, \quad (3.39)$$

where $\tilde{\mathbf{T}}^{[b]}$ contains the rows of $\tilde{\mathbf{T}}$ corresponding to the transmit antennas at BS b [44].

In this section, we consider the precoding matrix $\tilde{\mathbf{T}}$ as being generated by the existent linear precoding techniques known as zero-forcing (ZF) [3] and MMSE [63]. ZF precoding eliminates all interference at the user terminal, but suffers from transmit signal attenuation. Therefore it is a sub-optimal approach and might result in a significant performance degradation. MMSE precoder makes a trade-off between interference cancellation and transmit power efficiency. In the same way as the receive spatial MMSE filter, it approximates a matched filter at low SNRs and is near optimal. At high SNRs, the MMSE precoder converges to a ZF precoder.

3.2.1 Cell Scenario, Channel Model and Simulation Steps

The simulator follows a Monte-Carlo simulation approach. It simulates the downlink of a MU-MIMO CoMP system, where the base-station (BS)s can process data to be transmitted to users in a joint way to precancel the effect of interference among cochannel users.

The cell scenario consists of 3 coordinated cells characterizing a CoMP structure as already explained in chapter 2. The BSs are placed in the center of each cell and they can use more than one antenna element simultaneously. Users are placed randomly at the beginning of the simulation, where one user is located at each cell. It is considered that a scheduling algorithm selected the best user to transmit before this simulation. This scheduling algorithm is not the object of our interest. All users share the same radio resource and can be equipped with more than one cross-polarized antenna element.

The considered channel model is frequency-flat and is divided into blocks, where in each block the channel varies slowly following Jakes' model for time correlation. Thus, in each block, fast fading is present and is subject to the Doppler spread effect. The channel is considered stationary during the simulation time, so the channel statistics remain valid during the simulation.

Next, we enumerate the steps followed by the simulator:

- i.** At the beginning of the simulation, each user k sends to the BSs its channel mean $\bar{\mathbf{H}}_{k,b}$. Moreover, the channel covariance for each user k , $\mathbf{R}_k[0]$, is evaluated at the central processing unit, since we can consider that the channel covariance of the uplink is the same as for the downlink [64]. As we consider a stationary channel, such matrices remain valid during the whole simulation (see figure 3.2);
- ii.** At the beginning of each block, the users send the corresponding initial measurement channel $\mathbf{H}_k[0]$ to the BSs.
- iii.** For each symbol time interval T_S inside the block, the users send their corresponding parameter $\rho_k[n]$ to the BSs and their channels are estimated using equation (3.35) at the central processing unit. We consider that the block has a length of N_S symbols.

3.2.2 Simulation Parameters and Performance Metrics

In order to analyze the proposed channel model, we perform some simulations in which the main simulation parameters are listed in Table 3.1.

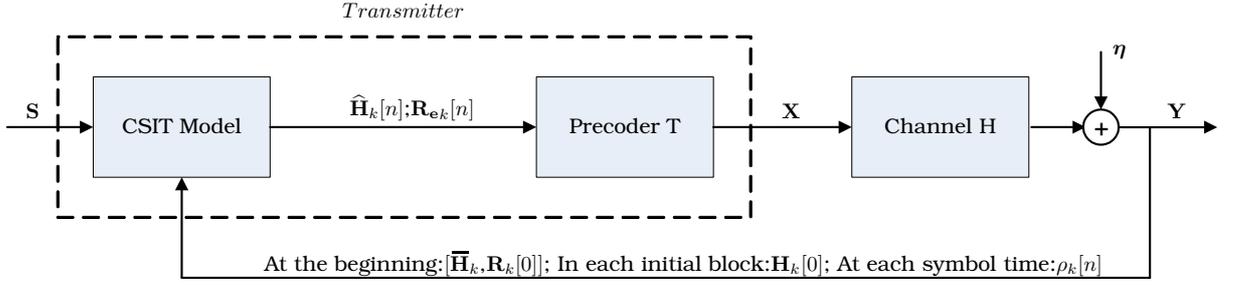


Figure 3.2: Modelling of the proposed channel model.

Table 3.1: Parameter of the Simulations.

Parameter	Value
Number of cells N_b	3
Number of users per cell	1
Cell radius	1 km
Number of Tx antennas per BS N_t	2 or 4
Number of Rx antennas per MS N_r	2 or 4
Carrier Frequency	2 GHz
Users velocity	30 km/h
Doppler frequency f_D	55.5 Hz
Coherence Time	$\frac{1}{2f_D} = 9$ ms
Noise power	30 dBm
Precoding techniques	ZF and MMSE

The spatial antenna correlations are fixed for all BSs and users as [65]:

$$\mathbf{R}_{\mathbf{t}_k} = \begin{cases} \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix} \otimes \mathbf{I}_{N_b} & \text{if } N_t = 2 \\ \begin{bmatrix} 1 & 0.4 & 0.3 & 0 \\ 0.4 & 1 & 0.4 & 0.3 \\ 0.3 & 0.4 & 1 & 0.4 \\ 0 & 0.3 & 0.4 & 1 \end{bmatrix} \otimes \mathbf{I}_{N_b} & \text{if } N_t = 4 \end{cases} \quad (3.40)$$

$$\mathbf{R}_{\mathbf{r}_k} = \begin{cases} \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix} & \text{if } N_r = 2 \\ \begin{bmatrix} 1 & 0.4 & 0.3 & 0 \\ 0.4 & 1 & 0.4 & 0.3 \\ 0.3 & 0.4 & 1 & 0.4 \\ 0 & 0.3 & 0.4 & 1 \end{bmatrix} \otimes \mathbf{I}_{N_b} & \text{if } N_{rt} = 4 \end{cases} \quad (3.41)$$

In order to analyze the results, the average spectral efficiency per user SE_{avg} is adopted as performance measure. Hence, we need to define the SINR_i of the receive antenna i as:

$$\text{SINR}_i = \frac{\|(\mathbf{HT})_{i,i}\|_F^2}{\sum_{j \neq i} \|(\mathbf{HT})_{i,j}\|_F^2 + \|\boldsymbol{\eta}\|_F^2}, \quad (3.42)$$

where \mathbf{T} is the joint precoding matrix given in (3.36) and \mathbf{H} is the joint channel matrix given by $\mathbf{H} = [\mathbf{H}_1^T \ \mathbf{H}_2^T \ \dots \ \mathbf{H}_K^T]^T_{KN_r \times N_t N_b}$. Moreover, $\mathbf{X}_{i,j}$ means the i, j -th element of matrix \mathbf{X} .

The average spectral efficiency per user is given by:

$$\text{SE}_{\text{avg}} = \frac{1}{K} \sum_{i=1}^{N_r K} \log_2 (1 + \text{SINR}_i). \quad (3.43)$$

3.2.3 Simulation Results

Figures 3.3(a) and 3.3(b) compare the SE_{avg} curves of linear precoding techniques using both perfect and the proposed CSIT model with $N_S = 5$ and considering that the Ricean factor κ varies. The scenario has 2 receive and 2 transmit antennas. In figure 3.3(a) we have a Ricean factor κ equal to 0 and we can observe that, at low SNRs, the performance obtained with the proposed channel model is similar to the case considering perfect knowledge of the channel at the transmitter. At high SNRs, the performance gap between the case considering perfect CSIT and the proposed channel model increases. The MMSE precoding has better performance in low signal-to-noise ratio (SNR)s than the ZF one, but at high SNRs the performance obtained with the MMSE precoding converges to the performance obtained with the ZF one. This is already expected from these precoding techniques [66]. From figure 3.3(b), we simulated the same scenario with the Ricean factor κ equal to 10. In this figure, we can observe the same behaviour than the top figure. When the SNR increases, the performance gap between the proposed channel model and the model with perfect CSIT also increases. Comparing the two figures 3.3(a) and 3.3(b) we can notice that, when the Ricean factor κ increases, the performance gap between the perfect and the proposed model becomes smaller. This happens because, when the Ricean factor κ is greater than 0, the channel has a LOS component which makes the channel more stable relative to variations of the channel temporal correlation. Moreover, we can note that the MMSE precoder technique is more robust when using the proposed channel model than the ZF technique. This can be seen since the performance gap between the perfect channel and the proposed CSIT one obtained with MMSE precoder technique is smaller compared to ZF precoder. Thus, we can conclude that the used precoding technique influences the performance results of the proposed channel model.

In order to evaluate the performance results when the number of symbols per block N_S varies, we simulate the MMSE precoder in scenarios where both the number of transmit and receive antennas are 2 and 4 and the parameter N_S varies. Figures 3.4(a) and 3.4(b) show the SE comparison when the number of transmit and receive antennas are both equal to 2. In figure 3.4(a) the Ricean factor κ is equal to 0 and we can note that when the number of symbols N_S increases, the performance obtained with the proposed channel model degrades. This is already expected since when the length of the block increases, the temporal channel correlation decreases and the channel estimation error in the proposed channel model becomes larger. From figure 3.4(a) we can also notice that the results obtained with all N_S values are similar in the SNR range $[-5, 10]$ dB. Only from the SNR value of 10 dB the results are degraded, mainly for $N_S = 10$. In figure 3.4(b) we have the same simulation but with the Ricean factor is equal to 10. We can note from this figure that the results obtained with all N_S values become similar until the SNR value is equal to 20 dB. And, only for a number of symbols $N_S = 10$, the performance degrades from the SNR value of 20dB. Hence, the loss of performance of the proposed channel model when the number of symbols in each block N_S increases is less significant for high Ricean factor values, except when $N_S = 10$. Figures 3.5(a) and 3.5(b) show the SE comparison obtained with the MMSE precoder when the number of transmit and receive antennas are both equal to 4 and the number of symbols N_S varies. From these figures we can also notice that when the Ricean factor κ increases, the performance variation of the proposed channel model becomes smaller.

3.3 Summary

In this chapter we proposed a statistical channel model for the MU-MIMO CoMP scenario which takes into account the channel temporal correlation. This proposed CSIT model was obtained from error estimation theory and compared to the case with perfect CSIT using the ZF and MMSE precoding. The results showed that the performance gap between the proposed CSIT model and the cases with perfect CSIT is negligible for low SNR values and moderate for medium to high SNR values.

We evaluated also the influence of the update frequency of the joint initial channel matrix $\mathbf{H}[0]$, which is inversely proportional to the number of symbols N_S , on the performance results and showed that, when the update frequency increases (N_S decreases), the results become better. But, since the increase of the update frequency causes more feedback information, it is necessary to determine a suitable trade-off value between performance and amount of feedback.

Another important conclusion about the proposed CSIT model is that their performance results are better or worse depending of the chosen precoder technique. Moreover, the precoding techniques used in the simulation were designed for the perfect CSIT case. Hence, designing a precoder that exploits the advantages of the proposed channel model is the next step of our work, and it will be discussed in the next chapter.

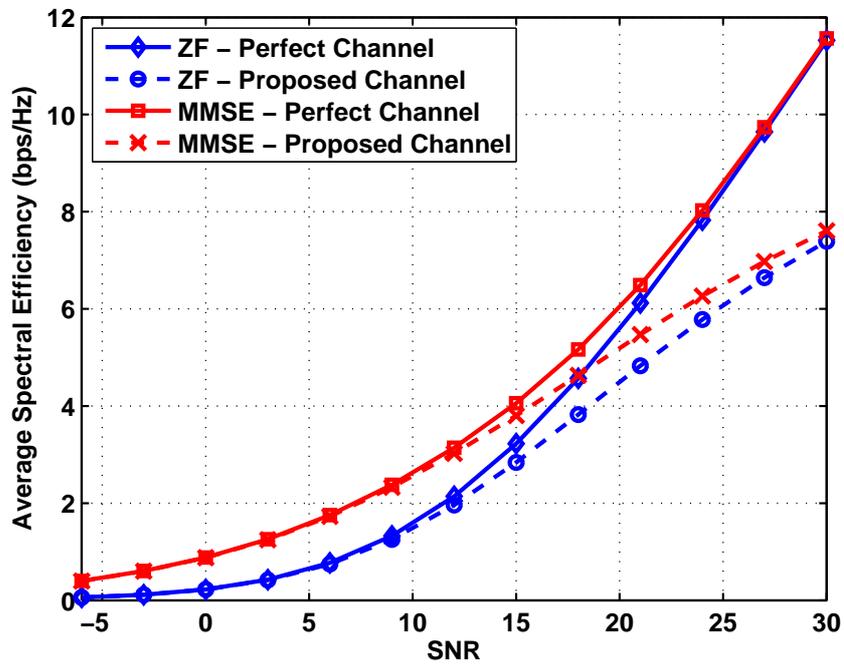
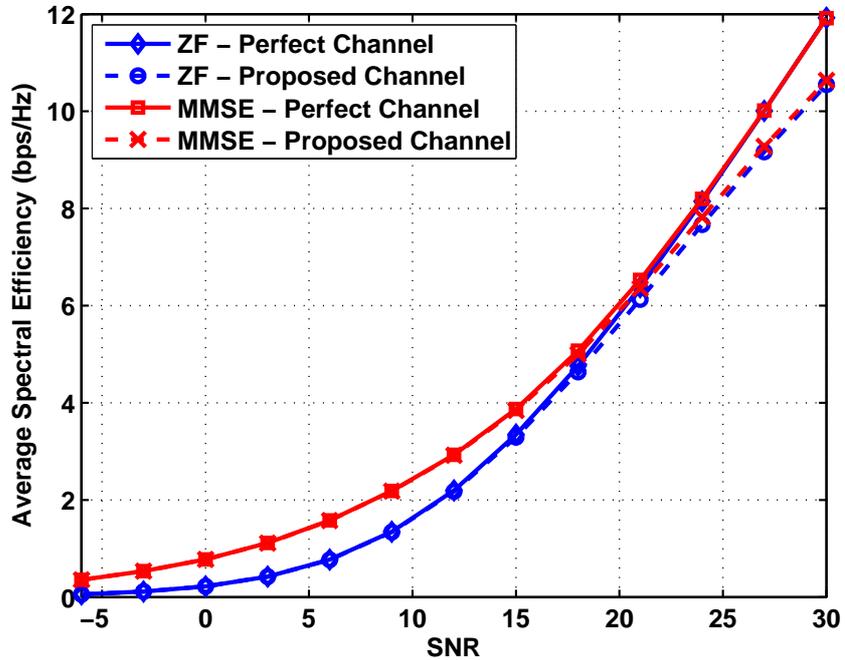
(a) Ricean factor $\kappa = 0$ (b) Ricean factor $\kappa = 10$

Figure 3.3: SE curves of linear precoding techniques using perfect and the proposed channel model with $N_S = 5$ and $N_t = N_r = 2$.

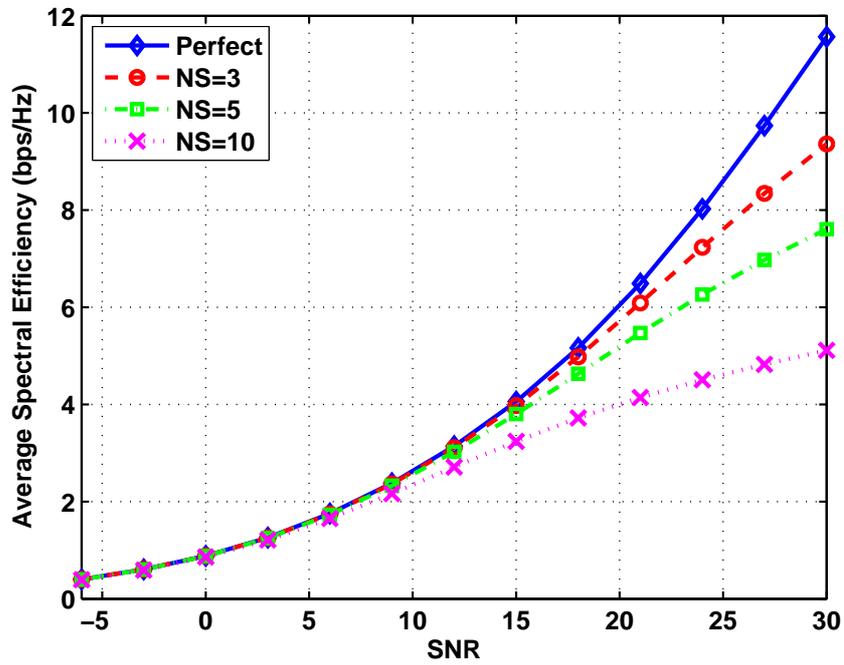
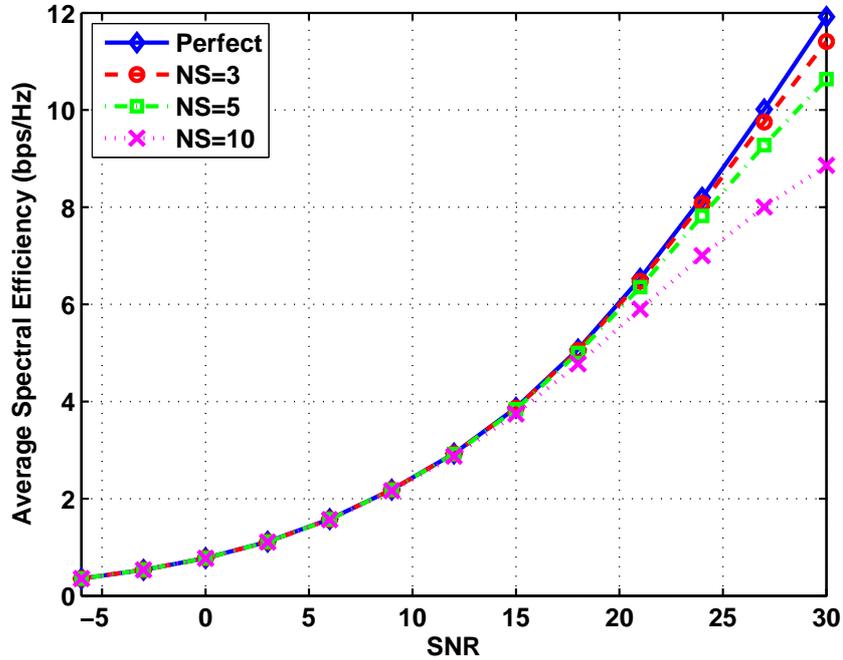
(a) Ricean factor $\kappa = 0$ (b) Ricean factor $\kappa = 10$

Figure 3.4: SE curves of MMSE precoding technique using the proposed channel model and varying the number of symbols in each block N_S for $N_t = N_r = 2$.

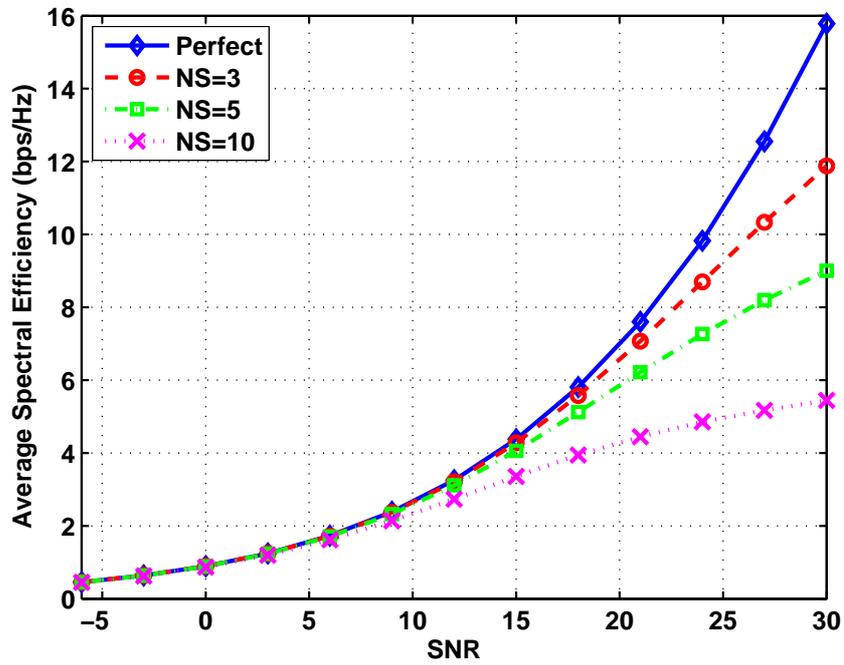
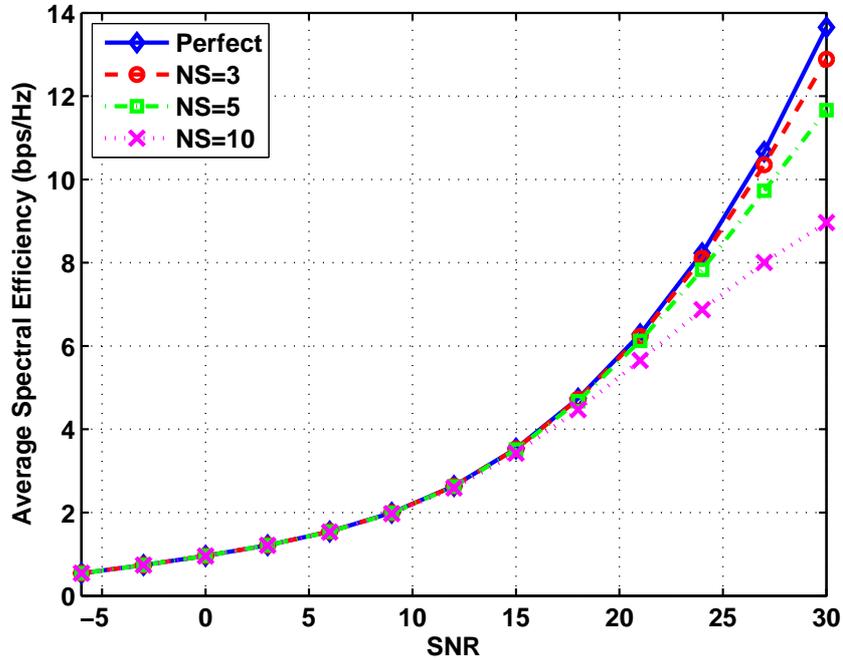
(a) Ricean factor $\kappa = 0$ (b) Ricean factor $\kappa = 10$

Figure 3.5: SE curves of MMSE precoding technique using the proposed channel model and varying the number of symbols in each block N_S for $N_t = N_r = 4$.

Chapter 4

Statistical Transmit Scheme for MU-MIMO CoMP Systems

In this chapter, we propose two transmit schemes that aim at maximizing the first and second-order approximations of the ergodic sum rate in the downlink of a multi-user MIMO (MU-MIMO) coordinated multipoint (CoMP) system. This considered system model has the mean and the covariance matrices of the channel as partial channel state information at the transmitter (CSIT). Firstly, we use the duality theory [14, 15] to compute the sum rate of the downlink MU-MIMO CoMP system and next find a first and second-order approximation of the ergodic sum rate of the considered system. In sequel, we derive input covariance matrices that maximize the found approximations and, using simulations, we present results showing that our proposed schemes are near-optimal and provide good results compared to the optimal iterative water-filling algorithm proposed by Jindal *et al* in [14].

