Analysis and Optimization of Massive MIMO Systems via Random Matrix Theory Approaches

Dissertation By
Ikram Boukhedimi

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia

(June, 2019)
The dissertation of Ikram Boukhedimi is approved by the examination committee

Committee Chairperson: Prof. Mohamed-Slim Alouini
External Committee: Prof. Yonghui Li
Committee Members: Prof. Meriem Taous Laleg, Prof. Tareq AlNaffouri, Dr.Abla Kammoun
ABSTRACT

Analysis and Optimization of Massive MIMO Systems via Random Matrix Theory Approaches
Ikram Boukhedimi

By endowing the base station with hundreds of antennas and relying on spatial multiplexing, massive multiple-input–multiple-output (MIMO) allows impressive advantages in many fronts. To reduce this promising technology to reality, thorough performance analysis has to be conducted. Along this line, this work is focused on the convenient high-dimensionality of massive MIMO’s corresponding model. Indeed, the large number of antennas allows us to harness asymptotic results from Random Matrix Theory to provide accurate approximations of the main performance metrics. The derivations yield simple closed-form expressions that can be easily interpreted and manipulated in contrast to their alternative random equivalents. Accordingly, in this dissertation, we investigate and optimize the performance of massive MIMO in different contexts. First, we explore the spectral efficiency of massive MIMO in large-scale multi-tier heterogeneous networks that aim at network densification. This latter is epitomized by the joint implementation of massive MIMO and small cells to reap their benefits. Our interest is on the design of coordinated beamforming that mitigates cross-tier interference. Thus, we propose a regularized SLNR-based precoding in which the regularization factor is used to allow better resilience to channel estimation errors. Second, we move to studying massive MIMO under Line-of-Sight (LoS) propagation conditions. To this end, we carry out an analysis of the uplink (UL) of a massive MIMO system with per-user channel correlation and Rician factor. We start by analyzing conventional processing schemes such as LMMSE and MRC under training-based imperfect-channel-estimates, and then, propose a statistical combining
technique that is more suitable in LoS-prevailing environments. Finally, we look into
the interplay between LoS and the fundamental limitation of massive MIMO systems,
namely, pilot contamination. We propose to analyze and compare the performance
using single-cell and multi-cell detection methods. In this regard, the single-cell schemes
are shown to produce higher SEs as the LoS strengthens, yet remain hindered by
LoS-induced interference and pilot contamination. In contrast, for multi-cell combining,
we analytically demonstrate that M-MMSE outperforms both single-cell detectors by
generating a capacity that scales linearly with the number of antennas, and is further
enhanced with LoS.
ACKNOWLEDGMENTS

First of all, Alhamdulillah. I would like to thank my supervisor Prof. Mohamed-Slim Alouini for his advice, guidance, availability, and endless support. I would like to thank Dr. Abla Kammoun for introducing me and teaching me everything I know about Random Matrix Theory, and also for her kindness, guidance, time and efforts to always push me for what is best. I will be forever grateful as it has been a true blessing and honor to work with Prof. Alouini and Dr. Abla. I would also like to thank Prof. Mohamed-Salem Sultan for his time, advice, and very informative lectures from which I have learned a great deal. I would like to express my gratitude to the committee members, Prof. Taous-Meriem Laleg, Prof Tareq AlNaffouri, and Prof. Yonghui Li, for accepting and assessing my work, and for their constructive comments that have helped improve the quality of this dissertation.

I would like to thank the amazing group of friends I have met at KAUST. My closest squad who have been there for me during my highest and lowest moments, (in no particular order): Rayan, Ghada, Hanin, Shahrazed, Konpal, Manal and family, Marwa and family, Wafa, Fatma, Rasha, Hajar, Tsiki, and Evgeniya.

I would like to express my deepest and utmost gratitude to my biggest fans and support system, my parents, Baba and Yemma, and siblings, Zora, Mohamed (Amira, Mouadh), and Amina. Thank you for putting my education above everything. Thank you for teaching me the importance of pursuing my dreams and ambitions. Thank you for the sacrifices, the love, support, and for always being there. I also extend my gratitude to my family-in-law for their support, encouragements, and prayers.

Last but not least, I would like to thank my beloved husband, Mehdi, for being there for me, every step of the way. Thank you for making my PhD journey a little more fun. I genuinely appreciate your love, support and most importantly, your understanding.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination Committee Page</td>
<td>2</td>
</tr>
<tr>
<td>Abstract</td>
<td>4</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>6</td>
</tr>
<tr>
<td>List of Acronyms</td>
<td>10</td>
</tr>
<tr>
<td>Notations</td>
<td>11</td>
</tr>
<tr>
<td>List of Figures</td>
<td>12</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>14</td>
</tr>
<tr>
<td>1.1 Massive MIMO Systems for Future Wireless Networks</td>
<td>14</td>
</tr>
<tr>
<td>1.1.1 Overview</td>
<td>14</td>
</tr>
<tr>
<td>1.1.2 Benefits and Limitations of Massive MIMO</td>
<td>17</td>
</tr>
<tr>
<td>1.2 Motivation and Contributions</td>
<td>19</td>
</tr>
<tr>
<td>1.2.1 Massive MIMO and Coordinated Heterogeneous Networks</td>
<td>19</td>
</tr>
<tr>
<td>1.2.2 Massive MIMO with Correlated Rician Fading</td>
<td>22</td>
</tr>
<tr>
<td>1.2.3 Pilot Contamination and Line of Sight Interplay on Massive</td>
<td>24</td>
</tr>
<tr>
<td>MIMO Systems</td>
<td>26</td>
</tr>
<tr>
<td>1.3 Outline</td>
<td>26</td>
</tr>
<tr>
<td><strong>2 Random Matrix Theory for Massive MIMO Systems</strong></td>
<td>28</td>
</tr>
<tr>
<td>2.1 Some Notions of Probability Theory</td>
<td>29</td>
</tr>
<tr>
<td>2.2 Important Results from RMT</td>
<td>31</td>
</tr>
<tr>
<td>2.2.1 Lemmas and Identities</td>
<td>31</td>
</tr>
<tr>
<td>2.2.2 Deterministic Equivalents</td>
<td>33</td>
</tr>
<tr>
<td>2.3 Chapter Summary</td>
<td>35</td>
</tr>
<tr>
<td>**3 Coordinated SLNR based Precoding in Large-Scale Heterogeneous</td>
<td>36</td>
</tr>
<tr>
<td>Networks</td>
<td>36</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>36</td>
</tr>
</tbody>
</table>
3.2 System Model ......................................................... 37
   3.2.1 Transmission Model ........................................... 37
   3.2.2 Channel Model ................................................ 39
3.3 Coordinated SLNR-MAX Design ........................................ 40
   3.3.1 Asymptotic Analysis of Different Coordination Strategies .... 42
   3.3.2 Numerical Results ............................................... 50
3.4 CSI Impact on Coordination ........................................... 52
   3.4.1 Deterministic Equivalent of the SLNR ....................... 52
   3.4.2 Setting of the Coordination Factor $\beta_j$ .................... 54
   3.4.3 Numerical Results ............................................... 57
3.5 Chapter Summary ................................................... 59

4 Statistical Combining for Massive MIMO Systems with Correlated
   Rician Fading ....................................................... 60
   4.1 Introduction ........................................................ 60
   4.2 Part I - Performance Analysis in a Single-Cell Setting .......... 61
      4.2.1 Channel Estimation .......................................... 63
      4.2.2 Detection and Achievable Uplink Spectral Efficiency ......... 63
      4.2.3 Single-cell Spectral Efficiency Analysis ..................... 66
      4.2.4 Numerical Results for the Single-cell setting ................. 74
   4.3 Part II - Performance Analysis in a Multi-cell scenario ........ 78
      4.3.1 Conventional combining in Multi-Cell Systems ............... 79
      4.3.2 Statistical Combining in Multi-Cell Systems ................ 83
      4.3.3 Numerical Results for the Multi-cell scenario ............... 84
   4.4 Chapter summary .................................................. 87

5 Pilot Contamination and Line of Sight Interplay on Multi-cell Mas-
   sive MIMO Systems ............................................... 89
   5.1 System Model ..................................................... 90
   5.2 Analysis of Achievable UL Spectral Efficiencies ................. 95
      5.2.1 MRC ......................................................... 96
      5.2.2 S-MMSE ..................................................... 97
      5.2.3 M-MMSE ................................................... 99
   5.3 Numerical Results ................................................. 100
   5.4 More Results on LoS effects and the M-MMSE scheme .......... 104
      5.4.1 Asymptotic Analysis ....................................... 105
      5.4.2 Numerical Results ....................................... 107
LIST OF ACRONYMS

MIMO  Multiple-Input Multiple-Output
RMT   Random Matrix Theory
SE    Spectral Efficiency
SINR  Signal-to-Interference-and-Noise Ratio
SLNR  Signal-to-Leakage-and-Noise Ratio
MMSE  Minimum Mean Square Error
ZF    Zero Forcing
MRC   Maximum Ratio Combining
CMT   Continuous Mapping Theorem
LoS   Line-of-Sight
NLoS  Non-Line-of-Sight
i.i.d independent identically distributed
RV    Random Variable
BS    Base Station
UE    User Equipment
CSI   Channel State Information
HetNet Heterogeneous Network
DL    Downlink
UL    Uplink
TDD   Time Division Duplex
FDD   Frequency Division Duplex
PSD   Positive Semi Definite
PD    Positive Definite
## NOTATIONS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS&lt;sub&gt;j&lt;/sub&gt;</td>
<td>BS in cell j</td>
</tr>
<tr>
<td>UE&lt;sub&gt;ℓk&lt;/sub&gt;</td>
<td>k-th UE in cell ℓ</td>
</tr>
<tr>
<td>h&lt;sub&gt;jk&lt;/sub&gt;</td>
<td>Channel linking UE&lt;sub&gt;ℓk&lt;/sub&gt; to BS&lt;sub&gt;j&lt;/sub&gt;</td>
</tr>
<tr>
<td>( \hat{h}_{jk} )</td>
<td>Estimate of channel h&lt;sub&gt;jk&lt;/sub&gt;</td>
</tr>
<tr>
<td>( \vec{H}_{jk} )</td>
<td>LoS component between UE&lt;sub&gt;ℓk&lt;/sub&gt; and BS&lt;sub&gt;j&lt;/sub&gt;</td>
</tr>
<tr>
<td>( R_{jk} )</td>
<td>Spatial correlation of channel h&lt;sub&gt;jk&lt;/sub&gt;</td>
</tr>
<tr>
<td>w&lt;sub&gt;jk&lt;/sub&gt;</td>
<td>DL precoding vector for UE&lt;sub&gt;jk&lt;/sub&gt;</td>
</tr>
<tr>
<td>g&lt;sub&gt;jk&lt;/sub&gt;</td>
<td>UL combining vector for UE&lt;sub&gt;jk&lt;/sub&gt;</td>
</tr>
<tr>
<td>( \gamma_{jk} )</td>
<td>SINR associated with transmission between UE&lt;sub&gt;jk&lt;/sub&gt; and BS&lt;sub&gt;j&lt;/sub&gt;</td>
</tr>
<tr>
<td>SE&lt;sub&gt;jk&lt;/sub&gt;</td>
<td>Spectral efficiency associated with transmission between UE&lt;sub&gt;jk&lt;/sub&gt; and BS&lt;sub&gt;j&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>scalar</td>
</tr>
<tr>
<td>a</td>
<td>vector</td>
</tr>
<tr>
<td>A</td>
<td>Matrix</td>
</tr>
<tr>
<td>A&lt;sup&gt;H&lt;/sup&gt;</td>
<td>Conjugate transpose of A</td>
</tr>
<tr>
<td>tr(A)</td>
<td>Trace operator of A</td>
</tr>
<tr>
<td>[A]&lt;sub&gt;ij&lt;/sub&gt;</td>
<td>the element (i, j) of matrix A</td>
</tr>
<tr>
<td>diag {a&lt;sub&gt;i&lt;/sub&gt;}_{i=1}^{N}</td>
<td>( N \times N ) diagonal matrix with a&lt;sub&gt;i&lt;/sub&gt; being its ( i )-th diagonal element</td>
</tr>
<tr>
<td>B(A, ℓ, j)</td>
<td>( ℓj )-th block of the block matrix A</td>
</tr>
<tr>
<td>|A|</td>
<td>Spectral Norm of A</td>
</tr>
<tr>
<td>|A|_F</td>
<td>Frobenius norm of matrix A</td>
</tr>
<tr>
<td>( x_n - x \xrightarrow{a.s.}{N \to \infty} 0 )</td>
<td>Almost sure convergence of ( x_n ) to ( x ) as ( N \to \infty )</td>
</tr>
<tr>
<td>log(.)</td>
<td>The natural logarithm</td>
</tr>
<tr>
<td>( \partial_\tau )</td>
<td>Partial derivative with respect to ( \tau )</td>
</tr>
<tr>
<td>E[.]</td>
<td>Statistical expectation</td>
</tr>
<tr>
<td>( I_N )</td>
<td>( N \times N ) identity matrix</td>
</tr>
<tr>
<td>( 0_N )</td>
<td>( N \times N ) matrix with all zero entries</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

3.1 Two-Tier HetNet with legitimate and different types of interfering channels. 37
3.2 (a) Macro-cell and Micro-cell (b) achievable per-user SE (3.6) vs. SNR for No, and Full coordination with $\beta_j = 1, \forall j$ and different levels of CSI. 51
3.3 Macro-cell - Accuracy of Asymptotic SLNR (3.14) with $\beta_j = 1, \forall j$ and different levels of CSI. 57
3.4 (a) Macro-cell and (b) micro-cell spectral efficiency vs. SNR, for No, Full, and Optimal Coordination, at different levels of CSI. Solid and dashed lines represent empirical and asymptotic results, respectively. 58
3.5 Macro-cell and micro-cell optimal regularization parameter $\beta^*$ with respect to the quality of CSI $(1 - \tau^2)$, for SNR=10dB. 58

4.1 UL of a TDD single-cell massive MIMO system with scattered and specular links. 62
4.2 UL of block-fading system of coherence length $T$ symbols with pilot/data transmissions scheme. 64
4.3 Single-cell setting: (a) Impact of $\tau$ on the UL SE using conventional combining (4.5), for different levels of Rician factor, s.t. $\kappa_k \sim U(0, \kappa_{\text{max}})$. (b) Optimal number of training symbols $\tau^*$ (4.15) for different levels of $\kappa_k$. 75
4.4 Single-cell setting: UL SE using conventional combining (4.5) with optimal training $\tau^*$ and statistical combining (4.21) for different levels of $\kappa$, with (a) ordinary and (b) favorable propagation conditions. 77
4.5 Multi-cell network setup with $L=3$ cells and $K=20$ cell-edge users. 85
4.6 Multi-Cell setting: UL SE using multi-cell conventional combining (4.32) with $\tau = K$ and statistical processing (4.38) with different levels of Rician factor, in (a) ordinary and (b) favorable propagation conditions. 86
4.7 UL SEs of conventional and statistical combining in both single-cell ($L = 1$) and multi-cell ($L = 3$) settings, for different levels of Rician factor, and $K = 20$ cell-edge users. Solid and dashed lines (and arrows) correspond to the single-cell and multi-cell cases, respectively. Dotted lines represent the asymptotic approximations given in Theorems 5 - 9. 87
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>UL transmissions in a Multi-cell massive MIMO network with Rician intra-cell channels and scattered inter-cell channels</td>
</tr>
<tr>
<td>5.2</td>
<td>Network setup with L=4 cells and K=2 cell-edge users. The UEs with the same marker use the same pilot to ensure pilot contamination.</td>
</tr>
<tr>
<td>5.3</td>
<td>UL SE using MRC, S-MMSE and M-MMSE in Rayleigh ($\kappa_{jk} = 0$) and Rician fading with $\kappa_{jk} \in [0, 2]$ for a different number of antennas $N$, using the exponential channel correlation model with $r = 0.5$. Solid and dashed lines depict empirical and asymptotic results, respectively.</td>
</tr>
<tr>
<td>5.4</td>
<td>UL SE using MRC, S-MMSE and M-MMSE in Rayleigh ($\kappa_{jk} = 0$) and Rician fading with $\kappa_{jk} \in [0, 2]$ for different large-scale fading variations over the array, with $N = 200$. Solid and dashed lines depict empirical and asymptotic results, respectively.</td>
</tr>
<tr>
<td>5.5</td>
<td>$L = 4$ Multi-cell network setup with $K = 5$ UEs each.</td>
</tr>
<tr>
<td>5.6</td>
<td>UL SE using M-MMSE and MRC with respect to the number of antennas $N$, for different levels of Rician factor $\kappa_{\text{max}} \in {0, 2}$. Dotted lines depict the asymptotic approximation given in Theorem 13.</td>
</tr>
<tr>
<td>5.7</td>
<td>UL SE using M-MMSE with respect to $\kappa_{\text{max}}$ for $N = 200$, in both ordinary (Ord.P) and favorable (Fav.P) propagation conditions, where: Case I indicates both NLoS and LoS inter-cell channels, while Case II represents fully scattered inter-cell links.</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Two perennial postulates are indisputable: the demand for higher wireless data-rates will never cease, and the quantity of available spectrum will never increase. In fact, statistics \[^1\] predict that mobile data-traffic will exceed 77 exabytes per month by 2022, a seven-fold increase over 2017, at a compound annual growth rate of 46%. Plus, in light of the spectral bandwidth (BW) becoming an ever more valuable commodity for communications systems, techniques that aim at increasing the spectral efficiency are paramount. One technology that has been receiving a lot of interest over the past years in this context is known as massive Multiple-Input Multiple-Output (massive MIMO) \[^2\] which relies on the use of hundreds of antennas and spatial-multiplexing to improve the capacity, energy efficiency and reliability of the channels \[^3\,4\]. It represents an advanced design of its fore versions, namely Point-to-Point MIMO and multi-user (MU) MIMO.

1.1 Massive MIMO Systems for Future Wireless Networks

1.1.1 Overview

MIMO technology has become fundamental in wireless networks for the many advantages it provides namely, the capacity gains and reliability of the network \[^3\,4\]. Initially, the technology involved the classical Point-to-Point MIMO wherein both end devices (transmitter and receiver) are equipped with a compact array of multiple antennas. It has been used in the 802.11ac standard with up to eight antennas at both ends.
Intuitively, the more antennas put to use, the higher the throughput gain. However, studies showed that in Point-to-Point MIMO, the amount of time spent in training is proportional to the number of antennas which represents an important limiting factor for this technology. Another major drawback is the complexity of the receiver and the involved signal processing required to achieve the promised performances. Additionally, Point-to-Point MIMO requires favorable propagation environments and a good signal-to-noise-ratio (SNR) and performs poorly under Line-of-Sight conditions [5,6].

Over the past years, MIMO matured significantly and the focus shifted towards multi-user MIMO (MU-MIMO) systems, which are equivalent to a Point-to-Point link on the transmitter side (BS with multiple antennas), however, communicating simultaneously with a set of single antenna users [7]. Interestingly, “breaking” the receiver’s compact array to several, geographically spaced, mono-antenna terminals yields many impressive advantages. First, the complexity of the UE is drastically reduced since it can be replaced by relatively cheap single-antenna devices [3]. Second, due to the multi-user diversity, the performance of MU-MIMO systems is less sensitive to the propagation environment and thus does not require rich scattering conditions in contrast to Point-to-Point MIMO. As a result, MU-MIMO technology was incorporated in most recent wireless communication standards such as 4G-LTE and LET-Advanced. On another note, the more antennas in the base station, the better performances in terms of achievable rates, reliability and energy efficiency. However, similarly to Point-to-Point MIMO, this comes at the cost of complex signal processing techniques. In fact, it requires the use of dirty paper coding/decoding whose computational burden grows exponentially with the dimensions of the system. Consequently, most current MIMO implementations are still modest with only up to eight antennas in the BS [5,6].

The ever-increasing amount of wireless data is expected to reach 492 exabytes by 2020 and exceed 900 exabytes by 2022 [1]. Mainly driven by video and real-time applications, the data overflow and the growing number of wireless devices cannot
be accommodated by current technologies. Consequently, for the next generation of wireless networks, research witnessed an extensive effort towards developing novel strategies to overcome these challenges along with the current limitations of the available systems such as coverage, energy efficiency, complex signal processing, to name a few. Among the different technologies in this context, massive MIMO is considered as one of the most promising. Also referred to as large-scale antenna systems, this technology is expected to guarantee significant gains in data rates, link reliability and energy efficiency by using a large number of antennas [8–12].

The basic premise behind massive MIMO is to reap all the benefits of conventional MIMO, but on a much greater scale. Accordingly, the BS is equipped with hundreds of antennas to simultaneously serve many tens of terminals in the same time-frequency resource, using spatial multiplexing. One should note that implementing hundreds of antennas does not imply a huge physical size of the array. In fact, the first designs managed to confine up to 128 antennas in configurations of size $28 \times 28 \text{ cm}$ [4,5]. To carry out the multiplexing, channel state information (CSI) in both the uplink and downlink is necessary. For the UL, this is usually achieved through the use of pilot sequences sent from the UEs to the BS. Likewise, the UL of a massive MIMO system can be estimated in a similar manner. On the other hand, since the number of antennas is much greater than that of served users, the same cannot be accomplished for the DL in massive MIMO since the amount of CSI feedback will be proportional to the number of antennas. Furthermore, noting that optimal DL pilots should be mutually orthogonal (between antennas), the number of time-frequency blocks needed for pilots in massive MIMO will be as large as the number of antennas in the BS. As a result, in contrast to conventional MIMO systems, a hundred times more resources would be needed to acquire proper CSI of the DL of a massive MIMO network. Therefore, the wide disparity between the number of antennas and users requires the reciprocity of the UL/DL channels yielding the necessity to operate in Time Division Duplexing
(TDD) mode. Although, most studies agree that TDD mode is a must for massive MIMO systems, [13] presents the possibility of operating with FDD under certain channel conditions.

1.1.2 Benefits and Limitations of Massive MIMO

The implementation of massive MIMO in wireless networks has the potential to achieve dramatic improvements in many fronts. First, basic asymptotic arguments such as the strong law of large numbers show that both noise and interference vanish as the number of antennas grows large by the use of rudimentary linear signal processing techniques like matched-filter (MF) precoding/decoding [2,14,15]. This comes in contrast to conventional MIMO systems that require computationally involved dirty paper coding. Plus, assuming channel reciprocity (TDD) for massive MIMO enables to add as many antennas as desired without any training cost (training overhead is proportional to number of UEs) [16]. Second, by equipping the BS with a large number of antennas and using spatial multiplexing, both spectral and energy efficiency can be greatly enhanced. In [17], the authors show that the transmit power in massive MIMO systems is inversely proportional to the number of antennas employed under perfect CSI conditions. Actually, scaling down the transmit power may open the door to solar or wind powered BSs yielding the reduction of BSs energy consumption which constitutes a worldwide growing concern [5]. Another interesting advantage of massive MIMO is its reliability. Indeed, using hundreds of antennas provides more resilience to system impairments or failure of one or a few antennas [3].

On the other hand, every technology has its limitations and every gain comes at a certain cost. Despite the substantial advantages of massive MIMO, many issues still need to be tackled. In this line, the main challenge that massive MIMO faces is the phenomenon of pilot contamination [3,18,21]. In all the above, we discussed massive MIMO at the BS level or in a single-cell scenario. In dense areas, multiple
BSs are deployed to ensure users QoS demands. Plus, due to spectrum limitations, many BSs operate in the same time-frequency blocks. In such settings, we speak of multi-cell networks. Recalling that in TDD mode, to properly perform the spatial multiplexing, massive MIMO depends highly on accurate CSI acquired through the use of orthogonal UL pilot sequences sent out by UEs. Ideally, every terminal is assigned an orthogonal uplink pilot sequence. However, the number of these latter is limited by the channel coherence interval (only 200 sequences for a one millisecond coherence interval \(2\)) which entails the reuse of pilot sequences among cells. Consequently, the obtained channel estimates are “contaminated” by pilots sent by other users that share the same pilot sequence, hence the denomination “pilot contamination”. Furthermore, when the contaminated estimates are used to design the precoding or combining vector, interference emanates at the other UEs using the same pilot sequences. Unlike the intra-cell interference caused by UEs in the same cell, increasing the number of antennas does not eliminate the damaging effects of pilot contamination. In this line, considerable efforts have been made to reduce these negative repercussions, namely blind channel estimation methods to eliminate the use of pilots \([19,20,22,23]\), novel precoding techniques \([21]\), or multi-cell processing schemes to mitigate the interference due to pilot contamination \([24]\), as shall be seen in further detail in Chapter 5.

Other concerns rise in the road to making massive MIMO a reality and are yet to be investigated and handled; among these, channel reciprocity. Notwithstanding its numerous advantages, the fact that TDD is indispensable for massive MIMO to meet the promised gains make of it a restriction and a limitation. To this day, many research efforts are being done to enable FDD massive MIMO techniques \([25,27]\). Finally, all potential advantages (and challenges) of massive MIMO cited thus far were discussed in an isotropic (rich) scattering environment which, mathematically, is modeled via Rayleigh fading channels. Nevertheless, such models do not incarnate the Line-of-Sight (LoS) components which are commonly present in practical wireless
propagation scenarios. Since massive MIMO is becoming an integral part of future communication networks, further analysis ought to be pursued to ensure its resilience and performance in such propagation conditions (as shall be explored in Chapters 4 and 5).

1.2 Motivation and Contributions

By equipping the base station (BS) with a massive number of antennas to simultaneously serve many tens of users, massive MIMO allows for orders of magnitude of improvement in spectral and energy efficiency, using fairly rudimentary linear signal processing. Considering the numerous advantages massive MIMO promises, thorough performance analysis needs to be conducted to exploit its full potential. Towards this aim, Random Matrix Theory (RMT) appears as a suitable solution in the heart of a “large” challenge. Indeed, as massive MIMO uses antennas in the order of hundreds or even thousands, its corresponding system model is of high dimensions. Accordingly, the resulting mathematical frame consists of large random terms that are quite arduous to investigate or manipulate and thus, can only be analyzed through heavy simulations. Such procedures not only require high computational complexity but also do not lend themselves to simple interpretation. On the other hand, the high-dimensional nature of the system renders RMT’s asymptotic based results applicable and close to reality. In this line, leveraging RMT tools to investigate the performance and attempt to achieve further enhancement for massive MIMO systems constitute the focus of our dissertation.

1.2.1 Massive MIMO and Coordinated Heterogeneous Networks

With the ever-increasing demand of mobile data traffic, future 5G systems are required to support higher data rates and wider coverages [8, 9, 28, 29]. Obviously, these
requirements cannot be accommodated by current technologies. In this context, there is a broad consensus that the demands of 5G systems can only be met through a substantial network densification \cite{8,10,28,30}, which can be reduced into practice through multi-antenna tier Heterogeneous networks (HetNets). As shall be seen in Chapter 3, such networks consist of a significant number of small cells deployed in the macro-cell range and coupled with massive MIMO technology \cite{10,31}. This latter aims to ensure a wide coverage and to support high-mobility by equipping macro-cell base stations (BSs) with very large antenna arrays. In addition, the dense deployment of low-cost and low-power micro-cells \cite{10,31} allows to bring the user equipment (UE) and its serving station close to one another. This has been shown to enhance indoor coverage and improve the cell-edge performance \cite{32,33}. However, if these small cells operate over the same spectrum as the macro-cell (which is the case for cellular networks), important intra-cell and cross-tier interference emanate \cite{30}, thereby affecting seriously the gain achieved by the aforementioned technologies \cite{33}.

To mitigate this interference, coordinated beamforming based techniques have received a lot of interest \cite{34-38}. In essence, the idea is to use precoding schemes that exploit information about the interference that each BS causes to the neighboring cells \cite{39}. In TDD systems, the coordinating BSs will acquire CSI by means of pilot sequences sent from UEs to their serving BS. This is arguably less demanding than joint processing techniques (see \cite{40} and the references therein), which require BSs to share both CSI and data. Usually, the design of coordinated beamforming is based on one of the following criteria. One can consider the minimization of the transmit power under some target constraints on the signal to noise ratio (SINR) \cite{41}, or the complete elimination of the interference through the use of Zero-Forcing beamforming (ZFBF), \cite{42}. The design criterion that is considered in this work is the Signal-to-Leakage-and-Noise ratio (SLNR), which uses the concept of leakage. This latter refers to the amount of interference caused by the signal intended for the UE of interest on
the remaining users; in contrast to conventional SINR based schemes which involve the interference suffered by the desired UE.

Prior to our work, the use of the SLNR notion for precoding design was mainly represented by [43–48]. However, most of these works consider single-cell multi-user systems or non-coordinated scenarios, and focus on the perfect CSI case. Under such settings, it was already shown in [43] that the resulting beamforming that maximizes the SLNR coincides with well-known transmit Wiener filter [42]. In [48], the authors carry out a simulation-based analysis of coordinated SLNR based precoding in HetNets. In the work presented in Chapter 3, we conduct a theoretical analysis of the design of SLNR based precoding for HetNets under the assumption of imperfect CSI. Specifically, we consider a HetNet in which each BS employs a regularized variant of the classical SLNR based precoding, which we refer to as SLNR-MAX precoding. The novelty of our proposed scheme lies in the use of a regularization parameter that weights the other cells’ information used in the precoding. Intuitively, the setting of this weight should depend on the reliability of the available CSI, taking higher values for good CSI qualities and lower values otherwise. A key question that we pursue in this respect is how to select a proper value for this regularization factor. To tackle this issue, we consider the asymptotic regime in which the number of UEs per cell and the number of antennas deployed in each BS increase with the same pace, thus allowing us to leverage results from random matrix theory [11,15,49]. The analysis provides accurate closed-form approximations for the SINR and SLNR for each BS-UE pair that depend solely on the system large-scale statistics and the power of the channel estimation errors. The results concerning the SLNR are then optimized in order to get a proper selection method for the regularization parameter.
1.2.2 Massive MIMO with Correlated Rician Fading

Over the past couple of years, massive MIMO has been extensively investigated as a key technology of future wireless networks, (see [2,3,8,11,49,50] and references therein). In general, most of these works, including the aforementioned analysis in Chapter 3, consider Rayleigh fading channels. Such channel models represent scattered signals, yet do not encompass the possibility of a Line-of-Sight (LoS) component which is commonly present in practical wireless propagation scenarios and modeled by Rician-fading. At the same time, in order to meet the 5G performance demands, massive MIMO is expected to be omnipresent and thus, all propagation conditions ought to be examined. This is the case for indoor applications or small areas operating over mmWave communications wherein the presence of LoS is conceivable [51]. In fact, research on the pairing of massive MIMO systems with mmWave communications to jointly reap their benefits is of growing importance [12,52,53]. Accordingly, understanding and enhancing the performance of massive MIMO systems operating under LoS conditions is the focus of Chapter 4.

Prior to our work, the main literature on massive MIMO with LoS components was represented by [54–66]. The authors in [54] investigate the power scaling laws and the uplink rates using zero-forcing (ZF) and maximum-ratio (MR) combiners, however assuming uncorrelated channels. In [56], an analytical study of the rates of the downlink of a MU-MIMO system assuming ZF beamforming and uncorrelated Rician fading channels is performed. Conversely, in [58], the authors use tools from random matrix theory to conduct an asymptotic analysis of the DL of a spatially correlated MIMO Rician fading model, yet under the assumption of perfect channel state information at the base station. A similar large system analysis relying on some recent RMT results is led in [61]. In this latter, the authors investigate regularized-ZF and MR-transmit schemes in the DL of a large-scale MIMO system under uncorrelated Rician fading system and assuming imperfect channel estimates. Moreover, in the
same line, the authors in [63] use similar tools to analyze the ergodic UL rates of a massive MIMO system and determine the optimal fraction of the coherence time used for channel training in a Rayleigh fading setting. Finally, the work in [66] analyzes transmit and receive LoS-based beamforming designs which treat the scattered signals as interference in Rician fading massive MIMO systems.

In the work of Chapter 4, we investigate the UL performance of a massive MIMO system wherein every user is allotted a distinct channel correlation matrix and Rician factor, and above all, assuming imperfect CSI. Furthermore, we consider two different combining schemes. The first method consists in the involved Linear-Minimum-Mean-Square-Error (LMMSE) receiver which we will refer to as the ‘conventional’ receiver in the sequel. Note that LMMSE’s performance under such an intricate system model renders the analysis comprehensive and unprecedented. As to the second technique, we propose a ‘statistical’ combiner that only utilizes the long-term statistics of the channels like the Rician factors and spatial correlation matrices. In essence, this technique is purposely designed for LoS-prevailing environments to circumvent training, and channel estimation with its associated errors, and exploit the presence of LoS components in a more efficient manner.

We first consider a single-cell network with the aforementioned comprehensive channel model, and carry out a theoretical analysis of the UL spectral efficiency achieved by each receiver. Assuming the large antenna-limit with a fixed number of users, we harness some elemental asymptotic tools such as the law of large numbers (LLN) and convergence of quadratic forms lemma (Chapter 2, lemma 2), to derive closed-form approximations of the SEs. These approximations provide insights on the impact of the system parameters such as the training sequence’s length, the Rician factor as well as the propagation conditions on the overall performance. Furthermore, we exploit them to determine the optimal number of training symbols that maximizes the UL SE whose value is shown to be particularly crucial at low Rician factors. A
relevant outcome of the study reveals that high Rician factors generate far better performances and that in such environments, longer training sequences are rather counterproductive since they degrade the achievable SEs. This result led us to propose the statistical receiver for systems with strong specular signals. As will be elaborated in Chapter 4, this scheme is obtained through the maximization of the corresponding UL SE and proven to outperform the conventional receiver in LoS-prevailing systems.

