MODELLING, ANALYSIS AND OPTIMIZATION OF NON-ORTHOGONAL MULTIPLE ACCESS IN NEXT-GENERATION WIRELESS COMMUNICATION SYSTEMS

A Thesis

Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

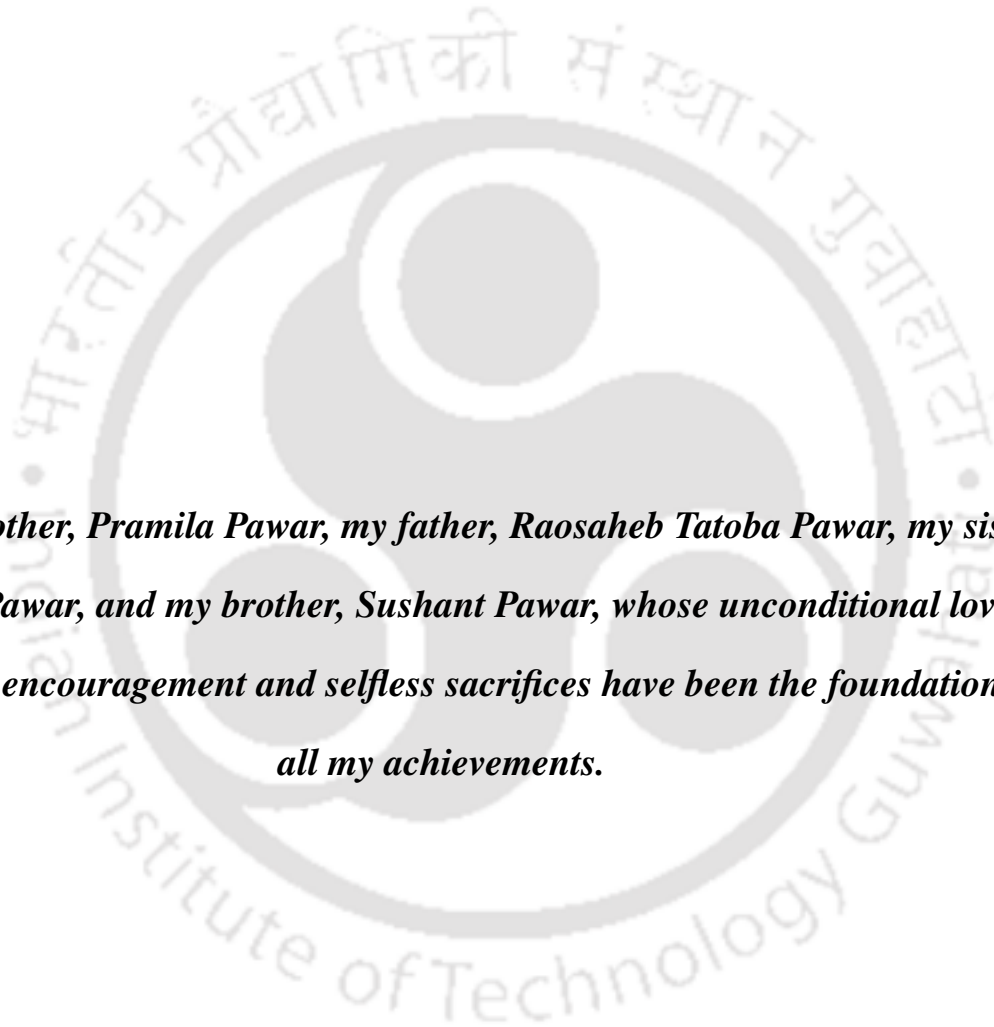
by

Aditya Raosaheb Pawar

(166302001)

Department of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati
Guwahati - 781039, Assam, India

December 2024



To my mother, Pramila Pawar, my father, Raosaheb Tatoba Pawar, my sister, Priya Pawar, and my brother, Sushant Pawar, whose unconditional love, support, encouragement and selfless sacrifices have been the foundation of all my achievements.

Certificate

This is to certify that the thesis entitled "**Modelling, Analysis and Optimization of Non-Orthogonal Multiple Access in Next-Generation Wireless Communication Systems**", submitted by **Aditya Raosaheb Pawar**, a research scholar in the *Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati*, for the award of the degree of **Doctor of Philosophy**, is a record of an original research work carried out by him under our supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the institute and in our opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Date:

Place: Guwahati

Dr. Sonali Chouhan
Dept. of Electronics and Electrical Engg.,
Indian Institute of Technology Guwahati,
Guwahati - 781 039, Assam, India.

Date:

Place: Guwahati

Dr. Salil Kashyap
Dept. of Electronics and Electrical Engg.,
Indian Institute of Technology Guwahati,
Guwahati - 781 039, Assam, India.

Acknowledgements

Firstly, I extend my heartfelt appreciation to my Supervisors, Dr. Sonali Chouhan and Dr. Salil Kashyap, for their exceptional guidance throughout my research journey. Their compassion, commitment, diligent work, and meticulousness have been incredibly motivating. I am deeply grateful for their unwavering support and patience with me. There were several challenging instances during my PhD journey and it was the constant support and appropriate direction from my Supervisors that helped me believe in myself and overcome the challenges. I am indebted to them for their thorough review and constructive feedback on every piece of work I submitted.

I am immensely grateful to my doctoral committee members, Prof. Prabin Kumar Bora, Dr. A. Rajesh, and Dr. Kalpana Dhaka, for taking time out of their hectic schedules to assess my progress and enhance my work with their priceless advice and critiques.

My gratitude also extends to the Head of the Department for the support during my academic journey. I must also acknowledge the Ministry of Education, Government of India, for granting me a fellowship during my Ph.D. studies.

I am immensely grateful to my friends from Communication Lab-I, Mohit Mishra, Arijit Bhattacharjee, Ashim Kumar and Rahul Sharma for their invaluable support, both technical and non-technical discussions, throughout this journey. Your willingness to go out of your way to help during critical moments made the difference. The lighter and humorous moments we shared were a breath of fresh air. Whether it was a chai break turned into a laughter marathon, random socio-political discussions, or simply venting out frustrations together, these moments not only reduced my stress but also kept me motivated and grounded. Without your camaraderie and encouragement, this journey would have been difficult to navigate.

To my dual-degree friends and cricket team friends, Sibasis, Nishant, Raju, Atanu and Birjit, your companionship has been one of the treasured aspects of this journey. The countless lighter moments we spent together, from our long, non-technical discussions on random topics to endless laughter, provided much-needed relief at times and made this experience so much

more enjoyable. Our cricket matches were truly special with a mix of competitiveness, hilarious mishaps, and unforgettable moments. The cooking parties were an absolute delight, and yes, while there will always be that anecdote to endlessly debate about why and how we could not win the PPL, the joy, the adrenaline rush, and the endless banter we shared on and off the field more than made up for it. Those cherished memories will forever bring a smile. In many ways, the story of my PhD journey would feel incomplete without acknowledging these moments that brought joy, friendship, and a sense of belonging.

I would also like to thank my friends from the research group, Chandan Kumar, Pothan, Gangesh, Faisal, Vikas, Nabajyoti and Chetan, whose company helped me a great deal at times to tackle the challenges of research life and otherwise.

I would also like to extend my gratitude to the beautiful state of Assam, which became my home during PhD journey. The beauty of its landscapes, the majestic flow of the mighty Brahmaputra, and the colors of its culture have provided a sense of peace. Assam's serenity played an undeniable role in helping me stay motivated throughout my research journey. I would like to thank Prarthana for amazing experiences in the state of Assam during my PhD journey.

Finally, I am profoundly grateful to my mother, my father, my sister, and my brother for their unwavering love and support, not just during this journey but throughout my life. Their endless encouragement, patience, and belief in my abilities have been my greatest sources of strength. Whether it was through their comforting words during challenging times, their constant motivation to push forward, or simply their presence that reminded me I was never alone, they have been my anchor. This achievement is as much theirs as it is mine, and I dedicate it to the love and sacrifices they have made for me.

(Aditya Raosaheb Pawar)

Abstract

In this thesis, we present detailed performance analysis and optimization of non orthogonal multiple access (NOMA) systems aided by massive multiple input multiple output (MIMO) and intelligent reflecting surface (IRS). To being with, we analyze the uplink of a massive MIMO-NOMA system in the second chapter, and deduce new lower bounds on the achievable spectral efficiency (SE) based on zero-forcing (ZF) decoding at the base station (BS). Users are grouped to employ NOMA. In order to cancel the inter-group interference, the ZF decoder is designed as a function of channel estimates acquired based on two low overhead channel estimation schemes, namely, Scheme-I and Scheme-S. Under these estimation schemes, pilots are shared among users in a group, while they are orthogonal across groups. User grouping and power allocation are two pivotal ways to regulate the performance of users in a NOMA system. Thus, to ensure uniform quality-of-service to all users, we obtain the max-min power control coefficients which maximize the minimum achievable SE. Furthermore, we investigate two different user grouping strategies, namely, the non-scalable near-far (NF) grouping and scalable neighbor (NB) grouping, and study their impact on the uplink max-min achievable SE. We benchmark against the performance obtained with maximum ratio (MR) decoder. Extensive analysis and simulations show that in a substantial portion of the over-loaded regime and in the entire under-loaded regime, ZF decoder designed using channel estimates acquired via Scheme-S with NF grouping gives the highest max-min achievable SE among all the considered decoding and grouping strategies.

The analysis and corresponding contributions made in the second chapter are based on the assumption of quasi-static fading, which implies that the channel remains

invariant over the transmission interval, which includes both channel estimation phase and data transmission phase. However, in a high-mobility scenario, the channel may vary at a faster rate and the acquired channel estimates may get outdated by the time those are used for processing. Therefore, in the third chapter, we quantify the joint impact of channel outdatedness and pilot contamination on the achievable SE in uplink of a massive MIMO-NOMA system. Specifically, we compute novel closed-form expressions for achievable SE (i) for both ZF and MR decoders based on outdated channel estimates obtained using Scheme-I and Scheme-S, and (ii) for ZF and MR decoders based on predicted channel obtained using Wiener linear predictor (WLP). The analysis accounts for time-variations in channel due to Doppler shift, correlation due to pilot sharing and imperfect successive interference cancellation (SIC). We formulate and solve optimization problems for max-min and proportional fairness power control using convex programming. We also study equal and inversion power control strategies. Intriguingly, among equal, inversion and max-min power control under channel aging, the ZF decoder based on Scheme-S (i) gives the highest per-user SE with max-min power control for smaller values of normalized Doppler shift ($f_D T_S \lesssim 0.25$) and with inversion power control for larger values of $f_D T_S$, and (ii) gives the highest minimum user SE with max-min power control. Proportional fairness power control provides higher per-user SE than max-min power control while maintaining fairness to a significant extent.

The performance gains provided by massive MIMO systems bank on the availability of a large number of active RF chains, which are costly and power hungry. Thus, to address this issue, we then consider an IRS-aided NOMA system as a cost-efficient alternative to large antenna systems. Here, we analyze the error performance of IRS-aided NOMA systems. We deduce novel approximate analytical expressions for average symbol error probability (SEP) for pulse amplitude modulation (PAM) and quadrature amplitude modulation (QAM). We also analyze the multi-user scenario by deriving novel approximate analytical expressions for (i) the

average SEP for multicast transmission and (ii) the average pairwise error probability (PEP) for unicast transmission. We consider Rayleigh faded direct path and independently Rician faded reflected paths from BS to users via IRS. We also formulate and solve a constrained optimization problem to determine the fraction of elements to be activated at any IRS in order to guarantee near-identical SEP performance for both the users. We present results to elucidate that adding few extra IRS elements yields signal-to-noise ratio (SNR) gain for both users. Specifically, SNR gain of around 4 dB is observed with increase in IRS elements from 128 to 256 for the target SEP of 10^{-3}. Further, we show that increasing the fraction of IRS elements activated to serve the weaker users allows significant reduction in the fraction of power allocated to the weaker user in order to match its SEP performance with that of the strong user. In addition, we show that adding a few extra IRS elements allows communication with higher-order constellation without significant increase in SEP/PEP. Furthermore, a 3-bit programmable IRS yields near-identical performance as an IRS with infinite precision phase-shift. Impact of estimation error on average SEP is also investigated.

Contents

List of Figures **xiii**

List of Tables **xvi**

List of Abbreviations **xviii**

1 Introduction **1**

1.1 Introduction 2

1.2 Potential Technologies 4

1.2.1 NOMA 4

1.2.1.1 Downlink NOMA 5

1.2.1.2 Uplink NOMA 8

1.2.2 Massive MIMO 10

1.2.2.1 Uplink Massive MIMO 10

1.2.2.2 Downlink Massive MIMO 11

1.2.3 IRS 12

1.3 Literature Review 14

1.3.1 Massive MIMO-NOMA Systems without Channel Aging 14

1.3.2 Massive MIMO Systems under Channel Aging and Massive MIMO-NOMA Systems under Channel Aging 15

1.3.3 IRS-Aided NOMA Systems 17

1.3.4 Research Gaps 19

1.4 Key Challenges 22

1.5 Problem Statements . . . 23
1.6 Thesis Objectives . . . 24
1.7 Thesis Contributions . . . 25
1.7.1 Massive MIMO-NOMA Systems . . . 25
1.7.1.1 Contributions . . . 25
1.7.2 Massive MIMO-NOMA Systems under Channel Aging . . . 26
1.7.2.1 Contributions . . . 26
1.7.3 IRS-Aided NOMA Systems . . . 27
1.7.3.1 Contributions . . . 28
1.8 Thesis Organization . . . 29
1.9 List of Publications . . . 30

2 Massive MIMO-NOMA Systems 32

2.1 System Model . . . 33
2.1.1 Channel Estimation . . . 34
2.1.1.1 Preliminaries on MMSE Estimation . . . 34
2.1.1.2 MMSE Estimates with Scheme-I . . . 37
2.1.1.3 MMSE Estimates with Scheme-S . . . 38
2.1.2 Uplink Data Transmission . . . 39
2.2 Spectral Efficiency Analysis . . . 39
2.2.1 Achievable SE with ZF based on Scheme-I . . . 40
2.2.2 Achievable SE with ZF based on Scheme-S . . . 44
2.3 Max-Min Power Control . . . 47
2.4 User Grouping Strategies . . . 50
2.5 Numerical and Simulation Results . . . 53
2.6 Summary . . . 59

3 Massive MIMO-NOMA Systems under Channel Aging 61

3.1 System Model . . . 63

3.1.1 Channel Estimation 64
3.1.1.1 MMSE Estimates with Scheme-I 66
3.1.1.2 MMSE Estimates with Scheme-S 67
3.1.2 Uplink Data Transmission 68
3.2 Spectral Efficiency Analysis 68
3.2.1 Achievable SE with Scheme ZF-I 69
3.2.2 Achievable SE with Scheme ZF-S 71
3.2.3 Achievable SE with Scheme MR-I 75
3.2.4 Achievable SE with Scheme MR-S 77
3.3 Mitigation of Impact of Channel Aging using Channel Prediction 79
3.3.1 Linear Prediction of Channel Vectors $\mathbf{g}_{nk}[t+1]$ 80
3.3.2 Achievable SE with ZF using Predicted CSI (ZF-I-P) 81
3.3.3 Achievable SE with MR using Predicted CSI (MR-I-P) 82
3.4 Power Control 84
3.4.1 Max-Min Power Control 86
3.4.2 Proportional Fairness Power Control 87
3.5 Numerical and Simulation Results 87
3.5.1 Impact of Channel Aging on the Per-User Achievable SE 88
3.5.1.1 Equal Power Control 88
3.5.1.2 Inversion Power Control 89
3.5.1.3 Max-Min Power Control 90
3.5.1.4 Ad-hoc Power Control vs Max-Min Power Control . . . 90
3.5.1.5 Impact of Users Per Group (K) 91
3.5.1.6 Impact of User Grouping Strategies 92
3.5.2 Impact of Channel Aging on Minimum Achievable User SE . . . 93
3.5.3 Proportional Fairness Power Control 94
3.5.4 Performance of ZF Designed using Predicted CSI 95
3.6 Summary 96

3.7 Comparison of Contributions in Chapters 2 and 3 97

4 IRS-Enabled NOMA Systems 100

4.1 System Model 101

4.2 Error Probability Analysis 104

4.2.1 SEP Analysis for Two-User Setup with PAM Symbols 105

4.2.2 SEP Analysis for Two-User Setup with QAM Symbols 108

4.2.3 Error Probability for Multi-User Setup with PAM 114

4.2.3.1 SEP Analysis for Multicast Transmission 115

4.2.3.2 PEP Analysis for Unicast Transmission 117

4.3 Numerical and Simulation Results 119

4.3.1 Two-User Setup with PAM 119

4.3.1.1 Imperfect CSI 121

4.3.1.2 IRS Activation Ratio 122

4.3.1.3 Discrete Phase-Shifts 124

4.3.2 Two-User Setup with QAM 124

4.3.3 Multi-User Setup: SEP Analysis for Multicast Transmission . . . 124

4.3.4 Multi-User Setup: PEP Analysis for Unicast Transmission 125

4.4 Summary 126

5 Conclusions 128

5.1 Research Contributions 129

5.1.1 Massive MIMO-NOMA Systems 129

5.1.2 Massive MIMO-NOMA Systems under Channel Aging 130

5.1.3 IRS-aided NOMA Systems 131

5.2 Quantification of the Contributions Towards the Objectives of the Thesis 131

5.3 Directions for Future Research 133

5.3.1 Massive MIMO Systems with RSMA Transmission 133

5.3.2 Error Probability Analysis of NOMA Systems under Imperfect CSI . . . 134
5.3.3 IRS-enabled Massive MIMO-NOMA Systems . . . 134
5.3.4 Multi-IRS-aided Massive MIMO mmWave/THz Systems . . . 135

A Appendix **136**
A.1 Proof of Theorem 2.1 . . . 136
A.2 Proof of Theorem 2.2 . . . 137

B Appendix **141**
B.1 Proof of Theorem 3.1 . . . 141
B.2 Proof of Theorem 3.2 . . . 142
B.3 Proof of Theorem 3.4 . . . 145
B.4 Proof of Theorem 3.5 . . . 147
B.5 Proof of Lemma 3.1 . . . 151
B.6 Proof of Theorem 3.6 . . . 152
B.7 Proof of Theorem 3.7 . . . 153

C Appendix **156**
C.1 Proof of Lemma 4.1 . . . 156
C.2 Proof of Theorem 4.1 . . . 157
C.3 Proof of Theorem 4.2 . . . 158
C.4 Proof of Theorem 4.3 . . . 160
C.5 Proof of Theorem 4.4 . . . 161
C.6 Proof of Theorem 4.5 . . . 162

Bibliography **163**

List of Figures

1.1 NOMA vs OMA. 5

1.2 Downlink NOMA. 6

1.3 SE comparison between NOMA and OMA (downlink): $|h_A| = \frac{\sqrt{5}}{10}$, $|h_B| = \sqrt{5}$, $P_{dl} = 40$ W, $\sigma^2 = 1$. 7

1.4 Uplink NOMA. 8

1.5 Capacity region of uplink NOMA: $|h_A| = 1, |h_B| = 0.7, P_{ul}^A = 0.1$ W, $P_{ul}^B = 0.1$ W, $\sigma^2 = 0.1$. 9

1.6 Single-cell massive MIMO. 10

1.7 IRS-aided wireless system. 13

1.8 Structure of the thesis. 30

2.1 System model: Uplink massive MIMO-NOMA. 34

2.2 Capacity region of uplink NOMA: $h_A = 1, h_B = 0.7, P_{ul}^A = 0.1, P_{ul}^B = 0.1, \sigma^2 = 0.1$. 44

2.3 User grouping strategies. 51

2.4 Max-min achievable SE comparison: over-loading, $M = 64$, $K = 2$. 53

2.5 Max-min achievable SE comparison: under-loading, $M = 256$, $K = 2$. 54

2.6 Max-min achievable SE comparison: over-loading, $M = 120$. 56

2.7 Max-min achievable SE comparison: under-loading, $M = 200$. 57

2.8 Max-min achievable sum SE comparison: $M = 120$. 58

2.9 Fairness comparison among SE achieved with different power allocation policies using proposed ZF decoder based on Scheme-S and neighbor grouping: $M = 8$, $L = 8$, $K = 2$. 58

3.1 System model: Uplink massive MIMO-NOMA under time-variant channel. 63
3.2 Equal power control: NF grouping, $N = 10$, $K = 2$, $M = 128$. 89
3.3 Inversion power control: NF grouping, $N = 10$, $K = 2$, $M = 128$. 89
3.4 Max-min power control: NF grouping, $N = 10$, $K = 2$, $M = 128$. 90
3.5 Ad-hoc vs max-min power control: NF grouping, $N = 10$, $K = 2$, $M = 128$. 90
3.6 ZF-I vs ZF-S under max-min power control with NB grouping: $K = 2, 3, 4$, $L = 24$, $M = 128$. 91
3.7 NF vs NB grouping under max-min power control: $K = 2, 3, 4$, $L = 24$, $M = 128$. 92
3.8 Minimum achievable user SE (in bits/s/Hz) with NF grouping: $N = 10$, $K = 2$, $M = 128$. 93
3.9 ZF-S with NF grouping: $N = 10$, $K = 2$, $M = 128$. 94
3.10 Max-min vs proportional fairness with NF grouping: $N = 10$, $K = 2$, $M = 128$. 95
3.11 Comparison of max-min and proportional fairness: $N = 10$, $K = 2$, $M = 128$, NF grouping. 95
3.12 ZF with predicted CSI under max-min power control: $N = 10$, $K = 2$, $M = 128$, NF grouping. 96

4.1 System model: IRS-aided NOMA, two-user case. 102
4.2 System model: IRS-aided NOMA, multi-user case. 115
4.3 SEP vs. Eb/N0: two-user PAM, perfect CSI, $\lambda = 0.7$. 120
4.4 SEP vs. Eb/N0: two-user PAM, $N_2 = 64$, $N_1 = 4$, $\lambda = 0.95$. 121
4.5 Impact of imperfect CSI, two-user PAM: $T = 128$, $\lambda = 0.7$. 121
4.6 Trade-off between γ^* and λ: $N_1 = 2$, $N_2 = 4$, SNR $= 10$ dB. 122
4.7 Discrete phase-shift, weak user: $N_1 = 2$, $N_2 = 4$, $T = 100$, $\lambda = 0.7$. 124

4.8 SEP vs. Eb/N0: two-user QAM, $N_2 = 16$, $N_1 = 4$, $\lambda = 0.95$. 125

4.9 SEP vs. Eb/N0: multi-user multicast transmission with PAM, $T = 256$, $\lambda = 0.7$. 126

4.10 PEP vs. Eb/N0 (dB): Multi-user unicast transmission, $K = 4$, $P_1 = 0.4P$, $P_2 = 0.3P$, $P_3 = 0.2P$, $P_4 = 0.1P$. 127

List of Tables

1.1 Comparison of 5G and 6G in terms of KPIs 3

1.2 Research gaps in the literature on uplink massive MIMO-NOMA systems. . 19

1.3 Research gaps in the literature on downlink IRS-aided NOMA systems. . . 21

2.1 Simulation parameters. 52

2.2 Comparison of SEs (in bps/Hz) of decoding schemes. 54

3.1 Simulation parameters. 88

3.2 Comparison of Per-User Achievable SEs (in bps/Hz). 91

3.3 Comparison of Per-User Achievable SEs (in bps/Hz) under Max-Min Power Control. 92

3.4 Comparison of Minimum Achievable User SEs (in bps/Hz) with scheme ZF-S under NF Grouping. 94

3.5 Comparison of Per-User Achievable SEs (in bps/Hz) under Max-Min and Proportional Fairness Power Control. 96

3.6 Comparison of Per-User Achievable SEs (in bps/Hz) ZF under NF Grouping and Max-Min Power Control. 97

3.7 Comparison of modelling, analysis and contributions in Chapter 2 and Chapter 3. 98

4.1 Simulation parameters. 119

4.2 Comparison of Average PEP for PAM and BPSK, SNR = 12 dB. 126

List of Abbreviations

AWGN	Additive white Gaussian noise
BER	Bit error rate
BF	Beamforming
BLER	Block error rate
BPSK	Binary phase shift keying
BS	Base station
CDF	Cumulative distribution function
CLT	Central limit theorem
CSI	Channel state information
FDD	Frequency-division duplex
FDMA	Frequency-division multiple access
IoT	Internet-of-Things
IRS	Intelligent reflecting surface
KPI	Key performance indicator
MIMO	Multiple input multiple output
MMSE	Minimum mean square error
mmWave	Millimeter-wave
MR	Maximum ratio
NB	Neighbor
NF	Near-far
NOMA	Non-orthogonal multiple access
OMA	Orthogonal multiple access

PAM	Pulse amplitude modulation
PDF	Probability density function
PEP	Pairwise error probability
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
RF	Radio-frequency
RSMA	Rate-splitting multiple access
SC	Superposition coding
SDMA	Space-division multiple access
SE	Spectral efficiency
SEP	Symbol error probability
SIC	Successive interference cancellation
SISO	Single input single output
SINR	Signal-to-interference-plus-noise ratio
SNR	Signal-to-noise ratio
TDMA	Time-division multiple access
WLP	Wiener linear predictor
ZF	Zero-forcing
5G	Fifth generation
5G NR	Fifth generation new radio
6G	Sixth generation

1

Introduction

Contents

1.1 Introduction 2
1.2 Potential Technologies 4
1.3 Literature Review 14
1.4 Key Challenges 22
1.5 Problem Statements 23
1.6 Thesis Objectives 24
1.7 Thesis Contributions 25
1.8 Thesis Organization 29
1.9 List of Publications 30

1.1 Introduction

Massive connectivity is one of the primary requirements of fifth generation (5G) and beyond wireless communication systems along with higher data rates, improved coverage, ultra-reliability and low-latency [1, 2]. Table 1.1 at the top of the next page presents the comparison between 5G and sixth generation (6G) technologies in terms of the key performance indicators (KPIs) [3]. It is estimated that more than 125 billion devices will have to be provided wireless connectivity by 2030 [1, 4, 5]. This is due to an exponential increase in the number of personal devices and applications such as smart cities, connected vehicles, industry 4.0, e-health and internet-of-things (IoT) devices. The connection density requirement of 5G is 10^6 devices/km^2 [4]. The 6G networks aim to support connection density of 10^7 devices/km^2 [1]. As technologies evolve further from 5G to 6G, the reliability performance in terms of block error rate (BLER) is anticipated to improve from 10^{-5} to 10^{-9}. In addition, high mobility support is also one of the KPIs which is expected to improve from 500 Km/h to 1000 Km/h. Along with the important KPIs, improving the spectral efficiency (SE) (measured in bits/s/Hz) is a persistent requirement of wireless communication systems.

To achieve these ambitious goals, several potential technologies are being thoroughly investigated by the research community in both industry and academia. For instance, there are several possible ways to increase wireless connectivity and improve SE. The connection density can be increased by techniques like (i) increasing the bandwidth and exploiting higher frequencies like millimeter waves (mmWaves) for communication, (ii) using extremely large number of antennas at the base station (BS), e.g. massive multiple input multiple output (MIMO), (iii) deploying smaller cells, and (iv) using multiple access technologies for non-orthogonal transmission, e.g. power-domain non-orthogonal multiple access (NOMA) [6], and rate-splitting multiple access (RSMA) [4]. Further, it is shown that the reliability in signal detection can be improved using BF gains achieved with the large multi-antenna systems, e.g. massive MIMO, intelligent reflecting surfaces (IRSs). Apart from this, wireless connectivity under high-mobility scenarios is challenging because with increase in mobility/speed, the coherence interval shortens.[1]

[1]The time duration over which the channel response can be assumed to be constant is termed as coherence

Table 1.1: Comparison of 5G and 6G in terms of KPIs

KPI	**5G**	**6G**
Peak data rate	20 Gbps	1 Tbps
User experienced data rate	100 Mbps	1 Gbps
Peak SE	30 bits/s/Hz	60 bits/s/Hz
User experienced SE	0.3 bits/s/Hz	3 bits/s/Hz
End-to-end latency	10 ms	1 ms
Radio-only latency	1 ms	100 μs
BLER	10^{-5}	10^{-9}
Connection density	10^6 devices/Km^2	10^7 devices/Km^2
Position accuracy	1 m	0.1 m
Mobility	500 Km/h	1000 Km/h
Maximum frequency	100 GHz	10 THz
Maximum bandwidth	1 GHz	100 GHz
Satellite integration	Partial	Full

Thus, the available channel state information (CSI) gets outdated very fast. This phenomenon is termed as channel aging [8]. Thus, in order to ensure seamless wireless connectivity under high-mobility scenarios, first and foremost, it is crucial to quantify the impact of channel aging on wireless system's performance.

In this thesis, we focus on NOMA for the design and analysis of next-generation wireless communication systems aiming to improve connection density, ensure fairness in performance in terms of achievable SE and improve reliability of signal detection.[2] We also focus on ensuring wireless connectivity in high mobility scenarios, where the channel gets outdated faster. It is well-known from the existing literature that NOMA holds the potential to improve the connection density with improvement in SE by providing simultaneous service to multiple users over the same time-frequency resource [9]. In addition, massive MIMO technology, which significantly improves SE and reliability due to BF gain achieved using large number of antennas, plays a vital role in 5G and promises to extend its role to future wireless communication systems. Therefore, it has been noted in the literature that NOMA, when aided with spatial multiplexing capabilities of massive MIMO systems, can boost the connection density and SE further [10]. Recently, IRSs are also being explored to understand the improvements in SE,

interval [7].

[2]Fairness in performance (or equivalently, user fairness) implies that all users experience similar performance irrespective of differences in users' channel conditions.

reliability and coverage that can be brought about through their use in next generation wireless communication systems. The passive nature of IRSs is an added advantage in terms of cost and energy efficiency. Therefore, it is of great interest to understand how IRSs can improve system performance when aided by NOMA transmission. Thus, in this thesis, we focus on design and analysis of massive MIMO-NOMA and IRS-enabled NOMA systems in the context of next-generation wireless communication. However, each of these technologies in combination with each other poses several challenges, and it is critical to address these challenges effectively. We will discuss these challenges in detail subsequently. Below we first discuss the principles of NOMA, massive MIMO and IRS followed by the detailed literature review in order to present the state-of-the-art and to identify the crucial research gaps therein.

1.2 Potential Technologies

As discussed above, NOMA is a key technology that has the potential to support massive connectivity. In addition, massive MIMO is an integral part of 5G new radio (5G NR) owing to its capability to significantly improve the system performance. Recently, IRS has shown the ability to provide significant BF gains and improve coverage, thereby emerging as a strong candidate technology for future wireless systems. Below we briefly discuss the operating principles of these technologies.

1.2.1 NOMA

NOMA is a promising technology for future wireless communication systems since it has the potential to provide massive connectivity by serving multiple users in any time-frequency resource [9, 11]. While in an orthogonal multiple access technique (e.g. time-division multiple access (TDMA), frequency-division multiple access (FDMA)), orthogonal resources are assigned to different users, NOMA uses all available resource to simultaneously serve multiple users, as shown in Fig. 1.1. To enable this, superposition coding (SC) is employed at the transmitter and successive interference cancellation (SIC) is employed at the receiver(s) [12]. Specifically, users receiving signals at higher power levels can decode their messages by treating the interfering signals from other users as noise. Meanwhile, users at lower power levels

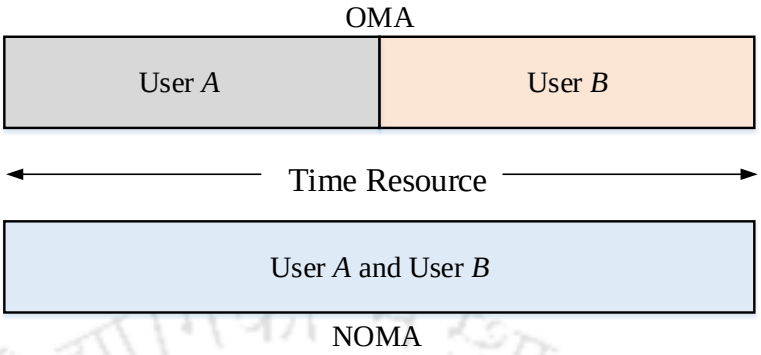


Fig. 1.1: NOMA vs OMA.

can utilize SIC to decode their messages by successively eliminating the signals transmitted by other users. There are a few key differences in the working principles of uplink and downlink NOMA, which we discuss below.

1.2.1.1 Downlink NOMA

To understand the details, consider the downlink communication in a network with two single-antenna users, user A and user B, and a single antenna BS as shown in Fig. 1.2. Let P_{dl} be the total transmit power available at the BS. Now, the messages corresponding to both users are linearly combined with different weights based on the transmit power constraint. Thus, the superposition coded signal x_{dl} transmitted on the downlink is given by

$$x_{dl} = \sqrt{\alpha P_{dl}}\, s_A^{dl} + \sqrt{(1-\alpha)P_{dl}}\, s_B^{dl}, \tag{1.1}$$

where $\alpha \in [0, 1]$ is the fraction of total downlink transmit power allocated to user A and s_k^{dl} is the message intended for user $k \in \{A, B\}$ such that $\mathbb{E}\left|s_k^{dl}\right|^2 = 1$. The signals y_A^{dl} and y_B^{dl} received by users A and B are respectively given by

$$y_A^{dl} = h_A' \left(\sqrt{\alpha P_{dl}}\, s_A^{dl} + \sqrt{(1-\alpha)P_{dl}}\, s_B^{dl} \right) + w_A, \tag{1.2}$$

$$y_B^{dl} = h_B' \left(\sqrt{\alpha P_{dl}}\, s_A^{dl} + \sqrt{(1-\alpha)P_{dl}}\, s_B^{dl} \right) + w_B, \tag{1.3}$$

where h_k' is the complex-conjugate of channel coefficient h_k from BS to user $k \in \{A, B\}$ and $w_k \sim \mathcal{CN}(0, \sigma^2)$ denotes independent additive white Gaussian noise (AWGN) at user k. Without loss of generality, let $|h_A'| \geq |h_B'|$, i.e., user A be the stronger user and user B be the weaker user.

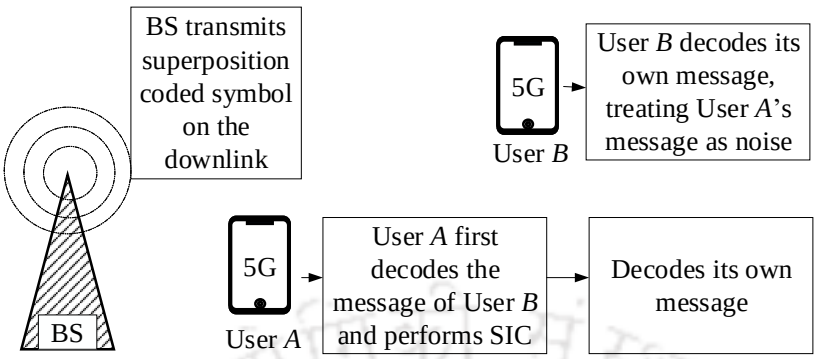


Fig. 1.2: Downlink NOMA.

On the downlink, the optimal decoding order to ensure successful SIC is - the stronger user (user A) decodes the message of the weaker user (user B) and cancels the interference from user B to itself (user A) before decoding its own message [9]. Thus, the onus is on user A to cancel the interference from user B. On the other hand, user B decodes its message from the received signal by treating the message of user A as noise. This causes intra-group interference to user B. Assuming availability of perfect CSI at each user, the achievable SE (in bits/s/Hz) for user B equals

$$R_{B,dl}^{\text{NOMA}} = \log_2\left(1 + \text{SINR}_{B,dl}^{\text{NOMA}}\right), \tag{1.4}$$

where the signal-to-interference-plus-noise-ratio (SINR) at user B equals

$$\text{SINR}_{B,dl}^{\text{NOMA}} = \frac{(1-\alpha)P_{dl}|h'_B|^2}{\sigma^2 + \alpha P_{dl}|h'_B|^2}. \tag{1.5}$$

And the achievable SE for user A equals

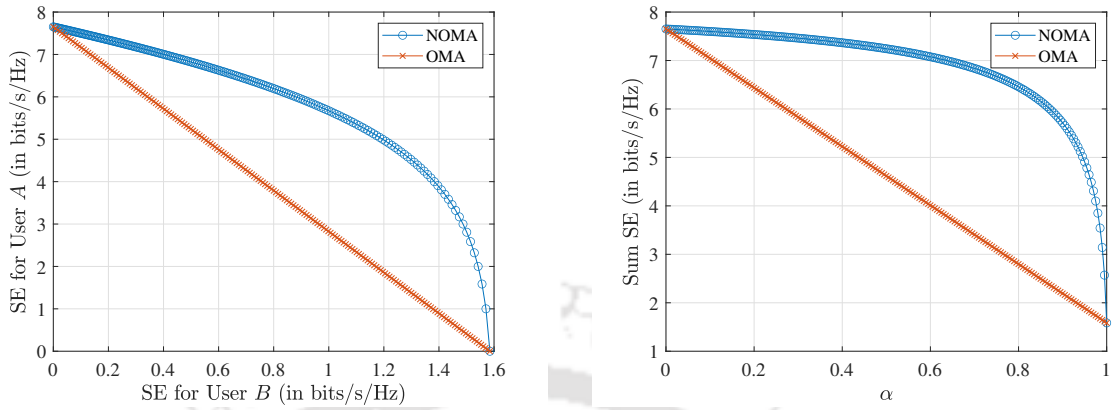
$$R_{A,dl}^{\text{NOMA}} = \log_2\left(1 + \text{SINR}_{A,dl}^{\text{NOMA}}\right), \tag{1.6}$$

where the SINR at user A equals

$$\text{SINR}_{A,dl}^{\text{NOMA}} = \frac{\alpha P_{dl}|h'_A|^2}{\sigma^2}. \tag{1.7}$$

By varying α, different rate-pairs in the capacity region can be achieved [9].

On the contrary, in orthogonal multiple access (OMA), resources are allocated orthogonally across users to avoid any interference. Examples of OMA are TDMA, FDMA and code-division



(a) SE of user A vs SE of user B **(b)** Sum SE vs α

Fig. 1.3: SE comparison between NOMA and OMA (downlink): $|h_A| = \frac{\sqrt{5}}{10}$, $|h_B| = \sqrt{5}$, $P_{dl} = 40$ W, $\sigma^2 = 1$.

multiple access (CDMA). In TDMA, users operate in non-overlapping time-slots, whereas in FDMA, users operate in orthogonal frequency bands. In CDMA, users are assigned orthogonal codes to avoid interference. When an OMA strategy, e.g., TDMA is employed, the achievable SE for user k equals

$$R_{k,dl}^{\text{OMA}} = \alpha_k \log_2\left(1 + \text{SINR}_{k,dl}^{\text{OMA}}\right), \tag{1.8}$$

where the SINR at user k equals

$$\text{SINR}_{k,dl}^{\text{OMA}} = \frac{P_{dl}|h_k'|^2}{\sigma^2}. \tag{1.9}$$

The multiplication factor α_k in the achievable SE accounts for the fraction of resource allocated to user $k \in \{A, B\}$, such that $\alpha_A + \alpha_B = 1$. Moreover, SE of each user is directly proportional to the amount of resource allocated to it. Similarly, the sum SE of both users is given by

$$R_{\text{sum},dl}^{\text{OMA}} = R_{A,dl}^{\text{OMA}} + R_{B,dl}^{\text{OMA}} = \alpha_A \log_2\left(1 + \text{SINR}_{A,dl}^{\text{OMA}}\right) + \alpha_B \log_2\left(1 + \text{SINR}_{B,dl}^{\text{OMA}}\right), \tag{1.10}$$

which is the sum of two linear functions of α. The performance gain achieved with NOMA over OMA can be observed in Fig. 1.3.[3] Specifically, in Fig. 1.3(a), we can observe the superior performance NOMA over OMA, in terms of achievable rate pairs as α is varied for NOMA and α_k is varied for OMA. Corresponding sum SE comparison is plotted in Fig. 1.3(b).

[3]Deterministic channel model is used for the ease of exposition.

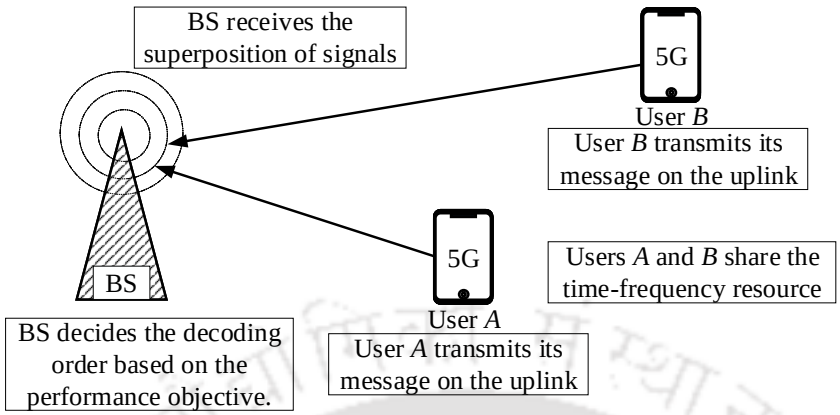


Fig. 1.4: Uplink NOMA.

1.2.1.2 Uplink NOMA

The uplink NOMA is in principle different from downlink NOMA, as explained in Fig. 1.4. In contrast to the downlink, on the uplink, each user transmits its message using the transmit power available at its end. Let P_{ul}^{k} be the uplink transmit power available at user $k \in \{A, B\}$. Therefore, the signals transmitted on the uplink by users A and B are, respectively, given by

$$x_A = \sqrt{P_{ul}^{A}} s_A^{ul}, \tag{1.11}$$

$$x_B = \sqrt{P_{ul}^{B}} s_B^{ul}, \tag{1.12}$$

where s_k^{ul} is the uplink message transmitted by user $k \in \{A, B\}$ and they are chosen such that $\mathbb{E}\left|s_k^{ul}\right|^2 = 1$. The signal received at the BS is given by

$$y_{BS}^{ul} = \sqrt{P_{ul}^{A}} h_A s_A^{ul} + \sqrt{P_{ul}^{B}} h_B s_B^{ul} + w_{BS}, \tag{1.13}$$

where $w_{BS} \sim \mathcal{CN}(0, \sigma^2)$ denotes the AWGN at BS.

On the uplink, the onus to do SIC is on the BS, and BS is free to decide the SIC decoding order depending on the desired rate-pair within the capacity region shown in Fig. 1.5 [7, 13]. For example, suppose the received power of the weaker user's signal is very less compared to the received signal power of the stronger user. In that case, the BS first decodes the message of the stronger user (user A) and cancels the interference that it creates to the weaker user (user B) before decoding the message of the weaker user. This decoding order guarantees fairness and maximizes the weaker user's achievable SE [7,13]. With this decoding order, point P in Fig. 1.5

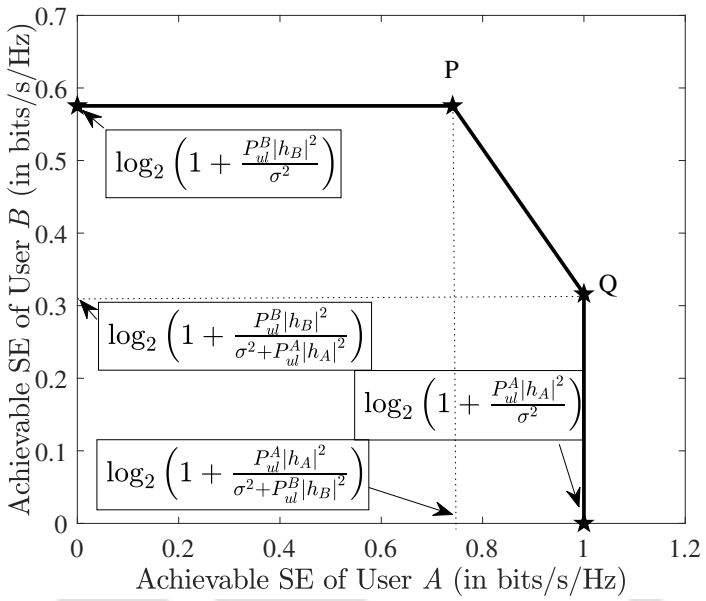


Fig. 1.5: Capacity region of uplink NOMA: $|h_A| = 1, |h_B| = 0.7, P_{ul}^A = 0.1$ W, $P_{ul}^B = 0.1$ W, $\sigma^2 = 0.1$.

is achieved. Therefore, if the objective is to improve fairness, the SINRs for users A and B are given by

$$\text{SINR}_A^{ul} = \frac{P_{ul}^A|h_A|^2}{\sigma^2 + P_{ul}^B|h_B|^2} \tag{1.14}$$

and

$$\text{SINR}_B^{ul} = \frac{P_{ul}^B|h_B|^2}{\sigma^2}, \tag{1.15}$$

respectively.

On the other hand, if the BS first decodes the message of the weaker user (user B) and cancels the interference that it creates to the stronger user (user A) before decoding the message of the stronger user, then the performance of the stronger user improves, while the performance of the weaker user deteriorates further [7], [13]. With this decoding order, point Q in Fig. 1.5 is achieved. Therefore, with the decoding order to achieve point Q in Fig. 1.5, the SINRs for users A and B are given by

$$\text{SINR}_A^{ul} = \frac{P_{ul}^A|h_A|^2}{\sigma^2} \tag{1.16}$$

and

$$\text{SINR}_B^{ul} = \frac{P_{ul}^B|h_B|^2}{\sigma^2 + P_{ul}^A|h_A|^2}, \tag{1.17}$$

respectively. These two strategies lead to achieving very different points in the achievable

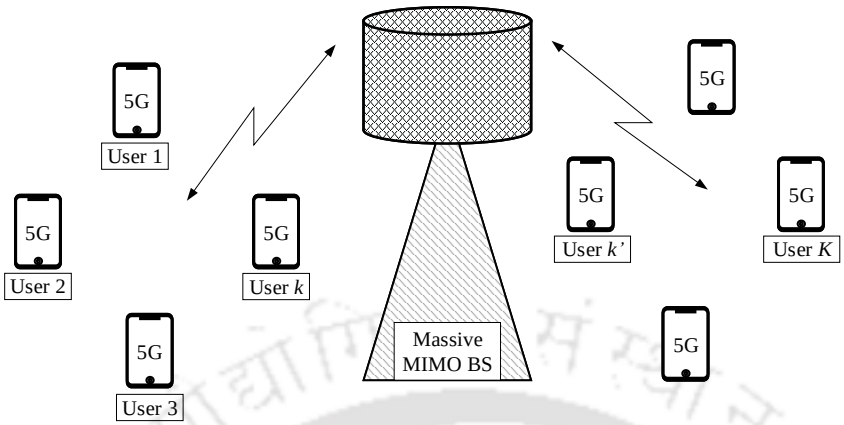


Fig. 1.6: Single-cell massive MIMO.

capacity region as can be seen in Fig. 1.5.