4.1 Duality and Multiuser MIMO Downlink CoMP Capacity

In this section, we describe the problem of maximizing the downlink sum rate of the MU-MIMO CoMP system subject to a per-base power constraint. From equations (2.26) and (2.28), and modifying the power constraint for the power restriction of our considered model (detailed in Section 2.4, equation (2.44)), we have that the sum rate of the system considered in this work can be rewritten as

$$C_{\text{DL}}(\mathbf{H}_1, \dots, \mathbf{H}_K, P_1, \dots, P_{N_b}) = \max_{\{\boldsymbol{\Omega}_k\}_{k=1}^K; \boldsymbol{\Omega}_k \succeq 0, \sum_{k=1}^K \text{Tr}(\boldsymbol{\Omega}_{k,b}) \leq P_b} \sum_k \log \frac{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j \geq k} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j > k} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}, \quad (4.1)$$

where P_b is the power constraint of base-station (BS) b , $\boldsymbol{\Omega}_{k,b} = \mathbf{T}_{k,b} \mathbf{T}_{k,b}^H$ is the input covariance matrix of user k and BS b and $\mathbf{H}_k = \begin{bmatrix} \mathbf{H}_{k,1} & \mathbf{H}_{k,2} & \dots & \mathbf{H}_{k,N_b} \end{bmatrix}_{N_r \times N_b N_b}$ is the joint channel matrix from all BSs to user k with $\mathbf{H}_{k,b}$ being the channel matrix from BS b to user k .

Since the optimization in (4.1) must be performed over all $\boldsymbol{\Omega}_k$ for $k = 1, \dots, K$ and the power constraint is applied only in a matrix that is a part of $\boldsymbol{\Omega}_k$, this optimization problem becomes hard to solve. An alternative solution for this impairment is, firstly, to compute the sum rate considering a global power restriction, given by the sum of the power restriction of each base, and later, to apply a power normalization matrix that satisfies the per-base power restrictions. Accordingly, in this section we will relax the per-base power constraints and consider a sum

power constraint. In a posterior section, the considered power normalization will be explained and, in spite of being only near-optimal, it will be shown to provide good results.

Now, considering a sum-power constraint and assuming perfect synchronization of the different delays from different BSs to each user, we can model the multiple BSs of our CoMP system as a single large BS. Then, our model can be viewed as the downlink of a large MU-MIMO system and the maximization of the sum rate in the downlink of a MU-MIMO system can be written as (equations (2.26) and (2.28)):

$$C_{\text{DL}}(\mathbf{H}_1, \dots, \mathbf{H}_K, P) = \max_{\{\boldsymbol{\Omega}_k\}_{k=1}^K; \boldsymbol{\Omega}_k \geq 0, \sum_{k=1}^K \text{Tr}(\boldsymbol{\Omega}_k) \leq P} \sum_k \log \frac{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j \geq k} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}{\left| \mathbf{I} + \mathbf{H}_k \left(\sum_{j > k} \boldsymbol{\Omega}_j \right) \mathbf{H}_k^H \right|}, \quad (4.2)$$

where $P = \sum_{b=1}^{N_b} P_b$ is the sum power constraint.

The maximization in (4.2) is performed over the downlink input covariance matrices $\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_K$, each of which is an $N_b N_t \times N_b N_t$ positive semidefinite matrix given by $\boldsymbol{\Omega}_k = \mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\}$, where $\mathbf{x}_k = [\mathbf{x}_{k,1} \ \mathbf{x}_{k,2} \ \dots \ \mathbf{x}_{k,N_b}]^T$ and $\mathbf{x}_{k,b}$ is the signal transmitted from BS b to user k . The main interest is to find the covariance matrices that achieve the maximum in (4.2). However, this is a nontrivial problem since the objective function in (4.2) is not a concave function of $\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_K$ [37]. One possible solution is to exploit the existing duality between the uplink and downlink channels of a MU-MIMO system [15], which has been already detailed for the MU-MIMO system case in Section 2.3. Figure 4.1 shows the downlink channel of the CoMP MU-MIMO system with K users along with the dual uplink channel. The dual uplink channel is a K -user MU-MIMO uplink channel where each of the dual uplink channels is the conjugate transpose of the corresponding downlink channel.

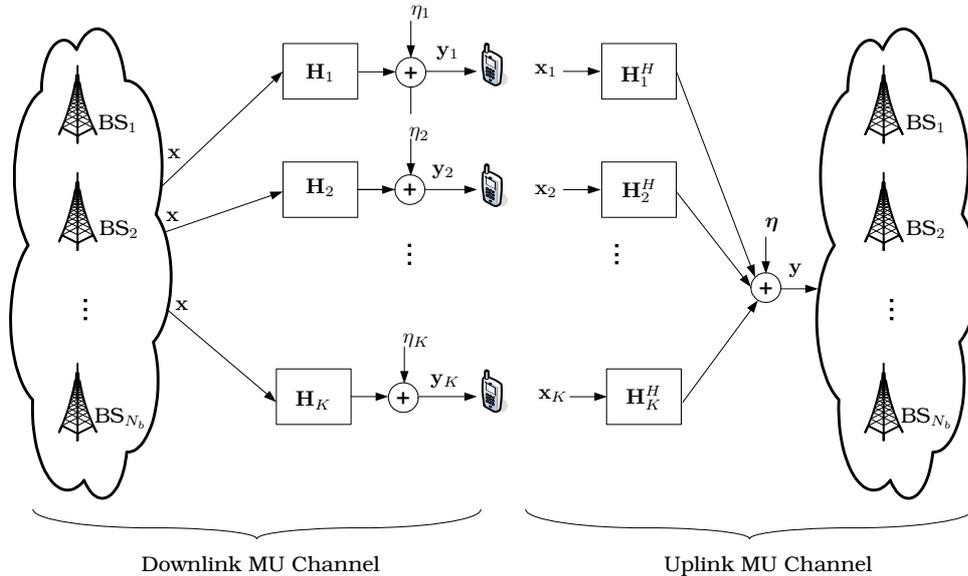


Figure 4.1: System model of the downlink MU-MIMO CoMP (left) and uplink MU-MIMO CoMP (right) channels.

Considering the duality theorem stated in [15], which affirms that the sum rate of the downlink MU-MIMO is equal to the sum rate of the dual uplink MU-MIMO, i.e.:

$$C_{\text{DL}}(\mathbf{H}_1, \dots, \mathbf{H}_K, P) = C_{\text{UL}}(\mathbf{H}_1^H, \dots, \mathbf{H}_K^H, P), \quad (4.3)$$

we can now rewrite our optimization problem as the maximization of the sum rate of the uplink MU-MIMO, which is given by the following expression (from (2.41)) [14]:

$$C_{\text{DL}}(\mathbf{H}_1^H, \dots, \mathbf{H}_K^H, P) = C_{\text{UL}}(\mathbf{H}_1^H, \dots, \mathbf{H}_K^H, P) = \max_{\{\mathbf{Q}_k\}_{k=1}^K; \mathbf{Q}_k \succeq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \log \left| \mathbf{I} + \sum_{k=1}^K \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right|, \quad (4.4)$$

where \mathbf{Q}_k is a $N_r \times N_r$ matrix given by $\mathbf{Q}_k = \mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\}$ with \mathbf{x}_k being the signal vector transmitted from user k to all BSs. The maximization is performed over the uplink input covariance matrices $\mathbf{Q}_1, \dots, \mathbf{Q}_K$ subject to sum power constraint P . This objective function is a concave function of the covariance matrices [15]. Since a transformation that maps uplink input covariance matrices to downlink input covariance matrices achieving the same rates and using the same sum power has already been presented in Section 2.3 and it is also provided in [15], we can find first the uplink input covariance matrices and, after that, obtain the corresponding downlink input covariance matrices using this mapping.

In order to solve (4.4) we perform a block-coordinated ascent algorithm, which consists of optimizing \mathbf{Q}_k while holding constant all other variables \mathbf{Q}_j for $j \neq k$ [38]. Accordingly, when optimizing \mathbf{Q}_1 , the matrices $\mathbf{Q}_2, \dots, \mathbf{Q}_K$ are assumed to remain constant, when optimizing \mathbf{Q}_2 , the matrices $\mathbf{Q}_1, \mathbf{Q}_3, \dots, \mathbf{Q}_K$ are assumed to remain constant, and so on. This iterative optimization is performed until the convergence of the sum rate is achieved. Moreover, since we have a global sum power constraint in our optimization problem in (4.4), the input covariance matrices of all K users must be updated in each iteration of the algorithm to maintain a constant sum power. In order to perform this, we consider in our algorithm that all K input covariance matrices are updated in each iteration based on the covariance matrices from the previous iteration. Thus we have:

$$\mathbf{Q}_k^{(i+1)} = \begin{cases} \arg \max_{\mathbf{Q}_k; \mathbf{Q}_k \succeq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \log \left| \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j + \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right|, & \text{if } k = [(i-1) \bmod K] + 1 \\ \mathbf{Q}_k^{(i)}, & \text{if } k \neq [(i-1) \bmod K] + 1 \end{cases} \quad (4.5)$$

where $\mathbf{Q}_k^{(i)}$ is the uplink input covariance matrix of user k evaluated in the i -th iteration of the algorithm and $[x \bmod y]$ is the remainder of the integer division of x by y .

Let us denote the objective function of (4.5) by $f(\cdot)$, i.e.,

$$f(\mathbf{Q}_1, \dots, \mathbf{Q}_K) \triangleq \log \left| \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j + \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right|. \quad (4.6)$$

Note that the function $f(\cdot)$ in (4.6) can be rewritten after some mathematical manipulations as

$$f(\mathbf{Q}_1, \dots, \mathbf{Q}_K) = \log \left| \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right| + \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1/2} \right| \quad (4.7)$$

for any k , where the derivation is shown in appendix A.1.

When optimizing $f(\cdot)$ for user k , we notice that the first term in (4.7) is constant. Hence,

maximizing (4.7) is equivalent to maximizing only its second term. Defining an auxiliary matrix \mathbf{C}_k as

$$\mathbf{C}_k = \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \quad (4.8)$$

and using this fact, we can rewrite (4.5) as

$$\mathbf{Q}_k^{(i+1)} = \begin{cases} \arg \max_{\mathbf{Q}_k: \mathbf{Q}_k \succeq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \log \left| \mathbf{I} + (\mathbf{H}_k \mathbf{C}_k^{-1/2})^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right|, & \text{if } k = [(i-1) \bmod K] + 1 \\ \mathbf{Q}_k^{(i)}, & \text{if } k \neq [(i-1) \bmod K] + 1 \end{cases} \quad (4.9)$$

Therefore, we can define that the sum rate of the uplink MU-MIMO CoMP channel given by (4.4) is numerically equivalent to the value C_G which is given by [14]:

$$C_G = \max_{\mathbf{Q}_k \succeq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \log \left| \mathbf{I} + (\mathbf{H}_k \mathbf{C}_k^{-1/2})^H \mathbf{Q}_k (\mathbf{H}_k \mathbf{C}_k^{-1/2}) \right|, \quad (4.10)$$

i.e., maximizing (4.4) is equivalent to maximizing the rate of the point-to-point MIMO channel \mathbf{G}_k described by

$$\mathbf{G}_k = \mathbf{H}_k \mathbf{C}_k^{-1/2}. \quad (4.11)$$

In the next section, we will obtain an approximation of the ergodic sum-rate of the uplink MU-MIMO CoMP channel. For this, we will use the fact previously shown that the sum rate of uplink MU-MIMO CoMP is numerically equal to the rate C_G of the point-to-point MIMO channel \mathbf{G}_k [14]. Moreover, we will consider that the transmitter has partial knowledge about the channel.

4.2 Approximations of the MU-MIMO CoMP Ergodic Sum-Rate

In some recent works, the authors assume that the channel is perfectly known at the transmitter [14–16]. This assumption can have a significant impact on the maximum ergodic sum rate that can be reliably communicated over the channel, but may not be realistic in many practical scenarios. In this thesis, we assume that the transmitter has access to statistical channel state information (CSI), while the receiver has access to instantaneous CSI. In this section, we derive a first- and second-order approximation of the ergodic sum rate for a MU-MIMO CoMP system.

Considering the sum rate of the MU-MIMO CoMP system with perfect CSI at the transmitter and its numerical equivalence with the sum rate of the point-to-point MIMO channel C_G , we have from (4.10) and (4.11) that the ergodic sum rate of the MU-MIMO CoMP channel is given by:

$$\bar{C}_{\text{CoMP}}(\mathbf{H}_1^H, \dots, \mathbf{H}_K^H, P) = \max_{\mathbf{Q}_k \succeq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \mathbb{E} \left\{ \log \left| \mathbf{I} + \mathbf{G}_k^H \mathbf{Q}_k \mathbf{G}_k \right| \right\}, \quad (4.12)$$

where $P = \sum_{b=1}^{N_b} P_b$ and the expectation is taken over the channel matrices \mathbf{H}_k , which are part of \mathbf{G}_k as defined in (4.11). Since we have considered in the optimization that the input covariance matrices of the other users $j \neq k$ are constant in the evaluation of the input covariance matrix for user k , the matrix \mathbf{C}_k of (4.8) can also be considered constant.

In [67] we observe that the equality $\log |\mathbf{X}| = \text{tr}(\log(\mathbf{X}))$ is considered true for positive definite matrices. Using this result for the matrix $\mathbf{I} + \mathbf{G}_k^H \mathbf{Q}_k \mathbf{G}_k$ in (4.12) and considering that the transmit power P is sufficiently small so that the maximum eigenvalue of $\mathbf{I} + \mathbf{G}_k^H \mathbf{Q}_k \mathbf{G}_k$

is less than 1 with high probability, we can use the Taylor expansion $\log(\mathbf{I} + \mathbf{A}) = \mathbf{A} - \frac{1}{2}\mathbf{A}^2 + \frac{1}{3}\mathbf{A}^3 - \dots$ on the objective function of (4.12).

After these considerations and considering only the two first terms on the Taylor expansion (second-order approximation), we rewrite (4.12) as

$$\begin{aligned}
\bar{C}_{\text{CoMP}} &= \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{tr}(\mathbf{Q}_k) \leq P} \mathbb{E} \left\{ \log \left| \mathbf{I} + \mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right| \right\} \\
&= \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{tr}(\mathbf{Q}_k) \leq P} \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right) - \frac{1}{2} \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right)^2 \right\} \\
&= \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{tr}(\mathbf{Q}_k) \leq P} \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right) \right\} - \frac{1}{2} \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right)^2 \right\}.
\end{aligned} \tag{4.13}$$

Following, we will evaluate the first- and second-order approximation of the MU-MIMO CoMP sum-rate.

4.2.1 Proposed Input Covariance Matrices using the First-Order Approximation of the Sum Rate - FOIC Approach

Considering only the first term of the Taylor expansion in equation (4.13), we obtain the first-order approximation of the ergodic sum-rate:

$$\bar{C}_{\text{CoMP_FO}} = \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} E \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right) \right\}. \tag{4.14}$$

If we assume that the transmitter has access to statistical channel state information (CSI), we can substitute the Kronecker channel model equation given by (2.42) in (4.14). For simplicity of notation, we will omit the factors $\sqrt{\frac{\kappa}{1+\kappa}}$ and $\sqrt{\frac{1}{1+\kappa}}$ of (2.42), which will be suitably reintroduced later. After some mathematical manipulations [68], we have:

$$\bar{C}_{\text{CoMP_FO}} = \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \text{tr} \left(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \right) + \text{tr} \left(\mathbf{R}_{\mathbf{r}_k} \mathbf{Q}_k \right) \text{tr} \left(\mathbf{R}_{\mathbf{t}_k} \mathbf{C}_k^{-1} \right) \tag{4.15}$$

In order to simplify notation, we introduce an additional auxiliary matrix \mathbf{X}_k defined as

$$\mathbf{X}_k = \left(\bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H + \text{tr} \left(\mathbf{R}_{\mathbf{t}_k} \mathbf{C}_k^{-1} \right) \mathbf{R}_{\mathbf{r}_k} \right), \tag{4.16}$$

which will be used in the sequel.

Therefore, the first-order approximation of the sum rate can be rewritten as:

$$\bar{C}_{\text{CoMP_FO}} = \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{tr}(\mathbf{Q}_k) \leq P} \text{tr} \left(\mathbf{X}_k \mathbf{Q}_k \right). \tag{4.17}$$

In order to use an already known solution, we have to perform a modification in the power constraint of the optimization problem in (4.17). We divide equally the sum power constraint P for all K users, thus we have that our optimization problem can be rewritten as:

$$\bar{C}_{\text{CoMP_FO}} = \max_{\mathbf{Q}_k \geq 0, \text{tr}(\mathbf{Q}_k) \leq \frac{P}{K}} \text{tr} \left(\mathbf{X}_k \mathbf{Q}_k \right). \tag{4.18}$$

This modification in the power constraint will not affect the solution of the original optimization problem since this sum power constraint is a relaxed one as we have already

explained in Section 4.1. Moreover, after the optimization, a power normalization matrix will be applied in order to satisfy the per-base power restrictions which is the intrinsic power constraint of the CoMP system.

The sub-optimal input covariance matrix for each user can be found optimizing this first-order approximation of the sum rate. In order to find the solution, we use the result in [69, Example 7.4.13] to consider as solution the matrix \mathbf{Q}_k with the same eigenbasis as \mathbf{X}_k . That is, if $\mathbf{X}_k = \mathbf{U}_X \Lambda_X \mathbf{U}_X^H$ denotes the ordered eigendecomposition of \mathbf{X}_k , it is sufficient to consider matrices \mathbf{Q}_k of the form $\mathbf{U}_X \Lambda_Q \mathbf{U}_X^H$, where Λ_Q denotes the diagonal matrix of (nonnegative) eigenvalues of \mathbf{Q}_k . If the maximum eigenvalue of \mathbf{X}_k is distinct, the optimal eigenvalues of \mathbf{Q}_k are $\lambda_{Q_1} = \frac{P}{K}$ and $\lambda_{Q_2} = \dots = \lambda_{Q_{N_r}} = 0$. That is, beamforming along the principal eigenvector of \mathbf{X}_k is sufficient for rate-optimal communication. For the case in which the largest eigenvalue of \mathbf{X}_k has multiplicity greater than 1, any partitioning of power in the direction of the eigenvectors corresponding to these eigenvalues is optimal up to the first-order approximation [56].

4.2.2 Proposed Input Covariance Matrices using the Second-Order Approximation of the Sum Rate - SOIC Approach

As we assume that the transmitter has access to statistical channel state information, we can introduce our channel model (2.42) into the second-order approximation shown in equation (4.13) and proceed with the derivations term-by-term. For simplicity of notation, we will again omit the factors $\sqrt{\frac{\kappa}{1+\kappa}}$ and $\sqrt{\frac{1}{1+\kappa}}$ of (2.42), which will be suitably reintroduced later. Let us consider initially the first term of (4.13), which we will denote by

$$\mathbf{\Pi} = \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right) \right\}. \quad (4.19)$$

Substituting the channel model given by (2.42) into (4.19), we have

$$\begin{aligned} \mathbf{\Pi} &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \left(\bar{\mathbf{H}}_k + \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right)^H \mathbf{Q}_k \left(\bar{\mathbf{H}}_k + \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right) \mathbf{C}_k^{-1/2} \right) \right\} \\ &= \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbb{E} \left\{ \bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k + \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} + \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k + \right. \right. \\ &\quad \left. \left. + \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right\} \mathbf{C}_k^{-1/2} \right) \\ &= \text{tr} \left(\mathbf{C}_k^{-1/2} \left(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k + \mathbb{E} \left\{ \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right\} \right) \mathbf{C}_k^{-1/2} \right). \end{aligned} \quad (4.20)$$

We show in Appendix A.2 that the term $\mathbb{E} \left\{ \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right\}$ reduces to $\text{tr} \left(\mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \right) \mathbf{R}_{t_k}$ after some mathematical manipulations. Hence, (4.20) can be written as

$$\mathbf{\Pi} = \text{tr} \left(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \right) + \text{tr} \left(\mathbf{R}_{r_k} \mathbf{Q}_k \right) \text{tr} \left(\mathbf{R}_{t_k} \mathbf{C}_k^{-1} \right). \quad (4.21)$$

Let us now study the second term of the Taylor expansion contained in (4.13), which we will denote by

$$\mathbf{\Xi} = \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_k^{-1/2} \right)^2 \right\}. \quad (4.22)$$

Substituting the channel model given by (2.42) into (4.22), we have

$$\begin{aligned} \Xi &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \left(\bar{\mathbf{H}}_k + \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^H \mathbf{Q}_k \left(\bar{\mathbf{H}}_k + \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \mathbf{C}_k^{-1/2} \right)^2 \right\} \\ &= \mathbb{E} \left\{ \text{Tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2} + \mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} + \right. \right. \\ &\quad \left. \left. + \mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2} + \mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right)^2 \right\}. \end{aligned} \quad (4.23)$$

Then, we rewrite (4.23) as

$$\begin{aligned} \Xi &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{A} + (\mathbf{B} + \mathbf{B}^H) + \mathbf{C} \right)^2 \right\} \\ &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{A}^2 + (\mathbf{B} + \mathbf{B}^H)^2 + \mathbf{C}^2 + 2\mathbf{A}(\mathbf{B} + \mathbf{B}^H) + 2\mathbf{A}\mathbf{C} + 2(\mathbf{B} + \mathbf{B}^H)\mathbf{C} \right) \right\} \\ &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{A}^2 \right) \right\} + \mathbb{E} \left\{ \text{tr} \left((\mathbf{B} + \mathbf{B}^H)^2 \right) \right\} + \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}^2 \right) \right\} + \mathbb{E} \left\{ \text{tr} \left(2\mathbf{A}(\mathbf{B} + \mathbf{B}^H) \right) \right\} + \\ &\quad + \mathbb{E} \left\{ \text{tr} \left(2\mathbf{A}\mathbf{C} \right) \right\} + \mathbb{E} \left\{ \text{tr} \left(2(\mathbf{B} + \mathbf{B}^H)\mathbf{C} \right) \right\}, \end{aligned} \quad (4.24a)$$

where

$$\mathbf{A} = \mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2}, \quad (4.24b)$$

$$\mathbf{B} = \mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}, \quad (4.24c)$$

$$\mathbf{C} = \mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}. \quad (4.24d)$$

We can notice that $\mathbb{E} \{ \text{tr} (\mathbf{A}\mathbf{B}) \} = 0$ and $\mathbb{E} \{ \text{tr} (\mathbf{A}\mathbf{B}^H) \} = 0$ since \mathbf{A} is a deterministic matrix and $\mathbb{E} \{ \mathbf{H}_{\mathbf{w}_k} \} = 0$. In Appendix A.3 we show that $\mathbb{E} \{ \text{tr} (\mathbf{B}\mathbf{B}) \} = 0$ and $\mathbb{E} \{ \text{tr} (\mathbf{B}^H\mathbf{B}^H) \} = 0$; and in Appendix A.4 we show that $\mathbb{E} \{ \text{tr} (\mathbf{B}\mathbf{C}) \} = 0$ and $\mathbb{E} \{ \text{tr} (\mathbf{B}^H\mathbf{C}) \} = 0$. This reduces (4.24a) to

$$\Xi = \mathbb{E} \left\{ \text{tr} \left(\mathbf{A}^2 \right) \right\} + 2\mathbb{E} \left\{ \text{tr} \left(\mathbf{B}\mathbf{B}^H \right) \right\} + \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}^2 \right) \right\} + 2\mathbb{E} \left\{ \text{tr} \left(\mathbf{A}\mathbf{C} \right) \right\}. \quad (4.25)$$

In Appendix A.5, we show the evaluation of $\mathbb{E} \{ \text{tr} (\mathbf{A}\mathbf{C}) \}$ and of $\mathbb{E} \{ \text{tr} (\mathbf{B}\mathbf{B}^H) \}$ and in Appendix A.6 we evaluated the term $\mathbb{E} \{ \text{tr} (\mathbf{C}^2) \}$. Using the results provided in the referred appendices, we rewrite (4.25) as

$$\begin{aligned} \Xi &= \text{tr} \left(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \right)^2 + 2 \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2} \right) + \\ &\quad + \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \right)^2 + \text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \right)^2 + \\ &\quad + 2 \text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \text{tr} \left(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k} \mathbf{C}_k^{-1} \right). \end{aligned} \quad (4.26)$$

After some mathematical manipulation, we have that (4.26) is

$$\begin{aligned} \Xi &= \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \right)^2 + 2 \text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \text{tr} \left(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k} \mathbf{C}_k^{-1} \right) + \\ &\quad + \text{tr} \left(\left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \mathbf{R}_{\mathbf{r}_k} + \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H \right) \mathbf{Q}_k \right)^2. \end{aligned} \quad (4.27)$$

Using (4.21) and (4.27) in (4.13), the second-order approximation of the ergodic sum rate

of the MU-MIMO CoMP system can be written as

$$\begin{aligned}
\bar{C}_{\text{CoMP_SO}} &= \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \text{tr}(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1}) + \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) - \\
&\quad - \frac{1}{2} \left(\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 (\text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k))^2 + 2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \text{tr}(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) + \right. \\
&\quad \left. + \text{tr}((\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) \mathbf{R}_{\mathbf{r}k} + \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H) \mathbf{Q}_k)^2 \right), \\
\bar{C}_{\text{CoMP_SO}} &= \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \text{tr}((\bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H + \text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) \mathbf{R}_{\mathbf{r}k}) \mathbf{Q}_k) - \\
&\quad - \frac{1}{2} \left(\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 (\text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k))^2 + 2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \text{tr}(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) + \right. \\
&\quad \left. + \text{tr}((\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) \mathbf{R}_{\mathbf{r}k} + \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H) \mathbf{Q}_k)^2 \right). \tag{4.28}
\end{aligned}$$

Using (4.16) in (4.28), we have that the second order approximation of the ergodic sum rate of the MU-MIMO CoMP system is given by

$$\begin{aligned}
\bar{C}_{\text{CoMP_SO}} &= \max_{\mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} \text{tr}(\mathbf{X}_k \mathbf{Q}_k) - \frac{1}{2} \left(\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 (\text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k))^2 + \right. \\
&\quad \left. + 2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \text{Tr}(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) + \text{tr}(\mathbf{X}_k \mathbf{Q}_k)^2 \right). \tag{4.29}
\end{aligned}$$

Since \mathbf{C}_k is considered a constant matrix in the optimization for user k , we can observe from (4.29) that the required channel information for user k in order to optimize \mathbf{Q}_k are only the channel mean matrix $\bar{\mathbf{H}}_k$ and the channel covariance matrix, which are represented by $\mathbf{R}_{\mathbf{t}k}$ and $\mathbf{R}_{\mathbf{r}k}$. In the next section, we will use optimization tools to obtain the input covariance matrices from the optimization problem given in (4.29).

