UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Carlos Daniel Altamirano Carrillo

Performance of Massive MIMO Under Interference Reduction and Imperfect Channel Estimation

Avaliação de Desempenho de Sistemas MIMO Massivos com Redução de Interferência e Estimação de Canal Imperfeita

Campinas
2021

UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Carlos Daniel Altamirano Carrillo

Performance of Massive MIMO Under Interference Reduction and Imperfect Channel Estimation

Avaliação de Desempenho de Sistemas MIMO Massivos com Redução de Interferência e Estimação de Canal Imperfeita

Thesis presented to the School of Electrical and Computer Engineering of the University of Campinas as partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering, in the area of Telecommunications and Telematics.

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na área de Telecomunicações e Telemática.

Supervisor/*Orientador*: Prof. Dr. Celso de Almeida

Este exemplar corresponde à versão final da tese defendida pelo aluno Carlos Daniel Altamirano Carrillo, e orientada pelo Prof. Dr. Celso de Almeida

---

Campinas
2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

AL79p

Altamirano Carrillo, Carlos Daniel, 1985-
Performance of massive MIMO under interference reduction and imperfect channel estimation / Carlos Daniel Altamirano Carrillo. – Campinas, SP : [s.n.], 2021.

Orientador: Celso de Almeida.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Sistemas MIMO. 2. Sistema de telefonia celular. 3. Probabilidade de erro (Matemática). 4. Estimativa de parâmetro. 5. Antenas (Eletrônica). I. Almeida, Celso de, 1957-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Avaliação de desempenho de sistemas MIMO massivos com redução de interferência e estimação de canal imperfeita
Palavras-chave em inglês:
MIMO systems
Cellular phone system
Probability of errors (Mathematics)
Parameter estimation
Antennas (Electronics)
Área de concentração: Telecomunicações e Telemática
Titulação: Doutor em Engenharia Elétrica
Banca examinadora:
Celso de Almeida [Orientador]
Renato Baldini Filho
Reginaldo Palazzo Júnior
Henry Ramiro Carvajal Mora
Juan Carlos Minango Negrete
Data de defesa: 26-02-2021
Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-3123-1667
- Currículo Lattes do autor: http://lattes.cnpq.br/7258077981778581

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: Carlos Daniel Altamirano Carrillo **RA:** 098422
Data da Defesa: 26 de fevereiro de 2021

Título da Tese: Performance of Massive MIMO Under Interference Reduction and Imperfect Channel Estimation (*Avaliação de Desempenho de Sistemas MIMO Massivos com Redução de Interferência e Estimação de Canal Imperfeita.*)

Prof. Dr. Celso de Almeida (Presidente, FEEC/UNICAMP)
Prof. Dr. Renato Baldini Filho (FEEC/UNICAMP)
Prof. Dr. Reginaldo Palazzo Júnior (FEEC/UNICAMP)
Prof. Dr. Henry Ramiro Carvajal Mora (FICA/UDLA)
Dr. Juan Carlos Minango Negrete (CINEMAROS SAC)

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Tese) e na secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

To my beloved family.
Diana & Felipe.

Acknowledgments

Thank you God for the blessings received during my stay in Brazil. These gave me the strength to successfully complete another stage of my life.

I would like to thank my beloved wife Diana for her love and patience during the time it took me to complete my doctorate, and for all the effort in raising our son Felipe during my absence.

I would also like to thank my mother Rosario for her tender love, dedication, and advice, my father Carlos for supporting my projects, and my sister Analia for the strength, encouragement, support and advice from a distance.

Thank you Prof. Dr. Celso de Almeida, for your patience, help, and continuous encouragement during my graduate studies, as well as providing the opportunity to study in Brazil. You have my sincere and eternal gratitude for the dedication and guidance in the preparation of this work.

My gratitude also goes to the members of my dissertation committee, Prof. Dr. Renato Baldini, Prof. Dr. Reginaldo Palazzo, Prof. Dr. Henry Carvajal, Dr. Juan Carlos Minango, Prof. Dr. Carlos Eduardo Câmara and Prof. Dr. Lee Luan Ling, for reading my dissertation and providing many valuable comments that improved the presentation and content of this document.

I thank my whole family, my uncles and aunts for their support and affection. Special thanks to my in-laws for their continued support during the execution of this goal.

Thanks to my friends for the shared knowledge, for the joy of our moments, for the companionship in Brazil. My special gratitude goes to Prof. Dr. Gonzalo Olmedo for his continued motivation and friendship.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Finally, I would like to thank Universidad de las Fuerzas Armadas ESPE for the financial support.

Abstract

Massive MIMO (M-MIMO) is one of the most promising technologies towards the fifth generation (5G) wireless systems and beyond. M-MIMO uses a large number of antennas at the base station (BS) serving many user terminals (UTs) in a same radio resource. Furthermore, it presents exceptional performance in terms of the ergodic capacity and bit error rate (BER) using simple linear detectors, such as maximal-ratio combining (MRC), zero-forcing (ZF), or minimum-mean-square error (MMSE) due to the interference reduction and the favorable propagation. However, this exceptional performance is reached when time-invariant channels (TIC) and perfect channel state information (CSI) are present.

Impairments introduced by time-variant channels (TVC) and imperfect channel estimation restrict the M-MIMO performance. Some works have evaluated these impairments in terms of the BER using only simulations. Besides, multiplexed pilot estimation (MPE) employing the minimum-mean-square error (MMSE) estimator is widely employed in the channel estimation. However, other channel estimation techniques, such as superimposed pilot estimation (SPE) and hybrid pilot estimation (HPE) that perform differently on TICs and TVCs can be used for M-MIMO.

In general, the uplink performance of single cell M-MIMO systems is evaluated in terms of interference reduction factor (IRF) and BER. For this purpose, the IRF of M-MIMO systems employing linear and planar antenna arrays in line-of-sight (LOS) and non-line-of-sight NLOS channels is evaluated. The normalized mean-square-error (NMSE) of the channel estimator and the system BER is then evaluated for MRC, ZF and MMSE detectors using MPE, SPE and MPE techniques for obtaining the channel estimation on TICs and TVCs models.

The contributions of this dissertation are summarized in the following. Closed-form expressions are derived to evaluate the IRF as a function of the spacing among antennas for linear and planar arrays. Additionally, the condition of favorable propagation for different channel models is presented. On the other hand, MPE, SPE and HPE techniques are described. These channel estimation techniques are evaluated for TICs and TVCs in terms of the NMSE. Furthermore, closed-form expressions of the BER are derived for linear detectors. For this aim, the signal-to-noise-plus-interference ratio (SNIR) is derived for TICs and TVCs, considering the aforementioned channel estimation techniques. Additionally, lower and upper bounds of the BER are also derived. Finally, a comparison among detectors and estimation techniques is presented. In all scenarios, multilevel quadrature amplitude modulation (M-QAM) modulations are considered. Monte Carlo simulations are employed to verify the accuracy of the presented mathematical expressions.

Keywords: Massive MIMO; channel estimation; bit error rate; time-variant channel; time-correlation; linear detection; antenna array; channel state information; inter-user interference; Rayleigh fading.

Resumo

MIMO massivo (M-MIMO) é uma das tecnologias mais promissoras para os sistemas sem fio de quinta geração (5G) e além. M-MIMO usa um grande número de antenas na estação radio base (BS), atendendo a vários aparelhos de usuário (UTs). Além disso, ele apresenta um desempenho excepcional em termos da capacidade ergódica e da taxa de erro de bit (BER) usando detectores lineares, como o combinador de razão máxima (MRC), forçamento de zero (ZF) ou erro quadrático médio mínimo (MMSE) devido à redução da interferência e ao mecanismo de propagação favorável. No entanto, esse desempenho excepcional é alcançado somente na presença de canais invariantes no tempo (TIC) e quando a informação de estado do canal (CSI) é perfeita.

As deficiências introduzidas por canais variantes no tempo (TVC) e por uma estimação de canal imperfeita limitam o desempenho do M-MIMO. Alguns trabalhos avaliaram essas deficiências em termos da BER usando apenas simulações. Além disso, a técnica de estiamção de canal com pilotos multiplexados (MPE) em conjunto com o estimador de erro médio quadrático mínimo (MMSE) são amplamente empregadas. No entanto, outras técnicas de estimação de canal, como a de pilotos sobrepostos (SPE) e a híbrida (HPE), que apresentam desempenhos diferentes em TICs e TVCs, podem também ser usadas em M-MIMO.

Em geral, este trabalho aborda o desempenho do enlace de subida de sistemas M-MIMO para uma única célula em termos do fator de redução de interferência (IRF) e da BER. Para isso, o IRF de arranjos de antenas lineares (unidimensionais) e planares em canais com linha de visada (LOS) e sem linha de visada (NLOS) é calculada. A seguir, o erro quadrático médio normalizado (NMSE) do estimador e a BER são avaliadas para os detectores lineares MRC, ZF e MMSE usando as técnicas de estimação MPE, SPE e HPE em canais TICs e TVCs.

As contribuições desta dissertação são resumidas a seguir. Expressões em forma fechada são derivadas para o IRF em função do espaçamento entre antenas de arranjos lineares e planares. Além disso, as condições de propagação favorável são apresentadas. As técnicas de estimação de canal MPE, SPE e HPE são apresentadas a seguir. Essas técnicas são avaliadas em termor do NMSE para TICs e TVCs. Além disso, expressões em forma fechada da relação sinal ruido mais interferência (SNIR) e da BER são derivadas para os detectores lineares em TICs e TVCs. Também, limitantes inferiores e superiores da BER são derivados. Por último, uma análise comparativa entre os detectores e as técnicas de estimação é apresentada. Em todos os cenários, modulações multinível de amplitude em quadratura (M-QAM) são consideradas. Simulações de Monte Carlo corroboram a precisão das expressões apresentadas.

Palavras-chave: MIMO Massivo, estimação de canal; taxa de erro de bit; canal variante no tempo; correlação temporal; deteção linear; arranjos de antenas; informação de estado do canal; interferência entre usuários; desvanecimento Rayleigh.

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Connotation
3G	third generation
3GPP	Third Generation Partnership Project
4G	fourth generation
5G	fifth generation
AWGN	additive white Gaussian noise
ASK	amplitude shift keying
BER	bit error rate
BPSK	binary phase shift keying
BS	base station
CCI	co-cell interference
CEEI	channel estimation error interference
CSI	channel state information
HPE	hybrid pilot estimation
i.i.d.	independent and identically distributed
IMT	International Mobile Telecommunications
ITU	International Telecommunications Union
IRF	interference reduction factor
LOS	line-of-sight
LS	least-squares
LTE	Long Term Evolution
LTE-A	Long Term Evolution advanced
MAI	multiple access interference
MIMO	multiple-input-multiple-output
MISO	multiple-input single-output
ML	maximum likelihood
M-MIMO	massive multiple-input-multiple-output
MMSE	minimum mean-square-error
M-PAM	multilevel pulse amplitude modulation
MPE	multiplexed pilot estimation
M-QAM	multilevel quadrature amplitude modulation
mmWave	millimeter wave
MRC	maximal ratio combining
MSE	mean-square-error
NLOS	non-line-of-sight
NMSE	normalized mean-square-error

LIST OF ABBREVIATIONS (cont.)

Abbreviation	**Connotation**
PAM	pulse amplitude modulation
PCE	perfect channel estimation
PDF	probability density function
PEP	pairwise error probability
PSK	phase shift keying
SD	sphere detector
SIMO	single-input multiple-output
SISO	single-input single-output
SNR	signal-to-noise ratio
SNIR	signal-to-noise-plus-interference ratio
SPE	superimposed pilot estimation
TDMA	time division multiple access
TIC	time-invariant channel
TVC	time-variant channel
ULA	uniform linear array
UPA	uniform planar array
UR-LOS	uniform random line-of-sight channel
UT	user terminal
ZF	zero-forcing

List of Symbols

Symbol	Connotation
$A(\ell)$	pulse amplitude for the ℓth time interval
B	system bandwidth
B_c	channel coherence bandwidth
B_D	Doppler spread
d	separation between antennas
E_B	received energy per block
E_b	received energy per bit
E_d	received data energy
E_p	received pilot energy
E_s	received energy per symbol
f_c	carrier frequency
f_d	Doppler shift
$f_{d,\max}$	maximum Doppler shift
$\mathbb{i}$	imaginary unit, $\sqrt{-1}$
K	total number of single antenna UTs
$\mathcal{K}$	propagation factor
L	total number of symbols
$\mathcal{L}$	array length
L_d	number of data symbols
$L_{d,m}$	number of multiplexed data for HPE
L_p	number of pilot symbols
L_s	number of superimposed symbols
M	modulation order
$\mathcal{M}$	multiple access signals
N	total number of BS antennas
N_0	unilateral noise power spectral density
P_r	received power
P_t	transmitted power
$p(x)$	PDF of the random variable x
$\mathrm{p}(t)$	unitary baseband pulse
$\mathcal{R}$	radiation pattern
$\mathcal{R}_{b,k}$	kth user bit rate
r	distance between a UT and the BS

LIST OF SYMBOLS (cont.)

Symbol	**Connotation**
T_B	block time interval
T_c	channel coherence time
T_d	channel delay spread
T_s	symbol time duration
v	UT speed
$\overline{\alpha^2}$	Rayleigh fading mean power
β	ratio K/N
Δ	normalized separation between antennas
ε	propagation path-loss exponent
η	ratio between the number of pilot and data symbols for MPE
η_h	ratio between the number of superimposed and multiplexed data symbols for HPE
γ_a	SNIR asymptote
$\gamma_s(\ell)$	average SNIR at the ℓth time interval
γ_s	average SNIR per block
λ	eigenvalue
λ_c	carrier wavelength
μ	ratio between the energy of pilot and data symbols for MPE
μ_s	ratio between the energy of pilot and data symbols for SPE
$\mu_{h,m}$	ratio between the energy of superimposed and multiplexed data symbols for HPE
$\mu_{h,m}$	ratio between the energy of superimposed pilot and data symbols for HPE
σ_w^2	complex noise variance
ϱ	variance of the estimated channel coefficient
$\tilde{\varrho}$	variance of the channel estimation error or MSE
Υ	interference reduction factor
ξ	cellular spectral efficiency
$\mathbf{x}$	vector
$\mathbf{X}$	matrix
$\mathbf{X}_{i,j}$	element (i,j) of matrix $\mathbf{X}$
$\underline{\mathbf{X}}_{i,j}$	matrix obtained removing the i-th row and j-th column of matrix $\mathbf{X}$
$\overline{x}$, $\mathrm{E}\{\cdot\}$	expectation, mean
$\lfloor\cdot\rfloor$	floor operation
$\lvert\cdot\rvert$	complex modulus
$\lVert\cdot\rVert$	Euclidean norm
$(\cdot)^*$	conjugate
$(\cdot)^T$	transpose
$(\cdot)^H$	Hermitian transpose
$(\cdot)^+$	Moore-Penrose pseudoinverse

LIST OF SYMBOLS (cont.)

Symbol	**Connotation**
$\otimes$	Kronecker product
$\forall$	for all
$\mathbf{C_{xy}}$	covariance matrix of vectors $\mathbf{x}$ and $\mathbf{y}$
$J_0(\cdot)$	Bessel function of the first kind of order zero
$\mathcal{O}(\cdot)$	upper bound of complex operations
$\Re\{\cdot\}$	real part of its argument
$\mathbf{R_{xy}}$	correlation matrix of vectors $\mathbf{x}$ and $\mathbf{y}$
$\det(\mathbf{X})$	determinant of matrix $\mathbf{X}$
$\mathrm{diag}(\mathbf{x})$	diagonal matrix of vector $\mathbf{x}$
$\mathrm{Var}\{\cdot\}$	variance

List of Publications

Journal Articles

- Carlos Daniel Altamirano and Celso De Almeida. Inter-user interference reductionin massive MIMO for linear and planar arrays. *International Journal on Communications Antenna and Propagation (IRECAP)*, 9(1):30, feb 2019.

- Carlos Daniel Altamirano, Juan Minango, Henry Carvajal Mora, and Celso De Almeida. BER evaluation of linear detectors in massive MIMO systems under imperfect channel estimation effects. *IEEE Access*, 7:174482–174494, 2019.

- Carlos Daniel Altamirano, Juan Minango, Celso de Almeida, and Nathaly Orozco. On the asymptotic BER of MMSE detector in massive MIMO systems. In *Communications in Computer and Information Science*, pages 57–68. Springer International Publishing, 2020.

- Carlos Daniel Altamirano, Henry Carvajal Mora, and Celso De Almeida. BER of massive MIMO in time-variant channels using multiplexed, superimposed and hybrid channel estimation techniques. *AEU - International Journal of Electronics and Communications,* 131:153594, jan 2021. .

Conference Articles

- Carlos Daniel Altamirano and C. de Almeida, Inter-user interference reduction factor for 3-D massive MIMO systems, In *2016 8th IEEE Latin-American Conference on Communications (LATINCOM)*, IEEE, nov 2016.

Contents

CHAPTER 1

Introduction

The development of wireless communication systems promises a continuous increase in mobile data traffic, allowing operators to offer several services that can range from a simple call, to data services or even newer applications or services [1, 2]. The continuous development motivated by the growing requirements has brought new challenges for both academic and industry researches to pursue innovative strategies that enable reliable communications with higher data rates, lower latency and bit error rate and greater spectral efficiency and coverage.

Wireless communication systems present several operational impairments such as path loss, multipath fading, co-channel interference, additive white Gaussian noise (AWGN), and others. In order to overcome these undesirable effects, several transmission-reception mechanism have been employed by wireless communication standards. In particular, the fourth generation (4G) of cellular networks has introduced a technique that uses multiple antennas at both the transmitter and the receiver, known as multiple-input multiple-output (MIMO) systems [3]. This solution provides spatial multiplexing and diversity to the wireless system, achieving higher data rates and reliability without any expansion of the bandwidth or increase in the system transmission power. Nowadays, MIMO systems employing at most 8 antennas at both ends are a mature and well understood technology [4]. Furthermore, MIMO is part of some wireless communication standards, such as 802.11n/Wi-Fi and 802.16/WiMAX [5, 6].

However, 4G systems are proving to be insufficient to meet future demands of services and applications [7]. The fifth generation (5G) of cellular networks has established new parameters in order to fulfill the International Mobile Telecommunications (IMT) 2020 requirements [8], with wide impact not only in the technology but also in the economics by creating new business models [9, 10]. In this context, 5G introduces a new network core and a new radio interface [11]. One of the new radio innovations is the deployment of MIMO systems with a large number of antennas at the base station (BS) serving a limited number of single antenna user terminals (UT), known as massive MIMO (M-MIMO). M-MIMO has emerged as one of the most promising technologies since it allows a much

higher spectral efficiency and, consequently, very high data rates without requiring extra bandwidth, or time resources, or additional transmit power [12]. M-MIMO improves the ordinary MIMO performance, because it exploits the advantages of extra degrees of freedom provided by using more antennas.

The remainder of this chapter is organized as follows. In Section 1.1, related works in the area of M-MIMO systems are summarized and the motivation of this study is also introduced. Finally, the contributions and the outline of this dissertation are presented in Section 1.2.

1.1 Related Work and Motivation

The concept of M-MIMO was introduced by Marzetta in his seminal paper, "Noncooperative cellular wireless with unlimited numbers of base station antennas" [13]. Succeeding works have shown that the fundamental concepts of M-MIMO has improved ordinary MIMO, allowing reliable communication with high throughput and power efficiency using simple linear processing [14–16].

In [17], it was established that M-MIMO takes advantage of the channel favorable propagation property, defined by the mutual orthogonality among the channel vectors of the UTs. Favorable propagation has the effect of reducing the inter-user interference [18], which can be assessed by the interference reduction factor (IRF) [19]. One condition to achieve favorable propagation is to employ very large antenna arrays, which could be impractical for real BS with limited physical space [20, 21]. Furthermore, a dense array of antennas in a limited space does not reduce the inter-user interference due to the spatial correlation [22, 23]. The large size of M-MIMO arrays is a fundamental problem for its practical deployment. Thus, a study that evaluates the IRF of several antennas array on different channels is necessary for the M-MIMO development.

The IRF is a basic performance indicator for M-MIMO, however, in order to evaluate the overall performance of M-MIMO, other performance indicators may be used. In most of the literature, the performance of M-MIMO is evaluated in terms of the ergodic capacity [12, 15, 24–28]. Moreover, other more practical performance evaluations measures could also be used, such as the bit error rate (BER), pair-wise error probability (PEP), or outage probability [29–32]. The BER is often used, once it assesses the full end to end performance of wireless systems, including the transmitter, receiver and channel impairments. Thus, it is interesting to evaluate the BER of M-MIMO systems.

M-MIMO also presents near-optimal performance for simple linear detectors such as zero-forcing (ZF) and minimum-mean-square error (MMSE), assuming perfect channel state information (CSI) [31, 33–36]. Furthermore, maximal ratio combining (MRC), which is the simplest linear detector, presents an exceptional performance [37]. In the literature,

the performance of M-MIMO systems is commonly evaluated considering perfect CSI, as shown in the works mentioned above.

CSI is obtained by channel estimation, which is a classical problem of communication systems. Usually, MIMO systems use pilot symbols to aid with channel estimation [38]. Thus, for the uplink, the CSI is obtained at the BS by using known orthogonal pilot sequences transmitted by each UT. Commonly, pilot symbols are time-multiplexed with data. In the literature, this technique is named multiplexed pilot estimation (MPE), conventional pilot (CP), or regular pilot (RP) [39]. Some studies suggest that it may be advantageous to superimpose the pilot and data, this technique is named superimposed pilot estimation (SPE), that has been studied in the channel estimation of MIMO systems and used to solve synchronization issues [40, 41]. It is considered an excellent technique to combat pilot contamination in multi-cell M-MIMO [42, 43]. Furthermore, a hybrid pilot estimation (HPE) that consist of combining MPE and SPE techniques is explored for M-MIMO in [44].

In [45], a survey on M-MIMO systems in presence of channel and hardware impairments is presented. In that work, in order to study the channel impairments, the channel dispersion is divided into four categories: In the first category, the channel is assumed time-invariant and the fading is flat. In the second category, the channel is time-variant and the fading is flat. In the third one, the channel is time-invariant and the fading is frequency-selective. Finally, the last category corresponds to a time-variant and frequency-selective fading. The authors conclude that channel issues need to be addressed, particularly the estimation and channel aging effects, especially in fast mobility environments. Therefore, the channel estimation effects on the performance of M-MIMO systems should be analyzed for some channel models and estimation techniques.

Most of the M-MIMO studies lie on the first category defined in [45], which considers a time-invariant channel (TIC) and flat fading. In general, the system performance of M-MIMO using MPE is evaluated in terms of the ergodic capacity in [15, 24, 26] and in terms of BER in [30, 46, 47]. On the other hand, the ergodic capacity of M-MIMO using SPE is evaluated in [42, 48, 49] and the BER based on simulation in [43] for an iterative detector employing the least-squares (LS) estimator, and in [50] by using estimation aided by second-order statistics. Finally, the performance of M-MIMO for HPE in [44, 51], in terms of the ergodic capacity and in [43] in terms of the BER. As a consequence, there are a limited number of studies that evaluates M-MIMO performance in terms of the BER using SPE and HPE.

The second category in [45] considers a time-variant channel (TVC), where the channel coefficients vary in a coherence time-interval due to the Doppler effects caused by the UTs motion [52, 53]. The effects of time-variant coefficients in the uplink performance of M-MIMO in terms of the channel capacity using both MRC and ZF detectors are presented in [54], where another reason for CSI inaccuracy, named as channel aging, is

also introduced. The channel aging problem is due to delay between the channel estimation instant and the detection usage in MPE and HPE techniques. The effects of channel aging on the sum-rate and on the spectral efficiency were investigated for a multi-cell M-MIMO system with regularized ZF precoder and ZF detector in [55] and [56]. The results show that channel aging degrades the system performance severely. However, using channel prediction techniques, the effects of channel aging can be minimized. It was also shown that for UT at higher speeds, the maximum achievable rates are reduced. Further analysis of M-MIMO performance using MPE can be found in [57–60]. On the other hand, the effects of TVCs on the channel capacity of M-MIMO systems using SPE technique for high speeds are shown in [61]. Based on the literature review, there is a lack of investigation for both SPE and HPE techniques on TVCs. Furthermore, the BER analysis of M-MIMO systems on TVCs is scarce.

The third and fourth categories described in [45] for frequency-selective fading goes beyond the proposal of this work. However, some useful references related to this topic are [62, 63].

By the aspects mentioned in previous paragraphs, evaluating the impairments introduced by TVCs on the channel estimation quality and on the BER of M-MIMO systems it is a relevant topic that deserves attention. Hence, evaluating the BER of M-MIMO systems employing MRC, ZF and MMSE detectors using MPE, SPE and HPE techniques on TICs and TVCs is one of the contributions of this thesis.

1.2 Contributions and Outline of the Dissertation

This thesis encompasses several results in the performance evaluation of M-MIMO systems. The contributions are addressed in separate chapters raised from the motivations described in Section 1.1. The first contribution is in terms of the IRF, as a result of the M-MIMO channel analysis. The second contribution in terms of the BER, for some linear detectors and some channel estimation techniques for time-variant and time-invariant channel models. For this purpose, some basic concepts of digital communications, cellular systems, MIMO systems, linear detection and channel estimation are presented. The remainder of the dissertation is structured as follows:

- **Chapter** 2. The aim of this chapter is to provide some basic concepts for the analysis performed in the next chapters. A brief review of fundamental concepts such as digital modulation, cellular systems, MIMO and M-MIMO systems and their mathematical representation is presented. Moreover, the principles of linear detection and channel estimation are introduced.

- **Chapter** 3. This chapter depicted the contributions of the published articles [64] and [65]. Hence, exact closed-form expressions to evaluate the IRF for linear and

planar antenna arrays are presented. Furthermore, based on the array geometry, a study of the relationship between IRF and favorable propagation is presented. This chapter also describes the channel models for the M-MIMO performance evaluation.

- **Chapter** 4. This chapter presents the contributions based on the results published in [47] and [66]. The M-MIMO system model is described and some channel estimation techniques for M-MIMO systems are presented. These channel estimation techniques are evaluated for TIC and TVC channel models in terms of the normalized mean-square-error (NMSE).

- **Chapter** 5. The contributions of the results published in [34, 47, 66, 67] are presented in this chapter. The performance of linear detectors in M-MIMO systems in terms of the BER is outlined, the signal-to-noise-plus-interference ratio (SNIR) of linear detectors is derived and the performance is evaluated for TIC and TVC channel models by using the expressions obtained in Chapter 4. Exact closed-form expressions are obtained to evaluate the BER. Moreover, lower and upper bounds on the BER are also derived. Finally, a comparison analysis between the detectors and estimation techniques is presented.

- **Chapter** 6. This chapter summarizes the main contributions of this work and some proposals for future works are presented.

CHAPTER 2

Basic Concepts

2.1 Introduction

In this chapter, the basic concepts of wireless communication systems are described. The wireless communication systems face several challenges, including the limited availability of the radio frequency spectrum and the distortion and degradation by the channel effects.

For improving the quality of the communication service to the users, splitting a geographic region into cells is a widely used technique. Furthermore, in order to increase data rates efficient modulation techniques are required.

Thus, an introduction to cellular systems and modulation techniques are presented in this chapter. Moreover, the characteristics of the propagation phenomena that distort and degraded the radio signals are also presented.

Multiple-input-multiple-output (MIMO) systems exploit some channel characteristics in order to increase system capacity and reliability. Moreover, the benefits of increasing the number of antennas are harnessed by massive multiple-input-multiple-output (M-MIMO) systems. Additionally, MIMO systems require efficient detection techniques. Therefore, low complexity linear detectors are presented in this chapter. Furthermore, as the channel state (CSI) information is required for linear detection, the channel estimation is also discussed.

For better understanding, this chapter is organized as follows. The basics of digital modulation is presented in Section 2.2. The cellular systems principles are described in Section 2.3. The mobile radio channel is analyzed in Section 2.4. MIMO systems are presented in Section 2.5. Linear detectors are detailed in Section 2.6. Finally, the channel estimation is described in Section 2.7.

2.2 Digital Modulation Schemes

Modulation shifts the spectrum of a baseband digital signal to a radio channel with frequency carrier f_c. Digital modulation performs the mapping of a block of bits into a waveform. The waveform can be obtained by changing one of the following parameters of a carrier: amplitude, phase or frequency. Furthermore, the set of waveforms can be represented by a constellation diagram as shown in Fig. 2.1 for a particular modulation.

At the transmitter, the modulation stage is a mapper followed by an up-converter. At the receiver, the signals undergo a down-converter, a matched-filter, a decision rule device and finally a demapper [68]. In the following, some concepts regarding the modulation process are presented and the digital modulation schemes used in this thesis are also described.

2.2.1 PAM Modulation

Pulse amplitude modulation (PAM) consists in transmitting digital information in baseband through the amplitudes of a pulse train. Hence, PAM signals can be written as [68]:

$$x(t) = \sum_{\ell=-\infty}^{\infty} A(\ell)\mathrm{p}(t - \ell T_s), \tag{2.1}$$

where $A(\ell)$ is the signal amplitude at the ℓth time interval, T_s is the symbol time duration and $\mathrm{p}(t)$ is a baseband pulse format that satisfies the Nyquist criterion. If the PAM signal has M different amplitudes, then the modulation scheme is known as M-PAM (multilevel pulse amplitude modulation), which transmits $\log_2 M$ bits in each symbol.

2.2.2 ASK Modulation

Including a sinusoidal carrier in the PAM signals, the bandpass modulation scheme is named amplitude shift keying (ASK) modulation. Hence, the ASK scheme can written as [68]:

$$s(t) = x(t)\cos(2\pi f_c t), \tag{2.2}$$

where f_c is the carrier frequency and $x(t)$ is the M-PAM signal. In this scheme, the digital information is transmitted on the amplitudes of the sinusoidal carrier, that is, a unique combination of bits is attributed to each amplitude. For example, if $M = 2$ and the set of amplitudes is $\{-A, A\}$, the modulation is called 2-ASK. For the binary case, amplitude modulation is equivalent to phase modulation, that is named 2-PSK or BPSK (binary phase shift keying) modulation [69].

2.2.3 M-QAM Modulation

Multilevel quadrature amplitude modulation (M-QAM) consist in transmitting information using two orthogonal $\sqrt{M}$-ASK modulation schemes. The ASK carrier waves have the same frequency and are out of phase with each other by $\pi/2$ radians. Thus, by the orthogonality, there is no interference between carriers despite of they use the same bandwidth. Both carriers are typically called quadrature carriers. Hence, M-QAM signal can be written as:

$$s(t) = x_i(t)\cos(2\pi f_c t) - x_q(t)\sin(2\pi f_c t), \tag{2.3}$$

where $x_i(t)$ and $x_q(t)$ are independent in-phase and in-quadrature PAM signals, respectively.

M-QAM scheme has $M = 2^k$ waveforms, where $k \geqslant 2$ represents the number of bits transmitted per symbol. If k is an even number, then the M-QAM constellation can be obtained as the Cartesian product of two $\sqrt{M}$-ASK constellations. Besides, the baseband or lowpass equivalent of the M-QAM signal can be obtained by the cartesian product of two $\sqrt{M}$-PAM constellations, which is given by:

$$x(t) = x_i(t) + \mathbb{i}x_q(t), \tag{2.4}$$

where $\mathbb{i} = \sqrt{-1}$. Furthermore, the PAM amplitudes are $\pm A, \pm 3A, \cdots, (\sqrt{M}-1)A$, as can be seen in Fig. 2.1a and 2.1b for modulations 4-QAM and 16-QAM, respectively. Thus, the mean power of the baseband M-QAM constellation is equal to:

$$\overline{|x|^2} = \frac{2}{3}(M-1)A^2. \tag{2.5}$$

On the other hand, the energy per symbol of a bandpass M-QAM signal is equal to:

$$E_s = \frac{1}{2}\overline{|x|^2}T_s. \tag{2.6}$$

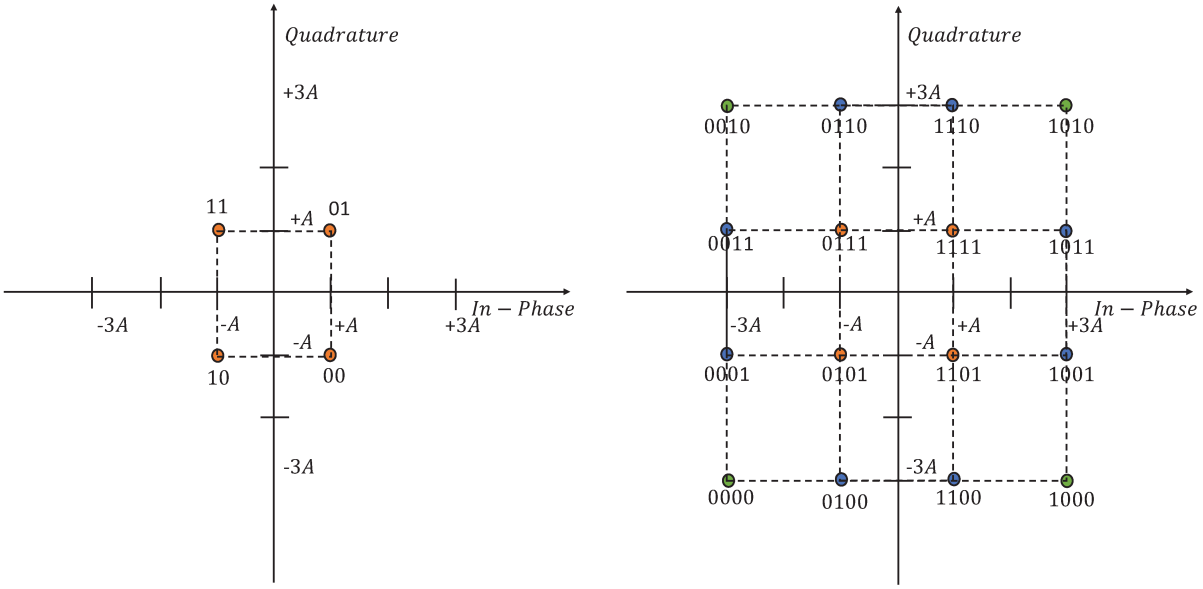
Furthermore, the bit energy is given by:

$$E_b = \frac{E_s}{\log_2 M}. \tag{2.7}$$

The M-QAM modulation has been widely used in the last generations of cellular systems because it uses the radio spectrum efficiently.

2.2.4 Gray Mapping

Gray mapping is a labeling technique of the symbols of modulation schemes in terms of bits, where the closest symbols differ in only one bit position, which minimizes the



(a) 4-QAM modulation.

(b) 16-QAM modualtion.

Figure 2.1: M-QAM modulation.

BER. Fig. 2.1 shows 4-QAM and 16-QAM constellations employing Gray mapping.

2.3 Cellular Systems

In cellular systems, a geographical region is ideally divided into small circular areas called cells, whose coverage is provided by a base station (BS) at the center of each circular area. Each BS communicates with the user terminals (UT) in its cell. The radio link for signals transmission from the BS to the UTs is known as downlink and the radio link for signals transmission from the UTs to the BS is called uplink [70].

The advantage of cellular systems is reduced transmission power, radio spectrum, congestion problems and increases the system capacity, providing voice and data services, even in crowded regions. As the uplink performance is poorler than the downlink, it is usual to consider it in the performance analysis. In this thesis, the uplink performance analysis is taken into consideration.

2.3.1 Interference

Interference is all unwanted signals arriving at a receiver. It is considered as the primary limiter of the cellular system capacity. In the uplink, the interference is produced by other UTs in the same cell and in co-cells. In the downlink, the interference is produced by the BSs in co-cells. Generally, the mean interference affecting the uplink is stronger than that affecting the downlink [71].

There are different sources of interference. The interference produced by users in the same cell is called multiple access interference (MAI). The interference produced by users in co-cells is known as co-cell interference (CCI). These types of interference are produced when users employ the same radio channels simultaneously. Notice that multi cell systems include both CCI and MAI. A simpler single cell model does not include CCI.

Errors during the channel estimation process can be seen as added interference named channel estimation error interference (CEEI). The CEEI is analyzed in this thesis.

Finally, the effects of thermal noise can be mitigated by increasing the transmission power, on the contrary, the interference effects can not be diminished as the transmission power increases. Consequently, techniques should be employed in order to reduce the interference effects.

2.3.2 Cellular Spectral Efficiency

Spectral efficiency refers to the total information rate that can be transmitted in a cell through a given channel bandwidth. In cellular systems, the cellular spectral efficiency (ξ) is evaluated in bits/s/Hz. Thus, the cellular spectral efficiency can be defined as the ratio of the total cell throughput and the total system bandwidth, that is [72]:

$$\xi = \frac{\sum_{k=1}^{K} \mathcal{R}_{b,k}}{B}, \tag{2.8}$$

where K is the total number of UTs in the cell, $\mathcal{R}_{b,k} = \log_2(M_k)/T_s$ is the k-th UT bit rate and $B = 1/T_s$ is the system minimum Nyquist bandwidth for passband modulation.

2.4 Mobile Radio Channel

The propagation of radio signals through wireless channels involves different phenomena that create distortion and attenuation. The propagation phenomena can be divided into large-scale and small-scale fading [73].

2.4.1 Large-Scale Fading

Large-scale fading is the result of signal attenuation due to signal propagation over large distances and diffraction around large objects between the transmitter and receiver. This type of fading is relatively slow and is responsible for long-term signal power variations. Furthermore, there are two large-scale fading phenomena: path-loss and shadowing [71].

2.4.1.1 Path-Loss

The received power depends on the transmitted power and it is a function of the distance between the transmitter and the receiver. Hence, the received power is given by:

$$P_r = \mathcal{K} P_t r^{-\varepsilon}, \tag{2.9}$$

where P_t is the transmitted power, r is the distance between a UT and the BS, $\mathcal{K}$ is the propagation factor and ε is the propagation path-loss exponent. These last two parameters depend on the employed propagation model. There are several propagation models in the literature, such as Okumura-Hata, COST231-Hata and the SUI model, which include the path-loss phenomena [70, 74].

2.4.1.2 Shadowing

Shadowing is produced by obstacles between the UTs and the BS. Thus, as the obstacles size and position are random, the shadowing produces fluctuations of the received power around its mean value given by the path-loss. The received power can increase or decrease considerably. Shadowing is typically modeled by a log-normal distribution [70].

2.4.1.3 Power Control

In cellular systems, the power levels transmitted by each UT are under constant control by the corresponding BS. In this way, power control contributes to several functionalities, such as connectivity, power economy and interference management. Moreover, the power control solves the near-far problem and allows to extend the UTs battery life [70].

The power control consists of adjusting the power transmitted by all UTs in the same cell, so that the received power from all UTs is the same at the BS. Thus, the transmitted power for each UT is given by:

$$P_t = \mathcal{K}^{-1} P_{r,0} r^{\varepsilon}, \tag{2.10}$$

where $P_{r,0}$ is the constant received power at the BS from each UT in the same cell, r is the distance between a UT and its BS, $\mathcal{K}$ is the propagation factor and ε is the propagation path-loss exponent. Regulation standards for wireless systems consider that $P_t \leqslant P_{t,\max}$, where $P_{t,\max}$ is the maximum transmission power of a UT. Notice that power control compensates the effects of large-scale fading.

Finally, the power control used in 3G and 4G cellular systems is considered almost perfect [75]. Thus, perfect power control is considered along this thesis.

2.4.2 Small-Scale Fading

The small-scale fading is due to the constructive or destructive combination of randomly delayed multipath signal components. These components may be originated by reflection, scattering and diffraction phenomena. Some factors influence on the small-scale fading, such as the the UT speed, surrounding objects speed and the signal bandwidth. This type of fading is relatively fast and is therefore responsible for the short-term signal variations. Depending on the nature of the radio propagation environment, there are different models describing the statistical behavior of the multipath fading envelope [73].

Small-scale variations of a radio signal is directly related to the channel impulse response. Thus, the radio channel can be modeled as a linear filter with a time-variant impulse response that is a function of t and τ, where t represents time variations due to the UT movement and τ denotes multipath delays for a fixed t. Thus, the low-pass equivalent of the channel impulse response is given by [69]:

$$h(t,\tau) = \sum_{l=1} \alpha_l(t) \exp\left[- \mathbb{i}\phi_l(t)\right] \delta\left[\tau - \tau_l(t)\right], \tag{2.11}$$

where $\alpha_l(t)$, $\phi_l(t) = 2\pi f_c \tau_l(t)$ and $\tau_l(t)$ are the amplitude, phase and delay of the l-th multipath at the instant of time t, respectively, and $\delta(\tau)$ denotes the delta of Dirac. Observe that the filter model sum signals with random amplitudes, phases and delays. Furthermore, the time-variant nature is produced by the UT movement.

2.4.2.1 Line-of-Sight Channel

In the line-of-sight (LOS) case, the received signal is composed by the direct path component between transmitter and receiver. Thus, the low-pass equivalent of the channel impulse response given by (2.11) can be rewritten as:

$$h(t,\tau) = \alpha(t) \exp\left[- - \mathbb{i}\phi(t)\right] \delta[\tau]. \tag{2.12}$$

By considering the fading amplitude $\alpha(t)$ as deterministic and modeling $\phi(t)$ as a uniform random variable over $[0, 2\pi)$, this channel is named uniform random line-of-sight (UR-LOS) [21].

2.4.2.2 Non-Line-of-Sight Channel

In the non-line-of-sight (NLOS) case, the received signal is composed by different components that are replicas of the transmitted signal, among which there is no dominant component. As the number of multipaths is high and the amplitudes, delays and phases of the multipaths are random, the low-pass equivalent of the channel impulse response can be modeled as a zero-mean complex Gaussian stochastic process. Thus, the envelope of

this process ($\alpha = |h(t,\tau)|$) can be modeled by a Rayleigh distribution, whose probability density function (PDF) is given by [69]:

$$f(\alpha) = \frac{\alpha}{\sigma^2} \exp\left(\frac{\alpha^2}{2\sigma^2}\right), \quad \alpha \geqslant 0, \tag{2.13}$$

where σ^2 is the Rayleigh parameter. Additionally, the resultant phase (ϕ) of the Gaussian process is a random variable assumed uniformly distributed over $[0, 2\pi)$.

2.4.2.3 Small-Scale Fading Parameters

The small-scale fading can be classified in different types according to some parameters that describe the dispersive nature of the channel. These parameters are detailed below.

Delay spread, T_d, is defined as the propagation time difference between the shortest and the longest path, counting only the paths with significant energy. Hence, the delay spread is a measure of the channel impulse response duration [76].

Channel coherence bandwidth, B_c, shows how quickly the channel changes in frequency. Furthermore, it is related to the inverse of the delay spread and can be written as [76]:

$$B_c \approx \frac{1}{2T_d}. \tag{2.14}$$

Two signals separated by a frequency range much greater than B_c are affected by independent fading [70].

Another important parameter is the channel temporal variation. In this case, when a UT moves, the carrier frequency shifts upward or downward, depending if the UT approaches or moves away from the BS. In fact, the power spectrum is not only shifted but spreaded in the interval $f_c - f_d \leqslant f \leqslant f_c + f_d$, where f_c is the carrier frequency and f_d is the Doppler shift. The power spectrum depends on the maximum Doppler shift $f_{d,\max} = v f_c/c$, where v is the UT speed and c is the light speed. Therefore, the channel time variation is described by the Doppler spread.

Doppler spread, B_D, is defined as the range of frequencies over which the power spectrum is non-zero. If the bandwidth of the transmitted signal is much greater than B_D, then the Doppler spread effects are negligible [70].

Channel coherence time, T_c, is the time interval within the received signals present high time correlation. The channel coherence time is inversely proportional to the Doppler spread, that is [76]:

$$T_c \approx \frac{1}{2B_D}, \tag{2.15}$$

where $B_D = f_{d,\max}$.

2.4.2.4 Types of Small-Scale Fading

By considering the parameters described in the previous subsection, the small-scale fading can be classified in the next types:

Non-frequency-selective or flat fading occurs when the coherence bandwidth is much larger than the bandwidth of the transmitted signal, that is, $B_c \gg B$. In this case, the spectral components of the signal are susceptible to the same magnitude of fading. Thus, the channel impulse response can be approximated by a Dirac delta function and there is no ISI.

Frequency-selective fading occurs when the coherence bandwidth is smaller than the transmitted signal bandwidth, that is, $B_c < B$. The spectral components of the signal are attenuated by different magnitudes of fading causing significant distortion in the signal.

Slow fading occurs when the coherence time interval of the channel is much greater than the symbol time interval, that is $T_c \gg T_s$. Thus, the channel can be considered roughly invariant during a symbol time interval T_s.

Fast fading occurs when the coherence time interval is less than the symbol duration, that is $T_c < T_s$. In this case, coherent detection can not be employed.

2.4.2.5 Slow-Flat Time-Variant Channel

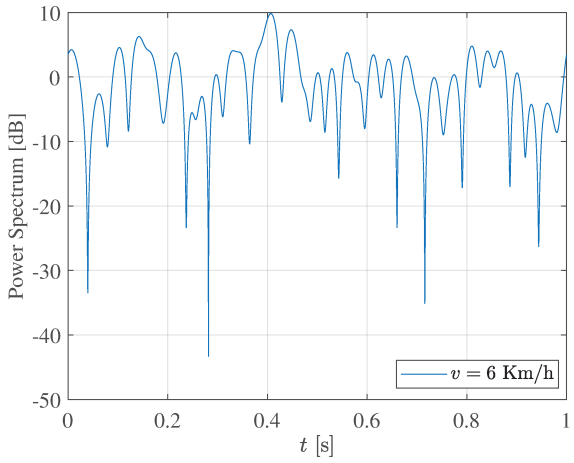
The slow-flat time-variant channel (TVC) considers the Doppler shift effects. Further, the Doppler effects are part of the channel impulse response, by considering that the phase shift is equal to $\phi_\ell(t) = 2\pi(f_c + f_d)\tau_\ell(t)$ in (2.11), where $f_d = v f_c \cos(\theta)/c$ is the Doppler shift that depends on the UT speed v, the light speed c, the frequency carrier f_c and the angle θ between the direction of UT motion and the direction of arrival of the electromagnetic wave.

By employing the Jakes' model [77–79], the channel can be modeled as a zero-mean complex wide-sense stationary Gaussian process with auto-correlation function given by [80]:

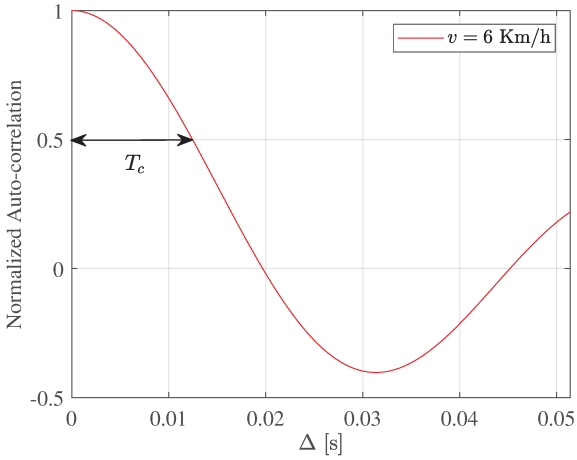
$$\mathrm{R}_h(\Delta) = \overline{\alpha^2} J_0\left(2\pi f_{d,\max}\Delta\right), \tag{2.16}$$

where $\overline{\alpha^2}$ is the Rayleigh fading mean power, $J_0(\cdot)$ is the Bessel function of the first kind of order zero, $f_{d,\max}$ is maximum Doppler shift and Δ is the time delay. Therefore, in this case, the transmitted symbols are affected by a time-correlated fading.

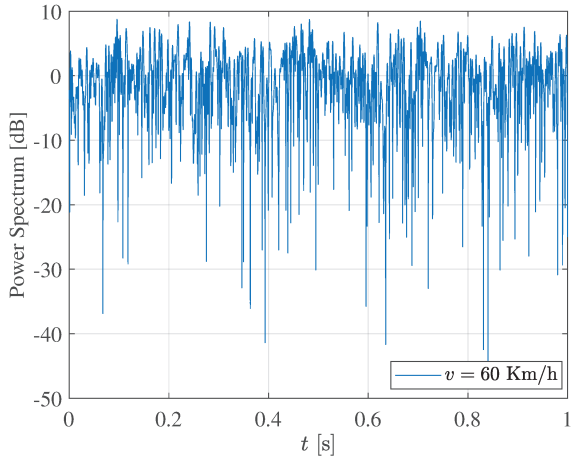
Fig. 2.2 presents an example of the fading power spectrum as a function of time and the fading normalized auto-correlation function as a function of the delay for $\overline{\alpha^2} = 1$ and by considering $f_c = 3.5$ GHz. In particular, Fig. 2.2a and 2.2c presents a snapshot of the fading time variation during one second, for $v = 6$ km/h and $v = 60$ km/h, respectively. Observe that as the UT speed increases the fading envelope variation becomes faster. Furthermore, Fig. 2.2b and 2.2d shows the fading normalized channel auto-correlation



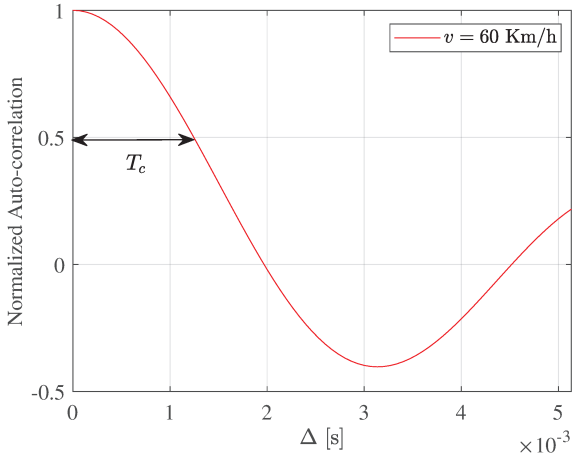
(a) Fading power spectrum for $v = 6$ km/h.



(b) Fading normalized auto-correlation function for $v = 6$ km/h.



(c) Fading power spectrum for $v = 60$ km/h.



(d) Fading normalized auto-correlation function for $v = 60$ km/h.

Figure 2.2: Time-variant channel example.

function for $v = 6$ km/h and $v = 60$ km/h, respectively. Notice that as the UT speed increases the channel coherence time T_c is reduced. The channel coherence time can also be considered as the time interval where the normalized channel auto-correlation function varies less than 50%.

2.4.2.6 Slow-Flat Time-Invariant Channel

The slow-flat time-invariant channel (TIC) is a special case of the slow-flat TVC, where $f_{d,\max} = 0$. Hence, the channel spectrum is flat and the fading is considered invariant during a coherence time interval T_c. Therefore, the fading remains invariant during a transmitted symbol, whose envelope has Rayleigh distribution given by (2.13) and the phase is uniformly distributed over $[0, 2\pi)$, that is, a zero-mean complex Gaussian random variable.