1.2.2 Massive MIMO

Massive MIMO employs space-division multiple access (SDMA) to achieve multiplexing gain and serve tens of users on the same time-frequency resource using hundreds of antennas at the BS [14]. The large number of antennas at the BS facilitates channel hardening and favorable propagation in the asymptotic regime [14]. While channel hardening implies that, when the number of antennas at the BS grows very large, the gain of the random fading channel tends to its mean value [15]. The favorable propagation entails that the directions of channel vectors corresponding to different users from the BS become pairwise orthogonal as the number of antennas grows very large [15]. Thus, massive MIMO systems reap the benefits of channel hardening and favorable propagation to improve the system performance. Next, we briefly discuss the signal models for uplink and downlink massive MIMO.

1.2.2.1 Uplink Massive MIMO

In uplink data transmission phase, the processing is done at the BS, while users adjust their symbol powers using power control coefficients and transmit their symbols in synchronism. We assume that these symbols are independent and have zero mean with unit variance. Each BS antenna receives the signal transmitted on the uplink, and then processes it using a linear decoder to detect the transmitted signal q_k corresponding to user k. The signal transmitted by

user k is given by

$$x_k = \sqrt{\gamma_k} q_k, \tag{1.18}$$

where $0 \leq \gamma_k \leq 1$ denotes the corresponding power control coefficient. Thus, the signal received by the BS is given by

$$\mathbf{y} = \sqrt{\rho_{ul}} \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1.19}$$

where ρ_{ul} denotes the uplink transmit power of each user, $\mathbf{H} \in \mathbb{C}^{M \times K}$ denotes the channel matrix whose $(m,k)^{\text{th}}$ element is the channel coefficient between user k and m^{th} BS antenna, $\mathbf{x} = [x_1, \ldots, x_K]^T$ is the transmit symbol vector, and $\mathbf{w} = [w_1, \ldots, w_K]^T$ is the noise vector. The BS applies linear decoder $\mathbf{L}$ to the received signal $\mathbf{y}$ and the processed signal is given by [14]

$$\overline{\mathbf{y}} = \sqrt{\rho_{ul}} \mathbf{L}'\mathbf{H}\mathbf{x} + \mathbf{L}'\mathbf{w}. \tag{1.20}$$

The linear decoder $\mathbf{L}$ is a function of the channel estimates and may take different forms depending on the linear processing technique used. For example, under maximum ratio (MR) processing, $\mathbf{L} \in \mathbb{C}^{M \times K}$ takes the form given by [14]

$$\mathbf{L} = \widehat{\mathbf{H}}, \tag{1.21}$$

and under zero-forcing (ZF) processing, $\mathbf{L}$ takes the form given by [14]

$$\mathbf{L} = \widehat{\mathbf{H}} \left(\widehat{\mathbf{H}}' \widehat{\mathbf{H}} \right)^{-1}, \tag{1.22}$$

where $\widehat{\mathbf{H}}$ denotes the channel estimate matrix available at the BS.

1.2.2.2 Downlink Massive MIMO

Prior to downlink data transmission, the BS linearly combines the data symbols for different users using a linear precoder $\mathbf{L} \in \mathbb{C}^{M \times K}$. Thus, the downlink transmit symbol vector $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is given by

$$\mathbf{x} = \mathbf{L}\mathbf{D}_\gamma^{\frac{1}{2}} \mathbf{q} \tag{1.23}$$

where $\mathbf{q} = [q_1, q_2, \ldots, q_K]^T$ is the data symbol vector and $\mathbf{D}_\gamma = \text{diag}\,[\gamma_1, \ldots, \gamma_K]$ is the diagonal power allocation coefficient matrix such that $\sum_{k=1}^{K} \gamma_k = 1$. The collective signal received at K

users is given by [14]

$$\mathbf{y} = \sqrt{\rho_{dl}}\mathbf{H}'\mathbf{x} + \mathbf{w} \tag{1.24}$$

$$= \sqrt{\rho_{dl}}\mathbf{H}'\mathbf{L}\mathbf{D}_{\gamma}^{\frac{1}{2}}\mathbf{q} + \mathbf{w}, \tag{1.25}$$

where ρ_{dl} is the downlink transmit power. User k decodes its data symbol from the observation y_k, which is essentially the k^{th} element of $\mathbf{y}$. For ZF processing, the linear precoder $\mathbf{L}$ takes the form given by [14]

$$\mathbf{L} = \sqrt{M-K}\widehat{\mathbf{H}}\left(\widehat{\mathbf{H}}'\widehat{\mathbf{H}}\right)^{-1}, \tag{1.26}$$

and for the MR processing, the linear precoder $\mathbf{L}$ takes the form given by [14]

$$\mathbf{L} = \frac{1}{\sqrt{M}}\widehat{\mathbf{H}}. \tag{1.27}$$

1.2.3 IRS

An IRS can be used to modify the behaviour of the wireless propagation environment. It is essentially a metasurface consisting of large number of reflecting elements and it is connected to the BS via a controller as shown in Fig. 1.7. The reflecting elements can be used to induce amplitude and phase changes in the incident signal so as to make the reflected multipath components add up in phase at the receiver. To configure the phases at each IRS element, PIN diodes are used. By adjusting the biasing voltages on the PIN diode through a direct-current (DC) feeding line, the diode can be made to transition between the "ON" and "OFF" states. This enables the device to induce a phase-shift of π radians on the signal that is incident on it. IRS can potentially be used to provide coverage to users experiencing blockage and to boost the received signal strength at the receiver. Essentially, the IRS can steer signals around obstacles, create stronger paths between nodes, and reduce unwanted noise or interference, all of which improve network performance [16].

IRS also has several practical benefits. Firstly, the IRS elements beamform the incident signals without requiring dedicated radio frequency (RF) chains. Thus, IRS can be implemented at much lower hardware cost and consumes lesser energy compared to standard active antennas

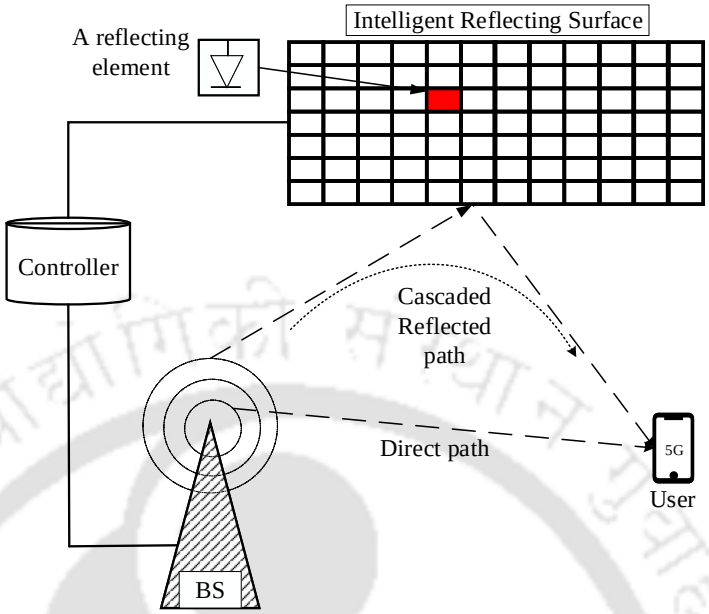


Fig. 1.7: IRS-aided wireless system.

with dedicated RF chains [17]. Also, IRS can work in a way where it receives and reflects signals at the same time and doesn't increase noise or create interference for itself. This makes it better than the relay systems which are prone to these issues.

To understand the performance gains that may be achieved using an IRS with T elements, consider a simple single input single output (SISO) system serving a single-antenna user on the downlink, as shown in Fig. 1.7. Let x denote the transmitted signal, then the received signal is given by

$$y = (h_{\mathrm{D}} + \mathbf{h}_{\mathrm{B\text{-}I}}' \mathbf{\Phi} \mathbf{h}_{\mathrm{I\text{-}U}})\, x + w, \tag{1.28}$$

where h_{D} denotes the channel coefficient for the direct path from BS to the user, $\mathbf{h}_{\mathrm{B\text{-}I}} \in \mathbb{C}^T$ denotes the channel coefficient vector from BS to IRS, $\mathbf{\Phi} \in \mathbb{C}^{T \times T}$ denotes the diagonal phase shift matrix corresponding to the IRS, $\mathbf{h}_{\mathrm{I\text{-}U}} \in \mathbb{C}^T$ denotes the channel coefficient vector from IRS to user and w denotes AWGN at receiver. Note that

$$\mathbf{\Phi} = \mathrm{diag}\left[e^{j\phi_1}, \ldots, e^{j\phi_T}\right], \tag{1.29}$$

where ϕ_t denotes the phase shift induced by the t^{th} IRS element, for $t \in \{1, \ldots, T\}$. Thus, (1.28)

can be expanded as

$$y = \left(|h_{\mathrm{D}}| \, e^{j\phi_{h_{\mathrm{D}}}} + \sum_{t=1}^{T} \left|h_{\mathrm{B\text{-}I}}^{t}\right| \left|h_{\mathrm{I\text{-}U}}^{t}\right| e^{j\left(\phi_{h_{\mathrm{B\text{-}I}}^{t}} + \phi_{h_{\mathrm{I\text{-}U}}^{t}} + \phi_{t}\right)} \right) x + w. \tag{1.30}$$

Assuming that the BS has access to perfect CSI, i.e., the BS accurately knows h_D, $\mathbf{h}_{\mathrm{B\text{-}I}}$, and $\mathbf{h}_{\mathrm{I\text{-}U}}$, then the BS can set ϕ_t as

$$\phi_t = \phi_{h_{\mathrm{D}}} - \left(\phi_{h_{\mathrm{B\text{-}I}}^{t}} + \phi_{h_{\mathrm{I\text{-}U}}^{t}}\right), \tag{1.31}$$

to ensure that the signal reflected from each IRS element adds up in phase with the signal reaching at the receiver via direct path [18]. Thus, in comparison to the system without IRS, now there exist T extra paths through which the signal reaches the receiver and by appropriate phase programming at IRS, the received signal strength can be boosted.

1.3 Literature Review

In this section, we present and discuss the existing literature to motivate new contributions of the thesis. NOMA, massive MIMO and IRS are investigated extensively in literature individually and in collaboration with other technologies. We first discuss the literature on massive MIMO-NOMA systems without channel aging followed by the literature on massive MIMO-NOMA systems under channel aging. Next, we discuss the literature on IRS-aided NOMA systems. Finally, we discuss the research gaps in the literature.

1.3.1 Massive MIMO-NOMA Systems without Channel Aging

Massive MIMO-NOMA systems are studied extensively in the literature. These studies can be found in [10, 19–40] and references therein. Specifically, in [10], authors proposed a transmission scheme for downlink massive MIMO-NOMA systems and obtained the results for the outage probability. Further, authors in [19] proposed antenna selection algorithm for downlink massive MIMO-NOMA systems considering the practical constraint of limited number of RF chains at the BS. Outage analysis was presented for the proposed algorithm. Next, in [21], authors analyzed achievable rate for a multicell downlink massive MIMO-NOMA system. Here, authors proposed a pilot sharing scheme to reduce the pilot overhead. They also proposed a

max-min power control scheme for a downlink massive MIMO-NOMA system in order to ensure fairness in users' performance.

At the same time, multicell downlink massive MIMO-NOMA systems were studied to analyze the impact of pilot contamination and imperfect SIC in [22]. Here, the authors re-established that massive MIMO-NOMA system's performance is significantly impacted by the particular power allocation algorithm in place. In addition, K-tier heterogeneous networks were analyzed for downlink in [23], when massive MIMO is employed for macro-cells and NOMA is employed for all small cells. Corresponding downlink SE expressions were derived. Impact of channel estimation error on the achievable downlink SE of massive MIMO-NOMA system was analyzed in [24]. A user pairing algorithm is proposed for downlink massive MIMO-NOMA systems in [26] and sum rate expressions were derived. Further, in [29], authors proposed a hybrid NOMA and multi-user BF scheme for downlink massive MIMO-NOMA systems, which was shown to perform better than either of the standalone schemes. A random user pairing scheme was proposed in [31] in order to minimize outage probability in downlink mmWave massive MIMO-NOMA systems.

Furthermore, the authors in [36, 41] explored the performance of mmWave NOMA-based massive MIMO system assuming availability of perfect CSI at the BS. Also, in [37], the authors proposed a precoder design for a two-user uplink MIMO-NOMA system assuming availability of perfect CSI. However, in practice, CSI will have to be estimated. In [38,40], authors analyzed the downlink SE of a NOMA-based massive MIMO system with MR precoding under imperfect CSI at the BS for independent Rayleigh fading channel, whereas in [39], the authors analyzed the downlink SE with ZF precoding based on imperfect CSI at the BS for spatially correlated Rayleigh fading channel.

1.3.2 Massive MIMO Systems under Channel Aging and Massive MIMO-NOMA Systems under Channel Aging

As discussed in detail above, there is a vast amount of literature on massive MIMO-NOMA systems without incorporating channel outdatedness. Likewise, the impact of channel aging on massive MIMO systems standalone (without employing NOMA transmission) has also been

studied quite well in the literature. For instance, performances of uplink and downlink of massive MIMO systems under outdated CSI due to channel aging and with predicted channel were first discussed in [8]. Further, the authors in [42] considered a massive MIMO system and derived closed-form lower bounds on the uplink and downlink SE under channel aging. Therein, the authors considered MR and ZF processing. In [43], the authors studied how channel outdatedness impacts downlink massive MIMO system aided by multi-way relay networks.

Next, the authors in [44] considered MR and regularized ZF precoders to derive the SINR's deterministic equivalent for downlink massive MIMO systems. Therein, the authors considered a channel aging model that captures the channel estimates' outdatedness due to Doppler shift and noisy local oscillators. Further, in [45], the authors addressed the throughput maximization problem for downlink massive MIMO systems under channel aging. At the same time, an opportunistic user scheduling algorithm for downlink massive MIMO was proposed in [46] to enhance SE using aged CSI. Next, the uplink massive MIMO system was analyzed in [47] using regularized ZF and ZF receivers designed based on outdated estimates. Similarly, the uplink ergodic rate with ZF receiver in massive MIMO system considering outdated channel was derived in [48].

Furthermore, in [49], the authors investigated frequency-division duplexing (FDD) massive MIMO systems under channel aging. They derived lower bounds on SE with MR, minimum mean square error (MMSE) decoders for uplink, and MR, regularized ZF precoders for downlink. Next, the authors in [50] optimized the frequency of updating the CSI to strike a balance between the estimation accuracy and estimation overhead. In a similar direction, the authors in [51] proposed a novel channel estimation scheme to enhance the SE of a downlink massive MIMO system under channel aging. Also, in [52], the authors proposed an intermittent channel estimation scheme to enhance the sum-SE of downlink massive MIMO system by exploiting the temporal correlation under channel aging. At the same time, the authors in [53] explored how channel outdatedness influences downlink massive MIMO systems in an urban environment using ray tracing.

However, there is very limited work that studies massive MIMO-NOMA systems under

channel aging. For instance, the authors in [54] studied downlink massive MIMO-NOMA systems with MR precoder under channel aging. They derived the SE for the strong and weak users and compared their performances under ad-hoc power control strategies, namely, equal and inversion power control.

1.3.3 IRS-Aided NOMA Systems

While IRS was being explored as an alternative technology to large antenna systems due to its passive nature, it was also being explored in collaboration with other candidate technologies like NOMA, which is of interest to us. For instance, sum SE maximization problems for IRS-aided NOMA systems were discussed in [55–59]. Specifically, in [55], the authors investigated the downlink communications of IRS-assisted NOMA systems. To maximize system capacity, a joint optimization problem was formulated to address the decoding order, power allocation, and reflection coefficients. To tackle this problem, a novel two-step iterative algorithm was proposed. In [56], the authors examined a two-user downlink IRS-assisted network over fading channels, focusing on NOMA and OMA schemes, specifically TDMA and FDMA. The objective was to maximize the system's average sum rate for delay-tolerant transmission while considering power budget, minimum average user rate, and discrete unit modulus reflection coefficient constraints. Low-complexity phase shifter adjustment algorithms were proposed.

Next, in [57], the study focused on an IRS-assisted uplink NOMA system, aiming to maximize the sum rate of all users under individual power constraints. The problem, which required joint power control at the users and BF design at the IRS, was non-convex. Semidefinite relaxation was employed to address it. Furthermore, in [59], the authors focused on joint user clustering, passive BF, and power allocation in a downlink IRS-assisted NOMA system, aiming to maximize energy efficiency. The optimization problem was solved iteratively. Further, the investigations to ascertain fairness in users' performance were provided in [60, 61] and outage probability analyses for IRS-aided NOMA systems were conducted in [62–66]. For instance, in [60] authors examined the deployment of IRS between a single-antenna BS and multiple single-antenna users to assist downlink NOMA transmission. To ensure fairness among users, the study aimed to jointly optimize power allocation, decoding order, and phase shifts to maxi-

mize the minimum user rate under a total power constraint.

Similarly, in [61] the authors analyzed an IRS-assisted NOMA in which a combined-channel-strength (CCS) based user ordering scheme was proposed. To optimize rate performance and ensure user fairness, the study maximized the minimum decoding SINR by jointly optimizing power allocation at the BS and phase shifts at the IRS. In addition, the authors in [62] examined the impact of hardware impairments on IRS-aided NOMA systems, focusing on performance metrics. Analytical expressions for outage probability and throughput were derived as KPIs. Likewise, in [63] authors studied a downlink IRS-aided NOMA system to enhance coverage by assisting a cell-edge user device (UD) in communicating with the BS. New channel statistics for the BS-IRS-UD link under Nakagami-*m* fading were investigated, and based on these statistics, a closed-form expression for the outage probability was derived.

In [64] the authors derived the best-case and worst-case scenarios of channel statistics to characterize effective channel gains in IRS-aided NOMA systems. Subsequently, the best-case and worst-case closed-form expressions for both outage probability and ergodic rate of the prioritized user were derived. In [65], the authors presented a simple design for IRS-assisted NOMA downlink transmission. Initially, conventional SDMA was employed at the BS to create orthogonal beams using the spatial directions of near users' channels. Then, IRS-assisted NOMA was utilized to serve additional cell-edge users by aligning their effective channel vectors with the established spatial directions. considered IRS-aided NOMA systems but did not analyze the error performance. Furthermore, in [66] the authors investigated the downlink performance of IRS-aided NOMA networks using stochastic geometry. The closed-form expressions for the coverage probabilities of paired NOMA users were derived. Similarly, in [67], a comparison between NOMA and OMA in terms of achievable rate in IRS-assisted downlink communication was conducted, focusing on transmit power minimization under the discrete unit-modulus reflection constraint for each IRS element. The study analyzed and numerically compared the minimum transmit powers required by different multiple access schemes.

Furthermore, the error probability for IRS-aided NOMA systems was analyzed in [68, 69] and recently in [70]. Specifically, the study in [68] examined bit error rate (BER) of an IRS-

Table 1.2: Research gaps in the literature on uplink massive MIMO-NOMA systems.

Paper	Research gaps
[34]	• Focus is on mmWave massive MIMO-NOMA systems • No analysis with imperfect CSI • No focus on improving the weaker users' data rate • No focus on channel outdatedness
[35]	• Focus on cell-free massive MIMO-NOMA systems • No analysis for improving weaker users' data rate • Employed MR decoder causes intra-group interference • Did not consider channel outdatedness
[27]	• No analysis with imperfect CSI • Power allocation does not ensure uniform quality of service • Did not consider channel outdatedness
[11]	• Considered MR decoder causes intra-group interference limiting the performance • Did not consider channel outdatedness
[54]	• Considered downlink massive MIMO-NOMA system • MR precoder is employed • Focus on ad-hoc power control

aided NOMA system using quadrature phase shift keying (QPSK)-binary phase shift keying (BPSK) modulation scheme, specifically when the direct link is blocked and the indirect links through the IRS experience Rayleigh fading. Additionally, in [69], the authors presented an approximate pairwise error probability (PEP) analysis for an IRS-aided NOMA system. Similarly, in [70], the authors analyzed the BER performance of IRS and Relay-assisted NOMA systems with interference cancellation at the relay node.

1.3.4 Research Gaps

As discussed above, very little attention has been paid to analyze uplink massive MIMO-NOMA and address the challenges therein. For instance, energy efficient power allocation policies were proposed in [34] and [35] for uplink mmWave massive MIMO-NOMA systems and uplink cell-free massive MIMO-NOMA systems, respectively. In [34], authors also proposed a hybrid analog-digital BF scheme considering limited number of RF chains. In [35], authors

proposed a power control algorithm to ensure maximization of energy efficiency in uplink cell-free massive MIMO-NOMA systems considering the impact of imperfect CSI. In [27], authors proposed a receive antenna selection algorithm and a power allocation scheme to maximize sum-rate in uplink massive MIMO-NOMA systems assuming perfect CSI. Further, achievable rate analysis and max-min power control for uplink massive MIMO-NOMA system under imperfect CSI was discussed in [11] for MR decoder. This is summarized in Table 1.2.

The SE of uplink massive MIMO-NOMA systems under ZF processing has not been analyzed in the literature with and without considering the impact of channel aging. Also, the impact of channel aging on the performance of MR processing in massive MIMO-NOMA systems has been studied solely for the downlink scenario as discussed above [54]. This is summarized in Table 1.2. However, the working principles of downlink NOMA are different from those of uplink NOMA [7, 13]. In addition, channel estimation error and channel outdatedness lead to imperfect SIC at the BS. Investigations on power control to ensure fairness in users' performance for uplink massive MIMO-NOMA systems are very limited in the literature.

It is also crucial to understand what counter measures we can take to mitigate the impact of channel outdatedness in uplink massive MIMO-NOMA systems, which is missing in the existing literature. Therefore, it is important to answer fundamental and open questions pertaining to modelling and performance analysis of uplink massive MIMO-NOMA systems under the collective impact of channel estimation errors, channel aging and pilot contamination.

Furthermore, we observe that there exists very limited literature on error performance analysis IRS-aided NOMA systems [68], [69], [70]. Although [68], [69] and [70] analyzed the error performance of IRS-aided NOMA systems, their analysis is applicable only when there is no direct link between the BS and users, and the reflected links via IRS are Rayleigh faded. Moreover, in [68], BER expressions were derived for a specific constellation pair (QPSK-BPSK) and were not in closed-form, in [69] approximate expression for the PEP was derived, and the analysis presented in [70] is not in closed-form and is applicable specifically for BPSK modulation. This is summarized in Table 1.3.

Based on the discussion above, we observe that the error performance analysis of NOMA

Table 1.3: Research gaps in the literature on downlink IRS-aided NOMA systems.

Paper	Research gaps
[68]	• No direct link between the BS and users • Indirect channels via IRS are assumed to be Rayleigh faded • Analysis specific to QPSK-BPSK modulation pair • No closed-form expressions for BER are provided
[69]	• No direct link between the BS and users • Indirect channels via IRS are assumed to be Rayleigh faded • Approximate expressions for PEP are derived
[70]	• No direct link between the BS and the far user • Indirect channels via IRS are assumed to be Rayleigh faded • A relay node is employed for SIC • Upper bound on BER is provided specifically for BPSK modulation

systems with IRS in wireless communication network is still in its infancy and remains an open problem. Employment of an IRS to aid communication in NOMA systems poses several new challenges as the signal reaches users via direct path and via reflected path. For instance, the phase configuration at the IRS becomes crucial while serving multiple users simultaneously.

In addition, it is crucial to account for the effects of practical constraints of IRS-aided NOMA systems such as discrete-phase shifts at IRS and imperfect CSI, which further lead to non-coherent reception at receiving users and imperfect SIC. Thus, it is crucial to analyze the error performance of IRS-aided NOMA systems in a two-user scenario and a multi-user scenario for any arbitrary modulation scheme. It is also essential to characterize the effects of non-coherent reception at users due to discrete-phase shifts at IRS and/or due to imperfect CSI leading to imperfect SIC. Apart from this, while serving users with widely varying channel conditions, it becomes imperative to design the system to ensure fairness in user performance, which could be obligatory in certain applications with stringent quality-of-service requirements.

1.4 Key Challenges

Although, the considered technologies hold the potential to achieve the desired KPIs, there exist several challenges in practice. First and foremost, owing to the fundamental differences between uplink and downlink NOMA, it is essential to address the challenges in uplink massive MIMO-NOMA separately. For instance, when the aim is to serve a large number of users under the practical constraint of limited coherence interval, it is vital to develop low-overhead channel estimation schemes.[4] In addition, it is essential to reduce inter-user interference to achieve the desired system performance. It is equally important to investigate the interplay among channel estimation, decoder, power control and user grouping strategies. Furthermore, it is critical to ensure fairness in service for different users regardless of the disparity in their channel conditions.

While designing a massive MIMO-NOMA system with the aim of serving a very large number of users, it is crucial to understand that the estimated CSI may get outdated by the time it is used for processing at the BS. This could potentially be due to dynamic wireless environments and high mobility users and this phenomenon is referred to as channel aging. Linear processing at the massive MIMO BS and SIC are functions of CSI available at the BS. Thus, to ensure wireless connectivity in high speed mobility scenario, it becomes critical to understand the impact of channel aging on the performance of uplink massive MIMO-NOMA systems. Furthermore, it is equally important to develop counter measures to mitigate the effects of channel aging on system performance.

As we know, the effective number of degrees-of-freedom offered by massive MIMO essentially depends on the number of active RF chains comprising of signal converters, mixers, and power amplifiers. Thus, the benefits of massive MIMO-enabled systems come at the cost of increased hardware complexity and power consumption. Therefore, it is essential to explore the cost-efficient alternative to large antenna systems, such as IRS-aided systems, and address the

[4] As we know that the channel estimation in massive MIMO systems requires sending uplink pilots, estimation takes part of the coherence interval and the remaining part of the coherence interval is used for data transmission. The portion of the coherence interval used for channel estimation is termed as the channel estimation overhead. Low-overhead schemes ensure that the channel estimation overhead is low.

challenges therein. With the objective of ensuring reliability in data detection, there is a strong motivation behind exploring error probability of IRS-aided NOMA systems, specifically, under the comprehensive and generic scenario where the direct path from BS to each user may be present in addition to the cascaded reflected path via IRS. In addition, it is crucial to consider the practical constraints of discrete phase shifts at IRS and imperfect CSI-induced non-coherent reception at users. Again, it is important to ensure uniform quality of service among users while designing a practically relevant IRS-aided NOMA system.

1.5 Problem Statements

Considering the research gaps in the literature on massive MIMO-NOMA systems as tabulated in Table 1.2, and considering the research gaps in the literature on downlink IRS-aided NOMA systems as tabulated in Table 1.3, we formulate the following problem statements:

- **Problem Statement 1:** Investigation of effects of channel estimation schemes and user grouping strategies on the SE of uplink massive MIMO-NOMA systems, and exploring the interplay between channel estimation schemes and user grouping strategies. Examination of how the decoding and power control schemes based on the estimated CSI affect the SE performance of uplink massive MIMO-NOMA systems.
- **Problem Statement 2:** Examining the effects of channel aging on pilot-sharing-based channel estimation schemes with different user grouping strategies on SE performance of uplink massive MIMO-NOMA systems. Further investigations on how various decoding and power control schemes based on the outdated estimated CSI affect SE of uplink massive MIMO-NOMA systems.
- **Problem Statement 3:** Examining the effectiveness of the decoders designed based on predicted channel as a counter measure to channel aging on the SE performance of uplink massive MIMO-NOMA systems and investigation of the performance under various power control schemes.
- **Problem Statement 4:** Investigating the role of IRSs in improving the error performance of downlink NOMA systems for pulse amplitude modulation (PAM)/quadrature

amplitude modulation (QAM) modulation of arbitrary order for two-user and multi-user scenarios. Also, understanding the effects of practical system constraints, like channel estimation errors and discrete phase shifts at IRS, on the error probability in downlink IRS-aided NOMA systems, and exploring how we can employ IRSs to ensure performance fairness among users.

1.6 Thesis Objectives

With the motivation of achieving improved connectivity, fairness in terms of SE, and reliability in terms of signal detection, the specific objectives of the thesis are as follows:

- Firstly, the objective is to investigate an uplink massive MIMO-NOMA system under practical constraints of limited coherence interval, imperfect CSI and imperfect SIC. Here, we explore two low-overhead channel estimation schemes to acquire CSI. Using the acquired channel estimates, we design ZF decoders to mitigate inter-group interference and analyze the SE performance. Here, we also explore power control to ensure fairness leading to uniform quality of service.
- Next, we aim to address the challenges involved in an uplink massive MIMO-NOMA system while serving a large number of users under high Doppler induced due to high speed mobility. In this direction, we design ZF decoders based on outdated channel estimates and we quantify the impact of channel outdatedness and imperfect SIC on the SE performance. We also explore channel prediction as a counter-measure to mitigate the impact of channel outdatedness and power control to ensure fairness while guaranteeing reasonable performance.
- Finally, we focus on downlink IRS-aided NOMA systems. Here, in order to characterize the interactions among SC at the BS, SIC at the strong user, IRS and the modulation technique employed, we analyze the reliability of data detection in terms of error probability for both two-user and multi-user scenarios for PAM/QAM of arbitrary order. Here, we also characterize the system performance under practical constraints of discrete phase shifts and imperfect CSI.

With these objectives, we state the contributions of the thesis in the next section.

1.7 Thesis Contributions

In this thesis, we focus on modelling, analysis and optimization of NOMA-based next-generation wireless communication systems with the aim of providing increased connectivity, fairness in terms of SE performance and reliability in terms of signal detection. Below, we provide a concise overview of the problems addressed, followed by an outline of our specific contributions.

1.7.1 Massive MIMO-NOMA Systems

We consider the uplink of a massive MIMO-NOMA system where, using the same time-frequency resources, a large number of users communicate with a massive MIMO BS equipped with hundreds of antennas. In order to employ receive BF, the BS needs to learn the uplink channel from each user. Since the aim is to serve a large number of users simultaneously and owing to the limited coherence interval in practical wireless systems, we employ non-orthogonal pilot assignment and develop two MMSE-based channel estimation schemes, namely, Scheme-I and Scheme-S, for estimating the channel coefficients on the uplink. To employ non-orthogonal pilot assignment and NOMA on the uplink, we form groups of users based on two practically realizable user grouping strategies, namely, near-far (NF) grouping and neighbor (NB) grouping. These user grouping strategies are based on large-scale fading coefficients. Given this setup, we make the following key contributions.

1.7.1.1 Contributions

- *Low Overhead Channel Estimation:* With the aim of strengthening the quality of channel estimates to weaker users, we analyze two different channel estimation schemes, namely, Scheme-I and Scheme-S. While in Scheme-I, channel vector of each user is estimated, in Scheme-S, sum of channel vectors of all users within a group is estimated. We also analyze two different user grouping strategies, namely, NF grouping and NB grouping.
- *Achievable SE Analysis:* We derive new lower bounds on the uplink achievable SE of a NOMA-based massive MIMO system using ZF decoders designed based on Scheme-I

and Scheme-S. Also, both the considered schemes provide correlated estimates for all the K users in a group which share the same pilot. Hence, the ZF decoders designed using these estimates allow $K\times$ connectivity, when compared against conventional ZF decoder designed using uncorrelated estimates obtained via orthogonal pilot signaling.

- *Max-Min Power Control:* To ensure uniform quality-of-service to all users in the network, for both the channel estimation schemes, we formulate and solve the max-min optimization problem where the objective is to determine the power allocation coefficients that maximize the minimum achievable SE.
- *Novel Insights:* Through our rigorous analysis and extensive simulations, we identify the operational regimes that favor the use of the ZF decoder to boost the performance in terms of the max-min achievable SE on the uplink and the number of users served. We observe that in the entire under-loaded regime and a large portion of the over-loaded regime, the ZF decoder designed based on channel estimation Scheme-S using NF grouping gives the highest max-min achievable SE.

1.7.2 Massive MIMO-NOMA Systems under Channel Aging

Next, we extend our study of massive MIMO-NOMA systems to incorporate the practically relevant case of Doppler-induced channel aging. As non-orthogonal pilot signal assignment, i.e., pilot sharing is required to keep the estimation overhead small, the system performance will be influenced by the combined effects of mobility-induced channel outdatedness and pilot-sharing-induced contamination in channel estimates. Given this model, our objective is to determine the combination of decoder, pilot control strategy and channel estimation scheme, which is robust to the cumulative effects of channel aging and pilot contamination. To this end, we make following key contributions.

1.7.2.1 Contributions

- *Channel Modelling with Outdated Estimates:* We consider a novel and practically relevant uplink massive MIMO-NOMA system in which channel outdatedness is characterized using Jakes model. As before, the channel estimates are acquired using two different

low-overhead schemes, namely, Scheme-I and Scheme-S. This modelling helps jointly capture the errors induced due to channel aging, estimation errors and pilot contamination.

- *SE Analysis under Channel Aging:* We derive novel lower bounds on the achievable SE for both ZF and MR decoders that are designed based on outdated channel estimates. This analysis accounts for the correlation due to pilot sharing, time-variant channel response due to mobility and the penalty incurred due to imperfect SIC. As before, we also investigate the impact of two different user grouping strategies, namely, NF grouping and NB grouping, on the per-user achievable SE.
- *Max-Min Fairness Power Control:* To ensure fairness in performance under channel aging, we formulate and solve optimization problems to compute max-min power control coefficients corresponding to ZF and MR decoders. For ZF decoder based on Scheme-S, the formulated problem is non-convex. We propose a tight approximation to convert it to a convex program.
- *Proportional Fairness Power Control:* In order to improve sum SE while maintaining fairness to a significant extent, we formulate optimization problems to derive coefficients for proportional fairness power control. The formulated problems are geometric programs and we solve them efficiently using convex programming.
- *SE Analysis using Predicted CSI:* Furthermore, we derive the p^{th} order Wiener linear predictor (WLP) for predicting the individual channel vectors to all users, using $(p+1)$ observations obtained through transmission of shared pilots. We derive the lower bounds on the achievable SE with ZF and MR decoders designed using the predicted CSI.

While we focused on comprehensive analysis of massive MIMO-NOMA systems so far, in the next part of the thesis, we investigate IRS-aided NOMA systems since these seem to be cost-efficient alternative to massive MIMO systems.

1.7.3 IRS-Aided NOMA Systems

In the last part of the thesis, we consider a single antenna BS serving single antenna users via NOMA, i.e., SC at the transmitter and SIC at the receivers, with the help of two IRSs. We

consider Rayleigh faded direct paths from the BS to each user and the links from the BS to each IRS and from each IRS to its corresponding user are assumed to be Rician faded. We analyze two-user scenario and then extend the analysis to multi-user scenario as well. With this comprehensive model, we make following contributions.

1.7.3.1 Contributions

- *Average SEP Analysis for Two-User PAM and QAM:* We derive the probability density function (PDF) of the effective channel gain in closed-form. Using the derived PDF, we derive novel approximate analytical expressions for the average symbol error probability (SEP) of the weak and strong users considering PAM constellation of arbitrary order for both users. As a special case, we also present average SEP expressions for both users when direct links from BS to each user are blocked. We further extend the SEP analysis to a scenario when both users adopt QAM constellations of arbitrary order for NOMA-based transmission.
- *Error Performance Analysis for Multi-User PAM:* For the multi-user scenario, we develop the average SEP expressions with multicast[5] transmission in IRS-aided NOMA system when PAM modulation of arbitrary order is used. Moreover, for the multi-user scenario, we also derive closed-form expression for the average PEP under unicast[6] transmission.
- *Impact of Discrete Phase Shifts and Imperfect CSI:* We characterize and elucidate the impact of discrete phase-shifts at IRS and imperfect CSI on the average SEP.
- *Fairness in SEP:* We formulate and solve a constrained optimization problem to determine the optimal fraction of elements to be activated at any IRS to guarantee near-identical SEP for both users in a two-user scenario.
- *Novel Insights:* Through extensive Monte Carlo simulations, we verify the tightness of derived novel approximate analytical expressions. Through our rigorous analysis, we show that the addition of a few passive elements at each IRS (i) helps achieve desired average SEP with significantly lower signal-to-noise ratio (SNR) for both users in a two-

[5]Under multicast transmission, we consider that the transmitted signal for all strong users is the same and similarly, the transmitted signal for all weak users is the same [71].

[6]Under unicast transmission, we consider that the transmitted signal for all users is different [71].

user scenario and for all users in a multi-user scenario, (ii) allows communication with higher-order constellation without substantial increase in average SEP/average PEP, and (iii) requires allocation of lower power to the weak user to attain the SEP identical to the strong user in a two-user scenario. We also show that a 3-bit programmable IRS is sufficient to obtain SEP very close to that obtained by an IRS with infinite precision phase-shift.

Connections Among Thesis Contributions: The first contribution of the thesis entails the performance analysis of uplink massive MIMO-NOMA systems to ensure increased connectivity and user fairness in terms of SE. Herein, we considered quasi-static fading channel which implies that the channel remains constant over the transmission interval. The channel estimation, uplink data transmission and decoding phases are performed within the coherence interval. However, considering mobility-induced Doppler shift, in the second contribution of the thesis, we assume channel aging, i.e., the channel estimates are outdated by the time they are used for processing at the BS. The decoders are designed based on the outdated channel which is obtained using Jakes model. Here, we quantify the impact of channel aging on the performance of uplink massive MIMO-NOMA systems. Next, as a counter-measure to channel aging, we investigate the performance of uplink massive MIMO-NOMA systems when decoders designed using predicted channel are employed. Noting that the large BF gain benefit of massive MIMO-NOMA systems is led by the large number of RF chains at the BS, we next study IRS-aided NOMA systems as a cost-efficient alternative to massive MIMO systems. Herein, we investigate the reliability performance of downlink massive MIMO-NOMA systems.

1.8 Thesis Organization

The outline of the thesis is summarized in Fig. 1.8 and is as follows. We present comprehensive analysis of uplink massive MIMO-NOMA systems and address the corresponding challenges in Chapter 2. In Chapter 3, we extend our study of uplink massive MIMO-NOMA systems to incorporate the effects of channel outdatedness due to dynamic wireless environments. Herein, we analyze the robustness of decoding, channel estimation and user grouping

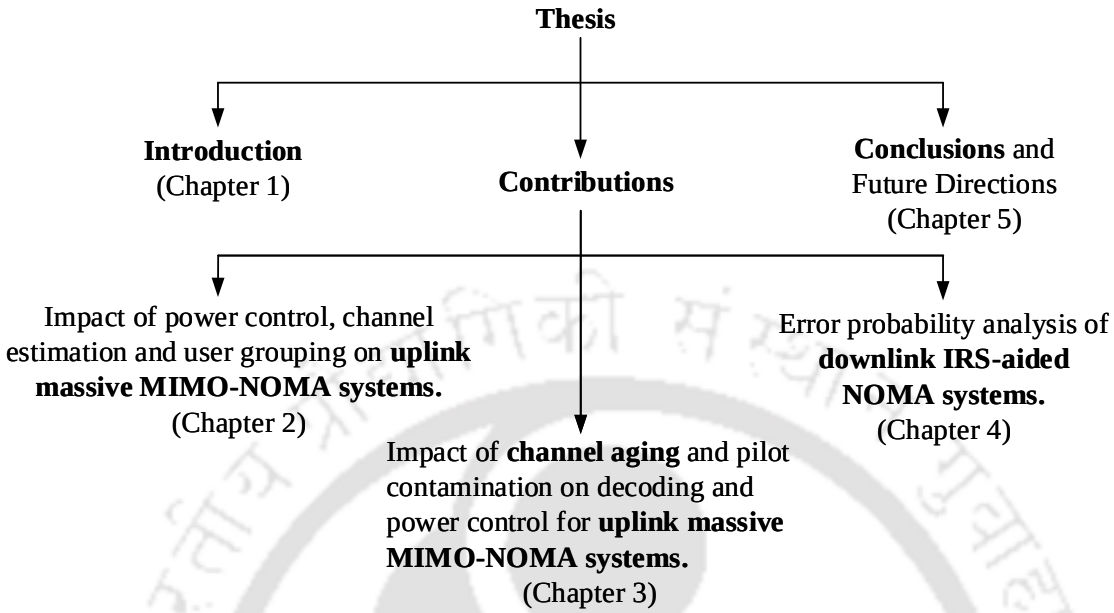


Fig. 1.8: Structure of the thesis.

strategies to channel aging. We then investigate IRS-aided NOMA systems in terms of reliability of data detection in Chapter 4. Finally, we discuss conclusions of the thesis and key directions for the future work in Chapter 5. The detailed proofs are given in the appendices.

1.9 List of Publications

- **Aditya Raosaheb Pawar**, Salil Kashyap and Sonali Chouhan, "Impact of Channel Aging and Pilot Contamination on Decoding and Power Control Strategies for Uplink Massive MIMO-NOMA Systems," in *IEEE Transactions on Communications*, Jul. 2024, doi: 10.1109/TCOMM.2024.3429160.
  - This work is discussed in detail in Chapter 3.
- **Aditya Raosaheb Pawar**, Salil Kashyap and Sonali Chouhan, "Error Probability Analysis of Intelligent Reflecting Surface-Enabled Non-Orthogonal Multiple Access Systems," in *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, pp. 9237-9251, Jul. 2023, doi: 10.1109/TVT.2023.3252083.
  - This work is discussed in detail in Chapter 4.
- **Aditya Raosaheb Pawar**, Salil Kashyap and Sonali Chouhan, "Impact of Max-Min Power Control, Channel Estimation and User Grouping Strategies on Uplink Massive

MIMO-NOMA Systems," in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7858-7869, Aug. 2021, doi: 10.1109/TVT.2021.3093716.

- This work is discussed in detail in Chapter 2.

2

Massive MIMO-NOMA Systems

Contents

2.1 System Model 33
2.2 Spectral Efficiency Analysis 39
2.3 Max-Min Power Control 47
2.4 User Grouping Strategies 50
2.5 Numerical and Simulation Results 53
2.6 Summary 59

In this chapter, we analyze uplink of a massive MIMO NOMA system. We first present uplink pilot signaling and develop expressions for MMSE-based channel estimates using two low-overhead estimation schemes, namely, Scheme-I and Scheme-S. While in Scheme-I, the channel vectors to all users are estimated individually, in Scheme-S, sum of channel vectors of users within a group is estimated. We then derive new bounds on achievable SE when the BS uses ZF decoder to decode signals from different users. The decoder is designed as a function of channel estimates acquired either via Scheme-I or via Scheme-S. Thereafter, we formulate an optimization problem to maximize the minimum achievable SE and obtain the corresponding max-min optimal power control coefficients to guarantee uniform quality of service to all users. We then present extensive numerical results to identify favorable operational regime for ZF decoder.

Chapter Organization: In Section 2.1, we discuss system model, channel estimation schemes and uplink data transmission. In Section 2.2, we present detailed SE analysis followed by max-min power control in Section 2.3 and discussion on user grouping strategies in Section 2.4. Extensive numerical results and the corresponding inferences are provided in Section 2.5. Finally, we state conclusions in Section 2.6.

2.1 System Model

We consider the uplink of a single-cell NOMA-based massive MIMO system where an M antenna BS serves L single-antenna users in the network.[1] To employ NOMA and perform SIC, we group users based on their large-scale fading coefficients.[2] Let N be the total number of groups and K be the number of users per group, then $L = N \times K$. We denote the k^{th} user in the n^{th} group by UE_{nk}, for $k \in \{1, 2, \cdots, K\}$ and $n \in \{1, 2, \cdots, N\}$. Let $\mathbf{g}_{nk} \in \mathbb{C}^{M\times 1}$ denote the channel vector from the BS to UE_{nk}. We consider independent Rayleigh fading. Thus, $\mathbf{g}_{nk} \sim \mathcal{CN}(\mathbf{0}, \beta_{nk}\mathbf{I}_M)$, where β_{nk} denotes the large-scale fading coefficient corresponding to UE_{nk}. Furthermore, the $M \times L$ composite channel matrix $\mathbf{G}$ from BS to all L users is given by $\mathbf{G} =$

[1]The notations and symbols used in each chapter are defined separately and uniquely in each chapter.

[2]User grouping based on large-scale fading is a very effective and practically feasible strategy because the large-scale fading coefficients vary very slowly and remain unchanged for several coherence intervals [15].

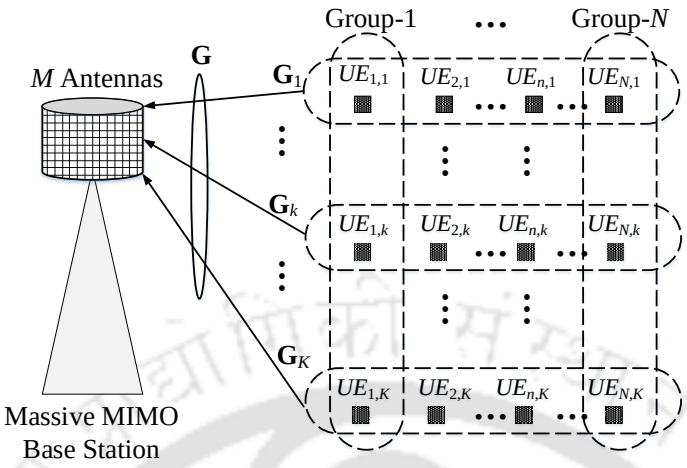


Fig. 2.1: System model: Uplink massive MIMO-NOMA.

$[\mathbf{G}_1\ \mathbf{G}_2\ \cdots\ \mathbf{G}_K]$, where $\mathbf{G}_i = [\mathbf{g}_{1i}\ \mathbf{g}_{2i} \cdots \mathbf{g}_{Ni}]$ denotes the $M \times N$ channel matrix to the i^{th} user in all the N groups, for $i \in \{1, 2, \cdots, K\}$, as shown in Fig. 2.1. The channel remains time-invariant over a coherence interval consisting of τ_c samples and varies independently across coherence intervals. Therefore, the channel needs to be estimated once every coherence interval and these estimates are then used to decode the signal received during uplink data transmission phase.

2.1.1 Channel Estimation

Next, we discuss the MMSE-based low-overhead channel estimation schemes. Prior to this discussion, we discuss the principles of MMSE estimation.

2.1.1.1 Preliminaries on MMSE Estimation

Let θ be a random variable representing the parameter or signal to be estimated, and let Y be the observed random variable. The mean square error (MSE) of an estimator $\hat{\theta}(Y)$ is defined as [72]:

$$\text{MSE} = \mathbb{E}\left[\left(\theta - \hat{\theta}(Y)\right)^2\right]. \tag{2.1}$$

The MMSE estimator $\hat{\theta}_{\text{MMSE}}(Y)$ is the one that minimizes this MSE and is given by

$$\hat{\theta}_{\text{MMSE}}(Y) = \arg\min_{\hat{\theta}} \mathbb{E}\left[\left(\theta - \hat{\theta}(Y)\right)^2\right]. \tag{2.2}$$

The error between the true parameter and the MMSE estimate is orthogonal to any function of the observation. This implies that

$$\mathbb{E}\left[\left(\theta - \hat{\theta}_{\text{MMSE}}(Y)\right) g(Y)\right] = 0, \tag{2.3}$$

for any function $g(Y)$. This property is useful for deriving MMSE estimators in practical scenarios.

