

**NOVEL SIGNALING SCHEMES TO IMPROVE THE
PERFORMANCE OF 5G CELLULAR NETWORKS AND BEYOND**

by

Abdelmalik Nasser Ali Aljalai

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2021

© Abdelmalik Nasser Ali Aljalai, 2021

The following individuals certify that they have read, and recommend to the **Faculty of Graduate and Postdoctoral Studies** for acceptance, the dissertation entitled:

**NOVEL SIGNALING SCHEMES TO IMPROVE THE PERFORMANCE OF 5G
CELLULAR NETWORKS AND BEYOND**

submitted by **Abdelmalik Nasser Ali Aljalai** in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY** in **Electrical and Computer Engineering**.

Examining Committee:

Prof. Victor C. M. Leung, Electrical and Computer Engineering, UBC

Supervisor

Prof. Chen Feng, School of Engineering, UBC

Co-Supervisor

Prof. Rabab K. Ward, Electrical and Computer Engineering, UBC

Supervisory Committee Member

Prof. Vijay K. Bhargava, Electrical and Computer Engineering, UBC

University Examiner

Prof. Hasan Cavusoglu, Business Administration, UBC

University Examiner

Abstract

Fifth Generation (5G) cellular networks will be the backbone of the telecommunications infrastructure for the next decade. Massive Multiple-Input Multiple-Output (MIMO) and Non-Orthogonal Multiple Access (NOMA) are two key technologies behind 5G that aim to make massive connectivity and green communications feasible.

This dissertation aims to improve the performance of massive MIMO and NOMA in 5G cellular networks and beyond with a particular focus on enhancing channel estimation, improving energy efficiency, and increasing the Quality-of-Service (QoS). Firstly, we tackle the well-known pilot contamination problem by developing a novel channel estimation scheme called the Dual Pilot Scheme (DPS). We show via mathematical analyses and simulations that this new scheme provides more accurate Channel State Information (CSI) and universally outperforms the conventional pilot scheme in 5G networks. Secondly, we develop the Extended Dual Pilot Scheme (EDPS) to handle both the inter-cell and intra-cell interference. Compared to state-of-the-art solutions for solving the pilot contamination problem, our DPS/EDPS are easier to integrate within the current 5G networks, while still achieving significant improvements for both massive MIMO and NOMA. Thirdly, we improve the energy efficiency in 5G systems employing the Discrete Fourier Transform-spread-Orthogonal Frequency Division Multiplexing (DFT-s-OFDM) waveform by developing a new scheme that combines DFT-s-OFDM with Barker Codes and DPS/EDPS. We show via extensive simulations that this new scheme improves energy efficiency, reduces Peak-to-Average Power Ratio (PAPR), and limits Out of Band (OOB) leakage in various realistic scenarios. Fourthly, we further enhance the QoS in 5G networks by developing a new decoding

scheme for uplink NOMA based on the Compute-and-Forward (CaF) framework. We show that this scheme achieves better fairness and smaller outage probabilities, while essentially keeping the same complexity as the conventional Successive Interference Cancellation (SIC) decoding. Finally, we enhance the performance of Integer-Forcing Linear Receiver (IFLR) for massive MIMO-NOMA by combining DPS/EDPS with CaF decoding to mitigate the imperfect CSI and lower the CaF sensitivity to estimation errors.

Overall, we demonstrate that the novel schemes proposed in this dissertation will improve the performance, provide valuable tools for tackling real-world technical problems, and enhance operations of 5G cellular networks and beyond.

Lay Summary

The tremendous and rapid growth in mobile devices, wireless sensors, and other intelligent devices that require wireless connectivity is becoming a significant challenge for cellular networks. To address the demands of the resulting massive wireless connectivity and support the exchange of colossal volumes of data traffic, the 5G cellular network has been envisioned. In this dissertation, we investigate two enabling technologies for 5G: Massive MIMO and NOMA. In particular, we have developed several novel schemes that are able to obtain better channel estimation, improve energy efficiency, enhance the quality of service, and increase the robustness and reliability of 5G technologies and user equipment. Through mathematical analyses and extensive simulations, we demonstrate that our proposed novel schemes offer significant advantages over conventional schemes in various realistic settings of massive MIMO and NOMA. The contributions of this dissertation will reshape the 5G technologies and facilitate in creating technical efficiencies to 5G cellular networks and beyond.

Preface

The work in this dissertation was completed under the joint supervision of Prof. Victor C. M. Leung and Prof. Rabab K. Ward in the Department of Electrical and Computer Engineering as well as Prof. Chen Feng in the School of Engineering at the University of British Columbia, Canada.

The material presented in all the chapters of this dissertation as well as the corresponding publications are based on my ideas that I developed and the research conducted by myself under the guidance with intensive discussions of my supervisors. I carried out the literature review, identified the research gaps, found the rooms of contributions, formulated the research problems, and developed the proposed solutions. The mathematical analyses, the assessments of the proposed solutions, and writing the manuscripts were carried out by me. My supervisors helped me by guiding the direction of my researches, to validate the analytical and simulation results, and by providing feedback to improve the presentation of the manuscripts with advising me on the style of writing in our academic field.

A co-author of published contribution in a publication related to chapter 4, Cheng Guo collaborated with me on the topic of NOMA. Inspired by the duality principle of DPS, we brought and exploited the CaF decoding scheme to use it for enhancing the NOMA uplink transmission.

Below is a list of the journal and conference publications related to the work presented in this dissertation.

Publications and Patent Related to Chapter 2:

- Abdelmalik Nasser Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, “Eliminating pilot contamination using Dual Pilot Sequences in Massive MIMO,” in IEEE 86th Vehicular Technology Conference VTC-Fall, Toronto, ON, Canada, 24-27 Sept. 2017, pp. 1–6.

- Abdelmalik Nasser Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, “Dual Pilot Scheme (DPS) and Its Application in Massive MIMO,” in *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1431-1444, November 2020.

Furthermore, the work in this dissertation resulted in filing a provisional patent to University Industry Lesion Office (UILO) at UBC and the Canadian Intellectual Property Office (CIPO).

- Abdelmalik Nasser Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, “The Dual Pilot Sequences Scheme for wireless Communications Systems,” filed on February 2017 to UBC-UILO office, accepted at UBC-UILO office as a provisional patent in September 2017, (Ref. # UBC-ID 18-047) and accepted at CIPO office as a provisional patent in October 2019.

Publications Related to Chapter 3:

- Abdelmalik Nasser Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, “Improving the Energy Efficiency of DFT-s-OFDM in Uplink Massive MIMO with Barker Codes,” in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, Big Island, Hawaii, USA, 17-20 Feb. 2020, pp. 731-735.
- Abdelmalik Nasser Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, “Improving Energy Efficiency of 5G Uplinks Using Barker Codes in DFT-s-OFDM,” Submitted to *IEEE Transactions on Signal Processing*, pp. 1–15, December 2020.

Publications Related to Chapter 4:

- C. Guo*, Abdelmalik Nasser Aljalai*, C. Feng, L. Zhao, V. C. M. Leung, and R. K. Ward, “Compute-and-Forward for Uplink Non-Orthogonal Multiple Access,” *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 986-989, December 2018. (*equal contribution)
- Abdelmalik Nasser Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, “Integer-Forcing Linear Receiver for Uplink Massive MIMO-NOMA with Imperfect CSI,” to be submitted.

Table of Contents

Abstract	iii
Lay Summary	v
Preface	vi
Table of Contents	viii
List of Tables	xiii
List of Figures	xiv
List of Abbreviations	xvii
Notations	xxii
Acknowledgments	xxv
Dedication	xxviii
Chapter 1: Introduction and Overview	1
1.1 Background and Overview	1
1.2 Dissertation Objectives	3
1.3 Research Motivations and Challenges	3
1.3.1 Imperfect CSI	3

Table of Contents

1.3.2	Massive Multiple Input Multiple Output	8
1.3.3	Non-Orthogonal Multiple Access	12
1.4	Literature Review of Related Works and Research Goals	13
1.5	Dissertation Contributions	23
1.6	Dissertation Outline and Organization	26
Chapter 2:	Dual Pilot Scheme and its Application in Massive MIMO	30
2.1	Synopsis	30
2.2	Introduction	31
2.3	Related Work	33
2.4	System Model	36
2.5	Dual Pilot Scheme	39
2.5.1	SPS with Orthogonal Pilots	40
2.5.2	DPS with Non-Orthogonal Pilots	40
2.5.3	Characterizing the Advantages of DPS	44
2.6	Correlated Rayleigh Fading	46
2.6.1	SPS with Orthogonal Pilots	47
2.6.2	DPS with Non-Orthogonal Pilots	47
2.6.3	Characterizing the Advantages of DPS	47
2.7	DPS in Massive MIMO Cellular Networks: Macroscopic View	49
2.7.1	Integration Plan of DPS into Massive MIMO	51
2.8	Additional Desirable Features of DPS	55
2.8.1	Interference Alignment and Interference Cancellation	56
2.8.2	Smart Pilot Assignment and User Scheduling	56
2.8.3	Periodically Updated the Second-Order Statistics	57
2.9	Numerical Results and Discussions	57
2.9.1	Normalized Mean Squared Error	59
2.9.2	Achievable Sum-Rate	59

2.9.3	Singular Values Spread	62
2.10	Conclusions	64
Chapter 3:	Barker Coded DFT-s-OFDM Waveforms for 5G Uplinks	66
3.1	Synopsis	66
3.2	Introduction	67
3.3	The Energy Efficiency Metrics	72
3.4	Barker Codes	73
3.5	System Model	76
3.5.1	Transmitter Configuration at the UE	78
3.5.2	The Base Station Receiver Configuration	80
3.6	NOMA Deployment in Massive MIMO	83
3.6.1	PD-NOMA Integration in Massive MIMO with Perfect CSI	83
3.6.2	PD-NOMA Integration in Massive MIMO with Imperfect CSI	87
3.7	Numerical Results and Discussion	94
3.7.1	Generating the Channel Matrix	94
3.7.2	Numerical Results for OMA Massive MIMO Cell Configuration	95
3.7.3	Numerical Results for PD-NOMA Deployment in Massive MIMO	101
3.7.4	Peak to Average Power Ratio	104
3.7.5	Out of Band Leakage	106
3.8	Conclusion	106
Chapter 4:	Compute-and-Forward for NOMA and Massive MIMO-NOMA Uplinks	108
4.1	Synopsis	108
4.2	Introduction	109
4.3	SIC-NOMA	111
4.4	CaF-NOMA	112
4.5	CaF in Integer-Forcing Linear Receiver	116

Table of Contents

4.5.1	Low Density Construction-A Lattices	116
4.5.2	Construction of a Linear Labeling	117
4.5.3	Integer Forcing Linear Receiver via Nested Lattices	119
4.5.4	Extension to the Complex Field	122
4.6	CaF for Massive MIMO-NOMA	124
4.6.1	Integer Forcing Linear Receiver for Massive MIMO-NOMA	124
4.7	The Role of Imperfect CSI in Massive MIMO-NOMA with CaF Decoding	127
4.7.1	Effects of Channel Estimation Errors	128
4.7.2	Employing DPS and EDPS to Mitigate Channel Estimation Errors	135
4.8	Numerical Results	135
4.8.1	CaF-NOMA with Perfect CSI	135
4.8.2	CaF-NOMA in Massive MIMO with Imperfect CSI	137
4.9	Conclusion	142
Chapter 5:	Conclusions and Future Works	144
5.1	Concluding Remarks	144
5.1.1	Summaries and Contributions	145
5.2	Future Works	148
5.2.1	Expected Future Extensions	148
Bibliography	152
Appendix A:	Related Proofs of Chapter 2	164
A.1	Derivation of Equation (2.8)	164
A.2	Derivation of Inequality (2.24)	165
A.3	Derivative of $g_{lk}^{\text{SUB}}(\gamma)$	166
Appendix B:	Related Proofs and Preliminaries of Chapter 4	168
B.1	Proof of Theorem 4	168

Table of Contents

B.2 Preliminaries for Lattices and Nested Lattice Codes.	171
B.3 Proof of Theorem 5	173
B.4 Proof of Theorem 6	174

List of Tables

Table 2.1	Dual PSs assignments and comparisons of total numbers of UEs served in a system-cluster for SPS (UC^{SPS}) and DPS (UC^{DPS}) with $\tau = N = K$ and $L = 7$ at different values of M . Where M and N are the numbers of BS's antennas and available orthogonal sequences, respectively. $N - L$ and ID are the numbers of repeatedly reused sequences (User-IDs) and Cell-IDs sequences in the system-cluster, respectively.	41
Table 2.2	The parameters used in simulations	58
Table 2.3	Comparisons of cell capacity C_{sum}^l in bps/Hz , for DPS and SPS at a given values of NMSE in (dB), for correlated Rayleigh fading channel when $M = 128$, $K = 8$, and $L = 7$	62
Table 3.1	System parameters used in simulations	95
Table 3.2	NMSE Performance metric in dB for DFT-s-OFDM and BC DFT-s-OFDM in PD-NOMA at different values of SNR with various BC length, using i.i.d. and correlated Rayleigh fading channels at $M = 128$ and active UE= 32.	102
Table 3.3	Cell's sum-rate performance metric in bps/Hz for DFT-s-OFDM and BC DFT-s-OFDM in PD-NOMA at different values of SNR with various BC length, using i.i.d. and correlated Rayleigh fading channels at $M = 128$ and active UE= 32.	103
Table 3.4	BER Performance metric for DFT-s-OFDM and BC DFT-s-OFDM in PD-NOMA at different values of SNR with various BC length, using i.i.d. and correlated Rayleigh fading channels at $M = 128$ and active UE= 32.	104

List of Figures

Figure 1.1	Expected data traffic volume and number of subscribers to be handled by 5G cellular networks [1]	2
Figure 1.2	Green communications and energy efficiency	23
Figure 1.3	A schematic view of the dissertation’s contributions.	26
Figure 1.4	A schematic view of the dissertation structure and relationship among chapters.	29
Figure 2.1	A cluster of seven non-cooperative adjacent cells with Cell-IDs assignment for DPS, the home cell is located in the center. $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_7\}$ are the assigned Cell-IDs sequences.	39
Figure 2.2	Performance comparisons between DPS and some exiting schemes for pilot contamination mitigation.	51
Figure 2.3	Dual Pilot Scheme with Cell-IDs and cluster reuse. Cells with the same color are using the same unique BS’s dual PS identifications (i.e., the same Cell-ID and User-IDs).	53
Figure 2.4	DPS Structural integration for a cluster of two non-cooperative adjacent cells where the BS is equipped with one antenna and serves K UEs.	55
Figure 2.5	The Normalized Mean Squared Error at $M = 32$ and using the correlated Rayleigh fading channel.	60
Figure 2.6	The Normalized Mean Squared Error at $M = 32$ and using the i.i.d. Rayleigh fading channel.	60

Figure 2.7	The achievable ergodic sum-rate per cell at $M = 32$ and using the correlated Rayleigh fading channel.	61
Figure 2.8	The Cumulative Distribution Function of the Singular Values Spread when the correlated Rayleigh fading channel is used.	63
Figure 2.9	The Cumulative Distribution Function of the Singular Values Spread when the i.i.d. Rayleigh fading channel is used.	63
Figure 3.1	Payload data symbols comparison in OFDM, localized DFT-s-OFDM, and distributed DFT-s-OFDM, when 12 sub-carriers are used.	68
Figure 3.2	Structures of BC DFT-s-OFDM when Barker code length=3, $b_i = \{+1, -1\}$, and 12 sub-carriers are used.	75
Figure 3.3	Single cell massive MIMO, the BS has M antennas and serves K users.	77
Figure 3.4	The proposed transmitter structural configuration for the k th UE.	79
Figure 3.5	The proposed structural configuration of the BS receiver.	82
Figure 3.6	Two cells of massive MIMO with PD-NOMA deployment, each cell is divided into two regions ($R=2$), each region has a set of K active users. The BS is equipped with M antennas and located at the center of each cell.	84
Figure 3.7	EDPS cluster reuse in massive MIMO PD-NOMA cellular network. Cells with the same color are using the same unique BS's extended dual PSs (i.e., the same Cell-ID (MCSG), same Region-IDs (NRSGs), and User-IDs (UESGs)).	94
Figure 3.8	The CSI-NMSE performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.	97
Figure 3.9	The uplink sum-rate capacity performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.	98
Figure 3.10	The BER performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.	99
Figure 3.11	The Energy Efficiency performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.	100

Figure 3.12	The CCDF of PAPR in a massive MIMO cell with PD-NOMA deployment using 32-QAM at different configurations.	105
Figure 3.13	The Out of Band Leakage for DFT-s-OFDM, G DFT-s-OFDM, and BC DFT-s-OFDM waveforms.	106
Figure 4.1	The uplink NOMA with more far users than near users.	111
Figure 4.2	<i>CaF</i> decoding offers two additional choices, namely, CaF_1 and CaF_2	113
Figure 4.3	(R_k, R_l) is better than (R'_k, R'_l) , and (R_k, R_l) is Pareto-optimal. The MAC region is the capacity region for multiple access channels.	114
Figure 4.4	The uplink of Massive MIMO-NOMA with more far users than near users. . .	125
Figure 4.5	Block diagram illustrating <i>CaF</i> decoding in a single antenna BS receiver.	128
Figure 4.6	The fairness comparison of SIC-NOMA and <i>CaF</i> -NOMA with different $\frac{D_1}{D_2}$. . .	136
Figure 4.7	The average outage probability comparison of SIC-NOMA and <i>CaF</i> -NOMA with different $\frac{D_1}{D_2}$ and $R_t = 2$ BPCU.	137
Figure 4.8	Gaussian MAC region for massive MIMO-NOMA IFLR, SIC corner points, and <i>CaF</i> dominant rate pairs when $M=16$	138
Figure 4.9	Comparison of achieved Gaussian MAC region for massive MIMO-NOMA in IFLR, at perfect CSI, using DPS/EDPS, and using SPS.	139
Figure 4.10	Comparison of achieved Gaussian MAC region for massive MIMO-NOMA in IFLR in a noisy environment, at perfect CSI, using EDPS/DPS, and using SPS.	139
Figure 4.11	Performance comparisons of the average sum-rate loss between EDPS/DPS and SPS over a range of $\sigma_{\epsilon_\alpha}^2$	140
Figure 4.12	Performance comparisons of the average sum-rate loss between EDPS/DPS and SPS over a range of $\sigma_{\epsilon_\beta}^2$	141
Figure 4.13	Performance comparisons of the average sum-rate loss between EDPS/DPS and SPS over a range of $\sigma_{\epsilon_c}^2$	141
Figure 5.1	Superimposed and multiplexed pilots configurations in TDD.	150

List of Abbreviations

2D	two dimensional
3D	three dimensional
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
AP	Access Point
AWGN	Additive White Gaussian Noise
BBS	Binary Barker Sequence
BC	Barker Code
BC DFT – s – OFDM	Barker Coded DFT-s-OFDM
BER	Bit-Error Rate
BPCU	Bits Per Channel Use
BS	Base Station
CaF	Compute-and-Forward
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CD – NOMA	Code Domain NOMA

List of Abbreviations

Cell – ID	Cell Identification
CMT	Cosine Modulated multi-Tone
CIR	Channel Impulse Response
CP	Cyclic Prefix
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DFT	Discrete Fourier Transform
DFT – s – OFDM	Discrete Fourier Transform -spread- OFDM
DPS	Dual Pilot Scheme
EDPS	Extended Dual Pilot Scheme
EE	Energy Efficiency
eMBB	Enhanced Mobile Broadband
FDD	Frequency Division Duplexing
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
G DFT – s – OFDM	Generalized DFT-s-OFDM
Hz	Hertz
ICI	Inter-Carrier Interference
ICT	Information and Communications Technologies
ID	Identification
IFFT	Inverse Fast Fourier Transform

List of Abbreviations

IFLR	Integer-Forcing Linear Receiver
i.i.d.	independent and identically distributed
IoT	Internet of Things
ISI	Inter-Symbol Interference
ITU	International Telecommunication Union
JD	Joint Detection
JT	Joint Transmission
LDPC	Low Density Parity Check
LMMSE	Linear MMSE
LTE	Long-Term Evolution
LTE – A	LTE-Advanced
MAC	Multiple Access Channel
MCC	Mission Critical Communications
MCSG	massive MIMO Cells-Clustering Sequences Group
MF	Merit Factor
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Squared Error
MRC	Maximum Ratio Combining
MSE	Mean Squared Error
mMTC	massive Machine Type Communications
MU	Multi-User
MU – MIMO	Multi-User MIMO
MUST	Multi-User Superposition Transmission

List of Abbreviations

NMSE	Normalized Mean Squared Error
NOMA	Non-Orthogonal Multiple Access
NRSG	NOMA Regions Sequences Group
OFDM	Orthogonal Frequency Division Multiplexing
OMA	Orthogonal Multiple Access
OOB	Out of Band
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
PD – NOMA	Power-Domain NOMA
PSD	Power Spectral Density
PS	Pilot Sequence
QAM	Quadrature Amplitude Modulation
QBS	Quaternary Barker Sequence
QoS	Quality-of-Service
RAN	Radio Access Network
RF	Radio Frequency
RX	Receiver
SE	Spectral Efficiency
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise-Ratio
SNR	Signal-to-Noise-Ratio
SPS	Single Pilot Scheme
SVD	Singular Value Decomposition

List of Abbreviations

SVS	Singular Value Spread
TDD	Time Division Duplexing
TX	Transmitter
UE	User Equipment
User – ID	User-Equipment Identification
UESG	Users Equipment Sequences Group
uRLLC	ultra-Reliable Low Latency Communications
UW DFT – s – OFDM	Unique-word DFT-s-OFDM
WPTN	Wireless Power Transfer Networks
ZF	Zero-Forcing
ZT DFT – s – OFDM	Zero-tail DFT-s-OFDM

Notations

a	Scalar, lower case
\mathbf{a}	Vector, bold lower case
\mathbf{A}	Matrix, bold upper case
$(\cdot)^T$	Transpose operation of a vector/matrix
$(\cdot)^*$	Complex conjugate operation
$(\cdot)_*$	Optimal value
$(\cdot)^H$	Complex conjugate transpose (Hermitian) operation of a vector/matrix
$(\cdot)^{-1}$	The inverse operation
$\ \cdot\ _2$	Second (Euclidean) norm
$ a , \mathbf{a} , \mathbf{A} $	The absolute value of a , the magnitude of \mathbf{a} , the determinant of \mathbf{A}
$\lceil \cdot \rceil$	Ceiling function
$tr(\cdot)$	Trace of a matrix
$\text{diag}(P_1, \dots, P_K)$	A diagonal matrix with size $K \times K$ and diagonal entries are P_1, \dots, P_K
$\mathbb{E}\{\cdot\}$	Statistical expectation operator
$\text{vec}(\cdot)$	Vectorization
\otimes	Kronecker matrix product-Tensor vector product
\mathbf{I}_K	Identity matrix of size $K \times K$
$\mathcal{CN}(\mu, \sigma^2)$	A complex Gaussian distribution with mean value μ and variance σ^2 .
\rightarrow	A linear map operator

\Leftrightarrow	Logical equality operator (i.e., if and only if)
$(\cdot)^{Inv}$	An inverse linear mapping
$\mathcal{L}(\Lambda_c, \Lambda_f)$	A <i>Nested Lattice</i> that represents a pair of Lattices (Λ_c, Λ_f)
$\max(\cdot)$	Maximum
$\min(\cdot)$	Minimum
mod	Modulo operator
$\text{round}(\cdot)$	Round operation over Gaussian integers
$\text{gcd}(\cdot)$	Greatest common divisor
\gg	Much greater than
\approx	Approximately equal to
$\mathbb{C}^{n \times m}$	Field of complex numbers with $n \times m$ dimensions
$\mathbb{R}^{n \times m}$	Field of real numbers with $n \times m$ dimensions
$\mathbb{Z}^{n \times m}$	Field of integer numbers with $n \times m$ dimensions
\mathbb{Z}_p	The p-Adic integers
$A \cup B$	Union of the sets A and B
$A \cap B$	Intersection of the sets A and B
$A \subset B$	A is a proper subset of B
$a \in A$	Element a belongs to the set A
(a, b)	Open interval between a and b
$[a, b]$	Closed interval between a and b
$f(\alpha, \beta \gamma)$	f is a function of α and β with given of γ as a sided information
\triangleq	By definition
\forall	For all
\max_i	Maximum with respect to i
\min_i	Minimum with respect to i
arg min	Argument that attains the minimum
arg max	Argument that attains the maximum

Notations

\bar{D}	Average of D
\hat{D}	Estimated value of D
\tilde{D}	Approximated value of D
\log_2	Logarithmic operator with base 2
$\log_2^+(\cdot)$	The maximum of $[0, \log_2(\cdot)]$.
\log_{10}	Logarithmic operator with base 10
$\Re\{\cdot\}$	Real part of a complex number
$\Im\{\cdot\}$	Imaginary part of a complex number

Acknowledgments

First and foremost, I would like to thank Allah, the Almighty, the Greatest for his countless blessings and for giving me the persistence, patience, knowledge, inspiration, strength, and energy to pursue the researches in this dissertation while I faced many obstacles and difficult conditions. The journey of the Ph.D. itself is understandably stressful for anyone pursuing it. However, my Ph.D. journey was extremely not easy, there were indefinite stress and huge challenges I faced. On top of the list, this Ph.D. started with the beginning of the aggressive war on my home country, which prevented me from seeing my family until now. This aggressive war resulted in many side/proxy wars, humanitarian crises, destroying infrastructures, and hidden occupations. The COVID-19 pandemic also had a profound impact on all of humanity in 2020.

I would like to express my profound thanks, sincere gratitude, and appreciation to my advisor Prof. Victor C. M. Leung. Prof. Leung taught Mobile Communications Networks, where I got the highest mark in the class and was one of the luckiest students to gain his magnificent knowledge and accumulated experience. Special profound thanks to my co-advisors Prof. Victor C. M. Leung, Prof. Rabab K. Ward, and Prof. Chen Feng for giving me the support and funding during my Ph.D. Their vast knowledge, magnificent experience, fruitful discussions, invaluable advice, continuous encouragement, ethical framework, and perseverance are truly inspiring, which made this endeavor productive and enriching. Their thoughtfulness, dedication, patience, and guidance helped me tremendously and have made this dissertation a reality. I gained many helpful insights and analytical perspectives from them. It has been a pleasure working with such supportive and accommodating advisors and I will follow and emulate their insightful approach to research

Acknowledgments

throughout my research and development career.

I would like to express my sincere thanks to Prof. Hong-Chuan Yang at the University of Victoria for serving as the external examiner. The time and efforts Prof. Yang dedicated to evaluating my work and providing feedback and constructive suggestions are greatly appreciated. Also, I am profoundly thankful to the university examiners in my doctoral examining committee at UBC - Prof. Vijay K. Bhargava and Prof. Hasan Cavusoglu - for their valuable time and efforts in evaluating my work. Besides, I am deeply thankful to Prof. Z. Jane Wang, the ECE department examiner, for her time and efforts in evaluating my work and providing insightful suggestions. My appreciation and thanks extend to Prof. William Hsieh - the chair of the examining committee- for his time and efforts. I am very honored to have had all of these distinguished professors on my examining committee.

Special and unique profound thanks to Prof. Susan Porter, the Dean and Vice-Provost of Graduate and Post Doctoral Studies at UBC for her unequivocal support. She is the first human rights and student rights defender at UBC who fought against corruption, a prime example of good governance, instrumental in advancing graduate studies into pioneer futuristic levels, while fighting racial discrimination and ensuring students' rights and wellbeing. I will always remember her as the shiny and bright face of UBC.

Unlimitedly and magnificently, I would like to express my utmost gratitude and profound thanks to my beloved mother, my siblings, my nephews, and my nieces for keeping their faith and trust in me, which has been my ultimate motivation and energized me to persevere during my Ph.D studies. I am deeply indebted for their endless and unconditional love and emotional support during my alone stay in Vancouver.

I am thankful to the Natural Science and Engineering Research Council of Canada (NSERC), the University of British Columbia (UBC), and the Islamic Development Bank (IsDB) for supporting this research partially through the Alexander Graham Bell Canada Graduate Scholarship, the Four Year Doctoral Fellowship, and the Merit Scholarship. In addition, I gratefully acknowledge the generous awards that I received: the President's Academic Excellence Initiative (PAEI) Ph.D.

Acknowledgments

award from UBC, the Public Scholarship Initiative (PSI) award from the Faculty of Graduate and Post-Doctoral Studies, the Graduate Support Initiative (GSI) award from the Department of Electrical and Computer Engineering (ECE), and the Teaching as Research (TAR) award from the Center for the Integration of Research, Teaching, and Learning (CIRTL) at UBC.

My thankfulness extends to my colleague Cheng Guo who was a visiting Ph.D. student in Prof. Feng's research group. Cheng – thank you for the valuable and constructive discussions we had, which resulted in a fruitful collaboration. Finally, I would like to express my sincere gratitude and appreciation for my friends Dr. Samer Al-Kiswany, Marta Lubaszewski, Alina Kunitskaya, and Dr. Mohamed Elgendi for their continued help and support.

Throughout my journey, there were hidden soldiers at UBC and around the globe who have been standing for justice and provided a shoulder to lean on when I needed them. I cannot thank you all enough. Whatever I will write in words would fall short to express my gratitude for their availability, courageous standing, sacrifice, and unlimited support. All of you will always be a source of light to brighten my path in life. I will never forget!

Palestine was, is still, and forever will be an Arabic land under Arabic Sovereignty

Jerusalem (Al-Quds) is the Permanent Capital of Palestine

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dedication

TO MY MOTHER

TO THE SOUL OF MY FATHER

TO MY ADVISOR PROF. VICTOR C.M. LEUNG

TO PROF. PULAN PORTER

TO MY SIBLINGS, NEPHEWS, AND NIECES

TO THE PEOPLE WHO I CARE ABOUT

TO "YEMEN"

TO THOSE WHO MARTYRED IN WARF THEY DID

NOT INVOLVE IN

TO THE HONORABLE FREE PEOPLE IN THE

WORLD

TO THOSE WHO ARE DEFENDING THEIR

VALUES, BELIEVES, ETHICS, MORALITY,

AND STAND FOR HUMAN RIGHTS

AND DIGNITY

ABDELMALIK...

Chapter 1

Introduction and Overview

1.1 Background and Overview

In next generation (Fifth Generation (5G) and beyond) wireless cellular communication networks, extremely large volumes of data will be exchanged through the cellular networks between large numbers of users, sensors, and machines. Recently, there is increasing interest in the integration of high-density wireless networks, such as Internet of Things (IoT). As a result, high demand for the technical infrastructure to deploy 5G cellular networks continues to grow. World-wide data traffic is growing at an unprecedented rate, with an average annual rate of 170% over 2018 and 2019, according to the Ericsson's biannual Mobile-Traffic Report [1]. Furthermore, with this explosive increase in demands on the data traffic, it is expected that approximately 4 billion additional mobile devices and wireless sensors will need to be served by cellular networks by 2025. Figure 1.1 shows these expectations graphically. Therefore, 5G and beyond Radio Access Network (RAN) should address the goals for achieving high capacity, exponential traffic growth, high spectral efficiency, and the increasing demand for high bandwidth with very low latency while offering enormous device connectivity in an energy-efficient manner.

Thus, 5G cellular network targets are to increase 1) the system connecting capacity by 1000 times, 2) the average cell throughput by 25 up to 100 times, and 3) the spectral efficiency by 10 times [2–5]. Furthermore, it is expected that the number of cells in a 5G cellular network will

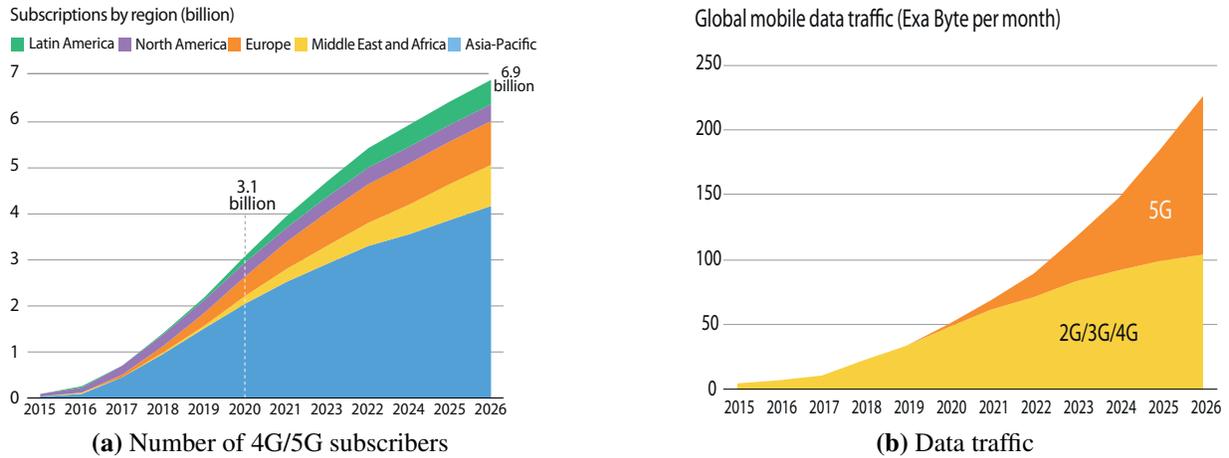


Figure 1.1: Expected data traffic volume and number of subscribers to be handled by 5G cellular networks [1]

increase significantly due to the high density of served users, which requires a shrinking in the cell coverage within a geographical area to serve as many users as possible. One of the requirements of 5G cellular networks is to offer enormous connectivity with reliable performance. Achieving this presents extreme challenges when designing 5G cellular networks. As a result, 5G solutions for cellular networks form a portfolio of technologies that will not be exclusive to one specific access or connectivity technology. Massive Multiple-Input Multiple-Output (MIMO) and Non-Orthogonal Multiple Access (NOMA) are the two promising candidates technologies for the 5G cellular networks deployment [6] which attracted a lot of attention. Massive MIMO, also called very large scale and full dimensions MIMO, has high capacity, spectral efficiency, and energy efficiency. Massive MIMO is the most feasible technology that offers a massive access connectivity [7]. On the other hand, NOMA is another 5G candidate technology which increases the access connectivity that can serve multiple users at the same time and frequency band to achieve low transmission latency, high spectral efficiency, and offer high capacity.

In mid 2019 Ericsson announced the commercializing of 5G new-radio for massive MIMO and the first deployment for testing started in 2019 with 64 antennas at the Base Station (BS) and supporting a variety of configurations including NOMA integration [8]. Massive MIMO and NOMA are still open research areas with capacity for further improvements. This Ph.D. dissertation ad-

dresses the critical limitations and impairments that affect the performance of these technologies with the aim to improve the performance of uplinks and downlinks, at the BS and the User Equipment (UE), that will result in an improvement to the overall Quality-of-Service (QoS), system capacity, and bring 5G targets a reality.

1.2 Dissertation Objectives

The ultimate objective of the dissertation is to enhance the cellular networks equipped with 5G technologies of massive MIMO and NOMA to ensure optimal performance and achieve the expectations targets of 5G. To this end, the overall users' traffic exchanged in the cellular network has been analyzed in both Orthogonal Multiple Access (OMA) massive MIMO and massive MIMO with NOMA deployment.

As a first step, this dissertation models, analyzes, and proposes novel improvements to the pilot signaling scheme used for channel estimation to obtain Channel State Information (CSI), in clusters consisting of groups of non-cooperative multiple adjacent cells. The performance of these architectures of the cellular networks have been evaluated mathematically and numerically using simulations to mimic the real environment scenarios. Furthermore, this dissertation investigates employing Barker codes to improve energy efficiency in 5G cellular networks and beyond. In addition, it examines the application of the Compute-and-Forward (CaF) method to improve the performance of NOMA technology in 5G and beyond. Lastly, it further explores the enhancement of Integer-Forcing Linear Receiver (IFLR) in massive MIMO-NOMA.

1.3 Research Motivations and Challenges

1.3.1 Imperfect CSI

Wireless communication channels are continuously changing over time and frequency, and are regularly subject to fading effects. The fading effects arise due to scattering, refraction, and diffraction of the transmitted signal due to obstacles and sharp edges in the propagation environment. CSI describes how these combined factors are represented and how the propagation of the signals be-

tween the transmitter and receiver is impacted by their resultant effects. Also, these factors cause the propagation environment to exhibit multipaths propagation and reception of multiple copies of the transmitted signal at the receiver. The reception of multiple copies of the transmitted signal causes fluctuations in its amplitude. This causes significant challenges in building reliable wireless communications.

CSI can be classified into two types, Channel State Information at the Receiver (CSIR) and Channel State Information at the Transmitter (CSIT), where CSIR in the uplink transmission is acquired at the BS. CSIR depends on the system duplexing mode that uses either Time Division Duplexing (TDD) or Frequency Division Duplexing (FDD) and the channel estimation approaches are varied accordingly. Accurate estimation of the communication channel is essential to obtain an accurate CSI. Accurate CSI is vital to carry out any required compensation to remove undesired effects, and precisely extract the originally transmitted signals. The performance of any communication system depends on its communication channel estimation accuracy. Therefore, the system's performance degrades significantly as the accuracy of channel estimation decreases.

Two main approaches are used for channel estimation: pilot aided and non-pilot aided. The pilot aided approach tends to be more practical and has become more popular, while the non-pilot aided approach suffers from higher complexity and has implementation limitations. In the pilot aided approach, it has been shown that the optimal number of pilot symbols in uplink transmission at each coherence interval is equal to the number of mobile station antennas under optimum power allocation for data transmission and pilot signaling transmission [9]. However, pilot based estimators require additional pilots for enhancing the CSI accuracy [10]. More pilot symbols in the coherence interval result in reduced spectral efficiency, (coherence interval (I_c) = coherence bandwidth (B_c) \times coherence time (T_c)). There is an unavoidable trade-off between how many and how often pilot symbols are used to estimate the CSI and both the accuracy of CSI and spectral efficiency. Therefore, in [10] an iterative space-alternating generalized expectation maximization estimator was proposed which uses data symbols instead of additional pilots to increase the CSI accuracy. However, this scheme increases the complexity of the BS estimator and requires two stages

of operations. Designing and using efficient pilot sequences in channel estimation for obtaining an accurate CSI with small overheads will lead to a smaller Normalized Mean Squared Error (NMSE) of the channel estimation and improved performance; this is a main goal of this dissertation.

Another approach is to use blind estimation by employing the Singular Value Decomposition (SVD), which also encounters computational and implementation complexities. Conventionally, to overcome the complexity of SVD at UE and its practical limits, it requires pre-knowledge of CSI at both Transmitter (TX) and Receiver (RX) and cooperative signal processing for all RX antennas and TX antennas respectively. The BS performs Joint Detection (JD) in the uplink and Joint Transmission (JT) in the downlink. Both JD and JT require accurate CSI and any degradation in the accuracy of CSI will lead to performance deterioration in the overall system, especially at low Signal-to-Noise-Ratio (SNR) [11]. The degradation gap in the performance increases as the number of BS antennas increases. Theoretically, we may increase the transmission power to compensate the reduction due to imperfect CSI but practically, the transmission power should be kept within certain constraints. A different approach was proposed in [12], where the CSI in massive MIMO system had been estimated based on applying the sparse Bayesian learning method. However, in addition to the complexity of the approach, the sparsity property was not enough as a condition, and the incoherence property is required to have a satisfying performance and valid assumptions.

In systems that employ TDD, there is channel reciprocity between uplink and downlink. Therefore, the downlink instantaneous CSI is obtained by estimating the uplink instantaneous CSI. However, to avoid any CSI error, a calibration for hardware differences for the uplink and downlink Radio Frequency (RF) chains is necessary. Thus, CSI imperfection can happen at the BS's receiver due to several reasons like RF chains mismatching. Also, the time delay between uplink channel estimation and downlink transmission may be greater than the coherence time of the channel. This leads to a random mismatch between uplink and downlink CSI accuracy. The effects of imperfect channel reciprocity (due to random mismatch and hardware differences in downlink and uplink channels in the RF chains) in TDD massive MIMO on CSI accuracy have been modeled and investigated in [13]. It has been found that this results in significant degradation in the CSI accuracy.

This consequently results in a reduction in the performance of linear pre-coding and the overall system performance, which in turn reduces the Signal-to-Interference-plus-Noise-Ratio (SINR) of the received signals. It has been also found that the channel estimation error is amplified by the reciprocity error [13]. On other hand, as the CSI will be used in the downlink precoder, it will determine the allocated directive antennas that should be used. This relates to the power allocations per antenna and affects the energy efficiency. More specifically, CSI inaccuracy may lead to the incorrect use of antennas (i.e., pointing in wrong direction or use of extra antennas which may cause interference and waste radiation power). Therefore, removing the Multi-User (MU) interference in a Multi-User MIMO (MU-MIMO) system will be impossible in the presence of the imperfect CSI due to the mismatch between the exact CSI and the pre-coder. In this Ph.D. dissertation, we propose Dual Pilot Scheme (DPS), which performs the uplink estimation in a very short time and ensures that the delay between the uplink estimation and the downlink transmission is less than the coherence time. This leads to improvements in the accuracy of CSI of the downlink pre-coders.

Furthermore, in TDD massive MIMO system, Orthogonal Frequency Division Multiplexing (OFDM) is adopted to mitigate the Inter-Symbol Interference (ISI) and exploits the space-frequency diversities; however, this is based on using multiple sub-carriers. Therefore, the co-channel interference (or Inter-Carrier Interference (ICI)) has a high probability of occurrence when the CSI lose their accuracy. On the other hand, there exists an inherent source that contributes to the CSI inaccuracy due to hardware impairments. The hardware impairments also contribute to the channel estimation errors and increase the inaccuracy of CSI. For example, the lack of perfect compensation in phase noise, imbalances between in-phase and quadrature signals at the front end of the analogue processing at the baseband, ICI, and inter-modulation due to the non-linearity in amplifiers may result in an inaccurate CSI [14]. Although some techniques and algorithms are used to compensate and mitigate these impairments, these deterioration effects can not be removed completely [15].

Imperfect CSI leads to a mismatch in decoding at the BS receiver [16]. An imperfect CSI significantly affects the data decoding of the uplink transmission, and consequently, the downlink data transmission and beam-forming as TDD adopts the channel reciprocity. Jose *et al.* have shown

the impact of imperfect CSI on the precoding of the downlink transmission [17]. Therefore, as the channel estimation error increases, the ergodic capacity decreases while the bit error rate, symbol error rate, and the outage probability increases. Outage probability occurs when the probability of the instantaneous SINR falls below a certain threshold. The capacity gain offered by the MIMO communication system is sensitive to the CSI errors where imperfect CSI results in reduced capacity gain. The effect of imperfect CSI on MIMO system's capacity and power allocation has been investigated under Rayleigh flat fading channels in [18] and it has been found that the degradation in capacity and the effects on power allocation vary, according to the characteristics of the communication channel and the fading conditions. The availability of an accurate CSI leads to a reduction in the required power radiated by UE and improve the energy efficiency.

On the other hand, accurate CSI in a time-varying channel is important to achieve the goals of the channel coding, enhance quantizers performance in the analogue-to-digital converter and analogue-to-digital converter, limiting the ISI, increasing the utilization of resources, and preventing co-channel interference. The channel estimation error in TDD massive MIMO depends on noise, number of antennas, the estimated channel matrix, and the length of pilot sequences. The channel matrix becomes a space-time matrix if multiple symbols are sent at the same time to a multiple antennas. It is considered semi-static over each coherence interval (fading block). The CSI accuracy increases and the channel estimation error decreases as the pilot signal power increases and the number of pilot symbols increases within each coherence interval. However, while the allocated power is subjected to several constraints, the length of the pilot sequence is limited which in turns limits the number of orthogonal pilot sequences. Also, as the length of pilot sequence becomes shorter, the channel estimation errors become slightly larger and lead to a bounded inaccuracy at CSI. The length of a pilot sequence should be equal to or greater than the total number of served users.

Therefore, reducing the channel estimation errors and increasing the accuracy of CSI involves various type of techniques. Most techniques require complex signal processing and iterative algorithms with a large number of iterations. This is a very challenging undertaking to be integrated in

real environments. In this Ph.D. dissertation, the CSI accuracy has been improved and the channel estimation error has been reduced by adopting a technique that uses simpler signal processing and a fewer number of iterations. Also, the accuracy of the channel estimation can be improved by designing a proper pilot scheme according to the propagation environment and system in use.

1.3.2 Massive Multiple Input Multiple Output

The idea of massive MIMO originated from the applications of random matrices theory. In a random matrix, if the number of rows increases excessively compared to the number of its columns, then the singular values of the matrix tend to be equal [19]. In this case, the columns tend to be pair-wise nearly orthogonal, which is a feature that can be exploited in the area of MIMO communications systems. Massive MIMO is an essential technology for the realization of 5G. It has the ability to offer tremendous access connectivity to serve a very large number of users and devices (e.g., as is the case in IoT).

As the number of BS antennas increases largely, the communication system converges to a massive MIMO where more users can be served and a higher sum rate throughput can be achieved for each cell. Although this leads to huge signal dimensions, massive MIMO due to hardening of the wireless communication channel offers to achieve high multiplexing gain with simple linear signal processing schemes. More antennas at the BS means higher spatial resolution that allows the BS to focus the energy toward the targeted users. Focusing energy results in reduction in both inter-cell and inter-user interference (intra-cell interference). However, achieving this is dependent on having an accurate CSI. Therefore, massive MIMO as a large scale MU-MIMO is a technology where the BS is equipped with a large number of antennas to achieve multiple orders of spectral efficiency, enhance the spatial degree of freedom, and increase the throughput of each user (i.e., lead to increase the sum-rate capacity of each cell). On the other hand, one of the fundamental benefits of massive MIMO is to harden the wireless communication channel to ensure reliable wireless communication and enhance the diversity and spatial multiplexing gains, even if the UE has only one antenna.

Moreover, deploying a large number of antennas (compared to the number of served users) at the BS improves the performance and reduces the complexity where simple linear signal processing like Minimum Mean Squared Error (MMSE) can be used and provides nearly optimal performance [20]. The MMSE receiver/precoder in uplink/downlink is the best amongst all types of linear signal processing based receivers/precoders, whereas such receiver/precoder maximizes the received SINR and outperforms all linear precoders over the entire range of SNRs [21]. This results in achieving high multiplexing gain and simplicity in signal processing. Also, it aids in improving the energy efficiency in both downlink and uplink. In the downlink, the BS focuses the energy beam into a very localized area. This enhances the spatial grid resolution and results in locating the targeted UE precisely and focuses the beam of the downlink transmitted signal. In the uplink, a very high array gain can be achieved by using coherent combining which results in significant reduction of the UE transmitted power, however, more improvement is needed. In downlink, as the number of antennas increases then lower-power rating amplifiers can be used to work in the linear region which results in improved energy efficiency.

In massive MIMO, in order to achieve high diversity and spatial multiplexing gains, the BS is required to employ coherent processing of the received signals. Accurate and timely acquisition of CSI is required, however, this is extremely challenging, especially in high mobility environments. However, in its ideal scenario, massive MIMO implementation grants a favorable propagation environment. Therefore, based on the law of large numbers (i.e., when the number of BS antennas are very high compared to the served users), the length of the vectors representing the CSI between the users and BS antennas increased and CSI vectors become pair-wise nearly orthogonal. The degree of orthogonality becomes near optimal as a huge number of scatterers exist in the propagation environment. As a result, provided the CSI are accurate, the received signals from multiple users are separable and result in maximizing the sum-rate capacity. The favorable propagation of a channel can be assessed either by the condition number of the channel matrix or by comparing the achievable sum-rate capacity with the one having favorable propagation, where the latter is more general. Then, under this favorable propagation, the wireless communication system can ex-

exploit the advantages of massive MIMO where the effect of noise and inter-user interference can be averaged out with simple linear signal processing like MMSE. The channel matrix also becomes well-conditioned. However, the cost of using the simple linear estimators in the uplink and linear pre-coders in the downlink at massive MIMO is the overhead requirement of acquiring accurate instantaneous CSI.

Antenna selection is another promising technology that may be used to get the advantages of massive MIMO while reducing its complexity. An accurate CSI is essential to achieve the target of antenna selection and is considered as a critical design objective [22]. In the reality, massive MIMO suffers from inherent impairment which is conventionally known as the pilot contamination. Pilot contamination reduces the CSI accuracy, the channel estimation performance, and spectral efficiency. Also, the failure in obtaining an accurate channel estimation can be considered, and seen from the opposite angle, as the main cause of the pilot contamination [12].

The classical multiple access used in the conventional massive MIMO is OMA. Also, TDD is the efficient and most widely feasible approach to be used in massive MIMO due to its unique feature of reciprocal channel [23–25]. TDD is more flexible in utilization of the frequency spectrum as the uplink and downlink signals use the same frequency channel in different time slots. As a consequence, the wireless communication channels of uplink and downlink are reciprocal under perfect radio frequency chain calibration and the channel estimation can be done only once.

Assuming the BS is equipped with M antennas and serves K UEs. Conventionally channel estimation is performed at the BS through pilot signals during the training phase of the uplink transmission where it requires a minimum of K channel uses (i.e., K symbols) to estimate the CSI of K served active users where the condition $2K < S$ ¹ is required which independent of the number of BS antennas, where S is the total number of symbols in the coherence interval. Therefore, reducing the pilot signaling overhead will improve the bandwidth utilization and increases the spectral efficiency. In contrast to TDD, FDD needs to do estimation for both downlink and uplink where

¹Using $2K$ symbols is the worst-case scenario. However, only K symbols are used for channel estimation if the UE estimates blindly the effective channel coefficient for decoding the downlink signal. Then, the remaining symbols in the coherence interval are used for effective data transmission.

the overall channel estimation requires $M + K$ symbols in the uplink transmission and M symbols in the downlink transmission. Therefore the coherence interval should satisfy the constraint of $(M + K) < S$ which compared to TDD appears not as efficient in terms of effective transmission usage of the coherence interval and it also limits the number of antennas equipped at the BS. Furthermore, in FDD, the number of downlink pilots are proportional to the number of BS antennas. Also, the bandwidth needed for the CSI feedback becomes very large. Both of these factors make the implementation of massive MIMO based on FDD that is pilot based extremely impractical.

In TDD, the coherence interval has a limited length which results in an insufficient number of orthogonal pilot sequences used in the training phase for estimating CSI. This leads to the pilot contamination problem in multiple adjacent cells, especially in fast fading environments where the channel has a shorter coherence time. As a result, this makes the number of available pilot orthogonal smaller (i.e., worsens the pilot contamination and enforces the use of the same pilot sequence, even in the same cell and creates both inter- and intra-cell interference). This dissertation follows the same assumptions of using the MMSE receiver/precoder and TDD approach in massive MIMO.

The main impairments that limits the performance of massive MIMO is pilot contamination [24]. The pilot contamination problem arises because of the existing correlation between different MU-MIMO channels due to: 1) insufficient antennas separation at the BSs and 2) the coherence time of the communication channel being limited. The limited coherence time leads to a limited number of users' orthogonal pilot sequences used for the uplink to acquire and estimate the CSI at the BS [20]; this pilot contamination problem has a strong presence in non-favorable propagation environments where the numbers of scatterers are small compared to the numbers of users being served [26]. To reduce the effects of pilot contamination several methods have been proposed, this Ph.D. dissertation analyzes and examines them in Section 1.4 and Chapter 2. Furthermore, we invented a new pilot scheme called DPS to mitigate the pilot contamination and offer additional features for the BS. This Ph.D. dissertation introduces DPS as a general tool that can be applied in any wireless communication system and demonstrates its application in massive MIMO and

NOMA.

1.3.3 Non-Orthogonal Multiple Access

In wireless communication, multiple access involves multiple users to a one-point transmission. In the uplink transmission of cellular networks, the one point is the BS that serves multiple users in its cell coverage area. The separation of users' received signals is achieved by using orthogonal multiple access schemes in which either frequency, code, or time resources are mutually exclusive allocated (orthogonal) for multiple users. NOMA is another type of multiple access that can be used in transmitting the signals of multiple users to a one-point receiver. Multiple users are superimposed on the same orthogonal resource in NOMA. The frequency and time resources in NOMA are shared among users that configuring a NOMA group, which are formed from multiple users in the cell. Therefore, the received signals at the BS that are transmitted by those multiple users are not orthogonal. The separation of users received signals in NOMA is achieved by using the Successive Interference Cancellation (SIC) at the BSs' receivers.

The challenges facing the realization and success of 5G cellular networks include the large volume of traffic exchanged and the high users' density with limited spectrum resources. These challenges can be tackled by integrating both NOMA and massive MIMO [27]. Furthermore, NOMA is considered as a 5G technology that aids massive MIMO to achieve low transmission latency, high spectral efficiency and high capacity, under a strong assumption of having perfect CSI [8, 27–30]. Compared to OMA, NOMA has many other features such as superior capacity and better channel utilization. This dissertation focuses on the enhancement of both the OMA massive MIMO and the NOMA massive MIMO in a practical approach that is feasible to implement.

One of the main impairments that limit the performance of NOMA is the error propagation during the decoding of the received signals in the SIC due to imperfect CSI. Accurate CSI in NOMA is essential and critical for achieving a reliable SIC [14], that will minimize the effects of error prorogation between consecutive SIC stages. NOMA is yet under further examinations to improve its performance, and rooms for more contributions are still available. One of the research gaps is to

integrate NOMA into massive MIMO and equip the combined system with an accurate CSI. The acquisition of an accurate CSI at the BS of a massive MIMO with NOMA integration is a very challenging task to achieve in the conventional system structure. Integrating NOMA into massive MIMO and finding effective ways to achieve this combination is also a very challenging and important topic that needs exploration [31]. Therefore, the Extended Dual Pilot Scheme (EDPS), we propose in this dissertation, plays a vital role in achieving this integration between NOMA and massive MIMO. Also, EDPS can provide an accurate estimate for CSI which is the key parameter to enhance NOMA's performance. This enhancement can facilitate proper resource scheduling with low system complexity and offers lower channel estimation errors. Additionally, this dissertation investigates CaF's application to improve the NOMA performance and propose new decoding schemes to be used in NOMA.