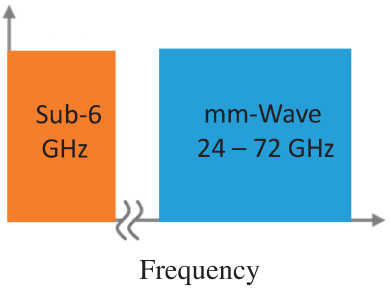


Figure 2.3: 5G Frequency Spectrum.

2.4.3 Radio Spectrum

The radio spectrum is the part of the electromagnetic spectrum with frequencies from 30 Hz to 300 GHz. In order to avoid interference between different services, the radio spectrum is strictly regulated by national laws and coordinated by the International Telecommunication Union (ITU). Besides, each wireless technology standard operates in a particular range of frequencies.

For 5G technology the radio spectrum is separated into two different frequency ranges as shown in Fig. 2.3. The first includes the sub-6GHz frequency band, part of which is currently being used by 3G and 4G cellular technologies, but according to IMT-2020 potential new spectrum is available. The second range includes millimiter waves (mmWaves) from 24 GHz to 72 GHz [11].

As aforementioned, the propagation characteristics depend on the carrier frequency. Consequently, in this work we consider the 3.5 GHz band because it provides a rich scattering that can be modeled as a Rayleigh fading. This bandwidth is a strong candidate for 5G networks deployment [81].

2.5 MIMO Systems

Depending on the number of transmit and receive antennas, wireless communication systems may have different configurations. Wireless systems with one transmit and one receive antenna are named as single-input single-output (SISO) systems. Systems having one transmit antenna and multiple receive antennas are said to be single-input multiple-output (SIMO). Systems having multiple transmit antennas and one receive antenna are said to be as multiple-input single-output (MISO). Finally, systems having multiple transmit antennas and multiple receive antennas are said to be multiple-input-multiple-output (MIMO).

In the past few decades, SISO systems were widely used. However, the spectral efficiency of these systems is limited by the modulation order, because if the modulation

order is increased, the system become more susceptible to fading effects [72]. Therefore, systems with high spectral efficiency are required nowadays.

Wireless systems with multiple transmit antennas present high spectral efficiency, once the antenas can transmit simultaneously on the same radio resource. On the other hand, wireless systems with multiple receive antennas present high performance, once the antennas provides spatial diversity and spatial multiplexing as detailed below.

2.5.1 Spatial Diversity

Spatial diversity is a technique for combating the undesirable fading effects, once it exploits multipath signals affected by independent fadings by combining different replicas of the received signal. As a consequence, the probability that combined replicas present significant degradation is much smaller than the probability of a single replica presenting the same degradation. The number of replicas is known as diversity order. Further, the higher the diversity the higher is the system performance. Several diversity techniques such as frequency diversity, time diversity, and spatial diversities are well established in the literature [76].

In particular, spatial diversity can be obtained by using multiple antennas at the reciever, which is also known as antenna diversity. It is also possible to obtain spatial diversity with multiple antennas at the transmitter. In this thesis we consider spatial diversity by employing multiple antennas only at the receiver. In Chapter 3, a study of the distance between antennas is carried out in order to make the fading independent at the different antennas. However, in a multipath channel the distance between two antennas is $d \approx \lambda_c/2$, where λ_c denotes wavelength, which is calculated as $\lambda_c = c/f_c$ [76].

Antenna diversity is the most employed diversity technique because its implementation is easy and presents superior performance than frequency or time diversity techniques [82].

2.5.2 Spatial Multiplexing

Multiplexing consists in combining signals from different sources (users) into one composite signal that is transmitted over a shared medium. In the literature, there are several multiplexing techniques, such as time-division multiplexing, frequency-division multiplexing, code-division multiplexing, and space-division multiplexing [69].

In space-division multiplexing many symbols are simultaneously transmitted by the different transmit antennas, thus, the data rate is increased without any change in frequency band or transmission power.

MIMO systems are attractive because it provides high performance reliability and spectral efficiency through spatial diversity and spatial multiplexing, respectively. However, in [76] it is shown that there is a trade-off between diversity and multiplexing if linear detectors are employed.

Finally, MIMO systems, in addition to providing spatial diversity and multiplexing, can also be used to suppress the interference.

2.5.3 MIMO Channel

MIMO systems employ K transmit antennas and N receive antennas to simultaneously transmit/receive symbols in the same radio resource, as shown in Fig. 2.4. As K transmit antennas are employed, the spectral efficiency is increased K times. Thus, the received signal vector $\mathbf{y}$ is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2.17}$$

where $\mathbf{y} = [y_1, y_2, ..., y_N]^T$ has dimension $N \times 1$, $\mathbf{x} = [x_1, x_2, ..., x_K]^T$ is the transmitted signal vector of dimension $K \times 1$, whose elements belong to the M-QAM modulation with order M, $\mathbf{w} = [w_1, w_2, ..., w_N]^T$ is the additive complex white Gaussian noise (AWGN) vector of dimension $N \times 1$, whose entries consists of independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and variance σ_w^2 and $\mathbf{H}$ is the slow-flat Rayleigh fading MIMO channel matrix of dimension $N \times K$, whose entries consist of i.i.d complex Gaussian random variables with zero mean and variance equal to $\overline{\alpha^2}$. The MIMO channel matrix is of the form:

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,K} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N,1} & h_{N,2} & \cdots & h_{N,K} \end{bmatrix}, \tag{2.18}$$

where $h_{i,j}$ denotes the entry of the link between the ith receive antenna and jth transmit antenna, as shown in Fig. 2.4. For a better understanding of the MIMO channel, some matrix properties are presented in Appendix A.

2.5.4 Massive MIMO Channel

M-MIMO considers a system with a great number of antennas, where the number of receive antennas is much greater than the number of transmit antennas, that is $N \gg K$. Therefore, the product $\mathbf{H}^H\mathbf{H}$, where H^H is the conjugate transpose of H, converges asymptotically to $\overline{\alpha^2}N\mathbf{I}_K$, as N goes to infinity for a fixed K, that is [24]:

$$\lim_{N\to\infty} \left(\mathbf{H}^H\mathbf{H}\right) = \overline{\alpha^2}N\mathbf{I}_K. \tag{2.19}$$

The condition given by (2.19) establish that all eigenvalues of $\mathbf{H}^H\mathbf{H}$ becomes equal,

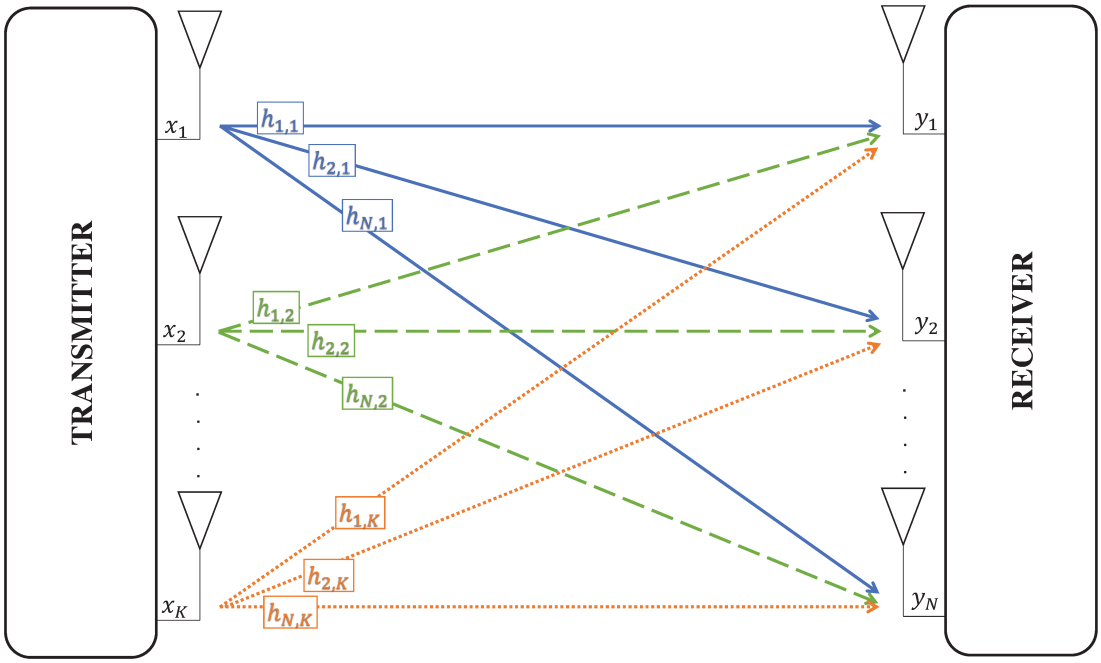


Figure 2.4: MIMO system with K transmit antennas and N receive antennas.

i.e., $\lambda_1 = \lambda_2 = \lambda_K$. Thus, the condition number, given by:

$$\kappa(\mathbf{H}^H\mathbf{H}) = \frac{\lambda_{\max}}{\lambda_{\min}}, \tag{2.20}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the greatest and the smallest nonzero eigenvalues of $\mathbf{H}^H\mathbf{H}$, respectively, is equal to one. This means that $\mathbf{H}^H\mathbf{H}$ is orthogonal [83].

The M-MIMO channel take advantage of two concepts known as channel favorable propagation regime [24] and channel hardening [84], which are explained below.

2.5.4.1 Favorable Propagation

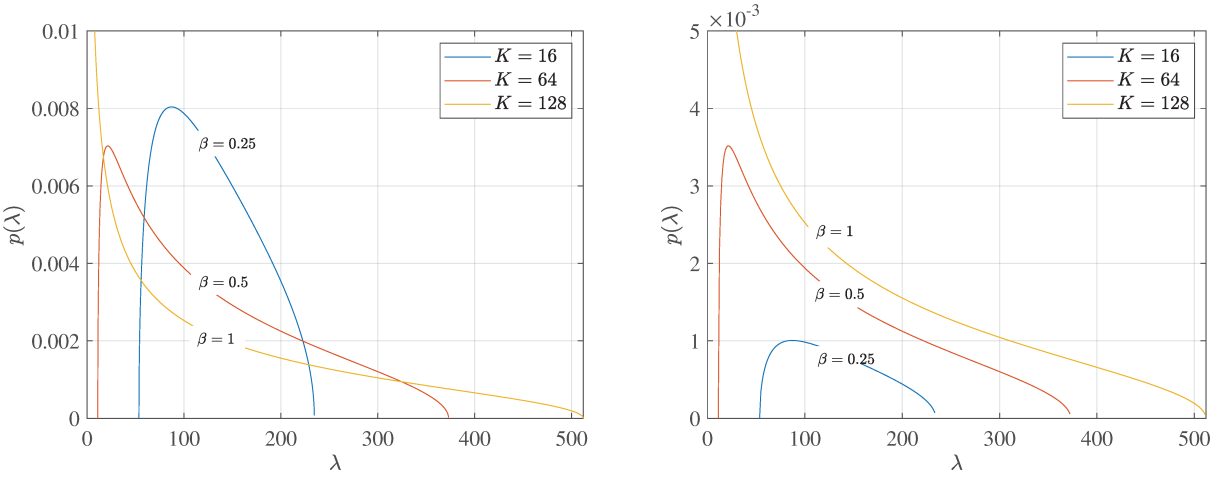
M-MIMO systems with large number of receive antennas allows the mutual orthogonality among channels, that is known as favorable propagation. Thus, the column channel vectors of $\mathbf{H}$ satisfy [22]:

$$\mathrm{E}\left\{\mathbf{h}_j^H\mathbf{h}_k\right\} = \begin{cases} 0 & \text{if} \quad j \neq k \\ \overline{\alpha^2}N & \text{if} \quad j = k \end{cases}. \tag{2.21}$$

This inner product is a measure of the interference reduction among users. For M-MIMO systems, it will be more detailed in Chapter 3.

2.5.4.2 Channel Hardening

In M-MIMO systems, the spatial diversity produces a phenomenon called channel hardening. Channel hardening means that a fading channel behaves as a AWGN channel,



(a) PDF of $\mathbf{H}^H\mathbf{H}$. (b) PDF of $\mathbf{H}\mathbf{H}^H$.

Figure 2.5: Empirical eigenvalues PDF for $N = 128$ and different K.

i.e, the fading impact on the communication is negligible [84].

Once the MIMO channel matrix $\mathbf{H}$, given by (2.18), becomes larger, that is, only N or both N and K increases with ratio $\beta = K/N$, then the eigenvalues distribution of $\mathbf{H}^H\mathbf{H}$ or $\mathbf{H}\mathbf{H}^H$ becomes less sensitive to the current distribution of its entries $h_{i,j}$ of the channel matrix $\mathbf{H}$.

This is a central result of random matrix theory based on the Marchenko-Pastur law, which establishes that for $N \gg 1$ and $\beta = K/N$, the empirical distribution of $\mathbf{H}^H\mathbf{H}$ converges to the probability density function given by [83] [1]:

$$p(\lambda) = \left(1 - \frac{1}{\beta}\right)_+ \delta(\lambda) + \frac{\sqrt{(\lambda - \tilde{a})_+ \left(\tilde{b} - \lambda\right)_+}}{2\pi N \beta \lambda}, \tag{2.22}$$

where $\delta(\cdot)$ is the Dirac delta function, $(x)_+ = \max(0, x)$, $\tilde{a} = Na$, $\tilde{b} = Nb$, $a = \left(1 - \sqrt{\beta}\right)^2$, and $b = \left(1 + \sqrt{\beta}\right)^2$, for $\tilde{a} \leqslant \lambda \leqslant \tilde{b}$ and $0 \leqslant \beta \leqslant 1$.

Analogously, the empirical distribution of $\mathbf{H}\mathbf{H}^H$ converges to the probability density function given by [83]:

$$p(\lambda) = (1 - \beta)_+ \delta(\lambda) + \frac{\sqrt{(\lambda - \tilde{a})_+ \left(\tilde{b} - \lambda\right)_+}}{2\pi N \lambda}. \tag{2.23}$$

In Fig. 2.5 the PDF of the eigenvalues of $\mathbf{H}^H\mathbf{H}$ and $\mathbf{H}\mathbf{H}^H$ for $N = 128$ and $K = 16$, 64, and 128 are presented. Observe that as β decreases, the eigenvalues ranges decreases.

[1] In [83] the entries of $\mathbf{H}$ are i.i.d. zero-mean complex Gaussian random variables with normalized variance $\overline{\alpha^2} = 1/N$.

Consequently, the eigenvalues becomes deterministic for $N \gg 1$ and $\beta \to 0$. This is equivalent to $N \gg K$, which is known as massive MIMO. This behavior produces the channel hardening phenomenon [12].

In [12] another channel hardening scenario is distinguished for large N and K, that is for $N = K$, which is known as symmetric MIMO channel. However, along this thesis only the massive MIMO channel scenario is considered.

2.5.4.3 Finite M-MIMO

The condition established in (2.19) produces favorable propagation and channel hardening, once the number of receive antennas N tends to infinite. However, the employment of an infinite number of receive antennas is impractical. Thus, a natural question arises: How many antennas do we need for M-MIMO? [25]. The answer is not straightforward because it depends on the system configuration required for a desired performance. However, an approach based on the condition number given by (2.20) can give us a practical answer. Thus, as the condition number of $\mathbf{H}^H\mathbf{H}$ goes to 1, the channel becomes asymptotically orthogonal. On the other hand, a large condition number, means that the channel is highly non-orthogonal or ill-conditioned.

Another reference for the channel orthogonality of M-MIMO systems is the Frobenius norm of the orthogonality error matrix, $\mathbf{E}_o = \mathbf{H}^H\mathbf{H} - \overline{\alpha^2}N\mathbf{I}_K$, given by:

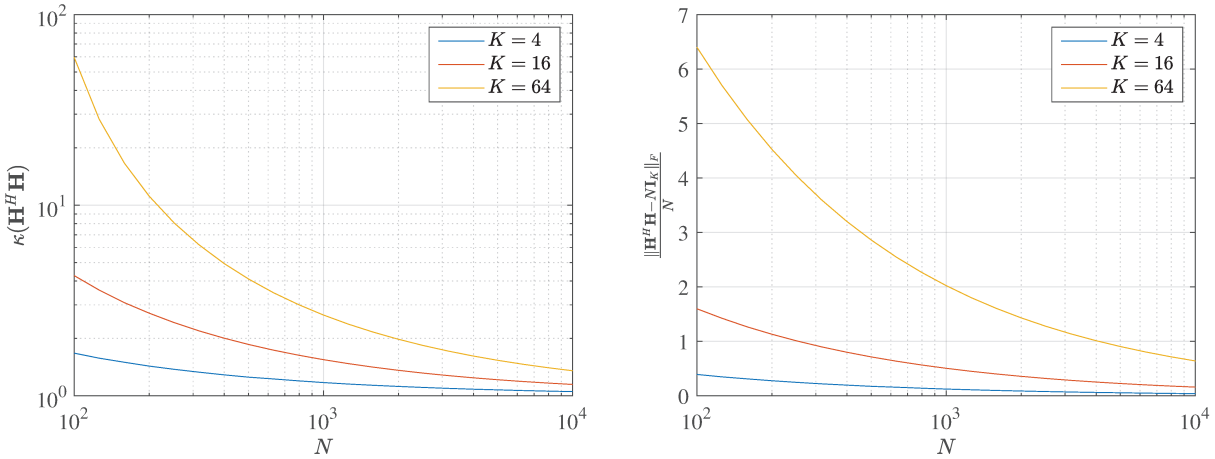
$$\|\mathbf{E}_o\|_F = \sqrt{\mathrm{Tr}(\mathbf{E}_o^H\mathbf{E}_o)}, \tag{2.24}$$

where $\mathrm{Tr}(\mathbf{E}_o^H\mathbf{E}_o)$ denotes the matrix trace.

In Fig. 2.6a and Fig. 2.6b, the condition number of $\mathbf{H}^H\mathbf{H}$ and the Frobenius norm of the orthogonality matrix error are obtained using Monte Carlo simulation. Both metrics are evaluated and averaged over 10^4 channel simulations.

Fig. 2.6a shows the condition number of $\mathbf{H}^H\mathbf{H}$ as a function of the number of receive antennas for $K = 4$, 16, and 64 transmit antennas. Notice that as N increases the condition number tends to one, that is $\kappa(\mathbf{H}^H\mathbf{H}) \approx 1$, which is more evident when $N \gg K$. For $N = 100$ the condition number is closer to one for $K = 4$ than for $K = 64$.

On the other hand, Fig. 2.6b shows the normalized Frobenius norm of the orthogonality error matrix $\mathbf{E}_o$ as a function of the number of receive antennas for $K = 4$, 16, and 64 transmit antennas $\overline{\alpha^2} = 1$. Notice that the normalized Frobenius norm of the orthogonality error matrix tends to zero as N increases, which shows that for $N \gg K$, the channel matrix $\mathbf{H}$ can be considered well conditioned (quasi-orthogonal) and behaves as deterministic. However, the condition number and the Frobenius norm show that the channel orthogonality is improved, but they do not show the loss in performance due to the channel non-orthogonality. Therefore, performance analysis such as the ergodic capacity or bit error rate are required.



(a) Condition Number.

(b) Normalized Frobenius norm of the orthogonality error matrix.

Figure 2.6: Condition number and Frobenius norm of finite M-MIMO.

Finally, practical finite M-MIMO systems are limited by the physical space used in the antenna array deployment, which is studied in Chapter 3. Furthermore, the performance of finite M-MIMO is limited by MAI and CEEI, which are studied in Chapter 4 and 5. Impairments such as pilot contamination, spatial correlation, hardware restrictions and other that limits the M-MIMO systems performance are not part of this thesis.

2.6 MIMO Linear Detectors

Data detection recovers the transmitted symbol at the receiver. The detection techniques vary from SISO to MIMO systems. MIMO systems due to the spatial multiplexing is suitable for multiuser detection. However, MIMO detectors requires high complexity algorithms for its effectiveness.

Given the received MIMO signal shown in (2.17), many detection techniques can be used at the receiver in order to detect the transmit signal under CSI knowledge. Consequently, efficient detection schemes in terms of performance and complexity should be considered. For uncoded systems, the spatial diversity order can be used as a performance metric for several MIMO detection methods, where the diversity gain is equal to the number of receive antennas.

From a performance perspective, optimal detection is always preferred once it achieves the full diversity gain given by N. However, its complexity grows exponentially with the number of transmit antennas, which becomes impractical even for low values of K. The maximum likelihood (ML) detector is optimal with complexity $\mathcal{O}\left(M^K\right)$, where M is the modulation order. On the other hand, the sphere detector (SD) has the same performance of ML, but smaller complexity given by $\mathcal{O}\left(M^{\nu K}\right)$ where $0 < \nu < 1$ [82].

In order to reduce the computational complexity, sub-optimal linear detectors can be an interesting approach. They consider the signals from the other antennas as MAI and detect the transmitted signal vector $\mathbf{x}$ by multiplying the received signal vector $\mathbf{y}$, given by (2.17), by a compensation matrix $\mathbf{A}$ as:

$$\widehat{\mathbf{x}} = \mathbf{A}\mathbf{y}. \tag{2.25}$$

Furthermore, the kth transmitted signal can be detected from:

$$\widehat{x}_k = \mathbf{a}_k\mathbf{y}, \tag{2.26}$$

where $\mathbf{a}_k$ is the kth row of $\mathbf{A}$.

Linear detectors cannot achieve reasonably good performance if the channel matrix $\mathbf{H}$ is ill-conditioned or non-orthogonal. Thus, they do not achieve the full diversity as the optimum detector. On the other hand, when the channel matrix $\mathbf{H}$ is well-conditioned or quasi-orthogonal, that is the case of M-MIMO, these detectors can achieve near optimum performance [24].

The most common linear detectors are the maximal-ratio-combining (MRC), zero-forcing (ZF) and minimum-mean-square-error (MMSE) detectors, and they are presented below under the assumption that $\widehat{\mathbf{H}}$ is the estimation of $\mathbf{H}$. Furthermore, their performance is presented in Chapter 5.

2.6.1 Maximal-Ratio-Combining

The MRC detector is the less complex linear detector which can be implemented by using (2.25) with:

$$\mathbf{A} = \widehat{\mathbf{H}}^H. \tag{2.27}$$

Therefore, by using (2.27) in (2.17), the detected symbol vector is given by:

$$\widehat{\mathbf{x}} = \widehat{\mathbf{H}}^H\mathbf{y}, \tag{2.28}$$

and the detected symbol transmitted by the kth antenna can be rewritten as:

$$\widehat{x}_k = \widehat{\mathbf{h}}_k^H\mathbf{y}, \tag{2.29}$$

where $\widehat{\mathbf{h}}_k^H$ is the kth row vector of $\widehat{\mathbf{H}}^H$.

Notice that the transpose-conjugate is an operation that requires complexity order of $\mathcal{O}(K)$, which is attractive. However, MRC considers MAI as pure noise. Thus, the MRC detector works properly only for the M-MIMO case, otherwise, the performance is severely degraded by MAI.

2.6.2 Zero-Forcing

The ZF detector eliminates the MAI at the cost of more complexity and noise enhancement. The ZF compensation matrix is given by:

$$\mathbf{A} = \widehat{\mathbf{H}}^{+} = \left(\widehat{\mathbf{H}}^{H}\widehat{\mathbf{H}}\right)^{-1}\widehat{\mathbf{H}}^{H}, \tag{2.30}$$

where $\widehat{\mathbf{H}}^{+}$ is the left-inverse or Moore-Penrose pseudo-inverse of the matrix $\widehat{\mathbf{H}}$ as shown in Appendix (A.4). Therefore, by using (2.30) in (2.17), the detected symbol vector is given by:

$$\widehat{\mathbf{x}} = \widehat{\mathbf{H}}^{+}\mathbf{y}, \tag{2.31}$$

and the kth detected symbol is given by:

$$\widehat{x}_k = \widehat{\mathbf{h}}_k^{+}\mathbf{y}, \tag{2.32}$$

where $\widehat{\mathbf{h}}_k^{+}$ is the kth row vector of $\widehat{\mathbf{H}}^{+}$. Observe that MAI is eliminated according to:

$$\widehat{\mathbf{h}}_k^{+}\widehat{\mathbf{h}}_j = \begin{cases} 1 & \text{if} \quad k = j \\ 0 & \text{if} \quad k \neq j \end{cases}. \tag{2.33}$$

The ZF detector complexity is in computing $\widehat{\mathbf{H}}^{+}$, which has cubic complexity, given by $\mathcal{O}(K^3)$. However, in [35], several inversion algorithms have been shown which at a performance cost present complexity of $\mathcal{O}(\nu K^2)$ or $\mathcal{O}(\nu K)$, where ν is a factor that depends on the inversion algorithm. On the other hand, the ZF performance achieves, on average, a diversity order of $N - K + 1$ [68]. For M-MIMO the diversity order is nearly optimal.

2.6.3 Minimum-Mean-Square-Error

The MMSE detector mitigates the noise enhancement of ZF detector by minimizing the mean-square-error (MSE) between the transmitted symbol vector $\mathbf{x}$ and the compensated received symbol vector $\mathbf{Ay}$, which consist in solving the minimization problem given by:

$$\min_{\mathbf{A}} \mathrm{E}\left[\|\mathbf{x} - \mathbf{Ay}\|^2\right]. \tag{2.34}$$

Thus, the solution of (2.34) is the compensation matrix given by:

$$\mathbf{A} = \mathbf{C}_{\mathbf{xy}}\mathbf{C}_{\mathbf{yy}}^{-1}, \tag{2.35}$$

where $\mathbf{C}_{\mathbf{xy}} = \overline{|x|^2}\ \widehat{\mathbf{H}}^{H}$ and $\mathbf{C}_{\mathbf{yy}} = \overline{|x|^2}\ \widehat{\mathbf{H}}\widehat{\mathbf{H}}^{H} + \sigma_w^2\mathbf{I}_N$ are covariance matrices, $\overline{|x|^2}$ is the constellation mean power and σ_w^2 is the noise variance. Notice that (2.35) requires the

inversion of a $N \times N$ matrix. However, by using the matrix inversion lemma given by (A.13) in (2.35), the compensation complexity can be reduced by the matrix inversion of $\mathbf{A} = \overline{|x|^2}\ \left(\overline{|x|^2}\ \widehat{\mathbf{H}}^H\widehat{\mathbf{H}} + \sigma_w^2\mathbf{I}_K\right)^{-1}\widehat{\mathbf{H}}^H$, which is $K \times K$. Therefore, the detected signal vector is given by:

$$\widehat{\mathbf{x}} = \overline{|x|^2}\ \left(\overline{|x|^2}\ \widehat{\mathbf{H}}^H\widehat{\mathbf{H}} + \sigma_w^2\mathbf{I}_K\right)^{-1}\widehat{\mathbf{H}}^H\mathbf{y}, \tag{2.36}$$

and the kth detected symbol is given by (2.26).

As for ZF detector, the complexity of the MMSE detector is cubic with the number of transmit antennas, due to the matrix inversion involved in (2.36). However, additional knowledge of the constellation mean power and noise variance is required.

MMSE detector has a diversity order of $N-K+1$. Besides, it has similar performance to the ZF detector in M-MIMO [47]. However, the benefits of the MMSE detector are well evident in MIMO systems with a few numbers of antennas [85].

2.7 Channel State Information

The CSI is the knowledge of channel matrix $\mathbf{H}$ at the receiver. Channel estimation is performed before the symbol detection is carried out [82]. For simplicity, perfect CSI is widely used in communication systems analysis. In this ideal scenario $\widehat{\mathbf{H}} = \mathbf{H}$.

In practice, CSI is obtained by transmitting pilot symbols which are known at the receiver. Thus, the received samples $\mathbf{y}_p$, of dimension $L_p \times 1$, used in the channel estimation can be modeled as:

$$\mathbf{y}_p = \mathbf{X}_p\mathbf{h} + \mathbf{w}, \tag{2.37}$$

where $\mathbf{X}_p$ is the pilot symbols matrix of dimension $L_p \times L_p$, $\mathbf{h}$ is the unknown channel vector of dimension $L_p \times 1$, $\mathbf{w}$ is the noise vector of dimension $L_p \times 1$ and L_p is the number of pilot symbols.

There are two techniques to estimate $\mathbf{h}$: the ML estimator and the MMSE [86]. Conventionally, the ML estimator is used only if the noise variance is known. On the other hand, if the noise variance and the a-priori distribution of $\mathbf{h}$ are known, the MMSE estimator can be used as shown below.

2.7.1 Minimum-Mean-Square-Error Estimator

The MMSE estimator, which is also called Bayesian mean-square-error (MSE) estimator, exploits the knowledge of the channel coefficients distribution in order to obtain an estimator that is better than the ML [86]. Thus, the Bayesian MSE estimator is given by the mean of the a-posteriori PDF $p(\mathbf{h}|\mathbf{y}_p)$ given by:

$$\hat{\mathbf{h}} = \mathrm{E}\left[\mathbf{h}|\mathbf{y}_p\right]. \tag{2.38}$$

The result of (2.38) is equivalent to minimize the MSE, that is:

$$\min_{\mathbf{h}} \mathrm{E}\left\| \left(\mathbf{h} - \hat{\mathbf{h}}\right)^2 \right\|, \tag{2.39}$$

where $\hat{\mathbf{h}}$ is the estimated channel vector.

By considering the observed data in (2.37) as a linear model, where $\mathbf{h}$ is a random vector with Gaussian PDF $\mathcal{CN}(0, \mathbf{C_{hh}})$, and $\mathbf{w}$ is the noise vector with Gaussian PDF $\mathcal{CN}(0, \mathbf{C_{ww}})$ that is independent of $\mathbf{h}$. Then, $\mathbf{y}$ has also a complex Gaussian distribution with PDF $\mathcal{CN}(0, \mathbf{C}_{\mathbf{y}_p\mathbf{y}_p})$ and therefore, $\mathbf{h}$ and $\mathbf{y}_p$ are jointly Gaussian. As a consequence the estimator in (2.38) can be rewritten as:

$$\hat{\mathbf{h}} = \mathbf{C}_{\mathbf{h}\mathbf{y}_p}\mathbf{C}^{-1}_{\mathbf{y}_p\mathbf{y}_p}\mathbf{y}_p, \tag{2.40}$$

where $\mathbf{C}_{\mathbf{h}\mathbf{y}_p} = \mathbf{C_{hh}}\mathbf{X}_p^H$, and $\mathbf{C}_{\mathbf{y}_p\mathbf{y}_p} = \mathbf{X}_p\mathbf{C_{hh}}\mathbf{X}_p^H + \sigma_w^2\mathbf{I}_{L_p}$ are covariance matrices. Furthermore, the MSE matrix is given by:

$$\mathbf{M_h} = \mathbf{C_{hh}} - \mathbf{C}_{\mathbf{h}\mathbf{y}_p}\mathbf{C}^{-1}_{\mathbf{y}_p\mathbf{y}_p}\mathbf{C}_{\mathbf{y}_p\mathbf{h}}, \tag{2.41}$$

which is obtained by applying the orthogonal principle which establishes that the optimal estimator is obtained when the error is orthogonal to the data sample.

Additionally, if the observed data in (2.37) is a wide-sense stationary Gaussian process and the MMSE estimator is used, this problem is known as Wiener filtering, that can be cast the functions of filtering, smoothing and prediction. The smoothing filter is given by (2.40) and the prediction filter is given by [86]:

$$\hat{h}(\ell) = \mathbf{C}_{h(\ell)\mathbf{y}_p}\mathbf{C}^{-1}_{\mathbf{y}_p\mathbf{y}_p}\mathbf{y}_p, \tag{2.42}$$

for all $\ell \geqslant L_p + 1$.

Finally, the CSI quality depends also on the training scheme, which consists on the usage of pilot and data symbols in a block with a duration T_B. Commonly, pilot and data symbols are multiplexed in a block. This technique is named as multiplexed pilot estimation (MPE) [87]. If pilot and data symbols are superimposed this technique is named superimposed pilot estimation (SPE). Furthermore, a hybrid pilot estimation (HPE) technique can be built by using both MPE and SPE [38]. These estimation techniques are detailed in Chapter 4.

CHAPTER 3

Inter-User Interference Reduction in M-MIMO

3.1 Introduction

The key idea in M-MIMO systems is the favorable propagation introduced in Section 2.5.4.1. Favorable propagation is defined as the mutual orthogonality among the channel column vectors of the channel matrix [21]. However, favorable propagation depends on the channel model [22, 88, 89]. In the literature, there are several channel models for M-MIMO systems [15, 90]. Although the benefits of M-MIMO systems are undoubtedly very attractive, its practical implementation, like the deployment of antenna arrays, is equally challenging [88].

The physical size of M-MIMO arrays is the fundamental problem of practical deployment. Once M-MIMO has been initially conceived for ordinary rich scattering in sub-6 GHz frequency channel [13], it is ideal for frequency bands in the range of 30-300 GHz, known as millimeter wave (mmWave). Despite conceptual similarities, the way M-MIMO can be exploited in these bands is radically different due to their specific propagation behaviors [91, 92]. Furthermore, for small wavelengths, the antenna elements can be packed in a finite volume, due to the highly directional nature of propagation at mmWave, line-of-sight (LOS) propagation dominates.

On the other hand, it is shown in [20] that the independent and identically distributed (i.i.d.) Rayleigh fading channel is a good model for non-line-of-sight (NLOS) channels, once it occurs in special cases with a rich scattering environment where multipaths are uniformly distributed which is a propagation characteristic of sub-6 GHz. However, the array size increases as the number of antennas grows [93]. For this purpose, recent works have also investigated the possibility of increasing the number of array elements in fixed physical spaces. However, this approach introduces correlation between antennas which limits the favorable propagation [94, 95]. Therefore, other properties of M-MIMO must be exploited [96].

This chapter addresses the favorable propagation issues through the interference reduction factor (IRF) presented in [19]. The IRF is the interference reduction of an array in relation to the single antennas case. It also can be used to evaluate the antenna correlation issues presented in fixed physical spaces. For this purpose, the uniform random line of sight (UR-LOS) and the Rayleigh fading channel models are used.

The channel models are introduced in Section 3.2, the IRF is derived in Section 3.3 and the numerical results and discussions are presented in Section 3.4.

3.2 Channel Model

The entries of the MIMO channel matrix defined in (2.18) depend on the channel impulse response given by (2.11). According to [90] and [46], the channel impulse response can be modeled by a geometric-based stochastic model (GBSM) or correlation-based stochastic models (CBSM). GBSM is used to evaluate the performance of practical wireless communication systems using simulation. On the other hand, CBSM is used as a theoretical model to evaluate the performance of M-MIMO systems based on the correlation of the channel impulse response [2].

In the context of theoretical analysis, this thesis approaches the UR-LOS channel model and the Rayleigh fading channel model, introduced in Section 2.4.2.1 and 2.4.2.2, respectively.

Actually, the $N \times K$ MIMO channel matrix given by (2.18), for the ℓth time interval can be rewritten as:

$$\mathbf{H}(\ell) = [\mathbf{h}_1(\ell)\ \mathbf{h}_2(\ell)\ \cdots \mathbf{h}_K(\ell)]\,, \tag{3.1}$$

where $\mathbf{h}_k(\ell)$ is the kth UT channel vector of dimension $N \times 1$.

3.2.1 UR-LOS Channel

UR-LOS channel model assumes that at each receive antenna arrives one LOS path as in (2.12), where the fading amplitude α is unitary, and the phase is a function of the UT position and the antenna array pattern.

In a three dimensional (3-D) scenario, the azimuthal angle of arrival (AoA) and the elevation angle of arrival (EoA) of the LOS path at each antenna depends on the position of each UT defined by ϕ_k and θ_k $\forall k \leqslant K$, which are modeled as independent random variables with probability density function $p(\phi_k) = \frac{1}{2\pi}$ distributed between $[0, 2\pi)$ and $p(\theta_k) = \sin\theta_k$ distributed over $[\pi/2, \pi)$, respectively. Furthermore, the 3-D scenario could

[2] The GBSM can model several channel parameters that are not included in the CBSM due to its increased complexity.

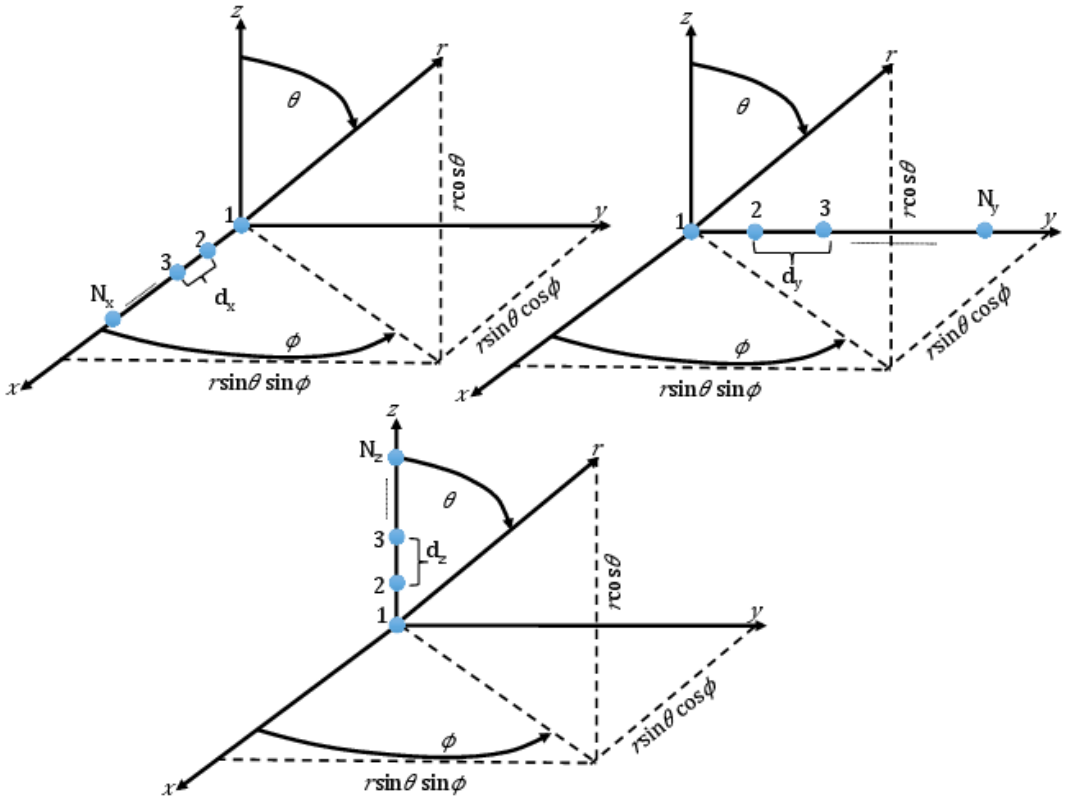


Figure 3.1: Uniform linear array along X-axis, Y-axis and Z-axis.

be reduced to a two dimensional (2-D) if $\theta_k = \pi/2$, which is called horizon or far-field scenario.

On the other hand, the antenna array radiation pattern, that is the radiated power distribution, depends on the array geometry and on each antenna radiation pattern. The most common array geometries are the linear, planar, circular and cylindrical. Furthermore, there are several antenna types, such as, the half-wave dipole antenna that is the very popular and the short dipole used when space is an issue [97].

For simplicity, no time-correlation is assumed in this channel model and consequently Doppler effects are not presented.

3.2.1.1 Uniform Linear Array

The uniform linear array (ULA) consists of equally spaced antenna elements placed in a straight line. Fig. 3.1 shows three uniform linear arrays using short dipoles with vertical polarization deployed along the X-axis, Y-axis and Z-axis, respectively. When the radio signal travels one wavelength λ_c, then the carrier phase changes 2π radians, therefore the phase difference between signals arriving at two adjacent antennas in each axis can be

written as:

$$\Phi_{k,x} = 2\pi\Delta_x \cos\phi_k \sin\theta_k, \tag{3.2}$$

$$\Phi_{k,y} = 2\pi\Delta_y \sin\phi_k \sin\theta_k, \tag{3.3}$$

$$\Phi_{k,z} = 2\pi\Delta_z \cos\theta_k, \tag{3.4}$$

where $\Delta_x = \frac{d_x}{\lambda_c} = \frac{\mathcal{L}_x}{\lambda_c(N_x-1)}$, $\Delta_y = \frac{d_y}{\lambda_c} = \frac{\mathcal{L}_y}{\lambda_c(N_y-1)}$ and $\Delta_z = \frac{d_z}{\lambda_c} = \frac{\mathcal{L}_z}{\lambda_c(N_z-1)}$, are the normalized separation between antennas, $\lambda_c = c/f_c$ is the carrier wavelength, $c \simeq 3 \times 10^8$ m/s is the light speed, f_c is the carrier frequency, d_x, d_y and d_z are the distance between antennas, $\mathcal{L}_x$, $\mathcal{L}_y$ and $\mathcal{L}_z$ are the array edge length and N_x, N_y and N_z are the number of antennas along the X-axis, Y-axis and Z-axis, respectively.

Notice that a steering vector can be formed by using the phase differences of the kth UT received signal at the antennas of the array [3] [72]. The first antenna is used as a phase of reference. Thus, for the ULA on each axis, the steering or channel vector for the kth UT is given by:

$$\mathbf{h}_{k,x} = \left[1 \quad \exp\left[-\mathbb{i}\Phi_{k,x}\right] \quad \cdots \quad \exp\left[-\mathbb{i}\left(N_x-1\right)\Phi_{k,x}\right]\right]^T, \tag{3.5}$$

$$\mathbf{h}_{k,y} = \left[1 \quad \exp\left[-\mathbb{i}\Phi_{k,y}\right] \quad \cdots \quad \exp\left[-\mathbb{i}\left(N_y-1\right)\Phi_{k,y}\right]\right]^T, \tag{3.6}$$

$$\mathbf{h}_{k,z} = \left[1 \quad \exp\left[-\mathbb{i}\Phi_{k,z}\right] \quad \cdots \quad \exp\left[-\mathbb{i}\left(N_z-1\right)\Phi_{k,z}\right]\right]^T. \tag{3.7}$$

Finally, in the far-field, the elevation angle is equal to $\theta_k = \pi/2$. Thus, the steering vector for the ULA on the Z-axis can be reduced to a single antenna element and the steering vectors for the ULA on X and Y axis do not depend on the elevation angle.

3.2.1.2 Uniform Planar Array

The uniform planar array (UPA) is composed by equally spaced antenna elements placed in a plane. A planar array could be deployed along the XY-plane, XZ-plane or YZ-plane. By using the ULA steering vectors for the ULAs in each axis, the UPA in each plane can be obtained as follows:

$$\mathbf{h}_{k,xy} = \mathbf{h}_{k,x} \otimes \mathbf{h}_{k,y}, \tag{3.8}$$

$$\mathbf{h}_{k,xz} = \mathbf{h}_{k,x} \otimes \mathbf{h}_{k,z}, \tag{3.9}$$

$$\mathbf{h}_{k,yz} = \mathbf{h}_{k,y} \otimes \mathbf{h}_{k,z}, \tag{3.10}$$

[3] The UR-LOS channel vector can also be named steering vector.

where $\otimes$ denotes the Kronecker product. For example, the steering vector $\mathbf{h}_{k,xy}$ of dimension $N_xN_y \times 1$, that represents the UPA on the XY-plane, can be rewritten as:

$$\mathbf{h}_{k,xy} = \begin{bmatrix} 1 \\ \exp\{-\mathbb{i}\Phi_{k,x}\} \\ \vdots \\ \exp\{-\mathbb{i}\left[N_x - 1\right]\Phi_{k,x}\} \\ \exp\{-\mathbb{i}\Phi_{k,y}\} \\ \exp\{-\mathbb{i}\left[\Phi_{k,x} + \Phi_{k,y}\right]\} \\ \vdots \\ \exp\{-\mathbb{i}\left[\left(N_x - 1\right)\Phi_{k,x} + \Phi_{k,y}\right]\} \\ \vdots \\ \exp\{-\mathbb{i}\left[\left(N_x - 1\right)\Phi_{k,x} + \left(N_y - 1\right)\Phi_{k,y}\right]\} \end{bmatrix}. \tag{3.11}$$

Additionally, for $N_x = N_y$ the UPA geometry is square and for $N_x \neq N_y$ is rectangular. Notice that the total number of antennas in the array is $N = N_xN_y$. A similar procedure can be performed in order to represent the UPA on the other planes.

Finally, in the far-field, the UPAs steering vectors $\mathbf{h}_{k,xz}$ and $\mathbf{h}_{k,yz}$, are reduced to the ULAs with steering vectors given by $\mathbf{h}_{k,x}$ and $\mathbf{h}_{k,y}$, respectively.

3.2.2 Rayleigh Fading Channel Model

Rayleigh fading channel occurs on an idealized rich scattering environment with a large number of multipaths uniformly distributed that rise a flat and slow fading. Thus, the entries of the channel matrix given by (3.1), for the ℓth time interval, can be written as:

$$h_{i,j}(\ell) = \alpha_{i,j}\exp\{-\mathbb{i}(2\pi f_{d_{i,j}}\ell T_s + \phi_{i,j})\}, \tag{3.12}$$

for $1 \leqslant i \leqslant N$ and $1 \leqslant j \leqslant K$. Therefore, the channel is modeled by a zero-mean complex wide-sense stationary Gaussian process with auto-correlation function $\mathrm{R}(\ell) = \mathrm{E}\{h_{i,j}(\ell)h^*_{i,j}(\ell - m)\}$, where $\alpha_{i,j}$ is the Rayleigh fading amplitude with mean power $\overline{\alpha^2}$ and PDF given by (2.13), $\phi_{i,j}$ is the uniformly distributed phase over the interval $[0, 2\pi)$, $f_{d_{i,j}} = \frac{v_j}{c} f_c \cos\varphi_{i,j}$ is the Doppler shift, v_j is the jth UT speed, and $\varphi_{i,j}$ is the angular position, uniformly distributed over the interval $[0, \pi)$. In a mobile communication environment that considers the Jakes' model, the time auto-correlation function is given by (2.16).

Once in (3.12) only the azimuthal angle is considered, the Rayleigh fading channel should be employed only on 2-D propagation scenarios. For 3-D scenarios, other channel models should be used.

Finally, notice that in the Rayleigh channel model there are multipaths, while in the UR-LOS channel model there is only one path.

3.3 Interference Reduction Factor

An antenna array can pinpoint the main beam of the radiation pattern in the direction of the desired UT [97]. Furthermore, as the number of antennas increases, the main beam becomes narrower and directive by diminishing the secondary lobes, which reduces the interference from other UTs.

The radiation pattern and the IRF are related definitions, both depend on the inner product between the channel steering vector fixed on a desired direction and channel steering vector from any direction. In particular, the IRF is defined by the mean value of the squared radiation pattern. These aspects are presented in the following sections.

3.3.1 IRF in UR-LOS Channels

The normalized radiation pattern of an antenna array in the direction of the kth UT for UR-LOS channels in 3-D scenarios is given by [97]:

$$\mathcal{R}_{3-\mathrm{D}} = \sin(\theta_k)\sin(\theta_j)\frac{\left|\mathbf{h}_j^H\mathbf{h}_k\right|}{N}, \tag{3.13}$$

where the steering vector $\mathbf{h}_k$ is a function of θ_k and ϕ_k, that is the desired direction and $\mathbf{h}_j$ is a function of θ_j and ϕ_j directions, N normalizes the radiation pattern and $\sin(\theta_j)$ and $\sin(\theta_k)$ are introduced by the use of vertical polarized short dipoles. Moreover, depending on the array geometry, the steering vectors $\mathbf{h}_k$ and $\mathbf{h}_j$ are given by (3.5), (3.6), and (3.7) for ULAs and (3.8), (3.9) and (3.10) for UPAs, respectively. For a simple notation of (3.13) the steering vectors axis or plane is not presented.

Once ϕ_k and ϕ_j may assume any angular position between $[0, 2\pi)$ and θ_k and θ_j any angular position between $[\pi/2, \pi)$, the array IRF can be obtained by averaging the squared normalized radiation pattern given by (3.13) as:

$$\begin{aligned}\Upsilon_{3-\mathrm{D}} &= \mathrm{E}\left\{\mathcal{R}_{3-\mathrm{D}}^2\right\}\\ \Upsilon_{3-\mathrm{D}} &= \frac{1}{N^2}\int_0^{2\pi}\int_{\pi/2}^{\pi}\int_0^{2\pi}\int_{\pi/2}^{\pi}\sin^2(\theta_j)\sin^2(\theta_k)\left|\mathbf{h}_j^H\mathbf{h}_k\right|^2 p(\phi_k)\,p(\theta_k)\,p(\phi_j)\,p(\theta_k)\,d\theta_k d\phi_k d\theta_j d\phi_j,\end{aligned} \tag{3.14}$$

where $p(\phi_j)$ and $p(\phi_k)$ are the azimuthal angle PDFs and $p(\theta_j)$ and $p(\theta_k)$ are the elevation angle PDFs.

In a 2-D scenario, i.e. fixing $\theta_k = \theta_j = \pi/2$, the factors introduced by the short dipoles in (3.13) are unitary. Thus, the radiation pattern for 2-D scenarios can be rewritten as

[64]:

$$\mathcal{R} = \frac{\left|\mathbf{h}_j^H \mathbf{h}_k\right|}{N}, \tag{3.15}$$

and the IRF is given by:

$$\begin{aligned} \Upsilon &= \mathrm{E}\left\{\mathcal{R}^2\right\} \\ \Upsilon &= \frac{1}{N^2} \int_0^{2\pi} \int_0^{2\pi} \left|\mathbf{h}_j^H \mathbf{h}_k\right|^2 p\left(\phi_k\right) p\left(\phi_j\right) d\phi_k d\phi_j. \end{aligned} \tag{3.16}$$

3.3.1.1 IRF for ULAs

The IRF expressions given by (3.14) and (3.16) depend on the inner product $\left|\mathbf{h}_k^H \mathbf{h}_\mathrm{j}\right|$ of the ULA steering vectors given by (3.5), (3.6) and (3.7). Thus, after some algebraic manipulation, the inner products for ULAs on each axis can be written as:

$$\left|\mathbf{h}_{j,x}^H \mathbf{h}_{k,x}\right| = \frac{\sin\left(N_x \frac{\psi_x}{2}\right)}{\sin\left(\frac{\psi_x}{2}\right)}, \tag{3.17}$$

$$\left|\mathbf{h}_{j,y}^H \mathbf{h}_{k,y}\right| = \frac{\sin\left(N_y \frac{\psi_y}{2}\right)}{\sin\left(\frac{\psi_y}{2}\right)}, \tag{3.18}$$

$$\left|\mathbf{h}_{j,z}^H \mathbf{h}_{k,z}\right| = \frac{\sin\left(N_z \frac{\psi_z}{2}\right)}{\sin\left(\frac{\psi_z}{2}\right)}, \tag{3.19}$$

where $\psi_x = 2\pi\Delta_x\left(\cos\phi_j \sin\theta_j - \cos\phi_k \sin\theta_k\right)$, $\psi_y = 2\pi\Delta_y\left(\sin\phi_j \sin\theta_j - \sin\phi_k \sin\theta_k\right)$ and $\psi_z = 2\pi\Delta_z\left(\cos\theta_j - \cos\theta_k\right)$.