In the case where the relationship between θ and Y is linear, i.e., $Y = \mathcal{H}\theta + \mathcal{W}$, where $\mathcal{H}$ is a known matrix and $\mathcal{W}$ is a zero-mean noise vector with known covariance matrix Σ_W, the MMSE estimator is given by:

$$\hat{\theta}_{\text{MMSE}} = \Sigma_{\theta Y}\Sigma_{YY}^{-1}Y, \tag{2.4}$$

where $\Sigma_{\theta Y} = \mathbb{E}[\theta Y^T]$ is the cross-covariance matrix between θ and Y, $\Sigma_{YY} = \mathbb{E}[YY^T]$ is the covariance matrix of Y.

Consider a massive MIMO system where the received pilot signal $\mathbf{Y} \in \mathbb{C}^{M\times\tau_p}$ at the BS during the pilot transmission phase is given by:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}, \tag{2.5}$$

where $\mathbf{H} \in \mathbb{C}^{M\times L}$ is the channel matrix, with $[H]_{ml}$ representing the channel gain between the m-th BS antenna and the l-th user, $\mathbf{X} \in \mathbb{C}^{L\times\tau_p}$ is the pilot matrix transmitted by the users, where $\tau_p \geq L$ is the length of the pilot sequence. Each row of $\mathbf{X}$ represents the pilot signal from one user, $\mathbf{W} \in \mathbb{C}^{M\times\tau_p}$ is the additive noise matrix at the BS, with each entry being an independent and identically distributed (i.i.d.) complex Gaussian noise with zero mean and variance σ_w^2. The goal is to estimate the channel matrix $\mathbf{H}$ from the received signal $\mathbf{Y}$. The MMSE estimator minimizes the mean square error between the true channel $\mathbf{H}$ and its estimate $\hat{\mathbf{H}}$.

Assume the elements of the channel matrix $\mathbf{H}$ are modelled as independent complex Gaussian random variables with zero mean and variance β_k, representing the large-scale fading:

$$\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}), \tag{2.6}$$

where $\mathbf{R} = \text{diag}(\beta_1, \beta_2, \ldots, \beta_L)$ is the $L \times L$ diagonal covariance matrix of the channel gains.

The linear MMSE (LMMSE) estimate of the channel matrix $\mathbf{H}$ given $\mathbf{Y}$ is given by:

$$\hat{\mathbf{H}} = \mathbf{R}_{HY}\mathbf{R}_{YY}^{-1}\mathbf{Y}, \tag{2.7}$$

where $\mathbf{R}_{HY} = \mathbb{E}[\mathbf{H}\mathbf{Y}^H]$ is the cross-covariance matrix between $\mathbf{H}$ and $\mathbf{Y}$. $\mathbf{R}_{YY} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^H]$ is the covariance matrix of $\mathbf{Y}$.

Having discussed the principles of MMSE, we consider a coherence interval of length τ_c symbols in which initial τ_p symbols are used for uplink pilot training for channel estimation at the BS and remaining $\tau_c - \tau_p$ symbols are used for data transmission. The channel is assumed to be constant over the duration of the coherence interval. We consider a scenario where every group is assigned an orthogonal pilot sequence while users within a group share the same pilot sequence for estimating the channel coefficients on the uplink. Let $\boldsymbol{\psi}_n \in \mathbb{C}^{\tau_p}$ denote the pilot sequence assigned to the n^{th} group. The pilot sequences are chosen such that $\boldsymbol{\psi}_i'\boldsymbol{\psi}_j = \delta_{ij}$, which equals 1 if $i = j$, and equals 0 otherwise. To ensure this, the length of the pilot sequence must be at least equal to the number of groups, i.e., $\tau_p \geq N$.

The signal $\mathbf{Y}_p \in \mathbb{C}^{M\times\tau_p}$ received at the BS during uplink pilot signaling is given by

$$\mathbf{Y}_p = \sqrt{\tau_p\rho_{ul}}\,[\mathbf{G}_1 + \mathbf{G}_2 + \cdots + \mathbf{G}_K]\,\boldsymbol{\Psi}' + \mathbf{W}_p\,, \tag{2.8}$$

where ρ_{ul} is the power with which any user transmits pilots on the uplink, $\boldsymbol{\Psi}$ is the $\tau_p \times N$ pilot matrix whose n^{th} column corresponds to the pilot sequence $\boldsymbol{\psi}_n$ assigned to the n^{th} group and $\mathbf{W}_p \in \mathbb{C}^{M\times\tau_p}$ is the additive noise matrix with $\mathcal{CN}(0,\sigma^2)$ entries which are independent and identically distributed (i.i.d.). We can write (2.8) in a simplified form as

$$\mathbf{Y}_p = \sqrt{\tau_p\rho_{ul}}\sum_{n=1}^{N}\sum_{i=1}^{K}\mathbf{g}_{ni}\boldsymbol{\psi}_n' + \mathbf{W}_p. \tag{2.9}$$

In order to estimate the channel of users in the n^{th} group, the BS correlates $\mathbf{Y}_p$ with $\boldsymbol{\psi}_n$ to obtain $\overline{\mathbf{y}}_{pn} \in \mathbb{C}^{M\times 1}$ as

$$\overline{\mathbf{y}}_{pn} \triangleq \mathbf{Y}_p\boldsymbol{\psi}_n = \sqrt{\tau_p\rho_{ul}}\sum_{i=1}^{K}\mathbf{g}_{ni} + \overline{\mathbf{w}}_{pn}\,, \tag{2.10}$$

where $\overline{\mathbf{w}}_{pn} = \mathbf{W}_p\boldsymbol{\psi}_n$ has identical distribution as $\mathbf{W}_p$ because $\overline{\mathbf{w}}_{pn}$ is obtained through a uni-

tary transformation of $\mathbf{W}_p$. Note that the components of $\mathbf{Y}_p$ in directions orthogonal to $\boldsymbol{\psi}_n$ are irrelevant for estimating the channels to users in the n^{th} group.

We next discuss two MMSE-based channel estimation schemes, namely, Scheme-I and Scheme-S. While in Scheme-I, we estimate the channel vectors to all users in the network individually, in Scheme-S, for each group of users, we estimate the sum of channel vectors to users within that group, based on $\overline{\mathbf{y}}_{pn}$.

2.1.1.2 MMSE Estimates with Scheme-I

Given the observable vector $\overline{\mathbf{y}}_{pn}$ in (2.10), the MMSE estimate $\hat{\mathbf{g}}_{nk}$ of the channel vector $\mathbf{g}_{nk}$ to UE_{nk} equals [72]

$$\hat{\mathbf{g}}_{nk} = \zeta_{nk}\,\overline{\mathbf{y}}_{pn} = \mathbf{g}_{nk} + \tilde{\mathbf{g}}^{\mathrm{I}}_{nk}\,, \tag{2.11}$$

where $\zeta_{nk} = \frac{\sqrt{\tau_p\rho_{ul}}\beta_{nk}}{\sigma^2+\tau_p\rho_{ul}\sum_{i=1}^{K}\beta_{ni}}$ and $\tilde{\mathbf{g}}^{\mathrm{I}}_{nk}$ is the corresponding channel estimation error which is given by

$$\tilde{\mathbf{g}}^{\mathrm{I}}_{nk} = \left(\sqrt{\tau_p\rho_{ul}}\zeta_{nk} - 1\right)\mathbf{g}_{nk} + \sqrt{\tau_p\rho_{ul}}\zeta_{nk}\sum_{j=1,j\neq k}^{K}\mathbf{g}_{nj} + \zeta_{nk}\overline{\mathbf{w}}_{pn}. \tag{2.12}$$

Note that $\hat{\mathbf{g}}_{nk} \sim \mathcal{CN}\left(\mathbf{0}, \eta_{nk}\mathbf{I}_M\right)$ and $\tilde{\mathbf{g}}^{\mathrm{I}}_{nk} \sim \mathcal{CN}\left(\mathbf{0}, \left(\beta_{nk}-\eta_{nk}\right)\mathbf{I}_M\right)$, where

$$\eta_{nk} = \frac{\tau_p\rho_{ul}\beta_{nk}^2}{\sigma^2 + \tau_p\rho_{ul}\sum_{i=1}^{K}\beta_{ni}}. \tag{2.13}$$

For MMSE estimation, $\tilde{\mathbf{g}}^{\mathrm{I}}_{nk}$ is uncorrelated to $\hat{\mathbf{g}}_{nk}$ [14]. We can write $\hat{\mathbf{g}}_{nk}$ in the normalized form as $\hat{\mathbf{g}}_{nk} = \sqrt{\eta_{nk}}\,\mathbf{z}_n$, where $\mathbf{z}_n \in \mathbb{C}^{M\times 1}$ is the normalized vector with i.i.d. $\mathcal{CN}(0,1)$ entries. From (2.11), we can see that the estimated channel vectors of all users within a group are perfectly correlated. Therefore, $\mathbf{z}_n$ is independent of the user index k. Under Scheme-I, let $\widehat{\mathbf{G}}^{\mathrm{I}} = \left[\widehat{\mathbf{G}}^{\mathrm{I}}_1\ \widehat{\mathbf{G}}^{\mathrm{I}}_2\cdots\widehat{\mathbf{G}}^{\mathrm{I}}_K\right]$ and $\widetilde{\mathbf{G}}^{\mathrm{I}} = \left[\widetilde{\mathbf{G}}^{\mathrm{I}}_1\ \widetilde{\mathbf{G}}^{\mathrm{I}}_2\cdots\widetilde{\mathbf{G}}^{\mathrm{I}}_K\right]$ respectively denote the composite estimated channel matrix and the composite channel estimation error matrix for all L users, where $\widehat{\mathbf{G}}^{\mathrm{I}}_i = \left[\hat{\mathbf{g}}_{1i}\ \ \hat{\mathbf{g}}_{2i}\ \ \cdots\ \ \hat{\mathbf{g}}_{Ni}\right]$ is the matrix of channel estimates of i^{th} user in all the N groups and $\widetilde{\mathbf{G}}^{\mathrm{I}}_i = \left[\tilde{\mathbf{g}}^{\mathrm{I}}_{1i}\ \ \tilde{\mathbf{g}}^{\mathrm{I}}_{2i}\ \ \cdots\ \ \tilde{\mathbf{g}}^{\mathrm{I}}_{Ni}\right]$ is the corresponding channel estimation error matrix, for $i \in \{1,2,\cdots,K\}$. Since, $\hat{\mathbf{g}}_{ni} = \sqrt{\eta_{ni}}\,\mathbf{z}_n$, for $n \in \{1,2,\cdots,N\}$,

$$\widehat{\mathbf{G}}^{\mathrm{I}}_i = \left[\sqrt{\eta_{1i}}\,\mathbf{z}_1\ \ \sqrt{\eta_{2i}}\,\mathbf{z}_2\ \cdots\ \sqrt{\eta_{Ni}}\,\mathbf{z}_N\right] = \mathbf{Z}\mathbf{D}^{\frac{1}{2}}_{\eta_i}, \tag{2.14}$$

where $\mathbf{Z} = [\mathbf{z}_1\ \mathbf{z}_2 \cdots \mathbf{z}_N]$ and $\mathbf{D}_{\eta_i} = \text{diag}\left[\eta_{1i}, \eta_{2i}, \cdots, \eta_{Ni}\right]$, for $i \in \{1, 2, \cdots, K\}$. Note that $\mathbf{z}_i$ and $\mathbf{z}_j$ are uncorrelated to each other for $i \neq j$.

2.1.1.3 MMSE Estimates with Scheme-S

Based on Scheme-S, we estimate $\sum_{i=1}^{K} \mathbf{g}_{ni}$, the sum of independent channel vectors of users within a group, based on $\overline{\mathbf{y}}_{pn}$. We can rewrite (2.10) as

$$\overline{\mathbf{y}}_{pn} = \sqrt{\tau_p \rho_{ul}}\mathbf{c}_n + \overline{\mathbf{w}}_{pn}, \tag{2.15}$$

where $\mathbf{c}_n = \sum_{i=1}^{K} \mathbf{g}_{ni}$. From (2.15), the MMSE estimate $\hat{\mathbf{c}}_n$ of $\mathbf{c}_n$ is given by

$$\hat{\mathbf{c}}_n = \mu_n \overline{\mathbf{y}}_{pn} = \mathbf{g}_{nk} + \tilde{\mathbf{g}}^{\text{S}}_{nk}, \tag{2.16}$$

where $\mu_n = \frac{\sqrt{\tau_p \rho_{ul}} \sum_{i=1}^{K} \beta_{ni}}{\sigma^2 + \tau_p \rho_{ul} \sum_{i=1}^{K} \beta_{ni}}$ and $\tilde{\mathbf{g}}^{\text{S}}_{nk}$ is the corresponding estimation error which is given by

$$\tilde{\mathbf{g}}^{\text{S}}_{nk} = \left(\sqrt{\tau_p \rho_{ul}}\mu_n - 1\right)\mathbf{g}_{nk} + \sqrt{\tau_p \rho_{ul}}\mu_n \sum_{j=1, j\neq k}^{K} \mathbf{g}_{nj} + \mu_n \overline{\mathbf{w}}_{pn}. \tag{2.17}$$

Note that $\hat{\mathbf{c}}_n \sim \mathcal{CN}(\mathbf{0}, \eta_n \mathbf{I}_M)$, $\tilde{\mathbf{g}}^{\text{S}}_{nk} \sim \mathcal{CN}(\mathbf{0}, \alpha_{nk}\mathbf{I}_M)$, where

$$\eta_n = \frac{\tau_p \rho_{ul} \left(\sum_{i=1}^{K} \beta_{ni}\right)^2}{\sigma^2 + \tau_p \rho_{ul} \sum_{i=1}^{K} \beta_{ni}} \tag{2.18}$$

and

$$\alpha_{nk} = \left(\sqrt{\tau_p \rho_{ul}}\mu_n - 1\right)^2 \beta_{nk} + \left(\sqrt{\tau_p \rho_{ul}}\mu_n\right)^2 \sum_{j=1, j\neq k}^{K} \beta_{nj} + \mu_n^2 \sigma^2. \tag{2.19}$$

Note that $\hat{\mathbf{c}}_n$ and $\tilde{\mathbf{g}}^{\text{S}}_{nk}$ are correlated. We can write $\hat{\mathbf{c}}_n$ in the normalized form as

$$\hat{\mathbf{c}}_n = \sqrt{\eta_n}\mathbf{v}_n, \tag{2.20}$$

where $\mathbf{v}_n$ is the normalized vector with i.i.d. $\mathcal{CN}(0, 1)$ entries. For Scheme-S, let $\widehat{\mathbf{G}}^{\text{S}} = \left[\widehat{\mathbf{G}}^{\text{S}}_1\ \widehat{\mathbf{G}}^{\text{S}}_2 \cdots \widehat{\mathbf{G}}^{\text{S}}_K\right]$ and $\widetilde{\mathbf{G}}^{\text{S}} = \left[\widetilde{\mathbf{G}}^{\text{S}}_1\ \widetilde{\mathbf{G}}^{\text{S}}_2 \cdots \widetilde{\mathbf{G}}^{\text{S}}_K\right]$ respectively denote the composite estimated channel matrix and the composite channel estimation error matrix for all L users, where $\widehat{\mathbf{G}}^{\text{S}}_i = \widehat{\mathbf{C}} = [\hat{\mathbf{c}}_1\ \hat{\mathbf{c}}_2 \cdots \hat{\mathbf{c}}_N]$ is the matrix of the estimates of channel to i^{th} user in all the N groups, for $i \in \{1, 2, \cdots, K\}$ and $\widetilde{\mathbf{G}}^{\text{S}}_i = \left[\tilde{\mathbf{g}}^{\text{S}}_{1i}\ \ \tilde{\mathbf{g}}^{\text{S}}_{2i}\ \ \cdots\ \ \tilde{\mathbf{g}}^{\text{S}}_{Ni}\right]$ is the corresponding estimation error. Note that the variance, η_n, of each entry of $\hat{\mathbf{c}}_n$ is independent of user index k and depends only on the

group index n. Therefore, using (2.20), $\widehat{\mathbf{G}}_i^{\mathrm{S}}$ can be equivalently written as

$$\widehat{\mathbf{G}}_i^{\mathrm{S}} = \widehat{\mathbf{C}} = \mathbf{V}\mathbf{D}_{\eta}^{\frac{1}{2}}, \tag{2.21}$$

where $\mathbf{V} = [\mathbf{v}_1\ \mathbf{v}_2 \cdots \mathbf{v}_N]$ and $\mathbf{D}_{\eta}^{\frac{1}{2}} = \mathrm{diag}\,[\eta_1, \eta_2, \cdots, \eta_N]$. The BS uses the estimated CSI obtained based on either Scheme-I or Scheme-S to decode the signal received during the uplink data transmission phase.

2.1.2 Uplink Data Transmission

Let $q_{nk} \sim \mathcal{CN}(0,1)$ and $\gamma_{nk} \in (0, \rho_{ul}]$, respectively denote the transmitted data symbol and the uplink transmit power for UE_{nk} for $k \in \{1, 2, \cdots, K\}$ and $n \in \{1, 2, \cdots, N\}$. We assume that all the symbols are independent of each other. The signal $\mathbf{y} \in \mathbb{C}^{M\times 1}$ received at the BS during uplink data signaling is given by

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w}, \tag{2.22}$$

where $\mathbf{w} \in \mathbb{C}^{M\times 1}$ is a noise vector with i.i.d. $\mathcal{CN}(0, \sigma^2)$ entries and the collective symbol vector $\mathbf{x} \in \mathbb{C}^{L\times 1}$ transmitted by L users equals

$$\mathbf{x} = \left[\mathbf{q}_1^T\mathbf{D}_{\gamma_1}^{\frac{1}{2}} \quad \mathbf{q}_2^T\mathbf{D}_{\gamma_2}^{\frac{1}{2}} \quad \cdots \quad \mathbf{q}_K^T\mathbf{D}_{\gamma_K}^{\frac{1}{2}}\right]^T, \tag{2.23}$$

where $\mathbf{q}_i = \left[q_{1i}\ q_{2i}\ \cdots\ q_{Ni}\right]^T$ is the symbol vector transmitted by the i^{th} user in all N groups and $\mathbf{D}_{\gamma_i} = \mathrm{diag}\left[\gamma_{1i}, \gamma_{2i}, \cdots, \gamma_{Ni}\right]$, for $i \in \{1, 2, \cdots, K\}$.

2.2 Spectral Efficiency Analysis

In an $M \times L$ massive MIMO system, the conventional ZF decoder takes up $L = N \times K$ degrees-of-freedom (DoF) to completely cancel the inter-user interference at the BS. This puts a constraint that maximum $M - 1$ users can be simultaneously served with a non-zero achievable SE. Furthermore, the conventional ZF decoder involves computation of the inverse of the Grammian matrix $\left(\widehat{\mathbf{G}}^u\right)'\left(\widehat{\mathbf{G}}^u\right)$, for $u \in \{\mathrm{I}, \mathrm{S}\}$. Therefore, the conventional ZF decoder cannot be designed based on correlated channel estimates acquired via shared pilot signaling. This mandates acquisition of CSI based on orthogonal pilots, which in turn scales the channel estimation

overhead proportional to L.

Another decoding strategy used in the literature for NOMA-based massive MIMO systems is MR decoding [11, 73]. The advantage of MR decoding is that its BF gain scales proportional to the number of antennas (M) at the BS and is independent of the number of users (L) in the network. While intra-group interference can be cancelled using SIC at the BS, the MR decoder fails to cancel IGI. Thus, every user experiences IGI and full/partial intra-group interference based on the order in which SIC is done. This heavy interference badly impacts users with weaker channel strengths.

In order to mitigate IGI with reduced channel estimation overhead, we first analyze the achievable SE on the uplink of a NOMA-based massive MIMO system using ZF decoder designed based on Scheme-I in Section 2.2.1 followed by that based on Scheme-S in Section 2.2.2.[3] Subsequently, to ensure user fairness, we formulate the max-min optimization problem and obtain the corresponding power control coefficients to ensure that the minimum achievable SE is maximized. After that, we investigate two different grouping strategies: the non-scalable NF grouping and the scalable NB grouping, and analyze their impact on the uplink achievable SE.

2.2.1 Achievable SE with ZF based on Scheme-I

The ZF decoder $\mathbf{A}' \in \mathbb{C}^{N \times M}$ designed based on Scheme-I takes the following form:

$$\mathbf{A}' = \left[\mathbf{Z}'\mathbf{Z}\right]^{-1} \mathbf{Z}', \tag{2.24}$$

where the n^{th} column of $\mathbf{Z} \in \mathbb{C}^{M \times N}$ with $N \leq L$ corresponds to the normalized channel estimate for all users in the n^{th} group. Note that the channel estimates for all users within a group are perfectly correlated and differ only in the scaling factor captured by different large-scale fading coefficients associated with different users within a group. However, the columns of $\mathbf{Z}$ are completely uncorrelated since they correspond to distinct groups.

[3] We would like to note that MMSE decoding can also be employed at the BS. The MMSE decoders can be designed based on the CSI estimates acquired with Scheme-I and Scheme-S. However, it is important to keep in mind that it is not possible to obtain closed-form SE expressions for MMSE decoders. Only numerical simulations can be done and SE can be presented in terms of expected values.

The decoder in (2.24) is used at the BS to process the received signal $\mathbf{y}$ given in (2.22). Expressing the true channel matrix in terms of its estimate and the estimation error and using (2.14), we can rewrite (2.22) as

$$\begin{aligned}\mathbf{y}^{\mathrm{I}} &= \widehat{\mathbf{G}}^{\mathrm{I}}\mathbf{x} + \mathbf{w} - \widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{x} \\ &= \left[\mathbf{Z}\mathbf{D}_{\eta_1}^{\frac{1}{2}} \cdots \mathbf{Z}\mathbf{D}_{\eta_K}^{\frac{1}{2}}\right]\left[\mathbf{q}_1^T\mathbf{D}_{\gamma_1}^{\frac{1}{2}} \cdots \mathbf{q}_K^T\mathbf{D}_{\gamma_K}^{\frac{1}{2}}\right]^T + \overline{\mathbf{n}}^{\mathrm{I}} \\ &= \sum_{i=1}^{K} \mathbf{Z}\mathbf{D}_{\eta_i}^{\frac{1}{2}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{q}_i + \overline{\mathbf{n}}^{\mathrm{I}}, \end{aligned} \tag{2.25}$$

where $\overline{\mathbf{n}}^{\mathrm{I}} = \mathbf{w} - \widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{x}$. The BS pre-multiplies $\mathbf{y}^{\mathrm{I}}$ by the decoder $\mathbf{A}'$ to give

$$\overline{\mathbf{y}}^{\mathrm{I}} = \mathbf{A}'\mathbf{y}^{\mathrm{I}} = \mathbf{A}' \sum_{i=1}^{K} \mathbf{Z}\mathbf{D}_{\eta_i}^{\frac{1}{2}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{q}_i + \mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}. \tag{2.26}$$

From (2.24), we know that $\mathbf{A}'\mathbf{Z} = \mathbf{I}_N$. Therefore,

$$\overline{\mathbf{y}}^{\mathrm{I}} = \sum_{i=1}^{K} \mathbf{I}_N\mathbf{D}_{\eta_i}^{\frac{1}{2}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{q}_i + \mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}. \tag{2.27}$$

Note that $\overline{\mathbf{y}}^{\mathrm{I}} \in \mathbb{C}^{N\times 1}$. Thus, for $n \in \{1, 2, \cdots, N\}$, the n^{th} entry of $\overline{\mathbf{y}}^{\mathrm{I}}$ corresponds to the signal received from users belonging to the n^{th} group and is given by

$$\overline{y}_n^{\mathrm{I}} = \sum_{i=1}^{K} \sqrt{\eta_{ni}\gamma_{ni}}q_{ni} + \left[\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right]_n. \tag{2.28}$$

Using $\overline{y}_n^{\mathrm{I}}$, the BS decodes the signal of UE_{nk} for $k \in \{1, 2, \cdots, K\}$. Before decoding the signal of UE_{nk}, the BS performs SIC. The only condition to ensure successful SIC for UE_{nk} ($k \in \{1, 2, \cdots, K\}$) is that the BS should be able to decode the message of UE_{ni} with a rate at least equal to the target rate of UE_{ni}, $\forall\, i \in \{1, 2, \cdots, k-1\}$ [74, 75].

Without loss of generality, we assume that, in the n^{th} group

$$\beta_{n1} \geq \beta_{n2} \geq \cdots \geq \beta_{nK}, \tag{2.29}$$

for $n \in \{1, 2, \cdots, N\}$. With this ordering, the BS decodes the signal of the 1^{st} user, i.e., the strongest user in the group by treating every other user's signal as noise. Then before decoding the 2^{nd} user's signal, BS cancels the interference from the 1^{st} user in the group and treats

every other user's signal as noise. In this way, the BS successively cancels the interference and decodes the users' signals in each group. Therefore, post-SIC, the processed signal for UE_{nk} equals

$$\bar{y}_{nk}^{\text{I-SIC}} = \sqrt{\eta_{nk}\gamma_{nk}}q_{nk} + \sum_{i=k+1}^{K}\sqrt{\eta_{ni}\gamma_{ni}}q_{ni} + \left[\mathbf{A}'\bar{\mathbf{n}}^{\text{I}}\right]_n . \tag{2.30}$$

The interference from UE_{ni} is successively cancelled prior to decoding the signal of UE_{nk}, for $i \in \{1, 2, \cdots, k-1\}$, $k \in \{1, 2, \cdots, K\}$ and $n \in \{1, 2, \cdots, N\}$. Therefore, BS uses (2.30) to decode the signal of UE_{nk}.

Justification for the SIC decoding order: To understand the role of SIC decoding order, let us consider a simple network with two users, user A and user B. Let P_{ul}^{t} be the uplink transmit power available at user $t \in \{A, B\}$. Each user transmits its message on the uplink using the transmit power available at its end. Therefore, the signals transmitted on the uplink by users A and B are given by

$$x_A = \sqrt{P_{ul}^{A}}s_A^{ul}, \tag{2.31}$$

$$x_B = \sqrt{P_{ul}^{B}}s_B^{ul}, \tag{2.32}$$

respectively, where s_t^{ul} is the uplink message transmitted by user $t \in \{A, B\}$ such that $\mathbb{E}\left|s_t^{ul}\right|^2 = 1$. The signal received at the BS is given by

$$y_{BS}^{ul} = \sqrt{P_{ul}^{A}}h_A s_A^{ul} + \sqrt{P_{ul}^{B}}h_B s_B^{ul} + z_{BS}, \tag{2.33}$$

where $z_{BS} \sim \mathcal{CN}(0, \sigma^2)$ is the additive Gaussian noise at BS.

On the uplink, the onus to do SIC is on the BS, and BS has the liberty to decide the SIC decoding order depending on the desired rate pair within the capacity region [7], [13]. For example, if the BS first decodes the message of the weaker user (user B) and cancels the interference that it creates to the stronger user (user A) before decoding the message of the stronger user, then the performance of the stronger user improves [7], [13] while the performance of the weaker user deteriorates further. With this decoding order, point Q in Figure 2.2 is achieved. Therefore, with the decoding order to achieve point Q in Figure 2.2, the instantaneous SINRs

for users A and B are given by

$$\text{SINR}_A^{ul} = \frac{P_{ul}^A |h_A|^2}{\sigma^2}, \tag{2.34}$$

$$\text{SINR}_B^{ul} = \frac{P_{ul}^B |h_B|^2}{\sigma^2 + P_{ul}^A |h_A|^2}. \tag{2.35}$$

On the other hand, suppose the received power of the weaker user's signal is very less compared to the stronger user's received signal power. In that case, the BS first decodes the message of the stronger user (user A) and cancels the interference that it creates to the weaker user (user B) before decoding the message of the weaker user. This decoding order facilitates fairness and maximizes the weaker user's achievable rate [7], [13]. With this decoding order, point P in Figure 2.2 is achieved. Therefore, if the objective is to improve fairness, the instantaneous SINRs for users A and B are given by

$$\text{SINR}_A^{ul} = \frac{P_{ul}^A |h_A|^2}{\sigma^2 + P_{ul}^B |h_B|^2}, \tag{2.36}$$

$$\text{SINR}_B^{ul} = \frac{P_{ul}^B |h_B|^2}{\sigma^2}. \tag{2.37}$$

These two strategies lead to achieving very different points in the achievable capacity region as it can be seen in Figure 2.2. As one of our primary objectives is to ensure performance fairness in terms of SE, we employ the discussed SIC decoding order where the BS decodes the stronger users' signals and cancels the interference prior to decoding weaker users' signals.

Theorem 2.1. *Given $\bar{y}_{nk}^{\text{I-SIC}}$, a lower bound on the achievable SE of UE_{nk} in NOMA-based massive MIMO system that uses ZF decoder designed based on channel estimation Scheme-I is given by*

$$R_{nk}^I = \delta \log_2 \left(1 + SINR_{nk}^I\right), \tag{2.38}$$

where $\delta = (1 - (\tau_p/\tau_c))$ accounts for the channel estimation overhead and $SINR_{nk}^I$ denotes the SINR for UE_{nk} and is given by

$$SINR_{nk}^I = \frac{(M-N)\,\eta_{nk}\,\gamma_{nk}}{\sigma^2 + \lambda^I + \sum_{k'=k+1}^{K} (M-N)\,\eta_{nk'}\gamma_{nk'}}, \tag{2.39}$$

where $\lambda^I = \sum_{m=1}^{N} \sum_{i=1}^{K} \{(\beta_{mi} - \eta_{mi})\,\gamma_{mi}\}$.

Proof. The proof is given in Appendix A.1. □

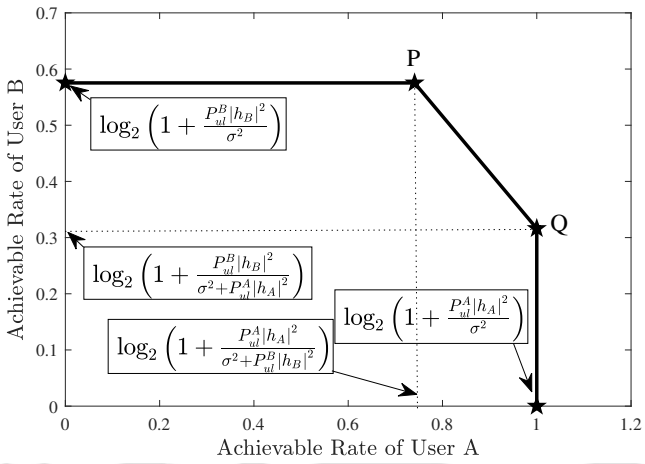


Fig. 2.2: Capacity region of uplink NOMA: $h_A = 1, h_B = 0.7, P_{ul}^A = 0.1, P_{ul}^B = 0.1, \sigma^2 = 0.1$.

Remark: The term $(M-N)$ in the numerator of (2.39) accounts for the DoF exhausted in IGI cancellation. The term λ^{I} in the denominator of (2.39) accounts for the variance of the channel estimation error with Scheme-I, and η_{mi} therein captures variance of the estimate of channel vector corresponding to UE_{mi} under pilot contamination. The third term in the denominator of (2.39) is essentially the variance of the residual intra-group interference. Therefore, the SINR expression in (2.39) captures the effects of (i) the DoF exhausted in IGI cancellation, (ii) the intra-group pilot contamination under Scheme-I, (iii) the intra-group interference and (iv) channel estimation errors under Scheme-I.

As we discussed the SE analysis of uplink massive MIMO-NOMA systems under ZF decoder designed using the estimates acquired with Scheme-I, we next discuss the SE analysis under ZF decoder designed using the estimates acquired with Scheme-S.

2.2.2 Achievable SE with ZF based on Scheme-S

The ZF decoder $\mathbf{B}' \in \mathbb{C}^{N\times M}$ designed based on Scheme-S takes the following form:

$$\mathbf{B}' = \left[\mathbf{V}'\mathbf{V}\right]^{-1}\mathbf{V}'. \tag{2.40}$$

Note that the n^{th} column of $\mathbf{V} \in \mathbb{C}^{M\times N}$ with $N \leq L$ corresponds to the normalized channel estimate for all users in the n^{th} group. Also, the scaling factor captured by large-scale fading is the same for all users within a group, unlike Scheme-I, and differs only across groups. The

columns of $\mathbf{V}$ are uncorrelated since they correspond to distinct groups.

The decoder in (2.40) is used at the BS to process the received signal $\mathbf{y}$ given in (2.22). Expressing the true channel matrix in terms of its estimate and the estimation error and using (2.21), we can rewrite (2.22) as

$$\begin{aligned}\mathbf{y}^{\mathrm{S}} &= \widehat{\mathbf{G}}^{\mathrm{S}}\mathbf{x} + \mathbf{w} - \widetilde{\mathbf{G}}^{\mathrm{S}}\mathbf{x} \\ &= \left[\mathbf{V}\mathbf{D}_{\eta}^{\frac{1}{2}} \cdots \mathbf{V}\mathbf{D}_{\eta}^{\frac{1}{2}}\right]\left[\mathbf{q}_1^T\mathbf{D}_{\gamma_1}^{\frac{1}{2}} \cdots \mathbf{q}_K^T\mathbf{D}_{\gamma_K}^{\frac{1}{2}}\right]^T + \overline{\mathbf{n}}^{\mathrm{S}} \\ &= \sum_{i=1}^{K} \mathbf{V}\mathbf{D}_{\eta}^{\frac{1}{2}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{q}_i + \overline{\mathbf{n}}^{\mathrm{S}}, \end{aligned} \tag{2.41}$$

where $\overline{\mathbf{n}}^{\mathrm{S}} = \mathbf{w} - \widetilde{\mathbf{G}}^{\mathrm{S}}\mathbf{x}$. The BS pre-multiplies $\mathbf{y}^{\mathrm{S}}$ by the decoder $\mathbf{B}'$ to give

$$\overline{\mathbf{y}}^{\mathrm{S}} = \mathbf{B}'\mathbf{y}^{\mathrm{S}} = \mathbf{B}'\sum_{i=1}^{K} \mathbf{V}\mathbf{D}_{\eta}^{\frac{1}{2}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{q}_i + \mathbf{B}'\overline{\mathbf{n}}^{\mathrm{S}}. \tag{2.42}$$

From (2.40), we know that $\mathbf{B}'\mathbf{V} = \mathbf{I}_N$. Therefore,

$$\overline{\mathbf{y}}^{\mathrm{S}} = \sum_{i=1}^{K} \mathbf{I}_N\mathbf{D}_{\eta}^{\frac{1}{2}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{q}_i + \mathbf{B}'\overline{\mathbf{n}}^{\mathrm{S}}. \tag{2.43}$$

Note that, $\overline{\mathbf{y}}^{\mathrm{S}} \in \mathbb{C}^{N\times 1}$. Thus, for $n \in \{1, 2, \cdots, N\}$,

$$\overline{y}_n^{\mathrm{S}} = \sum_{i=1}^{K} \sqrt{\eta_n\gamma_{ni}}q_{ni} + \left[\mathbf{B}'\overline{\mathbf{n}}^{\mathrm{S}}\right]_n. \tag{2.44}$$

Note that $\overline{y}_n^{\mathrm{S}}$ denotes the signal received from users belonging to the n^{th} group. Using $\overline{y}_n^{\mathrm{S}}$, the BS decodes the signal of UE_{nk} for $k \in \{1, 2, \cdots, K\}$. Before decoding the signal of UE_{nk}, the BS does SIC based on the decoding order explained in Section 2.2.1. Therefore, post SIC, the processed signal for UE_{nk} equals

$$\overline{y}_{nk}^{\text{S-SIC}} = \sqrt{\eta_n\gamma_{nk}}q_{nk} + \sum_{i=k+1}^{K} \sqrt{\eta_n\gamma_{ni}}q_{ni} + \left[\mathbf{B}'\overline{\mathbf{n}}^{\mathrm{S}}\right]_n. \tag{2.45}$$

Theorem 2.2. *Given the side information $\Omega = \mathbf{V}$ and $\overline{y}_{nk}^{S\text{-}SIC}$, a lower bound on the achievable SE of UE_{nk} in NOMA-based massive MIMO system that uses ZF decoder designed based on channel estimation Scheme-S is given by*

$$R_{nk}^{S} = \delta \log_2\left(1 + SINR_{nk}^{S}\right), \tag{2.46}$$

where $SINR_{nk}^{S}$ *denotes SINR for* UE_{nk} *and is given by*

$$SINR_{nk}^{S} = \frac{(M-N)\,\eta_n\,\gamma_{nk}}{\sigma^2 + \lambda^S + \sum_{k'=k+1}^{K}(M-N)\,\eta_n\gamma_{nk'}}, \tag{2.47}$$

where

$$\lambda^S = \sum_{i=1}^{K}\left[\left(\sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\right) - \frac{1}{N}\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right], \tag{2.48}$$

$$\epsilon_{ji} = \left(\sqrt{\tau_p\rho_{ul}}\mu_j - 1\right)\left(\sqrt{\tau_p\rho_{ul}}\mu_j\right)\beta_{ji} + \left(\sqrt{\tau_p\rho_{ul}}\mu_j\right)^2\sum_{l=1,l\neq i}^{K}\beta_{jl} + \mu_j^2\sigma^2, \tag{2.49}$$

with α_{ji} *as given in (2.19) and* η_j *as given in (2.18).*

Proof. The proof is given in Appendix A.2. □

Remark: The term $(M-N)$ in the numerator of (2.47) accounts for the DoF exhausted in IGI cancellation. The term λ^S in the denominator of (2.47) accounts for the variance of the channel estimation error with Scheme-S, and η_j therein captures variance of the estimate of sum of channel vectors corresponding to all users UE_{ji} in group i under pilot contamination. The third term in the denominator of (2.47) denotes the variance of the residual intra-group interference. Therefore, the SINR expression in (2.47) captures the effects of (i) the DoF exhausted in IGI cancellation, (ii) the intra-group pilot contamination under Scheme-S, (iii) the intra-group interference and (iv) channel estimation errors under Scheme-S.

In the next section, we formulate a max-min power allocation problem and solve using convex programming. Owing to the requirement of convex programming, we obtain an upper bound on λ^S so that the optimization problem gets converted to a set of linear feasibility problems. Recall that $\lambda^S = \sum_{i=1}^{K}\left[\left(\sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\right) - \frac{1}{N}\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right]$. Using the fact that for positive numbers, the square of their sum is lower bounded by the sum of their squares, we get

$$\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2 \geq \sum_{j=1}^{N}\left(\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2 = \sum_{j=1}^{N}\gamma_{ji}\left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2. \tag{2.50}$$

Therefore,

$$\lambda^S = \sum_{i=1}^{K}\left[\left(\sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\right) - \frac{1}{N}\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right] \tag{2.51}$$

$$\leq \sum_{i=1}^{K}\left[\left(\sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\right)-\frac{1}{N}\sum_{j=1}^{N}\gamma_{ji}\left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right] \tag{2.52}$$

$$= \sum_{i=1}^{K}\sum_{j=1}^{N}\gamma_{ji}\left(\alpha_{ji}-\frac{1}{N}\left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right) \tag{2.53}$$

$$= \lambda^{\text{S-UB}}, \tag{2.54}$$

where $\lambda^{\text{S-UB}}$ denotes an upper bound on λ^{S}. Using this upper bound $\lambda^{\text{S-UB}}$, we get a lower bound on $\text{SINR}_{nk}^{\text{S}}$ and then we use this lower bound to formulate a max-min power allocation problem which can be solved using convex programming.

Theorem 2.3. *Using the fact that $\lambda^{S\text{-}UB}$ is an upper bound on λ^{S}, a lower bound on the $SINR_{nk}^{S}$ is given by*

$$SINR_{nk}^{S\text{-}LB} = \frac{(M-N)\,\eta_n\,\gamma_{nk}}{\sigma^2+\lambda^{S\text{-}UB}+\sum_{k'=k+1}^{K}(M-N)\,\eta_n\gamma_{nk'}}, \tag{2.55}$$

where $\lambda^{S\text{-}UB} = \sum_{i=1}^{K}\sum_{j=1}^{N}\gamma_{ji}\left(\alpha_{ji}-\frac{1}{N}\left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right)$ with ϵ_{ji}, α_{ji} and η_j is given in (2.49), (2.19) and (2.18), respectively.

Remark: Notice that the coherent BF gain in the numerator of (2.39) and (2.55) is proportional to $(M-N)$, where $N \leq \tau_p < \tau_c$. Consequently, with ZF decoders $\mathbf{A}'$ and $\mathbf{B}'$ designed using the channel estimates acquired based on Scheme-I and Scheme-S, we can increase the total number of users served by increasing K and keeping $N\,(<\min\{M,\tau_c\})$ fixed, while maintaining a non-zero but reasonable achievable SE.

2.3 Max-Min Power Control

In this section, we investigate max-min power control to ensure that all users achieve a common optimal rate. The max-min power control implies that the power allocation ensures maximization of the minimum of the SINRs among all users' SINRs. Specifically, based on the result presented in Theorem 2.1, the max-min optimal power control coefficients $\{\gamma_{nk}\}$ would ensure that $\text{SINR}_{nk}^{I} = \overline{\text{SINR}}^{\text{I}}$, for all n and k, where $\overline{\text{SINR}}^{\text{I}}$ denotes a constant value. Similarly, based on the result presented in Theorem 2.3, the max-min optimal power control coefficients $\{\gamma_{nk}\}$ would ensure that $\text{SINR}_{nk}^{\text{S-LB}} = \overline{\text{SINR}}^{\text{S-LB}}$, for all n and k, where $\overline{\text{SINR}}^{\text{S-LB}}$ denotes a constant value.

In other words, the max-min power control entails setting all SINR values equal to a common target SINR value $\overline{\text{SINR}}^p$ and then finding the largest possible value of $\overline{\text{SINR}}^p$ while satisfying the constraints $0 \leq \gamma_{nk} \leq \rho_{ul}$ and $\text{SINR}_{nk} \geq \overline{\text{SINR}}^p,\ \forall n,k$ where $p \in \{\text{I}, \text{S-LB}\}$. While the former constraint is the transmit power constraint at every user, the latter constraint ensures that every user attains an SINR greater than or equal to $\overline{\text{SINR}}^p$. Note that ρ_{ul} is the maximum transmit power available at each user for uplink data transmission.

Mathematically, the max-min optimization problem can be stated in the epigraph form[4] as [11, 14, 76],

$$\mathcal{P}_1: \quad \text{maximize} \quad \overline{\text{SINR}}^p, \tag{2.56}$$

$$\text{with respect to} \quad \gamma_{11}, \cdots, \gamma_{nk}, \cdots, \gamma_{NK},$$

$$\text{subject to} \quad \text{SINR}_{nk}^p \geq \overline{\text{SINR}}^p,\ \forall n,k, \tag{2.57}$$

$$0 \leq \gamma_{nk} \leq \rho_{ul},\ \forall n,k, \tag{2.58}$$

where the power allocation coefficients $\{\gamma_{nk}\}$ are the optimization variables, the constraints in (2.57) and (2.58) are the common target SINR constraint and the power constraint at each user, respectively. The novel SINR expressions derived for the k^{th} user in the n^{th} group using ZF decoder based on Scheme-I or Scheme-S can be expressed in a generic form as

$$\text{SINR}_{nk}^p = \frac{(M-N)\,\eta_{nk}^p\,\gamma_{nk}}{\sigma^2 + \lambda^r + \sum_{k'=k+1}^{K}(M-N)\,\eta_{nk'}^p\,\gamma_{nk'}}, \tag{2.59}$$

where $p \in \{\text{I}, \text{S-LB}\}$, $r \in \{\text{I}, \text{S-UB}\}$, $\eta_{nk}^{\text{I}} = \eta_{nk}$, $\eta_{nk}^{\text{S-LB}} = \eta_n$ with $\lambda^{\text{I}} = \sum_{m=1}^{N}\sum_{i=1}^{K}\gamma_{mi}\,a_{mi}$ and $\lambda^{\text{S-UB}} = \sum_{m=1}^{N}\sum_{i=1}^{K}\gamma_{mi}\,b_{mi}$ where $a_{mi} = (\beta_{mi} - \eta_{mi})$ and $b_{mi} = \left(\alpha_{mi} - \frac{1}{N}\left(\frac{\epsilon_{mi}}{\sqrt{\eta_m}}\right)^2\right)$. Note that depending on the SIC decoding order at the BS, the number of terms in the intra-group interference $\sum_{k'=k+1}^{K}(M-N)\,\eta_{nk'}^p\gamma_{nk'}$, in the denominator of (2.59) may turn out to be different for every user. Therefore, for NOMA systems, the max-min optimization problem, as stated above, is analytically intractable. However, this optimization problem can be solved numerically using an optimization tool referred to as CVX [11, 14, 15, 77]. Using (2.59), we can rewrite the constraint

[4] In the epigraph form of an optimization problem, an auxiliary variable $\overline{\text{SINR}}^p$ is introduced which must satisfy the constraint that $\text{SINR}_{nk}^p \geq \overline{\text{SINR}}^p,\ \forall n,k$, and the objective is to maximize this auxiliary variable itself [76].

in (2.57) as

$$C_{nk} : \overline{\text{SINR}}^{p} \left[\sigma^2 + \lambda^r + \sum_{k'=k+1}^{K} (M-N)\, \eta_{nk'}^{p} \gamma_{nk'} \right] - (M-N)\, \eta_{nk}^{p} \gamma_{nk} \leq 0. \tag{2.60}$$

Therefore, the max-min optimization problem $\mathcal{P}_1$ can be rewritten as

$$\mathcal{P}_2 : \quad \text{maximize } \overline{\text{SINR}}^{p}, \tag{2.61}$$

$$\text{with respect to } \gamma_{11}, \cdots, \gamma_{nk}, \cdots, \gamma_{NK},$$

$$\text{subject to } C_{nk}, \ \forall n, k \tag{2.62}$$

$$0 \leq \gamma_{nk} \leq \rho_{ul}, \ \forall n, k, \tag{2.63}$$

where C_{nk} is given in (2.60). In this optimization problem, we observe that the constraints are linear with respect to the optimization variables $\{\gamma_{nk}\}$, i.e., the power allocation coefficients. Hence, we can solve $\mathcal{P}_2$ as a linear programming feasibility problem for a fixed $\overline{\text{SINR}}^{p}$. While solving this problem for a given $\overline{\text{SINR}}^{p}$, if the common target SINR constraint C_{nk} is satisfied $\forall n, k$, then the transmit power corresponding to each user must be increased as $\overline{\text{SINR}}^{p}$ increases to ensure that all the constraints are satisfied. And essentially, the problem reduces to performing a line search over $\overline{\text{SINR}}^{p}$ to find its maximum value for which the constraints are satisfied while solving the linear feasibility problem.