For a more thorough analysis, we extend our study to a multi-cell system to examine the performance of both combining schemes when they are subject to inter-cell interference. Therefore, we consider a multi-cell network and derive closed-form approximations of the achievable SEs, under the asymptotic antenna regime. Ultimately, the discussion unveils that in the multi-cell setup, the proposed combiner outperforms the conventional one to an even greater extent than it is the case for a single-cell system. Evidently, numerical results are provided to validate the accuracy of our analytical findings and better illustrate the efficiency of both schemes for finite system dimensions, albeit computed in the asymptotic regime.

1.2.3 Pilot Contamination and Line of Sight Interplay on Massive MIMO Systems

Despite the substantial advantages of massive MIMO, many issues remain to be tackled. In multi-cell networks, studies have shown that interference emanating from pilot contamination is a major impediment that limits the achievable capacity [3,11,21]. Interestingly, the authors in [67] show that when multi-cell combining schemes are employed, pilot contamination is no longer a fundamental asymptotic limitation, since the rates can in fact scale with the number of antennas as this latter grows large. However, the work in [67] considers Rayleigh fading channels and as discussed in the previous section, such models do not include specular signals which are common in practice. As a matter of fact, the impact of LoS propagation conditions on massive
MIMO systems has not been sufficiently investigated, despite the fact that several issues arise in such context. Particularly, establishing whether factors such as pilot contamination remain performance limiting even when using multi-cell combining of \cite{67}, or the circumstances under which LoS conditions become beneficial, will contribute to make massive MIMO more resilient to such elements and improve the overall performances. Accordingly, investigating these issues constitutes the main focus of the last chapter of this dissertation.

Although the performance of massive MIMO with Rician fading channels has, lately, received a lot of interest \cite{54, 66, 68, 71}, prior to our work, only few studies considered the multi-cell setting with pilot contamination \cite{61, 68, 71}. In \cite{68}, the authors examine channel estimation schemes for downlink transmissions of an uncorrelated massive MIMO system affected by pilot contamination. In the same line, \cite{69} proposes a novel technique that uses LoS estimates for channel estimation in order to minimize the pilot contamination effects. Furthermore, a large system analysis relying on some recent random matrix theory results is led in \cite{71} where the authors analyze both the UL and DL of a massive MIMO system, yet with uncorrelated Rician channels.

In our work, we consider the uplink (UL) of a multi-cell massive MIMO system with a comprehensive system model that includes imperfect channel estimation, pilot contamination, and distinct per-user Rician factor and channel correlation matrix. Furthermore, we conduct a comparative analysis of different combining schemes. Specifically, we first investigate the performances when using single-cell detectors wherein only local channels are used to process the received signal. In this context, we examine the UL spectral efficiencies generated by Maximum-Ratio-Combining and Minimum-Mean-Square-Error (which we shall denote S-MMSE to refer to single-cell detection) receivers. Second, we analyze multi-cell combining where both local and inter-cell channels are utilized in the design of the receiver. Accordingly, we consider multi-cell MMSE (M-MMSE), where we extend the study in \cite{67} to correlated Rician
fading. Ultimately, we aim at identifying the limitations of the three detectors with respect to propagation conditions and the different forms of undergone interference. To this end, assuming the large-antenna limit, we derive closed-form approximations of the UL SE using some basic asymptotic tools such as the Law of Large Numbers. These expressions are tractable and instructive as they allow to investigate the effects and interplay between the LoS and pilot contamination on the spectral efficiency. Our study reveals that in single-cell processing, the system remains limited by pilot contamination and that the specular signals enhance the achievable SE, notwithstanding some LoS-induced interference. Furthermore, we demonstrate that this latter dissipates under favorable propagation conditions for MRC, and propose certain design solutions to mitigate it for the S-MMSE. As to M-MMSE, we find that the UL rate scales with the number of antennas, assuming asymptotically linearly independent correlation matrices (Chapter 2, definition 2). Consequently, we validate, in multi-cell correlated Rician fading settings, the conclusions of [67] that M-MMSE outperforms both single-cell detection techniques. Finally, in light of the promising advantages of M-MMSE, we further investigate the effects of inter-LoS interfering links on its performance. Such investigation is peculiarly relevant for heterogeneous networks or emerging UAV systems, wherein the inter-cell LoS links are broadly present.

1.3 Outline

The remainder of the dissertation is organized as follows. In Chapter 2, we give an overview of random matrix theory tools and the probability theory lemmas that assist in the derivations of the results of all subsequent chapters. In chapter 3-5, we present our findings as follows. In each of these chapters, we first present the considered system model, then move to the asymptotic analysis to derive deterministic equivalents that allow to discuss the different factors that impact the performance. Plus, for each asymptotic result, we provide a selection of numerical illustrations that validate our
In essence, in Chapter 3 we investigate coordinated beamforming techniques that allow to mitigate the inter-tier interference in massive MIMO heterogeneous networks. Furthermore, we propose a regularized SLNR based precoding design in which the regularization factor is used to allow better resilience with respect to the channel estimation errors. Then, in Chapter 4 we shift our focus to the uplink of massive MIMO system with per-user channel correlation and Rician fading, using two processing approaches. The first technique consists of the linear-minimum-mean-square-error receiver under training-based imperfect channel estimates. As to the second approach, we propose a statistical combining technique that is more suitable in environments with strong Line-of-Sight (LoS) components. Lastly in Chapter 5 the effects of LoS signals and their interplay with pilot contamination on multi-cell massive MIMO systems are examined in depth. Finally, a summary of the dissertation is provided in Chapter 6 along with a brief discussion on future work directions for massive MIMO systems.
Chapter 2

Random Matrix Theory for Massive MIMO Systems

Large dimensions are becoming ubiquitous in modern day wireless communications applications wherein everyone and everything will be connected with the merge towards 5G networks. In this context, massive MIMO is undeniably one of the key enabling technologies to achieve 5G promises. The mathematical model underlying such systems involves matrices with hundreds of random entries, thus making the theoretical analysis challenging and computationally intractable. In fact, the main performance metrics generally used to examine such systems amount to large random compact terms which do not lend themselves to simple interpretations and can only be investigated through thousands of MonteCarlo simulations. With this in mind, Random Matrix Theory provides tools that allow the derivation of very tight asymptotic approximations of random quantities with infinite dimensions. Besides, since the essence of massive MIMO is the deployment of BSs with numbers of antennas in the range of hundreds, the use of RMT tools in such context is evidently justified and close to practice. Additionally, RMT enables another degree of freedom, as it assumes that both antennas and numbers of users increase unboundedly at the same speed. This comes in contrast to the first studies of massive MIMO’s performance that were conducted assuming a fixed number of users [2,6] which can be perceived as a special setting.

Our work belongs to the lineage of recent studies investigating the asymptotic performance of large dimensional MIMO systems using random matrix theory [15,49,58,72,74]. Formally, we investigate performance metrics like the spectral efficiency by using RMT tools to obtain simple expressions that depend solely on the long-term
system parameters such as number of BS antennas, SNR, length of training sequence, etc. Accordingly, these approximations enable insights on the effect of certain factors on the overall system performance (e.g., impact of channel correlation, CSI quality, pilot contamination, Line-of-Sight, to name a few.). In addition, they provide an efficient framework of closed-form expressions that can be leveraged for further operations such as optimizations on higher layers (scheduling, resource allocation, etc.). It is worth mentioning that conducting an asymptotic analysis does not imply that the parameters will be nearly-infinite in practice. In fact, its main purpose is to capture the behavioral tendencies of different performance metrics and establish, if any, the fundamental limitations of the overall performance. As shall be seen throughout the dissertation, selections of numerical results assert that the derived approximations using such tools are accurate for finite system dimensions, albeit computed assuming the asymptotic regime.

Our main contributions rely highly on some lemmas and theorems that are widely used to develop deterministic equivalents of the SINR and other network metrics. For a better understanding of the asymptotic results as well as the proofs given throughout the dissertation, we provide in this chapter the essential lemmas and definitions used for the derivations.

2.1 Some Notions of Probability Theory

Almost sure convergence

Definition 1 (Almost sure convergence). Let \( X_1, X_2, \ldots \) be a sequence of random variables. This sequence convergences almost surely to \( X \), if

\[
P \left( \lim_{n \to \infty} \sup |X_n - X| = 0 \right) = 1. \tag{2.1}
\]
In the sequel, this convergence is denoted $X_n - X \xrightarrow{a.s.}{n \to \infty} 0$.

The almost sure convergence was introduced to incorporate the measurability of the random variables by imposing the necessary condition that the set of convergence of the RVs has a probability equal to one. That is, the definition of “almost sure convergence” provides the random variables a certain flexibility not to converge on a set of zero probability. This flexibility, in addition to its strength (stronger than convergence in distribution and convergence in probability), made the “almost sure” convergence the most used type in RMT topics, especially for wireless communication applications such as $[24, 49, 72, 75]$.

**Continuous mapping theorem**

**Theorem 1.** [76, Theorem 2.3] Let $\{X_n\}$ be random elements defined on a metric space $\mathcal{S}$. Suppose a function $f : \mathcal{S} \mapsto \mathcal{S}'$ (where $\mathcal{S}'$ is another metric space) has the set of discontinuity points $D_f$ such that $P(X \in D_f) = 0$. Then

\[
(i) \quad X_n \xrightarrow{d} X \Rightarrow f(X_n) \xrightarrow{d} f(X);
\]

\[
(ii) \quad X_n \xrightarrow{p} X \Rightarrow f(X_n) \xrightarrow{p} f(X);
\]

\[
(iii) \quad X_n \xrightarrow{a.s.} X \Rightarrow f(X_n) \xrightarrow{a.s.} f(X);
\]

where $\xrightarrow{d}$, $\xrightarrow{p}$, and $\xrightarrow{a.s.}$, refer to convergence in distribution (or weak convergence), convergence in probability, and almost sure convergence, respectively.

In other words, the continuous mapping Theorem states that stochastic convergence is preserved for continuous functions. Accordingly, this theorem allows us to obtain approximations of involved continuous functions based on deterministic equivalent of the terms constituting them. For instance, finding approximations of the signal, interference, and noise terms leads to an approximation of the SINR.
2.2 Important Results from RMT

2.2.1 Lemmas and Identities

We first recall some lemmas and identities that will be extensively used throughout the dissertation.

**Lemma 1** (Woodbury Formula). [77, 78] Define the matrices $A \in \mathbb{C}^{N \times N}$, $B \in \mathbb{C}^{K \times K}$, $C \in \mathbb{C}^{N \times K}$, and $V \in \mathbb{C}^{K \times N}$. Then,

\begin{align}
(A + CBV) &= A^{-1} - A^{-1}C(B^{-1} + VA^{-1}C)^{-1}VA^{-1} \quad (2.5) \\
(A + CBC^H) &= A^{-1} - A^{-1}C(B^{-1} + C^HA^{-1}C)^{-1}C^HA^{-1} \quad (2.6)
\end{align}

if $A$ and $B$ are positive definite matrices, then

\[
(B^{-1} + C^HA^{-1}C)^{-1}C^HA^{-1} = BC^H(A + CBC^H) \quad (2.7)
\]

it follows that:

\[
(A + CC^H)C = A^{-1}C(I_K + C^HA^{-1}C)^{-1} \quad (2.8)
\]

**Lemma 2** (Inversion Lemma). [79, Lemma 2.2] Let $H$ be an $M \times K$ complex matrix and $h_k \in \mathbb{C}^{M \times 1}$ its $k$-th column. Denote by $H_k$, the matrix $H$ without $h_k$. For $t \in \mathbb{C}$, we define the resolvent matrices associated with $H$ and $H_k$ as:

\[
Q(t) = \left(\frac{i}{K}HH^H + I_M\right)^{-1} \quad (2.9)
\]

\[
Q_k(t) = \left(\frac{i}{K}H_kH_k^H + I_M\right)^{-1} \quad (2.10)
\]

It follows that:
\[ Q(t)h_k = \frac{Q_k(t)h_k}{1 + \frac{t}{K} h_k^\nu Q_k(t)h_k}. \tag{2.11} \]

and

\[ Q(t) = Q_k(t) - \frac{t}{K} h_k^\nu Q_k(t)h_k \frac{1}{1 + \frac{t}{K} h_k^\nu Q_k(t)h_k}. \tag{2.12} \]

**Corollary 1.** Let \( h, h_k, Q(t) \) and \( Q_k(t) \) be defined as in Lemma 2. We have then:

\[ \frac{Q(t)h_k}{\|Q(t)h_k\|} = \frac{Q_k(t)h_k}{\|Q_k(t)h_k\|}. \tag{2.13} \]

**Lemma 3** (Convergence of quadratic forms). \[ \text{[80]} \] Let \( A_M \in \mathbb{C}^{M \times M} \) be a sequence of matrices with bounded spectral norms, i.e. \( \limsup_M \|A\|_2 < \infty \). Let \( x_1, \ldots, x_K \), with \( x_i \in \mathbb{C}^{M \times 1} \), be \( K \) random vectors with i.i.d complex Gaussian entries having zero mean and unit variance, and independent from \( A_M \). It holds that, as \( M \) and \( K \) grow to infinity together

\[ \frac{1}{M} x_k^\mu A_M x_k - \frac{1}{M} \text{tr}(A_M) \xrightarrow{\text{a.s.}} 0. \tag{2.14} \]

We also have

\[ \max_{1 \leq k \leq K} \left| \frac{1}{M} x_k^\mu A_M x_k - \frac{1}{M} \text{tr}(A_M) \right| \xrightarrow{\text{a.s.}} 0. \tag{2.15} \]

**Corollary 2.** Let \( A_M \in \mathbb{C}^{M \times M} \) be defined as in Lemma 3. Let \( x_M \) and \( y_M \) be \( M \times 1 \) mutually independent vectors with i.i.d complex Gaussian entries and unit variance. Then,

\[ \frac{1}{M} x_M^\mu A_M y_M \xrightarrow{\text{a.s.}} 0. \tag{2.16} \]

**Lemma 4** (Rank one perturbation). \[ \text{[72]} \] Lemma 14.3/ Let \( Q(t) \) and \( Q_k(t) \) be defined as in Lemma 2. Then, for any sequence of matrices \( (A_M) \) with bounded spectral norm, we have:

\[ \frac{1}{M} \text{tr}(A_M Q(t)) - \frac{1}{M} \text{tr}(A_M Q_k(t)) \xrightarrow{\text{a.s.}} 0. \]

**Definition 2** (Asymptotically linearly independent correlation matrices). \[ \text{[24]} \] Consider
the correlation matrix \( \mathbf{R} \in \mathbb{C}^{N \times N} \). This matrix is asymptotically linearly independent of the correlation matrices \( \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_L \) if

\[
\lim \inf_{N \to \infty} \frac{1}{N} \left\| \mathbf{R} - \sum_{i=1}^{L} \lambda_i \mathbf{R}_i \right\|_F^2 > 0, \forall \lambda_1, \ldots, \lambda_L \in \mathbb{R}.
\] (2.17)

It is worth mentioning that the definition is not to be mistaken with the classic linear independence. In fact, asymptotic linear independence does not only require linear independence, but also that the subspace in which the matrices differ has a norm that grows at least linearly with \( N \). This definition will be repeatedly used to sustain the derivations in Chapter 5.

### 2.2.2 Deterministic Equivalents

In this section, let \( \mathbf{H} \) denote a \( M \times K \) complex matrix satisfying the assumptions:

**Assumption 1.**

\[
\mathbf{H} = \frac{1}{\sqrt{K}} \mathbf{X} \tilde{\mathbf{D}}^{1/2}
\] (2.18)

where \( \mathbf{X} \in \mathbb{C}^{M \times K} \) is a standard Gaussian matrix with zero mean and unit variance, and \( \tilde{\mathbf{D}} \) is a \( K \times K \) diagonal matrix.

**Assumption 2.** The diagonal elements of \( \tilde{\mathbf{D}} \) verify

\[
0 < \lim \inf_{K \to \infty} \sum_{i=1}^{K} \tilde{d}_i < \lim \sup_{K \to \infty} \sum_{i=1}^{K} \tilde{d}_i < \infty.
\] (2.19)

**Assumption 3.** Let \( c = \frac{K}{M} \). Then, we have:

\[
0 < \lim \inf c < \lim \sup c < \infty.
\] (2.20)
In the following, we review results concerning the behavior of quantities involving $H$ and its resolvent matrix $Q(t)$ in the asymptotic regime defined in Assumption 3. These results will be extensively used in Chapter 3. More specifically, we propose deterministic equivalents approximating the quantities $\frac{1}{K} \text{tr}(AQ(t))$ and $\frac{1}{K} \text{tr}(A(Q(t))^2)$, respectively, for any matrix $A$ with bounded spectral norm independent from $Q(t)$.

**Proposition 1** (Deterministic equivalent for $\frac{1}{K} \text{tr}(AQ(t))$). Let $H$ be as in Assumption 7. Define $Q(t)$ to be the resolvent matrix associated with $H$. Then, under Assumption 2, the following equation:

$$e(t) = 1_K \sum_{i=1}^{K} \frac{\bar{d}_i}{t + \frac{1}{c} \bar{d}_i \bar{e}(t)},$$

with $\bar{e}(t) = \frac{1}{1+ce(t)}$ admits a unique solution for $t > 0$.

Moreover, for any $t > 0$ and under Assumptions 2-3, the following convergence holds true:

$$\frac{1}{K} \text{tr}(AQ(t)) - t^{-1} \bar{e}(t) \frac{1}{K} \text{tr}(A) \xrightarrow{a.s.} 0, \quad \text{as } N \to \infty.$$  \hfill (2.21)

**Proposition 2** (Deterministic equivalent for $\frac{1}{K} \text{tr}(A(Q(t))^2)$). Consider the same setting as in Proposition 7, and define the quantities:

$$\Gamma(t,c) = \frac{1}{1 + \frac{tc}{d\bar{e}(t)}}, \quad \text{ (2.22)}$$

$$\tilde{\vartheta}(t) = \frac{c^2}{K} \sum_{i=1}^{K} (\Gamma(t,c))^2. \quad \text{ (2.23)}$$

Then, under Assumptions 2 and 3,

$$\frac{1}{K} \text{tr}(A(Q(t))^2) - \frac{(\bar{e}(t))^2}{t^2 \left(1 - c^{-1} \tilde{\vartheta}(t)\right)} \frac{1}{K} \text{tr}(A) \xrightarrow{a.s.} 0, \quad \text{as } N \to \infty.$$  \hfill (2.24)
Proof. Finding the deterministic quantity in (2.24) can be easily derived by observing that:

$$\partial_t \text{tr}(AQ(t)) = \frac{-t}{K} \text{tr}(AQ^2(t))$$

(2.25)

Therefore, according to the continuous mapping theorem, a deterministic equivalent of

$$\frac{1}{K} \text{tr}(A(Q(t))^2)$$

can be mainly obtained by differentiating the one of

$$\frac{1}{K} \text{tr}(A(Q(t)))$$

(2.21).

2.3 Chapter Summary

In this Chapter, we briefly explained the reason why Random Matrix Theory is a suitable tool to analyze massive MIMO systems. Furthermore, we provided some important lemmas from Probability Theory and RMT that will be extensively used in the derivation of all results given in the subsequent chapters of this dissertation. In the next chapter, we present our first main contribution in the context of massive MIMO systems and heterogeneous networks. As shall be seen, the theorems and proofs of the following chapter are a perfect framework to further grasp the usefulness of RMT tools in the analysis of massive MIMO networks in practical settings.
Chapter 3

Coordinated SLNR based Precoding in Large-Scale Heterogeneous Networks

3.1 Introduction

In this chapter, we propose to investigate the design of SLNR based precoding techniques which rely on acquired channel estimates to mitigate interference. The novelty of our proposed scheme lies in the use of coordination along with a regularization parameter that weights the inter-cells information implement in the precoding structure. To this end, we carry out a theoretical analysis of the Signal-to-Noise-and-Interference ratio (SINR) in a large-scale two-tier HetNet, taking into account imperfect CSI scenarios in the asymptotic regime. This latter assumes that the number of antennas at the BSs and that of users grow simultaneously large, thus enabling to harness results from random matrix theory to derive deterministic equivalents of the SINR and corresponding SE, for different coordination strategies. A selection of numerical results shows that these approximations are accurate for finite system dimensions, albeit computed assuming the asymptotic regime. Additionally, these expressions depend only on the statistics of the channels and the transmit power at the BSs, and as such, are instructive in how these parameters impact the overall performance. Furthermore, the results we obtain include explicit closed-form expressions that provide pertinent insights, such as the way coordinated beamforming mitigates the interference. Another interesting finding is the detrimental impact of poor CSI on the performance gains of coordinated beamforming. Finally, to ensure better resilience to such effects, we propose a more
Figure 3.1: Two-Tier HetNet with legitimate and different types of interfering channels.

suitable selection method for the regularization parameter through the optimization of the SLNR metric.

3.2 System Model

In this section we introduce the transmission system and the considered channel model.

3.2.1 Transmission Model

We consider the downlink of a two-tier HetNet composed of $L$ cells, as depicted in Fig. 3.1, wherein the main macro-cell, referred to as cell 1, is equipped with a central macro base station (MBS). The remaining cells, $2, \ldots, L$, are endowed with a single lower power micro BS (mBS). Each cell $j$ is assigned one base station denoted BS$_j$ having $M_j$ antennas to serve $K_j$ mono-antenna users. The set of UEs associated with the macro-cell, UEs$_1$, is uniformly distributed over the entire coverage area of the network, whereas the UEs of remaining cells, UEs$_j$ ($j \neq 1$), are uniformly spread over the coverage area of their corresponding cell $j$ delimited by a circle of radius $R$. We assume that all BSs use Gaussian codebooks and linear precoding. Accordingly, BS$_j$
transmits the signal

\[ \mathbf{x}_j = \sum_{k=1}^{K_j} \mathbf{w}_{jk}s_{jk}. \tag{3.1} \]

where \( s_{jk} \sim \mathcal{CN}(0, 1) \) is the downlink data symbol intended to UE \( k \) in cell \( j \), and \( \mathbf{w}_{jk} \) its associated precoding vector.

Let \( h_{\ell jk} \) denote the channel vector between BS \( \ell \) and UE \( j \) in cell \( k \). Therefore, the received signal by UE \( j \) in cell \( k \) can be expressed as:

\[ y_{jk} = \sum_{\ell=1}^{L} \sum_{k=1}^{K_\ell} h_{\ell jk}^H \mathbf{w}_{\ell i} s_{\ell i} + n_{jk}, \tag{3.2} \]

where \( n_{jk} \sim \mathcal{CN}(0, \sigma^2 I_{M_j}) \) accounts for the additive Gaussian noise. We can further decompose \( y_{jk} \) as:

\[ y_{jk} = h_{jjk}^H \mathbf{w}_{jk} s_{jk} + \sum_{i=1}^{K_j} \sum_{i \neq k} h_{jjk}^H \mathbf{w}_{ji} s_{ji} + \sum_{i=1}^{K_\ell} \sum_{\ell \neq j} h_{\ell jk}^H \mathbf{w}_{\ell i} s_{\ell i} + n_{jk}. \tag{3.3} \]

This decomposition distinguishes the desired signal term from the overall interference and noise. Hence, the SINR at UE \( k \) in cell \( j \), \( \gamma_{jk} \), writes:

\[ \gamma_{jk} = \frac{|h_{jjk}^H \mathbf{w}_{jk}|^2}{\sum_{i=1}^{K_j} |h_{jjk}^H \mathbf{w}_{ji}|^2 + \sum_{\ell=1}^{L} \sum_{i=1}^{K_\ell} |h_{\ell jk}^H \mathbf{w}_{\ell i}|^2 + \sigma^2}, \tag{3.4} \]

\[ = \frac{S_{jk}}{I_{\text{intra}}^{jk} + \sum_{\ell=1}^{L} I_{\ell jk}^{\text{inter}} + \sigma^2}, \tag{3.5} \]

where:

- \( S_{jk} = |h_{jjk}^H \mathbf{w}_{jk}|^2 \) is the received signal power at user \( k \) in cell \( j \),
- \( I_{\text{intra}}^{jk} = \sum_{i=1}^{K_j} |h_{jjk}^H \mathbf{w}_{ji}|^2 \) is the received intra-cell interference at user \( k \),
- \( I_{\ell jk} \) is the interference generated by cell \( \ell \) to the user \( k \) in cell
Ultimately, we are interested in the DL average achievable spectral efficiency per UE in each cell $j$, $j = 1, \cdots, L$, defined as:

$$\text{SE}_j = \frac{1}{K_j} \sum_{k=1}^{K_j} \mathbb{E} \left[ \log_2 \left( 1 + \gamma_{jk} \right) \right].$$ (3.6)

where interference is treated as noise and the expectation is taken over the channels $h_{j\ell k}$.

### 3.2.2 Channel Model

In this work, we consider the setting where the channels $\{h_{j\ell k}\}$ are spatially uncorrelated and obey the model:

$$h_{j\ell k} = \sqrt{P_{j\ell k}} z_{j\ell k},$$ (3.7)

where $z_{j\ell k} \in \mathbb{C}^{M_j \times 1}$ is a complex Gaussian vector with i.i.d entries of zero mean and unit variance to account for the small-scale fading. $P_{j\ell k}$ refers to the amount of power received at UE $k$ located in cell $\ell$ from BS$_j$. For each BS$_j$, we assume that the available power $P_j$ is equally allocated among its legitimate UEs. In addition, for each UE $k$ in cell $j$, the received power undergoes some channel attenuation which we denote by $\alpha_{j\ell k}$. Formally, $P_{j\ell k}$ is given by:

$$P_{j\ell k} = \frac{P_j}{K_j} \alpha_{j\ell k}.$$ (3.8)

In practice, only an imperfect estimate $\hat{h}_{j\ell k}$ of the true channel $h_{j\ell k}$ is available at the transmitter. This imperfection arises naturally from users mobility for instance. In the present work, we model imperfect CSI as follows:

$$\hat{h}_{j\ell k} = \sqrt{P_{j\ell k}} \left( \sqrt{1 - \tau_{j\ell k}^2} z_{j\ell k} + \tau_{j\ell k} q_{j\ell k} \right),$$ (3.9)
where $\sqrt{P_{j\ell k}} q_{j\ell k} \sim \mathcal{CN}(0, P_{j\ell k} I_{M_j})$ accounts for the channel estimation errors and is independent of $h_{j\ell k}$. The parameter $\tau_{j\ell k} \in [0, 1]$ reflects the accuracy of the channel estimate such that $\tau_{j\ell k} = 0$ refers to perfect CSI, whereas for $\tau_{j\ell k} = 1$, the estimate is completely independent of the real channel. Finally, we denote the aggregate channel matrix linking BS$_j$ to users in cell $\ell$ by $H_{j\ell} = [h_{j\ell 1}, h_{j\ell 2}, \ldots, h_{j\ell K_\ell}]$, and the aggregate channel estimates matrix:

$$\hat{H}_{j\ell} = [\hat{h}_{j\ell 1}, \hat{h}_{j\ell 2}, \ldots, \hat{h}_{j\ell K_\ell}]$$  \hspace{1cm} (3.10)

### 3.3 Coordinated SLNR-MAX Design

In a multi-cell system, the interference caused by the transmissions of neighboring cells can seriously affect the downlink performance, especially for cell-edge users. One possible solution to remedy this issue consists in using coordinated beamforming designs that mitigate the interference caused by each BS to non-intended users. Several levels of coordination can be envisioned. In this work, we consider the case in which each BS sends data to its own UEs, but knows by exchanging CSI with other BSs estimated values of channels linking it to UEs in other cells. If perfect CSI is available, complete elimination of both intra-cell and inter-cell interference can be accomplished by coordinated zero-forcing beamforming (ZFBF). However, this requires the BS to possess a sufficient number of antennas, and more importantly, is not necessarily optimal, as it may entail a considerable loss in the energy of the desired signal. To ensure a good balance between interference and signal, other coordinated precoding techniques should be used \[42\]. Of interest in this work is the class of precoding schemes that maximizes the signal to leakage and noise ratio, which we denote by SLNR-MAX.

Note that SLNR-MAX can also be used when coordination is not employed. In this case, the signal to leakage ratio is obtained by only considering the interference
leakage to the other UEs in the same cell, thus yielding:

\[
\text{SLNR}_{jk}^{\text{no-co}} = \frac{|h_{jjk}^H w_{jk}|^2}{\sum_{i=1}^{K_j} |h_{jji}^H w_{jk}|^2 + \rho \sigma^2},
\]

(3.11)

where \( \rho \) herein is a regularization parameter weighting the noise energy. It can be easily shown that the precoding vector maximizing (3.11) is given by:

\[
w_{jk}^{\text{no-co}} = \left( H_{jjj}^H H_{jjj}^H + \rho \sigma^2 I_{M_j} \right)^{-1} h_{jjk} \left[ \left( H_{jjj}^H H_{jjj}^H + \rho \sigma^2 I_{M_j} \right)^{-1} h_{jjk} \right]^{\frac{1}{2}},
\]

(3.12)

In the coordinated case, the total amount of interference leakage caused by the serving BS is taken into account. In particular, the SLNR corresponding to the communication between cell \( j \) and its UE \( k \) is given by:

\[
\text{SLNR}_{jk}^{\text{co}} = \frac{|h_{jjk}^H w_{jk}|^2}{\sum_{i=1}^{K_j} |h_{jji}^H w_{jk}|^2 + \sum_{\ell=1}^{L} \sum_{i=1}^{K_\ell} |h_{j\ell i}^H w_{jk}|^2 + \rho \sigma^2}
\]

(3.13)

\[
= \frac{S_{jk}}{L_{\text{intra}}^{jk} + \sum_{\ell=1}^{L} L_{\text{inter}}^{\ell,jk} + \rho \sigma^2},
\]

(3.14)

where:

- \( L_{\text{intra}}^{jk} \) is the interference leakage generated by the communication between BS \( j \) and UE \( k \) to all UEs \( i \neq k \) in cell \( j \).

- \( L_{\text{inter}}^{\ell,jk} \) is the interference leakage caused by the communication between BS \( j \) and its UE \( k \) to all UEs in cell \( \ell \).

Similarly, it can be easily proven that the beamforming vector maximizing (3.14) is
given by:

$$w_{jk}^{co} = \frac{v_{jk}^{co}}{\|v_{jk}^{co}\|}, \quad s.t., \quad v_{jk}^{co} = \left(H_{jj}H_{jj}^H + \sum_{\ell=1}^{\ell\neq j} H_{j\ell}H_{j\ell}^H + \rho \sigma^2 I_{M_j}\right)^{-1} h_{jjk}.$$  

(3.15)

If only imperfect CSI is available at the BSs, we propose to simply replacing the true channels $h_{\ell jk}$ by their estimates in the expressions of $w_{jk}^{no-co}$ and $w_{jk}^{co}$. In doing so, the beamforming vector $w_{jk}^{no-co}$ takes the form:

$$\hat{w}_{jk}^{no-co} = \frac{\hat{v}_{jk}^{no-co}}{\|\hat{v}_{jk}^{no-co}\|}, \quad s.t., \quad \hat{v}_{jk}^{no-co} = \left(\hat{H}_{jj}\hat{H}_{jj}^H + \rho \sigma^2 I_{M_j}\right)^{-1} \hat{h}_{jjk}.$$  

(3.16)

In the coordinated case, we pursue the same approach with the single difference that the channels with UEs in other cells are weighted by the coordination factor $\beta_j$.

$$\hat{w}_{jk}^{co} = \frac{\hat{v}_{jk}^{co}}{\|\hat{v}_{jk}^{co}\|}, \quad s.t., \quad \hat{v}_{jk}^{co} = \left(\hat{H}_{jj}\hat{H}_{jj}^H + \beta_j \sum_{\ell \neq j} \hat{H}_{j\ell}\hat{H}_{j\ell}^H + \rho \sigma^2 I_{M_j}\right)^{-1} \hat{h}_{jjk}.$$  

(3.17)

The reason behind using the factor $\beta_j$ is that channel estimates with UEs in other cells might be unreliable. The coordination factor $\beta_j$ is thus used to place more or less importance on these estimates. Another advantage of studying the performances of SLNR-MAX beamforming is that it encompasses other types of precoding designs by properly selecting the parameters $\beta_j$ and $\rho$. For instance, ZFBF can be obtained by setting $\beta_j = 1$ and taking the limit of (3.12) and (3.17) as $\rho \to 0$.

### 3.3.1 Asymptotic Analysis of Different Coordination Strategies

This section presents an asymptotic performance analysis of the SINR, and by the same token the DL spectral efficiency, using different coordination strategies. Our
results are valid in the asymptotic regime in which the number of antennas \( M_i \) and that of users \( K_i \) grow large at the same pace. More formally, this regime can be written as:

**Assumption 4.** We assume that \( L \) is fixed. Define \( c_i = \frac{K_i}{M_i} \). Then, we have:

\[
0 < \lim \inf c_i < \lim \sup c_i < \infty
\] (3.18)

Furthermore, we shall additionally assume that \( \forall (j, \ell, k), P_j \alpha_{j\ell k} \) remains bounded. This can be formulated as:

**Assumption 5.** \( \forall j, \ell \) and \( k \),

\[
0 < \lim \inf \min_{j, \ell, k} P_j \alpha_{j\ell k} \leq \lim \sup \max_{j, \ell, k} P_j \alpha_{j\ell k} < \infty
\] (3.19)

Let \( K = \sum_{i=1}^{L} K_i \). For simplicity, the asymptotic regime in Assumption 4 will be denoted by \( K \to \infty \).