A closed-form expression of the IRF in 3-D scenarios can be obtained after integrating (3.14). Thus, by using (3.17) and (3.19) in (3.14), the IRF for the ULA on X-axis and Z-axis are respectively given by:

$$\begin{aligned} \Upsilon_{3-\mathrm{D},x} = \frac{4}{9N_x} + \frac{2}{N_x^2} \sum_{m=1}^{N_x-1} \left(N_x - m\right) \frac{1}{\left(2\pi\Delta_x m\right)^6} \Big\{ &\left(2\pi\Delta_x m\right) \cos\left(2\pi\Delta_x m\right) \\ &+ \left[\left(2\pi\Delta_x m\right)^2 - 1\right] \sin\left(2\pi\Delta_x m\right) \Big\}^2, \end{aligned} \tag{3.20}$$

$$\begin{aligned} \Upsilon_{3-\mathrm{D},z} = \frac{4}{9N_z} + \frac{2}{N_z^2} \sum_{m=1}^{N_z-1} \left(N_z - m\right) \frac{1}{\left(2\pi\Delta_y m\right)^6} \Big\{ &8 + 8\left(2\pi\Delta_z m\right)^2 + \left(2\pi\Delta_z m\right)^4 \\ &- \left[8 + 4\left(2\pi\Delta_z m\right)^2\right] \left[\cos\left(2\pi\Delta_z m\right) + \left(2\pi\Delta_z m\right) \sin\left(2\pi\Delta_z m\right)\right] \Big\}. \end{aligned} \tag{3.21}$$

The IRF for the ULA on Y-axis, that is $\Upsilon_{3-\mathrm{D},y}$, is equivalent to the IRF for the ULA on X-axis, given by (3.20), with $\Delta_x = \Delta_y$ and $N_x = N_y$.

In 2-D scenarios, a closed-form expression for the IRF for the ULA on X-axis can be obtained by integrating (3.16), considering the steering vectors inner product in (3.17). Thus, the IRF is given by:

$$\Upsilon_x = \frac{1}{N_x} + \frac{2}{N_x^2} \sum_{m=1}^{N_x-1} (N_x - m)\, J_0^2 (2\pi\Delta_x m)\,. \tag{3.22}$$

As for the 3-D scenario, the IRF for the ULA on Y-axis is equivalent to the IRF on X-axis given by (3.22), with $\Delta_x = \Delta_y$ and $N_x = N_y$. Finally, there is no IRF for the ULA on Z-axis, once the elevation angle is not considered in 2-D scenario.

3.3.1.2 IRF for UPAs

Applying the same criteria used for ULAs, the inner product $\left|\mathbf{h}_k^H \mathbf{h}_\mathrm{j}\right|$ for the UPA on planes XY, XZ and YZ, in terms of the steering vectors given by (3.8), (3.9) and (3.10), can be written as:

$$\left|\mathbf{h}_{j,xy}^H \mathbf{h}_{k,xy}\right| = \frac{\sin\left(N_x \frac{\psi_x}{2}\right)}{\sin\left(\frac{\psi_x}{2}\right)} \frac{\sin\left(N_y \frac{\psi_y}{2}\right)}{\sin\left(\frac{\psi_y}{2}\right)}, \tag{3.23}$$

$$\left|\mathbf{h}_{j,xz}^H \mathbf{h}_{k,xz}\right| = \frac{\sin\left(N_x \frac{\psi_x}{2}\right)}{\sin\left(\frac{\psi_x}{2}\right)} \frac{\sin\left(N_z \frac{\psi_z}{2}\right)}{\sin\left(\frac{\psi_z}{2}\right)}, \tag{3.24}$$

$$\left|\mathbf{h}_{j,yz}^H \mathbf{h}_{k,yz}\right| = \frac{\sin\left(N_y \frac{\psi_y}{2}\right)}{\sin\left(\frac{\psi_y}{2}\right)} \frac{\sin\left(N_z \frac{\psi_z}{2}\right)}{\sin\left(\frac{\psi_z}{2}\right)}, \tag{3.25}$$

where $\psi_x = 2\pi\Delta_x (\cos\phi_j \sin\theta_j - \cos\phi_k \sin\theta_k)$, $\psi_y = 2\pi\Delta_y (\sin\phi_j \sin\theta_j - \sin\phi_k \sin\theta_k)$ and $\psi_z = 2\pi\Delta_z (\cos\theta_j - \cos\theta_k)$.

A closed-form expression of the IRF for the UPA on XY-plane can be obtained by

integrating (3.14) using (3.23), that is:

$$\Upsilon_{3-\mathrm{D},xy} = \frac{4}{9N_xN_y} + \frac{2N_y}{N_x^2N_y^2}\sum_{m=1}^{N_x-1}\frac{(N_x-m)}{(2\pi\Delta_x m)^6}\Big\{(2\pi\Delta_x m)\cos(2\pi\Delta_x m) + \left[(2\pi\Delta_x m)^2-1\right]\sin(2\pi\Delta_x m)\Big\}^2 + \frac{2N_x}{N_x^2N_y^2}\sum_{n=1}^{N_y-1}\frac{(N_y-n)}{(2\pi\Delta_y n)^6}\Big\{(2\pi\Delta_y n)\cos(2\pi\Delta_y n) + \left[(2\pi\Delta_y n)^2-1\right]\sin(2\pi\Delta_y n)\Big\}^2 + \frac{4}{N_x^2N_y^2}\sum_{m=1}^{N_x-1}\sum_{n=1}^{N_y-1}\frac{(N_x-m)(N_y-n)}{\left[\sqrt{(2\pi\Delta_x m)^2+(2\pi\Delta_y n)^2}\right]^6} \times\left\{\sqrt{(2\pi\Delta_x m)^2+(2\pi\Delta_y n)^2}\cos\left[\sqrt{(2\pi\Delta_x m)^2+(2\pi\Delta_y n)^2}\right] + \left[(2\pi\Delta_x m)^2+(2\pi\Delta_y n)^2-1\right]\sin\left[\sqrt{(2\pi\Delta_x m)^2+(2\pi\Delta_y n)^2}\right]\right\}^2. \tag{3.26}$$

On the other hand, the IRF for the UPA on XZ-plane is obtained by integrating (3.14) using (3.24). The result is a semi closed-form expression equal to:

$$\Upsilon_{3-\mathrm{D},xz} = \frac{4}{9N_xN_z} + \frac{2N_z}{N_x^2N_z^2}\sum_{m=1}^{N_x-1}\frac{(N_x-m)}{(2\pi\Delta_x m)^6}\Big\{(2\pi\Delta_x m)\cos(2\pi\Delta_x m) + \left[(2\pi\Delta_x m)^2-1\right]\sin(2\pi\Delta_x m)\Big\}^2 + \frac{2N_x}{N_x^2N_z^2}\sum_{l=1}^{N_z-1}\frac{(N_z-l)}{(2\pi\Delta_z l)^6}\Big\{8+8(2\pi\Delta_z l)^2 + (2\pi\Delta_z l)^4 - \left[8+4(2\pi\Delta_z l)^2\right]\left[\cos(2\pi\Delta_z l)+(2\pi\Delta_z l)\sin(2\pi\Delta_z l)\right]\Big\} + \frac{4}{N_x^2N_z^2}\sum_{m=1}^{N_x-1}\sum_{l=1}^{N_z-1}(N_x-m)(N_z-l)\int_{\pi/2}^{\pi}\int_{\pi/2}^{\pi}J_0(2\pi\Delta_x m\sin\theta_k)\,J_0(2\pi\Delta_x m\sin\theta_j) \times\cos\left[2\pi\Delta_z l(\cos\theta_k-\cos\theta_j)\right]\sin^3(\theta_k)\sin^3(\theta_j)\,d\theta_k d\theta_j. \tag{3.27}$$

The IRF for the UPA on YZ-plane is obtained by integrating (3.14) using (3.25) and the result is equal to that obtained for the UPA on YZ-plane given by (3.27), with $\Delta_x = \Delta_y$ and $N_y = N_x$.

For 2-D scenarios, a closed-form expression of the IRF for the UPA on XY-plane can be obtained by integrating (3.16) and considering that the steering vectors inner product given by (3.23) are function of $\psi_x = 2\pi\Delta_x(\cos\phi_j - \cos\phi_k)$ and $\psi_y = 2\pi\Delta_y(\sin\phi_j - \sin\phi_k)$.

Thus, the IRF is given by:

$$\Upsilon_{xy} = \frac{1}{N_x N_y} + \frac{2N_y}{N_x^2 N_y^2} \sum_{m=1}^{N_x-1} (N_x - m)\, J_0^2\left(2\pi\Delta_x m\right) + \frac{2N_x}{N_x^2 N_y^2} \sum_{n=1}^{N_y-1} (N_y - n)\, J_0^2\left(2\pi\Delta_y n\right) + \frac{4}{N_x^2 N_y^2} \sum_{m=1}^{N_x-1} \sum_{n=1}^{N_y-1} (N_x - m)(N_y - n)\, J_0\left(2\pi\Delta_x\sqrt{m^2+n^2}\right) J_0\left(2\pi\Delta_y\sqrt{m^2+n^2}\right). \tag{3.28}$$

Finally, for 2-D scenarios and for the UPAs on XZ-plane and YZ-plane, the IRF is equivalent to the ULA on X-axis and Y-axis, given by (3.22).

3.3.1.3 Physical Space Considerations

Depending on the available physical space to deploy the antenna array, there are two cases to be considered. In the first one physical space is unlimited, thus, there is enough room to increase the number of antennas and the separation between antennas can be greater than or equal to $\lambda_c/2$. In the second one the physical space is limited, thus, there is not enough room to increase the number of antennas and the separation between antennas should be less than $\lambda_c/2$ [22].

For the sake of simplicity, consider the IRF for the ULAs of 2-D scenarios given by (3.22). Therefore, in the first case, the optimal IRF is achieved by taking the array length $\mathcal{L}_x$ to infinity in (3.22), that is:

$$\lim_{\mathcal{L}_x \to \infty} \Upsilon = \frac{1}{N_x}. \tag{3.29}$$

On the other hand, the second case is relevant for M-MIMO, once the number of antennas in limited spaces is very large. Therefore, the IRF for a large number of antennas on (3.22) in a limited space, is given by:

$$\lim_{N_x \to \infty} \Upsilon = 2\Lambda - \Psi, \tag{3.30}$$

where $\Lambda = {}_2\mathrm{F}_3\left[1/2, 1/2; 1, 1, 3/2; (-2\pi\mathcal{L}_x/\lambda_c)^2\right]$ is defined by the generalized hypergeometric function [98] and $\Psi = J_0^2\left(2\pi\mathcal{L}_x/\lambda_c\right) + J_1^2\left(2\pi\mathcal{L}_x/\lambda_c\right)$, where J_1 is the first-order Bessel function of the first kind.

For M-MIMO systems, unlimited physical space should be desirable in order to avoid the IRF saturation and to reach the optimal IRF. In general, the optimal IRF of ULAs and UPAs for 2-D scenario is given by:

$$\Upsilon_{\text{opt}} = \frac{1}{N}, \tag{3.31}$$

while for 3-D scenario it is given by:

$$\Upsilon_{3-\mathrm{D,opt}} = \frac{4}{9N}, \tag{3.32}$$

where it is considered that $N = N_x = N_y = N_z$ for ULAs and that $N = N_x N_y = N_x N_z = N_y N_z$ for UPAs.

3.3.2 IRF in Rayleigh Fading Channels

For the Rayleigh fading channel, the normalized IRF is similar to that given by (3.16) for 2-D scenarios, that can be rewritten as:

$$\begin{aligned}\Upsilon &= \mathrm{E}\left\{\frac{\left|\mathbf{h}_j^H \mathbf{h}_k\right|^2}{\overline{\alpha^2}^2 N^2}\right\} \\ \Upsilon &= \frac{1}{\overline{\alpha^2}^2 N^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\mathbf{h}_j^H \mathbf{h}_k\right|^2 p\left(\mathbf{h}_j\right) p\left(\mathbf{h}_k\right) d\mathbf{h}_j d\mathbf{h}_k,\end{aligned} \tag{3.33}$$

where the entries of $\mathbf{h}_k$ and $\mathbf{h}_j$ are i.i.d. complex-Gaussian random variables given by (3.12). Therefore, $\mathbf{h}_k$ and $\mathbf{h}_j$ are multivariate complex-Gaussian random variables with PDFs, $p(\mathbf{h}_k)$ and $p(\mathbf{h}_j)$, given by $\mathcal{CN} \sim (\mathbf{0}, \overline{\alpha^2}\mathbf{I}_N)$. By integrating (3.33), the IRF is equal to:

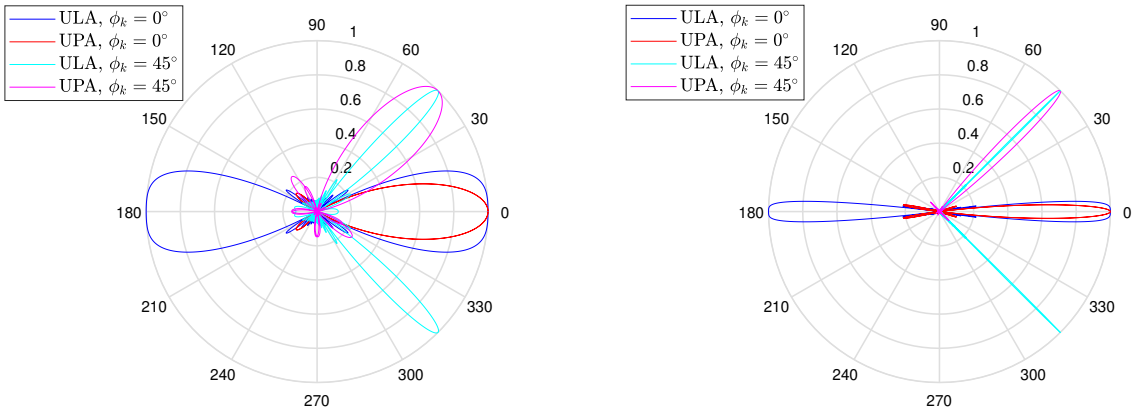
$$\Upsilon = \frac{1}{N}. \tag{3.34}$$

Notice that the IRF obtained in (3.34) is equal to the optimal IRF given by (3.31) for the UR-LOS channel. Therefore, in the Rayleigh fading channel model it is possible to assume that the antennas of the array are widely spaced in an unlimited physical space. Notice that in practical applications the separation between antennas is at least $\lambda_c/2$ [20]. In these cases, as N increases the IRF tends to zero, thus, the channel vectors becomes orthogonal and the favorable propagation established in Section 2.5.4.1 is reached.

Furthermore, in MIMO systems the Rayleigh fading is also named as spatially uncorrelated fading, once the channel coefficients between antennas are uncorrelated. In this particular scenario, it is reasonable to assume that the angle of arrival of each UT is uniformly distributed in [0, π) at each antenna. Therefore, the Jakes' model is suitable for modeling the channel time-correlations [99].

3.4 Numerical Results and Discussions

In this section, the IRF of ULAs and UPAs for UR-LOS channels and Rayleigh channels are evaluated using the proposed analytical expressions. Additionally, simulations



(a) Radiation pattern for $N = 16$.

(b) Radiation pattern for $N = 256$.

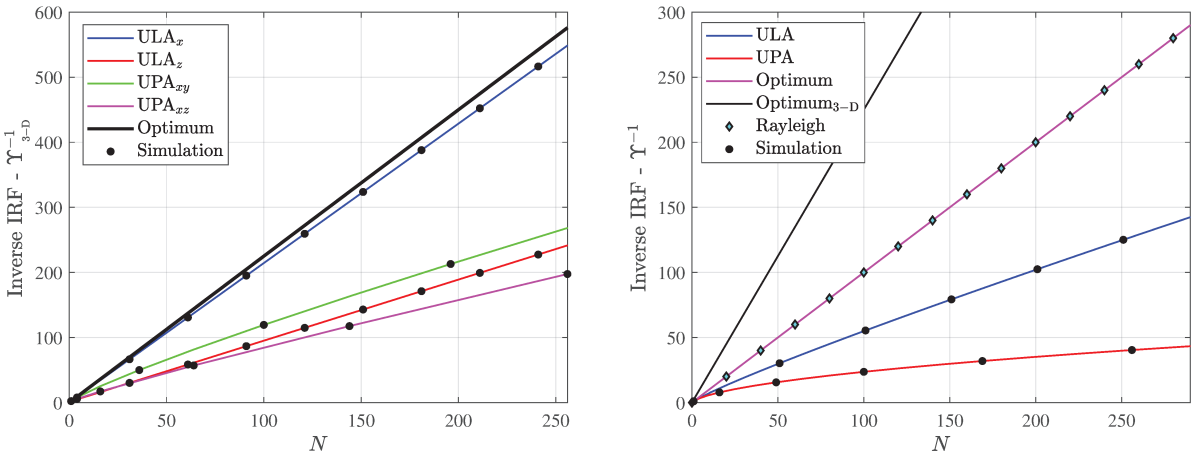
Figure 3.2: Radiation pattern of ULAs and square UPAs intended to an UT at position $\phi_k = 0°$ and $\phi_k = 45°$.

employing the Monte Carlo method are performed in order to support the validity of the proposed expressions. Furthermore, a comparison between ULAs and UPAs for unlimited and limited physical space is presented for 2-D and 3-D scenarios.

Fig. 3.2a and 3.2b show the BS azimuthal radiation pattern intended to an UT at position $\phi_k = 0°$ and $\phi_k = 45°$ using ULAs and square UPAs with $N = 16$ and $N = 256$ antennas spaced by $d = \lambda_c/2$. For plotting Fig. 3.2a and 3.2b, the equation (3.13) is used[4]. Notice that the radiation lobe is maximized on the direction of the kth UT, that is $\mathbf{h}_k = \mathbf{h}_j$ and small radiation lobes are presented for $\mathbf{h}_k \neq \mathbf{h}_j$. Besides, as expected, the radiation lobe becomes more directive as the number of antennas increases. As shown in Fig. 3.2a and 3.2b, the radiation pattern is a snapshot for a given position of the desired user. In particular, at 0° direction, the ULA lobe is wider than the UPA lobe, and in the direction of 45°, the UPA lobe is wider than the ULA which presents an additional lobe in the direction of −45°.

Fig. 3.3a shows the inverse of the IRF as a function of the number of antennas with separation between antennas of $d = d_x = d_y = d_z = \lambda_c/2$ for ULAs and UPAs in 3-D scenarios. The IRF for the ULA deployed along X-axis and Z-axis are given by (3.20) and (3.21), respectively. Besides, ULAs deployey along the X-axis or Y-axis have the same IRF. On the other hand, the IRF for square UPAs on XY-plane is given by (3.26) and for UPAs on XZ-plane it is given by (3.27). Additionally, the IRF of the UPA on YZ-plane is equivalent to UPA on XZ-plane. Notice that the ULA on X-axis and the UPA on XY-plane that are deployed on the horizontal plane reduce more interference than the arrays deployed on the vertical plane, which are the ULA on Z-axis and the UPA on XZ-plane. Furthermore, for the same number of antennas, the ULAs reduce more interference than the UPAs. Notice that the IRF of the ULA on X-axis is near

[4]The same results can be obtained by using (3.13) with $\theta_j = \theta_k = \pi/2$



(a) 3-D scenario.

(b) 2-D scenario.

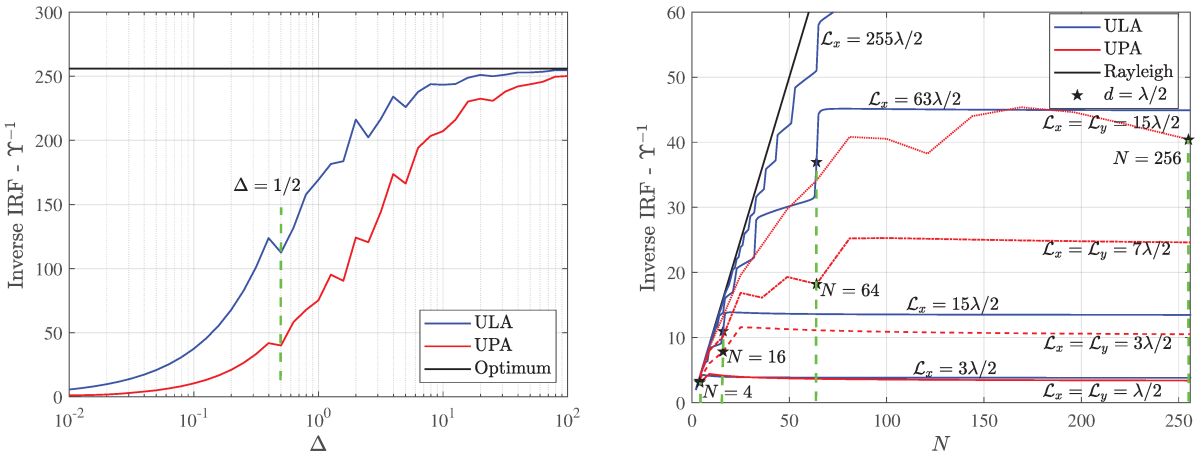
Figure 3.3: Inverse of the IRF as a function of the number of antennas for ULAs and UPAs for 2-D and 3-D scenarios with separation between antennas of $d = \lambda_c/2$.

to the optimum IRF. Finally, notice that the simulation results show that the derived theoretical expressions are very tight to simulation results.

Fig. 3.3b shows the inverse of the IRF as a function of the number of antennas with separation among them of $d = d_x = d_y = \lambda_c/2$ for ULAs and UPAs in 2-D scenarios. The IRF for the ULA and UPA are given by (3.22) and (3.28), respectively. Besides, the IRF for the Rayleigh channel given by (3.34) is also presented. As for the 3-D scenario, ULAs reduce more interference than UPAs for equal number of antennas. Observe in (3.28) that UPAs introduces more interference due to the cross antenna elements. Additionally, the optimum IRF for 2-D and 3-D scenarios are also plotted. Notice that the 3-D scenario reduce more interference than the 2-D scenario, once the elevation angle can be exploited. Furthermore, the IRF of the Rayleigh channel is equivalent to the optimum IRF for 2-D scenarios. Finally, the simulation results validate the derived theoretical expressions.

In Fig. 3.2a, 3.2b, 3.3a and 3.3b the separation between antennas is equal to $d = \lambda_c/2$. Thus, as the number of antennas increases, the array size grows. For example, considering a carrier frequency of $f_c = 3.5$ GHz and $d = \lambda_c/2$, for $N = 16$ the ULA size is $\mathcal{L}_x = 0.64$ m and the UPA edge length is $\mathcal{L}_x = \mathcal{L}_y = 0.13$ m, that has area of 0.017 m^2. Moreover, for $N = 256$ the ULA size is $\mathcal{L}_x = 10.98$ m and the UPA edge length is $\mathcal{L}_x = \mathcal{L}_y = 0.64$ m, that has area of 0.41 m^2. Notice that even increasing the IRF, the space for placing the same number of antennas in UPAs is lesser than the required for ULAs. Therefore, separation between antennas different from $d = \lambda_c/2$ should be analyzed. Furthermore, smaller antennas array can be deployed in the band of mmWaves. For example, for $N = 256$ antennas at a carrier frequency of $f_c = 28$ GHz and separation between antennas of $d = \lambda_c/2$ the ULA has size of $\mathcal{L}_x = 1.37$ m and the UPA of $\mathcal{L}_x = \mathcal{L}_y = 0.08$ m.

Fig. 3.4a shows the inverse of the IRF as a function of the normalized separation between antennas Δ for $N = 256$. Observe that $\Delta = \Delta_x$ for a ULA and that $\Delta = \Delta_x = \Delta_y$



(a) Inverse of the IRF as a function of the normalized separation between antennas for $N = 256$.

(b) Inverse of IRF as a function of the number of antennas for different edge length $\mathcal{L}_x$ and $\mathcal{L}_y$.

Figure 3.4: Inverse of the IRF for ULAs and UPAs in 2-D scenarios considering space constraints.

for a square UPA. As expected, increasing the distance between antennas also increases the IRF to the optimum value. However, the required room to deploy the array grows and could become unavailable in some practical scenarios. Thus, the IRF is limited by the available physical space. On the other hand, for $\Delta = 1/2$, the ULA IRF is near 50% of the optimum IRF and for the square UPA is near 25%. Notice that the IRF is an oscillating function, such that for $\Delta = 1/2$ the IRF is near the first minimum. In the first minimum spacing, the antennas are spatially uncorrelated. Therefore, $\Delta = 1/2$ is the normalized separation between antennas commonly used in practice.

Fig. 3.4b shows the inverse of the IRF as a function of N for different length edge $\mathcal{L}_x$ for ULAs and $\mathcal{L}_x = \mathcal{L}_y$ for square UPAs. By considering that the available physical space is equal to the array length edge, the arrays cannot grow without limit as the number of antennas N increases, resulting in a densely-spaced array with separation between antennas of $d_x < \lambda_c/2$ for ULAs and $d_x = d_y < \lambda_c/2$ for square UPAs. Fig. 3.4b shows the saturation of the IRF, which is due to the channel spatial correlation. In this case, the inter-user interference is not relieved and the favorable propagation required for M-MIMO is not reached. For ULAs, the saturation levels are given by (3.30).

The dots in Fig. 3.4b represent the IRF for separation between antennas of $d = d_x = d_y = \lambda_c/2$, which optimize the IRF before the channel saturation, i.e., more inter-user interference could be rejected for separation between antennas of $d = d_x = d_y < \lambda_c/2$. Tab. 3.1 shows the minimum distance between antennas $d_{x,\min}$, for a given length edge. For comparison purposes, the IRF for $d = d_x = d_y = \lambda_c/2$ is presented. Notice that the number of antennas can be increased, as the distance between antennas decreases, resulting in the increase of IRF.

Additionally, Fig. 3.4b shows the IRF for the Rayleigh fading channel model. It is a mistake to consider this model in environments with limited physical space, since

Table 3.1: Inverse of the IRF for ULAs and UPAs with different separation between antennas.

Array	Edge	Typical Spacing			Spacing Reduction		
	L_x	d_x	N_x	$\overline{\Upsilon}^{-1}$	$d_{x,\min}$	N_x	$\overline{\Upsilon}^{-1}_{\min}$
ULA	$3\lambda_c/2$	$\lambda_c/2$	4	3.32	$3\lambda_c/8$	5	4.27
	$15\lambda_c/2$		16	10.97	$15\lambda_c/38$	20	13.87
	$63\lambda_c/2$		64	36.96	$63\lambda_c/154$	78	45.16
UPA	$\lambda_c/2$	$\lambda_c/2$	4	3.09	$\lambda_c/4$	9	4.40
	$3\lambda_c/2$		16	7.83	$3\lambda_c/8$	25	11.59
	$7\lambda_c/2$		64	18.21	$7\lambda_c/18$	100	25.28

Table 3.2: Inverse of the IRF for different UPAs configuration.

N_x	N_y	$\mathcal{L}_x$	$\mathcal{L}_y$	Υ^{-1}
256	1	$255\lambda_c/2$	0	127.33
128	2	$127\lambda_c/2$	$\lambda_c/2$	81.32
64	4	$63\lambda_c/2$	$3\lambda_c/2$	54.87
32	8	$31\lambda_c/2$	$7\lambda_c/2$	43.70
16	16	$15\lambda_c/2$	$15\lambda_c/2$	40.38

in this channel the antennas arrays shows spatial uncorrelation. For this condition the separation between antennas must be at least $d = \lambda_c/2$. Besides, due to the propagation characteristics, the Rayleigh fading channel is not a good model for mmWaves [92].

Finally, the expressions of the IRF for UPAs given by (3.26), (3.27) and (3.28) are valid for square and rectangular arrays. Tab. 3.2 shows the IRF for $N = N_x N_y = 256$ antennas in 2-D scenarios for different edge lengths. Notice that more linear the arrays of antennas, the better the IRF.

CHAPTER 4

Channel Estimation

4.1 Introduction

The channel estimation is a fundamental process for obtaining the CSI in order to perform coherent detection, as indicated in Section 2.6. MPE, SPE, and HPE are described. Additionally, the perfect channel estimation (PCE) is also presented for comparison purposes. The estimation techniques are analyzed in terms of the average energy and spectral efficiency for TVCs and TICs.

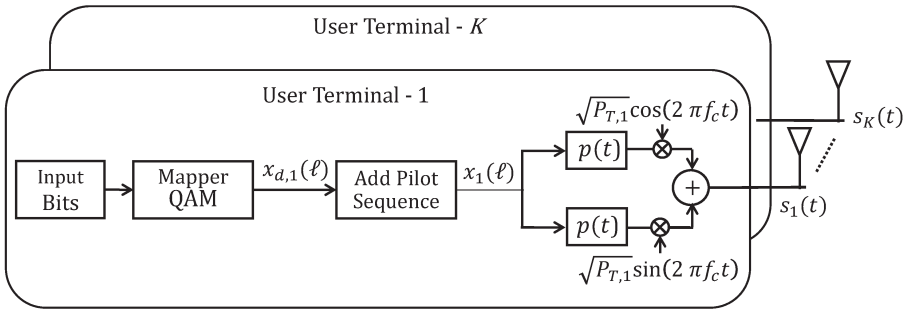
From the estimation theory, there are two main estimate process: ML and MMSE. The MMSE estimator is considered in this work due to its better performance [86]. Furthermore, the quality of the estimated channel is presented in terms of the NMSE.

In this chapter, the M-MIMO system model is also described. For this purpose, the transmitted and received signals are detailed for the uplink of a single cell cellular system considering M-MIMO.

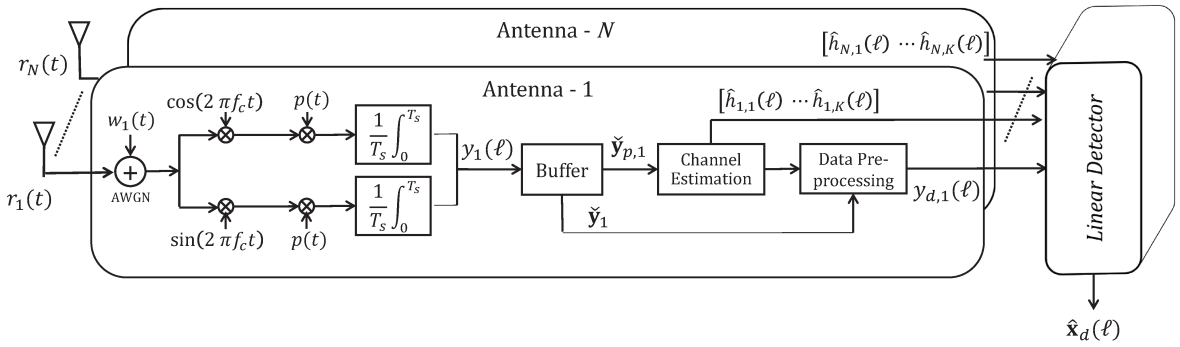
The system model is described in Section 4.2. The channel estimation techniques are presented in Section 4.3. The channel coefficients estimation is derived in Section 4.4 and the channel estimation error is presented in Section 4.5. Finally, the numerical results and discussion are carried out in Section 4.6.

4.2 System Model

The uplink of a single cell system employing M-MIMO with imperfect channel estimation is analyzed in this next chapter. Fig. 4.1 presents the system model including the UT and the receiver structure at the BS. Moreover, Fig. 4.2 shows a summarized representation of the cellular scenario.



(a) UT system model.



(b) BS system model.

Figure 4.1: M-MIMO system model.

4.2.1 Transmitted and Received Signals

The signal transmitted by the jth UT during a symbol time interval can be written as:

$$s_j(t) = \sqrt{P_{T,j}}\ \Re\{x_j \exp[\,\mathbb{i}2\pi f_c t]\mathrm{p}(t)\}\,, \tag{4.1}$$

where $\sqrt{P_{T,j}}$ is the jth UT transmitted power, x_j is the transmitted symbol by jth UT belonging to a M-QAM constellation with normalized mean power $\overline{x^2}$ and $\mathbb{i}=\sqrt{-1}$, whose in-phase and in-quadrature components are denoted as $x_{I,j}$ and $x_{Q,j}$, that is $x_j = x_{I,j} + \mathbb{i}x_{Q,j}$. Moreover, f_c is the carrier frequency and $\mathrm{p}(t)$ is a baseband pulse of duration T_s that satisfies the Nyquist criterion.

The matched filter output sample of the ith BS antenna for the ℓth symbol time interval can be written as:

$$y_i(\ell) = \frac{1}{T_s}\int_{\ell\tau}^{\ell\tau+T_s} [\mathcal{M}_i(t) + w_i(t)] \exp[\,-\mathbb{i}2\pi f_c t\,]\,\mathrm{p}(t-\ell\tau)dt, \tag{4.2}$$

where τ is the channel delay. For simplicity, the synchronous case is supposed, i.e., $\tau_j = \tau,\ \forall j$, where τ_j is the channel delay of each UT. This supposition performs worse than the asynchronous case and allows simpler analysis [100]. Additionally, $\mathcal{M}_i(t)$ represents

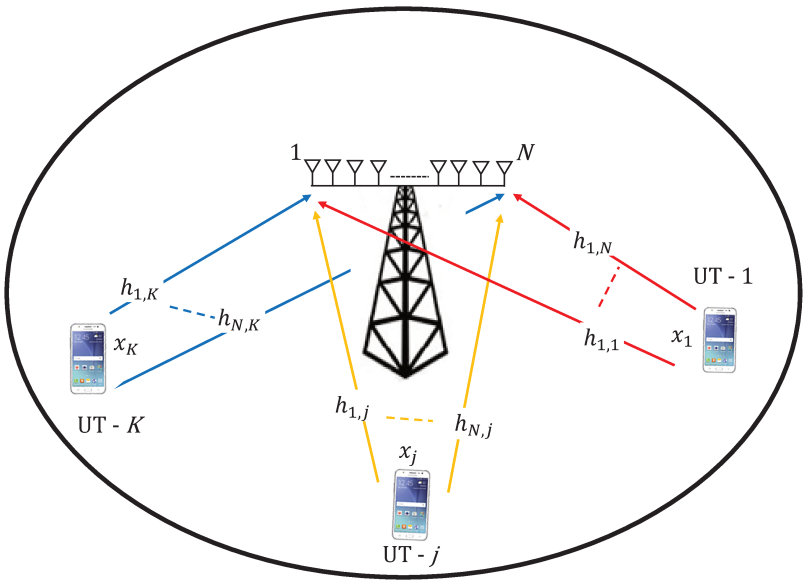


Figure 4.2: M-MIMO uplink cellular system.

the multiple access signals of all K UTs arriving to the ith BS antenna and $w_i(t)$ is the noise at the ith antenna. Notice that the multiple access signal $\mathcal{M}_i(t)$ can be written as:

$$\mathcal{M}_i(t) = h_{i,j}(t,\tau) * \sum_{j=1}^{K} \sqrt{P_{R,j}}\ \Re\{x_j \exp\left[\,\mathbb{i}2\pi\,(f_c)\,t\,\right] \mathrm{p}(t-\ell\tau)\}, \tag{4.3}$$

where $*$ denotes the convolution operation, $P_{R,j}$ is the received power from the jth UT and $h_{i,j}(t)$ is the channel impulse response between the jth user and the ith antenna. By considering that $f_c \gg \frac{1}{T_s}$, that the pulse shape is rectangular with unitary amplitude and applying some identities, it is easy to show that the received sample of the multiple access signal during the ℓth time slot at the ith antenna is given by:

$$\begin{aligned}\mathcal{M}_i(\ell) &= \frac{1}{T_s}\int_{\ell\tau}^{\ell\tau+T_s} \mathcal{M}_i(t) \exp\left[\,-\mathbb{i}2\pi f_c t\,\right] \mathrm{p}(t-\ell\tau)dt \\ &= \frac{1}{2}\sum_{j=1}^{K} h_{i,j}(\ell)x_j(\ell),\end{aligned} \tag{4.4}$$

where $h_{i,j}(\ell)$ is the sample of the channel impulse response defined in Section 3.2.2 for the ℓth time slot and $x_j(\ell)$ is the transmitted symbol by the jth UT. Notice that the 1/2 factor is introduced by modulation/demodulation process. Additionally, perfect power control is considered. Thus, $P_{R,j} = P_{T,j}\mathcal{K}r_j^{-\beta}$ and $P_{T,j} = \mathcal{K}^{-1}r_j^{\beta}$, where r_j is the distance between the jth UT and the BS array, $\mathcal{K}$ is the propagation factor and β is the path-loss exponent.

On the other hand, the noise sample at the ith BS antenna for the ℓth time slot is given by:

$$w_i(\ell) = \frac{1}{T_s}\int_{\ell\tau}^{T_s} w_i(t)\left\{\cos\left(2\pi f_c t\right) - \mathbb{i}\sin\left(2\pi f_c t\right)\right\}\mathrm{p}(t-\ell\tau)dt, \tag{4.5}$$

where the Euler identity is employed $\exp(\mathbb{i}2\pi f_c t) = \cos(2\pi f_c t) + \mathbb{i}\sin(2\pi f_c t)$. From (4.5), it is easy to show that noise samples have zero mean. Furthermore, the variance of the noise in-phase component samples can be obtained by:

$$\sigma_w^2 = \frac{1}{T_s^2}\int_{\ell\tau}^{\ell\tau+T_s}\int_{\ell\tau}^{\ell\tau+T_s}\mathrm{E}\left\{w_i(t)w_i(t')\right\}\mathrm{p}(t-\ell\tau)\mathrm{p}(t'-\ell\tau)\cos(2\pi f_c t)\cos(2\pi f_c t')dtdt', \tag{4.6}$$

where

$$\mathrm{E}\left\{w_i(t)w_i(t')\right\} = \frac{N_0}{2}\delta(t'-t), \tag{4.7}$$

is the white noise autocorrelation function, N_0 is the unilateral noise power spectral density and $\delta(x)$ is the Dirac delta function. By substituting (4.7) in (4.6), the variance of the noise in-phase component samples can be rewritten as:

$$\sigma_{w,I}^2 = \frac{N_0}{2T_s^2}\int_{\ell\tau}^{\ell\tau+T_s}\mathrm{p}^2(t-\ell\tau)\cos^2(2\pi f_c t)dt, \tag{4.8}$$

where it has been employed that $\int g(t')\delta(t'-t)dt' = g(t)$, where $g(t')$ is any function of t'. Thus, the integral result is given by:

$$\sigma_{w,I}^2 = \frac{N_0}{4T_s}. \tag{4.9}$$

Notice from (4.7) that the white noise is an uncorrelated stochastic process. Therefore, the noise samples are uncorrelated. Moreover, it is easy to show that the variance of the noise quadrature component samples are also given by (4.9), that is $\sigma_{w,Q}^2 = \sigma_{w,I}^2$. Therefore, the sample $\mathbf{w}_i(\ell)$ represents the additive complex white Gaussian noise, whose entries are independent and identically distributed random variables with distribution $\mathcal{CN}\left(0,\sigma_w^2\right)$, where the noise sample variance is equal to:

$$\sigma_w^2 = \frac{N_0}{2T_s}. \tag{4.10}$$

Finally, the sample at the output of the matched filter given by (4.2) can be rewritten as:

$$y_i(\ell) = \mathcal{M}_i(\ell) + w_i(\ell). \tag{4.11}$$

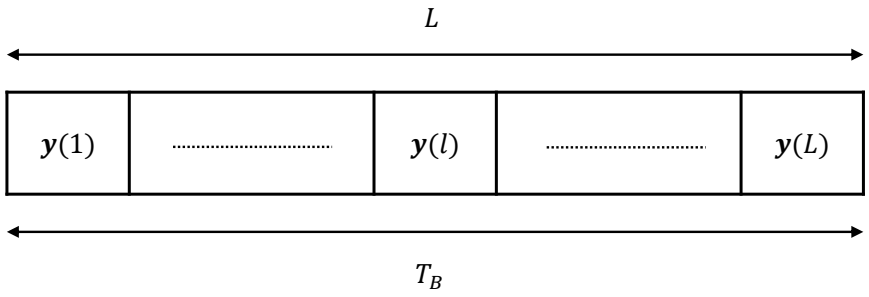


Figure 4.3: Temporal block structure of duration T_B consisting of L data and pilot symbols.

4.2.2 Matrix Form of the Received Samples

All the samples at the output of the matched filter bank can be written as an N dimensional column vector, that is:

$$\mathbf{y}(\ell) = \left[y_1(\ell)\ y_2(\ell)\ \ldots\ y_N(\ell)\right]^T. \tag{4.12}$$

From (4.11), (4.5), and (4.4), the received vector $\mathbf{y}(\ell)$ can be obtained as

$$\mathbf{y}(\ell) = \frac{1}{2}\mathbf{H}(\ell)\mathbf{x}(\ell) + \mathbf{w}(\ell), \tag{4.13}$$

where $\mathbf{H}(\ell) = [\mathbf{h}_1(\ell)\ \mathbf{h}_2(\ell)\ \cdots\ \mathbf{h}_K(\ell)]$ is the $N \times K$ channel matrix and $\mathbf{h}_k(\ell)$ is the jth UT channel vector with entries $h_{i,j}(\ell)$, for $i = 1,\ 2,\ \cdots,\ N$ and $j = 1,\ 2,\ \cdots,\ K$. Additionally, $\mathbf{x}(\ell) = [x_1(\ell)\ x_2(\ell)\ \cdots\ x_K(\ell)]^T$ is the $K \times 1$ vector of transmitted symbols and $\mathbf{w}(\ell) = [w_1(\ell)\ w_2(\ell)\ \cdots\ w_N(\ell)]^T$ is the $N \times 1$ vector of noise samples.

4.2.3 Block Structure

Fig. 4.3 shows the temporal block structure of duration T_B where L is the number of time slots consisting in data and pilot symbols. For N BS antennas, the received samples in a block can be written as a $N \times L$ matrix, that is:

$$\mathbf{Y} = [\mathbf{y}(1)\ \cdots\ \mathbf{y}(\ell)\ \cdots\ \mathbf{y}(L)], \tag{4.14}$$

where $\mathbf{y}(\ell)$ is the $N \times 1$ column vector of received samples at the ℓth time interval, given by (4.13). For a slow fading channel $T_B \leqslant T_c$, where $T_c = \frac{1}{2f_{d,\max}}$ is the channel coherence time interval.

The block structure allows to identify the pilot and data symbols position. Additionally, two different block structures can be defined: for TVCs and for TICs.

4.2.3.1 Time-Variant Channel

The received samples at the ith antenna during a block are given by the ith row of (4.14) that can be rewritten as:

$$\check{\mathbf{y}}_i = [y_i(1) \ \cdots \ y_i(\ell) \ \cdots \ y_i(L)]^T, \tag{4.15}$$

which is a $L \times 1$ vector with entries $y_i(\ell) = \sum_{j=1}^{K} h_{i,j}(\ell)x_j(\ell) + w_i(\ell)$, for $i = 1,\ 2, \cdots,\ N$ and $\ell = 1,\ 2,\ \cdots,\ L$. The elements of (4.15) are also obtained as:

$$\check{\mathbf{y}}_i = \sum_{j=1}^{K} \frac{1}{2}\check{\mathbf{X}}_j\check{\mathbf{h}}_{i,j} + \mathbf{w}_i, \tag{4.16}$$

where $\check{\mathbf{h}}_{i,j} = [h_{i,j}(1)\ h_{i,j}(2)\ \cdots\ h_{i,j}(L)]^T$ is the $L \times 1$ time-variant channel vector for the link between the jth UT and the ith BS antenna, $\check{\mathbf{X}}_j = \text{diag}(\mathbf{x}_j)$ is the vector of transmitted symbols by the jth UT in the form of a $L \times L$ diagonal matrix, and $\mathbf{w}_i = [w_i(1)\ w_i(2)\ \cdots\ w_i(L)]^T$ is the $L \times 1$ noise vector.

For TVCs, the temporal correlation between the channel coefficients is given by (2.16). Therefore, the channel autocorrelation matrix $\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} = \text{E}\{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}^H\}$, has dimension $L \times L$ whose (m,ℓ)th entry is given by:

$$\left(\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}\right)_{m,\ell} = \overline{\alpha^2} J_0\left[2\pi f_{d,\max} T_s(\ell - m)\right],\ \forall\ \ell, m = 1, 2, \cdots, L, \tag{4.17}$$

where $T_s = T_B/L$ is the symbol duration. Notice that, $\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}$ is a symmetric Toeplitz matrix [101].

4.2.3.2 Time-Invariant Channel

For the time-invariant channel, it is considerd that $\mathbf{H}(\ell) = \mathbf{H}$ for $\ell = 1, \cdots L$, i.e., it is assumed that the channel coefficients remain invariant during a block. It is a particular case of the time-variant channel, where the maximum Doppler shift $f_{d,\max} = 0$. Moreover, the received symbols at the ith antenna during a block are given by (4.15) and (4.16), with channel vector entries equal to $\check{\mathbf{h}}_{i,j} = h_{i,j}\mathbf{1}$, where $\mathbf{1}$ is a vector with ones with dimension $L \times 1$ and $h_{i,j}$ is the TIC channel coefficient.

For TICs the channel vector $\check{\mathbf{h}}_{i,j}$ has an autocorrelation matrix given by:

$$\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} = \overline{\alpha^2}\ \mathbf{J}_L, \tag{4.18}$$

where $\mathbf{J}_L = \mathbf{1}\mathbf{1}^T$ is an all-ones matrix of dimension $L \times L$.

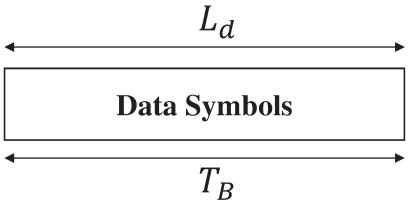


Figure 4.4: Block of $L = L_d$ transmitted data symbols by the kth UT during a block for PCE.

4.3 Estimation Techniques

In this section, the basis of MPE, SPE and HPE estimation techniques, which are also named as training techniques are presented. For this purpose, these techniques are compared to PCE in terms of average symbol energy and spectral efficiency. Furthermore, in the M-MIMO system model there are two blocks used for channel estimation as shown in Fig. 4.1b. The first block is the buffer that storage the samples required for channel estimation technique and the second block performs estimation of the channel coefficients.

4.3.1 Perfect Channel Estimation

PCE or perfect CSI is widely used in communications systems analysis due to its simplicity. In this utopic scenario, the BS has perfect knowledge of all channel matrix coefficients without transmitting pilot symbols to perform the channel estimation [73]. Fig. 4.4 shows a block of symbols for PCE. Notice that only data symbols are transmitted, that is $L = L_d$, where L_d is the number of data symbols in a block.

The received energy from the kth UT at the BS during a block is given by:

$$E_{B,k} = \sum_{\ell=1}^{L} \frac{1}{2}|x_k(\ell)|^2 T_s, \tag{4.19}$$

where $\frac{1}{2}|x_k(\ell)|^2$ is the power of the ℓth data symbol transmitted by the kth UT.

Furthermore, the received power at each BS antenna is the same for all UTs at each symbol time interval, i.e., $\frac{1}{2}\overline{|x_k(\ell)|^2} = \frac{1}{2}\overline{|x|^2}$ $\forall k, \ell$. Therefore, the mean received energy per block is given by:

$$E_B = \frac{1}{2}\overline{|x|^2} L T_s, \tag{4.20}$$

where $\frac{1}{2}\overline{|x|^2}$ is the constellation mean power.

For posterior comparison purposes, the average energy per data symbol is defined as:

$$E_s = \frac{E_B}{L_d}, \tag{4.21}$$

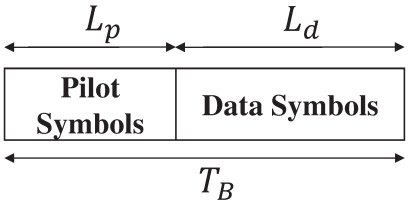


Figure 4.5: Block of $L = L_p + L_d$ transmitted symbols by the kth UT during a block for MPE.

where the average energy per data symbol is equal to the average symbol energy, that is, $E_s = \frac{1}{2}\overline{|x|^2}T_s$.

The cellular spectral efficiency, defined in Section 2.3.2, is given by (2.8), where $\mathcal{R}_{b,j}$ is the j-th UT bit rate and $B = 1/T_s$ is the system bandwidth. Moreover, the block duration can also be written as:

$$T_B = LT_s = L_dT_s. \tag{4.22}$$

Then, by using (4.22) in (2.8), the average cellular spectral efficiency is given by:

$$\xi = K \log_2 M, \tag{4.23}$$

where it is considered that all K UTs transmit the same bit rate $\mathcal{R}_b = \frac{\log_2 M}{T_s}$.

4.3.2 Multiplexed Pilot Estimation

In this technique, pilot symbols for the channel estimation are multiplexed with data symbols. Thus, $L = L_p + L_d$ symbols are transmitted during a block, where L_p is the number of pilot symbols and L_d is the number of data symbols, as shown in Fig. 4.5. For the channel estimation, all L_p pilot symbols are known by the receiver.

The received energy from the kth UT at BS during a block is given by:

$$E_{B,k} = \sum_{\ell=1}^{L_p} \frac{1}{2}|x_{p,k}(\ell)|^2 T_{s,m} + \sum_{\ell=L_p+1}^{L} \frac{1}{2}|x_{d,k}(\ell)|^2 T_{s,m}, \tag{4.24}$$

where $\frac{1}{2}|x_{p,k}(\ell)|^2$ is the power of the ℓth pilot symbol, $\frac{1}{2}|x_{d,k}(\ell)|^2$ is the power of the ℓth data symbol and $T_{s,m}$ is the symbol duration for MPE. Notice that the block duration can also be written as:

$$T_B = (L_p + L_d)T_{s,m}. \tag{4.25}$$

For a fair comparison between PCE and MPE, both systems transmit the same number of data symbols during a block. Thus, it is easy to show that the relation between the

symbol time intervals for PCE and MPE is given by:

$$T_s = (1+\eta)\, T_{s,m}, \tag{4.26}$$

where

$$\eta = \frac{L_p}{L_d}. \tag{4.27}$$

Notice that $\eta = 0$ for PCE and $\eta > 0$ for MPE. From (4.26), observe that $T_{s,m} < T_s$, which represents a bandwidth expansion for MPE.

Besides, the received energy per block can be rewritten as:

$$E_B = L_p E_p + L_d E_d, \tag{4.28}$$

where $E_p = \frac{1}{2}|x_p|^2 T_{s,m}$ is the pilot symbols energy[5] and $E_d = \frac{1}{2}\overline{|x_d|^2} T_{s,m}$ is the average energy of the data symbols. Additionally, equating the received energy per block E_B, for both techniques, it is easy to show that the relation between energies of PCE and MPE techniques is given by:

$$E_s = \eta E_p + E_d, \tag{4.29}$$

where E_s is the average symbol energy for the PCE scenario. By introducing the ratio:

$$\mu = \frac{E_p}{E_d}, \tag{4.30}$$

the pilot and data symbol energies for MPE are given by:

$$E_p = \frac{\mu}{(1+\eta\mu)} E_s, \tag{4.31}$$

$$E_d = \frac{1}{(1+\eta\mu)} E_s. \tag{4.32}$$

Finally, by considering the effects of bandwidth expansion, the MPE spectral efficiency is given by:

$$\xi = \frac{K \log_2 M}{1+\eta}, \tag{4.33}$$

where in (2.8) all K UTs transmit at the same bit rate $\mathcal{R}_b = \frac{\log_2 M}{T_s}$ and that the bandwidth is $B = \frac{(1+\eta)}{T_s}$. Notice that, as expected, the spectral efficiency of MPE is reduced, since $\eta > 0$.