The CVX solves any linear programming feasibility problem using the bisection search method [14, 15, 76]. Based on this method, a candidate value of $\overline{\text{SINR}}^{p}$ is computed through bisection over an initially chosen search space between $\overline{\text{SINR}}^{p}_{lower} = 0$ and $\overline{\text{SINR}}^{p}_{upper} = \frac{(M-N)\, \eta_{nk}^{p} \gamma_{nk}}{\sigma^2}$, which is the maximum value that can be assumed by the objective function in the absence of any interference and channel estimation errors. Specifically, $\overline{\text{SINR}}^{p}_{candidate} = \frac{\overline{\text{SINR}}^{p}_{lower} + \overline{\text{SINR}}^{p}_{upper}}{2}$. For this computed $\overline{\text{SINR}}^{p}_{candidate}$, a linear feasibility problem is then solved to obtain the power allocation coefficients $\gamma_{n,k}\ \forall\, n, k$. If the solution obtained is feasible, then the search space is updated with $\overline{\text{SINR}}^{p}_{lower} = \overline{\text{SINR}}^{p}_{candidate}$, else the search space is updated with $\overline{\text{SINR}}^{p}_{upper} = \overline{\text{SINR}}^{p}_{candidate}$. Thus, the search space of the max-min SINR value is halved in each iteration. And this reduction in the search space is continued until the search space, i.e., $\left[\overline{\text{SINR}}^{p}_{lower}, \overline{\text{SINR}}^{p}_{upper}\right]$ becomes

negligibly small. This is when the algorithm terminates and gives the optimal feasible set of optimization variables $\left\{\gamma_{n,k}^{\star}\right\}$, and the corresponding optimal value of the linear objective function, $\overline{\text{SINR}}^{p\star}$. Note that the set of linear programming feasibility problems can be classified as a convex program and the solution thus obtained based on bisection search is guaranteed to be the globally optimal solution [76].

2.4 User Grouping Strategies

Dividing users in the network into several groups with each group consisting of a smaller number of users is essential to reducing the complexity associated with SIC, reducing the negative impact of error propagation due to SIC, and controlling the impact of intra-group interference [7]. As we know from the NOMA literature, the users' performance varies significantly with user grouping and power allocation strategy adopted [78]. Thus, while ensuring uniform quality-of-service, we investigate two different user grouping strategies, namely, NF grouping and NB grouping[5]. The user with the largest large-scale fading coefficient value is called the strongest user, and the user with the smallest large-scale fading coefficient value is called the weakest user. In NF grouping, the strongest user is grouped with the weakest user; the second strongest user is grouped with the second weakest user, and so on. In NB grouping, the strongest user is grouped with the second strongest user; the weakest user is grouped with the second weakest user, and so on. We use an example to elaborate the user grouping strategies and illustrate their impact on the strength of the channel estimates.[6]

Example: Consider a 4-user system as shown in Fig. 2.3 with β_i denoting the large-scale fading coefficient value of user i for $i \in \{1, 2, 3, 4\}$. Let $\beta_1 \geq \beta_2 \geq \beta_3 \geq \beta_4$. Based on NF

[5]Although on the downlink, the difference in the channel gains of users within a group is crucial to ensure successful SIC, on the uplink, the only condition for successful SIC is that the BS should be able to decode the stronger user's message with a rate at least equal to its target rate [74], [75]. Therefore, considered user grouping strategies do not hamper successful SIC at the BS.

[6]We would like to point out that the considered user grouping strategies, NF grouping and NB grouping, are easy-to-implement and effective strategies. Although optimal user grouping strategies will optimize the given objective function, they will put a computational burden on the system. Furthermore, design of optimal user grouping strategies is an interesting avenue for research. However, it is not the focus of our current work. Our contributions are focused on employing easy-to-implement user grouping strategies and understand their interplay with low-overhead channel estimation schemes in the light of ensuring fairness in users' performance in terms of achievable SE.

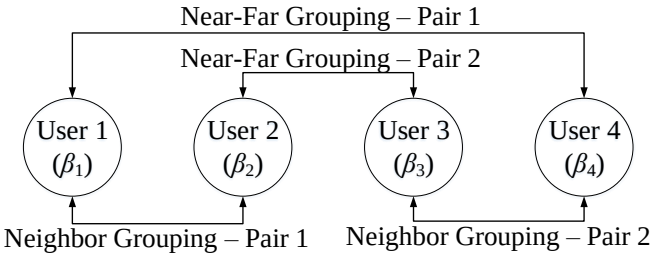


Fig. 2.3: User grouping strategies.

grouping strategy, user 1 is grouped with user 4, and user 2 is grouped with user 3. On the other hand, based on NB grouping strategy, user 1 is grouped with user 2, and user 3 is grouped with user 4. Let $\eta_i^{j,t}$ denote the variance of the channel estimate of user i for grouping strategy $j \in \{\text{NF}, \text{NB}\}$ using the channel estimation Scheme-t where $t \in \{\text{I}, \text{S}\}$ explained in Section 2.1.1. Therefore, with NF grouping,

$$\eta_1^{\text{NF,I}} = \frac{\tau_p \rho_{ul} \beta_1^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_1 + \beta_4)}, \tag{2.64}$$

$$\eta_4^{\text{NF,I}} = \frac{\tau_p \rho_{ul} \beta_4^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_1 + \beta_4)}, \tag{2.65}$$

$$\eta_1^{\text{NF,S}} = \frac{\tau_p \rho_{ul} (\beta_1 + \beta_4)^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_1 + \beta_4)}, \tag{2.66}$$

$$\eta_4^{\text{NF,S}} = \frac{\tau_p \rho_{ul} (\beta_1 + \beta_4)^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_1 + \beta_4)}. \tag{2.67}$$

Similarly, with NB grouping,

$$\eta_1^{\text{NB,I}} = \frac{\tau_p \rho_{ul} \beta_1^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_1 + \beta_2)}, \tag{2.68}$$

$$\eta_4^{\text{NB,I}} = \frac{\tau_p \rho_{ul} \beta_4^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_3 + \beta_4)}, \tag{2.69}$$

$$\eta_1^{\text{NB,S}} = \frac{\tau_p \rho_{ul} (\beta_1 + \beta_2)^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_1 + \beta_2)}, \tag{2.70}$$

$$\eta_4^{\text{NB,S}} = \frac{\tau_p \rho_{ul} (\beta_3 + \beta_4)^2}{\sigma^2 + \tau_p \rho_{ul}(\beta_3 + \beta_4)}. \tag{2.71}$$

Using the relationship that $\beta_1 \geq \beta_2 \geq \beta_3 \geq \beta_4$ and the fact that square $(\cdot)^2$ is a monotonically increasing function,

$$\eta_4^{\text{NB,I}} \geq \eta_4^{\text{NF,I}}, \tag{2.72}$$

Table 2.1: Simulation parameters.

Parameter	Description	Justification
Inner radius (d_o)	100 m	A typical urban micro-cell
Outer radius	300 m	A typical urban micro-cell
Path-loss model (β_{nk})	$\beta_{nk} = 10^{(\phi/10)}\beta_o\,(d_o/d_{nk})^{\nu}$, where $\phi \sim \mathcal{CN}(0,\sigma_s^2)$, $\sigma_s^2 = 4$ dB is the variance of log-normal shadowing	3GPP path loss model for a typical urban micro-cell [79]
Reference path loss (β_o)	−35.3 dB	Computed using the model in [79]
Path loss exponent (ν)	2.4	A widely used path loss exponent value in literature [11]
Length of a coherence interval (τ_c)	168	
Length of each pilot sequence $\left(\tau_p\right)$	N = number of groups	
Bandwidth	20 MHz	A typically used value in the literature [80]
BS noise figure	7 dB	
Noise power $\left(\sigma^2\right)$	−93.8 dBm	A typically used value in the literature [80]
$\mathcal{UE}$ transmit power (ρ_{ul})	20 dBm	A typically used value [81]
User locations	Users are dropped uniformly within the annular ring	

$$\eta_4^{\mathrm{NF,S}} \geq \eta_4^{\mathrm{NB,S}}, \tag{2.73}$$

$$\eta_4^{\mathrm{NB,S}} \geq \eta_4^{\mathrm{NB,I}}. \tag{2.74}$$

Based on (2.72)-(2.74), it is clear that, for channel estimation using Scheme-I, the NB grouping strengthens the channel estimate of the weakest user, i.e., user 4, while for channel estimation using Scheme-S, the NF grouping strengthens the channel estimate of the weakest user. Furthermore, unlike NF grouping, NB grouping, by design, is a scalable strategy in which it is feasible to have more than two users per group. Therefore, along with the ZF decoder, we can use NB grouping to serve a very large number of users by increasing K and keeping $N(< \min\{M, \tau_c\})$ constant while maintaining a non-zero data rate. From (2.72)-(2.74), it is clear that irrespective of grouping strategy, the weakest user's channel estimate is better with Scheme-S than with Scheme-I.

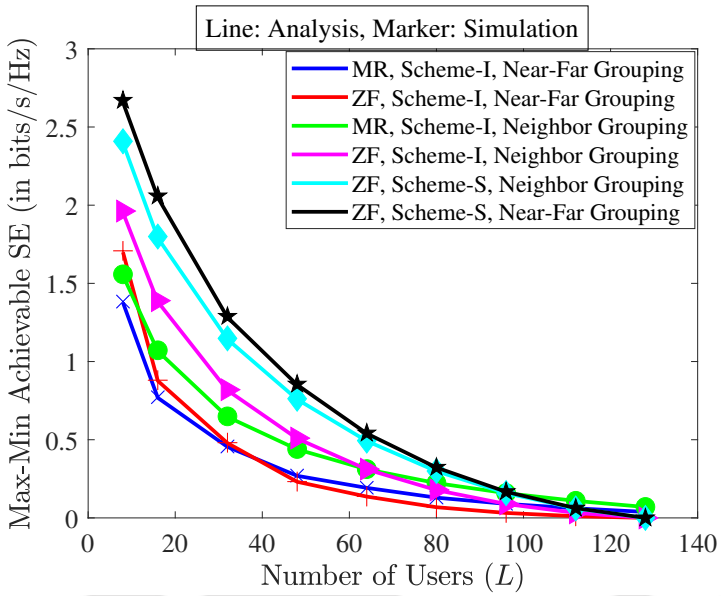


Fig. 2.4: Max-min achievable SE comparison: over-loading, $M = 64$, $K = 2$.

2.5 Numerical and Simulation Results

In this section, we present numerical results to quantify the improvement in the uplink max-min achievable SE of a NOMA-based massive MIMO system that uses a ZF decoder relative to that obtained using an MR decoder. We also present results to analyze the impact of two different grouping strategies, namely, NF grouping and NB grouping, on the max-min achievable SE with ZF decoder and prove that NOMA-based massive MIMO system with NB grouping can serve significantly more numbers of users with a reasonable max-min achievable SE when compared against conventional massive MIMO system with ZF decoder. The details of the simulation parameters are given in Table 2.1. Unless mentioned otherwise, we take $K = 2$, i.e., two users per group. We obtain the max-min power control coefficients to ensure that the minimum achievable SE is maximized and users get uniform quality-of-service.

Figure 2.4 plots the max-min achievable rate averaged over user locations for uplink of a massive MIMO-NOMA system as a function of the number of users L with ZF decoder designed using channel estimates obtained via Scheme-I and Scheme-S for both the grouping strategies, namely, NB grouping and NF grouping for $M = 64$. We present Monte Carlo simulations to corroborate our analysis. We also benchmark against the performance achieved with MR decoding [11], assuming perfect SIC at the BS. We observe that the ZF decoder based on Scheme-S

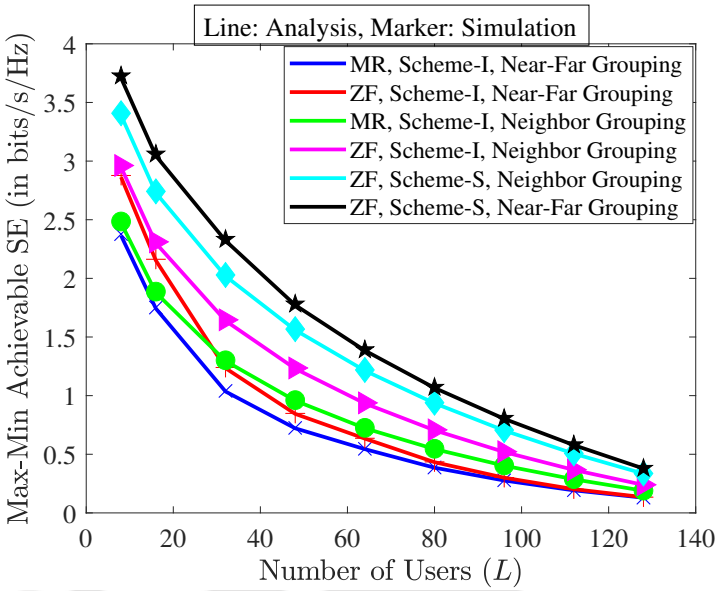


Fig. 2.5: Max-min achievable SE comparison: under-loading, $M = 256$, $K = 2$.

L	ZF, Scheme-S, NF	ZF, Scheme-I, NF	ZF, Scheme-S, NB	ZF, Scheme-I, NB
40	1.2087	0.3881	1.0808	0.7459
80	0.4695	0.0956	0.4347	0.2588

Table 2.2: Comparison of SEs (in bps/Hz) of decoding schemes.

gives a higher max-min achievable SE compared to that obtained under the ZF decoder based on Scheme-I for both grouping strategies. Specifically, in under-loaded regime, e.g., $L = 40$, when NF grouping is employed, the max-min achievable SE of ZF based on Scheme-S is 3 times that of ZF based on Scheme-I. On the other hand, in the over-loaded regime, e.g., $L = 80$, when NF grouping is employed, the max-min achievable SE of ZF based on Scheme-S is almost 5 times that of ZF based on Scheme-I. Similarly, when NB grouping is employed, the max-min achievable SE with ZF based on Scheme-S is approximately 1.5 times of that with ZF based on Scheme-I, for both under-loaded (e.g., $L = 40$) and over-loaded (e.g., $L = 80$) regimes. This is primarily because, Scheme-S boosts the strength of the channel estimate of the weakest user when compared against Scheme-I, as explained in Section 2.1.1. Also, under Scheme-S, NF grouping gives a better quality channel estimate to the weakest user when compared against NB grouping. Hence, the ZF decoder based on Scheme-S gives a higher max-min rate with NF grouping than NB grouping. Specifically, in the under-loaded regime, e.g., $L = 40$, the max-min achievable SE with ZF based on Schemes-S under NF grouping is almost 12% higher than that

with ZF based on Scheme-S under NB grouping.

Furthermore, for $L < M$ (under-loading), the ZF decoder based on Scheme-S and Scheme-I both outperform the MR decoder because of their ability to cancel IGI. For $L \geq M$ (over-loading), the ZF decoder based on Scheme-S gives the highest max-min achievable SE among all considered strategies for a substantial portion of the overloaded regime. However, as the number of users increases, the BF gain achieved with the ZF decoder reduces, but the BF gain with MR decoder is independent of the number of users and remains constant at M. Hence, the max-min SE achieved with the MR decoder is marginally higher than that achieved with ZF decoder for $L \gg M$.

Figure 2.5 shows the max-min achievable SE comparison for $M = 256 > L$, i.e., the under-loaded regime. As M increases to 256, the max-min SE increases relative to that illustrated in Fig. 2.4 for $M = 64$ because of the consequent increase in the BF gain for all strategies. In this under-loaded regime, we can see that the ZF decoder based on Scheme-S with NF grouping gives the highest max-min SE among all the channel estimation, user grouping and NOMA decoding strategies considered. This is essentially due to the complete cancellation of IGI by the ZF decoder based on Scheme-S and the fact that for Scheme-S, NF grouping boosts the strength of the estimated channel to the weakest user in the network.

Note that NF grouping is not a scalable strategy, whereas NB grouping is a scalable strategy. From Figs. 2.4 and 2.5, it is clear that the ZF decoder based on Scheme-S with scalable NB grouping outperforms every other scalable strategy over a substantial portion of the over-loaded regime. Therefore, we can serve a large number of users by employing the ZF decoder based on Scheme-S with NB grouping obtaining a significantly higher max-min SE in a large portion of the over-loaded regime.

In Fig. 2.6, we plot the max-min achievable rate versus the number of users (L) in the network for Scheme-S with ZF decoding and scalable NB grouping strategy in the over-loaded regime, i.e., for $L \geq M = 120$. We show results for three different values of K (number of users per group), namely, $K = 2$, 3, and 4. We also benchmark against the max-min rate achieved using conventional massive MIMO decoders in the over-loaded regime. From this

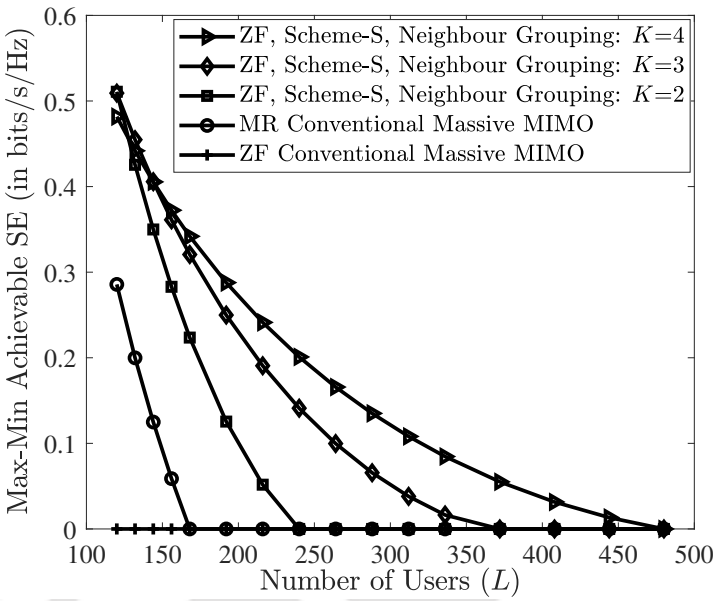


Fig. 2.6: Max-min achievable SE comparison: over-loading, $M = 120$.

plot, we observe that in the over-loaded regime, the BF gain $(M - N)$, where M is the number of antennas at the BS and N is the number of user groups, reduces to zero at a faster rate with $K = 2$ users per group when compared against $K = 4$ users per group. This is because for a given number of users $L = NK$ in the network, N is larger with smaller K. As expected, the max-min SE achieved by any user decreases as L increases.

Furthermore, with ZF decoder based on Scheme-S and with NB grouping strategy, a relatively larger number of users can be served while maintaining a reasonable max-min SE in contrast to the number of users served using conventional massive MIMO decoders. To be precise, under ZF decoding based on Scheme-S and NB grouping, one can serve $K\times$ number of users, i.e., $\min\{MK, \tau_c K\}$ users for $N < \min\{M, \tau_c\}$. Contrary to this, zero max-min SE is achieved with conventional ZF and conventional MR decoders for $L \geq \min\{M, \tau_c\}$ and $L \geq \tau_c$, respectively.

In contrast, in Fig. 2.7, we plot max-min achievable SE versus the number of users (L) in the network with ZF decoding based on Scheme-S and the scalable NB grouping strategy in the under-loaded regime, i.e., for $L < M = 200$. In the under-loaded regime, we observe a reversal in the performance trend. Specifically, in the under-loaded regime, as we increase the number of users in a group from $K = 2$ to $K = 4$, due to increase in the number of users that

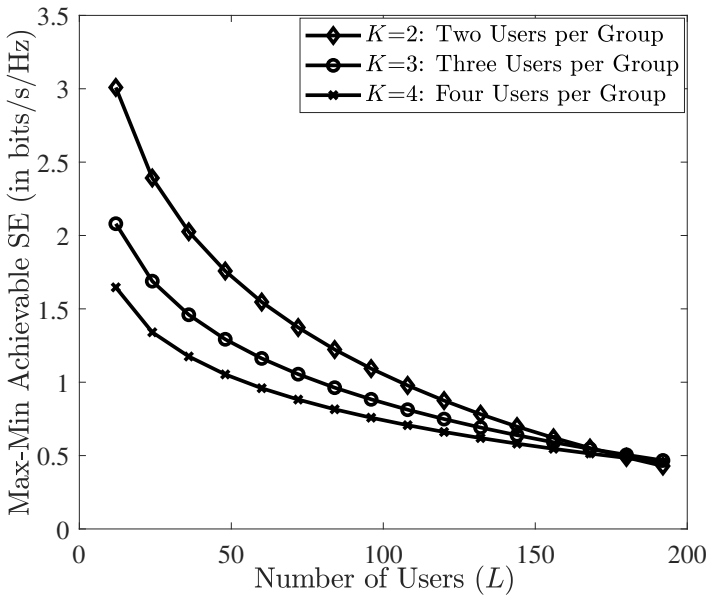


Fig. 2.7: Max-min achievable SE comparison: under-loading, $M = 200$.

share the same pilot sequence within a group, the intra-group pilot contamination increases. Consequently the achievable max-min SE decreases because the reduction in the max-min SE due to increased intra-group pilot contamination dominates over the increment in the max-min SE due to improved BF gain.

Figure 2.8 compares the max-min sum SE for the ZF decoder based on Scheme-S with scalable NB grouping for different number of users per group ($K = 2, 3, 4$). We observe that the sum-SE under max-min power allocation first increases as L increases and then reduces since the channel estimation overhead reduces the number of samples left for data transmission in a coherence interval. The ZF decoder based on Scheme-S with NB grouping strategy can provide a reasonable max-min sum-SE while serving a larger number of users, precisely $\min\{MK, \tau_c K\}$ for $N < \min\{M, \tau_c\}$ with K users per group.

Observe that the performance trend reverses from under-loaded regime ($L < M$) to the over-loaded regime ($L \geq M$). In the under-loaded regime, as the number of users within a group increases, the effect of intra-group pilot contamination increases and the max-min sum-SE decreases. Therefore, the max-min sum-SE performance for $K = 2$ is higher than that for $K = 4$ in the under-loaded regime. To increase $L = NK$ for a fixed value of K, we need to increase N. Consequently, as L increases, the BF gain $(M - N)$ decreases. Notice that as L

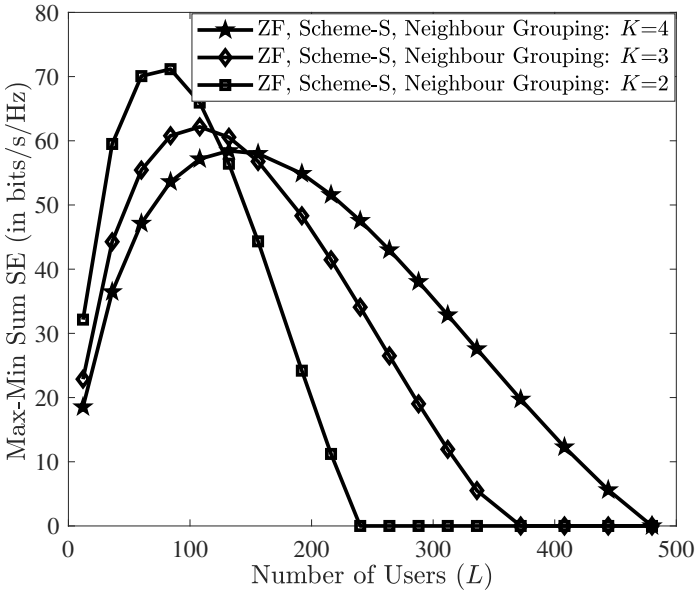
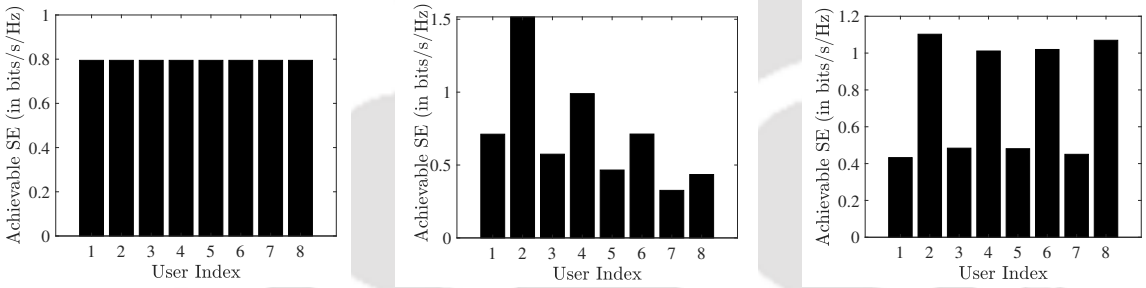


Fig. 2.8: Max-min achievable sum SE comparison: $M = 120$.



(a) Max-min power allocation. **(b)** Greedy power allocation. **(c)** Inversion power allocation.

Fig. 2.9: Fairness comparison among SE achieved with different power allocation policies using proposed ZF decoder based on Scheme-S and neighbor grouping: $M = 8$, $L = 8$, $K = 2$.

increases, the BF gain decreases the fastest for $K = 2$ and the BF gain decreases the slowest for $K = 4$ because for a fixed value of L, $N = L/2$ for $K = 2$ and $N = L/4$ for $K = 4$. Moreover, the impact of the BF gain variation on the max-min sum-SE is more prominent than that of the intra-group pilot contamination in the over-loaded regime. Therefore, we observe a reversal in the performance trend when we move from the under-loaded regime ($L < M$) to the over-loaded regime ($L \geq M$).

In Figs. 2.9(a), 2.9(b) and 2.9(c), we plot the achievable SE (with the ZF decoder designed based on Scheme-S and NB grouping) versus user index for three different power allocation policies. The plot in Fig. 2.9(a) is for max-min power allocation in which optimal power al-

location coefficients are derived to ensure that the minimum user SE is maximized and every user gets the same rate at optimality irrespective of user locations. The plot in Fig. 2.9(b) is for greedy power allocation in which every user transmits with full power equal to ρ_{ul} intending to maximize its own signal strength, and hence every user ends up getting a different achievable SE. And the plot in Fig. 2.9(c) is for inversion power allocation in which k^{th} user in n^{th} group transmits with power equal to

$$\gamma_{nk} = \rho_{ul} \frac{\min_{i,j}\{\eta_{ij}\}}{\eta_{nk}} \tag{2.75}$$

leading to the user with the minimum estimated channel strength transmitting with the full power and the user with the maximum estimated channel strength transmitting with the least power. As we can clearly see in Figs. 2.9(a), 2.9(b), and 2.9(c), the max-min power control is fairer than the greedy power control and inversion power control. Another precise way to quantify this fact is to observe Jain's fairness index for these three power allocation policies [82], [83]. The Jain's fairness index is given by

$$J = \frac{\left(\sum_{i=1}^{L} R_i\right)^2}{L \sum_{i=1}^{L} R_i^2}, \tag{2.76}$$

where L is the total number of users and R_i is the achievable rate of user i. Notice that $1/M \leq J \leq 1$. Also, $J = 1$ indicates the most fair policy, and $J = 1/M$ indicates the least fair policy. For the plots in Figs. 2.9(a), 2.9(b), 2.9(c), the Jain's fairness index values are $J = 1$, $J = 0.77$ and $J = 0.86$, respectively.

2.6 Summary

We analyzed the uplink of a NOMA-based massive MIMO system with a ZF decoder designed using estimated uplink CSI at the BS to cancel IGI. We considered two low overhead channel estimation schemes, namely, Scheme-I and Scheme-S, and two user grouping strategies, namely, NF grouping and NB grouping. We derived new expressions for SINR and lower bounds on the achievable SE corresponding to both the estimation schemes. To ensure uniform quality-of-service, we formulated a max-min transmit power allocation problem and obtained corresponding max-min power allocation coefficients using geometric programming and CVX.

Through extensive simulations based on derived achievable SE expressions, we draw the following inferences:

- In a substantial portion of the over-loaded regime ($L \geq M$) and in entire under-loaded regime ($L < M$), the designed ZF decoder based on Scheme-S with NF grouping strategy gives the highest max-min SE among all the considered NOMA decoding and grouping strategies.
- Also, for $L \geq M$, the designed ZF decoder based on Scheme-S with scalable NB grouping strategy can support a larger number of users, precisely $\min\{MK, \tau_c K\}$, with a reasonable max-min SE when compared against conventional ZF decoder and conventional MR decoder that cannot support more than $\min\{M, \tau_c\}$ and τ_c users, respectively.
- Furthermore, among all the considered scalable strategies, for $L < M, L \geq M$ but $L \gg M$, the designed ZF decoder based on Scheme-S with scalable NB grouping gives the highest max-min SE and can simultaneously serve a very large number of users, precisely $\min\{MK, \tau_c K\}$ for $N < \min\{M, \tau_c\}$.
- However, for $L \gg M$, MR decoder based on Scheme-I with NB grouping achieves marginally higher max-min SE when compared against ZF decoder, since the BF gain with MR decoder is M and is independent of L, whereas, with ZF decoder, the BF gain $(M - N)$ decreases as L increases keeping K fixed.

3

Massive MIMO-NOMA Systems under Channel Aging

Contents

3.1 System Model 63
3.2 Spectral Efficiency Analysis 68
3.3 Mitigation of Impact of Channel Aging using Channel Prediction 79
3.4 Power Control 84
3.5 Numerical and Simulation Results 87
3.6 Summary 96
3.7 Comparison of Contributions in Chapters 2 and 3 97

In the previous chapter, we investigated and discussed the challenges in uplink massive MIMO-NOMA systems. Therein, we considered quasi-static Rayleigh faded channel where channel is assumed to be invariant over a coherence interval and hence channel needs to be estimated once every coherence interval. However, in case of high mobility, length of the coherence interval reduces. With a significantly large number of high-speed mobile users in the network, there could be a significant delay between the time when the channel estimates are acquired and the time when the estimates are used for data processing. This phenomenon is called as channel aging [8].

In this chapter, we investigate the uplink massive MIMO-NOMA systems under channel aging to quantify the collective impact of channel outdatedness, pilot contamination and imperfect SIC. In this direction, we first present uplink pilot signaling and Jakes model to incorporate effects of channel aging and develop expressions for MMSE-based outdated channel estimates using two low-overhead estimation schemes, namely, Scheme-I and Scheme-S. We then design ZF and MR decoders using the outdated channel estimates obtained with Scheme-I and Scheme-S, and derive corresponding novel lower bounds on achievable SE. The presented analysis captures the impact of imperfect SIC along with estimation error and channel aging.

Thereafter, as a counter-measure to channel outdatedness, we derive an optimal WLP to predict the channel and develop corresponding achievable SE expressions for ZF and MR decoders designed using predicted CSI. With the aim of improving performance fairness among users in terms achievable SE, we formulate and solve max-min fairness and proportional fairness optimization problems, and obtain corresponding power allocation coefficients using convex programming. We also present extensive numerical results to elucidate the impact of aging, pilot contamination and imperfect SIC on power control and the type of decoding strategy used in these systems.

Chapter Organization: In Section 3.1, we discuss system model, channel estimation schemes and uplink data transmission. In Section 3.2, we present the detailed SE analysis for ZF and MR decoders based on outdated CSI. Therein, we also discuss max-min power control and proportional fairness power control. Next, in Section 3.3, we elucidate the SE analysis for

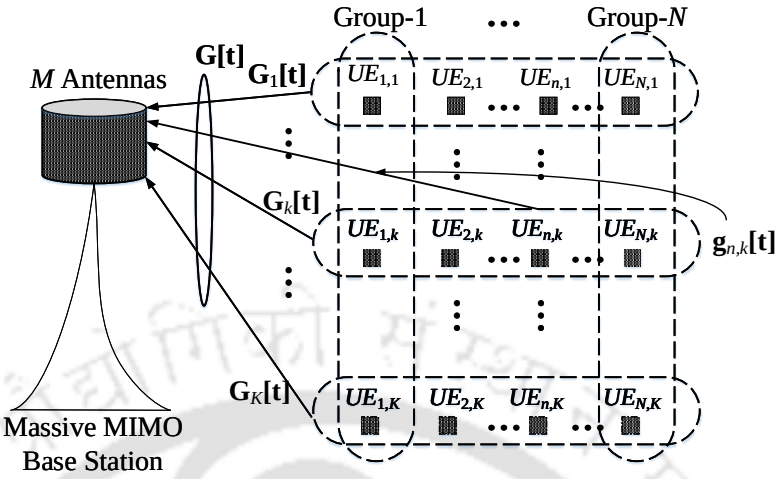


Fig. 3.1: System model: Uplink massive MIMO-NOMA under time-variant channel.

ZF and MR decoders based on the predicted CSI. Extensive numerical results and discussion are provided in Section 3.5. Finally, we state conclusions in Section 3.6.

3.1 System Model

We investigate a massive MIMO-NOMA system where L single-antenna users are served by a BS with M antennas. User grouping is necessary to employ NOMA. We form N groups of users with K users in each group, where $L = N \times K$. User grouping is based on users' channel strengths. Let UE_{nk} denote k^{th} user in the n^{th} group, where $k \in \mathcal{K} = \{1, 2, \cdots, K\}$ and $n \in \mathcal{N} = \{1, 2, \cdots, N\}$. We denote the channel vector from the BS to UE_{nk} at t^{th} symbol instant by $\mathbf{g}_{nk}[t] \in \mathbb{C}^{M\times 1}$. We consider that $\mathbf{g}_{nk}[t] \sim \mathcal{CN}(\mathbf{0}, \beta_{nk}\mathbf{I}_M)$ and it varies slowly across symbols. Note that β_{nk}, which denotes large-scale fading, varies at a much slower time scale compared to small scale fading coefficients [84]. Let $\mathbf{G}[t] \in \mathbb{C}^{M\times L}$ represent the concatenated channel matrix at t^{th} symbol instant to all users. Therefore, $\mathbf{G}[t]$ is given by $\mathbf{G}[t] = [\mathbf{G}_1[t]\ \mathbf{G}_2[t]\ \cdots\ \mathbf{G}_K[t]]$, where $\mathbf{G}_i[t] \in \mathbb{C}^{M\times N}$ represents the concatenated matrix of channel coefficients to the i^{th} user in all N groups at t^{th} symbol instant and is given by $\mathbf{G}_i[t] = [\mathbf{g}_{1i}[t]\ \mathbf{g}_{2i}[t] \cdots \mathbf{g}_{Ni}[t]]$, for $i \in \mathcal{K}$, as shown in Fig. 3.1. Based on the aging model in [8], the composite channel matrix at $(t+u)^{\text{th}}$ symbol instant is

$$\mathbf{G}[t+u] = \alpha(u)\mathbf{G}[t] - \mathbf{E}[t+u], \tag{3.1}$$

where $\alpha(u)$ captures the correlation between $\mathbf{G}[t]$ and $\mathbf{G}[t+u]$, and $\mathbf{E}[t+u]$ is the temporally uncorrelated complex white Gaussian noise matrix at the $(t+u)^{\text{th}}$ symbol instant. Note that $\mathbf{E}[t+u]$ is the error due to channel aging [8]. According to Jakes model, $\alpha(u) = J_0\left(2\pi f_D T_S |u|\right)$. Here J_0, T_S and $f_D = \frac{vf_c}{c}$ respectively denote the zeroth-order Bessel function of the first kind, the channel sampling interval, and the maximum Doppler frequency shift for velocity v and carrier frequency f_c (c denotes the speed of light). Note that $0 \leq |\alpha(u)| \leq 1$, and channel gets increasingly outdated as the value of $|\alpha(u)|$ decreases. Also, error due to channel aging is given by

$$\mathbf{E}[t+u] = [\mathbf{E}_1[t+u]\ \mathbf{E}_2[t+u]\ \cdots\ \mathbf{E}_K[t+u]]\,, \tag{3.2}$$

with $\mathbf{E}_i[t+u] = [\mathbf{e}_{1i}[t+u]\ \cdots\ \mathbf{e}_{Ni}[t+u]]$, where $\mathbf{e}_{nk}[t+u] \sim \mathcal{CN}\left(0, \left(1-\alpha(u)^2\right)\beta_{nk}\mathbf{I}_M\right)$, for $n \in \mathcal{N}$ and $k \in \mathcal{K}$. Using this model, we can rewrite (3.1) as

$$\mathbf{G}[t+u] = \alpha(u)\left(\widehat{\mathbf{G}}[t] - \widetilde{\mathbf{G}}[t]\right) - \mathbf{E}[t+u], \tag{3.3}$$

where $\widehat{\mathbf{G}}[t]$ and $\widetilde{\mathbf{G}}[t]$ denote the estimated CSI at t^{th} symbol instant at the BS and the corresponding error in estimation.[1] The channel estimation is performed at the BS in the uplink pilot training phase. The channel estimates age with time and are used to design the ZF and MR decoders at the BS.

3.1.1 Channel Estimation

We consider a frame of observation of T symbols, in which initial τ_p symbols are used for uplink pilot training for channel estimation at the BS[2] and remaining $T - \tau_p$ symbols are used for data transmission. During the data transmission phase, the channel is assumed to be constant for the period of one symbol and varying slowly from symbol to symbol [8, 48, 54].

To estimate the uplink channel coefficients, an orthogonal pilot sequence is allocated to every group of users. The τ_p length pilot sequence allocated to the n^{th} group is denoted by

[1] It is important to note that the channel aging emerges from the outdatedness of the estimated channel due to time variations in channel caused by mobility-induced Doppler shift. The model used for outdated channel is based on the time correlation between the estimated channel and the outdated channel. Therefore, modelling outdated channel without accounting for the channel estimation errors would not be justified.

[2] The underlying assumption is that the channel remains constant during estimation phase. During estimation phase of length τ_p, we neglect the error in estimation that would arise due to channel aging [8, 48, 54, 85].

$\boldsymbol{\psi}_n \in \mathbb{C}^{\tau_p}$. All users within the n^{th} group transmit $\boldsymbol{\psi}_n$ on the uplink for channel estimation. Note that the allocated pilot sequences are orthogonal across groups. Therefore, $\boldsymbol{\psi}_i'\boldsymbol{\psi}_j = \delta_{ij}$, where $\delta_{ij} = 1$ for $i = j$, and $\delta_{ij} = 0$ otherwise. This orthogonality is ensured only if $\tau_p \geq N$, where N denotes the total number of groups.

The signal $\mathbf{Y}_p[t] \in \mathbb{C}^{M\times\tau_p}$ received at the BS during the channel estimation phase is given by

$$\mathbf{Y}_p[t] = \sqrt{\tau_p\rho_{ul}}\,[\mathbf{G}_1[t] + \cdots + \mathbf{G}_K[t]]\,\boldsymbol{\Psi}' + \mathbf{W}_p[t]\,, \tag{3.4}$$

where ρ_{ul} denotes the transmit power of each user, $\boldsymbol{\Psi} = \begin{bmatrix}\boldsymbol{\psi}_1 & \boldsymbol{\psi}_2 & \cdots & \boldsymbol{\psi}_N\end{bmatrix} \in \mathbb{C}^{\tau_p\times N}$ denotes composite pilot matrix. Moreover, $\mathbf{W}_p[t] \in \mathbb{C}^{M\times\tau_p}$ denotes the AWGN matrix with entries which are i.i.d. and follow $\mathcal{CN}(0,\sigma^2)$ distribution. We can rewrite (3.4) as

$$\mathbf{Y}_p[t] = \sqrt{\tau_p\rho_{ul}}\sum_{n=1}^{N}\sum_{i=1}^{K}\mathbf{g}_{ni}[t]\boldsymbol{\psi}_n' + \mathbf{W}_p[t]. \tag{3.5}$$

The BS exploits the orthogonality of pilot sequences and obtains the de-spreaded signal $\overline{\mathbf{y}}_{pn}[t] \in \mathbb{C}^{M\times 1}$ given by

$$\overline{\mathbf{y}}_{pn}[t] \triangleq \mathbf{Y}_p[t]\boldsymbol{\psi}_n = \sqrt{\tau_p\rho_{ul}}\sum_{i=1}^{K}\mathbf{g}_{ni}[t] + \overline{\mathbf{w}}_{pn}[t]\,, \tag{3.6}$$

where $\overline{\mathbf{w}}_{pn}[t] = \mathbf{W}_p[t]\boldsymbol{\psi}_n \sim \mathcal{CN}(0,\sigma^2\mathbf{I}_M)$. The BS uses $\overline{\mathbf{y}}_{pn}[t]$ to estimate CSI corresponding to users in the n^{th} group.

Using $\overline{\mathbf{y}}_{pn}[t]$, below we explain two MMSE-based estimation schemes, Scheme-I and Scheme-S, employed at the BS. With Scheme-I, the aim is to obtain the CSI estimate for all users individually. On the other hand, with Scheme-S, the aim is to obtain the CSI estimate for each group. Therefore, for each group of users, an estimate is obtained for the sum of all users' channel vectors within the group.[3]

[3]The estimation overhead and type of pilot assignment for both schemes is essentially the same, and considered schemes reduce the overhead to just N symbols, unlike $N \times K$ symbols for conventional schemes.

3.1.1.1 MMSE Estimates with Scheme-I

Using $\overline{\mathbf{y}}_{pn}[t]$ in (3.6), the MMSE estimate $\hat{\mathbf{g}}_{nk}[t]$ of $\mathbf{g}_{nk}[t]$ to UE_{nk} is given by [72][4]

$$\hat{\mathbf{g}}_{nk}[t] = \zeta_{nk}\,\overline{\mathbf{y}}_{pn}[t] = \mathbf{g}_{nk}[t] + \widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t]\,, \tag{3.7}$$

where $\zeta_{nk} = \frac{\sqrt{\tau_p\rho_{ul}}\beta_{nk}}{\sigma^2+\tau_p\rho_{ul}\sum_{i=1}^{K}\beta_{ni}}$ and $\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t]$ is the error in estimate which is equal to

$$\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t] = \left(\sqrt{\tau_p\rho_{ul}}\zeta_{nk} - 1\right)\mathbf{g}_{nk}[t] + \sqrt{\tau_p\rho_{ul}}\zeta_{nk}\sum_{j=1,j\neq k}^{K}\mathbf{g}_{nj}[t] + \zeta_{nk}\overline{\mathbf{w}}_{pn}[t]. \tag{3.8}$$

Note that $\hat{\mathbf{g}}_{nk}[t] \sim \mathcal{CN}\left(\mathbf{0}, \eta_{nk}\mathbf{I}_M\right)$ and $\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t] \sim \mathcal{CN}\left(\mathbf{0}, (\beta_{nk}-\eta_{nk})\,\mathbf{I}_M\right)$, where

$$\eta_{nk} = \frac{\tau_p\rho_{ul}\beta_{nk}^2}{\sigma^2 + \tau_p\rho_{ul}\sum_{i=1}^{K}\beta_{ni}}. \tag{3.9}$$

With MMSE estimation, $\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t]$ and $\hat{\mathbf{g}}_{nk}[t]$ are uncorrelated to each other [14]. The estimate $\hat{\mathbf{g}}_{nk}[t]$ can be normalized as $\hat{\mathbf{g}}_{nk}[t] = \sqrt{\eta_{nk}}\,\mathbf{z}_n[t]$. Here, the elements of $\mathbf{z}_n[t] \in \mathbb{C}^{M\times 1}$ are i.i.d. and $\mathcal{CN}(0,1)$ distributed. From (3.7), it is evident that, for all users within a group, the estimates of channel vectors are perfectly correlated. Thus, $\mathbf{z}_n[t]$ represents the normalized CSI estimate for all K users in the n^{th} group. With Scheme-I, let $\widehat{\mathbf{G}}^{\mathrm{I}}[t] = \left[\widehat{\mathbf{G}}^{\mathrm{I}}_1[t]\ \widehat{\mathbf{G}}^{\mathrm{I}}_2[t]\cdots\widehat{\mathbf{G}}^{\mathrm{I}}_K[t]\right]$ and $\widetilde{\mathbf{G}}^{\mathrm{I}}[t] = \left[\widetilde{\mathbf{G}}^{\mathrm{I}}_1[t]\ \widetilde{\mathbf{G}}^{\mathrm{I}}_2[t]\cdots\widetilde{\mathbf{G}}^{\mathrm{I}}_K[t]\right]$, respectively, denote the concatenated CSI estimate matrix and the corresponding concatenated estimation error matrix for all L users. Here, $\widehat{\mathbf{G}}^{\mathrm{I}}_i[t] = \left[\hat{\mathbf{g}}_{1i}[t]\ \ \hat{\mathbf{g}}_{2i}[t]\ \ \cdots\ \ \hat{\mathbf{g}}_{Ni}[t]\right]$ and $\widetilde{\mathbf{G}}^{\mathrm{I}}_i[t] = \left[\widetilde{\mathbf{g}}^{\mathrm{I}}_{1i}[t]\ \ \widetilde{\mathbf{g}}^{\mathrm{I}}_{2i}[t]\ \ \cdots\ \ \widetilde{\mathbf{g}}^{\mathrm{I}}_{Ni}[t]\right]$, $\forall\, i \in \mathcal{K}$. As we know, $\hat{\mathbf{g}}_{ni}[t] = \sqrt{\eta_{ni}}\,\mathbf{z}_n[t]$, for $n \in \mathcal{N}$,

$$\widehat{\mathbf{G}}^{\mathrm{I}}_i[t] = \left[\sqrt{\eta_{1i}}\,\mathbf{z}_1[t]\ \cdots\ \sqrt{\eta_{Ni}}\,\mathbf{z}_N[t]\right] = \mathbf{Z}[t]\mathbf{D}^{\frac{1}{2}}_{\eta_i}, \tag{3.10}$$

where $\mathbf{Z}[t] = [\mathbf{z}_1[t]\ \mathbf{z}_2[t]\cdots\mathbf{z}_N[t]]$ and $\mathbf{D}_{\eta_i} = \mathrm{diag}\left[\eta_{1i}, \eta_{2i}, \cdots, \eta_{Ni}\right]$, for $i \in \mathcal{K}$. Note that $\mathbb{E}\left\{\mathbf{z}_i[t]\mathbf{z}_j[t]'\right\} = \mathbf{0}$ (an all zeros matrix) for $i \neq j$.

[4]Total computational complexity with Scheme-I involves $\left(4M\tau_p + 2MK + K + 1\right)N + 1$ real multiplications, $\left(2M\tau_p + 2M(\tau_p - 1) + K\right)N$ real additions and NK real divisions.