### 3.3.1.1 Deterministic Equivalents for the SINR

In this section, we derive closed-form approximations for the SINR, \( \gamma_{jk} \). Our main results are stated in two theorems, treating separately the cases in which the underlying BS coordinates or not. Without loss of generality, note that all the derivations are conducted considering cell \( j \) as cell of interest. In this line, we will provide in both theorems accurate approximations of the terms constituting \( \gamma_{jk} \) (3.4): signal power \( S_{jk} \) and intra-cell interference \( I^{\text{intra}}_{jk} \) received by \( \text{UE}_{jk} \), as well as the inter-cell interference \( I^{\text{inter}}_{j\ell k} \) that \( \{ \text{BS}_\ell \}_{\ell \neq j} \) causes to \( \text{UE}_{jk} \). Notwithstanding, this quantity is not to be mistaken with \( I^{\text{inter}}_{j\ell k}, k = 1, \ldots, K_\ell \), which refers to the leaked inter-cell interference that is caused by \( \text{BS}_j \) to the \( \text{UEs} \) of cell \( \ell \).

We find it the most informative to analyze separately each term rather than the overall expression of the SINR. As will be explained later, gathering these results will easily lead to the analysis of any possible network configuration. Plus, using
the continuous mapping theorem in [81], the deterministic equivalent of the SINR is obtained by the ratio between the respective approximations of the signal power and total interference plus noise derived in subsequent Theorems 1 and 2.

A. Single-cell processing setting

We first consider the case in which BS$_j$ does not perform any coordination. To distinguish this case from the coordination setting, we will use the superscript no-co in all our derivations. Prior to presenting our main results, we shall introduce deterministic equivalents that are commonly used within the framework of random matrix theory. These quantities will be needed to express deterministic approximations of the SINR and the SLNR.

For $t > 0$ and $j \in \{1, \cdots, L\}$, define $e_{j}^{\text{no-co}}(t)$ as the unique positive solution to the following fixed-point equation:

$$
\hat{e}_{j}^{\text{no-co}}(t) = \frac{P_j}{K_j} \sum_{i=1}^{K_j} \alpha_{jji} \left( t + \frac{1}{c_j} P_j \alpha_{jji} \hat{e}_{j}^{\text{no-co}}(t) \right)^{-1},
$$

with,

$$
\hat{e}_{j}^{\text{no-co}}(t) = \frac{1}{1 + e_{j}^{\text{no-co}}(t)}.
$$

The existence and uniqueness of $e_{j}^{\text{no-co}}(t)$ unfolds from standard results of random matrix theory (See proposition in Chapter 2). Furthermore, we will need the following deterministic quantities:

$$
\Gamma_{j\ell i}^{\text{no-co}}(t, c_j) = \left( 1 + \frac{t c_j}{P_j \alpha_{j\ell i} \hat{e}_{j}^{\text{no-co}}(t)} \right)^{-1},
$$

$$
\tilde{\vartheta}_{j}^{\text{no-co}}(t) = \frac{c_j^2}{K_j} \sum_{i=1}^{K_j} \left( \Gamma_{j\ell i}^{\text{no-co}}(t, c_j) \right)^2.
$$

Following is our first main result.

**Theorem 2** (No-Coordination). Let Assumptions 4 and 5 hold. Define for $t > 0$ and $j \in \{1, \cdots, L\}$, $e_{j}^{\text{no-co}}(t)$ as in (3.20) and consider the notations (3.21) - (3.23). Let
$S_{jk}^{\text{no-co}}$, $T_{jk}^{\text{intra,no-co}}$ and $T_{\ell jk}^{\text{inter,no-co}}$ be given by

$$S_{jk}^{\text{no-co}} = P_j \alpha_{jjk} \left( \frac{1 - \tau_{jjk}^2}{c_j} \right) \left[ 1 - \frac{1}{c_j} \tilde{\vartheta}_{j}^{\text{no-co}}(\rho \sigma^2) \right],$$  \hspace{1cm} (3.24)

$$T_{jk}^{\text{intra,no-co}} = P_j \alpha_{jjk} \left[ (1 - \tau_{jjk}^2) \left( 1 - \Gamma_{jjk}^{\text{no-co}}(\rho \sigma^2, c_j) \right)^2 + \tau_{jjk}^2 \right],$$  \hspace{1cm} (3.25)

$$T_{\ell jk}^{\text{inter,no-co}} = P_\ell \alpha_{\ell jk}. \hspace{1cm} (3.26)$$

Then,

$$\max_{1 \leq k \leq K_j} \left| S_{jk}^{\text{no-co}} - S_{jk}^{\text{no-co}} \right| \overset{\text{a.s.}}{\rightarrow} 0, \hspace{1cm} (3.27)$$

$$\max_{1 \leq k \leq K_j} \left| T_{jk}^{\text{intra,no-co}} - T_{jk}^{\text{intra,no-co}} \right| \overset{\text{a.s.}}{\rightarrow} 0, \hspace{1cm} (3.28)$$

$$\max_{1 \leq k \leq K_j} \left| T_{\ell jk}^{\text{inter,no-co}} - T_{\ell jk}^{\text{inter,no-co}} \right| \overset{\text{a.s.}}{\rightarrow} 0. \hspace{1cm} (3.29)$$

**Proof.** The proof of Theorem 2 is given in Appendix 7. \hfill \square

**Corollary 3.** If perfect CSI of channel $h_{jjk}$ is available, i.e. for $\tau_{jjk} = 0$, the asymptotic expressions $S_{jk}^{\text{no-co}}$, $T_{jk}^{\text{intra,no-co}}$ and $T_{\ell jk}^{\text{inter,no-co}}$ become:

$$S_{jk}^{\text{no-co}} = \frac{P_j \alpha_{jjk}}{c_j} \left( 1 - \frac{1}{c_j} \tilde{\vartheta}_{j}^{\text{no-co}}(\rho \sigma^2) \right),$$  \hspace{1cm} (3.30)

$$T_{jk}^{\text{intra,no-co}} = \frac{P_j \alpha_{jjk}}{1 + \frac{1}{\rho \sigma c_j} P_j \alpha_{jjk} \tilde{e}_{jj}^{\text{no-co}}(\rho \sigma^2) \tau_{jjk}^2},$$  \hspace{1cm} (3.31)

$$T_{\ell jk}^{\text{inter,no-co}} = P_\ell \alpha_{\ell jk}. \hspace{1cm} (3.32)$$

Some important insights can be readily extracted from Theorem 2 and corollary 3. First, we note that $S_{jk}^{\text{no-co}}$, $T_{jk}^{\text{intra,no-co}}$ and $T_{\ell jk}^{\text{inter,no-co}}$ depend only on the large-scale channel fading parameters, and thus are instructive in how these parameters affect the performance. In particular, we can easily obtain the asymptotic quantities when ZFBF
is used by simply taking the limit of the provided expressions as \( \rho \) tends to zero. In doing so, we obtain the following expressions\(^1\):

\[
\begin{align*}
[S_{no-co}]_{ZF}^{jk} &= \frac{1 - \tau_{jjk}^2}{c_j} P_j \alpha_{jjk} (1 - c_j), \\
[T_{intra,no-co}]_{ZF}^{jk} &= \tau_{jjk}^2 P_j \alpha_{jjk}, \\
[T_{inter,no-co}]_{ZF}^{jk} &= P_t r_{\ell jk}.
\end{align*}
\]

Comparing the SLNR-MAX and ZFBF signal and interference terms: (3.24) with (3.33), and (3.25) with (3.34), we can easily see that:

\[
\begin{align*}
[S_{no-co}]_{SLNR-MAX}^{jk} &> [S_{no-co}]_{ZF}^{jk}, \\
[T_{intra,no-co}]_{SLNR-MAX}^{jk} &> [T_{intra,no-co}]_{ZF}^{jk}.
\end{align*}
\]

This proves that ZFBF completely cancels the intra-cell interference in the perfect CSI case \( \tau_{jjk} = 0 \), but induces a loss in the signal power as compared to the SLNR-MAX.

**B. Coordinated Beamforming** We now consider the case wherein BS \( j \) uses estimates of inter-cell channels to perform coordination. Similarly, we use the superscript \( co \) to distinguish this case and introduce the following deterministic quantities that will serve later to express closed-form expressions of the associated SINR \( \gamma_{jk}^{co} \).

Let \( e_j^{co}(t) \) be the unique solution to the following fixed point equation:

\[
e_j^{co}(t) = \frac{P_j}{K} \left[ \sum_{i=1}^{K_j} \alpha_{jji} \frac{1}{t + \frac{1}{c_j} P_j \alpha_{jji} e_j^{co}(t)} + \sum_{\ell=1}^{K_\ell} \sum_{i=1}^{K_i} \frac{\beta_j \alpha_{j,\ell,i}}{t + \frac{\beta_j}{c_j} P_j \alpha_{j,\ell,i} e_j^{co}(t)} \right],
\]

\(^1\)This can be easily shown by observing that as \( t \) tends to zero, \( \Gamma_{j\ell}^{no-co}(t) \to 1 \) and \( \tilde{\vartheta}_j^{no-co}(t) \to c_j^2 \).
with \( \tilde{c}_j \triangleq \frac{K}{\tilde{M}_j} \) and \( \tilde{c}_j^{co}(t) = (1 + e_j^{co}(t))^{-1} \). Let

\[
t_j^{co} = \frac{K_j}{K} \rho \sigma^2.
\] (3.39)

Define also \( \Gamma_{j\ell}^{co}(t, c_j) \), \( \Psi_{j\ell}^{co}(t, c_j) \) and \( \tilde{\vartheta}_j^{co}(t) \) as:

\[
\Gamma_{j\ell}^{co}(t, c_j) = \left(1 + \frac{t c_j}{P_j \alpha_{j\ell} \tilde{c}_j^{co}(t)}\right)^{-1},
\] (3.40)

\[
\Psi_{j\ell}^{co}(t, c_j) = \left(1 + \frac{t c_j}{\beta_j P_j \alpha_{j\ell} \tilde{c}_j^{co}(t)}\right)^{-1},
\] (3.41)

\[
\tilde{\vartheta}_j^{co}(t) = \frac{c_j^2}{K} \left[ \sum_{i=1}^{K_j} \left( \Gamma_{jji}^{co}(t, \tilde{c}_j) \right)^2 + \sum_{\ell=1}^{L} \sum_{i=1}^{K_i} \left( \Psi_{j\ell i}^{co}(t, \tilde{c}_j) \right)^2 \right].
\] (3.42)

With these notations at hand, we are now ready to state our second main result, which provides deterministic equivalents for \( S_{jk}^{co} \), \( I_{jk}^{intra,co} \), and \( I_{\ell jk}^{inter,co} \).

**Theorem 3.** Let Assumptions 4 and 5 hold true. Let \( S_{jk}^{co} \), \( I_{jk}^{intra,co} \) and \( I_{\ell jk}^{inter,co} \) be given by

\[
S_{jk}^{co} = P_j \alpha_{jkk} \left( \frac{1 - \tau_{jkk}^2}{c_j} \right) \left[ 1 - \frac{1}{c_j} \tilde{\vartheta}_j^{co}(t_j^{co}) \right],
\] (3.43)

\[
I_{jk}^{intra,co} = P_j \alpha_{jkk} \left[ (1 - \tau_{jkk}^2) \left( 1 - \Gamma_{jkk}^{co}(t_j^{co}, \tilde{c}_j) \right)^2 + \tau_{jkk}^2 \right],
\] (3.44)

\[
I_{\ell jk}^{inter,co} = P_{\ell \alpha_{\ell jk}} \left[ (1 - \tau_{\ell jk}^2) \left( 1 - \Psi_{\ell jk}^{co}(t_{\ell j}^{co}, \tilde{c}_\ell) \right)^2 + \tau_{\ell jk}^2 \right].
\] (3.45)

Then,

\[
\max_{1 \leq k \leq K_j} \left| S_{jk}^{co} - \overline{S}_{jk}^{co} \right| \xrightarrow{a.s.} 0, \quad K \to \infty
\] (3.46)

\[
\max_{1 \leq k \leq K_j} \left| I_{jk}^{intra,co} - \overline{I}_{jk}^{intra,co} \right| \xrightarrow{a.s.} 0, \quad K \to \infty
\] (3.47)

\[
\max_{1 \leq k \leq K_j} \left| I_{\ell jk}^{inter,co} - \overline{I}_{\ell jk}^{inter,co} \right| \xrightarrow{a.s.} 0, \quad K \to \infty
\] (3.48)
Proof. The proof of Theorem 3 is given in Appendix 8.

Corollary 4. When the channel is perfectly known at all BSs, \((\forall (j, \ell, k), \tau_{j\ell k} = 0)\), we have:

\[
S_{jk} = \frac{P_j \alpha_{j\ell k} c_j}{c_j} \left(1 - \frac{1}{c_j} \tilde{c}_j\right), \quad (3.49)
\]

\[
T_{jk}^{\text{intra, co}} = \frac{P_j \alpha_{j\ell k}}{\left(1 + \frac{1}{\epsilon_{\text{co}}} P_j \alpha_{j\ell k} \tilde{c}_j (t^c_j)\right)^2}, \quad (3.50)
\]

\[
T_{j\ell k}^{\text{inter, co}} = \frac{P_j \alpha_{j\ell k}}{\left(1 + \frac{\beta_j}{\epsilon_{\text{co}}} P_j \alpha_{j\ell k} \tilde{c}_j (t^c_j)\right)^2}. \quad (3.51)
\]

Similar to above, we stress the fact that the closed-form approximations of the signal and interference terms provided by Theorem 3 depend solely on the large-scale fading parameters. Furthermore, our results can be used to analyze the case in which BS \(j\) employs coordinated ZFBF by taking the limit as \(\rho\) tends to zero. Therefore:

\[
[S_{jk}]_{ZF} = \frac{(1 - \tau_{j\ell k})}{c_j} P_j \alpha_{j\ell k} (1 - \tilde{c}_j), \quad (3.52)
\]

\[
[T_{jk}^{\text{intra, co}}]_{ZF} = \tau_{j\ell k}^2 P_j \alpha_{j\ell k}, \quad (3.53)
\]

\[
[T_{j\ell k}^{\text{inter, co}}]_{ZF} = \tau_{j\ell k}^2 P_j \alpha_{j\ell k}. \quad (3.54)
\]

Theorem 3 sheds light on the impact of performing coordination on the reduction of the inter-cell interference \textit{it causes to other cells}. Consider the interference that BS \(j\) causes to UE \(\ell k\) in cell \(\ell\), \(\ell \neq j\) (simply obtained by switching the indices \(j\) and \(\ell\) in \((3.45)):

\[
T_{j\ell k}^{\text{inter, co}} = P_j \alpha_{j\ell k} \left[(1 - \tau_{j\ell k}^2) (1 - \Psi_{j\ell k}^c (t_{j\ell k}^c, \tilde{c}_j))^2 + \tau_{j\ell k}^2\right]. \quad (3.55)
\]

Indeed, it follows from some basic mathematical manipulations\footnote{It is shown by noting: for \(0 < \tau_{j\ell k} < 1\), \((1 - \tau_{j\ell k}) (1 - \Psi_{j\ell k}^c (t_{j\ell k}^c, \tilde{c}_j))^2 + \tau_{j\ell k}^2 < 1\)} that \(T_{j\ell k}^{\text{inter, co}} < T_{j\ell k}^{\text{inter, no-co}}\). Also, from (3.41), (3.45) and (3.51), we can clearly see that increasing the
regularization parameter $\beta_j$ leads to the vanishing of the inter-cell interference. This, however, is not necessarily optimal since it can cause a considerable loss in the signal power ($S^\text{co}_{jk}$ decreasing function of $\beta_j$). Additionally, Theorem 3 reveals that there is little interest to performing coordination when the channel estimation quality is poor. This can be clearly shown by observing that as $\tau_{j\ell k} \to 1$, $I_{j\ell k}^{\text{inter,co}} \to I_{j\ell k}^{\text{inter,no-co}}$. It is thus of major importance to “regulate” the coordination level in the precoding design based on the quality of available CSI. This constitutes the main objective of the next section.

Lastly, we shall point out that the expressions in Theorems 2 and 3 can be used jointly to analyze any possible coordination configuration. Let $S_{\text{co}}$ be the set containing the indexes of BSs that perform coordination, and $S_{\text{no-co}}$ the set of the ones that do not. Accordingly, the asymptotic approximations of the SINRs for users associated with cells in $S_{\text{no-co}}$ and $S_{\text{no-co}}$ are, respectively, given by:

- $\text{UE}_{jk}$ with $j \in S_{\text{no-co}}$:
  \[
  \overline{\gamma}_{jk} = \frac{S^\text{no-co}_{jk}}{I_{j\ell k}^{\text{intra,no-co}}} + \sum_{\ell \in S_{\text{no-co}} - \{j\}} \frac{I_{j\ell k}^{\text{inter,no-co}}}{I_{j\ell k}^{\text{inter,no-co}}} + \sum_{\ell \in S_{\text{co}}} I_{j\ell k}^{\text{inter,co}} + \sigma^2. \tag{3.56}
  \]

- $\text{UE}_{jk}$ with $j \in S_{\text{co}}$:
  \[
  \overline{\gamma}_{jk} = \frac{S^\text{co}_{jk}}{I_{j\ell k}^{\text{intra,co}}} + \sum_{\ell \in S_{\text{no-co}}} \frac{I_{j\ell k}^{\text{inter,co}}}{I_{j\ell k}^{\text{inter,co}}} + \sum_{\ell \in S_{\text{co}} - \{j\}} I_{j\ell k}^{\text{inter,co}} + \sigma^2. \tag{3.57}
  \]

The heterogeneity of the network can reside in which tier adopts coordinated beamforming and which one performs single-cell processing. This statement is analytically evidenced in equations (3.56) and (3.57) where the non-coordinated and coordinated precoding designs are combined in order to establish different strategies of coordination among tiers.

**Case study**: Since the macro BS is usually the main source of interference, a natural configuration would be when only the macro BS uses coordinated BF in order to
minimize the interference it generates to the micro-cells’ users. In this case \( S_{co} = \{1\} \) and \( S_{no-co} = \{2, \ldots, L\} \).

### 3.3.2 Numerical Results

In this section, we present a selection of simulations to illustrate the analytical expressions presented thus far (Section 3.3.1). Note that in all the numerical simulations of this chapter, we utilize the following settings to validate our findings. We use Monte Carlo simulations with 10000 channel realizations to corroborate, for finite system dimensions, all the asymptotic expressions derived in sections 3.3.1 and the subsequent Section 3.4.1. To this end, we consider a two-tier HetNet with one MBS and two mBSs. Recalling the power model (3.8), the corresponding channel attenuation \( \alpha_{j\ell k} \) for all BSs follows the same model \[82\] :

\[
\alpha_{j\ell k} = \left( \frac{1}{d_{j\ell k}} \right)^{\eta_j} \Phi_{j\ell k},
\]

where \( d_{j\ell k} \) is the distance between the serving BS \( j \) and the relevant UE \( \ell k \). The exponent \( \eta_j \) accounts for the pathloss, and \( \Phi_{j\ell k} \) represents the correlated shadow fading.

Unless otherwise specified, we consider the following network configuration: The macro-cell is equipped with \( M_{macro} = 256 \) antennas and \( K_{macro} = 80 \) users and covers a circular area of radius \( R_{macro} = 250m \). The available power at the MBS is set to \( P_1 = 46dBm \) with a pathloss exponent \( \eta_1 = 4 \) and correlated shadow fading generated according to \[83\] equation.1] with standard deviation \( \sigma^2_{\Phi_1} = 8dB \). As for micro-cells, we set \( M_{micro} = 16 \) and \( K_{micro} = 5 \) with coverage area of radius \( R_{micro} = 35m \). The transmit power of the mBSs is fixed at 30dBm and the channel attenuation is configured with \( \eta_j = 3.5 \) and \( \sigma^2_{\Phi_{j\ell k}} = 8dB, j \neq 1 \). To prevent the cells from overlapping, we fix a minimum separation distance of \( 2R_{micro} \) between the BSs. We also set guard distances between the UEs and their serving BS. Denote by \( d_{min} \) and \( d_{max} \) the distances of
the closest and furthest users from the BS, respectively. For the MBS, we choose \((d_{\text{min}} = 40\,\text{m}, d_{\text{max}} = 249\,\text{m})\), and for the mBSs \((d_{\text{min}} = 1\,\text{m}, d_{\text{max}} = 25\,\text{m})\). In all simulations, we choose for SLNR based precoding \(\rho = 1\). We also impose the same quality of channel estimate for all UEs in the network, i.e., \(\tau_{i,\ell,j} = \tau, \forall (j, \ell, i)\). In all figures, we evaluate the performance of the MBs and one mBS in terms of the average achievable SE per UE defined in (3.6) and its deterministic equivalent 

\[
\overline{\text{SE}}_j = \frac{1}{K_j} \sum_{k=1}^{K_j} \mathbb{E} \left[ \log_2 (1 + \gamma_{jk}) \right].
\]

Figure 3.2: (a) Macro-cell and Micro-cell (b) achievable per-user SE (3.6) vs. SNR for No, and Full coordination with \(\beta_j = 1\), \(\forall j\) and different levels of CSI.

We propose in Fig. 3.2a and Fig. 3.2b to compare the average achievable spectral efficiency per UE when all BSs employ full coordination (F-coord. with \(\beta_j = 1\)), and when they do not coordinate (N-coord.), for different CSI levels in two precoding scenarios:

- **Scenario I**: The macro BS performs SLNR based precoding, while the micro BSs use ZFBF.

- **Scenario II**: SLNR based precoding schemes utilized at all tiers.

The dashed lines represent the deterministic equivalent of each associated curve. As seen, the simulation results match perfectly the theoretical analysis provided by
Theorems 2 and 3, thus validating the theoretical study in this Section 3.3.1. Moreover, as seen in Fig. 3.2b, the SLNR-based precoding outperforms the ZFBF especially when coordination is applied. Furthermore, as expected, in the perfect CSI case, coordination enhances the performance of both tiers, irrespective of the precoding design in use, as clearly depicted by the noteworthy increase in the SE (Fig. 3.2 all curves with $\tau^2 = 0$). Nevertheless, as explained in Section 3.3.1, having a good CSI quality is crucial to the well-functioning of coordination. This can be observed when investigating the performance results for low CSI levels, i.e. $\tau^2 = 0.1$, and particularly $\tau^2 = 0.3$, where the performance of coordination is very close to non-coordinated BF (even less than non-coordinated BF for ZFBF in Fig. 3.2b). Consequently, these results clearly assert the inefficiency of coordination when the CSI quality is poor, especially in high SNR levels in which coordinating can incur a performance loss.

3.4 CSI Impact on Coordination

As explained above, low CSI levels highly lessen the performance of coordinated beamforming which makes the value of the regularization parameter $\beta_j$ of crucial importance in the design of the coordinated precoding scheme. Therefore, to enable a better resilience to estimation errors, $\beta_j$ should be adjusted with the quality of the channel estimate. In this line, we propose and solve an optimization problem that aims at maximizing the SLNR to find the optimal $\beta_j^*$ that should be used in accordance to the level of available CSI. To do so, we first derive a deterministic equivalent of the SLNR as it provides a closed-form expression that is easier to manipulate than its random alternative.

3.4.1 Deterministic Equivalent of the SLNR

In this section, we derive closed-form approximations of the SLNR in (3.14). As shall be shown next, in addition to using them to determine a proper way of selecting
the regularization factor $\beta_j$, these approximations provide insights on the interference leakage terms $L_{jk}^\text{intra}$ and $L_{jk}^\text{inter}$ (3.14).

**Theorem 4.** Define, using the same notations as in Theorem 3, the following quantities:

$$S_{jk}^\text{co} = P_j \alpha_{jjk} \frac{1 - \tau_{jjk}^2}{c_j} \left[ 1 - \frac{1}{c_j} \tilde{\varrho}_{jk}^\text{co} (t_j^\text{co}) \right],$$

$$L_{jk}^\text{intra} = \sum_{i=1, i \neq k}^{K_j} \frac{P_j \alpha_{jji}}{K_j} \left[ (1 - \tau_{jji}^2) (1 - \Gamma_{jji}^\text{co} (t_j^\text{co}, \tilde{c}_j))^2 + \tau_{jji}^2 \right],$$

$$L_{jk}^\text{inter,}\ell = \sum_{i=1}^{K_\ell} \frac{P_j \alpha_{j,\ell,ji}}{K_j} \left[ (1 - \tau_{j,\ell,ji}^2) (1 - \Psi_{j,\ell,ji}^\text{co} (t_j^\text{co}, \tilde{c}_j))^2 + \tau_{j,\ell,ji}^2 \right].$$

Then, we have

$$\max_{1 \leq k \leq K_j} \left| S_{jk}^\text{co} - \overline{S}_{jk}^\text{co} \right| \xrightarrow{a.s.} 0,$$  

$$\max_{1 \leq k \leq K_j} \left| L_{jk}^\text{intra} - \overline{L}_{jk}^\text{intra} \right| \xrightarrow{a.s.} 0,$$

$$\max_{1 \leq k \leq K_j} \left| L_{jk}^\text{inter,}\ell - \overline{L}_{jk}^\text{inter,}\ell \right| \xrightarrow{a.s.} 0.$$

**Proof.** The proof relies on the same techniques used in Appendix 8 and is thus omitted.

**Corollary 5.** When the channel is perfectly known at all BSs ($\tau_{j\ell k} = 0, \forall j, \ell, k$), we have then:

$$S_{jk}^\text{co} = P_j \alpha_{jjk} \frac{1 - \tau_{jjk}^2}{c_j} \left[ 1 - \frac{1}{c_j} \tilde{\varrho}_{jk}^\text{co} (t_j^\text{co}) \right],$$

$$L_{jk}^\text{intra} = \frac{1}{K_j} \sum_{i=1, i \neq k}^{K_j} \frac{P_j \alpha_{jji}}{K_j} \left[ 1 + \frac{1}{c_j \rho \alpha} P_j \alpha_{jji} \tilde{e}_{jk}^\text{co} (t_j^\text{co}) \right]^2,$$

$$L_{jk}^\text{inter,}\ell = \frac{1}{K_\ell} \sum_{i=1}^{K_\ell} \frac{P_j \alpha_{j,\ell,ji}}{K_j} \left[ 1 + \frac{\beta_j}{c_j \rho \alpha} P_j \alpha_{j,\ell,ji} \tilde{e}_{jk}^\text{co} (t_j^\text{co}) \right]^2.$$
Again, we shall note that our asymptotic analysis allows to capture the impact of the large scale parameters on the SLNR performance. In this line, a close inspection of (3.60) and (3.61) reveals that curiously, the amount of interference leakage generated by the communication between BS\textsubscript{j} and UE\textsubscript{jk} to UEs in the other cells $\sum_{\ell=1}^{L} L_{\text{inter},\ell}^{\text{intra}}$, and to those in the same serving cell $j L_{jk}^{\text{intra}}$, is asymptotically the same for all UEs in cell $j$. In other words, the quantities $\sum_{\ell=1}^{L} L_{\text{inter},\ell}^{\text{intra}}$ and $L_{jk}^{\text{intra}}$ asymptotically, remain unchanged $\forall k = 1, \ldots, K_j$. Consequently, the overall interference leakage caused by BS\textsubscript{j} is independent of the underlying user to which it is transmitting. The reason behind this result lies in the precoding vector $\mathbf{w}_{jk}$ being of norm 1, which makes $L_{jk}^{\text{inter},\ell}$ independent of the channel statistics of UE\textsubscript{jk}. Additionally, one should note that this property holds true irrespective of the type of precoding being used. In fact, we can prove that this result is also valid for ZFBF when normalized to 1, by deriving its interference leakage expressions. Following the same steps as before, we find:

$$\begin{align*}
\begin{bmatrix}
L_{jk}^{\text{intra}} \\
L_{jk}^{\text{inter},\ell}
\end{bmatrix}
\end{align*}
_{ZF} = \sum_{i=1}^{K_j} P_j \alpha_{ji} \sum_{i \neq k} \frac{\tau_{ji} \tau_{ji}^2}{K_j}, \quad (3.68)
$$

$$\begin{align*}
\begin{bmatrix}
L_{jk}^{\text{intra}} \\
L_{jk}^{\text{inter},\ell}
\end{bmatrix}
\end{align*}
= \sum_{i=1}^{K_j} P_j \alpha_{j\ell} \sum_{i \neq k} \frac{\tau_{j\ell} \tau_{j\ell}^2}{K_j}, \quad (3.69)
$$

which are both independent of the channel statistics of UE\textsubscript{jk}.

### 3.4.2 Setting of the Coordination Factor $\beta_j$

In this section, we determine the optimal quantity $\beta_j^*$ in order to prevent the degradation that arises when the channel is poorly estimated, by maximizing the asymptotic equivalent of SLNR\textsuperscript{co}. However, prior to presenting the optimization problem, we first explain the motivation behind maximizing SLNR\textsuperscript{co} instead of the DL spectral efficiency.

To optimize the coordination factor $\beta_j$, one could consider the selection of the factor
\(\beta_j\) that maximizes the asymptotic average per UE SE, i.e,

\[
\beta_1^*, \ldots, \beta_L^* = \arg \max_{\beta_1, \ldots, \beta_L} \frac{1}{L} \sum_{j=1}^{L} \frac{1}{K_j} \sum_{k=1}^{K_j} \log_2(1 + \gamma_{jk}). \tag{3.70}
\]

However, this approach has several drawbacks. First, the optimization involves \(L\) variables \(\{\beta_j\}_{j=1}^{L}\), and thus characterizing the optimal values is not an easy task. Second, the optimization is coupled among the BSs and cannot be performed in a decentralized way, unless we assume that all BSs share their large scale statistics with a central node whose task is to evaluate the coordination factors \(\beta_j\).

To overcome these issues, we propose in the sequel a decentralized approach in which each BS selects independently its own coordination factor. More specifically, we propose in our approach that each BS selects the coordination factor that maximizes the asymptotic SLNR

\[\frac{S_{jk}^{\text{co}}}{\sum_{\ell=1, \ell \neq j}^{L} S_{jk}^{\text{inter}, \ell} + \rho \sigma^2}\]

given in Theorem 4. Note that this value is the same for all \(k = 1, \ldots, K_j\). As mentioned in the previous section, this independence is due to the fact that the interference leakage is the same for all UEs in cell \(j\), while the signal terms \(\{S_{jk}\}_{k=1}^{K_j}\) are equal up to a constant term independent of \(\beta_j\), thereby not affecting the optimization process. More formally, we propose to select the coordination level factor that solves the following optimization problem:

\[
\beta_j^* = \arg \max_{\beta_j} \text{SLNR}_{jk}, \tag{P2}
\]

\[s.t. \quad 0 < \beta_j < 1.\]

Note that in the perfect CSI case, the factor \(\beta_j\) solving (P2) should be equal to 1, as this coincides with the value that maximizes the instantaneous SLNR in (3.14). Thence, our approach is relevant exclusively in the imperfect CSI case. We expect that as the channel estimation quality deteriorates, the factor \(\beta_j\) solving (P2) should, in turn, decrease, unveiling the unreliability of the channel estimates used in coordination.
Using some mathematical manipulations\(^3\), we can show that the coordination level \(\beta_j\) solving (P2) is solution to the following fixed-point equation:

\[
\beta^*_j = \frac{I_j \left( \tilde{e}_j^{\text{co}} (t_j^{\text{co}}) \beta^*_j (e_j^{\text{co}} (t_j^{\text{co}}))^\prime - 1 \right)}{\left( \frac{\beta_j}{\tilde{e}_j^{\text{co}}} \right)^2 \left( \frac{\beta_j}{\tilde{e}_j^{\text{co}}} \right)^2 - \left( \frac{\beta_j}{\tilde{e}_j^{\text{co}}} \right)^2}.
\]

(3.71)

such that \(I_j\), \((e_j^{\text{co}} (t_j^{\text{co}}))^\prime\), \((S_j^{\text{co}}))^\prime\), \(\overline{\mathcal{N}}_{jk}^{\text{co}}\) and \(\mathcal{L}_{jk}^{\text{intra}}\) are given in (3.72)-(3.76), respectively.