1.4 Literature Review of Related Works and Research Goals

1. **Pilot Contamination Elimination in Massive MIMO:** As previously mentioned, TDD is the efficient and feasible way to implement the massive MIMO [23–25]. Accordingly and as aforementioned, adopting TDD requires the BS to estimate the CSI using pilots signals during the uplink training [23, 32, 33]. Pilot contamination during the uplink training is the main inherent limitation that degrades the performance of massive MIMO [17]. Interference arising from adjacent cells during the uplink training (due to reusing the pilot signals) is the main source of pilot contamination [34]. Atzeni *et al.* in [35] suggested a pilot allocation scheme that uses a fractional pilot reuse algorithm to minimize the effects of pilot contamination. Since spectral efficiency is one of the main requirements of 5G, therefore the use of a large frequency reuse factor to solve this problem is not an efficient approach because it will reduce the spectral efficiency and the Pre-log factor [34]. Frequencies bandwidth where the channel Doppler spectrum is equal to zero is known as the Pre-log factor. The reuse factor is determined by how many cells reuse the pilots simultaneously. On the one hand, soft-fractional pilot reuse incorporates power control (center users transmits less power than edge

users by $1/3$) and requires a fewer number of orthogonal pilots than the strict-fractional pilot reuse but it has several drawbacks. On the other hand, the coherence interval of the channel is limited, so the number of orthogonal pilots is limited [36].

Finding a creative solution for pilot contamination effects is a hot research topic. Several researchers have worked on mitigating pilot contamination effects and proposed different heuristic and suboptimal algorithms. Marzetta in [20] was the first to address this problem in TDD massive MIMO for multiple cells with a non-cooperative scenario. Since then, it has been extensively discussed in the literature [35, 37–52]. Xu *et al.* in [37] investigated the pilot contamination effects on the estimation of CSI in massive MIMO over frequency-selective channels and has shown that the MMSE estimation algorithm is more resistant to pilot contamination. Ngo *et al.* in [38] also analyzed the effects of pilot contamination for physical channel models in multi-cell multi-user massive MIMO and showed pilot contamination effects persist under finite dimensional channel models. Nguyen *et al.* in [43] proposed time-shifted pilots and adopted two uplink training methods by combining the conventional time-aligned pilots and time-shifted pilots to reduce the pilot contamination in multicell massive MIMO. Jin *et al.* in [44] exploited combat pilot contamination property of time shifted pilots and analyzed the performance of massive MIMO when a Zero-Forcing precoder and beam-former were used. Besides, Ngo and Larsson proposed in [53] to exploit eigenvalue-decomposition to perform the channel estimation directly from the received data under the assumption of the existence of asymptotic orthogonality between the channel vectors. To separate the signal subspace and the interference subspace, their approach assumes the existence of asymptotic orthogonality between the channel vectors and requires high computational complexity.

In contrast to the above, a semi-blind approach was proposed in [50] which has been developed based on the assumption of time shifted pilots and employs the independent component analysis with successive cancellation. This approach has a similar problem of the time-sifted pilots approach. Time shifted pilot signaling moves the BSs from operating in a synchronous

into an asynchronous TDD mode of operation. This results in huge interference in uplink transmitted signals, which is caused by the downlink transmitted signals. Downlink transmitted signals have higher transmission power compared to the uplink transmission power. The pilot decontamination performance will be a function of the pilot to data parts ratio. Elijah *et al.* in [39, 40] introduced an extensive survey of other sources of pilot contamination like hardware impairment, non-reciprocal transceivers, and presented suggested methods to mitigate pilot contamination. These suggested methods were based on different assumptions, different system configurations, and different channel models. Based on the channel estimation method, mitigation methods have been classified into either a subspace-based or pilot-based approach.

Sarker and Lee in [41] studied the reduction of the pilot contamination problem in massive MIMO based on a diagonal Jacket matrix. They proposed a diagonal Jacket based covariance aided channel estimation method, where perfect multipath fast channel estimation was assumed. Müller *et al.* in [42] analyzed a subspace projection to enhance the channel estimation and proposed a blind pilot decontamination algorithm. The algorithm had a polynomial complexity. Saxena *et al.* in [45] used an open loop power control and pilot reuse algorithms adopted in Long-Term Evolution (LTE) and developed techniques to mitigate pilot contamination. These techniques were effective for users moving at vehicular speeds and for estimating the channel using a least square estimator. Lee *et al.* in [46] proposed a partial sounding resource reuse strategy to mitigate pilot contamination where the cell was divided into the center and edges areas, assuming multi-cells operated cooperatively in the edges areas. Zhang *et al.* in [47] studied the pilot contamination (considering the OFDM system parameters) and mitigated pilot contamination by using training techniques for both uplink and downlink. Although the proposed techniques do not require the channel second order statistics, they have the drawback of expanding the training overhead by a factor equal to the number of interfering cells. In contrast, Farhang *et al.* in [48] replaced OFDM modulation by Cosine Modulated multi-Tone (CMT) modulation and extended the blind estimation capa-

bility of CMT to improve the accuracy of CSI and mitigate the pilot contamination in CMT based massive MIMO. Sørensen and De Carvalho in [49] converted the pilot contamination into randomized contamination by using pilot sequences hopping at each transmission slot and used the Kalman filter to reduce the effect of randomized contamination. The channel estimation was assumed to incorporate multiple time slots. This algorithm is efficient at low and moderate mobility users.

Yin *et al.* proposed in [54] to assign the pilots among cells by using a covariance aware pilot assignment scheme to mitigate the pilot contamination efficiently but this approach was limited and worked under certain conditions of the channel covariance. Hu *et al.* in [50] proposed a pilot decontamination algorithm based on semi-blind channel estimation, where a constrained minimization optimization problem based on an asynchronous pilot protocol was formulated for each user. The proposed algorithm required intensive computation and its efficiency depended on the data length. Zhu *et al.* in [51] exploited the large-scale characteristics of fading channels and proposed a smart pilot assignment based on measuring the inter-cell interference with the aim to maximize the minimum uplink SINR of all users in the served cell through an optimization algorithm. Users having the worst channel quality will assign the pilot sequences which have the smallest inter-cell interference. However, the proposed method requires the served cell to solve the optimization problem and perform a sequential procedure iteratively until a convergent solution is reached. Vu *et al.* in [52] proposed modified least square estimators that removed the pilot contamination based on consecutive pilot transmission phases algorithm which required coordination between the served cell and the adjacent cells.

In this dissertation, we develop an alternative novel method to eliminate the effects of pilot contamination, bridge the research gaps mentioned above, and hence improves massive MIMO performance through more accurate CSI estimation. The developed scheme has a lower complexity and can be generalized to any system configuration such as the 5G cellular networks and beyond.

2. **Enhancement of CSI Accuracy and SIC Performance in NOMA:** SIC plays an important role in NOMA technology. A reliable SIC that separates users' signals at the BS's receiver and does not result in a residual interference is a critical factor that determines the performance efficiency of the system. Reliable SIC can not be achieved without an accurate CSI [8]. Hardware impairments at transceivers is another factor that should be considered in the designing of SIC. The effects of these hardware impairments amplify and cause a significant impact on SIC performance in the presence of imperfect CSI.

The integration of both NOMA and massive MIMO brings more technical challenges to achieving reliable SIC because users pairing and scheduling into appropriate channels and selecting the best antenna subset at the BS need to be considered jointly. Moreover, massive MIMO with NOMA integration can consume excessive bandwidth resources to acquire accurate CSI [28]. Previous work done, that took these considerations into account, shows that a very high complexity was added to the SIC design and therefore this integration can only be applied to cells with small numbers of served users and antennas [27]. Using multi-polarized antennas has been proposed in [29]; however, this solution is only considered a slowly varying spatial correlation of the communication channels and also added complexity to the system. An approach for reducing the computational complexity of SIC in [28] by employing low feedback and decomposing the massive MIMO NOMA into a set of single-input single-output NOMA channels was proposed. This approach required a perfect knowledge of users ordering at the BS, which is more feasible only at the downlink.

Li *et al.* in [55] used a novel approach that employs the Forward Error Correction (FEC) and exploits the intrinsic diversity of code words to mitigate the imperfect CSI estimates in NOMA and improve the performance of SIC. However, this approach added more complexity to the SIC design. The SIC design in [56] considered an imperfect CSI which has been classified into two types a) the channel distribution information and b) the channel estimation uncertainty. However, this approach led to complexity in the SIC design, probabilistic constraints, and assumed non-practical approximations. The decoding with a fixed order was

replaced by decoding with a dynamic order in [57] to improve the performance of the SIC but this approach was feasible only in a slow fading environment. This approach has been extended in [58] to consider the existence of imperfect CSI. It has been shown that imperfect CSI led to incorrect decoding order in addition to the extra interference on the desired signal, which resulted in a degradation of the performance of the SIC.

Accordingly, based on the aforementioned, the performance of NOMA technology depends heavily on SIC proper functionality which can not be achieved and guaranteed without accurate CSI. A question that arises here, “how can we improve the SIC without increasing the system complexity?” In this dissertation, we address this question and provide the appropriate solution. The dissertation introduces EDPS for improving the accuracy of CSI estimation ensuring compatibility for practical implementation and lower complexity. Therefore, EDPS improves CSI accuracy and, as a consequence, accurate CSI facilitates the NOMA integration into massive MIMO. In addition, EDPS enhances the performance of SIC.

3. Improving NOMA by Employing CaF and Reducing CaF Sensitivity to CSI Estimation

Errors: The allocated power variations of Power-Domain NOMA (PD-NOMA) paired users play an important role in the performance of SIC. Keeping significant power differences ensures good performance of SIC, which can not be guaranteed all the time. Besides, fairness and outage probabilities are impacted when the paired users powers are close to each other. Therefore, in this dissertation, another approach is used in NOMA to decode the received uplink signal and de-multiplex the desired signals by employing the CaF decoding method. CaF also called physical-layer network coding. CaF started as a relaying strategy, where each relay computes a linear combination of transmitted signals and then forwards it to the ultimate receiver [59]. As a concrete example, suppose that there are two transmitters, two relays, and one destination. Each relay computes a linear combination and forwards it to the destination. As long as the two linear combinations are linearly independent, the destination can recover the transmitted signals. Besides, if the communication channel does not provide a suitable linear combination of the transmitted signals then CaF cannot work properly to

recover the desired signals.

Interestingly, the idea behind CaF goes beyond the relay scenario. It can be applied to other scenarios, such as the point-to-point channel and Multiple Access Channel (MAC). CaF is used for optimal utilization of interference [59], which can be employed in the uplink NOMA. CaF is the decoding method used as the core of IFLRs in MIMO wireless communications networks. Therefore, by employing CaF, the front end at the BS's receiver decodes and computes the linear combinations of joint signals transmitted from multiple served users in NOMA and then forwards separable signals to the appropriate stage(s) at the BS's receiver for further processing. The transmitters in the equipment of served users do not require the pre-knowledge of the CSI to employ the CaF in the uplink transmission [60]. However, accurate CSI is required at the BS's receiver.

Integrating CaF in a massive MIMO system improves the performance where CaF can be enhanced by using multiple antennas at the receiver [60, 61]. The aim of CaF is to tackle both the signal interactions introduced by the NOMA communication channel and the noise in the communication environment [62]. Moreover, the CaF exploits rather than combats the multiple access interference towards increasing the system throughput and, as a consequence, the sum-rate capacity [59, 63]. There are many approaches to implement the CaF but there is a trade-off between the reduction of implementation complexity and improving the achieved performance. However, the CaF multiple access strategy keeps the decoding complexity to grow linearly with the number of served users and the performance is improved by increasing the accuracy of CSI [64]. Furthermore, Niesen *et al.* in [62] shows the CSI plays an important role in the CaF performance and how it limits its degree of freedom. The DPS/EDPS mechanisms which we introduce in this dissertation tackles this efficiently.

On the other hand, CaF uses scaling coefficients before forwarding where it has been found that employing the MMSE scheme can uniquely maximize the computation rate and produce the optimal scaling coefficients [59]. However, the accomplished computation rate relies on the CSI accuracy. In [65] the sensitivity of CaF to the CSI errors has been investigated. It has

been found that imperfect CSI leads to errors in the MMSE scaling coefficients and errors in the integer coefficients used by CaF which are carefully chosen based on CSI. Therefore, CaF is sensitive to the errors in CSI and an accurate CSI is needed to reap the full advantages of CaF in practical implementation and avoid a significant rate loss and outage probability increment [65, 66]. Another approach to overcome the effects of CSI accuracy has been investigated in [66] where the blind CaF has been proposed which avoids the need for CSI. However, this has resulted in a trade-off between increasing computational complexity and only achieving a sub-optimal performance of CaF.

A question that arises here, “how can CaF guarantees improving NOMA performance without increasing the system complexity?” In this dissertation, we address this question and provide the robust solution. First, we propose new decoding schemes based on CaF that improves the decoding performance of NOMA and massive MIMO-NOMA that ensure higher fairness and smaller average outage probabilities while having the same SIC complexity. Second, we integrate DPS/EDPS into massive MIMO-NOMA, which both improve the CSI accuracy and enhance the overall performance.

- 4. Increasing Energy Efficiency and Promoting Green Communications:** Incorporating Energy Efficiency (EE) in the system design of 5G cellular networks and beyond is an essential goal and one of the 3rd Generation Partnership Project (3GPP) standards technical targets that should be adopted. The approach of acquiring green communications in 5G cellular networks and beyond would improve the EE in Information and Communications Technologies (ICT), and promote a healthy living environment [67]. Adopting this approach would result to reduce the energy consumption without affecting other performance factors in 5G cellular networks and beyond [67]. This can be achieved by maximizing the amount of data transmitted per unit energy without degradation of spectral efficiency and system capacity. Patcharamaneepakorn *et al.* in [68] investigated a generalized spatial modulation as a radio access technology to improve the throughput and EE in 5G massive MIMO, but this has been found to result in less spectral efficiency. This is contrary to the aim of 5G cellular networks

to improve spectral efficiency by one order of magnitude. Therefore, there is a trade-off between EE and Spectral Efficiency (SE); this dissertation devises an approach that improves EE without putting huge negative impacts on SE.

From another perspective, a statistical delay bounded QoS driven power allocation schemes have been developed in [69] to maximize the power efficiency for 5G cellular networks while assuming perfect CSI. However, the availability of perfect CSI at the BS is very challenging. This dissertation addresses this, and proposes DPS/EDPS to improve CSI accuracy. Also, pilot contamination with imperfect CSI limits the achievable EE, SE, and throughput [68]. Another approach to improve EE is to apply a perfect SIC at the receiver. This may show a noticeable improvement of the energy efficiency in a massive MIMO system [70], but as mentioned earlier, optimal SIC functionality can not be achieved without having an accurate CSI. Therefore, devoting efforts to addressing the pilot contamination problem and enhancing CSI is indirectly tackling the lack of EE, which has been achieved in this dissertation.

On the other hand, in [71] the EE for massive MIMO with a large number of antenna and radio frequency chains have been investigated and an energy-efficient hybrid precoding algorithm has been proposed to maximize the EE. The results show an improvement in EE performance for a low number of users and it has been only compared to the Zero-Forcing (ZF) precoder. However, MMSE is the recommended precoding technique used in massive MIMO as ZF suffers from a low performance at a low SNR as it amplifies noise and also massive MIMO supposed to provide high users connectivity. Also, it is highly expected the cells in 5G cellular networks and beyond will have high density and a huge number of users. Besides, the EE and green communications feasibility have been investigated in [72] and [73]. Results showed improvements only for cellular networks configured with small cells. This highlights the importance of EE as small cells with high-density of users are expected to be used in 5G cellular networks. However, 5G addresses a wide range of applications that make the adoption of small cell topology not only unrealistic, but also shows that other scenarios with different cell sizes should be considered when investigating EE. This disser-

tation explores in-depth into finding an optimum approach that can find an energy-efficient mechanism. A mechanism that can be generalized regardless of the cell size and number of served users used in 5G cellular networks and beyond.

Importantly, the explosive growth of subscribers and used devices in cellular networks entails a massive amount of energy consumption when aggregated, which inevitably leads to a larger carbon footprint, and greatly contributes to environmental pollution if the energy sources used are not renewable. Furthermore, the rapid growth of energy consumption is proportionally related to the rapid growth and evolutions in ICT, which occurs at a very fast rate as shown in Fig. 1.2. Moreover, according to recent statistics, the energy consumption in ICT occupied around 10% of the world's total energy consumption [74]. Recent statistics also show that the average current power consumption for the average UE is around 22.5 Watt daily. Accordingly, the expectation of having more than 9 billion UE to be served will lead to a power consumption of 202.5 GigaWatt daily. This is a massive number that should be reduced to promote green communications and achieve 5G technical targets. From a different perspective, 5G is expected to serve a huge number of sensors that embed into IoT and have limited power storage and some types depend on energy harvesting. To highlight the importance of EE in UEs, a Pre-Grant signaling scheme has been proposed in [75] to reduce the energy consumption by the UEs but only for downlink. Nevertheless, an energy-efficient scheme for uplink with reasonable cost to integrate is also highly needed especially in systems with TDD integration. Moreover, it has been shown that most energy efficient schemes suffer from high costs compared to the traditional approach [76]. In this dissertation, we propose a scheme that reduces power consumption and improves EE in UEs and 5G uplink transmissions without degradation of SE. The scheme we propose is an energy efficient, economically feasible, adaptive, easy to integrate, which can find its way into practical and real-life implementation.

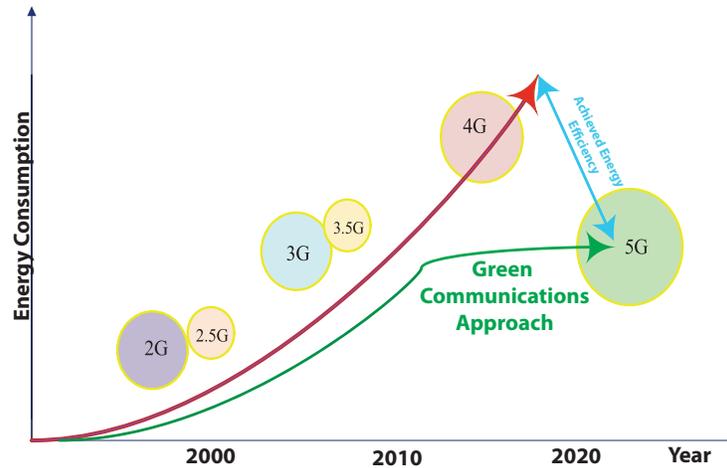


Figure 1.2: Green communications and energy efficiency

1.5 Dissertation Contributions

In this section, highlights of the major contributions of the dissertation are presented. A schematic diagram of the contributions of the dissertation is illustrated graphically in Figure 1.3, and are summarized as follows:

- ***Proposing the Novel Dual Pilot Scheme:***

The accuracy of the CSI is a critical factor in achieving all the technical targets of any wireless communication system. Therefore, we develop a new signaling scheme named DPS that would overcome the limitations of conventional schemes, promote simplicity, be easy to implement and get deployed into existing systems, improve the performance indicators and technical targets, and can be generalized to integrate into any wireless communication systems. DPS offer beneficial and unique features and can be used in a non-cooperative multi-cell cellular networks. Our evaluations show that DPS reduces the effects of pilot contamination substantially, increases the CSI accuracy, and consequently improves the massive MIMO performance. DPS works in one dimension in massive MIMO to mitigate inter-cell interference.

In contrast to the recent works that traditionally focused on the improvements in sum-rate

maximization with aim of optimizing the total network throughput, our DPS approach is different and maximizes the sum-rate and optimizes the total network throughput with ensuring the maximization and optimization for each served user. Thus, it provides improvements that convey the fairness and uniform QoS for each and every user is served. This guarantees to avoid service disruption or poorly serving users experiencing poor channel conditions, or those at the cell-edge whom under high attenuation path loss. Our contribution towards this innovative pilot signaling scheme has been accepted as a provisional patent and has been published in [77] and [78].

- ***Proposing the EDPS for NOMA and Massive MIMO-NOMA:***

As a second contribution we extend DPS and develop the novel EDPS to operate in two dimensions to mitigate both intra-cell and inter cell interference. EDPS provides the BS with more accurate CSI for multi-fold benefits. First, EDPS facilitates the integration of both NOMA and massive MIMO as availability of an accurate CSI is essential to achieve this integration. In addition, EDPS achieves this goal by a simple approach, which ensures easiness for deployment in practical systems. Second, EDPS enhances the performance of SIC in NOMA, as SIC performance depends on the accuracy of CSI. Third, EDPS equips the BS with interference management that can work in non-cooperative multi-cells environments. These highlights the importance of EDPS and result in enhancing several performance metrics. Our contributions on EDPS and its application in NOMA and massive MIMO-NOMA have been submitted for publication [79].

- ***Investigating the EE Performance of DFT-S-OFDM and Proposing Barker Coded DFT-S-OFDM for 5G Uplink Transmissions:***

We first examine the performance of Discrete Fourier Transform -spread- OFDM (DFT-s-OFDM) in 5G uplink transmissions with a focus on evaluating the EE. Our system model considers both massive MIMO with OMA and massive MIMO with PD-NOMA. This dissertation is the first to investigate both approaches. In the second contribution, we propose a novel mechanism that lowers power consumption at UEs, improves EE without impacting

the SE, and can be adopted in different applications of 5G using Barker Coded DFT-s-OFDM (BC DFT-s-OFDM) in 5G uplink transmissions. The proposed scheme adds a Barker code spreader at the UE and despreader at the BS, which results in lowering Peak-to-Average Power Ratio (PAPR) and significantly reduced power consumption at UEs. We confirm that BC DFT-s-OFDM enhances the performance of both OMA-massive MIMO and PD-NOMA-massive MIMO and increases the EE under different communication channel models. This dissertation is the first to open the doors for employing Barker code in 5G cellular networks and beyond. These contributions have been published in [80] and submitted for publication in [79].

- ***Employing and Improving CaF in NOMA and Massive MIMO-NOMA:*** Instead of using SIC to resolve interfering signals in NOMA, this dissertation proposes CaF in NOMA and finds new decoding schemes for PD-NOMA and massive MIMO-PD-NOMA based on CaF. These proposed schemes provide better performance in terms of increasing users' fairness and reducing the outage probabilities while preserving a low implementation complexity.

CaF is very sensitive to the CSI estimation error. Besides, the performance of CaF also improves as the number of available combinations of linearly independent equations increases at the BS. As a second contribution in this matter, this dissertation integrates DPS and EDPS with CaF to increase CaF robustness and lower the sensitivity of CaF to estimation errors. This integration improves the overall systems performance of PD-NOMA and massive MIMO-PD-NOMA in uplink transmissions. It also enhances the performance of IFLRs. We claim this dissertation is the first that proposes such an approach. Our contributions toward this innovative new decoding schemes and for employing the DPS/EDPS have been published in [81] and under submission in [82].



Figure 1.3: A schematic view of the dissertation’s contributions.

1.6 Dissertation Outline and Organization

The structure of the dissertation has been composed and formed according to the research flow and is connected to the list of contributions in Section 1.5. The dissertation is divided into five chapters and two appendices, it is organized as follows:

In **Chapter 1**, we provide an overview of the research problems and challenges, research gaps, research goals, and approaches to achieve these goals. We present background information on the 5G technologies and the impact of channel estimation errors on the performance of TDD systems. We then discuss the literature review and related previous work that have been done, focusing on articulating the gaps in research and identifying the areas of our contributions. We highlight the

practical challenges in the TDD interference-limited cellular communications systems and the importance of sustainability for green communications which motivated our research. This is followed by the introduction of our developed schemes and solutions to improve CSI accuracy, mitigate pilot contamination, ensure EE in 5G cellular networks, and employ and enhance CaF decoding in 5G.

In **Chapter 2**, we develop and apply the novel DPS in a massive MIMO system with OMA deployment to mitigate the pilot contamination and generate an accurate CSI. We use different wireless communication environments during our investigations, the independent and identically distributed (i.i.d.) and the correlated Rayleigh fading channels. We start by illustrating how the DPS is used in massive MIMO and how the pilot sequences codebook is formed. We show the performance of DPS in one dimension to mitigate the inter-cell interference and introduce γ as a design parameter. We derive the mathematical proofs for our theorems, which state the DPS is universally outperforming the conventional Single Pilot Scheme (SPS), provides the guarantee for a uniform quality of service for each served user, and generates lower MMSE of estimated CSI, where the MMSE is a function of γ . A group of detailed mathematical derivations for this chapter is provided in Appendix A. Numerical results depict superiority for the integration of DPS which increases the sum-rate capacity of massive MIMO cell that is located within a group of non-cooperative cells, where a BS is equipped with a given number of antennas and while varying the SNR. The materials presented in this chapter have been included in one accepted provisional patent, two publications, one journal article in [78], and one conference article and presentation in [77].

In **Chapter 3**, we dive deeper and extend the DPS to work in two dimensions to simultaneously mitigate inter-cell interference and intra-cell interference. We introduce two design parameters γ and μ for constructing the EDPS sequences. The sequences codebook has been divided into three groups. As a result, we gain advantages of EDPS and use it in massive MIMO with NOMA deployment to improve performance. Additionally, we investigate the feasibility of DFT-s-OFDM waveforms EE in both massive MIMO with OMA deployment and massive MIMO with PD-NOMA deployment. Then, we conduct an in-depth examination of Barker codes as an approach to achieve

the corresponding effective trade-off among spectral versus energy efficiency. Through this, we exploit Barker codes and propose their integration into the uplink waveforms in 5G. Numerical results demonstrate improved EE performance of the proposed schemes and the prove significant advantages of integrating BC DFT-s-OFDM in 5G. The materials presented in this chapter have been included in one published conference article and presentation [80] and another journal submission [79].

In **Chapter 4**, we begin by considering the main differences in the performance between the SIC-NOMA and CaF-NOMA when the NOMA paired users are close to each other in terms of allocated power. Then, we formulate the performance indicators in terms of users Jain's fairness index and outage probabilities. We propose new decoding methods for uplink PD-NOMA and massive MIMO-PD-NOMA which achieve better fairness and smaller average outage probabilities, while having the same decoding complexity as SIC decoding. We address the CaF sensitivity to the CSI estimation error and we employ the novel DPS/EDPS into our system model to exploit their unique features. We introduce the preliminaries of IFLR, and show its limitations and sensitivity to channel estimation errors. In addition to this, we derive the analytical expression of reduction loss in achievable rates (i.e., rate loss) and show numerically how DPS and EDPS improve the performance of CaF and consequently enhance IFLR. A group of detailed mathematical derivations for this chapter is provided in Appendix B. The materials presented in this chapter have been included in one published journal article [81] and another journal submission [82].

Finally, the concluding remarks, main research outcomes, drawn findings, avenues of approach more potential research, and some future work related to the topics covered in this dissertation are presented and summarized in **Chapter 5**. in inclusive communication scenarios of 5G cellular networks and beyond.

Appendices A and **B** present the mathematical derivations, proofs of the theorems, and preliminary materials relevant to Chapters 2 and 4, respectively. Figure 1.4 illustrates the schematic diagram of the dissertation structure and composition.

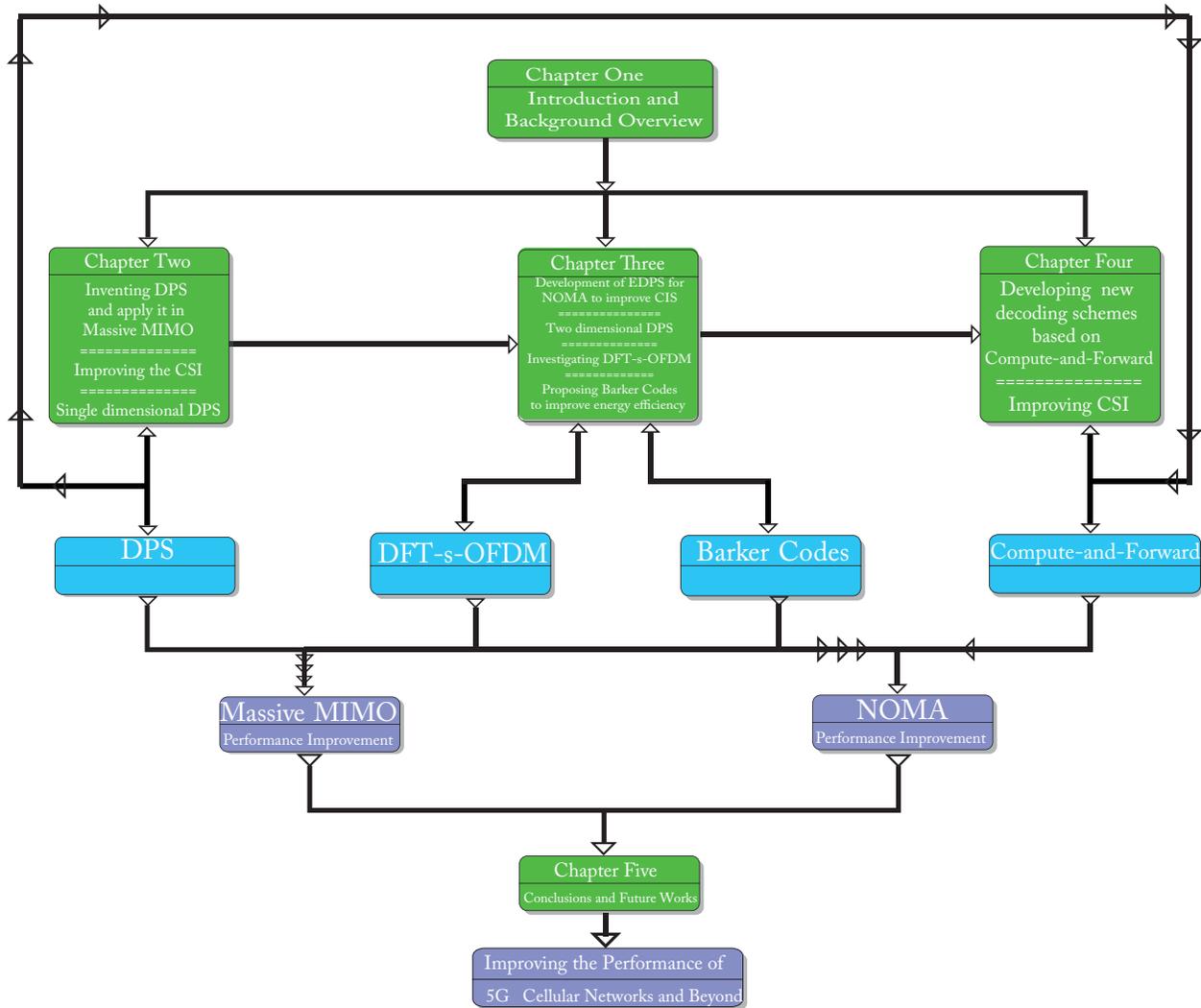


Figure 1.4: A schematic view of the dissertation structure and relationship among chapters.

Chapter 2

Dual Pilot Scheme and its Application in Massive MIMO

2.1 Synopsis

The pilot scheme currently used in 5G cellular networks assigns the same set of orthogonal pilot signals to all cells. This results in inter-cell interference, also known as pilot contamination, which can significantly degrade performance, especially in massive MIMO systems. To mitigate this interference, we propose a novel DPS that assigns a slightly modified set of nearly-orthogonal pilot signals. DPS is a general scheme that can be implemented in any wireless communication system, including 5G and beyond. We demonstrate the integration of DPS in a massive MIMO system in both microscopic and macroscopic levels and analytically prove that DPS enables more accurate estimates of the CSI in the minimum mean-squared error sense, under the i.i.d. and the correlated Rayleigh fading wireless communication channel models. We further validate and demonstrate the advantages of DPS over various channel models of massive MIMO 5G technology by extensive simulations.

2.2 Introduction

As mentioned in Chapter 1, accurate estimation of the characteristics of a wireless communication channel, i.e., CSI, is vital for reliable wireless communications in contemporary cellular networks [4, 5, 83]. Two main approaches for channel estimation are pilot aided and non-pilot aided, with the former being more practical and widely adopted in current standards especially in conjunction with TDD technology [4, 5, 84]. TDD is attractive for deployment in cellular networks due to its unique reciprocal channel property between uplink and downlink. Thus a BS can periodically perform channel estimation using the uplink Pilot Sequences (PSs) [20, 23, 24, 85, 86]. The duration between channel estimations should be less than the minimum coherence interval related to the specific application environment and system configuration, which together with the need to minimize overhead limit the length of the PSs and hence the number of available orthogonal PSs. In contemporary TDD cellular systems, the same set of PSs is used by all active users within each cell and also reused in all adjacent cells for channel estimation [20, 24]. This leads to the critical problem of pilot signal interference between adjacent cells, widely known as pilot contamination [21, 24, 86, 87], which is one of the main impairments that limit the performance of massive MIMO systems. Massive MIMO [7], which employs a large number of antennas at a BS to support beamed or parallel links to one or multiple users, is widely considered as one of the key technologies that enable the substantially higher system capacity in the emerging 5G cellular systems. As massive MIMO requires accurate CSI, it is commonly deployed together with TDD, and is thus particularly susceptible to the pilot contamination problem.

Pilot contamination is caused by inter-cell interference between pilot signals from adjacent cells, and can be a major cause of errors in the CSI of an intended user at the serving BS leading to sub-optimal beamforming. When this is compensated by an increase in transmit power at the user equipment or BS, it leads to reduction in the battery lifetime of mobile devices, or wasted energy at the BS. If this degradation is not sufficiently compensated, it causes a drop in the quality of service experienced by the users. Pilot contamination exhibits a strong presence in non-favorable propagation environments where the number of scatterers is not sufficiently large [26, 39]. Re-

ducing the effects of pilot contamination can improve the accuracy of channel estimation (or CSI) [37, 85, 86, 88]. This tends to improve the energy efficiency and enhance the sum-rate capacity at each cell, and hence improve the system capacity to serve more users [39].

Motivated by the needs to efficiently mitigate pilot contamination, especially in TDD massive MIMO systems, we propose a novel mechanism named DPS. Rather than using the same set of orthogonal PSs in every cell, adjacent cells in DPS use a different set of nearly-orthogonal PSs, thus avoiding pilot contamination. We prove, mathematically, that this scheme reduces the inter-cell interference compared to the conventional pilot scheme currently in use, leading to a smaller MMSE in channel estimation. Besides, our proposed novel DPS does not require any additional hardware configuration and is easy to integrate within current wireless communications systems including massive MIMO. It also consumes the same amount of energy and uses the same transmission overhead as in the conventional scheme. DPS requires only a few additional signal-processing operations that can be readily supported in contemporary systems.

The *main contributions* of this work include: (1) introduction and proposal of DPS as a novel mechanism to mitigate pilot contamination in cellular wireless networks, which is generally applicable in both TDD and FDD systems and can be employed in any wireless communications system including 5G and beyond (2) validation of the effectiveness of DPS through detailed mathematical analyses and simulations to show how DPS enhances TDD system performance with a guarantee for a uniform QoS for *each and every* user being served in the system, especially in systems employing massive MIMO, compared to the conventional pilot scheme.

As a first strategy, to validate our proposal, we start from a microscopic configuration of L cells, deriving simple conditions under which DPS outperforms the conventional scheme for *each and every* user under two-channel models: the i.i.d. and the correlated Rayleigh fading channels. In sharp contrast, prior work mostly focused on minimizing a weighted sum of Mean-Squared Errors without a per-user performance guarantee. Second, we move from a *microscopic* configuration of L cells to a *macroscopic* configuration of the whole cellular network, explaining why the advantages of our DPS still remain even if we keep the same overhead (i.e., the number of PSs is independent

of the number of cells).

The i.i.d. Rayleigh fading wireless communication channel model is conventionally used in the theoretical studies of massive MIMO. However, a study on massive MIMO channel measurements conducted at Lund University [23] has found that these channels are well-represented by the correlated Rayleigh fading wireless communication channel model. Therefore, in this chapter, to demonstrate the advantages of implementing DPS in any wireless communication system including massive MIMO, both the i.i.d. and the correlated Rayleigh fading wireless communication channel models are used in our analyses.

The rest of this chapter is organized as follows: In the next section, we review the related work. In Section 2.4 we present the system model. In Section 2.5 we introduce DPS as a general scheme that is broadly applicable in wireless communication systems. We show how DPS can be integrated in such systems and why it outperforms the conventional pilot scheme by enhancing CSI accuracy and consequently mitigating the pilot contamination, especially in massive MIMO systems. We prove mathematically that DPS achieves channel estimates with smaller MMSE than the conventional scheme, under the i.i.d. Rayleigh fading wireless channel model. In Section 2.6, these analyses are extended to the correlated Rayleigh fading wireless channel model, which has been shown to be appropriate for massive MIMO systems. In Section 2.7, we illustrate the application of DPS in massive MIMO by considering the entire cellular network at a *macroscopic* level. Section 2.9 presents and evaluates the obtained numerical results. Section 2.10 concludes the chapter.

2.3 Related Work

Existing solutions for the pilot contamination problem generally employ one or more of the following approaches: 1) adjacent cells coordination, 2) user grouping, 3) uplink and downlink cooperation, 4) interference alignment, and 5) blind estimation, which add complexity to the system in terms of signaling for coordination or computations for signal processing.

For instance, the scheme in [52] employs adjacent cell coordination to eliminate pilot contam-

ination. Pilot transmissions are organized in phases, such that users served by a BS stay idle in one phase while users served by an adjacent BS transmit their PSs. This scheme increases the pilot transmission overhead by a factor equal to the number of phases. Correspondingly, the approach proposed in [44] uses time-shifted PSs that requires multiple-cell coordination and asynchronous pilot transmissions in adjacent cells. Contrarily, DPS does not require users in adjacent cells to transmit their PSs in different phases or different times, and hence it does not require multiple-cell coordination or asynchronous pilots between adjacent cells.

To allow the use of a relatively small number of orthogonal training sequences to simultaneously train the links of all users, the works in both [89] and [90] followed the second approach by grouping users according to their spatial signatures. In contrast, DPS does not require the grouping and classification of users within the cells. Moreover, both works assumed the existence of a long preamble to enable extraction of the required model parameters. On the other hand, DPS does not have extra requirements on the preamble on top of what it is ordinarily used for, and it can easily accommodate a shorter preamble to reduce system overhead. As a consequence, saving in preamble's duration can be used in the actual transmission, thereby offering more enhancement to the system spectral efficiency. While a unified transmission strategy for TDD/FDD massive MIMO has been proposed based on the spatial basis expansion model in [89], one of the objectives of this strategy is to improve channel estimation for both uplink and downlink while reducing the training overhead. Consequently, more orthogonal training sequences can be allocated to the adjacent cells and be exploited to mitigate pilot contamination. To extend this approach, the uplink and downlink channel estimations for both TDD and FDD in time-varying massive MIMO networks were investigated in [90]. A sparse Bayesian learning framework is designed to estimate the spatial signatures and temporal varying characteristics of the channel model and exploit them to learn the spatial information instead of the channel covariance matrices to avoid computation complexity. It is important to highlight that DPS can also be integrated in both FDD and TDD massive MIMO systems, although this chapter only evaluates the performance of DPS in TDD massive MIMO.

The pilot contamination mitigation scheme proposed in [47], which exemplifies the third ap-

proach, is based on a downlink training stage followed by a scheduled uplink training stage, with the downlink training stage adopting phased coordination with adjacent cells as in [52]. However, DPS does not depend on coordination between the uplink and downlink. The uplink training in DPS is used for estimating the downlink CSI to enable beamforming and precoding of downlink transmissions, which does not add any additional complexity to conventional TDD system operation.

It has been shown recently that exploiting interference alignment allows the BS to use the radio resources more efficiently under certain conditions. However, to achieve the optimal performance, this approach needs to deal with a large number of signaling dimensions, which is proportional to the number of interferers under consideration. The pilot contamination mitigation scheme proposed in [91] uses interference alignment and a soft-space-reuse based cooperative transmission scheme to improve CSI estimation. However, to enable the use of low dimensional training matrices, this approach requires cooperation between adjacent cells, and thus exemplifies the combination of approaches one and four. Also, through optimizing and minimizing the Mean Squared Error (MSE) at each BS, an interference cancellation transceiver that reduces the overhead transmission has been proposed in [92] for multi-user cooperative multi-cell networks. However, it is found that there is a trade-off between the computational complexity and the number of users. In sharp contrast, the complexity associated with these approaches is not found in DPS while the latter is still capable of enhancing the system spectral efficiency. DPS also offers more signaling to the BS without any additional system complexity or trade-off between cooperation amongst multiple adjacent cells and computational complexity, as DPS does not require any cooperation between multiple adjacent cells to manage the interference.

As the fifth approach, blind pilot decontamination as proposed in [42] exploits singular value decomposition to perform a blind channel estimation directly from the received data without training. To separate the signal subspace and the interference subspace, this approach assumes the existence of asymptotic orthogonality between the channel vectors and requires a high computational complexity. To reduce the complexity of this approach, semi-blind pilot decontamination

was proposed in [50]. It treats the detected uplink data as pilot symbols such that a constrained optimization problem based on an asynchronous pilot protocol was formulated for each user. However, the proposed semi-blind algorithm also requires intensive computations and its efficiency depending on the data length. In comparison, DPS does not incur any additional computational complexity over existing pilot detection schemes and its performance is independent of the length of the transmitted data.

To summarize the above, the drawbacks and research gaps found commonly in currently suggested solutions that overcome the pilot contamination problem include, but are not limited to, the following: 1) limited performance in pilot contamination suppression; 2) the requirement of additional hardware configuration; 3) excessive computational complexity; 4) need of BS coordination resulting in additional signal processing and transmission overhead; 5) reduction in the spectral efficiency; 6) need of excessive backhaul integration; 7) increase in the length of the training period; and 8) dependency on the user's mobility speed. The work we present in this chapter addresses these drawbacks and research gaps which highlights the novelty and importance of DPS.

2.4 System Model

In this section, we present a system model that can be used to evaluate any pilot signaling design including our DPS. We consider an uplink transmission in a wireless communication system with L cells in each cluster¹. Figure 2.1 illustrates a cell cluster with $L = 7$. Each cell has one BS located at the center, or alternatively an Access Point (AP), and serves up to K user equipment or terminals in its coverage area. Assume that each BS is equipped with M antennas and each UE is equipped with a single antenna. If $M \gg K$, then this wireless communication system is conventionally known as a massive MIMO system. We use (CELL INDEX, USER INDEX) to uniquely identify a UE in the system-cluster. For example, UE (i, j) means the j th UE in the i th cell. Let $\mathbf{h}_{ij}^l \in \mathbb{C}^M$ be the vector of uplink channel coefficients from UE (i, j) to the BS in the l th cell. Following the notations in

¹The concept of clustering here is the same as that in cellular network planning for frequency-reuse. It is assumed that cells in adjacent clusters are sufficiently separated in distance that mutual interference can be neglected.

[21] we model \mathbf{h}_{ij}^l as

$$\mathbf{h}_{ij}^l = \sqrt{\beta_{ij}^l} \mathbf{g}_{ij}^l \quad (2.1)$$

in which the large-scale fading coefficient β_{ij}^l is known *a priori* and the vector of small-scale fading coefficients \mathbf{g}_{ij}^l is drawn as i.i.d. from the complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, where $\beta_{ij}^l = (d^{-\rho})_{ij}^l$, ρ is the path loss exponent, and d is the distance between the i th UE in j th cell and BS in l th cell. For presentation simplicity, we begin in this section with the i.i.d. Rayleigh fading wireless communication channel in our system model, and will consider the correlated Rayleigh fading wireless communication channel in Section 2.6.

In every coherence interval of the wireless communication channel, each UE transmits a PS that consists of τ symbols to enable the BS to estimate the CSI of the link between the BS and the UE. That is, the vector of channel coefficients \mathbf{h}_{ij}^l is assumed to be constant throughout the coherence interval, which length is dependent on the mobility speed of UEs and the propagation environment. Let $\phi_{ij} \in \mathbb{C}^\tau$ be the PS of UE (i, j) . Then, the pilot signals received at the BS of the l th cell from all active UEs being served synchronously in a cluster is denoted by the matrix $\mathbf{Y}_l \in \mathbb{C}^{M \times \tau}$ and can be written as:

$$\mathbf{Y}_l = \sum_{(i,j)} \mathbf{h}_{ij}^l \phi_{ij}^T + \mathbf{W}_l, \quad (2.2)$$

where $\mathbf{W}_l \in \mathbb{C}^{M \times \tau}$ is the matrix of the additive channel noise whose entries are drawn as i.i.d. from $\mathcal{CN}(0, \sigma^2)$.

Upon receiving the $M \times \tau$ size matrix \mathbf{Y}_l , the BS of the l th cell estimates the vectors of channel coefficients for all the K active UEs in its coverage area $\{\mathbf{h}_{l1}^l, \dots, \mathbf{h}_{lK}^l\}$. A particular goal of the BS of the l th cell in the cluster is, for each UE (l, k) in the cell, to find the optimal estimate of the channel coefficient vector for the UE, i.e., $\hat{\mathbf{h}}_{lk}^l$ that minimizes the MSE from the actual channel coefficient vector \mathbf{h}_{lk}^l . That is,

$$\hat{\mathbf{h}}_{lk}^l = \arg \min_{\mathbf{h} \in \mathbb{C}^M} \mathbb{E} \left\{ \|\mathbf{h}_{lk}^l - \mathbf{h}\|_2^2 \right\}, \quad (2.3)$$

where the expectation is over the Rayleigh fading distribution. It has been shown in [93] that the optimal estimate $\hat{\mathbf{h}}_{lk}^l$ using the MMSE estimator is given by:

$$\hat{\mathbf{h}}_{lk}^l = \left(\beta_{lk}^l \phi_{lk}^H \otimes \mathbf{I}_M \right) (\mathbf{D}_l \otimes \mathbf{I}_M)^{-1} \text{vec}(\mathbf{Y}_l), \quad (2.4)$$

where \mathbf{D}_l is the covariance matrix and is given by:

$$\mathbf{D}_l = \sigma^2 \mathbf{I}_\tau + \sum_{(i,j)} \beta_{ij}^l \phi_{ij} \phi_{ij}^H. \quad (2.5)$$

In (2.4), \otimes is the Kronecker (tensor) matrix (vector) product operation and the $\text{vec}(\cdot)$ operation stacks the columns of a matrix into a single column vector. The corresponding MSE is given by:

$$\text{MSE}_{lk} = M \beta_{lk}^l - M \left(\beta_{lk}^l \right)^2 \phi_{lk}^H \mathbf{D}_l^{-1} \phi_{lk}. \quad (2.6)$$

The above MSE formula shows that achieving MMSE is subject to choosing an appropriate PS, and this formula applies to *any* design of PSs $\{\phi_{ij}\}$. While it is straightforward to evaluate a particular pilot design by using (2.6), it is nontrivial to propose a simple design that is easy to implement and is *universally better*² than the conventional design that uses orthogonal pilots. Our proposed DPS enjoys simple conditions under which it outperforms the conventional scheme for all UEs in the network. In addition, our DPS design contains only one design parameter that facilitates the implementation.

On the other hand, as mentioned above the wireless communication system under consideration becomes a massive MIMO when $M \gg K$ (i.e., the number of BS antennas in massive MIMO is much greater than the number of UEs being served in the cell). So, based on this system feature and according to the law of large numbers, massive MIMO implementation grants the closeness to a favorable propagation environment. Thus, simple linear estimators are close to optimal under semi-favorable propagation conditions [24].

²By universally better, we mean that a scheme provides better channel estimation than other schemes for every UE being served in the whole cellular network.

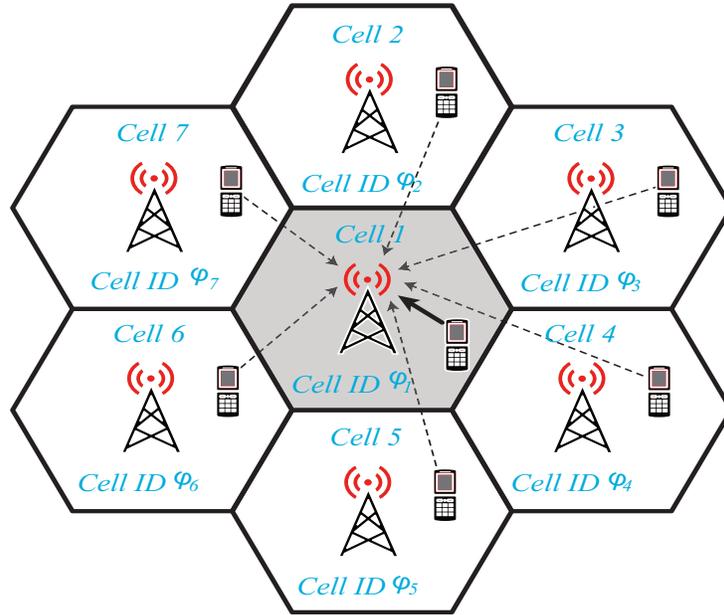


Figure 2.1: A cluster of seven non-cooperative adjacent cells with Cell-IDs assignment for DPS, the home cell is located in the center. $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_7\}$ are the assigned Cell-IDs sequences.

2.5 Dual Pilot Scheme

In this section, we present the proposed novel DPS as a general scheme that can be employed in any wireless communication system using pilot signaling for CSI estimation. Also, we show how DPS yields significantly better accuracy in estimating the CSI compared with the conventional scheme employing orthogonal pilot signals, referred as the SPS in this chapter, and would therefore lead to performance enhancement of a wireless communication system. This is proven through theoretical analyses under the i.i.d. Rayleigh fading wireless communication channel model, described in the previous section, and using the MMSE estimator, which is well-known to achieve the best performance in both low and high signal to interference-plus-noise ratios compared to other types of estimators [21, 93–95]. Although we consider only the MMSE estimator in our analyses which is used by the current state-of-the-art works, DPS is compatible with other types of estimators such as ZF and the Maximum Ratio Combining (MRC), and can achieve improved performance with these estimators as well.

2.5.1 SPS with Orthogonal Pilots

Consider N normalized orthogonal sequences, $\{\psi_1, \dots, \psi_k, \dots, \psi_N\}$, where each sequence $\psi_k \in \mathbb{C}^\tau$ is a normalized vector (i.e., $\|\psi_k\|_2 = 1$) and $\tau \geq N$. For SPS, the PS used by the intended UE (l, k) in the intended cell consists of a single code sequence, given by:

$$\phi_{lk} = \sqrt{P}\psi_k \quad (2.7)$$

where P represents the power constraint determined by the limited energy resources and system configuration. (Note that $\|\phi_{lk}\|_2^2 = P$ for any intended UE (l, k) .) Substituting (2.7) into (2.5) and (2.6), we obtain the MSE for SPS (with details given in Appendix A.1)

$$\text{MSE}_{lk}^{\text{SPS}} = M\beta_{lk}^l - M \left(\beta_{lk}^l \right)^2 \frac{P}{\sigma^2 + P \sum_i \beta_{ik}^l}. \quad (2.8)$$

2.5.2 DPS with Non-Orthogonal Pilots

The proposed DPS utilizes the same N normalized orthogonal code sequences as in SPS, but divide them into two groups as in $\{\psi_1, \dots, \psi_{N-L}, \varphi_1, \dots, \varphi_L\}$. The sequence φ_l is called the l th Cell Identification (Cell-ID), which is assigned to the l th cell in the cluster of L cells and is reused across system-clusters. The sequence ψ_k is called the User-Equipment Identification (User-ID), and is assigned to the k th UE in each cell in the cluster of L cells and is reused across system-clusters. Thus, the Cell-ID in-conjunction with the User-ID can uniquely identify any UE in the cluster of L cells. For our example of a seven-cells³ cluster, as shown in Figure 2.1, seven different Cell-IDs, given by sequences $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_7$ are used to identify the seven cells, the home cell is located at the center and has given number 1. The number of active UEs that can be served per system-cluster at this configuration is shown in Table 2.1, which compares the maximum numbers of UEs served per system-cluster for DPS, denoted as UC^{DPS} , and SPS, denoted as UC^{SPS} . Note that these numbers are limited by the number of PSs available for assignment. The actual

³This configuration can be generalized to any cluster size that has an arbitrary number of cells.

Table 2.1: Dual PSs assignments and comparisons of total numbers of UEs served in a system-cluster for SPS (UC^{SPS}) and DPS (UC^{DPS}) with $\tau = N = K$ and $L = 7$ at different values of M . Where M and N are the numbers of BS's antennas and available orthogonal sequences, respectively. $N - L$ and ID are the numbers of repeatedly reused sequences (User-IDs) and Cell-IDs sequences in the system-cluster, respectively.

M	$N = K$	$N - L$	$ID = L$	UC^{DPS}	UC^{SPS}
32	8	1	7	7	56
64	16	9	7	63	112
128	32	25	7	175	224
256	64	57	7	339	448
512	128	121	7	847	896
1024	256	249	7	1743	1792
2048	512	505	7	3535	3584

number of UEs that each cell would actually serve at any time depends on the SINR and might be lower than the upper limits shown in Table 2.1. It is important to mention that the duration of the training period within a coherence interval is the main factor that determines the number of available orthogonal sequences N . In practice, the value of N increases as the length of the training period increases, which implies longer coherence interval duration. The length of the training period is inversely proportional to spectral efficiency. Therefore, there is a trade-off between CSI accuracy, spectral efficiency, and system capacity. The optimal value of N should be chosen, in practice, to offer a higher system capacity subject to maintaining high spectral efficiency.