3.1.1.2 MMSE Estimates with Scheme-S

Under Scheme-S, the BS obtains the sum channel estimate for all users within a group. Specifically, for the n^{th} group, the estimate is obtained for $\sum_{i=1}^{K} \mathbf{g}_{ni}[t]$. We can modify (3.6) as

$$\overline{\mathbf{y}}_{pn}[t] = \sqrt{\tau_p \rho_{ul}}\mathbf{c}_n[t] + \overline{\mathbf{w}}_{pn}[t]\,, \tag{3.11}$$

where $\mathbf{c}_n[t] = \sum_{i=1}^{K} \mathbf{g}_{ni}[t]$. From (3.11), the MMSE estimate $\hat{\mathbf{c}}_n[t]$ of $\mathbf{c}_n[t]$ is given by[5]

$$\hat{\mathbf{c}}_n[t] = \mu_n\, \overline{\mathbf{y}}_{pn}[t] = \mathbf{g}_{nk}[t] + \widetilde{\mathbf{g}}^{\mathrm{S}}_{nk}[t], \tag{3.12}$$

where $\mu_n = \frac{\sqrt{\tau_p \rho_{ul}} \sum_{i=1}^{K} \beta_{ni}}{\sigma^2 + \tau_p \rho_{ul} \sum_{i=1}^{K} \beta_{ni}}$ and $\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk}[t]$ is corresponding error in the estimate which is given by

$$\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk}[t] = \left(\sqrt{\tau_p \rho_{ul}}\mu_n - 1\right)\mathbf{g}_{nk}[t] + \sqrt{\tau_p \rho_{ul}}\mu_n \sum_{j=1, j\neq k}^{K} \mathbf{g}_{nj}[t] + \mu_n \overline{\mathbf{w}}_{pn}[t]. \tag{3.13}$$

Here $\hat{\mathbf{c}}_n[t] \sim \mathcal{CN}(\mathbf{0}, \eta_n \mathbf{I}_M)$, $\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk}[t] \sim \mathcal{CN}(\mathbf{0}, \iota_{nk}\mathbf{I}_M)$, where

$$\eta_n = \frac{\tau_p \rho_{ul} \left(\sum_{i=1}^{K} \beta_{ni}\right)^2}{\sigma^2 + \tau_p \rho_{ul} \sum_{i=1}^{K} \beta_{ni}}, \tag{3.14}$$

and $\iota_{nk} = \left(\sqrt{\tau_p \rho_{ul}}\mu_n - 1\right)^2 \beta_{nk} + \left(\sqrt{\tau_p \rho_{ul}}\mu_n\right)^2 \sum_{j=1, j\neq k}^{K} \beta_{nj} + \mu_n^2 \sigma^2$. Also, there is correlation between $\hat{\mathbf{c}}_n[t]$ and $\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk}[t]$. The estimate vector $\hat{\mathbf{c}}_n[t]$ is normalized as

$$\hat{\mathbf{c}}_n[t] = \sqrt{\eta_n}\mathbf{v}_n[t], \tag{3.15}$$

where the entries of $\mathbf{v}_n[t]$ are i.i.d. and $\mathcal{CN}(0, 1)$ distributed. With Scheme-S, the concatenated CSI estimate matrix and the corresponding concatenated estimation error matrix for all L users are given by $\widehat{\mathbf{G}}^{\mathrm{S}}[t] = \left[\widehat{\mathbf{G}}^{\mathrm{S}}_1[t]\ \widehat{\mathbf{G}}^{\mathrm{S}}_2[t] \cdots \widehat{\mathbf{G}}^{\mathrm{S}}_K[t]\right]$ and $\widetilde{\mathbf{G}}^{\mathrm{S}}[t] = \left[\widetilde{\mathbf{G}}^{\mathrm{S}}_1[t]\ \widetilde{\mathbf{G}}^{\mathrm{S}}_2[t] \cdots \widetilde{\mathbf{G}}^{\mathrm{S}}_K[t]\right]$, respectively. Here, $\widehat{\mathbf{G}}^{\mathrm{S}}_i[t] = \widehat{\mathbf{C}}[t] = [\hat{\mathbf{c}}_1[t]\ \hat{\mathbf{c}}_2[t] \cdots \hat{\mathbf{c}}_N[t]]$ and $\widetilde{\mathbf{G}}^{\mathrm{S}}_i[t] = \left[\widetilde{\mathbf{g}}^{\mathrm{S}}_{1i}[t]\ \ \widetilde{\mathbf{g}}^{\mathrm{S}}_{2i}[t]\ \ \cdots\ \ \widetilde{\mathbf{g}}^{\mathrm{S}}_{Ni}[t]\right]$, $\forall\, i \in \mathcal{K}$. It is evident that each entry of $\hat{\mathbf{c}}_n[t]$ has variance η_n and is independent of user index k. Therefore, using (3.15), $\widehat{\mathbf{G}}^{\mathrm{S}}_i[t]$ can be modified as

$$\widehat{\mathbf{G}}^{\mathrm{S}}_i[t] = \widehat{\mathbf{C}}[t] = \mathbf{V}[t]\mathbf{D}_{\eta}^{\frac{1}{2}}\,, \tag{3.16}$$

[5]Total computational complexity with Scheme-S involves $\left(4M\tau_p + 2M + 2\right)N + 1$ real multiplications, $\left(2M\tau_p + 2M(\tau_p - 1) + K\right)N$ real additions and N real divisions.

where $\mathbf{V}[t] = [\mathbf{v}_1[t]\ \mathbf{v}_2[t] \cdots \mathbf{v}_N[t]]$ and $\mathbf{D}_{\eta}^{\frac{1}{2}} = \text{diag}\,[\eta_1, \eta_2, \cdots, \eta_N]$. The estimates acquired using Scheme-I or Scheme-S are used to design the decoder at the BS.

Remark: We observe that the channel estimates $\hat{\mathbf{g}}_{nk}[t]$ and $\hat{\mathbf{c}}_n[t]$ obtained using Scheme-I and Scheme-S, respectively, are functions of true channel vectors $\{\mathbf{g}_{n1}[t], \cdots \mathbf{g}_{nk}[t], \cdots \mathbf{g}_{nK}[t]\}$ of all users that share the same pilot signal. Thus, the effect of pilot contamination is well-captured.

3.1.2 Uplink Data Transmission

Let $q_{nk}[t+u] \sim \mathcal{CN}(0, 1)$ be the transmit symbol at time $(t+u)$ and $\gamma_{nk}[t+u] \in (0, \rho_{ul}]$ be the transmit power corresponding to UE_{nk}, for $k \in \mathcal{K}$ and $n \in \mathcal{N}$. All transmit symbols are assumed to be independent of each other. Let the data signal received on the uplink is given by

$$\mathbf{y}[t+u] = \mathbf{G}[t+u]\mathbf{x}[t+u] + \mathbf{w}[t+u]. \tag{3.17}$$

Here, $\mathbf{w}[t+u] \in \mathbb{C}^{M\times 1}$ is the noise vector with entries which are i.i.d. and $\mathcal{CN}(0, \sigma^2)$ distributed. Moreover, $\mathbf{x}[t+u] \in \mathbb{C}^{L\times 1}$ denotes the composite transmit symbol vector given by

$$\mathbf{x}[t+u] = \left[\mathbf{q}_1^T[t+u]\mathbf{D}_{\gamma_1}^{\frac{1}{2}}[t+u] \cdots \mathbf{q}_K^T[t+u]\mathbf{D}_{\gamma_K}^{\frac{1}{2}}[t+u]\right]^T, \tag{3.18}$$

where $\mathbf{q}_i[t+u] = \left[q_{1i}[t+u]\ q_{2i}[t+u]\ \cdots\ q_{Ni}[t+u]\right]^T$ comprises of the symbols transmitted by the i^{th} user in all N groups, and $\mathbf{D}_{\gamma_i}[t+u] = \text{diag}\left[\gamma_{1i}[t+u], \cdots, \gamma_{Ni}[t+u]\right]$, for $i \in \mathcal{K}$.

3.2 Spectral Efficiency Analysis

In this section, we derive novel closed-form expressions for the lower bound on the achievable SE of each decoding scheme considered. We use the notation $\bar{p} - \bar{q}$ to denote the decoding scheme, which uses the decoder $\bar{p}$ designed using the estimates acquired based on Scheme-$\bar{q}$, where $\bar{p} \in \{\text{ZF, MR}\}$ and $\bar{q} \in \{\text{I, S}\}$. For example, the notation ZF-I denotes the ZF decoder designed using the estimates acquired based on Scheme-I.

3.2.1 Achievable SE with Scheme ZF-I

With decoding Scheme ZF-I, the decoder $\left(\mathbf{A}^{\text{ZF}}[t+u]\right)' \in \mathbb{C}^{N\times M}$ for scheme ZF-I takes the form

$$\begin{aligned}\left(\mathbf{A}^{\text{ZF}}[t+u]\right)' &= \left[(\mathbf{Z}[t])'\,\mathbf{Z}[t]\right]^{-1}(\mathbf{Z}[t])' \\ &= \left[\mathbf{a}_1[t+u] \quad \cdots \quad \mathbf{a}_N[t+u]\right]', \end{aligned} \tag{3.19}$$

where $\mathbf{Z}[t] = [\mathbf{z}_1[t]\ \mathbf{z}_2[t]\cdots\mathbf{z}_N[t]]$ and $\mathbf{a}_n[t]$ is the n^{th} column of $\mathbf{A}^{\text{ZF}}[t+u]$. The BS processes the signal $\mathbf{y}[t+u]$ using the decoder $\left(\mathbf{A}^{\text{ZF}}[t+u]\right)'$. The n^{th} column of $\mathbf{A}^{\text{ZF}}[t+u]$ is used as a decoding vector to decode the signal corresponding to UE_{nk} by the BS. Therefore, the received signal given in (3.17) is pre-multiplied by $\mathbf{a}_n'[t+u]$ to give

$$\bar{y}_{nk}^{\text{ZF-I}}[t+u] = \sum_{n'=1}^{N}\sum_{k'=1}^{K}\sqrt{\gamma_{n'k'}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] + \mathbf{a}_n'[t+u]\mathbf{w}[t+u], \tag{3.20}$$

which can be expanded as

$$\begin{aligned}\bar{y}_{nk}^{\text{ZF-I}}[t+u] = {} & \sqrt{\gamma_{nk}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{nk}[t+u]q_{nk}[t+u] \\ & + \underbrace{\sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{nk'}[t+u]q_{nk'}[t+u]}_{\text{intra-group interference to be removed by SIC}} \\ & + \sum_{k''=k+1}^{K}\sqrt{\gamma_{nk''}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u] \\ & + \sum_{n'\neq n;n'=1}^{N}\sum_{k'=1}^{K}\sqrt{\gamma_{n'k'}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] \\ & + \mathbf{a}_n'[t+u]\mathbf{w}[t+u]. \end{aligned} \tag{3.21}$$

Prior to decoding the signal of UE_{nk}, the BS successively cancels interference from all other users which are relatively stronger than UE_{nk}. We consider that, for $n \in \mathcal{N}$, $\beta_{n1} \geq \cdots \geq \beta_{nK}$. Thus, prior to decoding the signal corresponding to UE_{nk}, the BS successively cancels the interference from $UE_{n1}, \cdots, UE_{n(k-1)}$. In (3.21), the first term denotes the desired signal, the second term denotes the intra-group interference (from relatively stronger users) to be cancelled

using SIC, the third term denotes the intra-group interference from relatively weaker users, the fourth term denotes the inter-group interference and the fifth term denotes noise.

The BS has outdated estimated CSI $\{\alpha(u)\widehat{\mathbf{g}}_{nk'}[t]\}$. Therefore, post-SIC, (3.21) can be written as

$$\begin{aligned}\overline{y}_{nk}^{\text{ZF-I-SIC}}[t+u] = & \sqrt{\gamma_{nk}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{nk}[t+u]q_{nk}[t+u] \\ & + \underbrace{\sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\,\mathbf{a}_n'[t+u]\left(\mathbf{g}_{nk'}[t+u]-\alpha(u)\widehat{\mathbf{g}}_{nk'}[t]\right)q_{nk'}[t+u]}_{\text{residual term due to imperfect SIC}} \\ & + \sum_{k''=k+1}^{K}\sqrt{\gamma_{nk''}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u] \\ & + \sum_{n'\neq n;n'=1}^{N}\sum_{k'=1}^{K}\sqrt{\gamma_{n'k'}[t+u]}\,\mathbf{a}_n'[t+u]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] \\ & + \mathbf{a}_n'[t+u]\mathbf{w}[t+u]. \quad (3.22)\end{aligned}$$

Substituting for $\{\mathbf{g}_{nk}[t+u]\}$, $\hat{\mathbf{g}}_{nk}[t] = \sqrt{\eta_{nk}}\,\mathbf{z}_n[t]$ and using the relationship $\mathbf{a}_n'[t+u]\mathbf{z}_{n'}[t] = \delta_{nn'}$, where $\delta_{nn'} = 1$ if $n = n'$, else $\delta_{nn'} = 0$, we get

$$\begin{aligned}\overline{y}_{nk}^{\text{ZF-I-SIC}}[t+u] = & \sqrt{\eta_{nk}\gamma_{nk}[t+u]}\alpha(u)q_{nk}[t+u] \\ & + \sum_{k''=k+1}^{K}\sqrt{\eta_{nk''}\gamma_{nk''}[t+u]}\alpha(u)q_{nk''}[t+u] + \mathbf{a}_n'[t+u]\overline{\mathbf{n}}^{\text{I}}[t+u], \quad (3.23)\end{aligned}$$

where $\overline{\mathbf{n}}^{\text{I}}[t+u]$ denotes the effective noise and is given by

$$\overline{\mathbf{n}}^{\text{I}}[t+u] = \sum_{n'=1}^{N}\sum_{k'=1}^{K}\sqrt{\gamma_{n'k'}[t+u]}\left(-\alpha(u)\widetilde{\mathbf{g}}_{n'k'}^{\text{I}}[t]-\mathbf{e}_{n'k'}[t+u]\right)q_{n'k'}[t+u]+\mathbf{w}[t+u]. \quad (3.24)$$

Note that the inter-group interference term from (3.22) goes to zero due to ZF decoding. Further, the term $\mathbf{a}_n'[t+u]\overline{\mathbf{n}}^{\text{I}}[t+u]$ can be represented in the matrix form as

$$\mathbf{a}_n'[t+u]\overline{\mathbf{n}}^{\text{I}}[t+u] = \left[\left(\mathbf{A}^{\text{ZF}}[t+u]\right)'\overline{\mathbf{n}}^{\text{I}}[t+u]\right]_n. \quad (3.25)$$

Therefore,

$$\overline{y}_{nk}^{\text{ZF-I-SIC}}[t+u] = \sqrt{\eta_{nk}\gamma_{nk}[t+u]}\alpha(u)q_{nk}[t+u]$$

$$+\sum_{k''=k+1}^{K}\sqrt{\eta_{nk''}\gamma_{nk''}[t+u]}\alpha(u)q_{nk''}[t+u]+\left[\left(\mathbf{A}^{\mathrm{ZF}}[t+u]\right)'\overline{\mathbf{n}}^{\mathrm{I}}[t+u]\right]_n. \quad (3.26)$$

The BS decodes the signal corresponding to UE_{nk} using (3.26).

Theorem 3.1. *In a massive MIMO-NOMA system that employs Scheme ZF-I, the SE of UE_{nk} under channel aging and imperfect SIC is lower bounded by*

$$R_{nk}^{ZF\text{-}I}[t+u]=\log_2\left(1+SINR_{nk}^{ZF\text{-}I}[t+u]\right), \quad (3.27)$$

where $SINR_{nk}^{ZF\text{-}I}[t+u]$ is the corresponding SINR, given by

$$SINR_{nk}^{ZF\text{-}I}[t+u]=\frac{\mathcal{N}_{nk}^{ZF\text{-}I}[t+u]}{\mathcal{D}_{nk}^{ZF\text{-}I}[t+u]}, \quad (3.28)$$

where

$$\mathcal{N}_{nk}^{ZF\text{-}I}[t+u]=(M-N)\,\alpha(u)^2\,\eta_{nk}\,\gamma_{nk}[t+u], \quad (3.29)$$

$$\mathcal{D}_{nk}^{ZF\text{-}I}[t+u]=\sigma^2+\lambda^{ZF\text{-}I}[t+u]+\sum_{k'=k+1}^{K}(M-N)\,\alpha(u)^2\,\eta_{nk'}\gamma_{nk'}[t+u], \quad (3.30)$$

$$\lambda^{ZF\text{-}I}[t+u]=\sum_{m=1}^{N}\sum_{i=1}^{K}\left\{\left(\beta_{mi}-\alpha(u)^2\,\eta_{mi}\right)\gamma_{mi}[t+u]\right\}. \quad (3.31)$$

Proof. The detailed proof is provided in Appendix B.1. □

Remark: The term $(M-N)$ in (3.29) accounts for the DoF exhausted in IGI cancellation. The term $\lambda^{\mathrm{ZF\text{-}I}}[t+u]$ in (3.30) accounts for the variance of the channel estimation error with Scheme-I under channel aging, and η_{mi} therein captures variance of the estimate of channel vector corresponding to UE_{mi} under pilot contamination. The third term in (3.30) is essentially the variance of the residual intra-group interference under channel aging. Therefore, the SINR expression in (3.28) captures the effects of (i) the DoF exhausted in IGI cancellation, (ii) the intra-group pilot contamination with Scheme-I, (iii) the intra-group interference under channel aging and (iv) channel estimation errors with Scheme-I under channel aging.

3.2.2 Achievable SE with Scheme ZF-S

With decoding scheme ZF-S, the decoder $\left(\mathbf{B}^{\mathrm{ZF}}[t+u]\right)'\in\mathbb{C}^{N\times M}$ for scheme ZF-S takes the form

$$\left(\mathbf{B}^{\mathrm{ZF}}[t+u]\right)'=\left[(\mathbf{V}[t])'\,\mathbf{V}[t]\right]^{-1}(\mathbf{V}[t])'$$

$$= \left[\mathbf{b}_1[t+u] \quad \cdots \quad \mathbf{b}_N[t+u]\right]', \tag{3.32}$$

where $\mathbf{V}[t] = [\mathbf{v}_1[t]\ \mathbf{v}_2[t] \cdots \mathbf{v}_N[t]]$ and $\mathbf{b}_n[t]$ is the n^{th} column of $\mathbf{B}^{\text{ZF}}[t+u]$. The BS processes the signal $\mathbf{y}[t+u]$ using the decoder $\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'$. The n^{th} column of $\mathbf{B}^{\text{ZF}}[t+u]$ is used as a decoding vector to decode the signal corresponding to UE_{nk} by the BS. Therefore, the received signal given in (3.17) is pre-multiplied by $\mathbf{b}_n'[t]$ to give

$$\bar{y}_{nk}^{\text{ZF-S}}[t+u] = \sum_{n'=1}^{N}\sum_{k'=1}^{K} \sqrt{\gamma_{n'k'}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] + \mathbf{b}_n'[t+u]\mathbf{w}[t+u], \tag{3.33}$$

which can be expanded as

$$\begin{aligned}\bar{y}_{nk}^{\text{ZF-S}}[t+u] = {} & \sqrt{\gamma_{nk}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{nk}[t+u]q_{nk}[t+u] \\ & + \underbrace{\sum_{k'=1}^{k-1} \sqrt{\gamma_{nk'}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{nk'}[t+u]q_{nk'}[t+u]}_{\text{intra-group interference to be removed by SIC}} \\ & + \sum_{k''=k+1}^{K} \sqrt{\gamma_{nk''}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u] \\ & + \sum_{n'\neq n;n'=1}^{N}\sum_{k'=1}^{K} \sqrt{\gamma_{n'k'}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] \\ & + \mathbf{b}_n'[t+u]\mathbf{w}[t+u]. \end{aligned} \tag{3.34}$$

The BS has outdated estimated CSI $\{\alpha(u)\widehat{\mathbf{c}}_n[t]\}$. Therefore, post-SIC, (3.34) can be written as

$$\begin{aligned}\bar{y}_{nk}^{\text{ZF-S-SIC}}[t+u] = {} & \sqrt{\gamma_{nk}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{nk}[t+u]q_{nk}[t+u] \\ & + \underbrace{\sum_{k'=1}^{k-1} \sqrt{\gamma_{nk'}[t+u]}\, \mathbf{b}_n'[t+u]\left(\mathbf{g}_{nk'}[t+u] - \alpha(u)\widehat{\mathbf{c}}_n[t]\right)q_{nk'}[t+u]}_{\text{residual term due to imperfect SIC}} \\ & + \sum_{k''=k+1}^{K} \sqrt{\gamma_{nk''}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u] \\ & + \sum_{n'\neq n;n'=1}^{N}\sum_{k'=1}^{K} \sqrt{\gamma_{n'k'}[t+u]}\, \mathbf{b}_n'[t+u]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] \\ & + \mathbf{b}_n'[t+u]\mathbf{w}[t+u]. \end{aligned} \tag{3.35}$$

Substituting for $\{\mathbf{g}_{nk}[t+u]\}$, $\hat{\mathbf{c}}_n[t] = \sqrt{\eta_n}\,\mathbf{v}_n[t]$ in (3.35) and using the relationship $\mathbf{b}'_n[t+u]\mathbf{v}_{n'}[t] = \delta_{nn'}$, where $\delta_{nn'} = 1$ if $n = n'$, else $\delta_{nn'} = 0$, we get

$$\bar{y}_{nk}^{\text{ZF-S-SIC}}[t+u] = \sqrt{\eta_n \gamma_{nk}[t+u]}\alpha(u) q_{nk}[t+u] + \sum_{k''=k+1}^{K} \sqrt{\eta_n \gamma_{nk''}[t+u]}\alpha(u) q_{nk''}[t+u] + \mathbf{b}'_n[t+u]\bar{\mathbf{n}}^{\text{S}}[t+u], \quad (3.36)$$

where $\bar{\mathbf{n}}^{\text{S}}[t+u]$ denotes the effective noise and is given by

$$\bar{\mathbf{n}}^{\text{S}}[t+u] = \sum_{n'=1}^{N}\sum_{k'=1}^{K} \sqrt{\gamma_{n'k'}[t+u]}\left(-\alpha(u)\widetilde{\mathbf{g}}^{\text{S}}_{n'k'}[t] - \mathbf{e}_{n'k'}[t+u]\right) q_{n'k'}[t+u] + \mathbf{w}[t+u]. \quad (3.37)$$

Note that the inter-group interference term from (3.35) reduces to zero due to ZF decoding. Further, the term $\mathbf{b}'_n[t+u]\bar{\mathbf{n}}^{\text{S}}[t+u]$ can be represented in the matrix form as

$$\mathbf{b}'_n[t+u]\bar{\mathbf{n}}^{\text{S}}[t+u] = \left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)' \bar{\mathbf{n}}^{\text{S}}[t+u]\right]_n. \quad (3.38)$$

Therefore,

$$\bar{y}_{nk}^{\text{ZF-S-SIC}}[t+u] = \sqrt{\eta_n \gamma_{nk}[t+u]}\alpha(u) q_{nk}[t+u] + \sum_{k''=k+1}^{K} \sqrt{\eta_n \gamma_{nk''}[t+u]}\alpha(u) q_{nk''}[t+u] + \left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)' \bar{\mathbf{n}}^{\text{S}}[t+u]\right]_n. \quad (3.39)$$

The BS decodes the signal corresponding to UE_{nk} using (3.39).

Theorem 3.2. *In a massive MIMO-NOMA system that employs scheme ZF-S, given the side information $\Omega[t+u] = \mathbf{V}[t]$, the SE of UE_{nk} under channel aging and imperfect SIC is lower bounded by*

$$\widetilde{R}_{nk}^{ZF\text{-}S}[t+u] = \log_2\left(1 + \widetilde{SINR}_{nk}^{ZF\text{-}S}[t+u]\right), \quad (3.40)$$

where $\widetilde{SINR}_{nk}^{ZF\text{-}S}[t+u]$ is the corresponding SINR, given by

$$\widetilde{SINR}_{nk}^{ZF\text{-}S}[t+u] = \frac{\mathcal{N}_{nk}^{ZF\text{-}S}[t+u]}{\mathcal{D}_{nk}^{ZF\text{-}S}[t+u]}, \quad (3.41)$$

where

$$\mathcal{N}_{nk}^{ZF\text{-}S}[t+u] = (M-N)\,\alpha(u)^2\,\eta_n\,\gamma_{nk}[t+u], \quad (3.42)$$

$$\mathcal{D}_{nk}^{ZF\text{-}S}[t+u] = \sigma^2 + \lambda^{ZF\text{-}S}[t+u] + \omega^{ZF\text{-}S}[t+u] + \sum_{k'=k+1}^{K} (M-N)\,\alpha(u)^2 \eta_n \gamma_{nk'}[t+u], \tag{3.43}$$

$$\lambda^{ZF\text{-}S}[t+u] = \alpha(u)^2 \sum_{i=1}^{K} \left[\left(\sum_{j=1}^{N} \gamma_{ji}[t+u] \iota_{ji} \right) - \frac{1}{N} \left(\sum_{j=1}^{N} \gamma_{ji}^{\frac{1}{2}}[t+u] \frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2 \right], \tag{3.44}$$

$$\omega^{ZF\text{-}S}[t+u] = \sum_{i=1}^{K} \sum_{j=1}^{N} \gamma_{ji}[t+u] \left(1 - \alpha(u)^2\right) \beta_{ji}, \tag{3.45}$$

$$\epsilon_{ji} = \mu_j^2 \sigma^2 + \left(\sqrt{\tau_p \rho_{ul}} \mu_j - 1 \right) \left(\sqrt{\tau_p \rho_{ul}} \mu_j \right) \beta_{ji} + \left(\sqrt{\tau_p \rho_{ul}} \mu_j \right)^2 \sum_{l=1, l \neq i}^{K} \beta_{jl}, \tag{3.46}$$

and η_j, ι_{ji} are given in (3.14).

Proof. The detailed proof is provided in Appendix B.2. □

Remark: The term $(M-N)$ in (3.42) accounts for the DoF exhausted in IGI cancellation. The term $\lambda^{\text{ZF-S}}[t+u]$ in (3.43) accounts for the variance of the channel estimation error with Scheme-S under channel aging, and η_j therein captures variance of the estimate of channel vector corresponding to users in j^{th} group under pilot contamination. The third term in (3.43) is essentially the variance of the residual intra-group interference under channel aging. Therefore, the SINR expression in (3.41) captures the effects of (i) the DoF exhausted in IGI cancellation, (ii) the intra-group pilot contamination with Scheme-S, (iii) the intra-group interference under channel aging and (iv) channel estimation errors with Scheme-S under channel aging.

In Section 3.4 we formulate the optimization problems for power control, namely, max-min power control and proportional fairness power control. For converting the optimization problems to convex programs, we now provide a lower bound on $\widetilde{\text{R}}_{nk}^{\text{ZF-S}}[t+u]$ given in (3.40). To do so we compute an upper bound on $\lambda^{\text{ZF-S}}[t+u]$ given in Theorem 3.2 as follows:

$$\left(\sum_{j=1}^{N} \gamma_{ji}^{\frac{1}{2}}[t+u] \frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2 \geq \sum_{j=1}^{N} \gamma_{ji}[t+u] \left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2. \tag{3.47}$$

Therefore,

$$\begin{aligned} \lambda^{\text{ZF-S}}[t+u] &= \alpha(u)^2 \sum_{i=1}^{K} \left[\left(\sum_{j=1}^{N} \gamma_{ji}[t+u] \iota_{ji} \right) - \frac{1}{N} \left(\sum_{j=1}^{N} \gamma_{ji}^{\frac{1}{2}}[t+u] \frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2 \right] \\ &\leq \alpha(u)^2 \sum_{i=1}^{K} \left[\left(\sum_{j=1}^{N} \gamma_{ji}[t+u] \iota_{ji} \right) - \frac{1}{N} \sum_{j=1}^{N} \gamma_{ji}[t+u] \left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2 \right] \end{aligned}$$

$$= \alpha(u)^2 \sum_{i=1}^{K} \sum_{j=1}^{N} \gamma_{ji}[t+u] \left(\iota_{ji} - \frac{1}{N} \left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2 \right)$$

$$\triangleq \lambda^{\text{ZF-S-UB}}[t+u], \tag{3.48}$$

where $\lambda^{\text{ZF-S-UB}}[t+u]$ is an upper bound on $\lambda^{\text{ZF-S}}[t+u]$.

Theorem 3.3. *Using the inequality in (3.48), $\widetilde{SINR}_{nk}^{ZF\text{-}S}[t+u]$ in (3.41) is lower bounded by*

$$SINR_{nk}^{ZF\text{-}S}[t+u] = \frac{\mathcal{N}_{nk}^{ZF\text{-}S}[t+u]}{\mathcal{D}_{nk}^{ZF\text{-}S\text{-}UB}[t+u]}, \tag{3.49}$$

where

$$\mathcal{D}_{nk}^{ZF\text{-}S\text{-}UB}[t+u] = \sigma^2 + \lambda^{ZF\text{-}S\text{-}UB}[t+u] + \omega^{ZF\text{-}S}[t+u] + \sum_{k'=k+1}^{K} (M-N)\,\alpha(u)^2 \eta_n \gamma_{nk'}[t+u], \tag{3.50}$$

$$\lambda^{ZF\text{-}S\text{-}UB}[t+u] = \alpha(u)^2 \sum_{i=1}^{K} \sum_{j=1}^{N} \gamma_{ji}[t+u] \left(\iota_{ji} - \frac{1}{N} \left(\frac{\epsilon_{ji}}{\sqrt{\eta_j}} \right)^2 \right), \tag{3.51}$$

with $\omega^{ZF\text{-}S}[t+u]$, ϵ_{ji} as given in Theorem 3.2 and η_j, ι_{ji} as given in (3.14). Hence, $\widetilde{R}_{nk}^{ZF\text{-}S}[t+u]$ in (3.40) is lower bounded by

$$R_{nk}^{ZF\text{-}S}[t+u] = \log_2 \left(1 + SINR_{nk}^{ZF\text{-}S}[t+u] \right). \tag{3.52}$$

3.2.3 Achievable SE with Scheme MR-I

With decoding scheme MR-I, the decoder $\left(\mathbf{A}^{\text{MR}}[t+u]\right)' \in \mathbb{C}^{NK \times M}$ takes the form

$$\left(\mathbf{A}^{\text{MR}}[t+u]\right)' = \left(\widehat{\mathbf{G}}[t]\right)'. \tag{3.53}$$

The BS processes the signal $\mathbf{y}[t+u]$ using the decoder $\left(\mathbf{A}^{\text{MR}}[t+u]\right)'$. Specifically, the signal corresponding to UE_{nk} is decoded using $\widehat{\mathbf{g}}_{nk}[t]$. Thus, post-processing, (3.17) can be rewritten as

$$y_{nk}^{\text{MR-I}}[t+u] = \sum_{n'=1}^{N} \sum_{k'=1}^{K} \sqrt{\gamma_{n'k'}[t+u]}\, \widehat{\mathbf{g}}_{nk}'[t] \mathbf{g}_{n'k'}[t+u] q_{n'k'}[t+u] + \widehat{\mathbf{g}}_{nk}'[t] \mathbf{w}[t+u]. \tag{3.54}$$

Prior to decoding the signal corresponding to UE_{nk}, the BS does SIC according to the NOMA decoding protocol as explained in Section 3.2.1. Therefore, employing "use and forget CSI" technique, the signal $y_{nk}^{\text{MR-I}}[t+u]$ after SIC can be rewritten as [14]

$$
\begin{aligned}
y_{nk}^{\text{MR-I-SIC}}[t+u] = & \sqrt{\gamma_{nk}[t+u]}\,\mathbb{E}\left\{\widehat{\mathbf{g}}_{nk}'[t]\mathbf{g}_{nk}[t+u]\right\} q_{nk}[t+u] \\
& + \sqrt{\gamma_{nk}[t+u]}\,\left(\widehat{\mathbf{g}}_{nk}'[t]\mathbf{g}_{nk}[t+u] - \mathbb{E}\left\{\widehat{\mathbf{g}}_{nk}'[t]\mathbf{g}_{nk}[t+u]\right\}\right) q_{nk}[t+u] \\
& + \underbrace{\sum_{k'=1}^{k-1} \sqrt{\gamma_{nk'}[t+u]}\widehat{\mathbf{g}}_{nk}'[t]\left(\mathbf{g}_{nk'}[t+u] - \alpha(u)\widehat{\mathbf{g}}_{nk'}[t]\right) q_{nk'}[t+u]}_{\text{residual term due to imperfect SIC}} \\
& + \sum_{k''=k+1}^{K} \sqrt{\gamma_{nk''}[t+u]}\widehat{\mathbf{g}}_{nk}'[t]\mathbf{g}_{nk''}[t+u] q_{nk''}[t+u] \\
& + \sum_{n'\neq n, n'=1}^{N} \sum_{k'''=1}^{K} \sqrt{\gamma_{n'k'''}[t+u]}\widehat{\mathbf{g}}_{nk}'[t]\mathbf{g}_{n'k'''}[t+u] q_{n'k'''}[t+u] \\
& + \widehat{\mathbf{g}}_{nk}'[t]\mathbf{w}[t+u]. \qquad (3.55)
\end{aligned}
$$

Note that due to availability of imperfect and aged CSI, the SIC is also imperfect. This shows up in the third term of (3.55). The BS decodes the signal corresponding to UE_{nk} using $y_{nk}^{\text{MR-I-SIC}}[t+u]$.

Theorem 3.4. *In a massive MIMO-NOMA system that employs scheme MR-I, the SE of UE_{nk} under channel aging is lower bounded by*

$$
\begin{aligned}
R_{nk}^{MR\text{-}I}[t+u] &= \log_2\left(1 + SINR_{nk}^{MR\text{-}I}[t+u]\right) \\
&= \log_2\left(1 + \frac{\mathcal{V}_1}{\sum_{i=2}^{6}\mathcal{V}_i}\right), \qquad (3.56)
\end{aligned}
$$

where

$$\mathcal{V}_1 = \gamma_{nk}[t+u]\alpha^2(u)M^2\eta_{nk}^2, \qquad (3.57)$$

$$\mathcal{V}_2 = \gamma_{nk}[t+u]\beta_{nk}M\eta_{nk}, \qquad (3.58)$$

$$\mathcal{V}_3 = \sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\left(\beta_{nk'} - \alpha(u)^2\eta_{nk'}\right)M\eta_{nk}, \qquad (3.59)$$

$$\mathcal{V}_4 = \sum_{k''=k+1}^{K}\gamma_{nk''}[t+u]\left(\alpha(u)^2 M\eta_{nk''} + \beta_{nk''}\right)M\eta_{nk}, \qquad (3.60)$$

$$\mathcal{V}_5 = \sum_{n'\neq n, n'=1}^{N}\sum_{k'''=1}^{K}\gamma_{n'k'''}[t+u]\beta_{n'k'''}M\eta_{nk}, \qquad (3.61)$$

$$\mathcal{V}_6 = \sigma^2 M\eta_{nk}. \qquad (3.62)$$

Proof. The detailed proof is provided in Appendix B.3. □

Remark: The effects of imperfect SIC and inter-group interference *under channel aging*

are captured in $\mathcal{V}_3$ and $\mathcal{V}_5$, respectively. Moreover, $\mathcal{V}_4$ captures the intra-group interference *under channel aging*.

3.2.4 Achievable SE with Scheme MR-S

With decoding scheme MR-S, the decoder $\left(\mathbf{B}^{\text{MR}}[t+u]\right)' \in \mathbb{C}^{N\times M}$ takes the form

$$\left(\mathbf{B}^{\text{MR}}[t+u]\right)' = \left(\widehat{\mathbf{C}}[t]\right)'. \tag{3.63}$$

The BS uses $\left(\mathbf{B}^{\text{MR}}[t+u]\right)'$ to process $\mathbf{y}[t+u]$ given in (3.17). Specifically, the signal corresponding to UE_{nk} is decoded using $\widehat{\mathbf{c}}_n[t]$. Thus, post-processing, (3.17) can be rewritten as

$$y_{nk}^{\text{MR-S}}[t+u] = \sum_{n'=1}^{N}\sum_{k'=1}^{K}\sqrt{\gamma_{n'k'}[t+u]}\,\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{n'k'}[t+u]q_{n'k'}[t+u] + \widehat{\mathbf{c}}_n'[t]\mathbf{w}[t+u]. \tag{3.64}$$

The BS employs SIC prior to decoding the signal corresponding to UE_{nk} as explained in Section 3.2.1. Thus, employing "use and forget CSI" technique, the signal $y_{nk}^{\text{MR-S}}[t+u]$ after SIC can be rewritten as [14]

$$\begin{aligned}
y_{nk}^{\text{MR-S-SIC}}[t+u] = {} & \sqrt{\gamma_{nk}[t+u]}\,\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\}q_{nk}[t+u] \\
& + \sqrt{\gamma_{nk}[t+u]}\left(\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u] - \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\}\right)q_{nk}[t+u] \\
& \underbrace{+\sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\widehat{\mathbf{c}}_n'[t]\left(\mathbf{g}_{nk'}[t+u] - \alpha(u)\widehat{\mathbf{c}}_n[t]\right)q_{nk'}[t+u]}_{\text{residual term due to imperfect SIC}} \\
& + \sum_{k''=k+1}^{K}\sqrt{\gamma_{nk''}[t+u]}\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u] \\
& + \sum_{n'\neq n, n'=1}^{N}\sum_{k'''=1}^{K}\sqrt{\gamma_{n'k'''}[t+u]}\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{n'k'''}[t+u]q_{n'k'''}[t+u] \\
& + \widehat{\mathbf{c}}_n'[t]\mathbf{w}[t+u].
\end{aligned} \tag{3.65}$$

Note that due to availability of outdated and estimated CSI, the SIC is also imperfect. This is reflected in the third term of (3.65). The BS decodes the signal corresponding to UE_{nk} using $y_{nk}^{\text{MR-S-SIC}}[t+u]$ given in (3.65).

Theorem 3.5. *In a massive MIMO-NOMA system that employs scheme MR-S, the SE of* UE_{nk}

under channel aging is lower bounded by

$$R_{nk}^{MR\text{-}S}[t+u] = \log_2\left(1 + SINR_{nk}^{MR\text{-}S}[t+u]\right) = \log_2\left(1 + \frac{\mathcal{X}_1}{\sum_{i=2}^{6}\mathcal{X}_i}\right), \tag{3.66}$$

where

$$\mathcal{X}_1 = \gamma_{nk}[t+u]\alpha^2(u)M^2\left|\eta_n - \bar{\epsilon}_{nk}\right|^2, \tag{3.67}$$

$$\mathcal{X}_2 = \gamma_{nk}[t+u]\left(\mathcal{A}_{nk} - \left|\mathcal{B}_{nk}\right|^2\right), \tag{3.68}$$

$$\mathcal{X}_3 = \sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\left[\mathcal{A}_{nk'}^3 + \left(1-\alpha(u)^2\right)\eta_n M\beta_{nk'}\right], \tag{3.69}$$

$$\mathcal{X}_4 = \sum_{k''=k+1}^{K}\gamma_{nk''}[t+u]\mathcal{A}_{nk''}, \tag{3.70}$$

$$\mathcal{X}_5 = \sum_{n'\neq n,n'=1}^{N}\sum_{k'''=1}^{K}\gamma_{n'k'''}[t+u]\beta_{n'k'''}M\eta_n, \tag{3.71}$$

$$\mathcal{X}_6 = \sigma^2 M\eta_n, \tag{3.72}$$

where, for $\hat{k} \in \{1, \cdots, K\}$*:*

$$\mathcal{A}_{n\hat{k}} = \mathbb{E}\left\{\left|\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{n\hat{k}}[t+u]\right|^2\right\} = \mathcal{A}_{n\hat{k}}^1 - \mathcal{A}_{n\hat{k}}^2 + \mathcal{A}_{n\hat{k}}^3 + \mathcal{A}_{n\hat{k}}^4, \tag{3.73}$$

$$\mathcal{B}_{n\hat{k}} = \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{n\hat{k}}[t+u]\right\} = \alpha(u)M\left(\eta_n - \bar{\epsilon}_{n\hat{k}}\right), \tag{3.74}$$

$$\bar{\epsilon}_{n\hat{k}} = \bar{a}\bar{b}\beta_{n\hat{k}} + \bar{b}^2\sum_{i\neq\hat{k}}^{K}\beta_{ni} + \bar{c}^2\sigma^2, \tag{3.75}$$

$$\mathcal{A}_{n\hat{k}}^1 = \alpha(u)^2 M\left(M+1\right)\eta_n^2,\ \mathcal{A}_{n\hat{k}}^4 = \left(1-\alpha(u)^2\right)\eta_n M\beta_{n\hat{k}}, \tag{3.76}$$

$$\begin{aligned}\mathcal{A}_{n\hat{k}}^2 = 2\alpha(u)^2\Re\Bigg\{&\left[\bar{a}\bar{b}^3\beta_{n\hat{k}}^2 + \bar{b}^4\sum_{j\neq\hat{k}}^{K}\beta_{nj}^2 + \bar{c}^4\sigma^4\right]M\left(M+1\right)\\ &+\left[\bar{a}\bar{b}^3\sum_{j\neq\hat{k}}^{K}\beta_{nj}\beta_{n\hat{k}} + \bar{b}^2\bar{c}^2\sum_{i=1}^{K}\beta_{ni}\sigma^2 + \bar{a}\bar{b}\bar{c}^2\beta_{n\hat{k}}\sigma^2\right.\\ &\left.+\bar{b}^2\bar{c}^2\sum_{j\neq\hat{k}}^{K}\beta_{nj}\sigma^2 + \bar{b}^4\sum_{j\neq\hat{k}}^{K}\beta_{nj}\beta_{n\hat{k}} + \bar{b}^4\sum_{i\neq j,\hat{k}}^{K}\sum_{j\neq i,\hat{k}}^{K}\beta_{ni}\beta_{nj}\right]M^2\\ &+\left[\bar{a}\bar{b}\bar{c}^2\beta_{n\hat{k}}\sigma^2 + \bar{b}^2\bar{c}^2\sum_{j\neq\hat{k}}^{K}\beta_{nj}\sigma^2 + \bar{b}^2\bar{c}^2\sum_{i=1}^{K}\beta_{ni}\sigma^2\right]M\Bigg\},\end{aligned} \tag{3.77}$$

$$\mathcal{A}^3_{n\hat{k}} = \alpha(u)^2 \Re\Bigg\{ \Bigg[\bar{a}^2\bar{b}^2\beta^2_{n\hat{k}} + \bar{b}^4 \sum_{j\neq\hat{k}}^{K} \beta^2_{nj} + \bar{c}^4\sigma^4 \Bigg] M(M+1) + \Bigg[2\bar{a}\bar{b}^3 \sum_{j\neq\hat{k}}^{K} \beta_{nj}\beta_{n\hat{k}} + 2\bar{a}\bar{b}\bar{c}^2\beta_{n\hat{k}}\sigma^2 + 2\bar{b}^2\bar{c}^2 \sum_{j\neq\hat{k}}^{K} \beta_{nj}\sigma^2 + \bar{b}^4 \sum_{i\neq j,\hat{k}}^{K} \sum_{j\neq i,\hat{k}}^{K} \beta_{ni}\beta_{nj} \Bigg] M^2 + \Bigg[\bar{a}^2\bar{b}^2 \sum_{j\neq\hat{k}}^{K} \beta_{nj}\beta_{n\hat{k}} + \bar{b}^4 \sum_{i\neq j,\hat{k}}^{K} \sum_{j\neq i,\hat{k}}^{K} \beta_{ni}\beta_{nj} + \bar{b}^2\bar{c}^2 \sum_{i=1}^{K} \beta_{ni}\sigma^2 + \bar{a}^2\bar{c}^2\beta_{nk}\sigma^2 + \bar{b}^2\bar{c}^2 \sum_{j\neq\hat{k}}^{K} \beta_{nj}\sigma^2 + \bar{b}^4 \sum_{j\neq\hat{k}}^{K} \beta_{nj}\beta_{n\hat{k}} \Bigg] M \Bigg\}, \quad (3.78)$$

where $\bar{a} = \sqrt{\tau_p\rho_{ul}}\mu_n - 1$, $\bar{b} = \sqrt{\tau_p\rho_{ul}}\mu_n$, $\bar{c} = \mu_n$.

Proof. The detailed proof is provided in Appendix B.4. □

Remark: The effects of imperfect SIC and inter-group interference *under channel aging* are captured in $\mathcal{X}_3$ and $\mathcal{X}_5$, respectively. Moreover, $\mathcal{X}_4$ captures the intra-group interference *under channel aging*.

3.3 Mitigation of Impact of Channel Aging using Channel Prediction

As we have seen so far, channel aging potentially affects quality of estimates. To deal with the time-varying channel, prediction is a widely used strategy. Thus, to negate the effect of channel aging, we now consider channel prediction using the p^{th} order WLP based on the observations $\left\{\overline{\mathbf{y}}_{pn}[t-\widetilde{p}]\right\}_{\widetilde{p}=0}^{p}$.

Specifically, using the same pilot assignment strategy as Scheme-I, we collect the channel observations for $(p+1)$ symbols. Using these $(p+1)$ observations, we design a WLP of order p, which minimizes the mean squared error. Using the designed WLP, we "predict" the channel vectors of all users individually, just like Scheme-I, where we "estimate" channel vectors of all users individually.[6] We use the predicted channel to design the ZF and MR decoders, and derive the expressions for the lower bound on the achievable SE.

[6] A WLP to "predict" the sum of channel vectors of users with each group, (just like Scheme-S, where we "estimate" the sum of channel vectors of users within each group) can be designed using a similar procedure.

3.3.1 Linear Prediction of Channel Vectors $\mathbf{g}_{nk}[t+1]$

We denote the p^{th} order optimal WLP by $\mathcal{P}_{nk} \triangleq \left[\mathcal{P}_{nk}^{0} \ \cdots \ \mathcal{P}_{nk}^{p}\right]$ (where $\mathcal{P}_{nk}^{\widetilde{p}} \in \mathbb{C}^{M\times M}$) that predicts $\mathbf{g}_{nk}[t+1]$ based on $\left\{\overline{\mathbf{y}}_{pn}[t-\widetilde{p}]\right\}_{\widetilde{p}=0}^{p}$, for $\widetilde{p} \in \{0,1,\cdots,p\}$. Also, let $\widetilde{\mathbf{y}}_{pn}[t] = \left[\overline{\mathbf{y}}'_{pn}[t] \ \cdots \ \overline{\mathbf{y}}'_{pn}[t-p]\right]'$, $\widetilde{\mathbf{y}}_{pn}[t] \in \mathbb{C}^{M(p+1)\times 1}$.