4.3 Input Covariance Matrix Maximizing the Second-Order Approximation of the Ergodic Sum Rate

In this section, we maximize the second-order approximation of the ergodic sum rate and find a near-optimal input covariance matrix for each user. Firstly, we analyze the convergence of the ergodic sum rate and then we derive an efficient algorithm for obtaining the input covariance matrices that maximize the second-order approximation of the ergodic sum rate.

4.3.1 Convergence Analysis

In order to analyze the convergence of the ergodic sum rate, we take the ergodic sum rate given by (4.12) assuming that an optimal input covariance matrix \mathbf{Q}_k^O is obtained. Moreover, denoting the eigen decomposition of $\mathbf{G}_k^H \mathbf{Q}_k^O \mathbf{G}_k$ as $\Lambda_k \Phi_k \Lambda_k^H$, where Λ_k is an $N_b N_t \times N_b N_t$ unitary matrix and Φ_k is an $N_b N_t \times N_b N_t$ ordered diagonal one, we can write the optimal ergodic sum rate of the MU-MIMO CoMP channel as

$$\begin{aligned}
\bar{C}_{\text{CoMP_Opt}} &= \mathbb{E} \{ \log |\mathbf{I} + \mathbf{G}_k^H \mathbf{Q}_k^O \mathbf{G}_k| \} \\
&= \mathbb{E} \{ \text{tr}(\log(\mathbf{I} + \Phi_k)) \} \tag{4.30}
\end{aligned}$$

where the expectation is taken over the diagonal elements of Φ_k $\{\theta_i\}_{i=1}^{N_b N_t}$.

We explore now the optimal ergodic sum rate aiming to guarantee its convergence. One

can write (4.30) as

$$\bar{C}_{\text{CoMP_Opt}} = \int_0^\infty \dots \int_0^\infty \text{tr}(\log(\mathbf{I} + \Phi_k)) p_{\Phi_k}(\theta_1, \dots, \theta_{N_b N_t}) d\theta_1 \dots d\theta_{N_b N_t} \quad (4.31)$$

where $p_{\Phi_k}(\Phi)$ is the joint probability density function of the diagonal elements of Φ_k . Partitioning the integral in (4.31), we have

$$\begin{aligned} \bar{C}_{\text{CoMP_Opt}} &= \int_0^\delta \dots \int_0^\delta \text{tr}(\log(\mathbf{I} + \Phi_k)) p_{\Phi_k}(\theta_1, \dots, \theta_{N_b N_t}) d\theta_1 \dots d\theta_{N_b N_t} \\ &\quad + \int_\delta^\infty \dots \int_\delta^\infty \text{tr}(\log(\mathbf{I} + \Phi_k)) p_{\Phi_k}(\theta_1, \dots, \theta_{N_b N_t}) d\theta_1 \dots d\theta_{N_b N_t} \end{aligned} \quad (4.32)$$

We observe from (4.32) that the convergence of the optimal ergodic sum rate will be guaranteed only if we choose a scalar δ sufficiently large so as to ensure that

$$\int_\delta^\infty \dots \int_\delta^\infty \text{tr}(\log(\mathbf{I} + \Phi_k)) p_{\Phi_k}(\theta_1, \dots, \theta_{N_b N_t}) d\theta_1 \dots d\theta_{N_b N_t} \leq \epsilon, \quad (4.33)$$

for some arbitrary small $\epsilon > 0$. Hence, if we show that (4.33) is true, we guarantee that the integral in (4.31) converges. Thus, we can also guarantee that our algorithm will converge. Assuming ϵ to be sufficiently small and choosing δ to satisfy (4.33), we can write:

$$\begin{aligned} \bar{C}_{\text{CoMP_Opt}} &\approx \mathbb{E} \left\{ \text{tr}(\log(\Phi_k - \delta \mathbf{I}_{N_b N_t} + (\delta + 1) \mathbf{I}_{N_b N_t})) \right\}, \\ &= N_t N_b \log(\delta + 1) + \mathbb{E} \left\{ \text{tr} \left(\log \left(\frac{1}{\delta + 1} (\Phi_k - \delta \mathbf{I}_{N_b N_t}) + \mathbf{I}_{N_b N_t} \right) \right) \right\}. \end{aligned} \quad (4.34)$$

Using the Taylor expansion $\log(\mathbf{I} + \mathbf{A}) = \mathbf{A} - \frac{1}{2} \mathbf{A}^2 + \frac{1}{3} \mathbf{A}^3 - \dots$ with $\mathbf{A} = \frac{1}{\delta + 1} (\Phi_k - \delta \mathbf{I}_{N_b N_t})$ in (4.34), we have

$$\bar{C}_{\text{CoMP_Opt}} = N_t N_b \log(\delta + 1) + \mathbb{E} \left\{ \text{tr} \left(\log \left(\frac{1}{\delta + 1} (\Phi_k - \delta \mathbf{I}_{N_b N_t}) - \frac{1}{2(\delta + 1)^2} (\Phi_k - \delta \mathbf{I}_{N_b N_t})^2 - \dots \right) \right) \right\}. \quad (4.35)$$

We can observe from (4.35) that the first and second terms of the expansion of the optimal ergodic sum rate have the factors $\frac{1}{\delta + 1}$ and $\frac{1}{2(\delta + 1)^2}$, respectively. These factors play a key role in determining the dominance of each term (first and second) of the expansion in the optimal input covariance solution. If δ is not sufficiently large, the Taylor expansion (4.35) may not converge, on the other hand, if δ is set too large, the expansion will converge too fast for the first term to capture the dominant components of the expansion. Accordingly, we need to consider the factors $\frac{1}{\delta + 1}$ and $\frac{1}{2(\delta + 1)^2}$ in the second-order approximation of the ergodic sum rate given by equation (4.29).

Now, we have that our optimization problem is

$$\max_{\{\mathbf{Q}_k\}_{k=1}^K: \mathbf{Q}_k \geq 0, \sum_{k=1}^K \text{Tr}(\mathbf{Q}_k) \leq P} F(\mathbf{Q}_k), \quad (4.36)$$

with

$$\begin{aligned} F(\mathbf{Q}_k) &= \frac{1}{\delta + 1} \text{tr}(\mathbf{X}_k \mathbf{Q}_k) - \frac{1}{2(\delta + 1)^2} \left(\text{tr}(\mathbf{R}_{t_k} \mathbf{C}_k^{-1})^2 \cdot (\text{tr}(\mathbf{R}_{r_k} \mathbf{Q}_k))^2 + \right. \\ &\quad \left. + 2 \text{tr}(\mathbf{R}_{r_k} \mathbf{Q}_k) \cdot \text{tr}(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{t_k} \mathbf{C}_k^{-1}) + \text{tr}(\mathbf{X}_k \mathbf{Q}_k)^2 \right) \end{aligned} \quad (4.37)$$

where \mathbf{X}_k comes from (4.16) and δ is appropriately chosen so as to ensure that (4.33) is satisfied for the sum power of interest.

4.3.2 Convex Optimization of the Second-Order Approximation of the Ergodic Capacity

It is possible to show that the objective function in (4.36) is convex and strictly feasible for any $P > 0$ [37]. Accordingly, we can use the Karush-Kuhn-Tucker (KKT) conditions to attain both the primal and dual solutions of the optimization problem [37]. The Lagrangian of (4.37) can be written as

$$\begin{aligned} \mathbf{L}(\mathbf{Q}_k, \mathbf{Z}_k, \nu) &= -F(\mathbf{Q}_k) - \text{tr}(\mathbf{Z}_k \mathbf{Q}_k) + \nu \left(\sum_{k=1}^K \text{tr}(\mathbf{Q}_k) - P \right) \\ &= \frac{1}{2(\delta+1)^2} \left(\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 (\text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k))^2 + 2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \text{tr}(\bar{\mathbf{H}}_k^H \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1}) \right) + \\ &\quad + \text{tr}(\mathbf{X}_k \mathbf{Q}_k)^2 - \frac{1}{\delta+1} \text{tr}(\mathbf{X}_k \mathbf{Q}_k) - \text{tr}(\mathbf{Z}_k \mathbf{Q}_k) + \nu \left(\sum_{k=1}^K \text{tr}(\mathbf{Q}_k) - P \right) \end{aligned} \quad (4.38)$$

where \mathbf{Z}_k and ν are dual variables.

The KKT conditions for the optimization problem are

$$\begin{aligned} \nabla_{\mathbf{Q}_k} \mathbf{L} = 0 \Rightarrow & \frac{1}{2(\delta+1)^2} \left(2 \text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \mathbf{R}_{\mathbf{r}k} + 2 \text{tr}(\bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H \mathbf{Q}_k) \mathbf{R}_{\mathbf{r}k} + \right. \\ & \left. + 2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H + 2 \mathbf{X}_k \mathbf{Q}_k \mathbf{X}_k \right) - \frac{1}{\delta+1} \mathbf{X}_k - \mathbf{Z}_k + \nu \mathbf{I} = 0, \end{aligned} \quad (4.39a)$$

$$\mathbf{Q}_k \succeq 0, \quad (4.39b)$$

$$\nabla_{\mathbf{Z}_k} \mathbf{L} = 0 \Rightarrow \text{tr}(\mathbf{Z}_k \mathbf{Q}_k) = 0, \quad (4.39c)$$

$$\mathbf{Z}_k \succeq 0, \quad (4.39d)$$

$$\nabla_{\nu} \mathbf{L} = 0 \Rightarrow \sum_{k=1}^K \text{tr}(\mathbf{Q}_k) - P = 0 \Rightarrow \sum_{k=1}^K \text{tr}(\mathbf{Q}_k) = P \quad (4.39e)$$

Assuming that \mathbf{X}_k is invertible, we obtain the following input covariance matrix \mathbf{Q}_k that maximizes the function $F(\mathbf{Q}_k)$:

$$\begin{aligned} \mathbf{Q}_k = & \mathbf{X}_k^{-1} \left((\delta+1)^2 \left(\mathbf{Z}_k + \frac{1}{\delta+1} \mathbf{X}_k - \nu \mathbf{I} \right) - \left(\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) + \text{tr}(\bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H) \right) \mathbf{R}_{\mathbf{r}k} - \right. \\ & \left. - \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H \right) \mathbf{X}_k^{-1}. \end{aligned} \quad (4.40)$$

Making some mathematical manipulations in (4.40), we have that the input covariance that maximizes $F(\mathbf{Q}_k)$ can be simplified as

$$\mathbf{Q}_k = (\delta+1)^2 \mathbf{X}_k^{-1/2} \left(\tilde{\mathbf{Z}}_k + \frac{1}{\delta+1} \mathbf{I} - \nu \mathbf{X}_k^{-1} - \Theta_1 \tilde{\mathbf{R}}_{\mathbf{r}k} - \Theta_2 \tilde{\mathbf{S}}_k \right) \mathbf{X}_k^{-1/2}, \quad (4.41a)$$

where

$$\tilde{\mathbf{Z}}_k = \mathbf{X}_k^{1/2} \mathbf{Z}_k \mathbf{X}_k^{1/2}, \quad (4.41b)$$

$$\tilde{\mathbf{R}}_{\mathbf{r}k} = \mathbf{X}_k^{1/2} \mathbf{R}_{\mathbf{r}k} \mathbf{X}_k^{1/2}, \quad (4.41c)$$

$$\tilde{\mathbf{S}}_k = \mathbf{X}_k^{1/2} \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H \mathbf{X}_k^{1/2}, \quad (4.41d)$$

$$\Theta_1 = (\delta+1)^2 \left(\text{tr}(\mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1})^2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k) + \text{tr}(\bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}k} \mathbf{C}_k^{-1} \bar{\mathbf{H}}_k^H) \right), \quad (4.41e)$$

$$\Theta_2 = (\delta+1)^2 \text{tr}(\mathbf{R}_{\mathbf{r}k} \mathbf{Q}_k). \quad (4.41f)$$

Note that the input covariance matrix \mathbf{Q}_k obtained in (4.41a) is found in an iterative way. Moreover, we can observe that the matrix \mathbf{Z}_k , which is a variable of the Langrangian problem, is not defined yet. The matrix \mathbf{Z}_k is found by verifying whether the solution for \mathbf{Q}_k given by (4.41a) satisfy the KKT conditions given in (4.39).

We first verify the condition $\mathbf{Q}_k \succeq 0$ given in (4.39b). From the solution found for \mathbf{Q}_k in (4.41a), we write that the condition $\mathbf{Q}_k \succeq 0$ is equivalent to state

$$\tilde{\mathbf{Z}}_k + \frac{1}{\delta+1}\mathbf{I} + \nu\mathbf{X}^{-1} - \Theta_1\tilde{\mathbf{R}}_{\mathbf{r}_k} - \Theta_2\tilde{\mathbf{S}}_k \succeq 0. \quad (4.42)$$

and thus we can easily verify the condition $\mathbf{Q}_k \succeq 0$.

Now, verifying the condition $\text{tr}(\mathbf{Z}_k\mathbf{Q}_k) = 0$ given in (4.39c) and substituting the solution (4.41a) for \mathbf{Q}_k into (4.39c), we have

$$\begin{aligned} \text{tr}(\mathbf{Z}_k\mathbf{Q}_k) &= 0 \\ \text{tr}\left(\mathbf{Z}_k\mathbf{X}_k^{-1/2}\left(\tilde{\mathbf{Z}}_k + \frac{1}{\delta+1}\mathbf{I} + \nu\mathbf{X}^{-1} - \Theta_1\tilde{\mathbf{R}}_{\mathbf{r}_k} - \Theta_2\tilde{\mathbf{S}}_k\right)\mathbf{X}_k^{-1/2}\right) &= 0, \\ \text{tr}\left(\tilde{\mathbf{Z}}_k\left(\tilde{\mathbf{Z}}_k + \frac{1}{\delta+1}\mathbf{I} + \nu\mathbf{X}^{-1} - \Theta_1\tilde{\mathbf{R}}_{\mathbf{r}_k} - \Theta_2\tilde{\mathbf{S}}_k\right)\right) &= 0. \end{aligned} \quad (4.43)$$

Denoting the eigen decomposition of $\tilde{\mathbf{Z}}_k$ and $(\frac{1}{\delta+1}\mathbf{I} + \nu\mathbf{X}^{-1} - \Theta_1\tilde{\mathbf{R}}_{\mathbf{r}_k} - \Theta_2\tilde{\mathbf{S}}_k)$ by $\mathbf{U}_Z\mathbf{\Lambda}_Z\mathbf{U}_Z^H$ and $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, respectively, we have that (4.43) can be written as

$$\begin{aligned} \text{tr}(\mathbf{U}_Z\mathbf{\Lambda}_Z\mathbf{U}_Z^H(\mathbf{U}_Z\mathbf{\Lambda}_Z\mathbf{U}_Z^H + \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H)) &= 0, \\ \text{tr}(\mathbf{\Lambda}_Z)^2 + \text{tr}(\mathbf{U}_Z\mathbf{\Lambda}_Z\mathbf{U}_Z^H\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H) &= 0. \end{aligned} \quad (4.44)$$

If we denote

$$\mathbf{U}_Z = \mathbf{U}, \quad (4.45)$$

the optimality of the input covariance matrix will not be affected since the optimality is guaranteed by satisfying the KKT conditions, and so we have that equation (4.44) can be rewritten as

$$\begin{aligned} \text{tr}(\mathbf{\Lambda}_Z)^2 + \text{tr}(\mathbf{\Lambda}_Z\mathbf{\Lambda}) &= 0, \\ \text{tr}(\mathbf{\Lambda}_Z(\mathbf{\Lambda}_Z + \mathbf{\Lambda})) &= 0, \\ \text{tr}\left(\begin{bmatrix} \mathbf{\Lambda}_{Z1} & 0 \\ 0 & \mathbf{\Lambda}_{Z2} \end{bmatrix}\left(\begin{bmatrix} \mathbf{\Lambda}_{Z1} & 0 \\ 0 & \mathbf{\Lambda}_{Z2} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda}^+ & 0 \\ 0 & \mathbf{\Lambda}^- \end{bmatrix}\right)\right) &= 0. \end{aligned} \quad (4.46)$$

where $\mathbf{\Lambda}^+$ and $\mathbf{\Lambda}^-$ are the matrices formed by non-negative and negative entries of $\mathbf{\Lambda}$, respectively. The matrices $\mathbf{\Lambda}_{Z1}$ and $\mathbf{\Lambda}_{Z2}$ have the same dimensions of $\mathbf{\Lambda}^+$ and $\mathbf{\Lambda}^-$, respectively.

In order to find the matrix \mathbf{Z}_k that satisfies the condition (4.46) we can first affirm, from both equations $\mathbf{Z}_k \succeq 0$ in (4.39d) and $\tilde{\mathbf{Z}}_k = \mathbf{X}_k^{1/2}\mathbf{Z}_k\mathbf{X}_k^{1/2}$ in (4.41b), that the matrix $\tilde{\mathbf{Z}}$ is definite positive, and thus the matrices $\mathbf{\Lambda}_{Z1}$ and $\mathbf{\Lambda}_{Z2}$ are also definite positive, i.e.,

$$\mathbf{\Lambda}_{Z1} \succeq 0 \quad \text{and} \quad \mathbf{\Lambda}_{Z2} \succeq 0. \quad (4.47)$$

In sequel, substituting the eigen decompositions of both $\tilde{\mathbf{Z}}_k$ and $(\frac{1}{\delta+1}\mathbf{I} + \nu\mathbf{X}^{-1} - \Theta_1\tilde{\mathbf{R}}_{\mathbf{r}_k} -$

$\Theta_2 \tilde{\mathbf{S}}_k$), given by $\mathbf{U}_Z \mathbf{\Lambda}_Z \mathbf{U}_Z^H$ and $\mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ respectively, in (4.42) we have

$$\begin{aligned} \tilde{\mathbf{Z}}_k + \frac{1}{\delta + 1} \mathbf{I} + \nu \mathbf{X}^{-1} - \Theta_1 \tilde{\mathbf{R}}_{\mathbf{r}_k} - \Theta_2 \tilde{\mathbf{S}}_k \succeq 0, \\ \mathbf{U}_Z \mathbf{\Lambda}_Z \mathbf{U}_Z^H + \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \succeq 0, \end{aligned} \quad (4.48)$$

and using the statement made in (4.45), we can rewrite the equation (4.48) by

$$\mathbf{U} \begin{bmatrix} \mathbf{\Lambda}_{Z1} + \mathbf{\Lambda}^+ & 0 \\ 0 & \mathbf{\Lambda}_{Z2} + \mathbf{\Lambda}^- \end{bmatrix} \mathbf{U}^H \succeq 0,$$

which is equivalent to state that

$$\mathbf{\Lambda}_{Z1} + \mathbf{\Lambda}^+ \succeq 0 \quad \text{and} \quad \mathbf{\Lambda}_{Z2} + \mathbf{\Lambda}^- \succeq 0. \quad (4.49)$$

The only solution that satisfies (4.46), (4.47) and (4.49) is

$$\mathbf{\Lambda}_{Z1} = 0, \quad (4.50)$$

$$\mathbf{\Lambda}_{Z2} = -\mathbf{\Lambda}^-. \quad (4.51)$$

Therefore, we formulate the solution for the input covariance matrices that maximizes the second-order approximation of the ergodic sum rate, given in equation (4.37). This formulation follows some steps, which are summarized in Algorithm 4.1.

Algorithm 4.1 Optimization of the Multi-Cell Multi-User MIMO Ergodic Capacity

- 1: The input covariance matrices are first initialized to a scaled version of the identity, $\mathbf{Q}_k^0 = (P/N_r K) \mathbf{I}_{N_r}$
 - 2: Set the parameters δ and ν
 - 3: **while** Capacity does not converge **do**
 - 4: **for** each user k **do**
 - 5: Evaluate the matrix $\mathbf{C}_k = \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i-1)} \mathbf{H}_j$
 - 6: Evaluate the parameters Θ_1 and Θ_2 using (4.41e) and (4.41f)
 - 7: Evaluate the matrices \mathbf{X}_k , $\tilde{\mathbf{R}}_{\mathbf{r}_k}$, $\tilde{\mathbf{S}}_k$ using (4.16), (4.41c) and (4.41d)
 - 8: Evaluate the matrices \mathbf{U} and $\mathbf{\Lambda}$, given that $\mathbf{U} \mathbf{\Lambda} \mathbf{U}^H = (\frac{1}{\delta+1} \mathbf{I} + \nu \mathbf{X}_k^{-1} - \Theta_1 \tilde{\mathbf{R}}_{\mathbf{r}_k} - \Theta_2 \tilde{\mathbf{S}}_k)$
 - 9: Evaluate $\tilde{\mathbf{Z}}_k = \mathbf{U}_Z \mathbf{\Lambda}_Z \mathbf{U}_Z^H$ using (4.45), (4.50) and (4.51)
 - 10: Evaluate $\mathbf{Q}_k^{(n)}$ using (4.41a)
 - 11: **end for**
 - 12: Normalize the matrices $\mathbf{Q}_k^{(i)}$ in order to obey the per-BS power constraint.
 - 13: Evaluate the sum rate and verify convergence
 - 14: **end while**
-

In order to normalize the input covariance matrices and obey the per-BS power constraints, we first map the uplink input covariance matrices $\mathbf{Q}_k^{(i)}$ of all users into the corresponding downlink input covariance matrices $\mathbf{\Omega}_k^{(i)}$ [15]. Next, we normalize these downlink input covariance matrices $\mathbf{\Omega}_k^{(i)}$ for all users by the trace of the downlink input covariance matrix with maximum trace. This normalization is made in a suboptimal way, so that only the BS satisfying the minimum value can transmit with full power and any other BS transmits with a power lower than its power constraint.