In DPS, as each UE (l, k) is uniquely identified by a pair of code sequences, we combine them into the so-called dual PS, which is given by:

$$\phi_{lk} = \sqrt{P} \left(\sqrt{\gamma} \psi_k + \sqrt{1 - \gamma} \varphi_l \right) \quad (2.9)$$

where $\gamma \in (0, 1)$ is a design parameter to be specified later. Note that $\|\phi_{lk}\|_2^2 = P$ for any intended UE (l, k) as before. In other words, DPS consumes the same amount of energy as SPS. However, as will be shown below, DPS enhances the CSI accuracy and thus it has a better energy efficiency

compared to SPS. Substituting (2.9) into (2.5), we obtain

$$\mathbf{D}_l^{\text{DPS}}(\gamma) = \sigma^2 \mathbf{I}_\tau + \gamma \sum_j \beta_j \boldsymbol{\psi}_j \boldsymbol{\psi}_j^H + (1 - \gamma) \sum_i \alpha_i \boldsymbol{\varphi}_i \boldsymbol{\varphi}_i^H + \sqrt{\gamma(1 - \gamma)} \sum_{(i,j)} P \beta_{ij}^l (\boldsymbol{\psi}_j \boldsymbol{\varphi}_i^H + \boldsymbol{\varphi}_i \boldsymbol{\psi}_j^H) \quad (2.10)$$

where $\alpha_i = P \sum_j \beta_{ij}^l$ and $\beta_j = P \sum_i \beta_{ij}^l$ are short-hand notations. Hence, the MSE for DPS is given by

$$\text{MSE}_{lk}^{\text{DPS}}(\gamma) = M \beta_{lk}^l - M \left(\beta_{lk}^l \right)^2 \boldsymbol{\phi}_{lk}^H \left(\mathbf{D}_l^{\text{DPS}}(\gamma) \right)^{-1} \boldsymbol{\phi}_{lk}. \quad (2.11)$$

Note that, when $\gamma = 1$, we have $\text{MSE}_{lk}^{\text{DPS}}(1) = \text{MSE}_{lk}^{\text{SPS}}$, which is as expected.

Next, we describe the CSI estimation procedure at the BS in the l th cell in order to provide our readers with more insights. Recall that the received signal is given by $\mathbf{Y}_l = \sum_{(i,j)} \mathbf{h}_{ij}^l \boldsymbol{\phi}_{ij}^T + \mathbf{W}_l$. Upon receiving \mathbf{Y}_l , the BS in the l th cell computes $\mathbf{Y}_l \boldsymbol{\psi}_j^*$ for $j \in \{1, \dots, N - L\}$ as well as $\mathbf{Y}_l \boldsymbol{\varphi}_i^*$ for $i \in \{1, \dots, L\}$ (where $\boldsymbol{\psi}_j^*$ and $\boldsymbol{\varphi}_i^*$ are complex conjugates of $\boldsymbol{\psi}_j$ and $\boldsymbol{\varphi}_i$, respectively). Note that

$$\mathbf{Y}_l \boldsymbol{\psi}_j^* = \sum_{(i,j')} \mathbf{h}_{ij'}^l \boldsymbol{\phi}_{ij'}^T \boldsymbol{\psi}_j^* + \mathbf{W}_l \boldsymbol{\psi}_j^* \quad (2.12)$$

$$= \sum_i \mathbf{h}_{ij}^l \sqrt{P} \left(\sqrt{\gamma} \boldsymbol{\psi}_j^T + \sqrt{1 - \gamma} \boldsymbol{\varphi}_i^T \right) \boldsymbol{\psi}_j^* + \mathbf{W}_l \boldsymbol{\psi}_j^* \quad (2.13)$$

$$= \sqrt{\gamma P} \sum_i \mathbf{h}_{ij}^l + \mathbf{W}_l \boldsymbol{\psi}_j^* \quad (2.14)$$

where $\mathbf{W}_l \boldsymbol{\psi}_j^* \in \mathbb{C}^M$ is the effective noise whose entries are drawn as i.i.d. from $\mathcal{CN}(0, \sigma^2)$. Similarly,

$$\mathbf{Y}_l \boldsymbol{\varphi}_i^* = \sum_{(i',j)} \mathbf{h}_{i'j}^l \boldsymbol{\phi}_{i'j}^T \boldsymbol{\varphi}_i^* + \mathbf{W}_l \boldsymbol{\varphi}_i^* \quad (2.15)$$

$$= \sum_j \mathbf{h}_{ij}^l \sqrt{P} \left(\sqrt{\gamma} \boldsymbol{\psi}_j^T + \sqrt{1 - \gamma} \boldsymbol{\varphi}_i^T \right) \boldsymbol{\varphi}_i^* + \mathbf{W}_l \boldsymbol{\varphi}_i^* \quad (2.16)$$

$$= \sqrt{(1 - \gamma) P} \sum_j \mathbf{h}_{ij}^l + \mathbf{W}_l \boldsymbol{\varphi}_i^* \quad (2.17)$$

where $\mathbf{W}_l \boldsymbol{\varphi}_i^* \in \mathbb{C}^M$ is the effective noise whose entries are drawn as i.i.d. from $\mathcal{CN}(0, \sigma^2)$. Then,

the BS in the l th cell applies the MMSE estimator for the intended UE (l, k) based on the N processed signals, namely $\{\mathbf{Y}_l \boldsymbol{\psi}_j^*\}_{j=1}^{N-L}$ and $\{\mathbf{Y}_l \boldsymbol{\varphi}_i^*\}_{i=1}^L$. The MSE of this estimator is given by (2.11) as explained before.

Finally, we turn our attention to finding a sub-optimal CSI estimator for two reasons. First, the sub-optimal estimator can be designed to be a “low-complexity approximation” of our optimal MMSE estimator described above. This would be of practical interest if we would like to reduce the computational complexity as well as the processing latency at the BS. Second, the sub-optimal estimator may allow us to derive an upper bound for the MSE given by (2.11). This bound will be used in the proofs of our main theoretical results.

Unlike the optimal MMSE estimator that relies on N processed signals, our proposed sub-optimal estimator only uses *two* processed signals $\mathbf{Y}_l \boldsymbol{\psi}_k^*$ and $\mathbf{Y}_l \boldsymbol{\varphi}_l^*$ given by

$$\mathbf{Y}_l \boldsymbol{\psi}_k^* = \sqrt{\gamma P} \sum_i \mathbf{h}_{ik}^l + \mathbf{W}_l \boldsymbol{\psi}_k^* \quad (2.18)$$

and

$$\mathbf{Y}_l \boldsymbol{\varphi}_l^* = \sqrt{(1-\gamma)P} \sum_j \mathbf{h}_{lj}^l + \mathbf{W}_l \boldsymbol{\varphi}_l^*. \quad (2.19)$$

Then, the sub-optimal estimator conducts the MMSE estimation for the intended UE (l, k) based on $\mathbf{Y}_l \boldsymbol{\psi}_k^*$ and $\mathbf{Y}_l \boldsymbol{\varphi}_l^*$ only. That is,

$$\hat{\mathbf{h}}_{lk}^l = \left(\boldsymbol{\beta}_{lk}^l [\sqrt{\gamma P}, \sqrt{(1-\gamma)P}] \otimes \mathbf{I}_M \right) (\mathbf{D}_l^{\text{SUB}}(\gamma) \otimes \mathbf{I}_M)^{-1} \begin{bmatrix} \mathbf{Y}_l \boldsymbol{\psi}_k^* \\ \mathbf{Y}_l \boldsymbol{\varphi}_l^* \end{bmatrix}, \quad (2.20)$$

where the covariance matrix $\mathbf{D}_l^{\text{SUB}}(\gamma)$ is given by

$$\mathbf{D}_l^{\text{SUB}}(\gamma) = \begin{bmatrix} \gamma P \sum_i \beta_{ik}^l + \sigma^2 & \sqrt{\gamma(1-\gamma)P} \beta_{lk}^l \\ \sqrt{\gamma(1-\gamma)P} \beta_{lk}^l & (1-\gamma)P \sum_j \beta_{lj}^l + \sigma^2 \end{bmatrix}. \quad (2.21)$$

The corresponding MSE is

$$\text{MSE}_{lk}^{\text{SUB}}(\gamma) = M\beta_{lk}^l - M\left(\beta_{lk}^l\right)^2 P \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}^T (\mathbf{D}_l^{\text{SUB}}(\gamma))^{-1} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}. \quad (2.22)$$

Clearly, we have $\text{MSE}_{lk}^{\text{SUB}}(\gamma) \geq \text{MSE}_{lk}^{\text{DPS}}(\gamma)$ since the above estimator is suboptimal (as it is based on only two processed signals).

2.5.3 Characterizing the Advantages of DPS

We will show that DPS is *universally better* than SPS under certain conditions. That is, there exists a parameter γ such that $\text{MSE}_{lk}^{\text{DPS}}(\gamma) < \text{MSE}_{lk}^{\text{SPS}}$ for each and every intended UE (l, k) in the l th cell. Therefore, DPS provides a better channel estimation for each and every user compared to the conventional SPS. DPS gives the BS the ability to do this minimization operation of estimation errors for each and every user individually. Consequently, the BS can focus on improving the CSI accuracy of users being served at cell's edges to enhance their QoS, as a goal that ensures promoting fairness and uniform QoS for the whole users being served. We have the following theorem.

Theorem 1 *If the parameter $\gamma \in (0, 1)$ is chosen such that*

$$\frac{1-\gamma}{\gamma} \frac{\sum_{j \neq k} \beta_{lj}^l}{\left(\sum_{i \neq l} \beta_{ik}^l\right)^2} < \frac{P}{\sigma^2} \text{ for all } k, \quad (2.23)$$

then $\text{MSE}_{lk}^{\text{DPS}}(\gamma) < \text{MSE}_{lk}^{\text{SPS}}$ for all k .

Proof: The proof consists of two steps. First, we have $\text{MSE}_{lk}^{\text{SUB}}(\gamma) \geq \text{MSE}_{lk}^{\text{DPS}}(\gamma)$ as explained before. Second, we show that our suboptimal estimator still achieves a smaller MSE than the SPS for each and every intended UE (l, k) under condition (2.23). Therefore, we conclude that $\text{MSE}_{lk}^{\text{DPS}}(\gamma) < \text{MSE}_{lk}^{\text{SPS}}$ for all k under condition (2.23).

We now focus on Step 2, showing that $\text{MSE}_{lk}^{\text{SUB}}(\gamma) < \text{MSE}_{lk}^{\text{SPS}}$ for all k under condition (2.23).

Comparing (2.22) with (2.8), it suffices for us to prove that

$$\begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}^T (\mathbf{D}_l^{\text{SUB}}(\gamma))^{-1} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix} > \frac{1}{\sigma^2 + P \sum_i \beta_{ik}^l}. \quad (2.24)$$

The details are given in Appendix A.2. \square

Theorem 2 *The optimal γ for the intended UE (l, k) under the suboptimal estimator is given by*

$$\gamma^* = \frac{\sum_{j \neq k} \beta_{lj}^l}{\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l}. \quad (2.25)$$

Proof: Taking the derivative of $g_{lk}^{\text{SUB}}(\gamma) = \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}^T (\mathbf{D}_l^{\text{SUB}}(\gamma))^{-1} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}$, we can show that $g_{lk}^{\text{SUB}}(\gamma)$ is unimodal over the interval $(0, 1)$ and γ^* is the only zero of its derivative. The details are given in Appendix A.3. Therefore, γ^* maximizes $g_{lk}^{\text{SUB}}(\gamma)$ for UE (l, k) . \square

Note that $\sum_{j \neq k} \beta_{lj}^l$ can be viewed as ‘‘intra-cell interference’’ since all the UEs in the l th cell share the same sequence φ_l . Similarly, $\sum_{i \neq l} \beta_{ik}^l$ can be viewed as ‘‘inter-cell interference’’ as UE (i, k) share the same sequence ψ_k (i.e., every k th UE in each cell in the system-cluster share ψ_k). Hence, Theorem 2 says that γ^* is the fraction of intra-cell interference over the total interference. This provides us with some design guidelines. For example, γ^* is relatively small when the cells are densely packed (which implies a relatively large $\sum_{i \neq l} \beta_{ik}^l$) and each cell supports a smaller number of UEs (which implies a relatively small $\sum_{j \neq k} \beta_{lj}^l$).

Nevertheless DPS advantages come with a price of serving L fewer UEs than with SPS, where L is the number of cells in the system-cluster, assuming the same N orthogonal sequences are used. This is because SPS uses all N orthogonal PSs to serve the K UEs in each cell, i.e., $K = N$, whereas DPS needs to use L of these sequences as the Cell-IDs and can only assign $N - L$ sequences to the UEs in each cell. However, as shown in Table 2.1, this disadvantage diminishes as K grows large. Nevertheless, DPS ensures a uniform high QoS for each and every UE and enhances

the system capacity indirectly by improving the CSI accuracy which is essential to maximize the data throughput. Moreover, despite the number of PSs available for assignment to UEs in a cell determines the maximum number of UEs that can be served simultaneously, in practice, the cell can only serve a smaller number of UEs due to pilot contamination effects, in the case of SPS, where an acceptable level of interference is the constraint.

2.6 Correlated Rayleigh Fading

In this section, we extend our discussion to the correlated Rayleigh fading channel model. We assume that the small-scale fading vector \mathbf{g}_{ij}^l is drawn from $\mathcal{CN}(\mathbf{0}, \mathbf{R}_{ij}^l)$ where $\mathbf{R}_{ij}^l \in \mathbb{C}^{M \times M}$ is the covariance matrix. The remaining settings are the same as in Section 2.4. Therefore, for a correlated Rayleigh fading channel, it has been shown in [93] that the optimal estimates of the vector of channel coefficients for the intended UE in the intended cell (i.e., $\hat{\mathbf{h}}_{lk}^l$) using the MMSE estimator at the BS in the l th cell is given by:

$$\hat{\mathbf{h}}_{lk}^l = \mathbf{Z}_{lk} \mathbf{U}_l^{-1} \text{vec}(\mathbf{Y}_l), \quad (2.26)$$

where $\mathbf{Z}_{lk} \in \mathbb{C}^{M \times \tau M}$ and $\mathbf{U}_l \in \mathbb{C}^{\tau M \times \tau M}$ are given by

$$\mathbf{Z}_{lk} = \beta_{lk}^l \phi_{lk}^H \otimes \mathbf{R}_{lk}^l \quad (2.27)$$

and

$$\mathbf{U}_l = \sigma^2 \mathbf{I}_{\tau M} + \sum_{(i,j)} \beta_{ij}^l \phi_{ij} \phi_{ij}^H \otimes \mathbf{R}_{ij}^l. \quad (2.28)$$

The corresponding MSE is

$$\text{MSE}_{lk} = \beta_{lk}^l \text{tr}(\mathbf{R}_{lk}^l) - \text{tr}(\mathbf{Z}_{lk} \mathbf{U}_l^{-1} \mathbf{Z}_{lk}^H). \quad (2.29)$$

Note that when $\mathbf{R}_{ij}^l = \mathbf{I}_M$ for all (i, j) , then (2.29) reduces to (2.6), i.e., the MMSE for the i.i.d. (uncorrelated) case.

2.6.1 SPS with Orthogonal Pilots

Following the steps in Section 2.5.1, we obtain the MSE for the SPS under correlated Rayleigh fading

$$\text{MSE}_{lk}^{\text{SPS}} = \beta_{lk}^l \text{tr}(\mathbf{R}_{lk}^l) - P \left(\beta_{lk}^l \right)^2 \text{tr} \left(\mathbf{R}_{lk}^l \left(\sigma^2 \mathbf{I}_M + \sum_i \beta_{ik}^l P \mathbf{R}_{ik}^l \right)^{-1} \left(\mathbf{R}_{lk}^l \right)^H \right). \quad (2.30)$$

In particular, if $\mathbf{R}_{ij}^l = \mathbf{I}_M$ for all (i, j) , then (2.30) reduces to (2.8), i.e., the MMSE of the SPS for the i.i.d. (uncorrelated) case.

2.6.2 DPS with Non-Orthogonal Pilots

Substituting the dual PS, given by (2.9), $\phi_{lk} = \sqrt{P} (\sqrt{\gamma} \psi_k + \sqrt{1-\gamma} \varphi_l)$ that used by the intended user (l, k) into (2.28), we obtain

$$\mathbf{U}_l^{\text{DPS}}(\gamma) = \sigma^2 \mathbf{I}_{\tau M} + P \sum_{(i,j)} \beta_{ij}^l \left(\gamma \psi_j \psi_j^H + (1-\gamma) \varphi_i \varphi_i^H + \sqrt{\gamma(1-\gamma)} (\psi_j \varphi_i^H + \varphi_i \psi_j^H) \right) \otimes \mathbf{R}_{ij}^l. \quad (2.31)$$

The corresponding MSE is given by

$$\text{MSE}_{lk}^{\text{DPS}}(\gamma) = \beta_{lk}^l \text{tr}(\mathbf{R}_{lk}^l) - \text{tr} \left(\mathbf{Z}_{lk} \left(\mathbf{U}_l^{\text{DPS}}(\gamma) \right)^{-1} \mathbf{Z}_{lk}^H \right). \quad (2.32)$$

2.6.3 Characterizing the Advantages of DPS

Recall that the BS in the l th cell applies the MMSE estimator for the intended UE (l, k) based on N processed signals, namely $\{\mathbf{Y}_l \psi_j^*\}_{j=1}^{N-L}$ and $\{\mathbf{Y}_l \varphi_i^*\}_{i=1}^L$. However, similar to the i.i.d. case, we will here construct a suboptimal estimator that uses only *two* processed signals as follows. Upon receiving the signals from all users \mathbf{Y}_l , the BS in l th cell computes $\mathbf{Y}_l \psi_k^*$ and $\mathbf{Y}_l \varphi_l^*$ and then applies

the MMSE estimator based on $\mathbf{Y}_l \boldsymbol{\psi}_k^*$ and $\mathbf{Y}_l \boldsymbol{\varphi}_l^*$ only. That is,

$$\hat{\mathbf{h}}_{lk}^l = \left(\beta_{lk}^l [\sqrt{\gamma P}, \sqrt{(1-\gamma)P}] \otimes \mathbf{R}_{lk}^l \right) (\mathbf{U}_l^{\text{SUB}}(\gamma))^{-1} \begin{bmatrix} \mathbf{Y}_l \boldsymbol{\psi}_k^* \\ \mathbf{Y}_l \boldsymbol{\varphi}_l^* \end{bmatrix}, \quad (2.33)$$

where the covariance matrix $\mathbf{U}_l^{\text{SUB}}(\gamma)$ is given by

$$\begin{aligned} \mathbf{U}_l^{\text{SUB}}(\gamma) &= \sigma^2 \mathbf{I}_{2M} + \sum_{i \neq l} \beta_{ik}^l \begin{bmatrix} \gamma P & 0 \\ 0 & 0 \end{bmatrix} \otimes \mathbf{R}_{ik}^l + \sum_{j \neq k} \beta_{lj}^l \begin{bmatrix} 0 & 0 \\ 0 & (1-\gamma)P \end{bmatrix} \otimes \mathbf{R}_{lj}^l \\ &+ \beta_{lk}^l \begin{bmatrix} \gamma P & \sqrt{\gamma(1-\gamma)P} \\ \sqrt{\gamma(1-\gamma)P} & (1-\gamma)P \end{bmatrix} \otimes \mathbf{R}_{lk}^l. \end{aligned} \quad (2.34)$$

The corresponding MSE is

$$\text{MSE}_{lk}^{\text{SUB}}(\gamma) = \beta_{lk}^l \text{tr}(\mathbf{R}_{lk}^l) - P \left(\beta_{lk}^l \right)^2 \text{tr} \left(\begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}^T \otimes \mathbf{R}_{lk}^l (\mathbf{U}_l^{\text{SUB}}(\gamma))^{-1} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix} \otimes (\mathbf{R}_{lk}^l)^H \right). \quad (2.35)$$

Clearly, in this case we will have $\text{MSE}_{lk}^{\text{SUB}}(\gamma) \geq \text{MSE}_{lk}^{\text{DPS}}(\gamma)$ since the above estimator is suboptimal.

We are particularly interested in a case when all the UEs have the same covariance matrix \mathbf{R}^l for the small-scale fading at the BS in the l th cell. (That is, $\mathbf{R}_{ij}^l = \mathbf{R}^l$ for all (i, j) .) This case allows us to have simplified expressions

$$\mathbf{U}_l^{\text{SUB}}(\gamma) = \sigma^2 \mathbf{I}_{2M} + \begin{bmatrix} \gamma P \sum_i \beta_{ik}^l & \sqrt{\gamma(1-\gamma)P} \beta_{lk}^l \\ \sqrt{\gamma(1-\gamma)P} \beta_{lk}^l & (1-\gamma)P \sum_j \beta_{lj}^l \end{bmatrix} \otimes \mathbf{R}^l \quad (2.36)$$

and

$$\text{MSE}_{lk}^{\text{SUB}}(\gamma) = \beta_{lk}^l \text{tr}(\mathbf{R}^l) - P \left(\beta_{lk}^l \right)^2 \text{tr} \left(\left[\begin{array}{c} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{array} \right]^T \otimes \mathbf{R}^l (\mathbf{U}_l^{\text{SUB}}(\gamma))^{-1} \left[\begin{array}{c} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{array} \right] \otimes (\mathbf{R}^l)^H \right). \quad (2.37)$$

Next, we introduce two short-hand notations:

$$f_{lk}(\lambda) = \text{tr} \left(\left[\begin{array}{c} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{array} \right]^T \otimes \mathbf{R}^l (\mathbf{U}_l^{\text{SUB}}(\gamma))^{-1} \left[\begin{array}{c} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{array} \right] \otimes (\mathbf{R}^l)^H \right), \quad (2.38)$$

and

$$g_{lk} = \text{tr} \left(\mathbf{R}^l \left(\sigma^2 \mathbf{I}_M + \sum_i \beta_{ik}^l P \mathbf{R}^l \right)^{-1} (\mathbf{R}^l)^H \right). \quad (2.39)$$

By comparing (2.37) and (2.30), we have the following theorem.

Theorem 3 When $\mathbf{R}_{ij}^l = \mathbf{R}^l$ for all (i, j) , the suboptimal estimator of DPS is universally better than the MMSE estimator of SPS if and only if $f_{lk}(\lambda) > g_{lk}$ for all k .

Proof: When $\mathbf{R}_{ij}^l = \mathbf{R}^l$ for all (i, j) , then $\text{MSE}_{lk}^{\text{SUB}}(\gamma) < \text{MSE}_{lk}^{\text{SPS}}$ if and only if $f_{lk}(\lambda) > g_{lk}$. \square

Note that if $\mathbf{R}^l = \mathbf{I}_M$, the condition $f_{lk}(\lambda) > g_{lk}$ reduces to the condition (2.24), which is proven in Appendix A.2. Theorem 3 allows us to numerically find the value of λ for any given covariance matrix \mathbf{R}^l such that our suboptimal estimator is *universally better* than the SPS.

2.7 DPS in Massive MIMO Cellular Networks: Macroscopic View

The previous sections proved that DPS is a superior mechanism that can be integrated into any wireless communication system, including massive MIMO, to enhance CSI estimation accuracy and consequently improve the system's performance. As it is highly expected that massive MIMO will be used in practical implementations of 5G cellular networks in the coming years, a question naturally arises: "How can we integrate DPS into the practical implementation of massive MIMO?"

In this section, we answer this question by highlighting the practical features of DPS, comparing DPS with other recently proposed schemes, and providing a *macroscopic* view of how the entire cellular network's planning can be configured with the DPS integration. Thereafter, we propose a system-cluster size that utilizes DPS assignments efficiently and provide a detailed illustration of its features in massive MIMO.

As aforementioned, pilot contamination is the main inherent impairment of massive MIMO. Different approaches have been presented in Section 2.3 that provide suggested solutions to overcome the pilot contamination. However, these solutions suffer drawbacks and their effectiveness is subjected to limited scenarios and certain conditions (i.e., specific environments and system configurations that cannot be generalized) [17, 34, 42, 44, 47, 50–52]. Clearly, the existence of these drawbacks forms obstacles that may prevent these recently suggested solutions to be adopted in practical implementations. In contrast, our proposed novel DPS mechanism fills these gaps and mitigates their limitations or shortcomings by offering the following features: 1) easy to implement; 2) incurs the same overhead and the same length of the training period as SPS; 3) does not require additional hardware or a change in the cellular network backhaul; 4) reduces the pilot contamination and resulting CSI errors significantly; 5) independent of user's mobility speed; and 6) improved spectral efficiency. Therefore, DPS offers the feasibility for implementation in practical environments of 5G cellular networks and beyond.

Figure 2.2 illustrates the performance comparisons between DPS and some recently proposed pilot contamination mitigation schemes for enhancing channel estimation. We set up the systems' parameters to form identical or closest environments to ensure fair comparisons. Besides, the NMSE has been selected as the performance metric. From the figure, it can be seen that DPS outperforms the compared schemes by achieving lower MMSEs over the considered range of SNR and also ensures a uniform QoS for each and every user as demonstrated in Figure 2.2, which shows how DPS outperforms other suggested solutions. DPS treats edge UEs uniformly as central UEs and ensures to achieve the smallest MMSE while estimating the CSI for every UE being served in the cell, not as other schemes where UEs either are classified into edge UEs and central UEs or BS

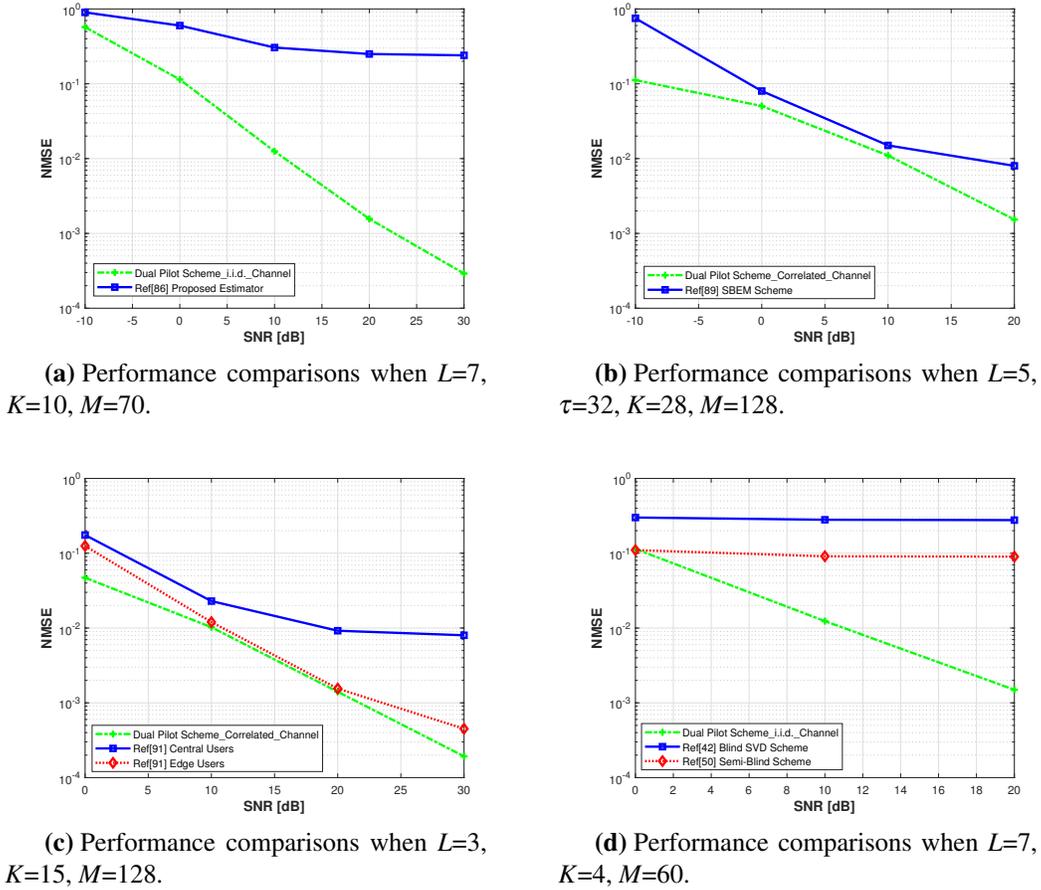


Figure 2.2: Performance comparisons between DPS and some existing schemes for pilot contamination mitigation.

focuses on achieving the smallest average MMSE not the smallest MMSE for each and every UE.

2.7.1 Integration Plan of DPS into Massive MIMO

Recall, DPS is based on using two sets of (orthogonal) sequences as demonstrated in the previous sections. The first set $\{\varphi_1, \dots, \varphi_L\}$ is used to identify the cells and is thus called Cell-ID sequences, where L is the cluster size based on the planning of the cellular network. The second set $\{\psi_1, \dots, \psi_{N-L}\}$ is the set of User-ID sequences, each assigned to a different UE in each cell, where in DPS $N \geq (K + L)$ and K is the maximum number of UEs being served simultaneously in each cell based on the capacity of the cellular network and $N \geq K$ in SPS. The User-ID sequences are reassigned (reused) in all the cells in a system-cluster and in all other clusters. The number of

orthogonal code sequences N is determined by the characteristics of the coherence interval in the TDD system as explained before. It is important to re-emphasize that with DPS these cells work in a non-cooperative mode of operation, which is a desirable approach in practical implementation.

Table 2.1 shows the dual PSs assignments in a system-cluster consisting of $L = 7$ non-cooperative adjacent cells, at different configurations of M . It also shows that the DPS becomes more efficient as we increase M and K . In Table 2.1, N is the number of available orthogonal sequences in the codebook, ID is the number of sequences used to identify the cells uniquely within a cluster (i.e., Cell-IDs), $N - L$ is the number of repeatedly used sequences (i.e., User-IDs), and UC^{DPS} is the number of UEs (per cluster) that being served and use the dual PSs in the system.

According to Figure 2.1, so far we have considered the performance of DPS within the center cell of a system-cluster, which would be the worst case performance when there is no other cell outside of the cluster. Nevertheless, in practice a cellular network may have any number of cells. In the following, we show that the above results remain valid when we add more clusters to expand the coverage area. While the Cell-ID assigned to each cell in a cluster is unique, when more clusters are added there will be multiple cells with the same Cell-ID. This calls for cell planning, which is similar in concept to frequency-reuse planning in legacy cellular networks. The purpose of cell planning is to ensure that two cells with the same Cell-ID are sufficiently far apart so that their signals do not interfere with each other.

Figure 2.3 illustrates cell planning with seven-cell clusters (i.e., $L = 7$), in which each hexagon represents a cell and seven colors are used to identify seven group of cells, each cell assigned with one of the seven Cell-IDs. We note that a dual PS constructed using (2.9) only guarantees its uniqueness within a cluster; i.e., dual PSs are reused in cells with the same colour in Figure 2.3, where BSs are ≈ 2.3 cell-diameters (referred as the reuse distance) away from each other. If the link budget guarantees that signal power received at any home cell from cells that are at least one reuse distance away is negligible, then we can assume that the dual PSs in one cluster do not contaminate the dual PSs in adjacent clusters or clusters farther away, and the mathematical analyses presented in the previous sections are still valid and can be extended in a straightforward

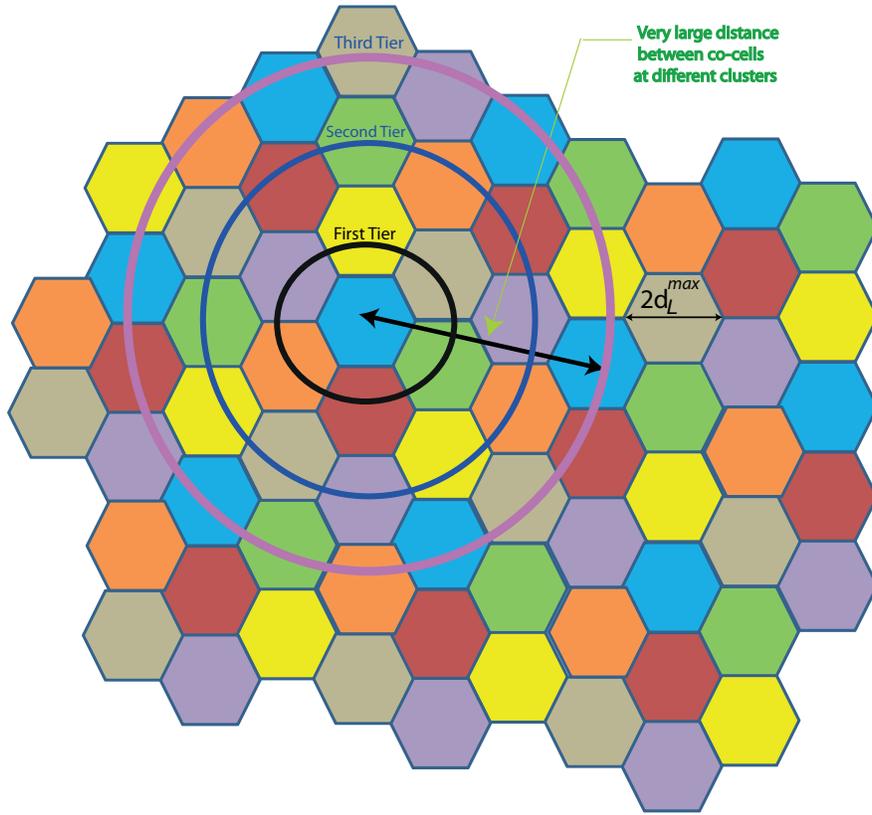


Figure 2.3: Dual Pilot Scheme with Cell-IDs and cluster reuse. Cells with the same color are using the same unique BS’s dual PS identifications (i.e., the same Cell-ID and User-IDs).

manner. The “inter-cell interference” $\sum_{i \neq l} \beta_{ik}^l$ remains the same because users (i, k) share the same sequence ψ_k . However, the “co-cell interference” which arises due to the reuse of Cell-IDs and the existence of co-cells (i.e., cells use the same Cell-IDs) contains two terms: one is still $\sum_{j \neq k} \beta_{ij}^l$, and the other is $\sum_{j \neq k} \beta_{l'j}^l$ for all other cells l' of the same color (i.e., Cell-ID) as cell l . Assuming the total number of cells the network needs to cover its coverage area is C . Then, as we will soon see, the second term is often small. So, setting $L = 7$ and here the number of clusters needed is $\lceil \frac{C}{7} \rceil$ gives us almost the same estimation accuracy as setting $L = C$ where the number of clusters needed here $\lceil \frac{C}{C} \rceil = 1$. In other words, the impact of reusing Cell-IDs on performance degradation is negligible.

Thus, the cluster size depends on the reuse distance, which in turn depends on the radio environ-

ment. The choice of a cluster size of seven is conservative. In this regard, SPS can be considered a special case of DPS with a cluster size $L = 1$, in which case the use of a Cell-ID sequence becomes unnecessary and all the available orthogonal PSs are assigned to different UEs in each cell.

Therefore, according to Figure 2.3, the pilot signals received at the BS of the l th home cell from all active UEs being served synchronously in its cell, first-tier cells, second-tier cells, and third-tier cells can be denoted by the matrix $\mathbf{Y}_l \in \mathbb{C}^{M \times \tau}$ and can be written as:

$$\mathbf{Y}_l = \sum_{j=1}^K (d_l^{-\frac{\rho}{2}})_{ij}^l \mathbf{g}_{ij}^l \phi_{ij}^T + \sum_{\substack{i=6 \\ j=K}}^{i=12} (d_l^{-\frac{\rho}{2}})_{ij}^l \mathbf{g}_{ij}^l \phi_{ij}^T + \sum_{\substack{i=12 \\ j=K}}^{i=18} (d_{t_2}^{-\frac{\rho}{2}})_{ij}^l \mathbf{g}_{ij}^l \phi_{ij}^T + \sum_{\substack{i=18 \\ j=K}}^{i=24} (d_{t_3}^{-\frac{\rho}{2}})_{ij}^l \mathbf{g}_{ij}^l \phi_{ij}^T + \mathbf{W}_l,$$

where $d_l < d_{t_1} < d_{t_2} < d_{t_3}$, d_l is the distance between UEs and BS in the l th home cell, d_{t_1} is the distance between UEs in any cell belongs to a first tier and BS of l th cell, d_{t_2} is the distance between UEs in any cell belongs to a second tier and BS of l th cell, d_{t_3} is the distance between UEs in any cell belongs to a third tier and BS of l th cell. Clearly, $0 \leq d_l < d_l^{max}$, where d_l^{max} equals to the cell radius. Also, $d_l^{max} \leq d_{t_1} \leq \sqrt{10} d_l^{max}$, $2 d_l^{max} \leq d_{t_2} \leq \sqrt{26} d_l^{max}$, and $\sqrt{15} d_l^{max} \leq d_{t_3} \leq \sqrt{50} d_l^{max}$. The co-Cell-IDs interference started to arise at the third tier, according to locations of cells having co-Cell-IDs of l th cell (i.e., same color) in Figure 2.3, the average distance between UEs in any of these cells to l th cell range from $\sqrt{28} - 1 \approx 4.3 d_l^{max}$ up to $\sqrt{28} + 1 \approx 6.3 d_l^{max}$. Consequently, the average increment in path loss of received co-Cell-IDs interfere signals at BS of a home cell will be in the range of $\{-6.33\rho \text{ dB}, -7.99\rho \text{ dB}\}$ added to the maximum path loss of the desired received signals.

Therefore, it is evident that the co-cell interference, due to Cell-IDs reuse, between the centric cell at any cluster and each of its co-cells in adjacent clusters is negligible due to a very large distance between them that result in a very strong path loss. Remarkably, reusing Cell-IDs in practice allows more active users to be served simultaneously which increases the capacity of the cellular network without any harmful interference penalty.

2.8 Additional Desirable Features of DPS

The previous sections have demonstrated and proved the main feature of the proposed novel DPS in enhancing the CSI of the UEs being served. Besides, the features of DPS are multi-fold, this section highlights these additional features that can be used in the practical improvement of wireless communications systems with the main focus on 5G cellular networks and beyond. Figure 2.4 illustrates the integration structure that will be used at any BS (Alternatively an AP) in any cell of the system to incorporate the DPS. For presentation simplicity, it shows the structure that is used at a BS of a system-cluster that consists of two non-cooperative adjacent cells that serve up to K active UEs in each cell. This structure can be generalized to a BS that is equipped with an M antennas and located within a cluster of L cells. The DPS additional features can include, but are not limited to, interference alignment, smart pilot assignments, and updating the second-order statistics of the communication channels at the BS in a periodically real-time manner.

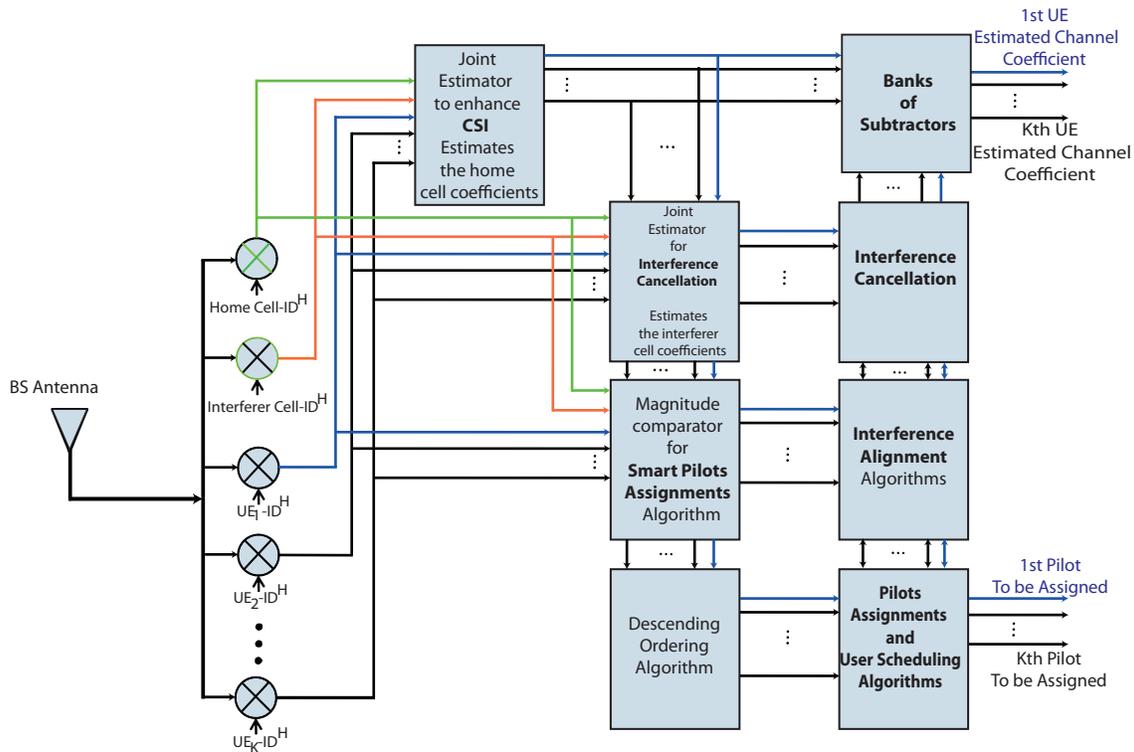


Figure 2.4: DPS Structural integration for a cluster of two non-cooperative adjacent cells where the BS is equipped with one antenna and serves K UEs.

2.8.1 Interference Alignment and Interference Cancellation

Interference alignment is a strategy that has been adopted recently to be integrated into a transceiver to increase the sum-rate capacity of the cell in cellular networks. Interference alignment uses linear precoding to align the interfering signals in code, frequency, space, or time such that the interference signal lies in a reduced dimensional subspace at each intended receiver. However, this transmission strategy requires coordination and feedback information which will result in many challenges such as excessive backhaul integration, overhead, and reduces spectral efficiency. According to Figure 2.4, DPS provides the BS's transceiver with a periodically real-time feedback CSI about all adjacent cells without any coordination or backhaul cooperation integration. Therefore, this available real-time information, at the BS of the home cell, about the adjacent cells can be exploited to build efficient interference alignment algorithms in the downlink transmissions and also interference cancellation algorithms in the uplink receptions at the BS.

2.8.2 Smart Pilot Assignment and User Scheduling

Smart pilot assignment algorithm is responsible to select which pilot in the code-book is suitable to serve a certain UE based on the real-time system configuration and communications environment. Such an algorithm is essential to maintain a lower bound of both intra-cell interference and inter-cell interference which consequently guaranteed the optimality of the system performance and can be integrated into efficient users scheduling algorithms. In contrast to SPS, where an efficient pilot assignment is very challenging, in DPS such an algorithm is easy to adopt in a methodical and real-time manner. As shown in Figure 2.4, the BS of the home cell in DPS are provided with real-time CSI periodically for both UEs belong to its coverage area and also for UEs being served in the adjacent cells, i.e., interfering cells. This available information can be used as inputs to smart algorithms to add a distinguishability feature to the BS that provides the knowledge to check if a User-ID sequence is currently in use at the adjacent cell/s or it is idle. Thus, a User-ID sequence that is not assigned at adjacent cells will be given a high priority of assignment to be used by one of UEs being served in the coverage area of the BS.

2.8.3 Periodically Updated the Second-Order Statistics

Theoretically, statistical CSI is usually assumed to be available at the BS and perfectly known. However, in practice this assumption is difficult to realize precisely and a solution is needed to overcome this challenge. Also, the large scale fading coefficients depend on the separation distance between the BS's antenna and the UE's antenna. Beside, MIMO technology is adopted in most 5Gs networks, with a large MIMO dimension in massive MIMO where the large scale fading has an importance as it is the main significant fading effect [24]. Therefore, it is very important for a BS to acquire a precise real-time second-ordered statistics and updating the large scale fading coefficients which should be done continuously. On the one hand DPS provides the BS with a more accurate CSI periodically. On the other hand, according to Figure 2.4, the statistical level of CSI is extracted over a long term and DPS enables the BS to do this extraction mechanism accurately in a real-time manner, not only for its home cell but also for its adjacent cells. The received Cell-IDs pilot signals at the BS can be exploited jointly with the other received signals to extract these second-order statistics. Thus, the proposed DPS in this dissertation, brings reality a practical system that can start with a transient second-order statistics and then reach a real-time second-order statistics after a specific period of employing DPS, which its length is determined based on the system configuration and communications environment.

2.9 Numerical Results and Discussions

To mimic the real environment and assess the performance of the proposed DPS in comparison with SPS, numerical results are presented and evaluated in this section. We consider a cluster of seven cells in a system with an urban outdoor multi-micro hexagonal cells massive MIMO configuration. The home cell under consideration is located at the center of the cluster and surrounded by six adjacent first-tier cells that form the source of pilot contamination as shown in Figure 2.1. Each cell has K active users. To generate the orthogonal PSs for SPS we used Walsh Hadamard matrices of size $\tau_p \times \tau_p$, where $\tau_p = K + 7$ with complex entries. Each column in this matrix represents a PS that is orthogonal to those represented by the other columns. Subsequently, to generate the dual

Table 2.2: The parameters used in simulations

Parameter Name	Setting
Number of BS Antennas M	32, 64, and 128
Carrier Frequency	2 GHz
BS Antenna Height	25 m
Cell Radius	600 m
Min UE-BS distance	75 m
Bandwidth	20 MHz
Number of UEs K	8 at each cell
UE Height	1.7 m
Pathloss and Lognormal Shadowing	From [96]
Noise Power Spectral Density	-174 dBm/Hz

PSs for use in DPS, the forward Walsh Hadamard transform is used to rearrange the columns of this matrix in ascending order with respect to the order of the sequences (i.e., number of zero-crossings or number of changes in the sign of entries in every column). The columns with the seven highest order of the sequences are assigned as Cell-IDs of the seven cells. The remaining sequences are used as User-IDs. A (Cell-ID, User-ID) pair of orthogonal sequences forms a dual PS based on (2.9) and uniquely identify a UE in the cluster. We assume the UEs in every cell are uniformly distributed around the cell's BS with equal power allocation. The simulation considers the worst-case scenario and uses synchronous transmissions of PSs in all cells in each configuration either when using a single PS for SPS or a dual PS for DPS. Table 2.2 shows the simulation parameters.

We also assume that the scatterers are distributed randomly in each cell and that a frequency flat fading channel during each frame, where each frame occupies the channel coherence interval. Each coherence interval is divided into four parts, the uplink training interval, the uplink data transmission interval, the downlink beam-forming (i.e., linear pre-coding) interval and the downlink data transmission interval. We use MATLAB as the simulation environment. Furthermore, we have done a calibration for the simulation environment by choosing simulation's configurations identical to those used by the existing MIMO LTE cellular system (i.e., at reducing the number of BS's antennas), the simulation has produced results that matched the benchmark results. Lastly, due to

the theoretical superiority that Linear MMSE (LMMSE) has shown amongst other estimation algorithms (as aforementioned) and for evaluating the impacts of DPS practical implementation, we assume the BS employs the LMMSE estimation method in the simulation environment.

2.9.1 Normalized Mean Squared Error

The NMSE is used to evaluate the proposed novel DPS and compare it with the SPS. Considering the m th element of \mathbf{h}_{ij}^l , then the NMSE of the estimates of the wireless communication channel coefficient between the j th user in the l th cell and the m th base station antenna at the l th cell h_{ij}^{lm} is given by $\text{NMSE}_{ij}^{lm} = \frac{1}{N} \sum_{n=1}^N \mathbb{E} \left\{ \frac{\|\hat{h}_{ij,n}^{lm} - h_{ij}^{lm}\|^2}{\|h_{ij}^{lm}\|^2} \right\}$ where \hat{h}_{ij}^{lm} is the estimated channel coefficient, h_{ij}^{lm} is the actual channel coefficient and n is the simulation index. Figure 2.5 shows the NMSE in the correlated Rayleigh fading channel environment of 1) the single cell scenario without the effects of pilot contamination (i.e., the case which is free from the effects of pilot contamination because adjacent cells do not exist), 2) the NMSE of the multiple non-cooperative adjacent cells scenario using the conventional SPS (i.e., single PSs are used), and 3) the NMSE of the multiple non-cooperative adjacent cells scenario using the DPS (i.e., dual PSs are used). Figure 2.6 shows the NMSE for the i.i.d. Rayleigh fading channel environment case. The curves show that for the SPS, as the SNR increases the inter-cell interference level increases. As a result, the NMSE will be saturated at high SNR. While in DPS, as the SNR increases the inter-cell interference level is minimal and fixed. As a result, the NMSE decreases proportionally, which leads to channel estimation enhancement and improving CSI accuracy.

2.9.2 Achievable Sum-Rate

The sum-rate of a cell provides information about the achieved average high data throughput. In particular, we follow the method proposed in [94] to compute the achievable ergodic sum-rate C_{sum}^l at the BS of the l th home cell. The sum-rate is given by $C_{sum}^l = \mathbb{E} \left\{ \log_2 \left(\left| \mathbf{I}_K + \xi \hat{\mathbf{H}}_l^H \hat{\mathbf{H}}_l \right| \right) \right\}$, where ξ is the average SNR per BS receiver's antenna and $\hat{\mathbf{H}}_l$ is the estimated channel matrix at the BS using the LMMSE estimator.

Figure 2.7 shows the achievable sum-rate of the following cases: a single cell, non-cooperative

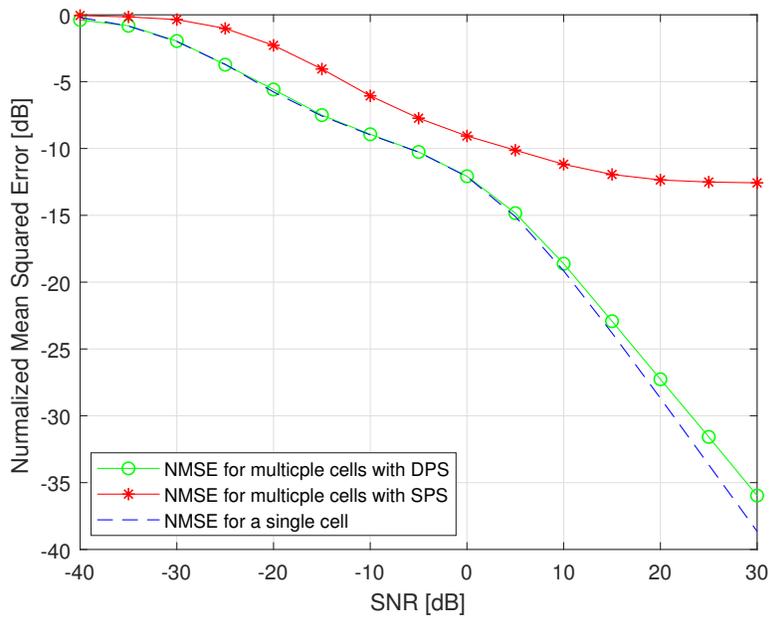


Figure 2.5: The Normalized Mean Squared Error at $M = 32$ and using the correlated Rayleigh fading channel.

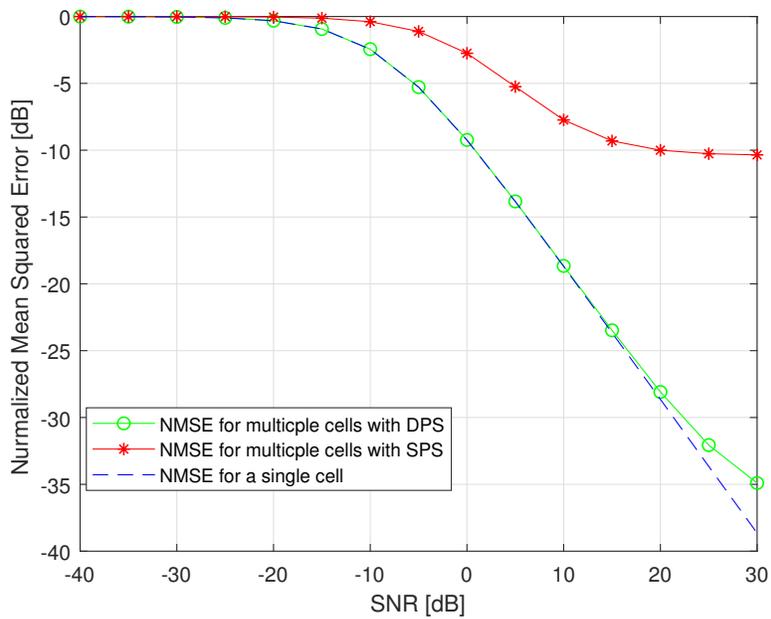


Figure 2.6: The Normalized Mean Squared Error at $M = 32$ and using the i.i.d. Rayleigh fading channel.

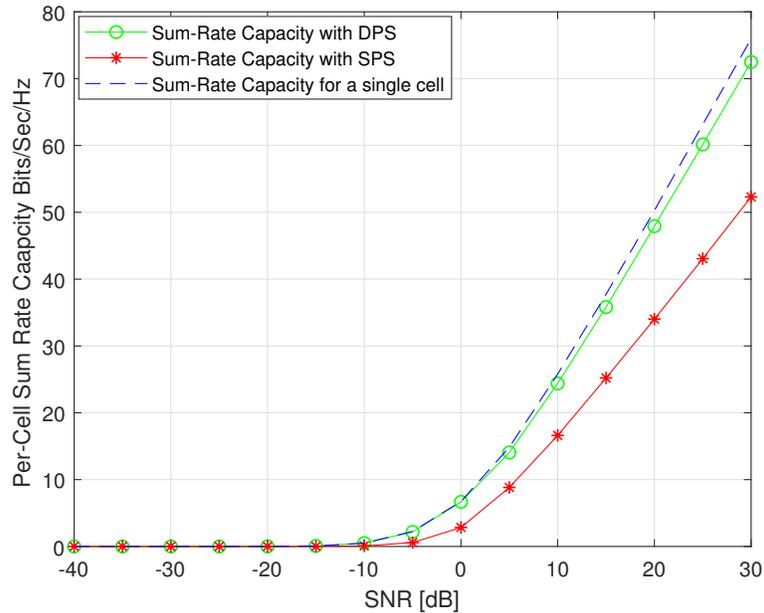


Figure 2.7: The achievable ergodic sum-rate per cell at $M = 32$ and using the correlated Rayleigh fading channel.

multiple adjacent cells using SPS, and non-cooperative multiple adjacent cells using the DPS. The performance achieved using the proposed novel DPS is very close to the performance of the single cell configuration (i.e., the case which is free from the effects of pilot contamination because adjacent cells do not exist). This shows that the proposed DPS significantly eliminates the effects of pilot contamination, improves the channel estimation, enhances the CSI accuracy, and consequently boosts the average high data throughput. Additionally, Table 2.3 presents the massive MIMO cell capacity C_{sum} at given values of NMSE for various SNRs when correlated Rayleigh fading channel is used. The table reveals a comparison between SPS and DPS where it demonstrates that DPS outperforms SPS and achieves a smaller NMSE, which resulted in improving the CSI and consequently improves the achieved sum-rate capacity. On the other hand, the BS can achieve the same sum-rate capacity that SPS offers at a lower SNR which implies improving the energy efficiency.

Table 2.3: Comparisons of cell capacity C_{sum}^l in bps/Hz , for DPS and SPS at a given values of NMSE in (dB), for correlated Rayleigh fading channel when $M = 128$, $K = 8$, and $L = 7$.