[5]Once the pilot symbols are deterministic (known at the receiver), their energy do not need to be expressed in terms of the average.

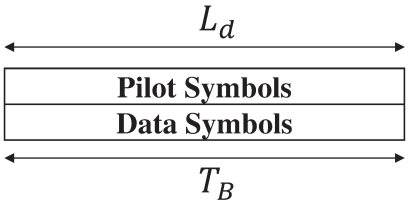


Figure 4.6: Block of $L = L_p = L_d$ transmitted symbols by the kth UT during a block for SPE.

4.3.3 Superimposed Pilot Estimation

This technique consists in superimposing pilot symbols with data symbols. Thus, $L = L_p = L_d$ symbols are transmitted during a block as shown in Fig. 4.6.

The received energy from the kth UT at BS during a block is given by:

$$E_{B,k} = \sum_{\ell=1}^{L} \frac{1}{2} \left[|x_{p,k}(\ell)|^2 + |x_{d,k}(\ell)|^2 \right] T_{s,s}, \tag{4.34}$$

where $T_{s,s}$ is the symbol duration for SPE. Furthermore, the block duration is given by:

$$T_B = LT_{s,s}. \tag{4.35}$$

For a fair comparison between PCE and SPE, the same block duration should be considered in both schemes. Thus, it is easy to show that the relation between the symbol time intervals is given by:

$$T_s = T_{s,s}. \tag{4.36}$$

Observe that the symbol time interval is the same. Thus, in this case, like for PCE there is no bandwidth expansion. Additionally, the spectral efficiency for the SPE is also given by (4.23). This is the reason for considering that SPE is bandwidth efficient at the cost of the interference. This impairment is evaluated later.

The received energy per block can be rewritten as:

$$E_B = L \left(E_p + E_d \right), \tag{4.37}$$

where $E_p = \frac{1}{2}|x_p|^2 T_{s,s}$ is the pilot symbol energy and $E_d = \frac{1}{2}\overline{|x_d|^2} T_{s,s}$ is the data symbol energy. Therefore, the received energy relation between PCE and SPE considering same energy per block, is given by:

$$E_s = E_p + E_d. \tag{4.38}$$

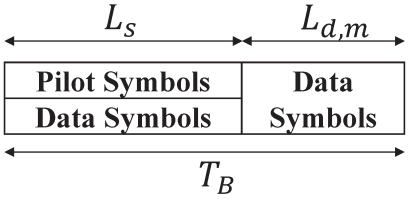


Figure 4.7: Block of $L = L_s + L_{d,m}$ transmitted symbols by the kth UT during a block for HPE.

By introducing the ratio between pilot and data energy,

$$\mu_s = \frac{E_p}{E_d}, \tag{4.39}$$

the pilot and data symbol energy are given respectively by:

$$E_p = \frac{\mu_s}{(1+\mu_s)} E_s, \tag{4.40}$$

$$E_d = \frac{1}{(1+\mu_s)} E_s. \tag{4.41}$$

Notice that E_s is the same average symbol energy of PCE.

4.3.4 Hybrid Pilot Estimation

HPE is a hybrid technique that employs both MPE and SPE. The block is divided into two groups, as shown in Fig. 4.7. The first group consists of superimposed symbols, while the second group of multiplexed data symbols. The block length is given by:

$$L = L_s + L_{d,m}, \tag{4.42}$$

where $L_s = L_{p,s} = L_{d,s}$ is the number of superimposed symbols, $L_{p,s}$ is the number of pilot symbols, $L_{d,s}$, is the number of superimposed data symbols and $L_{d,m}$ is the number of the multiplexed data symbols. Notice that the number of data symbols is $L_d = L_{d,s} + L_{d,m}$.

The received energy from the kth UT at BS during a block is given by:

$$\begin{aligned} E_{B,k} = &\sum_{\ell=1}^{L_s} \frac{1}{2} \left[|x_{p,s,k}(\ell)|^2 + |x_{d,s,k}(\ell)|^2 \right] T_{s,h} \\ &+ \sum_{\ell=L_s+1}^{L} \frac{1}{2} |x_{d,m,k}(\ell)|^2 T_{s,h}, \end{aligned} \tag{4.43}$$

where $\frac{1}{2}|x_{p,s,k}(\ell)|^2$ is the power of pilot symbols, $\frac{1}{2}|x_{d,s,k}(\ell)|^2$ is the power of superimposed data symbols, $\frac{1}{2}|x_{d,m,k}(\ell)|^2$ is the power of multiplexed data symbols and $T_{s,h}$ is the symbol duration for HPE.

For a fair comparison between PCE and HPE, in this case, the relation between the symbol duration is given by:

$$T_B = (L_s + L_{d,m})T_{s,h}. \tag{4.44}$$

For the spectral efficiency analysis, by using (4.44) and (4.42), it is easy to show that $T_s = T_{s,h}$. Thus, there is no bandwidth expansion for HPE. As a consequence, the spectral efficiency of HPE is also given by (4.23).

The received energy per block given by (4.43), can be rewritten as:

$$E_B = L_s(E_{p,s} + E_{d,s}) + L_{d,m}E_{d,m}, \tag{4.45}$$

where $E_{p,s} = \frac{1}{2}|x_{p,s}|^2 T_{s,h}$ is the pilot symbol energy, $E_{d,s} = \frac{1}{2}\overline{|x_{d,s}|^2}T_{s,h}$ is the average energy of superimposed data symbols and $E_{d,m} = \frac{1}{2}\overline{|x_{d,m}|^2}T_{s,h}$ is the average energy of multiplexed data symbols.

Considering that the block energy and the number of data symbols in a block are equal for PCE and HPE, then the relation between the energies is given by:

$$E_s = \frac{\eta_h}{1+\eta_h}E_{p,s} + \frac{\eta_h}{1+\eta_h}E_{d,s} + \frac{1}{1+\eta_h}E_{d,m}, \tag{4.46}$$

where

$$\eta_h = \frac{L_s}{L_{d,m}}. \tag{4.47}$$

Furthermore, by introducing the ratios

$$\mu_{h,m} = \frac{E_{p,s} + E_{d,s}}{E_{d,m}}, \tag{4.48}$$

and

$$\mu_{h,s} = \frac{E_{p,s}}{E_{d,s}}, \tag{4.49}$$

it is straightforward to show that the pilot, the superimposed data and multiplexed data symbol energy are given respectively by:

$$E_{p,s} = \frac{\mu_{h,s}\mu_{h,m}(1+\eta_h)}{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}E_s, \tag{4.50}$$

$$E_{d,s} = \frac{\mu_{h,m}(1+\eta_h)}{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}E_s, \tag{4.51}$$

$$E_{d,m} = \frac{(1+\eta_h)}{(1+\eta_h\mu_{h,m})}E_s. \tag{4.52}$$

Finally, notice that HPE can be considered the general estimation scenario and the other techniques as special cases. Notice that the received energy per block of HPE,

given by (4.45), can be reduced to the energy per block of SPE, given by (4.37), by using $L_{d,m} = 0$ in (4.45), or yet to the energy per block of MPE, given by (4.28), by using $E_{d,s} = 0$ in (4.45) or to the energy per block of PCE, given by (4.20), by using $E_{p,s} = 0$ in the same equation.

4.4 Channel Coefficients Estimation

In this section, the channel coefficients are estimated by using the MMSE estimator/predictor presented in Section 2.7.1 and the estimation techniques shown in the previous section. Moreover, the channel coefficients are estimated for TVCs and TICs. For this purpose, the next assumptions are made:

The N antennas at the BS are enough spaced and all K UTs are uniformly distributed in the cell. Thus, the spatially uncorrelated scenario is assumed [20, 65]. Therefore, the coefficients of the channel matrix are uncorrelated and they can be estimated independently at each antenna.

Furthermore, for a channel estimation of all K UTs, each user terminal must transmit at least K pilot symbols [18]. Besides, the pilot sequences transmitted by the kth UT must be pairwise orthogonal in order to eliminate the MAI during the channel estimation. Thus, at the output of the matched filter of each BS antenna, the received pilots matrix are stored in a buffer in order to perform the channel estimation as shown in Fig. 4.1b.

Additionally, the mean and variance of the estimated channel coefficients are derived in order to evaluate the channel estimation quality.

4.4.1 Channel Coefficients Estimation for TVCs

In this section, the channel coefficients are estimated/predicted assuming TVCs. The CSI is obtained for each one of the three estimation techniques.

4.4.1.1 CSI for MPE

The matrix of received pilot samples for MPE is obtained from the overall block given by (4.14), that is:

$$\mathbf{Y}_p = [\mathbf{y}(1) \ \cdots \mathbf{y}(\ell) \ \cdots \ \mathbf{y}(L_p)]\,. \tag{4.53}$$

Therefore, from (4.16), the vector of pilot samples of length L_p received at the ith antenna can be written as:

$$\check{\mathbf{y}}_{p,i} = \sum_{j=1}^{K} \frac{1}{2} \check{\mathbf{X}}_{p,j} \check{\mathbf{h}}_{i,j} + \mathbf{w}_i, \tag{4.54}$$

where $\check{\mathbf{h}}_{i,j} = [h_{i,j}(1)\ h_{i,j}(2)\ \cdots\ h_{i,j}(L_p)]^T$ and $\mathbf{w}_i = [w_i(1)\ w_i(2)\ \cdots\ w_i(L_p)]^T$ are the channel vector and the noise vector, respectively. Additionally, $\check{\mathbf{X}}_{p,j} = \text{diag}(\mathbf{x}_{p,j})$ is a diagonal matrix of dimension $L_p \times L_p$ that contains the vector of pilot symbols $\mathbf{x}_{p,j}$ of dimension $1 \times L_p$ transmitted by the jth UT, that is selected from the pilots matrix $\mathbf{X}_p$.

The pilots matrix can be defined as $\mathbf{X}_p = \mathbf{\Delta}\Phi$, where $\mathbf{\Delta}$ is a $K \times K$ diagonal matrix containing the amplitude of the transmitted pilot symbols. By assuming that the received power is the same for all K UTs at each symbol time interval, then, $x_{p,j}(1) = x_{p,j}(2) = \cdots = x_{p,j}(L_p) = x_p$ and $\mathbf{\Delta} = x_p\mathbf{I}_K$. Besides, Φ is a $K \times L_p$ matrix containing the first K rows of a $L_p \times L_p$ Hadamard matrix (see Appendix A.5 for the Hadamard matrix construction), such that $K \leqslant L_p$, where $K = 2^r$ for $r \geqslant 0$ and $L_p \geqslant 2^q K$ for $q \geqslant 1$. Notice that $\check{\mathbf{X}}_{p,j} = \text{diag}(\mathbf{x}_{p,j})$ can be rewritten as $\check{\mathbf{X}}_{p,j} = x_p\check{\Phi}_j$, by considering that $\mathbf{x}_{p,j} = x_p\boldsymbol{\varphi}_j$, where $\boldsymbol{\varphi}_j$ is the jth row vector of Φ which can be rewritten as a diagonal matrix, i.e., $\check{\Phi}_j = \text{diag}(\boldsymbol{\varphi}_j)$. In order to make the definition of $\check{\mathbf{X}}_{p,j}$ clearer, consider the next example. Given $K = 2$ and $L_p = 4$, the matrix of pilot symbols is given by:

$$\mathbf{X}_p = \begin{bmatrix} x_{p,1}(1) & x_{p,1}(2) & x_{p,1}(3) & x_{p,1}(4) \\ x_{p,2}(1) & x_{p,2}(2) & x_{p,2}(3) & x_{p,2}(4) \end{bmatrix}, \tag{4.55}$$

where $\mathbf{x}_{p,1} = [x_{p,1}(1)\ x_{p,1}(2)\ x_{p,1}(3)\ x_{p,1}(4)]$ and $\mathbf{x}_{p,2} = [x_{p,2}(1)\ x_{p,2}(2)\ x_{p,2}(3)\ x_{p,2}(4)]$. Therefore, for the second user, $\check{\mathbf{X}}_{p,2}$ can be written as:

$$\check{\mathbf{X}}_{p,2} = \begin{bmatrix} x_{p,2}(1) & 0 & 0 & 0 \\ 0 & x_{p,2}(2) & 0 & 0 \\ 0 & 0 & x_{p,2}(3) & 0 \\ 0 & 0 & 0 & x_{p,2}(4) \end{bmatrix}. \tag{4.56}$$

By considering that $\mathbf{X}_p = \mathbf{\Delta}\Phi$, (4.55) can be rewritten as:

$$\mathbf{X}_p = x_p \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \tag{4.57}$$

where $\Delta = x_p\mathbf{I}_2$ and Φ is composed by the first two rows of the Hadamard matrix:

$$\mathcal{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \tag{4.58}$$

Thus, for the second user, (4.56) can be rewritten as:

$$\check{\mathbf{X}}_{p,2} = x_p \check{\Phi}_2 = x_p \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{4.59}$$

where $\check{\Phi}_2$ is the diagonal matrix form of $\boldsymbol{\varphi}_2$ that is the second row vector of the Hadamard matrix given by (4.58).

By substituting the received vector given by (4.54) in a MMSE predictor given by (2.42), then the estimated coefficients of the channel between between the kth UT and the ith BS antenna at the ℓth symbol time interval are given by:

$$\hat{h}_{i,k}(\ell) = \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_{p,i}} \mathbf{R}^{-1}_{\check{\mathbf{y}}_{p,i}\check{\mathbf{y}}_{p,i}} \check{\mathbf{y}}_{p,i}, \tag{4.60}$$

where

$$\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_{p,i}} = \frac{1}{2} x_p^* \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \check{\Phi}_k^H, \tag{4.61}$$

and

$$\mathbf{R}_{\check{\mathbf{y}}_{p,i}\check{\mathbf{y}}_{p,i}} = \sum_{j=1}^{K} \frac{1}{4} |x_p|^2 \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}_j^H + \sigma_w^2 \mathbf{I}_{L_p}, \tag{4.62}$$

where $\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}$ is the channel autocorrelation matrix, whose entries $(\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}})_{m,\ell}$ are given by $\overline{\alpha^2} J_0 \left[2\pi f_{d,\max} \frac{T_s}{(1+\eta)} (\ell - m) \right]$ for $\ell, m = 1, 2, \cdots, L_p$. Moreover, $\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}$ of dimension $1 \times L_p$ is the channel correlation vector at the ℓth symbol time interval, which enables the prediction of the channel coefficients for $\ell \geqslant L_p + 1$. Notice that the MMSE estimator performs the channel coefficients prediction by using the pilot symbols transmitted during the first L_p time slots of the block.

The vector $\check{\mathbf{y}}_{p,i}$ is a complex Gaussian random variable. Therefore, the estimated channel coefficient $\hat{h}_{i,k}(\ell)$ has also complex Gaussian distribution with mean $\mathrm{E}\{\hat{h}_{i,k}(\ell)\} = 0$ and variance:

$$\begin{aligned} \varrho_{i,k}(\ell) &= \mathrm{E}\left\{ \hat{h}_{i,k}(\ell) \hat{h}^*_{i,k}(\ell) \right\} \\ &= \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_{p,i}} \mathbf{R}^{-1}_{\check{\mathbf{y}}_{p,i}\check{\mathbf{y}}_{p,i}} \mathbf{R}^H_{h_{i,k}(\ell)\check{\mathbf{y}}_{p,i}}, \end{aligned} \tag{4.63}$$

where $\hat{h}_{i,k}(\ell)$ is given by (4.60). By using (4.61) and (4.62) in (4.63), the variance can be

rewritten as:

$$\varrho_{i,k}(\ell) = \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \check{\Phi}_k^H \mathcal{M}^{-1} \check{\Phi}_k \mathbf{R}^H_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}, \tag{4.64}$$

where:

$$\begin{aligned}
\mathcal{M} &= \frac{4}{|x_p|^2} \mathbf{R}_{\check{\mathbf{y}}_{p,i}\check{\mathbf{y}}_{p,i}} \\
&= \sum_{j=1}^{K} \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}_j^H + \frac{4\sigma_w^2}{|x_p|^2} \mathbf{I}_{L_p} \\
&= \sum_{j=1}^{K} \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}_j^H + \frac{N_0}{E_p} \mathbf{I}_{L_p} \\
&= \sum_{j=1}^{K} \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}_j^H + \frac{1+\mu\eta}{\mu \frac{E_b}{N_0} \log_2(M)} \mathbf{I}_{L_p},
\end{aligned} \tag{4.65}$$

E_p for MPE, is given by (4.31), η is given by (4.27) and μ is given by (4.30). Notice that $\mathcal{M}$ is a symmetric Toeplitz matrix and its inverse is a bisymmetric matrix [102]. A bisymmetric matrix presents symmetries in relation to the main and secondary diagonals.

Additionally, due to the uncorrelated antennas assumption, all the estimated channel coefficients have the same statistics at each antenna. Thus, the variance obtained in (4.64) is the same for any channel coefficient, i.e., $\varrho_{i,j}(\ell) = \varrho(\ell)$ for $1 \leqslant i \leqslant N$ and $1 \leqslant j \leqslant K$.

Furthermore, by assuming that $L_p = K$ the variance can be approximated by:

$$\varrho(\ell) \leqslant \frac{\overline{\alpha^2}^2 \sum_{j=1}^{K} J_0^2 \left[2\pi f_{d,\max} T_s(\ell - j)\right]}{\overline{\alpha^2} K + \frac{(1+\mu\eta)}{\mu E_b/N_0 \log_2(M)}}, \tag{4.66}$$

where $\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \check{\Phi}_k^H \check{\Phi}_k \mathbf{R}^H_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \leqslant \overline{\alpha^2}^2 \sum_{j=1}^{K} J_0^2 \left[2\pi f_{d,\max} T_s(\ell - j)\right] \mathbf{I}_K$ and $\sum_{j=1}^{K} \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}_j^H = \overline{\alpha^2} K \mathbf{I}_K$. Once $L_p = K$ the spectral efficiency is maximized because the minimum number of pilot symbols for reliable estimation is used. Besides, (4.66) is a simple scalar expression where the matrix operations are not necessary in order to obtain the channel variance.

On the other hand, for $L_p > K$ there is no closed-form expression for $\mathcal{M}^{-1}$ and therefore for (4.64).

4.4.1.2 CSI for SPE

The matrix of received pilot and data samples at the matched filter output is also given by (4.14). Therefore, the vector of superimposed pilot and data symbols at the ith

antenna, given by (4.16), can be rewritten as:

$$\check{\mathbf{y}}_i = \sum_{j=1}^{K} \frac{1}{2} \left(\check{\mathbf{X}}_{p,j} + \check{\mathbf{X}}_{d,j} \right) \check{\mathbf{h}}_{i,j} + \check{\mathbf{w}}_i, \tag{4.67}$$

where $\check{\mathbf{X}}_{d,j} = \text{diag}(\mathbf{x}_{d,j})$ represents the transmitted data symbols that introduce interference during channel estimation. Notice that the definitions are similar to MPE but with different dimensions. Thus, $\check{\mathbf{y}}_i$, $\check{\mathbf{h}}_{i,j}$ and $\check{\mathbf{w}}_i$ are $L \times 1$ vectors. Besides, the pilot symbols matrix is $\check{\mathbf{X}}_{p,j} = x_p \check{\Phi}_j$ where $\check{\Phi}_j = \text{diag}(\boldsymbol{\varphi}_j)$ and $\boldsymbol{\varphi}_j$ is the jth row vector of the $K \times L$ matrix Φ.

The estimated channel coefficients can be obtained by substituting the vector of received samples, given by (4.67), in the MMSE smoothing filter, given by (2.40). Therefore, the estimated channel coefficient for the link between the the kth UT and the ith BS antenna for the ℓth symbol time interval is given by:

$$\hat{h}_{i,k}(\ell) = \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_i} \mathbf{R}^{-1}_{\check{\mathbf{y}}_i\check{\mathbf{y}}_i} \check{\mathbf{y}}_i, \tag{4.68}$$

where:

$$\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_i} = \frac{1}{2}(x_p + x_d)^* \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \check{\Phi}^H_k, \tag{4.69}$$

and

$$\mathbf{R}_{\check{\mathbf{y}}_i\check{\mathbf{y}}_i} = \sum_{j=1}^{K} \frac{1}{4} |x_p|^2 \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}^H_j + \sum_{j=i}^{K} \overline{\alpha^2}\ \frac{1}{4} \overline{|x_d|^2} \mathbf{I}_L + \sigma^2_w \mathbf{I}_L. \tag{4.70}$$

where $\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}$ is the $1 \times L$ channel correlation vector. Notice that $\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}$ is the channel autocorrelation matrix with dimension $L \times L$, given by (4.17) and $\sum_{j=i}^{K} \overline{\alpha^2}\ \overline{|x_d|^2} \mathbf{I}_L/4$ is the interference introduced by the superimposed data symbols.

The estimated channel coefficients are complex Gaussian random variables with mean $\text{E}\{\hat{h}_{i,k}(\ell)\} = 0$ and variance:

$$\begin{aligned} \varrho_{i,k}(\ell) &= \text{E}\left\{ \hat{h}_{i,k}(\ell) \hat{h}^*_{i,k}(\ell) \right\} \\ &= \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_i} \mathbf{R}^{-1}_{\check{\mathbf{y}}_i\check{\mathbf{y}}_i} \mathbf{R}^H_{h_{i,k}(\ell)\check{\mathbf{y}}_i} \end{aligned} \tag{4.71}$$

where $\hat{h}_{i,k}(\ell)$ is given by (4.68). By using (4.69) and (4.70) in (4.71), the variance can be rewritten as:

$$\varrho_{i,k}(\ell) = \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \check{\Phi}^H_k \mathcal{S}^{-1} \check{\Phi}_k \mathbf{R}^H_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}, \tag{4.72}$$

where:

$$\begin{aligned}\mathcal{S} &= \frac{4}{|x_p|^2}\mathbf{R}_{\check{\mathbf{y}}_i\check{\mathbf{y}}_i} \\ &= \sum_{j=1}^{K}\check{\Phi}_j\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}\check{\Phi}_j^H + \overline{\alpha^2}K\,\frac{\overline{|x_d|^2}}{|x_p|^2}\mathbf{I}_L + \frac{4\sigma_w^2}{|x_p|^2}\mathbf{I}_L \\ &= \sum_{j=1}^{K}\check{\Phi}_j\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}\check{\Phi}_j^H + \overline{\alpha^2}K\frac{E_d}{E_p}\mathbf{I}_L + \frac{N_0}{E_p}\mathbf{I}_L \\ &= \sum_{j=1}^{K}\check{\Phi}_j\mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}}\check{\Phi}_j^H + \overline{\alpha^2}\frac{K}{\mu_s}\mathbf{I}_L + \frac{(1+\mu_s)\mathbf{I}_L}{\mu_s\frac{E_b}{N_0}\log_2(M)},\end{aligned} \tag{4.73}$$

E_p and E_d are given by (4.40) and (4.41), respectively, for SPE, and μ_s is given by (4.39). The matrix $\mathcal{S}$ is also a symmetric Toeplitz and its inverse is bisymmetric. Additionally, $\varrho_{i,j}(\ell) = \varrho(\ell),\ \forall i,j$.

Furthermore, for $L = K$, the variance given by (4.72) can be approximated by:

$$\varrho(\ell) \leqslant \frac{\overline{\alpha^2}^2\sum_{j=1}^{K}J_0^2\left[2\pi f_{d,\max}T_s(\ell-j)\right]}{\overline{\alpha^2}K + \overline{\alpha^2}\frac{K}{\mu_s} + \frac{(1+\mu_s)}{\mu_s E_b/N_0\log_2(M)}}, \tag{4.74}$$

where the same considerations for obtaining (4.66) were made. Notice that the expression is simple, but it is restricted to short blocks.

4.4.1.3 CSI for HPE

From (4.14), the received pilot samples matrix for HPE is:

$$\mathbf{Y}_s = [\mathbf{y}(1)\ \cdots\mathbf{y}(\ell)\ \cdots\ \mathbf{y}(L_s)], \tag{4.75}$$

where L_s is the number of superimposed symbols.

Pilot symbols are superposed to data symbols, then, from (4.16) the vector of length L_s of received superimposed samples at the ith antenna, can be rewritten as:

$$\check{\mathbf{y}}_{s,i} = \sum_{j=1}^{K}\frac{1}{2}\left(\check{\mathbf{X}}_{p,j} + \check{\mathbf{X}}_{d,j}\right)\check{\mathbf{h}}_{i,j} + \check{\mathbf{w}}_i, \tag{4.76}$$

where $\check{\mathbf{h}}_{i,j}$ is the channel vector and $\check{\mathbf{w}}_i$ is the noise vector, both of dimension $L_s \times 1$. The pilot symbols matrix is equal to $\check{\mathbf{X}}_{p,j} = x_p\check{\Phi}_j$ where $\check{\Phi}_j = \mathrm{diag}(\boldsymbol{\varphi}_j)$ and $\boldsymbol{\varphi}_j$ is the jth row vector of the $K \times L_s$ matrix Φ.

Using (4.76) on the MMSE estimator, the estimated channel coefficient for the link

between the kth UT and the ith BS antenna for the ℓth time interval is given by:

$$\hat{h}_{i,k}(\ell) = \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{y}}_{s,i}} \mathbf{R}^{-1}_{\check{\mathbf{y}}_{s,i}\check{\mathbf{y}}_{s,i}} \check{\mathbf{y}}_{s,i}, \tag{4.77}$$

where the autocorrelation matrix and the correlation vector can be obtained using the same procedure of MPE. Notice that (4.77) estimates the channel for $1 \leqslant \ell \leqslant L_p$ and predicts the channel for $L_p + 1 \leqslant \ell \leqslant L$.

The estimated channel coefficient $\hat{h}_{i,k}(\ell)$ has complex Gaussian distribution with mean $\mathrm{E}\{\hat{h}_{i,k}(\ell)\} = 0$ and variance:

$$\begin{aligned} \varrho_{i,k}(\ell) &= \mathrm{E}\{\hat{h}_{i,k}(\ell)\hat{h}^*_{i,k}(\ell)\} \\ &= \mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}} \check{\Phi}_k^H \mathcal{H}^{-1} \check{\Phi}_k \mathbf{R}^H_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}, \end{aligned} \tag{4.78}$$

where

$$\mathcal{H} = \sum_{j=1}^{K} \check{\Phi}_j \mathbf{R}_{\check{\mathbf{h}}_{i,j}\check{\mathbf{h}}_{i,j}} \check{\Phi}_j^H + \overline{\alpha^2} \frac{K}{\mu_{h,s}} \mathbf{I}_{L_s} + \frac{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}{\mu_{h,s}\mu_{h,m}(1+\eta_h)\frac{E_b}{N_0}\log_2(M)} \mathbf{I}_{L_s}, \tag{4.79}$$

where η_h, $\mu_{h,m}$ and $\mu_{h,s}$ are given by (4.47), (4.48) and (4.49), respectively. Notice that $\mathcal{H}$, given by (4.79), can be obtained in a similar form of $\mathcal{S}$, given by (4.73), by assuming that $L = L_s$, that E_p is given by (4.50) and that E_d is given by (4.51). Finally, due to the uncorrelated antennas assumption, $\varrho_{i,j}(\ell) = \varrho(\ell)$..

Additionally, for $L_s = K$ the variance given by (4.78) can be approximated to:

$$\varrho_{i,k}(\ell) \leqslant \frac{\overline{\alpha^2}^2 \sum_{j=1}^{K} J_0^2\left[2\pi f_{d,\max} T_s(\ell - j)\right]}{\overline{\alpha^2}K + \overline{\alpha^2}\frac{K}{\mu_{h,s}} + \frac{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}{\mu_{h,s}\mu_{h,m}(1+\eta_h)E_b/N_0\log_2(M)}}, \tag{4.80}$$

where the same considerations of (4.66) were assumed.

4.4.2 Channel Coefficients Estimation in TICs

In this section, the channel coefficients are estimated/predicted for each estimation technique assuming TICs. Moreover, as the channel is time-invariant, only one channel coefficient is estimated for each block. Thus, the CSI for TICs is simpler than for TVCs.

4.4.2.1 CSI for MPE

The matrix and vector of pilot samples are similar to the TVC case, given by (4.53) and (4.54), respectively. However, for TICs, the vector of received pilot samples can be rewritten as:

$$\check{\mathbf{y}}_{p,i} = \sum_{j=1}^{K} \frac{1}{2} \check{\mathbf{X}}_{p,j} \mathbf{1} h_{i,j} + \mathbf{w}_i, \tag{4.81}$$

where $\mathbf{1}$ is an all-ones vector of dimension $L_p \times 1$, $h_{i,j}$ is the channel coefficient and $\mathbf{w}_i = [w_i(1)\ w_i(2)\ \cdots\ w_i(L_p)]^T$ is the noise vector. Additionally, $\check{\mathbf{X}}_{p,j} = \text{diag}(\mathbf{x}_{p,j})$ is the diagonal matrix that contains the pilot symbols transmitted by the jth UT.

By applying the received vector in (4.81) to the MMSE estimator given by (2.40), the estimated channel coefficient for the link between the kth UT and the ith BS antenna is given by:

$$\begin{aligned}\hat{h}_{i,k} &= \mathbf{R}_{h_{i,k}\check{\mathbf{y}}_{p,i}}\mathbf{R}^{-1}_{\check{\mathbf{y}}_{p,i}\check{\mathbf{y}}_{p,i}}\check{\mathbf{y}}_{p,i} \\ &= 2\overline{\alpha^2}\mathbf{x}^H_{p,k}\left(\sum_{j=1}^{K}\overline{\alpha^2}\mathbf{x}_{p,j}\mathbf{x}^H_{p,j} + 4\sigma_w^2\mathbf{I}_{L_p}\right)^{-1}\check{\mathbf{y}}_{p,i},\end{aligned} \tag{4.82}$$

where $\mathbf{R}_{h_{i,k}\check{\mathbf{y}}_{p,i}} = \frac{1}{2}\overline{\alpha^2}\mathbf{x}^H_{p,k}$, $\mathbf{R}_{\check{\mathbf{y}}_{p,i,j}\check{\mathbf{y}}_{p,i,j}} = \sum_{j=1}^{K}\frac{1}{4}\overline{\alpha^2}\mathbf{x}_{p,j}\mathbf{x}^H_{p,j} + \sigma_w^2\mathbf{I}_{L_p}$ and $\mathbf{x}_{p,j}$ of dimension $1 \times L_p$ is the vector of pilot symbols transmitted by the jth UT. Notice that there is no dependence on time and therefore this expression is simpler than that for TVC.

Furthermore, the estimated channel coefficient is a complex Gaussian random variable, due to the channel coefficient and the noise be complex Gaussian random variables. Therefore, the mean of the estimated channel coefficient is $\text{E}\{\hat{h}_{i,k}\} = 0$ and the variance is given by:

$$\begin{aligned}\varrho_{i,k} &= \text{E}\{\hat{h}_{i,k}\hat{h}^*_{i,k}\} \\ &= \overline{\alpha^2}^2\mathbf{x}_{p,k}\left(\sum_{j=1}^{K}\overline{\alpha^2}\mathbf{x}^H_{p,j}\mathbf{x}_{p,j} + 4\sigma_w^2\mathbf{I}_{L_p}\right)^{-1}\mathbf{x}^H_{p,k}.\end{aligned} \tag{4.83}$$

By applying the inversion lemma given by (A.14) and considering the orthogonality between the pilot sequences, the variance given by (4.83) can be written as a function of the average symbol as:

$$\varrho = \overline{\alpha^2}^2 L_p\left(\overline{\alpha^2}L_p + \frac{1+\mu\eta}{\mu\frac{E_b}{N_0}\log_2(M)}\right)^{-1}, \tag{4.84}$$

where η and μ are given by (4.27) and (4.30), respectively. Besides, the pilot symbol energy given by (4.31) was employed. Additionally, due to the uncorrelated antennas assumption, the variance obtained in (4.84) is the same for any estimated channel coefficient, i.e., $\varrho_{i,j} = \varrho,\ \forall i, j$.

4.4.2.2 CSI for SPE

The vector of received pilot samples given by (4.67) for TVCs, can be rewritten for TICs as:

$$\check{\mathbf{y}}_i = \sum_{j=1}^{K} \frac{1}{2} \left(\check{\mathbf{X}}_{p,j} + \check{\mathbf{X}}_{d,j} \right) \mathbf{1} h_{i,j} + \check{\mathbf{w}}_i, \tag{4.85}$$

where $\check{\mathbf{X}}_{d,j}$ are data symbols. The definitions are similar to MPE, but with different dimension. Therefore, the vector $\mathbf{1}$ has dimension of $L \times 1$ and $\check{\mathbf{X}}_{p,j} = \text{diag}(\mathbf{x}_{p,j})$ of dimension $L \times L$ represents the transmitted pilot sequences.

By applying the MMSE smoothing filter given by (2.40) to the received vector given by (4.85), the estimated channel coefficient for the link between the kth UT and the ith BS antenna is given by:

$$\begin{aligned} \hat{h}_{i,k} &= \mathbf{R}_{h_{i,k}\check{\mathbf{y}}_i} \mathbf{R}^{-1}_{\check{\mathbf{y}}_i\check{\mathbf{y}}_i} \check{\mathbf{y}}_i \\ &= 2\overline{\alpha^2} \mathbf{x}_{p,k} \left(\sum_{j=1}^{K} \overline{\alpha^2} \mathbf{x}^H_{p,j} \mathbf{x}_{p,j} + \overline{\alpha^2} K \overline{|x_d|^2} \mathbf{I}_L + 4\sigma_w^2 \mathbf{I}_L \right)^{-1} \check{\mathbf{y}}_i, \end{aligned} \tag{4.86}$$

where $\mathbf{R}_{h_{i,k}\check{\mathbf{y}}_i} = \overline{\alpha^2} \mathbf{x}_{p,k}$ and $\mathbf{R}_{\check{\mathbf{y}}_i\check{\mathbf{y}}_i} = \sum_{j=1}^{K} \overline{\alpha^2} \mathbf{x}^H_{p,j} \mathbf{x}_{p,j} + \overline{\alpha^2} K \overline{|x_d|^2} \mathbf{I}_L + 4\sigma_w^2 \mathbf{I}_L$.

Furthermore, the estimated channel coefficient $\hat{h}_{i,k}$ is a complex Gaussian random variable with mean $\text{E}\{\hat{h}_{i,k}\} = 0$ and variance given by:

$$\begin{aligned} \varrho_{i,k} &= \text{E}\{\hat{h}_{i,k} \hat{h}^*_{i,k}\} \\ &= \overline{\alpha^2}^2 \mathbf{x}_{p,k} \left(\sum_{j=1}^{K} \overline{\alpha^2} \mathbf{x}^H_{p,j} \mathbf{x}_{p,j} + \overline{\alpha^2} K \overline{|x_d|^2} \mathbf{I}_L + 4\sigma_w^2 \mathbf{I}_L \right)^{-1} \mathbf{x}^H_{p,k}, \end{aligned} \tag{4.87}$$

As for MPE, by applying the inversion lemma in (A.14) and considering the orthogonality among pilot sequences, the variance in (4.87) can be rewritten as:

$$\varrho = \overline{\alpha^2}^2 L \left(\overline{\alpha^2} L + \overline{\alpha^2} \frac{K}{\mu_s} + \frac{1 + \mu_s}{\mu_s \frac{E_b}{N_0} \log_2(M)} \right)^{-1}, \tag{4.88}$$

where μ_s is given by (4.39), the pilot symbols energy is given by (4.40) and the data symbol energy is given by (4.41).

4.4.2.3 CSI for HPE

The vector of received pilot samples given by (4.76) can be rewritten for TICs as:

$$\check{\mathbf{y}}_{s,i} = \sum_{j=1}^{K} \frac{1}{2} \left(\check{\mathbf{X}}_{p,j} + \check{\mathbf{X}}_{d,j} \right) \mathbf{1} h_{i,j} + \check{\mathbf{w}}_i. \tag{4.89}$$

where $\check{\mathbf{h}}_{i,j} = \mathbf{1}h_{i,j}$, $\mathbf{1}$ is an all-ones vector of dimension $L_s \times 1$, $h_{i,j}$ is the channel coefficient and $\check{\mathbf{w}}_i$ is the noise vector. The matrix $\check{\mathbf{X}}_{p,j} = \text{diag}(\mathbf{x}_{p,j})$ of dimension $L_s \times L_s$ contains the vector of transmitted pilot symbols. The received vector is similar to (4.85), but with different dimension. Thus, by applying MMSE channel estimation on the received vector, the estimated coefficients are similar to that of SPE, that is:

$$\begin{aligned}\hat{h}_{i,k} &= \mathbf{R}_{h_{i,k}\check{\mathbf{y}}_{s,i}}\mathbf{R}^{-1}_{\check{\mathbf{y}}_{s,i}\check{\mathbf{y}}_{s,i}}\check{\mathbf{y}}_{s,i} \\ &= 2\overline{\alpha^2}\mathbf{x}_{p,k}\left(\sum_{j=1}^{K}\overline{\alpha^2}\mathbf{x}^H_{p,j}\mathbf{x}_{p,j} + \overline{\alpha^2}K\overline{|x_d|^2}\mathbf{I}_{L_s} + 4\sigma_w^2\mathbf{I}_{L_s}\right)^{-1}\check{\mathbf{y}}_i. \end{aligned} \tag{4.90}$$

As the estimated channel coefficient is a complex Gaussian random variable, the mean is given by $\text{E}\{\hat{h}_{i,k}\} = 0$ and the variance is given by:

$$\begin{aligned}\varrho &= \text{E}\{\hat{h}_{i,k}\hat{h}^*_{i,k}\} \\ &= \overline{\alpha^2}^2 L_s\left(\overline{\alpha^2}L_s + \overline{\alpha^2}\frac{K}{\mu_{h,s}} + \frac{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}{\mu_{h,s}\mu_{h,m}(1+\eta_h)\frac{E_b}{N_0}\log_2(M)}\right)^{-1}, \end{aligned} \tag{4.91}$$

where η_h, $\mu_{h,m}$ and $\mu_{h,s}$ are given by (4.47), (4.48) and (4.49), respectively. Besides the pilot symbol energy given by (4.50), the superimposed data symbol energy given by (4.51) and the multiplexed data symbol energy given by (4.52) were employed to derive (4.91).

4.5 Channel Estimation Error

The channel estimation process is not perfect and errors are introduced. Therefore, the channel coefficient at the ℓth time interval is given by:

$$h_{i,k}(\ell) = \hat{h}_{i,k}(\ell) + \tilde{h}_{i,k}(\ell), \tag{4.92}$$

where $\hat{h}_{i,k}(\ell)$ is the estimated channel coefficient and $\tilde{h}_{i,k}(\ell)$ is the channel estimation error introduced by the channel estimation process. Moreover, the channel matrix can be rewritten as $\mathbf{H}(\ell) = \widehat{\mathbf{H}}(\ell) + \widetilde{\mathbf{H}}(\ell)$ or the kth channel vector as $\mathbf{h}_k(\ell) = \widehat{\mathbf{h}}_k(\ell) + \widetilde{\mathbf{h}}_k(\ell)$, whose entries are given by (4.92).

According to the MMSE estimator principles, $\hat{h}_{i,k}(\ell)$ and $\tilde{h}_{i,k}(\ell)$ are orthogonal random variables. Therefore, they are independent and uncorrelated and have the same distribution of $h_{i,k}(\ell)$. Furthermore, due to the orthogonality principle the MSE can be minimized. Once $h_{i,k}(\ell)$ is a complex Gaussian random variable, $\tilde{h}_{i,k}(\ell)$ has same distribution with

mean $\mathrm{E}\{\tilde{h}_{i,k}(\ell)\} = 0$ and variance:

$$\begin{aligned}\tilde{\varrho}_{i,j}(\ell) &= \mathrm{E}\{\tilde{h}_{i,j}(\ell)\tilde{h}^{*}_{i,j}(\ell)\} \\ &= \overline{\alpha^2} - \varrho(\ell),\end{aligned} \tag{4.93}$$

where $\varrho(\ell)$ is the variance of the estimated channel, which depends on the channel estimation technique. Besides, the variance of the channel estimation error is $\tilde{\varrho}_{i,j}(\ell) = \tilde{\varrho}(\ell)\ \forall i, j$.

Notice that the variance of the estimated channel is bounded by $0 < \varrho(\ell) \leqslant \overline{\alpha^2}$, where the upper bound is the perfect channel estimation and the lower bound is the completely imperfect channel estimation.

A similar analysis with resembling results can be made for TICs, without the temporal variable ℓ, i.e., $h_{i,k} = \hat{h}_{i,k} + \tilde{h}_{i,k}$, where the channel estimation error has complex Gaussian distribution with mean $\mathrm{E}\{\tilde{h}_{i,k}\} = 0$ and variance:

$$\begin{aligned}\tilde{\varrho}_{i,j} &= \mathrm{E}\{\tilde{h}_{i,j}\tilde{h}^{*}_{i,j}\} \\ &= \overline{\alpha^2} - \varrho,\end{aligned} \tag{4.94}$$

Finally, the channel estimation quality is measured by the variance of the channel estimation error given by (4.93) for TVCs and (4.94) for TICs, that is equivalent to the MSE [103]. Furthermore, by normalizing the MSE by $\overline{\alpha^2}$, the estimation quality is measured by the NMSE.

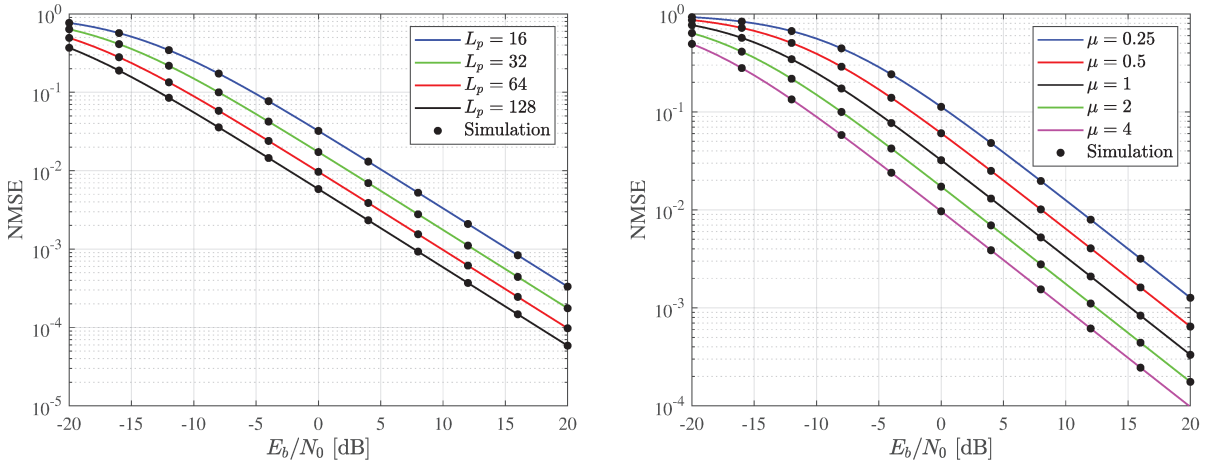
4.6 Numerical Results and Discussion

In this section, the channel estimation techniques are evaluated in terms of the NMSE, using the derived expressions. Additionally, for some cases, simulations employing the Monte Carlo method are performed in order to validate our expressions. The simulation parameters are detailed in Tab. 4.1, which have been chosen based on the long term evolution (LTE) and the proposals of the 5G standard [81].

Table 4.1: Simulation parameters.

Parameter	Value
Bandwidth - B	15 kHz
Carrier Frequency - f_c	3.5 GHz
Fading Mean Power - $\overline{\alpha^2}$	1
UT Speed - v	6 km/h and 60 km/h

As shown in Tab. 4.1, once the fading mean power is normalized, the NMSE is the metric employed for the estimation quality in the next subsections. Additionally, as the NMSE expressions are a function of the modulation order, $M = 4$ is used in this chapter.



(a) $\mu = 1$. (b) $L_p = 16$.

Figure 4.8: NMSE as a function of E_b/N_0 for MPE, $L_d = 256$.

4.6.1 NMSE for TICs

In this subsection, the NMSE for different channel estimation techniques on TICs is evaluated. The effects of varying some parameters are analyzed, such as the pilot sequence length, the ratio between the energy of pilot and data symbols and the number of UTs. For this purpose, the expressions obtained in Sections 4.4.2 and 4.5 are used.

4.6.1.1 MPE

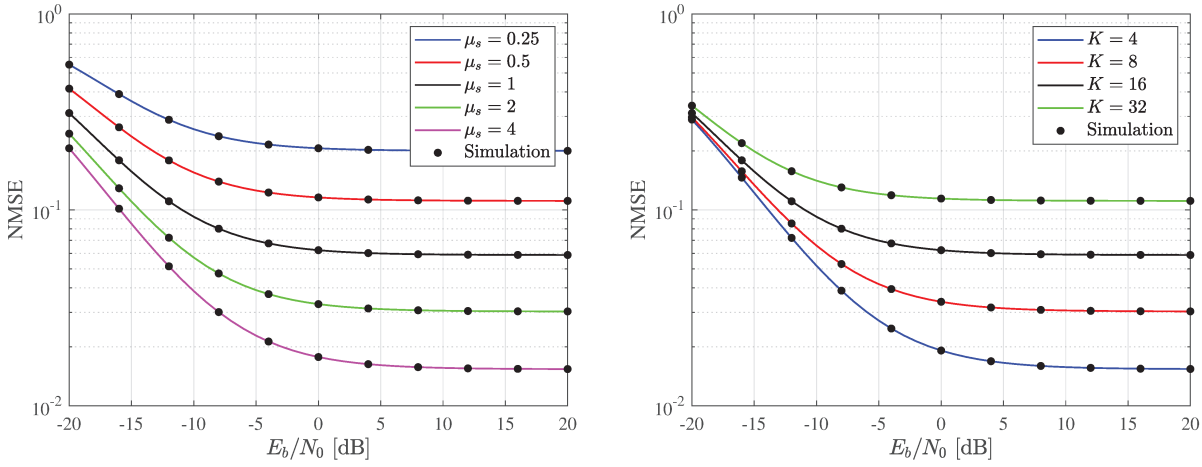
For evaluating the NMSE, the variance of the estimated channel of MPE given by (4.83) is used in (4.94). Therefore, Fig. 4.8a shows the NMSE as a function of E_b/N_0 parameterized by L_p, for $L_d = 256$, $K = 16$ and $\mu = 1$. Notice that as the number of pilot symbols increases the NMSE decreases at the cost of some loss in spectral efficiency. By using (4.33), the spectral efficiency of $\varepsilon = 30.1$ bits/s/Hz for $L_p = 16$ is reduced to $\varepsilon = 21.3$ bits/s/Hz for $L_p = 128$.

Fig. 4.8b shows the NMSE as a function of E_b/N_0 parameterized by μ, for $L_d = 256$, $K = 16$ and $L_p = 16$. Notice that the NMSE can be reduced by increasing the power of pilot symbols for $\mu > 1$. However, increasing the power of pilot symbols it reduces the power of data symbols, which can result in a poorer BER performance, as shown in the next chapter.

The effects of K and L on the NMSE will be presented later, in figures that compare the different estimation techniques and scenarios.

4.6.1.2 SPE

For evaluating the NMSE of SPE, the estimated channel variance given by (4.87) is used in (4.94). Fig. 4.9a shows the NMSE as a function of E_b/N_0 parameterized by μ_s, for $L_d = 256$ symbols and $K = 16$. Notice that SPE introduces a floor in the NMSE



(a) $K = 16$. (b) $\mu_s = 1$.

Figure 4.9: NMSE as a function of E_b/N_0 for SPE, $L_d = 256$.

due to the interference between pilot and data symbols, bringing on loss in performance. However, in this case, the spectral efficiency is maximized, that is $\varepsilon = 32$ bits/s/Hz. Additionally, pilot symbols with higher power than that of data symbols are desirable for a smaller NMSE. However, as for MPE, using too much power on pilot symbols can cause problems on the BER performance once the power of data symbols is reduced.

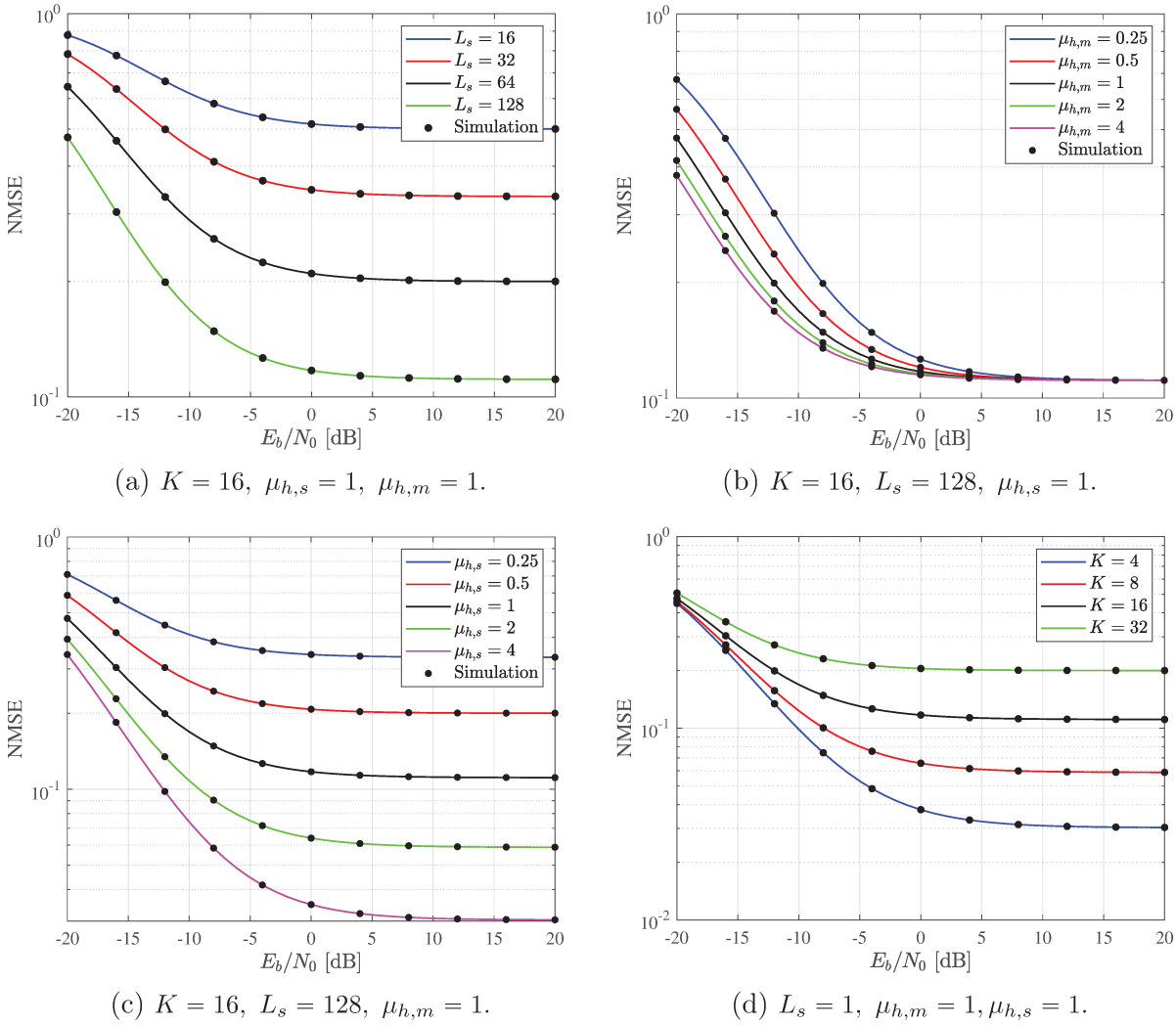
Fig. 4.9b shows the NMSE as a function of E_b/N_0 parameterized by K, for $L_d = 256$ symbols and $\mu_s = 1$. Notice that the pilot interference between UTs is eliminated by the orthogonality among pilot sequences. However, the interference from data symbols is not eliminated. Therefore, the NMSE has become worsen as the number of UTs increases.

