Universidade
Estadual de Londrina

Centro de Tecnologia e Urbanismo
Departamento de Engenharia Elétrica

João Lucas Negrão

Efficient Detection: from Conventional MIMO to Massive MIMO Communication Systems

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Estadual de Londrina para obtenção do Título de Mestre em Engenharia Elétrica.

Londrina, PR
2017

Universidade Estadual de Londrina

João Lucas Negrão

Efficient Detection: from Conventional MIMO to Massive MIMO Communication Systems

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Estadual de Londrina para obtenção do Título de Mestre em Engenharia Elétrica.

Área: Sistemas de Telecomunicações

Orientador:
Prof. Dr. Taufik Abrão

Londrina, PR
2017

Ficha Catalográfica

Negrão, João Lucas
Efficient Detection: from Conventional MIMO to Massive MIMO Communication Systems. Londrina, PR, 2017. 111 p.

Dissertação (Mestrado) – Universidade Estadual de Londrina, PR. Departamento de Engenharia Elétrica.

I.Sistemas com Múltiplas Antenas. II.MIMO Massivo. III.Programação Semidefinida. IV.Redução Treliça. V.Técnicas de Detecção e Pré-codificação
Departamento de Engenharia Elétrica

João Lucas Negrão

Efficient Detection: from Conventional MIMO to Massive MIMO Communication Systems

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Estadual de Londrina para obtenção do Título de Mestre em Engenharia Elétrica.

Área: Sistemas de Telecomunicações

Comissão Examinadora

Prof. Dr. Paulo Rogério Scalassara
Depto. Acadêmico de Elétrica
Universidade Tecnológica Federal do Paraná -
Câmpus Cornélio Procópio

Prof. Dr. Fábio Renan Durand
Depto. Acadêmico de Elétrica
Universidade Tecnológica Federal do Paraná -
Câmpus Cornélio Procópio

Prof. Dr. Taufik Abrão
Depto. de Engenharia Elétrica
Universidade Estadual de Londrina
Orientador

22 de março de 2018

Dedico este trabalho ao meu pai Rafael Robson Negrão (*in memorian*).

Agradecimentos

Agradeço primeiramente a Deus por iluminar as escolhas em minha vida e todo caminho percorrido.

Agradeço também ao Prof. Dr. Taufik Abrão por toda a paciência, discussões e orientações, sem as quais este trabalho não chegaria a sua conclusão e também por todos os conselhos e conversas as quais foram essenciais para o meu crescimento profissional e pessoal.

Gostaria de agradecer em especial aos meus pais, Rafael Robson Negrão (*in memorian*) e Magna Solanges Negrão por todo o apoio durante minha vida, por todos os sacrifícios e pela educação que me proporcionaram. Agradeço também a minha namorada, companheira e esposa Paula da Silva Hatadani, por todo o suporte, compreensão, amor e carinho com os quais sou presenteado todos os dias.

Agradeço também aos meus amigos, pelos bons momentos e risadas que proporcionam todos os dias. Em especial: João Antônio Brancalion e Vitor Alegro que estão sempre ao meu lado.

Agradeço também à todos os colegas do Lab. Telecom&DSP, Lucas Claudino, Aislan Hernandes, Edno Gentilho Júnior, Alex Miyamoto Mussi, Prof. Jaime Jacob Laelson e em especial gostaria de agradecer ao amigo Ricardo Kobayashi por toda as discussões construtivas, as quais foram essenciais para o desenvolvimento de meu trabalho, bem como as conversas e músicas compartilhadas.

Por fim, expresso minha gratidão à CAPES pelo auxilio financeiro durante esta etapa.

Resumo

Ao longo deste trabalho, problemas relacionados aos sistemas de comunicação equipados com múltiplas antenas no transmissor e receptor (MIMO - *Multiple-Input Multiple-Output)* são analisados sob o ponto de vista de detecção clássica, da otimização não-linear, bem como da pré-codificação linear, desde MIMO convencional (algumas antenas no Tx e Rx) até sistemas MIMO de larga-escala (massivo). Inicialmente, a eficiência de detecção de vários detectores MIMO foi analisada sob a prerrogativa de canais altamente correlacionados, situação em que sistemas MIMO apresentam elevada perda de desempenho, além de, em alguns casos, uma crescente complexidade. Diante deste cenário, foi estudado especificamente o comportamento em termos do compromisso complexidade × taxa de erro de bits (BER - *Bit Error Rate*), para diferentes técnicas de detecção, como o cancelamento de interferências sucessivo (SIC), redução treliça (LR), bem como a combinação de cada uma destas às técnicas lineares de detecção. Nessa análise, também foram considerados diferentes estruturas de antennas uniformes com arranjos geométricos lineares (ULA - *uniform linear array*) e de arranjo planar (UPA - *uniform planar array*) em ambos transmissor e receptor. Além disso, também foram considerados diferentes número de antenas e ordem de modulação. Em seguida, o problema de detecção MIMO foi estudado sob uma perspectiva de otimização não-linear, visando especificamente alcançar o desempenho ótimo. Foi analisada a solução de detecção com relaxação semi-definida (SDR - *semi-definite relaxation*). O detector SDR-MIMO é uma abordagem eficiente capaz de atingir o desempenho muito próximo ao ótimo, especialmente para baixas e médias ordens de modulação. Concentramos nossos esforços no desenvolvimento de uma aproximação computacionalmente eficiente para o algoritmo de detecção de máxima verossimilhança (ML - *Maximum Likelihood*) MIMO baseado na programação semi-definida (SDP - *Semidefinite Programming*) para as constelações M-QAM. Finalmente, estuda-se um problema de alocação de potência com o objetivo de maximizar a capacidade de um canal de *broadcasting* MIMO massivo em uma única celula equipada com pré-codificação forçagem à zero (ZFBF - *zero-forcing beamforming*) e inversão de canal regularizado (RCI - *regularized channel inversion*) na estação rádio base (BS). Nosso objetivo é investigar esse problema considerando um sistema massivo no limite, ou seja, quando o número de usuários, K, e antenas na BS, M, tendem ao infinito porém com uma razão constante, $\beta = \frac{K}{M}$. Primeiramente deriva-se a relação sinal-interferência mais ruído (SINR) para ambos os pré-codificadores escolhidos. Em seguida, investiga-se um esquemas de alocação de potência ótimo que maximiza a soma das capacidades por antena sob uma restrição de potência máxima disponível, conclui-se que o problema é convexo e que a alocação de potência ótima segue a estratégia de *watter-filling* (WF). Também estudou-se o problema relacionado à alocação de potência em um grupo finito de usuários separados em grupos e determinou-se o impacto desse esquema na capacidade total do sistema.

Palavras Chave: MIMO, MIMO Massivo , Detecção, Pré-codificação, Otimização

Abstract

Throughout this work, problems related to communication systems equipped with multiple antennas in the transmitter and receiver (MIMO - Multiple-Input Multiple-Output) are analyzed from the point of view of classical detection, nonlinear optimization, as well as linear pre-coding, from conventional MIMO (some Tx and Rx antennas) to large-scale (massive) MIMO systems. Initially, the detection efficiency of several MIMO detectors were analyzed under the prerogative of highly correlated channels, in which situation, MIMO systems present a high loss of performance, and, in some cases, an increasing complexity. Considering this scenario, we have specifically studied the behavior in terms of compromise complexity × bit error rate (BER), for different detection techniques, such as the successive interference cancellation (SIC), lattice reduction (LR), as well as the combination of each of these with linear detection techniques. In this analysis, different uniform antenna structures with uniform linear array (ULA) and planar array array (UPA) were also considered in both transmitter and receiver side. In addition, different number of antennas and order of modulation were also considered. Next, the MIMO detection problem was studied from a nonlinear optimization perspective, specifically aiming to achieve optimum performance. The detection solution with semi-defined relaxation (SDR - it semidefinite relaxation) were analyzed. The SDR-MIMO detector is an efficient approach capable of achieving near-optimal performance, especially for low and medium modulation orders. We focused our efforts on developing a computationally efficient approach for the maximum likelihood (ML) MIMO detection algorithm based on semi-definite programming (SDP) for M-QAM constellations. Finally, we study an optimal power allocation problem aiming to maximizes the sum-rate capacity of a single cell massive MIMO broadcast channel equipped with zero-forcing beamforming (ZFBF) and regularized channel inversion (RCI) precoding at the base station (BS). Our purpose is to investigate this problem in the large-scale system limit, i.e, when the number of users, K, and antennas at the BS, M, tend to infinity with a ratio $\beta = \frac{K}{M}$ being held constant. We first derive the signal to interference plus noise (SINR) ratio for both chosen precoders. Then we investigate optimal power allocation schemes that maximize the sum-rate per antenna under an average power constraint and we show that the problem is convex and the power allocation follows the well-known Water-Filling strategy. We also studied a problem related to an optimal power allocation at a finite group of clustered users and determine the impact of this scheme in the *ergodic* sum-rate capacity.

Keywords: MIMO, Massive MIMO, Detection, Precoding, Optimization

List of Contents

List of Figures

List of Tables

Abbreviations List

5G	Fifth Generation of mobile communications
AF	Array Factor
AoD	Angle-of-departure
a.s	Almost Sure Convergence
AWGN	Additive White Gaussian Noise
BC	MIMO Broadcast Channel
BER	Bit-Error rate
BF	Conjugated precoder
BPSK	Binary Phase Shift Keying
BS	Base Station
CSI	Channel State Information
CSIT	Channel State Information at the Transmitter
DPC	Dirty-Paper-Coding
EE	Energy Efficiency
EP	Equal Power allocation strategy
i.i.d.	Independent and Identically Distributed
IUI	Inter-user Interference
KKT	Karush-Kuhn-Tucker
LLL	Lenstra-Lenstra-Lovász
LLR	Log-Likelihood Ratio
LOS	Line-of-Sight

LR	Lattice-Reduction
MIMO	Multiple-Input Multiple-Output
MF	Matched Filter
ML	Maximum-Likelihood
MMSE	Minimum Mean Squared Error
MRC	Maximum Ratio Combining
MT	Mobile Terminal
MUI	Multi-user Interference
MU-MIMO	Multiuser Multiple-Input Multiple-Output
NLOS	Non-Line-of-Sight
OSIC	Ordered Successive Interference Cancellation
QAM	Quadrature Amplitude Modulation
QCQP	Quadratically Constrained Quadratic Programming
RCI	Regularized Channel Inversion
r.v.	Random variable
SD	Sphere Decoding
SDP	Semi-definite Programming
SDR	Semi-definite Relaxation
SE	Spectral Efficiency
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference plus-Noise Ratio
SNR	Signal-to-Noise Ratio
SQRD	Sorted QR Decomposition
s.t.	Subject to
ULA	Uniform Linear Array

UPA	Uniform Planar Array
ZF	Zero-Forcing
ZFBF	Zero-Forcing Beamforming
WF	Water-Filling
w.r.t.	with respect to

Conventions and Notations

The following mathematical notations where adopted in this work:

	Boldface lower case letters represent vectors;
	Boldface upper case letters denote matrices;
$(\cdot)^{-1}$	Inversion operator;
$(\cdot)^{H}$	Hermitian operator (transposition and conjugation);
$(\cdot)^{T}$	Transposition operator;
$(\cdot)^{\star}$	Conjugation operator;
$(\cdot)^{\dagger}$	Moore-Penrose pseudo-inverse;
$(\cdot)^{*}$	Optimal solution;
$\det(\cdot)$	Determinant of an Square Matrix;
$\lVert\cdot\rVert_n$	Norm of order n;
$\mathrm{tr}\,(\cdot)$	Trace operation;
$\widetilde{(\cdot)}$	Boldface lower case letter with tilde superscript represent a symbol vector estimation;
$\widehat{(\cdot)}$	Boldface lower case with hat superscript represents a symbol estimation after a slicer;
$\mathrm{diag}(\cdot)$	Diagonalization Operation;
$\lfloor\cdot\rceil$	Round Operation;
$\lceil\cdot\rceil$	Superior Round Operation;
$\lfloor\cdot\rfloor$	Inferior Round Operation;
$\mathbf{I}_m$	Identity matrix of order m;
$\mathbf{0}_{m\times n}$	All zero matrix of size $m \times n$;

$\mathbf{1}_{m\times n}$	All ones matrix of size $m \times n$;
$\mathcal{CN}\{\mu, \sigma^2\}$	Gaussian Random Variable circularly-symmetric with mean μ and variance σ^2;
$\mathcal{O}(\cdot)$	Complexity order of an operation or algorithm;
$\mathbb{E}\left[\cdot\right]$	Statistical Expectation;
$\mathbb{C}$	Complex numbers set;
$\mathbb{N}$	Natural numbers set;
$\mathbb{Z}$	Integer numbers set;
$\Re\left\{\cdot\right\}$	Real part of a complex number;
$\Im\left\{\cdot\right\}$	Imaginary part of a complex number;
$\in$	Belongs to the set;

Symbols List

Chapter 2

n_T	Number of antennas at the transmitter
n_R	Number of antennas at the receiver
$\mathbf{s}$	Transmitted symbols vector
$\mathbf{H}$	MIMO systems channel matrix
$\mathbf{n}$	Additive Gaussian Noise Vector
$\mathbf{x}$	Received vector after passing symbols through the channel
σ_n^2	Noise Variance
$\mathcal{S}$	Complex symbols Set
M	Modulation Order
E_s	Transmitted symbols mean energy
λ	Transmitted signal wave length
$AF_{\text{ula}}, AF_{\text{upa}}$	Array factor for both ULA and UPA array elements
d	Antenna elements spacing at ULA
d_x, d_y	Antenna element spacing in direction of x and y-axes respectively
N	Number of antenna elements placed at x-axes for ULA
N'	Number of antenna elements placed x-axes and y-axes for UPA
$I_n, I_{n'}$	Amplitude excitation for each antenna element in ULA and UPA respectively
ϕ, θ	Azimuth and Elevation antenna angles respectively
u/v	Azimuth and Elevation mapped in Cartesian coordinates.
$\mathbf{R}_{H,Rx}, \mathbf{R}_{H,Tx}$	Spacial correlation matrix seen by the receiver and transmitter respectively
ρ	Correlation index
$\mathbf{R}_{H,x}, \mathbf{R}_{H,y}$	Spacial correlation matrix along the x and y-axes, respectively for the UPA structure
N_v, N_h	Number of antenna element in the horizontal and vertical plane of an UPA used for Geometrical modeling
$\mathbf{R}_{\text{ula}}, \mathbf{R}_h$	Geometrical ULA and UPA correlation model respectively
$\mathbf{R}_{\text{az}}, \mathbf{R}_{\text{ez}}$	Azimuth and Elevation correlation matrix for the geometrical UPA correlation model
δ, ξ	Standard deviation for both azimuth and elevation AoD respectively

$\widehat{\mathbf{s}}$	Estimated symbol vector after the slicer.
$\mathbf{Q}, \mathbf{R}$	Orthogonal and upper triangular matrices provided by the QR decomposition
$\mathbf{H}^{\dagger}$	Channel pseudo-inverse or ZF equalization matrix
$\underline{\mathbf{H}}$	Channel matrix extended version
$\underline{\mathbf{x}}$	Received signal extended version
$\mathbf{\Pi}$	Permutation matrix for ordered detection
$\widetilde{\mathbf{H}}$	Channel matrix in LR domain
$\mathbf{T}$	Unimodular matrix generated from LLL algorithm
$\mathbf{z}$	Transmitted signal vector in LR domain
$\widetilde{\mathbf{z}}$	Estimated symbol vector in LR domain
$\widehat{\mathbf{z}}$	Estimated and quantized symbol vector in LR domain
β'	Variable used for quantization in LR domain
$\underline{\widetilde{\mathbf{H}}}$	Extended version of the channel matrix in LR domain
$\underline{\mathbf{T}}$	Unimodular $\mathbf{T}$ matrix extended version
$\widetilde{\mathbf{Q}}, \widetilde{\mathbf{R}}$	Orthogonal and upper triangular matrices provided by the QR decomposition of the channel matrix $\tilde{\mathbf{H}}$ in LR domain

Chapter 3

n_T	Number of antennas at the transmitter
n_R	Number of antennas at the receiver
$\mathbf{s}$	Transmitted symbols vector
$\mathbf{H}$	MIMO systems channel matrix
$\mathbf{n}$	Additive Gaussian Noise Vector
$\mathbf{x}$	Received vector after passing symbols through the channel
σ_n^2	Noise Variance
$\mathbb{S}$	Complex symbols Set
M	Modulation Order
E_s	Transmitted symbols mean energy
$\widehat{\mathbf{s}}$	Estimated symbol vector after the slicer
$\mathbf{L}$	Auxiliary matrix to closure the QCQP problem in the SDP form
$\mathbf{X}$	Relaxed solution set matrix
$\mathbf{e}$	All ones vector
I_L, S_L	Inferior and superior constellation limits on the first constraint of our SDP problem.
m, n, ϵ	Number of constraints, SDP problem size and solution accuracy for worst case complexity evaluation.

λ_i	Eigenvalue of the i-th line for the eigen-decomposition of $\mathbf{X}^*$
$\mathbf{U}$	Lower triangular matrix with real and positive diagonal entries
S_g	Number of randomization samples

Chapter 4

M	Number of antennas at the BS transmitter
K	Number of single-antenna users (receivers)
β	Ratio between M and N (Cell-loading)
$\mathbf{x}$	Transmitted symbols vector
$\mathbf{A}$	Path-loss coefficient matrix
$\mathbf{H}$	MIMO systems channel matrix
$\mathbf{n}$	Additive Gaussian Noise Vector
$\mathbf{y}$	Received vector after passing symbols through the channel
σ_n^2	Noise Variance
$\mathbf{G}$	Linear Precoding Matrix
$\mathbf{P}$	Vector representing the power allocated for each user
P	Total available transmit power
a_k	Path-Loss associated to user k
γ	SNR at the receiver
SINR_k	SINR per user
λ	Signal wave length
$\mathbf{R_h}$	Transmit correlation matrix
$\mathcal{R}_k$	Sum-Rate capacity of user k
$\mathcal{R}_\Sigma$	*Ergodic* sum-rate capacity
$\mathbf{G}_{\text{ZF}}$	ZFBF precoding matrix
α	Parameter ensuring the transmit power constraint (power normalization)
$\text{SINR}_{k,\text{ZF}}$	SINR of user k under ZFBF precoder
$\mathbf{G}_{\text{RCI}}$	RCI precoding matrix
ξ	Regularization parameter
$\text{SINR}_{k,\text{RCI}}$	SINR of user k under RCI precoder
$\mathbf{G}_{\text{BF}}$	MF precoding matrix
p_k^*	Optimal allocated power
μ	Water level
K_A	Number of active antennas before water-filling strategy
$\mathbf{X}$	Rectangular matrix with independent entries
$f_\beta(x)$	Probability density function of an Hermitian matrix $\mathbf{XX}^H$, which is given by the Marčenko-Pastur law

X	Real valued random value
F_X	Probability distribution function of a X
m_X	Stieltjes transform of F_X, or probability distribution function of F_X in the large limit.
$\mathbf{\Lambda}$	Diagonal matrix containing the eigenvalues of $\mathbf{X}$
$\mathbf{x}$	Random vector whose entries are i.i.d. with zero mean and variance $\frac{1}{N}$
$\mathbf{A}_N$	Any $\mathbb{C}^{N\times N}$ matrix with uniformly bounded spectral norm.
$\mathbf{B}$	Any $\mathbb{C}^{N\times n}$, $n < N$ matrix, where the entries are i.i.d. elements with zero mean and variance $\frac{1}{N}$
ν	Regularization parameter in the Large System limit
$g(\beta,\nu)$	Probability distribution of $g_N = \mathbf{x}\left(\mathbf{B}\mathbf{B}^H + \nu\mathbf{I}_N\right)\mathbf{x}^H$
$\mathsf{S}(\beta,\nu)$	Signal in the Large Limit
$\mathsf{I}(\beta,\nu)$	Interference in the Large Limit
$\mathrm{SINR}^{\infty}(\gamma,\beta,\nu)$	Deterministic limiting SINR
$\mathcal{R}_k^{\infty}$	Limiting sum rate capacity per-user
$\mathcal{P}$	Empirical mean of the users power or just average power.
$\mathfrak{b}$	Path-loss Exponent
L	Number of Clusters
R	Cell-Radius
$\mathcal{R}_{\Sigma}^{\infty}$	Limiting Achievable sum rate capacity
$\mathcal{L}$	Lagrangian for problem (4.46)
λ, μ_j, κ	Associated Lagrange multipliers

1 Introduction

Nowadays the telecommunications services and their technological advances present an indispensable role in our lives. We are surrounded by devices that provide multimedia services in real time, such as smartphones, tablets, cable TVs. They have caused great changes in the way that human beings interact to each other and with the world that surrounds them. The communications services and the use of devices are increasing every year, as they have the purpose of promoting convenience, security, leisure and connection to their users. As a result, there are growing demands for telecommunication services in terms of number of users served simultaneously occupying the same frequency spectrum, new services requiring higher transmission rates, and higher network speeds.

Such demands require more capacity and reliability from today's wireless systems. However, this technological demand arises in a scenario in which spectrum and energy availabilities become increasingly limited because of a variety of services sharing the wireless channel, while there is a growing concern about saving energy. In this context, one of the most promising solutions to this problem is the technology of multiple antennas in both transmission and reception, known as MIMO (multiple-input multiple-output). MIMO systems are one of the main foundations of modern wireless communication systems such as Long Term Evolution (LTE), 3GPP LTE-Advanced (LTE-A), IEEE 802.11 (Wi-Fi), IEEE 802.16e - Worldwide Interoperability for Microwave Access (WiMAX) (HANZO et al., 2010; LI et al., 2010), due to the high energy and spectral efficiency rates achieved.

Despite the first appearances of the MIMO systems in literature are in the early twentieth century (FOSCHINI; GANS, 1998), the applications in practical communication systems can be considered recent (LI et al., 2010). Since their inception, several technological advances has been proposed and implemented (at least in terms of proof-of-concept), focusing to close the demand gap that emerges every year. On the next generation of communication systems, namely the fifth generation (5G), one of the most potential technologies that has been considered for application is the massive or large-scale MIMO systems, which consists in the usage of a MIMO with very large antenna arrays at both transmission and receiver sides (RUSEK et al., 2013; BOCCARDI et al., 2014).

When it comes to point-to-point MIMO communications, the transmitted signals are linearly combined through the wireless link, which create possibilities to enhance the system features. There are techniques designed to achieve diversity gains and improve channel reliability, such as those related to space-time block coding (TAROKH et al., 1999; ZHENG; QIU; ZHU, 2004), and those that operate with spatial multiplexing, which are intended to maximize data rates by providing multiplexing gains (WOLNIANSKY et al., 1998).

For MIMO systems, there are two families of detectors with promising performance × complexity tradeoff: a) sphere decoding (SD) based MIMO detector; and b) the ones based on semidefinite relaxation (SDR). The SD detectors can achieve the same solution as the optimum maximum likelihood (ML) detector, however, they require lower complexity for solutions in high signal-to-noise (SNR) regime. On the other hand, the SD-based detectors are inefficient for large array size problems, with high order constellation or even in low SNR regime. In such situations, the SD complexity is expected to grow exponentially, i.e, to the same complexity order of the ML detector, becoming inefficient in those system configuration scenarios.

On the other hand, for the SDR-based MIMO detector, the result complexity becomes polynomial in association with promising performance results. The semidefinite relaxation is an optimization technique used to solve many different problems related to non-linear optimizations, specially applied to telecommunication problems. On MIMO systems, the SDR-based detection was first proposed for low constellation order problems, such as, binary/quadratic phase shift keying (BPSK/QPSK) (JALDEN; MARTIN; OTTERSTEN, 2003; MA; CHING; DING, 2004); in which very near-ML optimal performance was observed. These results suggest that even with higher constellations, there is a high probability that SDR will yield the true ML decision; so in Mao, Wang e Wang (2007) the SDR detection problem was generalized considering 4^q-QAM ($q \geq 1$) modulation orders demonstrating the potential of the SDR-based detection.

Lattice reduction (LR) technique has been deployed to enhance the detection performance, specially in MIMO systems. This technique is applied in the predetection phase, with the objective of aiding in the separation of the signal from interference plus noise. Since LR is used to improve channel conditions, it allows the use of simpler detection techniques (WUBBEN et al., 2011). As a consequence, with a small addition at the level of complexity, BER performance of the system is able to achieve substantial improvement. The LR technique is based on a mathematical concept developed to solve different problems involving points in a lattice or trellis. A lattice is an arrangement of discrete points, which can be described by infinite vector basis, with this wide range of vector basis options, we have the flexibility to choose the most interesting for the problem in question. When it comes to MIMO context, the closest the orthogonality, the better is the performance of linear detectors; and the smaller the basis, lesser interference between antennas; because of that features, in many applications the smallest vector basis are the ones with most interest in MIMO systems applications (LING; MOW; HOWGRAVE-GRAHAM, 2013).

There are many algorithms that can be used in order to implement the lattice reduction technique; among them, one that stands out is the Lenstra, Lenstra & Lovász (LLL) algorithm (LENSTRA; LENSTRA; LOVÁSZ, 1982). This algorithm has shown greater viability when it comes to computational cost, due to present polynomial complexity at any operative configuration of the system (LING; MOW; GAN, 2009). Generically speaking, the LLL algorithm generates an unimodular matrix, which transforms the coefficient channel matrix into a new equivalent matrix, but using a new reduced vector basis and nearest to the orthogonality condition.

Currently, the scenario of greatest interest in MIMO communications is the multi-user application, where each cell has its own radio base station (BS) eqquiped with multiple antennas, that serves a pre-determined number of users inside the cell, all of them equipped with a single-antenna. In this case, the interest is to increase simultaneously the energy efficiency (EE) and spectral efficiency (SE) of order of 10-100 and 100-1000 times, respectively, higher than those achieved with conventional small-scale MIMO. Indeed such results can be achieved deploying a large number of antennas in the BS, typically hundreds of antenna elements, hence the name massive or large-scale MIMO systems.

Even in MIMO point-to-point systems or in multi-user systems, work with large system dimensions can bring a lot of benefits, such as, reaching higher energy efficiency and/or spectral efficiency, as well as a greater number of users served and mainly the system becomes immune to additive noise. Although some problems arise with this configuration, as for instance, when the number of antennas grow, the available space to accommodate them hold the same, causing an increasing on the channel correlation effect. In general , the correlation effect emerges as the distance between both transmit or receive antennas decreases, and the channel for those systems with physically close antennas becomes increasingly similar. As the correlation between channels increase, spatial diversity and, consequently, performance in MIMO systems are reduced substantially.

There are some alternatives to circumvent the problem of correlation between antennas that can be listed. The first one, and simplest, is to increase the spacing between them. However, it will lead to a lower quantity of antennas in determined area. Another alternative is the use of one of the promising technologies for the next generation of communications, the millimeter-waves (ROH et al., 2014). As this technique is composed of extremely high frequencies, this leads to signals with millimeter wavelength, which allows to decrease the antenna spacing until a half wavelength ($\approx \lambda/2$) without incurs in substantial antenna/channel correlation. Finally, MIMO detection techniques able to deal with the spacial correlation between antennas can be applied; those techniques can attenuate the correlation effect at high SNR regime, which is the case of the lattice reduction (WUBBEN et al., 2004).

This work focus on the analysis of performance and complexity trade-off for MIMO systems equipped with large arrays in different scenarios of channel correlation and number of antennas, as well as modulation order. Different detection and precoding techniques, that are able to achieve efficient operation, are analyzed under the proposed scenarios in order to determine the characteristics of each one and identify its benefits.

In the sequel, the work contributions are summarized; also, more detailed bibliographical reviews can be found at the respective chapters as well.

1.1 Work Contribution

- **Chap. 2**: This chapter provides a BER performance (reliability) × complexity trade-off for a broad class of MIMO detectors operating under realistic scenarios, where we consider different antennas structures under different correlated channel and system scenarios. Lattice-reduction technique proves its strength under correlated MIMO channels, increasing significantly the BER performance; however, there is a concern with the complexity because it scales exponentially as the correlation index grows. The combination of both lattice-reduction and ordered successive cancellation technique aided linear detectors delivered the best trade-off on the reliability × complexity figures of merit; it is apparent that such combination techniques fits as the best choice for larger, correlated channels.

- **Chap. 3**: It also provides a BER performance (reliability) × complexity tradeoff based on semidefinite relaxation MIMO detection procedures, which are able to deliver a near optimum performance. The focus is the evaluation of such class of MIMO detector under near large-scale antenna condition, which consists in a unicellular scenario with high number of antennas under correlated channels. Our results have proved that as the number of antennas increase, the lattice-reduction aided detectors have a lack of performance. On the other hand the use of SDR strategy is able to hold the performance suitable. Two algorithms which aims reconstruct the optimal solution were analyzed, the Rank-1 approximation and the Gaussian randomization; at the massive channel cases, where the number of antennas is increased up to 128 for both transmit and receive ones, the Rank-1 approximation proved to be the best choice, since it provides a great performance improvement while keeping an affordable complexity.

- **Chap. 4**: In this chapter, we provide an optimal power allocation scheme aiming to maximize the sum-rate of a single cell massive MIMO broadcast channel equipped with zero-forcing beamforming (ZFBF) and regularized channel inversion (RCI) precoders at the base station (BS). We analyze the problem from the perspective of an uniform linear array (ULA) antenna structures at the BS, which is equipped with many antennas while mobile users are fitted with a single-antenna causing them to become uncorrelated. Our purpose is to investigate this problem in the large-scale limit, so it is necessary to know the behavior of the signal-to-interference-plus-noise ratio (SINR) in order to evaluate the system capacity. Knowing this, an investigation related to an optimal power allocation scheme which maximize the *ergodic* sum-rate capacity under the average power constraint is carried out. We prove that the problem is convex and that the power allocation follows the well-known Water-Filling (WF) strategy. Hence, the main contribution of this part of the work is related to an optimal power allocation scheme related to a finite group of clustered users and to determine the impact of this scheme on the *ergodic* sum-rate capacity. Using the WF strategy our goal is to ensure the best path-loss distribution over the cell which turns the capacity to get close enough to the one achieved by the RCI-WF precoder in a cell with uniform random user distribution.

2 Efficient Detection for Uniform Array MIMO Systems under Correlated Channels

The Multiple-Input Multiple-Output systems are recognized by the capacity to provide significant spectral efficiency and/or performance enhancements on wireless communication systems by the use of multiple antennas at both transmitter and receiver sides. In spatial multiplexing gain mode, the deployment of simultaneously transmit data streams through multiple antennas were developed to enhance the spectral efficiency at the cost of increasing data detection complexity at the receiver side (WUBBEN et al., 2011). The V-Blast architecture, proposed in the pioneer work Wolniansky et al. (1998), was able to exploit the communication channel capacity, providing spatial multiplexing gain and high data rates, which inspired so many works into multiple antenna systems. This spacial multiplexing gain on MIMO systems is achieved by dividing the total transmitted power over the antennas, taking advantage of the multi-path diversity to achieve a great array gain, in other words, more bits per second per Hertz of bandwidth are transmitted. Moreover, with MIMO systems, improvements can be considered on the transmitted energy efficiency, data rates and/or symbol error rates, being defined by the antennas disposal at array configuration and the transmission-detection techniques applied. In project meanings, it is necessary to balance these improvements with the available resources in the systems, this procedure is crucial, since energy and spectrum are a scarce resource. Thereby, the purpose is to provide solutions attaining to a performance improvement under a low or moderate complexity constraint. Hence, the goal of this work consists in study MIMO architectures equipped with low or moderate complexity detectors, keeping appropriate BER performance under full diversity condition. Moreover, linear MIMO detectors and their combinations with sub-optimal equalization techniques like ordering , interference cancellation (SIC) and LR were studied in terms of performance-complexity trade-off.

Another relevant consideration in our work is the correlated fading channels; as currently the physical size of communication devices are being greatly reduced, the space to accommodate the antennas in those device is reducing as well. In realistic MIMO systems, operating under ultra high frequency (UHF) ranges, the desired antenna element spacing to provide an uncorrelated channel state is reasonably great. Moreover, MIMO systems equipped with a great number of antennas, and exploit the maximum multiplexing gain (or even the maximum diversity gain) is a project challenge. Thereby, it is easy to conclude that a correlated MIMO channel scenario will cause degradation effects on the performance, as well as the achievable rates; and in practical conditions this will result in more transmitting power needed. Hence, efficient MIMO detectors operating under proper BER performance and transmit power limits, which is directly connected to the SNR, are of great interest.

Recently, large (or massive) MIMO systems have arrived as a technology for 5G systems carrying many promises (BOCCARDI et al., 2014), such as higher spectral and energy efficiency and mainly the immunity to additive noise provided by very large arrays (RUSEK et al., 2013). When the number of antennas becomes large some effects arise, such as, channel properties that were random before now appears deterministic; e.g, singular values of the channel matrix approach to deterministic functions; system is limited by interference from other transmitters because thermal noise is averaged out (RUSEK et al., 2013). Although, large arrays bring two main problems: correlation between antennas and the signal processing complexity. The first one comes from the fact that the accommodation area for the large arrays are small, causing the effect of the correlated channels. The second occurs due to an increasing demand of signal processing which arises from the large number of antennas, which requires more hardware and operations from the system. Therefore, the study on MIMO processing techniques is important to know the limitations of each scenario and to analyze the best choices, in terms of performance and complexity trade-off, to practical high efficiency communication systems.

The decoupling of a transmitted signal originated from a received signal sample, can be designated as the main problem of MIMO detection. As it is known, the MIMO systems send data over different antennas, that travel over different paths, then the signal at the receiver side, at each antenna, is a combination of every transmit antenna signal and the received signal is a combination of every transmit antenna. There are many techniques on MIMO structures capable to decouple the transmitted signal, each one offers an achieved performance and a complexity level, the design challenge is to balance the available resources into the project requirements.

The optimal algorithm that achieves a minimum joint probability of error, detecting all the symbols simultaneously, is the maximum likelihood (ML) detector, that is known to be NP-hard. It can be carried out with a brute force-search over all possibilities in the transmitted vectors set, searching for the one that minimizes the Euclidean distance from the received vector. However, the expected computation complexity of the ML receiver is unpractical for many applications. Another possibility when considered looking for near-optimum performance is the sphere decoding (SD), that is a promising approach on MIMO detection. The SD provides lower complexity when compared to the ML for small noise value, but remains complex under low or medium SNR regions for real communication systems, becoming of the same order of ML complexity for low SNR region (JALDEN, 2004; BARBERO; THOMPSON, 2008).

Moreover, classic linear MIMO detection approaches are considered, such as the zero-forcing (ZF) detector which is know by being able to completely remove inter-antenna interference, at the cost of a significantly increase at the additive noise for ill conditioned channel matrices. There is also the minimum mean squared error (MMSE)-based detector which can be considered as a better alternative, once it considers the noise power throughout the symbol detection procedure. Besides ZF and MMSE detectors when combined with SIC (BOHNKE et al., 2003) perform an layer-by-layer detection canceling the interference form the previous detected symbol. Since first layers detection errors can be propagated along the algorithm, the ordered SIC (OSIC) (WOLNIANSKY et al., 1998; WUBBEN et al., 2003) detector provides remarkable improvements on performance by detecting the most reliable antennas first. Both ZF and MMSE detectors when combined with OSIC turns into detection schemes able to provide lower complexity compared to the ML or even the SD detector, however present greater degradation in the BER performance. Furthermore, pre-processing techniques such as the lattice reduction (LR) (VALENTE; MARINELLO; ABRÃO, 2014; MA; ZHANG, 2008; WUBBEN et al., 2004) aided linear MIMO detectors can be used to simultaneously provide performance improvement and complexity reduction, since the transformed channel has quasi-orthogonality features it will improve the final quality of the detected signal, achieving, in some cases, near-ML performance. The LR computational complexity is recognized as polynomial in time; however, highly correlated channel scenarios result in devastating impacts on the MIMO channel matrix estimation while the LR-aided linear MIMO detectors may result in an undesirable additional complexity, especially when the system is equipped with a high number of antennas (VALENTE; MARINELLO; ABRÃO, 2014). However, those problems are part of the challenge to implement the applicability of large-MIMO systems.

2.1 System Model

Considering a point-to-point MIMO system composed by n_T transmit antennas and n_R receive antennas, where the transmitted data is demultiplexed over the n_T transmit antennas. A MIMO system topology is depicted in Fig.2.1.

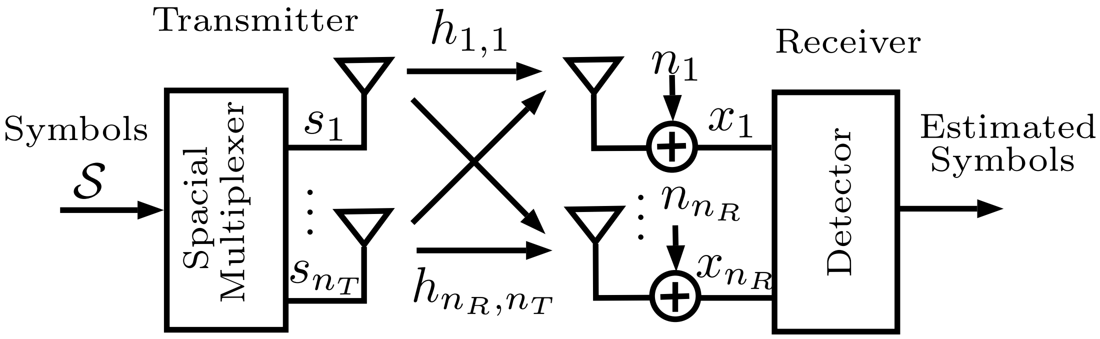


Figure 2.1: MIMO System with Spacial Multiplexing

The model is considered under an overdetermined MIMO system, i.e., $n_R \geq n_T$, working in spatial multiplexing mode. A classical problem in MIMO systems consists in reliably detect the transmitted symbol, despite the channel's distortion and noise (LARSSON, 2009). Thereby, the received signal can be described by:

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{s}_{n_T \times 1}$ symbols are transmitted through a channel which gain is represented by $\mathbf{H}_{n_R \times n_T}$ and additive noise $\mathbf{n}_{n_R \times 1}$. Each element of matrix $\mathbf{H}$ represents the channel gain for a selected path and these gains are known at the receiver. The column-vector $\mathbf{x}_{n_R \times 1}$ represents the received signal vector, formed by the symbols after passing through the channel. Furthermore, it is still possible to add a pre-processing block that is responsible to modulation and coding steps, the last one is not applied in this section, while the MIMO decoder works to recover the transmitted data from the received signal, that is corrupted by noise and inter-antenna interference.

It is also assumed that the noise vector $\mathbf{n}$, are samples of additive noise represented as circularly-symmetric Gaussian distribution, $\mathbf{n} \sim \mathcal{CN}\{0, \sigma_n^2 \mathbf{I}\}$, with variance σ_n^2. An alternative way to represent the noise statistics is through the covariance matrix $\mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right] = \sigma_n^2 \mathbf{I}_{n_R}$

In order to achieve better spectral efficiency and performance we will consider a M-QAM modulation, where the symbols are denoted by a complex number which real and imaginary part are limited to $\pm\left(\sqrt{M}-1\right)$ (BAI; CHOI; YU, 2014; WUBBEN et al., 2004; KOBAYASHI; CIRIACO; ABRÃO, 2015).

The structure of the complex set is represented by

$\mathcal{S} = \left\{a + jb \quad | \quad a, b \in \left\{-\sqrt{M}-1, -\sqrt{M}+3, \ldots, \sqrt{M}-1\right\}\right\}$

For this modulation, the average symbol energy is given by:

$$E_s = \frac{2(M-1)}{3} \tag{2.2}$$

Also, it will be adopted Gray coded symbols, where adjacent symbols differ only one bit, which minimize the BER performance.