SNR	$NMSE^{DPS}$	$NMSE^{SPS}$	C_{Sum}^{DPS}	C_{Sum}^{SPS}
-10	-11.4146	-9.1973	3.0072	2.0704
-5	-12.7374	-10.2721	4.1616	2.7570
0	-13.7779	-11.4327	8.7215	4.8596
5	-17.1431	-12.5563	16.3685	10.6665
10	-20.4576	-13.9362	26.6856	18.5142
15	-24.8243	-14.2946	38.5148	27.6051
20	-29.2082	-14.6852	49.8413	35.9962

2.9.3 Singular Values Spread

One of the important metrics for assessing the performance of massive MIMO and analyzing the propagation environment involves measuring differences between the channel coefficients of different UEs and the joint spatial correlation between columns of the channel matrix (where each column represents the parallel channels between each UE and the BS's antennas). This can be evaluated by finding the Cumulative Distribution Function (CDF) of the Singular Value Spread (SVS) Γ , the spread between the smallest and the largest singular values of the wireless communications channel matrix \mathbf{H}_l , where \mathbf{h}_{lj}^l is the j th column in the matrix \mathbf{H}_l . Assuming the ordered singular values of \mathbf{H}_l are $\sigma_1^l \geq \sigma_2^l \geq \dots \geq \sigma_K^l \geq 0$, then the SVS in dB can be computed by $\Gamma = 10 \log_{10} \left(\frac{\sigma_1^l}{\sigma_K^l} \right)$.

SVS, which also known as the condition number, shows whether the columns of the channel matrix tend to be orthogonal or highly correlated. If the singular values of the matrix \mathbf{H}_l are equal (i.e., the square roots of eigenvalues of the matrix $\mathbf{H}_l \times \mathbf{H}_l^H$, also known as a Wishart matrix, are equal), then the Euclidean norm condition number will be equal to one. In this case, the columns of the channel matrix are orthogonal and the users' signals will be separable and free of pilot contamination at the BS. Moreover, the variance of the CDF of the SVS provides information about the stability behavior of the system, (this stability increases as the variance decreases). Figure 2.8 and Figure 2.9 show that the median of the SVS is the largest at $M = 32$, this means the fading margin is high and the additional power required to get reasonable SINR will be the highest.

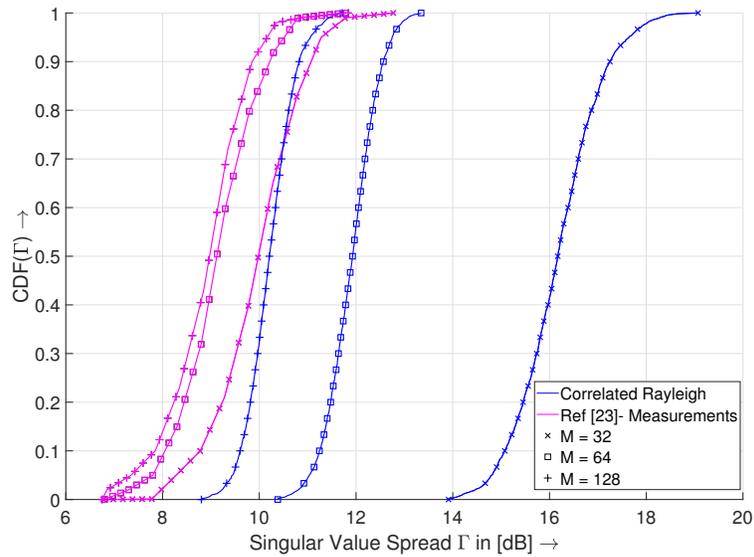


Figure 2.8: The Cumulative Distribution Function of the Singular Values Spread when the correlated Rayleigh fading channel is used.

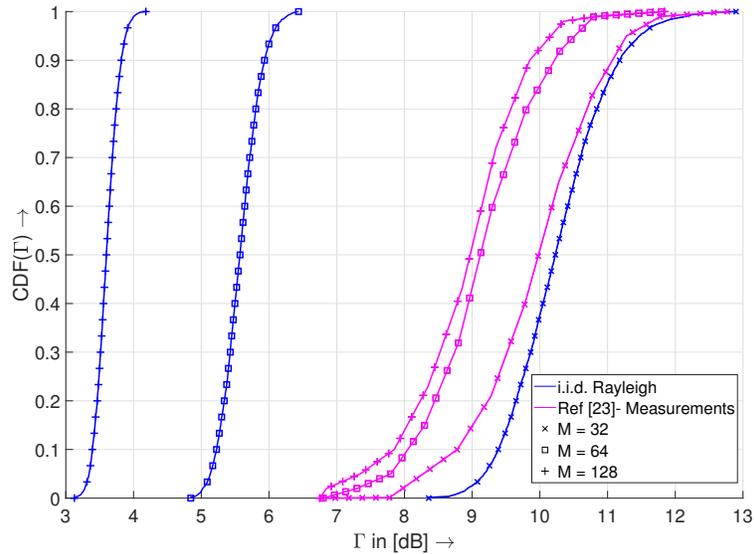


Figure 2.9: The Cumulative Distribution Function of the Singular Values Spread when the i.i.d. Rayleigh fading channel is used.

Also, Figure 2.8 shows the numerical CDF curves of the SVS for the channel matrix used in 1) the case of a single cell scenario 2) the non-cooperative multiple adjacent cells scenario when the DPS sequences are used and 3) the non-cooperative adjacent multiple cells scenario that

uses the conventional SPS when the correlated Rayleigh fading wireless communication channel is simulated using $M = 32$, $M = 64$, and $M = 128$. Similarly, Figure 2.9 shows the numerical CDF curves of the SVS for the channel matrix used in all the three cases when the i.i.d. Rayleigh fading wireless communication channel is simulated. Also, Figure 2.8 and Figure 2.9 illustrate the channel characteristics and SVS of the real wireless communications channel measurements performed in [23].

Besides, Figure 2.8 shows that the real measurements are most closest to the simulation environment settings when $M = 32$ for i.i.d. Rayleigh channel. So, we have evaluated the performance of DPS at $M = 32$ for i.i.d. and correlated Rayleigh fading channels. This shows how the performance of DPS has been assessed in a simulation environment that has a very good closeness to the real environment. Also, figures reveal implicitly, the performance of DPS in a real environment that has an existence of a high correlation will be better, as DPS was assessed under the worst conditions and considered a high fading margin compared to the real environment. Finally, in-depth mathematical derivations of the joint CDF for eigenvalues of the Wishart matrix $\mathbf{H}_l \times \mathbf{H}_l^H$ and the CDF for its condition number (i.e., the equivalent joint CDF for singular values of the channel matrix \mathbf{H}_l and CDF for its condition number, taking into account a square root transformation function between random variables) are available in [19].

2.10 Conclusions

In this chapter, we have proposed the novel DPS as an efficient mechanism for mitigation of pilot contamination, and shown that it offers unique desirable features for use in massive MIMO systems by improving system performance without cooperation between BSs. The advantages of DPS over the conventional SPS have been proven mathematically, under the i.i.d. and the correlated Rayleigh fading channel models. We have shown that by properly choosing a system parameter, DPS can be designed so that it has *universally better* performance than SPS. Unlike the state-of-the-art works, our DPS enjoys a uniform quality of service and per-user guarantees to enhance the performance.

Using proposed planning with a seven-cell cluster for pilot signal reuse with the proposed DPS,

we have shown how DPS can be practically applied in an entire cellular network that has any number of cells. This reuse concept can be generalized to any cluster size appropriate for maintaining the distance between two cells that use exactly the same PSs. Simulation results have been presented to show the advantages of DPS over SPS in terms of normalized mean squared error and the sum-rate cell capacity, due to much improved channel estimation at various signal-to-noise ratios. Remarkably, the numerical results have revealed that DPS reduces the normalized mean squared error of the channel estimates by about 10 dB at 10 dB SNR relative to SPS, and the sum-rate capacity of DPS approaches that of the single-cell scenario that is free from pilot contamination, showing the effectiveness of DPS in almost completely mitigating the harmful effects of pilot contamination. Beside, due to the existence of a high correlation in the communication channels in practice, the DPS has the tendency to be more efficient and beneficial for practical implementation in reality. Lastly, while we have demonstrated the effectiveness of DPS for TDD massive MIMO systems, DPS is a general design that can be applied in any wireless communications system that utilizes PSs for CSI, including 5G and beyond.

Chapter 3

Barker Coded DFT-s-OFDM Waveforms for 5G Uplinks

3.1 Synopsis

DFT-s-OFDM is used in LTE-Advanced (LTE-A) which is considered the Fourth Generation (4G). DFT-s-OFDM is expected to be used as the air interface waveform for uplink transmissions in 5G cellular networks according to the 3GPP recommendations. We investigate the performance of DFT-s-OFDM and the trade-off between its EE and SE in 5G uplinks employing massive MIMO antenna systems. We propose a novel method, using an adaptive length Barker Code (BC) for spreading in DFT-s-OFDM, to improve the EE-SE trade-off in 5G uplink transmissions. The performance of the proposed system is evaluated in terms of EE, Bit-Error Rate (BER), sum-rate capacity, and NMSE under the i.i.d. and the correlated Rayleigh fading wireless communications channel models. We extend DPS that we develop in Chapter 2 to operate in massive MIMO-NOMA. Furthermore, we investigate the scenario where PD-NOMA is deployed over 5G uplinks with the proposed waveform and examine the SE and EE of near/far user pairs. The numerical results show that the proposed air interface waveform results in a significant improvement in the uplink EE without degrading the SE over a range of SNRs for massive MIMO uplinks with either OMA or PD-NOMA.

3.2 Introduction

Uplink air interface waveforms need to be energy efficient to facilitate EE in uplink transmission. For designing the air interface waveforms in 5G cellular networks with MIMO technology, there are currently two possible approaches that provide backward compatibility with older generations of cellular networks. These are the OFDM [97] and the DFT-s-OFDM. Two main advantages of OFDM are the mitigation of ISI, and the ease of the channel equalization by converting a wide-band frequency-selective fading channel into a group of parallel narrow-band frequency-flat fading channels. However, it has been found that employing OFDM results in increasing PAPR which leads to ICI, increased power losses in power amplifiers, and produces inter-modulation distortions. For instance, uplink single-carrier NOMA and multiple-carrier NOMA face the drawbacks of high PAPR due to the OFDM [98]. All of these drawbacks of OFDM have huge negative impacts on EE. Therefore, the 3GPP has adopted DFT-s-OFDM for the uplink transmission in LTE networks and 5G uplink transmissions to maintain acceptable PAPR level. DFT-s-OFDM waveforms in LTE networks demonstrate improvements in reducing UE power consumption and in expanding the coverage range. However, the performance of employing DFT-s-OFDM in 5G cellular networks with both massive MIMO-OMA and massive MIMO-NOMA have not been investigated yet.

The differences between OFDM and DFT-s-OFDM is shown in Figure 3.1. In OFDM, time-domain data symbols of each UE are payloaed and mapped into all allocated sub-carriers. Therefore, the length of time-domain data symbols in OFDM without Cyclic Prefix (CP) equals the number of the allocated sub-carriers in OFDM. In contrast to OFDM, data symbols length in DFT-s-OFDM (i.e., length of Discrete Fourier Transform (DFT) outputs) is shorter than the number of allocated sub-carriers. Therefore, frequency-domain data symbols (i.e., DFT outputs) are spread amongst allocated sub-carriers, either in a distributed or localized configuration.

As 5G wireless networks aim to serve a variety of communications environments and services, investigating EE and the performance of DFT-s-OFDM in 5G cellular networks is challenging. Recall, 5G wireless networks promise to support a variety of enhanced and new services, including Enhanced Mobile Broadband (eMBB) for high speed Internet access, massive Machine Type Com-

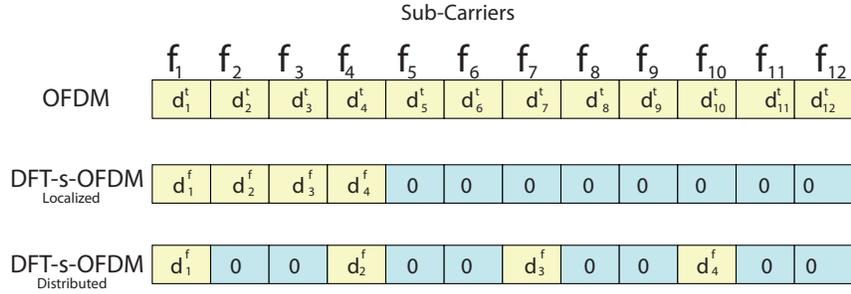


Figure 3.1: Payload data symbols comparison in OFDM, localized DFT-s-OFDM, and distributed DFT-s-OFDM, when 12 sub-carriers are used.

munications (mMTC) of pervasive IoT, and ultra-Reliable Low Latency Communications (uRLLC) for Mission Critical Communications (MCC) such as between self-driving cars, with substantial enhancements in system capacity and EE [2, 3, 99, 100]. These enhancements are only enabled by adopting advanced air interface waveforms that can *adapt* to different scenarios with improved EE [101], and utilize massive MIMO antenna systems [7], which are key technologies of 5G RANs. Thus, it is important to examine the performance of any proposed air interface waveforms in a massive MIMO environment.

In massive MIMO, as a BS is equipped with a large number of antennas. Intuitively, this architecture improves the downlink EE since it reduces the power loss in power amplifiers and uses focused beamforming. However, the EE of massive MIMO for the uplink scenario is less well understood [97] [102]. Since EE in uplink massive MIMO is important for ensuring green communications and meeting the 5G requirements, it becomes an essential factor in designing the 5G air interface waveforms [102].

On the one hand, the deployment of OMA technique into massive MIMO have the advantage that intra-cell interference does not arise. However, the use of OMA limits the number of served users, as this number is determined by the available orthogonal resources. On the other hand, NOMA technology has the promise of enhancing the SE capability that would increase the system capacity of 5G cellular networks. Multiple users in NOMA could be served using the same resource simultaneously. This, unfortunately, comes at the cost of introducing intra-cell interference.

Besides, NOMA offers other advantages and reduces both the signaling overhead and the latency in the uplink transmission. In addition, NOMA can offer a higher data rate, better cell coverage range, and thus promotes more the feasibility of higher massive connectivity [103]. Moreover, NOMA is compatible with current and future generation of cellular networks architecture, and it has been proposed for the downlink of 3GPP LTE-A networks with the name of multiuser superposition transmission [104]. In this regards, one of the promising applications of NOMA is the remote energizing of IoT sensors and Wireless Power Transfer Networks (WPTN) including their low-power profile devices with limited charging capabilities, which intuitively required an energy efficient uplink mechanism. These advantages encourage us to (1) explore the deployment of NOMA in massive MIMO cellular networks, as NOMA promises to be a 5G technology that aids massive MIMO to achieve the 5G technical targets, and (2) assess the proposed air interface waveforms in massive MIMO with NOMA deployment.

In NOMA, multiple users are served simultaneously in the same frequency-time band. This improves the spectrum efficiency and enables the simultaneous service of multiple users at various communication channel conditions. However, this also introduces intra-cell interference amongst users within the same cell and increases the receiver complexity due to the implementation of SIC. There are two types of NOMA, the first is the Code Domain NOMA (CD-NOMA), which requires more bandwidth, and the second is PD-NOMA. Improving the EE is our main goal in this chapter. Therefore, we focus on the PD-NOMA because it is easy to implement, and can be embedded into existing networks. In contrast to CD-NOMA, PD-NOMA does not require any additional bandwidth. In PD-NOMA, paired users are assigned distinct power levels that maximize the performance and ensure that a diverse range of power levels arrive at the BS [105].

From a NOMA perspective, the recent work on NOMA has focused on enhancing the SE and on reducing the complexity of SIC at the receiver, but there have been few works on EE. The EE optimization of downlink MIMO-NOMA deployments has been investigated in [106] but EE in uplink transmission is still an open research area for more avenue of contributions. Accordingly, in addition to a massive MIMO with OMA deployment, this chapter considers improving the EE

of the uplink transmissions in a massive MIMO with PD-NOMA deployment. Also, one critical factor that plays an important role in the realization of EE is the total consumed power in uplink transmission. Total consumed power include transmission power, circuit power, and power losses due to external factors (external power consumption). The International Telecommunication Union (ITU) recommends the circuit power and the transmission power to be reduced in any communication device to improve the EE [107], but this approach is limited due to the trade-off between SE and EE [108]. Therefore, we took into consideration to improve the EE-SE trade-off.

To the best of our knowledge, there is no previous work done on usage evaluation of DFT-s-OFDM (instead of OFDM) in the uplink transmission in massive MIMO-NOMA. This chapter aims to fill this research gap. On the one side, DFT-s-OFDM has the flexibility in supporting both the localized and the distributed OFDM. This capability enables its use in a wide range of applications, services, and environments [109]. However, DFT-s-OFDM has several issues such as poor spectral containment and noise enhancement [102]. On the other side, both OFDM and DFT-s-OFDM exhibit high Out of Band (OOB) leakage, but DFT-s-OFDM results in lower PAPR [100]. To reduce the OOB leakage a new waveform named Generalized DFT-s-OFDM (G DFT-s-OFDM) has been proposed in [99], where the data symbols are sandwiched between head and tail short sequences to offer a guard interval. A special case of this waveform is called Zero-tail DFT-s-OFDM (ZT DFT-s-OFDM), where the head and tail sequences are set to zeros [110]. However, the EE of the proposed waveforms has not been investigated. Also, ZT DFT-s-OFDM showed a small increase in its PAPR and suffers from ISI in high delay spread environments. In a similar approach to ZT DFT-s-OFDM, the Unique-word DFT-s-OFDM (UW DFT-s-OFDM) waveform has been proposed in [111] by replacing the zero-tail sequence in ZT DFT-s-OFDM with a unique fixed sequence. Although UW DFT-s-OFDM and its enhanced versions improve slightly the PAPR, it increases the transmitter's complexity significantly. Thus, motivated by the need to improve EE while preserving a high level of SE in 5G uplink transmissions, in this chapter, we propose a novel alternative approach that uses an adaptive length BC DFT-s-OFDM waveforms which reduce both OOB and PAPR instead of conventional DFT-s-OFDM. Notably BC DFT-s-OFDM waveforms

acquires the features of BCs. BCs are a special type of binary code that represents biphasic codes, which demonstrates advantages in producing the smallest achievable sidelobes and focusing the energy into the main lobe, thus reducing power losses and restraining ICI. An uplink waveform that has smaller side lobes improves the robustness of asynchronous transmission [99].

We first investigate the use of the conventional DFT-s-OFDM in massive MIMO uplink transmission and its EE feasibility, with OMA and NOMA techniques employment, and then to enhance EE we introduce adaptive length BC DFT-s-OFDM. We model and formulate the problem, consider the scenario involving a cell located within a cluster of non-cooperative multiple adjacent cells¹, and when BCs of different lengths are employed. More specifically, we consider a single cell with its BS (equipped with a very large number of antennas) located at the center of the cell. The CSI is estimated in uplink transmission using the DPS that has been proposed in Chapter 2 when OMA is used together with massive MIMO in non-cooperative multiple adjacent cells.

However, interference management becomes more complicated when NOMA is used together with massive MIMO in non-cooperative multiple adjacent cells due to existence of both inter-cell and intra-cell interference, as most proposed NOMA designs are agnostic to inter-cell interference [112]. Therefore, in continuation to the development of DPS that we proposed in Chapter 2, we propose and demonstrate in this chapter EDPS as an efficient scheme that significantly improves the CSI accuracy and mitigates simultaneously both intra-cell and inter-cell interference when NOMA is used together with massive MIMO in non-cooperative multiple adjacent cells. To demonstrate the effectiveness of our proposed BC DFT-s-OFDM method in the uplink transmissions of 5G massive MIMO, we compare its performance with that of the conventional DFT-s-OFDM. Our numerical results demonstrate that the proposed BC DFT-s-OFDM waveforms (with the adaptive length of BC) improves the performance and achieves better EE, lower BER, lower PAPR, and higher sum-rate capacity compared to the conventional DFT-s-OFDM. In particular, the MMSE estimator is used; and the NMSE of the estimated CSI of both the conventional DFT-s-OFDM and the proposed BC DFT-s-OFDM are compared.

¹This configuration can be easily generalized to any cellular network with any arbitrary cluster size, which consists of a group of non-cooperative cells that mitigate pilot contamination by using DPS.

The *main contributions* of this chapter include: (1) the first to investigate EE performance of DFT-s-OFDM in the uplink transmission of massive MIMO cellular networks with two main different configurations of multiple access a) OMA and b) PD-NOMA (2) the first which proposes the use of BC in DFT-s-OFDM to generate adaptive length BC DFT-s-OFDM waveforms, to obtain a better EE enhancement for 5G uplink transmissions while preserving a good level of SE-EE trade-off (3) proposal of an efficient method to integrate NOMA into massive MIMO by employing EDPS, which can significantly mitigate both intra-cell and inter-cell interferences either in a single cell scenario and/or non-cooperative multiple adjacent cells scenario.

The rest of this chapter is organized as follows: In the following section, we discuss the EE metrics. In Section 3.4, we review the theoretical background of BCs and illustrate how the BC spreading matrix used in our system model is formed. In Section 3.5 we present the system configuration of a single cell massive MIMO with OMA. In Section 3.6, we demonstrate the cell's system model of massive MIMO with PD-NOMA, and then introduce EDPS. Section 3.7 discusses the obtained numerical results. Section 3.8 concludes the chapter.

3.3 The Energy Efficiency Metrics

In cellular communication systems, EE of a UE can be defined as the number of bits reliably delivered to the final destination (i.e., the BS in uplink) per unit of energy consumed at the originating source (i.e., the UE), and is measured in bits per Joule (b/J). However, the system EE metric more widely used is the energy consumption required to achieve the target system throughput, i.e., the ratio of the achieved total system throughput to the total power consumed [113]. Another metric of measuring the EE relates the achieved capacity to the consumed power, where the capacity is measured in bits per channel use (simply in bps/Hz) and $EE = \frac{Capacity}{ConsumedPower}$ [114]. The asymptotic uplink Shannon capacity C_{UL} of a single cell massive MIMO under favorable conditions is given by $C_{UL} = \sum_{k=1}^K \log_2(1 + p_{ul}M\beta_k)$, where M is the number of BS antennas, p_{ul} is the SNR, K is the number of users, and β_k is the large scale fading coefficient of k th UE.

The consumed power can be splitted into three parts: the transmitted power, the circuit power,

and the external system power. SE depends on the transmitted power and is measured in bits per second per hertz (b/s/Hz) (i.e., the throughput per unit of the bandwidth) and $SE = \frac{Capacity}{Bandwidth}$. Then, $EE = \frac{SE \times Bandwidth}{ConsumedPower}$. It is important to emphasize that if SE increases in the numerator, the consumed power in the denominator also increases and vice versa. The aim is to increase SE with as little as a possible increment in the consumed power. Our system model is based on DFT-s-OFDM with massive MIMO, so we use a global sum EE metric that considers the multiple sub-carriers, multiple antennas, and multiple UEs in the system. Implementing the BCs at the UE's transmitter improves the uplink throughput and enhances SE, which results in EE improvements of the proposed system. It also significantly reduces the PAPR, which results in reducing the consumed power at the UE. However, using BCs introduces an incremental increase in circuit power as an additional circuit needs to be embedded into UE. Due to the simplicity of the additional required circuitry, this increment is small, and has a limited impact compared to the improvements offered by BC DFT-s-OFDM waveform in both EE and SE. As a result, the trade-off effect between the EE and SE is minimized while achieving our target for improving the EE.

3.4 Barker Codes

This section reviews the theoretical background of BCs and illustrates how the BC matrix \mathbf{B} in our system model is formed using BC sequences. BC plays a critical role in the proposed BC DFT-s-OFDM waveform and system architecture. The distinguished property of a BC waveform is that its autocorrelation function has side lobes that are the lowest possible. This property maximizes the peak-to-side lobes ratio of the autocorrelation function of the generated waveform. Thus, since BC is employed after the DFT operation, the Power Spectral Density (PSD) of the resulting waveform is concentrated in the desired frequency band while its OOB leakage is significantly reduced. This is a very special property that could be exploited advantageously in wireless communications in that it improves the SE, reduces the ICI, and improves EE. So, we recommend the deployment of BCs into uplinks transmissions of 5G wireless networks. Beside, Barker coded waveforms are easy to generate by a simple phase modulation circuit [115], where only very little additional circuit

power is required. Also, the use of Barker coded air interface waveforms have shown results in an improvement in the channel estimation of fading channels [116].

The metric that measures the EE in BCs sequences is called the Merit Factor (MF). As MF increases, the energy of the transmitted waveform tends to be more uniformly distributed in the frequency band [117]. As a consequence, Barker coded waveforms acquire immunity against OOB leakage effects and become more energy efficient. Let $F(\mathbf{b}_j)$ be the MF of the BC sequence \mathbf{b}_j of length N ; it is defined by:

$$F(\mathbf{b}_j) = \frac{N^2}{2\sum_{0 < u < N} [\sum_i b_j(i)b_j(i+u)]^2}. \quad (3.1)$$

The j th row in the BC matrix \mathbf{B} consists of a generalized BC sequence \mathbf{b}_j with length N and zero padding. A generalized BC sequence \mathbf{b}_j is a finite sequence with N entries. In this chapter, the value of N is adaptive and varies from 1 to 13, which adds a degree of freedom to be used in reality depending on the application of 5G. Therefore, N depends on the application type, the size of the transmitted data, the number of active users, number of available sub-carriers, and the strength of the wireless communication channel coefficients. The value of each entry of \mathbf{b}_j is a complex number $\pm b_i$ where b_i has a unity magnitude and can only take the value 1 or i , (i is a pure imaginary number). When $N = 1$ and b_i is a real positive value, then our proposed waveform and system becomes equivalent to the conventional DFT-s-OFDM. Any generalized BC sequence \mathbf{b}_j must satisfy a condition of having an autocorrelation function $C_b(\tau)$ such that $|C_b(\tau)| \leq 1, \tau \neq 0$ [115, 118]. The autocorrelation function is given by:

$$C_b(\tau) = \sum_{j=1}^{N-\tau} \{\mathbf{b}_j\} \{\mathbf{b}_{j+\tau}^H\}, \quad (3.2)$$

where $\mathbf{b}_{j+\tau}^H$ is the complex conjugate transpose of $\mathbf{b}_{j+\tau}$. When all the entry values of \mathbf{b}_j are real, the BC sequence is called a Binary Barker Sequence (BBS). If the entry values of \mathbf{b}_j can include both 1 and i then the BC sequence is called a Quaternary Barker Sequence (QBS). Our proposed structures can use any of these sequences.

3.4. Barker Codes

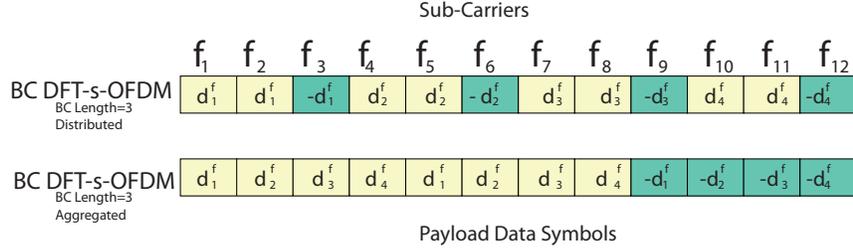


Figure 3.2: Structures of BC DFT-s-OFDM when Barker code length=3, $b_i = \{+1, -1\}$, and 12 sub-carriers are used.

Thus, a natural question is raised, “how to construct the spreader BC matrix \mathbf{B} ?”, as 5G wireless networks vary according to their RAN technology, the applications they are intended for, their target services, and backhaul architectures. Thus, in addition to satisfying the requirements mentioned in this section, the structure of the spreader BC matrix \mathbf{B} should also vary to capture and tackle these aspects and variations of 5G wireless networks. Hence, our proposed system has a flexible and adaptable structure of the BC matrix \mathbf{B} to accommodate these variations which either generate aggregated or distributed BC DFT-s-OFDM. Figure 3.2 shows the differences between aggregated BC DFT-s-OFDM and distributed BC DFT-s-OFDM.

Assume the DFT output frequency-domain data symbols are $\mathbf{d} = \{d_1^f, \dots, d_s^f, \dots, d_S^f\}$, which has a length of S and $(\cdot)^f$ is a data symbol in frequency domain. Also, assume N is the suitable BC length based on the factors mentioned previously that determine the used length. Then, the structure of the spreader BC matrix can be formed by:

$$\mathbf{B} = (\mathbf{I}_S \otimes \mathbf{b}_N^T)^T, \quad (3.3)$$

for the distributed mode, whereas for the aggregated mode it can be formed by:

$$\mathbf{B} = (\mathbf{b}_N^T \otimes \mathbf{I}_S)^T. \quad (3.4)$$

Consequently, BC DFT-s-OFDM sequence \mathbf{x}_{BC} can be generated by:

$$\mathbf{x}_{BC} = \mathbf{d} \times \mathbf{B}. \quad (3.5)$$

Thus, the sequences in Figure 3.2 are generated using $N = 3$, $S = 4$, and number of sub-carriers equals 12. Therefore, in a matrix format, (3.3) and spreader BC matrix \mathbf{B} in this case can be expressed as:

$$\mathbf{B} = \begin{bmatrix} b_1 & \cdots & b_3 & 0_4 & \cdots & \cdots & \cdots & \cdots & 0_{12} \\ 0_1 & \cdots & 0_3 & b_1 & \cdots & b_3 & 0_7 & \cdots & 0_{12} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0_1 & \cdots & \cdots & \cdots & \cdots & 0_9 & b_1 & \cdots & b_3 \end{bmatrix}. \quad (3.6)$$

3.5 System Model

We consider a single cell in a massive MIMO cellular network as described in [21], whose BS is equipped with an array of M antennas, uses OMA to serve K users in its coverage area as shown in Figure 3.3. Each UE is equipped with a single antenna. The cell under consideration is located at the center of a cluster consisting of multiple non-cooperative adjacent cells. Besides, we assume BS employs DPS to significantly mitigate pilot contamination².

Considering the uplink transmission of K active UEs, the signal y_m^l received at the m th antenna of the BS corresponding to the l th channel tap is a superposition of all the signals transmitted from the K UEs. We assume a favorable propagation environment with sufficient scatters that produce dense multipath components for the signals transmitted by each UE. Mathematically, we can write the received y_m^l signal as:

$$y_m^l = \sqrt{\rho_{ul}} \sum_{k=1}^K \sqrt{\eta_k} h_k^{m,l} x_k + w_m^l, \quad (3.7)$$

where ρ_{ul} denotes the SNR, η_k denotes the power control coefficient of k th UE, x_k is the signal

²The effectiveness of DPS only appears in the case of a non-cooperative multiple-cells scenario is operated.

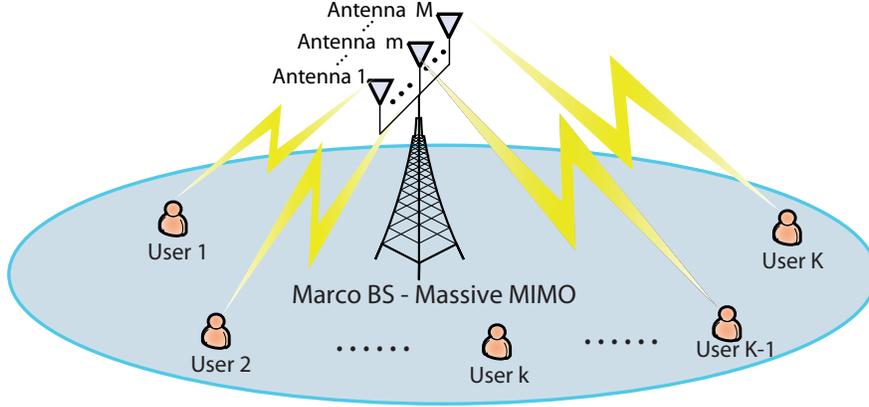


Figure 3.3: Single cell massive MIMO, the BS has M antennas and serves K users.

transmitted by the k th UE, $h_k^{m_l}$ is the channel coefficient of the l th transmission path (corresponding to average multipath components at the l th tap of the channel Finite Impulse Response (FIR)) between the k th UE in the cell and the m th BS's antenna, and w_m^l is the additive noise, which is assumed to be $\mathcal{CN}(0, 1)$ and independent across different l and m . In vector form, we rewrite (3.7) as:

$$\mathbf{y}^l = \sqrt{\rho_{ul}} \mathbf{H}^l \mathbf{D}_\eta^{1/2} \mathbf{x} + \mathbf{w}^l, \quad (3.8)$$

where $\mathbf{y}^l = [y_1^l, \dots, y_M^l]^T$, $\mathbf{w}^l = [w_1^l, \dots, w_M^l]^T$, $\mathbf{x} = [x_1, \dots, x_K]^T$ is formed of all the signals transmitted by all UEs in the cell to all BS's antennas, \mathbf{D}_η is a diagonal $K \times K$ matrix that has a diagonal vector $\boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_K]^T$, and \mathbf{H}^l is the $M \times K$ channel matrix corresponding to the l th transmission path (i.e., l th channel tap) and each of its columns represents the parallel channels between each UE and the BS's antennas, \mathbf{H}^l is expressed by:

$$\mathbf{H}^l = \begin{bmatrix} h_1^{1^l} & \cdots & h_K^{1^l} \\ \vdots & \ddots & \vdots \\ h_1^{M^l} & \cdots & h_K^{M^l} \end{bmatrix}. \quad (3.9)$$

The received signal given by (3.8) takes into account only one channel tap, then the total re-

ceived signal with taking into account all the channel taps L at a time index n is given by:

$$\mathbf{y}[n] = \sum_{l=1}^L \sqrt{\rho_{ul}} \mathbf{H}^l \mathbf{D}_\eta^{1/2} \mathbf{x}[n-l+1] + \mathbf{w}^l[n], \quad (3.10)$$

where L is the total number of channel taps (i.e., the length of the channel FIR), which is determined by the delay spread and the transmission bandwidth. Equation (3.10) shows there are an L channel matrices in the system model. Each one of them is a two dimensional (2D) matrix with a size of $M \times K$. These L channel matrices can be rearranged to form a three dimensional (3D) channel matrix. Therefore, the 3D channel matrix $\mathbf{H}_{M \times K \times L}^{3D}$ that is given by:

$$\mathbf{H}_{M \times K \times L}^{3D} = \begin{bmatrix} \begin{bmatrix} h_1^{1L} & \cdots & h_K^{1L} \\ h_1^{11} & \cdots & h_K^{11} \\ \vdots & \ddots & \vdots \\ h_1^{M1} & \cdots & h_K^{M1} \end{bmatrix} \end{bmatrix}, \quad (3.11)$$

is the $M \times K \times L$ matrix containing all the channel coefficients between all the UEs in the cell and all M antennas at the BS for all L channel taps (transmission paths).

Following the notations used in [21], we factor the channel coefficient as $h_k^{m^l} = \sqrt{\beta_k^{m^l}} g_k^{m^l}$, where $\beta_k^{m^l} > 0$ is called the *large-scale fading* coefficient (which characterizes the range-dependent path loss and shadow fading and its value is assumed to be known to the BS) [119] and $g_k^{m^l}$ is called the *small-scale fading* coefficient. We assume that the small-scale fading has a Rayleigh distribution and we consider two Rayleigh-fading channel models: i.i.d. and the correlated Rayleigh fading.

3.5.1 Transmitter Configuration at the UE

The UEs within the massive MIMO cell generate the uplink transmitted signals denoted by \mathbf{x} in (3.8). The generated uplink signal by the k th UE, is denoted by x_k in (3.7). Figure 3.4 shows the proposed structural configuration of the k th UE's transmitter, which used to generate the BC DFT-s-OFDM waveform x_k . For simplicity, similar to [120] we present how x_k is generated within one

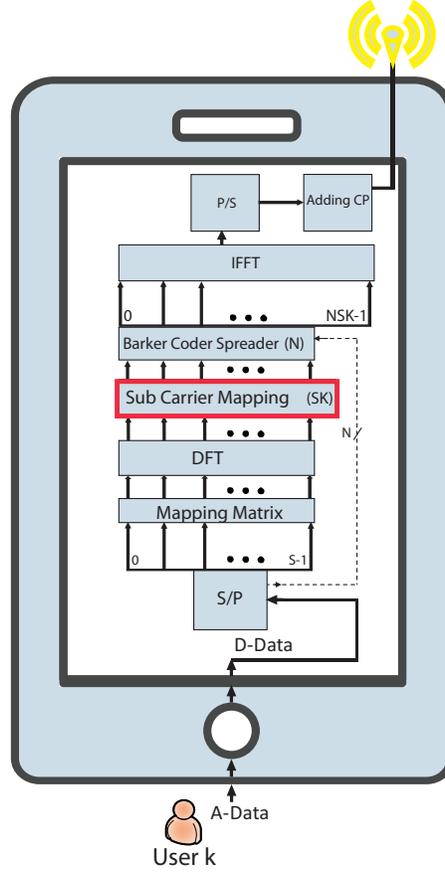


Figure 3.4: The proposed transmitter structural configuration for the k th UE.

BC DFT-s-OFDM symbol and we assume BC DFT-s-OFDM symbols are statistically distributed in an uncorrelated fashion with zero mean. The data symbol is denoted by the row vector \mathbf{d}_k and has a length of S data points, so the generated BC DFT-s-OFDM signal x_k becomes now a vector and is denoted by \mathbf{x}_k with length of NSK . The S data points of \mathbf{d}_k are mapped to the input of the DFT matrix through a time-domain mapping matrix denoted by \mathbf{M}_t with size of $S \times S$ to carry out a series to parallel conversion. The outputs of the DFT matrix \mathbf{T} are mapped to a subset of S sub-carriers (i.e., out of SK sub-carriers, only S sub-carriers are mapped either localized or interleaved) through a frequency-domain mapping matrix denoted by \mathbf{M}_f with size of $SK \times S$. The \mathbf{M}_f structure depends on the applications and services introduced by the cellular network. In contrast to the conventional DFT-s-OFDM, the output of matrix \mathbf{M}_f is spread using the BC matrix

denoted by \mathbf{B}^3 in our proposed scheme that has a size of $NSK \times SK$. Note that when the BC used to generate the \mathbf{B} matrix is BBS with $N = 1$, the generated BC DFT-s-OFDM waveform becomes identical to the conventional DFT-s-OFDM waveform. Then the BC spreader output is converted to the time-domain by the Inverse Fast Fourier Transform (IFFT) equivalent multiplication matrix denoted by \mathbf{F}^H that has a size of $NSK \times NSK$. This can be expressed mathematically by:

$$\mathbf{x}_k = \mathbf{F}^H \mathbf{B} \mathbf{M}_f \mathbf{T} \mathbf{M}_t \mathbf{d}_k^T, \quad (3.12)$$

where \mathbf{T} is the DFT equivalent multiplication matrix whose (m, n) entry is given by $T_{m,n} = \frac{1}{\sqrt{S}} e^{-j2\pi \frac{(m-1)(n-1)}{S}}$. The matrix \mathbf{H}^{3D} in (3.11) is a 3D matrix consisting also of K 2D matrices, each matrix has a size of $M \times L$. The channel matrix of the k th UE can be denoted by \mathbf{H}_k and expressed as:

$$\mathbf{H}_k = \begin{bmatrix} h_k^{1^1} & \cdots & h_k^{1^L} \\ \vdots & \ddots & \vdots \\ h_k^{M^1} & \cdots & h_k^{M^L} \end{bmatrix}_k, \quad (3.13)$$

where each row in the matrix has a length of L , which is the length of Channel Impulse Response (CIR) of the wireless communications channel between the UE and one antenna element at the BS [32]. In particular, L is assumed to be less than or equal to the size of the CP used in our system model.

3.5.2 The Base Station Receiver Configuration

Without loss of generality, the signal transmitted by the k th UE and received at the m th antenna of the BS propagates over the wireless communications channel whose CIR \mathbf{h}_{mk} is given by:

$$\mathbf{h}_{mk} = \left[h_k^{m^1}, h_k^{m^2}, \cdots, h_k^{m^{L-1}}, h_k^{m^L} \right]. \quad (3.14)$$

Then at the m th BS antenna, the received uplink signal that was transmitted by the k th UE is

³ \mathbf{B} here is the transpose of the matrix that is formed either by (3.3) or by (3.4) because the data points are also transposed and became a row vector.

expressed as:

$$\mathbf{y}_{mk} = \sqrt{\rho_{ul}\eta_k}\mathbf{C}_{mk}\mathbf{x}_k + \mathbf{w}_{mk}, \quad (3.15)$$

where \mathbf{x}_k is the transmitted uplink signal generated by the k th UE's transmitter using (3.12), \mathbf{C}_{mk} is the equivalent Toeplitz channel matrix that models the circular convolution to perform the linear convolution between \mathbf{x}_k and \mathbf{h}_{mk} . The structure of \mathbf{C}_{mk} is explained in [109]. The first column of \mathbf{C}_{mk} has the same length of \mathbf{x}_k such that its first L elements is equal to \mathbf{h}_{mk}^T and the remaining elements are zero paddings. In (3.15), \mathbf{w}_{mk} is the additive Gaussian noise, where each element of \mathbf{w}_{mk} is assumed to be $\mathcal{C}\mathcal{N}(0, 1)$. Using (3.12), we can rewrite (3.15) as:

$$\mathbf{y}_{mk} = \sqrt{\rho_{ul}\eta_k}\mathbf{C}_{mk}\mathbf{F}^H\mathbf{B}\mathbf{M}_f\mathbf{T}\mathbf{M}_t\mathbf{d}_k^T + \mathbf{w}_{mk}. \quad (3.16)$$

Therefore, taking into account all antennas, the signal received at the BS from the k th UE is given by:

$$\mathbf{Y}_k = \begin{bmatrix} \mathbf{y}_{1k}, & \mathbf{y}_{2k}, & \cdots & \mathbf{y}_{mk}, & \cdots & \mathbf{y}_{(M-1)k}, & \mathbf{y}_{Mk} \end{bmatrix}. \quad (3.17)$$

On the other hand, the received uplink signal at the m th BS antenna transmitted by all UEs in the massive MIMO cell is given by:

$$\mathbf{y}_m = \sum_{k=1}^K \sqrt{\rho_{ul}\eta_k}\mathbf{C}_{mk}\mathbf{F}^H\mathbf{B}\mathbf{M}_f\mathbf{T}\mathbf{M}_t\mathbf{d}_k^T + \mathbf{w}_m, \quad (3.18)$$

where \mathbf{y}_m is a column vector. Thus, the received signal across all the BS antennas can be expressed as matrix $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_m, \cdots, \mathbf{y}_M]$.

Figure 3.5 shows the proposed structure of the BS receiver. The receiver has M inputs, one input at each antenna, with an input at the m th antenna denoted by \mathbf{y}_m . The signals received at these inputs are processed simultaneously by applying the reverse of the operations that have been carried at the UE's transmitter, in the reverse order, and the transmitted uplink data are estimated jointly by using the LMMSE estimator.

The BS's receiver first removes the CP from the received \mathbf{Y} signal, then applies the Fast Fourier

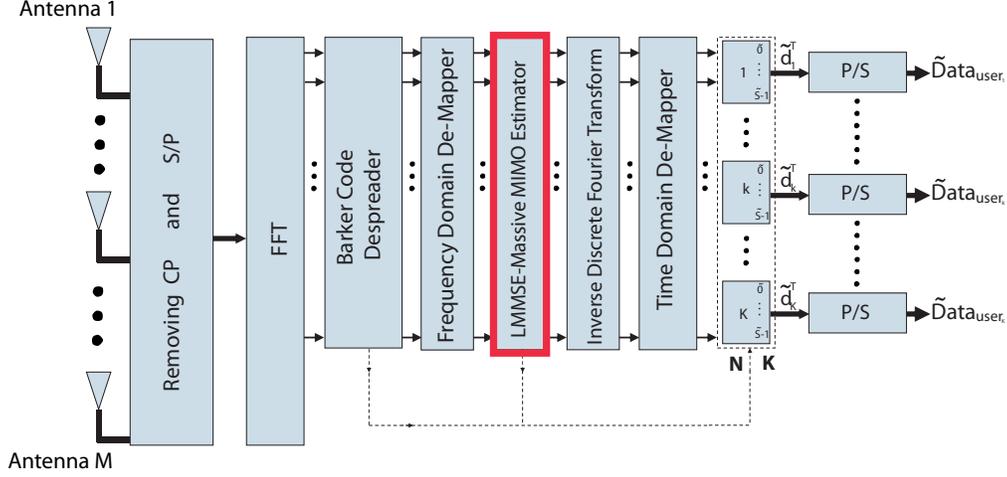


Figure 3.5: The proposed structural configuration of the BS receiver.

Transform (FFT), followed by the BC de-spreader before applying the LMMSE. We choose the LMMSE estimator because of its superior performance at both low and high SINR [93, 94, 121]. The signal at the front end of the LMMSE estimator can be written as:

$$Y_{\text{LMMSE}} = M_f^H B^H F Y, \quad (3.19)$$

where B^H and M_f^H are the Hermitian versions of B and M_f matrices respectively, and F is the FFT matrix. Equation (3.19) shows that we apply the FFT on the received signal Y to reverse the IFFT operation that has been carried out at the transmitter according to (3.15) and (3.16). However, by applying the FFT on the received signal Y , (3.19) shows that the wireless channel coefficients that are given in (3.15) and (3.16) (i.e., the entries of the C_{mk} matrix) are transformed into the frequency domain as there was no an IFFT operation carried out on these channel coefficients. Also, the circular convolution operation between the CIR and the transmitted signal becomes multiplication operation after applying the FFT operation on the received signal at the BS's receiver. Therefore, the decoding matrix $Q_{k,m}^L$ of the LMMSE estimator that considers the CIR of the wireless communications channel with length L between the k th UE and m th BS antenna is given by:

$$Q_{k,m}^L = \left(\tilde{D}_\eta^{1/2H} (\hat{H}_{k,m}^f)^H (\hat{H}_{k,m}^f) \tilde{D}_\eta^{1/2} + \rho_{ul} \mathbf{I} \right)^{-1} \tilde{D}_\eta^{1/2H} (\hat{H}_{k,m}^f)^H, \quad (3.20)$$

where $\hat{\mathbf{H}}_{k,m}^f$ is the estimated channel matrix in the frequency domain. The diagonal elements in $\hat{\mathbf{H}}_{k,m}^f$ are the FFT of the estimated version of the first column of \mathbf{C}_{mk} . Then the output of the LMMSE estimator is given by $\mathbf{Q}_{k,m}^L \mathbf{Y}_{\text{LMMSE}}$ and can be expressed as:

$$\mathbf{Q}_{k,m}^L \mathbf{Y}_{\text{LMMSE}} = \left(\tilde{\mathbf{D}}_{\eta}^{1/2H} (\hat{\mathbf{H}}_{k,m}^f)^H (\hat{\mathbf{H}}_{k,m}^f) \tilde{\mathbf{D}}_{\eta}^{1/2} + \rho_{ul} \mathbf{I} \right)^{-1} \tilde{\mathbf{D}}_{\eta}^{1/2H} (\hat{\mathbf{H}}_{k,m}^f)^H \mathbf{M}_f^H \mathbf{B}^H \mathbf{F} \mathbf{Y}. \quad (3.21)$$

Applying the reverse operations carried at the UE's transmitter in a reverse order according to Figure 3.5, yields the outputs (i.e., $\tilde{\text{Data}}_{User_1} \cdots \tilde{\text{Data}}_{User_K}$) of the BS's receiver which can be expressed as:

$$\begin{aligned} \mathbf{M}_t^H \mathbf{T}^H \mathbf{Q}_{k,m}^L \mathbf{Y}_{\text{LMMSE}} = \\ \mathbf{M}_t^H \mathbf{T}^H \left(\tilde{\mathbf{D}}_{\eta}^{1/2H} (\hat{\mathbf{H}}_{k,m}^f)^H (\hat{\mathbf{H}}_{k,m}^f) \tilde{\mathbf{D}}_{\eta}^{1/2} + \rho_{ul} \mathbf{I} \right)^{-1} \tilde{\mathbf{D}}_{\eta}^{1/2H} (\hat{\mathbf{H}}_{k,m}^f)^H \mathbf{M}_f^H \mathbf{B}^H \mathbf{F} \mathbf{Y}, \end{aligned} \quad (3.22)$$

where \mathbf{M}_t^H and \mathbf{T}^H are the Hermitian versions of \mathbf{M}_t and \mathbf{T} matrices respectively.

Based on (3.18) and (3.22) together with (3.11), the BS receiver should perform an M parallel processing in a quadrature mode of operation (i.e., in two dimensions), the first dimension represents L and the second dimension represents K .

3.6 NOMA Deployment in Massive MIMO

3.6.1 PD-NOMA Integration in Massive MIMO with Perfect CSI

In Section 3.5, we present the system model in a single cell massive MIMO with the assumption OMA is used. However, OMA techniques in cellular communications prevent the system from reaching its capacity limit in uplink transmissions due to the mutual exclusiveness of resource allocations [122]. So, in this section, we extend the system model to incorporate the PD-NOMA deployment in massive MIMO cells. Assume the cell under consideration in Figure 3.3 is divided into a number of R concentric circular regions whose center is the BS that serves up to K active UEs simultaneously in each region. For simplicity, we assume $R = 2$ in this chapter. Thus, the

number of active UEs served simultaneously in the cell is now doubled and becomes $2K$, K active UEs in each region.

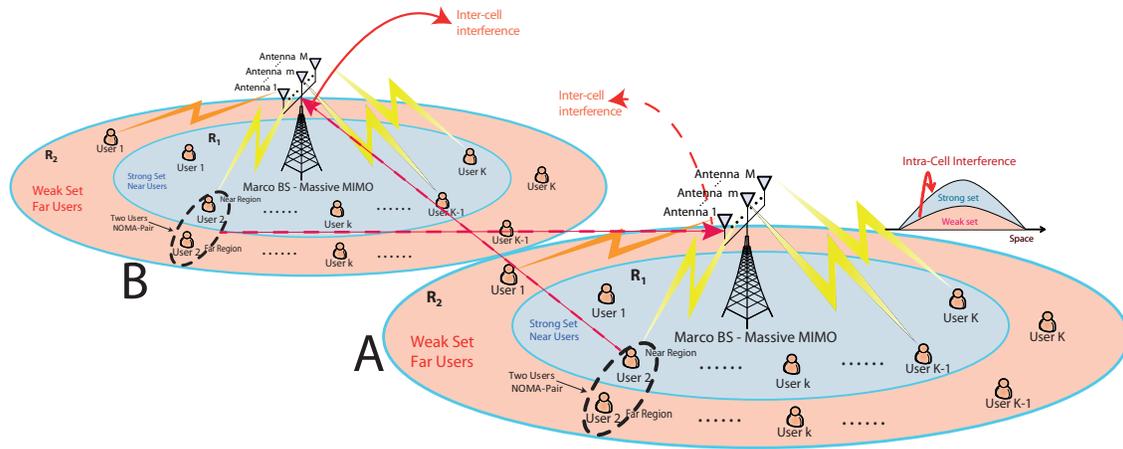


Figure 3.6: Two cells of massive MIMO with PD-NOMA deployment, each cell is divided into two regions ($R=2$), each region has a set of K active users. The BS is equipped with M antennas and located at the center of each cell.

Figure 3.6 shows the cell's coverage map where users are uniformly and randomly distributed inside the cell regions when $R = 2$, and it shows the formation of the PD-NOMA pairs. The users inside the intended cell are categorized into two groups: the near users in the central region and the far users in the edge region. Hence, the number of PD-NOMA-pairs, in this case, is equal to K . The number of PD-NOMA-pairs and consequently the number of users within each PD-NOMA-pair in a cell is determined by the acceptable level of intra-cell interference at its BS receiver, the number of R concentric circular regions of a cell, and the number of available PSs during the coherence interval for performing channel estimation. In our system model, each PD-NOMA pair consists of a far UE in the edge (outer) region and a near UE in the central (inner) region. Each PD-NOMA pair is formed by a random selection of the two UEs. The maximum distance between a near user (resp., a far user) and the BS is denoted by d_n (resp., d_f), where $d_f > d_n$. The two UEs in the PD-NOMA pair (i.e., the near UE and the far UE) transmit their uplink signals simultaneously and share the same sub-band. We assume a fully loaded cell where all $2K$ users⁴ are active and formed

⁴In massive MIMO $2K \leq \frac{M}{2}$ is recommended. The deployment of PD-NOMA in massive MIMO allows the number of users to exceed the number of available PSs in the coherence interval, this results in increasing the system capacity

K PD-NOMA pairs.

Consequently, (3.7) should be updated to take into account that PD-NOMA is now deployed in the massive MIMO cell instead of OMA. Therefore, in this case, the received signal at the m th BS antenna, through a favorable propagation environment that has sufficient scatters that produce dense multipath components for the l th channel tap is given by:

$$y_m^l = \sqrt{\rho_{ul}} \sum_{k=1}^K \left(h_{k_n}^{m,l} \sqrt{\eta_k} x_{k_n} + h_{k_f}^{m,l} \sqrt{\eta_k \delta} x_{k_f} \right) + w_m^l, \quad (3.23)$$

where δ denotes the back-off step for the far UE to ensure the signal power received at the BS from the far UE is less than the signal power received from the near UE, $\delta \in (0, 1]$ (in PD-NOMA, the system allocates more power to the UEs with weaker wireless channel coefficients). In the k th PD-NOMA pair, x_{k_n} denotes the uplink signal transmitted by the near UE (i.e., the user belongs to the inner region and closer to the cell center) and x_{k_f} denotes the uplink signal transmitted by the far UE (i.e., the user belongs to the outer region and closer to the cell edge). Taking into account only one BC DFT-s-OFDM symbol, both signals x_{k_n} and x_{k_f} become vectors denoted by \mathbf{x}_{k_n} and \mathbf{x}_{k_f} , respectively, which are generated at the transmitters of UEs using (3.12), where the near UE and the far UE at each PD-NOMA pair share the same OFDM sub-carriers during the uplink transmission. The additive noise which is assumed to be $\mathcal{CN}(0, 1)$ and independent across different l and m is denoted by w_m^l .