4.6.1.3 HPE

NMSE of HPE can be evaluated using the estimated channel variance given by (4.91) in (4.94).

Fig. 4.10a shows the NMSE as a function of E_b/N_0 parameterized by L_s, for $L_d = 256$, $K = 16$, $\mu_{h,m} = 1$ and $\mu_{h,s} = 1$. As for SPE, a floor on the NMSE is introduced due to the interference between pilot symbols and data symbols in the superimposed region. Notice that increasing the number of superimposed symbols reduces the NMSE floor. However, due to the pilot symbols orthogonality restriction, the maximum number of superimposed symbols is $L_s = 128$.

Fig. 4.10b shows the NMSE as a function of E_b/N_0 parameterized by $\mu_{h,m}$, for $L_d = 256$, $L_s = 128$, $K = 16$ UTs and $\mu_{h,s} = 1$. Kept L_s fixed, the NMSE floor remains the same for any $\mu_{h,m}$. However, the NMSE is improved in the low E_b/N_0 region as $\mu_{h,m}$ is increased. By increasing $\mu_{h,m}$, the energy of the multiplexed data symbols is reduced and the energy of the superimposed symbols for both pilot and data symbols is increased. Notice that, due to the increment of energy of superimposed data symbols, the NMSE improvement is limited.



(a) $K = 16,\ \mu_{h,s} = 1,\ \mu_{h,m} = 1.$

(b) $K = 16,\ L_s = 128,\ \mu_{h,s} = 1.$

(c) $K = 16,\ L_s = 128,\ \mu_{h,m} = 1.$

(d) $L_s = 1,\ \mu_{h,m} = 1, \mu_{h,s} = 1.$

Figure 4.10: NMSE as a function of E_b/N_0 for HPE, $L_d = 256$.

Fig. 4.10c shows the NMSE as a function of E_b/N_0 parameterized by $\mu_{h,s}$, for $L_d = 256$, $L_s = 128$, $K = 16$ and $\mu_{h,m} = 1$. Notice that the NMSE is improved as $\mu_{h,s}$ increases. For $\mu_{h,m} = 1$ the energy of the multiplexed data symbols is equal to the energy of superimposed symbols. However, as $\mu_{h,s}$ increases, the energy of pilot symbols becomes higher than energy of data symbols in the superimposed region. Thus, the interference of data symbols is reduced and the NMSE is improved. Therefore, it is recommendable to increase $\mu_{h,s}$ instead $\mu_{h,m}$ for lowering the NMSE. However, as for MPE, too much energy on pilot symbols may degrade the BER performance.

Fig. 4.10d shows the NMSE as a function of E_b/N_0 parameterized by K, for $L_d = 256$, $L_s = 128$, $\mu_{h,s} = 1$ and $\mu_{h,m} = 1$. As expected, increasing K results in more interference raising the floor in the NMSE.

Fig. 4.11 shows a comparison between the NMSE for all three estimation techniques for $K = 4$ and 16 UTs transmitting the same number of data symbols $L_d = 256$ with same energy ratios $\mu = \mu_s = \mu_{h,s}\mu_{h,m} = 1$. Additionally, the number of pilot symbols is $L_p = 16$ for MPE, $L_p = L = 256$ for SPE and $L_p = L_s = 128$ for HPE. As E_b/N_0 is increased, the

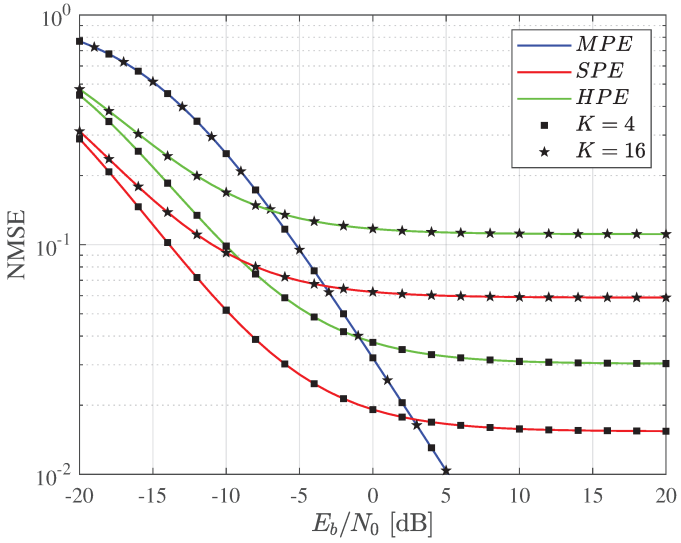


Figure 4.11: NMSE as a function of E_b/N_0 parameterized by the number of UTs K for $L_d = 256$, $\mu = \mu_s = \mu_{h,s}\mu_{h,m} = 1$, $L_p = 16$ for MPE, $L_p = 256$ for SPE and $L_p = 128$ for HPE.

best NMSE performance is obtained by MPE at a cost in spectral efficiency. Notice that for MPE, the NMSE is the same for $K = 4$ and $K = 16$. Thus, the NMSE depends on L_p, but not on K, once the pilot sequences orthogonality eliminates the interference during channel estimation. Furthermore, MPE's results are opossed to the results of SPE and HPE, because the NMSE does not present a floor. Observe that both techniques, SPE and HPE, presents a floor in the NMSE due to the interference between pilot and data symbols. Moreover, their NMSE performance is reduced as K increases. Finally, once $L_s < L$ the NMSE of SPE outperforms HPE on NMSE. For simplicity, in Fig. 4.11 only the theoretical results are plotted.

4.6.2 NMSE for TVCs

In this subsection, the NMSE is obtained for TVCs for the all three estimation techniques. Two scenarios are analyzed: the first for a pedestrian speed of 6 km/h and the second for a vehicular speed of 60 km/h. The effects on the NMSE of the number of UTs, the number of pilot symbols, and the energy ratio between pilot and data symbols are evaluated. For this purpose, the estimated channel variance for MPE, SPE and HPE obtained in Section 4.4.1 are used in (4.93). Furthermore, the generation of time-correlated channel is time-demanding, thus, the Monte Carlo simulation is only presented in few figures to validate the theoretical results.

4.6.2.1 Pedestrian Velocity

For the pedestrian speed of $v = 6$ km/h, the channel coherence time interval is $T_c = 26$ ms and the symbol bandwidth for SPE and HPE is $B = 15$ kHz. Thus, the number of time slots during this time interval is $L = 390$. Once $L = L_p$ for SPE and due to the

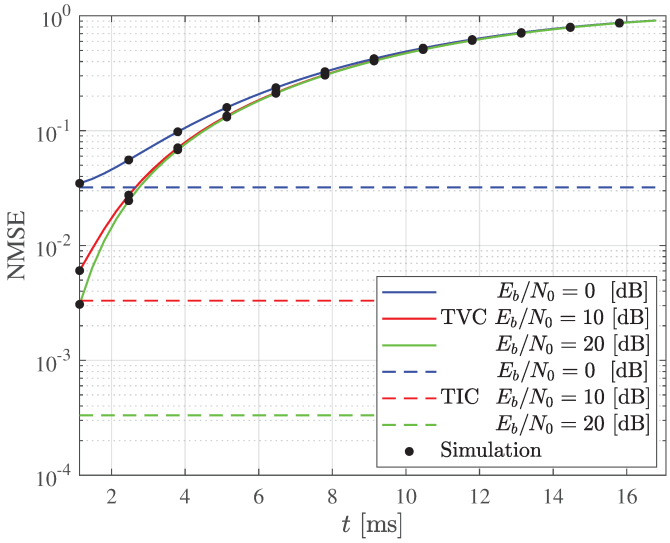


Figure 4.12: NMSE as a function of time for MPE, $L_d = 256$, $L_p = 16$, $\mu = 1$ and $K = 16$, parameterized by E_b/N_0.

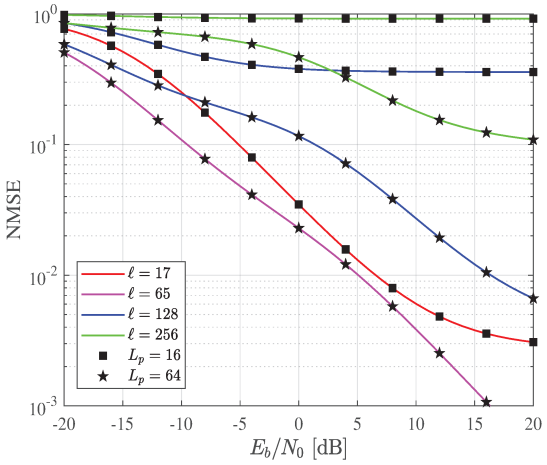
orthogonality of pilot sequences, L should be a power of two. Therefore, $L = 256$ time-slots should be employed by the system, that is, a block duration of $T_B = 17$ ms. For a fair comparison, HPE must also employ $L = 256$ time slots. Furthermore, MPE must employ $L = 256 + L_p$ time slots during $T_B = 17$ ms, but at a cost of bandwidth expansion.

Notice that the aforementioned reason gives a justification for using a maximum number of time slots per block in TICs. Otherwise, it could be possible to use an unlimited number of time slots per block in TICs.

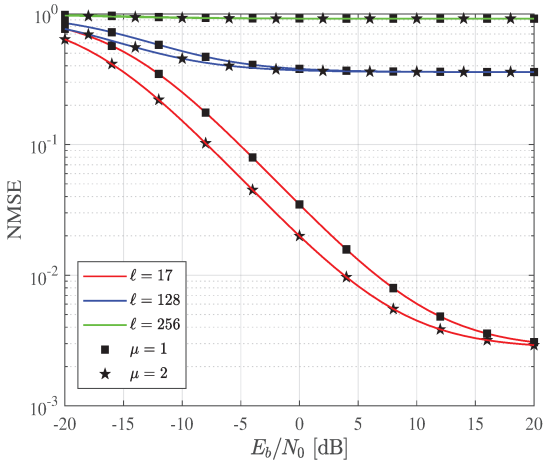
For comparison purposes, a block with duration $T_B = 17$ ms and $L_d = 256$ data symbols is used in the following analyses.

Fig. 4.12 shows the NMSE as a function of time for MPE, parameterized by E_b/N_0, for $L_d = 256$, $L_p = 16$, $\mu = 1$ and $K = 16$. Additionally, for comparison purposes, the NMSE on TICs is presented. For evaluating NMSE for MPE the estimated channel variance given by (4.64) is used in (4.93). Notice that (4.64) is a quadratic form that depends on the channel correlation vector $\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}$ and the bisymmetric matrix $\mathcal{M}^{-1}$, given by (4.65). Besides, the entries of the channel correlation vector are lower at the last time slot and higher at the time slot $\ell = L_p + 1$. Thus, the lower correlation as time elapses after L_p produces channel aging. Hence, the NMSE on TVCs increases with time. On the other hand, on TICs the NMSE does not change with the time. Moreover, on TICs as E_b/N_0 increases the NMSE improvement is more evident. In contrast, on TVCs the NMSE converges to the same value independently of E_b/N_0 because the channel aging dominates the performance.

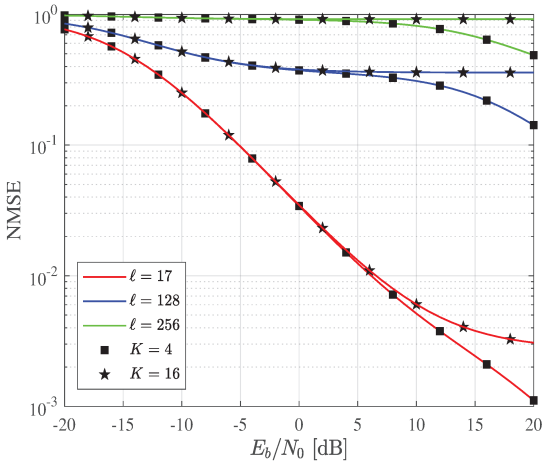
Fig. 4.13a shows the NMSE as a function of E_b/N_0 for MPE, $L_d = 256$, $\mu = 1$ and $K = 16$, parameterized by L_p and ℓ. As a consequence of the channel aging, the NMSE is worsen with the time and there is a NMSE floor. For relieving the floor, more pilot



(a) $K = 16,\ \mu = 1.$



(b) $K = 16,\ L_p = 16.$



(c) $L_p = 16,\ \mu = 1.$

Figure 4.13: NMSE as a function of E_b/N_0 for MPE and $L_d = 256$.

symbols should be used at the cost of decreasing the spectral efficiency, as can be seen for $L_p = 16$ and $L_p = 64$. Moreover, for $L_p = 16$ there is one inflection on the NMSE, while for $L_p = 64$ there are two inflections. From these results and other trials, the number of inflections (NoI) is given by $NoI = 1 + \log_2 L_p/K$. Therefore, the inflections are a result of increasing the number of pilot symbols and can be seen as E_b/N_0 increases.

Fig. 4.13b shows the NMSE as a function of E_b/N_0 for MPE, $L_d = 256$, $L_p = 16$ and $K = 16$, parameterized by μ and ℓ. Notice that there is only one inflection because L_p is equal to K. Furthermore, the number of inflections is independent of μ. The effect of μ is observed at the low E_b/N_0 region, when the NMSE is still dominated by noise. Therefore, by increasing μ, the NMSE performance is improved only for low E_b/N_0.

Fig. 4.13c shows the NMSE as a function of E_b/N_0 for MPE, $L_d = 256$, $L_p = 16$ and $\mu = 1$, parameterized by K and ℓ. Unlike for MPE on TICs, the NMSE on TVCs changes with K and introduces floor due to the channel aging impairments. Furthermore, for $L_p = K = 16$ the NMSE performance is limited at the first inflection, while for $K = 4$

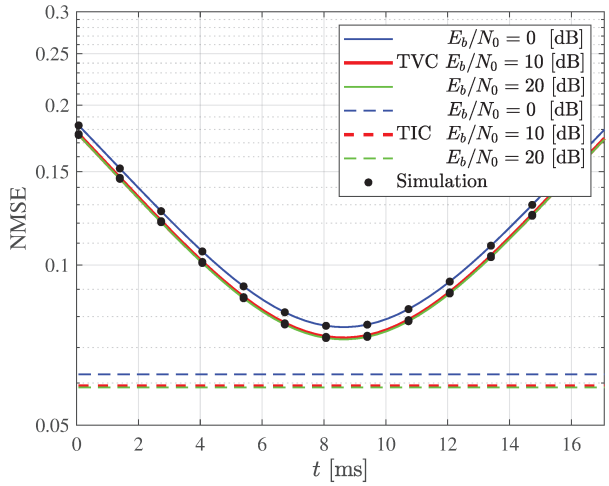
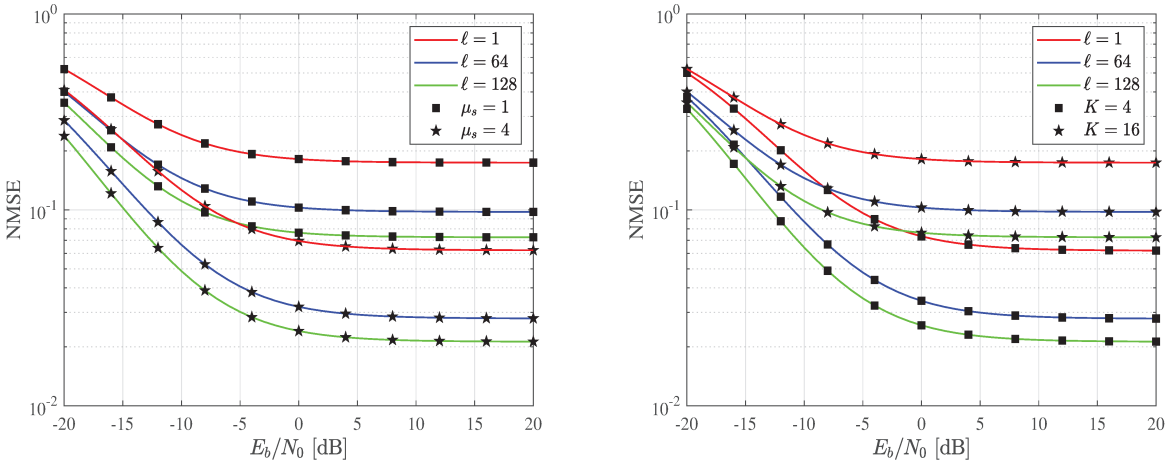


Figure 4.14: NMSE as a function of time for SPE, $L = 256$, $\mu_s = 1$ and $K = 16$, parameterized by E_b/N_0.

the surplus of pilots symbols relieves the NMSE floor by introducing more inflections as E_b/N_0 increases. However, for low E_b/N_0 the NMSE performance is the same for any $K \leqslant L_p$.

Fig. 4.14 shows the NMSE as a function of time for SPE, $L = 256$, $\mu = 1$ and $K = 16$, parameterized by E_b/N_0. For evaluating the NMSE of SPE, the estimated channel variance given by (4.72) is used in (4.93). The estimated channel variance has a quadratic form that depends on the channel correlation vector $\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}$ and the bisymmetric matrix $\mathcal{S}^{-1}$, given by (4.73). Furthermore, the correlation vector at the first time slot $\mathbf{R}_{h_{i,k}(1)\check{\mathbf{h}}_{i,k}}$ is the mirror of the correlation vector at the last time slot of the block $\mathbf{R}_{h_{i,k}(L)\check{\mathbf{h}}_{i,k}}$. Thus, the estimated channel variance of the first time slot $\varrho(1)$ is equal to the estimated channel variance of the last time slot $\varrho(L)$. Actually, the correlation vector at the instant of time ℓ, that is, $\mathbf{R}_{h_{i,k}(\ell)\check{\mathbf{h}}_{i,k}}$, is the mirror of the correlation vector at the instant of time $L+1-\ell$, that is, $\mathbf{R}_{h_{i,k}(L+1-\ell)\check{\mathbf{h}}_{i,k}}$, for $1 \leqslant \ell \leqslant L$. Therefore, the estimated channel variance is symmetric with respect to the middle of the block, that is $\varrho(\ell) = \varrho(L + 1 - \ell)$. Additionally, the channel correlation vector has entries with lower correlation in time slots at the beginning and at the end of the block and entries with higher correlation at the center of the block. Thus, the estimated channel variance is higher at the center of the block and lower at the beginning and at the end of the block. As a result, the filtering process of the MMSE estimator produces a channel correlation vector with U-shape on the NMSE. Besides, as the speed approaches zero, the U-shape flattens, as expected for TICs. Furthermore, the NMSE on TICs remains invariant at a value that is close to the minimum NMSE at the middle of the block on TVCs. For this case, the middle of the block occurs at the time interval $\ell = 128$. Additionally, increasing E_b/N_0 does not effective improves the NMSE, because the system is dominated by the interference introduced by the superimposed data symbols. On the other hand, SPE introduces more delay on the system than MPE, once



(a) $K = 16$. (b) $\mu_s = 1$.

Figure 4.15: NMSE as a function of E_b/N_0 for SPE and $L_d = 256$.

the block of $L = 256$ symbols must be received before the channel estimation process.

Fig. 4.15a shows the NMSE as a function of E_b/N_0 for SPE, $L = 256$ and $K = 16$, parameterized by μ_s and ℓ. As expected, by increasing the pilot symbol energy, i.e., $\mu_s > 1$ the NMSE is reduced. The NMSE is symmetric around $\ell = 128$, then, the performance at $\ell = 256$ is equivalent at $\ell = 1$ and the performance at $\ell = 192$ is equivalent at $\ell = 64$.

Fig. 4.15b shows the NMSE as a function of E_b/N_0 for SPE, $L = 256$ and $\mu_s = 1$, parameterized by K and ℓ. As expected, the interference is reduced, then, reducing the number of UTs it improves the NMSE.

For evaluating the NMSE for HPE, the estimated channel variance given by (4.78) is used in (4.93). The block of $L = 256$ symbols is divided into two parts, the superimposed region with L_s symbols and the multiplexed region with L_{md} symbols. Notice that there are pilot symbols only at the superimposed region. Therefore, there are L_s estimated channel coefficients and L_{md} predicted channel coefficients. In both cases the variance is given by (4.78).

Fig. 4.16 shows the NMSE as a function of time for HPE, $L = 256$, $L_s = 128$, $\mu_{h,s} = 1$, $\mu_{h,m} = 1$ and $K = 16$, parameterized by E_b/N_0. At the interval $1 \leqslant \ell \leqslant 128$ the NMSE presents a U-shape as for SPE, because this interval corresponds to the superimposed region. Moreover, for $\ell > 128$ the shape of NMSE is similar to MPE, because this interval corresponds to the multiplexed region. Therefore, the NMSE of HPE is a combination of both SPE and MPE. Additionally, the NMSE for HPE on TICs and TVCs are similar in the middle of the superimposed region, i.e., at $\ell = 64$, which is the best NMSE performance for TVCs.

Fig. 4.17a shows the NMSE as a function of E_b/N_0 for HPE, $L = 256$, $K = 16$, $\mu_{h,s} = 1$ and $\mu_{h,m} = 1$, parameterized by L_s, and ℓ. Notice that in order to maintain the orthogonality among pilot sequences, the maximum number of superimposed symbols is $L_s = 128$. Furthermore, as the NMSE is reduced by increasing L_s, the best NMSE

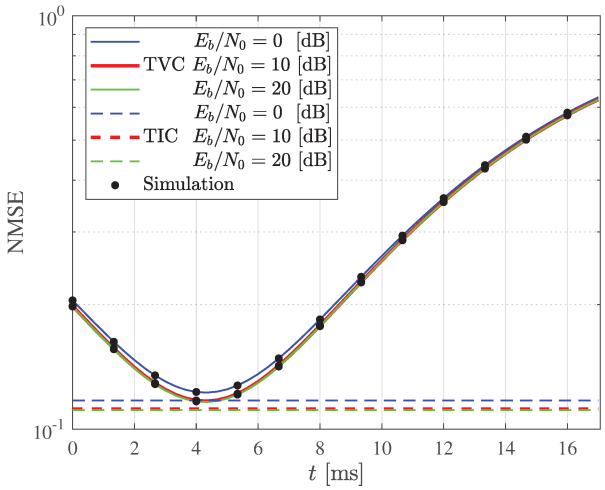
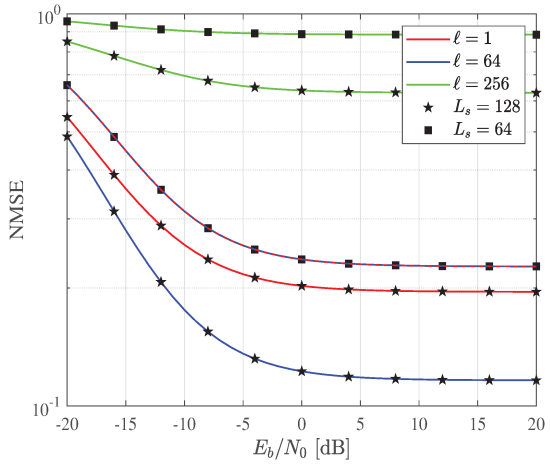
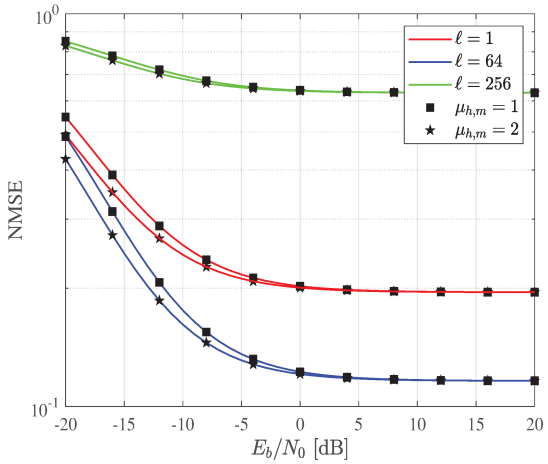


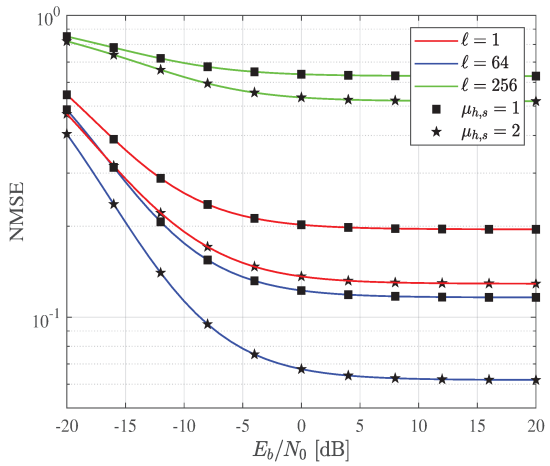
Figure 4.16: NMSE as a function of time for HPE, $L = 256$, $L_s = 128$, $\mu_{h,s} = 1$, $\mu_{h,m} = 1$ and $K = 16$, parameterized by E_b/N_0.



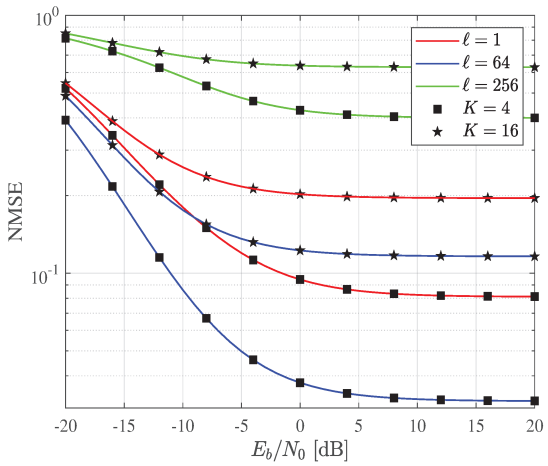
(a) $K = 16$, $\mu_{h,m} = 1$, $\mu_{h,s} = 1$.



(b) $L_s = 128$, $K = 16$, $\mu_{h,s} = 1$.



(c) $L_s = 128$, $K = 16$, $\mu_{h,m} = 1$.



(d) $L_s = 128$, $\mu_{h,s} = 1$, $\mu_{h,m} = 1$.

Figure 4.17: NMSE as a function of E_b/N_0 for HPE and $L_d = 256$.

performance is achieved for $L_s = 128$. Additionally, notice for $L_s = 64$ that the NMSE for $\ell = 1$ and $\ell = 64$ are coincident. This can be justified by the symmetry of the superimposed region.

Fig. 4.17b shows the NMSE as a function of E_b/N_0 for HPE, $L = 256$, $K = 16$, $L_s = 128$ and $\mu_{h,s} = 1$, parameterized by $\mu_{h,m}$, and ℓ. As on TICs, increasing $\mu_{h,m}$ has no effects over the NMSE floor and the NMSE is improved only in the low E_b/N_0 region. Furthermore, Fig. 4.17c shows the NMSE as a function of E_b/N_0 for $\mu_{h,m} = 1$, parameterized by $\mu_{h,s}$ and by the same parameters of the last figure. In this case, by increasing μ_s the NMSE floor is relieved because the pilot symbols energy is increased. Therefore, as on TICs, it is more attractive to increase $\mu_{h,s}$ instead $\mu_{h,m}$.

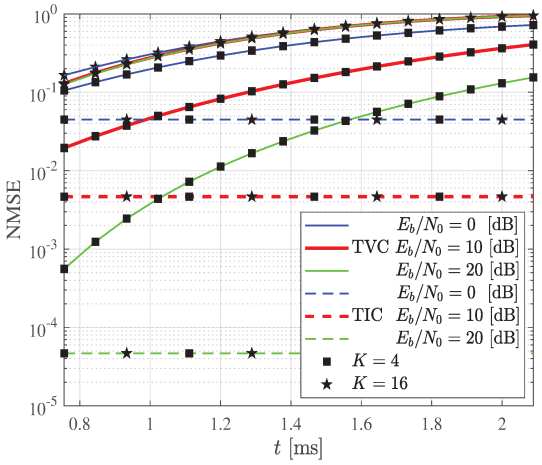
Fig. 4.17d shows the NMSE as a function of E_b/N_0 for $L = 256$, $L_s = 128$, $\mu_{h,s} = 1$ and $\mu_{h,m} = 1$, parameterized by K and ℓ. As expected, by increasing K the NMSE is reduced.

4.6.2.2 Vehicular Velocity

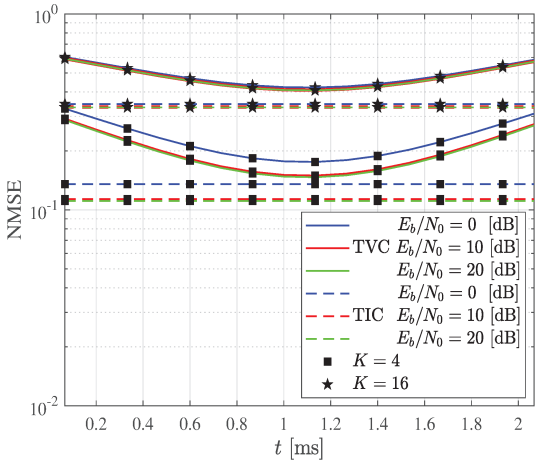
For the vehicular speed of 60 km/h, the channel coherence time interval is $T_c = 2.6$ ms. Thus, for a bandwidth of $B = 15$ kHz, the number of time slots during this time interval is $L = 39$. However, due to the orthogonality restriction of pilot sequences, $L = 32$ time slots should be employed for SPE and HPE. Thus, the block duration is $T_B = 2.1$ ms. For a fair comparison, $L = 32 + L_p$ time slots should be employed for MPE a the same block time interval of $T_B = 2.1$ ms, which introduces bandwidth expansion. The block duration should be shorter at higher speeds, which limits the number of pilot symbols. Therefore, the NMSE is worse for vehicular speed than for pedestrian speed.

Fig. 4.18a, 4.18b and 4.18c shows the NMSE as a function of time for MPE, SPE and HPE for $L = L_d = 32$ symbols and $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} = 1$, parameterized by E_b/N_0 and K. Furthermore, for MPE and HPE $L_p = L_s = 16$ is used. Notice that in order to improve the NMSE, it is recommendable to use the maximum number of pilot symbols and to reduce the number of UTs, as can be seen for $K = 4$ and $K = 16$. For comparison purposes the NMSE on TICs are also plotted. Notice that the results are similar to that of pedestrian speed.

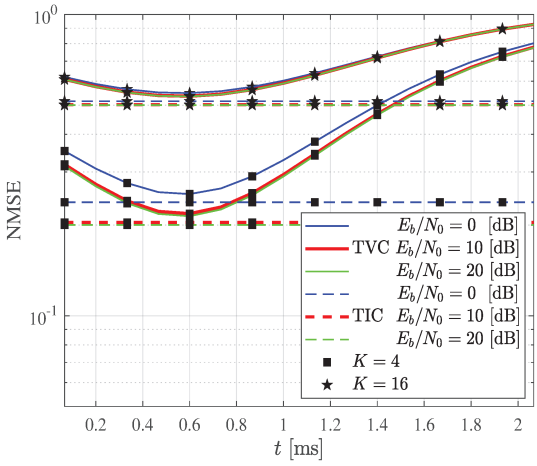
Additional results for vehicular speed can be obtained for NMSE by modifying μ, μ_s, $\mu_{h,s}$ and $\mu_{h,m}$. However, to avoid being redundant, these figures will not be presented.



(a) MPE for $L_d = 32,\ L_p = 16,\ \mu = 1.$



(b) SPE for $L = 32,\ \mu_s = 1.$



(c) HPE for $L = 32,\ L_s = 16,\ \mu_{h,s} = 1,\ \mu_{h,m} = 1.$

Figure 4.18: NMSE as a function of time for MPE, SPE and HPE on TICs and TVCs for $L_d = 32$.

CHAPTER 5

Performance of Linear Detectors in M-MIMO Systems

5.1 Introduction

The linear detection techniques presented in Section 2.6 are widely used in M-MIMO systems due to its low complexity and efficiency [24]. Furthermore, due to the M-MIMO channel properties, they present a near-optimal performance when perfect CSI is available at the receiver [33–35].

However, CSI can be obtained from an estimation process that is not perfect, thus, it introduces errors and limits the linear detectors performance. In Chapter 4, some practical channel estimation techniques were presented. They introduce different types of estimate errors as can be seen in TVCs and TICs.

In this chapter, the BER of M-MIMO systems is evaluated for linear detectors such as the MRC, ZF and MMSE using the MPE, SPE and HPE channel estimation techniques on TVCs and TICs. The uplink of a single cell system is considered, where the BS has N antennas serving K UTs that transmit M-QAM symbols.

The M-MIMO system model considering the estimation techniques is detailed in Section 5.2. The linear detectors and their SNIRs are analyzed for TICs and TVCs in Section 5.3. The exact average BER of M-MIMO systems is derived in Section 5.4 and the simpler BER bounds considering the M-MIMO properties are derived in Section 5.5. Finally, the numerical results and discussions are presented in Section 5.6.

5.2 System Model with Imperfect Estimation

The system model of M-MIMO systems was presented in Section 4.2. The three blocks between the matched filter and the linear detector shown in Fig. 4.1b are fundamental for the BER analysis. The first block is a buffer that stores the samples required on the channel estimation process. The second block performs the channel estimation itself, and

finally, the third block executes a pre-processing of the data and sends these samples to the linear detector.

In this Section, the interaction between these three blocks is detailed in order to define the samples sent to the linear detector for each estimation technique.

5.2.1 System Model for PCE

PCE is an ideal model where perfectly the channel is known. In this case, there are no pilot symbols and the samples are sent directly to the linear detector. Therefore, the linear detector input for the ℓth time interval is given by (4.13), that can be rewritten as:

$$\mathbf{y}_d(\ell) = \sum_{j=1}^{K} \mathbf{h}_j(\ell) \frac{x_{d,j}(\ell)}{2} + \mathbf{w}(\ell), \tag{5.1}$$

for $1 \leqslant \ell \leqslant L$. Notice that $\mathbf{y}_d$ represents the data samples vector.

5.2.2 System Model for MPE

For MPE, the data pre-processing block eliminates the pilot samples. Therefore, the linear detector input for the ℓth time interval is given by:

$$\mathbf{y}_d(\ell) = \sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell) \frac{x_{d,j}(\ell)}{2} + \sum_{j=1}^{K} \tilde{\mathbf{h}}_j(\ell) \frac{x_{d,j}(\ell)}{2} + \mathbf{w}(\ell), \tag{5.2}$$

for $L_p+1 \leqslant \ell \leqslant L$. Additionally, the channel vector at the ℓth time interval can be written as $\mathbf{h}_j(\ell) = \hat{\mathbf{h}}_j(\ell) + \tilde{\mathbf{h}}_j(\ell)$, where $\hat{\mathbf{h}}_j(\ell)$ is the estimated channel and $\tilde{\mathbf{h}}_j(\ell)$ is the channel estimation error. In this case before sending the data samples to the linear detector, L_p samples are discarded, introducing a delay of L_p time interval.

5.2.3 System Model for SPE

For SPE, the ℓth sample is given by:

$$\mathbf{y}(\ell) = \sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell) \frac{[x_{p,j}(\ell) + x_{d,j}(\ell)]}{2} + \sum_{j=1}^{K} \tilde{\mathbf{h}}_j(\ell) \frac{[x_{p,j}(\ell) + x_{d,j}(\ell)]}{2} + \mathbf{w}(\ell), \tag{5.3}$$

for $1 \leqslant \ell \leqslant L$. All the block samples are pre-processed. Furthermore, as $\hat{\mathbf{h}}_j(\ell)$ and $x_{p,j}(\ell)$ are known at the receiver, they can be subtracted by the pre-processing stage. Therefore,

the linear detector input at the ℓth time interval is given by:

$$\begin{aligned}\mathbf{y}_d(\ell) &= \mathbf{y}(\ell) - \sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell)\frac{x_{p,j}(\ell)}{2} \\ &= \sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2} + \sum_{j=1}^{K} \tilde{\mathbf{h}}_j(\ell)\frac{[x_{p,j}(\ell) + x_{d,j}(\ell)]}{2} + \mathbf{w}(\ell),\end{aligned} \tag{5.4}$$

for $1 \leqslant \ell \leqslant L$. Notice that the data pre-processing block introduces a delay of L time intervals.

5.2.4 System Model for HPE

For HPE, the samples at the linear detector input are divided into two intervals. The first interval, for $1 \leqslant \ell \leqslant L_s$, are the superimposed data samples, given by:

$$\begin{aligned}\mathbf{y}_{d,s}(\ell) &= \sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell)\frac{x_{d,s,j}(\ell)}{2} \\ &\quad + \sum_{j=1}^{K} \tilde{\mathbf{h}}_j(\ell)\frac{[x_{p,s,j}(\ell) + x_{d,s,j}(\ell)]}{2} + \mathbf{w}(\ell),\end{aligned} \tag{5.5}$$

where the term $\sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell)\frac{x_{p,s,j}(\ell)}{2}$ was subtracted. Notice that for obtaining (5.5), an analogous pre-processing stage to the SPE is implemented, resulting in an expression similar to (5.4).

The second interval, $L_s + 1 \leqslant \ell \leqslant L$, contains the multiplexed data symbols. This case is similar to the MPE scenario, given by (5.2), that is:

$$\mathbf{y}_{d,m}(\ell) = \sum_{j=1}^{K} \hat{\mathbf{h}}_j(\ell)\frac{x_{d,m,j}(\ell)}{2} + \sum_{j=1}^{K} \tilde{\mathbf{h}}_j(\ell)\frac{x_{d,m,j}(\ell)}{2} + \mathbf{w}(\ell). \tag{5.6}$$

The data pre-processing introduces a delay of L_s time intervals, which is less than the delay of SPE, once $L_s < L$.

5.3 Linear Detection and SNIR

The linear detectors presented in Section 2.6 are evaluated for M-MIMO systems with imperfect channel estimation. Hence, the SNIR for each linear detector on TICs and TVCs is derived.

5.3.1 Maximal Ratio Combining Detector

The data symbol of the kth UT at the output of the MRC detector, given by (2.29), can be rewritten as:

$$\hat{x}_{d,k}(\ell) = \hat{\mathbf{h}}_k^H(\ell)\mathbf{y}_d(\ell), \tag{5.7}$$

where $\mathbf{y}_d(\ell)$ is the ℓth data sample at the input of the MRC detector. The MRC performance depends on the estimation technique and the type of channel, therefore, several expressions of SNIR can be derived in the next sections.

5.3.1.1 SNIR for MPE on TVCs

By replacing (5.2) in (5.7), the detected data symbol of the kth UT is given by:

$$\begin{aligned}\hat{x}_{d,k}(\ell) = &\underbrace{\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1\\ j\neq k}}^{K}\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}}_{\text{MAI}}\\ &+\underbrace{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^H(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^H(\ell)\mathbf{w}(\ell)}_{\text{Noise}},\end{aligned} \tag{5.8}$$

where the first term is the kth user signal, the second term is the MAI, the third term is the CEEI and the last term is the noise.

The SNIR at the detector output conditioned on the kth user estimated channel for the ℓth time interval is given by [66]:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|\text{Signal}_k|^2}{\text{Var}\left\{\text{MAI} + \text{CEEI} + \text{Noise}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}}. \tag{5.9}$$

Notice that the elements in the denominator corresponds to variances of complex random variables.

From (5.8), the kth user signal power is given by:

$$\begin{aligned}|\text{Signal}_k|^2 &= \left|\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,k}(\ell)}{2}\right|^2\\ &= \left\|\hat{\mathbf{h}}_k(\ell)\right\|^4\frac{|x_{d,k}(\ell)|^2}{4}.\end{aligned} \tag{5.10}$$

Once perfect power control is considered, all UTs signals arrive with same power.

Therefore, the MAI variance is given by:

$$\begin{aligned}\mathrm{Var}\left\{\mathrm{MAI}|\hat{\mathbf{h}}_k(\ell)\right\} &= \mathrm{Var}\left\{\sum_{\substack{j=1\\ j\neq k}}^{K}\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}\\ &= \frac{\varrho(\ell)}{4}\left\|\hat{\mathbf{h}}_k(\ell)\right\|^2\overline{|x_d|^2}\,(K-1)\,,\end{aligned} \tag{5.11}$$

where $\varrho(\ell)$ is the variance of the estimated channel given by (4.64). The variance of the interference due to the imperfect channel estimation is given by:

$$\begin{aligned}\mathrm{Var}\left\{\mathrm{CEEI}|\hat{\mathbf{h}}_k(\ell)\right\} &= \mathrm{Var}\left\{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^H(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}\\ &= \frac{\left[\overline{\alpha^2}-\varrho(\ell)\right]}{4}\left\|\hat{\mathbf{h}}_k(\ell)\right\|^2\overline{|x_d|^2}K.\end{aligned} \tag{5.12}$$

Finally, the noise variance is given by:

$$\begin{aligned}\mathrm{Var}\left\{\mathrm{Noise}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\} &= \mathrm{Var}\left\{\hat{\mathbf{h}}_k^H(\ell)\mathbf{w}_k(\ell)\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}\\ &= \left\|\hat{\mathbf{h}}_k(\ell)\right\|^2\sigma_w^2,\end{aligned} \tag{5.13}$$

where σ_w^2 is the noise variance given by (4.10).

By using (5.10)-(5.13) in (5.9), the instantaneous SNIR conditioned on the kth user estimated channel vector is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|x_{d,k}(\ell)|^2\|\hat{\mathbf{h}}_k(\ell)\|^2}{\overline{|x_d|^2}\varrho(\ell)\,(K-1)+\left[\overline{\alpha^2}-\varrho(\ell)\right]\overline{|x_d|^2}K+4\sigma_w^2}, \tag{5.14}$$

where $\|\hat{\mathbf{h}}_k(\ell)\|^2$ and the SNIR follows a chi-square distribution with $2N$ degrees of freedom. See Appendix B.1 for more information about the chi-square distribution.

The SNIR can be unconditioned of $\hat{\mathbf{h}}_k(\ell)$ by applying the mean operation, that is:

$$\begin{aligned}\gamma_s(\ell) &= \mathrm{E}\left\{\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right\}\\ &= \int_0^\infty \gamma_{s|\hat{\mathbf{h}}_k}(\ell)p(\gamma_{s|\hat{\mathbf{h}}_k})d\gamma_{s|\hat{\mathbf{h}}_k},\end{aligned} \tag{5.15}$$

where $p(\gamma_{s|\hat{\mathbf{h}}_k})$ is the PDF of the SNIR conditioned on $\hat{\mathbf{h}}_k$. Observe that the mean of the numerator of (5.14) is given by:

$$\mathrm{E}\left\{|x_{d,k}(\ell)|^2\|\hat{\mathbf{h}}_k(\ell)\|^2\right\} = \overline{|x_d|^2}N\varrho(\ell). \tag{5.16}$$

Therefore, the average SNIR at the ℓth symbol time interval is given by:

$$\begin{aligned}\gamma_s(\ell) &= \frac{\overline{|x_d|^2} N \varrho(\ell)}{\overline{|x_d|^2}\varrho(\ell)\,(K-1) + \left[\overline{\alpha^2} - \varrho(\ell)\right]\overline{|x_d|^2}K + 4\sigma_w^2} \\ &= \frac{N\varrho(\ell)}{\varrho(\ell)\,(K-1) + \left[\overline{\alpha^2} - \varrho(\ell)\right]K + \frac{N_0}{E_d}} \\ &= N\rho(\ell)\left\{\overline{\alpha^2}K - \rho(\ell) + \frac{(1+\mu\eta)}{\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}. \end{aligned} \tag{5.17}$$

where the data symbol energy for the MPE scenario given by (4.32) was employed. Besides, η is given by (4.27) and μ is given by (4.30).

Notice that the SNIR varies with time, once the variance of the estimated channel is time-variant. Thus, the average SNIR in a block can be evaluated by:

$$\overline{\gamma_s} = \sum_{\ell=L_p+1}^{L_d} \frac{\gamma_s(\ell)}{L_d}. \tag{5.18}$$

For obtaining the average SNIR on TVCs, two averages were employed. The first average for unconditioning the SNIR of the estimated channel vector and the second for averaging the SNIR in a block.

5.3.1.2 SNIR for MPE in TICs

TICs are a special case of the TVCs, where $\hat{\mathbf{h}}_j(\ell) = \hat{\mathbf{h}}_j$ and $\tilde{\mathbf{h}}_j(\ell) = \tilde{\mathbf{h}}_j$, for $\ell \leqslant L$. Therefore, the detected data symbol given by (5.8), can be rewritten as:

$$\hat{x}_{d,k} = \underbrace{\hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_k \frac{x_{d,k}}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1 \\ j\neq k}}^{K} \hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_j \frac{x_{d,j}}{2}}_{\text{MAI}} + \underbrace{\sum_{j=1}^{K} \hat{\mathbf{h}}_k^H \tilde{\mathbf{h}}_j \frac{x_{d,j}}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^H \mathbf{w}}_{\text{Noise}}. \tag{5.19}$$

Furthermore, the SNIR conditioned on the kth user estimated channel is similar to (5.9), that is:

$$\gamma_{s|\hat{\mathbf{h}}_k} = \frac{|\text{Signal}_k|^2}{\text{Var}\left\{\text{MAI} + \text{CEEI} + \text{Noise}\,\middle|\,\hat{\mathbf{h}}_k\right\}}. \tag{5.20}$$

Hence, by employing the same procedure as that used for TVCs, the instantaneous SNIR conditioned on the kth user estimated channel is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k} = \frac{|x_{d,k}|^2 \|\hat{\mathbf{h}}_k\|^2}{\overline{|x_d|^2}\varrho\,(K-1) + \left[\overline{\alpha^2} - \varrho\right]\overline{|x_d|^2}K + 4\sigma_w^2}, \tag{5.21}$$

where ϱ is given by (4.84) and $\|\hat{\mathbf{h}}_k\|^2$ is a chi-square random variable with $2N$ degrees of freedom. Therefore, the average SNIR is given by:

$$\gamma_s = N\varrho\left\{\overline{\alpha^2}K - \varrho + \frac{(1+\mu\eta)}{\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}. \tag{5.22}$$

The SNIR is independent of time, then, the average SNIR per block is $\overline{\gamma_s} = \gamma_s$. Thus, only one average is required for the SNIR on TICs.

As the SNIR derivation for TICs is similar to TVCs, in the next sections, for simplicity, the average SNIR for TICs is presented directly without derivation. The average SNIR is necessary in the BER analysis.

5.3.1.3 SNIR for SPE on TVCs

By replacing (5.4) in (5.7), the detected data symbol of the kth user for the ℓth time interval can be rewritten as:

$$\begin{aligned}\hat{x}_{d,k}(\ell) = &\underbrace{\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1\\ j\neq k}}^{K}\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}}_{\text{MAI}}\\ &+\underbrace{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^H(\ell)\tilde{\mathbf{h}}_j(\ell)\left[\frac{x_{p,j}(\ell)+x_{d,j}(\ell)}{2}\right]}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^H(\ell)\mathbf{w}(\ell)}_{\text{Noise}}.\end{aligned} \tag{5.23}$$

The first term is the kth user signal, the second term is the MAI, the third term is CEEI and the last term is the noise. Notice that the CEEI is enhanced by the pilot symbols.

The instantaneous SNIR conditioned by the kth user estimated channel at the ℓth time interval was defined in (5.9). For SPE, it can be rewritten as:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|x_{d,k}(\ell)|^2\|\hat{\mathbf{h}}_k(\ell)\|^2}{\overline{|x_d|^2}\varrho(\ell)\,(K-1) + \left[\overline{\alpha^2} - \varrho(\ell)\right]\left(\overline{|x_p|^2} + \overline{|x_d|^2}\right)K + 4\sigma_w^2}, \tag{5.24}$$

where $|\text{Signal}_k|^2$, $\text{Var}\left\{\text{MAI}|\hat{\mathbf{h}}_k(\ell)\right\}$ and $\text{Var}\left\{\text{Noise}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}$ are given by (5.10), (5.11) and (5.13), respectively. Additionally, the CEEI variance is given by:

$$\text{Var}\left\{\text{CEEI}|\hat{\mathbf{h}}_k(\ell)\right\} = \left[\overline{\alpha^2} - \varrho(\ell)\right]\|\hat{\mathbf{h}}_k(\ell)\|^2\left(\overline{|x_p|^2} + \overline{|x_d|^2}\right)K, \tag{5.25}$$

where $\varrho(\ell)$ is given by (4.72) and $\|\hat{\mathbf{h}}_k(\ell)\|^2$ is a chi-square random variable with $2N$ degrees of freedom. Notice that the SNIR is also a chi-square random variable.

The average SNIR for the ℓth symbol time interval can be obtained by using (5.24) in (5.15), that is:

$$\begin{aligned}\gamma_s(\ell) &= \frac{\overline{|x_d|^2} N \varrho(\ell)}{\overline{|x_d|^2}\varrho(\ell)\,(K-1) + \left[\overline{\alpha^2} - \varrho(\ell)\right]\left(\overline{|x_p|^2} + \overline{|x_d|^2}\right) K + 4\sigma_w^2} \\ &= \frac{N\varrho(\ell)}{\varrho(\ell)\,(K-1) + \left[\overline{\alpha^2} - \varrho(\ell)\right]\left(\frac{E_p}{E_d} + 1\right) K + \frac{N_0}{E_d}} \\ &= N\varrho(\ell)\left\{\overline{\alpha^2} K\,(1+\mu_s) - \varrho(\ell)\,(1+K\mu_s) + \frac{(1+\mu_s)}{\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}, \end{aligned} \tag{5.26}$$

where the pilot symbols energy given by (4.40) and the data symbols energy given by (4.41) are employed and μ_s is given by (4.39). Besides, the average SNIR per block is given by:

$$\overline{\gamma_s} = \sum_{\ell=1}^{L_d} \frac{\gamma_s(\ell)}{L_d}. \tag{5.27}$$

5.3.1.4 SNIR for SPE on TICs

Using the same reasoning as for TVCs, the average SNIR is given by:

$$\gamma_s = N\varrho\left\{K\left(\overline{\alpha^2} + \mu_s\right) - \varrho\,(1+K\mu_s) + \frac{(1+\mu_s)}{\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}, \tag{5.28}$$

where ϱ is given by (4.88).