Lemma 3.1. *Based on the training observations $\left\{\overline{\mathbf{y}}_{pn}[t-\widetilde{p}]\right\}_{\widetilde{p}=0}^{p}$, the p^{th} order optimal WLP $\mathcal{P}_{nk}$ is given by [8]*

$$\mathcal{P}_{nk} = \sqrt{\tau_p \rho_{ul}}\alpha\beta_{nk}\left[\delta(\alpha,p)\otimes\mathbf{I}_M\right]\mathbf{T}_{nk}^{-1}(\alpha,p), \tag{3.79}$$

where $\delta(\alpha,p) = [1 \ \alpha \ \cdots \ \alpha^p]$ and $\mathbf{T}_{nk} = \left\{\tau_p\rho_{ul}\beta_n\left[\Delta(\alpha,p)\otimes\mathbf{I}_M\right] + \sigma^2\mathbf{I}_{M(p+1)}\right\}$ with $\beta_n = \sum_{i=1}^{K}\beta_{ni}$ and

$$\Delta(\alpha,p) = \begin{bmatrix} 1 & \alpha & \cdots & \alpha^p \\ \alpha & 1 & \cdots & \alpha^{p-1} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha^p & \alpha^{p-1} & \cdots & 1 \end{bmatrix}, \tag{3.80}$$

where $\alpha = \alpha(1) = J_0\left(2\pi f_D T_S\right)$

Proof. The detailed proof is given in Appendix B.5. □

Using the predictor $\mathcal{P}_{nk}$, the predicted channel vector $\mathbf{g}_{nk}[t+1]$ is given by

$$\bar{\mathbf{g}}_{nk}[t+1] = \mathcal{P}_{nk}\widetilde{\mathbf{y}}_{pn}[t] = \mathbf{g}_{nk}[t+1] + \check{\mathbf{e}}_{nk}[t+1], \tag{3.81}$$

where $\check{\mathbf{e}}_{nk}[t+1]$ is the channel prediction error. Thus, we can compute the mean square error in prediction as $\varpi = \text{tr}\left(\beta_{nk}\mathbf{I}_M - \mathbf{\Theta}_{nk}(\alpha,p)\right)$, where the covariance matrix of $\bar{\mathbf{g}}_{nk}[t+1]$ is given as follows [42]

$$\mathbf{\Theta}_{nk}(\alpha,p) = \alpha^2\beta_{nk}\tau_p\rho_{ul}\left[\delta(\alpha,p)\otimes\mathbf{I}_M\right]\mathbf{T}_{nk}^{-1}(\alpha,p)\left[\delta(\alpha,p)\otimes\mathbf{I}_M\right]'. \tag{3.82}$$

It can be shown that $\mathbf{\Theta}_{nk}(\alpha,p)$ is a scaled identity matrix and the variance of each element of $\bar{\mathbf{g}}_{nk}[t+1]$ is [42]

$$\theta_{nk}(\alpha,p) = \frac{1}{M}\text{tr}\left(\mathbf{\Theta}_{nk}(\alpha,p)\right). \tag{3.83}$$

Finally, note that $\bar{\mathbf{g}}_{nk}[t+1]$ and $\check{\mathbf{e}}_{nk}[t+1]$ are uncorrelated, and the covariance matrix of $\check{\mathbf{e}}_{nk}[t+1]$ is given by $\left(\beta_{nk} - \theta_{nk}(\alpha,p)\right)\mathbf{I}_M$.

The predicted channel $\bar{\mathbf{g}}_{nk}[t+1]$ can be normalized as $\bar{\mathbf{g}}_{nk}[t+1] = \sqrt{\theta_{nk}(\alpha,p)}\,\mathbf{h}_n[t+1]$. Here, the elements of $\mathbf{h}_n[t+1] \in \mathbb{C}^{M\times 1}$ are i.i.d. and $\mathcal{CN}(0,1)$ distributed. From (3.81), it is evident that, for all users within a group, the predicted channel vectors are perfectly correlated. Thus, $\mathbf{h}_n[t+1]$ represents the normalized predicted CSI for all K users in the n^{th} group. With channel

prediction, let $\overline{\mathbf{G}}[t+1] = \left[\overline{\mathbf{G}}_1[t+1]\cdots\overline{\mathbf{G}}_K[t+1]\right]$ and $\check{\mathbf{E}}[t+1] = \left[\check{\mathbf{E}}_1[t+1]\cdots\check{\mathbf{E}}_K[t+1]\right]$, respectively, denote the concatenated predicted CSI matrix and the corresponding concatenated channel prediction error matrix for all L users. Here, $\overline{\mathbf{G}}_i[t+1] = \left[\bar{\mathbf{g}}_{1i}[t+1] \quad \cdots \quad \bar{\mathbf{g}}_{Ni}[t+1]\right]$ and $\check{\mathbf{E}}_i[t+1] = \left[\check{\mathbf{e}}_{1i}[t+1] \quad \cdots \quad \check{\mathbf{e}}_{Ni}[t+1]\right]$, $\forall\, i \in \mathcal{K}$. As we know, $\bar{\mathbf{g}}_{ni}[t+1] = \sqrt{\theta_{ni}(\alpha,p)}\,\mathbf{h}_n[t+1]$, for $n \in \mathcal{N}$, therefore,

$$\overline{\mathbf{G}}_i[t+1] = \mathbf{H}[t+1]\mathbf{D}^{\frac{1}{2}}_{\theta_i(\alpha,p)}, \tag{3.84}$$

where $\mathbf{H}[t+1] = [\mathbf{h}_1[t+1]\cdots\mathbf{h}_N[t+1]]$ and $\mathbf{D}_{\theta_i(\alpha,p)} = \text{diag}\left[\theta_{1i}(\alpha,p),\cdots,\theta_{Ni}(\alpha,p)\right]$, for $i \in \mathcal{K}$. Note that $\mathbb{E}\left\{\mathbf{h}_i[t+1]\mathbf{h}_j[t+1]'\right\} = \mathbf{0}$ (an all zeros matrix) for $i \neq j$.

3.3.2 Achievable SE with ZF using Predicted CSI (ZF-I-P)

With predicted CSI, the ZF decoder $\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)' \in \mathbb{C}^{N\times M}$ takes the form

$$\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)' = \left[(\mathbf{H}[t+1])'\,\mathbf{H}[t+1]\right]^{-1}(\mathbf{H}[t+1])'. \tag{3.85}$$

Substituting for $\mathbf{G}[t+1]$, the BS processes the signal $\mathbf{y}[t+1]$ using the decoder $\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'$ to give

$$\overline{\mathbf{y}}^{\text{ZF-I-P}}[t+1] = \sum_{i=1}^{K}\mathbf{I}_N\mathbf{D}^{\frac{1}{2}}_{\theta_i(\alpha,p)}\mathbf{D}^{\frac{1}{2}}_{\gamma_i}[t+1]\mathbf{q}_i[t+1] + \left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\overline{\mathbf{n}}^{\text{P}}[t+1], \tag{3.86}$$

where $\overline{\mathbf{n}}^{\text{P}}[t+1] = \mathbf{w}[t+1] - \check{\mathbf{E}}[t+1]\mathbf{x}[t+u]$. Note that $\overline{\mathbf{y}}^{\text{ZF-I-P}}[t+1] \in \mathbb{C}^{N\times 1}$. Therefore, for $n \in \mathcal{N}$, $\left[\overline{\mathbf{y}}^{\text{ZF-I-P}}[t+1]\right]_n = \overline{y}^{\text{ZF-I-P}}_n[t+1]$ denotes the received signal corresponding to users in n^{th} group, and it is given by

$$\overline{y}^{\text{ZF-I-P}}_n[t+1] = \sum_{i=1}^{K}\sqrt{\theta_{ni}(\alpha,p)\gamma_{ni}[t+1]}\,q_{ni}[t+1] + \left[\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\overline{\mathbf{n}}^{\text{P}}[t+1]\right]_n. \tag{3.87}$$

Prior to decoding the signal corresponding to UE_{nk}, the BS does SIC according to the NOMA decoding protocol as explained in Section 3.2.1. Thus, after SIC at the BS, the processed signal for UE_{nk} is given by

$$\overline{y}^{\text{ZF-I-P-SIC}}_{nk}[t+1] = \sqrt{\theta_{nk}(\alpha,p)\gamma_{nk}[t+1]}\,q_{nk}[t+1]$$

$$+\sum_{i=k+1}^{K}\sqrt{\theta_{ni}(\alpha,p)\gamma_{ni}[t+1]}q_{ni}+\left[\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\overline{\mathbf{n}}^{\text{P}}[t+1]\right]_n. \quad (3.88)$$

The BS decodes the signal corresponding to UE_{nk} using (3.88).

Theorem 3.6. *In a massive MIMO-NOMA system that employs scheme ZF-I-P, the SE of UE_{nk} under channel aging and imperfect SIC is lower bounded by*

$$R_{nk}^{ZF\text{-}I\text{-}P}[t+1]=\log_2\left(1+SINR_{nk}^{ZF\text{-}I\text{-}P}[t+1]\right), \quad (3.89)$$

where $SINR_{nk}^{ZF\text{-}I\text{-}P}[t+1]$ is given by

$$SINR_{nk}^{ZF\text{-}I\text{-}P}[t+1]=\frac{\mathcal{N}_{nk}^{ZF\text{-}I\text{-}P}[t+1]}{\mathcal{D}_{nk}^{ZF\text{-}I\text{-}P}[t+1]}, \quad (3.90)$$

where

$$\mathcal{N}_{nk}^{ZF\text{-}I\text{-}P}[t+1]=(M-N)\;\theta_{nk}(\alpha,p)\,\gamma_{nk}[t+1], \quad (3.91)$$

$$\mathcal{D}_{nk}^{ZF\text{-}I\text{-}P}[t+1]=\sigma^2+\lambda^{ZF\text{-}I\text{-}P}[t+1]+\sum_{k'=k+1}^{K}(M-N)\;\theta_{nk'}(\alpha,p)\gamma_{nk'}[t+1], \quad (3.92)$$

$$\lambda^{ZF\text{-}I\text{-}P}[t+1]=\sum_{m=1}^{N}\sum_{i=1}^{K}\left\{(\beta_{mi}-\theta_{mi}(\alpha,p))\,\gamma_{mi}[t+1]\right\}. \quad (3.93)$$

Proof. The detailed proof is given in Appendix B.6. □

Remark: The first, second and third terms of (3.92) correspond to the noise variance, the variance of the channel estimation error and variance of the intra-group interference, respectively. Therefore, the SINR expression in (3.90) captures the effects of (i) the DoF exhausted in IGI cancellation, (ii) the intra-group pilot contamination, (iii) the intra-group interference and (iv) channel estimation errors, *under predicted channel*.

3.3.3 Achievable SE with MR using Predicted CSI (MR-I-P)

With predicted CSI, the MR decoder $\left(\mathbf{F}^{\text{MR-I-P}}[t+1]\right)'\in\mathbb{C}^{NK\times M}$ takes the form

$$\left(\mathbf{F}^{\text{MR-I-P}}[t+1]\right)'=\left(\overline{\mathbf{G}}[t+1]\right)'. \quad (3.94)$$

The BS processes the signal $\mathbf{y}[t+1]$ using the decoder $\left(\mathbf{F}^{\text{MR-I-P}}[t+1]\right)'$. Specifically, the signal corresponding to UE_{nk} is decoded using $\overline{\mathbf{g}}_{nk}[t+1]$. Thus, post-processing, (3.17) can be re-

written as

$$y_{nk}^{\text{MR-I-P}}[t+1] = \sum_{n'=1}^{N}\sum_{k'=1}^{K}\sqrt{\gamma_{n'k'}[t+1]}\,\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{n'k'}[t+1]q_{n'k'}[t+1] + \overline{\mathbf{g}}_{nk}'[t]\mathbf{w}[t+1]. \quad (3.95)$$

Prior to decoding the signal corresponding to UE_{nk}, the BS does SIC according to the NOMA decoding protocol as explained in Section 3.2.1. Therefore, employing "use and forget CSI" technique, the signal $y_{nk}^{\text{MR-I-P}}[t+1]$ after SIC can be rewritten as [14]

$$\begin{aligned} y_{nk}^{\text{MR-I-P-SIC}}[t+1] = {} & \sqrt{\gamma_{nk}[t+1]}\,\mathbb{E}\left\{\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{nk}[t+1]\right\}q_{nk}[t+1] \\ & + \sqrt{\gamma_{nk}[t+1]}\left(\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{nk}[t+1] - \mathbb{E}\left\{\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{nk}[t+1]\right\}\right)q_{nk}[t+1] \\ & + \underbrace{\sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\overline{\mathbf{g}}_{nk}'[t+1]\left(\mathbf{g}_{nk'}[t+1] - \overline{\mathbf{g}}_{nk'}[t]\right)q_{nk'}[t+1]}_{\text{residual term due to imperfect SIC}} \\ & + \sum_{k''=k+1}^{K}\sqrt{\gamma_{nk''}[t+1]}\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{nk''}[t+1]q_{nk''}[t+1] \\ & + \sum_{n'\neq n,n'=1}^{N}\sum_{k'''=1}^{K}\sqrt{\gamma_{n'k'''}[t+u]}\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{n'k'''}[t+1]q_{n'k'''}[t+1] \\ & + \overline{\mathbf{g}}_{nk}'[t+1]\mathbf{w}[t+1]. \quad (3.96) \end{aligned}$$

Note that due to availability of predicted CSI, the SIC is imperfect. This shows up in the third term of (3.96). The BS decodes the signal corresponding to UE_{nk} using $y_{nk}^{\text{MR-I-P-SIC}}[t+1]$.

Theorem 3.7. *In a massive MIMO-NOMA system that employs scheme MR-I-P, the SE of UE_{nk} under channel aging is lower bounded by*

$$\begin{aligned} R_{nk}^{MR\text{-}I\text{-}P}[t+u] &= \log_2\left(1 + SINR_{nk}^{MR\text{-}I\text{-}P}[t+1]\right) \\ &= \log_2\left(1 + \frac{\mathcal{W}_1}{\sum_{i=2}^{6}\mathcal{W}_i}\right), \end{aligned} \quad (3.97)$$

where

$$\mathcal{W}_1 = \gamma_{nk}[t+u]M^2\theta_{nk}^2, \quad (3.98)$$

$$\mathcal{W}_2 = \gamma_{nk}[t+u]\beta_{nk}M\theta_{nk}, \quad (3.99)$$

$$\mathcal{W}_3 = \sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\left(\beta_{nk'} - \theta_{nk'}\right)M\theta_{nk}, \quad (3.100)$$

$$\mathcal{W}_4 = \sum_{k''=k+1}^{K}\gamma_{nk''}[t+u]\left(M\theta_{nk''} + \beta_{nk''}\right)M\theta_{nk}, \quad (3.101)$$

$$\mathcal{W}_5 = \sum_{n' \neq n, n'=1}^{N} \sum_{k'''=1}^{K} \gamma_{n'k'''}[t+u]\beta_{n'k'''} M\theta_{nk}, \tag{3.102}$$

$$\mathcal{W}_6 = \sigma^2 M\theta_{nk}. \tag{3.103}$$

Proof. The detailed proof is provided in Appendix B.7. □

Remark: The effects of imperfect SIC and inter-group interference *under predicted channel* are captured in $\mathcal{W}_3$ and $\mathcal{W}_5$, respectively. Moreover, $\mathcal{W}_4$ captures the intra-group interference *under predicted channel*.

3.4 Power Control

Now, we formulate two power allocation optimization problems, namely, max-min power control and proportional fairness power control. The goal of max-min power control is to maximize the minimum SINR or SE among all users in the system. This ensures that the user experiencing the worst performance has their quality of service (QoS) improved as much as possible. In this scheme, the power levels of all users are adjusted such that the user with the lowest SINR or SE gets the best possible performance. This often involves reducing the power of users with higher SINR to decrease interference and boost the performance of users with lower SINR. Max-min power control provides a high level of fairness by ensuring that no user experiences significantly poorer performance than others. This is particularly important in scenarios with users having varying channel conditions, such as those at the cell edge or in high interference areas. It guarantees a minimum level of performance (SINR/SE) for all users, which is beneficial for applications that require reliable connections, such as emergency communications or voice services.

However, by focusing on the weakest user, the overall system capacity or sum SE may be compromised, as power is not allocated to maximize the total throughput but to equalize performance. Users with good channel conditions may have to lower their transmission power, leading to suboptimal utilization of their potential capacity. In cellular networks, max-min power control is ideal for improving the performance of users located at the edges of the cell who generally experience weaker signals and higher interference. It is suited for applications

where ensuring a minimum quality for all users is more important than maximizing the total throughput.

On the other hand, the objective of proportional fairness power control is to balance the trade-off between maximizing the total throughput (sum SE) and ensuring fairness among users. It aims to allocate resources such that the relative improvement in each user's SE is proportional to their current SE. This scheme works by assigning power levels to users in a way that maximizes the product of their SEs. It tries to maintain a balance between users, allowing those with good channel conditions to use higher power while not completely neglecting users with poorer conditions. The proportional fairness criterion is typically implemented by maximizing the sum of the logarithms of users' SEs. It provides a balanced trade-off between fairness and sum SE. Users with better channel conditions can still achieve high SE, while users with poorer conditions are not completely deprived of resources.

Compared to max-min power control, proportional fairness generally results in a higher total system capacity because it allows users with better channels to contribute more to the total sum SE. While proportional fairness power control ensures a level of fairness, it does not provide the strict fairness guarantee of max-min power control. Users with very poor channel conditions might still experience relatively low SE. Proportional fairness is often used in data-centric networks, where the goal is to maximize the total data throughput while maintaining an acceptable level of fairness among users. In scenarios with a mix of users having good and poor channel conditions, such as urban networks with varied user densities, proportional fairness provides a good balance between throughput and fairness.

Max-min power control focuses on fairness by equalizing the performance of all users, ensuring that the user with the worst performance is improved. Proportional fairness strikes a balance between fairness and total system throughput, allowing some users to perform better if it benefits the overall network performance. Therefore, max-min power control is preferable when fairness and QoS are critical, and it is essential that no user experiences very poor performance. And proportional fairness power control is suitable for environments where maximizing sum SE is important but without completely sacrificing fairness among users. Consequently, max-

min power control is ideal in environments with high variability in channel conditions, such as users at the edge of a cell or in high interference zones. And proportional fairness power control works better in more homogeneous environments or when the system can tolerate some level of performance inequality among users.

In summary, max-min power control is used when the objective is to ensure fairness and maintain a minimum acceptable performance level for all users, often at the expense of overall throughput. Proportional fairness power control is chosen when the goal is to maximize total system throughput while maintaining a balanced level of fairness among users. Selecting the appropriate power control strategy depends on the specific requirements of the network, such as whether fairness or sum SE is prioritized, and the nature of the user distribution and channel conditions.

3.4.1 Max-Min Power Control

The max-min power control maximizes the minimum achievable SE among all users, boosting the performance of users with weaker channels and ensuring fairness. This is also quantified in Section 2.5 using a widely used fairness metric called Jain's fairness index J, which is given by 2.76 [82]. It is shown that $J = 1$ for max-min power control, thus making it the fairest policy [86].

For scheme $\overline{p} \in \{\text{ZF-I}, \text{ZF-S}, \text{MR-I}, \text{MR-S}, \text{ZF-I-P}\}$, the max-min optimization problems can be formulated as

$$\mathcal{P}_1 : \text{maximize} \quad \overline{\text{SINR}}^{\overline{p}}[t+u], \tag{3.104}$$

$$\text{with respect to} \quad \gamma_{11}[t+u], \cdots, \gamma_{NK}[t+u],$$

$$\text{subject to} \quad \text{SINR}_{nk}^{\overline{p}}[t+u] \geq \overline{\text{SINR}}^{\overline{p}}[t+u], \; n \in \mathcal{N}, k \in \mathcal{K}, \tag{3.105}$$

$$0 \leq \gamma_{nk}[t+u] \leq \rho_{ul}, \; n \in \mathcal{N}, k \in \mathcal{K}, \tag{3.106}$$

where the optimization variables are power control coefficients $\{\gamma_{nk}[t+u]\}$. Also, (3.105) and (3.106) denote the common target SINR constraint and each user's transmit power constraint, respectively. The formulated problems are convex programs and we solve them numerically

with the help of CVX, which is a MATLAB-based optimization tool [14, 76].

3.4.2 Proportional Fairness Power Control

There could be scenarios where some users demand higher quality-of-service than others and complete fairness may not be the best strategy to employ. In this case, proportional fairness power control can strike a very good balance between fairness and sum SE [87].

For scheme $\overline{p} \in \{\text{ZF-I, ZF-S, MR-I, MR-S}\}$, the proportional fairness optimization problems can be formulated as

$$\mathcal{P}_2 : \text{maximize} \quad \prod_{n=1}^{N} \prod_{k=1}^{K} \overline{\text{SINR}}_{nk}^{\overline{p}}[t+u], \tag{3.107}$$

$$\text{with respect to} \quad \gamma_{11}[t+u], \cdots, \gamma_{NK}[t+u],$$

$$\text{subject to} \quad \text{SINR}_{nk}^{\overline{p}}[t+u] \geq \overline{\text{SINR}}_{nk}^{\overline{p}}[t+u], \; n \in \mathcal{N}, k \in \mathcal{K}, \tag{3.108}$$

$$0 \leq \gamma_{nk}[t+u] \leq \rho_{ul}, \; n \in \mathcal{N}, k \in \mathcal{K}, \tag{3.109}$$

where the optimization variables are power control coefficients $\{\gamma_{nk}[t+u]\}$. Also, (3.108) and (3.109) denote the common target SINR constraint and each user's transmit power constraint, respectively. The formulated problems are geometric programs and we solve them efficiently using CVX [14, 76].

3.5 Numerical and Simulation Results

In this section, we present numerical results to illustrate the performance of different decoding schemes (ZF-I, ZF-S, MR-I, MR-S, ZF-I-P) under channel aging. We plot the analytical results based on new expressions derived in Theorems 3.1-3.6 in Sections 3.2 and 3.3. We also present Monte Carlo simulations to corroborate our analysis. The presented numerical results are averaged over user locations.

Along with max-min and proportional fairness power control discussed in Section 3.4, we also consider two ad-hoc power control strategies, namely, equal and inversion power control. In equal power control, UE_{nk} transmits with power $\gamma_{nk}[t+u] = \rho_{ul}$, $\forall n, k$. On the other hand,

Table 3.1: Simulation parameters.

Parameter	Description
User distribution	Uniformly dropped in an annular ring
Inner radius of annular ring (d_o)	100 m
Outer radius of annular ring	300 m
Reference path loss (β_o)	−35.3 dB
Path loss exponent (ν)	2.4
Path-loss model (β_{nk})	$\beta_{nk} = 10^{(\phi/10)}\beta_o\,(d_o/d_{nk})^{\nu}$, where $\phi \sim \mathcal{CN}(0, \sigma_s^2)$, $\sigma_s^2 = 4$ dB is the variance of log-normal shadowing, d_{nk} is the path distance from BS to UE_{nk}
Bandwidth	20 MHz
Noise figure	7 dB
Noise power $\left(\sigma^2\right)$	−93.8 dBm
Transmit power at each user (ρ_{ul})	30 dBm

under inversion power control, UE_{nk} transmits with power $\gamma_{nk}[t+u]$ given by

$$\gamma_{nk}[t+u] = \rho_{ul}\frac{\min\left[\beta_{n,1}, \beta_{n,2} \cdots \beta_{n,k} \cdots \beta_{n,K}\right]}{\beta_{nk}}. \tag{3.110}$$

Table 3.1 lists the simulation parameters.

3.5.1 Impact of Channel Aging on the Per-User Achievable SE

3.5.1.1 Equal Power Control

To comprehend the effect of channel outdatedness when equal power control is employed, in Fig. 3.2, we plot the per-user achievable SE $\left(R_{\text{avg}}^{\overline{p}}[t+1]\right)$ vs normalized Doppler shift ($f_D T_S$) for schemes ZF-I, ZF-S, MR-I, MR-S that we have analyzed. Here, for scheme $\overline{p}$, $R_{\text{avg}}^{\overline{p}}[t+1]$ is defined as

$$R_{\text{avg}}^{\overline{p}}[t+1] = \frac{1}{L}\sum_{n=1}^{N}\sum_{k=1}^{K}\log_2\left\{1 + \text{SINR}_{nk}^{\overline{p}}[t+1]\right\}, \tag{3.111}$$

where $\overline{p} \in \{\text{ZF-I}, \text{ZF-S}, \text{MR-I}, \text{MR-S}\}$. We observe that Monte Carlo simulations closely match the analysis for each scheme.

Note that, as we increase $f_D T_S$, the channel becomes increasingly outdated. Clearly, if the impact of channel aging on a particular decoding scheme is high, then $R_{\text{avg}}^{\overline{p}}[t+1]$ dies out faster. We can observe that, $R_{\text{avg}}^{\overline{p}}[t+1]$ of scheme ZF-I dies out faster than that of scheme ZF-S. This is because the performance of ZF is susceptible to the quality of estimates [14]. As $f_D T_S$ increases, the estimates used for ZF become more and more outdated, leading to severe performance degradation for scheme ZF-I. On the other hand, the quality of estimates acquired

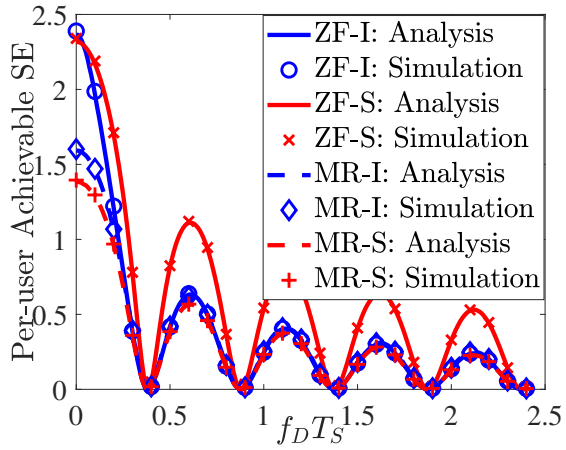


Fig. 3.2: Equal power control: NF grouping, $N = 10$, $K = 2$, $M = 128$.

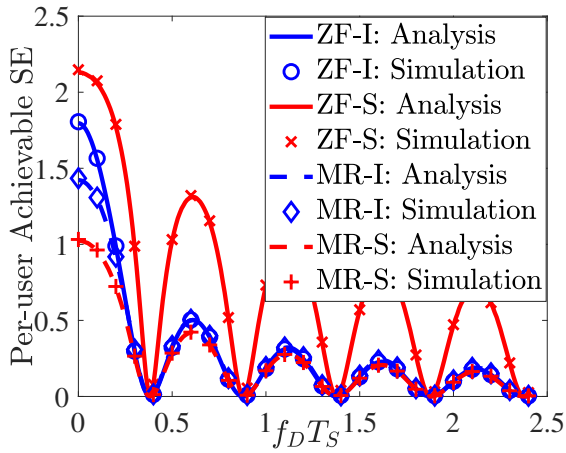


Fig. 3.3: Inversion power control: NF grouping, $N = 10$, $K = 2$, $M = 128$.

using Scheme-S is far better than those acquired using Scheme-I. Therefore, the performance degradation of scheme ZF-S is the least.

Further, we know that the MR decoder is less sensitive to the quality of channel estimates than the ZF decoder. The most dominant factor controlling the performance of MR decoder is the BF gain. Therefore, we see little difference in the rates at which the performances of schemes MR-I and MR-S die out as the outdatedness of the channel estimates increases.

3.5.1.2 Inversion Power Control

Next, in Fig. 3.3, we plot $R_{\text{avg}}^{\overline{p}}[t+1]$ vs $f_D T_S$ for inversion power control. Simulations closely match the analysis. We can observe similar performance trends with inversion power control as observed in Fig. 3.2 with equal power control.

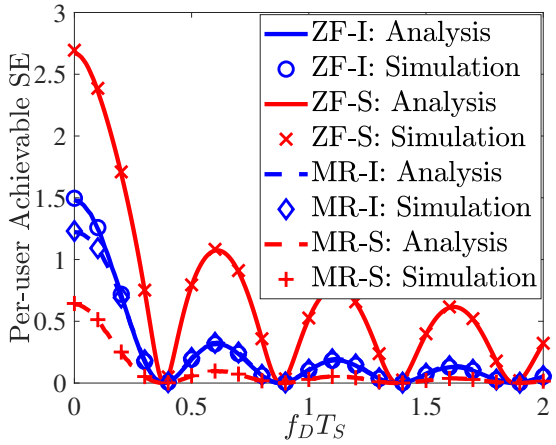


Fig. 3.4: Max-min power control: NF grouping, $N = 10$, $K = 2$, $M = 128$.

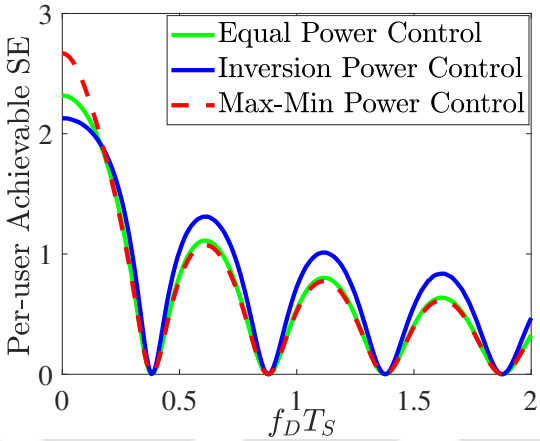


Fig. 3.5: Ad-hoc vs max-min power control: NF grouping, $N = 10$, $K = 2$, $M = 128$.

3.5.1.3 Max-Min Power Control

Next, we plot $R_{\text{avg}}^{\overline{p}}[t+1]$ vs $f_D T_S$ for max-min power control in Fig. 3.4. Monte Carlo simulations closely match the analysis. Trends similar to Figs. 3.2, 3.3 are observed. Specifically, with max-min power control as well, channel aging leads to severe degradation in performances of schemes ZF-I, MR-I and MR-S, and on the contrary, scheme ZF-S is the least vulnerable to channel aging.

3.5.1.4 Ad-hoc Power Control vs Max-Min Power Control

Observing the best performing scheme ZF-S in Figs. 3.2, 3.3 and 3.4, in Fig. 3.5, we plot $R_{\text{avg}}^{\text{ZF-S}}[t+1]$ vs $f_D T_S$ to compare equal, inversion and max-min power control strategies. We observe that, for smaller values of $f_D T_S$ ($\lesssim 0.25$), max-min power control provides higher $R_{\text{avg}}^{\text{ZF-S}}[t+1]$ and for higher values of $f_D T_S$, inversion power control gives higher $R_{\text{avg}}^{\text{ZF-S}}[t+1]$.

$f_D T_S$	Max-Min	Equal	Inversion
0.05	2.596	2.280	2.101
0.6	1.072	1.110	1.311

Table 3.2: Comparison of Per-User Achievable SEs (in bps/Hz).

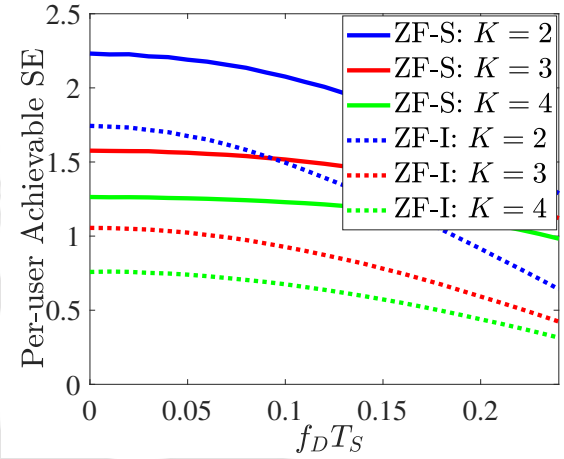


Fig. 3.6: ZF-I vs ZF-S under max-min power control with NB grouping: $K = 2, 3, 4$, $L = 24$, $M = 128$.

Specifically, for $f_D T_S = 0.05$, the per-user achievable SE with max-min power control is higher by 14% and 23% than that with equal power control and inversion power control, respectively. On the other hand, for $f_D T_S = 0.6$, the per-user achievable SE with inversion power control is higher by 18% and 22% than that with equal power control and max-min power control, respectively. This is primarily because inversion power control significantly boosts the SE of users with relatively weaker channel strengths. In addition, as illustrated in Section 3.2, weaker users also have an advantage of intra-group interference cancellation due to the order in which SIC is performed at the BS. However, low power allocation under inversion power control degrades stronger users' SE. We emphasize that the boost in weaker users' performance is more significant than the performance degradation of stronger users. Hence, we see that scheme ZF-S with inversion power control is more robust under channel aging.

3.5.1.5 Impact of Users Per Group (K)

Now, in Fig. 3.6, we plot $R_{\text{avg}}^{\overline{p}}[t+1]$ vs $f_D T_S$ for schemes ZF-S and ZF-I under max-min power control for $K = 2, 3, 4$, using scalable NB grouping, which allows more than two users in a group. Note that the number of groups $N = L/K$ decreases as K increases, leading to increase in the BF gain $(M - N)$ and hence increase in the variance of desired signal for both

$f_D T_S$	ZF-S: NF for $K = 2$	ZF-S: NF for $K = 3$
0.05	2.19	1.56

Table 3.3: Comparison of Per-User Achievable SEs (in bps/Hz) under Max-Min Power Control.

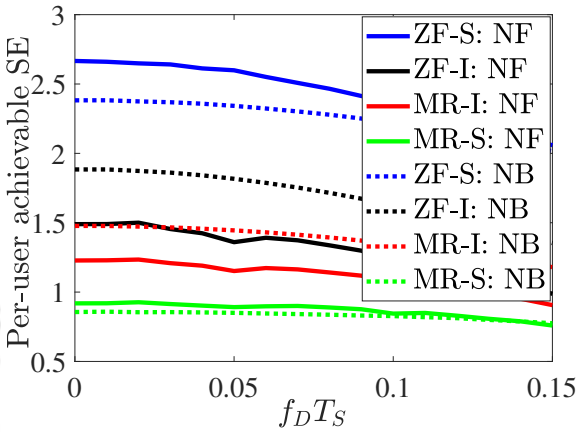


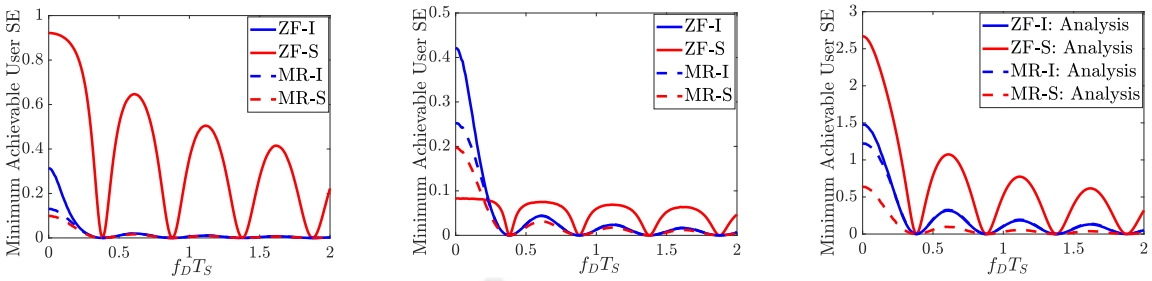
Fig. 3.7: NF vs NB grouping under max-min power control: $K = 2, 3, 4$, $L = 24$, $M = 128$.

schemes ZF-S and ZF-I, where M is the number of antennas. However, increase in K also leads to increased pilot contamination during estimation because in both Scheme-I and Scheme-S, all users within a group share the same pilot signal.

In addition, increase in K leads to increase in the impact of imperfect SIC [86]. In effect, the performance degradation due to (i) increased pilot contamination and (ii) increased impact of imperfect SIC, dominates over the performance improvement due to increased BF gain $(M-N)$. Thus, as we increase K, the performance of both schemes ZF-S and ZF-I degrades. Specifically, for scheme ZF-S under NF grouping and max-min power control, the per-user achievable SE with $K = 2$ is 40% greater than that with $K = 3$. Also, the performance of both schemes, for all values of K, degrades with increase in channel outdatedness $f_D T_S$.

3.5.1.6 Impact of User Grouping Strategies

Now, we investigate the performance of discussed schemes with two large-scale fading-based user grouping strategies, namely, NF and NB grouping under channel aging. These user grouping strategies are discussed in detail in Section 2.4. In Fig. 3.7, we plot $R_{\text{avg}}^{\overline{p}}[t+1]$ vs $f_D T_S$ for all considered schemes and grouping strategies under max-min power control. As channel outdatedness increases, as expected, the performance of all decoding schemes degrades. Interestingly, scheme ZF-S performs better with NF grouping and scheme ZF-I performs better



(a) Equal power control. **(b)** Inversion power control. **(c)** Max-min power control.

Fig. 3.8: Minimum achievable user SE (in bits/s/Hz) with NF grouping: $N = 10$, $K = 2$, $M = 128$.

with NB grouping.[7] Further, scheme ZF-S performs significantly better compared to scheme ZF-I for both grouping strategies, due to significantly boosted quality of estimates with Scheme-S.

3.5.2 Impact of Channel Aging on Minimum Achievable User SE

Now, we consider a metric called the minimum achievable user SE. For decoding scheme $\overline{p} \in \{\text{ZF-I, ZF-S, MR-I, MR-S}\}$, the minimum achievable user rate is defined as

$$R_{\text{min}}^{\overline{p}}[t+1] = \min\{R_{nk}^{p}[t+1],\, n \in \mathcal{N}, k \in \mathcal{K}\}. \tag{3.112}$$

To understand how channel outdatedness impacts $R_{\text{min}}^{\overline{p}}[t+1]$ for each decoding scheme, in Figs. 3.8(a), 3.8(b) and 3.8(c), we plot $R_{\text{min}}^{p}[t+1]$ vs $f_D T_S$ under equal, inversion and the derived max-min power control, respectively. We observe that, in terms of $R_{\text{min}}^{\overline{p}}[t+1]$ as well, scheme ZF-S is the most robust to channel aging under considered power control strategies. This is because of the improved quality of estimates acquired using Scheme-S.

In Fig. 3.9, we compare $R_{\text{min}}^{\text{ZF-S}}[t+1]$ vs $f_D T_S$ under equal, inversion and max-min power control. Here, the max-min power control significantly boosts $R_{\text{min}}^{\text{ZF-S}}[t+1]$ and hence is more robust to channel aging.[8] Specifically, $R_{\text{min}}^{\text{ZF-S}}[t+1]$ of Scheme ZF-S under NF grouping for $f_D T_S = 0.05$ with max-min power control is 2.8 times of that with equal power control and is 6.6 times of that with inversion power control. This is because, unlike equal and inversion power control, the max-min power control ensures that the minimum achievable SE is maximized and

[7] The impact of pilot contamination experienced by the weakest user is effectively reduced by using: (i) NF grouping with Scheme-S and (ii) NB grouping with Scheme-I.

[8] In case of max-min power control for a given decoding scheme, per-user achievable SE is the same as minimum achievable user SE.

$f_D T_S$	Max-Min	Equal	Inversion
0.05	2.596	0.918	0.390

Table 3.4: Comparison of Minimum Achievable User SEs (in bps/Hz) with scheme ZF-S under NF Grouping.

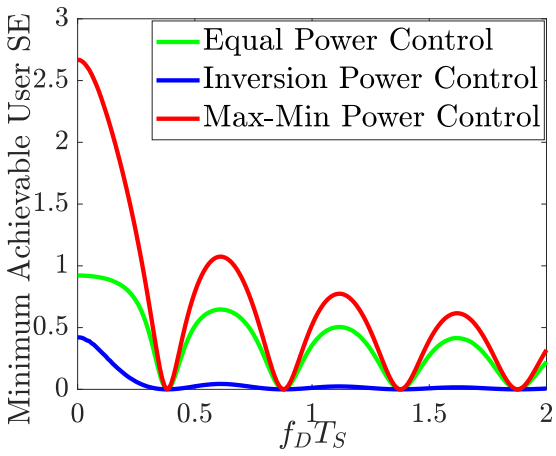


Fig. 3.9: ZF-S with NF grouping: $N = 10$, $K = 2$, $M = 128$.

all users get the same SE performance.

3.5.3 Proportional Fairness Power Control

Now, we discuss numerical results for the proportional fairness power control discussed in Section 3.4.2. In Fig. 3.10, we present $R_{\text{avg}}^{\overline{p}}[t+1]$ vs $f_D T_S$ to compare max-min and proportional fairness power control. We observe the improvement in $R_{\text{avg}}^{\overline{p}}[t+1]$ of each scheme when proportional fairness is employed instead of max-min power control. Specifically, $R_{\text{avg}}^{\overline{p}}[t+1]$ of ZF-S under NF grouping for $f_D T_S = 0.05$ with proportional fairness is 30% higher than that with max-min power control. This is due to improved sum SE with proportional fairness power control.

Furthermore, to quantify the improvement in sum SE at the cost slight degradation in fairness, we plot the SE for 20 *UE*s in Fig. 3.11. As presented, with max-min power control, $J = 1$ and with proportional fairness power control, $J = 0.8203$. At the cost of degradation in fairness, the proportional fairness power control provides the sum SE of 67.8084 bps/Hz, which is a significant improvement from sum SE of 58.8214 bps/Hz achievable with max-min power control.

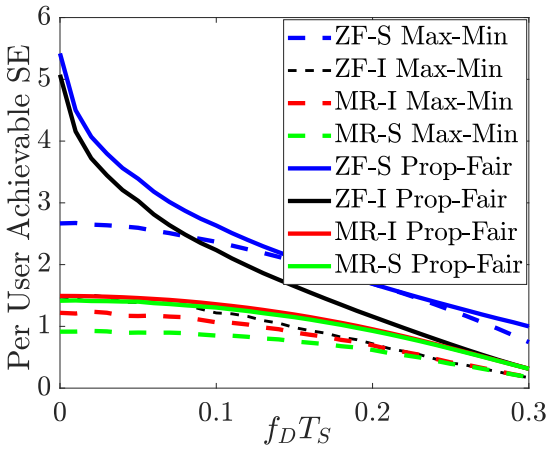


Fig. 3.10: Max-min vs proportional fairness with NF grouping: $N = 10$, $K = 2$, $M = 128$.

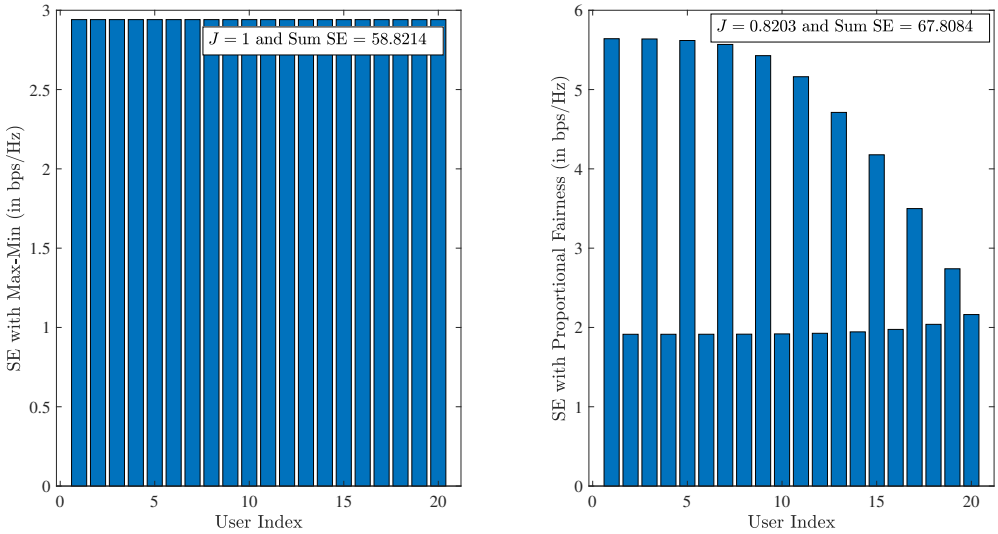


Fig. 3.11: Comparison of max-min and proportional fairness: $N = 10$, $K = 2$, $M = 128$, NF grouping.

3.5.4 Performance of ZF Designed using Predicted CSI

In Fig. 3.12, we compare $R^{\overline{p}}_{\text{avg}}[t+1]$ vs $f_D T_S$ under max-min power control for decoding schemes ZF-I, ZF-S and ZF-I-P, for the prediction filter of order $p = 0, 1, 10, 20$ and transmit SNR of 20 dB during pilot training phase.

The predicted channel with order $p = 0$ is equivalent to outdated estimated CSI. Thus, schemes ZF-I and ZF-I-P with $p = 0$ perform the same. As we increase the order p of the predictor, $R^{\overline{p}}_{\text{avg}}[t+1]$ improves when compared to $p = 0$.[9] For example, the per-user achievable

[9]It is important to note that the processing complexity increases significantly with increase in predictor order p [42].

$f_D T_S$	Max-Min	Proportional Fairness
0.05	2.596	3.390

Table 3.5: Comparison of Per-User Achievable SEs (in bps/Hz) under Max-Min and Proportional Fairness Power Control.

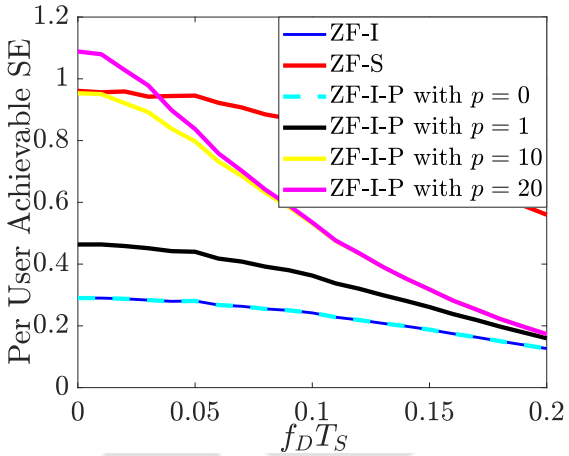


Fig. 3.12: ZF with predicted CSI under max-min power control: $N = 10$, $K = 2$, $M = 128$, NF grouping.

SE for scheme ZF-I-P with predictor of order $p = 20$ is 12.5% higher than for scheme ZF-S under NF grouping and max-min power control. However, for the considered setup, the predictor of order $p > 10$ is required to achieve better $R_{\text{avg}}^{\overline{p}}[t+1]$ than scheme ZF-S with outdated estimated CSI for lower values of $f_D T_S$ ($\lesssim 0.04$). However, note that, even $f_D T_S = 0.001$ corresponds to user speed of 150 meters per second for $f_c = 2$ GHz and $T_S = 1\,\mu$S, which is the typical speed foreseen in most high mobility applications in 6G.

3.6 Summary

We investigated the performance of a uplink massive MIMO-NOMA system under the influence of channel outdatedness in the estimates acquired using two different low-overhead schemes, namely, Scheme-I and Scheme-S. We also investigated two user grouping strategies, namely, NF and NB grouping. We derived novel lower bounds on the achievable SE for both ZF and MR decoders that are designed based on outdated channel estimates accounting for pilot contamination, time-variations in the channel response, and imperfect SIC. Furthermore, we derived the lower bounds on the achievable SE with ZF and MR decoders designed using predicted CSI obtained with p^{th} order WLP.

To ensure fairness in performance under channel aging, we formulated and solved max-min

$f_D T_S$	ZF-I-P with $p = 20$	ZF-S
0.01	1.08	0.96

Table 3.6: Comparison of Per-User Achievable SEs (in bps/Hz) ZF under NF Grouping and Max-Min Power Control.

power control optimization problems for both ZF and MR decoders based on outdated estimated CSI and also for ZF decoder based on predicted CSI. To strike a balance between fairness and sum SE, we formulated and solved proportional fairness power control optimization problems for both ZF and MR decoders based on outdated estimated CSI. We also analyzed equal and inversion power control. We make a few interesting inferences.