For ease of presentation, let us consider only the uplink transmission of the k th PD-NOMA pair in the cell through the l th transmission path with assuming both x_{k_n} and x_{k_f} become vectors, thus (3.23) becomes:

$$\mathbf{y}_{m_k}^l = \sqrt{\rho_{ul}} \left(h_{k_n}^{m,l} \sqrt{\eta_k} \mathbf{x}_{k_n} + h_{k_f}^{m,l} \sqrt{\eta_k \delta} \mathbf{x}_{k_f} \right) + \mathbf{w}_{m_k}^l, \quad (3.24)$$

where the wireless channel coefficient of the l th transmission path between the near UE in the k th PD-NOMA pair and the m th BS's antenna is $h_{k_n}^{m,l} = \sqrt{\beta_{k_n}^{m,l}} g_{k_n}^{m,l}$ whereas $h_{k_f}^{m,l} = \sqrt{\beta_{k_f}^{m,l}} g_{k_f}^{m,l}$ is the channel coefficient of the l th transmission path between the far UE in the k th PD-NOMA pair and to serve more users. In this case by a factor of 2.

the m th BS's antenna.

At the k th PD-NOMA pair, let $\beta_{k_n}^{m_l} = d_{nk}^{-\rho}$ and $\beta_{k_f}^{m_l} = d_{fk}^{-\rho}$ denote the large-scale fading coefficients of the near UE and of the far UE respectively, where d_{nk} and d_{fk} are the distances from the near UE and from the far UE to the BS respectively (i.e., $d_{fk} > d_{nk}$), and ρ is the path loss exponent of the cell environment. For each PD-NOMA pair, the BS receiver assumes the availability of perfect CSI and employs SIC. Thus, the BS's receiver structural configuration that is given in Figure 3.5 should be modified to accommodate the deployment of PD-NOMA by embedding the SIC circuitry block. Thus, in this scenario, the red block in the receiver diagram in Figure 3.5 includes a SIC algorithm in addition to LMMSE estimator. Similarly, the transmitter needs to be modified at the UE. Consequently, the the red block in the UE diagram in Figure 3.4 includes a power allocation algorithm in addition to the sub-carriers allocation mapper.

The BS's receiver decodes the signal of the near user first while considering the presence of the far user's signal as interference. The nearest user, which has the strongest channel, experiences interference from all other users in the PD-NOMA pair during decoding its uplink signal at the BS. Then, as a second step, the BS subtracts the decoded signal $\hat{\mathbf{x}}_{k_n}$ of the near user from the total received signal $\mathbf{y}_{m_k}^l$, and the remaining signal $\mathbf{y}_{m_k}^l - h_{k_n}^{m_l} \sqrt{\eta_k} \hat{\mathbf{x}}_{k_n}$ will be decoded as the received uplink signal of the far user. Therefore, the far user (i.e., the user served by the weakest channel) does not experience interference from other users in the PD-NOMA pair during the decoding of its uplink signal (i.e., assuming perfect SIC decoding is performed based on perfect CSI). Consequently, the achievable average rates per sub-band for the near and the far users in the uplink PD-NOMA massive MIMO are given by:

$$\bar{r}_{nUE}^{UL} = \mathbb{E} \left\{ \log_2 \left(1 + \frac{\rho_{ul} \eta_k \|h_{k_n}^{m_l} \mathbf{x}_{k_n}\|_2^2}{\|h_{k_f}^{m_l} \mathbf{x}_{k_f}\|_2^2 + \sigma_w^2} \right) \right\}. \quad (3.25)$$

$$\bar{r}_{fUE}^{UL} = \mathbb{E} \left\{ \log_2 \left(1 + \frac{\rho_{ul} \delta \eta_k \|h_{k_f}^{m_l} \mathbf{x}_{k_f}\|_2^2}{\sigma_w^2} \right) \right\}. \quad (3.26)$$

Equations (3.25) and (3.26) clearly demonstrate the useful advantage of PD-NOMA deployment in massive MIMO to increase the sum-rate capacity of the cell. Also, (3.25) and (3.26) imply that this improvement in the sum-rate capacity, which is directly proportional to the numerator of the EE metric, depends on the accuracy of the CSI.

3.6.2 PD-NOMA Integration in Massive MIMO with Imperfect CSI

In the previous sub-section, we assume the availability of perfect CSI at the BS receiver, where accurate CSI is essential for perfect operations of SIC in a system that adopted PD-NOMA integration. Moreover, accurate CSI is essential and very important to build a reliable wireless communication system. This sub-section focuses on highlighting the practical aspects of PD-NOMA deployment in massive MIMO, where the inherent interference impairs the accuracy of CSI.

As mentioned in Section 3.5 we consider a cell that is located at a center within a cluster consisting of non-cooperative multiple adjacent cells that use DPS for estimating CSI under this scenario of operating. DPS significantly mitigates the inter-cells interference caused by users in adjacent cells as explained in Chapter 2, where OMA has been used with massive MIMO. This new structure of PD-NOMA deployment in a massive MIMO cell creates an opportunity to achieve better EE but also introduces new design constraints that make the EE maximization challenging. Notwithstanding the existence of inter-cell⁵ interference is tackled by DPS, this new configuration introduces an intra-cell interference, a new source of interference that arises from all UEs within the same cell that share the same sub-band which needs to be mitigated.

Consequently, the accuracy of CSI in NOMA systems plays an additional and critical role in mitigating the error propagation in the SIC operation that is carried out at the BS receiver during the decoding of received uplink signals. Therefore, it is essential to tackle the new intra-cell interference. Moreover, the CSI is used in the pre-coding and digital beam-forming operations for the downlink signals as well. As a consequence, an inaccurate CSI causes imperfect beam-forming,

⁵The inter-cell interference of downlink in multiple cells with either PD-NOMA or OMA configurations have the same level. In contrast, the inter-cell interference of uplink in PD-NOMA configuration is different than OMA inter-cell interference, where it is directly proportional to the number of served users in PD-NOMA pairs at co-channel adjacent cells.

adds a new source of interference amongst the served users during the downlink, and causes a reduction in the system EE.

Definitely, the assumption of perfect CSI availability for the PD-NOMA pair at the BS's receiver is not practical. Instead, the BS estimates the CSI during the uplink training using the uplink PSs. Therefore, the accuracy of the CSI depends on the quality of the channel estimation at the BS. Although the system model in this chapter focuses on the performance of a single cell located at the center of non-cooperative multiple adjacent cells, where inter-cell interference is assumed to be tackled by DPS. In practice, it is important to adopt an interference mitigation scheme that considers both inter-cell and intra-cell interference. Therefore, to mitigate the effects of the intra-cell interference amongst PD-NOMA pairs and improve CSI accuracy at the intended cell, we introduce EDPS and use it in this sub-section. EDPS is an extension of DPS that operates on two levels in a cell of massive MIMO with NOMA deployment⁶. The first level to mitigate the inter-cell interference and the second mitigates the intra-cell interference. Therefore, using EDPS, PSs assignment in non-cooperative multiple adjacent cells for the NOMA massive MIMO is given by $\Phi_{R \times K \times C}$, as follows:

$$\Phi_{R \times K \times C} = \begin{bmatrix} \left[\begin{array}{ccc} \phi_1^{1^1} & \cdots & \phi_K^{1^1} \\ \vdots & \ddots & \vdots \\ \phi_1^{R^1} & \cdots & \phi_K^{R^1} \end{array} \right] & \cdots & \left[\begin{array}{ccc} \phi_1^{1^C} & \cdots & \phi_K^{1^C} \\ \vdots & & \vdots \\ \phi_1^{R^C} & \cdots & \phi_K^{R^C} \end{array} \right] \end{bmatrix}, \quad (3.27)$$

where C is the number of non-cooperative multiple adjacent cells in the NOMA massive MIMO cluster, R is the number of concentric circular regions inside each cell, K is the total number of users in each circular region, and $\phi_c^{r^k}$ is the PS used by the k th UE in the r th circular region within the c th cell. It represents a row vector and is normalized to 1, $\|\phi_c^{r^k}\|_2 = 1$ for all r, k , and c . Then

⁶EDPS can be employed in different types of NOMA like CD-NOMA.

$\phi_c^{r^k}$ can be decomposed as follows:

$$\phi_c^{r^k} = \sqrt{\mu} \mathbf{v}_r + \sqrt{\gamma(1-\mu)} \psi_k + \sqrt{(1-\gamma)(1-\mu)} \varphi_c, \quad (3.28)$$

where $\mu \leq 1$ is a NOMA pair design parameter, $\gamma \leq 1$ is a massive MIMO cells cluster design parameter, and \mathbf{v}_r , ψ_k , and φ_c are mutually orthogonal normalized energy sequences that belong to the orthogonal sequences codebook available in the system. Thus, the PSs codebook consists of $K + R + C$ mutual orthogonal sequences and can be categorized into three groups as follows:

$$\begin{aligned} \text{UESG} &= \{\psi_1, \psi_2, \dots, \psi_K\}, \\ \text{NRSG} &= \{\mathbf{v}_i, \mathbf{v}_{ii}, \dots, \mathbf{v}_R\}, \\ \text{MCSG} &= \{\varphi_a, \varphi_b, \dots, \varphi_C\}, \end{aligned} \quad (3.29)$$

where UESG stands for the Users Equipment Sequences Group, NRSG stands for the NOMA Regions Sequences Group, and MCSG stands for the massive MIMO Cells-Clustering Sequences Group. In the system model used in the previous sub-section, the analysis focused on one cell with $R = 2$, then the codebook used in the intended cell (denoted by cell A) given by (3.29) becomes:

$$\begin{aligned} \text{UESG} &= \{\psi_1, \psi_2, \dots, \psi_K\}, \\ \text{NRSG} &= \{\mathbf{v}_i, \mathbf{v}_{ii}\}, \\ \text{MCSG} &= \{\varphi_a\}, \end{aligned}$$

where φ_a is used as the Cell-ID, \mathbf{v}_i is used as the central region Identification (ID) (i.e., the inner region contains the near UEs inside the cell), \mathbf{v}_{ii} is used as the ID of the edge region (i.e., the outer region contains the far UEs inside the cell), and $\psi_1, \psi_2, \dots, \psi_K$ are the User-IDs and used in conjunction with the Cell-ID and NOMA regions IDs to identify the UEs within each NOMA

region inside the cell. As a consequence (3.27) can be rewritten as:

$$\Phi_A = \begin{bmatrix} \phi_1^{1^a} & \cdots & \phi_K^{1^a} \\ \phi_1^{2^a} & \cdots & \phi_K^{2^a} \end{bmatrix}_{2 \times K}. \quad (3.30)$$

Now, to illustrate the reasons why employing the conventional SPS will result in imperfect CSI and why the EDPS enhances the accuracy of CSI, let us consider the two scenarios in integrating PD-NOMA in massive MIMO.

PD-NOMA Deployment in Massive MIMO with Conventional SPS Integration

If SPS is used in the deployment of PD-NOMA in a cell of massive MIMO, then only one PS is used by each PD-NOMA pair. Therefore, using SPS, the two UEs in the k th PD-NOMA pair are assigned the same PS, $\phi_k = \psi_k$. Moreover, the same PS is reused at adjacent cells of the non-cooperative multiple cells PD-NOMA massive MIMO cluster which worsens the level of inter-cell interference due to the PD-NOMA deployment. Also, it leads to arising the intra-cell interference amongst PD-NOMA pairs. Then, considering a single cell, the uplink pilot signal received at the BS antenna m transmitted by UEs of the k th PD-NOMA pair through the transmission path l is given by:

$$\mathbf{y}_{m_k}^l = \sqrt{\rho_{ul}} \left(h_{k_n}^{m^l} \sqrt{\eta_k} + h_{k_f}^{m^l} \sqrt{\eta_k \delta} \right) \phi_k + \mathbf{w}_{m_k}^l. \quad (3.31)$$

The BS receiver performs the so-called de-spreading operation by right-multiplying $\mathbf{y}_{m_k}^l$ by ϕ_k^H , where $\phi_k \phi_k^H = 1$ (i.e., due to the unit energy normalization), obtaining:

$$\mathbf{y}_{m_k}^l \phi_k^H = \sqrt{\rho_{ul}} \left(h_{k_n}^{m^l} \sqrt{\eta_k} + h_{k_f}^{m^l} \sqrt{\eta_k \delta} \right) + \mathbf{w}_{m_k}^l \phi_k^H. \quad (3.32)$$

The BS uses the conventional SPS to estimate the CSI. Equation (3.32) is the only equation for the two unknowns $h_{k_n}^{m^l}$ and $h_{k_f}^{m^l}$ available at the BS using the conventional SPS. As a result, from a geometric perspective, the BS receiver's MMSE estimator is limited to generate three estimates that are parallel and aligned in the same channel direction for $h_{k_n}^{m^l}$, $h_{k_f}^{m^l}$, or the linear combination

$h_{k_n}^{n_l} + h_{k_f}^{n_l}$. Thus, the resulted estimated CSI will be imperfect. Beside, from a beam-forming perspective, the BS transceiver will not be able to distinguish the correct direction for the near UE and the far UE in the PD-NOMA pair due to imperfect estimation of the CSI. Although the best approach here is to use the linear combination of the estimated wireless channel coefficients and employ the joint beam-forming. However, this will cause an imperfect pre-coding and improper beam-forming in the downlink, reduction in the EE, and will add a new source of interference amongst UEs.

PD-NOMA Deployment in Massive MIMO with EDPS Integration

To illustrates the advantages of the EDPS for mitigating both the intra-cell and inter-cell interference, let us consider a PD-NOMA massive MIMO cluster consisting of two non-cooperative adjacent cells as shown in Figure 3.6 (denoted as A and B). Each cell is divided into two concentric circular regions that employ the PD-NOMA, the central region and the edge region (i.e., $R = 2$). According to this configuration, each cell has K PD-NOMA pairs where each pair consists of 2 UEs. As mentioned previously, all EDPS PSs have normalized unit energy. Then, in EDPS, the near UEs (i.e., those located in the central region) of the intended cell A are assigned with PSs expressed by:

$$\begin{aligned} \phi_a^{n^k} &= \sqrt{\mu} \mathbf{v}_i + \sqrt{\gamma(1-\mu)} \boldsymbol{\psi}_k + \sqrt{(1-\gamma)(1-\mu)} \boldsymbol{\varphi}_a. \\ &\text{for all } k=1,2,\dots,K. \end{aligned} \quad (3.33)$$

Whereas the far UEs (i.e., those located in the edge region) of the intended cell A are assigned with PSs expressed by:

$$\begin{aligned} \phi_a^{f^k} &= \sqrt{\mu} \mathbf{v}_{ii} + \sqrt{\gamma(1-\mu)} \boldsymbol{\psi}_k + \sqrt{(1-\gamma)(1-\mu)} \boldsymbol{\varphi}_a. \\ &\text{for all } k=1,2,\dots,K. \end{aligned} \quad (3.34)$$

Similarly, the near UEs of the adjacent cell B are assigned with PSs expressed by:

$$\phi_b^{n^k} = \sqrt{\mu} \mathbf{v}_i + \sqrt{\gamma(1-\mu)} \psi_k + \sqrt{(1-\gamma)(1-\mu)} \varphi_b. \quad (3.35)$$

for all $k=1,2,\dots,K$.

Whereas the far UEs of the adjacent cell B are assigned with PSs expressed by:

$$\phi_b^{f^k} = \sqrt{\mu} \mathbf{v}_{ii} + \sqrt{\gamma(1-\mu)} \psi_k + \sqrt{(1-\gamma)(1-\mu)} \varphi_b. \quad (3.36)$$

for all $k=1,2,\dots,K$.

At the base station of cell A , the received signal consists of all pilot signals transmitted by all UEs in the central region and the edge region of both cells A and B . Then, the received uplink pilot signal at the m th BS receiver antenna transmitted by the k th PD-NOMA pairs through the transmission path l is given by:

$$\begin{aligned} \mathbf{y}_{m_k}^l = & \sqrt{\rho_{ul}} \left(h_{k_{na}}^{m^l} \sqrt{\eta_k} + h_{k_{nb}}^{m^l} \sqrt{\eta_k} + h_{k_{fa}}^{m^l} \sqrt{\eta_k \delta} + h_{k_{fb}}^{m^l} \sqrt{\eta_k \delta} \right) \psi_k \\ & + \sqrt{\rho_{ul}} \left(h_{k_{na}}^{m^l} \sqrt{\eta_k} + h_{k_{nb}}^{m^l} \sqrt{\eta_k} \right) \mathbf{v}_i + \sqrt{\rho_{ul}} \left(h_{k_{fa}}^{m^l} \sqrt{\eta_k \delta} + h_{k_{fb}}^{m^l} \sqrt{\eta_k \delta} \right) \mathbf{v}_{ii} \\ & + \sqrt{\rho_{ul}} \left(h_{k_{na}}^{m^l} \sqrt{\eta_k} + h_{k_{fa}}^{m^l} \sqrt{\eta_k \delta} \right) \varphi_a + \sqrt{\rho_{ul}} \left(h_{k_{nb}}^{m^l} \sqrt{\eta_k} + h_{k_{fb}}^{m^l} \sqrt{\eta_k \delta} \right) \varphi_b + \mathbf{w}_{m_k}^l. \end{aligned} \quad (3.37)$$

Equation (3.37) has four unknowns: 1) $h_{k_{na}}^{m^l}$ and $h_{k_{fa}}^{m^l}$ which are the two channel coefficients for the near UE and the far UE at cell A , respectively, and 2) $h_{k_{nb}}^{m^l}$ and $h_{k_{fb}}^{m^l}$ which are the two channel coefficients for the near UE and the far UE at cell B , respectively.

The receiver of the BS at cell A performs the de-spreading operation by right-multiplying $\mathbf{y}_{m_k}^l$ by ψ_k^H , where $\psi_k \psi_k^H = 1$, obtaining:

$$\mathbf{y}_{m_k}^l \psi_k^H = \sqrt{\rho_{ul}} \left(h_{k_{na}}^{m^l} \sqrt{\eta_k} + h_{k_{nb}}^{m^l} \sqrt{\eta_k} + h_{k_{fa}}^{m^l} \sqrt{\eta_k \delta} + h_{k_{fb}}^{m^l} \sqrt{\eta_k \delta} \right) + \mathbf{w}_{m_k}^l \psi_k^H. \quad (3.38)$$

The PD-NOMA massive MIMO system integrated with the conventional SPS is limited to do one de-spreading operation given by (3.38). In contrast, the PD-NOMA massive MIMO system

integrated with the EDPS can do three additional de-spreading operations at the receiver of the BS at cell A. These de-spreading operations are accomplished by right-multiplying $\mathbf{y}_{m_k}^l$ by \mathbf{v}_i^H to generate (3.39), \mathbf{v}_{ii}^H to generate (3.40), and $\boldsymbol{\varphi}_a^H$ to generate (3.41). Where $\mathbf{v}_i \mathbf{v}_i^H = 1$, $\mathbf{v}_{ii} \mathbf{v}_{ii}^H = 1$, and $\boldsymbol{\varphi}_a \boldsymbol{\varphi}_a^H = 1$ as each sequence has a normalized unit energy. Thus, these additional equations are given by:

$$\mathbf{y}_{m_k}^l \mathbf{v}_i^H = \sqrt{\rho_{ul}} \left(h_{k_{na}}^{m^l} \sqrt{\eta_k} + h_{k_{nb}}^{m^l} \sqrt{\eta_k} \right) + \mathbf{w}_{m_k}^l \mathbf{v}_i^H. \quad (3.39)$$

$$\mathbf{y}_{m_k}^l \mathbf{v}_{ii}^H = \sqrt{\rho_{ul}} \left(h_{k_{fa}}^{m^l} \sqrt{\eta_k \delta} + h_{k_{fb}}^{m^l} \sqrt{\eta_k \delta} \right) + \mathbf{w}_{m_k}^l \mathbf{v}_{ii}^H. \quad (3.40)$$

$$\mathbf{y}_{m_k}^l \boldsymbol{\varphi}_a^H = \sqrt{\rho_{ul}} \left(h_{k_{na}}^{m^l} \sqrt{\eta_k} + h_{k_{fa}}^{m^l} \sqrt{\eta_k \delta} \right) + \mathbf{w}_{m_k}^l \boldsymbol{\varphi}_a^H. \quad (3.41)$$

Equations (3.38), (3.39), (3.40), and (3.41) are linearly independent and generated simultaneously at the BS receiver. Therefore, the BS receiver now has the capability to jointly and accurately estimate the unknown wireless channel coefficients, as the number of unknowns are equal to the number of linearly independent equations. Thus, the EDPS enhances the CSI accuracy and outperforms the conventional SPS in the PD-NOMA massive MIMO system. This can be generalized for a group of PD-NOMA massive MIMO non-cooperative adjacent cells that form a cluster and serve any arbitrary number of users, which has been proving mathematically in our work presented in [78], where dual PSs integration resulted in an enhancement of CSI accuracy compared to single PSs employment and CSI were estimated with a smaller MMSE during channel estimation. As a result of this CSI enhancement, the sum-rate capacity of the PD-NOMA massive MIMO system increases and as a consequence, the system EE is improved.

Similar to the macroscopic view presented in Section 2.7, the concept of cluster reuse is used as well in massive MIMO-PD-NOMA but with a smaller cluster size (i.e., number of cells L in the cluster). Figure 3.7 shows the reuse planning for the entire cellular network using a cluster size of

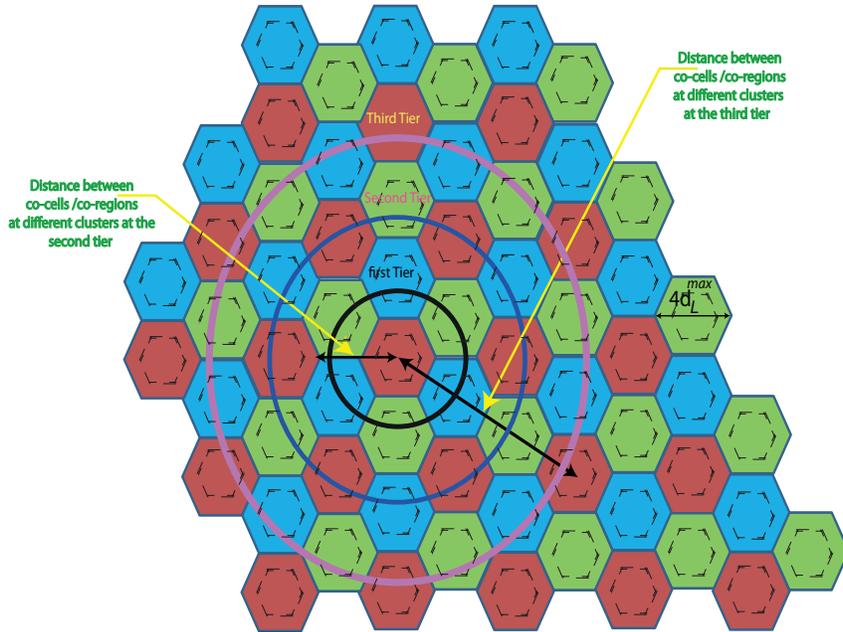


Figure 3.7: EDPS cluster reuse in massive MIMO PD-NOMA cellular network. Cells with the same color are using the same unique BS’s extended dual PSs (i.e., the same Cell-ID (MCSG), same Region-IDs (NRSGs), and User-IDs (UESGs)).

three cells (i.e., $L = 3$). The idea behind using a smaller cluster size is adopted to tackle the trade-off between CSI accuracy and system capacity. As aforementioned in Chapter 2, the number of available orthogonal sequences N is limited and mainly determines by the duration of the training period within a coherence interval. Hence, portions of these orthogonal sequences are dedicated for MCSGs and NRSGs within a cluster. Therefore, to increase the system capacity the orthogonal sequences dedicated for UESGs need to be increased which implies that a smaller cluster size is recommended. However, to increase the reuse distance the NRSGs will be used reversely between inner region and outer region of the co-cell every consecutive tier.

3.7 Numerical Results and Discussion

3.7.1 Generating the Channel Matrix

The 3D channel matrix \mathbf{H} of the i.i.d. Rayleigh fading channel, which we have used in simulations, was generated such that its entries are spatially and temporally independent and follow a circularly

Table 3.1: System parameters used in simulations

Parameter	Value
Number of BS Antennas M	128
Carrier Frequency	2 GHz
BS Antenna Height	25 m
UE Mobility Speed	80 km/h
Cell Radius	250 m
Bandwidth	20 MHz
Number of UEs K	32
UE Height	1.7 m
σ_{sh} [124]	8 dB
Maximum Delay Spread	10
QAM Constellation Size	32
Length of Barker Codes (N)	3, 5, and 7
Maximum Number of Sub-Carriers	1272

symmetric complex Gaussian random distribution with zero mean and unit variance. The channel coefficient of the l th transmission path between the k th UE and the m th BS antenna, $h_k^{m,l}$, is located at the m th row and k th column of the 2D matrix \mathbf{H}^l . The 3D channel matrix of the correlated Rayleigh fading channel is formed according to the model used in [123] based on the 3D i.i.d. channel matrix mentioned above and with taking into account both the temporal and the spatial correlations between the entries of the i.i.d. channel matrix.

3.7.2 Numerical Results for OMA Massive MIMO Cell Configuration

We consider a system with an urban outdoor macro-cell massive MIMO with a radius of 250 meters whose BS is located at the center as shown in Figure 3.3. The cell serves K active users simultaneously. The PSs are generated and assigned using the DPS described in Chapter 2 to estimate the CSI. We assume a uniform random distribution of all users around the BS with equal power allocation. Also, we assume the scatterers are distributed randomly in each cell. Furthermore, OFDM is converting a wide-band frequency-selective fading channel into a group of parallel narrow-band frequency-flat fading channels. Then, in a frequency-flat fading channel, we assume each frame

occupies the channel coherence interval. Each coherence interval ($I_c = B_c \times T_c$) is divided into four parts, the uplink training interval, the uplink data transmission interval, the downlink beam-forming (linear pre-coding) interval, and the downlink data transmission interval.

In all of our simulations for both conventional OMA massive MIMO configuration and massive MIMO with PD-NOMA deployment: We use MATLAB as the simulation environment. The length of the BC is adaptive and set to three, five, or seven. Table 3.1 shows the system parameters used in the simulations of the numerical results. In particular, we applied the large scale fading model used in [96, 124], and [119] in our simulations.

Normalized Mean Squared Error:

The uplink system throughput is defined as the rate of the successfully delivered data symbols transmitted by UEs to the BS over the wireless communication channel. To receive these transmitted data symbols successfully, the BS receiver should decode and estimate the received uplink signals accurately which depends on the accuracy of channel estimation. Thus, increasing the estimation accuracy of the CSI is essential to improve the uplink system throughput. This accuracy enhancement is measured by the NMSE. Lower NMSE implies a better channel estimation and a higher accuracy of the CSI which has a consequence of enhancing the system sum-rate capacity and reducing the BER, using the same transmission power. As a result, it leads to improving EE.

Therefore, we use the NMSE of the CSI estimation as a performance metric to evaluate the performance of the proposed BC DFT-s-OFDM waveform and compare it with the performance of the conventional DFT-s-OFDM waveform in the uplink transmission. The normalized mean squared channel estimation error is given by $NMSE_k^{m^l} = \frac{1}{N} \sum_{n=1}^N \mathbb{E} \left\{ \frac{\|\hat{h}_{k,n}^{m^l} - h_k^{m^l}\|^2}{\|h_k^{m^l}\|^2} \right\}$ where $\hat{h}_{k,n}^{m^l}$ is the estimated channel coefficient, $h_k^{m^l}$ is the actual channel coefficient and n is the simulation index. Figure 3.8 shows the NMSE of the uplink conventional DFT-s-OFDM and the NMSE of the uplink BC DFT-s-OFDM when the i.i.d. and the correlated Rayleigh fading channels are used. The curves show that the uplink BC DFT-s-OFDM waveform outperforms the conventional DFT-s-OFDM waveform at low SNR. Improving the NMSE for low SNR will result in using low power

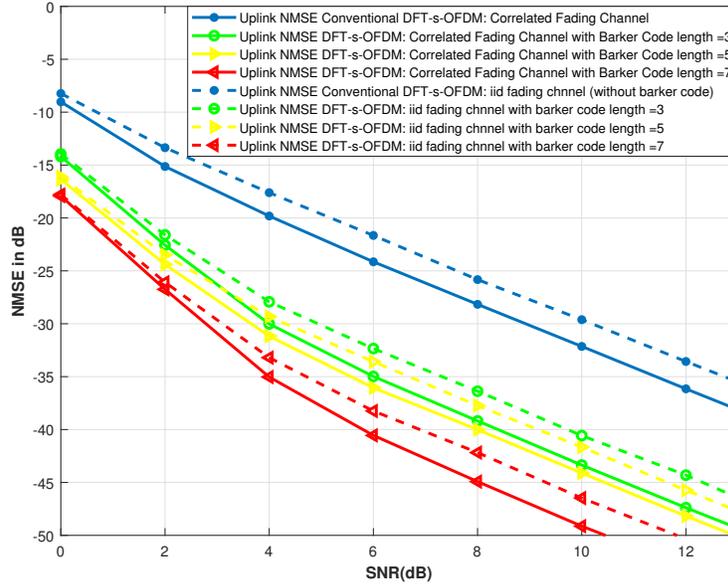


Figure 3.8: The CSI-NMSE performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.

transmission (i.e., the dominant part at the denominator of EE metric) and hence improve the EE. So, DFT-s-OFDM with BC spreader is shown to be an energy-efficient waveform for massive MIMO.

Sum-Rate Capacity:

The ergodic sum-rate capacity of a massive MIMO cell provides information about the achieved average data throughput. Throughput is the numerator of an EE metric, therefore increasing the sum-rate capacity enhances the system EE. Figure 3.9 shows the sum-rate capacity performance metric of uplink waveforms for both the conventional DFT-s-OFDM and the proposed BC DFT-s-OFDM when both i.i.d. and correlated Rayleigh fading communications channels are used. The curves show that the proposed scheme (i.e., the DFT-s-OFDM with adaptive length BCs) achieved better performance in both i.i.d and the correlated Rayleigh fading communications channels environments at low SNR and consequently outperforms the conventional DFT-s-OFDM in massive MIMO system. Thus, the proposed UE's transmitter and BS's receiver structures improve the sum-

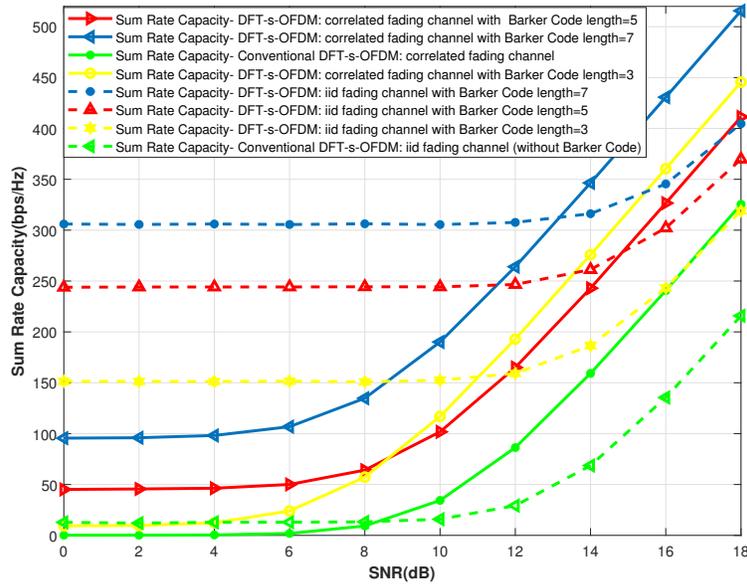


Figure 3.9: The uplink sum-rate capacity performance metric with $M=128$ (BS’s antennas) and $K=32$ (active UEs) using 32-QAM.

rate capacity performance and boost the average high data throughput at low SNR. In other words, these proposed structures improve EE significantly and make the green communications promise of 5G more feasible to implement by adopting the proposed structures.

Bit Error Rate:

BER in the uplink transmission is another important metric. Having an energy-efficient waveform implies that lower BER can be achieved at low SNR. On the other hand, achieving low BER leads to lower power consumption in the system due to the reduction of corrupted data symbols received at the BS, which implies a reduction in the re-transmission rate.

Figure 3.10 shows the BER performance metric of uplink waveforms for both the conventional DFT-s-OFDM and the BC DFT-s-OFDM when both i.i.d. and correlated Rayleigh fading communications channels are used. The curves show that the proposed BC DFT-s-OFDM scheme outperforms the conventional DFT-s-OFDM at low SNR. This results in improving the EE of the 5G massive MIMO system. On the other hand, Quadrature Amplitude Modulation (QAM) is the

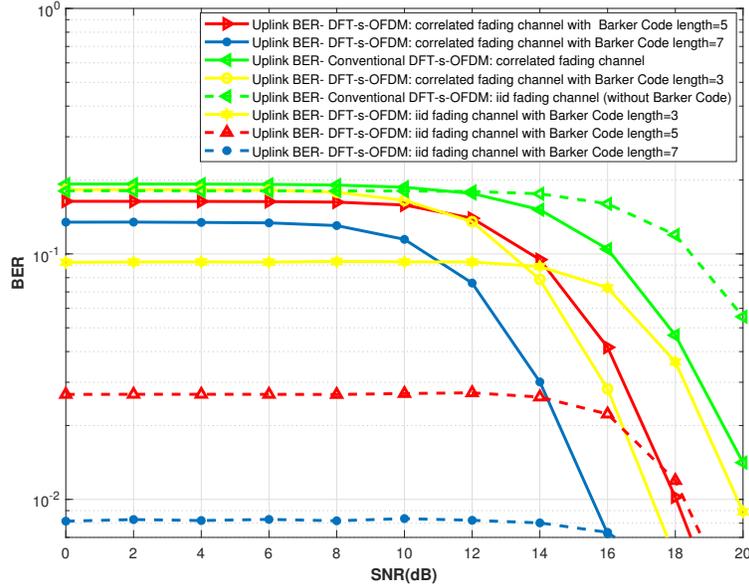


Figure 3.10: The BER performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.

modulation scheme used in the 5G and our proposed structure is shown to be energy efficient and to achieve low BER at low transmitted power when using this modulation scheme. The transmitted power is the dominant part of the denominator of the EE metric, therefore reducing the transmission power improves the system EE. We used QAM with a constellation size of 32 in the numerical results but this can be generalized to any constellation size of the QAM. The selection of the optimal QAM constellation size is an optimization problem where the required transmission power to maintain the same BER performance increases as the QAM constellation size increases. There have been works done on 5G cellular networks to be equipped with optimization algorithms to allocate adaptive QAM constellation size to optimize both EE and SE. Our proposed scheme enables these algorithms to work more efficiently.

Energy Efficiency:

The EE, in the uplink transmission, is an essential metric that should be used to assess the proposed scheme EE. Having an energy-efficient waveform implies that high data throughput and high SE

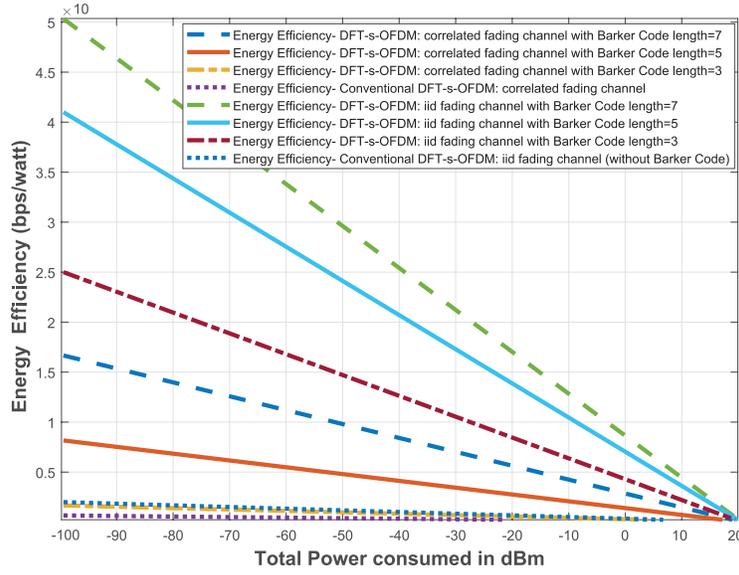


Figure 3.11: The Energy Efficiency performance metric with $M=128$ (BS's antennas) and $K=32$ (active UEs) using 32-QAM.

could be achievable with low power consumption at both UE's transmitter and BS's receiver. Our system model measured the received uplink SINR ratio at the BS and the EE metric is derived for the uplink massive MIMO according to the same mathematical modeling used in [125], and given by:

$$EE = \frac{\sum_{k=1}^K r_k}{\sum_{k=1}^K p_k + P_{CTx} + P_{MS}}, \quad (3.42)$$

where, P_{MS} is a power consumption adjustment factor at the UE, P_{CTx} is the circuit power consumed at the UE, p_k is the total average transmission power consumed by sub-carriers at k th UE, and r_k is the achievable average rate capacity of the k th user measured at the BS. The SINRs for both the conventional DFT-s-OFDM and the BC DFT-s-OFDM are measured at the BS where the LMMSE estimation algorithm is used to estimate the received uplink signals. Also, the LMMSE estimator employs the DPS during the phase of channel estimation which adds a second degree of improvement in EE due to DPS's capability to enhance the accuracy of CSI at low SNR.

Figure 3.11 shows the EE performance metric of uplink waveforms for both the conventional DFT-s-OFDM and the DFT-s-OFDM with adaptive length BCs when both i.i.d. and correlated

Rayleigh fading communications channels are used. The curves show that the proposed BC DFT-s-OFDM scheme outperforms the conventional DFT-s-OFDM with a significant gain at low power consumption. It also implies, employing BC DFT-s-OFDM results in improving the EE of the 5G massive MIMO system which is the technical infrastructure of many networks requiring low power consumption like IoT. So, implementation of the proposed BC DFT-s-OFDM scheme in 5G networks will result in prolonging the battery lifetime, especially of limited energy devices such as IoT sensors. This has huge economic, social, and environmental impacts that make 1) green communications and 2) affording the deployment of smart ICT infrastructures, are feasible to achieve.

3.7.3 Numerical Results for PD-NOMA Deployment in Massive MIMO

Similarly, we consider a system with an urban outdoor macro-cell PD-NOMA massive MIMO with a radius of 250 meters whose BS is located at the center as shown in Figure 3.6. The cell is divided into two concentric circular regions. The inner central region contains the near users and the outer edge region contains the far users. All users in each region are uniformly and randomly distributed around the BS. The BS controls the UEs transmitted power and controls the back-off power step to ensure different power levels between far UE and near UE during the reception of the uplink signals. The cell serves $2K$ active users simultaneously, K users at each region and forms K PD-NOMA pairs. The PSs are generated and assigned using the EDPS described and proposed in Section 3.6, to estimate the CSI. Similar to the conventional OMA massive MIMO cell configuration, we also assume the scatterers are distributed randomly in each call to ensure the existence of favorable propagation conditions at massive MIMO system with PD-NOMA deployment. Identically, we also assume a frequency flat fading channel where each frame occupies the channel coherence interval. Where each coherence interval ($I_c = B_c \times T_c$) is divided into four parts, the uplink training interval, the uplink data transmission interval, the downlink beam-forming (linear pre-coding) interval, and the downlink data transmission interval.

Normalized Mean Squared Error:

We have examined the performance of EDPS over a range of SNR values to assess its estimation accuracy when the proposed BC DFT-s-OFDM waveform is used in the uplink transmission of massive MIMO with PD-NOMA at various BC length. We have used NMSE as the performance metric. Then, we have compared the obtained results with the performance of the conventional DFT-s-OFDM waveform when it is used in the uplink transmission in massive MIMO with PD-NOMA. Table 3.2 shows that the EDPS achieves a high estimation accuracy and the BC DFT-s-OFDM waveform outperforms the performance of the conventional DFT-s-OFDM waveform when both are used in the uplink transmission in massive MIMO with PD-NOMA. We have performed the evaluations when both the i.i.d. and the correlated Rayleigh fading channels are used. Also, we found that the NMSE in EDPS is monotonically decreasing as the SNR increases. The obtained numerical results demonstrated that improving the NMSE for low SNR results in using low power transmission (i.e., the dominant part at the denominator of EE metric) and hence improve the EE, whereas at the same time it is very well known that PD-NOMA improves the SE. This improves the EE-SE trade-off in the massive MIMO with PD-NOMA uplink transmission.

Table 3.2: NMSE Performance metric in dB for DFT-s-OFDM and BC DFT-s-OFDM in PD-NOMA at different values of SNR with various BC length, using i.i.d. and correlated Rayleigh fading channels at $M = 128$ and active UE= 32.

SNR (dB)	i.i.d.				Correlated			
	DFT-s-OFDM	BC DFT-s-OFDM			DFT-s-OFDM	BC DFT-s-OFDM		
		BC=3	BC=5	BC=7		BC=3	BC=5	BC=7
0	-08.4615	-15.8462	-17.3462	-19.9808	-10.4231	-16.2462	-17.3462	-19.9885
2	-15.1823	-24.0231	-25.3692	-27.5793	-16.9789	-24.7923	-26.6191	-28.8693
4	-19.0371	-28.8646	-31.6992	-35.1706	-21.1673	-32.1705	-33.6221	-36.1775
6	-22.9885	-33.5577	-35.5769	-39.8846	-25.6731	-36.7961	-38.3654	-42.9233
8	-27.1192	-37.3115	-39.1002	-44.7497	-29.8269	-40.6038	-42.0497	-46.8477
10	-31.7415	-41.2646	-43.0962	-48.8454	-34.1608	-44.9112	-46.7108	-50.7101

Sum-Rate Capacity:

We have used the sum-rate as the performance metric to assess the performance of the proposed BC DFT-s-OFDM waveform over a range of SNR values when it is used in the uplink transmission of massive MIMO with PD-NOMA at various BC length. Then, we have compared the obtained results with the performance of the conventional DFT-s-OFDM waveform when it is used in the uplink transmission of massive MIMO with PD-NOMA. Table 3.3 shows that the BC DFT-s-OFDM waveform achieves a better performance and outperforms the performance of the conventional DFT-s-OFDM waveform when both are used in the uplink transmission of massive MIMO with PD-NOMA. We have performed the assessments when both the i.i.d. and the correlated Rayleigh fading channels are used. The obtained numerical results showed that a higher sum-rate capacity is achieved at using lower SNR. This implies a reduction in the transmission power is possible while maintaining the sum-rate offered by the conventional DFT-s-OFDM by using the proposed BC DFT-s-OFDM waveform in the uplink transmission of massive MIMO with PD-NOMA. Consequently this will result in improving EE.

Table 3.3: Cell's sum-rate performance metric in bps/Hz for DFT-s-OFDM and BC DFT-s-OFDM in PD-NOMA at different values of SNR with various BC length, using i.i.d. and correlated Rayleigh fading channels at $M = 128$ and active UE= 32.

SNR (dB)	i.i.d.				Correlated			
	DFT-s-OFDM	BC DFT-s-OFDM			DFT-s-OFDM	BC DFT-s-OFDM		
		BC=3	BC=5	BC=7		BC=3	BC=5	BC=7
8	016.2487	157.4781	254.7083	315.3701	013.5416	059.6251	068.7488	139.3774
10	018.8583	159.2355	258.3516	318.8621	036.4546	124.2577	104.7521	199.1073
12	032.7916	167.9273	263.7391	321.8901	095.8986	204.1250	174.9763	272.3089
14	075.2248	197.2470	275.8166	332.2477	170.8211	289.3758	249.7589	356.8735
16	152.2436	258.7081	313.3732	360.1034	257.1779	373.7511	334.7796	438.6702
18	229.5612	331.1539	389.3741	422.0811	337.5914	455.1241	429.1135	522.1206

Bit Error Rate:

We have used the BER as other performance metric to inspect the performance of the proposed BC DFT-s-OFDM waveform over a range of SNR values when it is used in the uplink transmis-

sion of massive MIMO with PD-NOMA at various BC length. Then, we have compared the obtained results with the performance of the conventional DFT-s-OFDM waveform when it is used in the uplink transmission of massive MIMO with PD-NOMA. Table 3.4 shows that the BC DFT-s-OFDM waveform achieves a lower BER and outperforms the performance of the conventional DFT-s-OFDM waveform when both are used in the uplink transmission of massive MIMO with PD-NOMA. We have performed the assessments when both the i.i.d. and the correlated Rayleigh fading channels are used. A contemplation into these results, we found that lowering the BER improves the transmission performance and reduces the number of resend trials at the UE's transmitter, which consequently implies a preserving of power usage at the UE and extend the battery life. Besides, it implies an energy conservation improvement at the system that embeds the proposed BC DFT-s-OFDM waveform in the uplink transmission of massive MIMO with PD-NOMA.

Table 3.4: BER Performance metric for DFT-s-OFDM and BC DFT-s-OFDM in PD-NOMA at different values of SNR with various BC length, using i.i.d. and correlated Rayleigh fading channels at $M = 128$ and active UE= 32.

SNR (dB)	i.i.d.				Correlated			
	DFT-s-OFDM	BC DFT-s-OFDM			DFT-s-OFDM	BC DFT-s-OFDM		
		BC=3	BC=5	BC=7		BC=3	BC=5	BC=7
8	$1.51 * 10^{-1}$	$8.25 * 10^{-2}$	$2.02 * 10^{-2}$	$8.06 * 10^{-3}$	$1.61 * 10^{-1}$	$1.49 * 10^{-1}$	$1.32 * 10^{-1}$	$9.69 * 10^{-2}$
10	$1.45 * 10^{-1}$	$8.19 * 10^{-2}$	$1.99 * 10^{-2}$	$7.45 * 10^{-3}$	$1.58 * 10^{-1}$	$1.41 * 10^{-1}$	$1.24 * 10^{-1}$	$8.89 * 10^{-2}$
12	$1.41 * 10^{-1}$	$8.05 * 10^{-2}$	$1.95 * 10^{-2}$	$7.00 * 10^{-3}$	$1.43 * 10^{-1}$	$1.12 * 10^{-1}$	$1.19 * 10^{-1}$	$6.07 * 10^{-2}$
14	$1.35 * 10^{-1}$	$7.93 * 10^{-2}$	$1.67 * 10^{-2}$	$6.23 * 10^{-3}$	$1.28 * 10^{-1}$	$6.91 * 10^{-2}$	$8.19 * 10^{-2}$	$2.11 * 10^{-2}$
16	$1.29 * 10^{-1}$	$6.61 * 10^{-2}$	$1.33 * 10^{-2}$	$5.04 * 10^{-3}$	$9.08 * 10^{-2}$	$2.05 * 10^{-2}$	$3.14 * 10^{-2}$	$5.01 * 10^{-3}$
18	$1.05 * 10^{-1}$	$3.03 * 10^{-2}$	$8.89 * 10^{-3}$	$2.11 * 10^{-3}$	$4.07 * 10^{-2}$	$5.13 * 10^{-3}$	$8.17 * 10^{-3}$	$1.06 * 10^{-3}$
20	$4.52 * 10^{-2}$	$7.64 * 10^{-3}$	$3.12 * 10^{-3}$	$5.71 * 10^{-4}$	$9.74 * 10^{-3}$	$1.07 * 10^{-3}$	$2.01 * 10^{-3}$	$1.86 * 10^{-4}$

3.7.4 Peak to Average Power Ratio

The circuit power is the second dominant part of the consumed power of the UE's transmitter. A significant portion of the circuit power is consumed at the Power Amplifier (PA). The power consumption at the PA in any wireless communications system is directly related to the PAPR of the transmitted signal. The PAPR measured in (dB) for a transmitted baseband uplink signal $x(t)$ is defined as $PAPR_{dB} = 10 \log_{10} \left(\frac{\max(|x(t)|^2)}{\mathbb{E}\{|x(t)|^2\}} \right)$. High PAPR leads to high power consumption at

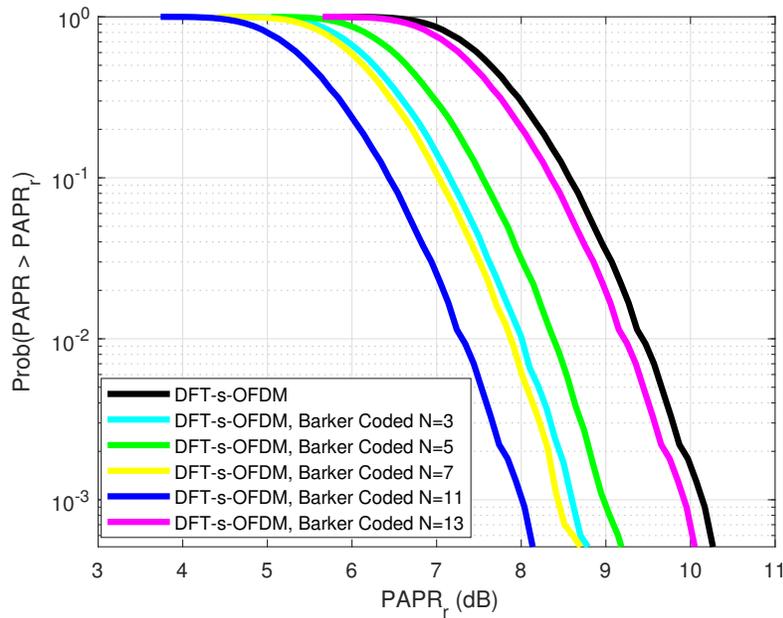


Figure 3.12: The CCDF of PAPR in a massive MIMO cell with PD-NOMA deployment using 32-QAM at different configurations.

the PA, which proportionally affects and decreases the EE of the wireless communications system. Furthermore, high PAPR pushes the PA to operate in the saturation region and causes ICI. The BCs reduce the side lobes power of the auto-correlation function and concentrate the power at the main lobe. As a result, the proposed system generates waveforms that have small PAPR, which facilitates the PA to operate in the linear region. Figure 3.12 shows the Complementary Cumulative Distribution Function (CCDF) of the PAPR in (dB) for the massive MIMO system with PD-NOMA deployment that employs the proposed DFT-s-OFDM with the adaptive length BCs. The CCDF is conventionally used for evaluating the probability of the proposed system to have higher PAPR greater than the maximum allowable reference PAPR (we call it in this chapter PAPR_r). The curves in Figure 3.12 show that the PAPRs of the BC DFT-s-OFDM waveforms is lower than the PAPR of the conventional DFT-s-OFDM waveform. This impels a UE equipped with BC spreader consumes less power compared with a UE uses the conventional DFT-s-OFDM waveform.

3.7.5 Out of Band Leakage

The leakage in the frequency band is measured by the power in dB per Hertz (Hz) that falls outside the allocated frequency bandwidth. Therefore, as this leakage increases, the EE decreases due to the dissipated power that is not exploited correctly and also caused extra interference. Figure 3.13 shows the OOB leakage performance comparison between DFT-s-OFDM, G DFT-s-OFDM, and BC DFT-s-OFDM during the usage of a bandwidth of 20 MHz. Figure 3.13 reveals that BC DFT-s-OFDM outperforms the conventional DFT-s-OFDM waveform recommended by 3GPP and G DFT-s-OFDM waveform proposed in [99], and achieves gains for suppressing the leakage by 50 dB and 30 dB, respectively.

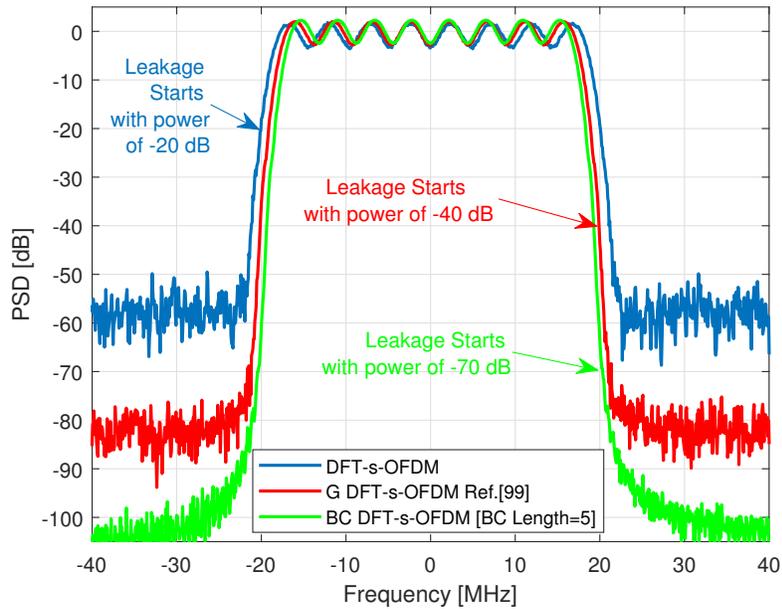


Figure 3.13: The Out of Band Leakage for DFT-s-OFDM, G DFT-s-OFDM, and BC DFT-s-OFDM waveforms.

3.8 Conclusion

In this chapter, we have proposed the BC DFT-s-OFDM waveform with an adaptive length BC spreader, which shows significantly better performance in the uplink transmission of both configurations 1) the conventional massive MIMO with OMA, and 2) the massive MIMO with PD-NOMA

deployment. We have proposed new structure for the UE's transmitter that adds and a BC spreader, and for the BS's receiver that adds a BC de-spreader. Our numerical results, of the conventional massive MIMO that employs the OMA, have demonstrated that the proposed BC DFT-s-OFDM with an adaptive length BC spreader achieves high EE, low NMSE with better CSI accuracy, low BER, and high sum-rate capacity at low SNRs. Consequently, the system reduces the required energy and consumes low transmission power while maintains to achieve the targets of performance metrics in terms of BER, NMSE, and sum-rate capacity. As a result the EE has been improved while preserving the average throughput. Thus, implementing our proposed schemes reduces the required transmission power without degrading the sum-rate capacity. This results in improving the system EE without reducing the SE, where the system EE can be improved by two approaches: 1) reduction in the consumed power at the UEs which reduces the denominator of the EE metric, and 2) increasing the cell sum-rate capacity (average data throughput) which increases the numerator of the EE metric.