5.3.1.5 SNIR for HPE on TVCs

From Section 5.2.4, the received data samples can be divided into two intervals. Therefore, two SNIRs should be evaluated, the first for the superimposed data and the second for multiplexed data.

For the superimposed data, the kth detected data symbol at the output of the MRC detector can be obtained by replacing (5.5) in (5.7) as:

$$\begin{aligned}\hat{x}_{d,s,k}(\ell) = &\underbrace{\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,s,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1\\ j\neq k}}^{K} \hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,s,j}(\ell)}{2}}_{\text{MAI}} \\ &+ \underbrace{\sum_{j=1}^{K} \hat{\mathbf{h}}_k^H(\ell)\tilde{\mathbf{h}}_j(\ell)\left[\frac{x_{p,s,j}(\ell) + x_{d,s,j}(\ell)}{2}\right]}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^H(\ell)\mathbf{w}(\ell)}_{\text{Noise}}. \end{aligned} \tag{5.29}$$

Observe that (5.29) is similar to (5.23), both represent the superimposed data. Thus, the conditioned SNIR for the superimposed data using HPE is similar to the SPE case, that is given by (5.24).

By applying the mean operator given by (5.15) in (5.24) and employing the symbol energy of superimposed pilots, given by (4.50) and the energy of superimposed data symbols, given by (4.51), the average SNIR for the ℓth symbol time interval is given by:

$$\gamma_{s_{d,s}}(\ell) = N\varrho(\ell)\left\{\overline{\alpha^2}K\left(1+\mu_{h,s}\right) - \varrho(\ell)\left(1+K\mu_{h,s}\right) + \frac{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}{\mu_{h,m}(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}, \tag{5.30}$$

where $\varrho(\ell)$, η_h, $\mu_{h,m}$ and $\mu_{h,s}$ are given by (4.78), (4.47), (4.48) and (4.49), respectively.

For the multiplexed data, the kth detected data symbol at the output of the MRC detector can be obtained by replacing (5.6) in (5.7) as:

$$\begin{aligned}\hat{x}_{d,m,k}(\ell) = \underbrace{\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,m,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1\\ j\neq k}}^{K}\hat{\mathbf{h}}_k^H(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,m,j}(\ell)}{2}}_{\text{MAI}} \\ + \underbrace{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^H(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,m,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^H(\ell)\mathbf{w}(\ell)}_{\text{Noise}}.\end{aligned} \tag{5.31}$$

The detected symbol, given by (5.31), is similar to the MPE scenario, given by (5.8). Thus, the conditioned SNIR for HPE multiplexed data is similar to the MPE case, and it is given by (5.14). Hence, the average SNIR for HPE multiplexed data can be obtained by (5.15) in (5.14) and employing the energy of multiplexed data symbols, given by (4.52). Therefore, the average SNIR for the ℓ symbol time interval is given by:

$$\gamma_{s_{d,m}}(\ell) = N\varrho(\ell)\left\{\overline{\alpha^2}K - \varrho(\ell) + \frac{(1+\eta_h\mu_{h,m})}{(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}, \tag{5.32}$$

where $\varrho(\ell)$, η_h and $\mu_{h,m}$ are given by (4.78), (4.47) and (4.48), respectively.

Furthermore, the average SNIR for HPE is given by:

$$\overline{\gamma_s} = \frac{1}{L_d}\left[\sum_{\ell=1}^{L_s}\gamma_{s_{d,s}}(\ell) + \sum_{\ell=L_s+1}^{L_d}\gamma_{s_{d,m}}(\ell)\right]. \tag{5.33}$$

5.3.1.6 SNIR for HPE on TICs

For the superimposed data symbols, the average SNIR is given by:

$$\gamma_{s_{d,s}} = N\varrho \left\{ \overline{\alpha^2} K \left(1+\mu_{h,s}\right) - \varrho \left(1+K\mu_{h,s}\right) + \frac{(1+\mu_{h,s})(1+\eta_h\mu_{h,m})}{\mu_{h,m}(1+\eta_h)\log_2(M)\frac{E_b}{N_0}} \right\}^{-1}, \tag{5.34}$$

where ϱ is given by (4.91). On the other hand, average SNIR for the multiplexed data symbols is given by:

$$\gamma_{s_{d,s}} = N\varrho \left\{ \overline{\alpha^2} K - \varrho + \frac{(1+\eta_h\mu_{h,m})}{(1+\eta_h)\log_2(M)\frac{E_b}{N_0}} \right\}^{-1}. \tag{5.35}$$

Notice that the estimated channel variance is the same for both superimposed and multiplexed data symbols.

The average SNIR in a block is given by:

$$\overline{\gamma_s} = \frac{L_s\gamma_{s_{d,s}} + L_{d,m}\gamma_{s_{d,m}}}{L_d}. \tag{5.36}$$

5.3.1.7 SNIR for PCE

The PCE is the simplest case, which can be seen as a special case of the aforementioned estimation techniques. For PCE, the detected data symbol of the kth UT at the MRC output can be obtained by replacing (5.1) in (5.7), that is:

$$\hat{x}_{d,k} = \underbrace{\hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_k \frac{x_{d,k}}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1 \\ j\neq k}}^{K} \hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_j \frac{x_{d,j}}{2}}_{\text{MAI}} + \underbrace{\hat{\mathbf{h}}_k^H \mathbf{w}}_{\text{Noise}}, \tag{5.37}$$

where the first term is the kth UT signal, the second is the MAI and the last term is the noise. Notice that there is no channel estimation interference, nor temporal channel variation issues, thus, $\hat{\mathbf{h}}_j = \mathbf{h}_j$.

For PCE, the SNIR conditioned on the kth UT channel vector can be obtained from (5.20), that is:

$$\gamma_{s|\mathbf{h}_k} = \frac{|x_{d,k}|^2 \|\mathbf{h}_k\|^2}{\overline{|x_d|^2}\ \overline{\alpha^2}\,(K-1) + 4\sigma_w^2}, \tag{5.38}$$

where $|\text{Signal}_k|^2 = \|\mathbf{h}_k\|^4 \frac{|x_{d,k}|^2}{4}$, $\text{Var}\{\text{MAI}|\mathbf{h}_k\} = \frac{\overline{\alpha^2}}{4}\|\mathbf{h}_k\|^2 \overline{|x_d|^2}\,(K-1)$, $\text{Var}\{\text{CEEI}|\mathbf{h}_k\} = 0$ and $\text{Var}\{\text{Noise}|\mathbf{h}_k\} = \|\mathbf{h}_k\|^2 \sigma_w^2$, was employed. By using (5.38) in (5.15), considering that $\|\mathbf{h}_k\|^2$ has chi-square distribution and employing the data symbol energy given by (4.20),

for the PCE, the average SNIR is given by:

$$\gamma_s = N\overline{\alpha^2}\left\{\overline{\alpha^2}(K-1) + \frac{1}{\log_2(M)\frac{E_b}{N_0}}\right\}^{-1}. \tag{5.39}$$

For PCE, the average SNIR per block is also given by (5.39). Additionally, the SNIR can be obtained from the SNIR expressions with imperfect channel estimation. For example, the SNIR for MPE given by (5.22) is equal to the SNIR for PCE given by (5.39), by considering that $\varrho = \overline{\alpha^2}$ and that $\eta = 0$

5.3.2 Zero Forcing Detector

The detected data symbol of the kth UT at the output of the ZF detector, given by (2.32) can be rewritten as:

$$\hat{x}_{d,k}(\ell) = \hat{\mathbf{h}}_k^{+}(\ell)\mathbf{y}_d(\ell), \tag{5.40}$$

where $\mathbf{y}_d(\ell)$ is the ℓth data sample at the input of the ZF detector that depends on the estimation technique.

5.3.2.1 SNIR for MPE on TVCs

By replacing (5.2) in (5.40), the detected data symbol of the kth UT for the ℓth symbol time interval can be rewritten as:

$$\hat{x}_{d,k}(\ell) = \underbrace{\frac{x_{d,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^{+}(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^{+}(\ell)\mathbf{w}(\ell)}_{\text{Noise}}, \tag{5.41}$$

where the MAI is eliminated according to (2.33), but the CEEI is still present.

In this case, the SNIR at the detector output conditioned on the estimated channel vector of the kth UT is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|\text{Signal}_k|^2}{\text{Var}\left\{\text{CEEI} + \text{Noise}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}}. \tag{5.42}$$

From (5.41), the kth user signal power is given by:

$$\begin{aligned}|\text{Signal}_k|^2 &= \left|\frac{x_{d,k}(\ell)}{2}\right|^2 \\ &= \frac{|x_{d,k}(\ell)|^2}{4},\end{aligned} \tag{5.43}$$

the variance of the CEEI is given by:

$$\begin{aligned}\mathrm{Var}\left\{\mathrm{CEEI}|\hat{\mathbf{h}}_k(\ell)\right\} &= \mathrm{Var}\left\{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^{+}(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\}\\ &= \frac{\left[\overline{\alpha^2}-\varrho(\ell)\right]}{4}\left\|\hat{\mathbf{h}}_k^{+}(\ell)\right\|^2\overline{|x_d|^2}K,\end{aligned} \tag{5.44}$$

where $\varrho(\ell)$ is the estimated channel variance and the noise variance is given by:

$$\begin{aligned}\mathrm{Var}\left\{\mathrm{Noise}\,\middle|\,\hat{\mathbf{h}}_k(\ell)\right\} &= \mathrm{Var}\left\{\hat{\mathbf{h}}_k^{+}(\ell)\mathbf{w}_k(\ell)\,\middle|\,\hat{\mathbf{h}}_k^{+}(\ell)\right\}\\ &= \left\|\hat{\mathbf{h}}_k^{+}(\ell)\right\|^2\sigma_w^2,\end{aligned} \tag{5.45}$$

where σ_w^2 is the noise variance given by (4.10).

By using (5.43)-(5.45) in (5.42), the instantaneous SNIR conditioned on the estimated channel vector of the kth UT is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|x_{d,k}(\ell)|^2}{\|\hat{\mathbf{h}}_k^{+}(\ell)\|^2\left[\left(\overline{\alpha^2}-\varrho(\ell)\right)\overline{|x_d|^2}K+4\sigma_w^2\right]}, \tag{5.46}$$

where $\frac{1}{\|\hat{\mathbf{h}}_k^{+}(\ell)\|^2}$ and the conditioned SNIR follows a chi-square distribution with $2(N-K+1)$ degrees of freedom for Rayleigh fading [68].

By employing the average given by (5.15) in (5.46), it is possible to show that:

$$\gamma_s(\ell) = \frac{\overline{|x_d|^2}(N-K+1)\varrho(\ell)}{\left[\left(\overline{\alpha^2}-\varrho(\ell)\right)\overline{|x_d|^2}K+4\sigma_w^2\right]}. \tag{5.47}$$

Hence, by employing the data symbol energy for the MPE scenario given by (4.32), the average SNIR for the ℓth symbol time interval is given by:

$$\gamma_s(\ell) = (N-K+1)\varrho(\ell)\left[K\left(\overline{\alpha^2}-\varrho(\ell)\right)+\frac{(1+\mu\eta)}{\log_2(M)\frac{E_b}{N_0}}\right]^{-1}, \tag{5.48}$$

where $\varrho(\ell)$ is given by (4.64). Furthermore, the average SNIR per block can be found by substituting (5.48) in (5.18).

5.3.2.2 SNIR for MPE on TICs

Following a similar procedure presented in the previous section, it is possible to show that the average SNIR for TICs is given by:

$$\gamma_s = (N-K+1)\varrho\left[K\left(\overline{\alpha^2}-\varrho\right)+\frac{(1+\mu\eta)}{\log_2(M)\frac{E_b}{N_0}}\right]^{-1}, \tag{5.49}$$

where ϱ is given by (4.84) and the average SNIR per block is also given by (5.49).

5.3.2.3 SNIR for SPE on TVCs

For SPE, by using (5.4) in (5.40), the detected data symbol of the kth user for the ℓth time interval can be rewritten as:

$$\hat{x}_{d,k}(\ell) = \underbrace{\frac{x_{d,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{j=1}^{K}\hat{\mathbf{h}}_k^{+}(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{p,j}(\ell)+x_{d,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^{+}(\ell)\mathbf{w}(\ell)}_{\text{Noise}}. \tag{5.50}$$

Notice in (5.50) that the MAI is eliminated by the channel inversion matrix given by (2.33) and that the CEEI has effects on the superimposed pilot symbols.

The instantaneous SNIR conditioned on the kth UT estimated channel vector is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|x_{d,k}(\ell)|^2}{\|\hat{\mathbf{h}}_k^{+}(\ell)\|^2\left[\left(\overline{\alpha^2}-\varrho(\ell)\right)\left(\overline{|x_p|^2}+\overline{|x_d|^2}\right)K+4\sigma_w^2\right]}, \tag{5.51}$$

that is obtained by substituting in (5.42), the signal power given by (5.43), the noise variance given by (5.45) and the CEEI variance given by:

$$\text{Var}\left\{\text{CEEI}|\hat{\mathbf{h}}_k(\ell)\right\} = \left[\overline{\alpha^2}-\varrho(\ell)\right]\|\hat{\mathbf{h}}_k^{+}(\ell)\|^2\left(\overline{|x_p|^2}+\overline{|x_d|^2}\right)K. \tag{5.52}$$

Moreover, in (5.51), $\frac{1}{\|\hat{\mathbf{h}}_k^{+}(\ell)\|^2}$ and the SNIR follows a chi-square distribution with $2(N-K+1)$ degrees of freedom. Thus, by using (5.51) in (5.15), the average SNIR is given by:

$$\gamma_s(\ell) = (N-K+1)\varrho(\ell)\left\{(1+\mu_s)\left[K\left(\overline{\alpha^2}-\varrho(\ell)\right)+\frac{1}{\log_2(M)\frac{E_b}{N_0}}\right]\right\}^{-1}, \tag{5.53}$$

where the pilot symbols energy given by (4.40) and the data symbols energy given by (4.41) are employed. Besides, $\varrho(\ell)$ is given by (4.72). Notice that the average SNIR per block can be obtained by substituting (5.53) in (5.27).

5.3.2.4 SNIR for SPE on TICs

The average SNIR for TICs is given by:

$$\gamma_s = (N-K+1)\varrho \left\{ (1+\mu_s) \left[K\left(\overline{\alpha^2} - \varrho\right) + \frac{1}{\log_2(M)\frac{E_b}{N_0}} \right] \right\}^{-1}, \tag{5.54}$$

where ϱ is given by (4.88) and the average SNIR is also given by (5.54).

5.3.2.5 SNIR for HPE on TVCs

As in the MRC detector analysis, there are two SNIRs to be evaluated, for the superimposed data and for the multiplexed data.

For the superimposed data the kth detected data symbol at the output of the ZF detector for the ℓth time interval is obtained by replacing (5.5) in (5.40) as:

$$\hat{x}_{d,s,k}(\ell) = \underbrace{\frac{x_{d,s,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{j=1}^{K} \hat{\mathbf{h}}_k^{+}(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{p,s,j}(\ell) + x_{d,s,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^{+}(\ell)\mathbf{w}(\ell)}_{\text{Noise}}. \tag{5.55}$$

Observe that (5.55) is similar to (5.50). Therefore, the conditioned SNIR for the superimposed data symbol of HPE is similar to the SPE given by (5.51). By using (5.51) in (5.15) and employing the energy of the superimposed data symbols given by (4.51) and the energy of superimposed pilot symbols given by (4.50), the average SNIR is given by:

$$\gamma_{s_{d,s}}(\ell) = (N-K+1)\varrho(\ell) \left\{ (1+\mu_{h,s}) \left[K\left(\overline{\alpha^2} - \varrho(\ell)\right) + \frac{1+\mu_{h,m}\eta_h}{\mu_{h,m}(1+\eta_h)\log_2(M)\frac{E_b}{N_0}} \right] \right\}^{-1}, \tag{5.56}$$

where $\varrho(\ell)$ is given by (4.78).

For the multiplexed data, the kth detected data symbol at the output of the ZF detector is given by replacing (5.6) in (5.40), that is:

$$\begin{aligned} \hat{x}_{d,m,k}(\ell) = {} & \underbrace{\hat{\mathbf{h}}_k^{H}(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,m,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1 \\ j\neq k}}^{K} \hat{\mathbf{h}}_k^{H}(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,m,j}(\ell)}{2}}_{\text{MAI}} \\ & + \underbrace{\sum_{j=1}^{K} \hat{\mathbf{h}}_k^{H}(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,m,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\hat{\mathbf{h}}_k^{H}(\ell)\mathbf{w}(\ell)}_{\text{Noise}}. \end{aligned} \tag{5.57}$$

The detected symbol, given by (5.57) is similar to the detected symbol of MPE, given by (5.41). Thus, the conditioned SNIR for the multiplexed data of HPE is similar to the

conditioned SNIR of MPE, given by (5.46). Furthermore, by employing the energy of the multiplexed data, given by (4.52), the average SNIR is given by:

$$\gamma_{s_{d,m}}(\ell) = (N-K+1)\varrho(\ell)\left[K\left(\overline{\alpha^2}-\varrho(\ell)\right)+\frac{1+\mu_{h,m}\eta_h}{(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}\right]^{-1}, \tag{5.58}$$

where $\varrho(\ell)$ is given by (4.78). The average SNIR per block can be obtained by substituting (5.56) and (5.58) in (5.33).

5.3.2.6 SNIR for HPE on TICs

For TICs, the average SNIR of the superimposed data is given by:

$$\gamma_{s_{d,s}} = (N-K+1)\varrho\left\{(1+\mu_{h,s})\left[K\left(\overline{\alpha^2}-\varrho\right)+\frac{1+\mu_{h,m}\eta_h}{\mu_{h,m}(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}\right]\right\}^{-1}, \tag{5.59}$$

where ϱ is given by (4.91). On the other hand, the average SNIR of the multiplexed data is given by

$$\gamma_{s_{d,m}} = (N-K+1)\varrho\left[K\left(\overline{\alpha^2}-\varrho\right)+\frac{1+\mu_{h,m}\eta_h}{(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}\right]^{-1}, \tag{5.60}$$

where ϱ is also given by (4.91). The average SNIR per block can be obtained by substituting (5.59) and (5.60) in (5.36).

5.3.2.7 SNIR for PCE

The detected data symbol of the kth UT at the ZF output can be obtained by replacing (5.1) in (5.40), that is:

$$\hat{x}_{d,k} = \underbrace{\frac{x_{d,k}}{2}}_{\text{Signal}} + \underbrace{\hat{\mathbf{h}}_k^{+}\mathbf{w}}_{\text{Noise}}, \tag{5.61}$$

where the MAI is eliminated according to (2.33) and in this case there is no CEEI. Thus, the SNIR conditioned on the kth UT channel vector is obtained from (5.42) as:

$$\gamma_{s|\hat{\mathbf{h}}_k} = \frac{|x_{d,k}|^2}{\|\mathbf{h}_k^{+}\|^2 4\sigma_w^2}, \tag{5.62}$$

where $|\text{Signal}_k|^2 = \frac{|x_{d,k}|^2}{4}$, $\text{Var}\{\text{MAI}|\mathbf{h}_k\} = 0$, $\text{Var}\{\text{CEEI}|\mathbf{h}_k\} = 0$ and $\text{Var}\{\text{Noise}|\mathbf{h}_k\} = \left\|\mathbf{h}_k^+\right\|^2 \sigma_w^2$, are employed. By using (5.62) in (5.15), the average SNIR is given by:

$$\gamma_s = (N - K + 1)\,\overline{\alpha^2}\log_2(M)\frac{E_b}{N_0}. \tag{5.63}$$

Notice that the average SNIR per block is also given by (5.63).

5.3.3 Minimum-Mean-Square Error Detector

The detected data symbol for the kth UT at the MMSE detector output is given by:

$$\hat{x}_{d,k}(\ell) = \mathbf{a}_k(\ell)\mathbf{y}_d(\ell), \tag{5.64}$$

where $\mathbf{a}_k(\ell)$ is the kth row of the compensation matrix $\mathbf{A}(\ell)$, given by (2.35). The matrix $\mathbf{A}(\ell) = \mathbf{C}_{x_{k,d}(\ell)\mathbf{y}_d(\ell)}\mathbf{C}^{-1}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}$, where $\mathbf{y}_d(\ell)$ is the data sample of the ℓth time interval at the input of the MMSE detector, given by (5.2).

5.3.3.1 SNIR for MPE on TVCs

By replacing (5.2) in (5.64), the detected data symbol of the kth UT during at the ℓth time interval is given by:

$$\hat{x}_{d,k}(\ell) = \underbrace{\mathbf{a}_k(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1\\ j\neq k}}^{K}\mathbf{a}_k(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}}_{\text{MAI}} + \underbrace{\sum_{j=1}^{K}\mathbf{a}_k(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{d,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\mathbf{a}_k(\ell)\mathbf{w}(\ell)}_{\text{Noise}}. \tag{5.65}$$

The compensation channel vector $\mathbf{a}_k(\ell)$ can be rewritten as [85]:

$$\mathbf{a}_k(\ell) = \hat{\mathbf{h}}_k^H(\ell)\left(\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}\right)^{-1}, \tag{5.66}$$

where $\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)} = \text{E}\left\{\underline{\mathbf{y}_{d,k}}(\ell)\underline{\mathbf{y}_{d,k}}^H(\ell)\right\}$ and $\underline{\mathbf{y}_{d,k}}(\ell)$ is the received signal vector $\mathbf{y}_d$ without the kth user entry given by:

$$\mathbf{y}_d(\ell) = \underline{\widehat{\mathbf{H}}_k}(\ell)\underline{\mathbf{x}_{d,k}}(\ell) + \underline{\widetilde{\mathbf{H}}_k}(\ell)\underline{\mathbf{x}_{d,k}}(\ell) + \mathbf{w}(\ell), \tag{5.67}$$

where $\underline{\widehat{\mathbf{H}}_k}(\ell)$ is the estimated channel matrix and $\underline{\widetilde{\mathbf{H}}_k}(\ell)$ is the channel estimation error matrix without the kth column vector. Besides, $\underline{\mathbf{x}_{d,k}}(\ell)$ is the vector of data symbols without the kth UT entry. Therefore, the covariance matrix $\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}$ can be rewritten

as:

$$\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)} = \frac{\overline{|x_d|^2}}{4}\underline{\widehat{\mathbf{H}}_k}(\ell)\,\underline{\widehat{\mathbf{H}}_k^H}(\ell) + \left[\frac{\overline{|x_d|^2}}{4}\left(\overline{\alpha^2} - \varrho(\ell)\right)K + \sigma_w^2\right]\mathbf{I}_N. \tag{5.68}$$

The product of matrices $\underline{\widehat{\mathbf{H}}_k}(\ell)\underline{\widehat{\mathbf{H}}_k^H}(\ell)$ can be eigen-decomposed [83]:

$$\underline{\widehat{\mathbf{H}}_k}(\ell)\underline{\widehat{\mathbf{H}}_k^H}(\ell) = \mathbf{U}^H(\ell)\mathbf{D}(\ell)\mathbf{U}(\ell), \tag{5.69}$$

where $\mathbf{U}(\ell)$ is an orthonormal matrix that contains the eigen-vectors of $\underline{\widehat{\mathbf{H}}_k}(\ell)\underline{\widehat{\mathbf{H}}_k^H}(\ell)$ and $\mathbf{D}(\ell) = \text{diag}[\lambda_1(\ell)\ \lambda_2(\ell)\ \cdots\ \lambda_{K-1}(\ell)\ \overbrace{0\ \cdots\ 0}^{N-K+1}]$ is a diagonal matrix that contains the eigenvalues of $\underline{\widehat{\mathbf{H}}_k}(\ell)\underline{\widehat{\mathbf{H}}_k^H}(\ell)$. Notice in the diagonal matrix that there are $K-1$ random eigenvalues and $N-K+1$ null eigenvalues [85]. Substituting the eigen-decomposition, given by (5.69), in (5.68), the covariance matrix can be rewritten as:

$$\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)} = \frac{\overline{|x_d|^4}}{4}\mathbf{U}^H(\ell)\mathbf{D}(\ell)\mathbf{U}(\ell) + \left[\frac{\overline{|x_d|^2}}{4}\left(\overline{\alpha^2} - \varrho(\ell)\right)K + \sigma_w^2\right]\mathbf{I}_N. \tag{5.70}$$

The conditioned SNIR is given by (5.9), where the kth UT signal power is given by:

$$|\text{Signal}_k|^2 = \frac{|x_{d,k}(\ell)|^2}{4}\hat{\mathbf{h}}_k^H(\ell)\left(\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}\right)^{-1}\hat{\mathbf{h}}_k(\ell)\hat{\mathbf{h}}_k^H(\ell)\left(\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}\right)^{-1}\hat{\mathbf{h}}_k(\ell), \tag{5.71}$$

that was obtained by using (5.66) in (5.65). Besides, the sum of the variances on the denominator is given by:

$$\text{Var}\left\{\text{MAI} + \text{CEEI} + \text{Noise}\ \middle|\ \hat{\mathbf{h}}_k(\ell)\right\} = \hat{\mathbf{h}}_k^H(\ell)\left(\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}\right)^{-1}\hat{\mathbf{h}}_k(\ell). \tag{5.72}$$

Therefore, by replacing (5.71) and (5.72) in (5.9), the conditioned SNIR is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|x_{d,k}(\ell)|^2}{4}\hat{\mathbf{h}}_k^H(\ell)\left(\mathbf{C}^{(k)}_{\mathbf{y}_d(\ell)\mathbf{y}_d(\ell)}\right)^{-1}\hat{\mathbf{h}}_k(\ell). \tag{5.73}$$

Furthermore, by using (5.70) in (5.73), and considering that the product of orthonormal matrices do not change the statistics of a random matrix, the conditioned SNIR can be rewritten as:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \sum_{j=1}^{N}\frac{|x_{d,k}(\ell)|^2\ |\hat{h}_{j,k}(\ell)|^2}{\overline{|x_d|^2}\lambda_j(\ell) + \left[\overline{\alpha^2} - \varrho(\ell)\right]\overline{|x_d|^2}K + 4\sigma_w^2}, \tag{5.74}$$

or as:

$$\gamma_{s|\hat{\mathbf{h}}_k,\lambda_j}(\ell) = \sum_{j=1}^{K-1} \frac{|x_{d,k}(\ell)|^2 \ |\hat{h}_{j,k}(\ell)|^2}{\overline{|x_d|^2}\lambda_j(\ell) + \left[\overline{\alpha^2} - \varrho(\ell)\right]\overline{|x_d|^2}K + 4\sigma_w^2} + \sum_{j=K}^{N} \frac{|x_{d,k}(\ell)|^2 \ |\hat{h}_{j,k}(\ell)|^2}{\left[\overline{\alpha^2} - \varrho(\ell)\right]\overline{|x_d|^2}K + 4\sigma_w^2}, \tag{5.75}$$

where $\lambda_j(\ell)$ is the jth eigenvalue of $\underline{\widehat{\mathbf{H}}_k}(\ell)\underline{\widehat{\mathbf{H}}_k^H}(\ell)$. In (5.75) notice that the first summation depends on the $K-1$ random eigenvalues and the second summation on the remaining $N-K+1$ eigenvalues that are equal to zero.

The SNIR, given by (5.74), can be averaged over the eigenvalues distribution. In [83], the authors have derived the eigenvalues distribution, that resulted in a tricky expression. Thus, averaging the SNIR is a quite complex task as shown in [33]. However, for M-MIMO, the eigenvalues distribution $p(\lambda_j)$ can be well approximated by the Marchenko-Pastur distribution, whose PDF is given by (2.23). Therefore, the SNIR can be well approximated by:

$$\begin{aligned}\gamma_{s|\hat{\mathbf{h}}_k}(\ell) &= \int_a^b \gamma_{s|\hat{\mathbf{h}}_k,\lambda_j} p(\lambda_j) d\lambda_j \\ &= \frac{|x_{d,k}(\ell)|^2}{\overline{|x_d|^2}} \left(\Omega(\ell) \sum_{j=1}^{K-1} |\hat{h}_{j,k}(\ell)|^2 + \frac{1}{\Lambda(\ell)} \sum_{j=K}^{N} |\hat{h}_{j,k}(\ell)|^2 \right),\end{aligned} \tag{5.76}$$

where $\Omega(\ell)$ is given by:

$$\begin{aligned}\Omega(\ell) &= \int_a^b \frac{1}{\lambda_j + \Lambda(\ell)} p(\lambda_j) d\lambda_j \\ &= \frac{1}{2\Lambda(\ell)} \left[\sqrt{[a + \Lambda(\ell)][b + \Lambda(\ell)]} - (\Lambda(\ell) + \sqrt{ab}) \right],\end{aligned} \tag{5.77}$$

and $\Lambda(\ell)$ is given by:

$$\Lambda(\ell) = \left[\overline{\alpha^2} - \varrho(\ell)\right] K + \frac{1 + \eta\mu}{\log_2(M)\frac{E_b}{N_0}}, \tag{5.78}$$

where the data symbols energy, given by (4.32), was employed. Besides, $\varrho(\ell)$ is given by (4.64), $a = \left(1 - \sqrt{\beta}\right)^2$, $b = \left(1 + \sqrt{\beta}\right)^2$ and $\beta = K/N$.

The SNIR conditioned on the kth UT estimated channel vector, given by (5.76), is the sum of two chi-square random variables with different variances. The sum of chi-square random variables with different variances produces a random variable with the generalized chi-square distribution, shown in Appendix B.2. The generalized chi-square PDF for two

different variances can written as:

$$p(\gamma_s) = \prod_{m=1}^{2} \frac{1}{\varsigma_m^{2r_m}} \sum_{k=1}^{2} \sum_{n=1}^{r_k} \frac{\psi_{k,n,\mathbf{r}}}{(r_k - n)!} (-\gamma_s)^{r_k - n} e^{-\frac{\gamma_s}{\varsigma_k^2}}, \tag{5.79}$$

where $\mathbf{r} = [r_1\ r_2]$ is a vector containing the degrees of freedom of each sum in (5.76). Thus, the first chi-square random variable has $r_1 = K - 1$ degrees of freedom with variance $\varsigma_1^2 = \Omega(\ell)$ and the second chi-square random variable has $r_2 = N - K + 1$ degrees of freedom with variance $\varsigma_2^2 = 1/\Lambda(\ell)$. Finally, $\psi_{1,n,\mathbf{r}}$ and $\psi_{2,n,\mathbf{r}}$ are given by:

$$\psi_{1,n,\mathbf{r}} = (-1)^{r_1 - 1} \begin{pmatrix} n + r_2 - 2 \\ n - 1 \end{pmatrix} \left(\frac{1}{\varsigma_2^2} - \frac{1}{\varsigma_1^2} \right)^{-(r_2 + n - 1)}, \tag{5.80}$$

$$\psi_{2,n,\mathbf{r}} = (-1)^{r_2 - 1} \begin{pmatrix} n + r_1 - 2 \\ n - 1 \end{pmatrix} \left(\frac{1}{\varsigma_1^2} - \frac{1}{\varsigma_2^2} \right)^{-(r_1 + n - 1)}. \tag{5.81}$$

By using (5.75) and (5.79) in (5.15), the average SNIR for the ℓth time interval is given by:

$$\gamma_s(\ell) = \varrho(\ell)\Omega(\ell)(K - 1) + \frac{\varrho(\ell)}{\Lambda(\ell)}(N - K + 1). \tag{5.82}$$

Moreover, as $N \gg K$, the contribution of the first term of (5.82) is negligible. As the second term of (5.82) is dominant, the SNIR of the MMSE detector is similar to the SNIR of the ZF detector, given by (5.48).

A similar result can be obtained in the following by considering that the first sum of (5.75) is negligible. Thus, by considering only the second summation, the SNIR has a chi-square distribution $\gamma_s \sim \chi^2[2(N - K + 1)]$, which is equal to the SNIR distribution of the ZF detector. Hence, the SNIR for the MMSE detector can be well approximated by the SNIR of the ZF detector, given by (5.48). Finally, the average SNIR per block for MPE can be obtained by substituting (5.82) in (5.18).

5.3.3.2 SNIR for MPE on TICs

The derivation of Λ is fundamental on the evaluation of the average SNIR. Hence, by performing a similar procedure as for TVCs, Λ is given by:

$$\Lambda = \left[\overline{\alpha^2} - \varrho \right] K + \frac{1 + \eta\mu}{\log_2(M)\frac{E_b}{N_0}}, \tag{5.83}$$

where ϱ is given by (4.84). Hence, the average SNIR is similar to (5.82), that can be rewritten as:

$$\gamma_s = \varrho\Omega(K-1) + \frac{\varrho}{\Lambda}(N-K+1), \tag{5.84}$$

where Λ is given by (5.83) and Ω is given by (5.77).

The average SNIR per block is also given by (5.84). Furthermore, the SNIR of the MMSE detector can also be approximated by the SNIR of the ZF, given by (5.49).

5.3.3.3 SNIR for SPE on TVCs

For SPE, the detected data symbol of the kth UT at the output of the MMSE detector is obtained by substituting (5.4) in (5.64), that is:

$$\begin{aligned}\hat{x}_{d,k}(\ell) = &\underbrace{\mathbf{a}_k(\ell)\hat{\mathbf{h}}_k(\ell)\frac{x_{d,k}(\ell)}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1\\ j\neq k}}^{K}\mathbf{a}_k(\ell)\hat{\mathbf{h}}_j(\ell)\frac{x_{p,j}(\ell)+x_{d,j}(\ell)}{2}}_{\text{MAI}}\\ &+\underbrace{\sum_{j=1}^{K}\mathbf{a}_k(\ell)\tilde{\mathbf{h}}_j(\ell)\frac{x_{p,j}(\ell)+x_{d,j}(\ell)}{2}}_{\text{CEEI}} + \underbrace{\mathbf{a}_k(\ell)\mathbf{w}(\ell)}_{\text{Noise}}.\end{aligned} \tag{5.85}$$

As for MPE, the SNIR conditioned on the kth user estimated channel vector and the eigenvalues is given by:

$$\begin{aligned}\gamma_{s|\hat{\mathbf{h}}_k,\lambda_j}(\ell) = &\sum_{j=1}^{K-1}\frac{|x_{d,k}(\ell)|^2\ |\hat{h}_{j,k}(\ell)|^2}{\left[|x_p|^2+\overline{|x_d|^2}\right]\lambda_j + \left[\overline{\alpha^2}-\varrho(\ell)\right]\left[|x_p|^2+\overline{|x_d|^2}\right]K + 4\sigma_w^2}\\ &+\sum_{j=K}^{N}\frac{|x_{d,k}(\ell)|^2\ |\hat{h}_{j,k}(\ell)|^2}{\left[\overline{\alpha^2}-\varrho(\ell)\right]\left[|x_p|^2+\overline{|x_d|^2}\right]K + 4\sigma_w^2}.\end{aligned} \tag{5.86}$$

By considering the Marchenko-Pastur approximation in (5.86) the average SNIR in relation to the eigenvalues can be written as:

$$\gamma_{s|\hat{\mathbf{h}}_k}(\ell) = \frac{|x_{d,k}(\ell)|^2}{|x_p|^2+\overline{|x_d|^2}}\left(\Omega(\ell)\sum_{j=1}^{K-1}|\hat{h}_{j,k}(\ell)|^2 + \frac{1}{\Lambda(\ell)}\sum_{j=K}^{N}|\hat{h}_{j,k}(\ell)|^2\right), \tag{5.87}$$

where $\Omega(\ell)$ is given by (5.77) and $\Lambda(\ell)$ is given by:

$$\Lambda(\ell) = \left[\overline{\alpha^2}-\varrho(\ell)\right]K + \frac{1}{\log_2(M)\frac{E_b}{N_0}}, \tag{5.88}$$

where $\varrho(\ell)$ is given by (4.72). In (5.88) it was used that the pilot symbols energy is given by (4.40) and the data symbols energy is given by (4.41).

Notice that (5.87) follows a generalized chi-square distribution. Therefore, by using (5.87) and (5.79) in (5.15), the average SNIR can be rewritten as:

$$\gamma_s(\ell) = \frac{1}{1+\mu_s}\left[\varrho(\ell)\Omega(\ell)(K-1) + \frac{\varrho(\ell)}{\Lambda(\ell)}(N-K+1)\right]. \tag{5.89}$$

The average SNIR per block can be obtained by substituting (5.89) in (5.27). The SNIR of the MMSE detector can be well approximated by the SNIR of the ZF detector, given by (5.53).

5.3.3.4 SNIR for SPE on TICs

For SPE on TICs, Λ is given by:

$$\Lambda = \left[\overline{\alpha^2} - \varrho\right]K + \frac{1}{\log_2(M)\frac{E_b}{N_0}}, \tag{5.90}$$

where ϱ is given by (4.88). Furthermore, the average SNIR is similar to (5.89), that can be rewritten as:

$$\gamma_s = \frac{1}{1+\mu_s}\left[\varrho\Omega(K-1) + \frac{\varrho}{\Lambda}(N-K+1)\right], \tag{5.91}$$

where Λ is given by (5.90) and Ω is given by (5.77).

The average SNIR per block is also given by (5.91). Furthermore, the average SNIR can be well approximated by the ZF SNIR, given by (5.54).

5.3.3.5 SNIR for HPE on TVCs

As for MRC and ZF detectors, the SNIR for HPE is evaluated in two intervals. For the superimposed data symbols, the kth detected data symbol at the output of the MMSE detector can be obtained by substituting (5.5) in (5.64), resulting in an equivalent expression to (5.85). The conditioned SNIR for the superimposed data is similar to SPE, given by (5.86). Therefore, the average SNIR is given by (5.89), by employing $\Omega(\ell)$ given by (5.77), $\Lambda(\ell)$ given by:

$$\Lambda_{d,s}(\ell) = \left[\overline{\alpha^2} - \varrho(\ell)\right]K + \frac{1+\mu_{h,m}\eta_h}{\mu_{h,m}(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}, \tag{5.92}$$

and $\varrho(\ell)$ given by (4.78). Furthermore, the pilot symbols energy given by (4.50) and the data symbols energy given by (4.51) were used in (5.92).

For the multiplexed data symbols, the conditioned SNIR is similar to the MPE, given by (5.75). Hence, the average SNIR is similar to (5.82), by employing $\Omega(\ell)$ given by (5.77), $\Lambda(\ell)$ given by:

$$\Lambda_{d,m}(\ell) = \left[\overline{\alpha^2} - \varrho(\ell)\right] K + \frac{1 + \mu_{h,m}\eta_h}{(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}, \tag{5.93}$$

and $\varrho(\ell)$ given by (4.78). For obtaining (5.92), the multiplexed data symbols energy, given by (4.52), was employed.

The average SNIR per block can be obtained by substituting (5.89), (5.92), (5.82) and (5.93) in (5.33). Besides, the SNIR can be approximated by the SNIR of the ZF detector, given by (5.56), for the superimposed data and given by (5.58), for the multiplexed data.

5.3.3.6 SNIR for HPE on TICs

For HPE on TICs, Λ for the superimposed data is given by:

$$\Lambda_{d,s} = \left[\overline{\alpha^2} - \varrho\right] K + \frac{1 + \mu_{h,m}\eta_h}{\mu_{h,m}(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}, \tag{5.94}$$

and for the multiplexed data is given by:

$$\Lambda_{d,m} = \left[\overline{\alpha^2} - \varrho\right] K + \frac{1 + \mu_{h,m}\eta_h}{(1+\eta_h)\log_2(M)\frac{E_b}{N_0}}, \tag{5.95}$$

where ϱ is given by (4.91).

The average SNIR for superimposed data symbols is similar to SPE, it can be obtained by employing (5.94) in (5.91). On the other hand, the average SNIR for multiplexed data is similar to MPE, it can be obtained by employing (5.95) in (5.84).Notice that the average SNIR in a block is obtained by using (5.36).

Moreover, the SNIR approximation by the SNIR of the ZF detector is given by (5.59) for the superimposed data symbols and (5.60) for the multiplexed data symbols.

5.3.3.7 SNIR for PCE

The detected data symbol of the kth UT at the MMSE detector output is obtained by substituting (5.1) in (5.64) that is:

$$\hat{x}_{d,k} = \underbrace{\mathbf{a}_k \hat{\mathbf{h}}_k \frac{x_{d,k}}{2}}_{\text{Signal}} + \underbrace{\sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{a}_k \hat{\mathbf{h}}_j \frac{x_{d,j}}{2}}_{\text{MAI}} + \underbrace{\mathbf{a}_k \mathbf{w}}_{\text{Noise}}, \tag{5.96}$$

where the MAI was compensated by $\mathbf{a}_k$. Notice that there is also no CEEI. It is easy to show that the SNIR conditioned on the kth UT estimated channel vector and on the

eigenvalues is given by:

$$\gamma_{s|\hat{\mathbf{h}}_k,\lambda_j} = \sum_{j=1}^{K-1} \frac{|x_{d,k}|^2\ |\hat{h}_{j,k}|^2}{\overline{|x_d|^2}\lambda_j + 4\sigma_w^2} + \sum_{j=K}^{N} \frac{|x_{d,k}|^2\ |\hat{h}_{j,k}|^2}{4\sigma_w^2}. \tag{5.97}$$

Therefore, by using the Marchenko-Pastur distribution, the average SNIR in relation to the eigenvalues is similar to that given by (5.76), by employing Ω given by (5.77) and Λ given by:

$$\Lambda = \frac{1}{\log_2(M)\frac{E_b}{N_0}}, \tag{5.98}$$

where the data symbol energy given by (4.20) was employed.

Thus, by considering the generalized chi-square distribution, the average SNIR is given by:

$$\gamma_s = \overline{\alpha^2}\Omega(K-1) + \frac{\overline{\alpha^2}}{\Lambda}(N-K+1). \tag{5.99}$$

The SNIR can be approximated by the SNIR of the ZF detector, given by (5.62) and the average SNIR per block is also given by (5.99).

5.4 Average BER for M-MIMO

An exact expression to evaluate the average BER per block is derived. Once the SNIR and its PDF are known, it is possible to obtain the BER for the ℓth time interval by [73]:

$$\text{BER}(\ell) = \text{E}\left\{P\left(b|\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right)\right\} = \int_0^\infty P\left(b|\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right) p\left(\gamma_{s|\hat{\mathbf{h}}_k}\right) d\gamma_{s|\hat{\mathbf{h}}_k}, \tag{5.100}$$

where $P\left(b|\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right)$ is the bit error probability for the ℓth time interval conditioned on the SNIR and $p\left(\gamma_{s|\hat{\mathbf{h}}_k}\right)$ is the PDF of the SNIR conditioned on the estimated channel.

Notice that $\text{BER}(\ell)$ changes at each time slot, cause the SNIR changes in time. Thus, the average BER per block can be defined as:

$$\overline{\text{BER}} = \frac{1}{L_d}\sum_{\ell=1}^{L_d} \text{BER}(\ell), \tag{5.101}$$

where L_d is the the number of data symbols[6] and $\text{BER}(\ell)$ is the BER at the ℓth time slot.

For TICs, the SNIR is the same in all time slots for PCE, MPE and SPE, and the BER does not change in a block, that is $\text{BER}(\ell) = \overline{\text{BER}}$. For HPE, there are two SNIRs. Hence,

[6]Notice that for SPE $L_d = L$ and for HPE $L_d = L_{sd} + L_{md} = L$.

$\overline{\text{BER}} = (L_{d,s}\text{BER}_{d,s} + L_{d,m}\text{BER}_{d,m})/L_d$, where $\text{BER}_{d,s}$ is the BER of the superimposed data symbols and $\text{BER}_{d,m}$ is the BER of the multiplexed data symbols.

For M-QAM modulation, the exact bit error probability using Gray mapping is presented in [104]. By considering that the noise, the MAI and the CEEI were modeled as complex Gaussians, the exact bit error probability conditioned to the SNIR can be written as:

$$P\left(b|\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right) = \frac{1}{\log_2\sqrt{M}} \sum_{\kappa=1}^{\log_2\sqrt{M}} \frac{1}{\sqrt{M}} \sum_{i=0}^{\left(1-2^{-\kappa}\right)\sqrt{M}-1} \left\{(-1)^{\left\lfloor \frac{i\cdot 2^{\kappa-1}}{\sqrt{M}} \right\rfloor} \left(2^{\kappa-1} - \left\lfloor \frac{i\cdot 2^{\kappa-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor\right) \right.$$
$$\left. \text{erfc}\left[(2i+1)\sqrt{\frac{3}{2\left(M-1\right)}\,\gamma_{s|\hat{\mathbf{h}}_k}(\ell)}\,\right]\right\}, \tag{5.102}$$

where $\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$ is the complementary error function and $\lfloor x \rfloor$ is the greatest integer less than or equal to x [69].

The BER for the ℓth time interval can be obtained by using (5.102) in (5.100), that is:

$$\text{BER}(\ell) = \frac{2}{\log_2\sqrt{M}} \sum_{\kappa=1}^{\log_2\sqrt{M}} \frac{1}{\sqrt{M}} \sum_{i=0}^{\left(1-2^{-\kappa}\right)\sqrt{M}-1} \left\{(-1)^{\left\lfloor \frac{i\cdot 2^{\kappa-1}}{\sqrt{M}} \right\rfloor} \left(2^{\kappa-1} - \left\lfloor \frac{i\cdot 2^{\kappa-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor\right) \right.$$
$$\left. \frac{1}{2}\int_0^\infty \text{erfc}\left[\sqrt{\frac{3\left(2i+1\right)^2}{2\left(M-1\right)}\gamma_{s|\hat{\mathbf{h}}_k}}\right] p\left(\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right) d\gamma_{s|\hat{\mathbf{h}}_k}\right\}. \tag{5.103}$$

Hence, the solution of (5.103) rely on the integral:

$$\mathcal{I}(\ell,\nu_i) = \frac{1}{2}\int_0^\infty \text{erfc}\left[\sqrt{\frac{\gamma_{s|\hat{\mathbf{h}}_k}(\ell)}{\nu_i}}\right] p\left(\gamma_{s|\hat{\mathbf{h}}_k}\right) d\gamma_{s|\hat{\mathbf{h}}_k}, \tag{5.104}$$

where $\nu_i = \frac{2(M-1)}{3(2i+1)^2}$.

5.4.1 BER of MRC Detector

The BER of the MRC detector can be obtained by substituting the PDF of the SNIR, with chi-square distribution with $2N$ degrees of freedom, in (5.104). Thus, the solution of the integral (5.104) is given by [105]:

$$\mathcal{I}(\ell,\nu_i) = \rho(\ell)^N \sum_{j=0}^{N-1} \binom{N-1+j}{j} \left[1-\rho(\ell)\right]^j, \tag{5.105}$$

where $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ is the binomial expansion, $\rho(\ell) = \frac{1}{2}\left(1+\sqrt{\frac{\gamma_s(\ell)}{N\nu_i+\gamma_s(\ell)}}\right)$ is the BER with no diversity and $\gamma_s(\ell)$ is the average SNIR for the PCE, MPE, SPE or HPE derived

in Section 5.3. Finally, the BER for the ℓth time interval is given by replacing (5.105) in (5.103).

5.4.2 BER of ZF Detector

The BER of the ZF detector can be obtained by substituting the PDF of the SNIR, that is a chi-square with $2(N-K+1)$ degrees of freedom, in (5.104). Thus, the solution of (5.104) is given by:

$$\mathcal{I}(\ell,\nu_i) = \rho(\ell)^{N-K+1} \sum_{j=0}^{N-K} \binom{N-K+j}{j} [1-\rho(\ell)]^j \,, \tag{5.106}$$

where $\rho(\ell) = \frac{1}{2}\left(1+\sqrt{\frac{\gamma_s(\ell)}{(N-K+1)\nu_i+\gamma_s(\ell)}}\right)$ is the BER with no diversity and $\gamma_s(\ell)$ is the average SNIR for the PCE, MPE, SPE or HPE, derived in Section 5.3. As for the MRC detector, the BER for the ℓth time interval is obtained by replacing (5.106) in (5.103).

5.4.3 BER of MMSE Detector

For the MMSE detector, the exact solution of (5.104) is quite complex, because the SNIR has a generalized chi-square distribution. In [33], the authors have obtained a exact closed-form expression. However, a simpler solution for (5.104) is presented in [106] using the channel reliability approach for MC-CDMA systems. Since the solution presented in [106] is simpler than that of [33], it can be modified for M-MIMO systems, that is:

$$\begin{aligned}\mathcal{I}(\ell,\nu_i) = \frac{1}{2}\Bigg\{1-\sqrt{\frac{1}{2\nu_i}}\Bigg[&\sum_{j=1}^{N-K+1}\frac{1}{\Gamma(j)}\left(\frac{N-K+1}{\gamma_s(\ell)}\right)^{j-1}\mathcal{J}\left(j-\frac{1}{2},\frac{1}{\nu_i}+\frac{N-K+1}{\gamma_s(\ell)}\right)\\ &+\sum_{j=N-K+2}^{N}\frac{1}{\Gamma(j)}\left(\frac{N-K+1}{\gamma_s(\ell)}\right)^{j-1}\sum_{l=0}^{N-j}\binom{K-1}{l}\\ &\times\mathcal{L}\left(j+l-\frac{1}{2},K-1,\frac{1}{\nu_i}+\frac{N-K+1}{\gamma_s(\ell)}\right)\Bigg]\Bigg\},\end{aligned} \tag{5.107}$$

where:

$$\mathcal{J}(x,y) = \Gamma(x)y^{-x}, \tag{5.108}$$

$$\mathcal{L}(x,y,z) = \Gamma(x)\left[\frac{\Gamma(y-x)}{\Gamma(y)}\,{}_1F_1(x;x-y+1;z)+z^{y-x}\frac{\Gamma(x-y)}{\Gamma(x)}\,{}_1F_1(y;y-x+1;z)\right], \tag{5.109}$$

$\Gamma(x)$ is the gamma function, ${}_1F_1(x;y;z)$ is the confluent hyper-geometric function [98] and $\gamma_s(\ell)$ is the average SNIR of the ZF detector for PCE, MPE, SPE and HPE, derived

in Section 5.3.

5.5 BER Bounds

Lower and upper bounds for the BER can be derived to avoid the complexity of the exact closed-form expression, obtained in the former section.