Better quality estimates are obtained for (a) Scheme-I with NB grouping and (b) Scheme-S with NF grouping. Under outdated estimated CSI and imperfect SIC, among equal, inversion and max-min power control, scheme ZF-S with NF grouping gives the highest per-user achievable SE under max-min power control for $f_D T_S \lesssim 0.25$ and with inversion power control for higher values of $f_D T_S$. Also, scheme ZF-S with max-min power control is the least vulnerable to channel aging in terms of minimum achievable user SE. We also identified that, for the considered setup, the WLP of order $p > 10$ is needed to beat per-user achievable SE of ZF-S with outdated estimated CSI for lower values of $f_D T_S$ ($\lesssim 0.04$). These inferences can be used by a wireless engineer to choose the most suitable decoding and power control strategy under channel aging and pilot contamination in massive MIMO-NOMA systems.

3.7 Comparison of Contributions in Chapters 2 and 3

As we investigated uplink massive MIMO-NOMA systems in depth in Chapters 2 and 3, it becomes important to identify the critical distinctions in modelling, analysis and results presented in these chapters. These key distinctions are presented in Table 3.7 and are discussed in detail below.

In Chapter 2, the channel is modeled as a quasi-static Rayleigh faded channel, which means it remains constant over the duration of a coherence interval. To estimate the channel, MMSE-based estimation schemes, referred to as Scheme-I and Scheme-S, are employed. In Chapter 3, we consider a more dynamic channel model compared to the channel model in Chapter 2.

	Chapter 2	Chapter 3
Channel Model	Quasi-static Rayleigh faded channel assumed to be constant over the transmission interval. The channel is estimated using MMSE-based estimation schemes, Scheme-I and Scheme-S.	Channel is assumed to be constant during the channel estimation phase and varies from one symbol to the next symbol during data transmission phase. The outdated estimates are modelled based on the time correlation between the channel estimates acquired with Scheme-I and Scheme-S. As a counter measure to channel aging, channel prediction is also proposed using linear prediction.
Decoders	ZF decoder designed using the channel estimates acquired with Scheme-I and Scheme-S.	ZF and MR decoders designed using the outdated estimates acquired with Scheme-I and Scheme-S.
Power Control	Max-min power control to ensure fairness in terms of SE	Max-min power control to ensure fairness in terms of SE and proportional fairness to improve sum SE while maintaining fairness to a significant extent. Performance under ad-hoc power control policies like equal power control and inversion power control are also studied.
Closed-form expressions	Lower bounds on the achievable SE with ZF decoders designed using Scheme-I and Scheme-S, incorporating the effects of channel estimation error and imperfect SIC.	Lower bounds on the achievable SE with MR and ZF decoders designed using outdated estimates acquired with Scheme-I and Scheme-S, incorporating the effects of channel aging, imperfect SIC, channel estimation errors. Lower bound on achievable SE with ZF decoder designed using the predicted channel is also presented.

Table 3.7: Comparison of modelling, analysis and contributions in Chapter 2 and Chapter 3.

Here, the channel is assumed to be constant only during the initial channel estimation phase but becomes outdated after this phase, as it varies from one symbol duration to the next during the data transmission phase. This variation introduces channel aging. The outdated channel estimates are modeled based on the time correlation between the channel estimates obtained using Scheme-I and Scheme-S. To counteract the adverse effects of channel aging, a channel prediction technique using linear prediction method is proposed, enhancing the reliability of the

channel estimates over time.

In Chapter 2, we discuss the design of a ZF decoder, which is based on the channel estimates derived from Scheme-I and Scheme-S. The ZF decoder is designed to nullify inter-group interference, thereby improving the quality of received signal. This approach is suitable for scenarios where accurate channel estimation can be achieved, providing a baseline for comparison with other decoder designs. In Chapter 3, the design of decoders is extended to accommodate the effects of channel aging. Both ZF and MR decoders are designed using the outdated channel estimates obtained through Scheme-I and Scheme-S. The ZF decoder design based on the predicted channel is also investigated.

In Chapter 2, a max-min power control strategy is employed to ensure fairness among users in terms of SE. This approach aims to maximize the minimum achievable SE among all users, preventing any single user from experiencing significantly worse performance compared to others. In Chapter 3, we expand on the power control strategies discussed in Chapter 2 by introducing additional techniques. Besides the max-min power control, proportional fairness is introduced to balance the overall sum SE while still maintaining a reasonable level of fairness among users. In this chapter, we also explore ad-hoc power control policies, such as equal power control and inversion power control, to examine their impact on system performance under different channel conditions.

In Chapter 2, we derive lower bounds on the achievable SE for systems using ZF decoders. These bounds take into account the channel estimation errors and the imperfections in SIC. In Chapter 3, the analysis is extended to consider the effects of channel aging. Lower bounds on the achievable SE are derived for both MR and ZF decoders designed using the outdated channel estimates obtained with Scheme-I and Scheme-S. This comprehensive analysis incorporates factors such as channel aging, imperfect SIC, and channel estimation errors. Additionally, in Chapter 3 we present lower bound on the SE for systems employing ZF decoders designed using predicted channel.

4

IRS-Enabled NOMA Systems

Contents

4.1 System Model 101
4.2 Error Probability Analysis 104
4.3 Numerical and Simulation Results 119
4.4 Summary 126

So far, we investigated massive MIMO-NOMA systems in depth and addressed several practically relevant challenges in Chapters 2 and 3 of the thesis. The performance benefits of massive MIMO-NOMA systems ensured increased connectivity, improved SE and improved minimum achievable SE. However, the potential of massive MIMO systems lies in the use of a large number of antennas and hence a large number of RF chains at the BS. This comes at a hardware cost. Thus, it is important to explore cost-efficient alternatives to large antenna systems. IRS is a recent technology which promises the gains similar to large antenna systems like massive MIMO without employing the active RF chains.

In this chapter, we analyze the error probability performance on the downlink of an IRS-enabled NOMA system. Firstly, we describe the system model incorporating practically relevant channel model. Then we discuss in detail the SEP analysis of a two-user setup for PAM and QAM modulation of arbitrary order. Next, we characterize the error probability performance of a multi-user setup, within which we discuss the SEP analysis for multicast transmission followed by PEP analysis for unicast transmission. To bring in the practical aspects of the system, we characterize the impact of discrete phase shifts at the IRS and imperfect CSI on the average SEP performance. We also formulate and solve a constrained optimization problem to determine the optimal fraction of elements to be activated at any IRS to guarantee near-identical SEP for both users in a two-user setup. Finally, we present extensive numerical results to derive novel design insights.

Chapter Organization: In Section 4.1, we discuss system model in detail. In Section 4.2, we present the detailed error probability analysis followed by extensive numerical results and discussion in Section 4.3. Finally, we summarize the chapter in Section 4.4.

4.1 System Model

We consider a single-antenna BS serving two single-antenna users, UE_1 and UE_2 on the downlink using power-domain NOMA. We refer to UE_1 as the *weak* user and UE_2 as the *strong* user.[1] IRS$_k$ with T elements is deployed to assist communication to UE_k as shown in Fig. 4.1.

[1]The user is classified as a weak user or strong user based on its effective channel strength compared to the other user taking the channel ordering into account, i.e., after phase configuration at each IRS.

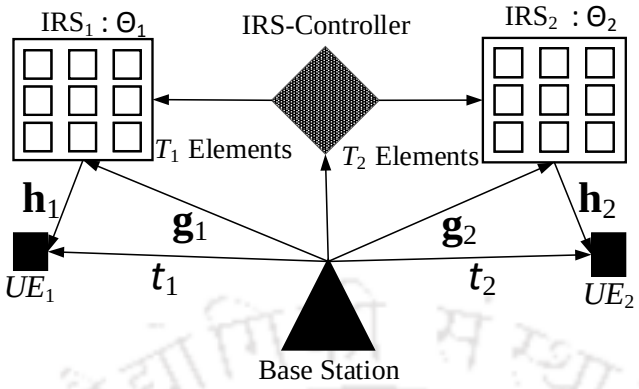


Fig. 4.1: System model: IRS-aided NOMA, two-user case.

The use of a separate IRS to serve each user enables constructive interference induced boost in the received signal energy simultaneously at both the users, since they can then be independently configured. Moreover, employing two different IRSs enables the system to provide simultaneous service to multiple users. On the contrary, use of a single IRS to serve both users leads to compromise in performance of one of the users or both users. Furthermore, we consider that the signal reflected by IRS$_k$ is too weak to cause any interference to $UE_{k'}$, where $k, k' \in \{1, 2\}$ and $k \neq k'$. We consider only first order reflections, since the reflections between the IRSs will further weaken the signal.[2] The BS broadcasts the superposition coded symbol which reaches both users via corresponding *direct paths* BS-UE_1 and BS-UE_2, and via corresponding *reflected paths* through the IRS, namely, BS-IRS$_1$-UE_1 and BS-IRS$_2$-UE_2.[3] The IRS is connected to BS

[2]The interference to UE_k from the signal reflected between the IRSs will be very weak because the signal reflected between IRSs will reach the undesired user in three hops, i.e., the first hop is from the BS to IRS$_i$, the second hop is from the IRS$_i$ to IRS$_k$ and the third hop is from IRS$_k$ to UE_i, for $i \neq k$ and $i, k \in \{1, 2\}$. Therefore, the effective path-loss corresponding to the interference signal reflected between IRSs is the product of the path-loss values corresponding to three hops reducing the strength of the interference signal to a negligibly small value. On top of that, the signal reflected between the IRS$_i$ and IRS$_k$ will not add up in phase with the desired signal at UE_i since the phase shifts at IRS$_k$ are programmed to beamform the signal to UE_k and not to UE_i. Hence the interference signals reaching UE_k after reflection from IRS$_i$ and the interference signal reaching UE_k after reflection between IRS$_k$ and IRS$_i$ can be neglected without affecting the effective SINR.

[3]The channel between the BS and each *UE* is composed of two components: (i) the path from BS to *UE* without any reflection from the IRS and (ii) the path from BS to *UE* with reflection from IRS. The term *direct path* signifies the first component of channel through which signal reaches from the BS to each *UE* without any reflection from IRS. The term *reflected path* signifies the second component of channel through which the signal reaches from the BS to each *UE* after reflection from IRS. Note that the term *direct path* does not necessarily signify the line-of-sight (LoS) channel. The channel fading via *direct path* may or may not contain an LoS component. We assume that, there is no LoS component in *direct path* channel between the BS and each UE. This assumption is based on the fact that the *UE* may be located either far from the BS or in a position with its LoS path blocked by the obstacles. Further, due to an NLoS component owing to rich scattering assumption, we consider the *direct path* channel as Rayleigh fading. Similarly, the channel fading via reflected path may or may not contain an LoS component. We assume that the *reflected path* component of the channel consists of a NLoS component due to rich scattering as

through a dedicated control link, and BS configures phase shifts of IRS_k to ensure coherent reception at UE_k.

Let $z_k \in \mathbb{C}$ denote the effective channel coefficient from the BS to UE_k. It is given by

$$z_k = t_k + \mathbf{g}_k^{\mathrm{T}} \Theta_k \mathbf{h}_k = t_k + \sum_{i=1}^{T} \left|g_{k,i}\right| e^{j\left(\theta_{g_{k,i}} + \theta_{k,i} + \theta_{h_{k,i}}\right)} \left|h_{k,i}\right|, \tag{4.1}$$

where $(\cdot)^{\mathrm{T}}$ denotes the transpose operator, $t_k = |t_k| e^{j\theta_{t_k}}$ is the channel coefficient of the Rayleigh fading direct path from the BS to UE_k, $\mathbf{g}_k = [g_{k,1} \cdots g_{k,T}]^{\mathrm{T}}$ is the Rician fading channel vector from the BS to IRS_k where $g_{k,i} = \left|g_{k,i}\right| e^{j\theta_{g_{k,i}}}$ is the channel coefficient from BS to the i^{th} element of IRS_k, $\mathbf{h}_k = [h_{k,1} \cdots h_{k,T}]^{\mathrm{T}}$ is the Rician fading channel vector from IRS_k to UE_k where $h_{k,i} = \left|h_{k,i}\right| e^{j\theta_{h_{k,i}}}$ is the channel coefficient from the i^{th} element of IRS_k to UE_k, and $\Theta_k = \text{diag}\left[e^{j\theta_{k,1}}, e^{j\theta_{k,2}}, \cdots, e^{j\theta_{k,T}}\right]$ is the phase-shift matrix corresponding to IRS_k, where $\theta_{k,i}$ denotes the phase-shift induced by the i^{th} element of IRS_k. Firstly, we assume that the BS has perfect knowledge of the CSI, which it uses to configure the phases at IRS. Later we also present results with imperfect CSI in Section 4.3.1.1 to assess how robust SEP is to channel estimation errors. Based on this, the BS configures the phase shift of the i^{th} element of IRS_k as $\theta_{k,i} = \theta_{t_k} - \left(\theta_{g_{k,i}} + \theta_{h_{k,i}}\right)$, for $i \in \{1, 2, \ldots, T\}$ to ensure coherent reception at UE_k. Therefore, (4.1) can be modified as

$$z_k = |z_k| e^{j\theta_{z_k}} = \left(|t_k| + \sum_{i=1}^{T} \left|g_{k,i}\right| \left|h_{k,i}\right|\right) e^{j\theta_{t_k}}. \tag{4.2}$$

This gives $|z_k| = |t_k| + \sum_{i=1}^{T} \left|g_{k,i}\right| \left|h_{k,i}\right|$ and $\theta_{z_k} = \theta_{t_k}$. Let $|z_k| = u_k + v_k$, where $u_k \triangleq |t_k|$ and $v_k \triangleq \sum_{i=1}^{T} \left|g_{k,i}\right| \left|h_{k,i}\right|$.

Based on the channel model, u_k is Rayleigh distributed with scale parameter $\sigma_{u_k} = \sqrt{\beta_{t_k}/2}$, where β_{t_k} is the path-loss corresponding to the direct channel t_k. Since $\left|g_{k,i}\right|$ and $\left|h_{k,i}\right|$ are independent random variables, for $i \in \{1, 2, \ldots, T\}$, v_k is the sum of T i.i.d. random variables. Therefore, invoking central limit theorem (CLT), the distribution of v_k approaches Gaussian distribution with mean μ_{v_k} and variance $\sigma_{v_k}^2$, where $\mu_{v_k} = T\mu_{\mathbf{g}_k}\mu_{\mathbf{h}_k}$ and $\sigma_{v_k}^2 = T\left(\left[\beta_{\mathbf{g}_k}\beta_{\mathbf{h}_k}\right] - \left[\left(\mu_{\mathbf{g}_k}\mu_{\mathbf{h}_k}\right)^2\right]\right)$.[4] Here,

well as an LoS component. In general, IRSs are expected to be placed at an elevation from ground such that there is a high probability of LoS between the BS and IRS and between the IRS and each *UE*. Thus, the *reflected path* channel component is modelled as Rician fading.

[4]We have verified that the mean squared error (MSE) between the cumulative density functions (CDFs) of

$\beta_{\mathbf{g}_k}$ and $\beta_{\mathbf{h}_k}$ are path-losses corresponding to the channels $\mathbf{g}_k$ and $\mathbf{h}_k$, respectively. Moreover, $\mu_{\mathbf{g}_k} = \sqrt{(\pi/4)\left[\beta_{\mathbf{g}_k}/\left(1+K_{\mathbf{g}_k}\right)\right]}\mathcal{L}_{1/2}\left\{-K_{\mathbf{g}_k}\right\}$, $\mu_{\mathbf{h}_k} = \sqrt{(\pi/4)\left[\beta_{\mathbf{h}_k}/\left(1+K_{\mathbf{h}_k}\right)\right]}\mathcal{L}_{1/2}\left\{-K_{\mathbf{h}_k}\right\}$, where $K_{\mathbf{g}_k}$, $K_{\mathbf{h}_k}$ are the corresponding Rice factors, and $\mathcal{L}_{1/2}\{\cdot\}$ denotes Laguerre polynomial of order $\frac{1}{2}$ [88].

Lemma 4.1. *The PDF of the effective channel gain $|z_k| = u_k + v_k$ is given by*

$$f_{|z_k|}(x) = f_{1,k}(x) - \left[f_{2,k,1}(x) + f_{2,k,2}(x)\right] - f_{3,k}(x), \tag{4.3}$$

where $f_{1,k}(x)$, $f_{2,k,1}(x)$, $f_{2,k,2}(x)$, and $f_{3,k}(x)$ are given by

$$f_{1,k}(x) = \sqrt{2/(\pi\, b_k)}(1/\bar{q}_k)e^{-\frac{(c_k-x)^2}{2b_k}}, \tag{4.4}$$

$$f_{2,k,1}(x) = \frac{a_k\sqrt{2/(\pi\, b_k)}}{(a_k+b_k)\,\bar{q}_k}e^{-\frac{(c_k-x)^2}{2b_k}}, \tag{4.5}$$

$$f_{2,k,2}(x) = \frac{\sqrt{(2b_k)/\pi}}{(a_k+b_k)\,\bar{q}_k}e^{-\frac{\left(a_k c_k^2+b_k x^2\right)}{2a_k b_k}}, \tag{4.6}$$

$$f_{3,k}(x) = \bar{q}_k{}^{-1}\left(\frac{(c_k-x)\,\sqrt{a_k}}{(a_k+b_k)^{3/2}}e^{-\frac{(c_k-x)^2}{2b_k}}\left\{Erf\left[\frac{\sqrt{a_k}\,(x-c_k)}{\sqrt{2\,a_k\,(a_k+b_k)}}\right]+Erf\left[\frac{a_k c_k+b_k x}{\sqrt{2\,a_k b_k\,(a_k+b_k)}}\right]\right\}\right), \tag{4.7}$$

where $a_k = \sigma_{u_k}^2$, $b_k = \sigma_{v_k}^2$, $c_k = \mu_{v_k}$, $\bar{q}_k = \left(1+Erf\left[\frac{c_k}{\sqrt{2b_k}}\right]\right)$, and $Erf[\cdot]$ denotes the error function.

Proof. The proof is given in Appendix C.1. □

The superposition coded symbol transmitted by the BS is given by $s = \sqrt{P_1}s_1 + \sqrt{P_2}s_2$, where s_1 and s_2 are symbols with unit average power drawn from N_1-PAM/N_1-QAM and N_2-PAM/N_2-QAM constellations, respectively. Note that $P_1 = \lambda P$ and $P_2 = (1-\lambda)P$, where $\lambda \in (0,1)$ is NOMA power allocation coefficient and P is total transmit power at BS. In case of PAM constellation for both users, the power allocation is done such that $P_1 > \frac{(N_2-1)\left(N_1^2-1\right)}{(N_2+1)}P_2$ to ensure non-overlapping decision regions of different constellation points [89].

4.2 Error Probability Analysis

The signal received at UE_k is given by

$$y_k = z_k\, s + w_k = |z_k|\, e^{j\theta_{z_k}}\, s + w_k = |z_k|\, e^{j\theta_{t_k}}\, s + w_k, \tag{4.8}$$

and $w_k \sim \mathcal{CN}(0,\sigma_k^2)$ is AWGN at UE_k. We assume that the receiver UE_k has phase information of the effective channel and does phase compensation before processing the received signal.

$v_k = \sum_{i=1}^{T}|g_{k,i}||h_{k,i}|$ (derived) and $v_k \sim \mathcal{N}\left(\mu_{v_k},\sigma_{v_k}^2\right)$ (CLT) decreases with increase in T. At $T = 200$, the MSE between the CDFs is of the order of 10^{-5}, which is negligible.

Therefore, after phase compensation, the received signal at UE_k reduces to

$$\widetilde{y}_k = y_k\, e^{-j\theta_{t_k}} = |z_k|\ s + \widetilde{w}_k\,, \tag{4.9}$$

where $\widetilde{w}_k = w_k e^{-j\theta_{t_k}}$ has same distribution as w_k owing to circular symmetry property.

4.2.1 SEP Analysis for Two-User Setup with PAM Symbols

The SEP of UE_1 conditioned on $|z_1|$, when N_1-PAM constellation is used for UE_1 and N_2-PAM constellation is used for UE_2, is given by [89]

$$\mathcal{P}_1^{\mathrm{C}}(e) \triangleq \mathcal{P}_1(e|\,|z_1|) = \frac{2\,(N_1-1)}{N_2 N_1}\sum_{m=1}^{N_2/2}\sum_{i=1}^{2}\mathbf{Q}\left\{L_i(m)\,|z_1|\right\}, \tag{4.10}$$

where $\mathbf{Q}\{\cdot\}$ denotes the Q-function, $L_1(m) = \frac{(d_1-(2m-1)d_2)}{\sqrt{\sigma_1^2/2}}$, $L_2(m) = \frac{(d_1+(2m-1)d_2)}{\sqrt{\sigma_1^2/2}}$. Also, $d_1 = \sqrt{\frac{3P_1}{(N_1-1)(N_1+1)}}$, $d_2 = \sqrt{\frac{3P_2}{(N_2-1)(N_2+1)}}$ are half the distances between neighboring constellation points of N_1-PAM and N_2-PAM constellations, respectively. Similarly, the SEP of UE_2 conditioned on $|z_2|$ is given by

$$\mathcal{P}_2^{\mathrm{C}}(e) \triangleq \mathcal{P}_2(e|\,|z_2|) = \mathcal{P}_{2,1}^{\mathrm{C}}(e) - \mathcal{P}_{2,2}^{\mathrm{C}}(e) \tag{4.11}$$

$$= \frac{2\,(N_2-1)}{N_2}\mathbf{Q}\left\{L_3\,|z_2|\right\} - \sum_{n=1}^{N_1-1}\frac{2\,(N_1-n)}{N_2N_1}\left[(N_2-1)\left(\sum_{i=4}^{5}(-1)^i\mathbf{Q}\left\{L_i(n)\,|z_2|\right\}\right)+\left(\sum_{i=6}^{7}(-1)^i\mathbf{Q}\left\{L_i(n)\,|z_2|\right\}\right)\right], \tag{4.12}$$

where

$$L_3 = \frac{d_2}{\sqrt{\sigma_2^2/2}}, L_4(n) = \frac{(2nd_1-d_2)}{\sqrt{\sigma_2^2/2}}, L_5(n) = \frac{(2nd_1+d_2)}{\sqrt{\sigma_2^2/2}}, \tag{4.13}$$

$$L_6(n) = \frac{((2n-1)d_1+(N_2-1)d_2)}{\sqrt{\sigma_2^2/2}}, L_7(n) = \frac{((2n-1)d_1-(N_2-1)d_2)}{\sqrt{\sigma_2^2/2}}. \tag{4.14}$$

Here, $\mathcal{P}_{2,1}^{\mathrm{C}}(e)$ and $\mathcal{P}_{2,2}^{\mathrm{C}}(e)$ are the first and the second terms of (4.12), respectively. We next analyze the average SEP of both users that employ NOMA decoding protocol.

Theorem 4.1. *In an IRS-aided NOMA system where symbols of UE_1 and UE_2 are drawn from N_1-PAM and N_2-PAM constellations, respectively, the average SEP of UE_1 is given by*

$$\mathcal{P}_1^{e}(T) \approx \mathcal{I}_{11} - (\mathcal{I}_{12,1} + \mathcal{I}_{12,2}) - \mathcal{I}_{13}\,, \tag{4.15}$$

where

$$\mathcal{I}_{11} = \sum_{m=1}^{N_2/2} \frac{e^{-\frac{c_1^2}{2b_1}}(N_1-1)\sum_{i=1}^{2}\left(\chi^1_{L_i(m)}+\Xi^1_{L_i(m)}\right)}{6N_2N_1\sqrt{b_1}\bar{q}_1}, \tag{4.16}$$

$$\mathcal{I}_{12,1} = \sum_{m=1}^{N_2/2} \frac{e^{-\frac{c_1^2}{2b_1}}(N_1-1)a_1\sum_{i=1}^{2}\left(\chi^1_{L_i(m)}+\Xi^1_{L_i(m)}\right)}{6N_2N_1\sqrt{b_1}\,(a_1+b_1)\,\bar{q}_1}, \tag{4.17}$$

$$\mathcal{I}_{12,2} = \sum_{m=1}^{M/2} \frac{e^{-\frac{c_1^2}{2b_1}}(N_1-1)\sqrt{b_1}\sum_{i=1}^{2}\sum_{j=1}^{2}\Omega^j_{L_i(m)}}{6N_2N_1(a_1+b_1)\bar{q}_1}, \tag{4.18}$$

where, for $i \in \{1,2\}$*:*

$$\chi^1_{L_i(m)} = \frac{e^{\frac{c_1^2}{2b_1+2L_i^2(m)b_1^2}}\left(1+Erf\left[\frac{c_1\sqrt{L_i^2(m)+\frac{1}{b_1}}}{\sqrt{2}\left(1+L_i^2(m)b_1\right)}\right]\right)}{\sqrt{L_i^2(m)+\frac{1}{b_1}}}, \tag{4.19}$$

$$\Xi^1_{L_i(m)} = \frac{3e^{\frac{3c_1^2}{6b_1+8L_i^2(m)b_1^2}}\left(1+Erf\left[\frac{c_1\sqrt{6L_i^2(m)+\frac{9}{2b_1}}}{3+4L_i^2(m)b_1}\right]\right)}{\sqrt{\frac{4L_i^2(m)}{3}+\frac{1}{b_1}}}, \tag{4.20}$$

$$\Omega^1_{L_i(m)} = \frac{1}{\sqrt{L_i^2(m)+\frac{1}{a_1}}}, \Omega^2_{L_i(m)} = \frac{3}{\sqrt{\frac{4L_i^2(m)}{3}+\frac{1}{a_1}}}, \tag{4.21}$$

and

$$\mathcal{I}_{13} = \int_{x=0}^{\infty} \mathcal{P}_1^C(e) f_{3,1}(x)dx, \tag{4.22}$$

where Erf$[\cdot]$ *denotes error function.*

Proof. The proof is given in Appendix C.2. □

Special Case I: When direct path channel t_1 is blocked, then the average SEP at UE_1 simplifies to: $\mathcal{P}_1^e(T) = \mathcal{I}_{11}$.

Theorem 4.2. *In an IRS-aided NOMA system where symbols of* UE_1 *and* UE_2 *are drawn from* N_1*-PAM and* N_2*-PAM constellations, respectively, average SEP of* UE_2 *is given by*

$$\mathcal{P}_2^e(T) \approx \sum_{j=1}^{2}(-1)^{(j+1)}\left[\mathcal{J}_2^{j1} - \left(\mathcal{J}_2^{j2,1}+\mathcal{J}_2^{j2,2}\right) - \mathcal{J}_2^{j3}\right], \tag{4.23}$$

where

$$\mathcal{J}_2^{11} = \frac{e^{-\frac{c_2^2}{2b_2}}(N_2-1)\left(\chi^1_{L_3}+\Xi^1_{L_3}\right)}{3N_2\sqrt{2\pi}\sqrt{b_2}\bar{q}_2}, \tag{4.24}$$

$$\mathcal{J}_2^{12,2} = \frac{e^{-\frac{c_2^2}{2b_2}}(N_2-1)\sqrt{b_2}\left(\frac{1}{\sqrt{L_3^2+\frac{1}{a_2}}}+\frac{3}{\sqrt{\frac{4L_3^2}{3}+\frac{1}{a_2}}}\right)}{(6N_2(a_2+b_2))\,\bar{q}_2}, \tag{4.25}$$

$$\mathcal{J}_2^{21} = \sum_{n=1}^{N_1-1}\frac{e^{-\frac{c_2^2}{2b_2}}(N_1-n)}{3N_2N_1\sqrt{2\pi}\sqrt{b_2}\bar{q}_2}\left[\left(\sum_{j=6}^{7}(-1)^j\left(\chi^2_{L_j(n)}+\Xi^2_{L_j(n)}\right)\right)+\left((M-1)\sum_{j=4}^{5}(-1)^j\left(\chi^2_{L_j(n)}+\Xi^2_{L_j(n)}\right)\right)\right], \tag{4.26}$$

$$\mathcal{J}_2^{22,2} = \sum_{n=1}^{N_1-1}\frac{e^{-\frac{c_2^2}{2b_2}}(N-n)\sqrt{b_2}}{6N_2N_1(a_2+b_2)\bar{q}_2}\left[\left(\sum_{i=3}^{4}\sum_{j=6}^{7}(-1)^j\Omega^i_{L_j(n)}\right)+(M-1)\left(\sum_{i=3}^{4}\sum_{j=4}^{5}(-1)^j\Omega^i_{L_j(n)}\right)\right], \tag{4.27}$$

$$\mathcal{J}_2^{12,1} = \frac{a_2}{a_2+b_2}\mathcal{J}_2^{11}, \mathcal{J}_2^{22,1} = \frac{a_2}{(a_2+b_2)}\mathcal{J}_2^{21}, \tag{4.28}$$

where, for $j \in \{4,5,6,7\}$*:*

$$\chi^1_{L_3} = \frac{e^{\frac{c_2^2}{2b_2+2L_3^2b_2^2}}\sqrt{\frac{\pi}{2}}\left(1+Erf\left[\frac{c_2\sqrt{L_3^2+\frac{1}{b_2}}}{\sqrt{2}\left(1+L_3^2b_2\right)}\right]\right)}{\sqrt{L_3^2+\frac{1}{b_2}}}, \tag{4.29}$$

$$\Xi^1_{L_3} = \frac{3e^{\frac{3c_2^2}{6b_2+8L_3^2b_2^2}}\sqrt{3\pi}\left(1+Erf\left[\frac{c_2\sqrt{6L_3^2+\frac{9}{2b_2}}}{3+4L_3^2b_2}\right]\right)}{\sqrt{8L_3^2+\frac{6}{b_2}}}, \tag{4.30}$$

$$\chi^2_{L_j(n)} = \frac{e^{\frac{c_2^2}{2b_2+2L_j^2(n)b_2^2}}\sqrt{\frac{\pi}{2}}\left(1+Erf\left[\frac{c_2\sqrt{L_j^2(n)+\frac{1}{b_2}}}{\sqrt{2}\left(1+L_j^2(n)b_2\right)}\right]\right)}{\sqrt{L_j^2(n)+\frac{1}{b_2}}}, \tag{4.31}$$

$$\Xi^2_{L_j(n)} = \frac{3e^{\frac{3c_2^2}{6b_2+8L_j^2(n)b_2^2}}\left(1+Erf\left[\frac{c_2\sqrt{6L_j^2(n)+\frac{9}{2b_2}}}{3+4L_j^2(n)b_2}\right]\right)}{\left(1/\sqrt{3\pi}\right)\sqrt{8L_j^2(n)+\frac{6}{b_2}}}, \tag{4.32}$$

$$\Omega^3_{L_j(n)} = \frac{1}{\sqrt{L_j^2(n)+\frac{1}{a_2}}}, \Omega^4_{L_j(n)} = \frac{3}{\sqrt{\frac{4L_j^2(n)}{3}+\frac{1}{a_2}}}, \tag{4.33}$$

and

$$\mathcal{J}_2^{13} = \int_{x=0}^{\infty}\mathcal{P}^C_{2,1}(e)f_{3,2}(x)dx, \tag{4.34}$$

$$\mathcal{J}_2^{23} = \int_{x=0}^{\infty} \mathcal{P}_{2,2}^{C}(e) f_{3,2}(x) dx. \tag{4.35}$$

Proof. The proof is given in Appendix C.3. □

Special Case II: When direct path channel t_2 is blocked, then average SEP at UE_2 simplifies to: $\mathcal{P}_2^e(T) = \mathcal{J}_2^{11} - \mathcal{J}_2^{21}$.

4.2.2 SEP Analysis for Two-User Setup with QAM Symbols

The SEP of UE_1 conditioned on $|z_1|$, when N_1-QAM constellation is used for UE_1 and N_2-QAM constellation is used for UE_2, is given by [89]

$$\mathcal{P}_1^{\text{C-QAM}}(e) \triangleq \mathcal{P}_1^{\text{QAM}}(e||z_1|) = 1 - \left(1 - \frac{2\left(\sqrt{N_1}-1\right)}{\sqrt{N_2 N_1}} \sum_{m=1}^{\sqrt{N_2}/2} \sum_{i=1}^{2} \mathbf{Q}\left\{\widetilde{L}_i(m)\,|z_1|\right\}\right)^2, \tag{4.36}$$

where $\widetilde{L}_1(m) = \frac{(\widetilde{d}_1 - (2m-1)\widetilde{d}_2)}{\sqrt{\sigma_1^2/2}}$, $\widetilde{L}_2(m) = \frac{(\widetilde{d}_1 + (2m-1)\widetilde{d}_2)}{\sqrt{\sigma_1^2/2}}$. Also, $\widetilde{d}_1 = \sqrt{\frac{3P_1}{2(N_1-1)}}$, $\widetilde{d}_2 = \sqrt{\frac{3P_2}{2(N_2-1)}}$ are half the distances between neighboring constellation points of N_1-QAM and N_2-QAM constellations, respectively. Similarly, the SEP of UE_2 conditioned on $|z_2|$ is given by [89].

$$\mathcal{P}_2^{\text{C-QAM}}(e) \triangleq \mathcal{P}_2^{\text{QAM}}(e||z_2|) = 1 - \left(\mathcal{P}_2^{\text{C-QAM}}(c)\right)^2, \tag{4.37}$$

where, $\mathcal{P}_2^{\text{C-QAM}}(c)$ is given by

$$\begin{aligned}\mathcal{P}_2^{\text{C-QAM}}(c) \triangleq\ & 1 - \frac{2\left(\sqrt{N_2}-1\right)}{\sqrt{N_2}} \mathbf{Q}\{L_3\,|z_2|\} \\ & + \sum_{n=1}^{\sqrt{N_1}-1} \frac{2\left(\sqrt{N_1}-n\right)}{\sqrt{N_2 N_1}} \left[\left(\sqrt{N_2}-1\right)\left(\sum_{i=4}^{5} (-1)^i \mathbf{Q}\{L_i(n)\,|z_2|\}\right) + \left(\sum_{i=6}^{7} (-1)^i \mathbf{Q}\{L_i(n)\,|z_2|\}\right)\right]. \end{aligned} \tag{4.38}$$

Therein, $\widetilde{L}_3 = \frac{\widetilde{d}_2}{\sqrt{\sigma_2^2/2}}$, $\widetilde{L}_4(n) = \frac{(2n\widetilde{d}_1 - \widetilde{d}_2)}{\sqrt{\sigma_2^2/2}}$, $\widetilde{L}_5(n) = \frac{(2n\widetilde{d}_1 + \widetilde{d}_2)}{\sqrt{\sigma_2^2/2}}$, $\widetilde{L}_6(n) = \frac{((2n-1)\widetilde{d}_1 + (N_2-1)\widetilde{d}_2)}{\sqrt{\sigma_2^2/2}}$, $\widetilde{L}_7(n) = \frac{((2n-1)\widetilde{d}_1 - (N_2-1)\widetilde{d}_2)}{\sqrt{\sigma_2^2/2}}$.

As it can be observed in (4.36), (4.37) and (4.38), the conditional SEP expressions for UE_1 and UE_2 involve the square of a summation term. The summations are over modulation orders N_1 and N_2. And the square of the summation cannot be expanded without fixing the values of N_1 and N_2. Therefore, it may not be possible to derive corresponding closed-form average SEP expressions. However, given the values of N_1 and N_2, one can expand the respective summations

in (4.36) and (4.37). Thereafter, the conditional SEP expressions with expanded summations can be used to derive the corresponding average SEP expressions by taking their expectation over the PDF of the effective channel gain. As an example, we consider $N_1 = 4$ and $N_2 = 16$ and expand the summation terms in $\mathcal{P}_1^{\text{C-QAM}}(e)$ and $\mathcal{P}_2^{\text{C-QAM}}(e)$. For $N_1 = 4$ and $N_2 = 16$,

$$\mathcal{P}_1^{\text{C-QAM}}(e) = 2\mathcal{A}_1 - (\mathcal{A}_1)^2, \tag{4.39}$$

where $\mathcal{A}_1 = \frac{1}{4}\sum_{m=1}^{2}\left(\mathbf{Q}\left\{\widetilde{L}_1|z_1|\right\} + \mathbf{Q}\left\{\widetilde{L}_2|z_1|\right\}\right)$. Similarly,

$$\mathcal{P}_2^{\text{C-QAM}}(e) = 2\mathcal{A}_2 - 2\mathcal{B}_2 + 2\mathcal{A}_2\mathcal{B}_2 - (\mathcal{A}_2)^2 - (\mathcal{B}_2)^2, \tag{4.40}$$

where $\mathcal{B}_2 = \frac{1}{4}\left(2\sum_{i=4}^{5}(-1)^i\mathbf{Q}\left\{\widetilde{L}_i|z_2|\right\} + \sum_{j=6}^{7}\mathbf{Q}\left\{\widetilde{L}_j|z_2|\right\}\right)$, $\mathcal{A}_2 = \frac{3}{2}\mathbf{Q}\left\{\widetilde{L}_3|z_2|\right\}$. We use these conditional SEP expressions for $N_1 = 4$ and $N_2 = 16$ to derive corresponding average SEP expressions in the following theorems.

Theorem 4.3. *In an IRS-aided NOMA system where symbols of UE_1 and UE_2 are drawn from 4-QAM and 16-QAM constellations, respectively, the average SEP of UE_1 is given by*

$$\mathcal{P}_1^{e,QAM}(T) \approx \mathcal{I}_{11}^{QAM} - \left(\mathcal{I}_{12,1}^{QAM} + \mathcal{I}_{12,2}^{QAM}\right) - \mathcal{I}_{13}^{QAM}, \tag{4.41}$$

where

$$\mathcal{I}_{12,1}^{QAM} = \frac{a_1}{a_1 + b_1}\mathcal{I}_{11}^{QAM}, \tag{4.42}$$

$$\mathcal{I}_{13}^{QAM} = \int_{x=0}^{\infty}\mathcal{P}_1^{C\text{-}QAM}(e)f_{3,1}(x)dx, \tag{4.43}$$

and $\mathcal{I}_{11}^{QAM}$ and $\mathcal{I}_{12,2}^{QAM}$ are given by

$$\begin{aligned}\mathcal{I}_{11}^{QAM} = \frac{e^{-\frac{c_1^2}{2b_1}}}{1152\sqrt{2\pi}\sqrt{b_1}\bar{q}_1}\Bigg\{&\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Gamma}_1\left(\widetilde{L}_i(j)\right)\right] - \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Gamma}_2\left(\widetilde{L}_i(j)\right)\right] + \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Gamma}_3\left(\widetilde{L}_i(j)\right)\right] \\ &- \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Gamma}_4\left(\widetilde{L}_i(j)\right)\right] - \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Gamma}_5\left(\widetilde{L}_i(j)\right)\right] \\ &- \sum_{\bar{k}=6}^{9}\left(\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Gamma}_{\bar{k}}\left(\widetilde{L}_2(i),\widetilde{L}_1(j)\right)\right] + \left[\overline{\Gamma}_{\bar{k}}\left(\widetilde{L}_2(2),\widetilde{L}_2(1)\right) + \overline{\Gamma}_{\bar{k}}\left(\widetilde{L}_1(2),\widetilde{L}_1(1)\right)\right]\right)\Bigg\},\end{aligned} \tag{4.44}$$

$$\mathcal{I}_{12,2}^{QAM} = \frac{e^{-\frac{c_1^2}{2b_1}}\sqrt{b_1}}{1152\sqrt{2\pi}(a_1+b_1)\bar{q}_1}\left\{\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Delta}_1\left(\widetilde{L}_i(j)\right)\right] - \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Delta}_2\left(\widetilde{L}_i(j)\right)\right] + \left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Delta}_3\left(\widetilde{L}_i(j)\right)\right]\right.$$

$$-\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Delta}_4\left(\widetilde{L}_i(j)\right)\right]-\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Delta}_5\left(\widetilde{L}_i(j)\right)\right]$$

$$-\sum_{\bar{k}=6}^{9}\left(\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\overline{\Delta}_{\bar{k}}\left(\widetilde{L}_2(i),\widetilde{L}_1(j)\right)\right]+\left[\overline{\Delta}_{\bar{k}}\left(\widetilde{L}_2(2),\widetilde{L}_2(1)\right)+\overline{\Delta}_{\bar{k}}\left(\widetilde{L}_1(2),\widetilde{L}_1(1)\right)\right]\right)\Bigg\}, \quad (4.45)$$

where

$$\overline{\Gamma}_1\left(\widetilde{L}_i(j)\right)=\frac{48e^{\frac{c_1^2}{2b_1+2\widetilde{L}_i(j)^2b_1^2}}\sqrt{2\pi}\left(1+Erf\left[\frac{c_1\sqrt{\widetilde{L}_i(j)^2+\frac{1}{b_1}}}{\sqrt{2}\left(1+\widetilde{L}_i(j)^2b_1\right)}\right]\right)}{\sqrt{\widetilde{L}_i(j)^2+\frac{1}{b_1}}}, \quad (4.46)$$

$$\overline{\Gamma}_2\left(\widetilde{L}_i(j)\right)=\frac{e^{\frac{c_1^2}{2b_1+4\widetilde{L}_i(j)^2b_1^2}}\sqrt{\frac{\pi}{2}}\left(1+Erf\left[\frac{c_1\sqrt{\widetilde{L}_i(j)^2+\frac{1}{2b_1}}}{1+2\widetilde{L}_i(j)^2b_1}\right]\right)}{\sqrt{2\widetilde{L}_i(j)^2+\frac{1}{b_1}}}, \quad (4.47)$$

$$\overline{\Gamma}_3\left(\widetilde{L}_i(j)\right)=\frac{288e^{\frac{3c_1^2}{6b_1+8\widetilde{L}_i(j)^2b_1^2}}\sqrt{3\pi}\left(1+Erf\left[\frac{c_1\sqrt{6\widetilde{L}_i(j)^2+\frac{9}{2b_1}}}{3+4\widetilde{L}_i(j)^2b_1}\right]\right)}{\sqrt{8\widetilde{L}_i(j)^2+\frac{6}{b_1}}}, \quad (4.48)$$

$$\overline{\Gamma}_4\left(\widetilde{L}_i(j)\right)=\frac{3e^{\frac{3c_1^2}{6b_1+14\widetilde{L}_i(j)^2b_1^2}}\sqrt{6\pi}\left(1+Erf\left[\frac{\sqrt{\frac{3}{2}}c_1\sqrt{7\widetilde{L}_i(j)^2+\frac{3}{b_1}}}{3+7\widetilde{L}_i(j)^2b_1}\right]\right)}{\sqrt{7\widetilde{L}_i(j)^2+\frac{3}{b_1}}}, \quad (4.49)$$

$$\overline{\Gamma}_5\left(\widetilde{L}_i(j)\right)=\frac{9e^{\frac{3c_1^2}{6b_1+16\widetilde{L}_i(j)^2b_1^2}}\sqrt{\frac{3\pi}{2}}\left(1+Erf\left[\frac{c_1\sqrt{12\widetilde{L}_i(j)^2+\frac{9}{2b_1}}}{3+8\widetilde{L}_i(j)^2b_1}\right]\right)}{\sqrt{8\widetilde{L}_i(j)^2+\frac{3}{b_1}}}, \quad (4.50)$$

$$\overline{\Gamma}_6\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{e^{\frac{c_1^2}{2b_1\left(1+\widetilde{L}_i(j)^2b_1+\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1\right)}}\sqrt{2\pi}\left(1+Erf\left[\frac{c_1\sqrt{\widetilde{L}_i(j)^2+\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{1}{b_1}}}{\sqrt{2}\left(1+\widetilde{L}_i(j)^2b_1+\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1\right)}\right]\right)}{\sqrt{\widetilde{L}_i(j)^2+\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{1}{b_1}}}, \quad (4.51)$$

$$\overline{\Gamma}_7\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{3e^{\frac{3c_1^2}{2b_1\left(3+4\widetilde{L}_i(j)^2b_1+3\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1\right)}}\sqrt{6\pi}\left(1+Erf\left[\frac{\sqrt{\frac{3}{2}}c_1\sqrt{4\widetilde{L}_i(j)^2+3\left(\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{1}{b_1}\right)}}{3+4\widetilde{L}_i(j)^2b_1+3\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1}\right]\right)}{\sqrt{4\widetilde{L}_i(j)^2+3\left(\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{1}{b_1}\right)}}, \quad (4.52)$$

$$\overline{\Gamma}_8\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{3e^{\frac{3c_1^2}{2b_1\left(3+3\widetilde{L}_i(j)^2b_1+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1\right)}}\sqrt{6\pi}\left(1+Erf\left[\frac{\sqrt{\frac{3}{2}}c_1\sqrt{3\widetilde{L}_i(j)^2+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{3}{b_1}}}{3+3\widetilde{L}_i(j)^2b_1+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1}\right]\right)}{\sqrt{3\widetilde{L}_i(j)^2+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{3}{b_1}}}, \quad (4.53)$$

$$\overline{\Gamma}_9\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{9e^{\frac{3c_1^2}{2b_1\left(3+4\widetilde{L}_i(j)^2b_1+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1\right)}}\sqrt{6\pi}\left(1+Erf\left[\frac{c_1\sqrt{6\widetilde{L}_i(j)^2+6\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{9}{2b_1}}}{3+4\widetilde{L}_i(j)^2b_1+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2b_1}\right]\right)}{\sqrt{4\widetilde{L}_i(j)^2+4\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{3}{b_1}}}, \tag{4.54}$$

$$\overline{\Delta}_1\left(\widetilde{L}_i(j)\right)=\frac{48\sqrt{2\pi}}{\sqrt{\widetilde{L}_i(j)^2+\frac{1}{a_1}}}, \tag{4.55}$$

$$\overline{\Delta}_2\left(\widetilde{L}_i(j)\right)=\frac{\sqrt{\pi}}{\sqrt{4\widetilde{L}_i(j)^2+\frac{2}{a_1}}}, \tag{4.56}$$

$$\overline{\Delta}_3\left(\widetilde{L}_i(j)\right)=\frac{288\sqrt{3\pi}}{\sqrt{8\widetilde{L}_i(j)^2+\frac{6}{a_1}}}, \tag{4.57}$$

$$\overline{\Delta}_4\left(\widetilde{L}_i(j)\right)=\frac{3\sqrt{6\pi}}{\sqrt{7\widetilde{L}_i(j)^2+\frac{3}{a_1}}}, \tag{4.58}$$

$$\overline{\Delta}_5\left(\widetilde{L}_i(j)\right)=\frac{9\sqrt{\frac{3\pi}{2}}}{\sqrt{8\widetilde{L}_i(j)^2+\frac{3}{a_1}}}, \tag{4.59}$$

$$\overline{\Delta}_6\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{\sqrt{2\pi}}{\sqrt{\widetilde{L}_i(j)^2+\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{1}{a_1}}}, \tag{4.60}$$

$$\overline{\Delta}_7\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{6\sqrt{3\pi}}{\sqrt{8\widetilde{L}_i(j)^2+6\left(\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{1}{a_1}\right)}}, \tag{4.61}$$

$$\overline{\Delta}_8\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{6\sqrt{3\pi}}{\sqrt{6\widetilde{L}_i(j)^2+8\left(\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{6}{a_1}\right)}}, \tag{4.62}$$

$$\overline{\Delta}_9\left(\widetilde{L}_i(j),\widetilde{L}_{\bar{i}}\left(\bar{j}\right)\right)=\frac{18\sqrt{3\pi}}{\sqrt{8\widetilde{L}_i(j)^2+8\left(\widetilde{L}_{\bar{i}}\left(\bar{j}\right)^2+\frac{6}{a_1}\right)}} \tag{4.63}$$