Furthermore, the channel model will be kept simple, however substantially adequate to the proposed systems. Specifically it is used a MIMO channel under Rayleigh fading and under the effect of spatial correlation between antennas. The Rayleigh fading is modeled as two random variables (r.v.) that follows circular complex Gaussian distribution, with zero-mean and unitary variance, i.e., $h_{ij} \sim \mathcal{CN}\{0, 1\}$, whose magnitude is represented by a Rayleigh r.v., while the phase is represented by a uniform distributed r.v. (CHO et al., 2010). Furthermore, is worth to note that the Rayleigh fading model is valid for environments that is rich in scattering, i.e., highly urbanized environments or with great number of obstacles. It means that the signal do not have a prevalence path, i.e., non-line-of-sight (NLOS) channels (GOLDSMITH, 2005).

2.1.1 Correlated MIMO Rayleigh-Fading Channels

This section discusses the MIMO channel correlation among different antenna structures. As we have already defined the channel basic characteristics, the next step is to evaluate the correlation effect and how to model it. The space for the accommodation of antennas elements in wireless systems is in many cases limited. Thus, the correlation of antennas appears as an aggravating fact in MIMO systems, and especially in large-scale MIMO systems. As the correlation between antennas increase, the channel between them gets more similar to each other and this is caused by decreasing the distance between antennas. Generally, the channels start to present correlation when the distance between antennas is lower then a half wave length ($\lambda/2$) (GOLDSMITH, 2005). With highly correlated channels, spatial diversity loss is expected and consequently, deterioration in system performance and capacity.

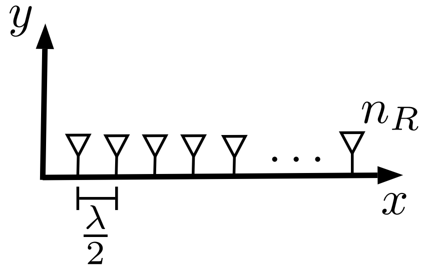


Figure 2.2: Uniform Linear Array (ULA)

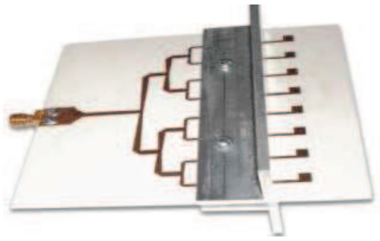


Figure 2.3: Example of an ULA implementation. Photo Source: (MANDEEP et al., 2010)

Commonly, the classical and simple configuration allowing us to analyze correlation for MIMO systems is the one where the distribution is organized as an uniform linear array(ULA) (ZELST; HAMMERSCHMIDT, 2002), Fig. 2.2, which simplifies the antenna model while allows a very good approximation for the correlation effect at MIMO systems with a low or moderate number of antennas. On the other hand, when the number of antennas are considerably increased, *i.e.* massive MIMO applications, another array structures are required in order to accommodate the transmit antenna elements in a feasible way. Different array possibilities and configurations have been proposed for the large MIMO channel; as a consequence, different correlated Massive MIMO channel models have arisen.

One of the first proposed antenna array arrangement is the uniform planar array (UPA), In Fig. 2.4, antenna elements are disposed in a two dimensional array. Accordingly to (BALANIS, 2005), planar array structures supply additional variables which can be used to control and shape the pattern array. Also, providing more versatility allowing more symmetrical patterns with lower side lobes at the total radiated power pattern. Additionally, they can be used as a scan mechanism for the main beam of the antenna towards any point in space.

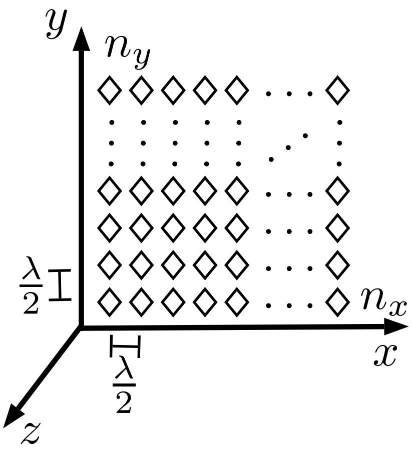


Figure 2.4: Uniform Planar Array (UPA)

Figure 2.5: Example of an UPA implementation. Photo Source: (GAO et al., 2011)

Another interesting antenna array structure is the MIMO cube, that is composed of a three dimension array (GETU; ANDERSEN, 2005); A cube is an attractive structure for building multiple antennas with low mutual coupling between antenna ports, because any two adjacent faces in a cube are perpendicular to each other. In addition, any two opposite faces in a cube have the farthest separation compared with other three dimensional structures with the same volume. An antenna cube, therefore, can take advantage of spacial and polarization orthogonality to implement a large number of antennas within a constrained volume.

In our work it will be considered two antenna array structures. The classic and simple ULA configuration will be taken as reference, and the UPA array, which can be identified as a promising candidate to compose the base station (BS) antenna structure in massive MIMO scenarios, will be considered as well. Those choices were made aiming to evaluate the UPA implementation impact at the BS, because theoretically the planar array structure has the potential to concentrate the downlink beamforming at the transmitter side. This characteristic can be showed through the Array Factor (AF), which is the factor by which the directivity function of an individual antenna must be multiplied to get the directivity of the entire array.

According to the antenna theory, the array factor of an ULA of N elements along the x-axis can be represented as (BALANIS, 2005):

$$AF_{\text{ula}} = \sum_{n=1}^{N} I_n e^{j(n-1)(kd \sin\theta \cos\phi)} \tag{2.3}$$

where, $\sin\theta\cos\phi$ is the directional cosine with respect to the x-axes, I_n is the amplitude excitation factor of each element and d is antenna element spacing. For simplification purposes, equation (2.3) can be written as:

$$AF_{\text{ula}} = \sum_{n=1}^{N} I_n e^{j(n-1)\psi} \tag{2.4}$$

where $\psi = (kd\sin\theta\cos\phi)$ and $k = \frac{2\pi}{\lambda}$.

According to (BALANIS, 2005), the AF in (2.4) can be expressed in an alternate, compact and closed form whose function and their distribution are more recognized. This is accomplished by multiplying both sides of (2.4) by $e^{j\psi}$, then we have:

$$(AF_{\text{ula}})e^{j\psi} = e^{j\psi} + e^{j2\psi} + e^{j3\psi} + \cdots + e^{j(N-1)\psi} + e^{jN\psi} \tag{2.5}$$

Subtracting (2.4) from (2.5) reduces to

$$(AF_{\text{ula}})\left(e^{j\psi} - 1\right) = \left(-1 + e^{jN\psi}\right) \tag{2.6}$$

which can also be written as

$$AF_{\text{ula}} = \left[\frac{e^{jN\psi} - 1}{e^{j\psi} - 1}\right] = e^{j[(N-1)/2]\psi}\left[\frac{\sin\left(\frac{N}{2}\psi\right)}{\sin\left(\frac{1}{2}\psi\right)}\right] \tag{2.7}$$

and according to (BALANIS, 2005), if we take as reference point the physical center of the array, the AF of (2.7) reduces to

$$AF_{\text{ula}} = \left[\frac{\sin\left(\frac{N}{2}\psi\right)}{\sin\left(\frac{1}{2}\psi\right)}\right] \tag{2.8}$$

In general lines, the array factor can be represented as a function of the number of elements, their geometrical disposal, corresponding magnitude, relative phases and element spacing. With those considerations, the AF should result in a simpler form if each element have identical amplitude, phase, and spacing related each other, which motivates a normalization in the AF expression, providing a fair comparison for different arrangements.

Substituting ψ into (2.8) and considering $I_n = 1$ we have the normalized version of AF for ULA, which is expressed as:

$$AF_{\text{ula}}(\theta,\phi) = \frac{1}{N}\frac{\sin(N\frac{kd\sin\theta\cos\phi}{2})}{\frac{kd\sin\theta\cos\phi}{2}} \tag{2.9}$$

which represents the directivity pattern of the ULA with $k = \frac{2\pi}{\lambda}$.

Now if $L = \frac{N}{2}$ antenna elements are placed in the x-axes and in the y-axes, a rectangular/planar array will be formed. Assuming again that all elements are equally spaced with intervals d_x and d_y in both axes, and all elements have the same amplitude excitation I_l, the UPA array factor can be represented as:

$$AF_{\text{upa}} = I_l \sum_{l=1}^{L} e^{j(l-1)(kd_x \sin\theta\cos\phi)} \sum_{l=1}^{L} e^{j(l-1)(kd_y \sin\theta\sin\phi)} \tag{2.10}$$

the normalized UPA array factor can be obtained as:

$$AF_{upa}(\theta,\phi) = \left\{ \frac{1}{L} \frac{\sin(L\frac{kd_x \sin\theta\cos\phi}{2})}{\frac{kd_x \sin\theta\cos\phi}{2}} \right\} \left\{ \frac{1}{L} \frac{\sin(L\frac{kd_y \sin\theta\sin\phi}{2})}{\frac{kd_y \sin\theta\sin\phi}{2}} \right\}, \tag{2.11}$$

where $k = \frac{2\pi}{\lambda}$.

The gain inside the array factor of a 5×5 UPA and 25 elements ULA has been plotted in Figure 2.6 aiming to identify their beam pattern and normalized power distribution over the azimuth and elevation directions, which directly impact on the array gain.

Further elaboration is depicted in Fig. 2.6, which is the case where the element spacing is defined as $d = 0.5\ \lambda$, and a frequency of 1GHz. The normalized beam pattern in polar coordinates and a cross section in the U-plane, where the normalized energy distribution is plotted as a function of the elevation angle variations projected onto the Cartesian plane, is also described. This coordinates projection over the Cartesian plane is known as UV mapping and it is commonly used in antenna theory, image processing and also 3D drawing. The UV mapping is a $\mathbb{R}^3$ to $\mathbb{R}^2$ projection which transform a 3D pattern on its 2D rectangular projection.

The $u-v$ coordinates can be easily derived from the ϕ and θ angles which are respectively the azimuth and elevation angles in spherical coordinates. The relationship between these two coordinates system is simply:

$$\begin{aligned} u &= \sin\theta\cos\phi \\ v &= \sin\theta\sin\phi \end{aligned} \tag{2.12}$$

the values of u and v satisfy the inequalities:

$$\begin{aligned} &-1 \le u \le 1 \\ &-1 \le v \le 1 \\ &u^2 + v^2 \le 1 \end{aligned} \tag{2.13}$$

An wider explanation regarding UV mapping can be found in the Appendix A.1.

Fig. 2.6 a) and d) represent the 3D pattern in terms of normalized power for

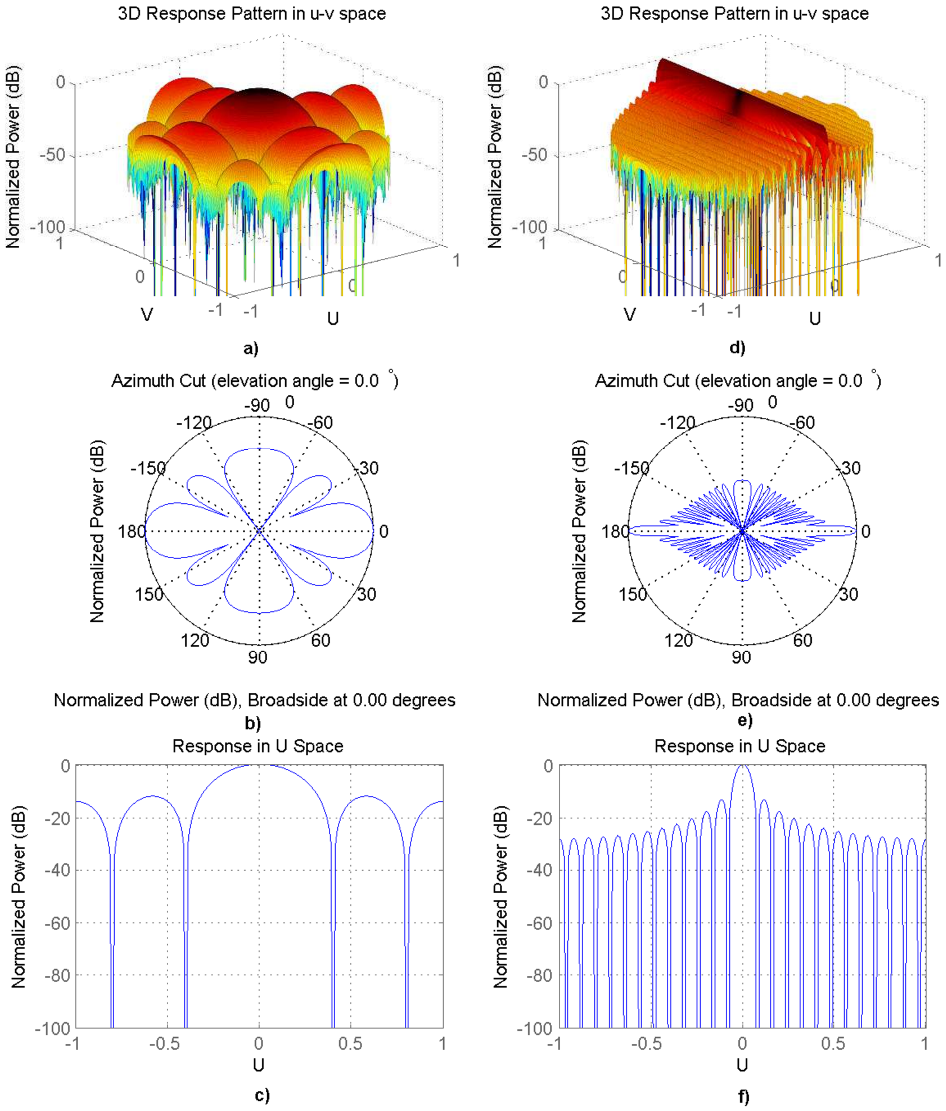


Figure 2.6: Array Factor for 0.5 λ element-spaced: [left] UPA of 5 × 5 elements; [right]: ULA with 25 elements

both UPA and ULA, respectively. It is simple to notice that the UPA structure presents a wider beam-width at the main lobe which provides larger beam gains that leads, at the BS, to lower transmit power and larger antenna coverage. On the other hand, the ULA power distribution is more heterogeneous presenting smaller beam-width and, a power concentration directly at the beam direction which provide a smaller coverage area with the total power transmitted. Another observed characteristic due to the UPA structure deployment, is that it provides larger side lobes when compared to the ULA side lobes, which implies in more transmit gain at the UPA side lobes benefiting the power distribution with this array structure. In order to manipulate the antenna beam pattern, especially the beam-width; there are two variables in the array structure that can modify the pattern distribution. The first is the number of antenna elements, which directly affects the beam-width, so that with increasing number of elements, the main lobe beam-width tends to become concentrated and the side lobes will suffer from higher attenuation. Another parameter impacting the beam pattern is the antenna element spacing; by decreasing the space between elements the beam-width become wider, providing higher normalized power along the array, directly impacting in less attenuation at the transmitted signal.

To exemplify those previous concepts and manipulations, we presented two modifications on the previous uniform arrays. Firstly, Figure 2.7 depicts the normalized power distributions along the azimuth-elevation directions, as well as u-v directions, relative to a larger array structure of 8×8 UPA and 64-element ULA both with the same as the previous $0.5\ \lambda$ element-spaced. Secondly, Figure 2.8 indicates the normalized power distribution for the same array dimension as in Fig. 2.6, i.e., 5×5 UPA and 25 elements ULA, but now with a small spacing between antenna-elements, i.e., $0.25\ \lambda$ element-spaced.

Thereby, the performance of MIMO systems equipped with both UPA and ULA arrays can be analyzed in a comparative meanings, aiming to determinate which is the best scenario to apply the planar array structure in comparison to the classical ULA approach. The following sections provide mathematical expressions which represent the spacial correlation function for both studied antenna array structures, also providing a comparison between the simplified version and the full geometrical correlation matrix for each structure.

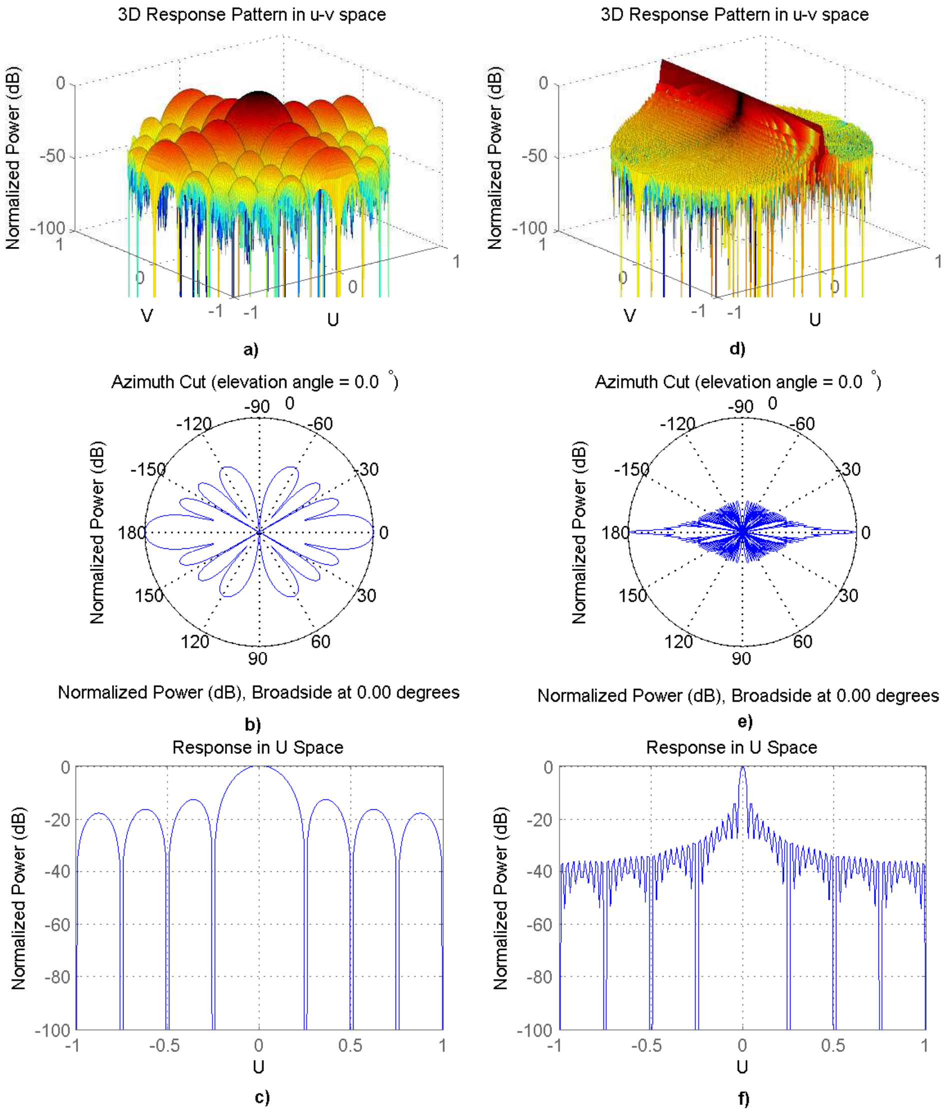


Figure 2.7: Array Factor for 0.5 λ element-spaced: [left]: 8×8 UPA; [right]: ULA with 64 elements

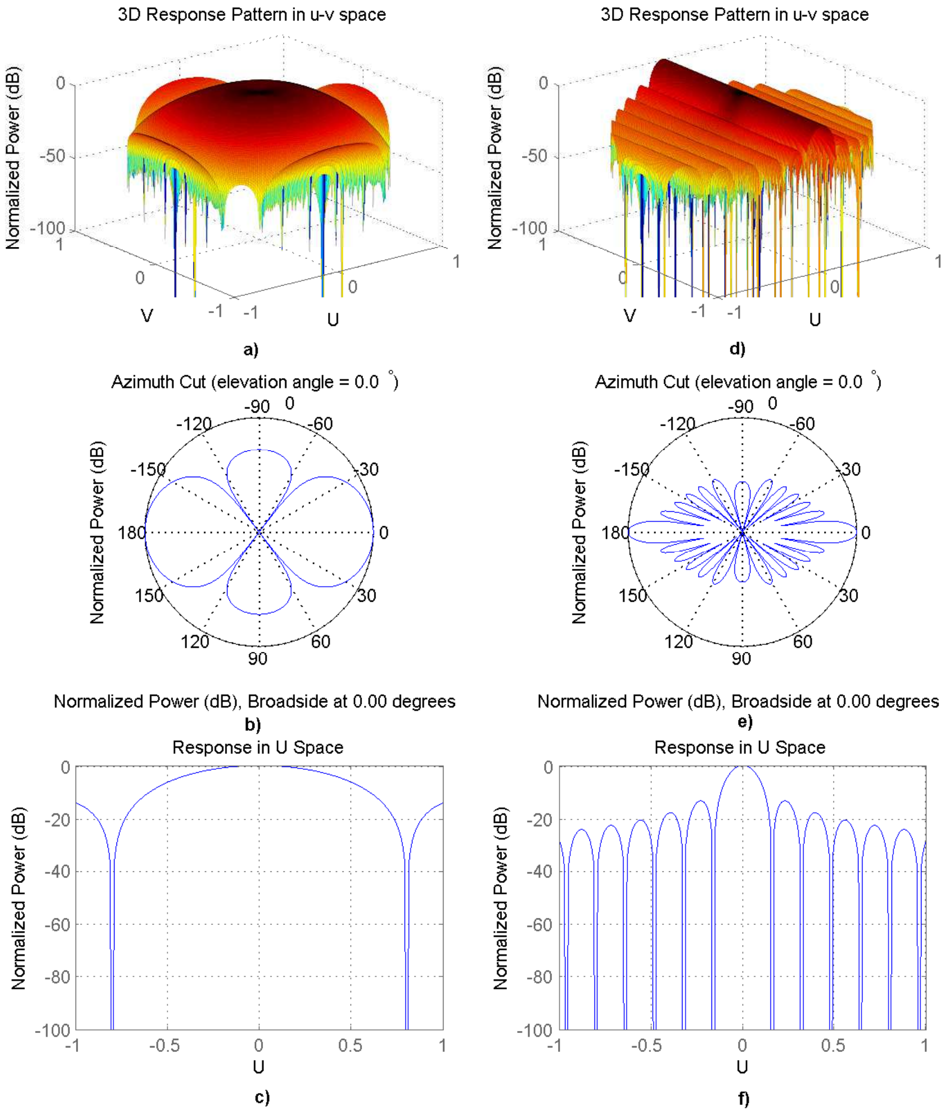


Figure 2.8: Array Factor for 0.25 λ element-spaced: [left] 5×5 UPA; [right]: ULA 25 elements

2.1.2 Uniform Linear Array

Several MIMO channel correlation models were proposed in the last decades; one of the most important yet simple class of MIMO channel models is the one that assume independence among the correlation between transmit antennas (TX) and receive antennas (RX) (and vice versa). Hence, a spatially correlated MIMO fading channel is decently modeled by flat Rayleigh distribution and the correlation among antennas elements will be determined over the Kronecker's correlation model (ZELST; HAMMERSCHMIDT, 2002), as follows:

$$\mathbf{H} = \sqrt{\mathbf{R}_{H,Rx}}\mathbf{H}'\sqrt{\mathbf{R}_{H,Tx}} \tag{2.14}$$

where $\mathbf{H}'(n_R \times n_T)$ is the uncorrelated MIMO channel which is represented with independent, identically distributed (i.i.d.) complex Gaussian with zero-mean and unitary variance, $g_{ij} \sim \mathcal{CN}\{0, 1\}$. The correlation matrices $\mathbf{R}_{H,Tx}(n_T \times n_T)$ and $\mathbf{R}_{H,Rx}(n_R \times n_R)$ denote the spatial channel correlation held among the transmitter and receiver side, respectively. Each element of those matrices are represented, in terms of a normalized correlation index ρ, by:

$$\begin{cases} r_{H,Rx\,ij} = \rho^{(i-j)^2} \\ r_{H,Tx\,ij} = \rho^{(i-j)^2} \end{cases} \tag{2.15}$$

Note that matrix $\mathbf{H}'$ in (2.14) is similar to matrix $\mathbf{H}$. Hence, for the rest of this work we assume that the Tx and Rx antenna elements are equidistant, with identical number of antennas $n_T = n_R = n$ and consequently the same correlation matrix $\mathbf{R}_{H,Rx} = \mathbf{R}_{H,Tx} = \mathbf{R}_H$, that is represented as:

$$\mathbf{R}_H = \begin{bmatrix} 1 & \rho & \rho^4 & \dots & \rho^{(n-1)^2} \\ \rho & 1 & \rho & \dots & \vdots \\ \rho^4 & \rho & 1 & \dots & \rho^4 \\ \vdots & \vdots & \vdots & \ddots & \rho \\ \rho^{(n-1)^2} & \dots & \rho^4 & \rho & 1 \end{bmatrix} \tag{2.16}$$

Also note that a fully uncorrelated channel means $\rho = 0$, while an entirely correlated scenario results in $\rho = 1$.

2.1.3 Uniform Planar Array

Traditionally, the MIMO systems adopt ULA setup as the simplest and standard structure. But considering the used space limitation, the ULA setup is not suitable for large-scale antenna arrays. Hence, for massive MIMO applications the necessity to adopt a two-dimensional array structure, such as UPA, is essential. A correlation matrix for the UPA structure was proposed by (LEVIN; LOYKA, 2010); In this paper a multidimensional array correlation structure is constructed for the UPA, based on a Kronecker product of two ULA correlation matrices. More specifically, considering a UPA constructed]with isotropic antenna elements lying on the XY plane with n_x and n_y antenna elements along x and y coordinates, respectively, so that $n_r = n_x \cdot n_y$.

Moreover, we can assume an approximation in which the correlation between elements along x coordinate does not depend on y and is given by matrix $\mathbf{R}_{H,x}$, and the correlation along y coordinate does not depend on x and is given by matrix $\mathbf{R}_{H,y}$. The following Kronecker-type approximation of the UPA correlation matrix is proposed by (LEVIN; LOYKA, 2010):

$$\begin{aligned} \mathbf{R}_{H,r} &= \mathbf{R}_{H,x} \otimes \mathbf{R}_{H,y} \\ \mathbf{R}_{H,r} &= \begin{bmatrix} r_{H,x\;1,1}\mathbf{R}_{H,y} & \dots & r_{H,x\;1,n_T}\mathbf{R}_{H,y} \\ \vdots & \ddots & \vdots \\ r_{H,x\;n_R,1}\mathbf{R}_{H,y} & \dots & r_{H,x\;n_R,n_T}\mathbf{R}_{H,y} \end{bmatrix} \end{aligned} \tag{2.17}$$

where $\otimes$ denotes the Kronecker product. The equation (2.17) indicates that the UPA correlation matrix $\mathbf{R}_{H,r}$ is the Kronecker product of two ULA correlation matrices $\mathbf{R}_{H,x}$ and $\mathbf{R}_{H,y}$, which are Toeplitz. Therefore, according to the authors, even tough $\mathbf{R}_{H,r}$ may not be a Toeplitz matrix, its approximation (2.17) has a Toeplitz structure. According to (LI et al., 2013), this approximation model is reasonably accurate, allowing the usage of the well-developed theory of Toeplitz matrices for the analysis of multidimensional antenna arrays.

Remembering that Toeplitz matrix or diagonal-constant matrix, is a matrix which each descending diagonal element from left to right is constant. For instance, a $n \times n$ Toeplitz matrix $\mathbf{A}$ is defined as (BAREISS, 1969):

$$A_{i,j} = A_{i+1,j+1} = a_{i-j}, \tag{2.18}$$

where

$$\mathbf{A} = [\mathbf{A}]_{i,j} = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \dots & \dots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \ddots & & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \dots & \dots & a_2 & a_1 & a_0 \end{bmatrix}$$

As a consequence, the impact of this structure is seen in a matrix equation of the form:

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{2.19}$$

which is called a Toeplitz system. If $\mathbf{A}$ is an $n \times n$ Toeplitz matrix, then the system has only $2n - 1$ degrees of freedom, rather than n^2; therefore, it produces an easier system to solve.

2.1.4 Geometrical Correlation Model

Another perspective to derive the correlation expression is made through the geometric properties of the problem. In (YING et al., 2014) the UPA analytical expression was derived based on a 3D channel model. The spacial correlation function was derived in a downlink transmission, where the BS is equipped with N_v vertical antenna elements spaced by d_1 wavelengths, and N_h horizontal antennas with d_2 wavelength spacing separation, as sketched in Figure 2.9.

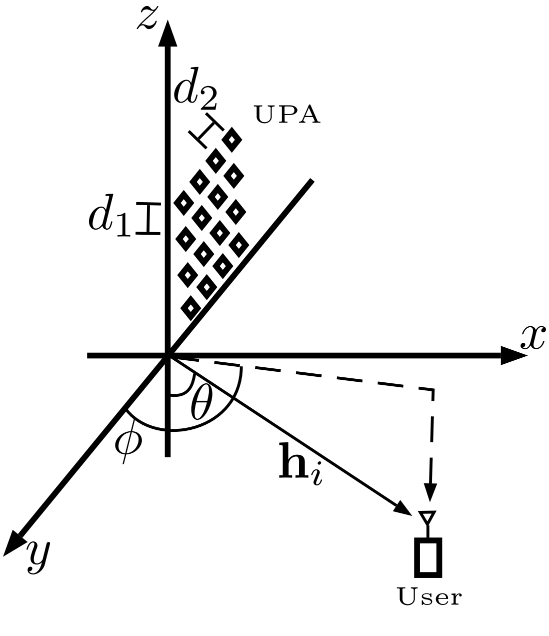


Figure 2.9: 3D Channel model adopted to derive the spacial correlation function. ϕ is the azimuth angle, and θ is the elevation angle.

The (a, b)-th antenna element denotes the antenna in the a-th row and the b-th column of the UPA, so the channel from the (a, b)-th element in the transmitter to the receiver is associated with the $b + N_h(a - 1)$-th element of $\mathbf{h}_i$, which is the channel vector related to the i-th fading block. As modeled by (ZHAO et al., 2016), the spacial correlation matrix $\mathbf{R}_h$ for the UPA is composed by the correlation element between the (a,b)-th and (p,q)-th antennas, given as:

$$[\mathbf{R_h}]_{(a,b),(p,q)} = \frac{D_1}{\sqrt{D_5}} e^{-\frac{D_7+(D_2(\sin\phi)\sigma)^2}{2D_5}} e^{j\frac{D_2D_6}{D_5}} \tag{2.20}$$

with the variables defined by:

$$\begin{aligned}
D_1 &= e^{j\frac{2\pi d_1}{\lambda}(p-a)\cos\theta} e^{-\frac{1}{2}(\xi\frac{2\pi d_1}{\lambda})^2(p-a)^2\sin^2\theta}, \\
D_2 &= \frac{2\pi d_2}{\lambda}(q-b)\sin\theta, \\
D_3 &= \xi\frac{2\pi d_2}{\lambda}(q-b)\cos\theta, \\
D_4 &= \frac{1}{2}(\xi\frac{2\pi}{\lambda})^2 d_1 d_2 (p-a)(q-b)\sin(2\theta), \\
D_5 &= (D_3)^2((\sin\phi)\sigma)^2+1, \\
D_6 &= D_4((sin\phi)\sigma)^2+\cos\phi, \\
D_7 &= (D_3)^2\cos^2\phi-(D_4)^2((\sin\phi)\sigma)^2-2D_4\cos\phi,
\end{aligned} \tag{2.21}$$

where λ is the carrier wavelength, ϕ is the azimuth angle-of-departure (AoD), θ is the elevation AoD, while σ is the standard deviation of horizontal AoD, and ξ is the standard deviation of vertical AoD. Finally, $\mathbf{R}_h$ is a $n_T \times n_T$ matrix and $[\mathbf{R}_h]_{(a,b),(p,q)}$ is the element at the $b+N_v(a-1)$-th row and the $q+N_h(p-1)$-th column.

Considering the above analytical expressions, it is clear that term D_1 is only associated with the elevation angle, containing only $(p-a)$ terms, while D_2, D_3 and D_5 are azimuth related containing only the $(q-b)$ terms. Variables D_4, D_6 and D_7 have the cross term $(p-a)(q-b)$, containing both elevation and azimuth correlations. However, D_6 and D_7 are functions of D_4. As proposed by (YING et al., 2014), if term D_4 could be neglected, i.e, $D_4 = 0$, the correlation term $[\mathbf{R}_h]_{(a,b),(p,q)}$ can be written as a simple product of elevation and azimuth correlations. Therefore, if $D_4 = 0$, the correlation matrix is separable:

$$\mathbf{R}_h = \mathbf{R}_{az} \otimes \mathbf{R}_{el}, \tag{2.22}$$

where the elements of elevation correlation matrix are expressed as:

$$[\mathbf{R}_{el}]_{(a,b)} = e^{j\frac{2\pi d_1}{\lambda}(p-a)\cos\theta} e^{-\frac{1}{2}(\xi\frac{2\pi d_1}{\lambda})^2(p-a)^2\sin^2\theta} \tag{2.23}$$

and the correlation elements in the azimuth direction are:

$$[\mathbf{R}_{az}]_{(p,q)} = \frac{1}{\sqrt{D_5}} e^{-\frac{D_3^2\cos^2\phi}{2D_5}} e^{j\frac{D_2\cos\phi}{D_5}} e^{-\frac{(D_2(\sin\phi)\sigma)^2}{D_5}} \tag{2.24}$$

It is demonstrated by (YING et al., 2014) that the Kronecker correlation model has very similar eigenvalues distribution as the correlation matrix, and thus is a good approximation for the original UPA correlation matrix.

The equation that express the ULA spacial correlation function is derived at (BUEHRER, 2002), and is defined as:

$$[\mathbf{R}_{\text{ula}}]_{(i,j)} = e^{j\frac{2\pi d}{\lambda}(i-j)\sin\theta} e^{-\frac{1}{2}(\xi\frac{2\pi d}{\lambda})^2(i-j)^2\cos^2\theta} \tag{2.25}$$

where d is the distance between antenna elements. This expression is similar to the elevation correlation, $[\mathbf{R}_{\text{el}}]_{(a,b)}$, in the UPA structure. From the previous expressions it is easy to conclude that the ULA and UPA models provided in subsection 2.1.2 and 2.1.3, respectively, are the simplified expressions to the above geometrical models.

As we have derived the correlated channel expressions for both antenna array structures, now we will introduce several MIMO detectors that will be deployed to detect the transmitted symbols from the BS to the mobile terminal (MT) (downlink); hence, the BS antenna correlation effect plays an important role on the system capacity/reliability reduction.

2.2 MIMO Detection Techniques

The present section recall the commonly approaches for MIMO detectors techniques, going through the maximum-likelihood (ML), sphere decoder (SD), zero-forcing (ZF) and minimum mean squared error (MMSE). Also, it will be provided a succinct discussion over the application of two techniques applicable to the MIMO detection context, *i.e.,* the sucessive interference cancellation (SIC) and lattice reduction (LR) The knowledge of each detector procedure is very important, in order to evaluate complexity and BER performance analysis.

2.2.1 Maximum Likelihood (ML)

The maximum-likelihood detector perform an exhaustive search over the whole set of possibles symbols $s \in \mathcal{S}^{n_T}$, of size M^{n_T}, in order to decide in favor of the one that minimizes the Euclidean distance, and therefore the lowest error, between the received signal $\mathbf{x}$ and the reconstructed signal $\mathbf{Hs}$:

$$\widehat{\mathbf{s}} = \underset{s\in\mathcal{S}^{n_T}}{\operatorname{argmin}} \left\|\mathbf{x} - \mathbf{Hs}\right\|^2. \tag{2.26}$$

It is well known that the ML detector ensures the lowest BER performance in all the spectrum of MIMO detectors, but the search complexity grows exponentially according to the number of antennas and the number of symbols. If we consider a M-ary modulation with n_T transmit antennas, each one transmitting a different symbol in a distinct time-slot system, the order of combinations is given by M^{n_T} This, way it becomes impractical in cases where the constellation order and number of antennas are considerably increased; for example, if $M = 16$ and $n_T = 8$, the number of candidates to be evaluated becomes incredibly large, more specifically, over $\approx$ 4 billion of candidate-symbols.

2.2.2 Sphere Decoder (SD)

Pursuing a reduction in the ML complexity, a similar approach has been proposed, namely the sphere-decoder detector, that searches only the candidates bounded in the hypersphere of radius d, causing it performance to be highly related to the SNR:

$$d^2 < \|\mathbf{x} - \mathbf{H}\mathbf{s}\|^2 \tag{2.27}$$

If the search radius were too high, the SD complexity get close to the ML one. In contrast, if the search radius is set too small, no candidate will be chosen upon hypersphere. Moreover, in order to obtain candidate-solution points to perform the sphere detection is necessary rewrite the eq.(2.1) evaluating the QR decomposition at the channel matrix, such that $\mathbf{H} = \mathbf{QR}$. The QR decomposition will ensure an orthogonal matrix $\mathbf{Q}$, where $\mathbf{I} = \mathbf{Q}^H\mathbf{Q}$, and an upper triangular matrix $\mathbf{R}$, then for detection purposes both matrices will have convenient properties. The procedure is performed as follows:

$$\mathbf{y} = \mathbf{Q}^H\mathbf{x} \;=\; \mathbf{Q}^H\mathbf{Q}\mathbf{R}\mathbf{s} + \mathbf{Q}^H\mathbf{n} \;=\; \mathbf{R}\mathbf{s} + \mathbf{n}'. \tag{2.28}$$

Since the matrix $\mathbf{Q}$ is orthogonal, the statistical properties of the additive noise, $\mathbf{n}^{'}$, remains unaltered and no noise increment is foreseen. Moreover, as $\mathbf{R}$ is an upper triangular matrix it enables noise estimation for each antenna independently. Hence, the points inside the hyper-sphere can be determined layer-by-layer, starting from the last row of $\mathbf{R}$, by evaluating:

$$d^2 < \|\mathbf{y} - \mathbf{R}\mathbf{s}\|^2 \tag{2.29}$$

Considering $\mathbf{R} = \begin{bmatrix}\mathbf{r}_1^T & \mathbf{r}_2^T & \mathbf{r}_3^T & \dots & \mathbf{r}_{n_T}^T\end{bmatrix}$, the noise norm is given by:

$$\left\|\mathbf{n}'\right\|^2 = \left\|\mathbf{y} - \mathbf{R}\mathbf{s}\right\|^2 = \sum_{k=1}^{n_T} \left|y_k - \mathbf{r}_k\mathbf{s}\right|^2,\ k = 1, 2, \dots, n_T. \qquad (2.30)$$

In fact, eq. (2.30) shows that the noise norm is the summation of each layers noise independently. This way, the noise norm can be updated as the symbols are tested in each layer, which avoids the evaluation of the estimated noise for every symbol combination.

A beneficial feature that emerges from the structure of the detection problem in (2.29) is that, due to the upper triangular properties of the $\mathbf{R}$ matrix, the tree search algorithm scan the symbol vector backwards, starting from the last antenna symbol to the first one, testing all candidates symbols recursively and independently, contrarily the ML. As this layer-by-layer procedure follows the radius restriction defined in (2.30), by finishing the SD detection, the most likely symbol-vector bounded by the hypersphere of radius d is the solution.