The flow chart of the solution steps of our algorithm is shown in figure 4.2.

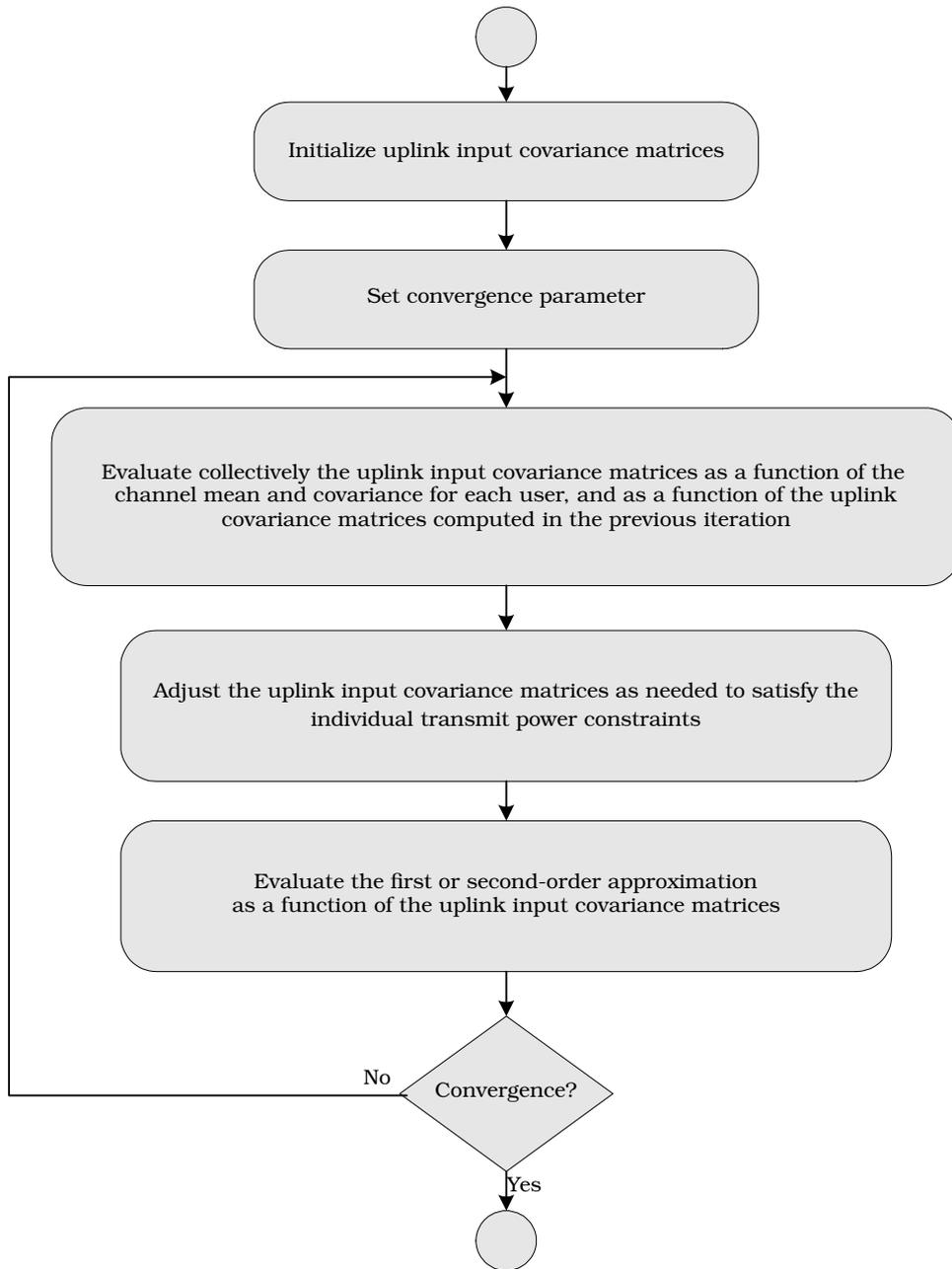


Figure 4.2: Flow chart of the proposed algorithm.

4.4 Simulations and Results

4.4.1 Scenario, Channel Model and Main Simulation Parameters

We assume the CoMP cell scenario consisting of 3 coordinated cells with BSs placed in the center of each cell. Initially, one user is placed randomly in each cell. These users are considered as the ones selected by a scheduling algorithm to transmit. The most relevant parameters are the same considered in the simulations of Chapter 3 on Table 3.1 and the spatial antenna correlations are fixed for all BSs and users as the same in (3.40) and (3.41) [65].

4.4.2 A Multi-user Analysis of the Proposed Transmit Scheme

In figures 4.3(a), 4.3(b), 4.4(a), 4.4(b), 4.5(a), 4.5(b), 4.6(a) and 4.6(b) we compare the average sum rate obtained with the proposed algorithms and the upper bound on the ergodic

sum rate which corresponds to the perfect CSIT and performs an iterative water-filling [14]. In figures both the number of transmit and receive antennas (N_t and N_r) and the Ricean factor κ vary. The proposed algorithms consider the case where the ergodic sum rate is approximated by only the first term (FOIC) and by the two first terms (SOIC) of the Taylor expansion. In order to have a better view, we divide the SNRs range in low and high values.

In figures 4.3(a) and 4.3(b) we have the scenario with 2 transmit and receive antennas and the Ricean factor κ equal to 0. We notice from figure 4.3(a) that, in low SNRs, the proposed algorithms have similar performance and that the difference gap between them and the upper bound is high. When the signal-to-noise ratio (SNR) increases, we note from figure 4.3(b), that the performance of the proposed algorithms improves and the SOIC algorithm performs close to the iterative water-filling algorithm. In figures 4.4(a) and 4.4(b) the Ricean factor κ is increased to 3. In low SNRs the proposed algorithms have the same performance and the difference gap between them and the upper bound is the same for the case with $\kappa = 0$. But in high SNRs (figure 4.4(b)) the performance is a little bit different than the case with $\kappa = 0$, we can note that the SOIC algorithm obtains good results but there is still a gap between it and the iterative water-filling algorithm. Thus, we conclude that, when the Ricean factor increases, the performance gap between the proposed algorithm FOIC and the iterative water-filling one keeps the same, but the performance gap between the proposed algorithm SOIC and the iterative water-filling one has a little decrease.

Figures 4.5(a), 4.5(b), 4.6(a) and 4.6(b) performs the same comparison when the number of transmit and receive antennas is equal to 4 and the Ricean factor varies. From those figures we observe the same behavior as for the case with 2 transmit and receive antennas. The proposed algorithms obtain good performance results when compared to the iterative water-filling algorithm and, when the Ricean factor increases, the performance gap between the SOIC algorithm and the iterative water-filling becomes slightly greater at high SNRs.

As mentioned earlier, in the SOIC algorithm, we have to choose appropriately the parameter δ in order to ensure the convergence of the algorithm. Figures 4.7(a) and 4.7(b) show the ergodic sum rate convergence when the convergence parameter δ varies and the SNR values are set to 12 dB and 15 dB, respectively. The considered scenario is $N_t = N_r = 2$ and Ricean factor $\kappa = 3$. We note that in each figure, we have divergent curve, rapid convergent curve and slow convergent curve. Moreover, we observe that the parameter δ must be chosen not too small in order to have the convergence, and also not too high since, in the fast convergence, the obtained ergodic sum rate after the convergence is smaller than the ergodic sum rate obtained in the slow convergence case.

From all figures 4.3(a), 4.3(b), 4.4(a), 4.4(b), 4.5(a), 4.5(b), 4.6(a) and 4.6(b), we can summarize our simulation analyzes in stating that the performance gap between the proposed algorithms and the bound-achieving algorithm is small and that the convergence rate of the proposed algorithm SOIC is high given that the parameter δ is rightly chosen.

4.4.3 A Single-user Analysis of the Proposed Transmit Scheme

We have compared our proposals with the iterative waterfilling algorithm proposed in [8]. This comparative algorithm is like a bound on the ergodic sum rate. The reason to compare only with this technique is that we could not find any multi-user CoMP technique based on statistical precoder.

In order to compare our proposals (FOIC and SOIC) with other techniques already existent, we simplify our scenario to a single-user one. Thus, after performing the scheduling algorithm to choose one user per cell, we select the user in which its joint channel given in (3.42) has

the highest Frobenius norm.

The single-user techniques used in the comparison are:

- ▶ Mean-optimal signalling [54]: This technique ignores the covariance information and treats the mean as if it were the true channel. The signalling is performed along the eigen basis of the channel mean and “water-fills” over its eigenvalues.
- ▶ Covariance-optimal signalling [54]: This technique ignores the mean information and transmits along the eigenvectors of the transmit covariance matrix.

Figures 4.8(a), 4.8(b), 4.9(a) and 4.9(b) show a comparison among the ergodic rate obtained using the proposed techniques (FOIC and SOIC), the mean-optimal signalling and the covariance-optimal signalling when the number of receive antennas and the Ricean factor vary. The considered channel model to estimate the channel at the transmitter is the Kronecker channel model given in equation (2.42).

Figure 4.8(a) shows the comparison with the number of transmit and receive antennas equal to 2 and the Ricean factor κ equal to 3. We notice that the proposed algorithms obtain quite similar performances and that these results are better than the result obtained with the comparative techniques. When can also note that the comparative technique based on the channel mean performs better than the other comparative technique, which is based on the channel covariance. Figure 4.8(b) shows the comparison among the techniques with a higher Ricean factor ($\kappa = 10$). We observe that the proposed algorithms also obtain a good performance, with the mean-optimal signalling performance being closer to the proposed techniques. This phenomenon is explained since this technique is based on the channel mean matrix, and when the factor κ increases, the strength of the channel mean matrix in the channel becomes greater. Thus, the results of the mean-optimal signalling are better when κ increases.

Figure 4.9(a) shows the comparison among the techniques with a higher number of receive antennas. The Ricean factor κ is set to 3. From this figure we note that the proposed techniques obtain better performance and that the performance gap between the proposed techniques FOIC and SOIC increases with a higher number of receive antennas. It is also interesting to observe that the optimal-covariance signalling obtains better results than the other comparative technique, except for very low SNR values. This happens because, when the number of receive antennas increases, the channel is more correlated and since the covariance-optimal technique is based on the channel correlation, it is expected that the results of this technique become better in this case.

In figure 4.9(b) we have the comparison among the techniques in the same scenario (2 transmit antennas and 4 receive antennas) and with the Ricean factor $\kappa = 10$. We observe that the performance gap between the proposed techniques is higher than in the scenario with 2 receive antennas and that the proposed techniques obtain better results. Moreover, as already seen in the case with 2 transmit and receive antennas, the performance of the mean-optimal signalling becomes better when the Ricean factor increases. The covariance-optimal signalling also obtains good results since we have a scenario with more antennas, and so more correlated channels.

Thus, we conclude that, in the single-user case scenario, the two proposed schemes perform closer and the performance gap between them increases when the number of receive antennas increases. Moreover, the proposed algorithms perform better than the other compared techniques in all simulated scenarios. Next, we simulate the same single-user scenario but considering the proposed statistical channel model.

4.4.4 A Single-user Analysis of the Proposed Transmit Scheme with the Proposed Channel Model

In chapter 3 we proposed a partial CSIT model. This model is based on the channel statistics and is obtained taking into account the temporal correlation parameter. Figures 4.10(a) and 4.10(b) show a comparison of the average rate obtained by the proposed technique SOIC using both the Kronecker model and the proposed partial channel model. In order to simulate the proposed CSIT model in the SOIC technique, we use the effective mean and covariance obtained through this model (equations (3.33) and (3.23)) as channel mean and covariance, respectively.

In figure 4.10(a) we have a scenario with 2 transmit and receive antennas and the Ricean factor $\kappa = 0$. From this figure we note that the proposed CSIT model obtains better sum rate results than the Kronecker channel model. This is explained because this model takes into account the channel temporal variation and thus can model the real channel in a more precise way.

Figure 4.10(b) shows the same comparison but now with a Ricean factor $\kappa = 3$. We note from this figure that the performance of the proposed CSIT is again better than the Kronecker model but the performance gap between them becomes smaller with a higher Ricean factor.

Figures 4.11(a), 4.11(b), 4.12(a) and 4.12(b) show the ergodic rate obtained using the proposed techniques (FOIC and SOIC), the mean-optimal signalling and the covariance-optimal signalling, in which the proposed techniques are considering the proposed CSIT model. Figure 4.11(a) shows the comparison for a scenario with 2 transmit and receive antennas and the Ricean factor $\kappa = 3$. We notice that the proposed techniques obtain better performance than the other ones and that the performance of the proposed techniques FOIC and SOIC are similar. Moreover, as the same case using Kronecker channel model, the optimal-mean signalling outperforms the optimal-covariance signalling. In figure 4.11(b) the Ricean factor is changed for a value $\kappa = 10$. The performance of the proposed techniques is superior again and the performance of the mean-optimal signalling improves with a higher Ricean factor. This has been already explained and it is due to the fact that, when the factor κ increase, the strength of the channel mean matrix in the channel becomes greater.

In figure 4.12(a) the number of receive antennas is increased to 4 and the Ricean factor is $\kappa = 3$. We note from the figure that the proposed techniques outperform the other compared techniques. Moreover, the performance gap between the FOIC and SOIC is increased when compared to the case with 2 receive antennas. As observed in the comparison using the Kronecker channel model, the optimal-covariance signalling obtains better results than the other compared technique, except for low SNR values. This has been already analyzed, and is because the channel is more correlated when the number of receive antennas increases.

In figure 4.12(b) we have the same scenario (2 transmit antennas and 4 receive antennas) and with the Ricean factor $\kappa = 10$. We also observe that the performance gap between the proposed techniques is higher than in the scenario with 2 receive antennas. Moreover, we observe that the mean-optimal signalling has improved performance when compared to the case with the same number of receive antennas and Ricean factor $\kappa = 3$.

4.4.5 Analysis of the Phase Mismatch between the Uplink and Downlink Input Covariance Matrices

In this thesis, based on the duality theory, we have assumed that the uplink channel is the conjugate transpose of the corresponding dual downlink channel. In practical situations, studies have shown that, in magnitude, the downlink and uplink channels differ only by a

gain in all frequencies range, which it is easier to be compensated. But these studies have also shown that the uplink and downlink channel phase plots deviate from each other in phase by a few radians [70]. In order to analyze this aspect, we introduced a phase mismatch in the channel model but consider that the proofs of our algorithm will not change, i.e., the complete derivation of our proposal considering phase mismatches is left for a future study. Accordingly, we consider that the uplink channel model is given by

$$\mathbf{H}_{\text{UL}} = \mathbf{H}_{\text{DL}}^H \exp(j\theta) \quad (4.52)$$

where θ is the phase mismatch given in the downlink channel model. In simulations, we consider that the parameter θ is a random variable generated uniformly within the interval $[-l, l]$, where l is given in degrees. Substituting the downlink channel model given by equation (2.42) in (4.52), we have

$$\begin{aligned} \mathbf{H}_{\text{UL}} &= \left(\sqrt{\frac{\kappa}{1+\kappa}} \bar{\mathbf{H}}_k + \sqrt{\frac{1}{1+\kappa}} \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}[n] \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^H \exp(j\theta) \\ &= \left(\sqrt{\frac{\kappa}{1+\kappa}} \bar{\mathbf{H}}_k \exp(-j\theta) + \sqrt{\frac{1}{1+\kappa}} \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}[n] \mathbf{R}_{\mathbf{t}_k}^{1/2} \exp(-j\theta) \right)^H \\ &= \left(\sqrt{\frac{\kappa}{1+\kappa}} \bar{\mathbf{H}}_k \exp(-j\theta) + \sqrt{\frac{1}{1+\kappa}} (\mathbf{R}_{\mathbf{r}_k} \exp(-j\theta))^{1/2} \mathbf{H}_{\mathbf{w}_k}[n] (\mathbf{R}_{\mathbf{t}_k} \exp(-j\theta))^{1/2} \right)^H \end{aligned}$$

Therefore, we simulate our proposed algorithm and consider as channel mean the matrix $\bar{\mathbf{H}}_k \exp(-j\theta)$ and as channel transmit and receive covariance the matrices $(\mathbf{R}_{\mathbf{t}_k} \exp(-j\theta))$ and $(\mathbf{R}_{\mathbf{r}_k} \exp(-j\theta))$, respectively.

Figures 4.13(a), 4.13(b) and 4.13(c) show the average sum rate obtained with our proposed algorithm SOIC using the Kronecker channel model which considers the phase mismatch between uplink and downlink. The considered scenario has 2 transmit and receive antennas and the interval of distribution of the phase mismatch parameter θ is varied. Figure 4.13(a) shows the comparison for the case with Ricean factor $\kappa = 0$. We note that when the phase mismatch parameter θ increases the performance of the algorithm degrades. Figures 4.13(b) and 4.13(c) show the comparison when the Ricean factor is equal to 3 and 10, respectively. From these figures we observe that when the Ricean factor increases, the influence of the phase mismatch in the performance decreases. This happens since the presence of line-of-sight (LOS) component in the channel, even a phase-mismatched one, makes the channel more stable against random perturbations.

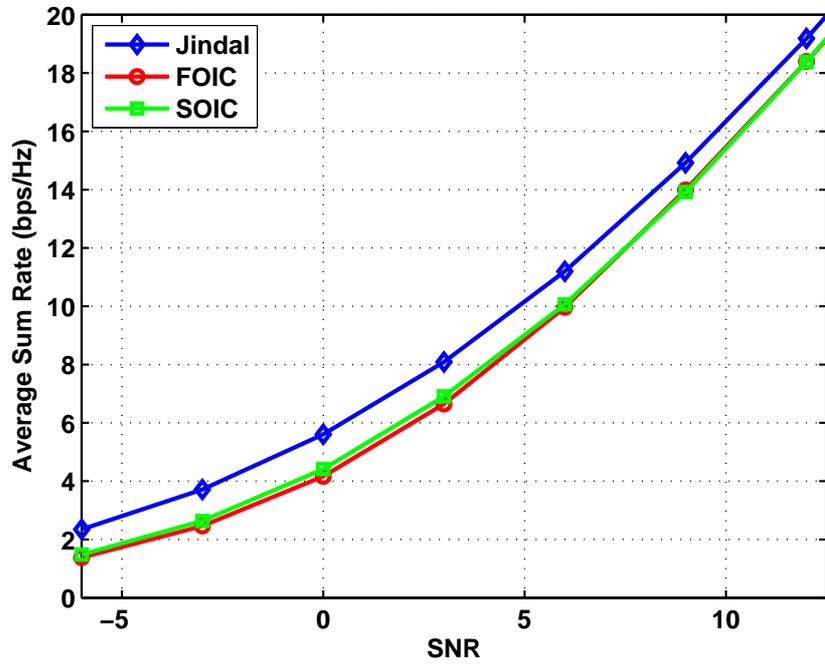
Figures 4.14(a), 4.14(b) and 4.14(c) show the average sum rate obtained with our proposed algorithm SOIC using the channel model proposed in Chapter 3, and considering the phase mismatch between uplink and downlink. The considered scenario has 2 transmit and receive antennas and the interval of distribution of the phase mismatch parameter θ is varied. Figure 4.14(a) shows the comparison for the case with Ricean factor $\kappa = 0$. Since the curves obtained are too close we show the curves in an SNR interval $(25.4995, 25.5)$ in order to have a better view. We note that when the phase mismatch parameter θ increases the performance of the algorithm degrades very slightly. Comparing with the Kronecker channel model, we notice that the proposed channel model is more robust to phase mismatch since the performance degradation obtained using the proposed channel model is very small when the phase mismatch parameter increases.

Figures 4.14(b) and 4.14(c) show the comparison when the Ricean factor is equal to 3 and 10, respectively. From these figures we observe, as with the Kronecker channel model, that when the Ricean factor increases, the influence of the phase mismatch in the performance decreases.

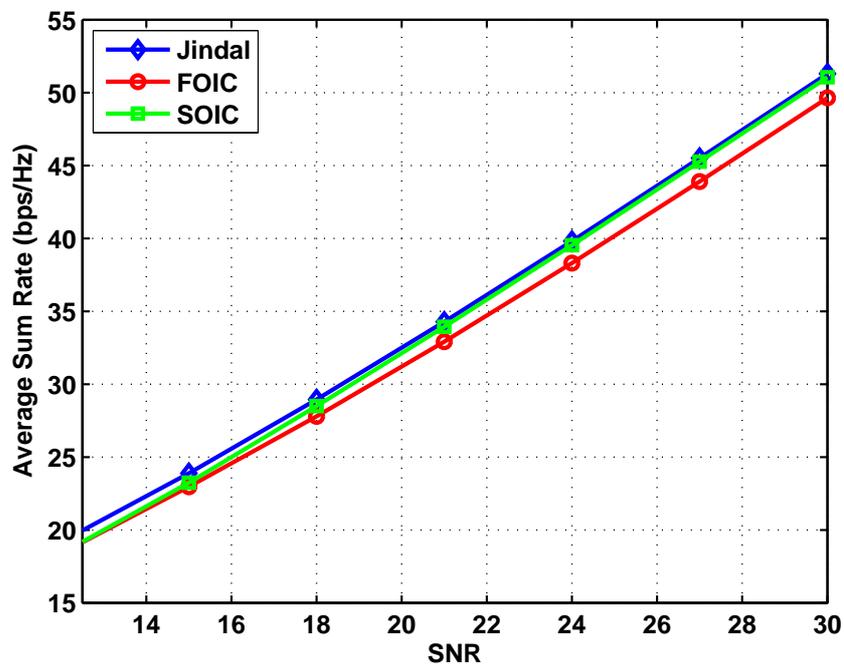
Hence, we summarize that the proposed statistical transmit scheme with the proposed channel model is a robust scheme against the possible phase mismatch between the uplink and downlink channels.

4.5 Summary

In this chapter, we optimized the input covariance matrix of a downlink multiuser MIMO CoMP system aiming at the maximization of an ergodic sum rate approximation. The scenario considers a multi-user CoMP system with joint processing, which is an architecture of great interest for future wireless systems due, e.g., to macrodiversity and good channel conditioning. Assuming that the transmitter has both the mean and covariance of the channel, first- and second-order approximations of the ergodic sum rate were found and the near-optimal input covariance matrix per user that maximizes the first- and second-order approximation of the ergodic sum rate was derived using convex optimization tools. The results shown that the performance gap between the proposed algorithms and the bound-achieving algorithm is small and that the convergence rate of the proposed algorithms is high given that the parameter δ is rightly chosen. Moreover, we analyzed the impact of a phase mismatch between the uplink and the dual downlink channel model, which is a more realistic model when considering the duality theorem. Simulation results shown that the statistical precoder is robust against the inserted perturbation. Moreover, from these simulation results, we could observe the channel model taking into account the channel temporal correlation is more powerful against the phase mismatch than the Kronecker channel model.