\(\bullet\) \(I_j = \frac{2 \left( \tilde{e}_j t_j^{\text{co}} \right)^2}{K_j} \sum_{\ell = 1}^{L} \sum_{i = 1}^{K} \frac{1 - \tau_{j,\ell,i}^2}{P_j \alpha_{j,\ell,i}} \left( \Psi_{j,\ell,i}^{\text{co}} (t_j^{\text{co}}, \tilde{q}_j) \right)^3,\)

(3.72)

\(\bullet\) \((e_j^{\text{co}} (t_j^{\text{co}}))^\prime = 1 - \frac{1}{\tilde{e}_j} \tilde{e}_j^{\text{co}} \left( \frac{\beta_j^{\text{co}}}{\tilde{e}_j^{\text{co}}} \right)^2 \left( \frac{\beta_j^{\text{co}}}{\tilde{e}_j^{\text{co}}} \right)^2 - \left( \frac{\beta_j^{\text{co}}}{\tilde{e}_j^{\text{co}}} \right)^2,\)

(3.73)

\(\bullet\) \((S_j^{\text{co}})^\prime = 2P_j \alpha_{j,jk} (1 - \tau_{j,jk}^2) \frac{\tilde{e}_j t_j^{\text{co}}}{K_j} \left[ \sum_{i = 1}^{K} \frac{\left( \Gamma_{j,\ell,j}^{\text{co}} (t_j^{\text{co}}, \tilde{q}_j) \right)^3}{P_j \alpha_{j,\ell,i}} + \sum_{\ell = 1}^{L} \sum_{i = 1}^{K} \frac{\left( \Psi_{j,\ell,i}^{\text{co}} (t_j^{\text{co}}, \tilde{q}_j) \right)^3}{P_j \alpha_{j,\ell,i}} \right] \times \left( e_j^{\text{co}} (t_j^{\text{co}}) \right)^\prime - \frac{1}{\tilde{e}_j^{\text{co}} \left( t_j^{\text{co}} \right)^\prime} \sum_{\ell = 1}^{L} \sum_{i = 1}^{K} \frac{\left( \Psi_{j,\ell,i}^{\text{co}} (t_j^{\text{co}}, \tilde{q}_j) \right)^3}{P_j \alpha_{j,\ell,i}} \right] \),

(3.74)

\(\bullet\) \(\overline{\mathcal{N}}_{jk}^{\text{co}} = \mathcal{L}_{jk}^{\text{intra}} + \mathcal{L}_{jk}^{\text{inter,}\ell} + \rho \sigma^2,\)

(3.75)

\(\bullet\) \((\mathcal{L}_{jk}^{\text{intra}}) = 2 \frac{\tilde{e}_j t_j^{\text{co}}}{K_j} \left( \frac{e_j^{\text{co}} (t_j^{\text{co}})}{\tilde{e}_j^{\text{co}} \left( t_j^{\text{co}} \right)^\prime} \right)^\prime \sum_{i = 1}^{K} \frac{1 - \tau_{j,jk}^2}{P_j \alpha_{j,\ell,i}} \left( \Gamma_{j,\ell,j}^{\text{co}} (t_j^{\text{co}}, \tilde{q}_j) \right)^3.\)

(3.76)

\(^3\)Maximizing a fraction \(\frac{N}{D}\), where \(N \neq 0\) and \(D \neq 0\) can be achieved by solving the equation \(\frac{N'}{N} - \frac{D'}{D} = 0.\)
3.4.3 Numerical Results

We provide in this section a selection of numerical results to verify the asymptotic derivations in Theorem 4 as well as the optimal solution of $\beta^*_j$ (3.71). To this end, the same network settings as in 3.3.2 are considered. First, Fig. 3.3 illustrates SLNR$^\text{co}$ (3.14) (solid lines) and its asymptotic approximation (dashed lines) given in Theorem 4 at the macro-cell, with $(\beta_j = 1, \forall j)$ and different CSI levels. As can be seen, this figure clearly asserts the accuracy of our derivations for finite sized networks. Second, we analyze in Fig. 3.4a and Fig. 3.4b the performance gained by employing (or not) coordination, for imperfect CSI levels with different settings of $\beta_j$, including the optimal $\beta^*_j$ (3.71). In this line, Fig 3.4 shows that the best performance at both tiers is attained in the setting $\beta_j = \beta^*_j$ (Fig 3.4 black curves). In particular, for low CSI levels and at high SNR, it can be seen that optimal coordination outperforms the non-coordinated scenario, the coordination with an arbitrary choice of $\beta$, $\forall \tau$, and the full-coordinated (conventional SLNR based precoding) with $\beta_j = 1, \forall j$. It is also important to note that the perfect CSI curves are not included in these figures since, for $\tau^2 = 0$, $\beta^*_j = 1$, which
Figure 3.4: (a) Macro-cell and (b) micro-cell spectral efficiency vs. SNR, for No, Full, and Optimal Coordination, at different levels of CSI. Solid and dashed lines represent empirical and asymptotic results, respectively.

Figure 3.5: Macro-cell and micro-cell optimal regularization parameter $\beta^*$ with respect to the quality of CSI $(1 - \tau^2)$, for SNR=10dB.

leads to the same plots as in figures 3.2a and 3.2b. Accordingly and as we mentioned in Section 3.4.2 the provided optimal coordination factor in (3.71) depends highly on the level of available CSI. Evidently, $\beta^*$ takes higher values with the increase of the level of CSI which is depicted by $(1 - \tau^2)$ and clearly shown in Fig.3.5.
3.5 Chapter Summary

In this chapter, we focused on interference mitigation in massive MIMO two-tier HetNets. Assuming imperfect CSI, we, essentially, investigated the coordinated SLNR-MAX beamforming and its impact on the spectral efficiency, through closed-form approximations derived using random matrix theory tools. Throughout this study, we analytically demonstrated in what way coordination enhances the SE performance, and manifested the crucial effects of CSI levels on the effectiveness of such schemes. In this context, we further proposed a modified design of the SLNR based precoding, by implementing a normalizing factor for each cell. The value of this latter is obtained via the maximization of the system metric “SLNR”, taking into account the quality of CSI. Ultimately, the proposed precoding design was shown to outperform both non-coordinated and conventional coordinated SLNR based precoding, as it provides better resilience to estimation errors.

All the results we presented thus far consider massive MIMO with scattered channels modeled by Rayleigh fading. In the subsequent chapters, we take into account a more realistic channel model that is closer to practical propagation scenarios, by including a combination of Line-of-Sight and correlated scattered components.
Chapter 4

Statistical Combining for Massive MIMO Systems with Correlated Rician Fading

4.1 Introduction

In this chapter, we focus on the uplink performance of massive MIMO when the channels comprehend both scattered and specular components. We examine a general setting assuming imperfect channel estimation, and a system wherein each link is assigned a distinct correlation matrix and Rician factor. Note that this latter sets the ratio between the scattered and LoS signals. Using RMT asymptotic tools, we first develop approximations of the SE using the conventional LMMSE receiver and analytically demonstrate the impact of different factors, such as LoS and training sequence length, $\tau$, on the overall performance. After that, these approximations are further harnessed to derive an expression of the optimal training sequence’s length that maximizes SE gains.

Since the LoS channels are hardly changing, the BS may estimate the LoS components during the previous transmission from the users to the BS, in contrast to the Rayleigh signals which must be estimated at every coherence interval. Besides, as we shall demonstrate when using the conventional LMMSE receiver, choosing the right number of training symbols is crucial since a small $\tau$ generates higher estimation-errors whereas a bigger $\tau$ implies less transmitted data. Motivated by this, we also attempt in this chapter to investigate the following: for high levels of Rician factors, can we exploit the long-term parameters of the fading channels and avoid
channel estimation altogether? More formally, is there a setting wherein using solely the quasi-deterministic LoS component and the statistics of the Rayleigh channels to process the received signals would outperform the conventional channel-estimate based schemes? To this end, we propose an approach that is more convenient for LoS-prevailing environments, which we refer to as ‘statistical processing’. By utilizing the long-term statistics of the channels, this scheme enables us to avoid training as well as the channel-estimation-errors-induced SE losses. Plus, as shall be seen, above a certain level of the Rician factor, the proposed statistical receiver yields better performance than its conventional alternative, especially in multi-cell systems.

The study in the chapter is presented in two main parts. In the first part, we consider a single-cell massive MIMO system in order to streamline the effect of LoS signals and estimation errors in a simple setting. Then, in Part II, we explore a multi-cell network, wherein we further investigate whether the conclusions on LoS impacts and efficiency of the statistical receiver still hold when inter-cell interference is taken into account. Evidently, in each part, simulation results are provided to illustrate and concur our findings.

4.2 Part I - Performance Analysis in a Single-Cell Setting

We consider uplink transmissions of a TDD single-cell system with $K$ mono-antenna users, and a BS equipped with $N$ antennas (Fig[4.1]). Assuming Gaussian codebooks, the vector of the transmitted data symbols sent by all the UEs is denoted $\sqrt{\frac{p_u}{N}} \mathbf{x} \sim \mathcal{CN}(0, \frac{p_u}{N} \mathbf{I}_K)$ and therefore, the received signal at the BS writes:

$$
\mathbf{y} = \sqrt{\frac{p_u}{N}} \mathbf{H} \mathbf{x} + \mathbf{n},
$$

(4.1)

where $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$ is the $N \times K$ aggregated MIMO channel matrix from all UEs to the BS, and $\mathbf{n}$ represents a zero-mean additive Gaussian noise with variance $\sigma^2$. 
Correlated Rician fading channels are considered such that the channel between the $k$–th UE and the BS is modeled as:

$$h_k = \sqrt{\beta_k} \left( \sqrt{\frac{1}{1+\kappa_k}} \Theta_{k}^{\frac{1}{2}} z_k + \sqrt{\frac{\kappa_k}{1+\kappa_k}} z_k \right), \quad (4.2)$$

where $\beta_k$ accounts for the large-scale channel fading of UE$_k$ and the second term represents the small-scale fading channel. This latter consists of the Rayleigh component $z_k \sim \mathcal{CN}(0, I_N)$ to depict scattered or Non-LoS signals and the deterministic component $z_k \in \mathbb{C}^N$ to represent the specular (LoS) signals. For each UE, the ratio between these components is depicted by the Rician factor $\kappa_k$. Plus, for each UE $k$, we consider a different channel correlation matrix $\Theta_k$. Throughout the section, $\forall k$, $\Theta_k$ is assumed to be slowly varying compared to the channel coherence time and thus is supposed to be perfectly known to the BS. Finally, for notational convenience, we let: $\bar{h}_k = \sqrt{\frac{\beta_k \kappa_k}{1+\kappa_k}} z_k$ and $R_k = \frac{\beta_k}{1+\kappa_k} \Theta_k$. Therefore $h_k \sim \mathcal{CN} (\bar{h}_k, R_k)$, and:

$$h_k = R_k^{\frac{1}{2}} z_k + \bar{h}_k, \quad (4.3)$$
4.2.1 Channel Estimation

In practice, prior to processing the received signal, the BS estimates the channel matrix $\mathbf{H}$. Let $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \ldots, \hat{\mathbf{h}}_K]$ denote the aggregate matrix of these estimates. In TDD systems, each UL channel coherence block of length $T$ is split into two phases starting by training and followed with data transmission. In the pilot training interval of $\tau \geq K$ symbols, all $K$ UEs broadcast orthogonal sequences of known pilot symbols with average power $\tau p_p$. It is important to note that in the considered Rician fading, since the specular components are hardly changing, it is reasonable to assume that both the LoS component and Rician factors of all UEs are known to both the transmitter and receiver. Accordingly, using single-cell LMMSE estimation, the estimate $\hat{\mathbf{h}}_k$ of the channel $\mathbf{h}_k$ is given by [11]:

$$
\hat{\mathbf{h}}_k = \mathbf{R}_k \Phi_k \left( \mathbf{y}_k^{tr} - \mathbf{n}_k \right) + \mathbf{h}_k,
$$

(4.4)

where, $\Phi_k = \left( \mathbf{R}_k + \frac{1}{\tau \rho_{tr}} \mathbf{I}_N \right)^{-1}$, $\mathbf{y}_k^{tr} = \mathbf{h}_k + \frac{1}{\sqrt{\tau \rho_{tr}}} \mathbf{n}_k^{tr}$, and $\rho_{tr} = \frac{p_p}{\sigma^2}$ is the SNR corresponding to the training phase. The higher value $\tau \rho_{tr}$ takes, the better quality of channel estimation becomes. In fact, as $\tau \rho_{tr} \to \infty$, $\hat{\mathbf{H}} \to \mathbf{H}$ which corresponds to the perfect CSI scenario. From (4.4), it can be shown that $\hat{\mathbf{h}}_k \sim \mathcal{CN} \left( \mathbf{h}_k, \tilde{\mathbf{R}}_k \right)$, with $\tilde{\mathbf{R}}_k = \mathbf{R}_k \Phi_k \mathbf{R}_k$. Plus, considering the orthogonality property of LMMSE estimation, the estimation error $\xi_k = \mathbf{h}_k - \hat{\mathbf{h}}_k$ follows the distribution $\mathcal{CN} \left( 0, \mathbf{R}_k - \tilde{\mathbf{R}}_k \right)$.

4.2.2 Detection and Achievable Uplink Spectral Efficiency

To process the signal $\mathbf{y}$ (4.1), the BS uses a linear receiver. In this work, we are interested in the conventional LMMSE receiver that relies on acquired channel estimates, and propose a statistical receiver that is mainly based on the long-term parameters of the system.
4.2.2.1 Conventional LMMSE Receiver

Uplink pilots

Coherence Period $T$

Uplink Data Transmission

Figure 4.2: UL of block-fading system of coherence length $T$ symbols with pilot/data transmissions scheme.

Let $g_k \in \mathbb{C}^{N \times 1}$ denote the conventional combining vector used to process the signal sent by UE $k$. Under imperfect channel estimation conditions, $g_k$ is defined as [11,84]:

$$
g_k = \left( \hat{H}^H + \sum_{i=1}^{K} \left( R_i - \tilde{R}_i \right) + \frac{N}{\rho_d} I_N \right)^{-1} \tilde{h}_k. \quad (4.5)$$

In order to retrieve useful data sent by UE $k$, the BS generates $r_k = g_k^H y$. As shown in (4.5), the design of $g_k$ leverages the channel estimates $\hat{H}$, thus making the processing gain sensitive to channel estimation errors. Therefore, $r_k$ can be decomposed as:

$$
r_k = \sqrt{\frac{p_u}{N}} \sum_{i=1}^{K} g_k^H \hat{h}_i x_i + \sqrt{\frac{p_u}{N}} \sum_{i=1}^{K} g_k^H \xi_i x_i + g_k^H n_k, \quad (4.6)$$

This expression respectively, separates the signal and intra-cell interference, channel estimation errors and noise terms. Additionally, when a pre-training phase of $\tau$ symbols is performed, only a fraction of the total coherence block is used for useful data transmission (Fig.4.2). Therefore, putting $\rho_d = \frac{p_u}{\sigma^2}$, the achievable UL SE for UE $k$ in case of channel-estimate based conventional processing is defined as [6]:

$$
SE_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \mathbb{E} \left[ \log \left(1 + \frac{|g_k^H \hat{h}_k|^2}{\mathbb{E} \left[ \sum_{i \neq k} |g_k^H \hat{h}_i|^2 + \sum_{i=1}^{K} |g_k^H \xi_i|^2 + \frac{N}{\rho_d} \|g_k\|^2 \right]} \right) \right]. \quad (4.7)
$$

\[\text{In the sequel, we add the superscripts } (.)^{\text{conv},S}, (.)^{\text{stat},S}, (.)^{\text{conv},M} \text{ and } (.)^{\text{stat},M} \text{ to, respectively, distinguish the relevant quantities corresponding to conventional combining and statistical combining, in single-cell and multi-cell schemes.}\]
4.2.2.2 Statistical LMMSE Receiver

Due to its slow varying pace, the LoS component $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$ can be easily estimated. For example, the BS may estimate the specular signals in a previous transmission from the UEs, in contrast to the Rayleigh signals which must be estimated at every $T$. In addition, choosing the right number of training symbols, $\tau$, is paramount to ensure the overall UL performance since a small $\tau$ entails significant estimation errors and a larger $\tau$ suggests less transmitted data. Motivated by these factors, we propose in this work a statistical receiver denoted $\mathbf{g}_k$, whose design exclusively exploits the presence of the quasi-deterministic LoS component $\mathbf{H}$ and the long-term parameters of the system, such as the spatial correlation matrices $\mathbf{R}_k$, the large-scale fading factors $\beta_k$, etc. Naturally, using such a receiver enables to avoid training and channel estimation altogether, thereby yielding the single-cell UL SE, $SE_{k,\text{stat}}$:

$$SE_{k,\text{stat}} = \mathbb{E} \left[ \log \left( 1 + \frac{|\mathbf{g}_k^H \mathbf{H}_k|^2}{\mathbb{E} \left[ \mathbf{g}_k^H \left( \sum_{i=1}^{K} \mathbf{h}_i \mathbf{h}_i^H - \mathbf{h}_k \mathbf{h}_k^H + \frac{N \rho}{\rho_d} \mathbf{I}_N \right) \mathbf{g}_k^H \right] } \right) \right]. \quad (4.8)$$

We propose to design $\mathbf{g}_k$ through the maximization of a deterministic equivalent of $SE_{k,\text{stat}}$ in the infinite antenna limit which we denote $SE_{k,\text{stat}}$. Formally, for $k = 1, \ldots, K$, $\mathbf{g}_k$ is defined as:

$$\mathbf{g}_k = \arg \max_{\mathbf{g}_k} \mathbb{E} SE_{k,\text{stat}}, \quad (P1)$$

$$\text{s.t } SE_{k,\text{stat}} - \mathbb{E} SE_{k,\text{stat}} \xrightarrow{a.s.} 0.$$

As shall be seen in the next section, on account that $\mathbf{g}_k$ is deterministic, one should note that $\mathbb{E} SE_{k,\text{stat}}$ is obtained by means of quite rudimentary asymptotic tools.
4.2.3 Single-cell Spectral Efficiency Analysis

In this section, we carry out a comparative theoretical analysis between the UL performance achieved by the conventional receiver, $g_k [4.5]$, and the proposed statistical combiner, $\mathbf{g}_k [P1]$. The study is conducted under the assumption of imperfect channel state information, and a distinct Rician factor as well as channel correlation per user. Ultimately, the objective is to determine conditions in which the statistical receiver outperforms the conventional one. Nonetheless, as can be seen from $\text{SE}_k^{\text{conv, S}} [4.7]$ and $\text{SE}_k^{\text{stat, S}} [4.8]$, these SE expressions involve random quantities that are rather compact and do not lend themselves to simple interpretations nor manipulations. Accordingly, we first derive closed-form asymptotic approximations of both $\text{SE}_k^{\text{conv, S}}$ and $\text{SE}_k^{\text{stat, S}}$ which we exploit thereafter for the comparison. To obtain these approximations, we consider the large-antenna limit with a fixed number of UEs. This can be formulated as:

Assumption 6. We assume that $K$ is fixed while $N$ grows large without bound. We also consider that as $N \to \infty$, $\forall k$, the channel correlation matrix has a bounded spectral norm (i.e. $\limsup N \| R_k \|_2 < \infty$) to ensure that the signal energy is spread over several spatial directions and not focused on only few dimensions. Additionally, assume that $\liminf N \frac{1}{N} \text{tr} R_k > 0$ which implies that the array gathers an energy proportional to $\sqrt{N}$.

For simplicity, this asymptotic regime will be denoted by $N \to \infty$.

4.2.3.1 Conventional Combining in Single-Cell Systems

Define the matrices:

$$Q = \left( \frac{1}{N} \mathbf{H}^{H} \mathbf{H} + \frac{1}{N} \text{diag} \left\{ \text{tr} \tilde{R}_\ell \right\}_{\ell=1}^{K} + \frac{1}{\rho_d} \mathbf{I}_K \right)^{-1}, \quad (4.9)$$

$$T_i = \mathbf{H}^{H} \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi \mathbf{H} + \text{diag} \left\{ \text{tr} (\tilde{R}_\ell \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi \mathbf{H}) \right\}_{\ell=1}^{K}, \quad (4.10)$$
and let \( q_k \) be the \( k \)-th column of the matrix \( Q \).

**Theorem 5** (Conventional combining in single-cell systems). Under Assumption 6, we have: \( \text{SE}^{\text{conv},S}_k \rightarrow \text{SE}^{\text{conv},S}_k \) \( \text{a.s.} \rightarrow 0 \), such that

\[
\text{SE}^{\text{conv},S}_k = \left( 1 - \frac{T}{T} \right) \times \\
\log \left( 1 + \frac{1}{\rho_d} \frac{1}{Q_{kk}} \right) = \frac{1}{\rho_d} \left( Q_{kk} - \frac{1}{\rho_d} |Q_{kk}|^2 \right),
\]

(4.11)

**Proof.** A proof is given in Appendix 9.

We provide in the closed-form expression (4.11) approximations of all the different terms constituting \( \text{SE}^{\text{conv},S}_k \). This allows to have some insights on the behavior of these signals and their impact on the achievable SE. Note nonetheless that further simplifications can be made in the infinite antenna limit. For instance, we can see that as \( N \) grows infinitely large and for a fixed \( K \): \( \frac{1}{N^2} \sum_{i=1}^{K} q_k^H T_i q_k \) \( \text{a.s.} \rightarrow 0 \), therefore implying that channel estimation errors vanish in the UL massive MIMO setting. Accordingly, \( \text{SE}^{\text{conv},S}_k \) (4.11) amounts to:

\[
\text{SE}^{\text{conv},S}_k = \left( 1 - \frac{T}{T} \right) \log \left( \frac{\rho_d}{|Q_{kk}|} \right) + \mathcal{O} \left( \frac{1}{N} \right).
\]

(4.12)

Another key point in the former \( \text{SE}^{\text{conv},S}_k \) expression (4.11) is the term \( \frac{1}{\rho_d} \sum_{i=1}^{K} |Q_{ki}|^2 \), which represents an approximation of the total intra-cell interference. If Rayleigh fading is considered (i.e. \( H = 0_{N \times K} \)), \( Q \) becomes diagonal and this term cancels out. Pursuant to [11], this result confirms that in the setting \( N \rightarrow \infty \) while \( K \) is fixed, intra-cell interference due to Rayleigh fading dissipates. However, as we can

\[2\] Easily obtained by observing that \( |Q|^2 \) is diagonal.
see here, in Rician fading, the specular signals generate intra-cell interference which is embodied by the inner products between the components of $\mathbf{H}$. In light of this outcome, one way to eliminate interference is to have LoS components that are mutually orthogonal between users. This circumstance can be accomplished under asymptotic favorable propagation conditions $[85]$ where, $\frac{1}{N} \mathbf{h}_i^H \mathbf{h}_j \xrightarrow{a.s.} 0$, $i \neq j$, therefore yielding

$$\sum_{i=1}^{K} \frac{1}{\rho_d} \left| Q_{ki} \right|^2 \xrightarrow{a.s.} 0. \quad \text{as} \quad N \to \infty.$$ 

Hence, with the elimination of interference, we can conclude that for Rician fading, better performances are achieved in favorable propagation environments, specifically:

Corollary 6 (Favorable Propagation). if $\frac{1}{N} \mathbf{h}_i^H \mathbf{h}_j \xrightarrow{a.s.} 0$, for $i \neq j$, we have:

$$\text{SE}^\text{conv.S}_k = \left( 1 - \frac{\tau}{T} \right) \log \left[ 1 + \frac{\rho_d}{N} \left( \text{tr} \, \mathbf{R}_k + \| \mathbf{h}_k \|^2 \right) \right] + \mathcal{O} \left( \frac{1}{N} \right). \quad (4.13)$$

Furthermore, under the same settings of corollary 6, we demonstrate in $[86]$, where we compare LMMSE and Matched Filters (MFs), that the UL SE (4.13) is in fact identical to when MF is used. Accordingly, we find that in massive MIMO with Rician fading channels, LMMSE and MF receivers attain comparable performances only under favorable propagation conditions. This is, however, different from Rayleigh fading, wherein similar performances are obtained by the receivers (see $[2$, Eq.(13)] and $[11$, Remark 3.4$]$). On another note, from the expression of channel estimates $\hat{\mathbf{h}}_k$ (4.4), it can be shown by a simple eigenvalue decomposition that for low CSI, (i.e. $\tau \rho_{tr} \to 0$), $\text{tr} \, \hat{\mathbf{R}}_i \to 0$, $\forall \, i$. In such a case, we can see from the SE expressions $\text{SE}^\text{conv.S}_k$ (4.12) and (4.13), that the UL SE degrades with the deterioration of the CSI quality, and $\text{SE}^\text{conv.S}_k$ will be mainly determined by the specular signals. Accordingly, we can state that the strength of the LoS component is peculiarly beneficial when the channels are poorly estimated. By the same token, having reliable CSI becomes of greater importance as the LoS component weakens. Consequently, good channel estimates highly impact the performance; however, in Rician fading channels, it is also of utmost
relevance to have a receiver that exploits the presence and strength of the specular signals in an efficient manner.

4.2.3.2 Optimal Training

Define the SINR $\gamma_k(\tau)$ such that (4.12) writes: $\overline{SE}_k^{\text{conv},S} = (1 - \frac{\tau}{T}) \log (1 + \gamma_k(\tau))$.

Following the aforementioned statement on the importance of CSI quality, we determine in the next Theorem the optimal value $\tau^*$ that maximizes the achievable average SE. The objective is to determine the optimal number of symbols out of the total coherence symbols to be dedicated for training, for a fixed power allocation. Therefore, we set $\tau^* \geq K$ to preserve orthogonality of the pilot sequences, and evidently, $\tau^* < T$. Accordingly, $\tau^*$ is solution to the optimization problem:

$$\tau^* = \arg \max_{\tau} \frac{1}{K} \sum_{k=1}^{K} \overline{SE}_k^{\text{conv},S},$$

$$s.t. \ K \leq \tau < T.$$ (P2)

**Theorem 6 (Optimal training).** Under imperfect channel estimates, the optimal training length is given by:

- If:
  $$\frac{1}{K} \sum_{k=1}^{K} \left[ (T - K) \frac{\gamma'_k(K)}{1 + \gamma_k(K)} - \log (1 + \gamma_k(K)) \right] \leq 0, \quad (4.14)$$
  then $\tau^* = K$.

- Otherwise: $\tau^*$ is the solution to the fixed point equation:
  $$\tau^* = T - \frac{1}{K} \sum_{k=1}^{K} \log (1 + \gamma_k(\tau)), \quad (4.15)$$
  where $\gamma'_k(\tau)$ is the derivative of $\gamma_k(\tau)$ with respect to $\tau$.

**Proof.** A proof of Theorem 6 is given in Appendix 10. \qed
Note that the derivation of $\tau^*$ relies on $\text{SE}_k^{\text{conv,S}}$ considering its explicit form relatively to the intractable alternative, $\text{SE}_k^{\text{conv,S}}$. The results of Theorem 6 will be validated by simulations. Nevertheless, they can be exploited to acquire some insights on the behavioral tendencies of the choice of $\tau$ and the overall uplink performance. In this context, an interesting direction is the impact of the Rician factor $\kappa$ on the value of $\tau$. In order to investigate this point, we consider the following case study.

**Case study**

Let us examine the case where if $K$ is kept fixed and $N$ grows without bound:

$$
\frac{1}{N} \bar{h}_i^H \bar{h}_j \xrightarrow{a.s.} N \rightarrow \infty \frac{1}{1 + \kappa_k} \delta_{ij}, \text{ where } \delta_{ij} \text{ is the Kronecker delta. Additionally, let } \forall k : R_k = \frac{\beta_k}{1 + \kappa_k} I_N, \beta_k = \beta \text{ and } \kappa_k = \kappa. \text{ Thus, according to corollary 6, the UL SINR becomes:}
$$

$$
\gamma(\tau) = \frac{\beta \rho_d}{1 + \kappa} \left( \frac{1}{1 + \kappa + \frac{1}{\tau \rho_d}} + \kappa \right). \quad (4.16)
$$

- **Low Rician factor**: Consider small values of $\kappa$. At a low SNR level ($\rho_d$ approaches 0), the solution (4.15) can be rewritten as:

$$
\lim_{\kappa \to 0} \tau^* = T - \frac{(1 + \beta \rho_d \tau)(1 + \beta \rho_d \tau + \beta^2 \rho_d \tau \rho_d) \log \left(1 + \frac{\beta^2 \rho_d \tau}{1 + \beta \rho_d \tau} \right)}{\beta^2 \rho_d \rho_d}. \quad (4.17)
$$

Using Taylor’s expansion in the low SNR regime yields:

$$
\lim_{\kappa \to 0} \tau^* = \max \left\{ K, \frac{-1 + \sqrt{1 + \beta \rho_d T}}{\beta \rho_d} \right\}. \quad (4.18)
$$

- **High Rician factor**: For high values of the Rician factor, $\kappa$, we find $\gamma'(K) = 0$. Therefore, (4.14) is always verified and:

$$
\lim_{\kappa \to \infty} \tau^* = K. \quad (4.19)
$$
This case study sheds some light on how the optimal number of training symbols depends on the large-scale fading parameters, the number of users, the UL SNR and the coherence interval. The first example represents the case wherein the Rayleigh fading is governing at poor SNR levels. As can be seen from (4.18), \( \tau^* \) depends on the system parameters and on the available SNR during training, \( \rho_{tr} \). For instance, if this latter is also low, (4.18) yields \( \lim_{\kappa \to 0} \tau^* = \max \{ K, \frac{T}{2} \} \). This result implies that in a network setting where \( T > 2K \), to ensure the best performance, half of the total transmitted symbols should be dedicated to training and the other half to useful data. Conversely, if more users are considered such that \( T \leq 2K \), then, the optimal number of training symbols should not go beyond the imposed minimum, \( K \).

In the second example, as the Rician factor takes higher values, we find that \( \tau^* \) always approaches \( K \) (4.19). Consequently, in such circumstances, there is no need to perform any optimization since the optimal number of training symbols is limited to the minimum possible value to ensure pilot orthogonality, namely \( K \). More importantly, we deduce that above a certain \( \kappa \), investing in more training samples is not optimal in terms of spectral efficiency as it will have a minor impact on the achievable UL SE. In fact, it might even induce performance losses when \( \tau \gg K \), as shall be illustrated in simulations. This result motivates us to analyse, in the next section, the potential outcomes of employing the statistical receiver \( \mathbf{g}_k \) (P1) in LoS-prevailing environments.

### 4.2.3.3 Statistical Combining in Single-Cell Systems

As previously mentioned, the objective of having a statistical receiver is to eliminate both training and channel estimation and to exploit the presence of the Rician component efficiently. To this end, the proposed combining vector \( \mathbf{g}_k \) is obtained through the maximization of a deterministic approximation of the UL spectral efficiency \( \text{SE}_{k}^{\text{stat},S} \) (4.8), as depicted in (P1). Taking into account that \( \mathbf{g}_k \) is deterministic itself, a direct application of the convergence of quadratic forms lemma (Chapter 2 lemma 3)
Theorem 7 (Statistical combining in single-cell systems). Under Assumption 6, 
\[ SE_k^{\text{stat},S} \xrightarrow{\text{a.s.}} 0, \quad N \rightarrow \infty, \] 
with:

\[
SE_k^{\text{stat},S} = \log \left[ 1 + \frac{\mathbf{g}_k^H \left( \frac{\mathbf{h}_k \mathbf{h}_k^H}{N} \right) \mathbf{g}_k}{\mathbf{g}_k^H \left( \frac{1}{N} \sum_{i=1}^K \mathbf{R}_i + \frac{1}{N} \mathbf{H}_k \mathbf{H}_k^H + \frac{1}{\rho_d} \mathbf{I}_N \right) \mathbf{g}_k} \right], \quad (4.20)
\]

where \( \mathbf{H}_k \) is obtained by removing the \( k \)-th column from the LoS-channels matrix \( \mathbf{H} \).

Using the expression \((4.20)\), we can now derive \( \mathbf{g}_k \) by solving the SE optimization problem \((P1)\). Note that the problem \((P1)\) is the sum of decoupled positive and increasing functions. Therefore, a sufficient condition to solve it is to find \( \forall \ k, \ k = 1, \ldots, K \), \( \mathbf{g}_k \) that satisfies:

\[ \mathbf{g}_k = \arg\max_{\mathbf{g}_k} \frac{\mathbf{g}_k^H \left( \frac{\mathbf{h}_k \mathbf{h}_k^H}{N} \right) \mathbf{g}_k}{\mathbf{g}_k^H \left( \frac{1}{N} \sum_{i=1}^K \mathbf{R}_i + \frac{1}{N} \mathbf{H}_k \mathbf{H}_k^H + \frac{1}{\rho_d} \mathbf{I}_N \right) \mathbf{g}_k}. \quad (P1') \]

It can be seen that \((P1')\) is equivalent to Rayleigh quotient and thus admits the solution:

\[ \mathbf{g}_k = \left( \sum_{i=1}^K \mathbf{R}_i + \mathbf{H}_k \mathbf{H}_k^H + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \mathbf{h}_k \quad (4.21) \]

Consequently, under assumption 6,

\[
SE_k^{\text{stat},S} = \log \left[ 1 + \frac{1}{N} \mathbf{h}_k^H \left( \frac{1}{N} \mathbf{H}_k \mathbf{H}_k^H + \frac{1}{\rho_d} \mathbf{I}_N \right)^{-1} \mathbf{h}_k \right] + \mathcal{O} \left( \frac{1}{N} \right). \quad (4.22)
\]

Furthermore, under favorable propagation conditions, using the woodbury matrix identity (Chapter 2, lemma 1), we find:

\[ SE_k^{\text{stat},S} = \log \left( 1 + \frac{\rho_d}{N} \| \mathbf{h}_k \|^2 \right) + \mathcal{O} \left( \frac{1}{N} \right). \quad (4.23) \]
As can be seen from (4.22)-(4.23), the UL SE generated by $g_k$ is mainly determined by the level of the specular component. That is, the proposed statistical processing is essentially beneficial in LoS-prevailing environments, thereby requiring a certain level of $\kappa$.

4.2.3.4 Comparative Analysis (Case Study)

As concluded above and will be illustrated in simulations, the statistical processing is convenient when the specular component is dominant over the scattered signals. For this reason, we aspire here to find a condition on the Rician factor $\kappa_k$ under which the proposed statistical processing outperforms the conventional processing. We examine a simple network setting where: $\frac{1}{N} \mathbf{h}_i^H \mathbf{h}_j \overset{a.s.}{\to} \frac{\beta_i \kappa_i}{1 + \kappa_i} \delta_{ij}$, where $\delta_{ij}$ is the Kronecker delta.