2.2.3 Zero-Forcing (ZF)

The Zero-Forcing detector, is a simple linear MIMO receiver, with low computational complexity. It is designed to suppress channel interference by multiplying the signal received by the Moore-Penrose pseudo-inverse of the channel matrix:

$$\mathbf{H}^{\dagger} = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H. \qquad (2.31)$$

With that, the estimated signal from the detector can be determined by:

$$\widehat{\mathbf{s}} = \mathbf{H}^{\dagger}\mathbf{x} = \mathbf{s} + \mathbf{H}^{\dagger}\mathbf{n} \qquad (2.32)$$

Considering a scenario without noise, the ZF detector has a identical ML performance, due to all channel interference suppression. Otherwise, in noise scenarios, ZF leads to noise enhancement. That problem inhibits the performance of the ZF algorithm due to ill-conditioned $\mathbf{H}$ matrices, i.e near to linearly dependent columns condition, which after the matrix inversion in (2.32) leads to enhancements in the thermal noise variance in $\widehat{s}$ when compared to $\mathbf{y}$ (CIRKIC, 2014).

2.2.4 Minimum Mean Squared Error (MMSE)

The MMSE detector can be seen as a particularly useful extension of the ZF detection, which by taking the noise and signal statistics into account the detector is able to improve the overall MIMO detection performance. The procedural difference to MMSE is that, instead of the pseudo-inverse, MMSE uses:

$$\mathbf{H}^{\dagger} = \left(\mathbf{H}^{H}\mathbf{H} + \sigma_n^2\mathbf{I}_{n_T}\right)^{-1}\mathbf{H}^{H}. \tag{2.33}$$

And the solution of the MMSE detector is:

$$\widehat{\mathbf{s}} = \left(\mathbf{H}^{H}\mathbf{H} + \sigma_n^2\mathbf{I}_{n_T}\right)^{-1}\mathbf{H}^{H}\mathbf{x}. \tag{2.34}$$

In another perspective, the MMSE detection can be fulfilled as:

$$\widehat{\mathbf{s}} = \underline{\mathbf{H}}^{\dagger}\underline{\mathbf{x}} \;=\; \mathbf{s} + \underline{\mathbf{H}}^{\dagger}\mathbf{n} \tag{2.35}$$

It is easy to note that the equation above has the same structure of (2.32), but the vector signal and the received vector are extended and respectively given by:

$$\underline{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \sigma_n\mathbf{I}_{n_T} \end{bmatrix}, \qquad \underline{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{0}_{n_T\times 1} \end{bmatrix} \tag{2.36}$$

The extended matrix model is more complex than the approach given in (2.34), but this model is required on successive interference cancellation (SIC) and can be used on lattice-reduction in order to achieve performance improvements (WUBBEN et al., 2004).

2.2.5 Successive Interference Cancellation (SIC)

The SIC detection technique can be performed by evaluating the QR decomposition of the matrix $\mathbf{H}$, which was addressed in section 2.2.2. An important observation is due to the fact that for ZF detectors, the QR decomposition should be executed on $\mathbf{H}$, while for MMSE cases it is applied on the matrix $\underline{\mathbf{H}}$. As we already know, the MIMO detection aided QR can be performed as follows:

$$\widehat{\mathbf{s}} = \mathbf{Q}^{H}\mathbf{x} = \mathbf{R}\mathbf{s} + \mathbf{Q}^{H}\mathbf{n}. \tag{2.37}$$

Since $\mathbf{Q}$ is an orthogonal matrix, when multiplied by the noise term, $\mathbf{Q}^{H}\mathbf{n}$, the statistical properties of the additive noise remains unaltered. As matrix $\mathbf{R}$ has an upper triangular structure, the n-th element of $\widehat{\mathbf{s}}$ is completely free of inter-antenna interference, and can be used to correctly estimate the received signal after the addition of an appropriate scale of $\text{f}rac1r_{ii}$, where, $i = n_T$ (WUBBEN et al., 2003). Hence, the linear system can be solved upwards by:

$$\widehat{\mathbf{s}} = \begin{cases} \dfrac{x_i}{r_{ii}}, & i = n_T \\ \dfrac{1}{r_{ii}}\left(x_i - \displaystyle\sum_{k=i+1}^{n_T} r_{ik}\widehat{s}_k\right), & i = n_T - 1, \ldots, 3, 2, 1 \end{cases} \tag{2.38}$$

It is important to note that each symbol must pass to the slicer before following to the interference cancellation and this step is applied at each symbol detection. The slicing is extremely important in order to provide a proper interference cancellation and attain fully detector performance. Hence, if we assume that the estimated symbol in a determined layer is correct, the furthest symbols can be detected as if there were no previous layers, in a simple equivalent system. However, if an error occur on the first layers, it will propagate until the end of the algorithm, resulting in performance deterioration.

2.2.6 Ordered Successive Interference Cancellation (OSIC)

Improvements related to the BER performance of SIC can be attained through a suitable ordering scheme (WUBBEN et al., 2003), preventing error propagation during interference cancellation computation. The ordering criteria has it's focus on minimizing the columns norm of $\mathbf{Q}$, which cause the detection process to start from the highest normalized power symbol to the weakest one.

The sorted decomposition can be expressed as:

$$\mathbf{H\Pi} = \mathbf{QR} \tag{2.39}$$

where $\mathbf{\Pi}$ is a permutation matrix that allows symbols reordering after executing the SIC detection. It is important to notice that the detection proceeds as a conventional SIC, is represented in (2.38). The only difference lay on the final step, where, by the end of the detection scheme, the reordering step is followed by multiplying the detected symbols vector with the permutation matrix.

The sorted QR decomposition (SQRD) is formalized at the pseudocode in Algorithm 2.1. The main difference between this algorithm and the conventional QR decomposition, lay at the lines 2 and 3 of the algorithm 2.1. Accordingly to (KOBAYASHI; ABRÃO, 2016), if these lines are ignored, the algorithm will perform a traditional QR decomposition with the Gram-Schmidt approach. Also, these lines do not carry out high complexity operation, causing the ordering complexity to be essentially negligible. For the rest of this work, this decomposition plus the detection scheme will be referred as ordered successive interference cancellation

(OSIC).

Algorithm 2.1 Sorted QR decomposition (KOBAYASHI; ABRÃO, 2016)

```
Input: $\mathbf{Q} = \mathbf{H}$, $\mathbf{R} = \mathbf{0}$, $\mathbf{\Pi} = \mathbf{I}_{n_T}$
Output: $\mathbf{Q}, \mathbf{R}$
for $i = 1$ to $n_T$ do
    $k = \underset{j=i \text{ to } n_T}{\text{argmin}} \|\mathbf{q}_j\|^2$
    Exchange columns $i$ and $k$ in $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{\Pi}$
    $r_{ii} = |\mathbf{q}_i|$
    $\mathbf{q}_i = \mathbf{q}_i / r_{ii}$
    for $j = i + 1 : n_T$ do
        $r_{ij} = \mathbf{q}_i^H \mathbf{q}_j$
        $\mathbf{q}_j = \mathbf{q}_j - r_{ij} \mathbf{q}_i$
    end for
end for
```

Recently, Kobayashi e Abrão (2016) have proved that the sorted QR decomposition based on the Gram-Schimid method is unable to operate satisfactorily at high SNR regime. It was also showed that SQRD algorithm based on Gram-Schimidt's was incapable to promote the orthonormalization of matrix channel $\mathbf{H}$ when the channel is highly correlated, which can make the $\mathbf{Q}$ matrix do not achieve the orthogonality and failing the OSIC requirements. Hence, the authors proposed a change in the norm update in the classic algorithm and numerically prove the stabilization of the OSIC in high SNR regime. Finally, the Algorithm 2.1 is the modified version that can achieve better BER performance in the high SNR regime, while the same performance of the classic algorithm was held in low and medium SNR regions.

2.2.7 Lattice Reduction (LR) aided MIMO Detector

As already mentioned, if the channel matrix has a strongly spacial correlation characteristic or even a strong line-of-sight (LOS) component, the channel matrix become ill conditioned; which disrupts the detection process and mainly deteriorates the MIMO system performance. An ill conditioned matrix causes a narrowing on the symbol decision regions, which makes the detection more vulnerable to even the smallest amount of noise. Hence, to circumvent this problem, we aim to turn the channel matrix as near-orthogonal as possible, looking for improve the MIMO detection process with a manageable complexity increase.

The LR can be efficiently carried out through the LLL algorithm, which was proposed by Lenstra-Lenstra-Lovaz in (LENSTRA; LENSTRA; LOVÁSZ, 1982). However, for this entire work, is recommended the usage of the Algorithm 2.2 to ensure the complex LR, which is known to be more robust for MIMO detection, furthermore presents less computational complexity (MA; ZHANG, 2008).

Algorithm 2.2 The Complex LLL Algorithm (MA; ZHANG, 2008) (Using MATLAB Notation)

```
Input: $\mathbf{H}$
Output: $\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}, \mathbf{T}$
$\delta = 0.75$
$m$ = columns number of $\mathbf{H}$
$\mathbf{T} = \mathbf{I}_m$
$\left[\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}\right] = \mathrm{QR}(\mathbf{H})$
$k = 2$
while $k \leq m$ do
    for $n = k - 1$ to $1$ do
        $u = \left\lfloor \frac{\tilde{\mathbf{R}}(n,k)}{\tilde{\mathbf{R}}(n,n)} \right\rceil$
        if $u \neq 0$ then
            $\tilde{\mathbf{R}}(1:n,k) = \tilde{\mathbf{R}}(1:n,k) - u\tilde{\mathbf{R}}(1:n,n)$
            $\mathbf{T}(:,k) = \mathbf{T}(:,k) - u\mathbf{T}(:,n)$
        end if
    end for
    Swap the $(k-1)$th and $k$th columns in $\tilde{\mathbf{R}}$ and $\mathbf{T}$
    if $\delta \left|\tilde{\mathbf{R}}(k-1,k-1)\right|^2 > \left|\tilde{\mathbf{R}}(k,k)\right|^2 + \left|\tilde{\mathbf{R}}(k-1,k)\right|^2$ then
        $\alpha = \frac{\tilde{\mathbf{R}}(k-1,k-1)}{\left\|\tilde{\mathbf{R}}(k-1:k,k-1)\right\|_2}$
        $\beta = \frac{\tilde{\mathbf{R}}(k,k-1)}{\left\|\tilde{\mathbf{R}}(k-1:k,k-1)\right\|_2}$
        $\Theta = \begin{bmatrix} \alpha^\star & \beta \\ -\beta & \alpha \end{bmatrix}$
        $\tilde{\mathbf{R}}(k-1:k,k-1:m) = \Theta\tilde{\mathbf{R}}(k-1:k,k-1:m)$
        $\tilde{\mathbf{Q}}(:,k-1:k) = \tilde{\mathbf{Q}}(:,k-1:k)\,\Theta^H$
        $k = \max(k-1,2);$
    else
        $k = k + 1$
    end if
end while
```

Basically, for detection purposes, the LLL algorithm decomposes the MIMO channel into a new base in a reduced domain:

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{T}, \tag{2.40}$$

where $\tilde{\mathbf{H}}$ is the reduced basis, offering improved properties regarding near-orthogonality when compared with the former $\mathbf{H}$, while $\mathbf{T}$ is a unimodular matrix with two properties: $\det(|\mathbf{T}|) = \pm 1$, and $\mathbf{T} \in \{\mathbb{Z}^{n_R \times n_T} + j\mathbb{Z}^{n_R \times n_T}\}$.

The new matrix $\tilde{\mathbf{H}}$ has better numerical conditioning properties, so the decision boundaries are enlarged and the noise amplification effect is reduced, which allow performance gain in the signal detection. The idea behind the LR-aided MIMO detection is to detect the symbols in the LR domain, so it is desirable to rewrite the MIMO transmit equation in the LR domain:

$$\begin{aligned}\mathbf{x} &= \mathbf{H}\mathbf{s} + \mathbf{n} \\ &= (\mathbf{H}\mathbf{T})\left(\mathbf{T}^{-1}\mathbf{s}\right) + \mathbf{n} \\ &= \widetilde{\mathbf{H}}\mathbf{z} + \mathbf{n}\end{aligned} \tag{2.41}$$

Applying the reworked system model, the detection scheme under LR domain can be performed by any linear MIMO detection technique, such as ZF and MMSE, optionally combined with the SIC or OSIC techniques. However, is extremely important to properly quantize the symbols in the reduced domain, this is performed through:

$$\widehat{\mathbf{z}} = \left\lfloor \frac{\widetilde{\mathbf{z}} - \beta'\mathbf{T}^{-1}\mathbf{1}_{1\times n_T}}{2} \right\rceil + \beta'\mathbf{T}^{-1}\mathbf{1}_{n_T\times 1} \tag{2.42}$$

where $\lfloor\cdot\rceil$ represents the round operator, $\mathbf{1}_{n_T\times 1}$ is an all ones column vector, $\widetilde{\mathbf{z}}$ is the estimated symbols after a MIMO detection strategy and β' is a constant controlled by the modulation order (MILFORD; SANDELL, 2011). For transmissions schemes that uses M-QAM modulation, we set $\beta' = (1+i)$ and for binary phase shift keying (BPSK) modulation we set $\beta' = 1$.

2.2.8 LR aided Linear Equalization

When linear detectors are taking into account, the equalization in the LR domain can be done in the exact same way as in sections 2.2.3 and 2.2.4, the only difference occurs in the quantization. Thus, for the ZF aided LR case, the solution is given as follows:

$$\begin{aligned}\widetilde{\mathbf{z}} &= \widetilde{\mathbf{H}}^{\dagger}\mathbf{x} \\ &= \mathbf{z} + \widetilde{\mathbf{H}}^{\dagger}\mathbf{n}.\end{aligned} \tag{2.43}$$

On the other hand, for the MMSE it is recommended the usage of the extended matrix due to its better performance. (WUBBEN et al., 2004). Thus, the LLL will be executed over the extended channel matrix, i.e,

$$\underline{\widetilde{\mathbf{H}}} = \underline{\mathbf{H}}\underline{\mathbf{T}}. \tag{2.44}$$

Then, the MMSE solution in the LR domain is given as:

$$\widetilde{\mathbf{z}} = \underline{\widetilde{\mathbf{H}}}^{\dagger}\underline{\mathbf{x}} \tag{2.45}$$

According to (WUBBEN et al., 2004, 2003), the noise term still corrupts the symbols on the LR domain and a proper decision must be made at the LR domain symbol. As the LR operations consists in scaling and shifting the lattice points, it is necessary to include a re-scaling and re-shifting operations, that is given by the LR quantization in eq (2.42).

Finally, the last step of the LR-assisted MIMO detection consists of converting the estimated symbol vector of the LR domain to the original signal space:

$$\widehat{\mathbf{s}} = \mathbf{T}\widehat{\mathbf{z}}. \tag{2.46}$$

2.2.9 LR and OSIC aided Linear Equalization

To perform the LR and OSIC aided detection it is necessary to realize some modifications in the procedure approach. Basically, it is necessary to change the QR decomposition in the Algorithm 2.2 to the sorted QR version that is described in Algorithm 2.1.

With that change, the equalization can be described in the LR domain as a upper triangular linear system as follows:

$$\begin{aligned}\mathbf{y} &= \widetilde{\mathbf{Q}}^H\mathbf{y}\\ &= \widetilde{\mathbf{Q}}^H\left(\widetilde{\mathbf{H}}\mathbf{z}+\mathbf{n}\right)\\ &= \widetilde{\mathbf{Q}}^H\left(\widetilde{\mathbf{Q}}\widetilde{\mathbf{R}}\mathbf{\Pi}^{-1}\mathbf{z}+\mathbf{n}\right)\\ &= \widetilde{\mathbf{R}}\mathbf{\Pi}^{-1}\mathbf{z}+\mathbf{n}\end{aligned} \tag{2.47}$$

From this point the SIC detection is proceeded, as described in section 2.2.5. Finally, the symbols in the LR domain are quantized, re-ordenated and converted to the original domain.

$$\widehat{\mathbf{s}} = \mathbf{\Pi}\mathbf{T}\widehat{\mathbf{z}} \tag{2.48}$$

The procedure described above can be applied to the ZF and MMSE equalization, is important to emphasize the usage of the extended channel matrix in the MMSE case. Besides, the ordering scheme can be by-passed, resulting in conventional QR decomposition which generates the LR and SIC aided linear detection.

2.3 Performance Analysis

Next, the simulated BER performance of the previously discussed MIMO detectors have been compared. The system analysis is performed as a data transmission from the BS, with n_T antennas to a MT equipped with n_R antennas, i.e, the downlink scenario. In order to have a rightful comparison among different MIMO transmission set-ups, even when particular modulation order and number of antennas are applied, all performances will be examined under a normalized SNR in terms of bit energy (E_b), as:

$$\frac{E_b}{N_0} = \frac{\text{SNR}}{\log_2 M},$$

where M is the constellation order and N_0 is noise power spectral density. Besides, the transmit power constraint must be adopted, with power equally distributed among the n_T antennas.

Firstly, we consider an ULA distribution on the transmit and receive antennas, which results in the spatial correlation modeled in Section 2.1.1. We have considered three different scenarios of modulation order and number of antennas as evaluation standard for the MIMO detectors performance that do not generate prohibitive computational effort for the SD detector. Those arrangements are listed as follows (modulation; $N_T \times N_R$):

a) (64-QAM; 4×4); *b*) (16-QAM; 8×8); *c*) (4-QAM; 20×20).

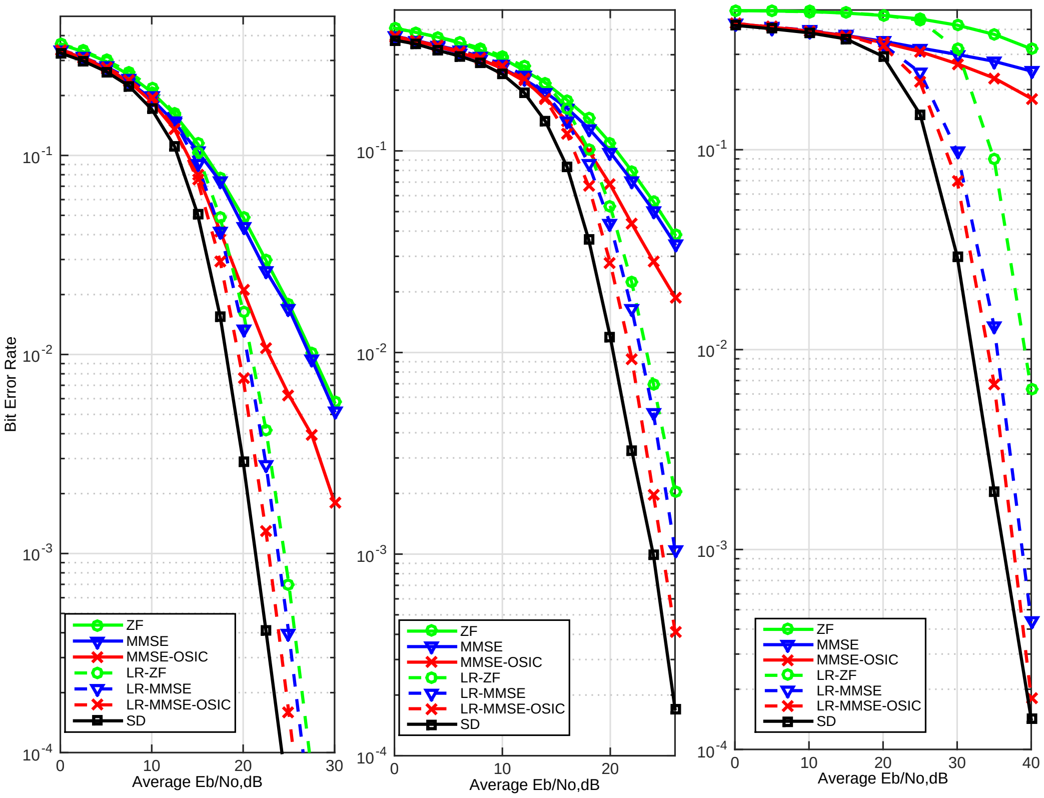
We also consider an UPA distribution for both transmit and receive antennas. In this case, as the structure is considered to work within massive MIMO systems, it was considered structures with high number of antennas. The studied arrangement are listed bellow:

a) (16-QAM; 8×8); *b*) (4-QAM; 64×64) (massive-MIMO);

Thus, three antenna correlation scenarios has been applied, specifically: $\rho = 0, 0.5$ and 0.9, which represents respectively no correlation, medium and strong correlation among antenna elements. Finally, in order of simplicity we have considered perfect knowledge of the channel gains in the receiver side, which means, the channel content $\mathbf{H}$ is available at the receiver, but unknown at the transmitter side.

The Figure 2.10 illustrates the first analyzed arrangement for the BER performance, which consists in 64-QAM modulation and 4×4 antennas format. We begin the analysis at low SNR regime, where all the analyzed MIMO detectors provide very similar performance. However, it is important to notice that the SD and the LR based detector can achieve full diversity, which means, in high SNR regime, where the SNR is negligible, there is a drop of 2^{n_T} in the BER for every 3 [dB] increase in the SNR. As the system is based on a small antenna array, the LR-aided detectors have an excellent performance showing a narrow gap in comparison with the SD, and also their BER curve remains parallel to the SD one, which implies in same diversity order.

Both ZF and MMSE detectors have similar performances in high SNR regions, this statement is verified by equations (2.32) and (2.34), where the difference is that the noise statistics are considered at the MMSE equation. Furthermore, by the application of interference cancellation, lattice-reduction techniques or the combination of both techniques leads to a great improvement in the MIMO detection performance, which is verified at Figure 2.10.



a) ρ = 0 (Uncorrelated)

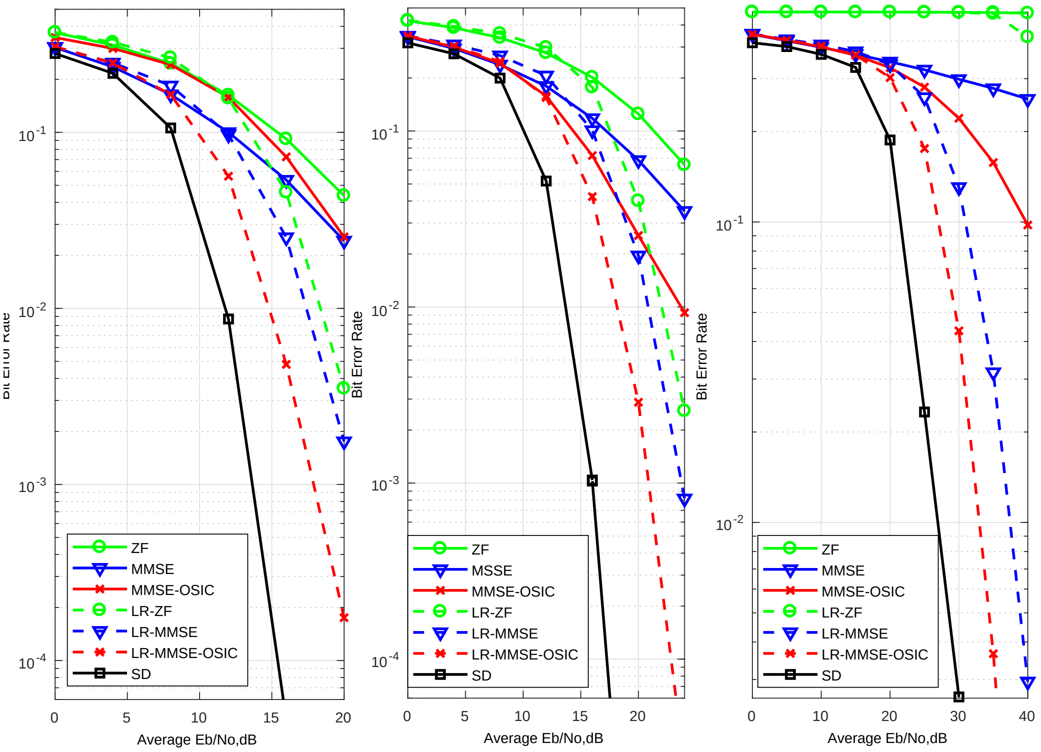
b) ρ = 0.5 (Medium Correlated)

c) ρ = 0.9 (High Correlated)

Figure 2.10: BER for the first arrangement ($64 - QAM$; 4×4)

Regarding the performance impact due to antenna correlation, is expedite confirm that as the correlation index increase the BER performance degenerates. At high correlation scenario, the non-LR-aided detectors require very high SNR to operate in suitable BER levels. This SNR demand for highly correlated scenarios directly impact in the energy efficiency, leading to undesirable rates. In fact, exclusively SD and LR-based MIMO detectors enables a great transmission energy efficiency and full diversity under high antenna correlation, which results in great BER performance, as seen in Figure 2.10(c).

Increasing the number of antennas, i.e. the (16-QAM; 8 × 8) case, will make the BER gap between the SD and the other MIMO detectors also to be increased, which is noticed in Figure 2.11. With this arrangement, differences in BER performance are evident; the most notable performance is achieved when the MMSE detector is combined with both LR and OSIC techniques, which was the closest to the optimal.



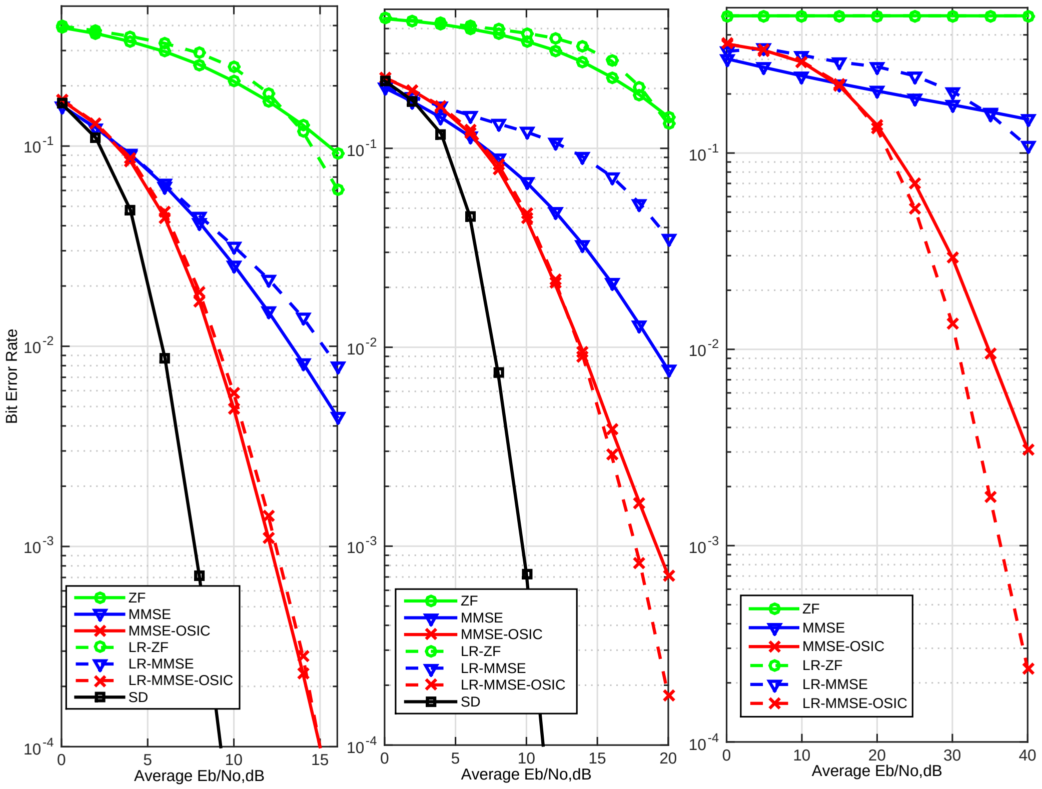
a) ρ = 0 (Uncorrelated)

b) ρ = 0.5 (Medium Correlated)

c) ρ = 0.9 (High Correlated)

Figure 2.11: BER for the second arrangement $(16 - QAM;\ 8 \times 8)$

It is important to notice that in large antenna arrays, such as Figure 2.12 with $n_R = n_T = 20$ antennas, the BER performance behaves different in each detection case. The first point is related to the ZF detector which becomes inefficient at low and medium space correlation scenarios, requiring high SNR regime to achieve reasonable BER performance. At high correlated scenarios, i.e $\rho = 0.9$, the ZF detector completely fails in decoupling the inter-antenna interference, also the MMSE detector loses diversity, while the LR-MMSE suffers from great BER performance degradation, particularly in high SNR regime. Furthermore, despite it extremely superior performance, under high spacial correlated channels the SD MIMO detector has showed an extremely exceeding computational complexity due to the vast number of branches that the SD algorithm needs to visit in order to detect the symbols in this configuration.



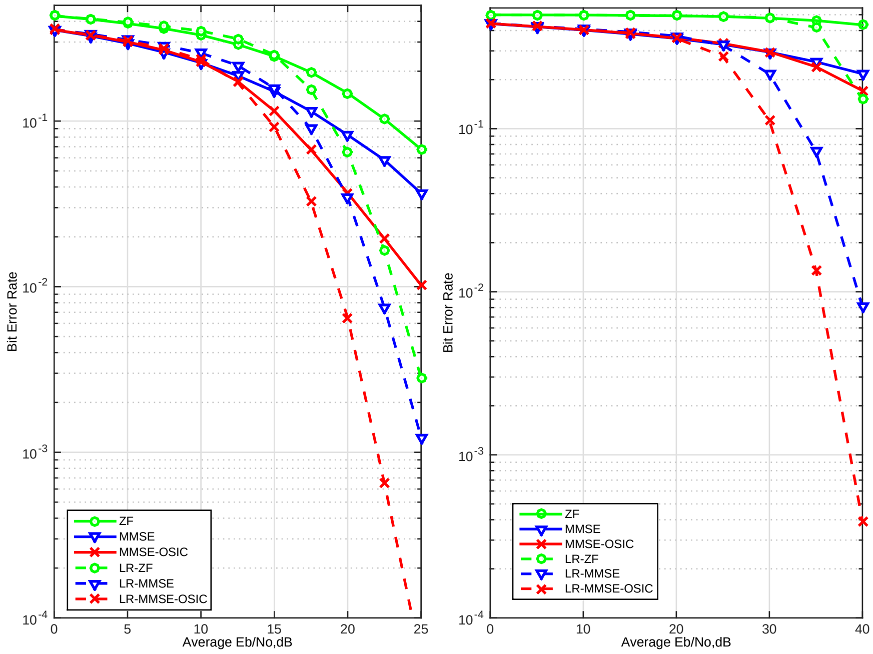
a) ρ = 0 (Uncorrelated) b) ρ = 0.5 (Medium Correlated) c) ρ = 0.9 (High Correlated)

Figure 2.12: BER for the third arrangement $(4 - QAM 20 \times 20)$

2.3.1 Spatial UPA × ULA Correlated Channels

Figure 2.13 depicts the BER performance for a 16-QAM MIMO with $n_R = n_T = 8$ with UPA antenna array (from Fig. 2.9) deployment and correlation index $\rho = 0.5$. Note that with the UPA array geometry applied the correlation effect becomes more severe due to the inner geometrical problem which is related to the antenna elements position. Notice that in uncorrelated channel scenarios the BER performance achieved for both ULA and UPA arrays will be the same for any system configuration, due to the Toeplitz structure of the correlation channel matrix. Otherwise, in correlated channel scenarios, performance losses will be expected for both UPA and ULA arrays, but with higher losses in the UPA structure, due to the cross-distances within antenna elements at both x and y-axes of the Euclidean plane. Such arrangement leads to higher interference in the received signal, leading to noise enhancement and consequently BER performance losses.

Figure 2.14 depicts the BER performance with the same previous configurations of modulation order and system size. The difference is that only the LR-MMSE detector is considered, in order to evaluate the performance gap between the system with different array structures. The detector choice was made based



a) ρ = 0.5

b) ρ = 0.9

Figure 2.13: BER for the first UPA arrangement (16-QAM; 8×8), with medium ($\rho = 0.5$) and very high ($\rho = 0.9$) spatial correlation.

on the LR-MMSE capacity to maintain the diversity at high SNR regions, and also because the characteristic of being able to deal with correlated channels while keeping a great performance. It is expedite conclude that, at medium and high correlation index there is a tiny performance gap between the ULA and UPA, in the order of approximately 2dB, which is introduced due to the correlation matrix condition. Specifically, as there is less interference coming from the neighbor antennas in the ULA correlated channel, a slightly better BER performance is observed.

Another interesting result depicted in 2.14 is the approximation between the Kronecker and the geometrical correlation models. The BER comparison were obtained for moderate correlation index, $\rho = 0.5$, which is straightforward calculated under the Kronecker model and the Geometrical-based correlation model following the expressions (2.25) and (2.22), respectively, for ULA and UPA antenna arrangements. The $\mathbf{R}_{\text{ula}}$ were simulated using the distance between elements $d = 0.5\lambda$ and the other variables are set to: $\theta = 3\pi/8$, $\xi = \pi/8$. For the UPA arrangement, the $\mathbf{R}_{\text{h}}$ were simulated using $d_1 = d_2 = 0.5\lambda$ and the other parameters were set to: $\theta = 3\pi/8$, $\phi = \pi/3$ $\xi = \pi/8$ and $\delta = \pi/6$ which represents a slightly greater angular spread when compared to the one used in (YING et al., 2014), because as larger the angular spread, the lower is the correlation index. So as we choose a moderated correlated case as our reference, $\rho = 0.5$, an enlargment of the angular spread is required to properly compare the correlated model with the geometrical one based on the refereed authors.

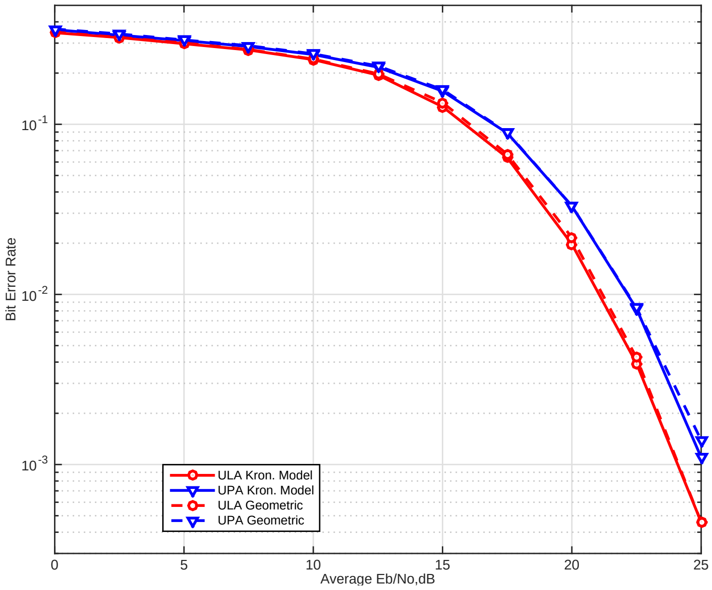


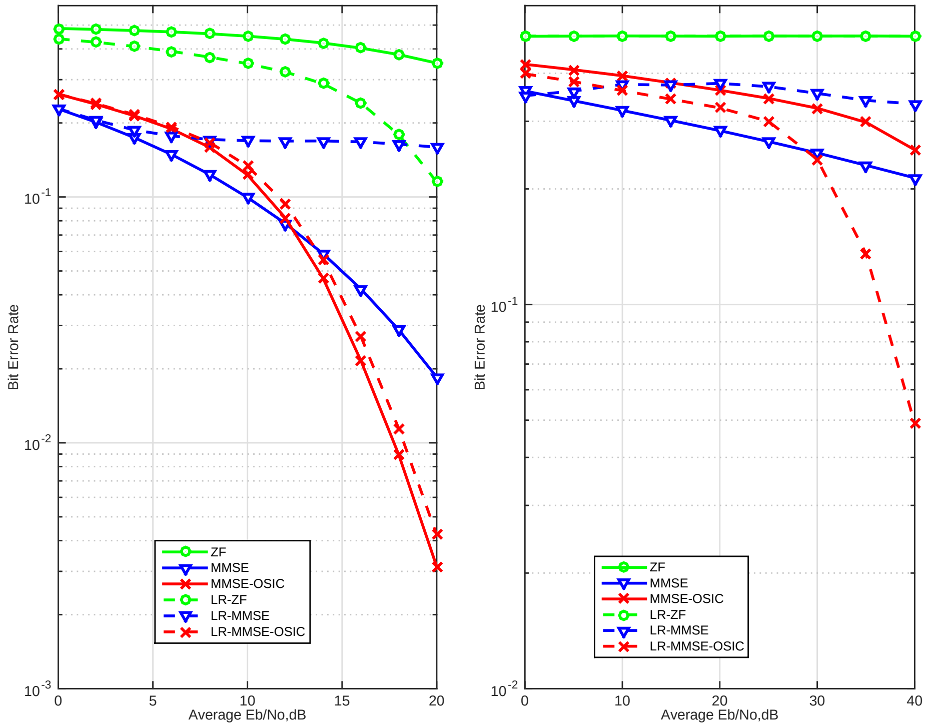
Figure 2.14: BER $\times E_b/N_0$ for the LR-MMSE detector with (16-QAM; 8×8) arrangement, correlation $\rho = 0.5$ with ULA and UPA arrays deployed at both Kronecker's approximation and Geometrical Model.

2.3.1.1 UPA Correlated Channels under Large-Scale Configuration

The UPA structure is proposed when large antennas array structures are deployed at the base station; following this perspective, Figure 2.15 depicts the BER performance for a 4-QAM 64×64 antennas systems. In this arrangement the SD detector performance is not depicted due to its impractical complexity over high number of antennas. With correlation index $\rho = 0.5$, all detectors tend to show greater degradation especially the LR-aided ones. This behavior can be explained due to the high size on the channel matrix, which makes more difficult to find a new orthogonal basis; as expected for high sized channel correlated matrix, the MMSE-OSIC detector presents a very similar performance of its LR-aided version. Furthermore, when the correlation index is incremented to a high correlated scenario, $\rho = 0.9$, all detectors suffer large diversity losses, except for the LR-MMSE-OSIC, which, despite the high-scale scenario, compared to other detectors, is still able to achieve greater diversity under high SNR regime.

2.3.2 Array Gain Impact on the Performance

The last analysis in this section is related to the array factor (AF), or array gain, which directly impacts the transmission gain, that eventually will impact over the SNR. As seen in section 2.1.1 the array factor for both structures will vary accordingly to the number of antennas and the spacing between them. It is also verified in that section that UPA structures are able to provide much more gain



a) ρ = 0.5 b) ρ =0.9

Figure 2.15: BER for the second UPA arrangement (4-QAM 64; ×64)

over the transmit direction regarding the uniform linear array. The impact of this characteristic is completely related to the BER performance, because the greater the normalized power loss, the worse will become the BER performance. In this perspective, it is important to emphasize that all previous BER performance results for ULA and UPA were conceived considering an array gain of 0dB, which implies that the beam gain from the BS is directed to the MT in a point-to-point MIMO link configuration. In terms of elevation and azimuth angles $\theta = 0°$ and $\phi = 0°$, and in this case, only the correlation effect will impact the BER performance.

A comparative analysis on the array gain for both array structures were made based on a 5×5 UPA and a 25 element ULA with 0.5λ element-spacing. The array gain comparison is provided in Table 2.1. To do such analysis, it is necessary to compare the UV response for both array structures. Figure 2.16 provide the UV response for both array structures in the azimuth cut condition, which means the azimuth angle is $\phi = 0°$, leaving only the normalized power response for elevation, θ, variations. Remembering that the x-axes follows the orthogonal projection given by equation (2.12), adopting $\phi = 0$, $\theta = \arcsin(u)$.