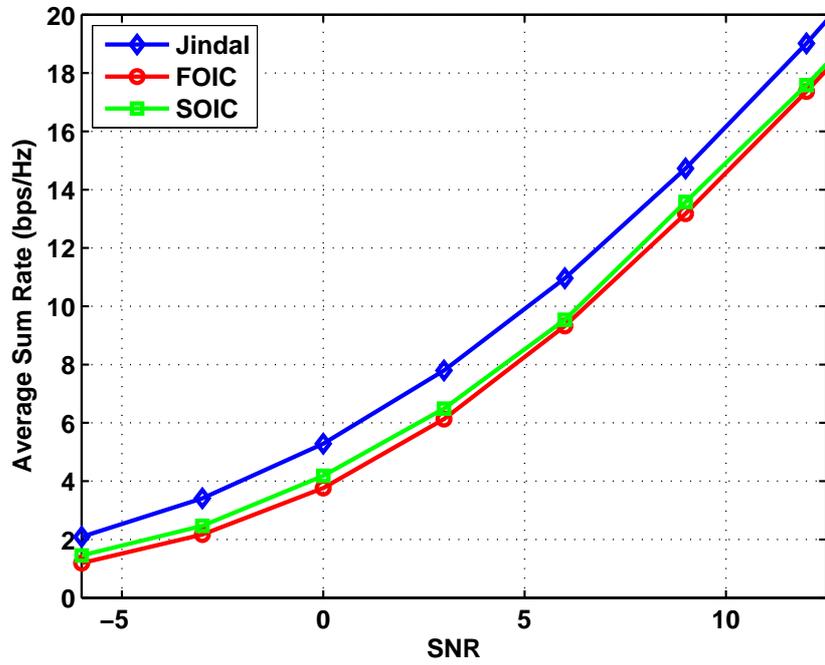


(a) Low SNRs

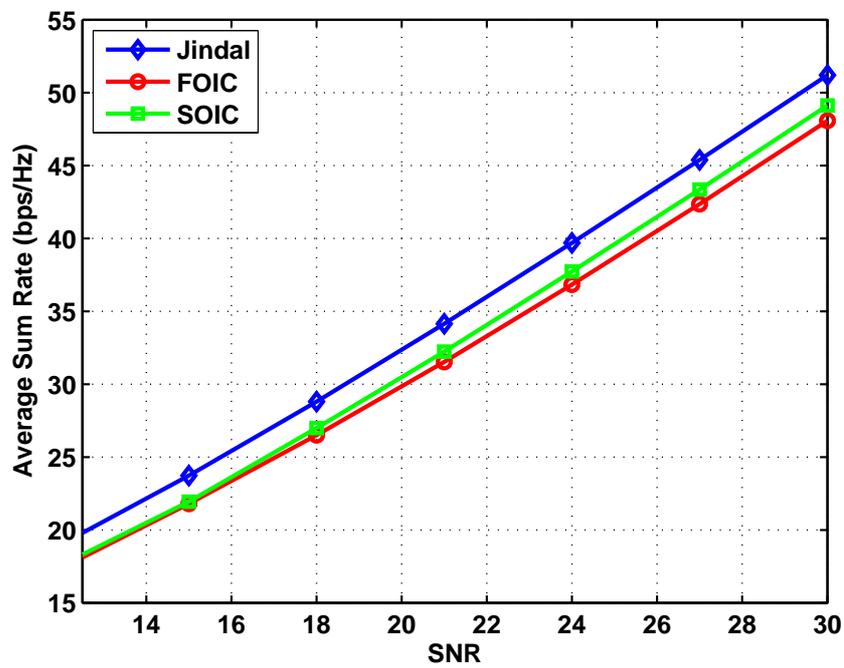


(b) High SNRs

Figure 4.3: Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 2$ and $\kappa = 0$.

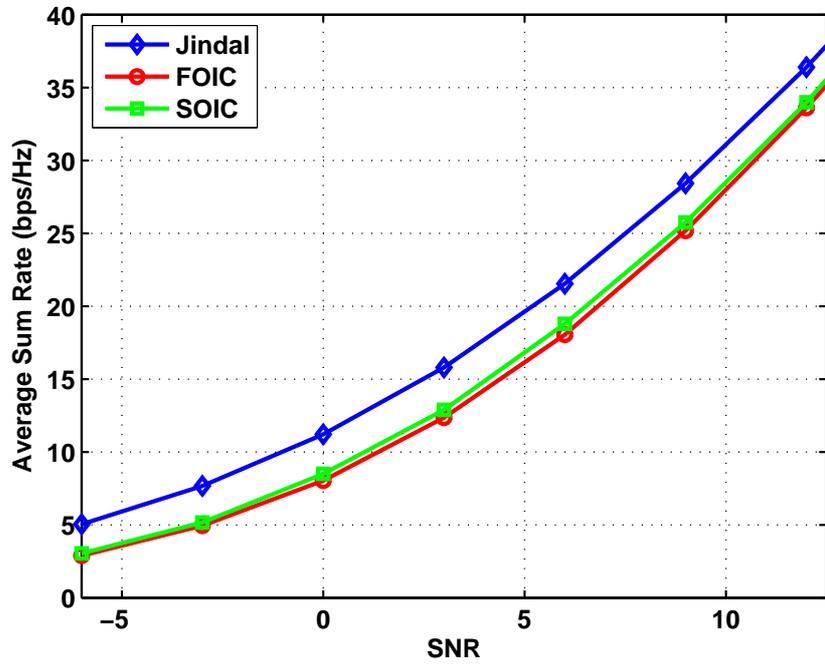


(a) Low SNRs

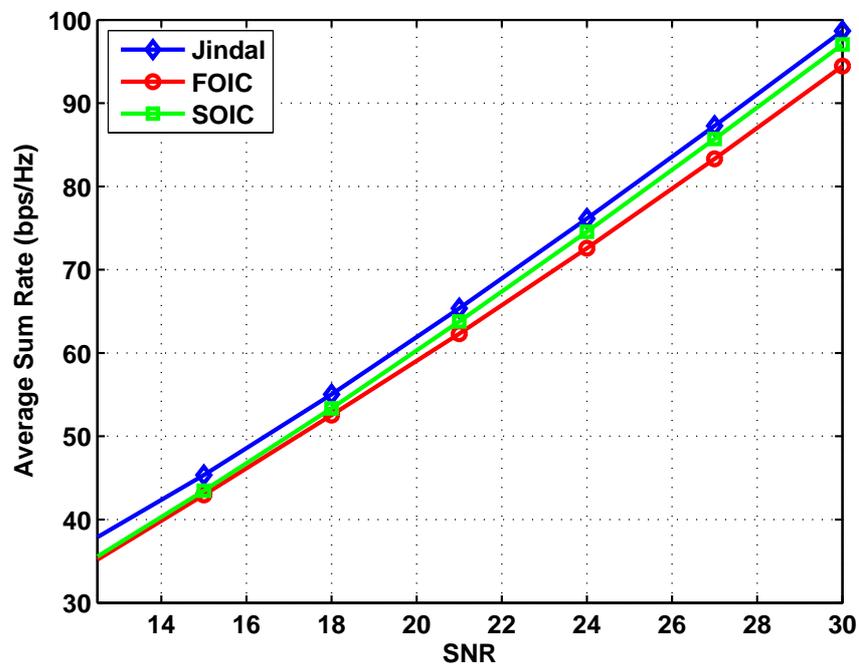


(b) High SNRs

Figure 4.4: Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 2$ and $\kappa = 3$.

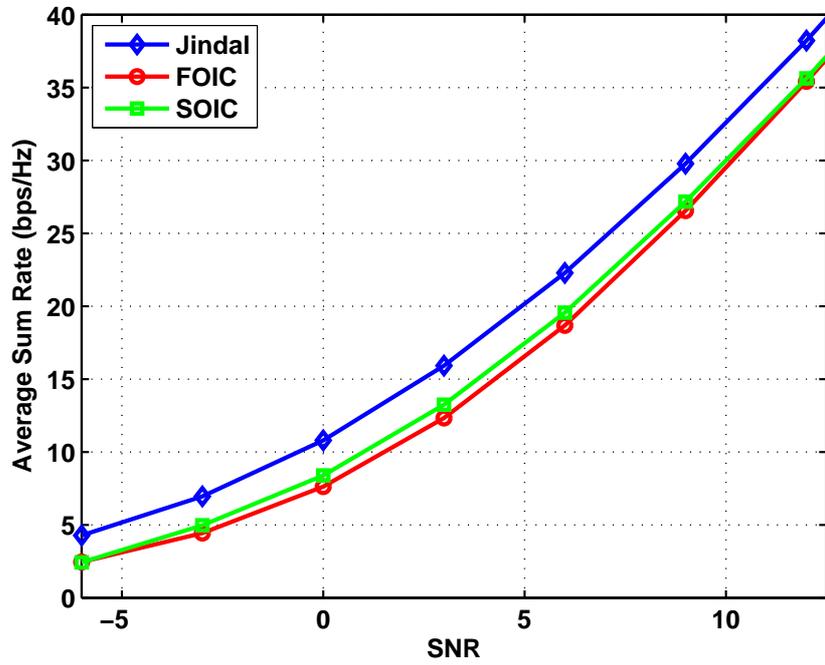


(a) Low SNRs

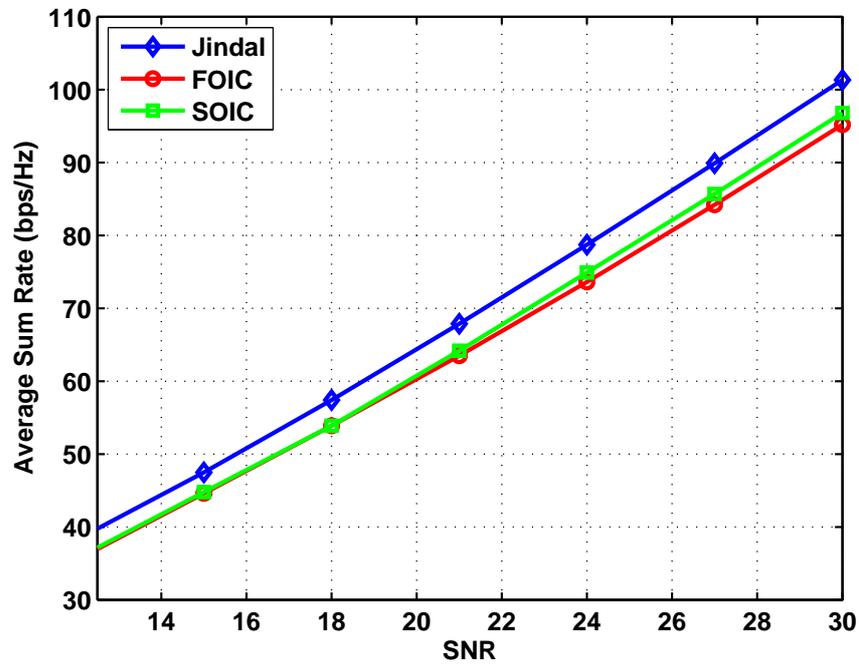


(b) High SNRs

Figure 4.5: Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 4$ and $\kappa = 0$.

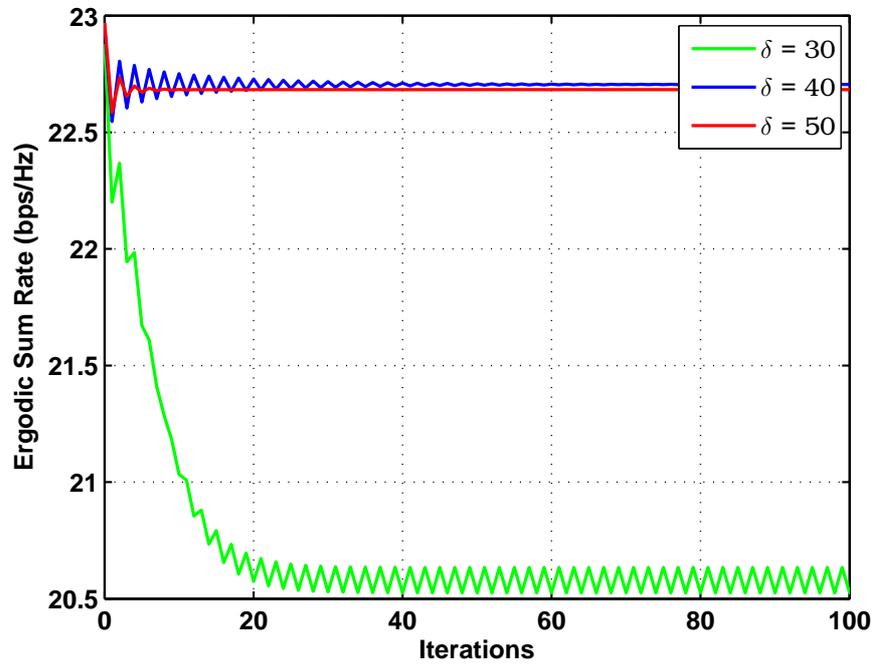


(a) Low SNRs

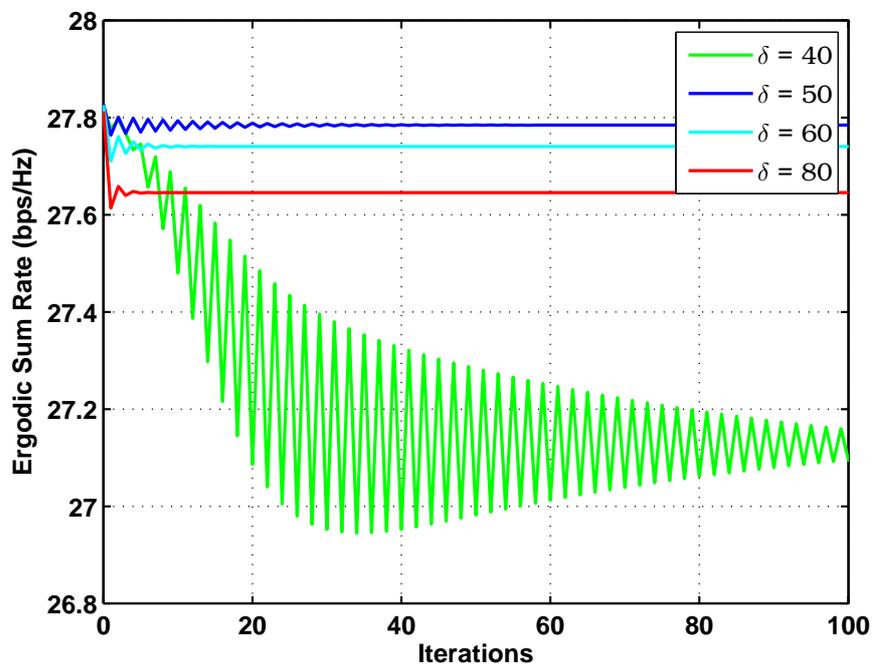


(b) High SNRs

Figure 4.6: Ergodic sum rate curves considering the Kronecker channel model, $N_t = N_r = 4$ and $\kappa = 3$.



(a) SNR = 12 dB



(b) SNR = 15 dB

Figure 4.7: Analysis of the convergence of the SOIC approach. $N_t = N_r = 2$ and $\kappa = 3$.

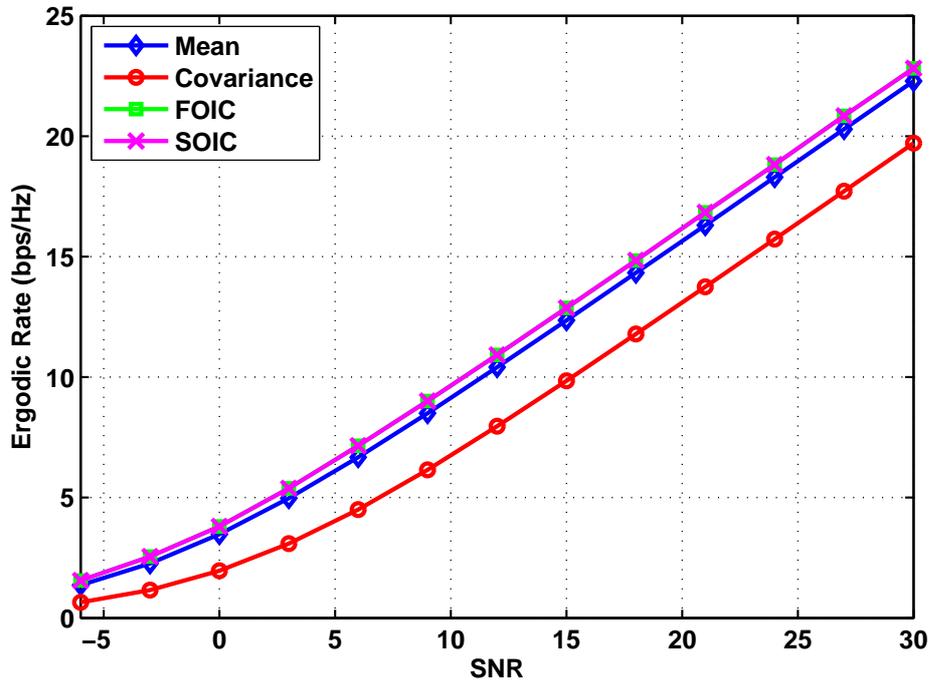
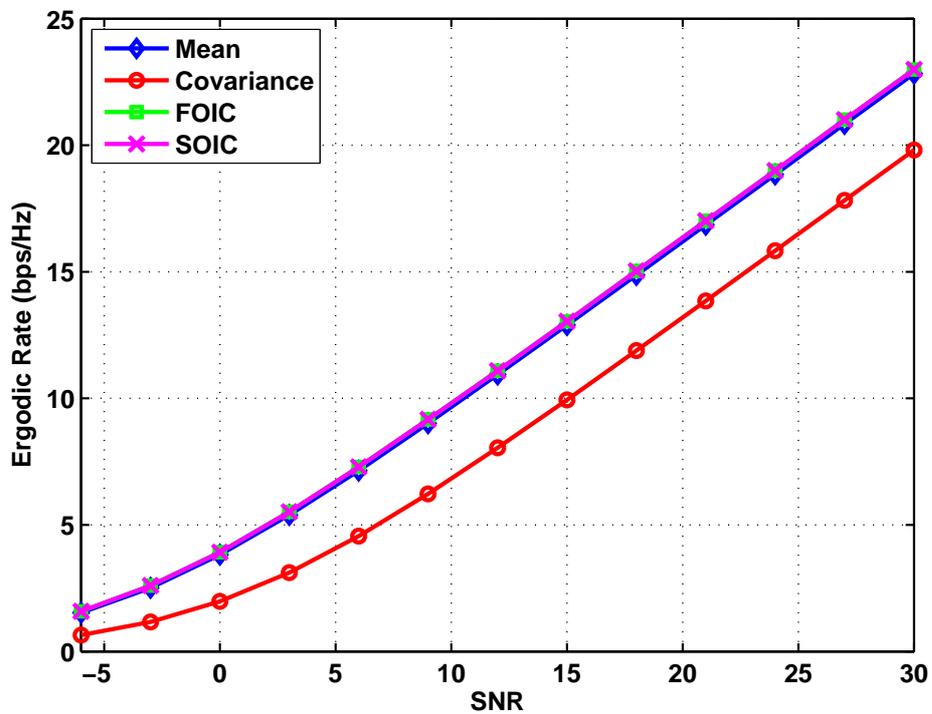
(a) Ricean factor $\kappa = 3$ (b) Ricean factor $\kappa = 10$

Figure 4.8: Ergodic rate curves of the single-user scenario considering the Kronecker channel model and $N_t = N_r = 2$.

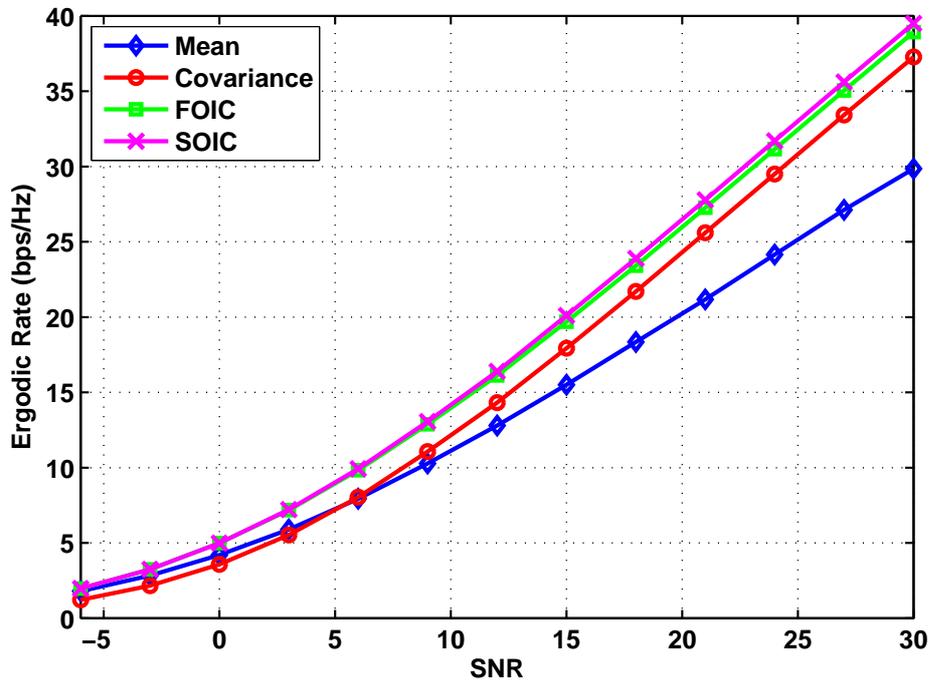
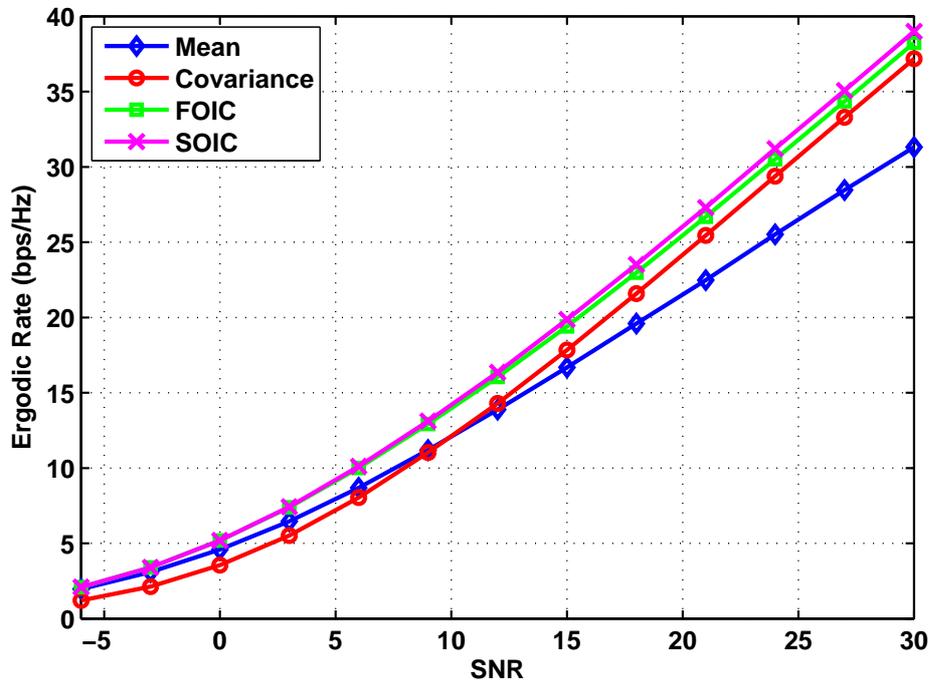
(a) Ricean factor $\kappa = 3$ (b) Ricean factor $\kappa = 10$

Figure 4.9: Ergodic rate curves of the single-user scenario considering the Kronecker channel model, $N_t = 2$ and $N_r = 4$.