Formally, we determine $\kappa_k$, $\forall k$, s.t.:

$$\log \left( 1 + \frac{\rho_d \beta_k \kappa_k}{1 + \kappa_k} \right) \geq \left( 1 - \frac{T}{T} \right) \log \left[ 1 + \rho_d \left( \frac{1}{N} \text{tr} \hat{R}_k + \frac{\beta_k \kappa_k}{1 + \kappa_k} \right) \right]$$  \quad (P3)

- **Result:** Under Assumption 6, taking into account that $\frac{1}{N} \mathbf{h}_i^H \mathbf{h}_j \overset{a.s.}{\to} \frac{\beta_i \kappa_i}{1 + \kappa_i} \delta_{ij}$ and $\tau \in [K, T)$, it can be shown that the statistical processing outperforms the conventional channel-estimate based processing, if $\kappa_k$ verifies the sufficient condition\textsuperscript{3}

$$\kappa_k \geq \frac{\text{tr} \Theta_k T - K}{K}. \quad (4.24)$$

*Proof.* A proof is given in Appendix 11

Furthermore, if the correlation matrix $\Theta_k$ follows the widely used one-ring model \textsuperscript{87} (introduced in the next section (4.25)) or the exponential correlation model \textsuperscript{88}, then $\frac{1}{N} \text{tr} \Theta_k = 1$, $\forall k$, therefore, (4.24) writes : $\kappa_k \geq \frac{T - K}{K}$. These inequalities provide a lower bound on a sufficient Rician factor above which the statistical combining is

\textsuperscript{3}As shown in proof, for mathematical convenience, we consider a higher bound than it is necessary.
more profitable than the conventional one. This bound is a function of the systems parameters, including the coherence length and the number of users. Moreover, as can be seen, for a fixed $T$, $\frac{T - K}{K}$ is a decreasing function of $K$. As a result, the higher is the number of users, the smaller is the required $\kappa_k$ to enable the use of the proposed statistical receiver with better UL performance and as such, avoid training along with channel estimation and its associated performance losses.

### 4.2.4 Numerical Results for the Single-cell setting

In this section, we carry out MonteCarlo simulations over 1000 channel realizations to validate, for finite system dimensions, the asymptotic results of the single-cell setting given in Section 4.2.

For this scenario, we consider a single-cell massive MIMO having one BS with $N = 150$ antennas, $K = 20$ users, and a coherence length $T_c = 500$ symbols. The inner cell-radius is 150m and the users are uniformly distributed around the BS at an arrival angle $\theta_k$. Furthermore, the pathloss is given $\beta_k = \frac{1}{x_k}$, where $x_k$ is the distance between UE $k$ and the BS and $\alpha = 2.5$. The specular component $z_k$ follows the model $[z_k]_n = e^{-j(n-1)\pi \sin(\theta_k)}$. Moreover, to ensure distinct Rician factors among the users, we assume throughout all the simulations that $\forall k$, $\kappa_k \sim \mathcal{U}[0, \kappa_{\text{max}}]$. As a result, varying $\kappa_{\text{max}}$ yields specular signals with different levels of strength which ultimately enables to epitomize both NLoS and LoS prevailing environments. Finally, the elements of the correlation matrix $\Theta_k$ of channel $h_k$ are given by the one ring model [87]:

$$
[\Theta_k]_{uv} = \frac{1}{\theta_{k,\text{max}} - \theta_{k,\text{min}}} \int_{\theta_{k,\text{min}}}^{\theta_{k,\text{max}}} e^{i\frac{2\pi}{\lambda} a_{uv} \cos(\theta)} d\theta,
$$

where $\lambda$ denotes the signal’s wavelength and $a_{uv}$ is the distance between receive antennas $u$ and $v$. We also choose $\frac{a_{uv}}{\lambda} = 0.5|u - v|$, $\theta_{k,\text{min}} = -\pi$ and $\theta_{k,\text{max}} = \theta_k - \pi$.

First, we focus on the conventional combining to validate the conclusions of
Figure 4.3: Single-cell setting: (a) Impact of $\tau$ on the UL SE using conventional combining \((4.5)\), for different levels of Rician factor, s.t. $\kappa_k \sim \mathcal{U}(0,\kappa_{\text{max}})$. (b) Optimal number of training symbols $\tau^*$ \((4.15)\) for different levels of $\kappa_k$.

In this line, we illustrate the effects of the LoS and the length of the training sequence on the SE when the conventional combining is utilized. To this end, we plot in Fig. 4.3a the achievable spectral efficiency $\text{SE}_{k}^{\text{conv},S}$ \((4.7)\) for different levels of the Rician factors including: $\kappa_{\text{max}} = 10, 4, 0.5$, and 0 (corresponding to Rayleigh fading). Moreover, to manifest the importance of the number of training symbols, we represent in the same figure $\text{SE}_{k}^{\text{conv},S}$ with different values of $\tau$, namely: the minimum $K$, the optimal $\tau^*$ \((4.15)\), and another arbitrary value ($\neq \tau^*$ nor $K$). Solid and dotted lines represent empirical and asymptotic SEs, respectively. In addition, for each one of these $\kappa_{\text{max}}$ settings, we plot in Fig. 4.3b the corresponding optimal values $\tau^*$ obtained by \((4.14)-(4.15)\), with respect to the SNR.

Overall, as expected, the LoS has a beneficial impact since increasing the Rician factor enables higher SEs, for any value of $\tau$. As to this latter, it can be seen that the best performance is clearly obtained when the optimal number of symbols given in Theorem 6 \((i.e. \tau = \tau^*)\) is considered (Fig. 4.3a, diamond-marked curves). Furthermore, note that the gap between the settings $\tau = \tau^*$ and $\tau = K$ is particularly noteworthy.
at small values of Rician factors, \( (\text{Fig. 4.3a}, \kappa_{\max} = 0 \text{ and } 0.5, \text{i.e. } 0 \leq \kappa_k \leq 0.5, \forall k) \). However, this difference in performance becomes less significant as \( \kappa_k \) takes higher values \( (\text{Fig. 4.3a}, \text{curves } \kappa_{\max} = 4 \text{ and } 10, \text{i.e. } 0 \leq \kappa_k \leq 10) \). These simulation results confirm that as \( \kappa \) takes higher values, \( \tau^* \to K \) as also displayed in \( \text{Fig. 4.3b} \). Indeed, \( \text{Fig. 4.3b} \) clearly asserts that, for small Rician factors \( (\kappa_{\max} = 0 \text{ and } 0.5) \), \( \tau^* \) takes increasingly higher levels with the increase of the SNR. Conversly, \( \text{Fig. 4.3a} \) indicates that \( \tau^* \to K \) when \( \kappa_{\max} = 10 \). Evidently, since \( \tau^* \to K \) for these cases, the SEs corresponding to \( \tau = K \) and \( \tau = \tau^* \) are almost identical \( (\text{Fig. 4.3a, overlapping diamond and circle-marked curves}) \); whereas, interestingly, for \( \tau > K \) (represented by the square-marked curve), we observe a decrease in the SE. Consequently, the plots in \( \text{Fig. 4.3a} \) and \( \text{Fig. 4.3b} \) validate the conclusions of Section 4.2.3 and the case study (in section 4.2.3.2) pointing the importance of assigning the optimal number of training symbols to attain the best performance. These simulations also manifest that above a certain level of \( \kappa_{\max} \), i.e. as the LoS gets stronger, investing in longer training sequences to enhance the spectral efficiency is actually counterproductive.

This interesting outcome inspired the proposed statistical combining that is a more opportune approach in such environments, as was previously demonstrated in Section 4.2, and is illustrated in what follows.

Under the aforementioned network setting and for different values of \( \kappa \), we compare in \( \text{Fig. 4.4} \) the UL spectral efficiency \( SE_{k,\text{stat}}^S (4.8) \) achieved using the statistical receiver with the one attained by the conventional technique \( SE_{k,\text{conv}}^S (4.7) \), assuming optimal training \( (i.e. \tau = \tau^*) \). Plus, we represent ordinary and favorable propagation conditions in \( \text{Fig. 4.4a} \) and \( \text{Fig. 4.4b} \) respectively. First, comparing \( \text{Fig. 4.4a} \) with \( \text{Fig. 4.4b} \) reveals that favorable propagation enables a better performance for both combining techniques. This consequence, as explained in Theorem 5, is due to the cancellation of LoS induced intra-cell interference when the specular signals are mutually orthogonal. Second, as can be seen in both propagation conditions, conventional LMMSE is more beneficial
than the statistical combiner at low ranges of the Rician factor, \( (Fig. 4.4, \kappa_{\text{max}} = 0.5) \).

Nonetheless, as the LoS component becomes stronger, \( SE_{\text{k,stat}} \) progressively approaches \( SE_{\text{k,conv}} \), up to generating exceeding gains starting at \( \kappa_{\text{max}} = 4 \) in ordinary conditions, and \( \kappa_{\text{max}} = 1.5 \) in favorable propagation. This consequence can be justified by the expression \( SE_{\text{k,stat}} \) that clearly demonstrates that the statistical receiver’s performance is mainly determined by the strength of the LoS components. Therefore, these results confirm our single-cell analysis by substantiating the existence of a \( \overline{\kappa} \) above which the statistical processing outperforms the conventional one. In the same line, this threshold value is fairly lower in favorable propagation compared to ordinary propagation environments. It is also important to note that these simulations concur the findings in Section 4.2.3.4 and even extend it to a more realistic scenario that accounts for different per-user correlations and Rician factors. Finally, Fig.4.3 and Fig.4.4 validate, for finite system dimensions, the accuracy of the asymptotic approximations derived in Theorems 5, 6 and 7.
4.3 Part II - Performance Analysis in a Multi-cell scenario

In this section, we extend our analysis of both conventional and statistical combining schemes to a multi-cell scenario. The objective is to examine whether the proposed statistical combining scheme remains profitable in LoS-prevailing conditions that include inter-cell interference. Therefore, we consider a multi-cell network with $L$ cells having each $K$ single-antenna users communicating with an $N$-antennas BS. Towards this aim, we follow the same notations as in the single-cell scenario in Part I(4.2), except that we adopt here a triple sub-script indexation to differentiate the receiving BS from the cell where the UE is located (similarly to Chapter [3]). Specifically, in the sequel, any quantity with the index $(.)_{j\ell k}$ is associated with channel $h_{j\ell k}$ that links BS$_j$ to UE$_{\ell k}$ ($k$–th UE in cell $\ell$). Plus, $(.)_{jk}$ depicts the associated quantity between BS$_j$ and its own $k$–th UE. For instance $\Theta_{j\ell k}$ refers to the correlation matrix of the channel $h_{j\ell k}$, and $SE_{jk}$ refers to the spectral efficiency associated with the communication between UE$_{jk}$ and its receiving BS$_j$.

Accordingly, the received signal at BS$_j$ is given by:

$$y_j = \sqrt{p} \sum_{\ell=1}^{L} \sum_{i=1}^{K} h_{j\ell i} x_{\ell i} + n_j,$$

where $n_j$ represents a zero-mean additive Gaussian noise with variance $\sigma^2 I_N$. Plus, we consider correlated Rician fading for intra-cell or local channels, and correlated Rayleigh fading for channels from other cells. This is a reasonable setting for inter-cell channels, owing to the longer distances between UEs and the BSs in other cells that would likely include scatterers and thus, significantly reduce the possibility of a Line-of-Sight
transmission. Formally, the channel linking UE$_{\ell k}$ to BS$_j$ is modeled as:

$$h_{jjk} = \sqrt{\beta_{jjk}} \left( \sqrt{\frac{1}{1 + \kappa_{jk}}} \Theta_{jjk}^{\frac{1}{2}} z_{jjk} + \sqrt{\frac{\kappa_{jk}}{1 + \kappa_{jk}}} z_{jk} \right), \quad (4.27)$$

$$h_{j\ell k} = \sqrt{\beta_{j\ell k}} \Theta_{j\ell k}^{\frac{1}{2}} z_{j\ell k}, \quad \ell \neq j, \quad (4.28)$$

where $\beta_{j\ell k}$ accounts for the large-scale fading. Finally, similarly to Part I, to simplify the presentation of the results, let $R_{j\ell k} = \frac{\beta_{j\ell k}}{1 + \kappa_{jk}} \Theta_{j\ell k}$, and the aggregate matrix of the LoS components in cell $j$ denoted $\tilde{H}_j = [\tilde{h}_{j1} \tilde{h}_{j2} \ldots \tilde{h}_{jK}]$, with $\tilde{h}_{jk} = \sqrt{\frac{\beta_{jk} \kappa_{jk}}{1 + \kappa_{jk}}} z_{jk}$. Hence,

$$h_{jjk} = R_{jjk}^{\frac{1}{2}} z_{jjk} + \tilde{h}_{jk}, \quad (4.29)$$

$$h_{j\ell k} = R_{j\ell k}^{\frac{1}{2}} z_{j\ell k}, \quad \ell \neq j, \quad (4.30)$$

In order to investigate the statistical receiver and draw rigorous conclusions on its performance in the multi-cell setting, we pursue the following approach. To better highlight the advantages behind the statistical scheme, we propose to compare its spectral efficiency denoted $\text{SE}^{\text{stat},M}$ with $\text{SE}^{\text{conv},M}$ that can be achieved by the conventional combining in similar systems. Therefore, we first introduce and analyze $\text{SE}^{\text{conv},M}$ in the next subsection, then in subsequent section 4.3.2, we delve into the study of $\text{SE}^{\text{stat},M}$ through a comparative analysis with $\text{SE}^{\text{conv},M}$.

### 4.3.1 Conventional combining in Multi-Cell Systems

**Detection**

Similarly to the single-cell scenario, to design the conventional receiver, we consider a pre-training phase of $\tau$ symbols in each cell. However, to account for pilot contamination, we assume that the same set of pilot sequences is reused in every cell. More specifically, the same pilot is assigned to every $k$-th UE in each cell, and as such $\forall (j, k)$, the
estimates of the channels \( h_{jk}, h_{j2k}, \ldots, h_{jLk} \), will be correlated. Accordingly, using the MMSE estimation, the estimate of \( h_{jjk} \) is given by [84]:

\[
\hat{h}_{jjk} = R_{jjk} \Phi_{jk} \left( \sum_{\ell'=1}^{L} h_{\ell'k} + \frac{1}{\sqrt{\tau \rho_{tr}}} n_{jk}^{tr} \right) + \delta_{j\ell} h_{jk}, \quad (4.31)
\]

where \( \Phi_{jk} = \left( \sum_{\ell'=1}^{L} R_{\ell'k} + \frac{1}{\tau \rho_{tr}} I_N \right)^{-1} \), and \( \delta_{j\ell} \) is the Kronecker delta. Therefore, \( \hat{h}_{jjk} \sim \mathcal{CN} \left( \tilde{h}_{jk}, \tilde{R}_{jjk} \right) \), with \( \tilde{R}_{jjk} = R_{jjk} \Phi_{jk} R_{jjk} \).

Let \( g_{jk} \in \mathbb{C}^{N \times 1} \) denote the conventional combining vector that BS \( j \) uses to process the signal sent by its UE \( k \). This vector is given by [11]:

\[
g_{jk} = \left( \sum_{i=1}^{K} \hat{h}_{jji} \hat{h}_{jji}^H + A_j + \frac{N}{\rho_d} I_N \right)^{-1} \hat{h}_{jjk}. \quad (4.32)
\]

where \( A_j \in \mathbb{C}^{N \times N} \) is an arbitrary hermitian positive semi-definite design parameter. For instance, it could contain the covariances of estimation errors and inter-cell interference as in [11], thus yielding:

\[
A_j = \sum_{i=1}^{K} (R_{jji} - \tilde{R}_{jji}) + \sum_{\ell' \neq j}^{L} \sum_{i=1}^{K} R_{j\ell'i}. \quad (4.33)
\]

**Asymptotic Analysis**

**Theorem 8** (Conventional combining in multi-cell systems). Under Assumption [6], we have:

\[
\text{SE}_{\text{conv,}M} = \left( 1 - \frac{T}{T} \right) \mathbb{E} \left[ \log \left( 1 + \frac{|g_{jk}^H \hat{h}_{jjk}|^2}{\mathbb{E} \left[ g_{jk}^H \left( \sum_{(i,i) \neq (j,k)} h_{\ell'i} h_{\ell'i}^H + \xi_{jk} \xi_{jk}^H + \frac{1}{\rho_d} I_N \right) g_{jk}^H \tilde{H}_j \right]} \right) \right].
\]

(4.33)

**Asymptotic Analysis**

**Theorem 8** (Conventional combining in multi-cell systems). Under Assumption [6], we have:

\[
\text{SE}_{\text{conv,}M} \rightarrow \text{SE}_{\text{conv,}M}, \quad \text{a.s.} \quad N \rightarrow \infty, \text{ such that:}
\]

\[
\begin{align*}
\text{SE}_{\text{conv,}M} & = \left( 1 - \frac{T}{T} \right) \mathbb{E} \left[ \log \left( 1 + \frac{|g_{jk}^H \hat{h}_{jjk}|^2}{\mathbb{E} \left[ g_{jk}^H \left( \sum_{(i,i) \neq (j,k)} h_{\ell'i} h_{\ell'i}^H + \xi_{jk} \xi_{jk}^H + \frac{1}{\rho_d} I_N \right) g_{jk}^H \tilde{H}_j \right]} \right) \right].
\end{align*}
\]
\[
\overline{\text{SE}_{jk}^{\text{conv},M}} = \left(1 - \frac{\tau}{\bar{T}}\right) \log \left[1 + \frac{1 - \frac{1}{\rho_d} [\overline{Q}_{j}]_{kk}^2}{\sum_{i=1, i \neq k}^K \left|\frac{1}{\rho_d} [\overline{Q}_{j}]_{ki}\right|^2 + \frac{1}{\rho_d} ([\overline{Q}_{j}]_{kk} - \frac{1}{\rho_d} [\overline{Q}_{j}]_{kk}^2) + \sum_{\ell=1}^L \sum_{i=1, i \neq \ell}^K \left|\frac{1}{N} [\overline{Q}_{j}]_{ki} \operatorname{tr}(\overline{R}_{j\ell i} \Phi_{ji} R_{jji})\right|^2}\right].
\]

(4.34)

with

\[
\overline{Q}_j = \left(\frac{1}{N} \overline{\mathbf{H}}_j \overline{\mathbf{H}}_j + \operatorname{diag} \left\{ \frac{1}{N} \operatorname{tr}(\overline{R}_{jji}) \right\}^K + \frac{1}{\rho_d} I_K \right)^{-1},
\]

(4.35)

**Proof.** The proof is given in Appendix 12

--

As can be seen from the approximation \(\overline{\text{SE}_{jk}^{\text{conv},M}}\) (4.34), the achievable UL SE in the multi-cell setting has an analogous expression to the single-cell scenario (4.11), apart from the inter-cell interference represented by the last term in the denominator of the above \(\overline{\text{SE}_{jk}^{\text{conv},M}}\) (4.34). As a result, most conclusions provided in section 4.2.3, Theorem 5 (conventional combining in single-cell systems) hold true in the multi-cell setup. These conclusions include the cancellation of the estimation errors as \(N\) grows large without bound; not to mention that intra-cell interference is also generated by the inner products between the LoS components, and as such, dissipates under favorable propagation conditions.

We now move on to studying the effect of inter-cell interference on this combining approach. In this line, by simplifying the denominator and expanding the inter-cell
interference, $\bar{SE}_{jk}^{\text{conv},M}$ writes:

$$
\bar{SE}_{jk}^{\text{conv},M} = \left( 1 - \frac{\tau}{T} \right) \log \left( 1 + \frac{\rho_d}{[Q_j]_{kk}} \left( \frac{1}{[Q_j]_{kk}} - 1 \right)^2 \right)
\frac{\rho_d}{[Q_j]_{kk}} - 1 + \sum_{\ell \neq j} \left( \frac{\rho_d}{N} |\text{tr}(R_{j\ell k} \Phi_j R_{jkk})|^2 \right) + \sum_{i \neq k} \left( \frac{\rho_d}{N} |[Q_j]_{ki}|^2 \text{tr}(R_{j\ell i} \Phi_j R_{jjj})^2 \right) \right).
$$

Expression (4.36) separates the pilot contamination induced interference from the remaining inter-cell interference, which we refer to as “uncorrelated” interference. Clearly, inter-cell interference limits the overall performance even at the infinite-antenna limit. Nevertheless, note that its impact can be alleviated through the mitigation of the uncorrelated interference by observing that this latter is eliminated when $Q_j$ becomes diagonal. In fact, this is achieved in favorable propagation conditions, wherein the SE will attain:

Corollary 7 (Favorable propagation in multi-cell). If $\frac{1}{N} \bar{h}_{ji} \bar{h}_{jk} \overset{a.s.}{\longrightarrow} 0$, for $i \neq k$, we have:

$$
\bar{SE}_{jk}^{\text{conv},M} = \left( 1 - \frac{\tau}{T} \right) \log \left( 1 + \frac{\rho_d}{N} \text{tr}(\bar{R}_{jkk} + \frac{\rho_d}{N} \|ar{h}_{jk}\|^2)^2 \right)
\frac{\rho_d}{N} \text{tr}(\bar{R}_{jkk}) + \frac{\rho_d}{N} \|ar{h}_{jk}\|^2 + \sum_{\ell \neq j} \left( \frac{\rho_d}{N} \text{tr}(R_{j\ell k} \Phi_j R_{jkk})^2 \right) \text{induced by pilot contamination}
$$

Expression (4.36) separates the pilot contamination induced interference from the remaining inter-cell interference, which we refer to as “uncorrelated” interference. Clearly, inter-cell interference limits the overall performance even at the infinite-antenna limit. Nevertheless, note that its impact can be alleviated through the mitigation of the uncorrelated interference by observing that this latter is eliminated when $Q_j$ becomes diagonal. In fact, this is achieved in favorable propagation conditions, wherein the SE will attain:

Corollary 7 (Favorable propagation in multi-cell). If $\frac{1}{N} \bar{h}_{ji} \bar{h}_{jk} \overset{a.s.}{\longrightarrow} 0$, for $i \neq k$, we have:

$$
\bar{SE}_{jk}^{\text{conv},M} = \left( 1 - \frac{\tau}{T} \right) \log \left( 1 + \frac{\rho_d}{N} \text{tr}(\bar{R}_{jkk} + \frac{\rho_d}{N} \|ar{h}_{jk}\|^2)^2 \right)
\frac{\rho_d}{N} \text{tr}(\bar{R}_{jkk}) + \frac{\rho_d}{N} \|ar{h}_{jk}\|^2 + \sum_{\ell \neq j} \left( \frac{\rho_d}{N} \text{tr}(R_{j\ell k} \Phi_j R_{jkk})^2 \right) \text{induced by pilot contamination}
$$

Another important outcome from Theorem 8 and corollary 7 lies in the interplay between the interference emanating from pilot-contamination and the LoS signals. In fact, consider the quantity $\frac{1}{N} \text{tr}(\sum_{\ell \neq j} R_{j\ell k} \Phi_j R_{jkk})$ in (4.36) and (4.37) which represents this type of correlated interference. As shown in the following proof, this
term is a decreasing function of the Rician factor $\kappa_{jk}$. Actually, in the limiting case $\kappa_{jk} \to \infty$, we have $\frac{1}{N} \text{tr}(\sum_{\ell \neq j}^{L} R_{j\ell k} \Phi_{j\ell k} R_{j\ell k}) \to 0$. Consequently, we can state that in such multi-cell systems, another advantage of having stronger LoS components is to reduce the adverse effects of pilot contamination, also known as the main limiting performance factor of massive MIMO systems [2,11].

Proof. The proof is given in Appendix 13.

4.3.2 Statistical Combining in Multi-Cell Systems

In this section, we focus on the performance of the statistical receiver defined in the single-cell scenario to see whether it is also beneficial in a multi-cell system with LoS-prevailing transmissions. In other words, we are interested in examining the resilience of this combining scheme when it is subject to inter-cell interference. In this line, let $\mathbf{g}_{jk}$ indicate the statistical combining vector associated with the communication between UE $k$ and its BS $j$, defined as:

$$
\mathbf{g}_{jk} = \left( \sum_{i=1}^{K} R_{jji} + \mathbf{H}_{j,\ell}^H \mathbf{H}_{j,\ell} + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \mathbf{h}_{jk}, \quad (4.38)
$$

where $\mathbf{H}_{j,\ell}$ is matrix $\mathbf{H}_{j}$ without the $k$–th column. Furthermore, since utilizing this receiver allows to circumvent training and estimation, the corresponding UL SE attains:

$$
\text{SE}_{\text{stat}}^{\text{statM}} = \mathbb{E} \left[ \log \left( 1 + \frac{|\mathbf{g}_{jk}^H \mathbf{h}_{jk}|^2}{\mathbb{E} \left[ \mathbf{g}_{jk}^H \left( \sum_{\ell=1}^{K} \sum_{i=1}^{K} \mathbf{h}_{j\ell i}^H \mathbf{h}_{j\ell i} - \mathbf{h}_{jk}^H \mathbf{h}_{jk} + \frac{N}{\rho_d} \mathbf{I}_N \right) \mathbf{g}_{jk}^H \right]} \right) \right]. \quad (4.39)
$$

Next, we provide an asymptotic approximation of the achievable SE generated by $\mathbf{g}_{jk}$. This constitutes the last main result of this chapter.

Theorem 9 (Statistical combining in multi-cell systems). Under Assumption [6].
\[ \text{SE}_{\text{stat,M}}^{jk} - \text{SE}_{\text{stat,S}}^{jk} \xrightarrow{N \rightarrow \infty} 0, \text{ with} \]

\[ \text{SE}_{\text{stat,M}}^{jk} = \log \left[ 1 + \mathbf{h}_{jk}^H \left( \mathbf{H}_{j,k} \mathbf{H}_{j,k}^H + \frac{N}{\rho_d} \mathbf{I}_N \right)^{-1} \mathbf{h}_{jk} \right], \quad (4.40) \]

**Proof.** Under assumption 6 since \( \mathbf{g}_{jk} \) is deterministic, a direct application of the convergence of quadratic forms lemma (Chapter 2, lemma 3) and the continuous mapping Theorem (Chapter 2, theorem 1) yields \( \text{SE}_{\text{stat,M}}^{jk} \).

First of all, we emphasize once again that the receiver \( \mathbf{g}_{jk} \) is purposely designed for environments with strong LoS components. Second, in such environments, comparing the above multi-cell spectral efficiency \( \text{SE}_{\text{stat,M}}^{jk} \) with the single-cell one \( \text{SE}_{\text{stat,S}}^{jk} \), reveals that employing the statistical combining scheme entails a similar asymptotic performance gain for both network settings, i.e., \( \text{SE}_{\text{stat,M}}^{jk} - \text{SE}_{\text{stat,S}}^{jk} \xrightarrow{N \rightarrow \infty} 0 \).

Therefore, we can conclude that, for LoS-prevailing communications, inter-cell interference can be mitigated through the use of \( \mathbf{g}_{jk} \). In other words, interference emanating from other cells is no longer a limitation of the achievable capacity when this processing approach is employed. This outcome is explained by the fact that the premise behind the statistical receiver is to “bypass” training and thereby, prevent pilot-contamination and its ensuing unfavorable effects. It is also important to note that this desirable feature comes in contrast to conventional combining, whose UL SE remains limited by pilot-contamination-induced interference, as previously shown in expressions (4.36)-(4.37), i.e., \( \text{SE}_{\text{conv,M}}^{jk} \leq \text{SE}_{\text{conv,S}}^{jk} \).

### 4.3.3 Numerical Results for the Multi-cell scenario

For the multi-cell scenario, we consider \( L = 3 \) adjacent cells having the same parameters as defined in the single-cell section. Besides, for each cell, we consider cell-edge users

---

\(^4\)Recall that the superscripts \((\cdot)^{\text{conv,S}}, (\cdot)^{\text{stat,S}}, (\cdot)^{\text{conv,M}} \) and \((\cdot)^{\text{stat,M}} \) are used to distinguish the relevant quantities corresponding to conventional and statistical combining, in single-cell and multi-cell schemes, respectively.
as shown in Fig. 4.5. Deploying the users in such a configuration generates high levels of inter-cell interference, and the close angles of arrival ensure considerable intra-LoS interference.

![Multi-cell network setup with L=3 cells and K=20 cell-edge users.](image)

Figure 4.5: Multi-cell network setup with L=3 cells and K=20 cell-edge users.

We illustrate in Fig. 4.6 the achievable SEs for both statistical and conventional combining schemes considering pilot contamination and correlated Rician fading. Fig. 4.6a and Fig. 4.6b account for ordinary and favorable propagation conditions, respectively. In accordance with the discussion in sections 4.3.1 and 4.3.2, these figures confirm that the multi-cell UL spectral efficiencies follow the same pattern as those observed for a single-cell system. That is, firstly, higher Rician factors entail increasingly better performances. Secondly, favorable propagation conditions further enhance the SE, due to the cancellation of the uncorrelated inter-cell interference for the conventional combining, and the intra-LoS interference for both receivers, as analytically demonstrated in Theorems 8 and 9.

Next, to highlight the multi-cell effect on the receivers, we consider cell 1 from the network setup in Fig. 4.5 as a cell of interest, and propose to compare its achievable UL SE in the cases where it is deployed:

1. in the multi-cell setting of Fig. 4.5
2. in a single-cell setting having the exact same system parameters as in (1).
Figure 4.6: Multi-Cell setting: UL SE using multi-cell conventional combining (4.32) with $\tau = K$ and statistical processing (4.38) with different levels of Rician factor, in (a) ordinary and (b) favorable propagation conditions.

Accordingly, we represent in Fig.4.7 the SEs corresponding to these cases for different levels of $\kappa_{\text{max}}$. Dashed and solid lines (and arrows) correspond to cases (1) and (2), respectively. As indicated for $\kappa_{\text{max}} = 3$, in the multi-cell setting, the statistical combining (dashed blue curve) achieves a 33% SE gain over the conventional one (dashed black curve); whereas in the single-cell case, the observed increase is by 8% only (solid blue vs solid black curves). As for smaller Rician factors (i.e., $\kappa_{\text{max}} = 1$), we can see that in the multi-cell plot, the statistical receiver outperforms the conventional one for a lower SNR (starting 17 dB) than it is the case for the single-cell scheme, wherein this is only achieved for SNRs above 25dB.

Consequently, Fig.4.7 validates that, in the infinite antenna-limit, compared to conventional combining, employing the statistical receiver works even better in a multi-cell network. As demonstrated in Section 4.3.2, this is explained by the fact that since it mitigates the inter-cell interference, this processing technique actually engenders a similar multi-cell spectral efficiency ($\text{SE}_{\text{stat,M}}$) to when it is used in a single-cell system ($\text{SE}_{\text{stat,S}}^k$). To summarize, Fig.4.7 asserts that for the same cell $j$: 
Figure 4.7: UL SEs of conventional and statistical combining in both single-cell ($L = 1$) and multi-cell ($L = 3$) settings, for different levels of Rician factor, and $K = 20$ cell-edge users. Solid and dashed lines (and arrows) correspond to the single-cell and multi-cell cases, respectively. Dotted lines represent the asymptotic approximations given in Theorems 5 - 9.

\[ \kappa_{\text{max}} = 3 \]

\[ \kappa_{\text{max}} = 1 \]

\[ SE_{\text{stat},S} > SE_{\text{conv},S} \]

\[ SE_{\text{stat},M} > SE_{\text{conv},M} \]

\[ SE_{\text{stat},S} - SE_{\text{stat},M} \xrightarrow{a.s.} 0 \] as $N \to \infty$; whereas, due to pilot contamination, \( SE_{k}^{\text{conv,S}} \geq SE_{k}^{\text{conv,M}} \).

Nonetheless, note that the gap observed in Fig.4.7 between \( SE_{k}^{\text{stat,S}} \) and \( SE_{k}^{\text{stat,S}} \) is due to the finite system dimension considered in the simulations which is reduced as $N$ grows larger.

### 4.4 Chapter summary

We studied in this chapter the UL performance of single and multi-cell massive MIMO systems underlying spatially correlated Rician channels. Considering the large-antenna limit, we derived closed-form approximations of the spectral efficiency achieved by the LMMSE conventional receiver and proposed a novel statistical combining scheme. These expressions are shown to be tight, and thus, can be applied for realistic scenarios involving different correlation matrix models, Rician factors, CSI errors.

For the conventional combining, the approximations were exploited to determine an explicit expression of the optimal number of training symbols which was proved to be
particularly crucial for small Rician factors. Nonetheless, it was also found that in LoS-prevailing environments, investing in longer training sequences to enhance the SE is ineffective. Furthermore, the multi-cell analysis unveiled that conventional processing is limited by pilot contamination, even under favorable propagation. However, it also demonstrated that stronger LoS signals reduce this correlated interference. On another note, the derivations indicated that statistical combining allows to mitigate inter-cell interference, and as such, outperforms the conventional receiver in a multi-cell system to an even greater extent.

In conclusion, the conventional receiver is limited by pilot contamination, and the statistical combiner is efficient for high LoS levels. Accordingly, in the next chapter we look into the interplay between the LoS and pilot contamination, and primarily investigate combining schemes that combat pilot contamination for any LoS propagation conditions.
Chapter 5

Pilot Contamination and Line of Sight Interplay on Multi-cell Massive MIMO Systems

In this chapter, we elaborate on the interplay between pilot contamination and LoS propagation environments in the context of massive MIMO. In this line, we first look into the different issues that arise when the common receivers such as MRC and MMSE are utilized. Considering UL transmissions of a massive MIMO system with Rician fading and pilot contamination effects, we derive an approximation of the spectral efficiency achieved by each receiver. Using these expressions, we propose some solutions to enhance the overall performance of both processing techniques in ordinary and favorable propagation. However, the study shows that both combining schemes generate a spectral efficiency that is always hindered by correlated interference emanating from pilot contamination, in all LoS propagation environments.