The array gain feature is directly related to the normalized power distribution, and each array structure provide its own power order. Analyzing the results in Table 2.1, it shows that a UPA structure provide greater gains, independently

Table 2.1: UPA and ULA array gain over various elevation angles θ under azimuth angle $\phi = 0°$ condition

Elevation angle		ULA	UPA
u	θ [°]	Array Gain (dB)	Array Gain (dB)
0	0	0	0
0.12	6.9	−13.41	−1.27
0.2	11.5	−17.8	−3.8
0.44	26	−24.05	−20.2
0.6	37	−26.12	−12.4
0.86	60	−30	−20

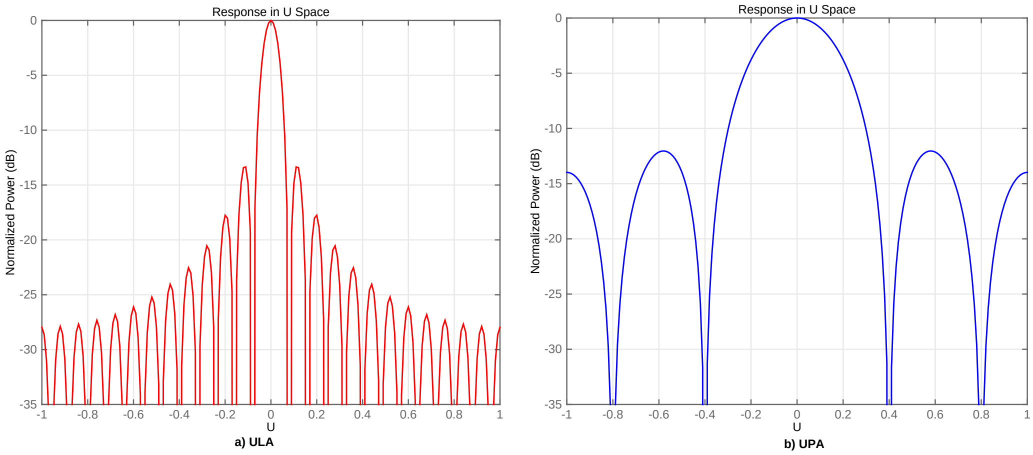


Figure 2.16: **a)** 25 antennas element ULA and **b)** 5×5 UPA with 0.5λ element-spacing UV response under azimuth angle $\phi = 0°$ condition.

of the MT position. The only exception occurs when $\theta = 0°$, because the BS will provide the same beam gain on the transmission for both ULA and UPA. In a MIMO point-to-point case, where the correlation effect is considered in both transmit and receiver side, the ULA structure will provide better performance in cases where the antenna beam pattern is focused on the MT. As the array gain directly impacts the SNR, and as consequence in the BER performance, cases where the antenna beam pattern is not focused directly on the MT, but in a region that has lower gain coverage, the UPA will provide better performances due to its higher gain lobe, especially under slightly deviations.

2.4 Complexity Analysis

The complexity analysis of MIMO detectors is of great importance, since all MT's should operate under strong signal processing and energy consumption limitations. With the combined analysis of BER performance and complexity it is possible to attain the best trade-off among the available detectors that comply with the system requirements.

With this objective in mind, this section presents a complexity comparison of those sub-optimum MIMO detectors. The complexity of each MIMO detector was measured in terms of flops (floating point operations), counting the total number of flops needed to perform the detection of a single transmitted symbol vector. For simplicity we have considered the flop counting for complex operations, specifically, one flops was considered for summations and three flops for complex product (WUBBEN et al., 2003). Furthermore, the flop counting for matrix operations were based on (GOLUB; LOAN, 1996), with the necessary modifications. Also, the complexity on the sorted QR decomposition were found in (WUBBEN et al., 2003), as well the complexity for the SD were based on the study carried out in (JALDEN; OTTERSTEN, 2005).

Through these methods, Table 2.2 presents the complexity in terms of number of flops for each used matrix operations, including matrix multiplying and inversion, approximated LLL complexity and the QR decomposition, which are procedures deployed in several MIMO detectors, specially those detectors treated herein, where, $n = n_R = n_T$ and M are the number of antennas and the M-QAM order of modulation, respectively.

Table 2.2: Number of flops for each operation/procedure

Operation	Number of flops
$\mathbf{C}_{n\times n} = \mathbf{A}_{n\times n} \times \mathbf{B}_{n\times n}$	$2n^3$
$\mathbf{y}_{n\times p} = \mathbf{A}_{n\times n} \times \mathbf{x}_{n\times p}$	$2n^2p$
$\mathbf{C}_{n\times n} = \mathbf{A}_{n\times n} + \mathbf{B}_{n\times n}$	n^2
$f_{\text{LLL}}(n, \rho)$ (KOBAYASHI; CIRIACO; ABRÃO, 2015)	$\approx (ae^{b\rho} + c)n^3$
SQRD (WUBBEN et al., 2003)	$16n^3/3 + 7n^2/3 + 25n/6$
$\mathbf{C}_{n\times n} = \mathbf{A}^{-1}_{n\times n}$	$2n^3/3$

When it comes to complexity, despite of its good BER performance, the LR aided detectors may present a growing complexity in certain scenarios. Through the simulations, it was observed that the complexity of the LLL algorithm does not only depend on the matrix size, but also on the correlation index. Naturally the dependence between complexity and matrix size is straightforward due to the number of operations evaluated. On the other hand, the increase of the correlation index leads to a quasi-singular matrix, which makes it difficult for LLL procedure to find an orthogonal basis, leading to an increase in computational complexity.

The exact LLL complexity cannot be easily evaluated due to all the variable dependencies. However, it is known that a good approximation for the LLL complexity can be evaluated as a $\mathcal{O}(n^3 \log n)$ order (LING; HOWGRAVE-GRAHAM, 2007). Aiming to provide a better expression that represent the LLL complexity, in (KOBAYASHI; CIRIACO; ABRÃO, 2015) a numerical experiment was conducted to determine, through the better surface fitting, the LLL complexity dependency w.r.t. the antenna correlation index and array dimension. With such experiment the most similar surface fitting the LLL complexity, were given by:

$$f_{\text{LLL}}\left(n, \rho\right) \approx \left(a e^{b\rho} + c\right) n^3 \tag{2.49}$$

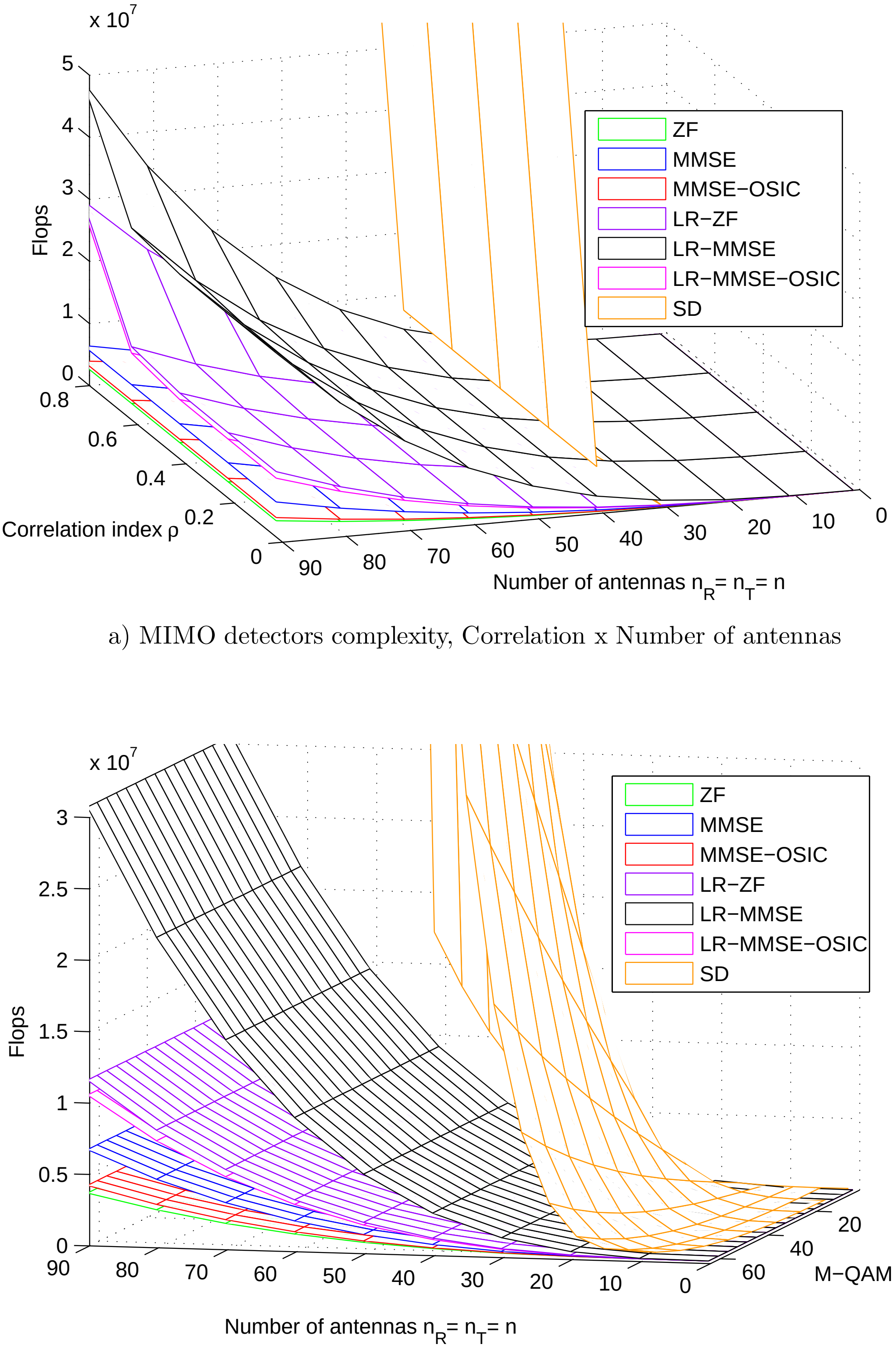
with $a = 5.018 \times 10^{-4}$, $b = 13.48$ and $c = 8.396$. Finally, for the LR-aided MIMO detectors, the necessary flop counting approximation for the LLL procedure given in (2.49) was included in the total complexity calculation.

A complexity evaluation of the previous analyzed MIMO detectors and the various combinations possibilities have been made. Table 2.3 summarizes the overall complexity for the most relevant combinations of sub-optimum MIMO detectors covered in this work. Notice that the ML complexity grows exponentially and it becomes prohibitive when the product number of antennas by modulation order $(n \cdot M)$ increases, which is the case of any practical MIMO case of interest (including small-medium $n \cdot M$ values) (KOBAYASHI; CIRIACO; ABRÃO, 2015). Moreover, the SD complexity is not trivial to obtain; as demonstrated in (JALDEN; OTTERSTEN, 2005) the SD complexity always present an exponential asymptotic behavior in low SNR and/or large $n \cdot M$ scenarios. This occurs because the algorithm needs to ensure certain probability to find a point inside the sphere, then if the problem size and/or noise power increase, the hyper-sphere radius grows and consequently the complexity.

Table 2.3: MIMO Detectors Complexity

MIMO Detector	**Total Complexity**
ZF (VALENTE; MARINELLO; ABRÃO, 2014)	$14n^3/3 + 2n^2$
MMSE (VALENTE; MARINELLO; ABRÃO, 2014)	$26n^3/3 + 4n^2$
MMSE-OSIC	$16n^3/3 + 13n^2/3 + 25n/6$
LR-ZF	$20n^3/3 + 10n^2 + 4n + f_{\text{LLL}}\left(n, \rho\right)$
LR-MMSE	$32n^3 + 14n^2 + 3n + f_{\text{LLL}}\left(n, \rho\right)$
LR-MMSE-OSIC	$22n^2/3 + 13n^2/3 + 25n/6 + f_{\text{LLL}}\left(n, \rho\right)$
SD Ref.(JALDEN; OTTERSTEN, 2005)	$4n^3 + 7n^2 + n/2 + (2n+2)\frac{M^{\gamma n}-1}{M-1}$, where $\gamma = 1/2\left[\frac{c^2(M^2-1)}{6N0} + 1\right]^{-1}$ and $c^2 = \mathbb{E}\left[\lVert\mathbf{h}_i\rVert^2\right]$, $\forall i \in [1, n]$
ML (KOBAYASHI; CIRIACO; ABRÃO, 2015)	$M^n(4n^2 + 2n)$

Regarding the ZF and MMSE detectors, their computational effort directly relates to Eq. (2.32) and (2.34) respectively, which can be calculated through a matrix inversion, a matrix summation, and multiplications. The main difference between their complexity is that the MMSE requires two multiplications instead of one, and four multiplications, instead of two needed for the ZF algorithm. The ZF and MMSE complexities comply with Ambrosio *et al* (VALENTE; MARINELLO; ABRÃO, 2014).



a) MIMO detectors complexity, Correlation x Number of antennas

b) MIMO detectors complexity, Modulation order x Number of antennas

Figure 2.17: MIMO detector complexity $E_b/N_0 = 20[dB]$

When it comes to the OSIC-aided detectors, the primary source of computational effort it is the SQRD algorithm, which offers a cubic complexity order

(WUBBEN et al., 2003). When the SQRD procedure is combined with the SIC algorithm, represented by Eq. (2.38), the result is an increment in the quadratic order, due to the SIC computational method. The complexity for the MMSE-OSIC in Table 2.3 is corroborated by that found in (KOBAYASHI; CIRIACO; ABRÃO, 2015) and (VALENTE; MARINELLO; ABRÃO, 2014).

Considering the LR aided detectors, the same procedure holds, but now with the addition of the LLL algorithm complexity. Following Algorithm 2.2, we notice that the LR is composed by a QR decomposition and the LLL, so the LR-ZF aided detector complexity can be determined from the equation, Eq. (2.43), in combination with the LLL algorithm and the matrix manipulations covered in section 2.2.7. Regarding the LR-MMSE the only difference is the usage of the extended matrix, which increases the computational effort by doubling the size of all operations. Finally, the LR-MMSE-OSIC is based on the addition of the SQRD algorithm, the LLL, and the SIC procedure, which reduce the complexity by eliminating a series of matrix multiplications. The complexity evaluation of the LR aided detectors relies upon (KOBAYASHI; CIRIACO; ABRÃO, 2015).

Figure 2.17 depicts the computational complexity for the various MIMO detectors studied in this chapter, divided in terms of flops as a function of:
a) normalized correlation index × number of antennas;
b) M-QAM order × number of antennas.
From Table 2.3 and Figure 2.17, one can notice that the OSIC detectors are capable to offer much better complexity-performance tradeoffs when compared to the versions with the pseudo-inverse. This is caused by the fact that the SQRD leaves an upper triangular systems which demands lower complexities then the pseudo inverse. Also, ordered version is preferable over the simple SIC, since, accordingly (WUBBEN et al., 2003), the first one requires $2n^2 - 2n$ flops in the overall complexity, while providing considerable performance improvements.

Moreover, Figure 2.17.a) shows that the LR-aided detectors presents a reasonable complexity under low to medium correlation index, besides it keeps full diversity for those scenarios, which makes these class of sub-optimal MIMO detectors one of the most promising in the context of this work. Finally, the exponential complexity on ML makes it prohibitive for any practical MIMO case of interest (including small-medium $n \cdot M$ product), while the SD detector can result in great complexity saving only for cases where the systems is under high SNR regime with low number of antennas and modulation order, otherwise it results in increasingly high computational complexity burden, mainly combining low SNR regime with high correlated channels and $n \cdot M$ products.

2.5 Conclusions

The initial MIMO detection performance analyses carried out in this chapter were based on the independent identically distributed and perfect estimated channels; this opened the possibility to analyze the performance × complexity trade-off of such MIMO detectors operating under more realistic correlated channels.

Lattice reduction technique has been proved to provide great BER performance improvements of linear sub-optimum MIMO detectors. The analysis of MIMO detectors under correlated channels indicates notable advantage in terms of reduction in BER degradation for the LR-aided MIMO detector due to the ability to deal with the near orthogonality of the channel matrix $\mathbf{H}$, besides it achieves full diversity. The LR-MMSE-OSIC MIMO detector presented the smaller degradation in terms of BER performance, even under high correlated MIMO channels. Linear detectors aided by the combination of both LR and OSIC techniques can provide a near optimum performance in some cases; however the LR aided detectors tend to present increasing complexity when high correlated scenarios are applied, resulting in lack of orthonormalization that the LLL algorithm present when a near singular matrix is given as an input. Therefore, the LR-MMSE-OSIC have achieved the best performance-complexity trade-off among the presented detectors.

When it comes to array structure and correlation effect, the ULA will always perform better when the antenna beam gain is focused on the MT, otherwise the UPA structure will provide greater BER performances, despite the correlation, due to its great inherent transmit power distribution pattern.

3 Semidefinite Relaxation for Large Scale MIMO Detection

This chapter provides an analysis of the MIMO detection under a non-linear optimization perspective, aiming the study of near optimum performance detectors. The analysis here provided is the comparison in terms of complexity-performance trade-off of the Semidefinite Relaxation (SDR) strategy detector with linear sub-optimum detectors. This comparison is provided in a single cell point-to-point MIMO equipped with the same number of transmit and receive antennas. The main contribution of the chapter is the SDR detection performance analysis under high number of antennas in both base station and the receiver, also providing an analysis of the most suitable technique to recovery the relaxed solution when the number of antennas are consistently increased.

3.1 Introduction

From the previous chapter, it is known that, the optimal detection solution in the sense of minimum joint probability of error for detecting all the symbols simultaneously is solved by the ML detector, which is known as NP-hard (BAI; CHOI; YU, 2014). It can be implemented by a brute force-search over all of the possible transmitted vectors set, searching for the one that minimizes the Euclidean distance from the received vector, or using more efficient search algorithms, i.e, the sphere decoder (SD)(BAI; CHOI; YU, 2014; JALDEN, 2004). However, the expected computational complexity of the ML receiver, even when SD is applied, is unpractical for many channel scenarios and applications. Consequently, there has been much interest in implementing sub-optimal or quasi-optimal MIMO detection algorithms, such as the linear receivers, i.e, the zero-forcing (ZF) and the minimum mean squared error (MMSE) MIMO detectors (BAI; CHOI; YU, 2014).

One of the most promising quasi-optimal MIMO detection strategies is the semi-definite relaxation (SDR), which provides a better BER performance than the linear and decision-feedback MIMO receivers (JALDEN; MARTIN; OTTERSTEN, 2003; WIESEL; ELDAR; SHAMAI, 2005; MA; CHING; DING, 2004; JALDEN, 2004; MA et al., 2002) while holds same order of complexity. The SDR attempts to approximate the solution for the ML problem using a convex program that can be efficiently solved in polynomial time. The usual approach of the SDR problem is first to formulate the ML problem in a higher dimension and then relax the non-convex constraints; such relaxation will result in a semi definite program (SDP), for which there are efficient tools to obtain solutions in polynomial time (CVX, 2012).

SDR was first proposed for signal detection problem on binary/quadratic phase shift keying (BPSK/QPSK) constellations,(JALDEN; MARTIN; OTTERSTEN, 2003; MA; CHING; DING, 2004), in which near-ML optimal performances were empirically observed. These results suggest that at high signal-to-noise ratios (SNRs), there is a high probability that SDR will yield the true ML decision. Another result that motivates the use of the SDR shows that many of the other conventional detectors, such as the MMSE, are relaxations of the SDR and are, therefore, inferior in performance ways (MA et al., 2002).

The success of SDR in demodulating BPSK signaling motivated its generalization to higher constellations, e.g., the generalization to M-ary quadrature amplitude modulation (M-QAM) signaling was intensively studied in order to conceive high data rate systems. An SDR detector scheme for high-order 16-QAM modulation was proposed in (WIESEL; ELDAR; SHAMAI, 2005), while an approximation of this detector was developed in (SIDIROPOULOS; LUO, 2006), aiming to achieve a high order 64-QAM constellation signaling. Moreover, in (MA et al., 2002; MAO; WANG; WANG, 2007) the SDR detection problem is generalized considering 4^q-QAM ($q \geq 1$) modulation orders. In (RAPOPORT et al., 2012) a large scale SDR-based detector is proposed for fast signal detection. The SDR problem is further reduced to the sequential linear programming by adding new form of cutting planes and column generation method. BER performance is compared with linear ZF and MMSE MIMO detectors, as well as the ML optimal detectors for 16×16 and 28×28 antennas. In (TRAN; HANIF; JUNTTI, 2014), authors suggest that the conventional SDR detector in a multi-casting problem, where the transmitter is equipped with a massive antenna array, the complexity of solving semi-definite problem (SDP) directly obtained can be prohibitively high. Authors devise the SDP in a dual domain, producing a more computationally efficient solution. Also, they proposed an iterative second-order cone programming solution that is free from employing any randomization step.

This work analyzes the performance-complexity trade-off of the SDR-MIMO detection algorithm, taking as reference both linear sub-optimal and ML optimal solutions; low signaling orders are adopted in comparison with ML while high order modulation schemes are adopted when comparing SDR-MIMO approaches with linear ZF or MMSE MIMO detectors.

3.2 Problem Statement

Considering a standard MIMO channel, the received signal can be described by:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{3.1}$$

where $n_T \times 1$ symbols $\mathbf{s}$ are transmitted simultaneously through a channel which gain is represented by a $n_R \times n_T$ matrix $\mathbf{H}$ and the additive noise $n_R \times 1$ vector samples $\mathbf{n}$. Each element of the channel matrix $\mathbf{H}$ represents the channel gain in the respective selected path; those gains are assumed known at the receiver side and represented by a Rayleigh distribution. The $n_T \times 1$ vector $\mathbf{y}$ represents the received signal samples in each symbol period, formed by the symbols after passing through the channel. It is also known that the noise vector $\mathbf{n}$, are samples of additive noise represented as circularly-symmetric Gaussian distribution, $\mathbf{n} \sim \mathcal{CN}\{0, \sigma_n^2\mathbf{I}\}$, with variance σ_n^2.

For the subsequent analysis and without loss of generality we assume $n_R = n_T$. The system model is fully defined by complex variables; however, since we focus on the optimization procedures, for simplicity and computational convenience, the complex variables are split into a double real-value structure. So, rewriting the received MIMO signal in (3.1) with imaginary and real part separately (BAI; CHOI; YU, 2014; KOBAYASHI; CIRIACO; ABRÃO, 2015):

$$\begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{s}\} \\ \Im\{\mathbf{s}\} \end{bmatrix} + \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix} \tag{3.2}$$

We consider a high order M-QAM modulation, where the symbols are denoted by a complex number with real and imaginary part are limited to $\pm\left(\sqrt{M}-1\right)$. The structure of the complex set can be represented by:

$$\mathbb{S} = \left\{a + jb \mid a, b \in \left\{-\sqrt{M}-1, -\sqrt{M}+3, \ldots, \sqrt{M}-1\right\}\right\}.$$

For this modulation, the average symbol energy is given by:

$$E_s = \frac{2(M-1)}{3} \tag{3.3}$$

3.2.1 Maximum Likelihood (ML)

The maximum-likelihood (ML) detector performs an exhaustive search over the whole set of possible symbols $s_i \in \mathbb{S}$, in order to decide in favor of the one that minimizes the Euclidean distance between the received signal $\mathbf{y}$ and the reconstructed signal $\mathbf{Hs}$:

$$\widehat{\mathbf{s}} = \arg\min_{\mathbf{s}\in\mathbb{S}} \left\|\mathbf{y} - \mathbf{Hs}\right\|^2 . \tag{3.4}$$

It is well known that the ML detector provides the lowest BER performance of all MIMO detectors, but the search complexity grows exponentially according to the number of antennas and the number of symbols, leading to a M^{n_T} symbol set combinations.

3.3 Relaxed ML Criterion by Semidefinite Programming

SDR is an efficient approximation tool for non-convex quadratically constrained quadratic programming (QCQP) problems and it has been shown to provide good approximation accuracy in the application of near-ML detection problem with BPSK (JALDEN; MARTIN; OTTERSTEN, 2003) and QPSK (WIESEL; ELDAR; SHAMAI, 2005; MA; CHING; DING, 2004) constellations. Like most relaxation methods, SDR consists of three steps: a) relax the feasible set of the original problem in order to ease the solution of the relaxed problem; b) solve the relaxed problem; c) convert the relaxation solution to an approximate solution of the original problem.

The main idea behind the SDR approach applied to hard decision MIMO detection is to first establish the finite constellation requirement as a low-rank (in this case rank one) constraint on a matrix whose diagonals belong to a finite constellation. After that, those two constrains are relaxed to a positive semidefinite constraint, which makes the resulting problem convex and enables to use semi-definite programming to solve it (CIRKIC, 2014). More specifically, we can rewrite the ML problem posed in (3.4) as follows:

$$\begin{aligned}\|\mathbf{y}-\mathbf{Hs}\|^2 &= \mathbf{s}^T\mathbf{H}^T\mathbf{Hs}-2\mathbf{y}^T\mathbf{Hs}+\|\mathbf{y}\|^2 \\ &= \mathbf{x}^T\mathbf{Lx}+\|\mathbf{y}\|^2,\end{aligned} \tag{3.5}$$

where $\mathbf{L} \triangleq \begin{bmatrix}\mathbf{H}^T\mathbf{H} & -\mathbf{H}^T\mathbf{y} \\ -\mathbf{y}^T\mathbf{H} & 0\end{bmatrix}$

Thus the $\widehat{\mathbf{s}}$ can equivalently be obtained through

$$\widehat{\mathbf{s}} = \underset{s\in\mathbb{S}}{\operatorname{argmin}}\, \mathbf{s}^T\mathbf{H}^T\mathbf{Hs}-2\mathbf{y}^T\mathbf{Hs} \tag{3.6}$$

since $\|\mathbf{y}\|^2$ does not depend on $\widehat{\mathbf{s}}$ (JALDEN, 2004). The function of the above problem can equivalently be written as

$$\begin{bmatrix}\mathbf{s}^T & 1\end{bmatrix}\begin{bmatrix}\mathbf{H}^T\mathbf{H} & -\mathbf{H}^T\mathbf{y} \\ -\mathbf{y}^T\mathbf{H} & 0\end{bmatrix}\begin{bmatrix}\mathbf{s} \\ 1\end{bmatrix} \tag{3.7}$$

and thus by letting $\mathbf{x} = \begin{bmatrix}\widehat{\mathbf{s}}^T & 1\end{bmatrix}^T$ the ML detection problem can be solved examining the equivalent problem in the second line of (3.5):

$$\begin{aligned}\min_{\mathbf{x}\in\mathbb{R}^{n_T+1}} \quad & \mathbf{x}^T\mathbf{Lx} \\ \text{s.t.} \quad & x_i^2 = 1 \quad i=1,\ldots,2n_T+1\end{aligned} \tag{3.8}$$

where x_i is the ith component of $\mathbf{x}$.

Then, SDR utilizes $\mathbf{x}^T\mathbf{Lx} = \operatorname{tr}(\mathbf{x}^T\mathbf{Lx})$ and $\mathbf{X} = \mathbf{xx}^T$, which lets the MIMO detection problem in (3.8) for high modulation order be equivalent to

$$\begin{aligned}\min_{\mathbf{X},\mathbf{x}} \quad & \operatorname{tr}(\mathbf{LX}) \\ \text{s.t.} \quad & \operatorname{diag}(\mathbf{X}) = \mathbf{e} \\ & \mathbf{X}\left(2n_T+1, 2n_T+1\right) = 1 \\ & \mathbf{X} \succeq 0; \quad rank(\mathbf{X}) = 1\end{aligned} \tag{3.9}$$

where $\mathbf{e}$ is the vector of all ones and where $\mathbf{X} \succeq 0$ means that $\mathbf{X}$ is symmetric and positive semi-definite.

We should observe that the optimization problem in (3.9) is not convex yet[1] and it is equivalent to (3.4) in the sense that if the solution of the first is known, the solution to the second can be easily computed and vice-versa. However, the component that makes (3.9) hard is more explicit than the constrains in (3.4). Accurately, the only difficult constraint in (3.9) is the rank constraint,

[1] Because of the rank constraint in $\mathbf{X}$ (JALDEN, 2004)

$rank(\mathbf{X}) = 1$, which is non-convex, the objective function and all the other constraints are convex in $\mathbf{X}$, thus we should drop the rank constraint in order to obtain the relaxed version of the problem (3.4):

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{tr}(\mathbf{LX}) \\ \text{s.t.} \quad & \text{diag}(\mathbf{X}) = \mathbf{e} \\ & \mathbf{X}\left(2n_T+1, 2n_T+1\right) = 1; \quad \mathbf{X} \succeq 0 \end{aligned} \tag{3.10}$$

The problem (3.10) is the SDR version for high modulation order of (3.4) and the difference between them is that the constraints on $\mathbf{X}$ has been replaced by $\mathbf{X} \succeq 0$. The problem in (3.10) is a semi-definite program and standard methods can be used to solve it in polynomial time (VANDENBERGHE; BOYD, 1996). The SDR problems can be handled very conveniently and effectively by readily available (and free) software packages; e.g, by using the convex optimization toolbox CVX (CVX, 2012), we can solve (3.10) in MATLAB with it's SDP mode.

Moreover, in order to solve a high modulation order problem one constraint must be modified and this modification is denominated as *bound constraint* SDR (BC-SDR). In this work the method for high order modulation problem was based on (SIDIROPOULOS; LUO, 2006). The convex optimization problem for high order modulation cases is rewritten on it relaxed version as:

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{tr}(\mathbf{LX}) \\ \text{s.t.} \quad & I_L\mathbf{I} \geq \text{diag}(\mathbf{X}) \geq S_L\mathbf{I} \\ & \mathbf{X}\left(2n_T+1, 2n_T+1\right) = 1; \quad \mathbf{X} \succeq 0 \end{aligned} \tag{3.11}$$

where, $I_L = \min \log_2(M)^2$; $S_L = \max \log_2(M)^2$ and $\mathbf{I}$ is the $2n_T + 1$ dimensional identity matrix.

In the backstage, most convex optimization toolboxes handle SDP with an interior point algorithm. Hence, the SDR problem (3.10) can be solved with a worst case complexity(LUO et al., 2010):

$$\mathcal{O}\left(\max\left\{m,n\right\}^4 n^{\frac{1}{2}} \log\left(\frac{1}{\epsilon}\right)\right) \tag{3.12}$$

where m is the number of constraints, n is the problem size and ϵ is a given solution accuracy. From the point of view of the MIMO equalization problem, the variables m and n are respectively represented by the number of transmit (n_T) and receive (n_R) antennas.

From (3.12), the SDR complexity scales slowly (logarithmically) with ϵ and most applications do not require a very high solution precision; hence, simply speaking, we can say that the SDR is a computationally efficient approximation approach to QCQP problems, in the sense that its complexity is just polynomial time. So, basically the SDR transforms a NP-hard combinatory problem (3.4) into a polynomial time solvable problem (3.10) and (3.11).

Furthermore, with the relaxation of the rank constraint, a fundamental issue that can be found while using SDR is how to convert a globally optimal solution $\mathbf{X}^*$ of (3.10) into a feasible solution $\tilde{\mathbf{x}}$ to (3.4). If $\mathbf{X}^*$ is already rank one, then there is nothing to do, and we can write $\mathbf{X}^* = \mathbf{x}^*\mathbf{x}^{*T}$, and $\mathbf{x}^*$ will be a feasible and optimal solution of (3.4). On the other hand, if the rank of $\mathbf{X}^*$ is larger than one we must extract from it, in an efficient manner, a vector $\tilde{\mathbf{x}}$ that is feasible for (3.4) (LUO et al., 2010).

There are many heuristic ways to extract the rank one solution, however, even though the extracted solution is feasible for (3.4), it is in general not an optimal solution. Different way to extract the optimal solution from the feasible solution include the rank one approximation and the Gaussian randomization. In this work both rank one approximation and the Gaussian randomization techniques have been deployed.

3.3.1 Rank One Approximation

The rank one approximation consists in the most simple technique to extract a solution $\mathbf{x}^*$ to the non-convex problem from the solution of the convex problem formulated, $\mathbf{X}^*$. With this procedure it is assumed that every solution of $\mathbf{X}^*$ is a rank one solution. Algorithm 3.1 describes the steps to perform the rank one approximation strategy; in step 3, the operator `slicer`$(\cdot)$ is an approximation to the nearest constellation value.

Algorithm 3.1 1-Rank Approximation SDR-MIMO Detection

```
Input: $\mathbf{X}^*$
Output: $\widehat{s}_i$
1: First we should take the eigen-decomposition of $\mathbf{X}^*$
       $\mathbf{X}^* = \sum_{i=1}^{r} \lambda_i q_i q_i^T$
2: Then we select the higher eigenvalue
       $I = \arg\max_i \lambda_i$
3: Take $\mathbf{x}^*$ as the slicer on the eigenvector constellation associated with the
   higher eigenvalue.
       $\mathbf{x}^* = \texttt{slicer}(\mathbf{q}_a)$
4: The estimation of the transmitted symbol in real form is obtained in $\mathbf{x}^*$,
   except from the last position of the vector
       $\widehat{s}_i = x_i^* \quad i = 1, \ldots, 2n_T$
```

3.3.2 Gaussian Randomization

The Gaussian randomization procedure is widely deployed; e.g., (LUO et al., 2010) has demonstrated excellent near-ML results under high number of antennas condition. Alternatively, in this work the Gaussian randomization process based in (WIESEL; ELDAR; SHAMAI, 2005) findings has been used. Algorithm 3.2 describes such procedure.

Algorithm 3.2 Gaussian Randomization SDR-MIMO Detection

```
Input: $\mathbf{X}^*, S_g, \mathbf{L}$,
Output: $\widehat{\mathbf{s}}_i$
Cholesky Factorization at the SDR solution matrix:
    $\mathbf{X}^* = \mathbf{U}^T\mathbf{U}$
Let $\mathbf{u}_i$ the i-th column of $\mathbf{U}$
for $i =$ to $S_g$ do
    Generate a random vector $\mathbf{r}$ with a uniform distributed over a unitary sphere of $(2n_T + 1)$ dimension.
    Let $\mathbf{x_g}$ be the:
    $\mathbf{x_g}_i = \texttt{slicer}\left(\frac{\mathbf{u}_i^T\mathbf{r}}{\mathbf{u}_{2n_T+1}^T\mathbf{r}}\right), \quad i = 1, 2, \ldots, 2n_T + 1$
    Calculate the the vector $\mathbf{k}$ as:
    $\mathbf{k}_i = \mathbf{x_g}^T\mathbf{L}\mathbf{x_g}, \quad i = 1, 2, \ldots, S_g$
end for
$\mathbf{x_g} = \min(\mathbf{k})$
$\widehat{\mathbf{s}}_i = \mathbf{x_g}, \quad i = 1, \ldots, 2n_T$
```

3.4 Numerical Results

In this section the BER *versus* E_b/N_0 performance analysis under perfect channel estimation, different number of antennas and modulation order have been considered. The performance and the complexity trade off is an important parameter to be defined; hence, the computation complexity was analyzed for each MIMO detector considered in this work. Furthermore, we have compared the SDR detector under both estimation approaches with the LR-ZF strategy. Specifically, on the SDR detection it was utilized the rank one approximation (SDR Rank One) and the Gaussian randomization (SDR Rand) in order to extract the feasible solution $\widehat{\mathbf{s}}$ from the globally optimum $\mathbf{X}^*$. Numerical simulations are performed in uncoded spatial multiplexing MIMO systems employing 16-QAM constellations for different antenna configurations, e.g., 8×8, 16×16, 64×64 and 128×128 antennas. As demonstrated in the following, the SDR Rand approach overcomes the SDR Rank One approximation for medium/high SNR regions and low size problems. On the other hand, when large MIMO was deployed, an inversion on the BER performance behavior emerges: SDR Rank One MIMO detector overcomes the SDR Rand MIMO detector performance because of its low complexity.

3.4.1 BER Performance

Fig. 3.1 depicts the BER performance for the SDR MIMO detector equipped with both rank approximation and gaussian randomization estimations in comparison with the LR aided linear detectors, namely LR-ZF and LR-MMSE. This procedure was performed in a scenario with 16-QAM constellation, $n_T = n_R = 8$ antennas (Fig. 3.1.a) and $n_T = n_R = 16$ antennas (Fig. 3.1.b), under non-line-of-sight (NLOS) Rayleigh propagation channels plus additive white Gaussian noise.

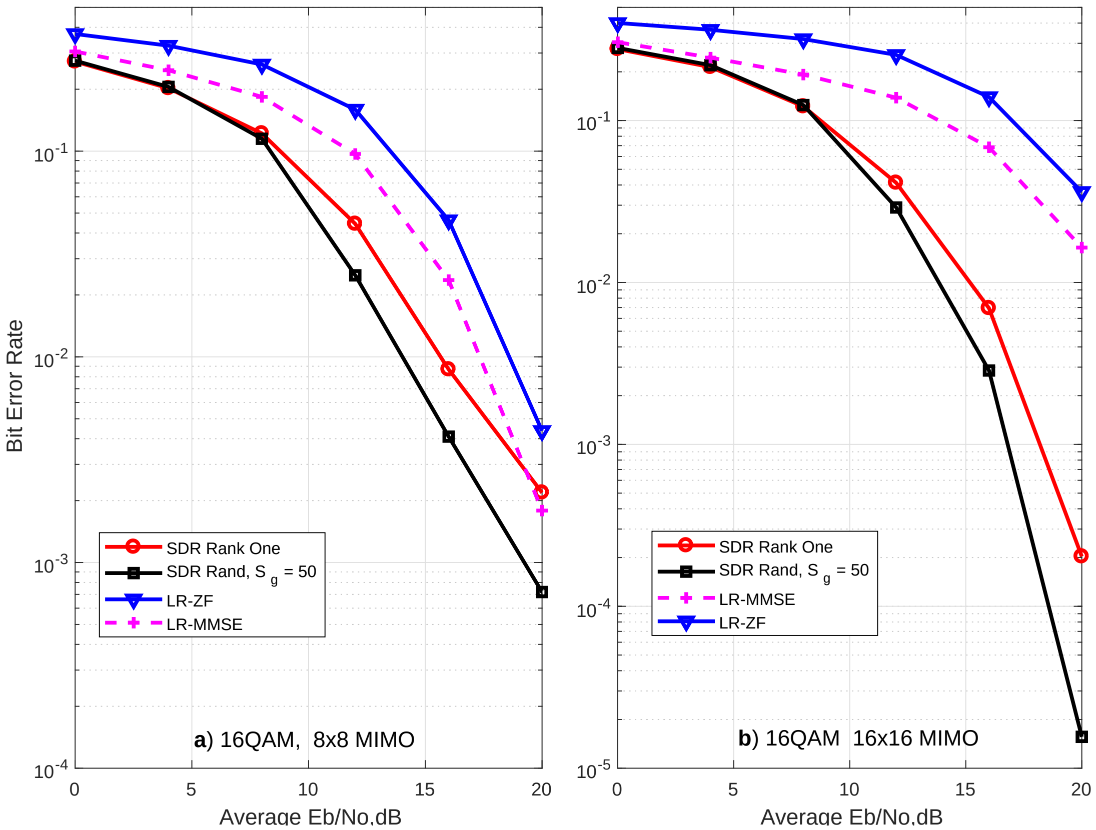


Figure 3.1: BER performance for 16-QAM SDR and LR-aided linear MIMO detectors equipped with a) 8 × 8 antennas and b) 16 × 16 antennas

Note that both SDR approaches result in better performance under low and high SNR regions; moreover, asymptotically speaking both SDR approximations (Rank One and Rand) tend to get close to each other but at the medium SNR regions the SDR with randomization approach has $4dB$ gain over the performance of the LR-ZF linear MIMO detector. Fig. 3.1.b depicts the BER performance for $n_T = n_R = 16$ antennas under the same 16-QAM constellation order and NLOS

Rayleigh channel. As the number of antennas grows the LR technique applied to MIMO systems makes them more sensitive to noise, what makes the BER performance be considerable in high SNR regions, where the additive noise is negligible. When the SDR is analyzed a diversity gain was directly observed. Moreover, a considerable performance gain is achieved in high SNR region, something ≈ 7dB higher.