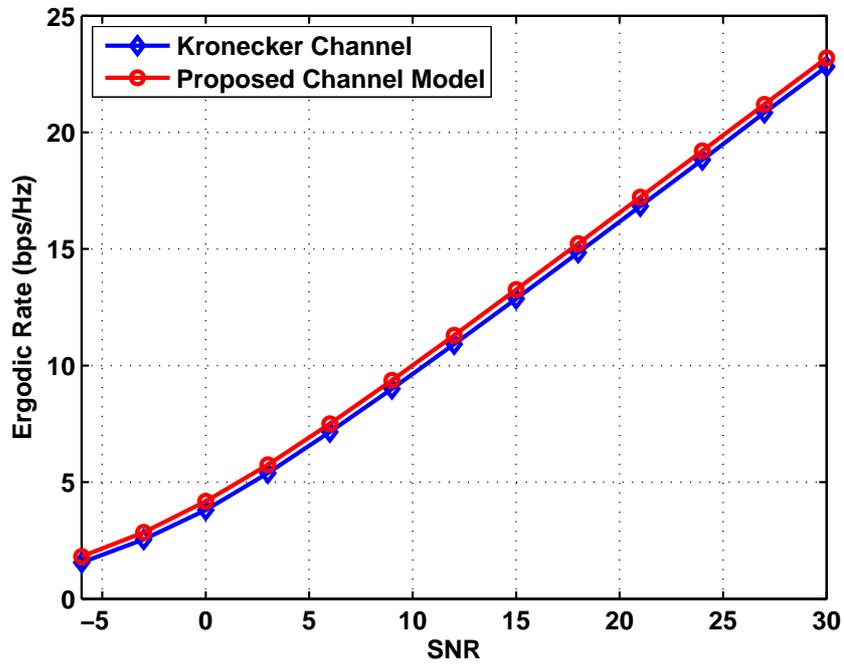
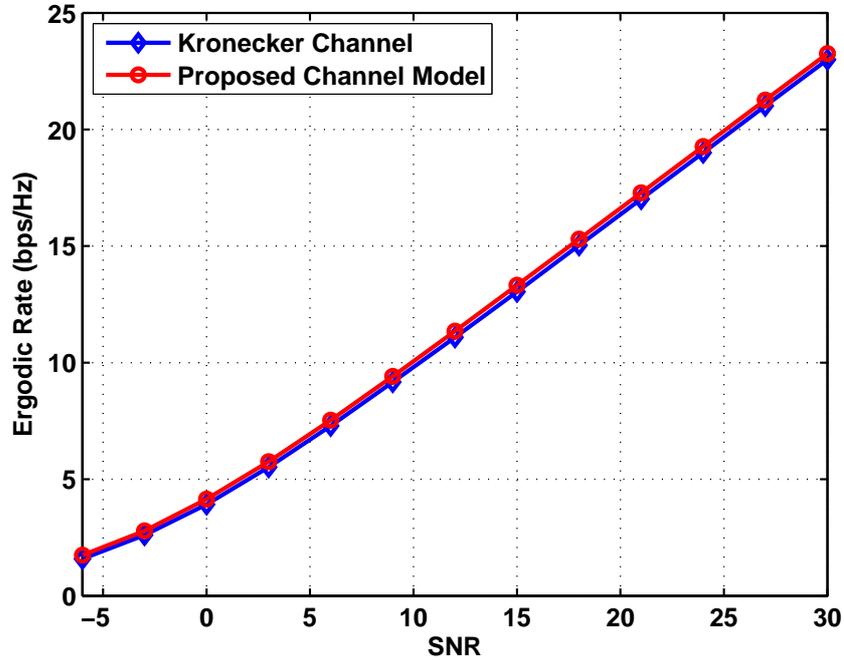
(a) Ricean factor $\kappa = 0$ (b) Ricean factor $\kappa = 3$

Figure 4.10: Comparison of the ergodic rate curves obtained by SOIC approach considering the Kronecker channel model and the proposed one. Scenario $N_t = N_r = 2$.

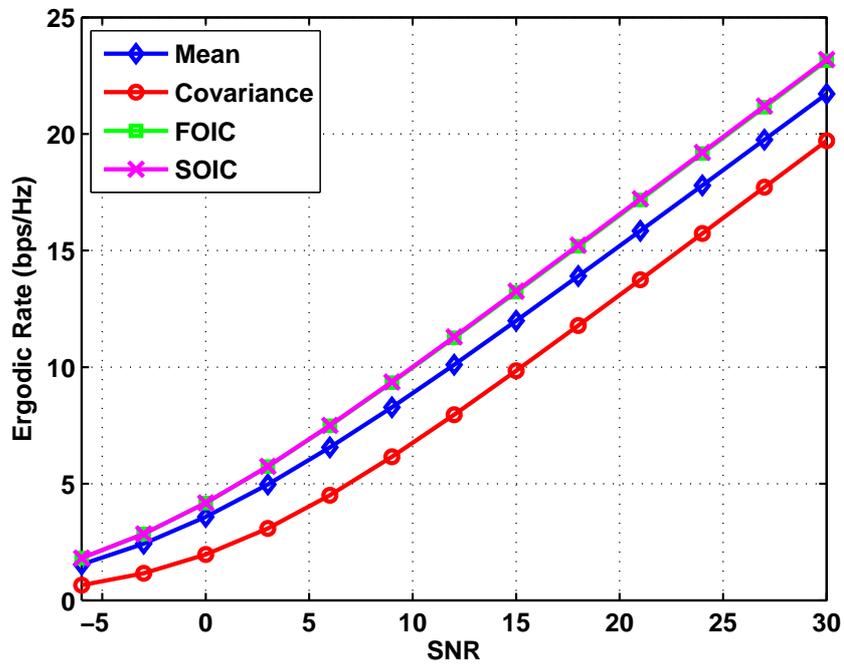
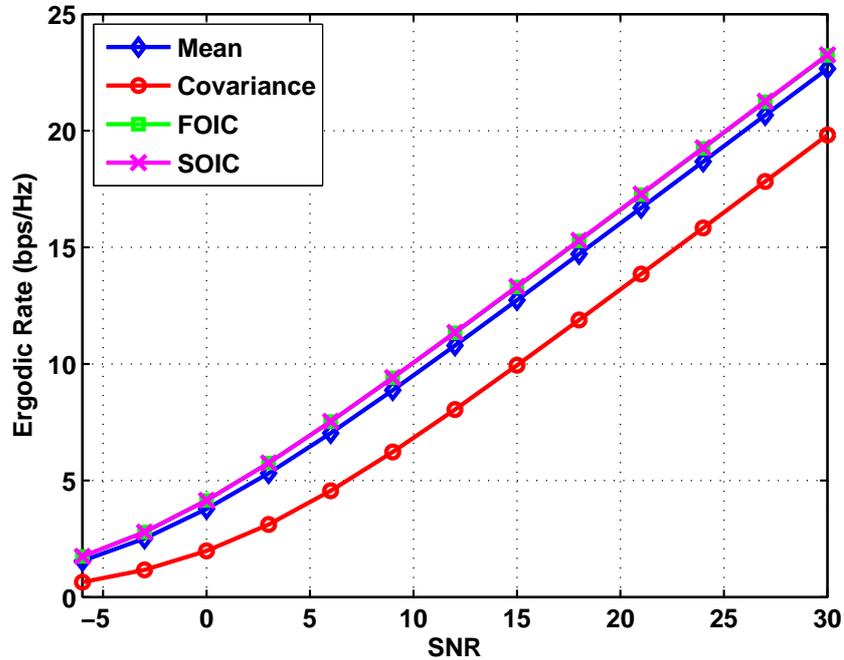
(a) Ricean factor $\kappa = 3$ (b) Ricean factor $\kappa = 10$

Figure 4.11: Ergodic rate curves of the single-user scenario considering the proposed channel model, $N_t = N_r = 2$.

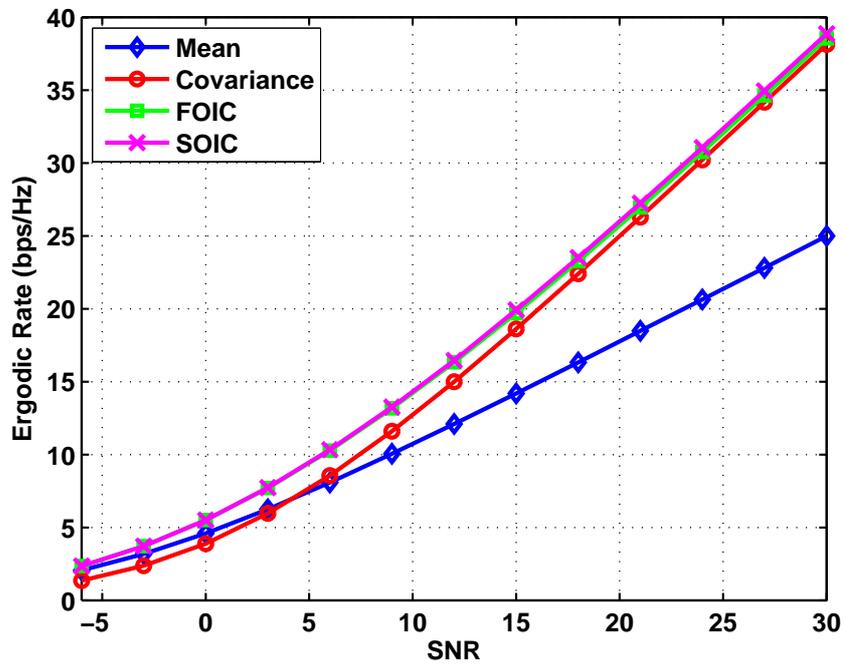
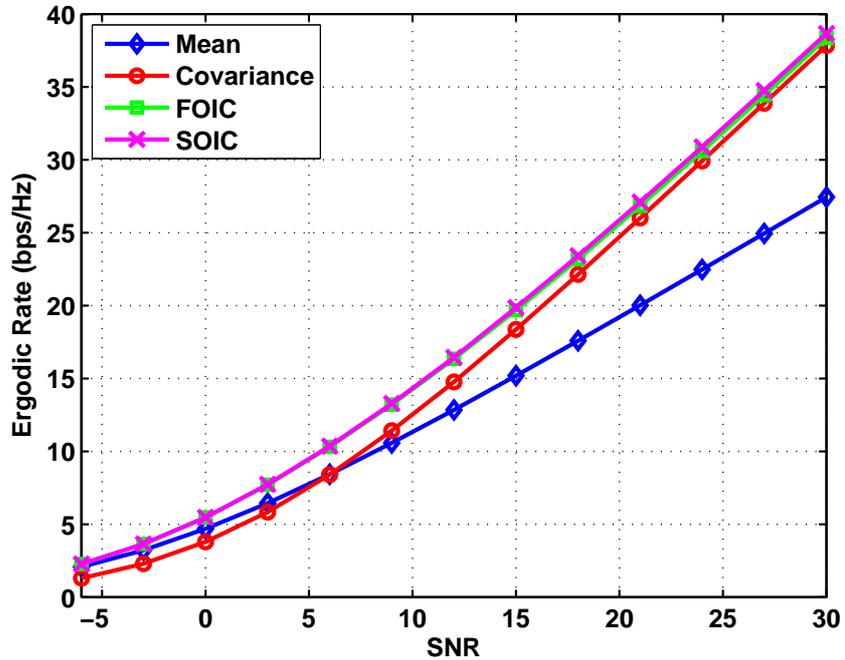
(a) Ricean factor $\kappa = 3$ (b) Ricean factor $\kappa = 10$

Figure 4.12: Ergodic rate curves of the single-user scenario considering the proposed channel model, $N_t = 2$ and $N_r = 4$.

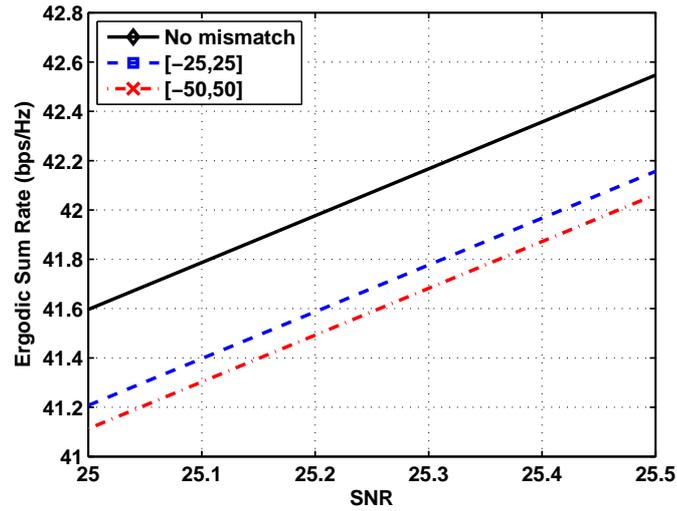
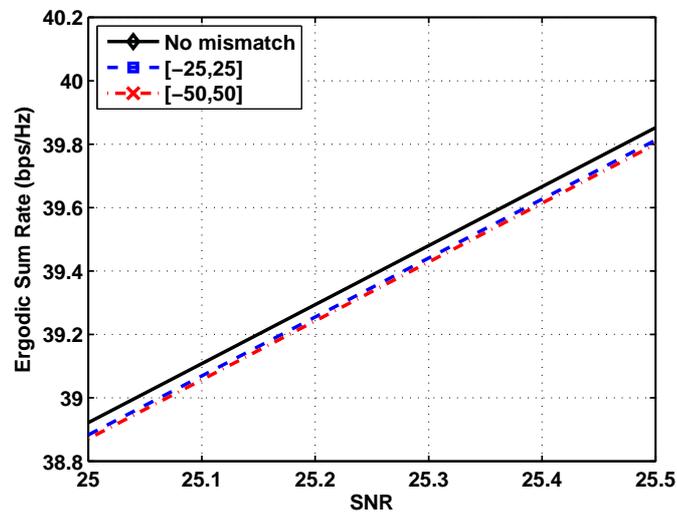
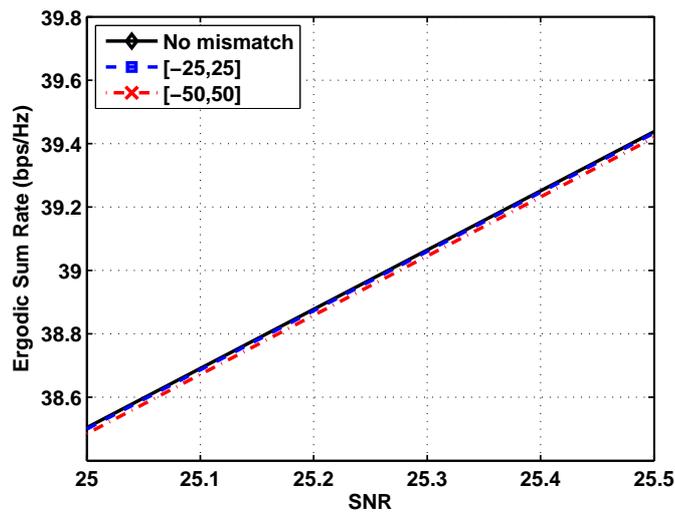
(a) Ricean factor $\kappa = 0$ (b) Ricean factor $\kappa = 3$ (c) Ricean factor $\kappa = 10$

Figure 4.13: Comparison of the ergodic sum rate curves obtained by SOIC approach considering the Kronecker channel model and a phase mismatch between uplink and downlink channels. Scenario $N_t = N_r = 2$.

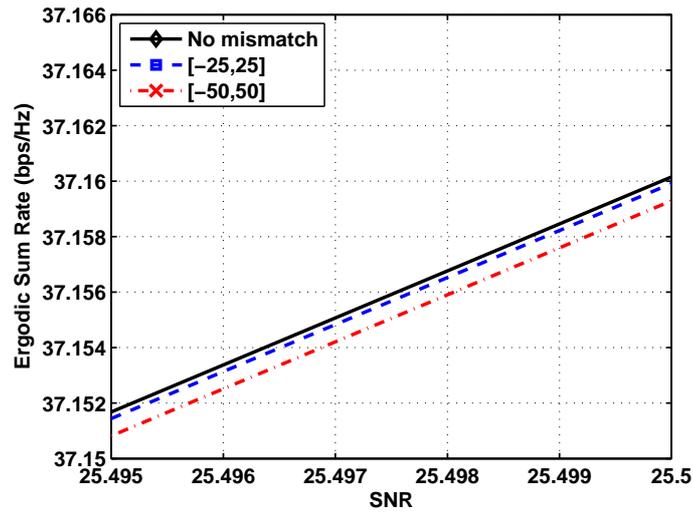
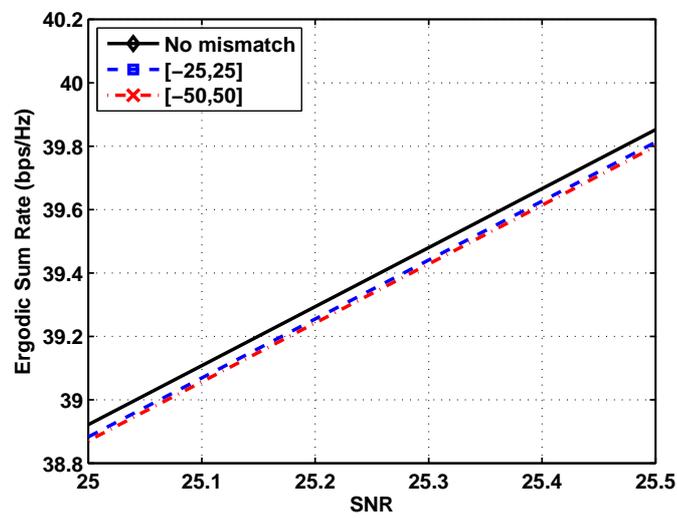
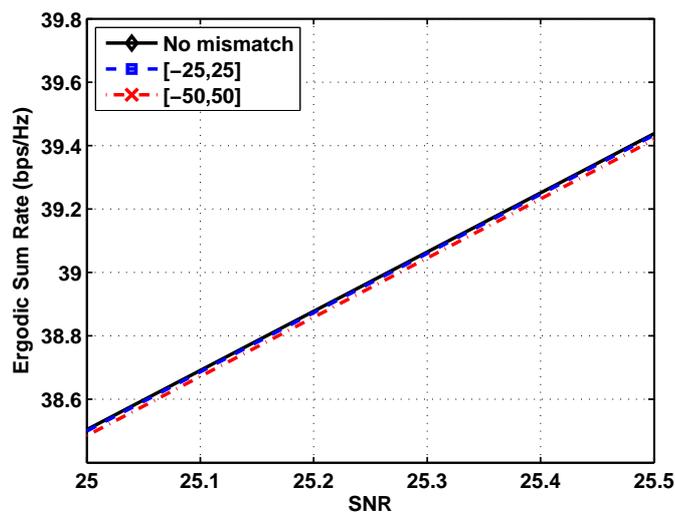
(a) Ricean factor $\kappa = 0$ (b) Ricean factor $\kappa = 3$ (c) Ricean factor $\kappa = 10$

Figure 4.14: Comparison of the ergodic sum rate curves obtained by SOIC approach considering the proposed channel model and a phase mismatch between uplink and downlink channels. Scenario $N_t = N_r = 2$.

Chapter 5

Conclusions and Perspectives

Channel state information at the transmitter (CSIT) can improve the performance of multi-user MIMO (MU-MIMO) wireless systems by means of precoding techniques. Most of them assume the transmitter knows the channel exactly. Accurate CSIT may be obtained reliably when the channel changes slowly but it is more difficult to acquire in situations where the users are highly mobile. Thus, partial CSIT has a great importance in such scenarios and has been widely studied in wireless communication.

We have considered a MU-MIMO coordinated multipoint (CoMP) scenario, in which a number of base-station (BS)s form a coordination group and the BSs from this group transmit to all users. In CoMP systems, besides the serving cell, the cells in the coordination group are normally chosen as the ones that create the highest interference to the user. By this approach, the received signal levels are improved and, at the same time, the inter-channel interference (ICI) level is decreased since part of it has been changed to useful signal. A gain from CoMP is obtained, not only by suppressing the interference sources but also by benefiting from it. Moreover, due to the higher number of transmit antennas involved in the joint transmission processing, better diversity gain can also be obtained. Hence, this scenario is of great interest for future wireless systems, due to, e.g., the macrodiversity and good channel conditioning.

In this thesis, we have proposed a generalization of a statistical CSIT model found in literature for our considered scenario. This partial CSIT model relies on the stochastic processes and estimation theories. Derived from a potentially outdated channel measurement and from the channel statistics (mean and covariance), this proposed CSIT model consists of a channel estimate and its error covariance, which act as the effective channel mean and covariance, respectively. Simulation results have shown that the proposed channel model obtains good results and that the performance gap between the proposed CSIT model and the case considering perfect knowledge of the channel is negligible for low signal-to-noise ratio (SNR) values and moderate for medium to high SNR values.

Moreover, from simulation results of the proposed CSIT, we have noticed that the proposed CSIT model obtains better or worse performance depending of the chosen precoder technique. Hence, our second contribution is to design a precoder that exploits the advantages of the statistical channel model. More specifically, our second contribution consists of deriving a first- and second-order approximation of the ergodic sum rate for a MU-MIMO CoMP system considering that the transmitter has access to statistical channel state information (CSI), while the receiver has access to instantaneous CSI. We use the duality theory [14, 15] to

compute the downlink MU-MIMO CoMP ergodic sum-rate since it states that the achievable sum-rate of the downlink MU-MIMO channel is equal to the achievable sum rate of the uplink MU-MIMO channel. After that, we use convex optimization tools to find covariance matrices of the transmitted signal (known as input covariance matrices) that maximize the first- and second-order approximation of the MU-MIMO CoMP ergodic sum-rate.

The simulation results have shown that the performance gap between the proposed algorithms and the bound-achieving algorithm is small and that the convergence rate of the proposed algorithms is high given that the convergence parameter δ is rightly chosen. Moreover, since in literature we only found techniques that consider a single-user scenario, we transform our multi-user scenario into a single-user one by choosing the user with channel of highest Frobenius norm and compare to these single-user techniques from the literature. Simulation results have shown that our proposed precoder techniques obtain good results compared to other techniques. Moreover, we have shown that the proposed CSIT model performs better than the Kronecker channel model and that this happens since the proposed model takes into account the channel temporal variation and thus it can model the real channel in a more precise way. We have also analyzed the phase mismatch between the uplink and the dual downlink channel model, which is a more realistic model when considering the duality theorem.