Accordingly, we shift our focus in the second part to the multi-cell MMSE processing technique as a solution to overcome the detrimental effects of pilot contamination. The denomination “multi-cell processing” is used because such a receiver uses not only estimates of the cell’s local channels, but also the channels from other cells. For more clarity, in the sequel, this scheme is referred to as “M-MMSE” to distinguish it from the conventional MMSE receiver that is investigated at the beginning of this chapter which we shall denote “S-MMSE”. Using the same system model as before, we analytically demonstrate that M-MMSE outperforms both S-MMSE and MRC, by generating a spectral efficiency that grows unboundedly large with the number of antennas.
Finally, for a more thorough analysis of M-MMSE, we provide in the last section of the chapter more results on this combining scheme. Indeed, we extend our results of the M-MMSE to the UL of a massive MIMO system underlying inter-cell LoS components.

5.1 System Model

We consider the UL of a TDD multi-cell multi-user system with $L$ cells having each $K$ single-antenna users communicating with an $N$–antennas BS. Let $h_{j\ell k}$ represent the channel linking the $k$–th UE in cell $\ell$ to BS$_j$. Throughout the chapter, we focus on the performance of cell $j$ and its $k$–th user, UE$_{jk}$. Assuming Gaussian codebooks, we denote by $\sqrt{p}x_j \sim \mathcal{CN}(0, p I_K)$ the vector of the transmitted data symbols sent by all the UEs in cell $j$ with the average power $p$. In this line, the received signal at BS$_j$ is given by:

$$y_j = \sqrt{p} \sum_{\ell=1}^{L} \sum_{i=1}^{K} h_{j\ell i}x_{\ell i} + n_j,$$

(5.1)

where $n_j$ represents a zero-mean additive Gaussian noise with variance $\sigma^2 I_N$. We assume that the system operates over a block-flat fading channel with coherence time $T$. In this part, we consider a similar system model to the previous chapter (section 4.3), i.e. correlated Rician fading for intra-cell or local channels and correlated Rayleigh fading for channels from other cells. Specifically, $h_{j\ell k} \in \mathbb{C}^{N \times 1}$ is given by:

$$h_{jjk} = R_{jjk}^{\frac{1}{2}} z_{jjk} + \tilde{h}_{jk},$$

(5.2)

$$h_{j\ell k} = R_{j\ell k}^{\frac{1}{2}} z_{j\ell k}, \quad \ell \neq j,$$

(5.3)

where $R_{j\ell k} = \frac{\beta_{j\ell k}}{1+\kappa_{jk} \delta_{\ell j}} \Theta_{j\ell k}$ modeling the total channel correlation with: $\beta_{j\ell k}$ and $z_{j\ell k} \sim \mathcal{CN}(0, I_N)$ to account for large and small scale fading of the scattered signals, respectively. Plus, for the intra-cell channels, i.e. $\ell = j$, $\tilde{h}_{jk} = \sqrt{\frac{\beta_{jjk} \kappa_{jk}}{1+\kappa_{jk}}} z_{jk}$ is considered to represent the specular (LoS) signals with Rician factor $\kappa_{jk} \geq 0$. Note that, for each
Figure 5.1: UL transmissions in a Multi-cell massive MIMO network with Rician intra-cell channels and scattered inter-cell channels

UE \( k \), we consider a different Rician factor \( \kappa_{jk} \), due to the distinct geographic locations of the UEs, and a different channel correlation matrix \( \Theta_{j\ell k} \), as well. Throughout this chapter, \( \forall (j, \ell, k) \), \( \Theta_{j\ell k} \) is assumed to be slowly varying compared to the channel coherence time and thus is supposed to be perfectly known to the BS.

**Channel Estimation**

During a dedicated uplink training phase of \( \tau \) symbols, the UEs in each cell transmit mutually orthogonal pilot sequences which allow the BSs to compute estimates of the channels. As shall be discussed next, depending on the type of receiver in use, BS\( j \) either solely estimates its local channels \( h_{jjk} \), \( \forall k \), or all the channels in the network linking it to UEs, i.e., BS\( j \) estimates \( h_{j\ell k} \), \( \forall (\ell, k) \). In addition, since it is a multi-cell system, we assume that the same set of pilot sequences is reused in every cell so that the channel estimates are corrupted by pilot contamination from adjacent cells. In fact, the same pilot is assigned to every \( k \)-th UE in each cell, and as such \( \forall (j, k) \), the estimates of the channels \( h_{j1k}, h_{j2k}, \ldots, h_{jLk} \), will be correlated. Accordingly, using
the MMSE estimation, the estimate of $h_{j\ell k}$ is given by [54, 84]:

$$\hat{h}_{jjk} = R_{jjk} \Phi_{jk} \left( \sum_{\ell' = 1}^{L} h_{j\ell'k} + \frac{1}{\sqrt{\tau \rho_{tr}}} n_{tr}^{\ell_k} \right) + \bar{h}_{jk},$$

(5.4)

$$\hat{h}_{j\ell k} = R_{j\ell k} \Phi_{jk} \left( \sum_{\ell' = 1}^{L} h_{j\ell'k} + \frac{1}{\sqrt{\tau \rho_{tr}}} n_{tr}^{\ell_k} \right), \quad \ell \neq j,$$

(5.5)

$$\Phi_{jk} = \left( \sum_{\ell' = 1}^{L} R_{j\ell'k} + \frac{1}{\tau \rho_{tr}} I_N \right)^{-1},$$

with $\rho_{tr} = \frac{p}{\sigma^2}$ and $n_{tr}^{\ell_k}$ are the SNR and noise during training, respectively. Therefore, it can be easily shown that $\hat{h}_{j\ell k} \sim \mathcal{CN} \left( \delta_{j\ell} \bar{h}_{jk}, \tilde{R}_{j\ell k} \right)$, with $\tilde{R}_{j\ell k} = R_{j\ell k} \Phi_{jk} R_{j\ell k}$. Plus, due to MMSE estimation orthogonality, the estimation error $\xi_{j\ell k} = h_{j\ell k} - \hat{h}_{j\ell k}$, follows the distribution $\xi_{j\ell k} \sim \mathcal{CN} \left( 0, R_{j\ell k} - \tilde{R}_{j\ell k} \right)$.

Detection and Uplink Spectral Efficiency

In order to retrieve the signal sent by UE $j_k$, BS $j$ uses a linear receiver $g_{jk} \in \mathbb{C}^{N \times 1}$ to process the received signal $y_j$ by generating $r_{jk} = g_{jk}^H y_j$. Therefore:

$$r_{jk} = \sqrt{p} g_{jk}^H \hat{h}_{jjk} x_{jk} + \sqrt{p} g_{jk}^H \xi_{jk} x_{jk} + \sqrt{p} \sum_{\ell = 1}^{L} \sum_{i=1, i \neq k}^{K} g_{jk}^H h_{j\ell i} x_{\ell i} + \sqrt{p} \sum_{\ell = 1, \ell \neq j}^{L} g_{jk}^H h_{j\ell k} x_{\ell k} + \sqrt{p} \sum_{\ell = 1, \ell \neq j}^{L} g_{jk}^H h_{j\ell k} x_{\ell k} + \sqrt{p} \sum_{\ell = 1, \ell \neq j}^{L} g_{jk}^H h_{j\ell k} x_{\ell k} + g_{jk}^H n_{jk}.$$

(5.6)

Let $\rho_d = \frac{p}{\sigma^2}$, denote the SNR during data transmission. Therefore, the SINR corresponding to the transmission between the UE $j_k$ and its BS, $\gamma_{jk}$, is defined as:

$$\gamma_{jk} = \frac{|g_{jk}^H \hat{h}_{jjk}|^2}{\mathbb{E} \left[ g_{jk}^H \left( \sum_{\ell = 1}^{L} \sum_{i=1, i \neq k}^{K} h_{j\ell i} h_{j\ell i}^H + \xi_{j\ell k} \xi_{j\ell k}^H + \frac{1}{\rho_d} I_N \right) g_{jk} \right]^{-1}}.$$

(5.7)
Finally, the achievable UL spectral efficiency at BS\(_j\) is defined as:

\[
\text{SE}_{j} = \left(1 - \frac{\tau}{T_c}\right) \frac{1}{K} \sum_{k=1}^{K} \mathbb{E} \left[ \log \left(1 + \gamma_{jk} \right) \right].
\]  

(5.8)

Signal Detection

We propose to investigate three designs for \(g_{jk}\). First, we consider the case wherein the BS uses only its legitimate channels to perform the signal detection. In other words, single-cell channel estimation is sufficient to build the detectors, and therefore such combining schemes are referred to as “single-cell processing techniques”. In this line, two linear receivers are of practical interest in massive MIMO systems, namely the matched filter or MRC and the MMSE detector. Note that we shall refer to MMSE detection in single-cell estimation by “S-MMSE”. Hence\(^1\)

\[
g_{jk}^{\text{MRC}} = \hat{h}_{jjk},
\]

(5.9)

\[
g_{jk}^{\text{S-MMSE}} = \left( \sum_{i=1}^{K} \hat{h}_{jji} \hat{h}_{jji}^H + (Z_{j}^{S})^{-1} \right)^{-1} \hat{h}_{jjk}.
\]

(5.10)

with:

\[
(Z_{j}^{S})^{-1} = \frac{1}{\rho_d} I_N + A,
\]

(5.11)

where \(A \in \mathbb{C}^{N \times N}\) is an arbitrary hermitian positive semi-definite design parameter. For instance, it could contain the covariances of estimation errors and inter-cell interference as in \([11]\), \(i.e.,\) we can put: \(A = \sum_{i=1}^{K} (R_{jji} - \tilde{R}_{jji}) + \sum_{\ell \neq j}^{L} \sum_{i=1}^{K} \tilde{R}_{j\ell i} \). We shall elaborate further on the structure of \(A\) in Rician fading, in the subsequent section.

Previous works have shown that MRC and S-MMSE receivers generate asymptotically the same performance over Rayleigh fading channels. Plus, since MRC has a basic structure, which is far less complex than S-MMSE, researchers concluded that

\(^1\)For sake of clarity, in the sequel, the superscripts \((.\text{MRC})\), \((.\text{S-MMSE})\) and \((.\text{M-MMSE})\) will be added to denote the corresponding quantity when using the MRC, S-MMSE and M-MMSE receivers, respectively.
there is no benefit from employing S-MMSE in massive MIMO networks. It is thus of interest to investigate whether this result still holds true in the presence of LoS components.

For the third detection technique, we consider the setting where the BS further exploits the observation received during training in order to estimate the interfering channels from other cells, and ultimately, uses them to design the MMSE detector. The authors in [67] refer to this technique as multi-cell-MMSE (M-MMSE) combining to differentiate from the above mentioned S-MMSE detector (5.10) obtained through single-cell estimation. Using M-MMSE implies that BS_{j} computes all the channel estimates $\hat{h}_{j\ell k}$, $\forall (\ell, k)$. Accordingly, based on the independence between the estimates and their associated estimation errors $\xi_{j\ell i}$, the conditional expectation in (5.7) can be simplified. Recalling that $\xi_{j\ell i} \sim \mathcal{CN}\left(0, R_{j\ell i} - \tilde{R}_{j\ell i}\right)$, the SINR expression (5.7) can be simply rewritten as:

$$\gamma_{M-MMSE}^{j k} = \frac{|g_{j k}^{\mathcal{H}} \hat{h}_{j jk}|^2}{g_{j k}^{\mathcal{H}} \sum_{\ell=1}^{L} \sum_{i=1}^{K} \hat{h}_{j\ell i} \tilde{h}_{j\ell i}^{\mathcal{H}} + (Z_{j}^{\mathcal{I}})^{-1}} g_{j k},$$

(5.12)

with:

$$(Z_{j}^{\mathcal{I}})^{-1} = \frac{1}{\rho_d} I_{N} + \sum_{\ell=1}^{L} \sum_{i=1}^{K} (R_{j\ell i} - \tilde{R}_{j\ell i}).$$

(5.13)

In fact, M-MMSE is formally defined as the optimal receiver that maximizes the UL SINR in multi-cell estimation (5.12), i.e.:

$$g_{j k}^{M-MMSE} = \arg\max_{g_{j k}} \gamma_{j k}^{M-MMSE}. \quad (5.14)$$

Clearly, finding (5.14) amounts to maximizing the Rayleigh-Quotient, whose well-known
solution can be:

\[
\mathbf{g}_j^{\text{MMSE}} = \left( \sum_{\ell=1}^{L} \sum_{i=1}^{K} \hat{h}_{j\ell i} \hat{h}_{j\ell i} + (\mathbf{Z}_j^M)^{-1} \right)^{-1} \hat{h}_{jjk}. \tag{5.15}
\]

Finally, plugging \( \mathbf{g}_j^{\text{MMSE}} \) (5.15) in the SINR expression \( \gamma_{jk}^{\text{MMSE}} \), thus yields:

\[
\gamma_{jk}^{\text{MMSE}} = \hat{h}_{jjk}^H \left( \sum_{(\ell,i) \neq (j,k)} \hat{h}_{j\ell i} \hat{h}_{j\ell i} + (\mathbf{Z}_j^M)^{-1} \right)^{-1} \hat{h}_{jjk}. \tag{5.16}
\]

5.2 Analysis of Achievable UL Spectral Efficiencies

We carry out a theoretical analysis of the UL spectral efficiency when using MRC, S-MMSE and M-MMSE detectors under the assumption of correlated Rician channels with imperfect CSI. To this end, we first derive closed-form expressions for the achievable SEs in the large-antenna limit. The obtained approximations are tractable, and as such, enable to apprehend the trends of the UL performance with respect to the different system parameters such as the Rician factors or the correlation matrices. Additionally, they allow to identify the factors that limit the performance in terms of spectral efficiency. Note that our results are tight and apply for systems with different structures of channel correlation matrices, as shall be illustrated with simulations. The main results of this analysis are stated in three theorems treating each a different receiver. In every theorem, we provide approximations of the SINR \( \gamma_{jk} \) which, according to the continuous mapping theorem (chapter 2 theorem 1), yields an approximation of the achievable capacity.

As \( N \to \infty \), while \( K \) is maintained fixed, we consider the following assumptions:

**Assumption 7.** \( \forall j, \ell, k, \lim \sup_N \| \mathbf{R}_{j\ell k} \|_2 < \infty, \text{ and } \lim \inf_N \frac{1}{N} \text{tr}(\mathbf{R}_{j\ell k}) > 0. \)

In the sequel, we shall refer to this regime as \( N \to \infty \). Let \( \mathbf{H}_j = [\mathbf{h}_j1, \ldots, \mathbf{h}_jK] \) denote the aggregate matrix of the LoS components between BS \( j \) and its UEs. Accordingly,
consider:

**Assumption 8.** \( \forall j, \limsup_N \frac{1}{\sqrt{N}} \| \mathbf{H}_j \|_2 < \infty, \liminf_N \frac{1}{N} \text{tr} (\mathbf{H}_j) > 0. \)

### 5.2.1 MRC

**Theorem 10.** Under Assumptions 7 and 8,

\[
\gamma_{jk}^{\text{MRC}} - \gamma_{jk}^{\text{MRC}} \xrightarrow{\text{a.s.}} 0, \quad N \to \infty, \text{with:}
\]

\[
\begin{align*}
\gamma_{jk}^{\text{MRC}} &= \frac{\frac{1}{N} \text{tr} \mathbf{R}_{jjk} + \frac{1}{N} \| \mathbf{H}_{jk} \|_2^2}{\frac{1}{N^2} \sum_{i=1, i \neq k}^K |\mathbf{h}_{jk}^H \mathbf{h}_{ji}|^2 + \sum_{\ell=1, \ell \neq j}^L \left( \frac{1}{N} \text{tr}(\mathbf{R}_{jjk} \Phi_{jk} \mathbf{R}_{j\ell k}) \right)^2}.
\end{align*}
\tag{5.17}
\]

\text{LoS Intra-cell interference} \quad \text{pilot contamination}

**Proof.** The proof is given in Appendix 14.

First, note that if only Rayleigh fading was considered, \textit{i.e.} Assumption 8 is not verified \( s.t. \kappa_{jk} = 0, \forall (j, k), \) the approximation \( \gamma_{jk}^{\text{MRC}} \) coincides with the findings in [2, 11]. Second, these works state that using MRC detection for massive MIMO systems leads to an SE that is only limited by pilot contamination. In this line, Theorem 10 demonstrates that this conclusion also holds in Rician fading, since \( \gamma_{jk}^{\text{MRC}} \) indicates that only pilot contamination remains out of the total interference inflicted by other cells. However, in contrast to [2, 11] and Rayleigh fading, our work shows that pilot contamination is not the sole hindrance of MRC. Indeed, \( \gamma_{jk}^{\text{MRC}} \) reveals that in Rician fading, the system undergoes LoS induced intra-cell interference as well. This latter is characterized by the inner product \( \frac{1}{N} \mathbf{h}_{jk}^H \mathbf{h}_{ji}, i \neq k \) and, thus would only dissipate under asymptotic favorable propagation conditions wherein: \( \frac{1}{N} \mathbf{h}_{jk}^H \mathbf{h}_{ji} \xrightarrow{\text{a.s.}} 0, N \to \infty, \) \( \forall i \neq k. \) Therefore, in such environments, better SE gains are attained:

**Corollary 8 (\( \gamma^{\text{MRC}} \) under favorable propagation).** Under Assumptions 7 and 8.
if \( \frac{1}{N} \overline{H}_{jk} \overline{H}_{ji} \xrightarrow{a.s.} 0, \forall i \neq k \):

\[
\gamma_{MRC,jk} = \frac{\left( \frac{1}{N} \text{tr} \overline{R}_{jjk} + \frac{1}{N} \| \overline{H}_{jk} \|^2 \right)^2}{\sum_{\ell=1}^{L} \left( \frac{1}{N} \text{tr} (R_{jjk} \Phi_{jk} R_{j\ell k}) \right)^2}.
\]  

(5.18)

### 5.2.2 S-MMSE

**Theorem 11 (S-MMSE).** Under Assumptions 7 and 8, we have:

\[
\gamma_{S-MMSE,jk} \xrightarrow{a.s.} 0, \quad \forall i \neq k
\]

such that:

\[
\gamma_{S-MMSE,jk} = \frac{1}{L} \sum_{\ell=1}^{L} \left| Q_{j\ell k} \beta_{jk,\ell j}^S \right|^2 + \sum_{\ell=1}^{L} \sum_{i=1}^{K} \left| Q_{j\ell i} \beta_{ji,\ell j}^S \right|^2
\]

\[\text{induced by pilot contamination} \quad \text{uncorrelated inter-cell interference}\]

(5.19)

with: \( \beta_{jk,\ell j}^S = \frac{1}{N} \text{tr} (R_{j\ell k} \Phi_{jk} R_{j\ell k} Z_j^S) \), and

\[
Q_j = \left( \frac{1}{N} \overline{H}_j^H \overline{H}_j + \text{diag} \{ \beta_{ji,\ell j}^S \}_{i=1}^{K} \right)^{-1}.
\]

(5.20)

**Proof.** The proof is given in Appendix 14

Let us examine the expression \( \gamma_{S-MMSE,jk} \) (5.19). In contrast to MRC, S-MMSE eliminates intra-cell interference completely (both LoS and scattered induced interference), and this for any propagation conditions. Furthermore, as can be seen by the denominator of (5.19), the entirety of inter-cell interference, including pilot contamination, remain, hence impeding S-MMSE from achieving higher UL SE gains. These outcomes can be explained by the structure of the S-MMSE detector \( \tilde{g}_{S-MMSE} \) (5.10) that includes all local channels \( \widehat{H}_{iji} \forall i \), and thus, cancels intra-cell interference; yet, does not mitigate inter-cell interference.
Nevertheless, note that we can alleviate the effects of this latter by mitigating the uncorrelated inter-cell interference \[^2\] which is depicted by the second term of the denominator. Indeed, examining this term unveils that uncorrelated inter-cell interference vanishes asymptotically when the matrix $Q_j$ (5.20) is diagonal. We propose to achieve this by a proper selection of $Z_j^S$ that designedly eliminates the off-diagonal elements of $Q_j$. Accordingly, we suggest two structures of $Z_j^S$ based on the propagation conditions. First, in favorable propagation conditions, it is sufficient to put:

$$Z_j^S = \rho_d I_N.$$  \hspace{1cm} (5.21)

Note that in such a setting, S-MMSE and MRC deliver the same SE performance given in corollary \[^8\]. Second, in non-favorable propagation, \(i.e.\ \lim \inf_N \frac{1}{N} \mathbf{h}_j^H \mathbf{h}_k > 0, \forall i \neq k\), one solution is:

$$Z_j^S = \mathbf{H}_j (\mathbf{H}_j^H \mathbf{H}_j)^{-1} D_j (\mathbf{H}_j^H \mathbf{H}_j)^{-1} \mathbf{H}_j^H,$$  \hspace{1cm} (5.22)

where $D_j$ is an arbitrary diagonal matrix that can be used for further optimization of the SE. Choosing $Z_j^S$ as such renders $Q_j$ diagonal, and hence, eliminates the uncorrelated inter-cell interference when using S-MMSE detection.

From Theorems \[^10\] and \[^11\] we deduce that for MRC and S-MMSE receivers, pilot contamination remains a tenacious limitation for massive MIMO systems with Rician fading channels, even under favorable propagation conditions. Additionally, although LoS components clearly enhance the received signal, they also cause interference. Normally, this latter lessens in favorable propagation conditions; however such environments are not constantly available. Therefore, more research should be done to fully leverage the LoS component so it becomes an enabling performance factor under all

\[^2\]Uncorrelated inter-cell interference refers to the total inter-cell interference minus the pilot contamination induced interference.
propagation scenarios. For instance, as has been shown to be in the previous chapter, using statistical processing schemes that exploit the LoS strength is more profitable in LoS-prevailing environments than common receivers.

### 5.2.3 M-MMSE

Prior to deriving a deterministic equivalent for $\gamma_{jk}^{M-MMSE}$, we consider the following assumption [67, Assumption 5] (see Chapter 2, definition 3):

**Assumption 9** (Asymptotic linear independence). $\forall j, \ell, k$, with $\lambda_{jk} = [\lambda_{j\ell 1}, \ldots, \lambda_{j\ell L}]^T$ and $\ell' = 1, \ldots, L$:

\[
\lim_{N \to \infty} \inf_{\{\lambda_{jk} : \lambda_{j\ell' k} = 1\}} \frac{1}{N} \left\| \sum_{\ell=1}^{L} \lambda_{j\ell k} R_{j\ell k} \right\|_F > 0.
\] (5.23)

Plus, define the quantities $T_{jk} \in \mathbb{R}^{L \times L}$ and block matrix $D_{jk} \in \mathbb{R}^{L(K-1) \times L(K-1)}$ with entries:

\[
[T_{jk}]_{uv} = \frac{1}{N} \text{tr} \left( R_{juk} \Phi_{jk} R_{jvk} Z_j^M \right)
\] (5.24)

\[
B(D_{jk}, u, v) = \text{diag} \left\{ \frac{1}{N} \text{tr} \left( R_{jui} \Phi_{ji} R_{jvi} Z_j^M \right) \right\}_{i=1}^{K \setminus \{i \neq k\}}.
\] (5.25)

**Theorem 12** (M-MMSE with intra and inter-LoS channels (5.5)). Under assumptions \tc{3.3} we have:

\[
\frac{1}{N} \gamma_{jk}^{M-MMSE} - \frac{1}{N} \tau_{jk}^{M-MMSE} \xrightarrow{a.s., N \to \infty} 0,
\] such that:

\[
\frac{\tau_{jk}^{M-MMSE}}{N} = \frac{1}{[T_{jk}]_{jj}} + \frac{N}{H_{jk}^H Q_{j,lk} H_{jk}},
\] (5.26)

with $Q_{j,lk} \in \mathbb{C}^{N \times N}$ defined as:

\[
Q_{j,lk} = \left( \frac{1}{N} H_{j,lk} D_{jk}^{-1} H_{j,lk}^H + (Z_j^M)^{-1} \right)^{-1},
\] (5.27)
and the $N \times L(K-1)$ LoS matrix:

$$
\mathbf{H}_{j,k} = [0_{(j-1)(K-1)}, \overline{h}_{j,1}, \ldots, \overline{h}_{j,k-1}, \overline{h}_{j,k+1}, \ldots, \overline{h}_{j,K}, 0_{(L-j)(K-1)}].
$$

Proof. The proof is given in Appendix 15.

First, Assumption 9 implies that for every UE $k$ in cell $\ell$, the correlation matrices $\mathbf{R}_{j\ell k}$, $\ell = 1, \ldots, L$, are asymptotically linearly independent, as explained in Chapter 2, definition 2. If this holds, as $N \to \infty$, the invertibility of $\mathbf{T}_{jk}$ is verified, as shown in [67, Appendix G]. Plus, if we consider Rayleigh fading, i.e., $\forall (j,k)$, $\kappa_{j,k} = 0$, the SINR approximation in Theorem 1 remains valid, and in fact, it can be easily shown that under such setting, (5.26) will coincide with the result given in [67]. Second, note that the provided expression in Theorem 12 is an approximation of the ‘$\frac{1}{N}$–scaled’ SINR, $\frac{1}{N}\gamma_{\text{M-MMSE}}$. In this line, Theorem 12 reveals that, under Assumptions 7–9, using M-MMSE gives rise to a SINR (and by the same token a capacity) that grows unboundedly as $N \to \infty$. This outcome was demonstrated in [67] for the correlated Rayleigh fading, and is validated here by Theorem 12 for correlated Rician fading systems. Moreover, the contribution of the specular signals is epitomized by the term $\frac{1}{N} \mathbf{h}^H_{jk} \mathbf{Q}_{j,k} \mathbf{R}_{jk}$, and as shall be seen with simulation results, stronger LoS components improve the SE generated by M-MMSE.

Accordingly, M-MMSE outperforms all the single-cell detectors including MRC and S-MMSE, whose performance remains limited by pilot contamination. This is due to the structure of $\mathbf{g}_{jk}^{\text{M-MMSE}}$ that involves estimates of all intra and inter-cell interfering channels $\mathbf{h}_{\ell i}$, $\forall (\ell, i)$, which mitigates both intra and inter-cell interference, and yields unbounded spectral efficiency.

### 5.3 Numerical Results

In this section, we carry out MonteCarlo simulations over 1000 channel realizations to validate, for finite system dimensions, the asymptotic results provided in Section

---

4Using the block matrix inversion lemma and by taking the limit as $\kappa_{j,k} \to 0$, $\forall (j, k)$
To this end, we consider a multi-cell massive MIMO with $L = 4$ cells and inner cell-radius of 150m. Each cell has $K = 2$ cell-edge users with similar distance and angle of arrival $\theta$, to ensure high levels of pilot contamination, as depicted in Fig.5.2. We fix $T = 200$ symbols and $\tau = K$. Furthermore, the SNR is the same during training and data transmission (i.e. recalling that $\beta_{j\ell k}$ is the large scale fading of $h_{j\ell k}$, $\beta_{j\ell k}\rho_d = \beta_{j\ell k}\rho_{tr}$). It is fixed at $-6$dB for intra-cell UEs and between $-6.3$dB and $-11.5$dB for the interfering channels from other cells. Additionally, for intra-cell channels, the specular component $z_{jk}$ follows the model $[z_{jk}]_n = e^{-j(n-1)\pi\sin(\theta_{jjk})}$.

Finally, the results are represented in terms of UL spectral efficiency (5.8) in two scenarios.

- **Scenario I**: In the first setting, we illustrate the performances as a function of the number of antennas, with different values of the Rician factors $\kappa_{jk}$, and considering the exponential correlation model [88], such that the elements of the correlation matrix $\Theta_{j\ell k}$ of channel $h_{j\ell k}$ are given by:

$$[\Theta_{j\ell k}]_{mn} = r^{|m-n|}e^{i(m-n)\theta_{j\ell k}}, \quad (5.28)$$

where $r = 0.5$. Accordingly, Fig.5.3 displays the corresponding UL achievable rate obtained by MRC, S-MMSE and M-MMSE. Solid and dashed lines depict empirical and asymptotic results, respectively.
As can be seen in Fig.5.3 for the three receivers, a better ergodic capacity is obtained in the presence of the LoS (curves with $0 < \kappa_{jk} \leq 2$), relatively to the case with only scattered signals ($\kappa_{jk} = 0$). Plus, S-MMSE provides increasingly better performances than MRC as $\kappa_{jk}$ takes higher values. However, both reach a plateau as $N$ grows large, which is due to pilot contamination, as demonstrated in Theorems 10 and 11. On the other hand, M-MMSE outperforms both single-cell combining schemes, MRC and S-MMSE, and clearly scales linearly with the number of antennas, thus confirming the results of Theorem 12.

• **Scenario II:** We now move on to illustrating the performance of the system with the uncorrelated fading model with independent log-normal large-scale variations over the array, such that:

$$\Theta_{jk} = diag \left\{ 10^{f_i/10} \right\}_{i=1}^{N}, \text{ where } f_i \sim \mathcal{N}(0, \sigma_c^2).$$

(5.29)
To this end, for \( N = 200 \) antennas and different values of \( \kappa_{jk} \), we represent in Fig. 5.4 the achievable UL SE with respect to the standard deviation of the fading variations over the array, \( \sigma_c \). As can be observed, S-MMSE and MRC generate comparable rates for all values of \( \kappa_{jk} \). Furthermore, M-MMSE yields a significantly larger capacity than with MRC and S-MMSE, except for the special case at \( (\sigma_c = 0) \) corresponding to Rayleigh fading \( (\kappa_{jk} = 0) \), wherein the correlation matrices become linearly dependent, as previously observed in [67]. Nevertheless, note that in this particular setting \( (\sigma_c = 0) \), for Rician fading (curves with \( \kappa_{jk} \in [0, 2] \)), M-MMSE still outperforms both single-cell combining schemes. Finally, Fig. 5.3 and Fig. 5.4 clearly validate the accuracy of the asymptotic approximations provided in Section 5.2 (dashed-lines curves quasi-overlapping with the solid ones).
5.4 More Results on LoS effects and the M-MMSE scheme

We analyzed in the previous sections the effects of pilot contamination and LoS signals on the performance achieved by MRC, S-MMSE and M-MMSE. As seen, both analytical and numerical results assert that the presence of specular components is beneficial in terms of spectral efficiency, for all three receivers. However, the Rician channel model that we examined thus far accounts for “intra-LoS” links only. As explained in section 5.1, such model is a reasonable setting in ‘split-multi-cell’ systems wherein the long distances between UEs and the BSs in other cells lower the possibility of inter-cell LoS links. Nevertheless, the ubiquity of massive MIMO in future networks makes its analysis in all propagation scenarios of utmost relevance. Accordingly, in this section, we consider a more general channel model that encompasses inter-cell specular links as well. Note that this setting is useful for the investigation of networks where the probability of both intra-cell and inter-cell LoS signals is prevalent like in heterogeneous networks, emerging UAV aided wireless systems, to name a few. Additionally, in light of the promising potential of M-MMSE compared to common single-cell processing techniques, especially in Rician fading, we propose to mainly focus on its achievable SE in such systems. Towards this aim, we consider the same system model as in section 5.1 except that now, all the channels in the system including the inter-cell links include specular components. That is, \( \forall (j, \ell, k) \):

\[
\begin{align*}
    h_{j\ell k} &= R_{j\ell k}^{\frac{3}{2}} z_{j\ell k} + \bar{h}_{j\ell k}, \\
    \hat{h}_{j\ell k} &= R_{j\ell k} \Phi_{jk} \left( \sum_{\ell' = 1}^{L} h_{j\ell' k} + \frac{1}{\sqrt{\rho_{tr} n_{tr}^{jk}}} n_{tr}^{jk} \right) + \bar{h}_{j\ell k},
\end{align*}
\]

and its MMSE estimate:

\[
\hat{h}_{j\ell k} = R_{j\ell k} \Phi_{jk} \left( \sum_{\ell' = 1}^{L} h_{j\ell' k} + \frac{1}{\sqrt{\rho_{tr} n_{tr}^{jk}}} n_{tr}^{jk} \right) + \bar{h}_{j\ell k},
\]

Plus, we consider the same system parameters as in (5.15)-(5.16), i.e. with:
\[ s_{jk}^{M-MMSE} = \left( \sum_{\ell=1}^{L} \sum_{i=1}^{K} \hat{h}_{j\ell i} \hat{h}_{j\ell i}^{H} + (Z_{j}^{M})^{-1} \right)^{-1} \hat{h}_{jjk}, \]  

(5.32)

with \((Z_{j}^{M})^{-1} = \frac{1}{\rho_{d}} I_{N} + \sum_{\ell=1}^{L} \sum_{i=1}^{K} (R_{j\ell i} - \tilde{R}_{j\ell i})\), and

\[ \gamma_{jk}^{M-MMSE} = \hat{h}_{jjk}^{H} \left( \sum_{(\ell,i)\neq(j,k)} \hat{h}_{j\ell i} \hat{h}_{j\ell i}^{H} + (Z_{j}^{M})^{-1} \right)^{-1} \hat{h}_{jjk}. \]  

(5.33)

### 5.4.1 Asymptotic Analysis

Similarly to the previous studies, we derive in this section closed-form deterministic equivalents of the SINR \( \gamma_{jk}^{M-MMSE} \) which, according to the continuous mapping theorem (CMT) \([81]\), leads to an approximation of the UL SE. These approximations, as shall be asserted by simulations, are tight for finite system dimensions, and thus provide an analytical framework that is valid for different channel correlation structures and Rician factors.