As a conclusion, the achieved performance of both approximations for the SDR detector in MIMO Rayleigh channels improves progressively with the number of both transmit and receive antennas. Such progressive improvement of SDR Rank One, depicted in Figs. 3.2.a and 3.2.b, reflects directly over the complexity the detection strategy. Finally for the Rand approach, as the problem size grows, the number of randomization samples, S_g, must be incremented for better BER performance. Indeed, under lower size problems, the lowest value for S_g on the SDR Rand algorithm have a better BER in comparison to the SDR Rank One approach.

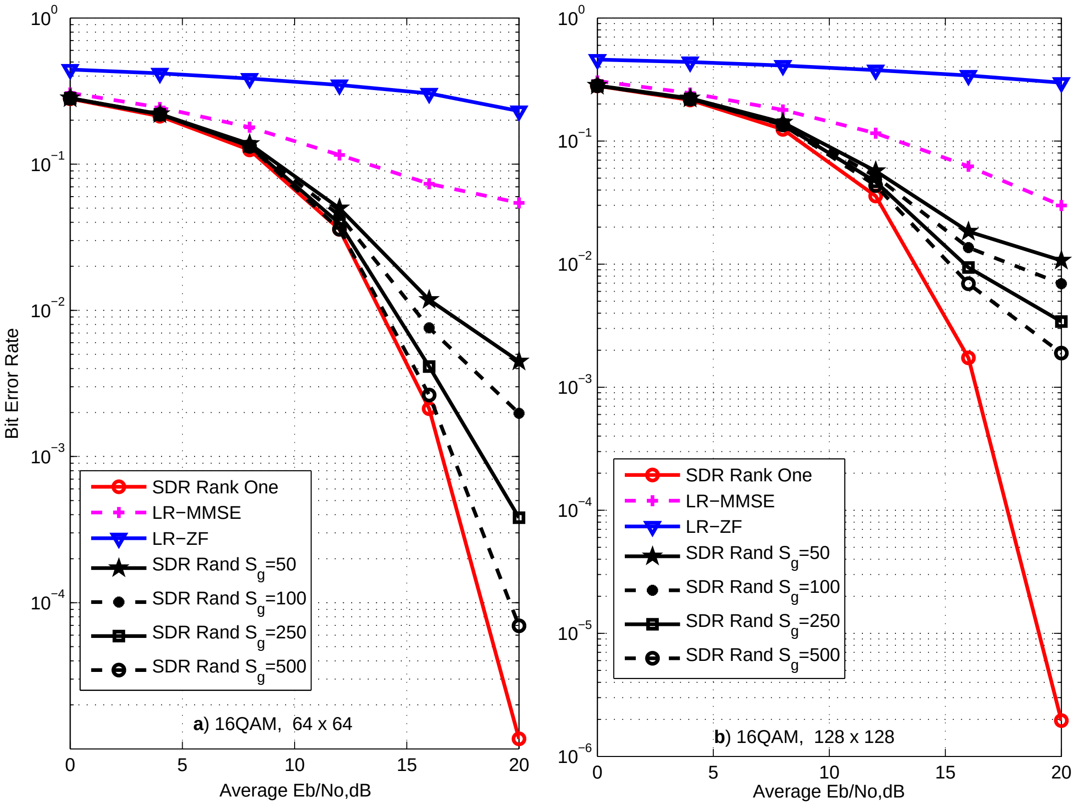


Figure 3.2: BER performance for 16-QAM SDR and LR-aided linear MIMO detectors equipped with a) 64 × 64 antennas and b) 128 × 128 antennas

3.4.2 Complexity

According to (GOLUB; LOAN, 1996), the algorithm complexity can be evaluated in terms of the total number of floating-point operations (flops), where one flop is defined as a unitary addition, subtraction, multiplication or division between two floating point numbers. Using this methodology, the complexity of MIMO detectors showed in Table 3.1 was determined, where $n = n_R = n_T$ is the number of receive and transmit antennas, respectively and M is the modulation order in M-QAM constellation.

Table 3.1: MIMO Detectors Complexity

Detector	**Total Complexity**
ML	$M^n(4n^2 + 2n)$
LR-ZF	$20/3n^3 + 10n^2 + 4n + f_{LLL}(n, \rho)$
LR-MMSE	$32n^3 + 14n^2 + 3n + f_{LLL}(n, \rho)$
SDR Rank One	$\frac{16}{3}n^3 + 12n^2 + \frac{32}{3}n + 1$
SDR Rand.	$\frac{16}{3}n^3 + 12n^2 + \frac{38}{3}n + 2 + (8n^2 + 26n + 10).S_g$

It was analyzed the complexity for the SDR by evaluating the number of real operations for the rank approximation and the Gaussian randomization, where S_g is the adopted number of generated symbols stored in vector $\mathbf{k}$, that is used to choose the nearest symbol from the original transmitted one. The computational complexity for the SDR detectors under both estimation techniques are placed near the order of $\mathcal{O}(n_t^3)$, which determines a cubic complexity for the SDR detectors.

It is important to emphasize that the order of constellation does not affect the complexity of both SDR algorithms. This characteristic is achieved by the limitations over the SDR constraints; in the literature it is called *bound constrained* SDR (SIDIROPOULOS; LUO, 2006). The specific procedure to determine the SDR detector complexity is detailed in (MUSSI; ABRAO, 2013) specifying the procedure and the auxiliary packages to perform the analysis.

For the ML approach it is simple to verify in Table 3.1 that the ML-MIMO detector is highly dependent on the constellation order (problem dimension) what requires a huge number of operations which makes it not feasible even for a low number of antennas. On the other hand, the LR-aided linear MIMO detectors approach the function $f_{\text{LLL}}(n_T)$ is an approximation for the flop count on the LLL algorithm presented at the lattice reduction procedure, this function turns out to become more and more complex to solve as the problem when the problem size gets higher which makes the BER for the LR-aided linear equalizer to shown a worst performance in comparison with the SDR approach. Moreover, a surface fitting for the flop count on LLL algorithm was suggested by (KOBAYASHI; CIRIACO; ABRÃO, 2015) and described by $f_{\text{LLL}}(n_T) = (a + c)\, n_t^3$, where $a = 5.08 \times 10^{-4}$ and $c = 8.396$. Remembering that this fitting is valid only for $n_R = n_T$ arrays.

The number of complex operations for all those considered MIMO detectors according to the number of antennas and modulation order is depicted in 3D-graphic of Fig 3.3. The SDR Rand algorithm is highly dependent on S_g which makes the complexity grows as higher as the number of samples. So as the number of antennas grow, the complexity grows proportionally leading to estimation errors. On the other hand, the SDR Rank One approach is suitable for high sized problems, leading to the lowest complexity and the best BER performance among the evaluated detection techniques.

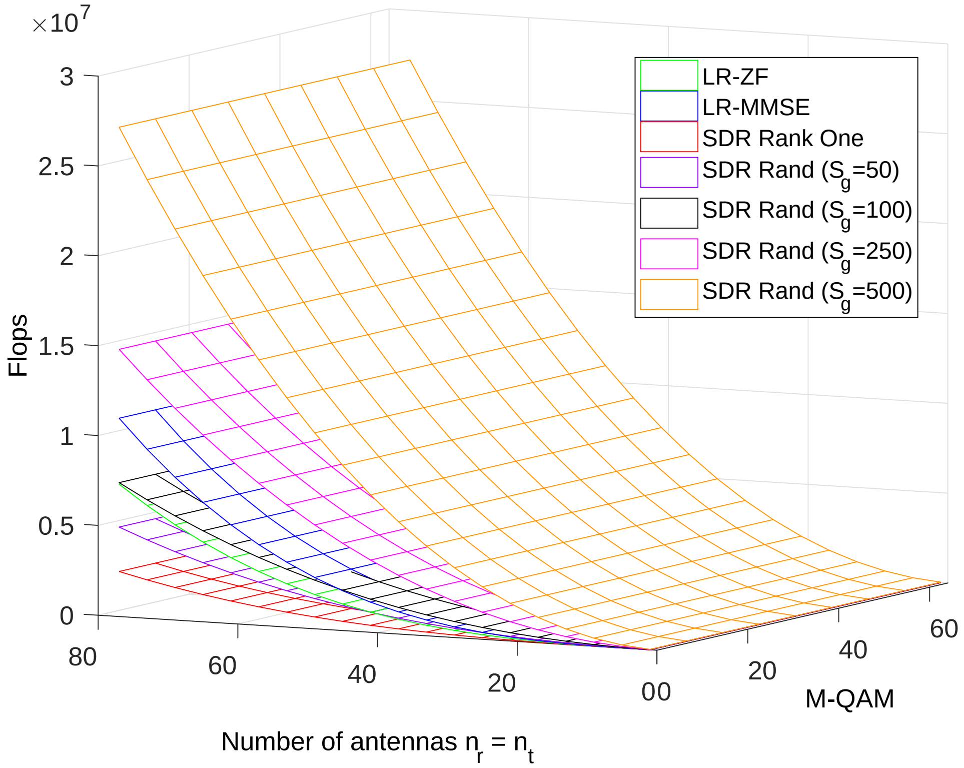


Figure 3.3: Complexity of the SDR and LR-aided linear MIMO detectors *versus* number of antennas and modulation order. For SDR Rand, S_g ranges from 50 to 500.

3.5 Conclusion

Semi-definite relaxation (SDR) technique has been applied to improve the MIMO detection performance in order to achieve near-ML performance on Rayleigh channels. The performance of SDR detectors and their respective computational complexity in term of number of operations under uncorrelated antennas were analyzed. As demonstrated, the SDR-MIMO detectors outperform the linear techniques, specially when the number of antennas increases. The lattice reduction aided MIMO detectors have an inherent advantage over the most suboptimal detectors, showing better BER performance over them, the SDR based detector outperform the LR based linear MIMO detectors, specially when the number of antennas increases substantially.

The complexity of the SDR based detectors was reduced by a semi-definite relaxation, which offers similar performance when compared with the conventional LR-aided linear MIMO detectors. As a consequence, the SDR approach presents considerable performance gain with a similar complexity, resulting in a promising solution for high order modulation MIMO systems equipped with a medium-high number of antennas.

4 Precoding and Beamforming for Large Scale MU-MIMO Downlink Channels

At the previous chapters, no channel state information (CSI) was assumed at the transmitter side (CSIT), and also equal power distribution were considered for all transmission antennas. In this chapter, assuming perfect channel knowledge at the transmitter, we will be able to analyze the beamforming for the downlink in a multi-user massive multiple-input multiple-output (MU-MIMO) channel with two different power allocation schemes: an equal power (EP) allocation and the water-filling (WF) strategy, which is the optimal solution in terms of maximizing the capacity of the communication system.

In a point-to-point or single-user (SU) MIMO, studied in the previous chapters, it was clearly demonstrated that the promised capacity and performance gains of a SU-MIMO systems with increasingly number of antennas are unachievable, mainly due to the antenna correlation. With a multi-user (MU) scenario deployment, the inherent SU-MIMO transmission problem can be largely surpassed with the MU diversity, i.e., by sharing the spacial dimension not only between the antennas of a single user, but among multiple (non-cooperative) users (WAGNER et al., 2012). The channel for an MU-MIMO transmission is commonly refereed to as the MIMO broadcast channel (BC) or MU downlink channel. Despite being much more robust to channel correlations than the SU-MIMO, the downlink MU-MIMO experiences inter-user interference (IUI) at the receiver which can only be efficiently mitigated by appropriate processing at the transmitter, with the channel awareness; hence, precoding design becomes essential in the IUI mitigation.

Furthermore, in the next generation of communication systems (5G), due to the increasing demand for higher data rates transmission and the exponential growth in the number of users, interference has turn into one of the major limiting factors for performance and capacity of wireless cellular systems. In downlink transmission of an MU-MIMO scenario, interference between users is a major source of system errors, and schemes that cancel it without the need a major detection collaboration are of great interest. These methods are commonly defined in the category of base station precoding and generally rely on the channel state information (CSI) knowledge.

Beamforming is a widely known technique for interference reduction and directed transmission of energy in the presence of noise and interference. In MIMO systems, the beamforming technique exploits channel knowledge at the transmitter side to maximize the SNR at the receiver by transmitting in the direction of the eigenvector corresponding to the largest eigenvalue of the channel (MORADI; DOOSTNEJAD; GULAK, 2011), while cancel the transmission in other directions. Furthermore, beamforming can also be used in the downlink of multi-user system aiming at maximizing the signal-to-interference-plus-noise (SINR) of a particular user (BENGTSSON; OTTERSTEN, 1999).

From information theory perspective it is proved that the sum capacity of the MIMO broadcast (spatial multiplexing mode) channel can be achieved through the technique known as dirty-paper-coding (DPC) (COSTA, 1983). However, DPC is a nonlinear precoding scheme and for most practical communication systems it is not feasible due to its very high computational complexity. Due to this reason, researches have focused on sub-optimal approaches. In contrast to the DPC, it has been showed in (YANG; MARZETTA, 2013; WIESEL; ELDAR; SHAMAI, 2008) that sub-optimal linear precoders, such as the matched filtering (MF) precoding, also known as conjugate beamforming, and the zero-forcing beamforming (ZFBF) can be applied to ensure much lower computational complexity and still providing good performance in terms of achievable sum-rate in the massive MIMO context.

Specifically, the ZFBF is a sub-optimal linear precoding or transmit beamforming strategy that is able to cancel the IUI simply by pre-multiplying data symbols with the inverse of the channel matrix. However, it is showed in (PEEL; HOCHWALD; SWINDLEHURST, 2005) that the sum-capacity of the ZFBF does not grow linearly with the number of users while channel inversion-based precoding strategies can become a serious concern when the channel becomes ill-conditioned. Hence, in order to handle this problem, a regularization parameter is introduced in the channel inversion and by that the corresponding sum-capacity scales linearly the number of users, but under a slower rate that achieved by the optimal DPC (PEEL; HOCHWALD; SWINDLEHURST, 2005). This beamforming technique is called *regularized channel inversion* (RCI) and it does not cancel the inter-user interference completely as the ZFBF, but also controls the amount of interference introduced to each user. Therefore, this regularization parameter should be optimally chosen to maximize some performance indicators, such as the SINR. The optimal regularization parameter when the number of BS antennas is equal to the number of users was derived in (PEEL; HOCHWALD; SWINDLEHURST, 2005), while in (NGUYEN; EVANS, 2008) a generic case was derived by using a large system analysis.

An alluring investigation about precoding techniques at single-cell MU-MIMO systems downlink is carried out in (YANG; MARZETTA, 2013). The authors compared the MF precoding and ZFBF with respect to spectral-efficiency and radiated energy-efficiency in a single cell scenario. It is showed that, for high spectral-efficiency and low energy-efficiency, ZFBF outperforms MF, while in the case at low spectral-efficiency and high energy-efficiency the opposite is true. An equivalent result for the uplink can be found in (NGO; LARSSON; MARZETTA, 2013); the authors have demonstrated that in a low SNR ratio, the simple maximum ratio combining (MRC) receiver outperforms the ZF receiver. This can be explained by the fact that, at low power levels, the inter-user interference introduced by the MRC receiver is occasionally less than the noise enhancement caused by the ZF detector, hence, the simple MRC detector becomes a better alternative. This result is analog for the downlink.

Another interesting analyses, which is the inspiration for this work, are presented by (COUILLET; DEBBAH, 2011) and (MUHARAR; EVANS, 2011). In these works, a large system limit for the SINR was derived under a single-cell MU-MIMO scenario, their goal is achieved by providing a deterministic expression that summarizes the SINR for a user in the asymptotic limit. In (COUILLET; DEBBAH, 2011), the main results are based on the large limit SINR derivation, considering a uniform circular array at the BS. On the other hand, (MUHARAR; EVANS, 2011) also provides a power allocation scheme under the large system limit SINR for the RCI precoder, providing a water-filling based resource allocation scheme.

Finally, in this Chapter we consider the linear sub-optimal ZFBF and RCI precoding techniques. In order to compare their achievable sum rates, we will also take advantage of a power allocation technique to maximize the sum rate capacity in both precoding approaches. It will be deployed the water-filling strategy (KHALIGHI et al., 2001) which is a power allocation technique that consists in increasing the transmission power for the streamers that experience better channel condition in order to maximize the overall capacity, with the price of an unfair resource (power) allocation for those users that experience worst channel conditions. Based on such assumptions, the objective of this Chapter consists in carrying out a comparative analysis between the ZFBF and the RCI precoding in terms of sum-rate capacity and BER performance, taking advantage of the water-filling power allocation strategy for both precoders in order to maximize the overall capacity of the channel, considering a 5G-like multi-user Massive-MIMO scenario.

The main contributions of this work is due to the aggregation of the user grouping power allocation scheme, where all users belong to the same group experience the same path-loss channel. Hence, the main objective is to determine how the users distribution in the cell could impact on the overall capacity and how the system can manipulate this information, in order to choose the best group user at the available configuration that maximizes the sum-rate capacity. We have also derived the large system limit SINR for the RCI precoder based on the assumptions of (MUHARAR; EVANS, 2011) in order to enable such analysis.

4.1 System Model

In this section, we consider the downlink of a single-cell MU-MIMO broadcast channel, depicted in Figure 4.1, where the BS is equipped with M antennas that transmit to K single-antenna user terminals. It is also considered that the slow-varying path-loss between the BS and the receiver user k is denoted by a_k.

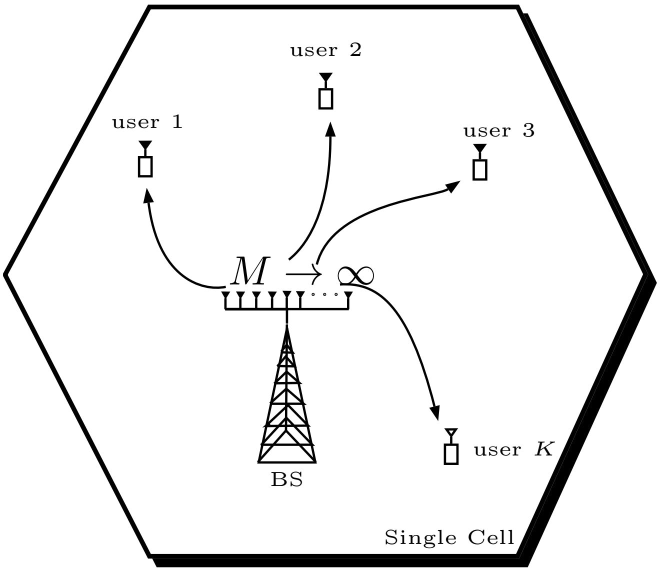


Figure 4.1: Single Cell MU-MIMO System

In the work context it will be considered an unbounded path-loss model which the transmitted signal power decays accordingly to $a_k = \dfrac{1}{r_k^{\mathfrak{b}}}$, where r is the distance between BS and the related user k and $\mathfrak{b}$ is the path-loss exponent that is related to the signal decay behavior in the cell which is scaled from $\mathfrak{b} = 2$ to $\mathfrak{b} = 6$ being respectively represented by the free-space attenuation and urban obstruction. Essentially, in order to summarize the path-loss exponent for outdoor environments we define in Table 4.1 the typical path-loss exponent ranges.

Table 4.1: Typical Path Loss Exponents

Environments	Path-Loss Exponent, $\mathfrak{b}$
Free Space	2
Urban Area cellular radio	$2.7 - 3.5$
Shadowed Urban cellular radio	$3 - 5$
In building LOS	$1.6 - 1.8$
Obstructed in building	$4 - 6$
Obstructed in factories	$2 - 3$

Source: Rappaport, Blankenship e Xu (1997).

With these considerations, the received signal vector $\mathbf{y} \in \mathbb{C}^K$ of a narrowband communication is given by

$$\mathbf{y} = \mathbf{A}\mathbf{H}\mathbf{x} + \mathbf{n} \tag{4.1}$$

where, $\mathbf{x} \in \mathbb{C}^M$ is the transmit vector, $\mathbf{A} \in \mathbb{R}^{K\times K} \triangleq \operatorname{diag}(a_1, \ldots, a_k)$ is the path-loss matrix, $\mathbf{H} \in \mathbb{C}^{K\times M}$ the channel matrix and noise vector $\mathbf{n} \sim \mathcal{CN}\{0, \sigma_n^2 \mathbf{I}_K\}$. The transmit signal vector $\mathbf{x}$ is obtained from the product of the symbol vector $\mathbf{s} \sim \mathcal{CN}\{0, \mathbf{I}_K\}$ which is normalized in power, i.e, $\mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right] = \mathbf{I}_k$. A linear precoding $\mathbf{G} \in \mathbb{C}^{M\times K} \triangleq [\mathbf{g}_1, \ldots, \mathbf{g}_k]$ and the power matrix $\mathbf{P} \in \mathbb{R}^{M\times M} = \operatorname{diag}(p_1, \ldots, p_k)$ allocated for each user define the transmit vector:

$$\mathbf{x} = \mathbf{P}^{1/2}\mathbf{G}\mathbf{s} \tag{4.2}$$

Indeed, the transmit vector $\mathbf{x}$ can also be expressed as a linear combination of the independent user symbols s_k:

$$\mathbf{x} = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{g}_k s_k \tag{4.3}$$

where $\mathbf{g}_k \in \mathbb{C}^{1\times K}$ and $p_k \geq 0$ are the precoding vector and the signal power of the k-th user. We are working under the assumption of perfect channel state information at the transmitter (CSIT), consequently the user k has perfect knowledge of $\mathbf{h}_k$ and the effective channel $\mathbf{h}_k^H \mathbf{g}_k$. The precoding vectors are normalized to satisfy the average power constraint

$$\operatorname{tr}\left(\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\mathrm{H}}\right]\right) = \operatorname{tr}\left(\mathbf{G}\mathbf{G}^{\mathrm{H}}\right) \leq P \tag{4.4}$$

where $P \geq 0$ is the total available transmit power. The SNR at the receiver is denoted as $\gamma = \frac{P}{\sigma^2}$.

With the previous considerations, the received DL signal y_k at the k-th user is given as:

$$y_k = a_k \sqrt{p_k} \mathbf{h}_k^H \mathbf{g}_k s_k + \sum_{i=1, i \neq k}^{K} a_k \sqrt{p_i} \mathbf{h}_k^H \mathbf{g}_i s_i + n_k \tag{4.5}$$

where $\mathbf{h}_k^H \in \mathbb{C}^{1 \times M}$ is here referenced as the k-th row of the channel matrix $\mathbf{H}$. Following that, the SINR per user is represented as follows (COUILLET; DEBBAH, 2011):

$$\text{SINR}_k = \frac{a_k^2 p_k \left|\mathbf{h}_k^H \mathbf{g}_k\right|^2}{\sum\limits_{j=1, j \neq k}^{K} a_k^2 p_j \left|\mathbf{h}_k^H \mathbf{g}_j\right|^2 + \sigma_n^2} \tag{4.6}$$

In the MIMO broadcast channel the distance between the users is supposed to be large enough compared to the signal wavelength λ; hence causing the users being assumed uncorrelated, i.e, $\sqrt{\mathbf{R}_h} = \mathbf{I}_K$. For the numerical simulation-based analyses, we will use the correlation models for ULA at the transmitter of the BS, considering a dense pack of antennas and following the correlation models derived at Sections 2.1.2.

Recovering the initial assumption, the normalized rate of user k is given as

$$\mathcal{R}_k = \log_2(1 + \text{SINR}_k) \qquad \text{[bits/s/Hz]} \tag{4.7}$$

Moreover, the *ergodic* sum-rate capacity with equal transmit power allocation (EP) is given by

$$\mathcal{R}_\Sigma = \mathbb{E}\left[\sum_{k=1}^{K} \log_2(1 + \text{SINR}_k)\right] \tag{4.8}$$

where the expectation is taken over the random channel realizations $\mathbf{h}_k$.

4.2 Linear Beamforming Schemes

In this section we derive the sum-rate capacity for each linear precoding case at the transmitter with equal power allocation. These results will be used in Section 4.3 to solve a power allocation problem with the perspective of maximize the sum-rate capacity of the correlated channel.

4.2.1 Zero-Forcing Beamforming

The ZFBF precoding, also refereed as channel inversion (CI) precoding, eliminates all the inter-user interference by performing an inversion of the channel matrix $\mathbf{H}$ at the transmitter side (WAGNER et al., 2012). ZFBF is largely applied to MU-MIMO networks with single antenna users (YANG; MARZETTA, 2013; COUILLET; DEBBAH, 2011); due to its simplicity on designing beamforming vectors $\mathbf{g}_k$ it makes the users to receive data free of interference, which can be attained thanks to the orthogonality imposed by the beamforming vectors for different users.

The precoding matrix for the ZFBF is given by (WAGNER et al., 2012)

$$\mathbf{G}_{\text{ZF}} = \alpha \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \tag{4.9}$$

where α is a parameter added to ensure the transmission power constraint. Indeed, α is chosen to satisfy the power constraint $\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right] = P$ and it is build only upon the channel realization $\mathbf{H}$, which is given by

$$\alpha^2 = \frac{P}{\text{tr}\left(\left[\mathbf{H}\mathbf{H}^{\text{H}}\right]^{-1}\right)} \tag{4.10}$$

The received vector can be represented as

$$\begin{aligned} \mathbf{y} &= \alpha \mathbf{A}\mathbf{H}\mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \mathbf{s} + \mathbf{n} \\ &= \alpha \mathbf{A}\mathbf{s} + \mathbf{n} \end{aligned} \tag{4.11}$$

where $\mathbf{n} = [n_1, \ldots, n_k]^T$ with $\mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right] = \sigma_n^2 \mathbf{I}_K$. With these considerations, the SINR of user k under ZFBF precoding is given as

$$\text{SINR}_{k,\text{ZF}} = \frac{\alpha^2 a_k^2}{\sigma_n^2}, \qquad \text{with } K < M \tag{4.12}$$

By using the ZFBF precoder, multi-user interference (MUI) is completely eliminated (zero-forced), in addition, the pre-coding vector is constructed to eliminate the interference that a particular user may cause to others, which is called inter-user interference (IUI).

Notice that the ZFBF has a limited number of users, which is bounded by the number of BS antennas, $K < M$. If the number of single-user antennas, K, increase beyond M the MUI is still present in the system and the SINR described in (4.12) does not hold. For cases where $K > M$, the received signal is given by

(MUHARAR, 2012)

$$\begin{aligned}\mathbf{y} &= \alpha \mathbf{A}\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{s} + n_k \\ &= \alpha a_k \mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{h}_k^H s_k + \alpha \sum_{j=1, j\neq k}^{K} a_k \sqrt{p_j}\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{h}_j^H s_j + n_k\end{aligned} \tag{4.13}$$

Leading to a SINR of user k, given by

$$\text{SINR}_{k,\text{ZF}}^{\text{ovl}} = \frac{\alpha^2 a_k^2 p_k \left|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{h}_k^H\right|^2}{\alpha^2 \sum_{j=1,j\neq k}^{K} a_k^2 p_j \left|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{h}_j^H\right|^2 + \sigma_n^2} \quad \text{with } K > M \tag{4.14}$$

where $\alpha^2 \sum_{j=1,j\neq k}^{K} a_k p_j \left|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{h}_j^H\right|^2$ represents the MUI.

Since the $\text{SINR}_{k,\text{ZF}}$ is proportional to α^2, a rank deficiency on the channel correlation matrix $\mathbf{H}\mathbf{H}^H$ will lead to a penalty in α^2 and consequently in the SINR. This motivates the addition of a regularization parameter, which is the main purpose of our next detector.

4.2.2 Regularized Channel Inversion

The RCI precoding can be faced as a generalization of the ZFBF precoding, in which the regularization parameter added to the pseudo-inverse has the ability to tune up the precoder between conventional ZF and matched filter (MF) schemes (BJORNSON; BENGTSSON; OTTERSTEN, 2014). Basically, to compensate a possibility of an ill-conditioned channel matrix $\mathbf{H}$, the regularization term is added within the pseudo-inverse of the ZF precoding matrix, as described in Eq. (4.9).

Accordingly to Muharar e Evans (2011), considering a system equipped with the RCI precoder, the precoding matrix is given as:

$$\mathbf{G}_{\text{RCI}} = \alpha \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \xi \mathbf{I}_K\right)^{-1} \tag{4.15}$$

or equivalently

$$\mathbf{G}_{\text{RCI}} = \alpha \left(\mathbf{H}^H\mathbf{H} + \xi \mathbf{I}_M\right)^{-1}\mathbf{H}^H \tag{4.16}$$

where α is used to normalize the transmit power constraint (4.4), and $\xi > 0$ is the regularization parameter. Using the representation (4.16) for the RCI precoder, the normalizing constant α is chosen to satisfy the power constraint $\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right] = P$. As stated in Wagner et al. (2012), by assuming independence between the data symbols, the total power constraint normalization is expressed as

$$\alpha^2 = \frac{P}{\mathrm{tr}\left(\mathbf{H}\left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \xi\mathbf{I}_{\mathrm{M}}\right)^{-2}\mathbf{H}^{\mathrm{H}}\right)} \tag{4.17}$$

Using the RCI precoder (4.16) the received signal for each user can be expressed as

$$y_k = \alpha a_k \sqrt{p_k}\mathbf{h}_k\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H s_k + \sum_{j=1, j\neq k}^{K} \alpha a_k \sqrt{p_j}\mathbf{h}_k\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_j^H s_j + n_k \tag{4.18}$$

where the first term in the right-hand side is the desired signal for each user k while the other terms are the interference introduced by the other users plus the receive thermal noise. Assuming single-user decoding at the receiver and treating interference as noise in the system, the SINR for each user is expressed as follows

$$\mathrm{SINR}_{k,\mathrm{RCI}} = \frac{\alpha^2 a_k^2 p_k \left|\mathbf{h}_k\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H\right|^2}{\alpha^2 \sum_{j=1, j\neq k}^{K} a_k^2 p_j \left|\mathbf{h}_k\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_j^H\right|^2 + \sigma_n^2} \tag{4.19}$$

considering $\mathbf{h}_k \in \mathbb{C}^M$ the $k-$th row of $\mathbf{H} \in \mathbb{C}^{K\times M}$. This result is based on (MUHARAR; EVANS, 2011), but in our system it is considered the path-loss term a_k.

Analyzing the RCI precoder expression (4.19), specially the regularization term, ξ, it is simple to verify that it represents a mid term between the ZFBF precoding and the conjugated beamforming (BF) precoder. Considering the case where $\xi \to \infty$, the RCI precoder converges to the BF precoder, which is given by

$$\mathbf{G}_{\mathrm{BF}} = \xi\mathbf{H}^H \tag{4.20}$$

since the term $\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{H}^H$ of the RCI will tend to $\xi\mathbf{H}^H$ as ξ tends to infinity. However, since the BF precoding is not one of our work effort, it will not be considered in our analysis; for more details over the BF precoding assumption the reader is refereed to (HOYDIS; BRINK; DEBBAH, 2013) and (YANG; MARZETTA, 2013). Now, in cases where $\xi \to 0$, the RCI precoder converges to the ZFBF precoder; this relationship is direct, since the term $\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{H}^H$ will tend to $\left(\mathbf{H}^H\mathbf{H}\mathbf{H}^H\right)$ as ξ tends to zero. With the intention to summarize the precoders presented in this chapter and to provide a better comparison between them, the precoding matrix and SINR of each precoder structure were organized in Table 4.2.

Table 4.2: Precoding structure, matrix and SINR relations

Precoding	**Matrix G**	**SINR$_k$**
ZFBF ($K < M$)	$\alpha \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1}$	$\dfrac{\alpha^2 a_k^2}{\sigma_n^2}$
ZFBF ($K > M$)	$\alpha \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1}$	$\dfrac{\alpha^2 a_k^2 p_k \left\|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1} \mathbf{h}_k^H\right\|^2}{\alpha^2 \sum_{j=1, j\neq k}^{K} a_k^2 p_j \left\|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H}\right)^{-1} \mathbf{h}_j^H\right\|^2 + \sigma_n^2}$
RCI	$\alpha \left(\mathbf{H}^H\mathbf{H} + \xi \mathbf{I}_M\right) \mathbf{H}^H$	$\dfrac{\alpha^2 a_k^2 p_k \left\|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_k^H\right\|^2}{\alpha^2 \sum_{j=1, j\neq k}^{K} a_k^2 p_j \left\|\mathbf{h}_k \left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_j^H\right\|^2 + \sigma_n^2}$

4.3 Power Allocation Scheme

In this section, we discuss the optimal power allocation scheme applied to the RCI precoding, we also discuss the ZFBF as an special case. Firstly, we must define the weighed sum-rate maximization problem with a generic precoder, subject to the power constraint, which is formulated as

$$\begin{aligned} \underset{\mathbf{g}_k,\, \forall k=1,\ldots,K}{\text{maximize}} \quad & \mathcal{R}_\Sigma = \mathbb{E}\left[\sum_{k=1}^{K} \log_2\left(1 + \frac{p_k \left|\mathbf{h}_k^H \mathbf{g}_k\right|^2}{\sum_{j=1,j\neq k}^{K} p_j \left|\mathbf{h}_k^H \mathbf{g}_j\right|^2 + \sigma_n^2}\right)\right] \\ \text{s.t.} \quad & \mathbb{E}\left[\sum_{k=1}^{K} p_k \left|\mathbf{g}_k\right|^2\right] \leq P \end{aligned} \tag{4.21}$$

where P is the total power available in transmission.

Considering the RCI precoding deployment as our beamforming matrix at the BS, $\mathbf{G} = \alpha \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \xi \mathbf{I}_K\right)$, the sum-rate maximization problem can be written as

$$\begin{aligned} \underset{\mathbf{p},\, \xi>0}{\text{maximize}} \quad & \mathcal{R}_\Sigma = \mathbb{E}\left[\sum_{k=1}^{K} \log_2\left(1 + \text{SINR}_{k,\text{RCI}}\right)\right] \\ \text{s.t.} \quad & \mathbb{E}\left[\sum_{k=1}^{K} p_k \mathbf{h}_k (\mathbf{H}^H\mathbf{H} + \xi \mathbf{I}_M)^{-2} \mathbf{h}_k^H\right] \leq P \\ & p_k \geq 0, \quad k = 1, \ldots, K \end{aligned} \tag{4.22}$$

where $\mathbf{p} = [p_1, p_2, \ldots, p_k]^T$.

For conventional MIMO, the RCI power allocated problem represented by Eq. (4.22), is still a non-convex problem; besides the concavity of the logarithm function, the constraint $\mathbb{E}\left[\sum_{k=1}^{K} p_k \mathbf{h}_k (\mathbf{H}^H\mathbf{H} + \xi \mathbf{I}_M)^{-2} \mathbf{h}_k^H\right] \leq P$ is not closed and accept infinity solutions, which for any given ξ a locally optimal power allocation scheme can be obtained. So, our effort is to derive a global optimum ξ which ensures a maximum value for the SINR function.

In Wagner et al. (2012), the solution to find a global optimum came by deriving the large system limit of the RCI precoder SINR, basically the authors have formulated an equation to ensure the optimal ξ based on finding the maximum argument of the *ergodic* sum-rate for each user, via 1-D line search. This motivate us to realize the large system analysis in section 4.4; such study enables us to find an asymptotically optimal value of ξ which maximizes the sum-rate capacity. The convexity analysis of the RCI is performed over section 4.4.1 and the *Proof* were derived in Appendix A.3.

Narrowing down the problem and considering the special case where $\xi = 0$, we have the power allocation problem applied in the ZFBF precoder. Now, respecting the limit number of user such as $K < M$ the problem can be expressed as

$$\begin{aligned} \underset{\mathbf{p},\, \xi>0}{\text{maximize}} \quad & \mathcal{R}_{\Sigma} = \mathbb{E}\left[\sum_{k=1}^{K} \log\left(1 + \text{SINR}_{k,\text{ZF}}\right)\right] \\ \text{s.t.} \quad & \mathbb{E}\left[\textstyle\sum_{k=1}^{K} p_k \mathbf{h}_k (\mathbf{H}^H\mathbf{H})^{-2} \mathbf{h}_k^H\right] \leq P \\ & p_k \geq 0, \quad k = 1, \ldots, K \end{aligned} \tag{4.23}$$

and even though the original sum-rate maximization problem is a non-convex optimization problem, the interference-free constraint of ZFBF precoder simplifies the SINR expression, because the multiplication $\mathbf{h}_k^H \mathbf{g}_k, \forall j \neq k = 0$ and all off-diagonal elements become zero. This feature guarantees the semi-positive definite characteristic for the precoded channel, enabling the optimal power allocation. Furthermore, the solution for the optimal power allocation scheme in (4.23) is given by the water-filing and can be easily solved by:

$$\begin{cases} p_k^* = \left[\mu - \dfrac{1}{k_i}\right]^+, & i = 1, 2, \ldots, K \\ \mu = \dfrac{1}{K_A}\left(P + \displaystyle\sum_{i=1}^{K_A} \dfrac{1}{k_i}\right) \end{cases} \tag{4.24}$$

where, k_i denotes the ith diagonal element of $\left(\mathbf{H}^H\mathbf{H}\right)^{-1}$. Notice that the solution allows K_A active antennas, so number of the active ones, K_A, must be determined to solve the problem properly.

Solution (4.24) it is the classical water-filling solution, where the parameter

μ is the water level. In this solution, the transmit antennas that are not capable to achieve a minimum SNR level are shut down and their power is redistributed among the other active remaining ones.

4.4 Large System SINR Analysis

In this section, our goal is to provide a large limit analysis for our system model. Specifically analyzing the SINR of the RCI in a large system limit where both number of users K and number of transmit antennas M approach infinity, we also assume $M \geq K$; and M and K are both large with their ratio $\beta = \frac{K}{M}$ being constant. Large system limit analysis expressions for the asymptotic sum-rate capacity where previously studied in (WAGNER et al., 2012; MUHARAR; EVANS, 2011). Our scheme is motivated by these works, however it will be applied in a downlink massive MIMO scenario equipped with uniform linear array antennas at the BS.

In this subject some random matrix theory tools have to be defined. Firstly, a random matrix can be defined as a matrix with elements being random variables entries or a matrix-valued random variable (COUILLET; DEBBAH, 2011). In this context, we are interested in the behavior of large random Hermitian matrices, and particularly in the asymptotic distribution of their eigenvalues. Several works provide analyses over the eigenvalue distribution of Hermitian matrices, one of the pioneers were the Wigner's work (TULINO; VERDU, 2004; WIGNER, 1958), which studied the empirical eigenvalue distribution of any Hermitian matrix whose upper triangular entries are independent and zero mean with identical variance. Wigner were able to prove that the eigenvalues distribution of those particular matrices structure converges to a semicircle law.