The next steps of this thesis are concentrated in:

- ▶ We can see from simulation results in Section 3.2 (Figures 3.4(a), 3.4(b), 3.5(a) and 3.5(b)) that the proposed CSIT model has a loss of performance as a cost to reduce the amount of feedback (to reduce the number of symbols N_S in each block). It is interesting to find a way of quantifying this reduction of feedback information and to make a trade-off between loss of performance and reduction of feedback. We judge this analysis to be important since it quantifies the loss of performance obtained with the proposed CSIT model compared to the economy in the amount of feedback information.
- ▶ As already mentioned in section 4.3, the power normalization used in the proposed algorithms is suboptimal so that only the BS satisfying the minimum value can transmit with full power and any other BS transmits with a power lower than its power constraint. We judge interesting to investigate other power allocation algorithms that can utilize the full power at each BS.
- ▶ In Section 4.4, Figures 4.7(a) and 4.7(b), we have seen that the parameter δ must be chosen not too small in order to have the convergence, but not too high since, in the fast convergence, the obtained ergodic sum rate after the convergence is smaller than the ergodic sum rate obtained in the slow convergence case. As a next step in this topic, we will investigate a way to obtain the parameter δ as a closed form, since in simulations of this thesis this parameter δ was found by exhaustive search for each scenario and each SNR value.
- ▶ Some studies have shown that the uplink and downlink channel phase plots deviate from each other in phase by a few radians when considering the duality theorem. In Subsection 4.4.5 we have analyzed the impact of this phase mismatch between the downlink and uplink channels and considered that the proofs of our algorithm will not change. As a future work, we will accomplish the complete derivation of this proposal considering phase mismatches

Mathematical Manipulations and Proofs

A.1 Proof of Equation (4.7)

In section 4.1 we have shown that the objective function in (4.6) can be written as equation (4.7). We show the mathematical manipulations to obtain this result:

$$\begin{aligned}
 f(\mathbf{Q}_1, \dots, \mathbf{Q}_K) &= \log \left| \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j + \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right| \\
 &= \log \left| \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j + \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right) \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right| \\
 &= \log \left| \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right) \left(\mathbf{I} + \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right) \right|
 \end{aligned} \tag{A.1}$$

If we use the propriety that $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$, we can write (A.1) as:

$$\begin{aligned}
 f(\mathbf{Q}_1, \dots, \mathbf{Q}_K) &= \log \left| \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right) \right| \left| \left(\mathbf{I} + \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right) \right| \\
 &= \log \left| \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right) \right| + \log \left| \left(\mathbf{I} + \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \right) \right| \\
 &= \log \left| \mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right| + \\
 &\quad + \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1/2} \mathbf{H}_k^H \mathbf{Q}_k \mathbf{H}_k \left(\mathbf{I} + \sum_{j \neq k} \mathbf{H}_j^H \mathbf{Q}_j^{(i)} \mathbf{H}_j \right)^{-1/2} \right|
 \end{aligned} \tag{A.2}$$

A.2 Proof of $\mathbb{E} \left\{ \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right\}$

In order to evaluate $\Delta = \mathbb{E} \left\{ \mathbf{R}_{t_k}^{1/2} \mathbf{H}_{w_k}^H \mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \right\}$, we denote the eigen decomposition of \mathbf{R}_{t_k} by $\mathbf{V}_k \Upsilon_k \mathbf{V}_k^H$ and of $\mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2}$ by $\mathbf{U}_k \Lambda_k \mathbf{U}_k^H$. We also verify that for any complex Gaussian matrix \mathbf{H}_{w_k} with zero mean and i.i.d. entries, and for any deterministic matrices \mathbf{U}_k and \mathbf{V}_k , we have

$$\mathbf{H}_{w_k} \stackrel{d}{=} \mathbf{U}_k^H \mathbf{H}_{w_k} \mathbf{V}_k, \quad (\text{A.3})$$

where $\stackrel{d}{=}$ denotes equality in distribution. Using these definitions, we have

$$\begin{aligned} \Delta &= \mathbb{E} \left\{ \mathbf{V}_k \Upsilon_k^{1/2} \mathbf{V}_k^H \mathbf{H}_{w_k}^H \mathbf{U}_k \Lambda_k \mathbf{U}_k^H \mathbf{H}_{w_k} \mathbf{V}_k \Upsilon_k^{1/2} \mathbf{V}_k^H \right\}, \\ &= \mathbf{V}_k \Upsilon_k^{1/2} \mathbb{E} \left\{ \mathbf{H}_{w_k}^H \Lambda_k \mathbf{H}_{w_k} \right\} \Upsilon_k^{1/2} \mathbf{V}_k^H. \end{aligned} \quad (\text{A.4})$$

In order to compute the expectation in (A.4), we use $[\mathbf{H}_{w_k}]_t$ to denote the t -th column of \mathbf{H}_{w_k} and λ_t the t -th diagonal element of Λ_k . Hence, we have

$$\mathbb{E} \left\{ \mathbf{H}_{w_k}^H \Lambda_k \mathbf{H}_{w_k} \right\} = \mathbb{E} \left\{ \sum_{t=1}^{N_r} [\mathbf{H}_{w_k}]_t \lambda_t [\mathbf{H}_{w_k}]_t^H \right\}. \quad (\text{A.5})$$

Substituting (A.5) into (A.4) we have

$$\begin{aligned} \Delta &= \mathbf{V}_k \Upsilon_k^{1/2} \mathbb{E} \left\{ \sum_{t=1}^{N_r} [\mathbf{H}_{w_k}]_t \lambda_t [\mathbf{H}_{w_k}]_t^H \right\} \Upsilon_k^{1/2} \mathbf{V}_k^H, \\ &= \mathbf{V}_k \Upsilon_k^{1/2} \sum_{r=1}^{N_r} \lambda_r \mathbb{E} \left\{ [\mathbf{H}_{w_k}]_r [\mathbf{H}_{w_k}]_r^H \right\} \Upsilon_k^{1/2} \mathbf{V}_k^H, \\ &= \mathbf{V}_k \Upsilon_k^{1/2} \sum_{r=1}^{N_r} \lambda_r \mathbf{I}_{N_b N_t} \Upsilon_k^{1/2} \mathbf{V}_k^H, \\ &= \sum_{t=1}^{N_r} \lambda_t \mathbf{V}_k \Upsilon_k \mathbf{V}_k^H, \\ &= \text{tr} \left(\mathbf{R}_{r_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \right) \mathbf{R}_{t_k}. \end{aligned} \quad (\text{A.6})$$

A.3 Proof of $\mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{B}) \} = 0$ and $\mathbb{E} \{ \text{tr}(\mathbf{B}^H \mathbf{B}^H) \} = 0$

In order to evaluate $\mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{B}) \}$, we denote the singular value decomposition of $\mathbf{R}_{t_k}^{1/2} \mathbf{C}_k^{-1/2} \mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2}$ by $\mathbf{U}_k \Upsilon_k \mathbf{V}_k^H$ and we have

$$\begin{aligned} \mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{B}) \} &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \mathbf{C}_k^{-1/2} \mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{H}_{w_k} \mathbf{R}_{t_k}^{1/2} \mathbf{C}_k^{-1/2} \right) \right\}, \\ &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_{w_k} \mathbf{U}_k \Upsilon_k \mathbf{V}_k^H \mathbf{H}_{w_k} \mathbf{U}_k \mathbf{U}_k^H \mathbf{R}_{t_k}^{1/2} \mathbf{C}_k^{-1/2} \right) \right\}. \end{aligned} \quad (\text{A.7})$$

Using the statement of (A.3) we have then

$$\mathbb{E} \{ \text{tr}(\mathbf{B}\mathbf{B}) \} = \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{r_k}^{1/2} \mathbf{V}_k \mathbb{E} \{ \mathbf{H}_{w_k} \Upsilon_k \mathbf{H}_{w_k} \} \mathbf{U}_k^H \mathbf{R}_{t_k}^{1/2} \mathbf{C}_k^{-1/2} \right). \quad (\text{A.8})$$

It is not difficult to note that the expectation $\mathbb{E} \{ \mathbf{H}_{w_k} \Upsilon_k \mathbf{H}_{w_k} \}$ is given by $\mathbb{E} \left\{ \sum_{t=1}^{\min(N_b N_t, N_r)} [\mathbf{H}_{w_k}]_t v_t [\mathbf{H}_{w_k}]_t^H \right\}$, where v_t is the t -th diagonal element of Υ_k . Hence, (A.8)

is written as

$$\mathbb{E}\{\text{tr}(\mathbf{BB})\} = \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{V}_k \sum_{t=1}^{\min(N_b, N_t, N_r)} \nu_t \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\} \mathbf{U}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right). \quad (\text{A.9})$$

In order to evaluate the expectation $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\}$, we firstly denote the (p, q) -th element of $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\}$ as being $\left[\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\} \right]_{p,q} = \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{p,t} [\mathbf{H}_{\mathbf{w}_k}^H]_{t,q}^H \right\}$. For the cases $p \neq q \neq t$, $p = q \neq t$, $p \neq q = t$ and $p = t \neq q$, we have $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{p,t} [\mathbf{H}_{\mathbf{w}_k}^H]_{t,q}^H \right\} = 0$ since $[\mathbf{H}_{\mathbf{w}_k}]_{p,t}$ and $[\mathbf{H}_{\mathbf{w}_k}]_{t,q}$ are i.i.d. random variables with zero mean. For the case $p = q = t$, we have $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{p,p} [\mathbf{H}_{\mathbf{w}_k}^H]_{p,p}^H \right\} = \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{p,p}^2 \right\}$ and since the real and imaginary parts of $[\mathbf{H}_{\mathbf{w}_k}]_{p,p}$ are i.i.d random variables with zero mean and have the propriety $\mathbb{E} \{ \text{Re} \{ [\mathbf{H}_{\mathbf{w}_k}]_{p,p} \} \} = \mathbb{E} \{ \text{Im} \{ [\mathbf{H}_{\mathbf{w}_k}]_{p,p} \} \} = 0$, so we have $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{p,p}^2 \right\} = 0$. Therefore, $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{p,t} [\mathbf{H}_{\mathbf{w}_k}^H]_{t,q}^H \right\} = 0$, for any p, q and t .

Using these results in (A.9), we can conclude that $\mathbb{E}\{\text{tr}(\mathbf{BB})\} = 0$.

In order to evaluate $\mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{B}^H)\}$ we denote the singular value decomposition of $\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}$ by $\mathbf{U}_k \mathbf{\Upsilon}_k \mathbf{V}_k^H$ and we have

$$\begin{aligned} \mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{B}^H)\} &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2} \right) \right\}, \\ &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_{\mathbf{w}_k}^H \mathbf{U}_k \mathbf{\Upsilon}_k \mathbf{V}_k^H \mathbf{H}_{\mathbf{w}_k}^H \mathbf{U}_k \mathbf{U}_k^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2} \right) \right\}, \\ &= \mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E} \left\{ \text{tr} \left(\mathbf{H}_{\mathbf{w}_k}^H \mathbf{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H \right) \right\} \mathbf{U}_k^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2}. \end{aligned} \quad (\text{A.10})$$

The expectation $\mathbb{E} \left\{ \text{tr} \left(\mathbf{H}_{\mathbf{w}_k}^H \mathbf{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H \right) \right\}$ can be evaluated in a similar way of $\mathbb{E}\{\text{tr}(\mathbf{H}_{\mathbf{w}_k} \mathbf{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k})\}$, which we obtain that $\mathbb{E} \left\{ \text{tr} \left(\mathbf{H}_{\mathbf{w}_k}^H \mathbf{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H \right) \right\} = 0$. Hence, $\mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{B}^H)\} = 0$.

A.4 Proof of $\mathbb{E}\{\text{tr}(\mathbf{BC})\} = 0$ and $\mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{C})\} = 0$

In order to evaluate $\mathbb{E}\{\text{tr}(\mathbf{BC})\}$, we first denote the eigen decomposition of $\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}$ by $\mathbf{V}_k \mathbf{\Upsilon}_k \mathbf{V}_k^H$ and of $\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2}$ by $\mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H$. Then we have

$$\begin{aligned} \mathbb{E}\{\text{tr}(\mathbf{BC})\} &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right) \right\}, \\ &= \mathbb{E} \left\{ \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{U}_k \mathbf{U}_k^H \mathbf{H}_{\mathbf{w}_k} \mathbf{V}_k \mathbf{\Upsilon}_k \mathbf{V}_k^H \mathbf{H}_{\mathbf{w}_k}^H \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H \mathbf{H}_{\mathbf{w}_k} \mathbf{V}_k \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right) \right\}, \\ &= \text{tr} \left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{U}_k \mathbb{E} \left\{ \mathbf{H}_{\mathbf{w}_k} \mathbf{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H \mathbf{\Lambda}_k \mathbf{H}_{\mathbf{w}_k} \right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right). \end{aligned} \quad (\text{A.11})$$

The expectation $\mathbb{E}\left\{\mathbf{H}_{\mathbf{w}k}\boldsymbol{\Upsilon}_k\mathbf{H}_{\mathbf{w}k}^H\boldsymbol{\Lambda}_k\mathbf{H}_{\mathbf{w}k}\right\}$ can be given by

$$\begin{aligned}\mathbb{E}\left\{\mathbf{H}_{\mathbf{w}k}\boldsymbol{\Upsilon}_k\mathbf{H}_{\mathbf{w}k}^H\boldsymbol{\Lambda}_k\mathbf{H}_{\mathbf{w}k}\right\} &= \mathbb{E}\left\{\sum_{t=1}^{N_bN_t} [\mathbf{H}_{\mathbf{w}k}]_t v_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k \mathbf{H}_{\mathbf{w}k}\right\}, \\ &= \sum_{t=1}^{N_bN_t} v_t \mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k \mathbf{H}_{\mathbf{w}k}\right\}, \\ &= \sum_{t=1}^{N_bN_t} v_t \mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}k}]_t \left([\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k \left[[\mathbf{H}_{\mathbf{w}k}]_1, \dots, [\mathbf{H}_{\mathbf{w}k}]_q, \dots, [\mathbf{H}_{\mathbf{w}k}]_{N_bN_t}\right]\right)\right\}, \\ &= \sum_{t=1}^{N_bN_t} v_t \mathbb{E}\left\{\left[[\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k [\mathbf{H}_{\mathbf{w}k}]_1, \dots, [\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k [\mathbf{H}_{\mathbf{w}k}]_q, \dots, \right.\right. \\ &\quad \left.\left. [\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k [\mathbf{H}_{\mathbf{w}k}]_{N_bN_t}\right]\right\}.\end{aligned}\tag{A.12}$$

We can note that each term $[\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k [\mathbf{H}_{\mathbf{w}k}]_q$ can be written as

$$[\mathbf{H}_{\mathbf{w}k}]_t \sum_{r=1}^{N_r} \lambda_r [\mathbf{H}_{\mathbf{w}k}]_{r,t}^* [\mathbf{H}_{\mathbf{w}k}]_{r,q} = \sum_{r=1}^{N_r} \lambda_r [\mathbf{H}_{\mathbf{w}k}]_{r,t}^* [\mathbf{H}_{\mathbf{w}k}]_{r,q} [\mathbf{H}_{\mathbf{w}k}]_t \tag{A.13}$$

for all q .

Considering (A.13) and analyzing the case when $q = t$, we have that equation (A.12) can be rewritten as

$$\sum_{t=1}^{N_bN_t} v_t \left(\sum_{r=1}^{N_r} \lambda_r \mathbb{E}\left\{[|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_t\right\} \right) = \sum_{t=1}^{N_bN_t} v_t \left(\sum_{r=1}^{N_r} \lambda_r \mathbb{E}\left\{ \begin{array}{c} [|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{1,t} \\ \vdots \\ [|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{p,t} \\ \vdots \\ [|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{N_r,t} \end{array} \right\} \right). \tag{A.14}$$

For $r \neq p$ we have to evaluate $\mathbb{E}\left\{[|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{p,t}\right\}$. It is not difficult to show that $\mathbb{E}\left\{[|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{p,t}\right\} = 0$, since $\mathbb{E}\left\{[|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{p,t}\right\} = \mathbb{E}_{[\mathbf{H}_{\mathbf{w}k}]_{r,t}} \left\{ \mathbb{E}_{[\mathbf{H}_{\mathbf{w}k}]_{p,t}} \left\{ [|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{p,t} | [\mathbf{H}_{\mathbf{w}k}]_{r,t} \right\} \right\}$. And when $r = p$, we have that $\mathbb{E}\left\{[|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{r,t}\right\}$ is proportional to the 3-rd order distribution of $[\mathbf{H}_{\mathbf{w}k}]_{r,t}$ and, since $[\mathbf{H}_{\mathbf{w}k}]_{r,t}$ is a Gaussian random variable, its moments of odd order are zero. Therefore, $\mathbb{E}\left\{[|\mathbf{H}_{\mathbf{w}k}]_{r,t}|^2 [\mathbf{H}_{\mathbf{w}k}]_{r,t}\right\} = 0$.

Analyzing now the case $q \neq t$ in (A.12), we have:

$$\begin{aligned}\sum_{t=1}^{N_bN_t} v_t \left(\sum_{r=1}^{N_r} \lambda_r \mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_{r,t}^* [\mathbf{H}_{\mathbf{w}k}]_{r,q}\right\} \right) &= \sum_{t=1}^{N_bN_t} v_t \left(\mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k [\mathbf{H}_{\mathbf{w}k}]_q\right\} \right) \\ &= \sum_{t=1}^{N_bN_t} v_t \left(\mathbb{E}_{[\mathbf{H}_{\mathbf{w}k}]_t} \left\{ [\mathbf{H}_{\mathbf{w}k}]_t [\mathbf{H}_{\mathbf{w}k}]_t^H \boldsymbol{\Lambda}_k \mathbb{E}_{[\mathbf{H}_{\mathbf{w}k}]_q} \left\{ [\mathbf{H}_{\mathbf{w}k}]_q | [\mathbf{H}_{\mathbf{w}k}]_t \right\} \right\} \right).\end{aligned}\tag{A.15}$$

The expectation $\mathbb{E}_{[\mathbf{H}_{\mathbf{w}k}]_q} \left\{ [\mathbf{H}_{\mathbf{w}k}]_q | [\mathbf{H}_{\mathbf{w}k}]_t \right\} = 0$. Therefore, we can conclude that the expectation $\mathbb{E}\left\{\mathbf{H}_{\mathbf{w}k}\boldsymbol{\Upsilon}_k\mathbf{H}_{\mathbf{w}k}^H\boldsymbol{\Lambda}_k\mathbf{H}_{\mathbf{w}k}\right\}$ is equal to zero and then $\mathbb{E}\{\text{tr}(\mathbf{B}\mathbf{C})\} = 0$.

In order to prove $\mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{C})\} = 0$, we use the same eigen decomposition used to prove

$\mathbb{E}\{\text{tr}(\mathbf{BC})\} = 0$. This leads us to the following result:

$$\begin{aligned}\mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{C})\} &= \mathbb{E}\left\{\text{tr}\left(\mathbf{C}_k^{-1/2}\mathbf{R}_{t_k}^{1/2}\mathbf{H}_{w_k}^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\mathbf{H}_{w_k}\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}\mathbf{H}_{w_k}^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \mathbb{E}\left\{\text{tr}\left(\mathbf{C}_k^{-1/2}\mathbf{R}_{t_k}^{1/2}\mathbf{V}_k\mathbf{V}_k^H\mathbf{H}_{w_k}^H\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^H\mathbf{H}_{w_k}\mathbf{V}_k\mathbf{\Upsilon}_k\mathbf{V}_k^H\mathbf{H}_{w_k}^H\mathbf{U}_k\mathbf{U}_k^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \text{tr}\left(\mathbf{C}_k^{-1/2}\mathbf{R}_{t_k}^{1/2}\mathbf{V}_k\mathbb{E}\left\{\mathbf{H}_{w_k}^H\mathbf{\Lambda}_k\mathbf{H}_{w_k}\mathbf{\Upsilon}_k\mathbf{H}_{w_k}^H\right\}\mathbf{U}_k^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1/2}\right).\end{aligned}\quad (\text{A.16})$$

We can show in a similar way than $\mathbb{E}\left\{\mathbf{H}_{w_k}\mathbf{\Upsilon}_k\mathbf{H}_{w_k}^H\mathbf{\Lambda}_k\mathbf{H}_{w_k}\right\}$ that $\mathbb{E}\left\{\mathbf{H}_{w_k}^H\mathbf{\Lambda}_k\mathbf{H}_{w_k}\mathbf{\Upsilon}_k\mathbf{H}_{w_k}^H\right\} = 0$. Hence, $\mathbb{E}\{\text{tr}(\mathbf{B}^H\mathbf{C})\} = 0$.

A.5 Evaluation of $\mathbb{E}\{\text{tr}(\mathbf{AC})\}$ and $\mathbb{E}\{\text{tr}(\mathbf{BB}^H)\}$

In order to evaluate $\mathbb{E}\{\text{Tr}(\mathbf{AC})\}$ we denote the eigen decomposition of $\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}$ by $\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^H$. Hence, we have

$$\begin{aligned}\mathbb{E}\{\text{Tr}(\mathbf{AC})\} &= \mathbb{E}\left\{\text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}\mathbf{H}_{w_k}^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\mathbf{H}_{w_k}\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \mathbb{E}\left\{\text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}\mathbf{V}_k\mathbf{V}_k^H\mathbf{H}_{w_k}^H\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^H\mathbf{H}_{w_k}\mathbf{V}_k\mathbf{V}_k^H\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}\mathbf{V}_k\mathbb{E}\left\{\mathbf{H}_{w_k}^H\mathbf{\Lambda}_k\mathbf{H}_{w_k}\right\}\mathbf{V}_k^H\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1/2}\right).\end{aligned}\quad (\text{A.17})$$

The expectation $\mathbb{E}\left\{\mathbf{H}_{w_k}^H\mathbf{\Lambda}_k\mathbf{H}_{w_k}\right\}$ can be written as

$$\begin{aligned}\mathbb{E}\left\{\mathbf{H}_{w_k}^H\mathbf{\Lambda}_k\mathbf{H}_{w_k}\right\} &= \mathbb{E}\left\{\sum_{r=1}^{N_r}\lambda_r[\mathbf{H}_{w_k}]_r[\mathbf{H}_{w_k}]_r^H\right\}, \\ &= \sum_{t=1}^{N_r}\lambda_t\mathbb{E}\left\{[\mathbf{H}_{w_k}]_t[\mathbf{H}_{w_k}]_t^H\right\}, \\ &= \sum_{r=1}^{N_r}\lambda_r\mathbf{I}_{N_b N_t}, \\ &= \text{tr}\left(\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\right)\mathbf{I}_{N_b N_t}.\end{aligned}\quad (\text{A.18})$$

Substituting the result of (A.18) into (A.17), we have

$$\begin{aligned}\mathbb{E}\{\text{Tr}(\mathbf{AC})\} &= \text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}\mathbf{V}_k\text{tr}\left(\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\right)\mathbf{I}_{N_b N_t}\mathbf{V}_k^H\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1/2}\right), \\ &= \text{tr}\left(\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\right)\text{Tr}\left(\bar{\mathbf{H}}_k^H\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1}\mathbf{R}_{t_k}\mathbf{C}_k^{-1}\right).\end{aligned}\quad (\text{A.19})$$

Then, $\mathbb{E}\{\text{tr}(\mathbf{AC})\} = \text{tr}\left(\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\right)\text{tr}\left(\bar{\mathbf{H}}_k^H\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1}\mathbf{R}_{t_k}\mathbf{C}_k^{-1}\right)$. In order to evaluate $\mathbb{E}\{\text{Tr}(\mathbf{BB}^H)\}$ we use the eigen decomposition of $\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}$ by $\mathbf{V}_k\mathbf{\Upsilon}_k\mathbf{V}_k^H$. Therefore, we have

$$\begin{aligned}\mathbb{E}\{\text{Tr}(\mathbf{BB}^H)\} &= \mathbb{E}\left\{\text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\mathbf{H}_{w_k}\mathbf{R}_{t_k}^{1/2}\mathbf{C}_k^{-1}\mathbf{R}_{t_k}^{1/2}\mathbf{H}_{w_k}^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \mathbb{E}\left\{\text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\mathbf{U}_k\mathbf{U}_k^H\mathbf{H}_{w_k}\mathbf{V}_k\mathbf{\Upsilon}_k\mathbf{V}_k^H\mathbf{H}_{w_k}^H\mathbf{U}_k\mathbf{U}_k^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \text{Tr}\left(\mathbf{C}_k^{-1/2}\bar{\mathbf{H}}_k^H\mathbf{Q}_k\mathbf{R}_{r_k}^{1/2}\mathbf{U}_k\mathbb{E}\left\{\mathbf{H}_{w_k}\mathbf{\Upsilon}_k\mathbf{H}_{w_k}^H\right\}\mathbf{U}_k^H\mathbf{R}_{r_k}^{1/2}\mathbf{Q}_k\bar{\mathbf{H}}_k\mathbf{C}_k^{-1/2}\right).\end{aligned}\quad (\text{A.20})$$

The expectation $\mathbb{E}\{\mathbf{H}_{\mathbf{w}_k} \boldsymbol{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H\}$ can be written as

$$\begin{aligned} \mathbb{E}\{\mathbf{H}_{\mathbf{w}_k} \boldsymbol{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H\} &= \mathbb{E}\left\{\sum_{t=1}^{N_b N_t} v_t [\mathbf{H}_{\mathbf{w}_k}]_t [\mathbf{H}_{\mathbf{w}_k}]_t^H\right\}, \\ &= \sum_{t=1}^{N_b N_t} v_t \mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}_k}]_t [\mathbf{H}_{\mathbf{w}_k}]_t^H\right\}, \\ &= \sum_{t=1}^{N_b N_t} v_t \mathbf{I}_{N_r}, \\ &= \text{tr}\left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}\right) \mathbf{I}_{N_r}. \end{aligned} \quad (\text{A.21})$$

Substituting the result of (A.21) into (A.20), we have

$$\begin{aligned} \mathbb{E}\{\text{Tr}(\mathbf{B}\mathbf{B}^H)\} &= \text{Tr}\left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{U}_k \text{tr}\left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}\right) \mathbf{I}_{N_r} \mathbf{U}_k^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2}\right), \\ &= \text{tr}\left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}\right) \text{Tr}\left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2}\right). \end{aligned} \quad (\text{A.22})$$

Therefore, $\mathbb{E}\{\text{Tr}(\mathbf{B}\mathbf{B}^H)\} = \text{tr}\left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}\right) \text{Tr}\left(\mathbf{C}_k^{-1/2} \bar{\mathbf{H}}_k^H \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k} \mathbf{Q}_k \bar{\mathbf{H}}_k \mathbf{C}_k^{-1/2}\right)$.