5.5.1 Lower Bounds

For M-MIMO systems, due to N is large, the channel becomes deterministic as established by the channel hardening property, shown in Section 2.5.4.2. The BER becomes a negative exponential function and the Jensen's inequality[7] can be employed to obtain a simple lower bound of the BER. By applying the Jensen's inequality in (5.100), the BER lower bound for the ℓth time interval is given by:

$$\mathrm{BER}_{LB}(\ell) = P\left[b|\mathrm{E}\left\{\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right\}\right] \leqslant \mathrm{E}\left\{P(b|\gamma_{s|\hat{\mathbf{h}}_k}(\ell))\right\}. \tag{5.110}$$

Hence, by considering the lower bound in (5.102), the BER lower bound of M-MIMO systems can be rewritten as:

$$\begin{aligned}\mathrm{BER}_{LB}(\ell) = \frac{1}{\log_2\sqrt{M}} \sum_{\kappa=1}^{\log_2\sqrt{M}} \frac{1}{\sqrt{M}} \sum_{i=0}^{\left(1-2^{-\kappa}\right)\sqrt{M}-1} \Bigg\{ (-1)^{\left\lfloor \frac{i\cdot 2^{\kappa-1}}{\sqrt{M}} \right\rfloor} \\ \left(2^{\kappa-1} - \left\lfloor \frac{i\cdot 2^{\kappa-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor\right) \\ \mathrm{erfc}\left[(2i+1)\sqrt{\frac{3}{2\left(M-1\right)}\,\gamma_s(\ell)}\right]\Bigg\},\end{aligned} \tag{5.111}$$

where $\gamma_s(\ell) = \mathrm{E}\left\{\gamma_{s|\hat{\mathbf{h}}_k}(\ell)\right\}$ is the average SNIR, derived in Section 5.3.

For M-MIMO, the BER lower bound of the ZF detector is a good approximation to the BER lower bound of the MMSE detector.

5.5.2 Upper Bound

The BER upper bound is obtained by considering that $E_b/N_0 \to \infty$ in (5.103).

[7] $f\left(\mathrm{E}\{x\}\right) \leqslant \mathrm{E}\{f(x)\}$.

5.5.2.1 MRC Detector

For the MRC detector, the upper bound of (5.103), depends on the solution of (5.104), that can be approximated by:

$$\mathcal{I}(\ell,\nu_i) \approx \left(\frac{N}{4\nu_i\gamma_s(\ell)}\right)^N \binom{2N-1}{N}, \tag{5.112}$$

where series expansion of $\rho(\ell)$ is employed, that is[8]:

$$\rho(\ell) = \frac{1}{2}\left(1-\sqrt{\frac{\gamma_s(\ell)}{N\nu_i+\gamma_s(\ell)}}\right) \approx \frac{N}{4\nu_i\gamma_s(\ell)}, \tag{5.113}$$

and it was considered that $1-\rho(\ell) \simeq 1$ and that $\sum_{j=0}^{N-1}\binom{N-1+j}{j} = \binom{2N-1}{N}$.

By using (5.112) in (5.103) and taking the first term of its sum over i, the BER upper bound is given by:

$$\mathrm{BER}_{UB}(\ell) = \frac{2(\sqrt{M}-1)}{\sqrt{M}\log_2(M)}\binom{2N-1}{N} \times \left[\frac{(M-1)}{6}\frac{N}{\gamma_s(\ell)}\right]^N. \tag{5.114}$$

5.5.2.2 ZF Detector

By applying similar procedure used for the MRC detector, the solution of the integral expression given by (5.104), can be approximated to:

$$\mathcal{I}(\ell,\nu_i) \approx \left(\frac{N-K+1}{4\nu_i\gamma_s(\ell)}\right)^{N-K+1} \binom{2(N-K+1)-1}{N-K+1}, \tag{5.115}$$

Hence, by using (5.115) in (5.103) and considering the first term of its sum over i, the BER upper bound is given by:

$$\mathrm{BER}_{UB}(\ell) = \frac{2(\sqrt{M}-1)}{\sqrt{M}\log_2(M)}\binom{2(N-K+1)-1}{N-K+1} \times \left[\frac{(M-1)}{6}\frac{(N-K+1)}{\gamma_s(\ell)}\right]^{N-K+1}. \tag{5.116}$$

5.5.2.3 MMSE Detector

The upper bound of the MMSE detector can be well approximated by the ZF upper bound. However, other tighter upper bound can be obtained by using the Marchenko-Pastur PDF [67].

[8] Maclaurin series of the function $\sqrt{\frac{1}{1+x}} = 1 - \frac{x}{2} + \frac{3x^2}{8} - \frac{5x^3}{16} + \cdots$.

The solution of (5.104) is obtained by using the SNIR and the PDF approximated by the Marchenko-Pastur distribution, derived in Section 5.3.3. Thus, (5.104) can be rewritten as:

$$\mathcal{I}(\ell,\nu_i) = \prod_{m=1}^{2} \frac{1}{\varsigma_m^{2r_m}(\ell)} \sum_{k=1}^{2} \sum_{n=1}^{r_k} \psi_{k,n,\mathbf{r}}(\ell)(-1)^{r_k-n} \varsigma_k^{2(r_k-n+1)}(\ell) \rho_k^{r_k-n+1}(\ell) \sum_{j=0}^{r_k-1} \binom{r_k-n+j}{j} [1-\rho_k(\ell)]^j , \tag{5.117}$$

where $\rho_k(\ell) = \frac{1}{2}\left(1 - \sqrt{\frac{\varsigma_k^2(\ell)}{\nu_i+\varsigma_k^2(\ell)}}\right)$ is the BER without diversity and $\varsigma_k^2(\ell)$ are the variances of the chi-square random variables of the MMSE detector.

5.6 Numerical Results and Discussions

The SNIR and the BER of M-MIMO systems for TICs and TVCs considering PCE, MPE, SPE and HPE, employing MRC, ZF and MMSE detectors are evaluated. Furthermore, Monte Carlo simulations are performed in order to check the accuracy of the SNIR and BER expressions[9].

Subsection 5.6.1 presents the SNIR curves and Subsection 5.6.2 presents the BER curves. The parameters of Tab. 4.1 are used in the evaluations.

5.6.1 SNIR Evaluation

The SNIR is a useful parameter that gives an insight of the system performance. Besides, BER is a function of the SNIR. Hence, as a first approach, all SNIRs obtained in Section 5.3 are evaluated.

5.6.1.1 SNIR on TICs

Fig. 5.1, shows the average SNIR as a function of E_b/N_0 for TICs by considering that $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. PCE is also considered for comparison purposes. MRC, ZF and MMSE detectors are considered. Notice that the SNIR of the MMSE detector is slightly better than that of the ZF detector and both outperform the SNIR of the MRC detector. These performance results are expected, because MMSE and ZF detectors eliminate the MAI. Besides, the complexity of MMSE detector is higher than the other detectors. In fact, the SNIR expressions of MMSE detector are more tricky than the SNIR expressions of MRC and ZF detectors. On the other hand, as E_b/N_0 increases, MPE presents better SNIR

[9]For simplicity, only BER simulations where realized.

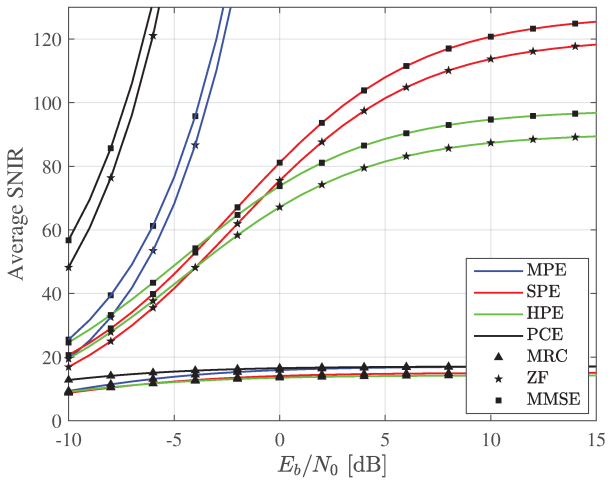


Figure 5.1: Average SNIR as a function of E_b/N_0, parameterized by the estimation techniques and the linear detectors for TICs, employing $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Besides, $L_p = 16$ and $\mu = 1$ is used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

than SPE or HPE, because there is no interference during the channel estimation process. Additionally, SPE outperforms the HPE, due to its more reliable channel estimation, once more pilot symbols are used in the channel estimation. PCE is an ideal technique and outperforms the other estimation techniques.

In Fig. 5.1 the SNIR increases with E_b/N_0. However, due to the MAI and CEEI, the SNIR may present a ceiling. The SNIR of the MRC detector is limited by the MAI. Thus, by considering that $E_b/N_0 \to \infty$ in (5.22) it is easy to show that the ceiling is equal to $N/(K-1)$ for MPE, as can be seen in Fig. 5.1. Furthermore, for PCE the ceiling is also equal to $N/(K-1)$.

Notice that the ceiling can be depicted by an asymptote that represents the maximum SNIR as $E_b/N_0 \to \infty$. In the following, the SNIR asymptotes for ZF and some channel estimation techniques are obtained.

In Fig. 5.1 it can be observed that MPE has no asymptote. On the other hand, for SPE superimposed symbols introduce interference during the channel estimation and data detection, which limits the SNIR. Thus, by considering that $E_b/N_0 \to \infty$ in (5.54), the SNIR asymptote is given by:

$$\gamma_{a,s} = \frac{(N-K+1)L\mu_s}{(1+\mu_s)K^2}. \tag{5.118}$$

Notice that the asymptote is a function of L, μ_s, N and K. Furthermore, the SNIR can be increased by augmenting L and N. Observe also that it is limited by $1/K^2$, i.e., it is limited by the number of UTs.

For HPE, there are two SNIR asymptotes, one for the superimposed data and other for the multiplexed data. Hence, by considering that $E_b/N_0 \to \infty$ in (5.59), the asymptote

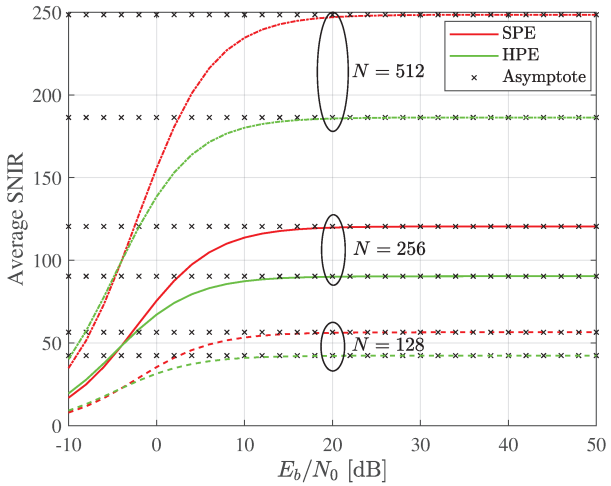


Figure 5.2: Average SNIR and its asymptote as a function of E_b/N_0 parameterized by N, employing $M = 4$, $K = 16$ and $L_d = 256$. Besides, $L_d = 256$ and $\mu_s = 1$ for SPE and $L_s = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

for the superimposed data is given by:

$$\gamma_{a_{d,s}} = \frac{(N - K + 1)L_s\mu_{h,s}}{(1 + \mu_{h,s})K^2}, \tag{5.119}$$

and by considering that $E_b/N_0 \to \infty$ in (5.60), the asymptote for the multiplexed data is given by:

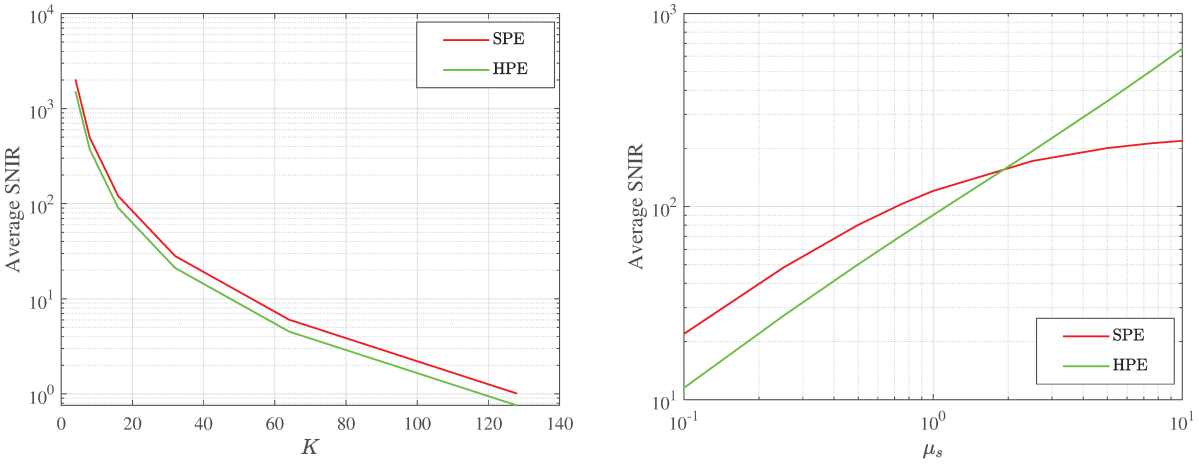
$$\gamma_{a_{d,m}} = \frac{(N - K + 1)L_s\mu_{h,s}}{K^2}. \tag{5.120}$$

Notice that the SNIR asymptote of the superimposed part of HPE is similar to SPE but as a function of L_s. On the other hand, the denominator of the SNIR asymptote of the multiplexed part does not depend on $\mu_{h,s}$. From (5.119) and (5.119), observe that the SNIR of the multiplexed part is higher than the SNIR of the superimposed part and the average of the SNIR asymptote is can be obtained by:

$$\gamma_{a,h} = \frac{(N - K + 1)L_s\mu_{h,s}}{(1 + \mu_{h,s})LK^2}\left[L(1 + \mu_{h,s}) - L_s\mu_{h,s}\right]. \tag{5.121}$$

Fig. 5.2 shows the average SNIR and its asymptote for the ZF detector, parameterized by N. Besides, $K = 16$, $L_d = 256$, $L_p = 128$ and $\mu_s = \mu_{h,s} = \mu_{h,m} = 1$ are used. Notice, that as expected the SNIR asymptotes are tighter to the SNIR as $E_b/N_0 \to \infty$. Furthermore, as the number of antennas increases, the ceiling is improved but not eliminated. Additionally, increasing the number of antennas is more effective for SPE, because the difference between SPE and HPE SNIRs increases with the number of antennas.

The SNIR asymptotes are close to the real SNIRs for high E_b/N_0, in the next two figures only the SNIR asymptotes are plotted. Fig. 5.3a shows the SNIR for high E_b/N_0



(a) Average SNIR as a function of K.

(b) Average SNIR as a function of $\mu_s = \mu_{h,s}$.

Figure 5.3: Average SNIR for high E_b/N_0 employing $N = 256$, $K = 16$, $L_d = 256$ and $L_s = 128$.

as a function of K, for $N = 256$, $L = 256$, $L_s = 128$ and $\mu_s = \mu_{h,s} = \mu_{h,m} = 1$. As expected, the SNIR increases as the number of UTs reduce.

Fig. 5.3b shows the average SNIR for high E_b/N_0 as a function of $\mu_s = \mu_{h,s}$, for $N = 256$, $K = 16$, $L_d = 256$ and $L_s = 128$. Notice that $\mu_s = \mu_{h,s}$ is the ratio between the energy of superimposed pilot and data symbols. Hence, the SNIR for high E_b/N_0 can be increased by using more energy on pilot symbols. Furthermore, for SPE, μ_s is in the denominator, which limits the SNIR. For HPE, note that $\mu_{h,s}$ is in the denominator of the superimposed region, but not in the denominator of the multiplexed region. Thus, higher average SNIR is obtained for HPE. This parameter that weigh the energies will be analyzed in the BER evaluation.

5.6.1.2 SNIR on TVCs

The SNIR on TVCs is evaluated by considering all estimation techniques and detectors.

Fig. 5.4 shows the SNIR as a function of time for $E_b/N_0 = 0$ dB, parameterized by the estimation techniques and the linear detectors. Besides, $M = 4$, $N = 256$, $K = 16$, $L_d = 256$ and UTs with a speed of $v = 6$ km/h. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. SNIR is a function of time, then, different values are obtained at each time slot. Furthermore, the SNIR is limited by the CEEI effects, which depends on the channel estimation NMSE. Therefore, the SNIR curve shape is similar to the NMSE but inverted.

Fig. 5.5 shows the average SNIR as a function of E_b/N_0, parameterized by the estimation techniques and the detectors. Furthermore, the same parameters of Fig. 5.4 are employed. As for TICs, the SNIR of the MMSE detector is superior than the ZF detector. Moreover, the SNIR of the ZF and MMSE detectors outperform the SNIR of the MRC detector because the MAI is eliminated for the ZF detector and minimized by the MMSE detector. Furthermore, as E_b/N_0 increases, the SNIR of all three estimation

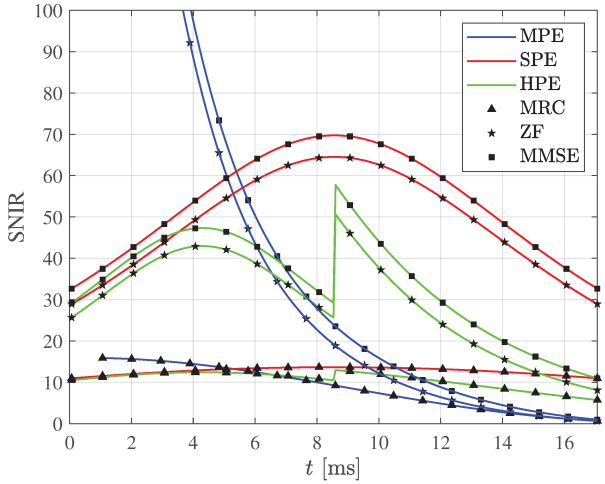


Figure 5.4: SNIR as a function of time parameterized by the estimation technique and the linear detector, employing $M = 4$, $N = 256$, $K = 16$, $L_d = 256$ and $v = 6$ km/h. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

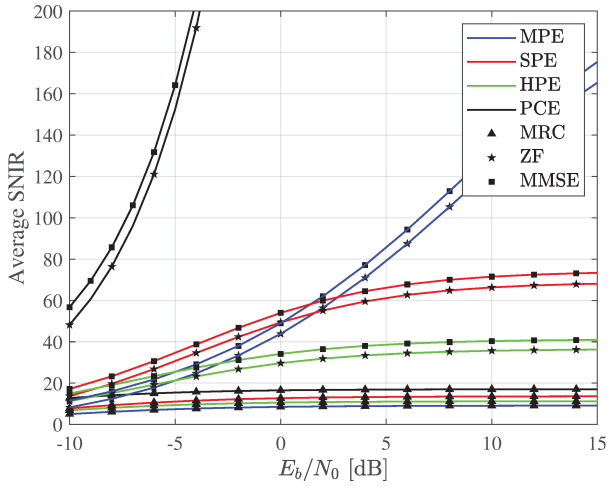


Figure 5.5: Average SNIR as a function of E_b/N_0, parameterized by the estimation techniques and linear detectors for TVCs, employing $M = 4$, $N = 256$, $K = 16$, $L_d = 256$ and $v = 6$ km/h. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

techniques presents a ceiling. This is a difference between TICs and TVCs. For TICs, ZF and MMSE detectors using MPE do not present SNIR ceiling, but for TVCs they present for high E_b/N_0, e.g., for the ZF detector the SNIR is equal to 209.5 for $E_b/N_0 = 40$ dB and equal to 209.7 for $E_b/N_0 = 50$ dB. For MRC detector, the SNIR using SPE and HPE outperforms the SNIR using MPE. These results for MPE are a consequence of the channel aging produced by the channel estimation process on TVCs. Finally, by comparing Fig. 5.5 with Fig. 5.1, the average SNIR is worse on TVCs than on TICs.

Unfortunately, the derivation of the SNIR asymptotes on TVCs is more complex than on TICs and there is no simple closed-form. However, the SNIR floor can be evaluated

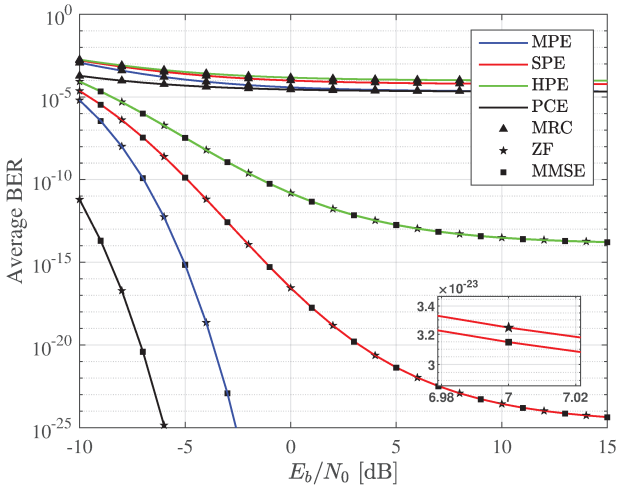


Figure 5.6: Average BER as a function of E_b/N_0, parameterized by the estimation techniques and linear detectors for TICs, employing $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

by plotting the SNIR expression. Moreover, similar conclusions to TICs can be obtained, in terms of the energy ratios and the number of UTs.

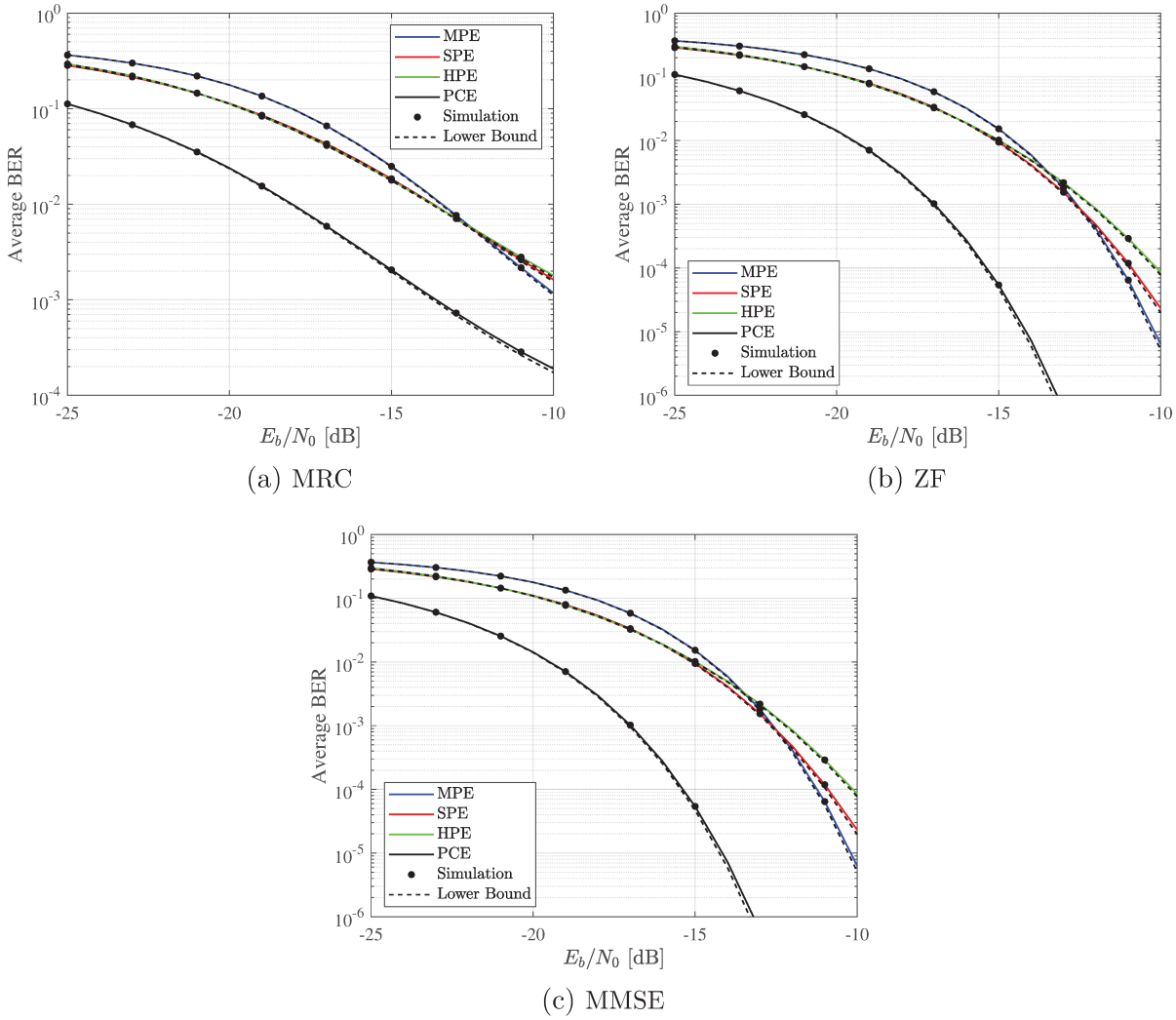
5.6.2 BER Evaluation

The BER is evaluated for M-MIMO systems on TICs and TVCs. For convenience, the BER on TICs is presented before the BER on TVCs.

5.6.2.1 BER on TICs

Fig. 5.6 shows the average BER as a function of E_b/N_0 parameterized by the estimation techniques and the linear detectors, for $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. As expected and in agreement to the SNIR results, MMSE and ZF detectors outperforms the MRC detector in terms of average BER. MMSE has a slightly better performance than ZF, as shown in the zoom inside Fig. 5.6, for SPE and $E_b/N_0 = 7$ dB. This result is expected because the SNIR of the MMSE detector is slightly higher than the SNIR of the ZF detector and the performance gain is negligible. On the other hand, MPE is better than SPE and HPE in terms of average BER and SPE outperforms HPE. However, at the cost of performance, SPE and HPE present higher spectral efficiency. Finally, in order to evaluate the loss of channel estimation techniques, observe in the same figure the BER of PCE.

In Fig. 5.6 only theoretical expressions are plotted because simulation of BER values less than 10^{-4} are difficult to obtain. Fig. 5.7a, Fig. 5.7b and Fig. 5.7c show the average BER as a function of E_b/N_0 parameterized by the estimation techniques and detectors



(a) MRC (b) ZF

(c) MMSE

Figure 5.7: Average BER as a function of E_b/N_0, parameterized by the estimation techniques for linear detectors, employing $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. BER lower bounds and simulations are included.

employing the same parameters of Fig. 5.6. Therefore, in order to validate the BER expressions using Monte Carlo simulation, smaller values of E_b/N_0 are employed in Fig. 5.7a, Fig. 5.7b and Fig. 5.7c. These figures show that the exact closed-form BER derived in Section 5.4 is accurate and that the lower bounds derived in Section 5.5 are tight to the exact closed-form BER, mainly in low E_b/N_0. In Fig. 5.7c the ZF detector lower bound is used as the lower bound of the MMSE detector. Notice that this lower bound is a good approximation for M-MIMO, once MMSE and ZF detectors present similar performance.

Fig. 5.8 shows the average BER upper bounds as a function of E_b/N_0, employing the same parameters used for Fig. 5.6. As expected, the upper bounds are tight as $E_b/N_0 \to \infty$. In M-MIMO, the BER upper bounds are not too useful for two reasons. First, the upper bound merges with the exact BER only for too small BER values. The ZF upper bound merges for a BER less than 10^{-250}, and the MMSE upper bound for a BER less than 10^{-50}. Second, the upper bounds are accurate only when the BER does

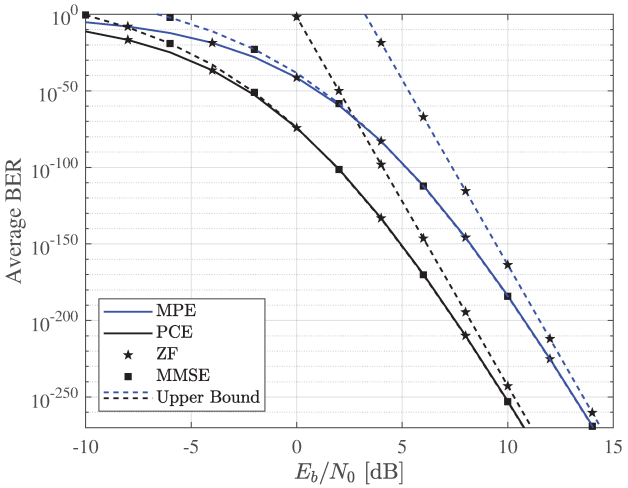


Figure 5.8: Average BER upper bound as a function of E_b/N_0 for ZF and MMSE detectors, employing $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE.

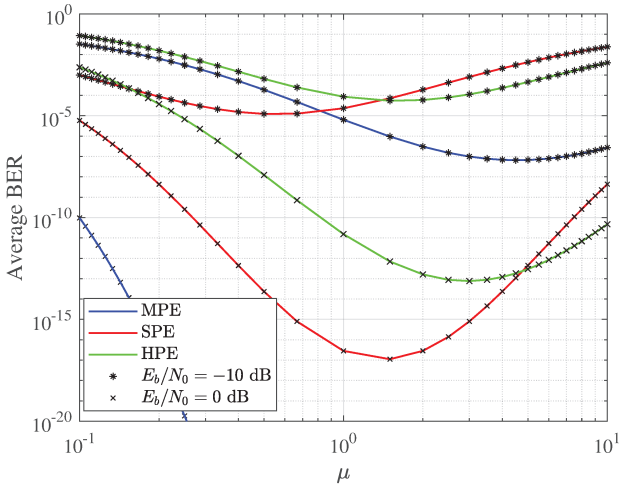


Figure 5.9: Average BER as a function of $\mu = \mu_s = \mu_{h,s} = \mu_{h,m}$ parameterized by E_b/N_0 and the estimation techniques, employing $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Besides, $L_p = 16$ is used for MPE, $L_p = 256$ for SPE and $L_p = 128$ for HPE.

not present a floor, that is, just the ZF and MMSE detector using PCE or MPE. For the cases where there is a BER floor, using the SNIR asymptotes is a good approximation, as will be shown later. In the following, only the ZF detector is analyzed because it presents similar performance to the MMSE detector and outperforms MRC detector.

Fig. 5.9 shows the average BER as function of $\mu = \mu_s = \mu_{h,s} = \mu_{h,m}$ parameterized by $E_b/N_0 = -10$ dB and $E_b/N_0 = 0$ dB, for $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Notice that μ is a parameter for MPE, μ_s is for SPE and $\mu_{h,s}$ and $\mu_{h,m}$ for HPE. Furthermore, $L_p = 16$ is used for MPE, $L_p = 256$ for SPE and $L_p = 128$ for HPE. For MPE, $\mu > 1$ is a good choice in order to improve the BER. However, there is an optimum μ, thus, increasing μ unlimitedly it is not recommended. For MPE, the optimum μ is independent

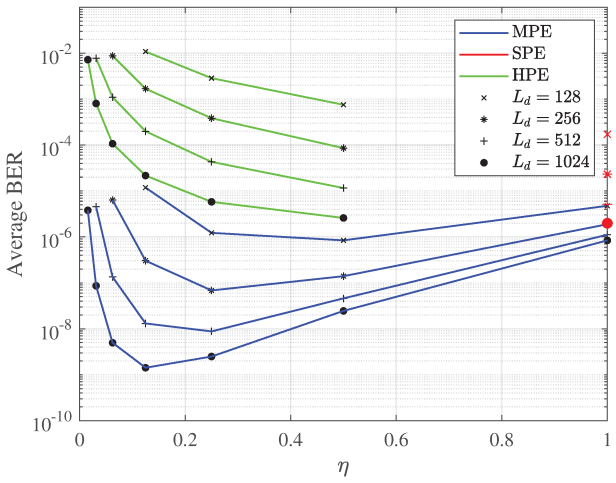


Figure 5.10: Average BER as a function of $\eta = \eta_h$, parameterized by the estimation techniques and L_d, employing $M = 4$, $N = 256$, $K = 16$, $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} = 1$, and $E_b/N_0 = -10$ dB.

of E_b/N_0 and for SPE the optimum μ_s depends on E_b/N_0. For SPE and $E_b/N_0 = -10$ dB the optimum shows up for $\mu_s < 1$ and for $E_b/N_0 = 0$ dB the optimum shows up for $\mu_s > 1$. Thus, μ_s should be selected depending on E_b/N_0 and the target BER. For HPE, the optimum shows up for $\mu_{h,s} = \mu_{h,m} > 1$, that depends on E_b/N_0. This is expected because HPE is a hybrid of both MPE and SPE.

Fig. 5.10 shows the average BER as a function of $\eta = \eta_h$, parameterized by L_d, for $M = 4$, $N = 256$, $K = 16$, $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} = 1$, and $E_b/N_0 = -10$ dB. Notice that η is the parameter for MPE and η_h is the parameter for HPE. For SPE, as $L_p = L_d$, $\eta = 1$ can be assumed. For MPE, the BER improves as L_d increases. Besides, increasing L_p is also beneficial for the BER performance, e.g., $L_d = 1024$ and $L_p = 128$ outperforms $L_d = 1024$ and $L_p = 64$ in terms of BER. However, the spectral efficiency, given by (4.33), decays as L_p increases, e.g., the spectral efficiency is 28.4 bit/s/Hz for $L_d = 1024$ and $L_p = 128$ and 30.1 bit/s/Hz for $L_d = 1024$ and $L_p = 64$. There is an optimum η and consequently L_p is bounded. For SPE, the BER improves as L_d increases. However, the improvement from $L_d = 512$ to $L_d = 1024$ is smaller than from $L_d = 256$ to $L_d = 512$. Thus, increasing unlimitedly L_d increases the complexity, delay and interference in the channel estimation and the BER decreases. For SPE, the spectral efficiency is 32 bits/s/Hz. For HPE, the BER improves as L_d increases. Additionally, in Fig. 5.10, $\eta_h = 0.5$ is the maximum value, due to the orthogonal restriction of the pilot sequences, e.g., for $L_d = 1024$ symbols, the maximum number of pilot symbols should be $L_p = 512$, because there is no Hadamard sequences of length between 512 and 1024. For HPE, the spectral efficiency is 32 bits/s/Hz. Increasing L_d is beneficial on TICs. However, on TVCs, L_d is bounded by the UTs speed and the system bandwidth.

Fig. 5.11 shows the average BER as a function of E_b/N_0 parameterized by M and the estimation techniques, for $N = 256$, $K = 16$ and $L_d = 256$. Furthermore, $L_p = 16$ and

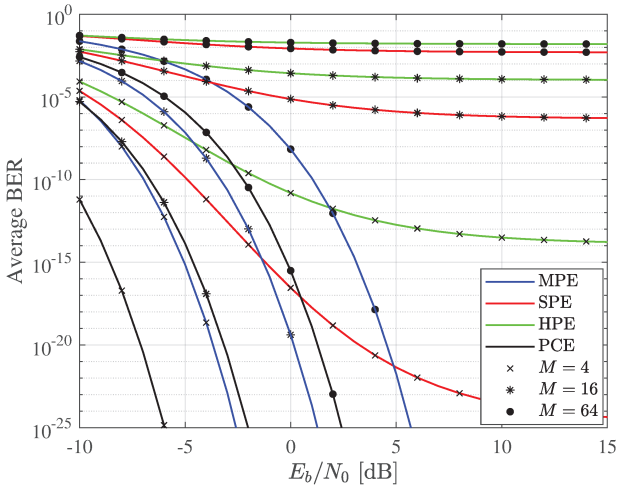


Figure 5.11: Average BER as a function of E_b/N_0, parameterized by the estimation techniques and M, employing $N = 256$, $K = 16$ and $L_d = 256$. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

$\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. As expected, as M increases, more E_b/N_0 is required to improve the BER. Notice that MPE presents the same difference in E_b/N_0 to PCE for $M = 4$, $M = 16$ and $M = 64$, e.g., for average BER of 10^{-25} the difference is 3.5 dB. For SPE, the difference in E_b/N_0 to PCE depends on the modulation order. Furthermore, there is remarkable change on the BER floor depending on the modulation order, for example, for $M = 4$ and $M = 16$, the BER difference is 10^{19}. This is a consequence of the interference introduced by the superimposed symbols in the channel estimation and data detection. For HPE, the results are similar to those observed for SPE. Despite the loss in performance, the advantage of increasing M is the gain in the spectral efficiency, e.g., for SPE with 64-QAM, the spectral efficiency is 96 bits/s/Hz for a BER near to 10^{-4}, which is quite reliable by considering that channel encoding can also be employed.

In Section 5.6.1, asymptotes of the SNIR were derived for SPE and HPE. Fig. 5.12 shows the average BER as a function of E_b/N_0 parameterized by N, for $M = 16$, $K = 16$ and $L_d = 256$. Furthermore, $L_p = 16$ and $\mu = 1$ is used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. Besides, the BER asymptotes are plotted substituting (5.118) in (5.111) for SPE and substituting (5.119) and (5.120) in (5.111) and then averaging for HPE. MPE has not BER floor, but it is included for comparison purposes. As expected, as N increases, the BER improves. Notice that 16-QAM is used to show that reliable BER can be obtained for M-MIMO. However, as shown in Chapter 3, increasing the number of antennas requires more physical space, which could be a major constraint.

Fig. 5.13 shows the average BER as a function of K for SPE and HPE, employing $M = 16$, $N = 256$ and $L_d = 256$. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE,

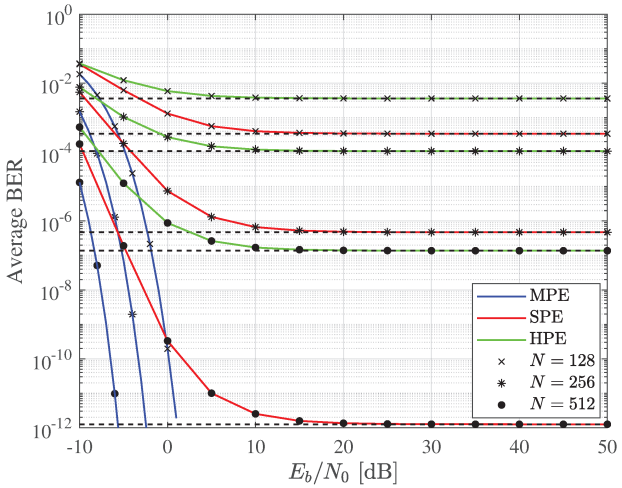


Figure 5.12: Average BER as a function of the E_b/N_0 parameterized by the estimation techniques and N, employing $M = 16$, $K = 16$ and $L_d = 256$. Besides, $L_p = 16$ and $\mu = 1$ is used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. BER asymptotes are included.

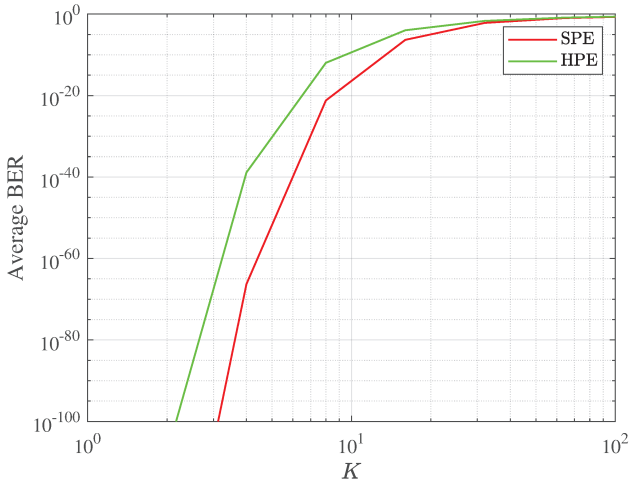


Figure 5.13: Average BER as a function of K employing the SNIR asymptotes for $M = 16$, $N = 256$ and $L_d = 256$. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

$L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. For simplicity, the SNIR asymptotes are used for evaluating the BER. Besides, MPE is not presented, once it has no BER floor. Notice that, as expected, as K decreases, the BER increases and an outstanding BER is obtained for $N \gg K$. However, there is a penalty in the spectral efficiency, given by (2.8), once it is a function of K, e.g., for $K = 4$ the spectral efficiency is 16 bits/s/Hz and for $K = 16$ is 64 bits/s/Hz.

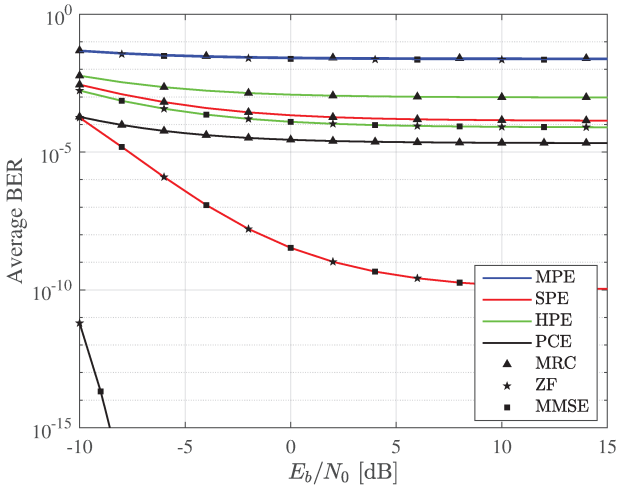


Figure 5.14: Average BER as a function of E_b/N_0, parameterized by the estimation techniques and linear detectors on TVCs, employing $M = 4$, $N = 256$, $K = 16$, $L_d = 256$ and $v = 6$ km/h. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE.

5.6.2.2 BER on TVCs

For evaluating the BER on TVCs, the speeds of Tab. 4.1 are employed. Fig. 5.14 shows the average BER as a function of E_b/N_0 parameterized by the estimation technique and detectors, for $M = 4$, $N = 256$, $K = 16$, $L_d = 256$ and $v = 6$ km/h. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE, and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. Notice that the same parameters employed in Fig. 5.6 that evaluates the average BER for TICs, are employed in Fig. 5.14. including the speed of $v = 6$ km/h. Comparing both figures, notice that the BER on TVCs is worst than on TICs. This is evidenced by the BER difference between the PCE and the other non ideal estimation techniques. As expected, the impairments introduced by TVCs during the channel estimation and data detection restrict the BER performance. Moreover, an insight of the performance restriction can be observed in the SNIR of Fig. 5.4. Observe in Fig. 5.14 that MRC detector presents the worst average BER of all the estimation techniques. Some differences between TICs and TVCs performances for ZF detector deserves attention. Notice that the average BER of MPE is widely affected by the channel aging, as it is near to the BER of MRC detector and SPE outperforms the BER of the other estimation techniques. As HPE is a hybrid of both, it presents an intermediate performance. The performance of MMSE detector is similar to ZF detector, as shown on TICs. Simulating the average BER per block is a complex task, as the BER at each time slot has to be evaluated. Therefore, in Fig. 5.14 only the theoretical BER expressions are evaluated.

Fig. 5.15 shows the BER as a function of the time, for the ZF detector, parameterized by the estimation techniques, employing $N = 256$, $K = 16$, $L_d = 256$, $M = 4$, $E_b/N_0 =$

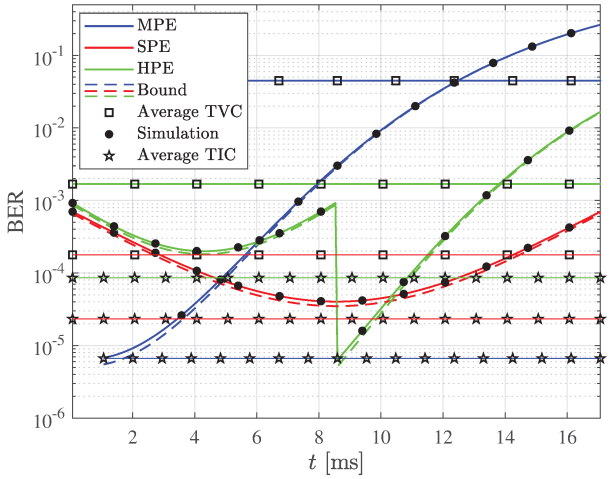
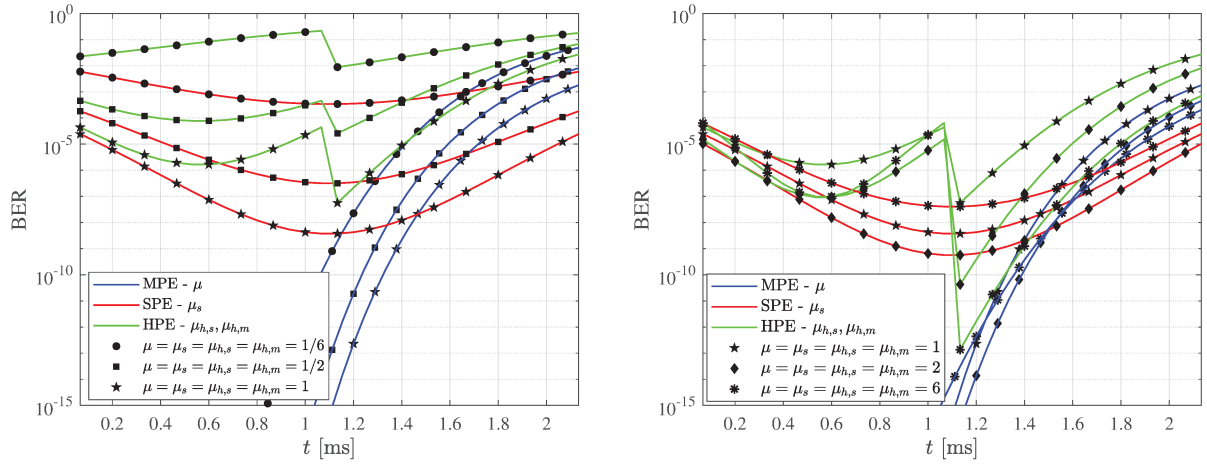


Figure 5.15: BER as a function of time, parameterized by the estimation techniques for ZF detector, employing $N = 256$, $K = 16$, $L_d = 256$, $M = 4$, $E_b/N_0 = -10$ dB and $v = 6$ km/h. Besides, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. BER lower bounds, average BER and simulation for TICs and TVCs are included.

-10 dB and $v = 6$ km/h. Furthermore, $L_p = 16$ and $\mu = 1$ are used for MPE, $L_p = 256$ and $\mu_s = 1$ for SPE and $L_p = 128$ and $\mu_{h,s} = \mu_{h,m} = 1$ for HPE. Fig. 5.15 shows the BER, given by (5.103), the average BER, given by (5.101) and the Monte Carlo simulations of the BER and average BER. The average BER on TICs is presented for comparison purposes. Notice that the simulations are restricted only to BER values above 10^{-6}, due to the simulation complexity. Despite this, the simulation confirms the precision of the derived expressions. As in Fig. 5.14, SPE presents the best average BER performance and has superior spectral efficiency than MPE. Notice that the CEEI introduced by the MMSE estimation in SPE is less harmful than the channel aging of MPE. In fact, MPE can be improved by using more pilot symbols at a cost of reducing the spectral efficiency. Moreover, the M-MIMO lower bound, given by (5.111), is also plotted. Observe that the lower bound is tight for $N \gg K$ and that there is a simpler expression than the exact one, given by (5.103). Fig. 5.15 shows that the worst performance for all three estimation techniques occurs at the end of the block, that is a consequence of the estimated channel NMSE. Notice that the average BER is near to the BER at the last time slots. Thus, the BER at the last time slots is a good reference for the average BER. Finally, notice that the average BER on TICs is far from the average BER on TVCs.

For simplicity, only the lower bounds are plotted in the following figures. Fig. 5.16a and Fig. 5.16b show the BER as a function of time, parameterized by μ for MPE, μ_s for SPE and by $\mu_{h,s} = \mu_{h,m}$ for HPE, employing $N = 256$, $K = 8$, $M = 4$, $E_b/N_0 = 0$ dB and $v = 60$ km/h. If the speed increases, it is necessary to reduce the block duration for a reliable performance. For MPE, $L_d = 32$ and $L_p = 16$ symbols are transmitted, for SPE, $L_d = L_p = 32$ and for HPE, $L_d = 32$.



(a) $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} \leqslant 1$. (b) $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} \geqslant 1$.

Figure 5.16: BER as a function of time parameterized by the estimation techniques and the ratios μ, μ_s, $\mu_{h,s}$, and $\mu_{h,m}$, employing $N = 256$, $K = 8$, $L_d = 32$, $M = 4$, $E_b/N_0 = 0$ dB and $v = 60$ km/h. Besides, $L_p = 16$ is used for MPE, $L_p = 32$ for SPE and $L_p = 16$ for HPE.

Fig. 5.16a shows the BER for ratios $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} \leqslant 1$. Observe that for $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} < 1$, the energy of pilot symbols is smaller than the energy of data symbols. Thus, the estimated channel coefficients are less reliable, degrading the BER. Therefore, for $E_b/N_0 = 0$ dB, using more energy in data symbols than in pilot symbols is not a good choice. On the other hand, Fig. 5.16b shows the BER for ratios $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} \geqslant 1$. Observe that for $\mu = \mu_s = \mu_{h,s} = \mu_{h,m} > 1$, the energy of pilot symbols is greater than the energy of data symbols. Notice that for $\mu_s = \mu_{h,s} = \mu_{h,m} = 2$, the BER of SPE and the superimposed part of HPE is improved because the estimated channel coefficients are more reliable than for $\mu_s = \mu_{h,s} = \mu_{h,m} = 1$. However, for $\mu_s = \mu_{h,s} = \mu_{h,m} = 6$, although the reliability of the estimated channel coefficients increases, the BER degrades because the energy of data symbols is excessively reduced, increasing the susceptibility to noise and interference in the detection. Besides, for MPE and the multiplexed part of HPE, increasing the ratios μ, μ_s, $\mu_{h,s}$ and $\mu_{h,m}$ is advantageous for the BER, as it effectively relieves the channel aging.

The optimal energy ratio between pilot and data symbols is not easy to be obtained and depends on the system configuration. After a complex optimization process, the best μ, μ_s, $\mu_{h,s}$ and $\mu_{h,m}$ ratios that minimize the BER can be obtained. This optimization process goes beyond the focus of this work.

In order to mitigate the effects of high UTs speed on the BER, it is not recommendable to overload the system with many users. It is desirable to maintain that $K < L_p$. Thus, in Fig. 5.15, for $K = 16$ and $L_p = 16$ for $v = 6$ km/h and in Fig. 5.16, $K = 8$ and $L_p = 16$ for $v = 60$ km/h.

Fig. 5.17 shows the BER as function of time for SPE, parameterized by L, for $N = 256$, $K = 16$, $M = 4$, $E_b/N_0 = 0$ dB, $\mu_s = 1$ and $v = 60$ km/h. As the speed is 60 km/h, the coherence time interval is $T_c = 2.6$ ms and consequently $L = 38$ symbols should be

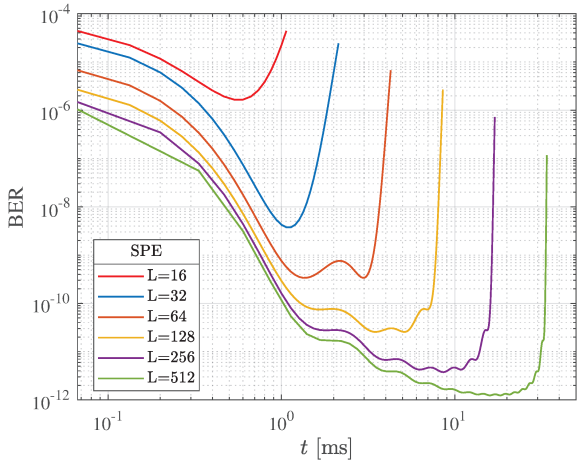


Figure 5.17: BER as a function of time for SPE parameterized by L, for $N = 256$, $K = 16$, $M = 4$, $E_b/N_0 = 0$ dB, $\mu_s = 1$ and $v = 60$ km/h.

transmitted. However, due to the pilot symbols orthogonality restriction, the block length L must be a power of two. For $L < 38$, the largest solutions are $L = 16$ and $L = 32$, where in these cases there is only one minimum in the BER curve, but for $L > 38$, the BER becomes wavier as L increases. As the block size increases, more interference is introduced in the channel estimation and data detection and the BER gain is limited. Moreover, the channel estimation process becomes more complex as the block length increases. Notice that the block length can be even larger than the coherence time interval without raising the BER, e.g., for $L = 128$, the block duration is $T_B = 8.5$ ms, which is more than three times the coherence time interval.