In the field of wireless communications the application of matrices with semicircle distributions are rather limited. Generally, an objective of interest in this area is in the structure of sample covariance matrices $\mathbf{XX}^H$, where $\mathbf{X}$ is a rectangular matrix with independent entries. Those matrices have their many applications in wireless communication, such as in the ergodic capacity per antenna that is provided in previous sections. The work proposed by Marcenko e Pastur (1967) was the first considering the limiting spectral distribution of sample covariance matrices. One of the major results is known as the Marčenko-Pastur law, as defined below:

Theorem 4.4.1. *(Marčenko-Pastur Law) (TULINO; VERDU, 2004). Consider an*

$n \times N$ matrix $\mathbf{X}$ whose entries are i.i.d. complex or real random variables with zero mean and variance $\frac{1}{N}$. As $n, N \to \infty$ with $\frac{n}{N} \to \beta > 0$, the empirical spectral density of $\mathbf{X}\mathbf{X}^H$ converges almost surely to a nonrandom limiting distribution F_β with probability density function (pdf) given by:

$$f_\beta(x) = \left(1 - \frac{1}{\beta}\right)^+ \delta(x) + \frac{1}{2\pi\beta x}\sqrt{(x-a)^+(b-x)^+}, \tag{4.25}$$

where $(y)^+ = \max\{0, y\}$, $a = \left(1+\sqrt{\beta}\right)^2$, $b = \left(1-\sqrt{\beta}\right)^2$ and $\delta(\cdot)$ is the Dirac delta function.

Equivalently, the empirical spectral density of $\mathbf{X}^H\mathbf{X}$ converges almost surely to a nonrandom distribution $\tilde{F}_\beta$ whose probability density function is (TULINO; VERDU, 2004):

$$\begin{aligned} \tilde{f}_\beta(x) &= (1-\beta)\delta(x) + \beta f_\beta(x) \\ &= (1-\beta)^+\delta(x) + \frac{1}{2\pi x}\sqrt{(x-a)^+(b-x)^+}. \end{aligned} \tag{4.26}$$

The Marčenko-Pastur law provides the probability density function of singular values of large Hermitian random matrices, when the dimension of the matrix tend to infinity. As an example, Figure 4.2 illustrates the behavior of the Marčenko-Pastur law.

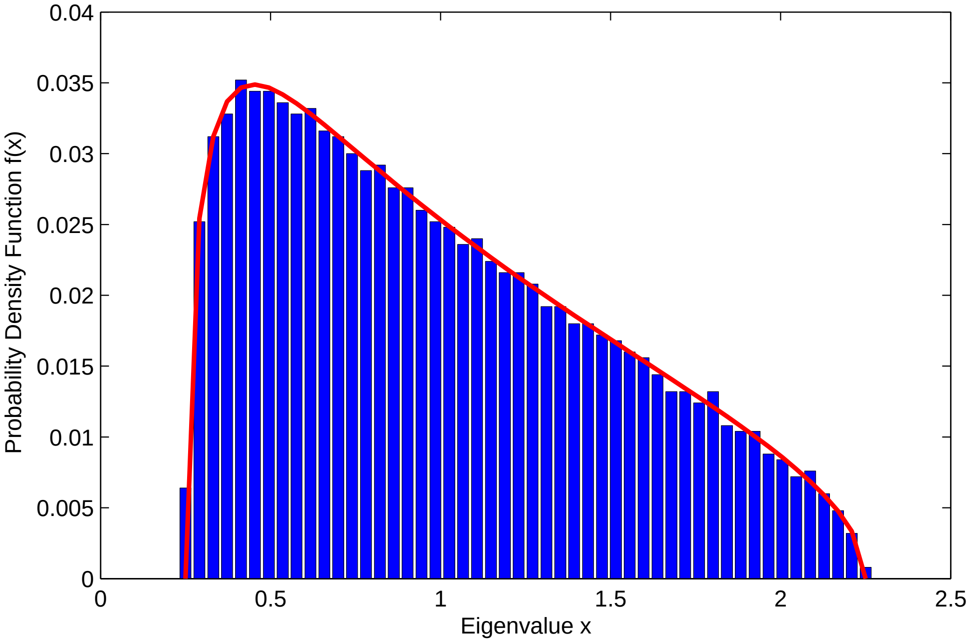


Figure 4.2: The histogram of the eigenvalues of $\mathbf{X}\mathbf{X}^H$ ($N = 2500$) vs. its Marčenko-Pastur law density for $\beta = 0.25$

For cases, where the Hermitian matrix were represented by a generic form, such as $\mathbf{Y} = \mathbf{L} + \mathbf{XTX}^H$, with $\mathbf{L}$ as a deterministic Hermitian matrix and $\mathbf{T}$ is a diagonal matrix, the analysis of the probability in the large limit may take advantage of a random matrix theory tool, known as the Stieltjes transform, which is defined in the sequel.

Definition 4.4.1. *Let X be a real-valued random variable with distribution F_X. The Stieltjes transform of F_X is defined as*

$$m_X(z) = \int_{-\infty}^{\infty} \frac{1}{\lambda - z} F_X(\lambda) d\lambda = \mathbb{E}\left[\frac{1}{X - z}\right], \tag{4.27}$$

for $z \in \mathbb{C}^+ = \{z \in \mathbb{C}, \Im(z) > 0\}$

Notice that the Stieltjes transform uniquely determines the probability distribution of a function in the large limit. A rigorous representation on the Stieltjes transform can be found in (COUILLET; DEBBAH, 2011), chapter 3.

Besides the mathematical definition of the Stieltjes transform, our development will consider an alternative form, which is related to the trace function. Now, considering $\mathbf{X} \in \mathbb{C}^{N \times N}$ as an Hermitian matrix, the Stieltjes transform of F_X, denoted by m_X, can be alternatively described as, (COUILLET; DEBBAH, 2011):

$$\begin{aligned} m_X(z) &= \textstyle\int_0^{\infty} \frac{1}{\lambda - z} F_X(\lambda) d\lambda = \frac{1}{N} \mathrm{tr}\left(\mathbf{\Lambda} - \mathrm{z}\mathbf{I}_\mathrm{N}\right)^{-1} \\ &= \frac{1}{N} \mathrm{tr}\left(\mathbf{X} - \mathrm{z}\mathbf{I}_\mathrm{N}\right)^{-1} \end{aligned} \tag{4.28}$$

where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{X}$. This directly implies that the evaluation of the normalized trace of $\left(\mathbf{X} - z\mathbf{I}_N\right)^{-1}$ is equivalent to evaluating the Stieltjes transform of F_X and vice-versa. This relation is extremely important in order to derive the limiting SINR expression.

In addition to those random matrices tools and theorems, the following lemmas also plays an important role in the analysis of the large limiting SINR. The lemmas are concentrated in provide asymptotic results for particular quadratic matrices represented in the quadratic form $\mathbf{x}^H \mathbf{A}_N \mathbf{x}$. In this context we consider $\mathbf{x} \in \mathbb{C}^{N \times 1}$ as a random row vector with i.i.d entries and $\mathbf{A}_N \in \mathbb{C}^{N \times N}$ as a deterministic complex matrix with uniform bounded spectral radius[1]. The first lemma provides the relation between matrices in the quadratic form and the trace of $\mathbf{A}_N$.

Lemma 4.4.2. *(COUILLET; DEBBAH, 2011; MUHARAR, 2012). Let $\mathbf{x} \in \mathbb{C}^{1 \times N}$ be a random row vector whose entries are i.i.d. with zero mean and variance*

[1] Spectral radius of a square matrix or a bounded linear operator is the largest absolute value of its eigenvalues, *i.e. supremum* among the absolute values of the elements in its spectrum, denoted by $\rho(\lambda)$.

$\frac{1}{N}$. *Let* $\mathbf{A}_N \in \mathbb{C}^{N\times N}$ *be any matrix with uniformly bounded spectral norm*[2] *and independent of* $\mathbf{x}$. *Then, as* $N \to \infty$,

$$\mathbf{x}\mathbf{A}_N\mathbf{x}^H - \frac{1}{N}\mathrm{tr}\left(\mathbf{A}_{\mathrm{N}}\right) \xrightarrow{\text{a.s.}} 0 \tag{4.29}$$

Secondly, an important quadratic expression, that is commonly used in precoding systems, is in the form $g_N = \mathbf{x}\left(\mathbf{B}\mathbf{B}^H + \nu\mathbf{I}_N\right)^{-1}\mathbf{x}^H$, which appears many times in the SINR expressions of a communication system.

Lemma 4.4.3. *(MUHARAR, 2012). Let* $g_N = \mathbf{x}\left(\mathbf{B}\mathbf{B}^H + \nu\mathbf{I}_N\right)^{-1}\mathbf{x}^H$, *where* $\mathbf{B} \in \mathbb{C}^{N\times n}$, *with* $n < N$ *and* $\nu > 0$. *Suppose that the elements of* $\mathbf{B}$ *are i.i.d. with zero mean and variance* $\dfrac{1}{N}$, $\mathbf{x}$ *and* $\mathbf{B}$ *are independent. Then, as* $n, N \to \infty$ *with* $\dfrac{n}{N} \to \beta$,

$$g_N - g \xrightarrow{a.s.} 0,$$

where g is given by

$$g(\beta,\nu) = \left(\nu + \frac{\beta}{1+g}\right)^{-1} = \frac{1}{2}\left(\sqrt{\frac{(1-\beta)^2}{\nu^2} + \frac{2(1+\beta)}{\nu} + 1} + \frac{1-\beta}{\nu} - 1\right). \tag{4.30}$$

The proof of this Theorem in held in (NGUYEN; EVANS, 2008).

In general, the above expression is the evaluation of Lemma 4.4.2 over $g_N = \mathbf{x}\left(\mathbf{B}\mathbf{B}^H + \nu\mathbf{I}_N\right)^{-1}\mathbf{x}^H$ which is related to the trace of $\dfrac{1}{N}\left(\mathbf{B}\mathbf{B}^H + \nu\mathbf{I}_N\right)^{-1}$, that leads to the Stietjes transform of the Hermitian matrix, $\left(\mathbf{B}\mathbf{B}^H + \nu\mathbf{I}_N\right)^{-1}$, whose probability density function obey the Marčenko-Pastur law. Thereby, the expression (4.30) can be seen as the Stieltjes transform of the Marčenko-Pastur probability distribution law.

By evaluating the SINR expression of the RCI precoder (4.19) and considering the previous theorems and lemmas, we derive the SINR at the large system limit with M and K tending to infinity with a fixed ratio $\beta = \frac{K}{M}$. The result is indicated by theorem 4.4.4.

Theorem 4.4.4. *(MUHARAR, 2012) Considering all users with the same power and same path-losses, let* $\nu = \dfrac{\xi}{M}$ *be the normalized regularization parameter,* $\gamma = \dfrac{P}{\sigma^2}$ *be the received SNR and* $g(\beta,\nu)$ *be the function defined in* (4.30). *Then, the desired signal converges almost surely to*

$$\mathsf{S}(\beta,\nu) = P\frac{g(\beta,\nu)}{(1+g(\beta,\nu))^2}\left(1 + \frac{\nu}{\beta}(1+g(\beta,\nu))^2\right) \tag{4.31}$$

[2] Spectral norm of a square matrix is the square root of the maximum eigenvalue of $\mathbf{A}^H\mathbf{A}$.

and the interference converges almost surely to

$$\mathsf{I}(\beta,\nu) = \frac{P}{(1+g(\beta,\nu))^2}. \tag{4.32}$$

Consequently, the SINR converges almost surely to a deterministic limiting SINR value, given by

$$\mathrm{SINR}^{\infty}(\gamma,\beta,\nu) = g(\beta,\nu)\frac{\gamma + \frac{\gamma\nu}{\beta}(1+g(\beta,\nu))^2}{\gamma+(1+g(\beta,\nu))^2}. \tag{4.33}$$

Prof: See Appendix A.2

From the SINR expression above, it can be seen that in the large limit the SINR is the same for all users and converges to a deterministic value that depends only upon the system parameters, such as the regularization parameter ν, cell loading β and the received SNR γ (MUHARAR, 2012).

4.4.1 Optimal System Parameters

At first sight the limiting SINR expression shows us that it is user independent and depends only on the regularization parameter ν, cell-loading β and received SNR γ. Firstly, our major concern is to evaluate how the regularization parameter affects the limiting SINR. It is important to ensure that the RCI precoding is configured in order to maximize the limiting SINR and consequently to provide the maximum limiting sum-rate capacity per user. Based on (4.33), the limiting sum-rate capacity per antenna (or single-antenna user) can be defined as

$$\mathcal{R}_k^{\infty} = \beta\log_2(1+\mathrm{SINR}^{\infty}(\gamma,\beta,\nu)). \tag{4.34}$$

One can verify that there is an one-to-one monotonic mapping between the limiting sum-rate capacity and the limiting SINR. Maximizing the limiting SINR via ν and γ will result in an equivalently increase at the maximum sum rate capacity. Analyzing the limiting interference equation, (4.32), it is clear that the interference energy is decreasing as $g(\beta,\nu)$ increases. So increasing the regularization parameter ν will increase the level of interference. Now, deriving the signal expression, (4.31), w.r.t. ν we have the signal behavior in terms of the regularization parameter variations. From (4.31), it is evident that $\mathsf{S}(\beta,\nu)$ is an increasing function with ν; for brevity, we denote $g(\beta,\nu)$ as g. Deriving the signal w.r.t. ν we have:

$$\frac{d\mathsf{S}(\beta,\nu)}{d\nu} = \frac{2g^2}{(1+g)} \cdot \frac{1}{\beta+\nu(1+g)^2} > 0, \tag{4.35}$$

Now, we must evaluate how ν affect the limiting sum rate capacity. As verified, both signal and interference increase with ν. To achieve improvements at the SINR and consequently to the limiting sum rate capacity, we should reduce ν to suppress the interference and, at the same time, we should increase ν to reinforce the desired signal. Hence, ν provides a trade-off between decreasing the interference level and increasing the signal energy and it must be precisely chosen the optimal regularization parameter value ν^* in order to maximize the system capacity. Figure 4.3, illustrates the behavior of the limiting sum rate capacity as a function of ν, considering two cell-loading configuration, the first one with 16×16 antennas and the second system with 16×12 antennas, which can be respectively related to $\beta = 1$ and $\beta = 0.75$; both scenarios were evaluated under a medium SNR regime, $\gamma = 10$dB.

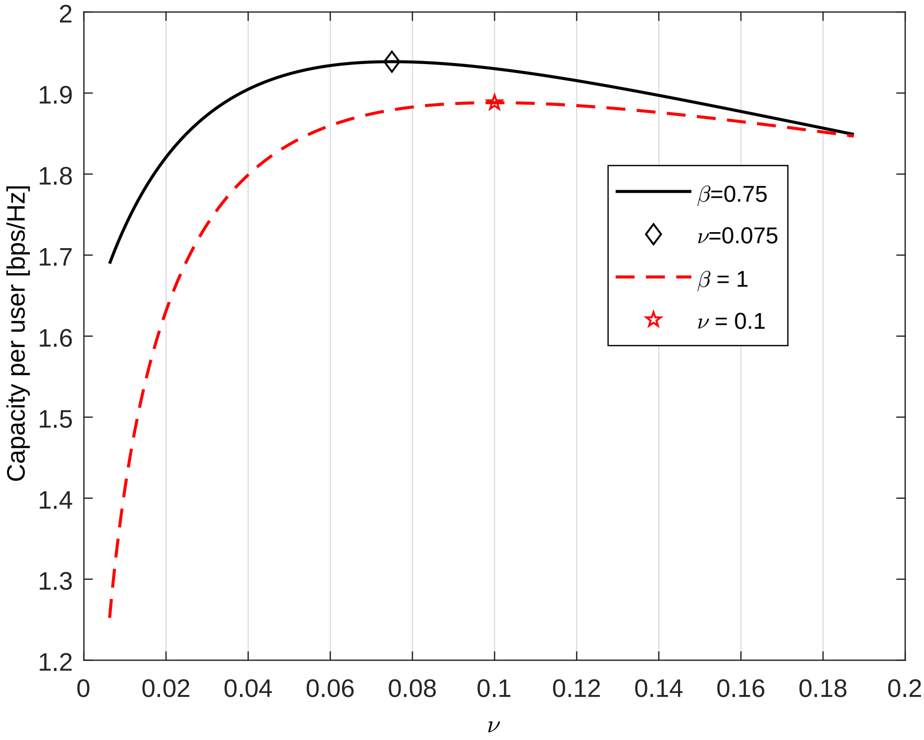


Figure 4.3: Limiting sum rate capacity for different values of ν in scenario limited to $\gamma = 10$dB of SNR.

One can observe that the limiting capacity per user have a maximum point w.r.t. ν and the determination of this optimal point is easily realized through a line-search over the objective function. The behavior of the limiting SINR w.r.t ν was studied by (MUHARAR, 2012), which concludes that the limiting SINR is a quasi-concave function of ν and it is maximized by setting $\nu = \nu^*$ where ν^* is unique and given by:

$$\nu^* = \frac{\beta}{\gamma} \tag{4.36}$$

Hence, assuming the *optimum* regularization parameter ν^*, one can conclude from (4.33) that the SINR converges almost surely to a deterministic optimum limiting SINR value, given by:

$$\mathrm{SINR}^{*\infty}(\gamma, \beta, \nu^*) \;\equiv\; g(\beta, \nu^*) \tag{4.37}$$

Consequently the maximum limiting sum-rate capacity per-user is obtained as:

$$\mathcal{R}_k^\infty = \beta \log_2(1 + \mathrm{SINR}^{*\infty}(\gamma, \beta, \nu^*)). \tag{4.38}$$

Proof: See Appendix A.3

Following this result, the expression for the optimal ν that maximizes the limiting sum rate capacity becomes very simple and is given by the ratio between the cell loading and the received SNR. This result implies that, at high SNR regime, ν^* becomes very small and the RCI precoding matrix tends to behave like the ZFBF one due to the low impact of the regularization parameter in the channel inversion. On the other hand, for cases where the cell loading is very high, *i.e*, an over loaded cell condition, the impact of ν^* in the channel inversion becomes more evident, and we should expect that the RCI precoding matrix will performs as the BF one or a matched filter. This analysis on the ν^* confirms that the RCI is a mid term between the ZFBF and the BF precoders and must be tuned properly in order to provide performance and capacity gains. This results are also discussed and confirmed by (WAGNER et al., 2012) but using a distinct approach for the large system analysis.

Another major concern related to the limiting sum-rate capacity achieved by the RCI precoder is related to the cell-loading β. Fig. 4.3 indicates that it is possible to achieve different limiting sum-rate capacities by varying β. Considering an optimal regularization parameter ν^*, we must verify how the cell-loading affects the maximum limiting SINR and consequently the maximum limiting sum rate capacity (4.38). The solution of $g(\beta, \nu^*)$ comes from (4.30), and considering the optimal regularization parameter it can given by:

$$g(\beta, \nu^*) = \left(\frac{\beta}{\gamma} + \frac{\beta}{1 + g(\beta, \nu^*)}\right)^{-1}. \tag{4.39}$$

From the above expression it is clearly that, increasing the SNR will result in improvements at the limiting SINR and consequently in the limiting sum rate. It is easy to check that (MUHARAR, 2012):

$$\frac{dg(\beta,\nu^*)}{d\beta} = -\frac{1}{\beta^2}\left(\frac{1}{\gamma} + \frac{1}{\left(1+g(\beta,\nu)\right)^2}\right)^{-1} < 0, \tag{4.40}$$

which indicates that the maximum limiting SINR is a decreasing function of the cell-loading. Analyzing the limiting sum-rate capacity, eq. (4.38), increasing β will increase the pre-log factor but decrease the log term.

The following equation defines the characterization of the sum-rate with respect to optimum cell-loading β. Accordingly to Muharar (2012), for $\gamma > 1$, the limiting sum-rate is a quasi-concave (unimodal) function of β. The unique stationary point, β^*, is given by the solution

$$\beta^* = \gamma\frac{(1+g(\beta^*,\nu^*))}{\left(\gamma+(1+g(\beta^*,\nu^*))^2\right)\log\left(1+g(\beta^*,\nu^*)\right)}. \tag{4.41}$$

Since, $\mathcal{R}_k^\infty$ is unimodal w.r.t. β, the optimum cell-loading β^* can be found effectively by using the bisection method of even a line search method to find the root of equation (VANDENBERGHE; BOYD, 1996).

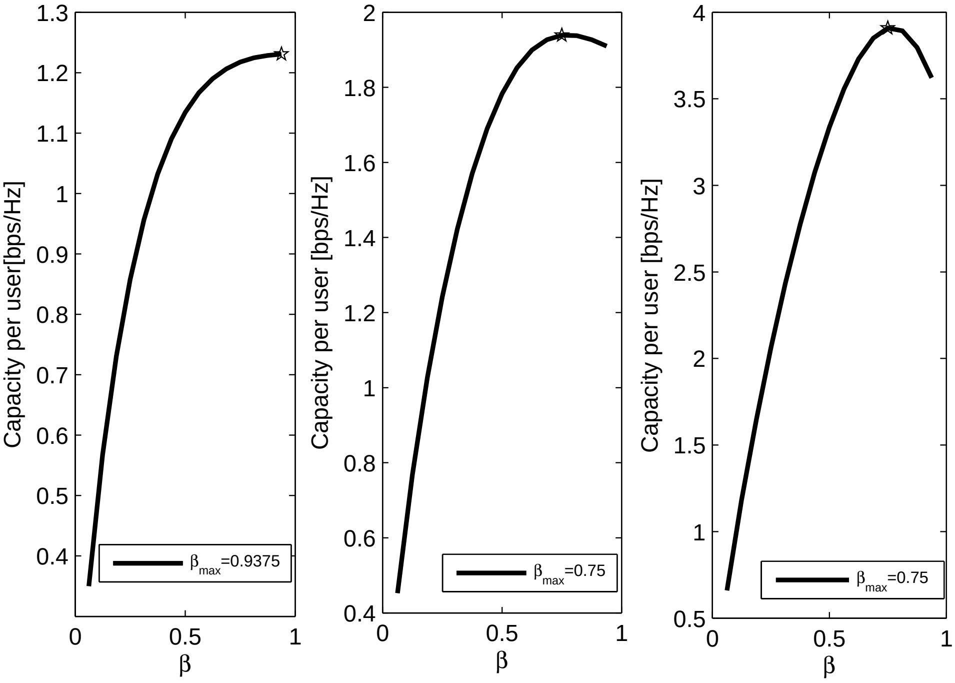
With these results it is easy to note that the limiting SINR and consequently the sum rate capacity per user have a maximum point w.r.t. the regularization parameter and the cell-loading. When it comes to the cell-loading, accordingly to (4.41), the maximum point is also function of the SNR output, and in order to evaluate the behavior of the limiting sum rate capacity w.r.t. the cell loading, it is important to consider various SNR outputs, in order to provide the best value of β for the desired SNR regime.

To confirm the previous metrics related to the limiting sum rate capacity, we simulate the RCI precoder for various cell-loading conditions, with the perspective to ensure the maximum regularization parameter, which is given by $\nu^* = \frac{\beta}{\gamma}$, and the maximum cell-loading. The simulations are presented in section 4.5 which also discuss the implementation of the water-filling strategy to maximize the limiting sum-rate capacity in different cell organization schemes.

4.5 Numerical Analysis and Results

This chapter developments are related to the simulation results for the SINR and consequently the limiting sum rate capacity of the ZFBF and the RCI precoders for different system arrangements, in the sense of number of antennas at the BS and number of users in the cell. Furthermore, it will be evaluated and compared the proposed precoding schemes under equal power distribution and under an optimal power allocation scheme, that will be given by the water-filling technique for a given correlated channel. Besides, the deterministic large scale limit SINR and sum rate capacity will be analyzed and compared aiming to verify its behavior at the maximum regularization parameter and cell-loading choices.

Firstly, for analysis purpose we consider a scenario consisting of 16 antennas in the radio base station and an increasing cell-loading until a condition of full loading. Firstly, Figure 4.4 a) characterize a system working in a low SNR regime, which is represented by the condition of $\gamma = 5$dB.



a) Low SNR $\gamma = 5$dB b) Medium SNR $\gamma = 10$dB c) High SNR $\gamma = 20$dB

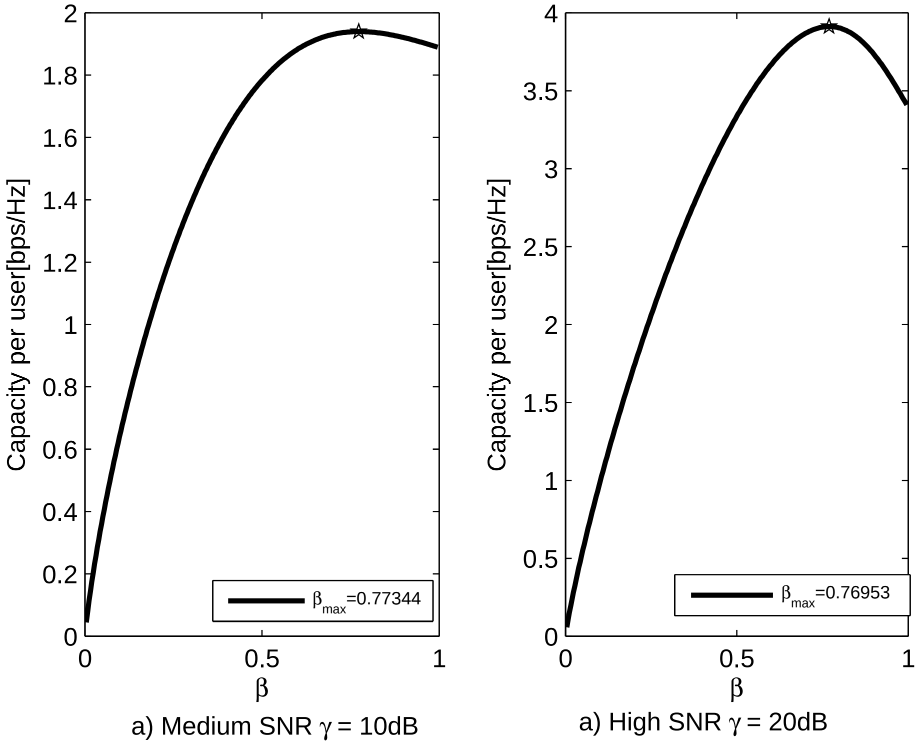
Figure 4.4: Limiting sum rate capacity for different values of β, considering ν^* in scenario with various SNR regimes and 16 antennas at the BS.

In a low SNR regime, the sum rate capacity per user increase with the cell-loading. This behavior is observed due to the high noise present in the system, so the limiting capacity per user is clearly poor and the interference effect over the increasing number of users is not verified.

Figure 4.4 b) gave us the sum rate capacity variations w.r.t. β considering a scenario of medium SNR, $\gamma = 10$dB. In this context, the signal power is considerably higher than the noise power and the interference between users effect can be verified by the decrease in capacity observed as the cell-loading increase beyond $\beta_{\max} = 0.75$. In this scenario choosing $\beta^* = \beta_{\max}$ is the best cell-loading in terms of capacity per users maximization.

Finally, Figure 4.4 c) represents the high SNR regime, $\gamma = 20$dB, in the following scenario we have a great increase in capacity due to the great signal power. One can observe that the maximum cell-loading for this case is also provided by $\beta^* = 0.75$, this result indicates that for an increasingly SNR scenarios, we may limit the number of users in the cell in order to ensure the maximum capacity per user. We also observe that for high SNR regime the decay in capacity is more intense, when compared to the medium SNR regime. This behavior is explained due to the increasingly interference between users in the cell that is observed with strong signal. So, in order to maximize the limiting sum rate capacity per user, on the RCI precoder, it is important to ensure the maximum cell-loading and regularization parameter as previously defined.

As our main objective is to provide an analysis for a large scale scenario, lets evaluate the cell-loading impact on the limiting sum-rate capacity considering a high sized system, specifically, composed by a $M = 256$ antennas at the BS.



a) Medium SNR $\gamma = 10$dB

a) High SNR $\gamma = 20$dB

Figure 4.5: Limiting sum rate capacity for different values of β, considering ν^* in scenario with various SNR regimes and 256 antennas at the BS.

Evaluating the limiting sum rate capacity in a high sized system, Eq. (4.34), it does not change much the asymptotic achievable capacity per user as seen in Figure 4.5. This is due to the fact that it is a theoretical result based on the limiting SINR, Eq. (4.33), which by its turn does not depend directly on the number of users, but mainly is in the cell-loading and in the system SNR output. One can verify that the optimal cell-loading is still near 75% ($\beta^* \approx 0.75$) even for a high sized system, presenting a minor deviation of this value. Hence, this behavior is expected asymptotically which lead us to conclude that in a single cell system the optimal cell-loading in terms of maximizing the limiting sum rate capacity is around 75% of the full loaded condition, while the respective number of users and the attainable sum-rate system capacity are:

$$K^* = \beta^* \cdot M \;\approx\; \lfloor 0.75 \cdot M \rfloor \quad \text{[users]} \tag{4.42}$$

$$\mathcal{R}_{\Sigma}^{*\infty} = \sum_{k=1}^{K} \beta^* \log_2(1 + \mathrm{SINR}^{*\infty}(\gamma, \beta^*, \nu^*)) = \sum_{k=1}^{K} \beta^* \log_2(1 + g(\beta^*, \nu^*)) \tag{4.43}$$

In this context, we already have modeled the limiting SINR and limiting sum rate capacity of a broadcast RCI channel and characterized its behavior w.r.t. the regularization parameter and the cell-loading. Figure 4.6 depicts a surface combining both the limiting sum rate capacity variations w.r.t. ν and β in a medium SNR regime scenario of $\gamma = 10dB$.

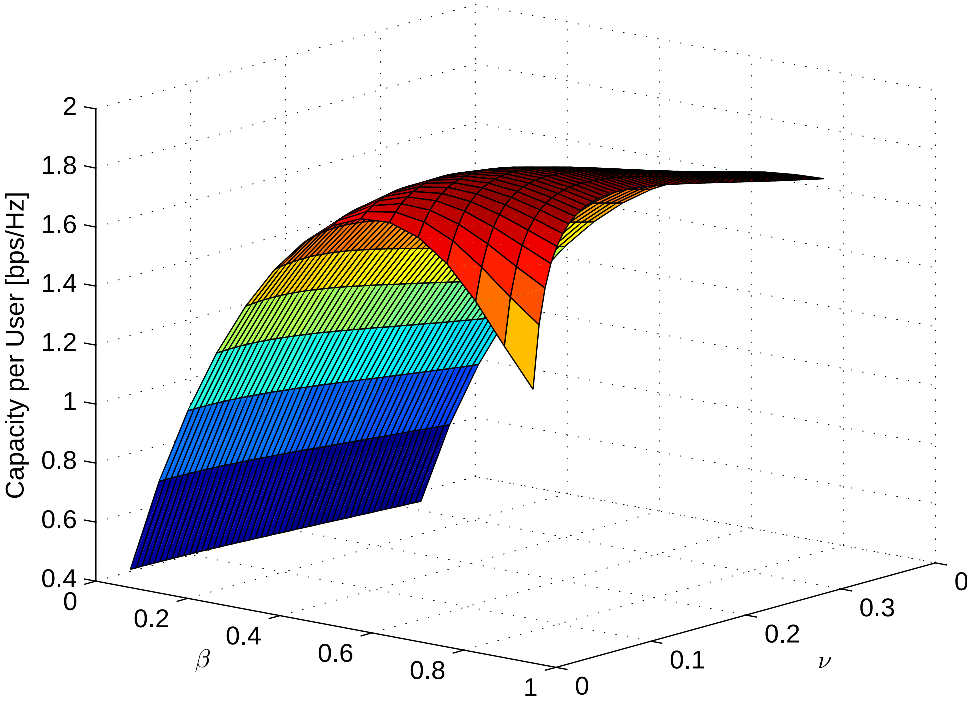


Figure 4.6: Surface representing the Limiting Capacity per user variations on the RCI precoder w.r.t. β and ν, limited to $\gamma = 10$dB

One can verify that the surface represented in Figure 4.6 have a maximum point w.r.t. ν and β, that when combined it maximizes the limiting sum rate capacity per-user. As our goal is to maximize the total sum-rate capacity of the system, ensuring the best per-user sum rate capacity is a first step towards achieve the best case scenario for the RCI precoding scheme.

As we have modeled the optimal parameter which maximizes the sum rate capacity of an RCI precoder, it was granted deterministic values for ν and β each one assuming the optimal value for a specific system configuration. The fact of having a deterministic value for the regularization parameter at the RCI precoder impacts directly on the convexity of a power allocation problem developed for the RCI precoder, represented by equation (4.22). The next section analyzes the optimal power allocation schemes for the RCI and ZFBF precoders in a scenario related to a group-wise user selection, where each group represents a specific path-loss related to each user and the impacts of power allocation scheme on this cell configuration.

4.5.1 Power Allocation and User Grouping in LS-MIMO Systems

In this section it will be considered the original SINR formulation for the RCI precoder, Eq. (4.19), considering different power and path-loss coefficients to each user which introduces a slightly difference in the limiting SINR per user, Eq. (4.33). Now, each user will have a related power and path-loss, consequently, the limiting SINR for each user is given by

$$\begin{aligned} \mathrm{SINR}_k^{\infty} &= \overline{p}_k g(\beta,\nu)\frac{\gamma_k + \dfrac{\gamma_k \nu}{\beta}(1+g(\beta,\nu))^2}{\gamma_k + (1+g(\beta,\nu))^2}. \\ &= \overline{p}_k \mathfrak{f}_k(\beta,\nu), \end{aligned} \tag{4.44}$$

where $\gamma_k = \dfrac{Pa_k^2}{\sigma_n^2}$ is defined as the effective SNR and $\overline{p}_k = \dfrac{p_k}{\mathcal{P}}$ is the normalized power w.r.t. $\mathcal{P}$. We define $\mathcal{P} = \lim\limits_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} p_k$. Different from equation (4.33), now we can observe that the limiting SINR is different for each user and depends on a_k and p_k. Also one can verify that $\mathfrak{f}_k$ is independent from $\overline{p}_k$ in (4.44), where $\mathfrak{f}_k$ represents the limiting SINR and $\overline{p}_k$ the average power allocated to each user in the cell. Such conditions will facilitate the analysis in finding the optimal power allocation that maximizes the limiting sum rate capacity.

Furthermore, in the context of this work, it will be consider a scenario where all K users are divided into a finite number of L groups and all users in the same group are assumed to have a similar path-loss. Hence, without loss of generality, we consider the path-loss power under an unbounded model where $\mathfrak{b}$ is the path-loss exponent. Under this assumption the path-loss assumes the form:

$$a_j = \frac{1}{r_j^{\mathfrak{b}}}, \quad \text{where} \quad r_j = j\frac{R}{L} \;\; \text{for} \;\; j = 1, \dots, L;$$

R is the radius of the cell. The number of users in group j is denoted by K_j, with $\sum_{j=1}^{L} K_j = K$. Since all user in the same jth group have the same path-loss and system parameters, such as β, ν and the SNR, based on Eq.(4.44), it is reasonable consider that the power allocated to each user in that group is also the same. Figure 4.7 exemplify the case for a 16 × 16 system in a $R = 4\ km$ macro-cell with user grouping (four clusters); an urban cellular radio environment has been considered, where all user in determined group will have the same associated path-loss.

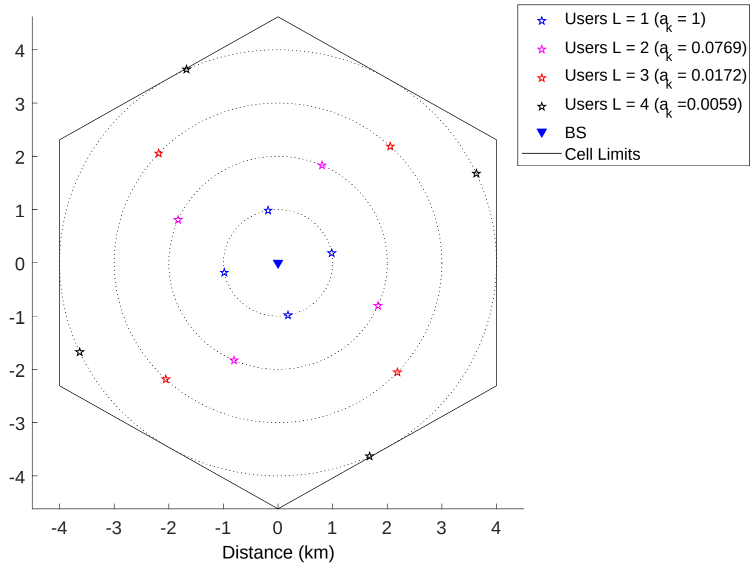


Figure 4.7: 16 × 16 system in a $R = 4\ km$ macro-cell with user grouping (four clusters), where all user in a group have the same path-loss.

Based on this scenario, we can characterize the limiting achievable sum rate capacity per antenna as follows:

$$\mathcal{R}_{\Sigma}^{\infty} = \sum_{j=1}^{L} \beta_j \log_2(1 + \text{SINR}_j^{\infty}) \tag{4.45}$$

where $\beta_j = \frac{K_j}{N}$ represents the cell-loading of the jth group. Considering the goal of finding the optimal power allocation that maximizes $\mathcal{R}_{\Sigma}$ we define $\overline{\mathbf{p}} =$

$[\overline{p}_1, \overline{p}_2, \ldots, \overline{p}_L]^T$, then a joint optimization problem can be formulated

$$\begin{aligned} \underset{\overline{\mathbf{p}},\, \nu}{\text{maximize}} \quad & \mathcal{R}_\Sigma^\infty \\ \text{s.t.} \quad & \textstyle\sum_{j=1}^{L} \beta_j \overline{p}_j \leq P \\ & \overline{p}_j > 0, \quad j = 1, \ldots, L \\ & \nu > 0 \end{aligned} \tag{4.46}$$

Note that the proposed power allocation problem is similar to the one defined in Eq.(4.22), the difference is that Eq. (4.46) represents the large scale version of the previous one. Also note that the first constraint represents the large system average power limitation and the second one guarantee that the normalized power are non-negative.

Before solving the problem (4.46), it is important to characterize the concavity of the objective function. To do so, we evaluate the concavity of a single group, considering the limiting SINR as a function of $\overline{p}_j$. The limiting sum rate for a single group j is denoted by $\mathcal{R}_{\Sigma,j}^\infty = \beta_j \log_2\left(1 + \text{SINR}_j^\infty\right)$ and one can verify that it is an increasing function in p_j. Moreover, one can observe that the sum rate per antenna $\mathcal{R}_\Sigma^\infty$ is concave in $\overline{\mathbf{p}}$.

In order to prove this concavity condition, the second derivative of the SINR_j^∞ w.r.t $\overline{p}_j$ is evaluated:

$$\frac{\partial^2 \text{SINR}_j^\infty}{\partial^2 \overline{p}_j} = -\frac{\mathfrak{f}_j^2(\beta, \nu)}{\left(1 + \overline{p}_j \mathfrak{f}_j^2(\beta, \nu)\right)^2} < 0. \tag{4.47}$$

The second derivative is negative which implies that SINR_j^∞ is concave in $\overline{p}_j$. Since the logarithm operation does not change the concavity of a determined function, $\mathcal{R}_{\Sigma,j}^\infty$ is also concave in $\overline{p}_j$. Moreover, $\mathcal{R}_\Sigma^\infty$ is a linear combination of SINR_j^∞ and this operator holds the convexity condition. From this analysis, we observe that for a fixed ν at the RCI precoder, problem (4.46) is a convex program due to the concavity of $\mathcal{R}_\Sigma^\infty$ w.r.t $\overline{\mathbf{p}}$, and because there are only linear constrains, resulting in a convex set.