**Theorem 13** (M-MMSE with intra and inter-LoS channels (5.30)). Under Assumptions \( \mathbb{B}\mathbb{R} \), we have:

\[ \frac{1}{N_{j}} \gamma_{jk}^{M-MMSE} - \frac{1}{N_{j}} \gamma_{jk}^{M-MMSE} \xrightarrow{a.s.} 0, \text{ such that:} \]

\[ \frac{1}{N_{j}} \gamma_{jk}^{M-MMSE} \frac{1}{[\left( T_{jk} + \frac{1}{N_{j}} \overline{H}_{j,k}^{H} \overline{P}_{jk} \overline{H}_{j,k} \right)^{-1}]_{jj}} \]  

(5.34)

where, \( T_{jk} \) is given in (5.24), and \( \overline{P}_{jk} \) is a \( N_{j} \times N_{j} \) matrix defined as:

\[ \overline{P}_{jk} = \left( \frac{1}{N_{j}} H_{j,k}^{-1} D_{jk}^{-1} H_{j,k}^{H} + (Z_{j}^{M})^{-1} \right)^{-1}, \]  

(5.35)

with \( D_{jk} \) (5.25). Finally, the matrix \( \overline{H}_{j,k} \in \mathbb{C}^{N_{j} \times L} \) collects the LoS components between
$BS_j$ and the $k$–th UE of each cell in the system, i.e. $H_{j,k} = [h_{j1k}, \ldots, h_{jLk}]$; and $H_{j,\backslash{k}} = [h_{j\ell i} : \forall(\ell, i), i \neq k]$ (the total specular matrix of $BS_j$ aside from the $k$–th UEs).

**Proof.** A proof is given in Appendix 16.

As can be seen, the achievable SINR (5.34) is fairly analogous to (5.26) attained when only intra-cell LoS components exist. In this line, Theorem 13 shows that, under Assumptions 6-9, M-MMSE still gives rise to a SE that grows unboundedly with $N$, even in the context inter-LoS interference. In contrast, considering this same channel model, the authors in [70] indicate that MRC generates bounded spectral efficiency due to pilot contamination. In fact, they note that MRC can only combat pilot contamination if two conditions hold: asymptotic favorable propagation and “asymptotic spatial orthogonality” between the correlation matrix of the UE of interest $R_{jkk}$, and all the other correlation matrices related to $BS_j$. Evidently, favorable propagation conditions are not always available, especially in the context of heterogeneous networks. Most importantly, it is worth mentioning that asymptotic spatial orthogonality is also unlikely to occur in practice [24] and is a much stronger condition than the asymptotic linear independence (assumption 9) considered in this work.

In conclusion, by generating a spectral efficiency that scales with the number of antennas, and this for any level of strength of LoS, M-MMSE outperforms all conventional single-cell detectors whose performance remains limited by pilot contamination in both NLoS and LoS propagation scenarios.

The result in Theorem 13 provides an analytical framework that can be harnessed to examine the various trends of the spectral efficiency with respect to the different elements composing the network. In this line, since we consider Rician channels throughout the entire network, it is relevant to examine the impact of assuming inter-cell channels with both NLoS and LoS components on the overall spectral efficiency.
other words, we are interested in comparing the SEs of the sections achieved by both channel models (5.5) and (5.30). For the sake of simplicity, we propose to investigate this aspect through the following case study, notwithstanding that the outcomes shall be validated by simulations.

Case Study

Consider a systems with two-cells \{1, 2\}, each assigned \( K = 1 \) UE that uses the same pilot during training. Plus, \( \forall(j, \ell, k) \), let the correlation matrices \( \Theta_{j\ell k} = I_{N_j} \), and assume favorable propagation conditions between the channels corresponding the same BS, i.e. for cell 1, \( \frac{1}{N} \bar{h}_{111}^H \bar{h}_{121} \xrightarrow{a.s. N \to \infty} 0 \). Plugging these settings in the SINR approximation (5.26) leads to the following expressions:

\[
\frac{\gamma_{11}^{M-MMSE}}{N} = \frac{\alpha_{11} (\alpha_{11} + \kappa_{111} \phi_{11} - \frac{\alpha_{11} \alpha_{12}}{\alpha_{12} + \kappa_{121} \phi_{11}})}{\phi_{11} \left( \frac{1}{\rho} + \alpha_{11} + \alpha_{12} - \frac{\alpha_{11} + \alpha_{12}}{\phi_{11}} \right)}
\]

such that \( \alpha_{1v} = \frac{\beta_{11v}}{1 + \kappa_{11v}} \), and \( \phi_{11} = \alpha_{11} + \alpha_{12} + \frac{1}{\tau \rho} \). Therefore, the objective is to investigate the impact of \( \kappa_{121} \) on the UL spectral efficiency. To this end, a basic derivation of (5.36) with respect to \( \kappa_{121} \), unfolds that \( \frac{\gamma_{11}^{M-MMSE}}{N} \) is an increasing function of \( \kappa_{121} \). In other words, this result demonstrates that when the inter-cell channels include a combination of diffuse and specular components is more beneficial than when they are fully scattered. Plus, the performance gain in terms of SE is increasingly higher as the specular signals become prevalent. Consequently, this case study unveils that in favorable propagation conditions, LoS enhances the overall performance, even when it exists in shapes of interfering inter-cell channels.

5.4.2 Numerical Results

We propose in this section a selection of numerical results using the \( L = 4 \) multi-cell setting in Fig 5.5 with \( K = 5 \) cell-edge users, and considering the same system
parameters as in section 5.3 with exponential correlation model \[88\] with correlation factor 0.5. The figures illustrate the spectral efficiency, using the approximations given in Theorem 13 (dotted lines), in addition to MonteCarlo plots carried over a 1000 channel realizations (solid and dashed lines).

![Diagram](image)

Figure 5.5: \( L = 4 \) Multi-cell network setup with \( K = 5 \) UEs each.

Fig.5.6 compares the M-MMSE and MRC respective SEs as functions of the number of antennas. Accordingly, taking into account that all channels have both NLoS and LoS components, the figure confirms that M-MMSE still outperforms single-cell MR combining irrespective of the levels of the Rician factors. Plus, as can be seen, the plots validate the accuracy of the approximations given in Theorem 13 for finite system dimensions and different values of \( \kappa_{\text{max}} \). Since the presence of LoS links enables better SEs for both receivers as clearly depicted by the solid curves in Fig.5.6, we illustrate in the following its impact as “interfering inter-cell” channels. To this end, we consider in Fig.5.7 two scenarios of inter-cell channels, wherein case I considers both NLoS and LoS inter-cell channels \( (i.e. \ \forall (j, \ell, k), \kappa_{j\ell,k} > 0) \); while case II represents fully scattered inter-cell links \( (i.e. \ \forall \ell \neq j, \kappa_{j\ell k} = 0) \). Note that the SE in this figure is plotted with respect to \( \kappa_{\text{max}} \) in order to span all ranges of practical Rician factors, and ultimately, highlight the impact of inter-cell specular links in all (prevailing and non-prevailing) LoS propagation environments. As can be seen, Fig.5.7 shows that a higher SE is
Figure 5.6: UL SE using M-MMSE and MRC with respect to the number of antennas $N$, for different levels of Rician factor $\kappa_{\text{max}} \in \{0, 2\}$. Dotted lines depict the asymptotic approximation given in Theorem 13.

Figure 5.7: UL SE using M-MMSE with respect to $\kappa_{\text{max}}$ for $N = 200$, in both ordinary (Ord.P) and favorable (Fav.P) propagation conditions, where: Case I indicates both NLoS and LoS inter-cell channels, while Case II represents fully scattered inter-cell links.

achieved when the inter-cell channels are composed of both NLoS and LoS signals (star marked lines as compared to the diamond ones), and this in both favorable and ordinary propagation conditions. In fact, up to a 15% SE performance gain is observed between the two scenarios. Accordingly, Fig. 5.7 confirms our conclusion in Section 5.4.1 and the case study, which stated that having both NLoS and LoS inter-cell channels culminates to a better SE than when said links are fully scattered.

5.5 Chapter Summary

We studied in this work the UL performance of multi-cell massive MIMO systems with correlated Rician fading channels, under the assumption of imperfect channel estimates. Considering the large-antenna limit, we derived closed-form approximations of the achievable spectral efficiency for different single-cell and multi-cell combining schemes. For the single-cell detection case, we analyzed the MRC and S-MMSE receivers that were shown to produce higher gains as the LoS signals became stronger; yet remain...
limited by LoS induced interference and pilot contamination as the number of antennas grows large. In contrast, for multi-cell combining, we analytically demonstrated that M-MMSE outperforms any single-cell combining technique. In fact, it provides unlimited capacity that scales linearly with the number of antennas, and is further enhanced in the presence of LoS.
Chapter 6

Concluding Remarks and Future Work

The high-dimensional nature of massive MIMO renders random matrix theory a suitable tool to elucidate its principles. In fact, RMT enables to derive tractable deterministic approximations of high dimensional performance metrics such as the spectral efficiency which can be exploited to gain explicit insights on the impact of different factors on the overall system performance. As such, throughout this dissertation, we harnessed RMT results and other asymptotic tools to investigate and further optimize the performance of massive MIMO systems in different contexts.

First, in chapter 3 we considered the application involving massive MIMO and heterogeneous networks. In these systems, the BS is equipped with a large number of antennas and densely overlaid with small-cells constituting thus a multiple-tier architecture. The small-cells being deployed within the coverage of the macro-cell causes high levels of interference. In this regard, we investigated SLNR based coordinated precoding techniques for interference mitigation in large-scale HetNets. The analysis revealed that there is little interest to perform coordination when the channel is poorly estimated, and in fact, can even lead to performance losses. To remedy this issue, we proposed a modified SLNR precoding design which includes a normalizing factor obtained via the maximization of SLNR respectively with the quality of CSI. Numerical results showed that our precoding design outperforms both non-coordinated and conventional coordinated SLNR based precoding.

Then, we moved to exploring the issues that rise when the LoS component is present in massive MIMO systems. We began the study in Chapter 4 wherein we focused on
the conventional LMMSE receiver which relies on training-based channel estimates. In the analysis we highlighted both positive and negative effects of the LoS on the spectral efficiency generated by the above receiver. Furthermore, we emphasized the importance of the number of training symbols on the overall performance when the scattered signals dominate the environment, and derived an optimal value of this parameter. Nonetheless, we also demonstrated that investing in longer training sequences to improve the SE can be counterproductive as the LoS component becomes stronger. In this line, we proposed a statistical receiver purposely designed for LoS-prevailing environments, that better exploits the presence of specular links and precludes from using training sequences.

Finally, in Chapter 5, we investigated the main limitation of massive MIMO systems, namely, pilot contamination under LoS propagation conditions. We evaluated the performance of such systems with different single-cell and multi-cell detection methods. Overall, the analysis showed that stronger LoS links improve the spectral efficiency generated by all single and multi-cell combining schemes. However, for the single-cell approaches, we demonstrated that MRC is limited by intra-cell LoS induced interference that would only vanish in favorable propagation conditions. Conversely, we showed that single-cell MMSE combining cancels all LoS induced interference, yet suffers from uncorrelated inter-cell interference which we proposed to combat through certain MMSE designs. Most importantly, the analysis asserted that pilot contamination remains a limiting factor for both processing techniques, irrespective of the level of LoS in the system. On the other hand, under certain conditions on the correlation matrices, multi-cell MMSE (M-MMSE) combining was shown to overcome pilot contamination and produce a SE that scales with the number of antennas.
Future Work

The theoretical analysis we have provided throughout the dissertation is general and offers a framework for further analysis of similar systems. For instance, in light of the positive impact of LoS components, more research efforts should be pursued to fully exploit these signals. In the same spirit as in Chapter 4, an interesting direction is to look into beamforming designs that would designedly aim at ensuring “favorable-propagation-like-performance” in all scenarios, especially in LoS-prevailing environments.

While extensive research has been conducted in the road to making massive MIMO a reality, several directions can be envisioned to further avail from the premise of this technology. Most of today’s wireless systems operate at microwave frequencies below 6GHz. As a result, this band is extremely crowded and is turning into a luxurious commodity. However, another trending technology for achieving dramatic improvements in capacity and spectral efficiency is communications over millimeter-wave (mmWave) (in the frequency bands above 30 GHz) [3,12,53], where the spectrum is underutilized and available bandwidths are broader. One of the interesting concepts is to pair this promising technology with massive MIMO in the framework of HetNets. In such systems, the idea is to operate over mmWaves in small areas relying on the wide BW to achieve “Gbps” rates, together with massive MIMO that exploits efficient spatial multiplexing to extend coverage and overcome the high mmWave RF pathloss. In fact, a serious pursuit arose to study these systems and their feasibility. While energy and spectral efficiency analysis of networks employing mmWave massive MIMO are ongoing, performance in real-life deployments still remains an open issue. Additionally, more channel measurement campaigns are still needed to enable the development of more accurate representative channel models, and thus aid in the process of channel estimation and precoding design algorithms for mmWave massive MIMO systems.

Another trending research topic for massive MIMO technology is “cell-free massive
MIMO". Unlike conventional cellular networks, in cell-free massive MIMO, the network is not partitioned into cells and the users are not assigned to a particular base station. Instead, the coverage area is overlaid with a massive number of single-antenna access points (APs) to simultaneously serve a smaller number of users \([89,90]\). Accordingly, this distributed feature provides additional macro-diversity and the co-processing eliminates entirely the inter-cell interference \([91]\). Additionally, the APs are connected via fronthaul connections to central processing units (CPUs) interconnected by backhaul links, which are responsible for the coordination. This cell-free approach, combined with the inherent system scalability of the massive MIMO technology, constitutes a paradigm shift compared to the conventional centralized and distributed wireless communication systems \([91]\). By any standards, every technology comes at a cost, which opens up many interesting research directions. For cell-free massive MIMO, three major issues need to be investigated and solved to enable practical deployment, namely, the cost and complexity of deployment, limited capacity of back and front-haul connections, and network synchronization \([91,92]\).

Finally, there is an emerging penchant for applying machine learning techniques in communication systems; massive MIMO is one of the potential technologies to avail of such methods \([92]\). Some researchers are already looking into this direction, like the authors in \([93]\) who propose to exploit UL CSI data sets to predict DL CSI, or the work in \([26]\) aiming at harnessing deep learning tools to enable massive MIMO over FDD operations.
REFERENCES


APPENDICES

7 Proof of Theorem 2

The objective here is to proof the validity of the deterministic equivalents given in Theorems 2. We consider herein the notations and the statistical assumptions given in sections 3.2-3.3.1.

Define \( \tilde{H}_{jj} \) as in (3.10) and let \( \tilde{H}_{jj_k} \) be the matrix obtained after removing its \( k \)-th column. The resolvent matrices associated with \( \tilde{H}_{jj} \) and \( \tilde{H}_{jj_k} \) are defined as:

\[
Q_j(\rho \sigma^2) = \left( \tilde{H}_{jj} \tilde{H}_{jj}^H + \rho \sigma^2 I_{M_j} \right)^{-1},
\]

(7.1)

and

\[
Q_j^{(k)}(\rho \sigma^2) = \left( \tilde{H}_{jj_k} \tilde{H}_{jj_k}^H + \rho \sigma^2 I_{M_j} \right)^{-1},
\]

(7.2)

respectively.

When no coordination is applied, consider \( \tilde{e}_{j^no-co} (\rho \sigma^2) \), \( \tilde{e}_{j^no-co} (\rho \sigma^2) \) and \( \tilde{d}_{j^no-co} (\rho \sigma^2) \) defined in (3.20) and (3.21) and (3.23). As per Propositions 1 and 2, for any sequence of matrices \( A_M \) with bounded spectral norm, the following convergences hold true:

\[
\frac{1}{K_j} \text{tr}(A_M Q_j(\rho \sigma^2)) - \frac{1}{\rho \sigma^2} \tilde{e}_{j^no-co} (\rho \sigma^2) \frac{1}{K_j} \text{tr} (A_M) \xrightarrow{a.s.} 0, \quad (7.3)
\]
and
\[
\frac{1}{K_j} \text{tr} \left( A_M \left( Q_j (\rho \sigma^2) \right)^2 \right) - \frac{(\epsilon_j^{\text{no-co}} (\rho \sigma^2))^2}{(\rho \sigma^2)^2 \left( 1 - \frac{1}{c_j} \tilde{g}_j^{\text{no-co}} (\rho \sigma^2) \right)} \frac{1}{K_j} \text{tr} (A_M) \xrightarrow{a.s.} 0, \quad (7.4)
\]

**Deterministic Equivalent for \( S_{jk}^{\text{no-co}} \)**

First, we write \( S_{jk}^{\text{no-co}} \) as:
\[
S_{jk}^{\text{no-co}} = \left| h_{jjk}^H \tilde{w}_{jk}^{\text{no-co}} \right|^2 = \left| h_{jjk}^H Q_j (\rho \sigma^2) \tilde{h}_{jk} \right|^2 \frac{\tilde{h}_{jjk}^H \left( Q_j (\rho \sigma^2) \right)^2 \tilde{h}_{jjk}}{h_{jjk}^H \left( Q_j (\rho \sigma^2) \right)^2 \tilde{h}_{jjk}}
\]

Using Corollary 1, we can substitute \( Q_j (t) \) by \( Q_j^{(k)} (t) \), thus yielding:
\[
S_{jk}^{\text{no-co}} = \left| h_{jjk}^H Q_j^{(k)} (\rho \sigma^2) \tilde{h}_{jk} \right|^2 \frac{h_{jjk}^H \left( Q_j^{(k)} (\rho \sigma^2) \right)^2 \tilde{h}_{jjk}}{\tilde{h}_{jjk} \left( Q_j^{(k)} (\rho \sigma^2) \right)^2 \tilde{h}_{jjk}}
\]

For simplicity, we derive the deterministic equivalents of the numerator and denominator separately. We start by the numerator \( h_{jjk}^H Q_j^{(k)} (\rho \sigma^2) \tilde{h}_{jk} \). Due to the real channel being independent of the estimation error (3.9), a direct application of Lemmas 3 and 2 implies that:
\[
h_{jjk}^H Q_j^{(k)} (\rho \sigma^2) \tilde{h}_{jk} - P_j \alpha_{jjk} \left( 1 - \tau^2 \right) \text{tr} \left( Q_j^{(k)} (\rho \sigma^2) \right) \xrightarrow{a.s.} 0. \quad (7.5)
\]

Finally, using the rank-one perturbation Lemma (Lemma 4) along with (7.3), we ultimately obtain:
\[
h_{jjk}^H Q_j^{(k)} (\rho \sigma^2) \tilde{h}_{jk} - P_j \alpha_{jjk} \frac{1 - \tau^2}{c_j \rho \sigma^2} E_j^{\text{no-co}} (\rho \sigma^2) \xrightarrow{a.s.} 0. \quad (7.6)
\]
Similarly, a deterministic equivalent of the denominator \( \hat{h}^u_{jjk} \left( Q_j^{(k)}(\rho \sigma^2) \right)^2 \hat{h}_{jjk} \) can be obtained using again the convergence in Lemma 3 along with the rank-one perturbation Lemma, thus yielding:

\[
\hat{h}^u_{jjk} \left( Q_j^{(k)}(\rho \sigma^2) \right)^2 \hat{h}_{jjk} - \frac{P_j \alpha_{jjk}}{K_j} \text{tr} \left( \left( Q_j(\rho \sigma^2) \right)^2 \right) \xrightarrow{a.s.} 0. \quad (7.7)
\]

Finally, invoking (7.4), we obtain:

\[
\hat{h}^u_{jjk} \left( Q_j^{(k)}(\rho \sigma^2) \right)^2 \hat{h}_{jjk} - \frac{P_j \alpha_{jjk}}{K_j} \left( c_j (\rho \sigma^2) \right)^2 \left( 1 - \frac{1}{c_j} \tilde{e}_{no-co} \right) \xrightarrow{K \to \infty} 0. \quad (7.8)
\]

**Deterministic Equivalent for \( I_{jk}^{\text{intra,no-co}} \)**

To begin with, we define \( D \) and \( \overline{D} \) as:

\[
D = \text{diag} \left\{ \left( \hat{h}^u_{jjj} (Q_j(\rho \sigma^2))^2 \hat{h}_{jjj} \right)^{-1} \right\}_{i=1}^{K_j}
\]

and

\[
\overline{D} = \text{diag} \left\{ \left( 1 + \frac{P_j \alpha_{jjj} \text{tr} \left( Q_j(\rho \sigma^2) \right)}{K_j} \right)^2 \right\}_{i=1}^{K_j}
\]

We express \( I_{jk}^{\text{intra,no-co}} \) as:

\[
I_{jk}^{\text{intra,no-co}} = \sum_{i=1}^{K_j} \left| \hat{h}^u_{jjk} \tilde{e}_{no-co} \hat{h}_{jjj} \right|^2
\]

\[
= \sum_{i=1}^{K_j} \frac{h_{jjk}^u Q_j(\rho \sigma^2) \hat{h}_{jjj}^u \hat{h}_{jjj}^u Q_j(\rho \sigma^2) \hat{h}_{jjj}^u}{h_{jjj}^u (Q_j(\rho \sigma^2))^2 \hat{h}_{jjj}^u}
\]

\[
= h_{jjk}^u Q_j(\rho \sigma^2) \tilde{H}_{jjk} D_k \tilde{H}_{jjk}^u Q_j(\rho \sigma^2) h_{jjk},
\]
\[
I_{jk}^{\text{intra,no−co}} - h_{jjk}^H Q_j^k (\rho \sigma^2) \hat{H}_{j,jk} \hat{D}_k \hat{H}_{j,jk}^H Q_j^k (\rho \sigma^2) h_{jjk} \\
+ \frac{h_{jjk}^H Q_j^k (\rho \sigma^2) \hat{H}_{j,jk} D_k \hat{D}_k \hat{H}_{j,jk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk} \hat{H}_{j,jk}}{1 + \hat{h}_{jjk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk}} \\
+ \frac{h_{jjk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk} \hat{h}_{jjk}^H Q_j^k (\rho \sigma^2) \hat{H}_{j,jk} D_k \hat{D}_k \hat{H}_{j,jk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk} h_{jjk}}{1 + \hat{h}_{jjk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk}} \\
- \left( 1 + \hat{h}_{jjk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk} \right)^2 \\
a.s. \xrightarrow{K \to \infty} 0.
\] (7.11)

where \( D_k \) is \( D \) with row \( k \) and column \( k \) being removed. We also form \( \overline{D}_k \) by removing the \( k \)-th row and the \( k \)-th column of matrix \( D \).

Using (2.15), we obtain:

\[
\| \overline{D}_k - D_k \| \xrightarrow{a.s.} 0, \quad K \to \infty
\] (7.9)

This allows us to approximate \( I_{jk}^{\text{intra,no−co}} \) as:

\[
I_{jk}^{\text{intra,no−co}} - h_{jjk}^H Q_j (\rho \sigma^2) \hat{H}_{j,jk} \hat{D}_k \hat{H}_{j,jk}^H Q_j (\rho \sigma^2) h_{jjk} \xrightarrow{a.s.} 0
\] (7.10)

For the results of Chapter 2 to be applicable, we need to handle the dependence of \( Q_j (\rho \sigma^2) \) on \( h_{jjk} \). One way to achieve that is to use the property given in (2.12), thereby yielding (7.11). A careful analysis of this latter shows that we need to obtain deterministic equivalents for the following terms to conclude:

- \( A_1 (\rho \sigma^2) = \frac{1}{K} \text{tr} \left( Q_j^k (\rho \sigma^2) \hat{H}_{j,jk} \hat{D}_k \hat{H}_{j,jk}^H Q_j^k (\rho \sigma^2) \right) \),
- \( A_2 (\rho \sigma^2) = h_{jjk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk} \),
- \( A_3 (\rho \sigma^2) = \hat{h}_{jjk}^H Q_j^k (\rho \sigma^2) \hat{h}_{jjk} \).

The asymptotic limit of \( A_2 (\rho \sigma^2) \) is already given in (7.6). As for \( A_3 (\rho \sigma^2) \) a deterministic
$\frac{P_{j,\alpha_{j}}\alpha_{j}}{K_{j}} - 2 \left(1 - \tau_{j,j}^{2}\right) \frac{\left(P_{j,\alpha_{j}}\alpha_{j}\right)^{2} \text{tr} \left(Q_{j} \left(\rho \sigma^{2}\right)\right)}{1 + \frac{\rho_{j} \sigma_{j}^{2}}{\text{tr} \left(Q_{j} \left(\rho \sigma^{2}\right)\right)}} + \left(1 - \tau_{j,j}^{2}\right) \frac{\left(P_{j,\alpha_{j}}\alpha_{j}\right)^{3} \left(\text{tr} \left(Q_{j} \left(\rho \sigma^{2}\right)\right)^{2}\right)}{1 + \frac{\rho_{j} \sigma_{j}^{2}}{\text{tr} \left(Q_{j} \left(\rho \sigma^{2}\right)\right)}}$

\[\xrightarrow{a.s.} K \to \infty 0. \]  

(7.16)

equivalent unfolds from a direct application of Lemma 3 along with (7.3), thus yielding:

$$A_{3} \left(\rho \sigma^{2}\right) - P_{j,\alpha_{j}}\alpha_{j} \frac{\tilde{c}_{j}^{\text{no-co}} \left(\rho \sigma^{2}\right)}{c_{j} \rho \sigma^{2}} \xrightarrow{a.s.} 0. \quad (7.12)$$

On the other hand, $A_{1} \left(\rho \sigma^{2}\right)$ has a variance decreasing at a rate $M^{-2}$ [94, Proposition 4], and as such satisfies:

$$A_{1} \left(\rho \sigma^{2}\right) - \mathbb{E} \left[A_{1} \left(\rho \sigma^{2}\right)\right] \xrightarrow{a.s.} 0, \quad (7.13)$$

Now, expanding $\mathbb{E} \left[A_{1} \left(\rho \sigma^{2}\right)\right]$, we obtain:

$$\mathbb{E} \left[A_{1} \left(\rho \sigma^{2}\right)\right] = \frac{1}{K_{j}} \sum_{i=1}^{K_{j}} \mathbb{E} \left[\text{tr} \left(\tilde{h}_{j,j}^{\text{h}} \left(Q_{j}^{(k)} \left(\rho \sigma^{2}\right)\right)^{2} \tilde{h}_{j,j}\right)\right]. \quad (7.14)$$

Again, the dependence of $Q_{j}^{(k)} \left(\rho \sigma^{2}\right)$ on $\tilde{h}_{j,j}^{\text{h}}$ can be controlled using (2.11).

$$\mathbb{E} \left[A_{1} \left(\rho \sigma^{2}\right)\right] - 1 \xrightarrow{a.s.} 0 \quad (7.15)$$

Plugging the deterministic equivalents of $A_{2} \left(\rho \sigma^{2}\right)$ (7.6), $A_{3} \left(\rho \sigma^{2}\right)$ (7.12) and $A_{1} \left(\rho \sigma^{2}\right)$ (7.15) into (7.11) yields the asymptotic expression in (7.16). This latter is exactly the same expression (3.25) of $I_{j,k}^{\text{intra,no-co}}$ given in Theorem 2 once we substitute $\frac{1}{K_{j}} \text{tr} \left(Q_{j} \left(\rho \sigma^{2}\right)\right)$ and $\frac{1}{K_{j}} \text{tr} \left(\left(Q_{j} \left(\rho \sigma^{2}\right)\right)^{2}\right)$ by their respective asymptotic limits (7.3) and (7.4).
Deterministic Equivalent for $I_{j\ell k}^{\text{inter,no-co}}$

To begin with, we write $I_{j\ell k}^{\text{inter,no-co}}$ as:

$$I_{j\ell k}^{\text{inter,no-co}} = \sum_{i=1}^{K_j} \sum_{i \neq k} |h_{j\ell k}^H \hat{w}_{ji}^{\text{no-co}}|^2 = \sum_{i=1}^{K_j} h_{j\ell k}^H Q_j(\rho \sigma^2) \hat{H}_{jji} h_{j\ell k}^H Q_j(\rho \sigma^2) h_{j\ell k}$$

Following the same technicalities leading to (7.10), we have:

$$I_{j\ell k}^{\text{inter,no-co}} - h_{j\ell k}^H Q_j(\rho \sigma^2) \hat{H}_{j,Jk} \hat{D}_k \hat{H}_{j\ell k}^H Q_j(\rho \sigma^2) h_{j\ell k} \xrightarrow{a.s.} 0 \quad K \to \infty$$

(7.17)

where $\hat{D}_k$ has the same expression as in (7.10).

Note that here, the result (7.15) can be applied directly since $h_{j\ell k}$ is independent of $Q_j(\rho \sigma^2)$, which is in contrast to (7.11) where lemma 2 had to be used. Therefore, we can easily find that

$$I_{j\ell k}^{\text{inter,no-co}} - P_j \alpha_{j\ell k} \xrightarrow{a.s.} 0 \quad K \to \infty$$

(7.18)

8 Sketch of Proof of Theorem 3

In this appendix, we sketch the proof of Theorem 3 as the derivations rely most often on the same arguments as those presented in Theorem’s 2 proof above.

Considering the design of the coordinated SLNR-MAX precoding (3.17), one should
note that the beamforming vectors involve now matrix $\hat{\mathbf{H}}_j$ given by:

$$\hat{\mathbf{H}}_j = \left[ \beta_j \hat{\mathbf{h}}_{j1}, \beta_j \hat{\mathbf{h}}_{j2}, \ldots, \hat{\mathbf{h}}_{jj}, \ldots, \beta_j \hat{\mathbf{h}}_{jL} \right], \quad (8.1)$$

This must be compared to the non-coordinated SLNR-MAX (3.12) in which the beamforming vectors involve only $\hat{\mathbf{h}}_{jj}$. Furthermore, some care must be paid when applying propositions 1 and 2 as the number of columns of matrix $\hat{\mathbf{H}}_j$ differs from the variance scaling $\frac{1}{K_j}$. To handle this issue, we need to consider the following relation:

$$\left( \hat{\mathbf{H}}_j \hat{\mathbf{H}}^\mathbf{H}_j + \rho \sigma^2 \mathbf{I}_{M_j} \right)^{-1} = \frac{K_j}{K} \mathbf{Q}_j \left( \frac{K_j}{K} \rho \sigma^2 \right), \quad (8.2)$$

where

$$\mathbf{Q}_j \left( \frac{K_j}{K} \rho \sigma^2 \right) = \left( \frac{K_j}{K} \hat{\mathbf{H}}_j \hat{\mathbf{H}}^\mathbf{H}_j + \frac{K_j}{K} \rho \sigma^2 \mathbf{I}_{M_j} \right)^{-1}. \quad (8.3)$$

Considering $e_j^\mathbf{co} (\rho \sigma^2)$ in (3.38) and $\tilde{e}_j^\mathbf{co} (\rho \sigma^2)$ in (3.42) and letting $t_j^\mathbf{co}$ be as in (3.39), a direct application of Propositions 1 and 2 yields:

$$\frac{1}{K} \text{tr} \left( \mathbf{A}_M \mathbf{Q}_j \left( t_j^\mathbf{co} \right) \right) - \frac{\tilde{e}_j^\mathbf{co} \left( t_j^\mathbf{co} \right)}{t_j^\mathbf{co}} \frac{1}{K} \text{tr} \left( \mathbf{A}_M \right) \xrightarrow{\text{a.s.}} 0, \quad (8.4)$$

and

$$\frac{1}{K} \text{tr} \left( \mathbf{A}_M \left( \mathbf{Q}_j \left( t_j^\mathbf{co} \right) \right)^2 \right) - \frac{\left( \tilde{e}_j^\mathbf{co} \left( t_j^\mathbf{co} \right) \right)^2}{\left( t_j^\mathbf{co} \right)^2 \left( 1 - \frac{1}{e_j^\mathbf{co} \tilde{e}_j^\mathbf{co} \left( t_j^\mathbf{co} \right)} \right)} \frac{1}{K} \text{tr} \left( \mathbf{A}_M \right) \xrightarrow{\text{a.s.}} 0, \quad (8.5)$$

where $\mathbf{A}_M \in \mathbb{C}^{M_j \times M_j}$ is an arbitrary sequence of matrices with bounded spectral norm.

As mentioned before, the procedure to derive the asymptotic expressions of Theorem 3 is in essence similar to the proof of Theorem 2. In fact, two differences are worthy of mentioning and have to be considered. First, the convergences (8.2)-(8.5) have to be taken into account and applied wherever necessary to preserve the coherence with regard to the Random Matrix Theory tools in use. Apart from this fact, the
working out of $S_{jk}^{co}$ and $T_{jk}^{intra,co}$, obeys exactly the same arguments as for $S_{jk}^{no-co}$ and $T_{jk}^{intra,no-co}$, respectively. The second difference between the two theorems takes place when dealing with the term $I_{jk}^{inter,co}$. As explained in section 3.3.1, when coordination is implemented among base stations, the latter will have knowledge of both intended and non-intended channels which will be used later on in the beamforming. Hence, regardless of the parameter $\beta_j$, the inter-cell interference and the intra-cell interference takes conceptually the same form. Therefore, in Theorem 3, the derivation of $T_{jk}^{inter,co}$ follows the same steps that led to finding $I_{jk}^{intra,no-co}$. Details are thus omitted.

9 Proof of Theorem 5 (Conventional Combining in Single-Cell Systems)

We demonstrate in this section the results of Theorem 5. As the derivations rely most often on the same arguments, we mention the pertinent steps that allow to obtain the asymptotic approximation $\overline{SE}_{k}^{conv,S}$ (4.1).