We have observed that the deployment of PD-NOMA in massive MIMO improves the overall system performance and enhances both EE and SE, where the sum-rate capacity increases and the PAPR has small level. Moreover, we have demonstrated that integrating our EDPS in massive MIMO with NOMA deployment improves the CSI accuracy and adds a second degree of improvements to EE. This also enhances the downlink beam-forming and improves the EE of the whole system. Although we have used the linear MMSE estimator at the BS receiver in this chapter, the proposed schemes and structures could be generalized to other types of detectors at the BS receiver. Finally, these findings can be easily generalized to any cellular network which uses any arbitrary cluster size that consists of a group of non-cooperative adjacent cells of massive MIMO with OMA/NOMA that employ DPS/EDPS to estimate the CSI.

Chapter 4

Compute-and-Forward for NOMA and Massive MIMO-NOMA Uplinks

4.1 Synopsis

In Chapter 3 we have used SIC in NOMA configuration. SIC has been regarded as the *de facto* decoding method in NOMA system. However, SIC requires the paired users to have significantly different received power levels at the BS receiver for effective operation. This is problematic in terms of fairness and outage probabilities when the paired users are close to each other in terms of allocated power. To address this challenge, we propose in this chapter a new decoding method for uplink NOMA based on CaF. Then, we extend the proposed CaF decoding into massive MIMO-NOMA. Extending the applications of CaF into uplink MIMO and use it at the BS's receiver is also widely referred as IFLR for the MIMO uplink scenario. In particular, we show that our decoding method achieves better fairness and smaller average outage probabilities while enjoying essentially the same complexity as SIC decoding. Then, we explore the benefits and limitations of IFLR in uplink of massive MIMO-NOMA. First, we prove that IFLR can be made more efficient by using practical code construction and successive cancellation. Second, we show that the IFLR achieves better performance compared to the standard approach based on SIC. Finally, we investigate the effects of channel estimation errors on the achievable sum-rate when the proposed CaF core decod-

ing is used at IFLR in massive MIMO with NOMA deployment. Consequently, we demonstrate how DPS and EDPS are effective to reduce channel estimation errors and enhance the performance of IFLR.

4.2 Introduction

NOMA is envisioned as an enabling technology to satisfy the requirements of the 5G wireless networks [126]. For example, a simplified version of NOMA, called Multi-User Superposition Transmission (MUST), has been proposed in 3GPP LTE [127]. Compared with OMA, NOMA can significantly improve the SE and scale up the number of users [128]. The main idea of NOMA is to exploit the power/code domain to realize the multiple access. Especially, in uplink PD-NOMA, multiple users non-orthogonally transmit their messages to a single BS at the same frequency, time and code resource. The BS employs SIC to decode the signal of the strong user and then subtracts it in order to decode that of the weak user. Therefore, in a typical uplink massive MIMO-PD-NOMA scenario, the separation of user signals is achieved by using SIC at the BS.

However, SIC decoding does not work well in terms of fairness when two users' signals have comparable received power levels at the BS receiver. For this reason, most prior work assumes that there is a balance between strong users and weak users so that a strong user can always be paired with a weak user in order to ensure power difference. Contrarily, in the unbalanced scenario, the power levels of users who belong to the same NOMA pair can not be guaranteed to have significant differences. It is only recently that Pan, Lu, and Liew have studied the unbalanced scenario [129] where far users outnumber near users as shown in Figure 4.1. In particular, they proposed a network-coded multiple-access scheme that combines physical-layer network coding and multiuser decoding to handle the challenging unbalanced scenario.

In this chapter, we take a new approach to addressing the unbalanced scenario based on CaF. As a starting point, we propose a new decoding method for uplink NOMA, which has the potential to achieve better performance at essentially the same computational cost of SIC. In the first half of the chapter, for presentation simplicity, our system model starts by considering the case where

the BS is equipped with a single antenna only, assumes perfect CSI are available at the BS, and its underlying code construction is based on Loeliger's random ensemble of lattice codes [130]. Then, to generalize our system model and move forward for real world implementation, we move from Loeliger's random ensemble of lattice codes to low density Construction-A lattice codes [131].

On the one hand, in Section 4.3 and Section 4.4 of this chapter, compared to [129] our CaF-NOMA decoding algorithm focuses on improving the fairness and average outage probability with CaF decoding, whereas [129] mainly considers the performance of throughput with different decoding strategy. Beside, compared to conventional SIC-NOMA, our CaF-NOMA decoding algorithm achieves better performance at essentially the same computational cost. More specifically, our simulation results show that CaF-NOMA can increase relatively Jain's fairness index and decrease the average outage probability significantly up to 15% and 46%, respectively.

On the other hand, in the second half of this chapter, we scale up the number of BS antennas, forming a massive MIMO configuration. In MIMO wireless communications, the BS usually utilizes a linear receiver architecture to avoid implementation complexity of maximum likelihood receiver. CaF applications in MIMO systems generalize the conventional linear receivers used in MIMO systems into a class that is well known as IFLRs. IFLR recovers linearly independent combinations of the transmitted codewords by multiple UEs, each integer-forcing decoder recovering a different linear combination. Therefore, a full rank matrix of the coefficients of these linear combinations is required. However, in conventional linear receivers like MMSE, the receiver recovers the codewords transmitted by multiple UEs directly. Linear MMSE receiver, which is widely used in massive MIMO is a special case of IFLR where the integer coefficients matrix of IFLR is the identity matrix. Linear MMSE receiver performance can be significantly improved via SIC. However, this approach fails to provide an optimal diversity multiplexing trade-off compared to IFLR [132]. Therefore, we start in Section 4.5 by introducing CaF decoding as the core of IFLR. Subsequently, we highlight in Section 4.5 how IFLR can be constructed via nested lattices codes.

Then, Section 4.6 expands the proposed decoding scheme to the field of massive MIMO-NOMA systems. In Section 4.7, we demonstrate a different approach by answering the following question:

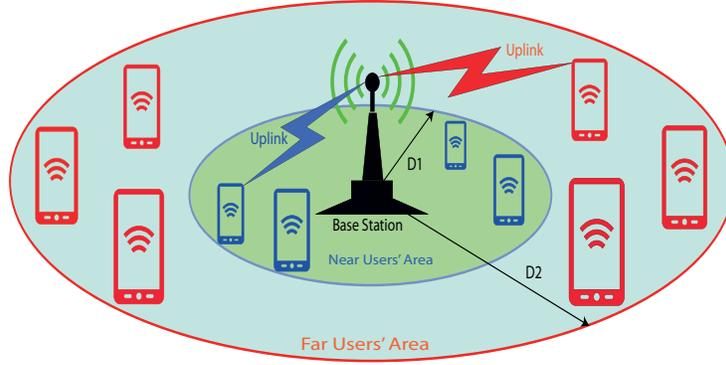


Figure 4.1: The uplink NOMA with more far users than near users.

“How can we improve the CSI of CaF decoding in IFLR without increasing the system complexity?” The answer is by exploiting the ideas and techniques developed and proposed in the previous chapters of this dissertation that include DPS and EDPS. Besides, the answer is supported by analytical derivations and numerical results presented in Section 4.8, which prove the achievement of improving CSI accuracy and consequently reducing the loss of achievable sum-rate. Finally, Section 4.9 concludes the chapter.

4.3 SIC-NOMA

We consider an uplink NOMA scenario where K users, each with a single antenna, are transmitting signals simultaneously to one BS equipped with a single antenna as shown in Figure 4.1. The number of near users (resp., far users) is denoted by K_1 (resp., K_2). The number of NOMA pairs that can be formed is equal to $\frac{K_1+K_2}{2}$. The maximum distance between a near user (resp., a far user) and the BS is denoted by $D1$ (resp., $D2$).

For the channel model, we consider both large-scale and small-scale fading. Let $\sqrt{d_k^{-\rho}}$ denotes the large-scale fading coefficient of a user k , where d_k is the distance from user k to the BS and ρ is the path loss exponent. Let g_k denotes the small-scale fading coefficient of the user k , where each g_k is assumed to be an i.i.d. complex Gaussian random variable with zero mean and unit variance. Then, the channel-gain coefficient for the user k is $h_k = g_k \sqrt{d_k^{-\rho}}$. The channel noise at the BS’s receiver is assumed to be Additive White Gaussian Noise (AWGN) which is represented

by a complex random variable with zero mean and unit variance. Under the above channel model, the received signal at the BS is given by:

$$y = \sum_{k=1}^K h_k \sqrt{p_k} x_k + w, \quad (4.1)$$

where x_k is the message transmitted by the user k , p_k is the transmission power for the user k , and w is the channel noise described above.

We assume, without loss of generality, that the users in uplink NOMA are ordered by their distances $d_1 \leq d_2 \leq \dots \leq d_K$. We consider a particular pairing method as follows: user 1 is paired with user K , user 2 is paired with user $K - 1$, and so on¹. Additionally, different pairs of NOMA work in OMA mode, so that it suffices to focus on only one pair.

We now consider a pair consisting of user k and user l where $d_k < d_l$ (i.e., user k is a strong user and user l is a weak user). Under SIC decoding, the BS first decodes x_k , then subtracts it from y in order to decode x_l . It is well known that under SIC decoding [133], the following data rates R_k and R_l in Bits Per Channel Use (BPCU) are achievable:

$$\begin{cases} R_k = \log_2 \left(1 + \frac{p_k |h_k|^2}{p_l |h_l|^2 + 1} \right) \\ R_l = \log_2 (1 + p_l |h_l|^2) \end{cases} \quad (4.2)$$

4.4 CaF-NOMA

In this section, we introduce CaF decoding and explain why it outperforms SIC decoding for uplink NOMA.

To introduce CaF decoding, we need to describe the set of Gaussian integers. A complex number $a + bi$ is called a Gaussian integer if a and b are both integers. Formally, Gaussian integers are the set $\mathbb{Z}[i] \triangleq \{a + bi : a, b \in \mathbb{Z}\}$ ².

¹Note that our main conclusions apply to any pairing method. We consider a particular method in this chapter for ease of presentation.

²In addition to Gaussian integers, we can use Eisenstein integers or more generally a principal ideal domain. See, e.g., [60] for details.

We consider again a pair consisting of user k and user l . CaF decoding is similar to SIC decoding in that it first decodes a Gaussian-integer combination $a_1x_k + a_2x_l$ (where the coefficients $a_1, a_2 \in \mathbb{Z}[i]$ are chosen by the decoder based on CSI as explained later in Algorithm 1) and then “subtracts” it from y to decode a second Gaussian-integer combination $b_1x_k + b_2x_l$ (which is linearly independent of the first combination). Similarly, the coefficients $b_1, b_2 \in \mathbb{Z}[i]$ are chosen by the decoder based on CSI as explained later in Algorithm 1. With two linearly independent combinations, the CaF decoder is able to recover both signals x_k and x_l . (Note that SIC decoding can be viewed as a special case of CaF decoding with $\mathbf{a} \triangleq [a_1, a_2] = [1, 0]$ and $\mathbf{b} \triangleq [b_1, b_2] = [0, 1]$.) Indeed, CaF decoding can be extended to the case of more than two users, where the BS decodes linearly independent combinations by applying a lattice reduction algorithm [60, 130].

Figure 4.2 illustrates the advantage of CaF decoding over SIC decoding with respect to the Gaussian MAC capacity region, which is defined as the set of rate pairs (R_k, R_l) such that

$$R_k \leq \log_2(1 + |h_k|^2 p_k)$$

$$R_l \leq \log_2(1 + |h_l|^2 p_l)$$

$$R_k + R_l \leq \log_2(1 + |h_k|^2 p_k + |h_l|^2 p_l)$$

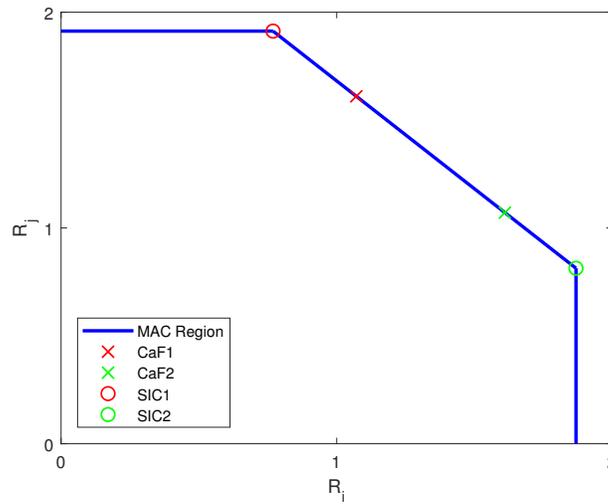


Figure 4.2: CaF decoding offers two additional choices, namely, CaF_1 and CaF_2 .

Operationally, the MAC capacity region represents the set of rate pairs achieved by all possible encoding/decoding methods [133]. Figure 4.2 shows that CaF decoding achieves not only the two corner points of the MAC capacity region as SIC does (i.e., SIC_1 and SIC_2) but also two additional “boundary” points (i.e., CaF_1 and CaF_2). These additional points make CaF decoding a better choice than SIC decoding in terms of fairness and outage probabilities, as shown in Section 4.8. Interestingly, CaF decoding has essentially the same complexity as SIC decoding as we will explain in Remark 4.

Finally, a question naturally arises: How to choose the “best possible” Gaussian-integer coefficients (\mathbf{a}, \mathbf{b}) ? To answer this question, we consider two rate pairs (R_k, R_l) and (R'_k, R'_l) induced by (\mathbf{a}, \mathbf{b}) and $(\mathbf{a}', \mathbf{b}')$, respectively. We say (R_k, R_l) is better than (R'_k, R'_l) if $R_k > R'_k$ and $R_l > R'_l$. We say (R_k, R_l) is a *Pareto-optimal solution* if there exists no (R'_k, R'_l) such that (R'_k, R'_l) is better than (R_k, R_l) . See Figure 4.3 for an illustration.

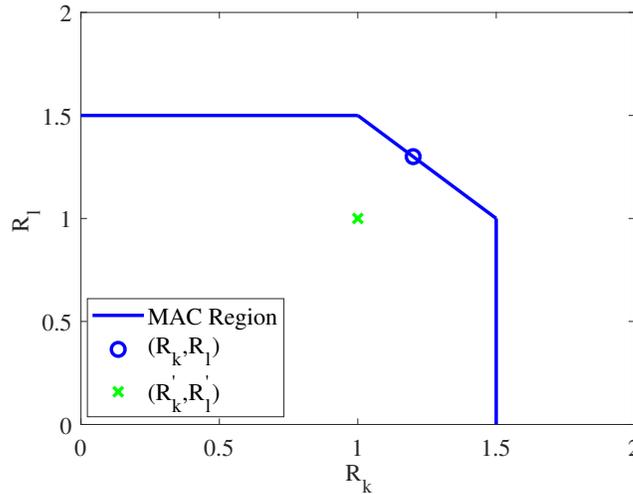


Figure 4.3: (R_k, R_l) is better than (R'_k, R'_l) , and (R_k, R_l) is Pareto-optimal. The MAC region is the capacity region for multiple access channels.

As our first technical contribution of this chapter, Theorem 4 characterizes all Pareto-optimal solutions under CaF decoding. The proof of Theorem 4 is provided in Appendix B.1.

Theorem 4 Any Pareto-optimal solution for CaF decoding is induced by $([1, 0], [0, 1])$, or $([0, 1], [1, 0])$, or $(\mathbf{a}_*, \mathbf{b}_*)$, where $(\mathbf{a}_*, \mathbf{b}_*)$ is the output of Algorithm 1.

Algorithm 1 Compute \mathbf{a}_* and \mathbf{b}_*

Input: p_k, p_l, h_k, h_l .

Output: \mathbf{a}_* and \mathbf{b}_*

1. Let $\mathbf{h} = [h_k, h_l]$, $\mathbf{P} = \begin{bmatrix} p_k & 0 \\ 0 & p_l \end{bmatrix}$ and $\mathbf{M} = (\mathbf{P}^{-1} + \mathbf{h}^H \mathbf{h})^{-1}$.
 2. SVD decomposition for $\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^H$ and $\mathbf{G}_{\text{old}} \triangleq \sqrt{\mathbf{S}}\mathbf{V}^H$.
 3. Let \mathbf{b}_1 and \mathbf{b}_2 be the first and second row of \mathbf{G}_{old} , respectively.
 4. **if** $\|\mathbf{b}_1\|_2 > \|\mathbf{b}_2\|_2$ **then**
 5. Swap \mathbf{b}_1 and \mathbf{b}_2 .
 6. **end if**
 7. OK=1.
 8. **while** OK > 0 **do**
 9. **if** $|\Re\{\mathbf{b}_1^H \mathbf{b}_2\}| \leq \frac{1}{2}\|\mathbf{b}_1\|_2^2$ and $|\Im\{\mathbf{b}_1^H \mathbf{b}_2\}| \leq \frac{1}{2}\|\mathbf{b}_1\|_2^2$ **then**
 10. OK=0.
 11. **else**
 12. $\mathbf{b}_2 = \mathbf{b}_2 - \text{round}\left(\frac{\mathbf{b}_1^H \mathbf{b}_2}{\|\mathbf{b}_1\|_2^2}\right) \mathbf{b}_1$.
 13. **if** $\|\mathbf{b}_2\|_2 > \|\mathbf{b}_1\|_2$ **then**
 14. OK=0.
 15. **else**
 16. Swap \mathbf{b}_1 and \mathbf{b}_2 .
 17. **end if**
 18. **end if**
 19. **end while**
 20. $\mathbf{G}_{\text{new}} = [\mathbf{b}_{1n}, \mathbf{b}_{2n}]$.
 21. Output $\mathbf{a}_* \triangleq \text{round}(\mathbf{G}_{\text{old}}^{-1} \mathbf{b}_{1n})$ and $\mathbf{b}_* \triangleq \text{round}(\mathbf{G}_{\text{old}}^{-1} \mathbf{b}_{2n})$.
-

Remark 1 CSI and users' power matrix \mathbf{P} are the two inputs of Algorithm 1. First, Algorithm 1 computes the matrix \mathbf{M} where $\mathbf{M} = (\mathbf{P}^{-1} + \mathbf{h}^H \mathbf{h})^{-1}$. In the second step, Algorithm 1 executes SVD algorithm on \mathbf{M} . The SVD is done only once in Algorithm 1 (i.e., not every iteration) over a low dimensional square matrix, where the matrix's dimension is determined by the number of UEs in the NOMA pair. Then, Algorithm 1 performs a lattice reduction algorithm used in [134], which is usually known as Gauss' algorithm and is very similar to Euclid's algorithm for computing the greatest common divisor [135]. The outputs of Algorithm 1 are two vectors \mathbf{a}_* and \mathbf{b}_* that have some important properties as stated in Appendix B.1. These properties allow us to complete the proof of Theorem 4.

Remark 2 Theorem 4 says that the BS can find out all Pareto-optimal solutions under CaF de-

coding by evaluating three coefficient vectors $(\mathbf{a}_*, \mathbf{b}_*)$, $([1, 0], [0, 1])$, and $([0, 1], [1, 0])$. Once these solutions are found, the BS can choose the “best possible” solution based on a particular criterion such as Jain’s fairness index [136]. In other words, CaF decoding improves SIC decoding by offering the BS more choices.

Remark 3 Theorem 4 is based on successive computation proposed in [130], the most general form of CaF decoding. Theorem 4 can be applied to scenarios other than uplink NOMA where SIC decoding is replaced by CaF decoding (based on successive computation). Thus, Theorem 4 may be of general interest.

Remark 4 Theorem 4 implies that CaF decoding has essentially the same complexity as SIC decoding. To see this, note that CaF decoding needs to find the values of $(\mathbf{a}_*, \mathbf{b}_*)$ before decoding two Gaussian-integer combinations of signals x_k and x_l . It turns out that decoding one Gaussian-integer combination of signals is as hard as decoding one signal, as explained in [130]. Therefore, the only additional computational cost is computing $(\mathbf{a}_*, \mathbf{b}_*)$, which can be done very efficiently via Algorithm 1 that involves a fast computational lattice reduction algorithm as explained in [134].

4.5 CaF in Integer-Forcing Linear Receiver

IFLR is a type of receiver that is used often in multiple-antenna systems to reduce the implementation complexity. CaF decoding is used as the core of IFLR [132]. In this section, we explain how to construct an IFLR based on low density Construction-A lattices. As shown below, a key ingredient is a linear labeling that enables efficient encoding and decoding operations. Appendix B.2 provides the necessary preliminaries for the basics of lattices and nested lattice codes.

4.5.1 Low Density Construction-A Lattices

Following the steps in [131, 137], we can construct a nested lattice code $\mathcal{L}(\Lambda_c, \Lambda_f)$ from nested Low Density Parity Check (LDPC) codes $C_2 \subset C_1 \subset \mathbb{Z}_p^n$, where C_i is specified by a parity-check matrix H_i of size $m_i \times n$, and p is a prime number. In particular, H_1 consists of the first m_1 rows of H_2 . The number of codewords in $\mathcal{L}(\Lambda_c, \Lambda_f)$ is $p^{m_2 - m_1}$. As shown in [137], if H_2 is chosen so that

the Tanner graphs corresponding to both C_1 and C_2 have the required expansion properties, then both Λ_f and Λ_c have the goodness properties with high probability.

4.5.2 Construction of a Linear Labeling

We would like to map messages over \mathbb{Z}_p to codewords in the nested lattice code $\mathcal{L}(\Lambda_c, \Lambda_f)$ in order to construct the IFLR. To this end, we make use of the linear labeling framework proposed in [60].

Definition 1 *Given a positive scaling factor γ which is applied to obtain a fine lattice, a mapping $\varphi : \gamma\mathbb{Z}^n \rightarrow \mathbb{Z}_p^n$ is called a linear labeling if it satisfies the following three properties.*

1. *A lattice point λ belongs to Λ_f if and only if the first m_1 components of its label $\varphi(\lambda)$ are equal to 0.*
2. *A lattice point λ belongs to Λ_c if and only if the first m_2 components of its label $\varphi(\lambda)$ are equal to 0.*
3. *For all $a_1, a_2 \in \mathbb{Z}$ and $\lambda_1, \lambda_2 \in \Lambda_f$, we have*

$$\varphi(a_1\lambda_1 + a_2\lambda_2) = q_1\varphi(\lambda_1) + q_2\varphi(\lambda_2), \quad (4.3)$$

where $q_i = a_i \bmod p$.

We are ready to construct a linear labeling. Recall that H_1 consists of the first m_1 rows of H_2 . By the Basis Extension Theorem [138], there exists some matrix H' of size $(n - m_2) \times n$ so that

$$H = \begin{bmatrix} H_2 \\ H' \end{bmatrix}$$

is an invertible matrix of size $n \times n$. Then, we define the mapping $\phi : \gamma\mathbb{Z} \rightarrow \mathbb{Z}_p$ as

$$\phi(v) \triangleq \gamma^{-1}v \bmod p$$

along with an inverse operation $\phi^{Inv} : \mathbb{Z}_p \rightarrow \gamma\mathbb{Z}$ as

$$\phi^{Inv}(u) \triangleq \gamma u.$$

These mappings can be naturally extended to vectors via element-wise operations.

Theorem 5 Let $\phi : \gamma\mathbb{Z}^n \rightarrow \mathbb{Z}_p^n$ be a map given by

$$\phi(\boldsymbol{\lambda}) \triangleq \phi(\boldsymbol{\lambda})\mathbf{H}^T.$$

Then, ϕ is a linear labeling.

See Appendix B.3 for the proof of Theorem 5. Next, we construct an inverse map $\phi^{Inv} : \mathbb{Z}_p^n \rightarrow \gamma\mathbb{Z}^n$ for the linear labeling ϕ from Theorem 5. Specifically, define

$$\phi^{Inv}(\mathbf{u}) = \phi^{Inv}(\mathbf{u}(\mathbf{H}^T)^{-1}).$$

Clearly, we have $\phi(\phi^{Inv}(\mathbf{u})) = \mathbf{u}$ and the first m_1 components of \mathbf{u} are equal to 0 if and only if $\phi^{Inv}(\mathbf{u})$ belongs to Λ_f .

With the linear labeling ϕ and its inverse ϕ^{Inv} , we can map messages over $\mathbb{Z}_p^{m_2-m_1}$ to codewords in the nested lattice code $\mathcal{L}(\Lambda_c, \Lambda_f)$ as follows. First, each message \mathbf{u} in $\mathbb{Z}_p^{m_2-m_1}$ is embedded into \mathbb{Z}_p^n using m_1 leading zeros and $n - m_2$ trailing zeros

$$(\mathbf{0}_{m_1}, \mathbf{u}, \mathbf{0}_{n-m_2}).$$

Then, it is mapped to a lattice point $\boldsymbol{\lambda}$ in Λ_f via

$$\boldsymbol{\lambda} \triangleq \phi^{Inv}(\mathbf{0}_{m_1}, \mathbf{u}, \mathbf{0}_{n-m_2}).$$

Finally, we apply the mod operation

$$\boldsymbol{\lambda} \bmod \Lambda_c \triangleq \boldsymbol{\lambda} - Q_{\Lambda_c}(\boldsymbol{\lambda}_f)$$

to ensure that $\boldsymbol{\lambda}_f \triangleq \boldsymbol{\lambda} \bmod \Lambda_c$ is indeed a codeword in the nested lattice code $\mathcal{L}(\Lambda_c, \Lambda_f)$.

4.5.3 Integer Forcing Linear Receiver via Nested Lattices

We now have all the ingredients (i.e., the nested low density Construction-A lattice code $\mathcal{L}(\Lambda_c, \Lambda_f)$ and the linear labeling φ) to construct the IFLR. It suffices to consider the case of two users, since the extension to the case of more than two users is straightforward. We use the parameters $m_1^{(\ell)}$ and $m_2^{(\ell)}$ for user $\ell \in \{1, 2\}$. Let $m_2 \triangleq \max m_2^{(\ell)}$. Construct a parity-check matrix \mathbf{H} of size $m_2 \times n$ following the steps of [131, 137]. Let $\mathbf{H}_i^{(\ell)}$ denote the matrix consisting of the first $m_i^{(\ell)}$ rows of \mathbf{H} . Let $\Lambda_f^{(\ell)}$ and $\Lambda_c^{(\ell)}$ be the resulting Construction-A lattices. If \mathbf{H} is chosen so that the Tanner graphs associated with $\{\mathbf{H}_i^{(\ell)}\}$ have the required expansion properties, then both $\Lambda_f^{(\ell)}$ and $\Lambda_c^{(\ell)}$ have the goodness properties with high probability.

With the nested lattice code $\mathcal{L}(\Lambda_c^{(\ell)}, \Lambda_f^{(\ell)})$, we can proceed to the encoding and decoding operations.

Encoding: User ℓ maps its message \mathbf{u}_ℓ to a codeword $\boldsymbol{\lambda}_f^{(\ell)}$ in the nested lattice code $\mathcal{L}(\Lambda_c^{(\ell)}, \Lambda_f^{(\ell)})$

$$\boldsymbol{\lambda}_f^{(\ell)} = \varphi^{Imv} \left(\mathbf{0}_{m_1^{(\ell)}}, \mathbf{u}_\ell, \mathbf{0}_{n-m_2^{(\ell)}} \right) \bmod \Lambda_c^{(\ell)} \quad (4.4)$$

and then produces its transmitted signal \mathbf{x}_ℓ

$$\mathbf{x}_\ell = \left(\boldsymbol{\lambda}_f^{(\ell)} + \mathbf{d}_\ell \right) \bmod \Lambda_c^{(\ell)}. \quad (4.5)$$

Here, \mathbf{d}_ℓ is a random dither for user ℓ , which is generated uniformly over the Voronoi region of the coarse lattice $\mathcal{V}(\Lambda_c^{(\ell)})$. The random dither \mathbf{d}_ℓ is introduced in [130] only to simplify the achievability proof, which can be replaced by a fixed dither (see Appendix H of [130] for a proof)

or even removed in practice under some mild condition (see [131] for a proof). We keep the random dither \mathbf{d}_ℓ in this chapter so that we can reuse the proof steps in [130]. It should be removed in practical implementation.

Channel model: Consider an uplink scenario with two single-antenna users and one BS with M antennas. The BS receives

$$\mathbf{Y} = \mathbf{h}_1 \mathbf{x}_1 + \mathbf{h}_2 \mathbf{x}_2 + \mathbf{W}, \quad (4.6)$$

where $\mathbf{x}_\ell \in \mathbb{R}^{1 \times n}$ is the transmitted signal for user ℓ , $\mathbf{h}_\ell \in \mathbb{R}^{M \times 1}$ is the channel-coefficients vector for user ℓ , and $\mathbf{W} \in \mathbb{R}^{M \times n}$ is the channel noise. Note that $\{\mathbf{h}_\ell\}$ is a column vector whereas the original message $\{\mathbf{u}_\ell\}$ and its codeword $\boldsymbol{\lambda}_f^{(\ell)}$, which associated with the transmitted signal $\{\mathbf{x}_\ell\}$ are row vectors³. We assume that the channel vectors ($\{\mathbf{h}_\ell\}, \forall \ell$) are known to the BS but unknown to the users.

Decoding: From a high-level view, the BS first computes one integer-linear combination $a_1 \boldsymbol{\lambda}_f^{(1)} + a_2 \boldsymbol{\lambda}_f^{(2)}$ and then subtracts it from the received signal \mathbf{Y} in order to compute another integer-linear combination $b_1 \boldsymbol{\lambda}_f^{(1)} + b_2 \boldsymbol{\lambda}_f^{(2)}$. The BS can recover the transmitted messages \mathbf{u}_1 and \mathbf{u}_2 as long as these two combinations are linearly independent. Clearly, this decoding strategy generalizes the standard SIC decoding, where $\mathbf{a} \triangleq (a_1, a_2)^T = (1, 0)^T$ and $\mathbf{b} \triangleq (b_1, b_2)^T = (0, 1)^T$, or $\mathbf{a} = (0, 1)^T$ and $\mathbf{b} = (1, 0)^T$.

The detailed decoding steps can be found in [130]. Here, we only present the achievable rates of this decoding strategy. We need to introduce some additional notation. Let \mathbf{P} be the diagonal matrix of the power constraints

$$\mathbf{P} \triangleq \text{diag}(P_1, P_2)$$

where P_ℓ is the average power constraint for user ℓ . Let \mathbf{H} be the channel-coefficient matrix

$$\mathbf{H} \triangleq \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 \end{bmatrix}.$$

³This is because messages and codewords are often row vectors in coding theory and channel vectors are usually column vectors in information theory.

Let

$$\sigma^2(\mathbf{H}, \mathbf{a}) \triangleq \mathbf{a}^T (\mathbf{P}^{-1} + \mathbf{H}^T \mathbf{H})^{-1} \mathbf{a}$$

and

$$\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) \triangleq \frac{P_1 P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \cdot \frac{1}{|\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T|}.$$

Operationally, $\sigma^2(\mathbf{H}, \mathbf{a})$ is the effective noise variance associated with the first integer-linear combination $a_1 \boldsymbol{\lambda}_f^{(1)} + a_2 \boldsymbol{\lambda}_f^{(2)}$ and $\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})$ is the effective noise variance associated with the second integer-linear combination $b_1 \boldsymbol{\lambda}_f^{(1)} + b_2 \boldsymbol{\lambda}_f^{(2)}$ (by using the first combination as side information), as explained in [130].

Proposition 1 *The following rate pairs are achievable by the above encoding and decoding operations.*

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p = 0$,

$$R_1 = \frac{1}{2} \log_2 \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a})} \right) = \frac{1}{2} \log_2 (1 + P_1 \mathbf{h}_1^T (\mathbf{I}_M + P_2 \mathbf{h}_2 \mathbf{h}_2^T)^{-1} \mathbf{h}_1)$$

$$R_2 = \frac{1}{2} \log_2 \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) = \frac{1}{2} \log_2 (1 + P_2 \|\mathbf{h}_2\|_2^2)$$

- When $a_1 \bmod p = 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \frac{1}{2} \log_2 \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) = \frac{1}{2} \log_2 (1 + P_1 \|\mathbf{h}_1\|_2^2)$$

$$R_2 = \frac{1}{2} \log_2 \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \right) = \frac{1}{2} \log_2 (1 + P_2 \mathbf{h}_2^T (\mathbf{I}_M + P_1 \mathbf{h}_1 \mathbf{h}_1^T)^{-1} \mathbf{h}_2)$$

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \min \left\{ \frac{1}{2} \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a})} \right), \frac{1}{2} \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) \right\}$$

$$R_2 = \frac{1}{2} \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \right)$$

or

$$R_1 = \frac{1}{2} \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a})} \right),$$

$$R_2 = \min \left\{ \frac{1}{2} \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \right), \frac{1}{2} \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) \right\}$$

where $\log_2^+(\cdot) = \max\{0, \log_2(\cdot)\}$. In particular, when $\sigma^2(\mathbf{H}, \mathbf{a}) \leq \sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})$,

$$R_1 + R_2 = \frac{1}{2} \log_2 |\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T|. \quad (4.7)$$

The proof of Proposition 1 is essentially the same as the proof of Theorem 5 in [130]. The key observation is that Theorem 5 in [130] holds for *any* nested lattices with goodness properties and *any* linear labeling satisfying Definition 1. Compared to the statement of Theorem 5, we have managed to avoid the concept of complicated admissible mappings by *explicitly* evaluating all the possibilities of successive cancellation for the two-user case. These explicit expressions in Proposition 1 may be of independent interest for those who would like to apply the IFLR with low density Construction-A lattices for the two-user case.

4.5.4 Extension to the Complex Field

As stated in [130], all of its results can be extended from real-valued models of channel coefficients to complex-valued models of channel coefficients either via a real-valued decomposition of the channel or by building nested lattices directly over the complex field using the algebraic framework in [60]. This extension leads to the following achievable rates for the channel-coefficient matrix \mathbf{H} over the complex field.

Proposition 2 *The following rate pairs are achievable:*

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p = 0$,

$$R_1 = \log_2 \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a})} \right) = \log_2 \left(1 + P_1 \mathbf{h}_1^H (\mathbf{I}_M + P_2 \mathbf{h}_2 \mathbf{h}_2^H)^{-1} \mathbf{h}_1 \right)$$

$$R_2 = \log_2 \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) = \log_2 \left(1 + P_2 \|\mathbf{h}_2\|_2^2 \right)$$

- When $a_1 \bmod p = 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \log_2 \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) = \log_2 \left(1 + P_1 \|\mathbf{h}_1\|_2^2 \right)$$

$$R_2 = \log_2 \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \right) = \log_2 \left(1 + P_2 \mathbf{h}_2^H (\mathbf{I}_M + P_1 \mathbf{h}_1 \mathbf{h}_1^H)^{-1} \mathbf{h}_2 \right)$$

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \min \left\{ \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a})} \right), \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) \right\}$$

$$R_2 = \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \right)$$

or

$$R_1 = \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a})} \right),$$

$$R_2 = \min \left\{ \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \right), \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})} \right) \right\}.$$

In particular, when $\sigma^2(\mathbf{H}, \mathbf{a}) \leq \sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})$, we have

$$R_1 + R_2 = \log_2 |\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H|. \quad (4.8)$$

Here, \mathbf{a} and \mathbf{b} are over Gaussian integers $\mathbb{Z}[i] \triangleq \{a + bi : a, b \in \mathbb{Z}\}$ and p as mentioned should be a Gaussian prime (i.e., $p \in \mathbb{Z}$ is a prime number that can be written in a form of $4m + 3$, where m is an integer number), as explained in [60]. The new expressions of $\sigma^2(\mathbf{H}, \mathbf{a})$ and $\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})$ are

given below:

$$\sigma^2(\mathbf{H}, \mathbf{a}) \triangleq \mathbf{a}^H (\mathbf{P}^{-1} + \mathbf{H}^H \mathbf{H})^{-1} \mathbf{a}$$

and

$$\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) \triangleq \frac{P_1 P_2}{\sigma^2(\mathbf{H}, \mathbf{a})} \cdot \frac{1}{|\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H|}.$$

Again, the explicit expressions in Proposition 2 may be of independent interest.

4.6 CaF for Massive MIMO-NOMA

4.6.1 Integer Forcing Linear Receiver for Massive MIMO-NOMA

In this subsection, we characterize the best possible achievable rates of IFLR for uplink massive MIMO-NOMA. We consider an uplink transmission scenario, as shown in Figure 4.4, for massive MIMO-NOMA in which up to $\frac{3K}{2}$ NOMA pairs⁴ of UEs, each UE equipped with a single antenna, are transmitting signals to a BS equipped with M antennas, where as mentioned previously in massive MIMO $M \gg K$. We assume, without loss of generality, that the users are ordered by their distances to the BS so that $d_1 \leq d_2 \leq \dots \leq d_{3K}$. We allow *any* pairing method in our system model as long as different pairs work in OMA mode⁵ so that a pair does not interfere another pair. Hence, it suffices for us to focus on a particular pair with a near user, say user k , and a far user, say user l (where $d_k < d_l$). Under this model, the received signal $\mathbf{Y} \in \mathbb{C}^{M \times n}$ observed by the BS across the M antennas over n channel uses can be written as

$$\mathbf{Y} = \mathbf{h}_k \mathbf{x}_k + \mathbf{h}_l \mathbf{x}_l + \mathbf{W}, \quad (4.9)$$

where $\mathbf{x}_k \in \mathbb{C}^{1 \times n}$ and $\mathbf{x}_l \in \mathbb{C}^{1 \times n}$ are the transmitted signals of length n at k th UE and l th UE respectively, $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_l \in \mathbb{C}^{M \times 1}$ are the channel coefficients vectors for user k and user l respectively, and $\mathbf{W} \in \mathbb{C}^{M \times n}$ is the channel noise. Following to the same channel model presented

⁴A NOMA pair can involve one user from the near region and up to two users from the far region.

⁵It is important to mention that user pairing methods are a key research approach in energy efficient NOMA cellular networks in 5G and beyond, and our concluding outcomes of this section can apply to any of them. However, for consistency we follow the same particular pairing method we have used in Section 4.3.

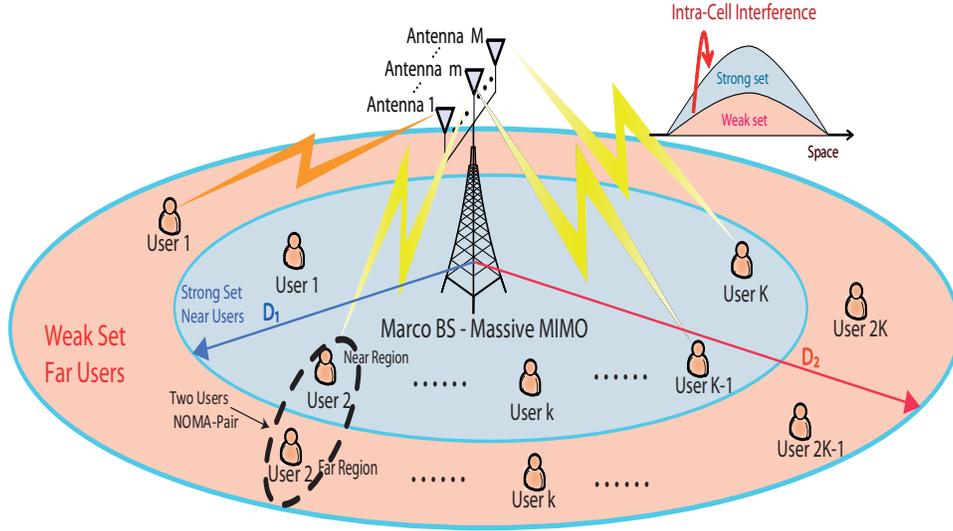


Figure 4.4: The uplink of Massive MIMO-NOMA with more far users than near users.

in Section 4.3, then the $M \times 1$ channel-coefficients vector for the user k is $\mathbf{h}_k = \mathbf{g}_k \sqrt{d_k^{-\rho}}$. As before, $\mathbf{P} \triangleq \text{diag}(P_k, P_l)$ where P_k and P_l are the average power constraints for user k and user l , respectively, and $\mathbf{H} \triangleq \begin{bmatrix} \mathbf{h}_k, \mathbf{h}_l \end{bmatrix}$.

Both UEs apply the encoding strategy described in Section 4.5.3 without random dithers. That is, $\mathbf{x}_k = \lambda_f^{(k)}$ and $\mathbf{x}_l = \lambda_f^{(l)}$. The BS applies the decoding strategy described in Section 4.5.3. As stated before, the BS as shown in Figure 4.5 first computes one Gaussian-integer combination (say, $a_k \mathbf{x}_k + a_l \mathbf{x}_l$ where $a_k, a_l \in \mathbb{Z}[i]$) and then subtracts it from the received signal \mathbf{Y} in order to compute another Gaussian-integer combination (say, $b_k \mathbf{x}_k + b_l \mathbf{x}_l$) which is linearly independent of the first combination. With two linearly independent combinations, the BS can recover the transmitted signals \mathbf{x}_k and \mathbf{x}_l . The achievable rates of this decoding strategy are given in Proposition 2, which clearly includes the standard SIC decoding as a special case. In other words, CaF decoding offers the same MAC capacity region as SIC decoding, however, CaF decoding enlarges the achievable rates of user k and user l with SIC decoding⁶, which increases the degree of freedom at the BS's

⁶This decoding strategy can be extended to the case of more than two users, where the BS decodes linearly independent combinations by applying a lattice reduction algorithm [60, 130].

IFLR receiver. This leads to better fairness and smaller outage probabilities for uplink massive MIMO-NOMA as we will show in our simulation results.

Next, a question naturally arises: *How to choose the “best possible” Gaussian-integer coefficients $\mathbf{a} = (a_k, a_l)^T$ and $\mathbf{b} = (b_k, b_l)^T$?* To answer this question, we consider two rate pairs (R_k, R_l) and (R'_k, R'_l) associated with (\mathbf{a}, \mathbf{b}) and $(\mathbf{a}', \mathbf{b}')$, respectively. We say (R_k, R_l) is *better than* (R'_k, R'_l) if $R_k > R'_k$ and $R_l > R'_l$. We say (R_k, R_l) is a *Pareto-optimal solution* if there exists no (R'_k, R'_l) such that (R'_k, R'_l) is better than (R_k, R_l) . It turns out that there are at most three Pareto-optimal solutions.

Theorem 6 *Let \mathbf{a}_* be an optimal solution to*

$$\min_{\mathbf{a} \neq \mathbf{0}} \mathbf{a}^H (\mathbf{P}^{-1} + \mathbf{H}^H \mathbf{H})^{-1} \mathbf{a}. \quad (4.10)$$

Let \mathbf{b}_ be an optimal solution to (4.10) under the constraint that \mathbf{a}_* and \mathbf{b}_* are linearly independent. Then, any Pareto-optimal solution is associated with $[(1, 0)^T, (0, 1)^T]$, or $[(0, 1)^T, (1, 0)^T]$, or $[\mathbf{a}_*, \mathbf{b}_*]$.*

See Appendix B.4 for the proof of Theorem 6. Theorem 6 states that the BS only needs to evaluate three rate pairs (associated with $[(1, 0)^T, (0, 1)^T]$, $[(0, 1)^T, (1, 0)^T]$, and $[\mathbf{a}_*, \mathbf{b}_*]$) and then pick the best one based on a particular criterion such as Jain’s fairness index and targeted outage probability. If the rate pair associated with $[\mathbf{a}_*, \mathbf{b}_*]$ is better than others, then the new decoding strategy outperforms the standard SIC decoding.

This brings us to a final question: How to compute $[\mathbf{a}_*, \mathbf{b}_*]$ efficiently when the BS is quipped with a massive number of antennas? It turns out that we can exploit the lattice reduction algorithm for MIMO system used in [134] to update Algorithm 1 to compute $[\mathbf{a}_*, \mathbf{b}_*]$. As shown in [134], the algorithm only requires one or two iterations in most cases, where each iteration involves a few simple operations such as rounding. Channel matrix and users’ power matrix \mathbf{P} are the two inputs of Algorithm 2. Algorithm 2 demonstrates how $[\mathbf{a}_*, \mathbf{b}_*]$ are computed, which is nothing other than a generalization of Algorithm 1 from a single antenna system to MIMO systems. Therefore, Algorithm 2 follows the same logical steps as Algorithm 1 yet with a higher dimension that is

determined by the number of BS's antennas.

Algorithm 2 Compute \mathbf{a}_* and \mathbf{b}_*

Input: $P_k, P_l, \mathbf{h}_k, \mathbf{h}_l$.

Output: \mathbf{a}_* and \mathbf{b}_*

1. Let $\mathbf{H} = [\mathbf{h}_k, \mathbf{h}_l]$, $\mathbf{P} = \text{diag}(P_k, P_l)$ and $\mathbf{M} = (\mathbf{P}^{-1} + \mathbf{H}^H \mathbf{H})^{-1}$.
 2. SVD decomposition for $\mathbf{M} = \mathbf{U} \mathbf{S} \mathbf{V}^H$ and $\mathbf{G}_{\text{old}} \triangleq \sqrt{\mathbf{S}} \mathbf{V}^H$.
 3. Let \mathbf{b}_1 and \mathbf{b}_2 be the first and second column of \mathbf{G}_{old} , respectively.
 4. **if** $\|\mathbf{b}_1\|_2 > \|\mathbf{b}_2\|_2$ **then**
 5. Swap \mathbf{b}_1 and \mathbf{b}_2 .
 6. **end if**
 7. OK=1.
 8. **while** OK > 0 **do**
 9. **if** $|\Re\{\mathbf{b}_1^H \mathbf{b}_2\}| \leq \frac{1}{2} \|\mathbf{b}_1\|_2^2$ and $|\Im\{\mathbf{b}_1^H \mathbf{b}_2\}| \leq \frac{1}{2} \|\mathbf{b}_1\|_2^2$ **then**
 10. OK=0.
 11. **else**
 12. $\mathbf{b}_2 = \mathbf{b}_2 - \text{round}\left(\frac{\mathbf{b}_1^H \mathbf{b}_2}{\|\mathbf{b}_1\|_2^2}\right) \mathbf{b}_1$.
 13. **if** $\|\mathbf{b}_2\|_2 > \|\mathbf{b}_1\|_2$ **then**
 14. OK=0.
 15. **else**
 16. Swap \mathbf{b}_1 and \mathbf{b}_2 .
 17. **end if**
 18. **end if**
 19. **end while**
 20. $\mathbf{G}_{\text{new}} = [\mathbf{b}_{1n}, \mathbf{b}_{2n}]$.
 21. Output $\mathbf{a}_* \triangleq \text{round}(\mathbf{G}_{\text{old}}^{-1} \mathbf{b}_{1n})$ and $\mathbf{b}_* \triangleq \text{round}(\mathbf{G}_{\text{old}}^{-1} \mathbf{b}_{2n})$.
-

4.7 The Role of Imperfect CSI in Massive MIMO-NOMA with CaF Decoding

In the previous sections, we have assumed perfect CSI (i.e., the BS knows the channel-coefficients matrix \mathbf{H}). In reality, the BS can only estimate \mathbf{H} , which leads to some estimation errors. Estimation errors of channel coefficients lead to sub-optimal choices of CaF multiplying (i.e., equalization) coefficients and consequently inefficient decoding, which result in a reduction in achievable rates and hence the sum-rate capacity of a cell. On the one hand, as demonstrated in previous sections, efficient decoding at BS to retrieve back transmitted signals depends on the optimal computation of CaF multiplying coefficients, which also depends on the accuracy of channel coefficients as il-

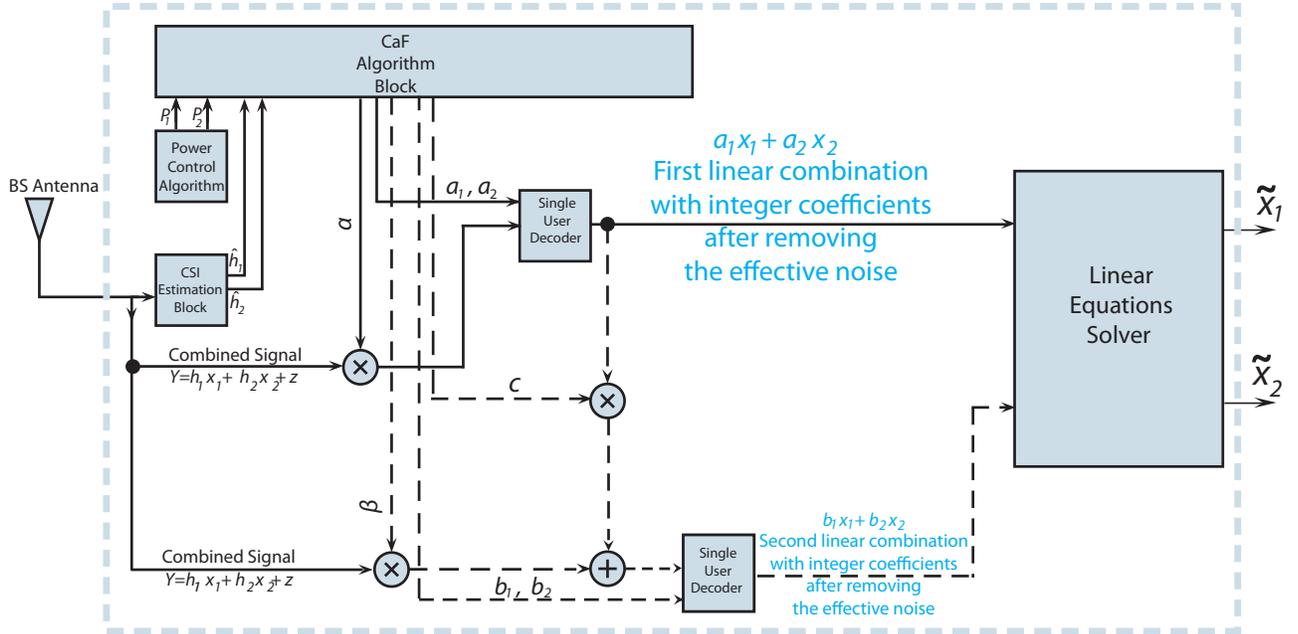


Figure 4.5: Block diagram illustrating CaF decoding in a single antenna BS receiver.

illustrated in Algorithm 2. The accuracy of estimated channel coefficients is vital to produce optimal multiplying CaF coefficients. On the other hand, CaF is sensitive to channel estimation errors, as explained in [65]. In particular, imperfect CSI may lead to a significant rate loss [65, 66]. An approach to overcoming the effects of CSI accuracy has been investigated in [66] where blind CaF has been proposed to avoid the need for CSI. However, this approach has an inherent trade-off between computational complexity and achievable rates. In this section, we dive deep to characterize the effects of channel estimation errors on the achievable rates, and we then propose a solution to mitigate these effects.

4.7.1 Effects of Channel Estimation Errors

To capture the effects of imperfect CSI, we need to introduce some additional details of integer-forcing decoding. Again, for ease of presentation, we start from the case of real-valued models of channel coefficients and then extend it to complex-valued models of channel coefficients. Recall that the receiver first computes an integer combination $a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2$. This step is done as follows. The receiver applies a so-called *equalization vector* $\alpha \in \mathbb{R}^M$ to its received signal \mathbf{Y} in order to

obtain the effective channel

$$\begin{aligned}
 \tilde{\mathbf{y}}_1 &= \boldsymbol{\alpha}^T \mathbf{Y} \\
 &= \boldsymbol{\alpha}^T (\mathbf{h}_1 \mathbf{x}_1 + \mathbf{h}_2 \mathbf{x}_2) + \boldsymbol{\alpha}^T \mathbf{W} \\
 &= a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + (\boldsymbol{\alpha}^T \mathbf{h}_1 - a_1) \mathbf{x}_1 + (\boldsymbol{\alpha}^T \mathbf{h}_2 - a_2) \mathbf{x}_2 + \boldsymbol{\alpha}^T \mathbf{W}
 \end{aligned}$$

where $(\boldsymbol{\alpha}^T \mathbf{h}_1 - a_1) \mathbf{x}_1 + (\boldsymbol{\alpha}^T \mathbf{h}_2 - a_2) \mathbf{x}_2 + \boldsymbol{\alpha}^T \mathbf{W}$ is the *effective noise*. The receiver as shown in Figure 4.5 then decodes $\tilde{\mathbf{y}}_1$ to the closest lattice codeword. Intuitively, the decoding of $a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2$ is successful if the effective noise is small enough. As shown in [130], the effective noise variance is given by

$$\begin{aligned}
 \sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}) &= \frac{1}{n} \mathbb{E} \left\{ \left\| (\boldsymbol{\alpha}^T \mathbf{h}_1 - a_1) \mathbf{x}_1 + (\boldsymbol{\alpha}^T \mathbf{h}_2 - a_2) \mathbf{x}_2 + \boldsymbol{\alpha}^T \mathbf{W} \right\|_2^2 \right\} \\
 &= \|\boldsymbol{\alpha}\|_2^2 + \|(\boldsymbol{\alpha}^T \mathbf{h}_1 - a_1, \boldsymbol{\alpha}^T \mathbf{h}_2 - a_2) \mathbf{P}^{1/2}\|_2^2 \\
 &= \boldsymbol{\alpha}^T \boldsymbol{\alpha} + (\boldsymbol{\alpha}^T \mathbf{H} - \mathbf{a}^T) \mathbf{P} (\mathbf{H}^T \boldsymbol{\alpha} - \mathbf{a})
 \end{aligned}$$

which can be expanded as

$$\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}) = \boldsymbol{\alpha}^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\alpha} - \mathbf{a}^T \mathbf{P} \mathbf{H}^T \boldsymbol{\alpha} - \boldsymbol{\alpha}^T \mathbf{H} \mathbf{P} \mathbf{a} + \mathbf{a}^T \mathbf{P} \mathbf{a}.$$

Hence, $\boldsymbol{\alpha}$ should be chosen to minimize the effective noise variance $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha})$. Its optimal value $\boldsymbol{\alpha}_*$ can be found by setting the derivative of $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha})$ to zero:

$$2\boldsymbol{\alpha}^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) - 2\mathbf{a}^T \mathbf{P} \mathbf{H}^T = \mathbf{0}. \quad (4.11)$$

The corresponding $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}_*)$ is given by $\mathbf{a}^T (\mathbf{P}^{-1} + \mathbf{H}^T \mathbf{H})^{-1} \mathbf{a}$, which is equal to $\sigma^2(\mathbf{H}, \mathbf{a})$ as expected. Note that $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}_*)$ depends on \mathbf{H} , where only estimated version of \mathbf{H} is available at the BS. Therefore, estimation accuracy is vital for determining the optimal value of $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}_*)$ and also to prevent any error prorogation to the computation of second integer combination.