Now, we are able to evaluate the Lagrangian for non-linear programming of (4.46), defined as follows

$$\mathcal{L} = \sum_{j=1}^{L} \beta_j \log_2\left(1 + \overline{p}_j \mathfrak{f}_j(\beta, \nu)\right) - \lambda \sum_{j=1}^{L} \beta_j \left(\overline{p}_j - 1\right) + \mu_j \overline{p}_j + \kappa \nu, \tag{4.48}$$

where λ and μ_j are the Lagrangian multipliers respectively related to the average and non-negative power constraints, and κ is the Lagrange multiplier for the regularization parameter constraint $\nu \geq 0$. Then, the associated Karush-Kuhn-

Tucker (KKT) necessary conditions are given by

$$\frac{\partial \mathcal{L}}{\partial \nu} = \sum_{j=1}^{L} \frac{\beta_j \overline{p}_j}{1 + \overline{p}_j \mathfrak{f}_j(\beta, \nu)} \frac{\partial \mathfrak{f}_j(\beta, \nu)}{\partial \nu} + \kappa = 0 \tag{4.49}$$

$$\frac{\partial \mathcal{L}}{\partial \overline{p}_j} = \beta_j \left(\frac{\mathfrak{f}_j(\beta, \nu)}{1 + \overline{p}_j \mathfrak{f}_j(\beta, \nu)} - \lambda \right) + \mu_j = 0, \tag{4.50}$$

$$\text{and} \qquad \begin{aligned} \lambda \textstyle\sum_{j=1}^{L} \beta_j \left(\overline{p}_j - 1 \right) &= 0 \\ \mu_j \overline{p}_j &= 0 \\ \kappa \nu &= 0 \qquad \forall j = 1, \ldots L \end{aligned} \tag{4.51}$$

Remembering that for a given ν, problem (4.46) reduces to a convex program, ensuring the necessary KKT conditions leads to an optimal power allocation strategy which maximizes the limiting sum rate. The solution of the problem is classical and is given by the Water-Filing algorithm as stated bellow.

Accordingly to Muharar (2012), for a fixed ν, the optimal power allocation for MIMO systems with user grouping, in which the optimization structure is defined as (4.46), follows the water-filling (WF) scheme and is given by

$$\overline{p}_j = \left[\frac{1}{\lambda} - \frac{1}{\mathfrak{f}_j(\beta, \nu)} \right]^{+} \tag{4.52}$$

where $(x)^{+} = \max[0, x]$. The Lagrange multiplier, λ, is the solution of

$$\sum_{j=1}^{L} \beta_j \overline{p}_j = P,$$

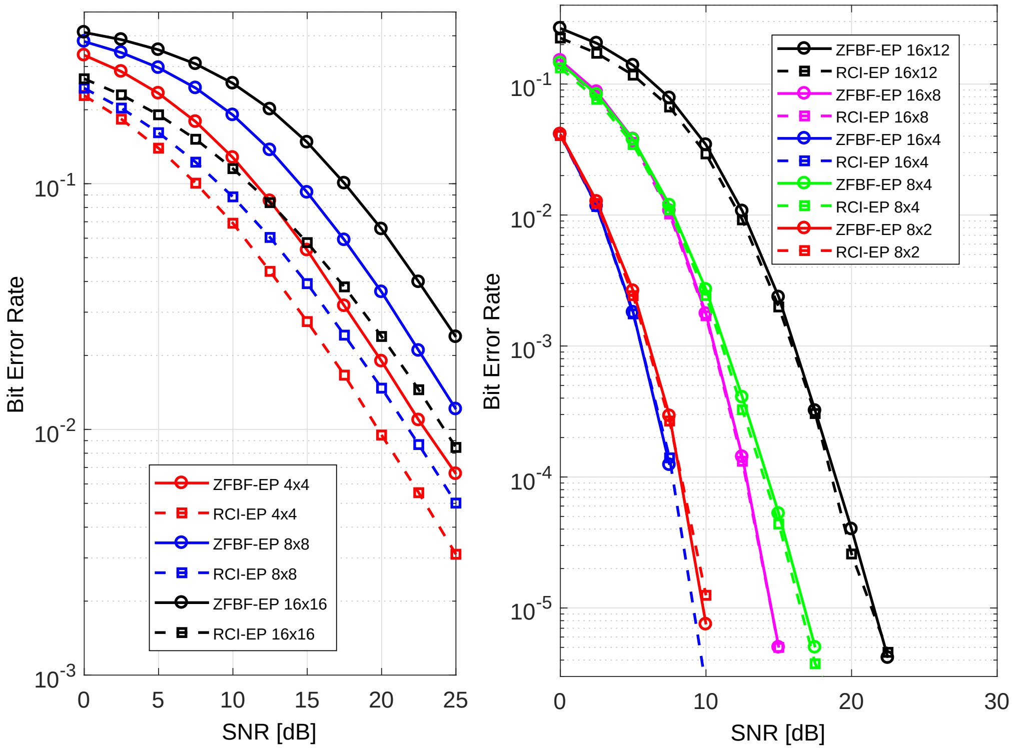
which guarantees that the average power constraint is satisfied.

Moreover, in the WF scheme for a large-scale MIMO systems the term $\frac{1}{\lambda}$ can be viewed as the water level, which determines how power level is poured among users and is based on the value of $\mathfrak{f}_j(\beta, \nu)$. With this conditions and recalling that the SINR of an user belongs to the jth group is determined by $\overline{p}_j \mathfrak{f}_j(\beta, \nu)$. It can be verified that $\mathfrak{f}_j(\beta, \nu)$ is increasing with a_j. Therefore, more power will be allocated for the user with better channel conditions which over this scenario can be determined by its path-losses.

4.5.2 Numerical Simulation Results

In this section, illustrative numerical BER, limiting sum-rate and per-user capacity of the previously discussed MIMO precoders are presented and compared. The system analysis is performed as a data transmission from a BS, with M antennas to K users (MT's) equipped with a single antenna which is the downlink scenario of a MIMO single-cell. In order of simplicity we have considered for all cases a 4-QAM modulation and perfect knowledge of the channel at the transmitter side, which means, the content of $\mathbf{H}$ is available at the transmitter side.

The first analysis consists in a comparison between the ZFBF and RCI precoding schemes. For such analysis we considered the BER and an *Ergodic* sum rate capacity evaluation, considering a Full loaded cell ($\beta = 1$) and various underloaded cell conditions ($\beta < 1$). Figure 4.8 depicts the BER for these scenarios.



a) Full-Loaded Cell (β = 1)

b) Underloaded Cell (β <1)

Figure 4.8: **a**) BER for the ZFBF and RCI precoders in a fully loaded condition ($\beta = 1$) and **b**) BER for the ZFBF and RCI precoders in an underloading condition ($\beta < 1$)

Regarding the Full-loaded Cell, it is easy to conclude that the RCI provides better performance in terms of BER. The RCI precoder introduces significantly gains in high SNR regime when compared to the ZFBF precoder, providing a performance gain in the order of approximately 5dB in the 16 × 16 condition.

Notice that, as the number of user in both BS and MT side increase, it will also be seen a slightly increase at the performance gap between the ZFBF and the RCI. Considering Figure 4.8 b), we have a representation of an under-loaded cell conditioning. In this scenario we verify that as lower the cell loading the higher is the performance gain, this behavior is given by the spatial diversity gain that is introduced due to the lower number of antennas at the MT side. Near a $\beta = 0.75$ cell-loading condition performance gains for the RCI are verified, specially under low SNR region.

Besides the performance analysis, the most relevant merit figure for the precoding scenario is the *Ergodic* Capacity of the system. Figure 4.9 a) represents the sum rate capacity of the ZFBF and the RCI precoders under a full-loaded cell conditioning.

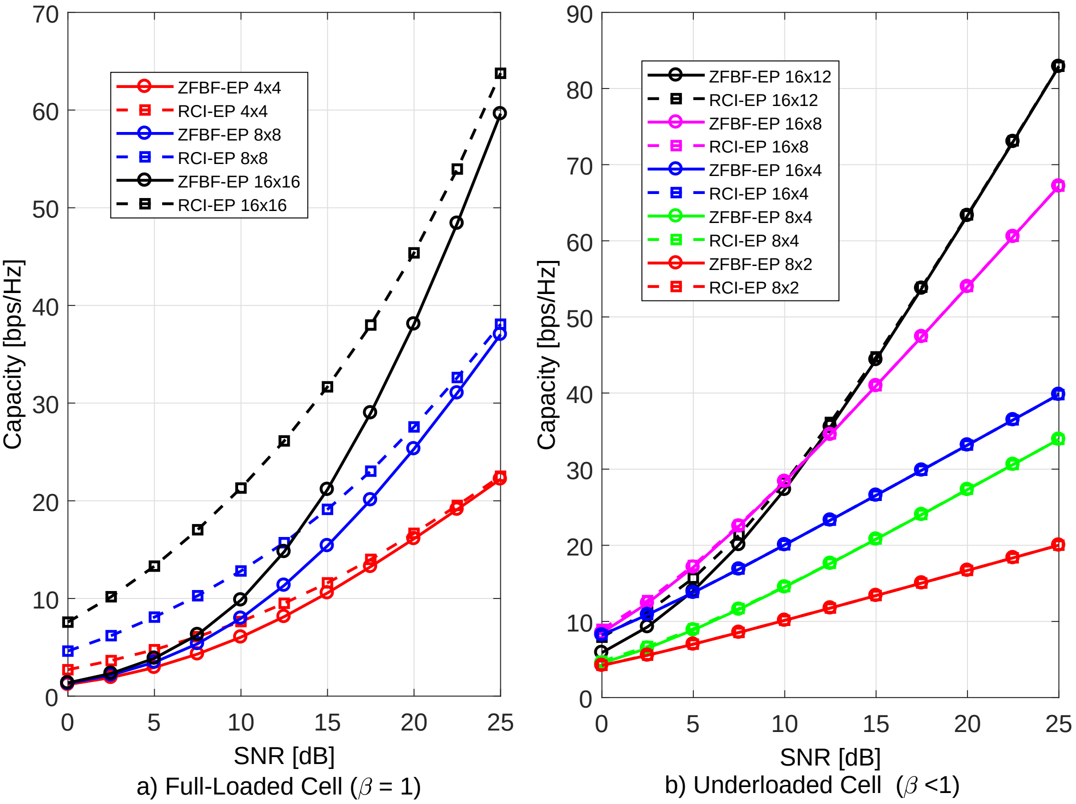


Figure 4.9: a) *Erdogic* Capacity for the ZFBF and RCI precoders in a full loaded condition ($\beta = 1$) and **b**) *Erdogic* Capacity for the ZFBF and RCI precoders in an under-loaded condition ($\beta < 1$)

In this context, it was verified that the RCI provides better capacity, specially under a low or medium SNR regime, when compared to the ZFBF scheme. As the number of antennas and users in the cell increase, the gap between both precoders capacity also increases. This behavior occurs due to the addition of the regularization parameter on the RCI precoder, which controls the amount of interference that is acceptable for each user in the cell. This parameter also considers the noise variance in the channel inversion which turns out to reduce the AWGN noise impact over the system capacity and BER performance. This can be verified in equation (4.36) when the optimal regularization parameter is derived as a ratio between the cell-loading and the SNR, which is directly related to the noise variance.

Now, regarding Figure 4.9 b) it gave us the sum rate capacity of the ZFBF and the RCI precoders under an under-loaded cell loading condition. Under this assumption, it is only observed a gain in capacity for the RCI precoder, in the 16×12 configuration, at a low SNR regime. This behavior is again due to the compensation in the noise variance introduced by the RCI regularization parameter, that compensates ill conditioned channel plus noise matrices that are commonly known to deteriorates capacity in low SNR. The lack of difference between the ZFBF and RCI capacities in the sub-loaded cell condition is due to the fact that there is a great spatial diversity, which implies in a lower interference between users in this condition and also makes the regularization parameter less effective in channel inversion, because the stronger tern in this condition is the thermal noise, not the interference.

In Figure 4.10, an interesting analysis of RCI performance and capacity were performed by comparing it with the MMSE strategy for channel inversion, where the regularization parameter in large scale system ν and the noise variance σ_n^2. In this context, the interference effect is not considered in the precoder project, leading to greater performance and capacity in low to medium SNR regime, where the noise effect is more relevant. The regularization parameter of the RCI precoder was designed to maximize the capacity in an asymptotic SINR condition ensuring the RCI to have better performance and capacity in high SINR regime, where the interference is the major concern. One can observe that, for optimal cell-loading condition $\beta^* \approx 0.75$, the MMSE and the RCI will achieve similar performance and capacity. This condition emerges due to the minimization of the interference effect that is introduced by the optimal cell-loading condition, which cause the noise variance turns out to be the main source of error in the system under this assumption.

As seen before, the RCI precoder provides better performance and capacity when compared to the ZFBF precoder. Also, as it considers the interference in the regularization parameter as one of the major concerns for the channel inversion, it ensures in high SNR regime, or asymptotic SINR regime, better performance and capacity when compared to the MMSE channel inversion strategy.

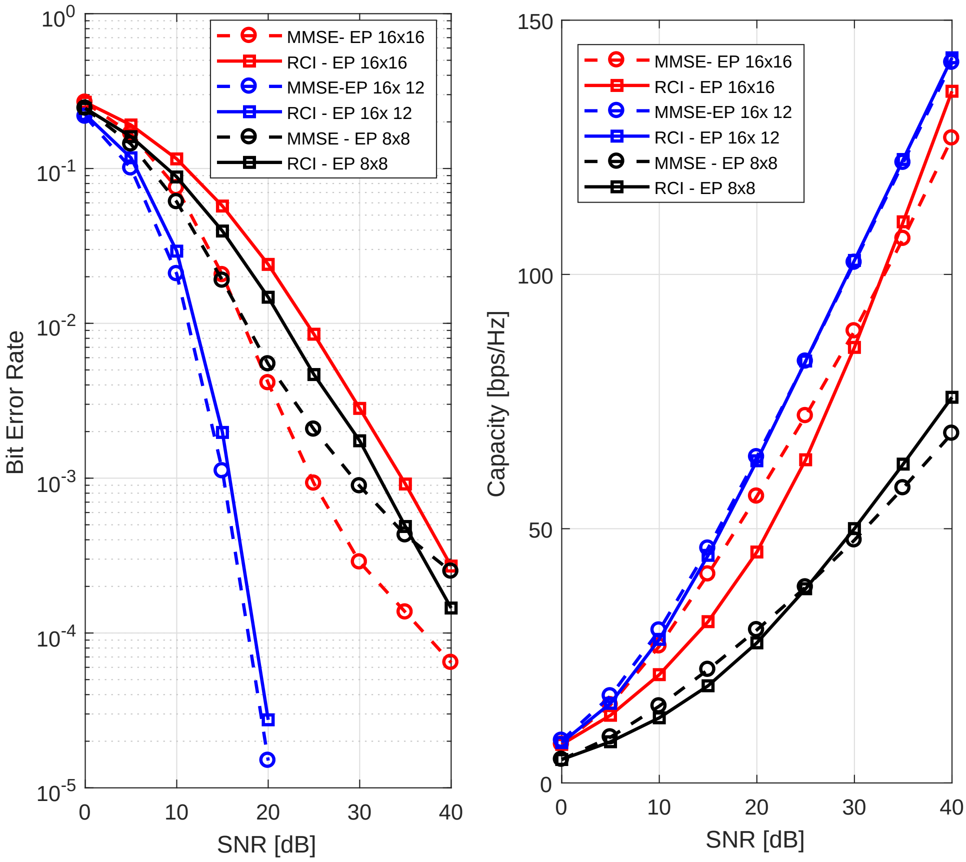


Figure 4.10: BER and *Erdogic* Capacity comparison between the MMSE and the RCI precoders in a full loaded condition ($\beta = 1$) and optimal under-loaded condition ($\beta \approx 0.75$)

4.5.2.1 Incorporating Power Allocation Scheme

Now, focusing on maximizing the achievable *Ergodic* Capacity we incorporate an optimal power allocation scheme at the system, which is given by the water-filling strategy that was defined in Equation (4.46). In our first scenario it was considered a LS-MIMO in a macro-cell of $R = 10\ km$ radius, with one BS equipped with $M = 64$ antennas and different cell-loading conditions. As an example, Figure 4.11 depicts an 64×64 system in a $R = 10\ km$ macro-cell, considered uniform distributed random path-loss for each user. Following Table 4.1, we choose an urban cellular radio environment as our main condition, which leads to a path-loss exponent $\mathfrak{b} = 3.5$.

In this context, Figure 4.12 shows the performance and capacity comparison between the RCI and RCI-WF under several cell-loading conditions. Analyzing the BER performance in Figure 4.12, greater spacial diversity, such as 64×16, will directly imply in better performance. It is also found that the difference in performance between the RCI and the RCI-WF will have a narrowing of the gap

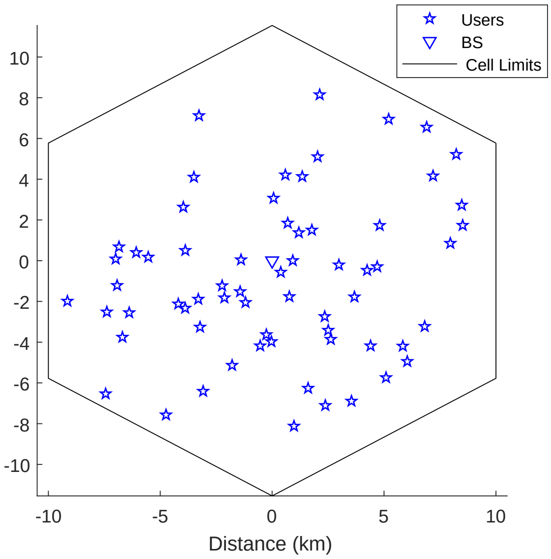


Figure 4.11: 64×64 system in a $R = 10\ km$ macro-cell with random path-loss for each user, considering an urban cellular radio environment ($\mathfrak{b} = 3.5$).

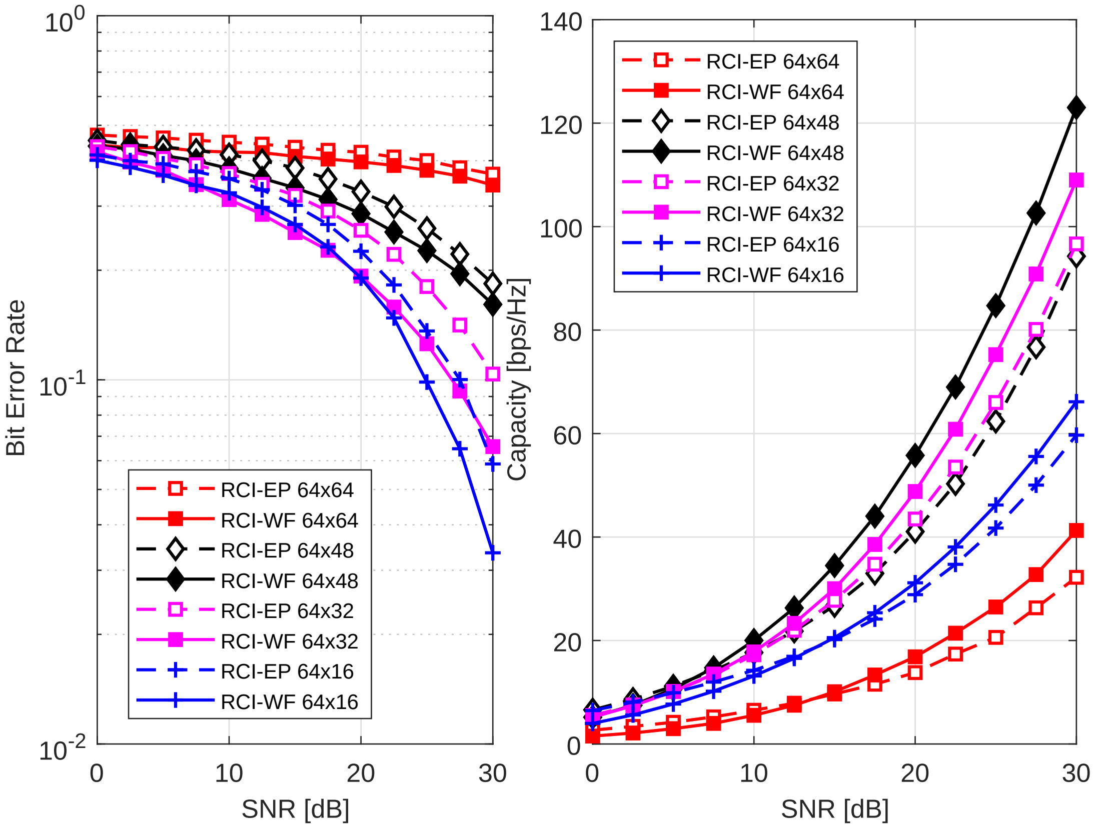


Figure 4.12: BER and *Erdogic* Capacity comparison between the RCI and RCI-WF strategy precoders in a full loaded condition ($\beta = 1$) and under-loaded conditions ($\beta < 1$), considering $M = 64$ antennas at the BS.

as the loading of cells approaches the full load condition. An interesting analysis can be done by studying the greater performance offered by the RCI-WF; this behavior can be explained by two different factors. The major reason is due to the path-loss and channel condition of each user in the cell, both information are used by the WF algorithm to determine which users will communicate with the BS. Hence, in poor channel conditions and/or higher path-losses, the performance gap between both precoders may increase, due to the active number of users. Knowing that the ones with better channel and aggregated path-loss will, consequently, have greater associated power which enhance the achieving capacity.

Now, considering the *ergodic* capacity, it is simple verify that the RCI-WF with an optimal-loading condition is the one that provides the greater achievable capacity, due to the interference minimization offered by optimal cell-loading and also the capability of eliminating some users with poor channel conditions that is provided by the WF. An interesting observation is that, as we are considering an uniform random path-loss distribution, there is the possibility that the users are placed near the BS and the achievable capacity will be even better then expected, which is the case of the RCI-EP 64×32 when compared with the RCI-EP 64×48 in an optimal cell-loading.

In order to demonstrate the effectiveness of the WF-based precoders to ensure better capacity we analyze a second scenario, consisting again in a macro-cell of $R = 10$ km radius with one BS, but now, considering a system MIMO with a BS equipped with $M = 16$ antennas and different cell-loading conditions. Figure 4.13 depict a 16×16 cell configuration in the $R = 10$ km macro-cell with uniform random path-loss for each user. The path-loss exponent were kept as $\mathfrak{b} = 3.5$.

Figure 4.14 provides a performance and capacity comparison between the RCI and RCI-WF, in a conventional MIMO condition, over different cell-loading. In this context, it was verified a small BER performance increase at an under-loaded cell condition for the RCI-WF and negligible performance gains under full-loaded scenarios.

Considering the *ergodic* capacity, it is simple verify that in a LS-MIMO scenario the achievable capacity is many times greater the conventional MIMO system, this behavior is expected as the capacity of a MIMO system will increase with the factor $\min(M, K)$. There is such greater discrepancy as a full-loaded LS-MIMO RCI-WF achievable capacity achieves the same 40 bps/Hz as the optimal-loading RCI-WF in a conventional MIMO scenario. Also, the tendency of the WF aided precoders to ensure greater capacity can be verified as the achievable capacity of

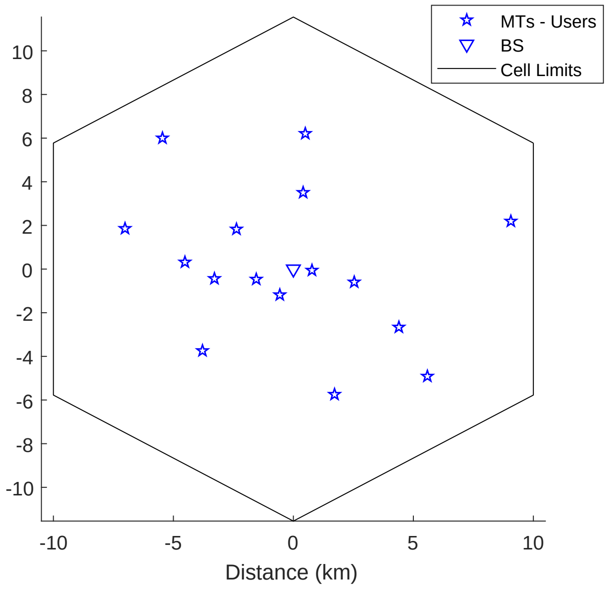


Figure 4.13: Example of random user's placement (uniform random path-loss for each user) in a 16×16 system with a macro-cell of $R = 10\ km$ radius, considering an urban cellular radio environment ($\mathfrak{b} = 3.5$).

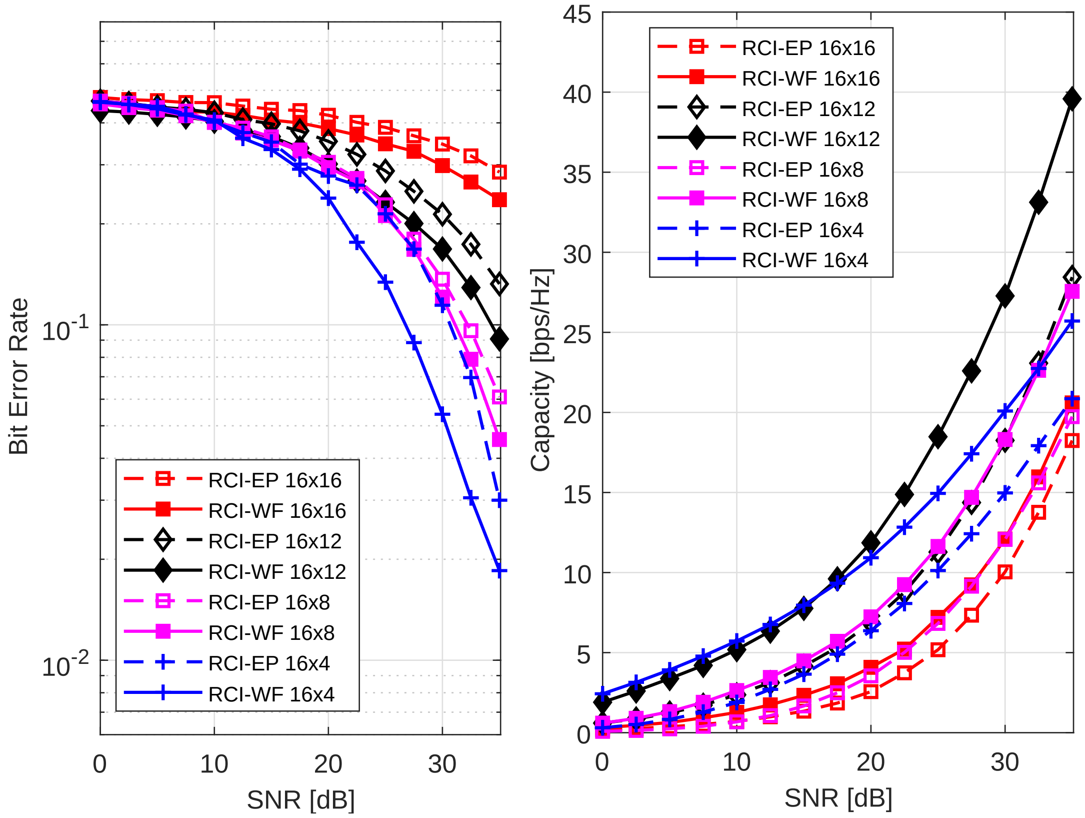


Figure 4.14: BER and *Erdogic* Capacity comparison between the RCI and RCI-WF strategy precoders in a full loaded condition ($\beta = 1$) and under-loaded conditions ($\beta < 1$), considering $M = 16$ antennas at the BS.

a RCI-WF 16 × 8 is the same capacity of an optimal loading RCI-EP 16 × 12, so the conclusion here is that the WF aided precoding strategies can provide an increase over the system *ergodic* capacity even in conventional MIMO scenarios.

4.5.2.2 User Grouping in LS-MIMO Systems

The last analysis of this work is related to the user grouping in LS-MIMO systems and how the user distribution over the cell will impact the capacity. Now, users in determined areas of the cell will be selected to be part of certain groups or, in other words, clusters. Figure 4.15, exemplify the case for a 64 × 64 system in a $R = 4$ km macro-cell with user grouping; an urban cellular radio environment has been considered, where all user in determined group will have the same associated path-loss.

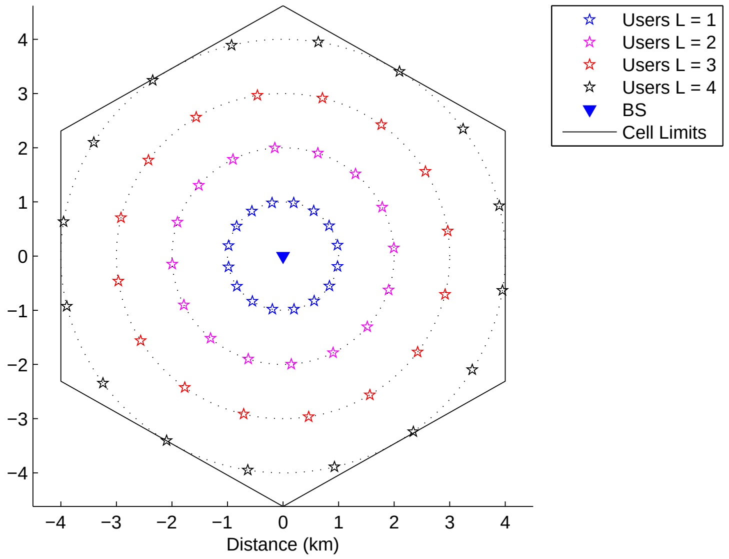


Figure 4.15: 64 × 64 system in a $R = 4$ km macro-cell with user grouping (four clusters), where all user in a group have the same path-loss.

The main goal of this analysis is to show, the best case scenario of user grouping in the cell configuration which maximizes the system capacity, under an uniform distribution of path-losses a_k. From the previous RCI-WF results we know that users with best channel conditions and lower associated path-loss will be favorable under an uniform distribution of users in the cell. Now, considering the RCI-WF precoder in a single-cell sectioned by $L = 4$ groups, where each group have the same number of users, our goal is to compare this distribution achievable capacity with a uniform user's placement distribution. Figure 4.16 depicts this first scenario, with $L = 4$ clusters and $\beta \in [0.75;\ 1.00]$.

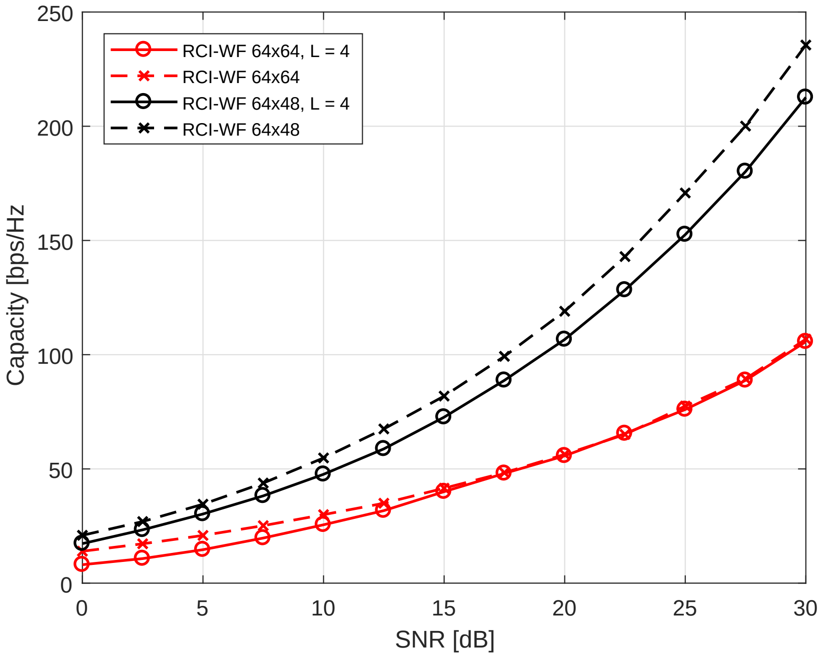


Figure 4.16: Achievable *Erdogic* Capacity of the RCI-WF precoder considering $L = 4$ group of users compared with an uniform distributed Path Loss Scenario.

It is known that in a uniform random distribution, we have a lower density of users as we move from BS. In this context, the first clustered scheme considers the same number of users in each cluster, and this condition was chosen to keep the distribution of the users density as close as possible to the uniform one. The greater difference is that in the uniform distribution, users can be placed closer to the BS when compared to the distance of the first cluster ($L = 1$km), leading to grater associated power to those users, and consequently, lesser users will be eliminated by the WF. This effect can be verified in Figure 4.16, where there is a gap between the RCI-WF 64×48 with random distribution and the clustered version. Under the same cell conditioning this gap also increases as the SNR is incremented. Now considering a full loaded scheme, 64×64, the gap between both strategies decreases as the SNR increases.

In the optimal-loading condition there is an interference balance, in the sense of a best trade-off, and as the SNR increases more groups will be activated by the WF. As these users have the same associated path-loss, there will be a fair distribution of power among users in the same group, but still serving, in the mean, a lower number of users when compared to the uniform distributed one. Also, this power distribution leads to a detriment of the over all achievable capacity of the clustered scheme in compare with the uniform distributed one. Now, in the full-loaded case, the behavior is the opposite because there is a higher amount of interference due to the greater number of users in the cell. So in low SNR regime only the first groups will be active leading to a lower capacity, and at high SNR regime the clustered version will tend to behave as an uniform distribution. This facts suggest that is better the BS to communicate only with the near users, which means the ones in the first and/or second groups, but with a cost of having a decrease in the total achievable capacity at high SNR regime.

With this idea in mind, the second scenario is based on the assumption that as close are the user from de BS, the better the achievable capacity. With the goal of nearing the gap between the clustered and the uniform distributions in a cell equipped with the RCI-WF, we redistribute the user's placement and the respective path-loss associated to each group by manipulating the path-loss exponent, $\mathfrak{b}$, decreasing from the first and second groups and increasing the fourth group associated path-loss, resulting:

$$\mathfrak{b}_1 = 2; \qquad \mathfrak{b}_2 = 2.5; \qquad \mathfrak{b}_3 = 3.5; \qquad \mathfrak{b}_4 = 5.5$$

This means that users placed in the first and second group will be seen much closer then the previous scenario, near a LOS condition, and users in the outer group will be considered in an obstructed path-loss condition and will probably be disconnected by the WF procedure.

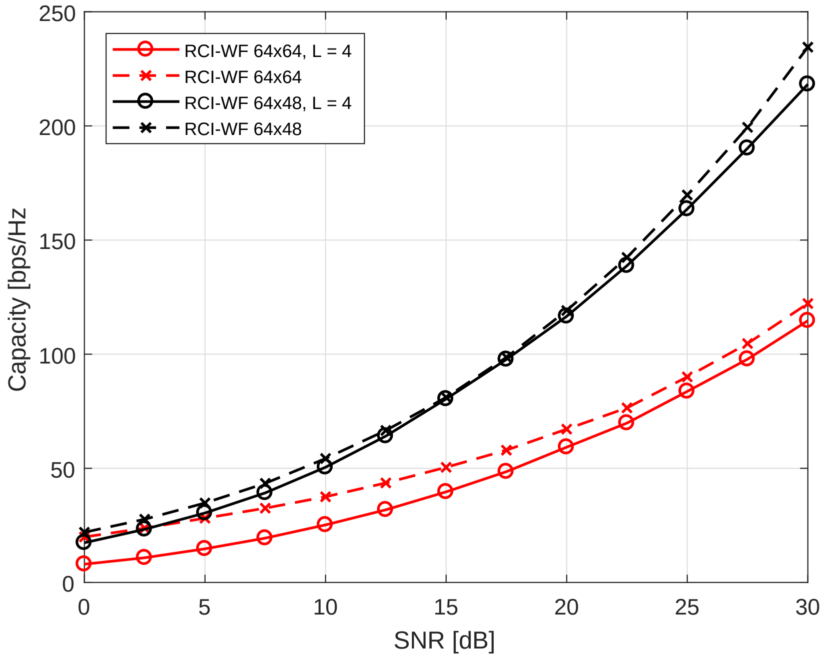


Figure 4.17: Achievable *Erdogic* Capacity of the RCI-WF precoder considering $L = 4$ groups with redistributed path-losses compared with an uniform distributed Path Loss Scenario.

Figure 4.17 illustrates how the system capacity behaves as users placed in groups $L = 1$ and 2 get closer to the BS, consequently their associated path-loss decrease, and the group $L = 4$ is seen farther from the BS, as a measure to provide a balance in the user distribution. Under this scenario, considering an optimal cell-loading, (64×48), in a medium SNR region, the gap between the uniform distributed version and the clustered one can be narrowed to a condition where they achieve near the same capacity.

In this clustered cell condition, the active number of users, under medium SNR, will be limited to the ones in groups $L = 1$ and 2, where in the mean, the number of active users in both clusters become very close to the one in the uniform distributed scenario. So, with these new path loss distribution the clustered version will behave as an uniform one in medium SNR regime, but with a fairness in the power distributed to each active users in the first two groups. Due to this condition, with an SNR increment to a high regime, an increase in the gap will be again verified. This behavior in its majority is due to the unequal distributed power in the uniform random scenario which prioritizes the closest users leading to and increase in the overall capacity, but also due to the deactivation of the fourth group in the clustered scheme.

Finally, in the full-loaded case (64×64) where it is considered a maximum interference condition, the redistributed path-loss will only decrease the capacity in all SNR range. As users are considered to be closer, there will be a greater amount of users activation by the WF, and a fair power distribution among them in their following groups, but now, as the channel conditions are more devastated, due to the interference, this greater amount of active users will lead to an over all capacity decrease over the entire SNR range.

This analysis show us that, in the sense of sum rate capacity maximization, it is interesting to the BS to maintain only a communication with the near users, activating the farthest ones as the power available at the BS (SNR) increases. The clustered scheme were proposed with the finality of diminishing the computational loading aggregated to the path-loss estimation for each user, which is fundamental in a uniform random distribution scheme, by considering the same path-loss for each cluster. Under this assumption, it was verified that, in an optimal cell loading, (64×48), considering the first two groups ($L = 1$ and 2) being closer to the BS, it was possible to achieve a near optimal WF capacity that is provided by an uniform distribution, but now, with a fairness in the power allocated to the active user.

4.6 Conclusions

This chapter carried out an investigation of the problem related to the optimal power allocation of a finite group of clustered users which maximizes the capacity in MIMO broadcast channel where all users are equipped with single antennas. The large system analysis of the broadcast channel with the RCI precoder were performed in order to determine the optimal regularization parameter and cell-loading which maximizes the limiting SINR and also their effect on it. In this condition, we show that the RCI precoder ensures a superior BER performance and capacity when compared to the ZFBF scheme. We also compare the RCI with the MMSE method which considers only noise variance in the channel inversion, showing the superiority of the RCI in a full-loaded cell under a large SNR regime.

Even though the analysis was performed over the large system limit, the simulations proved that it is also valid for finite size system. Related to the power allocation problem, we show that the optimal solution is achieved trough a water-filling resource allocation scheme at the BS, which determines the power associated to each active user in the cell.

We also provide the KKT necessary conditions for the optimal cell-loading allocation scheme when the BS is allowed to communicate only with a subset of users. Under this assumption we show that in a clustered conditions, it is convenient that the BS communicate only with users in the first groups which will lead to a capacity loss under an optimal cell-loading condition. Now, with a restructure in the path-loss distribution over the clusters, it was possible to diminish the gap between the RCI-WF with clustered path-loss version and the uniform distribution one, by the cost of considering users near the BS and decreasing the capacity in a full-loaded scenario.

5 Conclusions & Future Work

Through the studies carried out during the graduate program and presented in this text, it was possible to perform a deeper investigation around themes that have recently grown in interest to modern communication systems area, particularly MIMO and Massive MIMO systems.