Fig. 5.18 shows the BER of MPE and SPE techniques as a function of E_b/N_0 parameterized by the number of BS antennas N, for $K = 8$ and $M = 4$. For convenience, the BER is evaluated at the last time slot of a block of 32 symbols for $v = 60$ km/h. Furthermore, for MPE, $\mu = 1$, $L_p = 8$ and $L_d = 32$ are used and for SPE, $\mu_s = 1$ and $L = 32$. As expected, the BER is improved as N increases due to the high diversity provided by the large number of antennas. However, there is an evident difference between MPE and SPE. As aforementioned, the channel aging introduced by MPE limits the BER severely. Moreover, increasing the number of antennas from 256 to 1024 is not enough to reach a BER lower than $\leqslant 10^{-1}$. Thus, for MPE, increasing N or E_b/N_0 is not effective due to the channel aging effects. In order to lower the BER, it is recommendable to increase the number of pilot symbols to $L_p = 16$, which is more effective than increasing further the number of antennas, as shown in Fig. 5.16b. On the other hand, increasing N is advantageous for SPE, once a significant BER improvement is obtained. Notice that ZF detector is highly susceptible to unreliable CSI due to the NMSE is less reliable for MPE and increasing the number of antennas does not solve the channel aging issues.

Fig. 5.19 shows the BER as a function of $f_{d,\max T_s}$ for $E_b/N_0 = -10$ dB, $M = 4$, $N = 256$ and $K = 16$. Furthermore, $\mu = 1$, $L_p = 16$ and $L_d = 256$ are employed for MPE

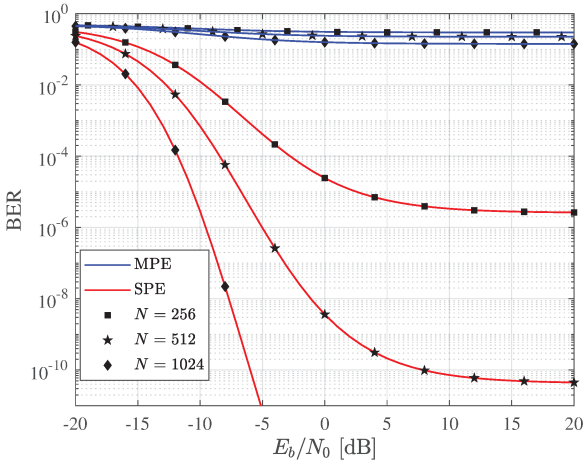


Figure 5.18: BER as a function of E_b/N_0 for the last time slot of a block, parameterized by MPE, SPE and the number of BS antennas N, for $L_d = 31$, $K = 8$, $M = 4$ and $v = 60$ km/h. Besides, $\mu = 1$ and $L_p = 8$ are used for MPE, and $\mu_s = 1$ and $L_p = 32$ for SPE.

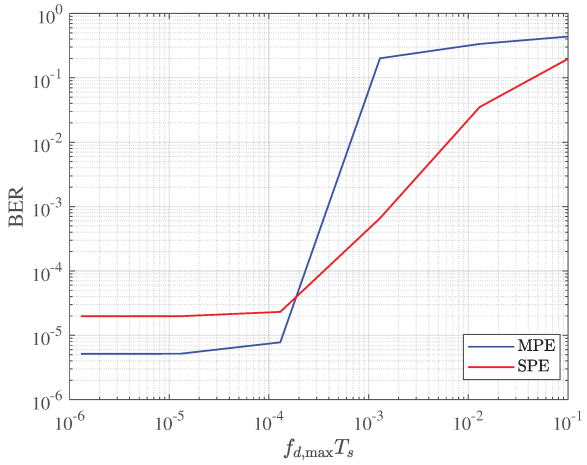


Figure 5.19: BER as a function of $f_{d,\max}T_s$, for $E_b/N_0 = -10$ dB, $M = 4$, $N = 256$, $K = 16$ and $L_d = 256$. Besides, $\mu = 1$ and $L_p = 16$ are used for MPE and $\mu_s = 1$ and $L_p = 256$ for SPE.

and $\mu_s = 1$ and $L = L_d = 256$ for SPE. For comparison purposes, the BER at the last time slot is evaluated. Notice that for $f_{d,\max T_s} = 0$ the channel becomes time-invariant and the performance is similar to TICs, as can be seen in Fig. 5.6 for $E_b/N_0 = -10$ dB. On the other hand, as $f_{d,\max T_s}$ increases, the BER worsen due to the bandwidth expansion that depends on T_s or due to the speed increase that influences on the maximum Doppler shift $f_{d,\max}$. Once the bandwidth is fixed at $B = 15$ kHz, $f_{d,\max}$ increases with speed. For 6 km/h, $f_{d,\max T_s} = 1.3 \times 10^{-3}$. Hence, Fig. 5.15 shows the BER, that for the last time slot is 6×10^{-4} for SPE and $2 \times 10{-}1$ for MPE. Notice that as the decreases, MPE outperforms SPE. Thus, MPE is more effective for TICs and SPE for TVCs.

CHAPTER 6

Conclusions and Future Works

This final chapter summarizes the conclusions of this thesis and provides the research directions for future works.

6.1 Conclusions

In this thesis, the IRF and the BER of the uplink of single cell M-MIMO systems are investigated. The first topic is the IRF, which relates the interference reduction of antennas array to the M-MIMO channel properties and links the antennas array to theoretical channel models. The second topic focuses on the BER of M-MIMO systems with imperfect channel estimation. Some channel estimation techniques and linear detectors were employed at the BS receiver by considering time-invariant and time-variant channel models.

For both topics, closed-form expressions were obtained. For IRF, the expressions are function of the number of BS antennas N, the number of UTs K and the array length $\mathcal{L}$, that depends on the carrier frequency f_c. For BER, the expressions are function of the number of BS antennas N, the number of UTs K, the modulation order M, the block length L, the number of pilot symbols L_p, the number of data symbols L_d and the ratio of pilot and data energies. Moreover, for TVCs, the BER depends also on the maximum Doppler shift $f_{D,\max}$ and the system bandwidth B. The accuracy of the derived expressions was verified by Monte Carlo simulations in representative scenarios.

The main contributions and conclusions obtained in each chapter of this thesis are summarized.

- In **Chapter** 2, the basic concepts were presented. The basis of M-MIMO systems and their properties were detailed. The principles of channel models, linear detection and channel estimation were introduced as a framework for the system model presented in the next chapters.

- In **Chapter** 3, exact closed-form expressions of the IRF for ULAs and UPAs were derived in 2-D and 3-D scenarios. The results show that better IRF is presented in 3-D scenarios, because the elevation angle can be exploited as an additional degree of freedom. If the number of antennas goes to infinity in unlimited physical space, then, the IRF goes to zero on UR-LOS and Rayleigh fading channel models. Therefore, the IRF is effectively related to the favorable propagation, because as the interference goes to zero, the channel vectors become asymptotically orthogonal.

  However, in a limited physical space, if the number of antennas increases, the spacing among antennas can be less than $\lambda_c/2$, resulting in saturation of the IRF. This is due the spatial correlation among antennas. Of course, in this case favorable propagation can not be reached. The lost of this condition depends on the separation among antennas in any frequency band (e.g., sub-6 GHz or mmWave frequency bands). Finally, the favorable propagation of M-MIMO depends on the channel model.

- In **Chapter** 4, the estimation techniques MPE, SPE and HPE were analyzed on TICs and TVCs using the MMSE estimator/predictor. These techniques are compared to PCE in terms of average symbol energy and spectral efficiency. Furthermore, the estimation quality of each technique is evaluated in terms of the NMSE, which is a function of the variance of the estimated channel coefficients.

  For TICs, MPE outperforms SPE and HPE, because it does not present a floor in the NMSE. Furthermore, SPE outperforms HPE, since it uses more pilot symbols. For TVCs, two analysis were made: for pedestrian speeds and for vehicular speeds. In general, as the speed increases, the block duration T_B should be shortened, due to the reduction of the coherence time interval. The increased speeds reduces the number of pilot symbols worsening the NMSE. The shortened block duration and the bandwidth B limit the number of time slots L. On the other hand, TICs could use infinite block size, but for a fair comparison, TICs should use the same TVCs block size.

  Furthermore, for a fair comparison of all estimation techniques, L should be a power of two in order to fulfill the requirement of orthogonal pilot symbol sequences. Additionally, they should transmit the same number of data symbols L_d in a block, reducing the spectral efficiency of MPE by a factor of $1/(1+\eta)$. SPE and HPE do not experience reduction in spectral efficiency and then they are considered bandwidth efficient estimation techniques. Moreover, the channel estimation techniques introduces a delay. The delay of MPE is of L_p symbol intervals, which is smaller than the delay of SPE, which is L symbol intervals. For HPE, the delay is L_s symbol intervals.

The derived expressions on TVCs are more complex than the expressions in TICs, because they consider the time-correlation between the channel coefficients of a block. The simplicity of the expressions on TICs is justified because only one channel coefficient must be estimated during each block, while L_d must be estimated on TVCs. Moreover, for TVCs, each estimation technique presents a particular NMSE shape, that is, a result of the MMSE estimation. As a consequence of the temporal channel correlation in TVCs, channel aging is introduced in the NMSE of MPE and HPE. For SPE and HPE, a U-shape is introduced in the NMSE. To improve the NMSE, the inequality $L_p > K$ has to be satisfied. This condition limits the spectral efficiency. An alternative to overcome this restriction, without loss in spectral efficiency, is to use $\mu > 1$, $\mu_s > 1$ and $\mu_{h,s} > 1$, which is more effective because more energy is available for channel estimation. Notice for $\mu_{h,m} > 1$ that there is a limited improvement in low E_b/N_0.

The relationships between the parameters of the estimation techniques and their effects on the system performance can not be seen only in terms of the NMSE. Therefore, the estimation error effects on the BER of M-MIMO are analyzed in Chapter 5. Notice that channel estimation is a significant challenge, further stressed in M-MIMO systems, because at the BS NK channel coefficients should be estimated on TICs and NKL_d on TVCs.

- In **Chapter** 5, the BER of M-MIMO systems was evaluated on TICs and TVCs for MPE, SPE and HPE techniques, employing MRC, ZF and MMSE detectors. For this purpose, exact-closed form expressions of the BER were derived. BER lower bounds considering the M-MIMO ideal channel properties and BER upper bounds in the high E_b/N_0 region were obtained. The BER is a function of the SNIR of each linear detector. Furthermore, the SNIR includes the interference produced by the channel estimation process, which is different for each estimation technique. The SNIR is a useful reference for the system performance, whose results are reflected on the BER.

  For TICs, the BER and average BER per block are equal, because the same estimated channel coefficients are employed for the linear detection in a block. On the other hand, the BER and the average BER per block are different for TVCs because the estimated channel coefficients are different at each time slot of a block. Notice that both the BER and NMSE have a similar shape as a function of time during a block. Moreover, the average BER per block of TVCs is near to the worse BER of a block, while the average BER of TICs is near the best BER of a block on TVCs. Hence, the average BER of M-MIMO on TVCs is far from the average BER on TICs. However, as $f_{d,\max T_s} \to 0$, the channel becomes slow and the average BER is similar to the average BER of the TICs.

Among the analyzed detectors, MMSE and ZF detectors outperform the MRC detector in terms of the BER on TICs and TVCs for any estimation technique. As expected MMSE detector outperforms slightly ZF detector and the BER difference is negligible for M-MIMO. Thus, the BER of ZF detector is a good approximation to MMSE detector in M-MIMO. Lower bounds of the BER were derived for the linear detectors considering the channel hardening property of M-MIMO. The lower bounds are tight for low E_b/N_0 and, upper bounds in the high E_b/N_0. The upper bounds are not attractive, because the BER is extremely small on the region of convergence to the exact expression due to the high diversity of M-MIMO. Moreover, an upper bound using the Marchenko Pastur distribution was derived for MMSE detector.

The average BER on TICs outperforms the average BER on TVCs. However, there are some differences between TICs and TVCs that deserve attention. Among the estimation techniques, the best average BER for TICs, is shown by MPE, while on TVCs is shown by SPE. HPE presents the worst average BER on TICs, but an intermediate average BER on TVCs. For MPE the channel aging on TVCs is so detrimental that ZF and MRC detectors have similar average BER. On the other hand, there is no channel aging on TICs, then, the average BER of ZF detector is superior than MRC. Furthermore, MPE does not have a BER floor on TICs, but has on TVCs. Indeed, the BER of MPE is strongly affected by the channel aging, even a huge increment in the number of BS antennas does not improve the BER. This observation shows that the channel aging effects limit the favorable propagation of M-MIMO systems for MPE and HPE. On the other hand, the effects of channel aging for SPE on TVCs are limited. Thus, in this case the block duration can even be larger than the coherence time interval without increasing the BER, which is not possible for MPE or HPE.

Many parameters can be modified in order to improve the performance of M-MIMO. The spectral efficiency can be improved by increasing the modulation order at the cost of reducing the BER. On the other hand, the BER can be improved by reducing the number of UTs, which reduces the spectral efficiency. For MPE, it is desirable to use $L_p > K$, which also reduces the spectral efficiency. However, the BER can be improved without effects on the spectral efficiency by modifying the energy ratio between pilot and data symbols, that is given by μ for MPE, μ_s for SPE and $\mu_{h,s}$ and $\mu_{h,m}$ for HPE. There is a trade-off between channel estimation and BER because increasing the energy of pilot symbols could be good for the channel estimation but not for BER. Hence, there is an optimum η and μ for each one of the estimation techniques, that depends on each system configuration and can be obtained after an

optimization process. The optimization process goes beyond the scope of this work but is an interesting topic for future works.

6.2 Future Works

In this section, some proposals for future research on M-MIMO systems and related topics, based on the contributions and results of this thesis, are presented.

- **Generalized Channel Models.** The loss of favorable propagation condition opens new questions about M-MIMO, such as the requirement to evaluate the effects of spatial correlation and interference among users on the performance of wireless networks [107]. The extension from 2-D to 3-D M-MIMO systems through the elevation angle deserves attention due to various strategies that can be implemented to exploit this additional degree of freedom [108].There is no channel model for M-MIMO in the literature that includes all impairments of signal propagation. Therefore, several M-MIMO models for particular scenarios could be analyzed [109]. Among them, there are the mmWaves channel model and the spatially correlated scenarios. Bands in mmWaves have been proposed for 5G systems and beyond [93]. On the other hand, it is attractive to study the spatially correlated scenario, once a recent work shows that some advantages can be obtained on the performance of M-MIMO [96].

- **Antennas Arrays.** ULAs and UPAs were analyzed, but other uniform arrays deserve attention, such as circular and cylindrical arrays [110]. Nonuniform and mimic arrays are useful for avoiding the visual pollution introduced by the antennas at the BS and their characteristics in M-MIMO should be studied [111, 112]. Moreover, there is a new topic in M-MIMO named extremely large aperture arrays or large intelligent surfaces [113], which could be analyzed in terms of the IRF.

- **Channel Estimation.** Since channel estimation is critical in M-MIMO, this topic deserves attention to further studies. One of the options is to exploit some channel properties, such as the channel sparsity and the spatial correlation [96]. Another option is to study suboptimal channel estimation in order to reduce the overload of this process, which increases with N. Finally, an option that requires much attention due to its potential is machine learning for the channel estimation [114].

- **M-MIMO Downlink.** The performance of the uplink of M-MIMO systems was presented in this dissertation. However, M-MIMO is also attractive for the downlink, once simple linear precoding and the channel reciprocity of time-division duplexing systems can be exploited. Therefore, evaluating the BER of precoders, such as the maximal-ratio-transmitter or the regularized zero-forcing considering imperfect channel estimation are also interesting options for future works [31].

- **Multicell M-MIMO.** A single cell scenario was considered in this work. In the literature, many studies claim that channel estimation is contaminated by pilots in multicell systems. They have even determined that the pilot contamination can be a performance limiting factor of M-MIMO systems [39]. However, a recent work considers that pilot contamination can be overcome in M-MIMO systems with channel correlation [96]. Hence, the study of multicell systems considering correlated channels is an interesting extension for a future development of this thesis.

- **Error Correcting Codes.** In this work, error correcting codes were not employed in the analysis. However, in practical wireless systems, these codes are essential to improve the performance. Furthermore, error correcting codes could be advantageous for both channel estimation and data detection [115]. Hence, analyzing channel estimation techniques with error correcting codes is an attractive topic for future works.

M-MIMO is a vast with high-impact topic for 5G systems and beyond. Hence, it requires further development in different areas. In the future, M-MIMO will be implemented in several wireless communication standards enabling many practical applications [116].

References

[1] Cisco. Cisco annual internet report (2018–2023). *White Paper*, pages 1–35, feb 2018.

[2] Nisha Panwar, Shantanu Sharma, and Awadhesh Kumar Singh. A survey on 5G: The next generation of mobile communication. *Physical Communication*, 18:64 – 84, 2016. Special Issue on Radio Access Network Architectures and Resource Management for 5G.

[3] Farooq Khan. *LTE for 4G Mobile Broadband*. Cambridge University Press, 2009.

[4] A. Chockalingam and B. Sundar Rajan. Introduction. In *Large MIMO Systems*, pages 1–15. Cambridge University Press, 2014.

[5] Ioannis Pefkianakis, Suk-Bok Lee, and Songwu Lu. Towards MIMO-aware 802.11n rate adaptation. *IEEE/ACM Transactions on Networking*, 21(3):692–705, jun 2013.

[6] Sassan Ahmadi. IEEE 802.16m system operation and state diagrams. In *Mobile WiMAX*, pages 97–147. Elsevier, 2011.

[7] Afif Osseiran, Federico Boccardi, Volker Braun, Katsutoshi Kusume, Patrick Marsch, Michal Maternia, Olav Queseth, Malte Schellmann, Hans Schotten, Hidekazu Taoka, Hugo Tullberg, Mikko A. Uusitalo, Bogdan Timus, and Mikael Fallgren. Scenarios for 5G mobile and wireless communications: the vision of the METIS project. *IEEE Communications Magazine*, 52(5):26–35, may 2014.

[8] Giovanni Romano. IMT-2020 requirements and realization, dec 2019.

[9] GSMA. The mobile economy 2020, 2019.

[10] Ahmad Rostami. Private 5g networks for vertical industries: Deployment and operation models. In *2019 IEEE 2nd 5G World Forum (5GWF)*. IEEE, sep 2019.

[11] Vincent W. S. Wong, Robert Schober, Derrick Wing Kwan Ng, and Li-Chun Wang, editors. *Key Technologies for 5G Wireless Systems*. Cambridge University Press, 2017.

[12] F. Rusek, D. Persson, Buon Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1):40–60, jan 2013.

[13] Thomas L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, nov 2010.

[14] Erik G. Larsson, Ove Edfors, Fredrik Tufvesson, and Thomas L. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, feb 2014.

[15] Kan Zheng, Long Zhao, Jie Mei, Bin Shao, Wei Xiang, and Lajos Hanzo. Survey of large-scale MIMO systems. *IEEE Communications Surveys & Tutorials*, 17(3):1738–1760, 2015.

[16] Emil Björnson, Erik G. Larsson, and Thomas L. Marzetta. Massive MIMO: ten myths and one critical question. *IEEE Communications Magazine*, 54(2):114–123, feb 2016.

[17] Hien Quoc Ngo, Erik G. Larsson, and Thomas L. Marzetta. Aspects of favorable propagation in massive mimo. In *European Signal Processing Conference (EUSIPCO) :*, pages 76–80. EURASIP, 2014.

[18] Thomas L. Marzetta, Erik G. Larsson, Hong Yang, and Hien Quoc Ngo. *Fundamentals of Massive MIMO*. Cambridge University Press, nov 2016.

[19] S. Kiomi Teshima and C. de Almeida. Lower bound on normalised interference mean and standard deviation for CDMA systems using antenna arrays. *Electronics Letters*, 36(21):1761, 2000.

[20] Xiang Gao, Ove Edfors, Fredrik Rusek, and Fredrik Tufvesson. Massive MIMO performance evaluation based on measured propagation data. *IEEE Transactions on Wireless Communications*, 14(7):3899–3911, jul 2015.

[21] Thomas L. Marzetta, Erik G. Larsson, Hong Yang, and Hien Quoc Ngo. The massive mimo propagation channel. In *Fundamentals of Massive MIMO*, pages 139–156. Cambridge University Press, 2016.

[22] Christos Masouros and Michail Matthaiou. Space-constrained massive MIMO: Hitting the wall of favorable propagation. *IEEE Communications Letters*, 19(5):771–774, may 2015.

[23] J. Choi and D. J. Love. Bounds on eigenvalues of a spatial correlation matrix. *IEEE Communications Letters*, 18(8):1391–1394, 2014.

[24] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta. Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. Commun.*, 61(4):1436–1449, April 2013.

[25] J. Hoydis, S. ten Brink, and M. Debbah. Massive MIMO in the UL/DL of cellu-

lar networks: How many antennas do we need? *IEEE J. Sel. Areas Commun.*, 31(2):160–171, February 2013.

[26] Ang Yang, Chengwen Xing, Zesong Fei, and Jingming Kuang. Performance analysis for uplink massive MIMO systems with a large and random number of UEs. *Science China Information Sciences*, 59(2):1–9, jan 2016.

[27] Antonios Pitarokoilis, Saif Khan Mohammed, and Erik G. Larsson. Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise. *IEEE Transactions on Wireless Communications*, 14(2):711–723, feb 2015.

[28] Sudarshan Mukherjee and Saif Khan Mohammed. Impact of frequency selectivity on the information rate performance of CFO impaired single-carrier massive MU-MIMO uplink. *IEEE Wireless Communications Letters*, 5(6):648–651, dec 2016.

[29] Juan Minango and Andrea Carolina Flores. Low-complexity MMSE detector based on refinement Jacobi method for massive MIMO uplink. *Physical Communication*, 26:128–133, feb 2018.

[30] José Carlos Marinello and Taufik Abrão. Pilot distribution optimization in multicellular large scale MIMO systems. *AEU - International Journal of Electronics and Communications*, 70(8):1094–1103, aug 2016.

[31] Long Zhao, Kan Zheng, Hang Long, and Hui Zhao. Performance analysis for downlink massive MIMO system with ZF precoding. *Transactions on Emerging Telecommunications Technologies*, 25(12):1219–1230, nov 2013.

[32] Mohammed A. Abuibaid and Sultan Aldırmaz Çolak. Energy-efficient massive MIMO system: Exploiting user location distribution variation. *AEU - International Journal of Electronics and Communications*, 72:17–25, feb 2017.

[33] Kai Zhai, Zheng Ma, and Xianfu Lei. Closed-formed distribution for the SINR of MMSE-detected MIMO systems and performance analysis. *AEU - International Journal of Electronics and Communications*, 97:16–24, dec 2018.

[34] J. Minango, C.D. Altamirano, and C. de Almeida. Performance difference between zero-forcing and maximum likelihood detectors in massive MIMO systems. *Electronics Letters*, 54(25):1464–1466, dec 2018.

[35] Mahmoud A. Albreem, Markku Juntti, and Shahriar Shahabuddin. Massive MIMO detection techniques: A survey. *IEEE Communications Surveys & Tutorials*, 21(4):3109–3132, 2019.

[36] Hien Quoc Ngo, Michail Matthaiou, Trung Q. Duong, and Erik G. Larsson. Uplink performance analysis of multicell MU-SIMO systems with ZF receivers. *IEEE Transactions on Vehicular Technology*, 62(9):4471–4483, nov 2013.

[37] Jamal Beiranvand and Hamid Meghdadi. Analytical performance evaluation of MRC receivers in massive MIMO systems. *IEEE Access*, 6:53226–53234, 2018.

[38] M. Coldrey and P. Bohlin. Training-based MIMO systems—part i: Performance comparison. *IEEE Transactions on Signal Processing*, 55(11):5464–5476, nov 2007.

[39] Olakunle Elijah, Chee Yen Leow, Tharek Abdul Rahman, Solomon Nunoo, and Solomon Zakwoi Iliya. A comprehensive survey of pilot contamination in massive MIMO—5g system. *IEEE Communications Surveys & Tutorials*, 18(2):905–923, 2016.

[40] Peter Hoeher and Fredrik Tufvesson. Channel estimation with superimposed pilot sequence applied to multi-carrier systems. In *Multi-Carrier Spread Spectrum & Related Topics*, pages 295–302. Springer US, 2000.

[41] Haidong Zhu, B. Farhang-Boroujeny, and C. Schlegel. Pilot embedding for joint channel estimation and data detection in MIMO communication systems. *IEEE Communications Letters*, 7(1):30–32, jan 2003.

[42] Han Zhang, Shan Gao, Dong Li, Hongbin Chen, and Liang Yang. On superimposed pilot for channel estimation in multicell multiuser MIMO uplink: Large system analysis. *IEEE Transactions on Vehicular Technology*, 65(3):1492–1505, mar 2016.

[43] Karthik Upadhya, Sergiy A. Vorobyov, and Mikko Vehkapera. Superimposed pilots are superior for mitigating pilot contamination in massive MIMO. *IEEE Transactions on Signal Processing*, 65(11):2917–2932, jun 2017.

[44] Changlin Guo, Jiaming Li, and Han Zhang. On superimposed pilot for channel estimation in massive MIMO uplink. *Physical Communication*, 25:483–491, dec 2017.

[45] Zahra Mokhtari, Maryam Sabbaghian, and Rui Dinis. A survey on massive MIMO systems in presence of channel and hardware impairments. *Sensors*, 19(1):164, jan 2019.

[46] Haiquan Wang, Meijun Zhou, Ruiming Chen, and Wei Zhang. On Error Rate Performance of Multi-cell Massive MIMO Systems with Linear Receivers. *PHYCOM*, 20:123 – 132, 2016.

[47] Carlos Daniel Altamirano, Juan Minango, Henry Carvajal Mora, and Celso De Almeida. BER evaluation of linear detectors in massive MIMO systems under imperfect channel estimation effects. *IEEE Access*, 7:174482–174494, 2019.

[48] Karthik Upadhya, Sergiy A. Vorobyov, and Mikko Vehkapera. Downlink performance of superimposed pilots in massive MIMO systems. *IEEE Transactions on Wireless Communications*, 17(10):6630–6644, oct 2018.

[49] Daniel Verenzuela, Emil Bjornson, and Luca Sanguinetti. Spectral and energy efficiency of superimposed pilots in uplink massive MIMO. *IEEE Transactions on Wireless Communications*, 17(11):7099–7115, nov 2018.

[50] Fei Li, Haiquan Wang, Mengyun Ying, and Wei Zhang. Channel estimation using superimposed pilots and second-order statistics for massive MIMO networks. In *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, jul 2017.

[51] Jiaming Li, Chau Yuen, Dong Li, Xianda Wu, and Han Zhang. On hybrid pilot for channel estimation in massive MIMO uplink. *IEEE Transactions on Vehicular Technology*, 68(7):6670–6685, jul 2019.

[52] J.W. Wallace and M.A. Jensen. Time-varying MIMO channels: Measurement, analysis, and modeling. *IEEE Transactions on Antennas and Propagation*, 54(11):3265–3273, nov 2006.

[53] Tatsuhiko Iwakuni, Kazuki Maruta, Atsushi Ohta, Yushi Shirato, Satoshi Kurosaki, Takuto Arai, and Masataka Iizuka. Massive MIMO effect for multiuser spatial multiplexing in time varying channel. *IEICE Communications Express*, 4(8):270–275, 2015.

[54] Kien T. Truong and Robert W. Heath. Effects of channel aging in massive MIMO systems. *Journal of Communications and Networks*, 15(4):338–351, aug 2013.

[55] Anastasios Papazafeiropoulos and Tharmalingam Ratnarajah. Linear precoding for downlink massive MIMO with delayed CSIT and channel prediction. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, apr 2014.

[56] Anastasios K. Papazafeiropoulos, Hien Quoc Ngo, and Tharmalingam Ratnarajah. Performance of massive MIMO uplink with zero-forcing receivers under delayed channels. *IEEE Transactions on Vehicular Technology*, 66(4):3158–3169, apr 2017.

[57] Chuili Kong, Caijun Zhong, Anastasios K. Papazafeiropoulos, Michail Matthaiou, and Zhaoyang Zhang. Sum-rate and power scaling of massive MIMO systems with channel aging. *IEEE Transactions on Communications*, 63(12):4879–4893, dec 2015.

[58] Wei Peng, Meng Zou, and Tao Jiang. Channel prediction in time-varying massive MIMO environments. *IEEE Access*, 5:23938–23946, 2017.

[59] Talha Younas, Jiandong Li, Muluneh Mekonnen Tulu, Hafiz Mudassir Munir, and Mahrukh Liaqat. Study of single-cell massive MIMO systems with channel aging and prediction. *Wireless Networks*, may 2019.

[60] Yuanxue Xin, Pengfei Shi, Xinjiang Xia, and Xinnan Fan. Spectral efficiency analy-

sis of multi-cell multi-user massive MIMO over channel aging. *IET Communications*, 14(5):811–817, mar 2020.

[61] Xianda Wu, Jiapeng Qiu, Yili Sun, and Han Zhang. Superimposed training-aided channel estimation for massive MIMO uplink: High mobility case. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, aug 2015.

[62] Gokhan M. Guvensen and Ender Ayanoglu. Beamspace aware adaptive channel estimation for single-carrier time-varying massive MIMO channels. In *2017 IEEE International Conference on Communications (ICC)*. IEEE, may 2017.

[63] Zahra Mokhtari, Maryam Sabbaghian, and Rui Dinis. Downlink sum-rates of SC-FDP and OFDM massive MIMO systems in doubly dispersive channels. *IEEE Transactions on Vehicular Technology*, pages 1–1, 2019.

[64] Carlos Daniel Altamirano and Celso de Almeida. Inter-user interference reduction factor for 3-d massive MIMO systems. In *2016 8th IEEE Latin-American Conference on Communications (LATINCOM)*. IEEE, nov 2016.

[65] Carlos Daniel Altamirano and Celso De Almeida. Inter-user interference reduction in massive MIMO for linear and planar arrays. *International Journal on Communications Antenna and Propagation (IRECAP)*, 9(1):30, feb 2019.

[66] Carlos Daniel Altamirano, Henry Carvajal Mora, and Celso De Almeida. BER of massive mimo in time-variant channels using multiplexed, superimposed and hybrid channel estimation techniques. *AEU - International Journal of Electronics and Communications*, XX(XX):to appear, jan 2021.

[67] Carlos Daniel Altamirano, Juan Minango, Celso de Almeida, and Nathaly Orozco. On the asymptotic BER of MMSE detector in massive MIMO systems. In *Communications in Computer and Information Science*, pages 57–68. Springer International Publishing, 2020.

[68] John R. Barry, David G. Messerschmitt, and Edward A. Lee. *Digital Communication*. Kluwer Academic Publishers, Norwell, MA, USA, third edition, 2003.

[69] J. Proakis, and D. Manolakis. *Digital Communications*. McGraw, NY, Fourth edition, 2007.

[70] Theodore S Rappaport. *Wireless communications: Principles and practice*. Prentice Hall communications engineering and emerging technologies series. Prentice Hall, Second edition, 2002.

[71] Michel Daoud Yacoub. *Wireless Technology*. CRC Press, dec 2017.

[72] Celso de Almeida. *Comunicações Digitais*. Livro não publicado, 2014.

[73] Marvin K. Simon and Mohamed-Slim Alouini. *Digital Communication over Fading Channels*. John Wiley & Sons, Inc., nov 2004.

[74] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz, A.F. Molisch, C. Oestges, and H. Özcelik. Survey of channel and radio propagation models for wireless MIMO systems. *EURASIP Journal on Wireless Communications and Networking*, 2007(1), feb 2007.

[75] 3GPP. LTE - Evolved Universal Terrestrial Radio Access (E-UTRA): Physical layer procedures. Technical report, 3GPP TS 36.213 version 13.2.0 Release 13, 2016.

[76] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, New York, NY, USA, 2005.

[77] William C. Jakes. *Microwave Mobile Communications*. IEEE, 1994.

[78] Jie Gao and Huaping Liu. Decision-directed estimation of MIMO time-varying rayleigh fading channels. *IEEE Transactions on Wireless Communications*, 4(4):1412–1417, jul 2005.

[79] Xinmin Deng, A.M. Haimovich, and J. Garcia-Frias. Decision directed iterative channel estimation for MIMO systems. In *IEEE International Conference on Communications*. IEEE, 2003.

[80] D.J. Young and N.C. Beaulieu. The generation of correlated rayleigh random variates by inverse discrete fourier transform. *IEEE Transactions on Communications*, 48(7):1114–1127, jul 2000.

[81] TS 38.101-1: NR. User equipment (ue) radio transmission and reception; part 1: Range 1 standalone. Technical report, 3GPP, 2020.

[82] Lin Bai and Jinho Choi. *Low Complexity MIMO Detection*. Springer US, 2012.

[83] Antonio M. Tulino and Sergio Verdú. Random matrix theory and wireless communications. *Commun. Inf. Theory*, 1(1):1–182, June 2004.

[84] B.M. Hochwald, T.L. Marzetta, and V. Tarokh. Multiple-antenna channel hardening and its implications for rate feedback and scheduling. *IEEE Transactions on Information Theory*, 50(9):1893–1909, sep 2004.

[85] Namshik Kim, Yusung Lee, and Hyuncheol Park. Performance Analysis of MIMO System with Linear MMSE Receiver. *IEEE Transactions on Wireless Communications*, 7(11):4474–4478, nov 2008.

[86] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, second edition, 1993.

[87] T. L. Marzetta. How much training is required for multiuser mimo? In *Proc. IEEE Asilomar Conference on Signals, Systems and Computers*, pages 359–363, Oct 2006.

[88] Xiaoyong Wu, N. C. Beaulieu, and Danpu Liu. On favorable propagation in massive MIMO systems and different antenna configurations. *IEEE Access*, pages 1–1, 2017.

[89] Sergey Loyka and Mahdi Khojastehnia. Comments on "On favorable propagation in massive MIMO systems and different antenna configurations". *IEEE Access*, 7:185369–185372, 2019.

[90] Kan Zheng, Suling Ou, and Xuefeng Yin. Massive MIMO channel models: A survey. *International Journal of Antennas and Propagation*, 2014:1–10, 2014.

[91] Theodore S. Rappaport, George R. MacCartney, Mathew K. Samimi, and Shu Sun. Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Transactions on Communications*, 63(9):3029–3056, sep 2015.

[92] S. Mumtaz, J. Rodriguez, and L. Dai. Introduction to mmWave massive MIMO. In *mmWave Massive MIMO*, pages 1–18. Elsevier, 2017.

[93] Emil Bjornson, Liesbet Van der Perre, Stefano Buzzi, and Erik G. Larsson. Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences. *IEEE Wireless Communications*, 26(2):100–108, apr 2019.

[94] Sudip Biswas, Christos Masouros, and Tharmalingam Ratnarajah. Performance analysis of large multiuser MIMO systems with space-constrained 2-d antenna arrays. *IEEE Transactions on Wireless Communications*, 15(5):3492–3505, may 2016.

[95] Adrian Garcia-Rodriguez and Christos Masouros. Exploiting the increasing correlation of space constrained massive MIMO for CSI relaxation. *IEEE Transactions on Communications*, 64(4):1572–1587, apr 2016.

[96] Luca Sanguinetti, Emil Bjornson, and Jakob Hoydis. Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination. *IEEE Transactions on Communications*, 68(1):232–257, jan 2020.

[97] Constantine A. Balanis, editor. *Modern Antenna Handbook*. Wiley, nov 2007.

[98] George Gasper and Mizan Rahman. *Basic Hypergeometric Series*. Cambridge University Press, oct 2004.

[99] T.L. Fulghum, K.J. Molnar, and A. Duel-Hallen. The jakes fading model for antenna arrays incorporating azimuth spread. *IEEE Transactions on Vehicular Technology*, 51(5):968–977, sep 2002.

[100] Daniel Altamirano and Celso Almeida. Evaluation of the effects of the co-channel interference on the bit error rate of cellular systems for BPSK modulation. In

Anais de VII International Telecommunications Symposium. Sociedade Brasileira de Telecomunicações, 2010.

[101] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.

[102] Leiba Rodman and Tamir Shalom. On inversion of symmetric toeplitz matrices. *SIAM Journal on Matrix Analysis and Applications*, 13(2):530–549, 1992.

[103] Bernard C. Levy. *Principles of Signal Detection and Parameter Estimation*. Springer US, 2008.

[104] Kyongkuk Cho and Dongweon Yoon. On the general ber expression of one- and two-dimensional amplitude modulations. *IEEE Trans. Commun.*, 50(7):1074–1080, July 2002.

[105] H. Carvajal, N. Orozco, and C. De Almeida. Mean Spectral Efficiency Evaluation of the Uplink of MC-CDMA Cellular Systems. *Wirel. Pers. Commun.*, 96(3):4595–4611, October 2017.

[106] Henry Carvajal Mora, Nathaly Orozco Garzón, and Celso de Almeida. Performance analysis of MC-CDMA cellular systems employing MMSE multiuser detector in presence of own-cell and co-cell interference. *AEU - International Journal of Electronics and Communications*, 80:19–28, oct 2017.

[107] Michail Matthaiou, Hien Quoc Ngo, Peter J. Smith, Harsh Tataria, and Shi Jin. Massive MIMO with a generalized channel model: Fundamental aspects. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, jul 2019.

[108] Hao Jiang, Yujie Xue, Jie Zhou, Zhen Chen, Jian Dang, and Liang Wu. A 3d cylinder MIMO channel model for 5g macrocell mobile-to-mobile communication systems. *IEEE Access*, 7:119350–119356, 2019.

[109] Ping Zhang, Jianqiao Chen, Xiaoli Yang, Nan Ma, and Zhi Zhang. Recent research on massive MIMO propagation channels: A survey. *IEEE Communications Magazine*, 56(12):22–29, dec 2018.

[110] Sara Gunnarsson, Jose Flordelis, Liesbet Van der Perre, and Fredrik Tufvesson. Channel hardening in massive MIMO-a measurement based analysis. In *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, jun 2018.

[111] Xiaohu Ge, Ran Zi, Haichao Wang, Jing Zhang, and Minho Jo. Multi-user massive MIMO communication systems based on irregular antenna arrays. *IEEE Transactions on Wireless Communications*, 15(8):5287–5301, aug 2016.

[112] Daniele Pinchera, Marco Migliore, Fulvio Schettino, and Gaetano Panariello. Antenna arrays for line-of-sight massive MIMO: Half wavelength is not enough. *Electronics*, 6(3):57, aug 2017.

[113] Abolfazl Amiri, Marko Angjelichinoski, Elisabeth de Carvalho, and Robert W. Heath. Extremely large aperture massive MIMO: Low complexity receiver architectures. In *2018 IEEE Globecom Workshops (GC Wkshps)*. IEEE, dec 2018.

[114] Eren Balevi, Akash Doshi, and Jeffrey G. Andrews. Massive MIMO channel estimation with an untrained deep neural network. *IEEE Transactions on Wireless Communications*, 19(3):2079–2090, mar 2020.

[115] Meixiang Zhang and Sooyoung Kim. Evaluation of MMSE-based iterative soft detection schemes for coded massive MIMO system. *IEEE Access*, 7:10166–10175, 2019.

[116] Emil Björnson, Luca Sanguinetti, Henk Wymeersch, Jakob Hoydis, and Thomas L. Marzetta. Massive MIMO is a reality—what is next? *Digital Signal Processing*, 94:3–20, nov 2019.

[117] E. Bjornson, D. Hammarwall, and B. Ottersten. Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated MIMO systems. *IEEE Transactions on Signal Processing*, 57(10):4027–4041, oct 2009.

APPENDIX A

Matrix Definitions

A.1 Hermitian and Positive Definite Matrix

A $N \times N$ complex square matrix $\mathbf{A}$ is Hermitian, if it is equal to its own conjugate transpose, that is [83]:

$$\mathbf{A} = \mathbf{A}^H, \tag{A.1}$$

where $\mathbf{A}^H$ is the conjugate transpose of matrix $\mathbf{A}$. A Hermitian matrix is said to positive-definite if:

$$\mathbf{b}^H \mathbf{A} \mathbf{b} > 0, \tag{A.2}$$

for any complex vector $\mathbf{b} \neq \mathbf{0}$ of dimension $N \times 1$.

A.2 Eigen-decomposition

The eigen-decomposition of a $N \times N$ square matrix $\mathbf{A}$ is given by [76, 83]:

$$\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}, \tag{A.3}$$

where the matrix $\mathbf{V}$ is composed by the eigenvectors of $\mathbf{A}$, and $\mathbf{\Lambda} = \text{diag}[\lambda_1,\ \lambda_2,\ \cdots\ \lambda_N]$ is a diagonal matrix that contains the eigenvalues of $\mathbf{A}$. If $\mathbf{A}$ is Hermitian, then $\mathbf{V}$ is invertible. Furthermore, if $\mathbf{A}$ is Hermitian and positive-definite (HPD), then all of its eigenvalues are positive and real.

A.3 Inverse

The inverse $\mathbf{A}^{-1}$ of $\mathbf{A}$ of dimension $N \times N$ satisfies that:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_N \tag{A.4}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix. If $\mathbf{A}^{-1}$ exists, $\mathbf{A}$ is said to be invertible. Otherwise, $\mathbf{A}$ is said to be singular or non-invertible.

A.4 Pseudoinverse

The pseudoinverse or Moore-Penrose inverse $\mathbf{A}^+$ of $\mathbf{A}$ of dimension $N \times K$ assumes that $\mathbf{A}$ is full rank. Thus, if $\text{rank}(\mathbf{A}) = N$, then it is called right-inverse that is given by:

$$\mathbf{A}^+ = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}. \tag{A.5}$$

If $\text{rank}(\mathbf{A}) = K$, then it is named left-inverse that is given by:

$$\mathbf{A}^+ = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H. \tag{A.6}$$

Finally, if $N = K$ the pseudoinverse is equal to the inverse given by (A.4).

A.5 Orthogonal Matrix

A $N \times N$ matrix $\mathbf{A}$ is said to be orthogonal, if [83]:

$$\mathbf{A}^H\mathbf{A} = m\mathbf{I}_N \tag{A.7}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix and $m > 0$ is a scalar. If $m = 1$, $\mathbf{A}$ is said to be orthonormal. To satisfy the orthogonality, all column vectors of $\mathbf{A}$ must be pairwise orthogonal; likewise for the row vectors. Moreover, the eigenvalues of an orthogonal matrix are equal, i.e., $\lambda_1 = \lambda_2 = \cdots = \lambda_K$. Therefore, an orthogonal matrix is always invertible, where its inverse is given by:

$$\mathbf{A}^{-1} = \mathbf{A}^H. \tag{A.8}$$

In particular, orthogonal matrices can be constructed from Hadamard matrices. Hadamard matrices $\mathbf{X}_n$ have dimension $2^n \times 2^n$ for $n = 1,\ 2,\ \cdots$. They are defined by the following recursive relation [69]:

$$\begin{aligned} \mathbf{X}_0 &= \begin{bmatrix} 1 \end{bmatrix}, \\ \mathbf{X}_{n+1} &= \begin{bmatrix} \mathbf{X}_n & \mathbf{X}_n \\ \mathbf{X}_n & -\mathbf{X}_n \end{bmatrix}. \end{aligned} \tag{A.9}$$

A.6 Condition Number

The condition number is a measure of how singular a matrix is. The condition number of a matrix $\mathbf{A}$ is given by [83]:

$$\kappa(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}} \tag{A.10}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the greatest and the smallest nonzero eigenvalues of $\mathbf{A}$, respectively. A condition number of one, $\kappa(\mathbf{A}) = 1$, means that $\mathbf{A}$ is orthogonal and therefore $\lambda_{\max} = \lambda_{\min}$. On the other hand, a large condition number, $\kappa(\mathbf{A}) \gg 1$, means that $\mathbf{A}$ is highly non-orthogonal or ill-conditioned.

A.7 Vector Norm

The Euclidean norm of a vector $\mathbf{b}$ of dimension $N \times 1$ is given by:

$$\|\mathbf{b}\| = \left(\mathbf{b}^H\mathbf{b}\right)^{1/2} \tag{A.11}$$

A.8 Matrix Norm

A $N \times K$ complex matrix $\mathbf{A}$ has matrix norm $\|\mathbf{A}\|$ that is a non-negative number associated with $\mathbf{A}$. In the literature some kind of matrix norms are discussed, among them, the maximum absolute column sum norm or 1-norm $\|\mathbf{A}\|_1$, the maximum absolute row sum norm or infinite-norm $\|\mathbf{A}\|_\infty$, the spectral norm or 2-norm $\|\mathbf{A}\|_2$ and the Frobenius norm $\|\mathbf{A}\|_F$. In particular, the Frobenius norm is given by:

$$\|\mathbf{A}\|_F = \sqrt{\mathrm{Tr}(\mathbf{A}^H\mathbf{A})}, \tag{A.12}$$

where $\mathrm{Tr}(\cdot)$ denotes the matrix trace.

A.9 Inversion Lemma

Let a matrix $\mathbf{A}$ of dimension $N \times K$, then the following identity holds:

$$\mathbf{A}^H\left(\mathbf{A}\mathbf{A}^H + \mathbf{I}_N\right)^{-1} = \left(\mathbf{A}^H\mathbf{A} + \mathbf{I}_K\right)^{-1}\mathbf{A}^H, \tag{A.13}$$

where $\mathbf{I}_N$ and $\mathbf{I}_K$ are identity matrices of dimension $N \times N$ and $K \times K$, respectively. An special case of the inversion lemma for vectors is given by:

$$\mathbf{z}\left(\mathbf{z}^H\mathbf{z} + \mathbf{I}_N\right)^{-1}\mathbf{z}^H = \left(\mathbf{z}\mathbf{z}^H + 1\right)^{-1}\mathbf{z}\mathbf{z}^H, \tag{A.14}$$

where $\mathbf{z}$ is a vector of dimension $1 \times N$ and $\mathbf{I}_N$ is a identity matrix of dimension $N \times N$.

APPENDIX B

Chi-square Random Variable

In this appendix, the chi-square random variable and the generalized chi-square random variable are described. Additionally, their PDFs are written based on the results of [69] and [117].

B.1 Chi-square Random Variable

Let X be a chi-square random variable, given by:

$$\begin{aligned} X &= \varsigma^2 \left(|h_1|^2 + \cdots + |h_i|^2 + \cdots |h_u|^2 \right), \\ &= \varsigma^2 \sum_{i=1}^{u} |h_i|^2 \end{aligned} \tag{B.1}$$

where ς^2 is a real number and h_i are zero-mean complex Gaussian random variables with unit variance. Thus, the factor ς^2 modifies the variance of each complex Gaussian random variables. Moreover, X has $2u$ degrees of freedom.

From [69], the PDF of this random variable can be written as:

$$p(x) = \frac{1}{\Gamma(u)\varsigma^{2u}} x^{u-1} \exp\left(-\frac{x}{\varsigma^2}\right) \tag{B.2}$$

where $\Gamma(u) = (u-1)!$ is the gamma function of u.

B.2 Generalized Chi-square Random Variable

Let X be a generalized chi-square random variable. Thus, it can be obtained from:

$$\begin{aligned} X &= \varsigma_1^2 \sum_{i=1}^{u_1} |h_i|^2 + \cdots \varsigma_\kappa^2 \sum_{i=1}^{u_\kappa} |h_i|^2 + \cdots + \varsigma_K^2 \sum_{i=1}^{u_K} |h_i|^2 \\ &= \sigma_1^2 \chi_{u_1}^2 + \cdots + \sigma_\kappa^2 \chi_{u_\kappa}^2 + \cdots + \sigma_K^2 \chi_{u_K}^2, \end{aligned} \tag{B.3}$$

where ς_κ is a real number and h_i are zero-mean complex Gaussian random variables with unit variance. Thus, the factors ς_κ^2 modify the variances of the complex Gaussian random variables. Moreover, $\chi^2_{u_\kappa}$ is a chi-square random variable with $2u_\kappa$ degrees of freedom.

From [117], the PDF of X can be written as:

$$p(x) = \Omega_K \sum_{\kappa=1}^{K} \sum_{\ell=1}^{u_\kappa} \frac{\Psi(\kappa,\ell)}{(u_\kappa-\ell)!}(-1)^{u_\kappa-1}(-x)^{u_\kappa-\ell}\exp\left(-\frac{x}{\varsigma_\kappa^2}\right), \quad x \geqslant 0, \tag{B.4}$$

where

$$\Omega_K = \prod_{\kappa=1}^{K} \varsigma_\kappa^{-2u_\kappa} \tag{B.5}$$

and

$$\Psi(\kappa,\ell) = \sum_{\forall \mathbf{p}\in\zeta_{\kappa,\ell}} \prod_{\substack{m=1 \\ m\neq\kappa}}^{K} \binom{\rho_m + u_m - 1}{\rho_m} \left(\frac{1}{\varsigma_m^2} - \frac{1}{\varsigma_\kappa^2}\right)^{-(u_m+\rho_m)}, \tag{B.6}$$

where $\zeta_{\kappa,\ell}$ is a set defined by $\zeta_{\kappa,\ell} = \left\{\mathbf{p} \in \mathbb{Z}^K; \sum_{n=1}^{K} \rho_n = \ell - 1, \rho_\kappa = 0, \rho_n \geqslant 0, \forall n\right\}$, where ρ_n is the nth element of $\mathbf{p}$. Hence, $\zeta_{\kappa,\ell}$ is the set of all partitions of length K of the integer $\ell - 1$, where the κ-th element is equal to zero. The integer partition of a positive integer x, is a way of writing x as a sum of positive integers. As example consider $K = 3$, $\kappa = 2$ and $\ell = 3$. The set of partitions of length $K = 3$ of the integer $\ell - 1 = 2$ is the following:

Position $\kappa = 1$	Position $\kappa = 2$	Position $\kappa = 3$
0	0	2
0	1	1
0	2	0
1	0	1
1	1	0
2	0	0

Therefore, the partitions whose second element ($\kappa = 2$) is equal to zero are the partitions of the first, fourth and sixth rows of the above table. Consequently, the set $\zeta_{2,3}$ can be written as $\zeta_{2,3} = \Big\{[0,0,2],[1,0,1],[2,0,0]\Big\}$.