First, define $\bar{Q} = \left( \frac{1}{N} \hat{H}^{H}Z\hat{H} + \frac{1}{\rho_d} I_K \right)^{-1}$, with $Z^{-1} = \frac{\rho_d}{N} \sum_{i=1}^{K} \left( R_i - \hat{R}_i \right) + I_N$. Second, Woodbury matrix identity (1) enables to express all the signals constituting $\overline{SE}_{k}^{conv,S}$ in terms of the entries of matrix $\bar{Q}$. For instance, the signal term $|g_k^{H}\hat{h}_k|^2$, can be written as:

$$
\left| \frac{1}{N} \hat{h}_k^{H} \left( \hat{H}^{H} + \frac{1}{\rho_d} I_N \right)^{-1} \hat{h}_k \right|^2 = \left| 1 - \frac{1}{\rho_d} [\bar{Q}]_{kk} \right|^2. \quad (9.1)
$$
Accordingly, $\text{SE}_k^{\text{conv},S}$ in (4.7) can be rewritten as:

$$
\text{SE}_k^{\text{conv},S} = \left(1 - \frac{\tau}{T}\right) \log \left(1 + \frac{1}{\rho_d} \left[ \hat{Q}_{kk} \right] \right)^2 + \frac{1}{\rho_d} \left[ \hat{Q}_{kk} \right] + \frac{1}{\rho_d} \left[ \hat{Q}^2_{kk} \right] \right),
$$

(9.2)

In fact, putting the spectral efficiency $\text{SE}_k^{\text{conv},S}$ in this format (9.2) facilitates the derivation of the deterministic equivalent of this latter since we can simply use the law of large numbers as follows:

- Under assumption 6, the LLN allows us to put $\frac{1}{N} \left[ \hat{H}^H \hat{H} \right]_{ij} - \frac{1}{N} \mathbb{E} \left[ \hat{h}_i^H \hat{h}_j \right] \xrightarrow{a.s. N \to \infty} 0$. Therefore, under Assumption 6 using the continuous mapping theorem (Chapter 2, Theorem 1), we have:

$$
\left[ \tilde{Q} \right]_{ij} - \left[ Q \right]_{ij} \xrightarrow{a.s. N \to \infty} 0,
$$

(9.3)

where the matrix $Q$ is given in Theorem 5 equation (4.9) as:

$$
Q = \left( \frac{1}{N} \hat{H}^H \hat{H} + \frac{1}{N} \text{diag} \left\{ \text{tr} \hat{R}_\ell \right\} \right)^{-1}.
$$

(9.4)

Therefore, a direct application of (9.3) with the continuous mapping theorem enables us to find asymptotic approximations of most of the terms in (9.2). For instance, for the signal term, we find:

$$
\left| 1 - \frac{1}{\rho_d} \left[ \tilde{Q} \right]_{kk} \right|^2 - \left| 1 - \frac{1}{\rho_d} \left[ Q \right]_{kk} \right|^2 \xrightarrow{a.s. N \to \infty} 0.
$$

(9.5)

Likewise, the same steps allow to derive approximations of the intra-cell interference term and processed noise. As to the estimation error term, $\mathcal{E}_k = \frac{1}{N^2} \sum_{i=1}^K \hat{q}_i^H \hat{H}^H \xi_i \xi_i^H \hat{H} \tilde{q}_k$, we mainly adopt the same reasoning except for the following step. Indeed, in order to find a deterministic equivalent for $\mathcal{E}_k$, we first exploit the
orthogonality property of LMMSE channel estimation by observing that \( \forall \{k, i\}, \hat{h}_k \) and \( \xi_i \) are independent. After that, applying the convergence of quadratic forms lemma (Chapter 2, lemma 3) yields:

\[
E_k - \left\{ \frac{1}{N^2} \sum_{i=1}^{K} q_k^H \left( \hat{H}^H \left( \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi_i \right) \hat{H} + diag \{ \text{tr} \left( \tilde{R}_\ell \frac{1}{\tau \rho_{tr}} \mathbf{R}_i \Phi_i \right) \} \right) q_k \right\} \xrightarrow{a.s.} 0 \quad (9.6)
\]

Finally, putting all the above terms together yields the asymptotic approximation of the spectral efficiency in Theorem 5 and as such, concludes the proof.

10 Proof of Theorem 6 (Optimal training)

Denote \( \partial_\tau F \) and \( \partial^2_\tau F \) as the first and second derivatives of any function \( F(\tau) \) with respect to \( \tau \). To prove the results of Theorem 6, we use the following approach:

First, we show that \( \forall \tau, \partial_\tau \text{SE}_{\text{conv}} \) is monotonically decreasing and that \( \exists \tau_0 \) such that \( \partial_\tau \text{SE}_{\text{conv}}|_{\tau=\tau_0} = 0 \). If held, these two properties lead to the unique \( \tau_0 \) which maximizes the SE, i.e., \( \tau_0 = \arg\max_{\tau} \text{SE}_{\text{conv}} \). Accordingly, if this step is verified, considering the constraint in (P2), finding \( \tau^* \) is simply obtained as:

- If \( \tau_0 \leq K \), then \( \tau^* = K \), which is depicted by the first solution (4.14).
- On the other hand, if \( K < \tau_0 \leq T \), then \( \tau^* = \tau_0 \) which is represented by the solution (4.15).

**Proof that \( \forall \tau, \partial_\tau \text{SE}_{\text{conv}} \) is monotonically decreasing** To establish this, a sufficient condition is to have: \( \forall \tau \in [K, T], \partial^2_\tau \text{SE}_{\text{conv}} < 0 \). In this line, recalling the SE expression (4.12), \( \text{SE}_{\text{conv}}^{k, S} = \frac{1}{T} \log \left( 1 + \frac{\gamma_k}{\tau} \right) \), with:

\[
1 + \gamma_k = \frac{\rho_d}{Q_{kk}}, \quad (10.1)
\]
we have:

\[ \partial_\tau \mathbf{SE}^{\text{conv}} = \frac{1}{K} \sum_{k=1}^{K} \left[ (1 - \frac{\tau}{T}) \frac{\gamma_k'(\tau)}{1 + \gamma_k(\tau)} - \frac{1}{T} \log (1 + \gamma_k(\tau)) \right], \quad (10.2) \]

\[ \partial^2_\tau \mathbf{SE}^{\text{conv}} = \frac{1}{K} \sum_{k=1}^{K} \left[ - \frac{2}{T} \gamma_k'(\tau) + \left( 1 - \frac{\tau}{T} \right) \frac{\gamma_k''(\tau)(1 + \gamma_k(\tau)) - (\gamma_k'(\tau))^2}{(1 + \gamma_k(\tau))^2} \right], \quad (10.3) \]

with \( \gamma_k'(\tau) \) and \( \gamma_k''(\tau) \) given by:

\[ \gamma_k'(\tau) = \rho_d \frac{q_k^H D_2 q_k}{([Q]_{kk})^2}, \quad (10.4) \]

\[ \gamma_k''(\tau) = -2 \rho_d \frac{q_k^H D_2 (Q [Q]_{kk} - q_k q_k^H) D_2 q_k - [Q]_{kk} q_k q_k^H D_2 q_k}{([Q]_{kk})^3}, \quad (10.5) \]

and \( D_\alpha = \frac{(-1)^\alpha}{\rho_d \tau^\alpha} \text{diag} \left\{ \frac{1}{N} \text{tr} (R_\ell^\alpha \Phi_\ell^\alpha) \right\}_{\ell=1}^{K} \), with \( \alpha \) being an integer. Note that \( D_\alpha \) is a positive semi-definite matrix for all even values of \( \alpha \), and negative semi-definite otherwise. With this in mind, from (10.4), we can see that \( \gamma_k'(\tau) \geq 0, \forall \tau \in [K,T[. \]

Plus, since \( \forall \tau, \gamma_k(\tau) \) is evidently positive, it can be seen from (10.3) that \( \gamma_k''(\tau) \leq 0, \forall \tau \in [K,T[ \) is a sufficient condition to obtain : \( \partial^2_\tau \mathbf{SE}^{\text{conv}} \leq 0. \) With this in mind, using the fact that \( a^H a I_M - a a^H \) is a positive semi-definite matrix, we can easily show that

\[ \gamma_k''(\tau) \leq 0, \forall \tau. \]

This concludes the proof that \( \partial_\tau \mathbf{SE}^{\text{conv}} \) is monotonically decreasing.
with respect to $\tau$, and validates the results given in Theorem 6.

### 11 Proof of Corollary 4.2.3.4

For simplicity, note that the index "$k$" will be dropped in the sequel. Accordingly, denoting $\alpha_i$ the $i-$th eigenvalue of $\Theta$, our objective is to find $\kappa$ such that:

$$
\left(1 - \frac{T}{T'}\right) \log \left[ 1 + \frac{\beta d}{1 + \kappa} \left( \frac{1}{N} \sum_{i=1}^{N} \frac{\alpha_i^2}{\alpha_i + \frac{1 + \kappa}{\rho_{tr}\beta \tau}} + \kappa \right) \right] \leq \log \left( 1 + \rho_d \beta \frac{\kappa}{1 + \kappa} \right). \quad (P3')
$$

Since "log" is an increasing function and $\alpha_i \geq 0, \forall i$, we consider the upper bound:

$$
\frac{\alpha_i^2}{\alpha_i + \frac{1 + \kappa}{\rho_{tr}\beta \tau}} \leq \alpha_i, \text{ for all positive values of } \kappa, \tau \text{ and } \rho_{tr}. \text{ Therefore, } (P3') \text{ is satisfied whenever } \kappa \text{ verifies:}
$$

$$
\left(1 - \frac{T}{T'}\right) \log \left[ 1 + \frac{\beta d}{1 + \kappa} \left( \frac{1}{N} \text{tr } \Theta + \kappa \right) \right] \leq \log \left( 1 + \rho_d \beta \frac{\kappa}{1 + \kappa} \right). \quad (11.1)
$$

Applying "exp" on both sides of (11.1) yields the lower bound $\kappa \geq f(\kappa)$, with:

$$
f(\kappa) = -1 + \frac{\frac{1}{N} \text{tr } \Theta}{-\frac{\kappa}{\kappa+1} + \frac{1}{\beta d} \left( -1 + \left( 1 + \frac{\beta d \kappa}{1 + \kappa} \right) \frac{1}{T'-T} \right)}. \quad (11.2)
$$

Next, we use the following result: \textit{Let } $x, \alpha \in \mathbb{R}$, \textit{and consider } $y(x)$, \textit{a positive increasing function of all } $x \geq 0$ \textit{with } $y(0) = 0$. \textit{If } $\alpha > 1$, \textit{the following inequality holds:}

$$
(1 + y(x))^\alpha \geq (1 + \alpha y(x)), \quad (11.3)
$$

\textit{Remark.} This result can be proved by showing that the function $g(x) = (1 + y(x))^\alpha - (1 + \alpha y(x))$ verifies, $\forall x \geq 0$: $g(0) = 0$ and $g'(x) \geq 0$, thus yielding: $g(x) \geq 0$, $\forall x \geq 0$.

Accordingly, applying (11.3) on (11.2) allows us to obtain the following condition.
on $\kappa$:

$$\kappa \geq -1 + \frac{\text{tr} \Theta (T - \tau) (1 + \kappa)}{\tau \kappa}.$$ 

This latter admits the solution: $\kappa \geq \frac{\text{tr} \Theta (T - \tau)}{\tau}$. Finally, since the function $\frac{T - \tau}{\tau}$ is decreasing in $\tau$, plus the fact that $\tau \in [K, T)$, we can simply consider the lower bound given in (4.24), therefore concluding the proof.

12 Proof of Theorem 8

The same steps and arguments given in appendix 9 can be used to find the asymptotic approximation $\overline{SE}^{\text{conv}, M}_{jk}$ in Theorem 8 and are thus omitted. Nevertheless, the main difference lies in the inter-cell interference term, where it is imperative to take into account the correlation between the estimates and the interfering channels that share the same pilot, s.t. $\forall \ell \neq j : \frac{1}{N} \mathbb{E} [\hat{h}_{jji}^H \hat{h}_{jli}] = \frac{1}{N} \text{tr}(R_{jli} \Phi_{jk} R_{jj}) \xrightarrow{a.s.} 0$.

13 Proof that pilot contamination is decreasing with respect to $\kappa_{jk}$

First we recall in the following lemma some properties on positive-semi-definite matrices.

**Lemma 5.** [95] For any positive semi-definite (PSD) $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$, the matrices $\mathbf{ABA}$, $\mathbf{BAB}$ and $\mathbf{A} + \mathbf{B}$ are positive semi-definite, and $\text{tr}(\mathbf{AB}) \geq 0$. Plus, if $\mathbf{AB} = \mathbf{BA}$, then $\mathbf{AB}$ is also PSD. Finally, if $\mathbf{A}$ is positive definite (PD), then $\mathbf{A}^{-1}$ is also PD.

Second, let $f(\kappa_{jk}) = \frac{1}{N} \text{tr}(\sum_{\ell \neq j}^{L} R_{j\ell k} \Phi_{jk} R_{jj})$. A straightforward differentiation of
\(f(\kappa_{jk})\) yields:

\[
f'(\kappa_{jk}) = \frac{-1}{1 + \kappa_{jk} N} \text{tr} \left( \sum_{\ell \neq j}^{L} R_{j\ell k} \Phi_{jk} (R_{jjk} - \tilde{R}_{jjk}) \right).
\] (13.1)

Accordingly, we prove in what follows that:

\[
\text{tr} \left( \sum_{\ell \neq j}^{L} R_{j\ell k} \Phi_{jk} (R_{jjk} - \tilde{R}_{jjk}) \right) \geq 0 \Rightarrow f'(\kappa_{jk}) \leq 0.
\] (13.2)

- In this line, we assume that the correlation matrix \(R_{jjk}\) is positive definite. That is, we add the ‘perturbation’ \(\epsilon I_N\), \(\epsilon > 0\) as follows:

Denote \(R_{jjk} = R_{jjk} + \epsilon I_N\), and \(\Phi_{jk} = \left( R_{jjk} + \sum_{\ell \neq j} R_{j\ell k} + \frac{1}{\tau_{pr}} I_N \right)^{-1}\). Therefore, under this assumption, define:

\[
g(\kappa_{jk}, \epsilon) = \text{tr} \left( \sum_{\ell \neq j}^{L} R_{j\ell k} \Phi_{jk} (R_{jjk} - \tilde{R}_{jjk}) \right).
\] (13.3)

Evidently, \(g(\kappa_{jk}, 0) = \text{tr} \left( \sum_{\ell \neq j}^{L} R_{j\ell k} \Phi_{jk} (R_{jjk} - \tilde{R}_{jjk}) \right)\). Plus, since \(g(\kappa_{jk}, \epsilon)\) is continuous, if:

\[
\forall \epsilon > 0, g(\kappa_{jk}, \epsilon) > 0 \Rightarrow \lim_{\epsilon \to 0} g(\kappa_{jk}, \epsilon) \geq 0
\] (13.4)

Therefore, proving this statement will conclude the proof.

- \(\forall \epsilon > 0\), let \(B_{jk} = \sum_{\ell \neq j}^{L} R_{j\ell k}\), \(A_{jk} = \left( B_{jk} + \frac{1}{\tau_{pr}} I_N \right)^{-1}\).

Therefore \(\Phi_{jk} = \left( R_{jjk} + A_{jk}^{-1} \right)^{-1}\). Now, using the Woodbury matrix identity, we can put: \(R_{jjk} \Phi_{jk} \tilde{R}_{jjk} = \left( R_{jjk}^{-1} + A_{jk} \right)^{-1}\). Therefore:

\[
g(\kappa_{jk}, \epsilon) = \text{tr} \left( B_{jk} \left( R_{jjk} + A_{jk}^{-1} \right)^{-1} \left( R_{jjk}^{-1} + A_{jk} \right)^{-1} \right)
\]

\[
= \text{tr} \left( B_{jk} A_{jk} \left( 2A_{jk} + R_{jjk}^{-1} + A_{jk} R_{jjk} A_{jk} \right)^{-1} \right). \quad \text{(13.5)}
\]
Plus, based on lemma 5 the product \( B_{jk}A_{jk} \) is also a PSD matrix by observing that \( B_{jk}A_{jk} = A_{jk}B_{jk} \).

Consequently, we find that \( g(\kappa_{jk}, \epsilon) \) amounts to the trace of the product of two PSD matrices which is always a positive quantity. With in mind, and considering (13.2)-(13.4), \( f'(\kappa_{jk}) \leq 0 \) and thus, the pilot contamination induced interference is decreasing with \( \kappa_{jk} \).

14 Proof of Theorems 10 and 11 (Single-cell Processing - MRC and S-MMSE)

For both Theorems 10 and 11, the derivations rely most often on the same arguments. Plus, we mention in this appendix the steps to obtain \( \gamma_{j,k}^{S-MMSE} \) since it is more involved than \( \gamma_{j,k}^{MRC} \).

The approach we pursue is similar to the one used in Appendix 9. In this line, define the \( N \times K \) matrix \( \hat{H}_j \) and the \( K \times K \) matrix:

\[
\bar{Q}_j = \left( \frac{1}{N} \hat{H}_j^H Z_j \hat{H}_j \right)^{-1}.
\] (14.1)

Under assumption 7, the LLN allows us to put:

\[
\frac{1}{N} \left[ \hat{H}_j^H Z_j \hat{H}_j \right]_{ik} - \frac{1}{N} \mathbb{E} \left[ \hat{h}_{J,j}^H \hat{h}_{J,j} \right] \xrightarrow{a.s. \ N \to \infty} 0.
\]

Therefore, using the continuous mapping theorem 81, we have:

\[
[\bar{Q}_j]_{ik} - [Q_j]_{ik} \xrightarrow{a.s. \ N \to \infty} 0,
\]

where the matrix \( Q_j \) is given in (5.20) as:

\[
Q_j = \left( \frac{1}{N} \hat{H}_j^H Z_j \hat{H}_j + \text{diag} \{ \beta_{j,j} \} \right)^{-1}.
\] (14.2)
Now, applying the Woodbury identity on the signal term yields:

\[
\left| (g_{jk}^{S-MMSE})^* \hat{h}_{jjk} \right|^2 = \left| 1 - \frac{1}{N} [\tilde{Q}_j]_{kk} \right|^2 \tag{14.3}
\]

Accordingly, rewriting each term in the SINR expression (5.7) using \( \tilde{Q}_j \) (14.1) enables easily to find an asymptotic equivalent of \( \gamma_{jk}^{S-MMSE} \). For instance: the signal:

\[
\left| (g_{jk}^{S-MMSE})^* \hat{h}_{jjk} \right|^2 - \left| 1 - \frac{1}{N} [Q_j]_{kk} \right|^2 \xrightarrow{a.s.; N \to \infty} 0.
\]

The main challenge towards obtaining (5.19) lies in the the inter-cell interference where one should be careful to take into account the correlation between the estimates and the interfering channels that share the same pilot (all channels \( \{ \hat{h}_{j\ell k} : \ell = 1, \ldots, L \} \) are correlated). In this line, we have \( \forall \ell \neq j \):

\[
\frac{1}{N} \mathbb{E} \left[ \hat{h}_{jji}^H Z_j^s \hat{h}_{j\ell i} \right] = \frac{1}{N} \text{tr} \left( R_{j\ell i} \Phi_{jk} R_{jjj} Z_j^s \right) = \beta_{ji,\ell j}^s. \tag{14.4}
\]

Finally, based on the continuous mapping Theorem [81], putting all these deterministic equivalents together yields the asymptotic approximation of the SINR given in Theorem [11].

15 Proof of Theorem [12] (Multi-cell Processing - M-MMSE)

We summarize in this appendix the main steps and arguments to obtain the approximation given in Theorem [12]. Prior to that, define the following quantities which will simplify the derivation and presentation of the results:

- \( \hat{H}_{j/k} = \{ \hat{h}_{j\ell i} : \ell = 1, \ldots, L \}, \{ i = 1, \ldots, K, i \neq k \} \), i.e. \( \hat{H}_{j/k} \) includes all the
estimates $\hat{h}_{j\ell i}$, $\forall (\ell, i)_{i \neq k}$, except for the $k$–th UE. This quantity is necessary to define:

$$A_{j,k} = \sum_{\ell=1}^{L} \sum_{i=1}^{K} \hat{h}_{j\ell i} \hat{h}_{ji} + (Z_j^M)^{-1}$$

$$= \hat{H}_{j,k} \hat{H}_{j,k}^H + (Z_j^M)^{-1} \quad (15.1)$$

- $\hat{H}_{j,k} = [\hat{h}_{j1k}, \ldots, \hat{h}_{jj-1k}, \hat{h}_{jj+1k}, \ldots, \hat{h}_{jLk}]$, i.e., $\hat{H}_{j,k} \in \mathbb{C}^{N \times (L-1)}$ contains all inter-cell pilot contamintors of channel $h_{jjk}$, and does not include any LoS components.

Accordingly, using the Woodburry identity (chapter 2, lemma 1), we can rewrite $\gamma_j^{M-MMSE}$ as:

$$\frac{\gamma_j^{M-MMSE}}{N} = \left( \frac{1}{N} \hat{H}_{j,k} \hat{H}_{j,k}^H A_{j,k}^{-1} \hat{H}_{j,k} \hat{H}_{j,k}^H \right)^{-1} \hat{H}_{j,k} \hat{H}_{j,k}^H Z_j^M \quad (15.2)$$

Next, we derive approximations for $\Pi_1, \Pi_2, \Pi_3$.

- First, let us focus on the term $\Pi_1$. Using the Woodburry identity enables to rewrite $\frac{1}{N} A_{j,k}^{-1}$ as:

$$= \frac{1}{N} \hat{h}_{j,k}^H Z_j^M \hat{h}_{j,k} + \frac{1}{N} Z_j^M \hat{H}_{j,k} \hat{H}_{j,k}^H \left( \frac{1}{N} \hat{H}_{j,k}^H Z_j^M \hat{H}_{j,k} \hat{H}_{j,k}^H + \frac{1}{N} I_{L(K-1)} \right)^{-1} \frac{1}{N} \hat{H}_{j,k}^H Z_j^M \quad (15.3)$$

Therefore, $\Pi_1$ can be rewritten as:

$$\Pi_1 = \left( \frac{1}{N} \hat{h}_{j,k}^H Z_j^M \hat{h}_{j,k} \right)_{\Pi_{1,1}} - \left( \frac{1}{N} \hat{h}_{j,k}^H Z_j^M \hat{H}_{j,k} \left( \frac{1}{N} \hat{H}_{j,k}^H Z_j^M \hat{H}_{j,k} \hat{H}_{j,k}^H + \frac{1}{N} I_{L(K-1)} \right)^{-1} \frac{1}{N} \hat{H}_{j,k}^H Z_j^M \right)_{\Pi_{1,2}} \hat{H}_{j,k} \hat{H}_{j,k}^H Z_j^M \hat{h}_{j,k} \quad (15.4)$$
Recalling that $\hat{H}_{j,k}$ includes all the estimates $\hat{h}_{j,i}$, $\forall (\ell, i)$, with $i \neq k$, it is thus, fully uncorrelated with $\hat{h}_{jjk}$, and only has Rician components for $\ell = j$. Consequently, recalling the expressions $D_{jk}(5.25)$ and $\bar{H}_{j,k}$ defined in Theorem 12, a direct application of the convergence of quadratic forms lemma [80], yields:

$$\Pi_{1.1} = \frac{1}{N} \text{tr} \left( R_{jjk} \Phi_{jk} R_{jjk} Z_j^M \right) + \frac{1}{N} \bar{h}_{jk} \bar{Z}_j^M \bar{h}_{jk} \xrightarrow{a.s.} 0, \quad (15.5)$$

$$\Pi_{1.2} = \frac{1}{N} \bar{h}_{jk} \bar{Z}_j^M \bar{H}_{j,k} \xrightarrow{a.s.} 0, \quad (15.6)$$

$$\Pi_{1.3} = \left\{ D_{jk} + \frac{1}{N} \bar{h}_{jk} \bar{Z}_j^M \bar{H}_{j,k} \right\} \xrightarrow{a.s.} 0. \quad (15.7)$$

Then, putting together the deterministic equivalents in (15.5)-(15.7) according to (15.4) provides an approximation of $\Pi_1$.

Second, we consider $\Pi_2$. Considering the expression of $A_{j,k}^{-1}(15.3)$, $\Pi_2$ can be expanded to:

$$\Pi_2 = \frac{1}{N} \hat{h}_{jk}^H Z_j^M \hat{H}_{j,k} \xrightarrow{n_{2.1}} - \frac{1}{N} \hat{h}_{jk}^H Z_j^M \hat{H}_{j,k} \xrightarrow{n_{2.2}} \left( \frac{1}{N} \hat{H}_{jk}^H Z_j^M \hat{H}_{j,k} + \frac{1}{N} I_{L(K-1)} \right)^{-1} \frac{1}{N} \hat{H}_{jk}^H Z_j^M \hat{H}_{j,k} \xrightarrow{n_{2.3}} \frac{1}{N} \hat{H}_{jk}^H Z_j^M \hat{H}_{j,k} \xrightarrow{n_{2.4}}$$

(15.8)

Since $\hat{H}_{j,k} = [\hat{h}_{j1k} \ldots \hat{h}_{j(j-1)k} \hat{h}_{j(j+1)k} \ldots \hat{h}_{jLk}]$ does not include any Rician components, obtaining an approximation of $\Pi_{2.1} \in \mathbb{C}^{1 \times (L-1)}$ is achieved by a direction application of the law of large numbers. Therefore $\Pi_{2.1} - \Pi_{2.1} \xrightarrow{a.s.} 0$, such that $\Pi_{2.1} \in \mathbb{C}^{1 \times (L-1)}$ has the $\ell -$th entry:

$$[\Pi_{2.1}]_\ell = \frac{1}{N} \text{tr}(R_{jlk} \Phi_{jk} R_{jjk} Z_j), \quad \ell \neq j. \quad (15.9)$$

Now, approximations of $\Pi_{2.2}$ and $\Pi_{2.3}$ can be easily obtained by observing that
these two quantities coincide with those in (15.6) and (15.7), respectively. On another note, recalling that \( \hat{H}_{j,k} \) and \( \hat{H}_{j,k,j} \) are completely uncorrelated, and that \( \hat{H}_{j,k,j} \) does not involve any LoS components, then, \( \Pi_{2.4} \xrightarrow{a.s.} 0 \). Thus, as \( \Pi_{2.2}, \Pi_{2.3} \) are finite size quantities, and since \( \Pi_{2.4} \xrightarrow{a.s.} 0 \), the entire second term of (15.8) tends to zero, as \( N \) grows large without bound. Accordingly, \( \Pi_2 - \Pi_{2.1} \xrightarrow{a.s.} 0 \) with \( \Pi_{2.1} \) in (15.9).

- Third, we focus on \( \Pi_3 \): Treating this term amounts to the entries of finite dimension matrix \( \Pi_3 = \frac{1}{N} \hat{H}^H_{j,k,j} A^{-1}_{j,k} \hat{H}_{j,k,j} \) which is fairly similar to \( \Pi_2 \). Accordingly, \( \Pi_3 \) can be rewritten as:

\[
\Pi_3 = \frac{1}{N} \hat{H}^H_{j,k,j} Z^M_{j} \hat{H}_{j,k,j} - \Pi_{3.1} - \Pi_{3.3} - \Pi_{3.4} \tag{15.10}
\]

\( \hat{H}_{j,k,j} \in \mathbb{C}^{(L-1)^2} \) is of finite size and contains only scattered components. Therefore, applying the LLN leads to: \( \Pi_{3.1} - \Pi_{3.1} \xrightarrow{a.s.} 0 \), s.t.:

\[
[\Pi_{3.1}]_{mn} = \frac{1}{N} \text{tr} \left( R_{jmk} \Phi_{jk} R_{jnk} Z^M_{j} \right). \tag{15.11}
\]

Additionally, we have \( \Pi_{3.2} = \Pi^H_{2.4} \), and thus \( \Pi_{3.2} \xrightarrow{a.s.} 0 \). Similarly, \( \Pi_{3.3} = \Pi_{1.3} \), thus \( \Pi_{3.3} \) can be approximated by (15.6). However, since \( \Pi_{3.2} \xrightarrow{a.s.} 0 \), we find \( \Pi_{3} - \Pi_{3.1} \xrightarrow{a.s.} 0 \) with \( \Pi_{3.1} \) in (15.11).

Finally, adding the obtained asymptotic approximations of the terms \( \Pi_1, \Pi_2 \) and
16 Proof of Theorem 13 (M-MMSE with inter-cell LoS)

We provide in this appendix the main steps to obtain the approximation given in Theorem 13. Note that all the notations and quantities in this appendix follow the quantities and channel model of Section 5.4 with the fully correlated Rician fading given in (5.30), (5.31)-(5.33), i.e. all the links in the network include a combination of scattered and specular components.

First, define the following matrix: 

\[ \hat{H}_{j,k} = [\hat{h}_{j1}, \ldots, \hat{h}_{jLk}] \in \mathbb{C}^{N \times L}, \] 

matrix that gathers all the estimates \( \hat{h}_{j\ell k} \), \( \forall \ell \), (i.e. \( \hat{h}_{j,jk} \) and all its pilot contaminants.). Therefore, using the Woodburry identity followed by the block-matrix inversion lemma, we can rewrite \( \gamma_{j,k}^{M-MMSE} \) as:

\[
\gamma_{j,k}^{M-MMSE} = \left[ \begin{array}{c} 1 \\ \hat{H}_{j,k} \hat{B}^{-1}_{j,/k} \hat{H}_{j,k} + \begin{bmatrix} 0 & 0_{1 \times (L-1)} \\ 0_{(L-1) \times 1} & I_{(L-1)} \end{bmatrix} \end{array} \right]_{jj}^{-1}
\]  

with \( B_{j,/k} \) defined as:

\[
B_{j,/k} = \sum_{\ell=1}^{L} \sum_{i=1}^{K} \hat{h}_{j\ell i}^H \hat{h}_{j\ell i} + (Z_j^M)^{-1} \\
= \hat{H}_{j,/k} \hat{H}_{j,/k}^H + (Z_j^M)^{-1}
\]  

Then, according to the CMT, deriving a deterministic equivalent for \( \gamma_{j,k}^{M-MMSE} \) boils down to the \( L \times L \) matrix: 

\[ \hat{H}_{j,k} \hat{B}^{-1}_{j,/k} \hat{H}_{j,k}, \] 

whose \( mn \)-th entry can be rewritten using...
the Woodburry inverse as:

\[
\frac{1}{N} \hat{h}_{jmk}^{H} B_{j,k}^{-1} \hat{h}_{jnk} = \frac{1}{N} \hat{h}_{jmk}^{H} Z_{j,k} \hat{h}_{jnk} - \frac{1}{N} \hat{h}_{jmk}^{H} Z_{j,k} \hat{H}_{j,k} \left( \hat{H}_{j,k}^{H} Z_{j,k} \hat{H}_{j,k} + I_{L(K-1)} \right)^{-1} \hat{H}_{j,k} Z_{j,k} \hat{h}_{jnk}.
\]

(16.3)

Second, since \( \forall (j, \ell, k), \hat{h}_{j\ell k} \sim CN(\hat{h}_{j\ell k}, \hat{R}_{j\ell k} \Phi_{jk} \hat{R}_{j\ell k}) \), a direct application of the CMT [80] can yield an approximation of first term in (J.3), i.e.:

\[
\frac{1}{N} \hat{h}_{jmk}^{H} Z_{j} \hat{h}_{jnk} - \left\{ \frac{1}{N} \text{tr} \left( R_{jmk} \Phi_{jk} R_{jmk} Z_{j} \right) + \frac{1}{N} \hat{h}_{jmk}^{H} Z_{j} \hat{h}_{jnk} \right\} \xrightarrow{a.s.} 0.
\]

(16.4)

Third, we recall that \( \hat{H}_{j,k} \) includes all the estimates \( \hat{h}_{j\ell i}, \forall \ell \), with \( i \neq k \), thus implying that it is fully uncorrelated with \( \hat{H}_{j,k} \) defined above. Consequently, we have:

\[
\frac{1}{N} \hat{h}_{jmk}^{H} Z_{j} \hat{H}_{j,k} - \left\{ \frac{1}{N} \hat{h}_{jmk}^{H} Z_{j} \hat{H}_{j,k} \right\} \xrightarrow{a.s.} 0,
\]

(16.5)

\[
\frac{1}{N} \hat{h}_{j,k}^{H} Z_{j} \hat{H}_{j,k} - \left\{ D_{jk} + \frac{1}{N} \hat{H}_{j,k}^{H} Z_{j} \hat{H}_{j,k} \right\} \xrightarrow{a.s.} 0,
\]

(16.6)

where \( D_{jk} \) as defined in (5.25), and \( \hat{H}_{j,k} = [\hat{h}_{j\ell i} : \forall (\ell, i), i \neq k] \) (the total specular matrix of BS \( j \) aside from the \( k \)-th UEs).

Putting together the deterministic equivalents in (J.4)-(J.5) provides an approximation of \( \frac{1}{N} \hat{h}_{jmk}^{H} B_{j,k}^{-1} \hat{h}_{jnk} \). Finally, applying a reverse Woodburry on this deterministic equivalent leads to the expression \( \frac{1}{N} \gamma_{jk}^{MMSE} \) given in Theorem 13, and thus concludes the proof.
17 Publications

Journals


Conferences