Briefly, this master's dissertation work made extensive analysis in MIMO systems, specifically in the detection processes (Chapter 2). Also in this chapter we have studied the application of two different antenna array structures on both the transmit and receive sides, and their effect on the detection process. Specifically, we studied the effect of the array factor on the correlated channel, and also the performance-complexity tradeoff for each detection technique has been verified in order to determine the best structure for MIMO environments. Subsequently, in Chapter 3 the detection problem was studied under the perspective of non-linear optimization, aiming to achieve a near-optimum BER performance. It was studied the semidefinite relaxed version of the ML detection problem which was able to deliver improved BER performances for high sized number of antenna systems, with a polynomial time computational complexity load.

Despite the focus of this investigation be on the MIMO detection techniques implemented at the receiver side, in Chapter 4 we draw our attention to the transmitter, studying linear precoding techniques in a system equipped with multi antennas in the BS and single-antenna users. Hence, this chapter aims to provide an investigation related to an optimal power allocation scheme which maximize the *ergodic* sum-rate capacity under an average power constraint. We prove that the problem is convex and that the power allocation follows the well-known Water-Filling strategy. It was also studied an optimal power allocation scheme related to a finite group of clustered users and to determine the impact of this scheme in the *ergodic* sum-rate capacity. Using the WF strategy our goal is to ensure the best path-loss distribution over the cell which turns the capacity to get close enough to the one achieved by the RCI-WF precoder in a cell with uniform random user distribution.

5.1 Future Work

As a future work, one of the major possibilities is to study an hybrid version of the RCI precoder with the MMSE strategy. The idea behind this scheme is to combine the advantages of both precoding schemes. As the MMSE provides better BER performance and capacity in a low SNR regime and the RCI works better in high SNR regime, the combination of both schemes can provide a better version of the channel inversion strategy. Also, we aim to study the behavior of the analyzed precoders under a multi-cellular broadcast channel, where the effect of each BS needs to be considered in order to determine which BS the user will communicate with. This scenario will add a multiuser intercellular interference that needs to be considered, so a pilot allocation scheme will be necessary in order to ensure the downlink transmition an to mitigate the multiuser interference at the BS.

5.2 Work Disseminations

As a result, the analyses developed along this work have resulted so far in the following disseminations, which have achieved trough the realization period of the MSc. Program:

5.2.1 Conference Papers

[C1] **Title:** Semidefinite Relaxation for Large Scale MIMO Detection.
Authors: João Lucas Negrão, Alex Myamoto Mussi, Taufik Abrão.
Status: Work presented and published in the proceedings of The XXXIV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (2016). Theme developed in the Chapter 3.
Abstract: *The semi-definite relaxation (SDR) is a high performance efficient approach to MIMO detection especially for low modulation orders. We focus on developing a computationally efficient approximation of the maximum likelihood detector (ML) algorithm based on semi-definite programming (SDP) for M-QAM constellations. The detector is based on a convex relaxation of the ML problem. A comparative analysis including the performance-complexity trade-off of the SDR and the lattice reduction (LR) aided linear MIMO detectors considering high number of antennas is carried out aiming to demonstrate the effectiveness of the SDR-based conventional and large scale MIMO detector. SDR-MIMO detector can provide a close, and under high order antennas cases, a better performance than the LR-aided linear MIMO detectors.*

5.2.2 Journal Papers

[J1] **Title:** Efficient Detection in Uniform Linear and Planar Arrays MIMO Systems under Spatial Correlated Channels.

Authors: João Lucas Negrão, Taufik Abrão.

Status: Submitted to *Wiley International Journal of Communication Systems, on May 2017*

Theme developed in the Chapter 2.

Abstract: *In this paper, the efficiency of various MIMO detectors was analyzed from the perspective of highly correlated channels, where MIMO systems have a lack of performance, besides in some cases an increasing complexity. Considering this hard, but a useful scenario, various MIMO detection schemes were accurately evaluated concerning complexity and bit error rate (BER) performance. Specifically, successive interference cancellation (SIC), lattice reduction (LR) and the combination of them were associated with conventional linear MIMO detection techniques. To demonstrate effectiveness, a wide range of the number of antennas and modulation formats have been considered aiming to verify the potential of such MIMO detection techniques according to their performance-complexity trade-off. We have also studied the correlation effect when both transmit and receiver sides are equipped with uniform linear array (ULA) and uniform planar array (UPA) antenna configurations. The performance of different detectors is carefully compared when both antenna array configurations are deployed considering a different number of antennas and modulation order, especially under near-massive MIMO condition. We have also discussed the relationship between the array factor (AF) and the BER performance of both antenna array structures.*

[J2] **Title:** Sum-Rate Maximization in Downlink Massive MIMO

Authors: João Lucas Negrão, Taufik Abrão.

Status: Pre-Submission to Transactions on Emerging Telecommunications Technologies - Wiley on March 2018.

Theme developed in the Chapter 4.

Abstract: *In this paper, we analyse the power allocation problem aiming to maximize the sum-rate capacity of a single cell massive MIMO broadcast channel equipped with zero-forcing beamforming (ZFBF) and regularized channel inversion (RCI) precoding at the base station (BS). We analyze the problem over the perspective of uniform linear array (ULA) antenna structures at the BS, which is equipped with many antennas while mobile terminals (MT) are equipped with uncorrelated single-antennas. The power allocation problem is investigated in the large-scale system limit, i.e, when the number of users, K, and antennas at the BS, M, tend to infinity with a constant ratio $\beta = \frac{K}{M}$. We first derive the signal to interference plus noise (SINR) ratio for both chosen precoders. Then* *we investigate optimal power allocation schemes that maximize the sum-rate per antenna under an average power constraint and we show that the problem is convex while the power distribution follows the well-known water-filling (WF) strategy. We also studied the power allocation problem considering a finite group of clustered MT's and determine the impact of this kind of distribution on the ergodic sum-rate capacity.*

Bibliography

BAI, L.; CHOI, J.; YU, Q. **Low Complexity MIMO Receivers**. : Springer Publishing Company, Incorporated, 2014. ISBN 3319049836, 9783319049830.

BALANIS, C. A. **Antenna Theory: Analysis and Design**. : Wiley-Interscience, 2005. ISBN 0471714623.

BARBERO, L. G.; THOMPSON, J. S. Fixing the complexity of the sphere decoder for mimo detection. **IEEE Transactions on Wireless Communications**, v. 7, n. 6, p. 2131–2142, June 2008. ISSN 1536-1276.

BAREISS, E. H. Numerical solution of linear equations with toeplitz and vector toeplitz matrices. **Numer. Math.**, Springer-Verlag New York, Inc., Secaucus, NJ, USA, v. 13, n. 5, p. 404–424, out. 1969.

BENGTSSON, M.; OTTERSTEN, B. Uplink and downlink beamforming for fading channels. In: **1999 2nd IEEE Workshop on Signal Processing Advances in Wireless Communications (Cat. No.99EX304)**, 1999. p. 350–353.

BJORNSON, E.; BENGTSSON, M.; OTTERSTEN, B. Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]. **IEEE Signal Processing Magazine**, v. 31, n. 4, p. 142–148, July 2014. ISSN 1053-5888.

BOCCARDI, F.; HEATH, R. W.; LOZANO, A.; MARZETTA, T. L.; POPOVSKI, P. Five disruptive technology directions for 5g. **IEEE Communications Magazine**, v. 52, n. 2, p. 74–80, February 2014. ISSN 0163-6804.

BOHNKE, R.; WUBBEN, D.; KUHN, V.; KAMMEYER, K. D. Reduced complexity mmse detection for blast architectures. In: **Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE**, 2003. v. 4, p. 2258–2262 vol.4.

BOYD, S.; VANDENBERGHE, L. **Convex Optimization**. New York, NY, USA: Cambridge University Press, 2004. ISBN 0521833787.

BUEHRER, R. M. Generalized equations for spatial correlation for low to moderate angle spread. In: **Wireless Personal Communications: Bluetooth and Other Technologies**. Boston, MA: Springer US, 2002. p. 101–108. ISBN 978-0-306-46986-2.

CHO, Y. S.; KIM, J.; YANG, W. Y.; KANG, C. G. **MIMO-OFDM Wireless Communications with MATLAB**. : Wiley Publishing, 2010. ISBN 0470825618, 9780470825617.

CIRKIC, M. **Efficient MIMO Detection Methods**. 43 p. Tese (Doutorado) — Linkping University, Communication Systems, The Institute of Technology, 2014.

COSTA, M. Writing on dirty paper (corresp.). **IEEE Transactions on Information Theory**, v. 29, n. 3, p. 439–441, May 1983. ISSN 0018-9448.

COUILLET, R.; DEBBAH, M. **Random Matrix Methods for Wireless Communications**. New York, NY, USA: Cambridge University Press, 2011. ISBN 1107011639, 9781107011632.

CVX, R. I. **CVX: Matlab Software for Disciplined Convex Programming, version 2.0**. ago. 2012. `http://cvxr.com/cvx`.

FOSCHINI, G.; GANS, M. On limits of wireless communications in a fading environment when using multiple antennas. **Wireless Personal Communications**, v. 6, n. 3, p. 311–335, 1998. ISSN 1572-834X. Disponível em: <http://dx.doi.org/10.1023/A:1008889222784>.

GAO, X.; EDFORS, O.; RUSEK, F.; TUFVESSON, F. Linear pre-coding performance in measured very-large mimo channels. In: **2011 IEEE Vehicular Technology Conference (VTC Fall)**, 2011. p. 1–5. ISSN 1090-3038.

GETU, B. N.; ANDERSEN, J. B. The mimo cube - a compact mimo antenna. In: . Piscataway, NJ, USA: IEEE Press, 2005. v. 4, n. 3, p. 1136–1141. ISSN 1536-1276. Disponível em: <http://dx.doi.org/10.1109/TWC.2005.846997>.

GOLDSMITH, A. **Wireless Communications**. New York, NY, USA: Cambridge University Press, 2005. ISBN 0521837162.

GOLUB, G. H.; LOAN, C. F. V. **Matrix Computations**. Third. Baltimore, USA: JH Univ. Press, 1996. 694 p.

HANZO, L. L.; AKHTMAN, Y.; WANG, L.; JIANG, M. **MIMO-OFDM for LTE, WiFi and WiMAX: Coherent versus non-coherent and cooperative turbo transceivers**. : John Wiley & Sons, 2010.

HOYDIS, J.; BRINK, S. ten; DEBBAH, M. Massive mimo in the ul/dl of cellular networks: How many antennas do we need? **IEEE Journal on Selected Areas in Communications**, v. 31, n. 2, p. 160–171, February 2013. ISSN 0733-8716.

JALDEN, J. **Maximum Likelihood Detection for the Linear MIMO Channel**. Tese (Doutorado) — Royal Institute of Technology, 2004.

JALDEN, J.; MARTIN, C.; OTTERSTEN, B. Semidefinite programming for detection in linear systems - optimality conditions and space-time decoding. In: **Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on**, 2003. v. 4, p. IV–9–12 vol.4. ISSN 1520-6149.

JALDEN, J.; OTTERSTEN, B. On the complexity of sphere decoding in digital communications. **IEEE Transactions on Signal Processing**, v. 53, n. 4, p. 1474–1484, April 2005. ISSN 1053-587X.

KHALIGHI, M. A.; BROSSIER, J. M.; JOURDAIN, G. V.; RAOOF, K. Water filling capacity of rayleigh mimo channels. In: **12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. PIMRC 2001. Proceedings (Cat. No.01TH8598)**, 2001. v. 1, p. A–155–A–158 vol.1.

KOBAYASHI, R. T.; ABRÃO, T. Ordered mmse—sic via sorted qr decomposition in ill conditioned large-scale mimo channels. **Telecommun. Syst.**, Kluwer Academic Publishers, Norwell, MA, USA, v. 63, n. 2, p. 335–346, out. 2016. ISSN 1018-4864. Disponível em: <http://dx.doi.org/10.1007/s11235-015-0123-5>.

KOBAYASHI, R. T.; CIRIACO, F.; ABRÃO, T. Efficient near-optimum detectors for large mimo systems under correlated channels. **Wireless Personal Communications**, v. 83, n. 2, p. 1287–1311, 2015. ISSN 1572-834X. Disponível em: <http://dx.doi.org/10.1007/s11277-015-2450-y>.

LARSSON, E. G. Mimo detection methods: How they work [lecture notes]. **IEEE Signal Processing Magazine**, IEEE, v. 26, n. 3, p. 91–95, 2009.

LENSTRA, A.; LENSTRA, H.; LOVÁSZ, L. Factoring polynomials with rational coefficients. **Mathematische Annalen**, Springer New York, v. 261, n. 4, p. 515–534, 12 1982. ISSN 0025-5831.

LEVIN, G.; LOYKA, S. On capacity-maximizing angular densities of multipath in mimo channels. In: **2010 IEEE 72nd Vehicular Technology Conference - Fall**, 2010. p. 1–5. ISSN 1090-3038.

LI, J.; SU, X.; ZENG, J.; ZHAO, Y.; YU, S.; XIAO, L.; XU, X. Codebook design for uniform rectangular arrays of massive antennas. In: **2013 IEEE 77th Vehicular Technology Conference (VTC Spring)**, 2013. p. 1–5. ISSN 1550-2252.

LI, Q.; LI, G.; LEE, W.; LEE, M. i.; MAZZARESE, D.; CLERCKX, B.; LI, Z. Mimo techniques in wimax and lte: a feature overview. **IEEE Communications Magazine**, v. 48, n. 5, p. 86–92, May 2010. ISSN 0163-6804.

LING, C.; HOWGRAVE-GRAHAM, N. Effective lll reduction for lattice decoding. In: **2007 IEEE International Symposium on Information Theory**, 2007. p. 196–200. ISSN 2157-8095.

LING, C.; MOW, W. H.; GAN, L. Dual-lattice ordering and partial lattice reduction for sic-based mimo detection. **IEEE Journal of Selected Topics in Signal Processing**, v. 3, n. 6, p. 975–985, Dec 2009. ISSN 1932-4553.

LING, C.; MOW, W. H.; HOWGRAVE-GRAHAM, N. Reduced and fixed-complexity variants of the lll algorithm for communications. **IEEE Transactions on Communications**, v. 61, n. 3, p. 1040–1050, March 2013. ISSN 0090-6778.

LUO, Z.-Q.; MA, W.-K.; SO, A.-C.; YE, Y.; ZHANG, S. Semidefinite relaxation of quadratic optimization problems. **Signal Processing Magazine, IEEE**, v. 27, n. 3, p. 20–34, May 2010. ISSN 1053-5888.

MA, W.-K.; CHING, P.-C.; DING, Z. Semidefinite relaxation based multiuser detection for m-ary psk multiuser systems. **Signal Processing, IEEE Transactions on**, v. 52, n. 10, p. 2862–2872, Oct 2004. ISSN 1053-587X.

MA, W.-K.; DAVIDSON, T.; WONG, K. M.; LUO, Z.-Q.; CHING, P.-C. Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous cdma. **Signal Processing, IEEE Transactions on**, v. 50, n. 4, p. 912–922, April 2002. ISSN 1053-587X.

MA, X.; ZHANG, W. Performance analysis for mimo systems with lattice-reduction aided linear equalization. **Communications, IEEE Transactions on**, IEEE, v. 56, n. 2, p. 309–318, 2008.

MANDEEP, J.; MISRAN, N.; ABDULLAH, H.; HOW, T. Patch array antenna serves satcom needs. **Microwaves and RF**, Penton Publishing Co., v. 49, n. 4, 4 2010. ISSN 0745-2993.

MAO, Z.; WANG, X.; WANG, X. Qam-mimo signal detection using semidefinite programming relaxation. In: **Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE**, 2007. p. 4232–4236.

MARCENKO, V. A.; PASTUR, L. A. Distribution of eigenvalues for some sets of random matrices. **Mathematics of the USSR-Sbornik**, v. 1, n. 4, p. 457–483, abr. 1967. Disponível em: <http://dx.doi.org/10.1070/sm1967v001n04abeh001994>.

MILFORD, D.; SANDELL, M. Simplified quantisation in a reduced-lattice mimo decoder. **Communications Letters, IEEE**, v. 15, n. 7, p. 725–727, July 2011. ISSN 1089-7798.

MORADI, S.; DOOSTNEJAD, R.; GULAK, G. Downlink beamforming for fdd systems with precoding and beam steering. In: **2011 IEEE Global Telecommunications Conference - GLOBECOM 2011**, 2011. p. 1–6. ISSN 1930-529X.

MUHARAR, R. **Multiuser Precoding in Wireless Communication Systems**. 233 p. Tese (Doutorado) — The University of Melbourne, Department of Electrical and Electronic Engineering, 2012.

MUHARAR, R.; EVANS, J. Optimal power allocation for multiuser transmit beamforming via regularized channel inversion. In: **2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)**, 2011. p. 1393–1397. ISSN 1058-6393.

MUSSI, A. M.; ABRAO, T. Sdr lattice-reduction-aided detector. **IEEE Latin America Transactions**, v. 11, n. 4, p. 1007–1014, June 2013. ISSN 1548-0992.

NGO, H. Q.; LARSSON, E. G.; MARZETTA, T. L. Energy and spectral efficiency of very large multiuser mimo systems. **IEEE Transactions on Communications**, v. 61, n. 4, p. 1436–1449, April 2013. ISSN 0090-6778.

NGUYEN, V. K.; EVANS, J. S. Multiuser transmit beamforming via regularized channel inversion: A large system analysis. In: **IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference**, 2008. p. 1–4. ISSN 1930-529X.

PEEL, C. B.; HOCHWALD, B. M.; SWINDLEHURST, A. L. A vector-perturbation technique for near-capacity multiantenna multiuser communication-part i: channel inversion and regularization. **IEEE Transactions on Communications**, v. 53, n. 1, p. 195–202, Jan 2005. ISSN 0090-6778.

RAPOPORT, L.; YANXING, Z.; IVANOV, V.; JIANQIANG, S. Implementation of quasi-maximum-likelihood detection based on semidefinite relaxation and linear programming. In: **Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of**, 2012. p. 1–5.

RAPPAPORT, T. S.; BLANKENSHIP, K.; XU, H. Propagation and radio system design issues in mobile radio systems for the glomo project. **Virginia Polytechnic Institute and State University**, 1997.

ROH, W.; SEOL, J. Y.; PARK, J.; LEE, B.; LEE, J.; KIM, Y.; CHO, J.; CHEUN, K.; ARYANFAR, F. Millimeter-wave beamforming as an enabling technology for 5g cellular communications: theoretical feasibility and prototype results. **IEEE Communications Magazine**, v. 52, n. 2, p. 106–113, February 2014. ISSN 0163-6804.

RUSEK, F.; PERSSON, D.; LAU, B. K.; LARSSON, E. G.; MARZETTA, T. L.; EDFORS, O.; TUFVESSON, F. Scaling up mimo: Opportunities and challenges with very large arrays. **IEEE Signal Processing Magazine**, v. 30, n. 1, p. 40–60, Jan 2013. ISSN 1053-5888.

SHERMAN, J.; MORRISON, W. J. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. **Ann. Math. Statist.**, The Institute of Mathematical Statistics, v. 21, n. 1, p. 124–127, 03 1950. Disponível em: <http://dx.doi.org/10.1214/aoms/1177729893>.

SIDIROPOULOS, N. D.; LUO, Z. Q. A semidefinite relaxation approach to mimo detection for high-order qam constellations. **IEEE Signal Processing Letters**, v. 13, n. 9, p. 525–528, Sept 2006. ISSN 1070-9908.

TAROKH, V.; NAGUIB, A.; SESHADRI, N.; CALDERBANK, A. R. Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths. **IEEE Transactions on Communications**, v. 47, n. 2, p. 199–207, Feb 1999. ISSN 0090-6778.

TRAN, L. N.; HANIF, M. F.; JUNTTI, M. A conic quadratic programming approach to physical layer multicasting for large-scale antenna arrays. **IEEE Signal Processing Letters**, v. 21, n. 1, p. 114–117, Jan 2014. ISSN 1070-9908.

TULINO, A. M.; VERDU, S. Random matrix theory and wireless communications. **Commun. Inf. Theory**, Now Publishers Inc., Hanover, MA, USA, v. 1, n. 1, p. 1–182, jun. 2004. ISSN 1567-2190. Disponível em: <http://dx.doi.org/10.1516/0100000001>.

VALENTE, R. A.; MARINELLO, J. C.; ABRÃO, T. Lr-aided mimo detectors under correlated and imperfectly estimated channels. **Wireless personal communications**, Springer, v. 77, n. 1, p. 173–196, 2014.

VANDENBERGHE, L.; BOYD, S. Semidefinite programming. **SIAM Review**, p. 38:49–96, 1996.

WAGNER, S.; COUILLET, R.; DEBBAH, M.; SLOCK, D. T. M. Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback. **IEEE Transactions on Information Theory**, v. 58, n. 7, p. 4509–4537, July 2012. ISSN 0018-9448.

WIESEL, A.; ELDAR, Y.; SHAMAI, S. Semidefinite relaxation for detection of 16-qam signaling in mimo channels. **Signal Processing Letters, IEEE**, v. 12, n. 9, p. 653–656, Sept 2005. ISSN 1070-9908.

WIESEL, A.; ELDAR, Y. C.; SHAMAI, S. Zero-forcing precoding and generalized inverses. **IEEE Transactions on Signal Processing**, v. 56, n. 9, p. 4409–4418, Sept 2008. ISSN 1053-587X.

WIGNER, E. P. On the distribution of the roots of certain symmetric matrices. **Annals of Mathematics**, JSTOR, p. 325–327, 1958.

WOLNIANSKY, P.; FOSCHINI, G.; GOLDEN, G.; VALENZUELA, R. V-blast: an architecture for realizing very high data rates over the rich-scattering wireless channel. p. 295–300, 1998.

WUBBEN, D.; BOHNKE, R.; KUHN, V.; KAMMEYER, K.-D. Mmse extension of v-blast based on sorted qr decomposition. In: **Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th**, 2003. v. 1, p. 508–512 Vol.1. ISSN 1090-3038.

WUBBEN, D.; BOHNKE, R.; KUUHN, V.; KAMMEYER, K.-D. Near-maximum-likelihood detection of mimo systems using mmse-based lattice reduction. In: **Communications, 2004 IEEE International Conference on**, 2004. v. 2, p. 798–802 Vol.2.

WUBBEN, D.; SEETHALER, D.; JALDÉN, J.; MATZ, G. Lattice reduction. **IEEE Signal Processing Magazine**, IEEE, v. 28, n. 3, p. 70–91, 2011.

YANG, H.; MARZETTA, T. L. Performance of conjugate and zero-forcing beamforming in large-scale antenna systems. **IEEE Journal on Selected Areas in Communications**, v. 31, n. 2, p. 172–179, February 2013. ISSN 0733-8716.

YING, D.; VOOK, F. W.; THOMAS, T. A.; LOVE, D. J.; GHOSH, A. Kronecker product correlation model and limited feedback codebook design in a 3d channel model. In: **2014 IEEE International Conference on Communications (ICC)**, 2014. p. 5865–5870. ISSN 1550-3607.

ZELST, A. V.; HAMMERSCHMIDT, J. A single coefficient spatial correlation model for multiple-input multiple-output (mimo) radio channels. In: **Proc. of URSI General Assembly**, 2002. p. 17–24.

ZHAO, Y.; WANG, X.; YANG, J.; ZHAO, B. Downlink closed-loop training sequence design for massive mimo systems with uniform planar arrays. In: **2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)**, 2016. p. 1–5.

ZHENG, X.; QIU, L.; ZHU, J. A simple diversity technique using spread space-time block coding in mimo systems. In: **2004 IEEE 59th Vehicular Technology Conference. VTC 2004-Spring (IEEE Cat. No.04CH37514)**, 2004. v. 1, p. 384–388 Vol.1. ISSN 1550-2252.

Appendix A

A.1 UV Mapping

The UV mapping is a 3D modeling process aiming to provide a 2D image representation of a 3D model.

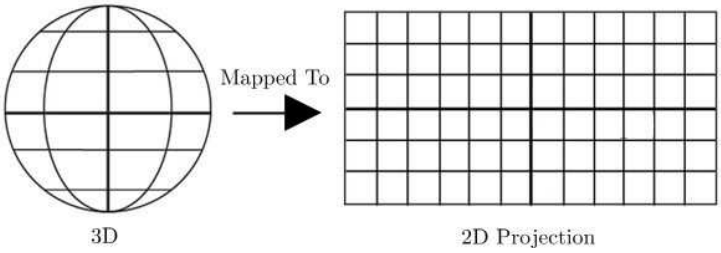


Figure A.1: Spherical Coordinates System

Basically, instead of having the image in the conventional Cartesian plane we will take advantage of spherical coordinate system, by defining the Azimuth and the Elevation angles, represented by ϕ and θ respectively.

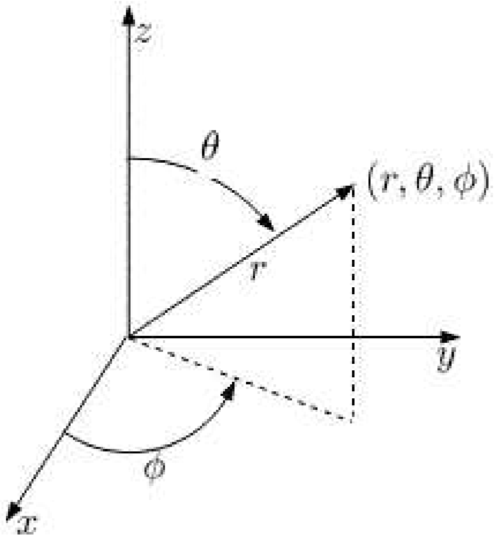


Figure A.2: Spherical Coordinates System

In the spherical coordinate system parameters are defined as:

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \\ \phi &= \arctan \frac{y}{x} \end{aligned} \tag{A.1}$$

Considering an unitary radius and having the Azimuth and Elevation angles, the u/v coordinates can be easily derived from the ϕ and θ. By definition, the azimuth angle of a vector is the angle between the x-axis and the orthogonal projection of the vector onto the xy plane and the elevation angle is the angle between the vector and its orthogonal projection onto the xy plane. The relationship between these two coordinates system is:

$$\begin{aligned} u &= \sin\theta\cos\phi \\ v &= \sin\theta\sin\phi \end{aligned} \tag{A.2}$$

the values of u and v satisfy the inequalities:

$$\begin{aligned} -1 \leq u \leq 1 \\ -1 \leq v \leq 1 \\ u^2 + v^2 \leq 1 \end{aligned} \tag{A.3}$$

In the antenna array context, we usually search for the azimuth cut result. Which means that the Azimuth angle $\phi = 0°$. In this context, we can evaluate how the Array Factor (AF) variation as a function of the Elevation angle θ.

To exemplify this procedure we will take the azimuth cut spherical response of a 3×3 UPA with 0.5λ element spacing and then provide the response un the U space.

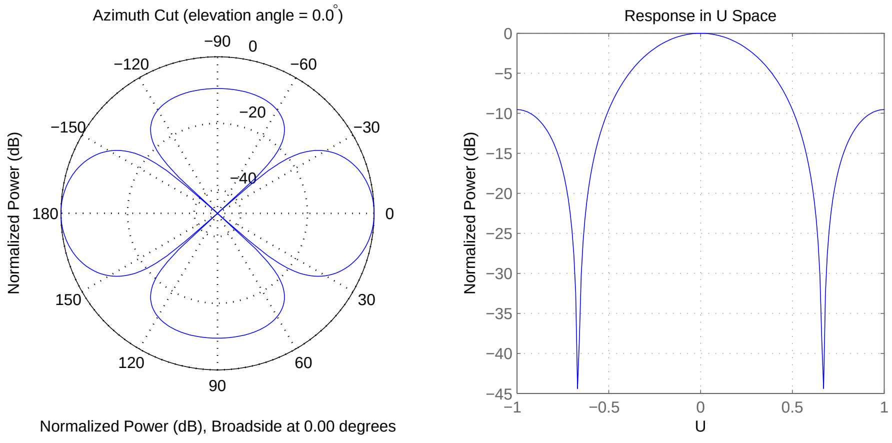


Figure A.3: Spherical Coordinates System

A.2 Proof of Theorem 4.4.4

Considering the RCI precoder SINR equation (4.19), we can write this expression as

$$\gamma_{k,rci} = \frac{c^2 S_k}{\sigma^2 + c^2 I_k} \tag{A.4}$$

where c^2S_k and c^2I_k represent the signal power and interference energy of user k respectively. The large limit SINR can be obtained by deriving the asymptotic limit of each term in $\gamma_{k,rci}$. Starting with the signal component, observe that $\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right) = \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M + \mathbf{h}_k^H\mathbf{h}_k\right)$, where $\mathbf{H}_k$ is $\mathbf{H}$ with the $k-$th row removed. Now, we can apply the matrix inversion *Lemma* (Sherman-Morrison Formula) (SHERMAN; MORRISON, 1950), we have:

$$\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} - \frac{\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H\mathbf{h}_k\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}}{1 + \mathbf{h}_k\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H} \tag{A.5}$$

which straightforwardly can be represented as

$$\left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_K^H = \frac{\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H}{1 + \mathbf{h}_k^H\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H}. \tag{A.6}$$

Now, the signal power S_k can be rewritten as

$$S_k = \frac{\left|\mathbf{h}_k\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H\right|^2}{\left(1 + \mathbf{h}_k\left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\mathbf{h}_k^H\right)^2} = \frac{|A_k|^2}{\left(1 + A_k\right)^2}, \tag{A.7}$$

where $A_k = \frac{1}{M}\mathbf{h}_k\left(\frac{1}{M}\mathbf{H}_k^H\mathbf{H}_k + \nu\mathbf{I}_M\right)\mathbf{h}_k^H$ and $\nu = \frac{\xi}{M}$. Thus, using *Lemma* 4.4.2, it is almost certain that $A_k - \dfrac{1}{M}\text{tr}\left[\left(\mathbf{H}^H\mathbf{H} + \nu\mathbf{I}_M\right)^{-1}\right] \xrightarrow{a.s.} 0$ and approximately we have

$$S_k = \frac{\dfrac{1}{M}\text{tr}\left[\left(\mathbf{H}^H\mathbf{H} + \nu\mathbf{I}_M\right)^{-1}\right]^2}{\left(1 + \dfrac{1}{M}\text{tr}\left[\left(\mathbf{H}^H\mathbf{H} + \nu\mathbf{I}_M\right)^{-1}\mathbf{h}\right]\right)^2}, \tag{A.8}$$

because, accordingly to (COUILLET; DEBBAH, 2011) *Lemma* 5, the removal of a single column does not affect the normalized trace in the asymptotic scenario.

Now, using *Lemma* 4.4.3, we have

$$A_k - g(\beta, \nu) \xrightarrow{a.s.} 0,$$

where $g(\beta, \nu)$ is the function defined in (4.30). Thus, its is expedite that:

$$S_k = \frac{g(\beta, \nu)^2}{(1 + g(\beta, \nu))^2}. \tag{A.9}$$

Now, moving to the interference term. I_k can be expressed as

$$\begin{aligned} I_k &= \textstyle\sum_{j\neq k} \mathbf{h}_k \left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_j^H \mathbf{h}_j \left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_k^H \\ &= \textstyle\sum_{j\neq k} \mathbf{h}_k \left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1} \mathbf{H}_k^H \mathbf{H}_k \left(\mathbf{H}^H\mathbf{H} + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_k^H \end{aligned} \tag{A.10}$$

again applying the Matrix inversion *Lemma* we have:

$$\begin{aligned} I_k &= \frac{\mathbf{h}_{\mathbf{k}} \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \mathbf{H}_k^H \mathbf{H}_k \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_k^H}{\left(1 + \mathbf{h}_k \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \mathbf{h}_k^H\right)^2} \\ &= \frac{B_k}{\left(1 + A_k\right)^2} \end{aligned} \tag{A.11}$$

Considering $B_k = \mathbf{h}_k \mathbf{B}_k \mathbf{h}_k^h$, we can show that

$$\begin{aligned} \mathbf{B}_k &= \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \mathbf{H}_k^H \mathbf{H}_k \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \\ &= \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M - \xi\mathbf{I}_M\right) \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \\ &= \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \left[\mathbf{I}_M - \xi \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1}\right] \\ &= \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} - \xi \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-2} \\ &= \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} + \xi \frac{\partial}{\partial \xi} \left(\mathbf{H}_k^H\mathbf{H}_k + \xi\mathbf{I}_M\right)^{-1} \end{aligned} \tag{A.12}$$

Hence, $B_k = A_k + \dfrac{\partial A_k}{\partial \xi}$, with $A_k - \dfrac{1}{M}\mathrm{tr}\left[\left(\mathbf{H}^H\mathbf{H} + \nu\mathbf{I}_M\right)^{-1}\right] \xrightarrow{a.s.} 0$ and applying the *Lemma* 4.4.3,

$$B_k - \left(g(\beta,\nu) + \frac{\partial g(\beta,\nu)}{\partial \nu}\right) \xrightarrow{a.s.} 0. \tag{A.13}$$

Consequently,

$$I_k - \frac{g(\beta,\nu) + \dfrac{\partial g(\beta,\nu)}{\partial \nu}}{\left(1 + g(\beta,\nu)\right)^2} \xrightarrow{a.s.} 0. \tag{A.14}$$

To complete the proof, we consider now the normalizing constant α^2. Using equation (4.17) and disregarding the power allocated term $\mathbf{P}^{1/2}$; the denominator of α^2 can be expressed as

$$\frac{1}{M}\mathrm{tr}\left[\left(\frac{1}{M}\mathbf{H}^H\mathbf{H} + \nu\mathbf{I}_M\right)^{-2} \frac{1}{M}\mathbf{H}^H\mathbf{H}\right].$$

Following the same steps that was used to derive the large system limit of A_k and B_k, we obtain

$$\alpha^2 - \frac{P}{\left(g(\beta,\nu) + \nu\dfrac{\partial g(\beta,\nu)}{\partial \nu}\right)} \xrightarrow{a.s.} 0 \tag{A.15}$$

Now, combining the large system results of S_k and α^2, given by equations (A.9) and (A.15), the signal energy converges almost surely to a deterministic value given by:

$$\begin{aligned}\alpha^2 S_k &= \frac{P}{\left(g(\beta,\nu)+\nu\dfrac{\partial g(\beta,\nu)}{\partial\nu}\right)}\frac{g(\beta,\nu)^2}{(1+g(\beta,\nu))^2}\\ &= P\frac{g(\beta,\nu)}{(1+g(\beta,\nu))^2}\left(1+\frac{\nu}{\beta}(1+g(\beta,\nu))^2\right)\end{aligned}$$

where

$$\frac{\partial g(\beta,\nu)}{\partial\nu} = -\frac{g(\beta,\nu)(1+g(\beta,\nu))^2}{\beta+\nu(1+g(\beta,\nu))^2}. \tag{A.16}$$

Similarly, combining the large system results of I_k and α^2, which are given by equations (A.14) and (A.15) respectively, the interference energy converges almost surely to a deterministic quantity expressed by:

$$\alpha^2 I_k = \frac{P}{(1+g(\beta,\nu))^2}. \tag{A.17}$$

Now, combining the previous results as (A.4) we can finally express the limiting SINR of the RCI precoder as the equation (4.33), with $\gamma = \dfrac{P}{\sigma_n^2}$ and this complete the proof.

A.3 Proof of optimal regularization parameter

First, we should rewrite the Limiting SINR as

$$\text{SINR}^\infty = \frac{\gamma}{\beta} g \Upsilon$$

where

$$\Upsilon = \frac{\beta+\nu(1+g)^2}{\gamma+(1+g)^2}.$$

The first derivative of SINR$^\infty$ over ν was taken ad is given by:

$$\frac{\partial\text{SINR}^\infty}{\partial\nu} = \frac{\gamma}{\beta}\left(\frac{\partial g}{\partial\nu}\Upsilon + g\frac{\partial\Upsilon}{\partial\nu}\right), \tag{A.18}$$

where

$$\frac{\partial\Upsilon}{\partial\nu} = \frac{\left[(1+g)^2+2\nu\frac{\partial g}{\partial\nu}(1+g)\right]\left[\gamma+(1+g^2)\right]-\left[\beta+\nu(1+g)^2\right]\left[2\frac{\partial g}{\partial\nu}(1+g)\right]}{\left[\gamma+(1+g^2)\right]^2} \tag{A.19}$$

Performing further elaborations in (A.18), we have the following steps

$$
\begin{aligned}
\frac{\partial \mathrm{SINR}^{\infty}}{\partial \nu} &= \frac{\gamma}{\beta}\left(\frac{\partial g}{\partial \nu}\Upsilon + g\frac{\partial \Upsilon}{\partial \nu}\right) \\
&= \frac{\gamma}{\beta} g \Upsilon \left[\frac{\frac{\partial g}{\partial \nu}}{g} + \frac{\frac{\partial \Upsilon}{\partial \nu}}{\Upsilon}\right] \\
&= \frac{\gamma}{\beta} g \Upsilon \left[\frac{\frac{\partial g}{\partial \nu}}{g} + \frac{\left[(1+g)^2 + 2\nu \frac{\partial g}{\partial \nu}(1+g)\right]\left[\gamma + \left(1+g^2\right)\right] - \left[\beta + \nu(1+g)^2\right]\left[2\frac{\partial g}{\partial \nu}(1+g)\right]}{\left[\gamma + (1+g)^2\right]\left[\beta + \nu(1+g)^2\right]}\right] \\
&= \frac{\gamma}{\beta} g \Upsilon \left[\frac{\frac{\partial g}{\partial \nu}}{g} + \frac{(1+g)^2}{\beta + \nu(1+g)^2} + \frac{2\nu\frac{\partial g}{\partial \nu}(1+g)}{\beta + \nu(1+g)^2} + \frac{2\frac{\partial g}{\partial \nu}(1+g)}{\gamma + \nu(1+g)^2}\right] \\
&= \frac{\gamma}{\beta} g \Upsilon \left[\frac{\frac{\partial g}{\partial \nu}}{g} - \frac{\frac{\partial g}{\partial \nu}}{g} + \frac{2\nu\frac{\partial g}{\partial \nu}(1+g)}{\beta + \nu(1+g)^2} + \frac{2\frac{\partial g}{\partial \nu}(1+g)}{\gamma + \nu(1+g)^2}\right] \\
&= \frac{2\gamma^2 g(1+g)^2}{\beta\left[\gamma + (1+g)^2\right]^2}\frac{\partial g}{\partial \nu}\left[\nu - \frac{\beta}{\gamma}\right] \\
&= \mathcal{K}\frac{\partial g}{\partial \nu}\left[\nu - \frac{\beta}{\gamma}\right]
\end{aligned}
\tag{A.20}
$$

where (A.20) is obtained from (A.18), i.e,

$$
-\frac{\frac{\partial g}{\partial \nu}}{g} = \frac{(1+g)^2}{\beta + \nu(1+g)^2}. \tag{A.21}
$$

Since $\mathcal{K} > 0$ and $\frac{\partial g}{\partial \nu} < 0$, then the stationary point is given by

$$
\nu^* = \frac{\beta}{\gamma}. \tag{A.22}
$$

Moreover, since $\mathcal{K}\frac{\partial g}{\partial \nu} < 0$, then (A.20) is positive for $\nu < \nu^*$ and it is negative for $\nu > \nu^*$. Thereby, the limiting SINR is increasing over ν until reaching $\nu = \nu^*$ and decreasing after that. With such behavior, it concludes that the limiting SINR is a quasi-concave function of ν (BOYD; VANDENBERGHE, 2004, p. 99) and ν^* is the global optimizer.