Accordingly after computing the first integer combination, the receiver then computes a second integer combination $b_1\mathbf{x}_1 + b_2\mathbf{x}_2$ by using the first combination as side information. This step is done as follows. The receiver forms the effective channel

$$\begin{aligned}\tilde{\mathbf{y}}_2 &= \boldsymbol{\beta}^T \mathbf{Y} + c(a_1\mathbf{x}_1 + a_2\mathbf{x}_2) \\ &= b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + (\boldsymbol{\beta}^T \mathbf{h}_1 + ca_1 - b_1)\mathbf{x}_1 + (\boldsymbol{\beta}^T \mathbf{h}_2 + ca_2 - b_2)\mathbf{x}_2 + \boldsymbol{\beta}^T \mathbf{W}\end{aligned}$$

where both $\boldsymbol{\beta} \in \mathbb{R}^M$ and $c \in \mathbb{R}$ are for the purpose of equalization. Here, the effective noise is $(\boldsymbol{\beta}^T \mathbf{h}_1 + ca_1 - b_1)\mathbf{x}_1 + (\boldsymbol{\beta}^T \mathbf{h}_2 + ca_2 - b_2)\mathbf{x}_2 + \boldsymbol{\beta}^T \mathbf{W}$ with the effective noise variance given by

$$\begin{aligned}\sigma^2(\mathbf{H}, \mathbf{b} | \mathbf{a}; \boldsymbol{\beta}, c) &= \frac{1}{n} \mathbb{E} \left\{ \left\| (\boldsymbol{\beta}^T \mathbf{h}_1 + ca_1 - b_1)\mathbf{x}_1 + (\boldsymbol{\beta}^T \mathbf{h}_2 + ca_2 - b_2)\mathbf{x}_2 + \boldsymbol{\beta}^T \mathbf{W} \right\|_2^2 \right\} \\ &= \boldsymbol{\beta}^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\beta} - (\mathbf{b}^T - c\mathbf{a}^T) \mathbf{P}\mathbf{H}^T \boldsymbol{\beta} - \boldsymbol{\beta}^T \mathbf{H}\mathbf{P}(\mathbf{b} - c\mathbf{a}) + (\mathbf{b}^T - c\mathbf{a}^T) \mathbf{P}(\mathbf{b} - c\mathbf{a}).\end{aligned}$$

Hence, $\boldsymbol{\beta}$ and c should be chosen to minimize the effective noise variance $\sigma^2(\mathbf{H}, \mathbf{b} | \mathbf{a}; \boldsymbol{\beta}, c)$. The optimal value $\boldsymbol{\beta}_*$ can be found by setting the derivative of $\sigma^2(\mathbf{H}, \mathbf{b} | \mathbf{a}; \boldsymbol{\beta}, c)$ (with respect to $\boldsymbol{\beta}$) to zero:

$$2\boldsymbol{\beta}^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) - 2(\mathbf{b}^T - c\mathbf{a}^T) \mathbf{P}\mathbf{H}^T = \mathbf{0}. \quad (4.12)$$

Similarly, the optimal value c_* can be found by setting the derivative of $\sigma^2(\mathbf{H}, \mathbf{b} | \mathbf{a}; \boldsymbol{\beta}, c)$ (with respect to c) to zero:

$$2(\mathbf{a}^T \mathbf{P}\mathbf{a})c - 2\mathbf{a}^T \mathbf{P}\mathbf{b} + 2\mathbf{a}^T \mathbf{P}\mathbf{H}^T \boldsymbol{\beta} = 0. \quad (4.13)$$

It is important to mention that (4.11), (4.12), and (4.13) are used to generate the optimal values of $\boldsymbol{\alpha}_*$, $\boldsymbol{\beta}_*$, c_* respectively. They all depend on the estimation accuracy of the channel matrix \mathbf{H} to obtain the optimal values. Therefore, estimation errors of the channel matrix lead to errors in the obtained optimal multiplying (i.e., equalization) coefficients, where ϵ_α , ϵ_β , and ϵ_c represent the effects of resulting errors on $\boldsymbol{\alpha}_*$, $\boldsymbol{\beta}_*$, c_* , respectively. Thus, the sub-optimal multiplying coefficients can be modeled as $\boldsymbol{\alpha}_* + \epsilon_\alpha$, $\boldsymbol{\beta}_* + \epsilon_\beta$, and $c_* + \epsilon_c$ respectively, all of which result in generating sub-optimal effective noise variances (i.e., the effective noise variance in each case

becomes greater than the optimal effective noise variance). These effective noise variances are given by:

$$\begin{aligned}
 \sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha) &= (\boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha)^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) (\boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha) \\
 &\quad - \mathbf{a}^T \mathbf{P}\mathbf{H}^T (\boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha) - (\boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha)^T \mathbf{H}\mathbf{P}\mathbf{a} + \mathbf{a}^T \mathbf{P}\mathbf{a} \\
 &= \sigma^2(\mathbf{H}, \mathbf{a}) + (\boldsymbol{\alpha}_*)^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\epsilon}_\alpha + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\epsilon}_\alpha \\
 &\quad - \mathbf{a}^T \mathbf{P}\mathbf{H}^T \boldsymbol{\epsilon}_\alpha - \boldsymbol{\epsilon}_\alpha^T \mathbf{H}\mathbf{P}\mathbf{a} \\
 &= \sigma^2(\mathbf{H}, \mathbf{a}) + ((\boldsymbol{\alpha}_*)^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) - \mathbf{a}^T \mathbf{P}\mathbf{H}^T) \boldsymbol{\epsilon}_\alpha + \boldsymbol{\epsilon}_\alpha^T ((\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\alpha}_* - \mathbf{H}\mathbf{P}\mathbf{a}) \\
 &\quad + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\epsilon}_\alpha \\
 &= \sigma^2(\mathbf{H}, \mathbf{a}) + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\epsilon}_\alpha,
 \end{aligned}$$

since both $(\boldsymbol{\alpha}_*)^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) - \mathbf{a}^T \mathbf{P}\mathbf{H}^T$ and $(\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\alpha}_* - \mathbf{H}\mathbf{P}\mathbf{a}$ are equal to $\mathbf{0}$ based on (4.11).

Similarly, we have

$$\begin{aligned}
 \sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}; \boldsymbol{\beta}_* + \boldsymbol{\epsilon}_\beta, c_* + \boldsymbol{\epsilon}_c) &= \sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) + \boldsymbol{\epsilon}_\beta^T (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^T) \boldsymbol{\epsilon}_\beta + \boldsymbol{\epsilon}_c^2 \mathbf{a}^T \mathbf{P}\mathbf{a} \\
 &\quad + \boldsymbol{\epsilon}_c \mathbf{a}^T \mathbf{P}\mathbf{H}^T \boldsymbol{\epsilon}_\beta + \boldsymbol{\epsilon}_\beta^T \mathbf{H}\mathbf{P}\mathbf{a} \boldsymbol{\epsilon}_c.
 \end{aligned}$$

Clearly, when $\boldsymbol{\epsilon}_\alpha = \mathbf{0}$, we have $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha) = \sigma^2(\mathbf{H}, \mathbf{a})$. Similarly, when $\boldsymbol{\epsilon}_\beta = \mathbf{0}$ and $\boldsymbol{\epsilon}_c = 0$, we have $\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}; \boldsymbol{\beta}_* + \boldsymbol{\epsilon}_\beta, c_* + \boldsymbol{\epsilon}_c) = \sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a})$. Substituting the expressions of $\sigma^2(\mathbf{H}, \mathbf{a}; \boldsymbol{\alpha}_* + \boldsymbol{\epsilon}_\alpha)$ and $\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}; \boldsymbol{\beta}_* + \boldsymbol{\epsilon}_\beta, c_* + \boldsymbol{\epsilon}_c)$ into Proposition 1, we have the following theorem which extends the work of Pappi, Karagiannidis, and Schober [65] from the single-antenna case to multi-antenna case. This may be of independent interest.

Theorem 7 *The following rate pairs are achievable under imperfect CSI.*

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p = 0$,

$$R_1 = \frac{1}{2} \log_2 \left(\frac{P_1}{P_1 - P_1^2 \mathbf{h}_1^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T)^{-1} \mathbf{h}_1 + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\alpha} \right)$$

$$R_2 = \frac{1}{2} \log_2 \left(\frac{P_2}{P_2 (1 + P_2 \|\mathbf{h}_2\|_2^2)^{-1} + \boldsymbol{\epsilon}_\beta^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\beta + P_1 \boldsymbol{\epsilon}_c^2 + \boldsymbol{\epsilon}_c P_1 \mathbf{h}_1^T \boldsymbol{\epsilon}_\beta + \boldsymbol{\epsilon}_\beta^T \mathbf{h}_1 P_1 \boldsymbol{\epsilon}_c} \right)$$

- When $a_1 \bmod p = 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \frac{1}{2} \log_2 \left(\frac{P_1}{P_1 (1 + P_1 \|\mathbf{h}_1\|_2^2)^{-1} + \boldsymbol{\epsilon}_\beta^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\beta + P_2 \boldsymbol{\epsilon}_c^2 + \boldsymbol{\epsilon}_c P_2 \mathbf{h}_2^T \boldsymbol{\epsilon}_\beta + \boldsymbol{\epsilon}_\beta^T \mathbf{h}_2 P_2 \boldsymbol{\epsilon}_c} \right)$$

$$R_2 = \frac{1}{2} \log_2 \left(\frac{P_2}{P_2 - P_2^2 \mathbf{h}_2^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T)^{-1} \mathbf{h}_2 + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\alpha} \right)$$

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \min \left\{ \frac{1}{2} \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\alpha} \right), \frac{1}{2} \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b} | \mathbf{a}) + g(\boldsymbol{\epsilon}_\beta, \boldsymbol{\epsilon}_c)} \right) \right\}$$

$$R_2 = \frac{1}{2} \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\alpha} \right)$$

or

$$R_1 = \frac{1}{2} \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\alpha} \right),$$

$$R_2 = \min \left\{ \frac{1}{2} \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \boldsymbol{\epsilon}_\alpha^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\alpha} \right), \frac{1}{2} \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b} | \mathbf{a}) + g(\boldsymbol{\epsilon}_\beta, \boldsymbol{\epsilon}_c)} \right) \right\}$$

where

$$g(\boldsymbol{\epsilon}_\beta, \boldsymbol{\epsilon}_c) = \boldsymbol{\epsilon}_\beta^T (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^T) \boldsymbol{\epsilon}_\beta + \boldsymbol{\epsilon}_c^2 \mathbf{a}^T \mathbf{P} \mathbf{a} + \boldsymbol{\epsilon}_c \mathbf{a}^T \mathbf{P} \mathbf{H}^T \boldsymbol{\epsilon}_\beta + \boldsymbol{\epsilon}_\beta^T \mathbf{H} \mathbf{P} \mathbf{a} \boldsymbol{\epsilon}_c.$$

Finally, by following the algebraic framework in [60] and taking Proposition 2 into account, we can obtain the following achievable rates for the channel-coefficient matrix \mathbf{H} over the complex field.

Theorem 8 *The following rate pairs are achievable for complex-valued \mathbf{H} under imperfect CSI.*

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p = 0$,

$$R_1 = \log_2 \left(\frac{P_1}{P_1 - P_1^2 \mathbf{h}_1^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H)^{-1} \mathbf{h}_1 + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\alpha} \right)$$

$$R_2 = \log_2 \left(\frac{P_2}{P_2 (1 + P_2 \|\mathbf{h}_2\|_2^2)^{-1} + \epsilon_\beta^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\beta + |\epsilon_c|^2 P_1 + \epsilon_c^H P_1 \mathbf{h}_1^H \epsilon_\beta + \epsilon_\beta^H \mathbf{h}_1 P_1 \epsilon_c} \right)$$

- When $a_1 \bmod p = 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \log_2 \left(\frac{P_1}{P_1 (1 + P_1 \|\mathbf{h}_1\|_2^2)^{-1} + \epsilon_\beta^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\beta + |\epsilon_c|^2 P_2 + \epsilon_c^H P_2 \mathbf{h}_2^H \epsilon_\beta + \epsilon_\beta^H \mathbf{h}_2 P_2 \epsilon_c} \right)$$

$$R_2 = \log_2 \left(\frac{P_2}{P_2 - P_2^2 \mathbf{h}_2^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H)^{-1} \mathbf{h}_2 + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\alpha} \right)$$

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p \neq 0$,

$$R_1 = \min \left\{ \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\alpha} \right), \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) + g(\epsilon_\beta, \epsilon_c)} \right) \right\}$$

$$R_2 = \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\alpha} \right)$$

or

$$R_1 = \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\alpha} \right),$$

$$R_2 = \min \left\{ \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\alpha} \right), \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) + g(\epsilon_\beta, \epsilon_c)} \right) \right\}$$

where

$$g(\epsilon_\beta, \epsilon_c) = \epsilon_\beta^H (\mathbf{I}_M + \mathbf{H} \mathbf{P} \mathbf{H}^H) \epsilon_\beta + |\epsilon_c|^2 \mathbf{a}^H \mathbf{P} \mathbf{a} + \epsilon_c^H \mathbf{a}^H \mathbf{P} \mathbf{H}^H \epsilon_\beta + \epsilon_\beta^H \mathbf{H} \mathbf{P} \mathbf{a} \epsilon_c. \quad (4.14)$$

Proof: This theorem follows immediately from Theorem 7 through a real-valued decomposition technique as used in [59] or applying the algebraic framework developed in [60] over Gaussian integers. \square

Remark 5 *Theorem 8 highlights the impact of channel coefficients estimation errors, which result in increasing the effective noise variances and a reduction of the rates that CaF decoding achieves. Besides, it can be used in conjunction with Proposition 2 to compute the rate loss that results due to imperfect CSI. It is clear that subtracting the sum-rate (i.e., $R_1 + R_2$ in each case) provided by Theorem 8 from the sum-rate provided by (4.8) gives the rate loss. Clearly the sum-rate loss depends on the multiplying coefficients \mathbf{a}_* and \mathbf{b}_* . The sum-rate loss L_{SR} can be evaluated by:*

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p = 0$,

$$L_{SR} = \log_2 \left\{ \left(|\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H| \right) (P_1 - P_1^2 \mathbf{h}_1^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H)^{-1} \mathbf{h}_1 + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha) \right. \\ \left. (P_2 (1 + P_2 \|\mathbf{h}_2\|_2^2)^{-1} + \epsilon_\beta^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\beta + |\epsilon_c|^2 P_1 + \epsilon_c^H P_1 \mathbf{h}_1^H \epsilon_\beta + \epsilon_\beta^H \mathbf{h}_1 P_1 \epsilon_c) \right\} \\ - \log_2(P_1 P_2)$$

- When $a_1 \bmod p = 0$ and $a_2 \bmod p \neq 0$,

$$L_{SR} = \log_2 \left\{ \left(|\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H| \right) (P_1 (1 + P_1 \|\mathbf{h}_1\|_2^2)^{-1} + \epsilon_\beta^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\beta + |\epsilon_c|^2 P_2 \right. \\ \left. + \epsilon_c^H P_2 \mathbf{h}_2^H \epsilon_\beta + \epsilon_\beta^H \mathbf{h}_2 P_2 \epsilon_c) (P_2 - P_2^2 \mathbf{h}_2^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H)^{-1} \mathbf{h}_2 + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha) \right\} \\ - \log_2(P_1 P_2)$$

- When $a_1 \bmod p \neq 0$ and $a_2 \bmod p \neq 0$,

$$\begin{aligned}
 L_{SR} = & \min \left\{ \log_2 |\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H| - \min \left\{ \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha} \right) \right. \right. \\
 & + \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha} \right), \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) + g(\epsilon_\beta, \epsilon_c)} \right) \\
 & \left. \left. + \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha} \right) \right\}, \log_2 |\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H| \right. \\
 & - \min \left\{ \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha} \right) + \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha} \right), \right. \\
 & \left. \left. \log_2^+ \left(\frac{P_2}{\sigma^2(\mathbf{H}, \mathbf{b}|\mathbf{a}) + g(\epsilon_\beta, \epsilon_c)} \right) + \log_2^+ \left(\frac{P_1}{\sigma^2(\mathbf{H}, \mathbf{a}) + \epsilon_\alpha^H (\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H) \epsilon_\alpha} \right) \right\} \right\}
 \end{aligned}$$

where $g(\epsilon_\beta, \epsilon_c)$ is given by (4.14).

4.7.2 Employing DPS and EDPS to Mitigate Channel Estimation Errors

As aforementioned and proven in the previous chapters, the employing of DPS and EDPS in massive MIMO with OMA and NOMA integration, respectively, results in improving the channel estimation accuracy. In other words, the employment of DPS and EDPS minimizes the channel estimation error's matrix, i.e., $\mathbf{E} = \mathbf{H} - \hat{\mathbf{H}}$, and reduces its effects on generating and/or increasing ϵ_α , ϵ_β , and ϵ_c which further impacts on the overall system performance. Therefore, DPS and EDPS improve the performance of CaF by reducing channel estimation errors, promoting the optimal computation of CaF multiplying coefficients, which prevent a significant reduction-loss in achievable rates and maintain higher sum-rate capacity of the cell.

4.8 Numerical Results

4.8.1 CaF-NOMA with Perfect CSI

In this sub-section, we conduct Monte Carlo simulations to evaluate the advantage of CaF decoding over SIC decoding. In our simulations, we have 100 users including 40 near users whose distance to the BS is less than D_1 and 60 far users whose distance to the BS is between D_1 and D_2 . All the

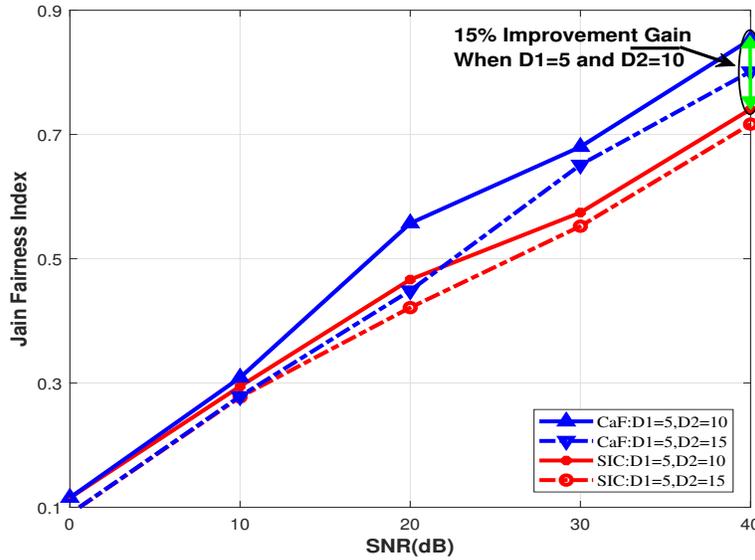


Figure 4.6: The fairness comparison of SIC-NOMA and CaF-NOMA with different $\frac{D1}{D2}$.

users are randomly deployed in the two areas. We set the path-loss exponent $\rho = 2$. For simplicity, our simulations apply equal power for the paired users. We vary the values of $D1$ and $D2$ to demonstrate the advantage further. We employ *Jain's fairness index*, which is defined in [136], and given by:

$$J = \frac{(\sum_{k=1}^K R_k)^2}{K \sum_{k=1}^K (R_k)^2} \quad (4.15)$$

as our fairness measure, and the *average outage probability* (i.e., the standard outage probability of having a rate less than a targeted rate R_t averaged by all the users) as our performance metric.

In Figure 4.6, the performance of Jain's fairness index is shown as a function of SNR. The proposed CaF-NOMA can outperform the conventional SIC-NOMA, especially in high SNR regimes. CaF-NOMA can increase the Jain's fairness 15% when SNR equals to 40 dB. This is because CaF-NOMA can balance the two users' rates compared to SIC-NOMA. This conclusion also holds when the ratio of $D1$ to $D2$ changes from 1/2 to 1/3.

In Figure 4.7, the performance of average outage probability is shown as a function of SNR. Once again, the proposed CaF-NOMA performs better than SIC-NOMA, especially in high SNR regime. CaF-NOMA can decrease the average outage probability 46% when SNR equals to 40 dB.

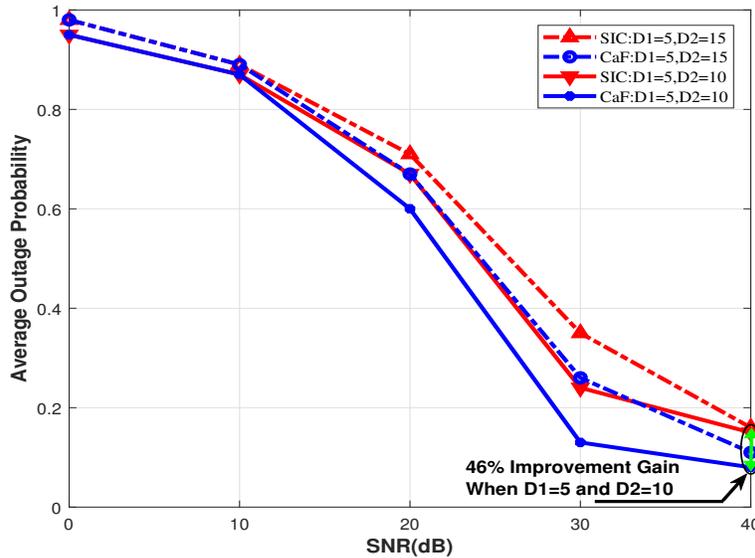


Figure 4.7: The average outage probability comparison of SIC-NOMA and CaF-NOMA with different $\frac{D_1}{D_2}$ and $R_t = 2$ BPCU.

4.8.2 CaF-NOMA in Massive MIMO with Imperfect CSI

In this sub-section, we conduct MATLAB simulations to demonstrate the effectiveness of employing DPS/EDPS to mitigate the CaF sensitivity to channel estimation errors in IFLR of a BS located at the center of a single cell; the BS is equipped with massive MIMO with PD-NOMA. We show that the impact of imperfect CSI on the reduction of the achievable rates by two methods, namely, channel estimation employing SPS, channel estimation employing DPS/EDPS, and compare them with the achievable rates of having perfect CSI. The considered single cell scenario is shown in Figure 4.4, under which both DPS/EDPS provide the same performance and work similarly. Further, the advantage of EDPS over DPS appears in a non-cooperative multiple cells scenario of massive MIMO-NOMA.

Rate Region:

The MAC capacity region represents the closure of the set of all possible achievable rate-pairs of multiple users. Figure 4.8 shows the MAC capacity region as well as four dominant rate pairs using CaF as the core of an IFLR of a BS that is equipped by a massive MIMO with PD-NOMA.

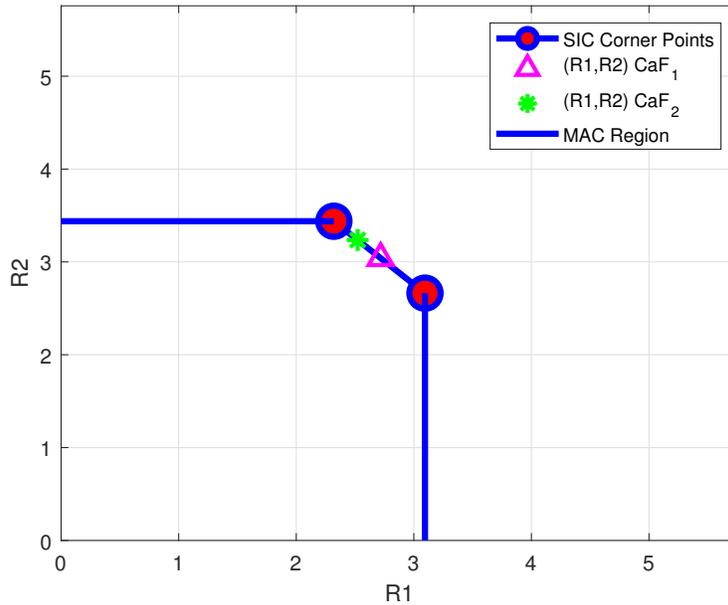


Figure 4.8: Gaussian MAC region for massive MIMO-NOMA IFLR, SIC corner points, and CaF dominant rate pairs when $M=16$.

Thereafter, under the same configurations, we assess the performance of DPS/EDPS and SPS, and we computed the achievable rate region at each case. Then, we compare the obtained rate regions with the capacity region of the perfect channel estimation.

Figure 4.9 shows the obtained rate regions using DPS/EDPS, SPS, and perfect CSI. It also shows the implications of channel estimation errors, where estimation errors resulted in shrinking the rate region in a proportional relation where higher estimation errors leads to a higher shrinking in size. Figure 4.9, also shows that DPS/EDPS have better performance than SPS, where the estimation errors are minimized and consequently the resulted shrinking is minimized. Therefore, DPS/EDPS outperform SPS and provide the CaF a robustness against the sensitivity of estimation errors, suppress the effects of the consequences and maintain the rate region. Afterwards, we assess the reliability of the robustness provided by DPS/EDPS. Therefrom, we investigate the performance of DPS/EDPS and SPS in a noisy environment. Figure 4.10 demonstrates that the robustness provided by DPS/EDPS against the sensitivity of estimation errors is still significant in a noisy environment, which is an additional desirable feature to promote the implementation of

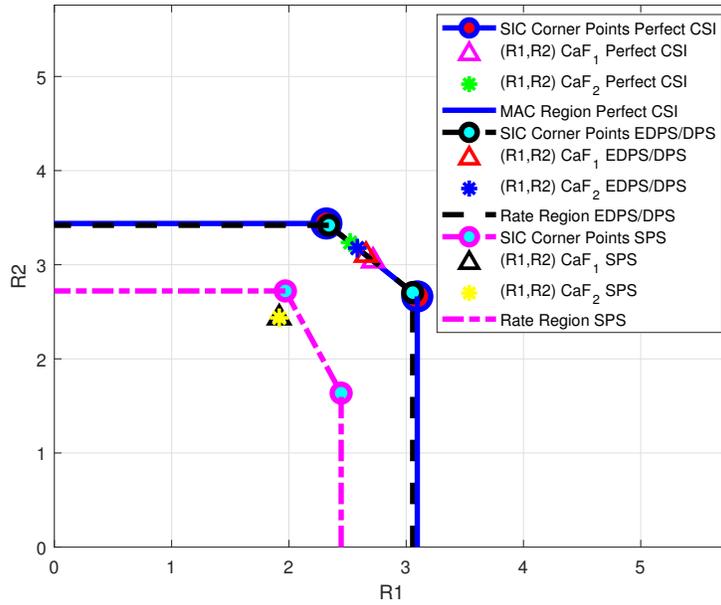


Figure 4.9: Comparison of achieved Gaussian MAC region for massive MIMO-NOMA in IFLR, at perfect CSI, using DPS/EDPS, and using SPS.

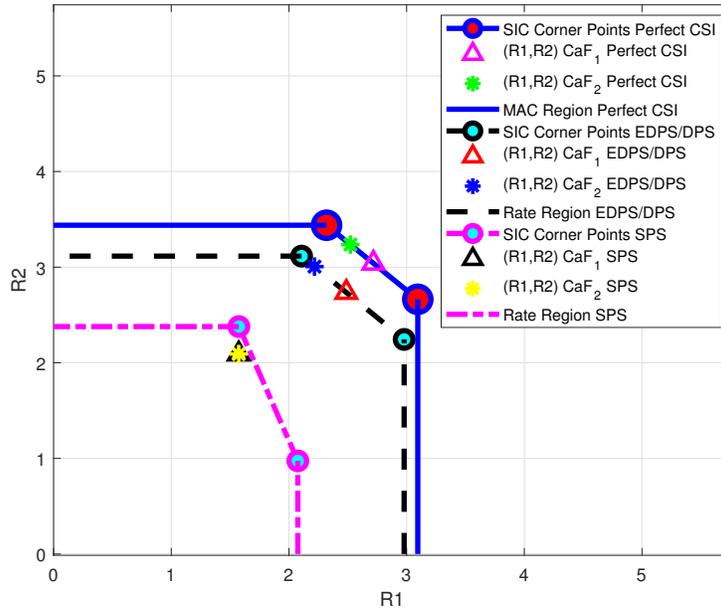


Figure 4.10: Comparison of achieved Gaussian MAC region for massive MIMO-NOMA in IFLR in a noisy environment, at perfect CSI, using EDPS/DPS, and using SPS.

DPS/EDPS in practical systems.

Average Sum-Rate Loss:

The mathematical derivations in Subsection 4.7.1 demonstrate analytically the effects of imperfect CSI on the achievable rates. Also, in Subsection 4.7.1, we introduce ϵ_α , ϵ_β , and ϵ_c to represent the consequences of the channel estimation errors and to capture the effects of imperfect CSI on computing the optimal multiplying coefficients (i.e., equalization vectors) of CaF in IFLR of massive MIMO with PD-NOMA. Therefore, we choose the average sum-rate loss, which is measured in BPCU, as the performance metric to highlight the effects of ϵ_α , ϵ_β , and ϵ_c on the overall performance.

Figure 4.11 illustrates the effects of varying the variance $\sigma_{\epsilon_\alpha}^2$ of ϵ_α on the average sum-rate loss at different configurations, whereas, Figure 4.12 illustrates the effects of varying the variance $\sigma_{\epsilon_\beta}^2$ of ϵ_β on the average sum-rate loss at different configurations. Further, Figure 4.13 illustrates the effects of varying the variance $\sigma_{\epsilon_c}^2$ of ϵ_c on the average sum-rate loss at different configurations.

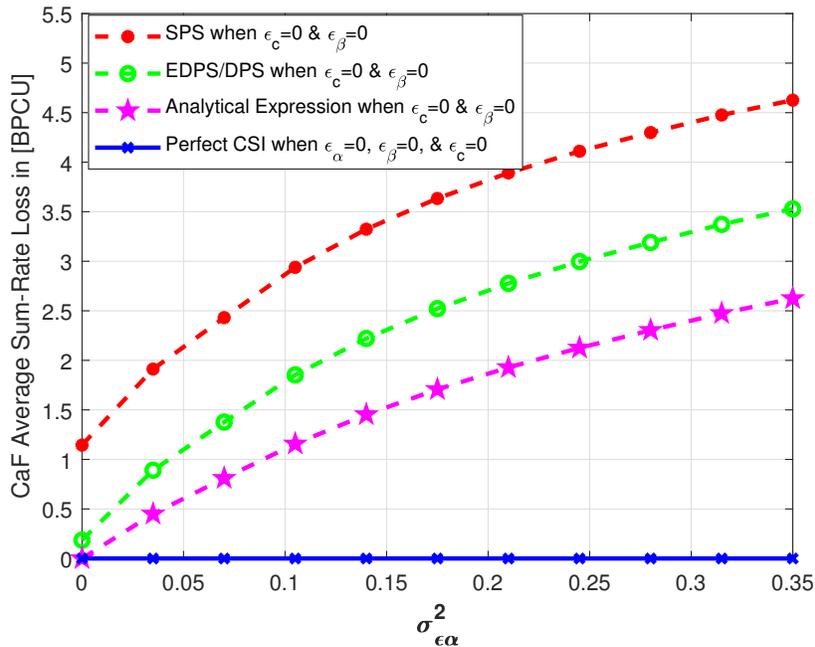


Figure 4.11: Performance comparisons of the average sum-rate loss between EDPS/DPS and SPS over a range of $\sigma_{\epsilon_\alpha}^2$.

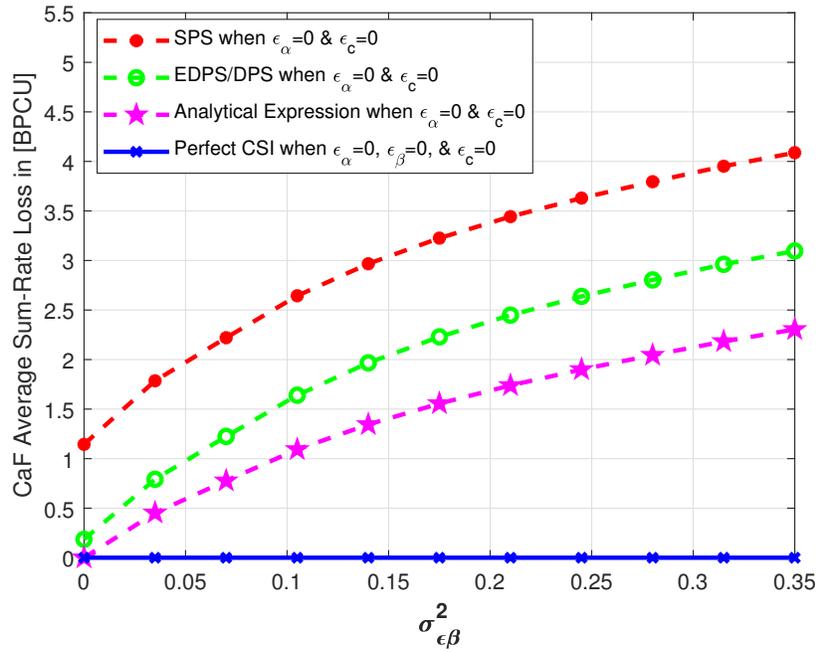


Figure 4.12: Performance comparisons of the average sum-rate loss between EDPS/DPS and SPS over a range of $\sigma_{\epsilon\beta}^2$.

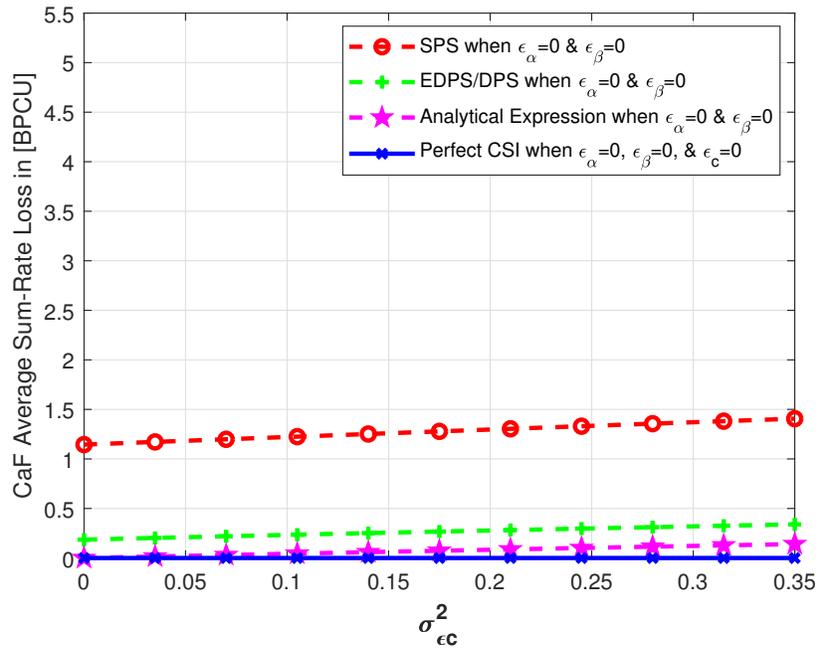


Figure 4.13: Performance comparisons of the average sum-rate loss between EDPS/DPS and SPS over a range of $\sigma_{\epsilon_c}^2$.

The results demonstrate that as ϵ_α has the only dominant effects on computing the first multiplying CaF coefficient to decode the first integer combination in IFLR, so $\sigma_{\epsilon_\alpha}^2$ has significant effects on reducing the average sum-rate loss. As $\sigma_{\epsilon_\alpha}^2$ increases the average sum-rate loss increases rapidly. Similarly, as ϵ_β has a significant dominant effects on computing the second multiplying CaF coefficient to decode the second integer combination in IFLR, so $\sigma_{\epsilon_\beta}^2$ has significant effects on reducing the average sum-rate loss as well. Then, in a similar way to the attribute of change while varying $\sigma_{\epsilon_\alpha}^2$, as $\sigma_{\epsilon_\beta}^2$ increases the average sum-rate loss increases rapidly. Additionally, ϵ_c affects the computation of the second multiplying CaF coefficients to decode the second linear combination in an IFLR, so $\sigma_{\epsilon_c}^2$ has a semi-constant effects (i.e., almost stable) over a range of $\sigma_{\epsilon_c}^2$ on reducing the average sum-rate loss in IFLR that employs DPS/EDPS. In contrast, in an IFLR that employs SPS, as $\sigma_{\epsilon_c}^2$ increases the average sum-rate loss increases slowly for SPS in a linear fashion.

Furthermore, in a comparison with the presence of ϵ_β and ϵ_c , the results show that the presence of ϵ_α has more significant effects to reduce the sum-rate and increases the average sum-rate loss because its effects go beyond the decoding of the first linear combination. Actually, the high average sum-rate loss resulted in the presence of ϵ_α is due to the propagating effects of ϵ_α to the decoding of the second linear combination, which matches the theoretical expectation. It is important to mention that the joint presences of ϵ_α with either ϵ_β , ϵ_c , or both worsens the performance and increase the average sum-rate loss. Finally, the results have shown that DPS/EDPS outperforms SPS and as a consequence IFLR in a massive MIMO with PD-NOMA that employs DPS/EDPS achieves higher sum-rate capacity, mitigates the sensitivity of CaF for imperfect CSI in a robustness way, and suffers a lower average sum-rate loss.

4.9 Conclusion

In this chapter, we have shown that CaF-NOMA can achieve better fairness and outage probability than SIC-NOMA. In particular, we have explained how the BS can find out all the Pareto-optimal solutions and why CaF decoding has essentially the same complexity as SIC decoding. In addition, we have conducted various simulations to demonstrate the advantages of CaF-NOMA over SIC-

NOMA in terms of Jain's fairness and average outage probability. We have shown how an efficient IFLR can be constructed via nested lattice codes. We have derived analytically the loss in the achievable sum-rate, due to imperfect CSI.

We have shown that the sensitivity of CaF decoding in massive MIMO with PD-NOMA integration to the channel estimation errors. We have demonstrated the impacts of this sensitivity on the performance of IFLR. Through extensive simulations, we have shown how DPS/EDPS mitigate this sensitivity and provide robustness against it, and consequently achieve better sum-rate capacity of the cell and enhances the performance. Consequently, this adds desirable features for the proposed construction of IFLR, which promotes its practical implementation into real systems. In particular, we have found that DPS/EDPS achieve significantly larger rate regions than SPS as DPS/EDPS improve the estimation accuracy of CSI. Furthermore, we have found that as channel estimation errors increase, the average sum-rate loss increases, which shrinks the size of the rate region rapidly.

Chapter 5

Conclusions and Future Works

In this chapter, we summarize the contributions of this dissertation, provide concluding remarks on the accomplished works, and suggest several topics for future research.

5.1 Concluding Remarks

The technical targets of the 5G cellular networks are very ambitious yet essential to offer mobile users a variety of services and to support a wide range of applications. However, several challenges and obstacles need to be overcome to reach these technical targets. In this dissertation, we have focused on some of these challenges and developed novel solutions to improve the performance of 5G cellular networks and beyond. These solutions can be utilized in various state-of-the-art wireless systems, as we have considered different architectures in developing and providing the proposed solutions to improve 5G technologies. Distinguishably from most research in improving 5G technologies, the schemes developed in this dissertation followed different approaches that ensure promising innovative economical-engineering solutions. The proposed schemes have considered the practical implementation feasibility that have low complexity and backward compatibility and can be generalized for a wide range of applications in 5G cellular networks and beyond. Therefore, the research work we have developed throughout the dissertation provides valuable timely contributions that can be taken as a stepping stone and transforming into tools for tackling real-world problems of 5G and beyond to advance communication technologies to facilitate supporting a wide

range of applications, achieved the technical targets, and offering new services. The main concluding remarks we draw from our work in this dissertation are briefly illustrated in the following list and sorted by each chapter.

5.1.1 Summaries and Contributions

- In Chapter 2, we have introduced and proposed the novel DPS. DPS is a general pilot signaling scheme that can be used in any wireless communication system. We have demonstrated that using DPS at massive MIMO with OMA integration significantly suppresses pilot contamination, which results in an accurate CSI and enhanced channel estimation at the BS during the uplink training. This way, besides enhancing the uplink transmission, employing DPS results in a more efficient use in precoding by the BS during downlink transmission and therefore improving the cell sum-rate. Moreover, DPS offers multifold unique desirable features and provides additional useful information to the BSs, which results in more opportunities and in enhancing the overall system performance without cooperation between BSs. DPS is easy to implement, uses the same overhead during the training period as the conventional schemes, does not necessitate additional hardware or a change in the cellular network backhaul, not reliant on user's mobility speed, and improves SE. Unlike state-of-the-art works, we have shown that DPS ensures a uniform high QoS for each and every UE. However, DPS advantages come with a price of serving L fewer UEs than with SPS, where L is the number of cells in the system-cluster. This is because SPS uses all orthogonal PSs to serve the UEs in each cell, whereas DPS needs to use L of these sequences as the Cell-IDs and can only assign $N - L$ sequences to the UEs in each cell (N is the number of available orthogonal sequences). However, we have shown this disadvantage diminishes as number of UEs grows large. Moreover, despite the number of PSs available for assignment to UEs in a cell determines the maximum number of UEs that can be served simultaneously, in practice, the cell can only serve a smaller number of UEs due to pilot contamination effects, in the case of SPS, where an acceptable level of interference is the constraint.

We have provided insights on how the DPS cluster configuration can be planned and designed to cover the entire coverage area of massive MIMO by using a cluster size of seven adjacent non-cooperative cells. It has shown that in SPS other inter-cell interference that arise from the second and third tiers on a BS at the home cell, worsen the pilot contamination problem. Therefore, in comparison with SPS, this clustering configuration of the DPS allows a further greater enhancement, improves SE, and enhances system capacity. Thereby, DPS with the proposed clustering configuration preserves a low level of interference and facilitates scalability which is an essential feature to be offered in 5G cellular networks and beyond.

- In Chapter 3, we have deepened our analyses of the uplink transmission in massive MIMO presented in Chapter 2 to include NOMA into the system architecture, and we have extended our investigation to include EE. The novel EDPS that we have proposed in Chapter 3 is an efficient scheme to obtain an accurate CSI in massive MIMO with NOMA integration. EDPS is an extended version of DPS that tackles both intra-cell and inter-cell interference in an environment of non-cooperative multiple cells cellular networks. We have found that massive MIMO and NOMA aid each other and achieve better performance when integrated, provided that an accurate CSI is available at the BS. EDPS facilitates this integration.

The EE of DFT-s-OFDM waveforms in 5G uplinks has been examined for both massive MIMO with OMA and massive MIMO with PD-NOMA. We have found that improving CSI results in improving EE. We have clarified the impact of improving uplink EE on the EE of downlink and overall system. In fact, we have found that the enhancement in uplink leads to performance enhancement of downlink as well, as TDD is the adopted duplexing scheme. We have proposed a new waveform based on an adaptive length BC DFT-s-OFDM and shown that it significantly enhances EE of uplink transmissions in massive MIMO with OMA or PD-NOMA. To realize the performance improvements in practice, accurate CSI is needed at the BS. We have shown that this can be accomplished by integrating DPS/EDPS with BC DFT-s-OFDM. We have shown that these strategies can achieve an effective trade-off between SE, power consumption, and EE, where the enhancements in EE do not cause

a substantial degradation of SE. The numerical results have depicted the benefits of the proposed BC DFT-s-OFDM to achieve high EE, low NMSE with better CSI accuracy, low BER, and high sum-rate capacity at low SNRs for massive MIMO with OMA and PD-NOMA. Also, we have demonstrated that BC DFT-s-OFDM reduces PAPR to better preserve the signal's peak power at the UEs within the dynamic range enabling linear operation of the power amplifier to further prevent inter-carrier interference, intermodulation distortions, and high power losses. This contributes to a noticeable reduction in the power consumption at the UE, which supports green communications. However, these benefits come at a cost in changing the UEs structures by adding a BC spreader and the BS's receiver by adding a BC de-spreader, with a minimal increase in circuit power consumption. We have illustrated the necessary modification for the UE's transmitter and the BS's receiver structures.

- In Chapter 4, we have addressed the problem that SIC decoding results in poor performance when served users in NOMA pairs have relatively equal power levels in their received signals at the BS. As an effective solution, we have proposed novel decoding schemes based on CaF. The proposed schemes work in both PD-NOMA stand-alone and integrated with massive MIMO. We have shown that the proposed scheme provides better fairness and lowers outage probability without significantly increasing complexity compared to SIC decoding. We have found that the proposed decoding schemes is general and include SIC decoding as a special case.

Simulation results have shown that CaF-NOMA increases Jain's fairness index and decreases the average outage probability significantly compared to SIC. Furthermore, we have combined both CaF decoding mechanism and EDPS/DPS in massive MIMO with NOMA to promote practical implementation. This enhanced IFLR to tackle imperfect CSI, maintain the MAC rate region, and reduce the average sum-rate loss. Numerical results have confirmed that our approach adds appreciable robustness against CSI estimation errors, and significantly enhances the performance of IFLR in massive MIMO with PD-NOMA. Consequently, this paves the way for practical implementation of IFLR in massive MIMO with PD-NOMA.

However, CaF decoding adds a minimum cost compared with its benefits where an additional block needs to be added into the BS's receiver.

5.2 Future Works

In this section, we provide several interesting possibilities for expanding the research work presented in this dissertation. The main idea is to extend the applications of our proposed schemes and investigate their performance in other 5G technologies or in different network scenarios. We believe, based on our findings, that there are still many valuable ways in which the research work we have demonstrated here could be extended further. In what follows, we illustrate several potential future extensions by providing summarize research topics, that can be considered as enriching starts towards the first step.

5.2.1 Expected Future Extensions

The work in this dissertation has built on five pillars which are 1) DPS, 2) EDPS, 3) BC DFT-s-OFDM, 4) CaF, and 5) EE enhancement. The technical performance of these pillars has been investigated under various state-of-the-art cellular networks configurations of emerging 5G technologies which are 1) Massive MIMO with OMA, 2) NOMA, 3) Massive MIMO with NOMA. Therefore, there are still several paths that can be considered as natural extensions of the presented research work. Besides, the schemes we have developed show the attributes of additional research work and analyses that could be done either using the same approaches as we have illustrated or new envision methodologies. Further development can be realized by building on our schemes and either applying them into new combined models of novel technologies to acquire further enhancement or using them to devise efficient signal processing algorithms. Specifically, we are suggesting the following research topics:

- **Investigating the Improvement of SIC in NOMA by Applying DPS**

In Chapter 2, we have developed DPS which offers unique attributes in non-cooperative multiple adjacent cells configurations, and applied the technique throughout this dissertation to

enhance channel estimation and improve CSI accuracy. However, as we stated in Section 2.8, DPS can also be used in improving the performance of interference cancellation. Designing an effective SIC is a current research area in wireless cellular communications with still many open problems, especially in multi-cell environments. Such research direction can be both challenging and promising, due to the intensive in frequency reusing encountered in multi-cells environments at the next generations of cellular networks in 5G and beyond. Addressing the interference resulted from adjacent cells in SIC is very critical as it can degrade the performance of the overall system. However, one of the main opportunities that DPS is offering is to facilitate interference management and cancellation mechanisms with avoiding coordination between the adjacent cells, where entities' involvement is very challenging and adds overheads on SE. On the other hand, rather than exploiting the primal decomposition method in designing the precoder, DPS offers an alternative and easier mechanism to enable the BS to design the precoders for its own served users independently. Besides, DPS provides the BS accurate CSI, which can enhance interference alignment further, where accurate CSI is required to implement effective interference alignment and efficient beam-former. Therefore, an interesting natural extension is to investigate and examine the integration of DPS for improving the performance of SIC could be a future research topic.

- **Performance Analysis of DPS Using Superimposed Pilot Sequences**

In this dissertation, we assumed a configuration of multiplexed pilots is used, where part of the coherence interval in the uplink transmission is dedicated to PSs to be time-multiplexed with the data in one coherence time block. Thereby, we have evaluated the performance of the proposed DPS and EDPS under this configuration. However, another configuration is the superimposed pilots, where PSs are added at low power to that data symbols prior to the transmission. Figure 5.1 shows both configurations. Although this approach requires more computational power and increases signal processing complexity to separate the data symbols from pilots, it may lead to exploiting the dedicated part of pilots in the coherence interval to be used for actual data transmission which may improve the efficiency of resource utiliza-

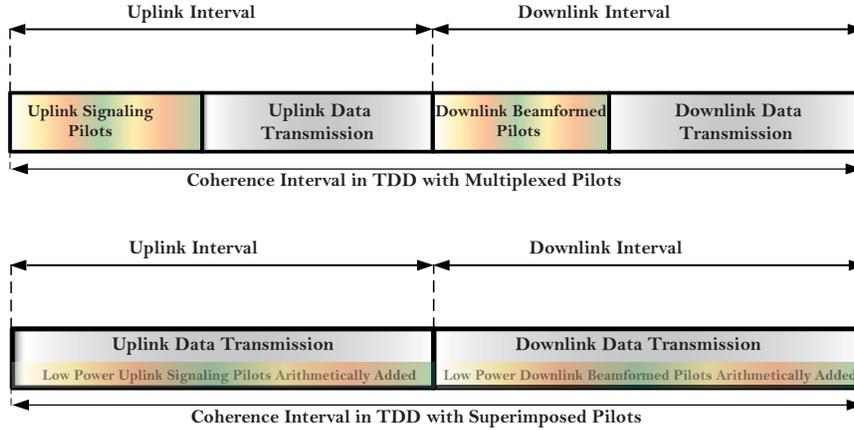


Figure 5.1: Superimposed and multiplexed pilots configurations in TDD.

tion. Therefore, exploring the performance of DPS and EDPS under the configuration of the superimposed pilots' approach and examine both OMA and NOMA massive MIMO systems performance can provide a new potential research direction that needs further investigations and developments. Moreover, given the existing variations of communication technologies in 5G and beyond, exploring the feasibility of a hybrid configuration that trades off multiplexed pilots against superimposed pilots to reach an optimal operating point that provides adaptive and elastic services on-demand while acquiring the advantages of both approaches is an interesting direction that has yet to be investigated.

- **Performance Analysis of Joint DPS and Machine Learning**

The recent trends of utilizing Machine Learning (ML) in many applications have gained attention from both industry and academia, which provides attractive solutions that should be explored more. Besides, various ML techniques have rekindled the interest in developing novel solutions for wireless communications and are considered an inevitable trend. One of the outcomes of this trend is to add smart capabilities to the BSs and evolve cellular networks. Given the success of the DPS in adding additional knowledge to the BS, we have applied it to improve CSI accuracy and eliminate pilot contamination as illustrated in Chapter 2. However, DPS can be exploited to add intelligence capability to the BS, which permits

the continuous updated extraction and characterization of large scale fading channel coefficients patterns, provides adaptive user scheduling, and provides better adaptable interference alignment. Several ML techniques can be exploited in this future research direction, including support vector machine, neural network, and deep learning. Therefore, exploring the performance of joint ML and DPS approach can leverage advantages of both DPS and ML, which is considered an interesting avenue for future research. One open problem is how to build this joint model while ensuring high efficacy.

- **Performance Analysis of Massive MIMO Serving UEs Equipped with Multiple-Antennas**

Throughout the entire dissertation, we took a conventional assumption that previous researchers have followed, which is that each UE is supplied with one antenna. It is a valid assumption but the scenario that considers UEs equipped with multiple antennae may be of interest in the future and has yet to be studied. Apparently, the advantages of deploying UEs with multiple antennas is twofold. First, it substantially enhances the data rate, and second, its feasibility to support full-duplex operation [139]. However, for both OMA and NOMA massive MIMO there are several challenges in the realization of this scenario. One of them is the increased complexity of system architecture and hence the complexity of channel estimation. A relevant future research direction is to exploit the schemes proposed in this dissertation to analyze the asymptotic effect of increasing the number of UEs antennas. Obtaining the optimal number of antennae that can be used by each UE to achieve maximum performance of the whole system is another factor to be considered. To this ends, a scenario that starts with UEs having two antennae and considers using Alamouti space-time block code [140] could be investigated.

Bibliography

- [1] F. Jejdling, P. Cerwall, A. Lundvall, P. Jonsson, S. Carson, R. Möller, S. Davis, P. Linder, A. Gomroki, A. Zaidi, A. Carlsson P, M. Opsenica, I. Sorlie, S. Elmgren, G. Blennerud, H. Baur, R. Svenningsson, B. Heath, J. Bugel, S. John, and S. Schwartz, “Ericsson Mobility Report,” *Ericsson Reports*, vol. Q4, pp. 1–36, November 2020. [Online]. Available: <https://www.ericsson.com/en/mobility-report/reports>
- [2] S. Mattisson, “An overview of 5G requirements and future wireless networks: Accommodating scaling technology,” *IEEE Solid-State Circuits Magazine*, vol. 10, no. 3, pp. 54–60, August 2018.
- [3] A. A. Ateya, A. Muthanna, M. Makolkina, and A. Koucheryavy, “Study of 5G services standardization: Specifications and requirements,” in *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Moscow, Russia, 5-9 Nov. 2018, pp. 1–6.
- [4] W. Xiang, K. Zheng, and X. S. Shen, *5G Mobile Communications*. Springer, 2016.
- [5] A. Osseiran, J. F. Monserrat, and P. Marsch, *5G Mobile and Wireless Communications Technology*. Cambridge University Press, 2016.
- [6] B. Bangerter, S. Talwar, R. Arefi, and K. Stewart, “Networks and devices for the 5G era,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 90–96, February 2014.
- [7] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.
- [8] K. Senel, H. V. Cheng, E. Björnson, and E. G. Larsson, “What role can NOMA play in Massive MIMO?” *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 597–611, June 2019.
- [9] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links?” *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, April 2003.
- [10] K. Mawatwal, D. Sen, and R. Roy, “A semi-blind channel estimation algorithm for Massive MIMO systems,” *IEEE Wireless Communications Letters*, vol. 6, no. 1, pp. 70–73, February 2017.