A.6 Evaluation of $\mathbb{E}\{\text{tr}(\mathbf{C}^2)\}$

In order to evaluate $\mathbb{E}\{\text{Tr}(\mathbf{C}^2)\}$ we denote the eigen decomposition of $\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2}$ by $\mathbf{U}_k \boldsymbol{\Lambda}_k \mathbf{U}_k^H$ and of $\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2}$ by $\mathbf{V}_k \boldsymbol{\Upsilon}_k \mathbf{V}_k^H$. Hence, we have

$$\begin{aligned} \mathbb{E}\{\text{tr}(\mathbf{C}^2)\} &= \mathbb{E}\left\{\text{tr}\left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k}^H \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{H}_{\mathbf{w}_k} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \mathbb{E}\left\{\text{tr}\left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbf{V}_k^H \mathbf{H}_{\mathbf{w}_k}^H \mathbf{U}_k \boldsymbol{\Lambda}_k \mathbf{U}_k^H \mathbf{H}_{\mathbf{w}_k} \mathbf{V}_k \boldsymbol{\Upsilon}_k \mathbf{V}_k^H \mathbf{H}_{\mathbf{w}_k}^H \mathbf{U}_k \boldsymbol{\Lambda}_k \mathbf{U}_k^H \mathbf{H}_{\mathbf{w}_k} \mathbf{V}_k \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}\right)\right\}, \\ &= \text{tr}\left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E}\left\{\mathbf{H}_{\mathbf{w}_k}^H \boldsymbol{\Lambda}_k \mathbf{H}_{\mathbf{w}_k} \boldsymbol{\Upsilon}_k \mathbf{H}_{\mathbf{w}_k}^H \boldsymbol{\Lambda}_k \mathbf{H}_{\mathbf{w}_k}\right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}\right), \\ &= \text{tr}\left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E}\left\{\sum_{t=1}^{N_r} \lambda_t [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k \sum_{p=1}^{N_r} \lambda_p [\mathbf{H}_{\mathbf{w}_k}^H]_p [\mathbf{H}_{\mathbf{w}_k}^H]_p^H\right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}\right), \\ &= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{tr}\left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H\right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}\right) \\ &\quad + \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{tr}\left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E}\left\{[\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_p [\mathbf{H}_{\mathbf{w}_k}^H]_p^H\right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2}\right). \end{aligned} \quad (\text{A.23})$$

Firstly, evaluating the first term of (A.23), we have

$$\begin{aligned}
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right), \\
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \right), \\
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\} \mathbf{V}_k^H \mathbf{V}_k \boldsymbol{\Upsilon}_k \mathbf{V}_k^H \mathbf{V}_k \right), \\
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\mathbb{E} \left\{ \left([\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_t \right)^2 \right\} \right), \\
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\mathbb{E} \left\{ \sum_{z=1}^{N_b N_t} v_z [\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* \sum_{y=1}^{N_b N_t} v_y [\mathbf{H}_{\mathbf{w}_k}]_{t,y} [\mathbf{H}_{\mathbf{w}_k}]_{t,y}^* \right\} \right), \\
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\sum_{z=1}^{N_b N_t} \sum_{y=1}^{N_b N_t} v_y v_z \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* [\mathbf{H}_{\mathbf{w}_k}]_{t,y} [\mathbf{H}_{\mathbf{w}_k}]_{t,y}^* \right\} \right), \\
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{Tr} \left(\sum_{z=y=1}^{N_b N_t} v_z^2 \mathbb{E} \left\{ \left([\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* \right)^2 \right\} + \sum_{z \neq y, z=1, y=1}^{N_b N_t} v_y v_z \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* [\mathbf{H}_{\mathbf{w}_k}]_{t,y} [\mathbf{H}_{\mathbf{w}_k}]_{t,y}^* \right\} \right). \tag{A.24}
\end{aligned}$$

We can note that the expectation $\mathbb{E} \left\{ \left([\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* \right)^2 \right\}$ of (A.24) is the second-order moment of random variable $|[\mathbf{H}_{\mathbf{w}_k}]_{t,z}|$ and thus we can find that $\mathbb{E} \left\{ \left([\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* \right)^2 \right\} = 2$. On the other hand, the expectation $\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}]_{t,z} [\mathbf{H}_{\mathbf{w}_k}]_{t,z}^* [\mathbf{H}_{\mathbf{w}_k}]_{t,y} [\mathbf{H}_{\mathbf{w}_k}]_{t,y}^* \right\}$ can be found to be equal to 1 since $[\mathbf{H}_{\mathbf{w}_k}]_{t,z}$ and $[\mathbf{H}_{\mathbf{w}_k}]_{t,y}$ are i.i.d random variables. Therefore, (A.24) is written as

$$= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{tr} \left(\sum_{z=y=1}^{N_b N_t} 2v_z^2 + \sum_{z \neq y, z=1, y=1}^{N_b N_t} v_y v_z \right). \tag{A.25}$$

It is not difficult to show that (A.25) is

$$\begin{aligned}
&= \sum_{t=p=1}^{N_r} \lambda_t^2 \text{tr} \left(\sum_{z=y=1}^{N_b N_t} v_z^2 + \left(\sum_{z=1}^{N_b N_t} v_z \right)^2 \right), \\
&= \sum_{t=1}^{N_r} \lambda_t^2 \text{tr} \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 + \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \right)^2 \right), \\
&= \sum_{t=1}^{N_r} \lambda_t^2 \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 + \sum_{t=1}^{N_r} \lambda_t^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \right)^2. \tag{A.26}
\end{aligned}$$

Now, evaluating the second term of (A.23), we have

$$\begin{aligned}
&= \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{Tr} \left(\mathbf{C}_k^{-1/2} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_p [\mathbf{H}_{\mathbf{w}_k}^H]_p^H \right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1/2} \right), \\
&= \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{Tr} \left(\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_p [\mathbf{H}_{\mathbf{w}_k}^H]_p^H \right\} \mathbf{V}_k^H \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{V}_k \right), \\
&= \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{Tr} \left(\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \boldsymbol{\Upsilon}_k [\mathbf{H}_{\mathbf{w}_k}^H]_p [\mathbf{H}_{\mathbf{w}_k}^H]_p^H \right\} \boldsymbol{\Upsilon}_k \right). \tag{A.27}
\end{aligned}$$

Since $[\mathbf{H}_{\mathbf{w}_k}^H]_t[\mathbf{H}_{\mathbf{w}_k}^H]_t^H$ and $[\mathbf{H}_{\mathbf{w}_k}^H]_p[\mathbf{H}_{\mathbf{w}_k}^H]_p^H$ are independent we can write (A.27) as

$$\begin{aligned}
&= \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{Tr} \left(\mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_t [\mathbf{H}_{\mathbf{w}_k}^H]_t^H \right\} \mathbf{\Upsilon}_k \mathbb{E} \left\{ [\mathbf{H}_{\mathbf{w}_k}^H]_p [\mathbf{H}_{\mathbf{w}_k}^H]_p^H \right\} \mathbf{\Upsilon}_k \right), \\
&= \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{tr} \left(\mathbf{I}_{N_b N_t} \mathbf{\Upsilon}_k \mathbf{I}_{N_b N_t} \mathbf{\Upsilon}_k \right), \\
&= \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2. \tag{A.28}
\end{aligned}$$

Therefore, (A.23) can be written as

$$\begin{aligned}
\mathbb{E} \{ \text{Tr}(\mathbf{C}^2) \} &= \sum_{t=1}^{N_r} \lambda_t^2 \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 + \sum_{t=1}^{N_r} \lambda_t^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \right)^2 + \\
&+ \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2. \tag{A.29}
\end{aligned}$$

It can be noted that

$$\begin{aligned}
&\sum_{t=1}^{N_r} \lambda_t^2 \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 + \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \\
&= \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\sum_{t=1}^{N_r} \lambda_t^2 + \sum_{t \neq p; t=1; p=1}^{N_r} \lambda_t \lambda_p \right), \\
&= \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\sum_{t=1}^{N_r} \lambda_t \right)^2, \\
&= \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \right)^2. \tag{A.30}
\end{aligned}$$

After some mathematical manipulations, (A.29) can be written as

$$\begin{aligned}
\mathbb{E} \{ \text{Tr}(\mathbf{C}^2) \} &= \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \right)^2 + \sum_{t=1}^{N_r} \lambda_t^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \right)^2, \\
&= \text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right) \right)^2 + \text{tr} \left(\mathbf{R}_{\mathbf{r}_k}^{1/2} \mathbf{Q}_k \mathbf{R}_{\mathbf{r}_k}^{1/2} \right)^2 \left(\text{tr} \left(\mathbf{R}_{\mathbf{t}_k}^{1/2} \mathbf{C}_k^{-1} \mathbf{R}_{\mathbf{t}_k}^{1/2} \right) \right)^2. \tag{A.31}
\end{aligned}$$

Bibliography

- [1] G. J. Foschini, "Layered Space-time Architecture for Wireless Communication in a Fading Environment When using Multi-element Antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996. [Online]. Available: <http://dx.doi.org/10.1002/bltj.2015>
- [2] I. E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Trans. on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
- [3] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.
- [4] 3GPP, "Further Advancements for E-UTRA Physical Layer Aspects," 3rd Generation Partnership Project, Tech. Rep. TR 36.814 V1.5.0 - Release 9, 2009.
- [5] G. Dietl and G. Bauch, "Linear Precoding in the Downlink of Limited Feedback Multiuser MIMO," *IEEE GLOBECOM'07*, pp. 4359–4364, Nov. 2007.
- [6] V. Stankovic and M. Haardt, "Generalized Design of Multi-User MIMO Precoding Matrices," *IEEE Trans. on Wireless Comm.*, vol. 7, no. 3, pp. 953–961, Mar. 2008.
- [7] P. Ding, D. Love, and M. Zoltowski, "Multiple Antenna Broadcast Channels with Shape Feedback and Limited Feedback," *IEEE Trans. on Signal Processing*, vol. 55, no. 7, pp. 3417–3428, Jul. 2007.
- [8] N. Jindal and A. Goldsmith, "Dirty-paper Coding Versus TDMA for MIMO Broadcast Channels," *IEEE Trans. on Inf. Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.
- [9] F. Shu, L. Lihua, C. Qimei, and Z. Ping, "Non-Unitary Codebook Based Precoding Scheme for Multi-User MIMO with Limited Feedback," in *IEEE Wireless Comm. and Net. Conference, 2008. WCNC 2008.*, Mar. 2008, pp. 678 –682.
- [10] F. Shu, W. Gang, X. Yue, and L. Shao-qian, "Multi-User MIMO Linear Precoding with Grassmannian Codebook," in *WRI Int. Conf. on Comm. and Mobile Computing, 2009. CMC '09.*, vol. 1, 6-8 2009, pp. 250 –255.
- [11] Y. Fu, W. Krzymien, and C. Tellambura, "Precoding for Multiuser Orthogonal Space-Time Block-Coded OFDM: Mean or Covariance Feedback?" in *IEEE Int. Conf. on Communications, 2007. ICC '07.*, 24-28 2007, pp. 4269 –4274.
- [12] G. Dartmann, X. Gong, and G. Ascheid, "Jointly Optimized Transmit Beamforming and Temporal User Scheduling in Multiuser Multicell Scenarios Based on Correlation Knowledge," in *Sarnoff Symposium, 2010 IEEE*, 12-14 2010, pp. 1–5.

- [13] M. Vu and A. Paulraj, "A Robust Transmit CSI Framework with Applications in MIMO Wireless Precoding," *Conference Record of the 39th ACSSC'05.*, pp. 623–627, Nov. 2005.
- [14] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, "Sum Power Iterative Water-filling for Multi-antenna Gaussian Broadcast Channels," *IEEE Trans. on Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.
- [15] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, Achievable Rates, and Sum-rate Capacity of Gaussian MIMO Broadcast Channels," *IEEE Trans. on Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
- [16] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative Water-filling for Gaussian Vector Multiple-access Channels," *IEEE Trans. on Information Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.
- [17] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.
- [18] K. Feher, *Wireless Digital Communications: Modulation & Spread Spectrum Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.
- [19] Y. Okumura, E. Ohmuri, T. Kawano, and K. Fukuda, "Field Strength and its Variability in VHF and UHF Land Mobile Radio Service," in *Review of the Electrical Communication Laboratory*, vol. 16, Sep. 1968, pp. 825–873.
- [20] M. Hata and T. Nagatsu, "Mobile Location using Signal Strength Measurements in a Cellular System," *Vehicular Technology, IEEE Transactions on*, vol. 29, no. 2, pp. 245 – 252, May 1980.
- [21] E. C. in the Field of Scientific and E.-C. . Technical Research, "Field Strength and its Variability in VHF and UHF Land Mobile Radio Service," in *Urban Transmission Loss Models for Mobile Radio in the 900 and 1800 MHz Bands*, vol. COST 231 TD (91) 73 Rev 2, Sep. 1991.
- [22] W. Jakes, *Microwave Mobile Communications*, 2nd ed. Piscataway, NJ: Wiley-IEEE Press, May 1994.
- [23] A. Paulraj, D. Gore, R. Nabar, and H. Bolcskei, "An Overview of MIMO Communications - a Key to Gigabit Wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198 – 218, Feb. 2004.
- [24] R. Bellman, *Introduction to Matrix Analysis (2nd ed.)*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1997.
- [25] D. Baum, D. Gore, R. Nabar, S. Panchanathan, K. Hari, V. Erceg, and A. Paulraj, "Measurement and Characterization of Broadband MIMO Fixed Wireless Channels at 2.5 GHz," in *Personal Wireless Communications, 2000 IEEE International Conference on*, 2000, pp. 203 –206.
- [26] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO Paradigm," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 36 –46, Sep. 2007.
- [27] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink Channel Capacity of Space-Division-Multiple-Access Schemes," *IEEE Trans. on Inf. Theory*, vol. 44, no. 4, pp. 1468 –1476, Jul. 1998.

- [28] M. Varanasi and T. Guess, "Optimum Decision Feedback Multiuser Equalization with Successive Decoding Achieves the Total Capacity of the Gaussian Multiple-access Channel," in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems Computers*, vol. 2, Nov. 1997, pp. 1405–1409.
- [29] N. Jindal, S. Jafar, S. Vishwanath, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna gaussian broadcast channels," *IEEE Trans. on Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.
- [30] H. Sato, "An Outer Bound to the Capacity Region of Broadcast Channels," *IEEE Trans. on Inf. Theory*, vol. IT-24, pp. 374–377, May 1978.
- [31] G. Caire and S. Shamai, "On the Achievable Throughput of a Multiantenna Gaussian Broadcast Channel," *IEEE Trans. on Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [32] M. Costa, "Writing on Dirty Paper (Corresp.)," *IEEE Trans. on Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [33] P. Viswanath and D. Tse, "Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality," *IEEE Trans. on Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
- [34] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit Beamforming and Power Control for Cellular Wireless Systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.
- [35] E. Visotsky and U. Madhow, "Optimum Beamforming using Transmit Antenna Arrays," in *IEEE 49th Vehicular Technology Conference*, vol. 1, Jul. 1999, pp. 851–856.
- [36] S. Vishwanath, N. Jindal, and A. Goldsmith, "On the Capacity of Multiple Input Multiple Output Broadcast Channels," in *IEEE International Conference on Communications, 2002. ICC 2002*, vol. 3, 2002, pp. 1444–1450 vol.3.
- [37] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [38] W. K. Ma, T. Davidson, K. M. Wong, and P. Ching, "Multiuser Detection for Asynchronous CDMA using Block Coordinate Ascent and Semi-definite Relaxation," in *IEEE International Conf. on Acoustics, Speech, and Signal Proc., 2002. Proceedings. (ICASSP '02)*, vol. 3, 2002, pp. III-2309–III-2312.
- [39] A. Saleh, A. Rustako, and R. Roman, "Distributed Antennas for Indoor Radio Communications," *IEEE Transactions on Communications*, vol. 35, no. 12, pp. 1245–1251, Dec. 1987.
- [40] H. Zhang, N. Mehta, A. Molisch, J. Zhang, and H. Dai, "Asynchronous Interference Mitigation in Cooperative Base Station Systems," *IEEE Trans. on Wireless Communications*, vol. 7, no. 1, pp. 155–165, Jan. 2008.
- [41] J. Zhang and J. Andrews, "Distributed Antenna Systems with Randomness," *IEEE Trans. on Wireless Communications*, vol. 7, no. 9, pp. 3636–3646, Sep. 2008.
- [42] H. Zhang, N. Mehta, A. Molisch, J. Zhang, and H. Dai, "On the Fundamentally Asynchronous Nature of Interference in Cooperative Base Station Systems," in *IEEE International Conference on Communications, 2007. ICC '07*, Jun. 2007, pp. 6073–6078.

- [43] B. L. Ng, J. S. Evans, S. V. Hanly, and D. Aktas, "Transmit Beamforming with Cooperative Base Stations," in *International Symposium on Information Theory, 2005. ISIT 2005.*, Sep. 2005, pp. 1431–1435.
- [44] H. Zhang and H. Dai, "Cochannel Interference Mitigation and Cooperative Processing in Downlink Multicell Multiuser MIMO Networks," *EURASIP J. Wirel. Comm. Netw.*, vol. 2004, no. 2, pp. 222–235, 2004.
- [45] P. Ma and W. Wang, "A Multiuser Beamforming Scheme for Downlink Spatial Multiplexing MIMO," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, Sep. 2007, pp. 1–5.
- [46] Y. Jing and H. Jafarkhani, "Network Beamforming using Relays with Perfect Channel Information," *IEEE Trans. on Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.
- [47] T. Liu, S. Lu, and M. Zheng, "Beamforming for Per-Antenna Power Constrained Downlink SINR Optimization," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, May 2010, pp. 1–5.
- [48] T. Sada, J. Webber, T. Nishimura, T. Ohgane, and Y. Ogawa, "A Generalized Approach to Block Diagonalization for Multiuser MIMO Downlink," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, Sep. 2010, pp. 504–509.
- [49] W. Dziunikowski, "Multi-antenna Transceiver Techniques for 3G and Beyond [Book Review]," *IEEE Communications Magazine*, vol. 42, no. 2, pp. 8–10, Feb. 2004.
- [50] A. Paulraj and C. Papadias, "Space-time Processing for Wireless Communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, Nov. 1997.
- [51] A. Narula, M. Lopez, M. Trott, and G. Wornell, "Efficient Use of Side Information in Multiple-antenna Data Transmission over Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.
- [52] D. Love, R. H. Jr., and T. Strohme, "Quantized Maximum Ratio Transmission for Multiple-input Multiple-output Wireless System," in *Conf. Record of the 36th ACSSC'06*, vol. 1, pp. 531–535, Nov. 2002.
- [53] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On Beamforming with Finite Rate Feedback in Multiple-antenna Systems," *IEEE Trans. on Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.
- [54] E. Visotsky and U. Madhow, "Space-time Transmit Precoding with Imperfect Feedback," *IEEE Trans. on Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.
- [55] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel Capacity and Beamforming for Multiple Transmit and Receive Antennas with Covariance Feedback," in *Proc. of IEEE ICC'01*, vol. 7, pp. 2266–2270, Jun. 2001.
- [56] R. Gohary and T. Davidson, "On Rate-optimal MIMO Signalling with Mean and Covariance Feedback," *IEEE Trans. on Wireless Communications*, vol. 8, no. 2, pp. 912–921, Feb. 2009.

- [57] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A Stochastic MIMO Radio Channel Model with Experimental Validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211 – 1226, Aug. 2002.
- [58] W. Weichselberger, M. Herdin, H. Ozelik, and E. Bonek, "A Stochastic MIMO Channel Model with Joint Correlation of both Link Ends," *IEEE Trans. on Wireless Communications*, vol. 5, no. 1, pp. 90 – 100, Jan. 2006.
- [59] A. H. S. T. Kailath and B. Hassibi, *Linear Estimation*, 1st ed. Prentice Hall, 2000.
- [60] S. Haykin, *Adaptive Filter Theory (4th ed.)*. Prentice Hall, 2001.
- [61] C. Peel, B. Hochwald, and L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-antenna Multi-user Communication," in *Proc. of the 41st Allerton Conf. on Comm., Control, and Computing.*, Oct. 2003.
- [62] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-reduction-aided Broadcast Precoding," in *Proc. 5th International ITG Conf. on Source and Channel Cod.*, Erlangen, Germany, Jan. 2004, pp. 403–408.
- [63] M. Joham, W. Utschick, and J. Nosssek, "Linear Transmit Processing in MIMO Communications Systems," *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.
- [64] M. Bengtsson and B. Ottersten, *Handbook of antennas in wireless communications*, 2002, ch. Optimum and suboptimum transmit beamforming.
- [65] 3GPP, "User Equipment (UE) Radio Transmission and Reception," Technical Specification Group Radio Access Networks - Evolved Universal Terrestrial Radio Access (E-UTRA), Tech. Rep. TR 36.101 V10.1.1 - Release 10, Jan. 2011.
- [66] C. B. Peel, "Studies in Multiple-antenna Wireless Communications," Ph.D. dissertation, Brigham Young University, Provo, Utah, United States, Apr. 2004.
- [67] J. M. Lee, *Introduction to Smooth Manifolds*. Seattle, WA: Draft Version, 2000.
- [68] L. M. C. Sousa, C. C. Cavalcante, and F. R. P. Cavalcanti, "Cooperative MIMO Precoding with Limited Feedback," *GTEL-UFC Ericsson UFC.25; Final Technical Report*, 2010.
- [69] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 1st ed. Cambridge Univ. Press., 1999.
- [70] S. Haile, "Investigation of Channel Reciprocity for OFDM TDD Systems," Ph.D. dissertation, University of Waterloo, Ontario, Canada, October 2009.