- [11] T. Weber, A. Sklavos, and M. Meurer, "Imperfect channel-state information in MIMO transmission," *IEEE Transactions on Communications*, vol. 54, no. 3, pp. 543–552, March 2006.
- [12] C. Wen, S. Jin, K. Wong, J. Chen, and P. Ting, "Channel estimation for Massive MIMO using Gaussian-mixture bayesian learning," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1356–1368, March 2015.
- [13] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, "Massive MIMO performance with imperfect channel reciprocity and channel estimation error," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3734–3749, September 2017.
- [14] N. I. Miridakis and T. A. Tsiftsis, "On the joint impact of hardware impairments and imperfect CSI on successive decoding," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 4810–4822, June 2017.
- [15] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, November 2014.
- [16] T. T. Do, T. J. Oechtering, S. M. Kim, M. Skoglund, and G. Peters, "Uplink waveform channel with imperfect channel state information and finite constellation input," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1107–1119, February 2017.
- [17] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, August 2011.
- [18] Taesang Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.
- [19] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [20] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.
- [21] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [22] W. A. Al-Hussaibi and F. H. Ali, "A closed-form approximation of correlated multiuser MIMO ergodic capacity with antenna selection and imperfect channel estimation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5515–5519, June 2018.

- [23] A. Bourdoux, C. Desset, L. van der Perre, G. Dahman, O. Edfors, J. Flordelis, X. Gao, C. Gustafson, F. Tufvesson, F. Harrysson, and J. Medbo, "D1.2 Massive MIMO for efficient transmission," *Project of MaMi Channel Characteristics : Measurement Results*, no. 619086, pp. 1–93, June 2015.
- [24] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, January 2013.
- [25] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *IEEE 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, 28-30 Sep. 2011, pp. 545–550.
- [26] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, February 2016.
- [27] X. Liu, Y. Liu, X. Wang, and H. Lin, "Highly efficient 3-D resource allocation techniques in 5G for NOMA-enabled Massive MIMO and relaying systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2785–2797, December 2017.
- [28] Z. Ding and H. V. Poor, "Design of Massive-MIMO-NOMA with limited feedback," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 629–633, May 2016.
- [29] A. Sousa de Sena, D. Benevides da Costa, Z. Ding, and P. H. J. Nardelli, "Massive MIMO–NOMA networks with multi-polarized antennas," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5630–5642, December 2019.
- [30] Y. Liu, G. Pan, H. Zhang, and M. Song, "On the capacity comparison between MIMO-NOMA and MIMO-OMA," *IEEE Access*, vol. 4, pp. 2123–2129, May 2016.
- [31] F. Fang, "Energy efficient resource allocation for non-orthogonal multiple access (NOMA) systems," Ph.D. dissertation, University of British Columbia, December 2017. [Online]. Available: <https://open.library.ubc.ca/collections/ubctheses/24/items/1.0362411>
- [32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.
- [33] F. Khan, *LTE for 4G Mobile Broadband: Air Interface Technologies and Performance*. Cambridge University Press, 2009.
- [34] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, February 2016.
- [35] I. Atzeni, J. Arnau, and M. Debbah, "Fractional pilot reuse in Massive MIMO systems," in *IEEE International Conference on Communication Workshop (ICC Workshop)*, London, UK, 8-12 June 2015, pp. 1030–1035.

- [36] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Massive MIMO transmission with pilot reuse in single cell," in *IEEE International Conference on Communications (ICC)*, Sydney, NSW, Australia, 10-14 June 2014, pp. 4783–4788.
- [37] P. Xu, J. Wang, and J. Wang, "Effect of pilot contamination on channel estimation in Massive MIMO systems," in *IEEE International Conference on Wireless Communications & Signal Processing (WCSP)*, Hangzhou, China, 24-26 Oct. 2013, pp. 1–6.
- [38] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, 22-27 May 2011, pp. 3464–3467.
- [39] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in Massive MIMO - 5G system," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 905–923, Secondquarter 2016.
- [40] O. Elijah, C. Y. Leow, A. R. Tharek, S. Nunoo, and S. Z. Iliya, "Mitigating pilot contamination in Massive MIMO system-5G: An overview," in *IEEE 10th Asian Control Conference (ASCC)*. Kota Kinabalu, Malaysia, 31 May - 3 June 2015, pp. 1–6.
- [41] M. A. L. Sarker and M. H. Lee, "A fast channel estimation and the reduction of pilot contamination problem for Massive MIMO based on a diagonal Jacket matrix," in *IEEE 4th International Workshop on Fiber Optics in Access Network (FOAN)*, Almaty, Kazakhstan, 11-12 Sep. 2013, pp. 26–30.
- [42] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773–786, October 2014.
- [43] V. Nguyen, H. V. Nguyen, and O. Shin, "Channel estimation based on time-shifted pilots in multicell Massive MIMO systems," in *IEEE International Conference on Communications, Management and Telecommunications (ComManTel)*, DaNang, Vietnam, 28-30 Dec. 2015, pp. 83–87.
- [44] S. Jin, X. Wang, Z. Li, K.-K. Wong, Y. Huang, and X. Tang, "On Massive MIMO zero-forcing transceiver using time-shifted pilots," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 59–74, January 2016.
- [45] V. Saxena, G. Fodor, and E. Karipidis, "Mitigating pilot contamination by pilot reuse and power control schemes for Massive MIMO systems," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, Glasgow, UK, 11-14 May 2015, pp. 1–6.
- [46] T. Lee, H.-S. Kim, S. Park, and S. Bahk, "Mitigation of sounding pilot contamination in Massive MIMO systems," in *IEEE International Conference on Communications (ICC)*, Sydney, NSW, Australia, 10-14 June 2014, pp. 1191–1196.
- [47] J. Zhang, B. Zhang, S. Chen, X. Mu, M. El-Hajjar, and L. Hanzo, "Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 759–772, October 2014.

- [48] A. Farhang, A. Aminjavaheri, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Pilot decontamination in CMT-based Massive MIMO networks," in *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, Barcelona, Spain, 26-29 Aug. 2014, pp. 589–593.
- [49] J. H. Sørensen and E. De Carvalho, "Pilot decontamination through pilot sequence hopping in Massive MIMO systems," in *IEEE Global Communications Conference (GLOBECOM)*, Austin, TX, USA, 8-12 Dec. 2014, pp. 3285–3290.
- [50] D. Hu, L. He, and X. Wang, "Semi-blind pilot decontamination for Massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 525–536, January 2016.
- [51] X. Zhu, Z. Wang, L. Dai, and C. Qian, "Smart pilot assignment for Massive MIMO," *IEEE Communications Letters*, vol. 19, no. 9, pp. 1644–1647, September 2015.
- [52] T. X. Vu, T. A. Vu, and T. Q. Quek, "Successive pilot contamination elimination in multi-antenna multicell networks," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 617–620, December 2014.
- [53] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 25-30 March 2012, pp. 3249–3252.
- [54] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, February 2013.
- [55] P. Li, Z. Ding, and K. Feng, "Enhanced receiver based on FEC code constraints for uplink NOMA with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4790–4802, October 2019.
- [56] J. Cui, Z. Ding, and P. Fan, "Outage probability constrained MIMO-NOMA designs under imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8239–8255, December 2018.
- [57] Y. Gao, B. Xia, K. Xiao, Z. Chen, X. Li, and S. Zhang, "Theoretical analysis of the dynamic decode ordering SIC receiver for uplink NOMA systems," *IEEE Communications Letters*, vol. 21, no. 10, pp. 2246–2249, October 2017.
- [58] Y. Gao, B. Xia, Y. Liu, Y. Yao, K. Xiao, and G. Lu, "Analysis of the dynamic ordered decoding for uplink NOMA systems with imperfect CSI," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6647–6651, July 2018.
- [59] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, October 2011.

- [60] C. Feng, D. Silva, and F. R. Kschischang, "An Algebraic approach to physical-layer network coding," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7576–7596, November 2013.
- [61] J. Zhan, B. Nazer, M. Gastpar, and U. Erez, "MIMO compute-and-forward," in *IEEE International Symposium on Information Theory*, Seoul, South Korea, 28 June-3 July 2009, pp. 2848–2852.
- [62] U. Niesen and P. Whiting, "The degrees of freedom of compute-and-forward," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5214–5232, August 2012.
- [63] M. Hejazi and M. Nasiri-Kenari, "Simplified compute-and-forward and its performance analysis," *IET Communications*, vol. 7, no. 18, pp. 2054–2063, December 2013.
- [64] E. Sula, J. Zhu, A. Pastore, S. H. Lim, and M. Gastpar, "Compute-forward multiple access (CFMA): Practical implementations," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1133–1147, February 2019.
- [65] K. N. Pappi, G. K. Karagiannidis, and R. Schober, "How sensitive is compute-and-forward to channel estimation errors?" in *IEEE International Symposium on Information Theory*, Istanbul, Turkey, 7-12 July 2013, pp. 3110–3114.
- [66] C. Feng, D. Silva, and F. R. Kschischang, "Blind compute-and-forward," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1451–1463, April 2016.
- [67] H. Ghazzai, E. Yaacoub, A. Kadri, H. Yanikomeroğlu, and M. Alouini, "Next-Generation environment-aware cellular networks: Modern green techniques and implementation challenges," *IEEE Access*, vol. 4, pp. 5010–5029, September 2016.
- [68] P. Patcharamaneepakorn, S. Wu, C. Wang, e. M. Aggoune, M. M. Alwakeel, X. Ge, and M. D. Renzo, "Spectral, energy, and economic efficiency of 5G multicell Massive MIMO systems with generalized spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9715–9731, December 2016.
- [69] W. Cheng, X. Zhang, and H. Zhang, "Statistical-QoS driven energy-efficiency optimization over green 5G mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3092–3107, December 2016.
- [70] T. Liu, J. Tong, Q. Guo, J. Xi, Y. Yu, and Z. Xiao, "Energy efficiency of Massive MIMO systems with low-resolution ADCs and successive interference cancellation," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3987–4002, August 2019.
- [71] R. Zi, X. Ge, J. Thompson, C. Wang, H. Wang, and T. Han, "Energy efficiency optimization of 5G radio frequency chain systems," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 758–771, April 2016.
- [72] M. M. Mowla, I. Ahmad, D. Habibi, and Q. V. Phung, "A green communication model for 5G systems," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 3, pp. 264–280, September 2017.

- [73] H. Çelebi, Y. Yapıcı, I. Güvenç, and H. Schulzrinne, "Load-based On/Off scheduling for energy-efficient delay-tolerant 5G networks," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 4, pp. 955–970, December 2019.
- [74] D. Jiang, P. Zhang, Z. Lv, and H. Song, "Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1437–1447, December 2016.
- [75] S. Rostami, K. Heiska, O. Puchko, K. Leppanen, and M. Valkama, "Pre-Grant signaling for energy-efficient 5G and beyond mobile devices: Method and analysis," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 2, pp. 418–432, June 2019.
- [76] Y. Al-Dunainawi, R. S. Alhumaima, and H. S. Al-Raweshidy, "Green network costs of 5G and beyond, expectations vs reality," *IEEE Access*, vol. 6, pp. 60 206–60 213, October 2018.
- [77] A. N. Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, "Eliminating pilot contamination using Dual Pilot Sequences in Massive MIMO," in *IEEE 86th Vehicular Technology Conference VTC-Fall*, Toronto, ON, Canada, 24-27 Sept. 2017, pp. 1–6.
- [78] A. N. Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, "Dual Pilot Scheme (DPS) and its application in Massive MIMO," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1431–1444, November 2020.
- [79] A. N. Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, "Improving energy efficiency of 5G uplinks using Barker codes in DFT-s-OFDM," *Submitted to IEEE Transactions on Signal Processing*, pp. 1–15, December 2020.
- [80] A. N. Aljalai, C. Feng, V. C. M. Leung, and R. K. Ward, "Improving the energy efficiency of DFT-s-OFDM in uplink Massive MIMO with Barker codes," in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, Big Island, Hawaii, USA, 17-20 Feb. 2020, pp. 731–735.
- [81] C. Guo, A. N. Aljalai, C. Feng, L. Zhao, V. C. M. Leung, and R. K. Ward, "Compute-and-Forward for uplink non-orthogonal multiple access," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 986–989, December 2018.
- [82] A. N. Aljalai, C. Feng, V. Leung, and R. Ward, "Integer-Forcing receiver for uplink massive MIMO-NOMA with imperfect CSI," *To be submitted*, pp. 1–12.
- [83] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-hill New York, 2007.
- [84] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, May 2014.

- [85] F. A. de Figueiredo, F. A. Cardoso, I. Moerman, and G. Fraidenraich, "Channel estimation for Massive MIMO TDD systems assuming pilot contamination and frequency selective fading," *IEEE Access*, vol. 5, pp. 17 733–17 741, September 2017.
- [86] F. A. de Figueiredo, F. A. Cardoso, I. Moerman, and G. Fraidenraich, "Channel estimation for Massive MIMO TDD systems assuming pilot contamination and flat fading," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 14, pp. 1–10, January 2018.
- [87] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, February 2013.
- [88] L. Wu, Z. Zhang, J. Dang, J. Wang, H. Liu, and Y. Wu, "Channel estimation for multicell multiuser Massive MIMO uplink over Rician fading channels," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 8872–8882, October 2017.
- [89] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD Massive MIMO systems with spatial basis expansion model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, April 2017.
- [90] J. Ma, S. Zhang, H. Li, F. Gao, and S. Jin, "Sparse bayesian learning for the time-varying Massive MIMO channels: Acquisition and tracking," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 1925–1938, March 2019.
- [91] J. Ma, S. Zhang, H. Li, N. Zhao, and V. C. M. Leung, "Interference-alignment and soft-space-reuse based cooperative transmission for multi-cell Massive MIMO networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1907–1922, March 2018.
- [92] R. F. Ustok, P. A. Dmochowski, P. J. Smith, and M. Shafi, "Interference cancellation with jointly optimized transceivers in multiuser multicellular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7219–7229, August 2018.
- [93] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall PTR, 1993.
- [94] H. Q. Ngo, *Massive MIMO: Fundamentals and System Designs*. Linköping University Electronic Press, 2015, vol. 1642.
- [95] A. Assalini, E. Dall'Anese, and S. Pupolin, "Linear MMSE MIMO channel estimation with imperfect channel covariance information," in *IEEE International Conference on Communications (ICC)*, Dresden, Germany, 14-18 June 2009, pp. 1–5.
- [96] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology and Techniques*, 2nd ed. Springer Science & Business Media, 2000.

- [97] L. Zhao, K. Li, K. Zheng, and M. O. Ahmad, "An analysis of the tradeoff between the energy and spectrum efficiencies in an uplink Massive MIMO-OFDM system," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 3, pp. 291–295, March 2015.
- [98] I. Baig, U. Farooq, N. U. Hasan, M. Zghaibeh, U. M. Rana, M. Imran, and M. Ayaz, "On the PAPR reduction: A novel filtering based Hadamard transform precoded uplink MC-NOMA scheme for 5G cellular networks," in *IEEE 1st International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia, 4-6 April 2018, pp. 1–4.
- [99] G. Berardinelli, K. I. Pedersen, T. B. Sorensen, and P. Mogensen, "Generalized DFT-Spread-OFDM as 5G waveform," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 99–105, November 2016.
- [100] A. Sahin, R. Yang, E. Bala, M. C. Beluri, and R. L. Olesen, "Flexible DFT-S-OFDM: Solutions and challenges," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 106–112, November 2016.
- [101] T. Nakamura, A. Benjebbour, Y. Kishiyama, S. Suyama, and T. Imai, "5G Radio access: Requirements, concept and experimental trials," *IEICE Transactions on Communications*, vol. 98, no. 8, pp. 1397–1406, August 2015.
- [102] A. Farhang, N. Marchetti, F. Figueiredo, and J. P. Miranda, "Massive MIMO and waveform design for 5th generation wireless communication systems," in *IEEE 1st International Conference on 5G for Ubiquitous Connectivity*, Akaslompolo, Finland, 26-28 Nov. 2014, pp. 70–75.
- [103] M. Aldababsa, M. Toka, S. Gökçeli, G. K. Kurt, and O. Kucur, "A tutorial on nonorthogonal multiple access for 5G and beyond," *Wireless Communications and Mobile Computing*, vol. 2018, pp. 1–24, June 2018.
- [104] H. V. Cheng, E. Björnson, and E. G. Larsson, "Performance analysis of NOMA in training-based multiuser MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 372–385, October 2017.
- [105] Z. Zhang, H. Sun, and R. Q. Hu, "Downlink and uplink non-orthogonal multiple access in a dense wireless network," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2771–2784, July 2017.
- [106] P. Wu, J. Zeng, X. Su, H. Gao, and T. Lv, "On energy efficiency optimization in downlink MIMO-NOMA," in *IEEE International Conference on Communications Workshops (ICC Workshops)*, Paris, France, 21-25 May 2017, pp. 399–404.
- [107] M. Series, "IMT vision–framework and overall objectives of the future development of IMT for 2020 and beyond," *Recommendation ITU-R M.2083-0*, pp. 1–21, September 2015.

- [108] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28–35, December 2011.
- [109] M. W. Chia, B. S. Thian, and T. T. Tjhung, "Distributed DFT-Spread OFDM," in *IEEE 10th Singapore International Conference on Communication Systems*, Singapore, Singapore, 30 Oct.- 1 Nov. 2006, pp. 1–5.
- [110] G. Berardinelli, F. M. L. Tavares, T. B. Sørensen, P. Mogensen, and K. Pajukoski, "Zero-tail DFT-spread-OFDM signals," in *2013 IEEE Globecom Workshops (GC Wkshps)*, 2013, pp. 229–234.
- [111] A. Sahin, R. Yang, M. Ghosh, and R. L. Olesen, "An improved unique word DFT-Spread OFDM scheme for 5G systems," in *IEEE Global Communications Conference Workshops (GLOBECOM Workshop)*, San Diego, CA, USA, 6-10 Dec. 2015, pp. 1–6.
- [112] K. S. Ali, H. Elsayy, A. Chaaban, and M. Alouini, "Non-orthogonal multiple access for large-scale 5G networks: Interference aware design," *IEEE Access*, vol. 5, pp. 21 204–21 216, September 2017.
- [113] R. S. Prabhu and B. Daneshrad, "Energy-efficient power loading for a MIMO-SVD system and its performance in flat fading," in *IEEE Global Communications Conference (GLOBECOM)*, Miami, FL, USA, 6-10 Dec. 2010, pp. 1–5.
- [114] Y. Zhang, Y. Xu, Y. Sun, Q. Wu, and K. Yao, "Energy efficiency of small cell networks: Metrics, methods and market," *IEEE Access*, vol. 5, pp. 5965–5971, April 2017.
- [115] B. Bedzhev, Z. Tasheva, and B. Stoyanov, "A method for synthesis of generalized Barker Codes," in *XL International Scientific Conference on Information, Communication and Energy Systems and Technologies Icest*, Serbia and Montenegro, 29 June- 1 July 2005, pp. 23–26.
- [116] A. Wang, Y. Qiu, and L. Lin, "A low complexity channel estimation algorithm based on Barker Code for IEEE 802.11," in *IEEE International Colloquium on Computing, Communication, Control, and Management (ISECS-CCCM)*, vol. 2, Guangzhou, China, 3-4 Aug. 2008, pp. 416–420.
- [117] J. Jedwab, "What can be used instead of a Barker sequence?" *Contemporary Mathematics 461*, American Mathematical Society, pp. 153–178, December 2008.
- [118] S. Golomb and R. Scholtz, "Generalized Barker sequences," *IEEE Transactions on Information Theory*, vol. 11, no. 4, pp. 533–537, October 1965.
- [119] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free Massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, March 2017.
- [120] A. Sahin, E. Bala, R. Yang, and R. L. Olesen, "DFT-Spread OFDM with frequency domain reference symbols," in *IEEE Global Communications Conference (GLOBECOM)*, Singapore, Singapore, 4-8 Dec. 2017, pp. 1–6.

- [121] B. Gupta, G. Gupta, and D. S. Saini, "BER Performance improvement in OFDM system with ZFE and MMSE equalizers," in *IEEE 3rd International Conference on Electronics Computer Technology*, vol. 6, Kanyakumari, India, 8-10 April 2011, pp. 193–197.
- [122] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, Barcelona, Spain, 26-29 Aug. 2014, pp. 781–785.
- [123] K. Zheng, S. Ou, and X. Yin, "Massive MIMO channel models: A survey," *International Journal of Antennas and Propagation*, vol. 2014, pp. 1–10, June 2014.
- [124] M. Abdulla and Y. R. Shayan, "Large-scale fading behavior for a cellular network with uniform spatial distribution," *Wireless Communications and Mobile Computing*, vol. 16, no. 7, pp. 748–764, May 2016.
- [125] Y. Huang, S. He, J. Wang, and J. Zhu, "Spectral and energy efficiency tradeoff for Massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 6991–7002, April 2018.
- [126] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. ElKashlan, C. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, February 2017.
- [127] 3GPP, "Study on downlink multiuser superposition transmission for LTE," in *3rd Generation Partnership Project Release 13*, no. 36.859, Sitges, Spain, 7-10 Dec. 2015, pp. 1–48.
- [128] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, Secondquarter 2017.
- [129] H. Pan, L. Lu, and S. C. Liew, "Practical power-balanced non-orthogonal multiple access," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2312–2327, October 2017.
- [130] B. Nazer, V. R. Cadambe, V. Ntranos, and G. Caire, "Expanding the Compute-and-Forward Framework: Unequal powers, signal levels, and multiple linear combinations," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 4879–4909, September 2016.
- [131] N. di Pietro, G. Zemor, and J. J. Boutros, "LDA lattices without dithering achieve capacity on the gaussian channel," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1561–1594, March 2018.
- [132] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-Forcing linear receivers," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7661–7685, December 2014.
- [133] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

- [134] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *IEEE Global Communications Conference (GLOBECOM)*, Taipei, Taiwan, 17-21 Nov. 2002, pp. 424–428.
- [135] P. Q. Nguyen and D. Stehlé, "Low-dimensional lattice basis reduction revisited," *ACM Transactions on algorithms (TALG)*, vol. 5, no. 4, pp. 1–48, October 2009.
- [136] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *ACM Transaction on Computer Systems, Tech. Rep. 301, Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA, USA*, pp. 1–37, September 1984.
- [137] S. Vatedka and N. Kashyap, "Some "Goodness" properties of LDA lattices," *Problems of Information Transmission*, vol. 53, no. 1, pp. 3–33, April 2017.
- [138] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed. John Wiley & sons Inc., 2004.
- [139] D. Bharadia, E. McMillin, and S. Katti, "Full duplex radios," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 375–386, August 2013.
- [140] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on selected areas in communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.
- [141] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory*. Cambridge University Press, 2014.
- [142] F. R. Kschischang and C. Feng, "An introduction to lattices and their applications in communications," 2014. [Online]. Available: <https://www.itsoc.org/australian-school-2014/pictures/Frank.pdf>

Appendix A

Related Proofs of Chapter 2

A.1 Derivation of Equation (2.8)

It suffices to prove that $\phi_{lk}^H \mathbf{D}_l^{-1} \phi_{lk} = \frac{P}{\sigma^2 + P \sum_i \beta_{ik}^l}$ when $\phi_{lk} = \sqrt{P} \psi_k$ for all k . First, note that

$$\phi_{lk}^H \mathbf{D}_l^{-1} \phi_{lk} = P \psi_k^H \mathbf{D}_l^{-1} \psi_k. \quad (\text{A.1})$$

Hence, it suffices to show that $\psi_k^H \mathbf{D}_l^{-1} \psi_k = \frac{1}{\sigma^2 + P \sum_i \beta_{ik}^l}$. Second, note that

$$\mathbf{D}_l = \sigma^2 \mathbf{I}_\tau + \sum_{(i,j)} P \beta_{ij}^l \psi_j \psi_j^H \quad (\text{A.2})$$

$$= \sigma^2 \mathbf{I}_\tau + \sum_j \sum_i P \beta_{ij}^l \psi_j \psi_j^H \quad (\text{A.3})$$

$$= \sigma^2 \mathbf{I}_\tau + \sum_j \beta_j \psi_j \psi_j^H \quad (\text{A.4})$$

where $\beta_j = P \sum_i \beta_{ij}^l$ is a short-hand notation. Hence, it suffices to show that

$$\psi_k^H \left(\sigma^2 \mathbf{I}_\tau + \sum_j \beta_j \psi_j \psi_j^H \right)^{-1} \psi_k = \frac{1}{\sigma^2 + \beta_k}. \quad (\text{A.5})$$

By the Sherman-Morrison formula, we have

$$\left(\sigma^2 \mathbf{I}_\tau + \sum_j \beta_j \psi_j \psi_j^H \right)^{-1} = \frac{1}{\sigma^2} \left(\mathbf{I}_\tau - \sum_j \frac{\beta_j}{\sigma^2 + \beta_j} \psi_j \psi_j^H \right). \quad (\text{A.6})$$

It follows that $\psi_k^H \left(\sigma^2 \mathbf{I}_\tau + \sum_j \beta_j \psi_j \psi_j^H \right)^{-1} \psi_k = \frac{1}{\sigma^2 + \beta_k}$. This proves (A.5).

A.2 Derivation of Inequality (2.24)

First, we have

$$(\mathbf{D}_l^{\text{SUB}}(\gamma))^{-1} = \frac{1}{\Delta_l^{\text{SUB}}(\gamma)} \begin{bmatrix} (1-\gamma)P\sum_j \beta_{lj}^l + \sigma^2 & -\sqrt{\gamma(1-\gamma)}P\beta_{lk}^l \\ -\sqrt{\gamma(1-\gamma)}P\beta_{lk}^l & \gamma P\sum_i \beta_{ik}^l + \sigma^2 \end{bmatrix} \quad (\text{A.7})$$

where $\Delta_l^{\text{SUB}}(\gamma)$ is the determinant of $\mathbf{D}_l^{\text{SUB}}(\gamma)$. Hence, we have

$$\begin{aligned} & \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}^T (\mathbf{D}_l^{\text{SUB}}(\gamma))^{-1} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix} \\ &= \frac{1}{\Delta_l^{\text{SUB}}(\gamma)} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix}^T \begin{bmatrix} (1-\gamma)P\sum_j \beta_{lj}^l + \sigma^2 & -\sqrt{\gamma(1-\gamma)}P\beta_{lk}^l \\ -\sqrt{\gamma(1-\gamma)}P\beta_{lk}^l & \gamma P\sum_i \beta_{ik}^l + \sigma^2 \end{bmatrix} \begin{bmatrix} \sqrt{\gamma} \\ \sqrt{1-\gamma} \end{bmatrix} \\ &= \frac{1}{\Delta_l^{\text{SUB}}(\gamma)} \left(\gamma(1-\gamma)P\sum_{j \neq k} \beta_{lj}^l + \gamma(1-\gamma)P\sum_{i \neq l} \beta_{ik}^l + \sigma^2 \right) \\ &= \frac{1}{\Delta_l^{\text{SUB}}(\gamma)} \left(\gamma(1-\gamma)P \left(\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l \right) + \sigma^2 \right). \end{aligned} \quad (\text{A.8})$$

Next, we would like to show that

$$\frac{\gamma(1-\gamma)P(\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l) + \sigma^2}{\Delta_l^{\text{SUB}}(\gamma)} > \frac{1}{\sigma^2 + P\sum_i \beta_{ik}^l} \quad (\text{A.9})$$

or equivalently,

$$\left(\sigma^2 + \gamma(1-\gamma)P \left(\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l \right) \right) \left(\sigma^2 + P \sum_i \beta_{ik}^l \right) > \Delta_l^{\text{SUB}}(\gamma) \quad (\text{A.10})$$

under the condition $\frac{1-\gamma}{\gamma} \frac{\sum_{j \neq k} \beta_{lj}^l}{(\sum_{i \neq l} \beta_{ik}^l)^2} < \frac{P}{\sigma^2}$. To see this, we separate terms with and without σ^2 in the above inequality, obtaining

$$\gamma(1-\gamma)P^2 \left(\sum_{i \neq l} \beta_{ik}^l \right)^2 > (1-\gamma)P \left(\sum_{j \neq k} \beta_{lj}^l - \sum_{i \neq l} \beta_{ik}^l - \gamma \left(\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l \right) \right) \sigma^2, \quad (\text{A.11})$$

if $\frac{1-\gamma}{\gamma} \frac{\sum_{j \neq k} \beta_{lj}^l}{(\sum_{i \neq l} \beta_{ik}^l)^2} < \frac{P}{\sigma^2}$.

We have two cases as follows.

1. $\sum_{i \neq l} \beta_{ik}^l \geq \frac{1-\gamma}{1+\gamma} \sum_{j \neq k} \beta_{lj}^l$. In this case, the above inequality always holds because the right-hand-side is no greater than 0.
2. $\sum_{i \neq l} \beta_{ik}^l < \frac{1-\gamma}{1+\gamma} \sum_{j \neq k} \beta_{lj}^l$. In this case, we need the following inequality

$$\sigma^2 < \frac{\gamma P (\sum_{i \neq l} \beta_{ik}^l)^2}{(1-\gamma) \sum_{j \neq k} \beta_{lj}^l - (1+\gamma) \sum_{i \neq l} \beta_{ik}^l}, \quad (\text{A.12})$$

which holds as long as $\sigma^2 < P \frac{\gamma}{1-\gamma} \frac{(\sum_{i \neq l} \beta_{ik}^l)^2}{\sum_{j \neq k} \beta_{lj}^l}$ or equivalently $\frac{1-\gamma}{\gamma} \frac{\sum_{j \neq k} \beta_{lj}^l}{(\sum_{i \neq l} \beta_{ik}^l)^2} < \frac{P}{\sigma^2}$.

A.3 Derivative of $g_{lk}^{\text{SUB}}(\gamma)$

Let $g'(\gamma)$ denote the derivative of $g_{lk}^{\text{SUB}}(\gamma)$. Then, we can show that

$$g'(\gamma) = h(\gamma)(A\gamma^2 + B\gamma + C) \quad (\text{A.13})$$

where $h(\gamma) > 0$ for all $\gamma \in (0, 1)$ and

$$A = \left(\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l \right) \left(\sum_{j \neq k} \beta_{lj}^l - \sum_{i \neq l} \beta_{ik}^l \right) \sigma^2 \quad (\text{A.14})$$

$$B = -2 \left(\left(\sum_{j \neq k} \beta_{lj}^l \right)^2 + \left(\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l \right) \sigma^2 \right) \sigma^2 \quad (\text{A.15})$$

$$C = \left(\left(\sum_{j \neq k} \beta_{lj}^l \right)^2 + 2 \left(\sum_{j \neq k} \beta_{lj}^l \right) \sigma^2 \right) \sigma^2 \quad (\text{A.16})$$

Hence, we have $g'(0) > 0$ (since $h(0) > 0$ and $C > 0$) and $g'(1) < 0$ (since $h(1) > 0$ and $A + B + C < 0$). Indeed, we can show that $g'(\gamma)$ has two zeros $\gamma_1 = \frac{\sum_{j \neq k} \beta_{lj}^l}{\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l}$ and $\gamma_2 = \frac{\sum_{j \neq k} \beta_{lj}^l + 2\sigma^2}{\sum_{j \neq k} \beta_{lj}^l - \sum_{i \neq l} \beta_{ik}^l}$.

Note that $\gamma_1 \in (0, 1)$ and $\gamma_2 > 1$. Hence, we have $g'(\gamma) > 0$ when $\gamma \in (0, \gamma_1)$ and $g'(\gamma) < 0$ when $\gamma \in (\gamma_1, 1)$. This implies that $g(\gamma)$ is unimodal over the interval $(0, 1)$ and the optimal γ is given by

$$\gamma^* = \frac{\sum_{j \neq k} \beta_{lj}^l}{\sum_{j \neq k} \beta_{lj}^l + \sum_{i \neq l} \beta_{ik}^l}, \quad (\text{A.17})$$

which does not depend on σ^2 .

Appendix B

Related Proofs and Preliminaries of Chapter 4

B.1 Proof of Theorem 4

In this appendix, we will prove Theorem 4 through two steps. First, we will show that the rate pairs induced by $([1,0], [0,1])$, $([0,1], [1,0])$, and $(\mathbf{a}_*, \mathbf{b}_*)$ are indeed Pareto-optimal solutions. Second, we will show that any Pareto-optimal solution is induced by $([1,0], [0,1])$, $([0,1], [1,0])$, or $(\mathbf{a}_*, \mathbf{b}_*)$.

Before conducting these two steps, we need to introduce some properties for $(\mathbf{a}_*, \mathbf{b}_*)$ proven in [130].

Property 1: The matrix $\begin{bmatrix} \mathbf{a}_* \\ \mathbf{b}_* \end{bmatrix}$ is invertible over $\mathbb{Z}[i]$. In particular, the greatest common divisor $\gcd(a_{1*}, a_{2*}) = 1$.

Property 2: \mathbf{a}_* minimizes $\mathbf{a} (\mathbf{P}^{-1} + \mathbf{h}^H \mathbf{h})^{-1} \mathbf{a}^H$ where $\mathbf{a} \neq \mathbf{0}$;

Property 3: \mathbf{b}_* minimizes $\mathbf{a} (\mathbf{P}^{-1} + \mathbf{h}^H \mathbf{h})^{-1} \mathbf{a}^H$ where \mathbf{a} is linearly independent of \mathbf{a}_* .

Step 1 of the proof. When the coefficient vector $(\mathbf{a}, \mathbf{b}) = ([1,0], [0,1])$, the induced rate pair is given by (4.2), which is a Pareto-optimal solution because it is a “corner point” of the capacity region. Similarly, the induced rate pair by $([0,1], [1,0])$ is also a Pareto-optimal solution.

Next, we consider the rate pair induced by $(\mathbf{a}_*, \mathbf{b}_*)$. By Property 1, we have the following three

cases.

1. $\mathbf{a}_* \in \{[1, 0], [-1, 0], [i, 0], [-i, 0]\}$. In this case, the induced rate pair is given by (4.2), which is a Pareto-optimal solution.
2. $\mathbf{a}_* \in \{[0, 1], [0, -1], [0, i], [0, -i]\}$. Similarly, the induced rate pair is Pareto optimal.
3. $a_{1*} \neq 0$ and $a_{2*} \neq 0$. In this case, the induced rate pair is given by

$$\begin{cases} R_1 = \min \left\{ \log_2 \left(\frac{p_1}{\sigma^2(\mathbf{a}_*)} \right), \log_2 \left(\frac{p_1}{\sigma^2(\mathbf{b}_*|\mathbf{a}_*)} \right) \right\} \\ R_2 = \log_2 \left(\frac{p_2}{\sigma^2(\mathbf{a}_*)} \right) \end{cases} \quad (\text{B.1})$$

or

$$\begin{cases} R_1 = \log_2 \left(\frac{p_1}{\sigma^2(\mathbf{a}_*)} \right), \\ R_2 = \min \left\{ \log_2 \left(\frac{p_2}{\sigma^2(\mathbf{a}_*)} \right), \log_2 \left(\frac{p_2}{\sigma^2(\mathbf{b}_*|\mathbf{a}_*)} \right) \right\} \end{cases} \quad (\text{B.2})$$

where $\sigma^2(\mathbf{a}_*) = \mathbf{a}_*(\mathbf{P}^{-1} + \mathbf{h}^H\mathbf{h})^{-1}(\mathbf{a}_*)^H$ is the effective noise variance associated with the first integer-linear combination (with \mathbf{P} and \mathbf{h} defined in Algorithm 1), and $\sigma^2(\mathbf{b}_*|\mathbf{a}_*) = \frac{p_1 p_2}{\sigma^2(\mathbf{a}_*)} \cdot \frac{1}{1 + \mathbf{h}\mathbf{P}\mathbf{h}^H}$ is the effective noise variance associated with the second integer-linear combination with the first combination as side information. Note that when $\sigma^2(\mathbf{a}_*) \leq \sigma^2(\mathbf{b}_*|\mathbf{a}_*)$, we have

$$R_1 + R_2 = \log_2 (1 + \mathbf{h}\mathbf{P}\mathbf{h}^H). \quad (\text{B.3})$$

That is, the induced rate pair in (B.1) or (B.2) achieves the sum capacity when $\sigma^2(\mathbf{a}_*) \leq \sigma^2(\mathbf{b}_*|\mathbf{a}_*)$. Hence, we consider two sub-cases as follows.

Sub-case 1) $\sigma^2(\mathbf{a}_*) \leq \sigma^2(\mathbf{b}_*|\mathbf{a}_*)$. The induced rate pair (R_1, R_2) is indeed Pareto-optimal as explained before.

Sub-case 2) $\sigma^2(\mathbf{a}_*) > \sigma^2(\mathbf{b}_*|\mathbf{a}_*)$. In this case, the induced rate pair (R_1, R_2) is given by

$$\begin{cases} R_1 = \log_2 \left(\frac{p_1}{\sigma^2(\mathbf{a}_*)} \right), \\ R_2 = \log_2 \left(\frac{p_2}{\sigma^2(\mathbf{a}_*)} \right). \end{cases} \quad (\text{B.4})$$

We will show that (R_1, R_2) is Pareto-optimal by using “proof by contradiction.” Consider a rate pair (\bar{R}_1, \bar{R}_2) induced by $(\bar{\mathbf{a}}, \bar{\mathbf{b}})$ with $\bar{a}_1 \neq 0$ and $\bar{a}_2 \neq 0$. We have

$$\begin{cases} R_1 \leq \log_2 \left(\frac{p_1}{\sigma^2(\bar{\mathbf{a}})} \right) \\ R_2 \leq \log_2 \left(\frac{p_2}{\sigma^2(\bar{\mathbf{a}})} \right) \end{cases} \quad (\text{B.5})$$

which is the rate constraint for decoding the first integer-linear combination associated with $\bar{\mathbf{a}}$. Now suppose that (\bar{R}_1, \bar{R}_2) is better than (R_1, R_2) in (B.4). Then we have

$$\log_2 \left(\frac{p_1}{\sigma^2(\bar{\mathbf{a}})} \right) > \log_2 \left(\frac{p_1}{\sigma^2(\mathbf{a}_*)} \right)$$

and

$$\log_2 \left(\frac{p_2}{\sigma^2(\bar{\mathbf{a}})} \right) > \log_2 \left(\frac{p_2}{\sigma^2(\mathbf{a}_*)} \right).$$

This contradicts Property 2 (i.e., \mathbf{a}_* minimizes $\sigma^2(\mathbf{a})$).

Step 2 of the proof. Consider a Pareto-optimal rate pair $(\tilde{R}_1, \tilde{R}_2)$ induced by $(\tilde{\mathbf{a}}, \tilde{\mathbf{b}})$. We will show that $(\tilde{R}_1, \tilde{R}_2)$ can also be induced by $([1, 0], [0, 1])$, $([0, 1], [1, 0])$, or $(\mathbf{a}_*, \mathbf{b}_*)$. We consider the following cases.

1. The rate pair $(\tilde{R}_1, \tilde{R}_2)$ is a “corner point” of the capacity region. In this case, $(\tilde{R}_1, \tilde{R}_2)$ can be induced by $([1, 0], [0, 1])$ or $([0, 1], [1, 0])$.
2. The rate pair $(\tilde{R}_1, \tilde{R}_2)$ is induced by $(\tilde{\mathbf{a}}, \tilde{\mathbf{b}})$ with $\tilde{a}_1 \neq 0$ and $\tilde{a}_2 \neq 0$. In this case, we will show that $(\tilde{R}_1, \tilde{R}_2)$ can be induced by $(\mathbf{a}_*, \mathbf{b}_*)$. We assume, without loss of generality, that the greatest common divisor $\gcd(\tilde{a}_1, \tilde{a}_2) = 1$, because otherwise, we can achieve a strictly better rate pair than $(\tilde{R}_1, \tilde{R}_2)$ by first decoding the integer-linear combination associated with $\frac{1}{d}\mathbf{a}_*$, where $d = \gcd(\tilde{a}_1, \tilde{a}_2)$ and $|d| > 1$. We now consider two sub-cases as follows.

Sub-case 1) $\tilde{\mathbf{a}}$ and \mathbf{a}_* are linearly dependent. In this case, we have $\tilde{\mathbf{a}} = g\mathbf{a}_*$ for some $|g| = 1$, because $\gcd(a_{1*}, a_{2*}) = 1$ and $\gcd(\tilde{a}_1, \tilde{a}_2) = 1$. Hence, we have $\sigma^2(\tilde{\mathbf{a}}) = \sigma^2(\mathbf{a}_*)$ and $\sigma^2(\tilde{\mathbf{b}}|\tilde{\mathbf{a}}) = \sigma^2(\mathbf{b}_*|\mathbf{a}_*)$. That is, $(\tilde{R}_1, \tilde{R}_2)$ can be induced by $(\mathbf{a}_*, \mathbf{b}_*)$ as well.

Sub-case 2) $\tilde{\mathbf{a}}$ and \mathbf{a}_* are linearly independent. Recall that \mathbf{b}_* is an optimal solution by Property 3, we have $\sigma^2(\tilde{\mathbf{a}}) \geq \sigma^2(\mathbf{b}_*)$. Hence,

$$\begin{cases} \tilde{R}_1 \leq \log_2 \left(\frac{p_1}{\sigma^2(\tilde{\mathbf{a}})} \right) \leq \log_2 \left(\frac{p_1}{\sigma^2(\mathbf{b}_*)} \right) \\ \tilde{R}_2 \leq \log_2 \left(\frac{p_2}{\sigma^2(\tilde{\mathbf{a}})} \right) \leq \log_2 \left(\frac{p_2}{\sigma^2(\mathbf{b}_*)} \right). \end{cases} \quad (\text{B.6})$$

On the other hand, we have $\sigma^2(\mathbf{b}_*) \geq \sigma^2(\mathbf{a}_*)$ (by Property 1) and $\sigma^2(\mathbf{b}_*) \geq \sigma^2(\mathbf{b}_*|\mathbf{a}_*)$ (as shown in [130]). Hence, if we compare (B.6) with (B.1) and (B.2), we conclude that the rate pair $(\tilde{R}_1, \tilde{R}_2)$ is either the same as the pair induced by $(\mathbf{a}_*, \mathbf{b}_*)$ or strictly worse than it. This completes the proof.

B.2 Preliminaries for Lattices and Nested Lattice Codes.

Here, we review some basics of lattices and nested lattice codes. More details can be found in a recent book [141] and a tutorial presentation [142].

A *lattice* is a discrete subgroup (under vector addition) of \mathbb{R}^n . Any (full-rank) lattice Λ in \mathbb{R}^n can be expressed in terms of some (full-rank) $n \times n$ generator matrix $\mathbf{G}_\Lambda \in \mathbb{R}^{n \times n}$ as

$$\Lambda = \{\mathbf{a}\mathbf{G}_\Lambda : \mathbf{a} \in \mathbb{Z}^n\}.$$

That is, Λ is the set of all integer combinations of the rows of \mathbf{G}_Λ .

A *nearest neighbour quantizer* $Q_\Lambda : \mathbb{R}^n \rightarrow \Lambda$ associated with the lattice Λ maps a vector in \mathbb{R}^n to the closest lattice point

$$Q_\Lambda(\mathbf{x}) = \arg \min_{\boldsymbol{\lambda} \in \Lambda} \|\mathbf{x} - \boldsymbol{\lambda}\|_2, \quad (\text{B.7})$$

where ties in (B.7) are broken systematically. The *Voronoi region* of Λ , denoted by $\mathcal{V}(\Lambda)$, is the set of all vectors in \mathbb{R}^n which are quantized to $\mathbf{0}$, i.e., $\mathcal{V}(\Lambda) = \{\mathbf{x} \in \mathbb{R}^n : Q_\Lambda(\mathbf{x}) = \mathbf{0}\}$. The volume of the Voronoi region is denoted by $V(\Lambda)$.

A *nested lattice* is a pair of lattices (Λ_c, Λ_f) such that $\Lambda_c \subset \Lambda_f$, where Λ_f is called the *fine lattice* and Λ_c is called the *coarse lattice*. A *nested lattice code* $\mathcal{L}(\Lambda_c, \Lambda_f)$ consists of the lattice

points of Λ_f in the Voronoi region $\mathcal{V}(\Lambda_c)$, i.e.,

$$\mathcal{L}(\Lambda_c, \Lambda_f) = \Lambda_f \cap \mathcal{V}(\Lambda_c).$$

For this reason, $\mathcal{L}(\Lambda_c, \Lambda_f)$ is also known as a Voronoi codebook. The number of codewords in $\mathcal{L}(\Lambda_c, \Lambda_f)$ is

$$|\mathcal{L}(\Lambda_c, \Lambda_f)| = \frac{V(\Lambda_c)}{V(\Lambda_f)}.$$

Intuitively, each lattice point of Λ_f “occupies” a Voronoi region of volume $V(\Lambda_f)$, and so the number of lattice points inside $\mathcal{V}(\Lambda_c)$ is $V(\Lambda_c)/V(\Lambda_f)$.

A nested lattice code can be constructed from a nested linear code. Consider two linear codes C_1 and C_2 over the field $\mathbb{Z}_p = \{0, 1, \dots, p-1\}$, where $C_2 \subset C_1 \subset \mathbb{Z}_p^n$, and p should be a Gaussian prime (i.e., $p \in \mathbb{Z}$ is a prime number of the form $4m+3$). By “lifting” these linear codes to \mathbb{Z}^n via Construction A, we obtain two lattices

$$\Lambda_1 = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x} \bmod p \in C_1\}$$

and

$$\Lambda_2 = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x} \bmod p \in C_2\}$$

with $\Lambda_2 \subset \Lambda_1 \subset \mathbb{Z}^n$.

Finally, we can apply some positive scaling factor γ to obtain a fine lattice

$$\Lambda_f = \gamma\Lambda_1 \triangleq \{\gamma\boldsymbol{\lambda} : \boldsymbol{\lambda} \in \Lambda_1\}$$

and a coarse lattice

$$\Lambda_c = \gamma\Lambda_2 \triangleq \{\gamma\boldsymbol{\lambda} : \boldsymbol{\lambda} \in \Lambda_2\}$$

with $\Lambda_c \subset \Lambda_f \subset \gamma\mathbb{Z}^n$.

B.3 Proof of Theorem 5

Proof: We will show that the map φ satisfies the three properties.

1. Note that

$$\begin{aligned}
 \boldsymbol{\lambda} \in \Lambda_f &\iff \gamma^{-1}\boldsymbol{\lambda} \in \Lambda_1 \\
 &\iff \gamma^{-1}\boldsymbol{\lambda} \bmod p \in C_1 \\
 &\iff \phi(\boldsymbol{\lambda}) \in C_1 \\
 &\iff \phi(\boldsymbol{\lambda})\mathbf{H}_1^T = \mathbf{0} \\
 &\iff \phi(\boldsymbol{\lambda})\mathbf{H}^T \text{ has } m_1 \text{ zeros at the beginning} \\
 &\iff \varphi(\boldsymbol{\lambda}) \text{ has } m_1 \text{ zeros at the beginning.}
 \end{aligned}$$

2. The proof is essential the same as above.

3. Note that

$$\begin{aligned}
 &\varphi(a_1\boldsymbol{\lambda}_1 + a_2\boldsymbol{\lambda}_2) \\
 &= \phi(a_1\boldsymbol{\lambda}_1 + a_2\boldsymbol{\lambda}_2)\mathbf{H}^T \\
 &= (\gamma^{-1}(a_1\boldsymbol{\lambda}_1 + a_2\boldsymbol{\lambda}_2) \bmod p)\mathbf{H}^T \\
 &= ((a_1\gamma^{-1}\boldsymbol{\lambda}_1 + a_2\gamma^{-1}\boldsymbol{\lambda}_2) \bmod p)\mathbf{H}^T \\
 &= ((q_1\gamma^{-1}\boldsymbol{\lambda}_1 + q_2\gamma^{-1}\boldsymbol{\lambda}_2) \bmod p)\mathbf{H}^T \\
 &= (q_1\phi(\boldsymbol{\lambda}_1) + q_2\phi(\boldsymbol{\lambda}_2))\mathbf{H}^T \\
 &= q_1\varphi(\boldsymbol{\lambda}_1) + q_2\varphi(\boldsymbol{\lambda}_2).
 \end{aligned}$$

Therefore, the map φ is a linear labeling. This completes the proof of Theorem 5. □

B.4 Proof of Theorem 6

Proof: First, we show that the rate pairs associated with $[(1,0)^T, (0,1)^T]$, $[(0,1)^T, (1,0)^T]$, and $[\mathbf{a}_*, \mathbf{b}_*]$ are indeed Pareto-optimal solutions. When $\mathbf{a} = (1,0)^T$ and $\mathbf{b} = (0,1)^T$, by Proposition 2, the associated rate pair (R_k, R_l) satisfies

$$R_k + R_l = \log_2 |\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H|, \quad (\text{B.8})$$

which is the sum capacity. Hence, (R_k, R_l) is a Pareto-optimal solution. Similarly, when $\mathbf{a} = (0,1)^T$ and $\mathbf{b} = (1,0)^T$, the associated rate pair (R_k, R_l) is a Pareto-optimal solution. We proceed to the rate pair (R_k, R_l) associated with $[\mathbf{a}_*, \mathbf{b}_*]$. Without loss of generality, we assume that $a_k \bmod p \neq 0$ and $a_l \bmod p \neq 0$, because otherwise it reduces to the previous two cases. We consider two sub-cases as follows:

1. $\sigma^2(\mathbf{H}, \mathbf{a}_*) \leq \sigma^2(\mathbf{H}, \mathbf{b}_* | \mathbf{a}_*)$. In this case, by Proposition 2, we have $R_k + R_l = \log_2 |\mathbf{I}_M + \mathbf{H}\mathbf{P}\mathbf{H}^H|$. Hence, (R_k, R_l) is a Pareto-optimal solution.
2. $\sigma^2(\mathbf{H}, \mathbf{a}_*) > \sigma^2(\mathbf{H}, \mathbf{b}_* | \mathbf{a}_*)$. In this case, the rate pair (R_k, R_l) is given by

$$\begin{aligned} R_k &= \log_2^+ \left(\frac{P_k}{\sigma^2(\mathbf{H}, \mathbf{a}_*)} \right) \\ R_l &= \log_2^+ \left(\frac{P_l}{\sigma^2(\mathbf{H}, \mathbf{a}_*)} \right). \end{aligned}$$

We can show that (R_k, R_l) is a Pareto-optimal solution by using the “proof by contradiction” technique developed in the proof of Theorem 4 in Appendix B.1.

Next, we consider a Pareto-optimal rate pair $(\tilde{R}_k, \tilde{R}_l)$ associated with $[\tilde{\mathbf{a}}, \tilde{\mathbf{b}}]$ and we show that $(\tilde{R}_k, \tilde{R}_l)$ is also associated with $[(1,0)^T, (0,1)^T]$, $[(0,1)^T, (1,0)^T]$, or $[\mathbf{a}_*, \mathbf{b}_*]$. To this end, we consider the following cases.

1. The rate pair $(\tilde{R}_k, \tilde{R}_l)$ is a “corner point” of the capacity region. In this case, $(\tilde{R}_k, \tilde{R}_l)$ is also associated with $[(1,0)^T, (0,1)^T]$ or $[(0,1)^T, (1,0)^T]$.

2. The rate pair $(\tilde{R}_k, \tilde{R}_l)$ is associated with $[\tilde{\mathbf{a}}, \tilde{\mathbf{b}}]$ with $\tilde{a}_k \bmod p \neq 0$ and $\tilde{a}_l \bmod p \neq 0$. In this case, we will show that $(\tilde{R}_k, \tilde{R}_l)$ is also associated with $[\mathbf{a}_*, \mathbf{b}_*]$. We assume, without loss of generality, that the greatest common divisor $\gcd(\tilde{a}_k, \tilde{a}_l) = 1$, because otherwise, we can achieve a strictly better rate pair than $(\tilde{R}_k, \tilde{R}_l)$ by first decoding the linear combination associated with $\frac{1}{d}\mathbf{a}_*$, where $d = \gcd(\tilde{a}_k, \tilde{a}_l)$ and $|d| > 1$. We now consider two sub-cases as follows.

Sub-case 1) $\tilde{\mathbf{a}}$ and \mathbf{a}_* are linearly dependent. In this case, we have $\tilde{\mathbf{a}} = g\mathbf{a}_*$ for some $|g| = 1$, because $\gcd(a_{k*}, a_{l*}) = 1$ and $\gcd(\tilde{a}_k, \tilde{a}_l) = 1$. Hence, we have $\sigma^2(\mathbf{H}, \tilde{\mathbf{a}}) = \sigma^2(\mathbf{H}, \mathbf{a}_*)$ and $\sigma^2(\mathbf{H}, \tilde{\mathbf{b}}|\tilde{\mathbf{a}}) = \sigma^2(\mathbf{H}, \mathbf{b}_*|\mathbf{a}_*)$. That is, $(\tilde{R}_k, \tilde{R}_l)$ is associated with $[\mathbf{a}_*, \mathbf{b}_*]$ as well.

Sub-case 2) $\tilde{\mathbf{a}}$ and \mathbf{a}_* are linearly independent. Recall that \mathbf{b}_* is an optimal solution to (4.10) under the constraint that \mathbf{a}_* and \mathbf{b}_* are linearly independent. Hence, we have $\sigma^2(\mathbf{H}, \tilde{\mathbf{a}}) \geq \sigma^2(\mathbf{H}, \mathbf{b}_*)$ and

$$\begin{cases} \tilde{R}_k \leq \log_2 \left(\frac{P_k}{\sigma^2(\mathbf{H}, \tilde{\mathbf{a}})} \right) \leq \log_2 \left(\frac{P_k}{\sigma^2(\mathbf{H}, \mathbf{b}_*)} \right) \\ \tilde{R}_l \leq \log_2 \left(\frac{P_l}{\sigma^2(\mathbf{H}, \tilde{\mathbf{a}})} \right) \leq \log_2 \left(\frac{P_l}{\sigma^2(\mathbf{H}, \mathbf{b}_*)} \right). \end{cases} \quad (\text{B.9})$$

On the other hand, we have $\sigma^2(\mathbf{H}, \mathbf{b}_*) \geq \sigma^2(\mathbf{H}, \mathbf{a}_*)$ and $\sigma^2(\mathbf{H}, \mathbf{b}_*) \geq \sigma^2(\mathbf{H}, \mathbf{b}_*|\mathbf{a}_*)$ (as shown in [130]). Hence, if we compare (B.9) with the rate expressions in Proposition 2, we conclude that the rate pair $(\tilde{R}_k, \tilde{R}_l)$ is either the same as the pair associated with $[\mathbf{a}_*, \mathbf{b}_*]$ or strictly worse than it.

This completes the proof of Theorem 6. □