Proof. The proof is given in Appendix C.4. □

Theorem 4.4. *In an IRS-aided NOMA system where symbols of UE_1 and UE_2 are drawn from 4-QAM and 16-QAM constellations, respectively, average SEP of UE_2 is given by*

$$\mathcal{P}_2^{e,QAM}(T)\approx\mathcal{J}_{11}^{QAM}-\left(\mathcal{J}_{12,1}^{QAM}+\mathcal{J}_{12,2}^{QAM}\right)-\mathcal{J}_{13}^{QAM}, \tag{4.64}$$

where

$$\mathcal{J}_{11}^{QAM}=48\mathcal{T}_1+288\mathcal{T}_2-\mathcal{T}_3-9\mathcal{T}_4-3\mathcal{T}_5+\mathcal{T}_6+3\mathcal{T}_7+3\mathcal{T}_8+9\mathcal{T}_9, \tag{4.65}$$

$$\mathcal{J}_{12,2}^{QAM} = 48\mathcal{U}_1 + 288\mathcal{U}_2 + \mathcal{U}_3 + 3\mathcal{U}_4 + \mathcal{U}_5 + 6\mathcal{U}_6 + \mathcal{U}_7 + 6\mathcal{U}_8 + 18\mathcal{U}_9, \tag{4.66}$$

$$\mathcal{J}_{12,1}^{QAM} = \frac{a_2}{a_2 + b_2}\mathcal{J}_{11}^{QAM}, \tag{4.67}$$

$$\mathcal{J}_{13}^{QAM} = \int_{x=0}^{\infty} \mathcal{P}_2^{C\text{-}QAM}(e) f_{3,2}(x) dx, \tag{4.68}$$

and $\mathcal{T}_{\bar{k}}$ *and* $\mathcal{U}_{\bar{k}}$*, for all* $\bar{k}$*, are given by*

$$\mathcal{T}_{\bar{k}} = \left[6\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1)\right) - 2\Upsilon^{\bar{k}}\left(\widetilde{L}_4(1)\right) + 2\Upsilon^{\bar{k}}\left(\widetilde{L}_5(1)\right) - \Upsilon^{\bar{k}}\left(\widetilde{L}_6(1)\right) + \Upsilon^{\bar{k}}\left(\widetilde{L}_7(1)\right)\right], \text{ for } \bar{k} = 1, 2 \tag{4.69}$$

$$= \left[18\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1)\right) + 2\Upsilon^{\bar{k}}\left(\widetilde{L}_4(1)\right) + 2\Upsilon^{\bar{k}}\left(\widetilde{L}_5(1)\right) + \Upsilon^{\bar{k}}\left(\widetilde{L}_6(1)\right) + \Upsilon^{\bar{k}}\left(\widetilde{L}_7(1)\right)\right], \text{ for } \bar{k} = 3, 4 \tag{4.70}$$

$$= \left[36\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1)\right) + 4\Upsilon^{\bar{k}}\left(\widetilde{L}_4(1)\right) + 4\Upsilon^{\bar{k}}\left(\widetilde{L}_5(1)\right) + \Upsilon^{\bar{k}}\left(\widetilde{L}_6(1)\right) + \Upsilon^{\bar{k}}\left(\widetilde{L}_7(1)\right)\right], \text{ for } \bar{k} = 5 \tag{4.71}$$

$$= 12\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_4(1)\right) - 12\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_5(1)\right) + 6\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_6(1)\right) - 6\Upsilon^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_7(1)\right) \tag{4.72}$$

$$+ 4\Upsilon^{\bar{k}}\left(\widetilde{L}_4(1), \widetilde{L}_4(5)\right) - 2\Upsilon^{\bar{k}}\left(\widetilde{L}_4(1), \widetilde{L}_4(6)\right) + 2\Upsilon^{\bar{k}}\left(\widetilde{L}_4(1), \widetilde{L}_4(7)\right) + 2\Upsilon^{\bar{k}}\left(\widetilde{L}_5(1), \widetilde{L}_4(6)\right)$$

$$- 2\Upsilon^{\bar{k}}\left(\widetilde{L}_5(1), \widetilde{L}_4(7)\right) + \Upsilon^{\bar{k}}\left(\widetilde{L}_6(1), \widetilde{L}_4(7)\right), \text{ for } \bar{k} = 6, 7, 8, 9, \tag{4.73}$$

$$\mathcal{U}_{\bar{k}} = -\left[6\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1)\right) - 2\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_4(1)\right) + 2\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_5(1)\right) - \overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_6(1)\right) + \overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_7(1)\right)\right], \text{ for } \bar{k} = 1, 2 \tag{4.74}$$

$$= \left[36\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1)\right) + 4\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_4(1)\right) + 4\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_5(1)\right) + \overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_6(1)\right) + \overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_7(1)\right)\right], \text{ for } \bar{k} = 3, 4 \tag{4.75}$$

$$= \left[162\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1)\right) + 18\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_4(1)\right) + 18\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_5(1)\right) + \frac{9}{16}\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_6(1)\right) + \frac{9}{16}\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_7(1)\right)\right], \text{ for } \bar{k} = 5 \tag{4.76}$$

$$= -12\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_4(1)\right) + 12\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_5(1)\right) - 6\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_6(1)\right) + 6\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_3(1), \widetilde{L}_7(1)\right)$$

$$- 4\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_4(1), \widetilde{L}_4(5)\right) + 2\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_4(1), \widetilde{L}_4(6)\right) - 2\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_4(1), \widetilde{L}_4(7)\right) - 2\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_5(1), \widetilde{L}_4(6)\right)$$

$$+ 2\overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_5(1), \widetilde{L}_4(7)\right) - \overline{\Upsilon}^{\bar{k}}\left(\widetilde{L}_6(1), \widetilde{L}_4(7)\right), \text{ for } \bar{k} = 6, 7, 8, 9, \tag{4.77}$$

where, for $i, j = 1, 2, 3, 4, 5$*:*

$$\Upsilon^1\left(\widetilde{L}_i(1)\right) = \frac{e^{\frac{c_2^2}{2b_2 + 2\widetilde{L}_i(1)^2 b_2^2}}\sqrt{2\pi}\left(1 + Erf\left[\frac{c_2\sqrt{\widetilde{L}_i(1)^2 + \frac{1}{b_2}}}{\sqrt{2}\left(1 + \widetilde{L}_i(1)^2 b_2\right)}\right]\right)}{\sqrt{\widetilde{L}_i(1)^2 + \frac{1}{b_2}}}, \tag{4.78}$$

$$\Upsilon^2\left(\widetilde{L}_i(1)\right) = \frac{e^{\frac{3c_2^2}{6b_2 + 8\widetilde{L}_i(1)^2 b_2^2}}\sqrt{3\pi}\left(1 + Erf\left[\frac{c_2\sqrt{6\widetilde{L}_i(1)^2 + \frac{9}{2b_2}}}{3 + 4\widetilde{L}_i(1)^2 b_2}\right]\right)}{\sqrt{8\widetilde{L}_i(1)^2 + \frac{6}{b_2}}}, \tag{4.79}$$

$$\Upsilon^3\left(\widetilde{L}_i(1)\right) = \frac{e^{\frac{c_2^2}{2b_2+4\widetilde{L}_i(1)^2b_2^2}}\sqrt{2\pi}\left(1+Erf\left[\frac{c_2\sqrt{\widetilde{L}_i(1)^2+\frac{1}{2b_2}}}{1+2\widetilde{L}_i(1)^2b_2}\right]\right)}{\sqrt{2\widetilde{L}_i(1)^2+\frac{1}{b_2}}}, \tag{4.80}$$

$$\Upsilon^4\left(\widetilde{L}_i(1)\right) = \frac{e^{\frac{3c_2^2}{6b_2+16\widetilde{L}_i(1)^2b_2^2}}\sqrt{6\pi}\left(1+Erf\left[\frac{c_2\sqrt{12\widetilde{L}_i(1)^2+\frac{9}{2b_2}}}{3+8\widetilde{L}_i(1)^2b_2}\right]\right)}{\sqrt{8\widetilde{L}_i(1)^2+\frac{i}{b_2}}}, \tag{4.81}$$

$$\Upsilon^5\left(\widetilde{L}_i(1)\right) = \frac{e^{\frac{3c_2^2}{6b_2+14\widetilde{L}_i(1)^2b_2^2}}\sqrt{6\pi}\left(1+Erf\left[\frac{\sqrt{\frac{i}{2}}c_2\sqrt{7\widetilde{L}_i(1)^2+\frac{i}{b_2}}}{3+7\widetilde{L}_i(1)^2b_2}\right]\right)}{\sqrt{7\widetilde{L}_i(1)^2+\frac{i}{b_2}}}, \tag{4.82}$$

$$\Upsilon^6\left(\widetilde{L}_i(1),\widetilde{L}_j(1)\right) = \frac{e^{\frac{c_2^2}{2b_2\left(1+\widetilde{L}_i(1)^2b_2+\widetilde{L}_j(1)^2b_2\right)}}\sqrt{2\pi}\left(1+Erf\left[\frac{c_2\sqrt{\widetilde{L}_i(1)^2+\widetilde{L}_j(1)^2+\frac{1}{b_2}}}{\sqrt{2}\left(1+\widetilde{L}_i(1)^2b_2+\widetilde{L}_j(1)^2b_2\right)}\right]\right)}{\sqrt{\widetilde{L}_i(1)^2+\widetilde{L}_j(1)^2+\frac{1}{b_2}}}, \tag{4.83}$$

$$\Upsilon^7\left(\widetilde{L}_i(1),\widetilde{L}_j(1)\right) = \frac{e^{\frac{3c_2^2}{2b_2\left(3+4\widetilde{L}_i(1)^2b_2+3\widetilde{L}_j(1)^2b_2\right)}}\sqrt{6\pi}\left(1+Erf\left[\frac{\sqrt{\frac{3}{2}}c_2\sqrt{4\widetilde{L}_i(1)^2+3\left(\widetilde{L}_j(1)^2+\frac{1}{b_2}\right)}}{3+4\widetilde{L}_i(1)^2b_2+3\widetilde{L}_j(1)^2b_2}\right]\right)}{\sqrt{4\widetilde{L}_i(1)^2+3\left(\widetilde{L}_j(1)^2+\frac{1}{b_2}\right)}}, \tag{4.84}$$

$$\Upsilon^8\left(\widetilde{L}_i(1),\widetilde{L}_j(1)\right) = \frac{e^{\frac{3c_2^2}{2b_2\left(3+3\widetilde{L}_i(1)^2b_2+4\widetilde{L}_j(1)^2b_2\right)}}\sqrt{6\pi}\left(1+Erf\left[\frac{\sqrt{\frac{3}{2}}c_2\sqrt{3\widetilde{L}_i(1)^2+4\widetilde{L}_j(1)^2+\frac{3}{b_2}}}{3+3\widetilde{L}_i(1)^2b_2+4\widetilde{L}_j(1)^2b_2}\right]\right)}{\sqrt{3\widetilde{L}_i(1)^2+4\widetilde{L}_j(1)^2+\frac{3}{b_2}}}, \tag{4.85}$$

$$\Upsilon^9\left(\widetilde{L}_i(1),\widetilde{L}_j(1)\right) = \frac{e^{\frac{3c_2^2}{2b_2\left(3+4\widetilde{L}_i(1)^2b_2+4\widetilde{L}_j(1)^2b_2\right)}}\sqrt{6\pi}\left(1+Erf\left[\frac{c_2\sqrt{6\widetilde{L}_i(1)^2+6\widetilde{L}_j(1)^2+\frac{9}{2b_2}}}{3+4\widetilde{L}_i(1)^2b_2+4\widetilde{L}_j(1)^2b_2}\right]\right)}{\sqrt{4\widetilde{L}_i(1)^2+4\widetilde{L}_j(1)^2+\frac{3}{b_2}}}, \tag{4.86}$$

$$\overline{\Upsilon}^1\left(\widetilde{L}_i(1)\right) = \frac{\sqrt{2\pi}}{\sqrt{\widetilde{L}_i(1)^2+\frac{1}{a_2}}}, \tag{4.87}$$

$$\overline{\Upsilon}^2\left(\widetilde{L}_i(1)\right) = \frac{\sqrt{3\pi}}{\sqrt{8\widetilde{L}_i(1)^2+\frac{6}{a_2}}}, \tag{4.88}$$

$$\overline{\Upsilon}^3\left(\widetilde{L}_i(1)\right) = \frac{\sqrt{\pi}}{\sqrt{4\widetilde{L}_i^2+\frac{2}{a_2}}}, \tag{4.89}$$

$$\overline{\gamma}^4\left(\widetilde{L}_i(1)\right) = \frac{\sqrt{6\pi}}{\sqrt{7\widetilde{L}_i(1)^2 + \frac{3}{a_2}}}, \tag{4.90}$$

$$\overline{\gamma}^5\left(\widetilde{L}_i(1)\right) = \frac{\sqrt{6\pi}}{\sqrt{8\widetilde{L}_i(1)^2 + \frac{3}{a_2}}}, \tag{4.91}$$

$$\overline{\gamma}^6\left(\widetilde{L}_i(1), \widetilde{L}_j(1)\right) = \frac{\sqrt{3\pi}}{\sqrt{8\widetilde{L}_i(1)^2 + 6\left(\widetilde{L}_j(1)^2 + \frac{1}{a_2}\right)}}, \tag{4.92}$$

$$\overline{\gamma}^7\left(\widetilde{L}_i(1), \widetilde{L}_j(1)\right) = \frac{\sqrt{2\pi}}{\sqrt{\widetilde{L}_i(1)^2 + \widetilde{L}_j(1)^2 + \frac{1}{a_2}}}, \tag{4.93}$$

$$\overline{\gamma}^8\left(\widetilde{L}_i(1), \widetilde{L}_j(1)\right) = \frac{\sqrt{3\pi}}{\sqrt{6\widetilde{L}_i(1)^2 + 8\widetilde{L}_j(1)^2 + \frac{6}{a_2}}}, \tag{4.94}$$

$$\overline{\gamma}^9\left(\widetilde{L}_i(1), \widetilde{L}_j(1)\right) = \frac{\sqrt{3\pi}}{\sqrt{8\widetilde{L}_i(1)^2 + 8\widetilde{L}_j(1)^2 + \frac{6}{a_2}}}. \tag{4.95}$$

Proof. The proof is given in Appendix C.5. □

Insights: The expressions given in Theorems 1-4 bring out the dependence of the average SEP on system and channel parameters like transmit power (P), PAM and QAM constellation sizes (N_1 and N_2), path-loss values, Rice-factor (K), noise variance, number of IRS elements (T), NOMA power allocation coefficient (λ), and distance between adjacent constellation points (d_1, d_2, $\widetilde{d}_1$, $\widetilde{d}_2$).

4.2.3 Error Probability for Multi-User Setup with PAM

Now, we extend the error performance analysis to a multi-user scenario with K users. The system model is given in Fig. 4.2. We discuss two separate cases. In Section 4.2.3.1, we discuss the SEP analysis of a multi-user system with multicast transmission. Herein, we consider that the transmitted signal for all strong users is the same and similarly, the transmitted signal for all weak users is the same. In Section 4.2.3.2, we discuss the PEP analysis of a multi-user setup with unicast transmission.[5] Therein, we consider only one group of all users to employ NOMA.

[5]Unicast transmission refers to the scenario when different information symbols (messages) are to be transmitted to different users [90].

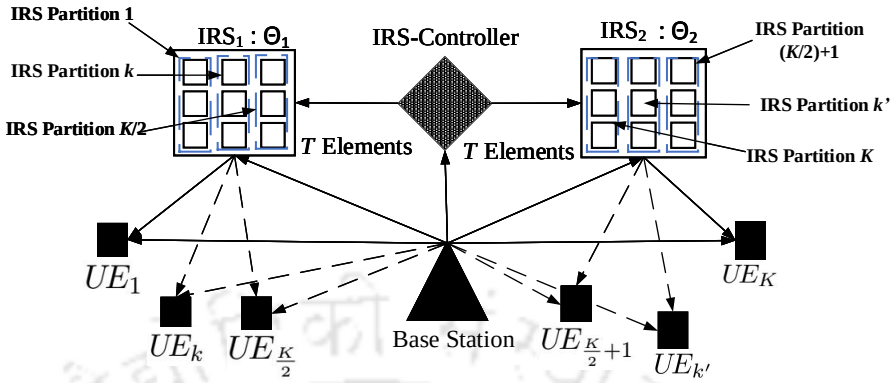


Fig. 4.2: System model: IRS-aided NOMA, multi-user case.

4.2.3.1 SEP Analysis for Multicast Transmission

Multicast transmission refers to the scenario when a common information symbol (message) is to be transmitted to a set of users. According to 3GPP standards Release 17, multicast and broadcast applications are very important use cases of 5G services [91]. Therefore, in order to extend the SEP analysis to multi-user case, we consider a multicast application and the corresponding system model is shown in Fig. 4.2. Here we consider $\frac{K}{2}$ pairs of users that will be simultaneously served on the downlink using two IRSs. IRS$_1$ will serve all $\frac{K}{2}$ weak users and IRS$_2$ will serve all $\frac{K}{2}$ strong users. In fact, we consider that IRS$_1$ will be partitioned into $\frac{K}{2}$ parts and phase of the k^{th} part will be configured to beamform the reflected signal to UE_k, for $k \in \mathcal{K}^{\text{W}} = \left\{1, \cdots, \frac{K}{2}\right\}$. Similarly, IRS$_2$ will be partitioned into $\frac{K}{2}$ parts and phase of the $(k')^{\text{th}}$ part will be configured to beamform the reflected signal to $UE_{k'}$, for $k' \in \mathcal{K}^{\text{S}} = \left\{\frac{K}{2}+1, \cdots, K\right\}$. Let $z_{\widetilde{k}} \in \mathbb{C}$ denote the effective channel coefficient from the BS to $UE_{\widetilde{k}}$, for $\widetilde{k} \in \mathcal{K}^{\text{W}} \cup \mathcal{K}^{\text{S}}$, where $\cup$ denotes the union operator. Note that $z_{\widetilde{k}}$ is the same as z_k given in (4.1) after replacement of T by $\frac{2T}{K}$. Therefore, the PDF $f_{|z_{\widetilde{k}}|}(x)$ of $z_{\widetilde{k}}$ is the same as the PDF $f_{|z_k|}(x)$ of $z_{\widetilde{k}}$ given in Lemma 4.1 after the replacement of T by $\frac{2T}{K}$.

Under multicast transmission, we consider that the transmitted signal for all strong users is the same and similarly, the transmitted signal for all weak users is the same. Moreover, each strong user is paired with a weak user for NOMA transmission. Specifically, UE_k and $UE_{k'}$ denote the weak and strong users, for $k \in \mathcal{K}^{\text{W}}$ and $k' \in \mathcal{K}^{\text{S}}$. The symbol s_1 is intended for each weak user UE_k and symbol s_2 is intended for each strong user $UE_{k'}$, for $k \in \mathcal{K}^{\text{W}}$ and $k' \in \mathcal{K}^{\text{S}}$.

The superposition coded symbol transmitted by the BS is given by $s = \sqrt{P_1}s_1 + \sqrt{P_2}s_2$, where s_1 and s_2 are symbols with unit average power drawn from N_1-PAM and N_2-PAM constellations, respectively.

With this setting, the signal received at $UE_{\widetilde{k}}$ is given by

$$y_{\widetilde{k}} = z_{\widetilde{k}}\, s + w_{\widetilde{k}} = \left|z_{\widetilde{k}}\right| e^{j\theta_{z_{\widetilde{k}}}}\, s + w_{\widetilde{k}} = \left|z_{\widetilde{k}}\right| e^{j\theta_{t_{\widetilde{k}}}}\, s + w_{\widetilde{k}}, \tag{4.96}$$

for $\widetilde{k} \in \mathcal{K}^{\mathrm{W}} \cup \mathcal{K}^{\mathrm{S}}$, where $w_{\widetilde{k}} \sim \mathcal{CN}\left(0, \sigma^2_{\widetilde{k}}\right)$ is the AWGN at $UE_{\widetilde{k}}$. We assume that the receiver $UE_{\widetilde{k}}$ has phase information of the effective channel and does phase compensation before processing the received signal. Therefore, after phase compensation, the received signal at $UE_{\widetilde{k}}$ reduces to

$$\widetilde{y_{\widetilde{k}}} = y_{\widetilde{k}}\, e^{-j\theta_{t_{\widetilde{k}}}} = \left|z_{\widetilde{k}}\right|\, s + \widetilde{w_{\widetilde{k}}}\,, \tag{4.97}$$

where $\widetilde{w_{\widetilde{k}}} = w_{\widetilde{k}} e^{-j\theta_{t_{\widetilde{k}}}}$ has same distribution as $w_{\widetilde{k}}$ owing to circular symmetry property.

For $k \in \mathcal{K}^{\mathrm{W}}$ and $k' \in \mathcal{K}^{\mathrm{S}}$, the SEP of UE_k conditioned on $|z_k|$, when N_1-PAM constellation is used for UE_k and N_2-PAM constellation is used for $UE_{k'}$, is given by [89]

$$\mathcal{P}^{\mathrm{C}}_k(e) \triangleq \mathcal{P}_k(e|\,|z_k|) = \frac{2\,(N_1-1)}{N_2 N_1} \sum_{m=1}^{N_2/2} \sum_{i=1}^{2} \mathbf{Q}\left\{L_i(m)\,|z_k|\right\}, \tag{4.98}$$

where $L_1(m) = \frac{(d_1-(2m-1)d_2)}{\sqrt{\sigma_1^2/2}}$, $L_2(m) = \frac{(d_1+(2m-1)d_2)}{\sqrt{\sigma_1^2/2}}$. Similarly, the SEP of $UE_{k'}$ conditioned on $|z_{k'}|$ is given by

$$\mathcal{P}^{\mathrm{C}}_{k'}(e) \triangleq \mathcal{P}_{k'}(e|\,|z_{k'}|) = \mathcal{P}^{\mathrm{C}}_{k',1}(e) - \mathcal{P}^{\mathrm{C}}_{k',2}(e)\,, \tag{4.99}$$

where

$$\mathcal{P}^{\mathrm{C}}_{k',1}(e) = \frac{2\,(N_2-1)}{N_2}\mathbf{Q}\left\{L_3\,|z_{k'}|\right\}\,, \tag{4.100}$$

$$\mathcal{P}^{\mathrm{C}}_{k',2}(e) = \sum_{n=1}^{N_1-1} \frac{2\,(N_1-n)}{N_2 N_1} \left[(N_2-1)\left(\sum_{i=4}^{5}(-1)^i \mathbf{Q}\left\{L_i(n)\,|z_{k'}|\right\}\right) + \left(\sum_{i=6}^{7}(-1)^i \mathbf{Q}\left\{L_i(n)\,|z_{k'}|\right\}\right)\right], \tag{4.101}$$

with $L_3 = \frac{d_2}{\sqrt{\sigma_2^2/2}}$, $L_4(n) = \frac{(2nd_1-d_2)}{\sqrt{\sigma_2^2/2}}$, $L_5(n) = \frac{(2nd_1+d_2)}{\sqrt{\sigma_2^2/2}}$, $L_6(n) = \frac{((2n-1)d_1+(N_2-1)d_2)}{\sqrt{\sigma_2^2/2}}$, $L_7(n) = \frac{((2n-1)d_1-(N_2-1)d_2)}{\sqrt{\sigma_2^2/2}}$. Here, $\mathcal{P}^{\mathrm{C}}_{k',1}(e)$ and $\mathcal{P}^{\mathrm{C}}_{k',2}(e)$ are the first and the seconds terms of (4.99), respectively [89]. We next state the average SEP of UE_k and $UE_{k'}$ that employ NOMA decoding protocol.

Corollary 4.1. *In a multi-user IRS-aided NOMA system of K users where symbols of UE_k and $UE_{k'}$, for $k \in \mathcal{K}^W$ and $k' \in \mathcal{K}^S$, are drawn from N_1-PAM and N_2-PAM constellations, respectively, the average SEP of UE_k under multicast transmission can be obtained by replacing T with $\frac{2T}{K}$ in Theorem 4.1.*

Corollary 4.2. *In a multi-user IRS-aided NOMA system of K users where symbols of UE_k and $UE_{k'}$, for $k \in \mathcal{K}^W$ and $k' \in \mathcal{K}^S$, are drawn from N_1-PAM and N_2-PAM constellations, respectively, the average SEP of $UE_{k'}$ under multicast transmission can be obtained by replacing T with $\frac{2T}{K}$ in Theorem 4.2.*

4.2.3.2 PEP Analysis for Unicast Transmission

Now we derive the PEP for multi-user setup. The corresponding system model is shown in Fig. 4.2. However, here we consider unicast transmission and consider that all K users form a single group for employing NOMA. Therefore, the BS transmits the symbol s given by

$$s = \sum_{k=1}^{K} \sqrt{P_k} s_k, \tag{4.102}$$

where s_k is the information symbol intended for UE_k and P_k is the corresponding power allocation coefficient, for $k \in \{1, \cdots, K\}$. Note that $\sum_{k=1}^{K} P_k = 1$ and the information symbols s_k for all, k, are taken from an arbitrary constellation. Furthermore, $\frac{K}{2}$ weak users are served by IRS_1 and $\frac{K}{2}$ strong users are served by IRS_2. Moreover, we partition $\text{IRS}_{\overline{k}}$, where $\overline{k} \in \{1, 2\}$, into $\frac{K}{2}$ parts such that the phase of k^{th} part will be configured to beamform the reflected signal to UE_k.

With this setting, the signal received at UE_k is given by

$$y_k = z_k\, s + w_k = |z_k|\, e^{j\theta_{z_k}}\, s + w_k = |z_k|\, e^{j\theta_{t_k}}\, s + w_k, \tag{4.103}$$

where $w_k \sim \mathcal{CN}\left(0, \sigma_k^2\right)$ is the AWGN at UE_k. We assume that the receiver UE_k has phase information of the effective channel and does phase compensation before processing the received signal. Therefore, after phase compensation, the received signal at UE_k reduces to

$$\widetilde{y}_k = y_k\, e^{-j\theta_{t_k}} = |z_k|\; s + \widetilde{w}_k\,, \tag{4.104}$$

where $\widetilde{w}_k = w_k e^{-j\theta_{t_k}}$ has same distribution as w_k owing to circular symmetry property.

The PEP of UE_k is defined by the probability of the event that symbol s_k is transmitted and symbol $\overline{s}_k$ is detected, where $s_k \neq \overline{s}_k$. Using maximum-likelihood decision rule, the conditional

PEP for UE_k is given by

$$\mathcal{P}\left(s_k \rightarrow \overline{s}_k \,|z_k|\right) = \Pr\left\{\left|\widetilde{y}_k - |z_k|\,\sqrt{P_k}\overline{s}_k\right|^2 \leq \left|\widetilde{y}_k - |z_k|\,\sqrt{P_k}s_k\right|^2\right\}. \tag{4.105}$$

Note that the PEP in (4.105) is conditioned on the effective channel gain $|z_k|$, transmitted symbols and detected symbols of all users. By substituting $\widetilde{y}_k$ in (4.105), we get

$$\mathcal{P}\left(s_k \rightarrow \overline{s}_k \,|z_k|\right) = \Pr\left\{\left||z_k|\left(\sqrt{P_k}\overline{\Delta}_k + X_k\right) + \widetilde{w}_k\right|^2 \leq \left||z_k|\,X_k + \widetilde{w}_k\right|^2\right\}, \tag{4.106}$$

where $\overline{\Delta}_k = s_k - \overline{s}_k$ and $X_k = \sum_{i=1}^{k-1}\sqrt{P_i}\delta_i + \sum_{j=k+1}^{K}\sqrt{P_j}s_j$. Here, $\delta_i = s_i - \widehat{s_i}$, where $\widehat{s_i}$ is the symbol detected at UE_i, for all i. Simplifying (4.106) further, we get

$$\mathcal{P}\left(s_k \rightarrow \overline{s}_k \,|z_k|\right) = \Pr\left\{2\sqrt{P_k}\Re\left\{\overline{\Delta}_k\widetilde{w}_k^*\right\} \leq -|z_k|\left(\left|\sqrt{P_k}\overline{\Delta}_k + X_k\right|^2 + |X_k|^2\right)\right\}, \tag{4.107}$$

which after further simplification gives

$$\mathcal{P}\left(s_k \rightarrow \overline{s}_k \,|z_k|\right) = \mathbf{Q}\left\{\frac{|z_k|\,\upsilon_k}{\iota_k}\right\}, \tag{4.108}$$

where $\upsilon_k = \sqrt{P_k}\,|\Delta_k|^2 + 2\Re\left\{\overline{\Delta}_k X_k^*\right\}$, $\iota_k = |\Delta_k|\sqrt{2\sigma_k^2}$. Now, we derive the PEP of UE_k.

Theorem 4.5. *In an IRS-aided multi-user NOMA system with K users, where users' symbols are drawn from any arbitrary constellation, the average PEP of UE_k is given by*

$$\mathcal{P}^{PEP}\left(s_k \rightarrow \overline{s}_k\right) \approx \mathcal{P}_{k,1}^{PEP} - \left(\mathcal{P}_{k,21}^{PEP} + \mathcal{P}_{k,22}^{PEP}\right) - \mathcal{P}_{k,3}^{PEP}, \tag{4.109}$$

where

$$\mathcal{P}_{k,21}^{PEP} = \frac{a_k}{a_k + b_k}\mathcal{P}_{k,1}^{PEP}, \tag{4.110}$$

$$\mathcal{P}_{k,31}^{PEP} = \int_{x=0}^{\infty}\mathcal{P}\left(s_k \rightarrow \overline{s}_k \,|z_k|\right) f_{3,1}(x)dx, \tag{4.111}$$

and $\mathcal{P}_{k,1}^{PEP}$ and $\mathcal{P}_{k,22}^{PEP}$ are given by

$$\mathcal{P}_{k,1}^{PEP} = \frac{e^{-\frac{c_k^2}{2b_k}}}{6\sqrt{2\pi}\sqrt{b_k}\bar{q}_k}\left(\frac{e^{\frac{\iota_k^2 c_k^2}{2\iota_k^2 b_k + 2\upsilon_k^2 b_k^2}}\sqrt{\frac{\pi}{2}}\left(1 + Erf\left[\frac{\iota_k^2 c_k\sqrt{\frac{\upsilon_k^2}{\iota_k^2} + \frac{1}{b_k}}}{\sqrt{2}\left(\iota_k^2 + \upsilon_k^2 b_k\right)}\right]\right)}{\sqrt{\frac{\upsilon_k^2}{\iota_k^2} + \frac{1}{b_k}}} + \frac{3e^{\frac{3\iota_k^2 c_k^2}{6\iota_k^2 b_k + 8\upsilon_k^2 b_k^2}}\sqrt{\frac{3\pi}{2}}\left(1 + Erf\left[\frac{\iota_k^2 c_k\sqrt{\frac{6\upsilon_k^2}{\iota_k^2} + \frac{9}{2b_k}}}{3\iota_k^2 + 4\upsilon_k^2 b_k}\right]\right)}{\sqrt{\frac{4\upsilon_k^2}{\iota_k^2} + \frac{3}{b_k}}}\right), \tag{4.112}$$

Table 4.1: Simulation parameters.

Parameter	Description
Two-user path distances (UE_1): (i) BS-UE_1, (ii) BS-IRS$_1$, and (iii) IRS$_1$-UE_1	(i) 500 m, (ii) 400 m, (iii) 200 m
Two-user path distances (UE_2): (i) BS-UE_2, (ii) BS-IRS$_2$, and (iii) IRS$_2$-UE_2	(i) 300 m, (ii) 200 m, (iii) 150 m
Multi-user path distances ($UE_{k\in\mathcal{K}^{\mathrm{w}}}$): (i) BS-$UE_k$, (ii) BS-IRS$_1$, and (iii) IRS$_1$-$UE_k$ path distances	Uniformly distributed in annual rings of radii $[r_{\mathrm{in}}, r_{\mathrm{out}}]$ given by (i) [490 m, 510 m], (ii) [390 m, 410 m], and (iii) [190 m, 210 m]
Multi-user path distances ($UE_{k'\in\mathcal{K}^{\mathrm{s}}}$): (i) BS-$UE_{k'}$, (ii) BS-IRS$_2$, and (iii) IRS$_2$-$UE_{k'}$ path distances	Uniformly distributed in annual rings of radii $[r_{\mathrm{in}}, r_{\mathrm{out}}]$ given by (i) [290 m, 310 m], (ii) [190 m, 210 m], and (iii) [140 m, 160 m]
Frequency of operation (f_c)	3 GHz [92]
LoS Path-loss: β_{LoS} (dB)	$-22\log_{10}(d) - 28 - 20\log_{10}(f_c)$, where d is the distance in meters [93]
NLoS Path-loss: β_{NLoS} (dB)	$-36.7\log_{10}(d) - 22.7 - 26\log_{10}(f_c)$ [93]
Rice-factor (K) model (dB)	$13 - (0.03d)$ [93]
Total power (P)	10 W [81]
Per bit SNR: Eb/N0	$(\beta_{\text{End-To-End}}P)/\left(\log_2(MN)\sigma_k^2\right)$, $\beta_{\text{End-To-End}} = \beta_{t_k}(1-[\pi/4]) + T_k\left(\left[\beta_{\mathbf{g}_k}\beta_{\mathbf{h}_k}\right] - \left[\left(\mu_{\mathbf{g}_k}\mu_{\mathbf{h}_k}\right)^2\right]\right)$

$$\mathcal{P}_{k,22}^{PEP} = \frac{e^{-\frac{c_k^2}{2b_k}}\sqrt{b_k}}{6\sqrt{2\pi}(a_k+b_k)\bar{q}_k}\left(\frac{\sqrt{\frac{\pi}{2}}}{\sqrt{\frac{v_k^2}{\iota_k^2}+\frac{1}{a_k}}} + \frac{3\sqrt{\frac{3\pi}{2}}}{\sqrt{\frac{4v_k^2}{\iota_k^2}+\frac{3}{a_k}}}\right). \tag{4.113}$$

Proof. The proof is given in Appendix C.6. □

4.3 Numerical and Simulation Results

We present numerical results to validate our rigorous mathematical analysis and to elucidate how the error performance of an IRS-aided NOMA system is influenced by different system parameters. Simulation parameters are listed in Table 4.1.

4.3.1 Two-User Setup with PAM

Based on the average SEP analysis presented in Theorems 1 and 2, in Fig. 4.3 we plot the average SEP versus Eb/N0 for UE_1 (weak user) and UE_2 (strong user), for $T \in \{128, 256\}$, $N_2 \in \{2, 4\}$, and $N_1 = 2$. The results obtained based on mathematical analysis match well with the simulation results, thus validating the tightness of the derived expressions. We benchmark against the average SEP performance of conventional NOMA systems that do not use IRS

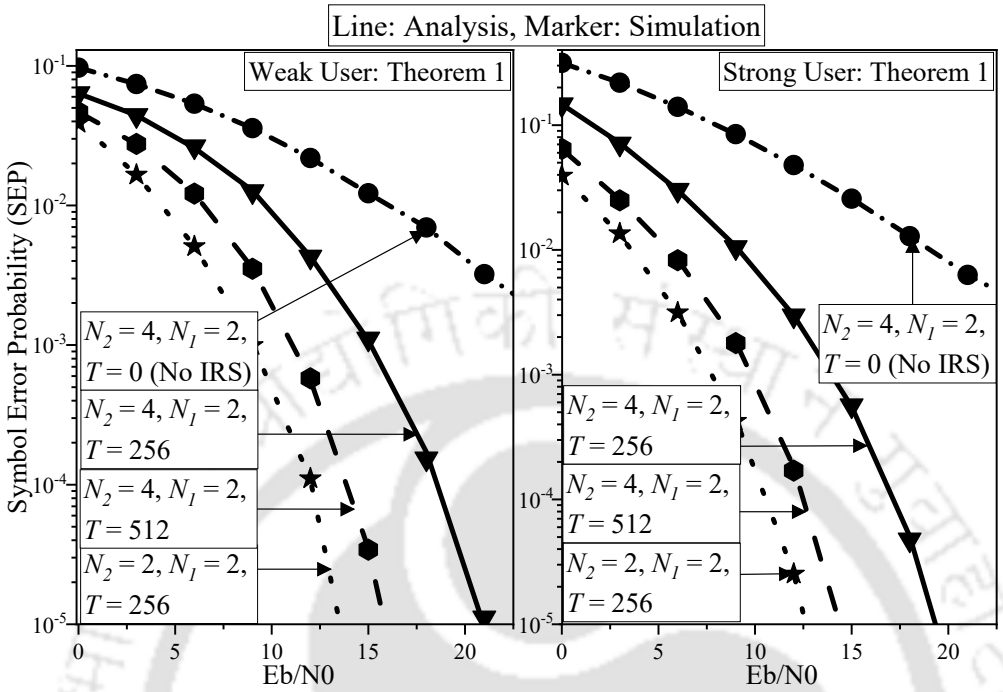


Fig. 4.3: SEP vs. Eb/N0: two-user PAM, perfect CSI, $\lambda = 0.7$.

($T = 0$). And we observe that the target SEP can be attained at much lower SNR with IRS-aided NOMA systems for both users. To be specific, if the target SEP of weak user is 10^{-3} (for $N_2 = 4$, $N_1 = 2$), then it can be achieved at SNR of 15 dB and 11 dB with $T = 256$ and $T = 512$, respectively. Similarly, if the target SEP of strong user is 10^{-3} (for $N_2 = 4$, $N_1 = 2$), then it can be achieved at SNR of 14 dB and 10 dB with $T = 256$ and $T = 512$ IRS elements, respectively. This implies that, by adding 256 elements at each IRS, we can obtain the average SEP of 10^{-3} with 4 dB lower SNR for both users. The SEP performance improves as the number of IRS elements increases due to constructive interference-induced increase in the received signal power.

Furthermore, we can observe that, increasing the PAM modulation order (N_2) of UE_2 from 2 to 4, and keeping T fixed at 128, causes significant increase in SEPs of both users. If we also simultaneously increase T from 128 to 256 along with increase in N_2 from 2 to 4, then we can significantly reduce SEPs of both users relative to the case with $N_2 = 4$, $N_1 = 2$ and $T = 128$. In other words, communication using a higher order constellation is supported by adding 128 extra IRS elements without significant increase in the SEP. In Fig. 4.4, we also plot SEP vs.

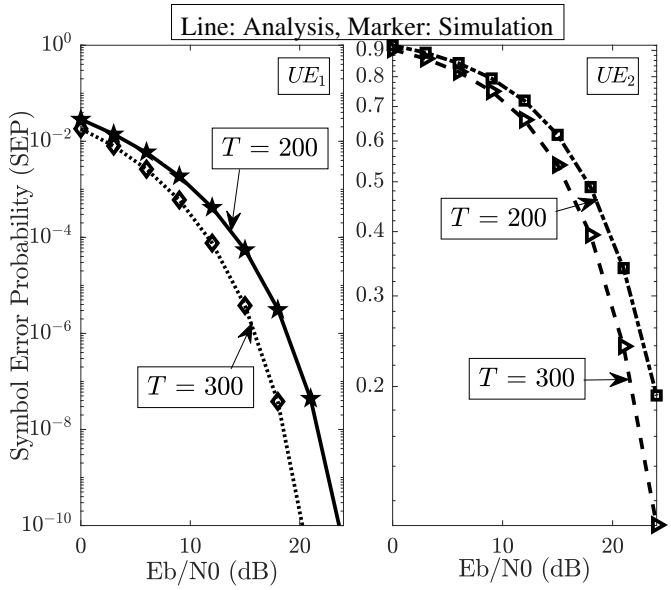
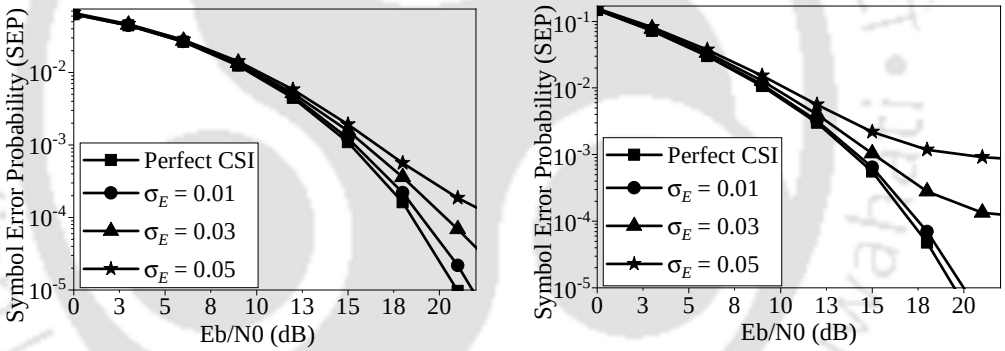


Fig. 4.4: SEP vs. Eb/N0: two-user PAM, $N_2 = 64$, $N_1 = 4$, $\lambda = 0.95$.



(a) Weak user **(b)** Strong user

Fig. 4.5: Impact of imperfect CSI, two-user PAM: $T = 128$, $\lambda = 0.7$.

Eb/N0 for higher order constellation, i.e., $N_2 = 64$, $N_1 = 4$, $T \in \{200, 300\}$. We can observe that the average SEP performance can be improved for both users by adding 100 extra IRS elements at both IRSs. These results in Fig. 4.4 re-establish the fact that we can employ higher order constellation without performance degradation by using extra IRS elements.

4.3.1.1 Imperfect CSI

We model imperfect CSI based on a generic Gauss-Markov uncertainty model presented in [94].[6] Based on this model, the estimates of the direct link and the cascaded reflected link

[6]The average SEP analysis with imperfect CSI is considerably more involved and beyond the scope of this work due to the presence of uncompensated residual phase component at the *UE*s. And it is an interesting avenue for future research.

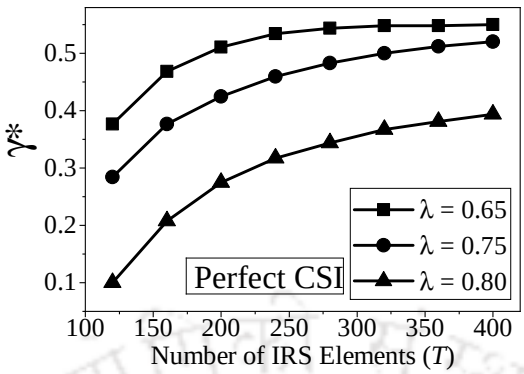


Fig. 4.6: Trade-off between γ^* and λ: $N_1 = 2$, $N_2 = 4$, SNR = 10 dB.

are respectively given by

$$\widehat{t_k} = \sqrt{1-\sigma_E}t_k + \sqrt{\sigma_E}\widetilde{t_k}, \tag{4.114}$$

$$\widehat{\breve{v}_{k,i}} = \sqrt{1-\sigma_E}\breve{v}_{k,i} + \sqrt{\sigma_E}\widetilde{\breve{v}}_{k,i}, \tag{4.115}$$

where t_k is true direct channel from BS to UE_k, $\widetilde{t_k}$ is corresponding estimation error, $\breve{v}_{k,i}$ is true cascaded reflected channel from BS to UE_k via i^{th} element of IRS_k, $\widetilde{\breve{v}}_{k,i}$ is corresponding estimation error in this link, for $i \in \{1, 2, \ldots, T\}$. The parameter σ_E captures variance of estimation error.

Based on (4.114) and (4.115), the BS configures phases at IRS_k. Also, it feeds back an estimate of the direct channel phase for compensation at the respective *UE*s. Figs. 4.5(a) and 4.5(b) plot the SEP vs. Eb/N0 for UE_1 and UE_2, respectively, with imperfect CSI for three different values of σ_E, namely, 0.01, 0.03 and 0.05. As we decrease σ_E, the SEP performance improves for both users and approaches the performance with perfect CSI, since reflected signals align better with the signals via direct path. Also, as variance of estimation error increases, the SEP performance of UE_2 that performs SIC at its end, degrades relatively more compared to the SEP of UE_1. Essentially the SEP of UE_2 flattens out as σ_E increases, due to the joint impact of estimation error and imperfect SIC at UE_2.

4.3.1.2 IRS Activation Ratio

Next, we formulate and solve a constrained optimization problem to determine the fraction of elements that must be activated at each IRS such that average SEP of both users is near-

identical subject to the constraint that sum total of active IRS elements at both IRSs is equal to T. Mathematically, this can be stated as:

$$\begin{aligned} &\underset{0<\gamma\leq 1}{\text{minimize}} \quad \left|\mathcal{P}_1^e(T_1) - \mathcal{P}_2^e(T_2)\right| \\ &\text{subject to} \quad T_1 + T_2 = T, \end{aligned} \tag{4.116}$$

where T_k is number of elements activated at IRS_k such that $T_1 = \gamma T$, $T_2 = (1-\gamma)T$ where $\gamma \in (0,1)$ is the fraction of IRS elements activated at IRS_1, T is the number of elements at each IRS, and $\mathcal{P}_k^e(T_k)$ is the average SEP at UE_k, for $k \in \{1,2\}$. Coherence interval would be limited in dynamic wireless environment with high speed users and fraction of coherence interval spent on channel estimation scales with number of IRS elements that are activated. Hence, the constraint $T_1 + T_2 = T$ in (4.116) is well-motivated to limit the channel estimation overhead while meeting the uniform QoS requirements of both users irrespective of their locations.

Optimization can be performed at BS using brute-force search based on CSI estimated on the uplink. Hence, no CSI feedback is needed. The number of search operations needed to obtain γ^* depends on feasible range of γ and step-size for search, and is independent of T. Smaller the step-size, closer will be the SEP values of both users. We present the relationship between NOMA power allocation λ and the fraction γ of T that must be activated at IRS_1 for fixed values of path-loss, N_1, N_2, K, and noise variance.

Fig. 4.6 plots optimal γ (γ^*) versus T for different values of the fraction λ of total power P allocated to UE_1, i.e., $\lambda \in \{0.65, 0.75, 0.8\}$. We observe that as we decrease λ, a larger fraction γ of total IRS elements T needs to be activated at IRS_1 such that the average SEP of both users is near-identical. In other words, if we decrease the fraction λ of total power P allocated to UE_1, then we need to increase the fraction γ of total number of IRS elements T that must be activated at IRS_1 to obtain the near-identical performance for both users. For example, as λ reduces from 0.80 to 0.65, the power allocated to UE_1 reduces from $0.8P$ to $0.65P$, where P denotes the total power at any UE. Therefore, for $T = 200$, the fraction of total elements that must be activated at IRS_1 (which is serving UE_1) increases from 0.275 to 0.51.

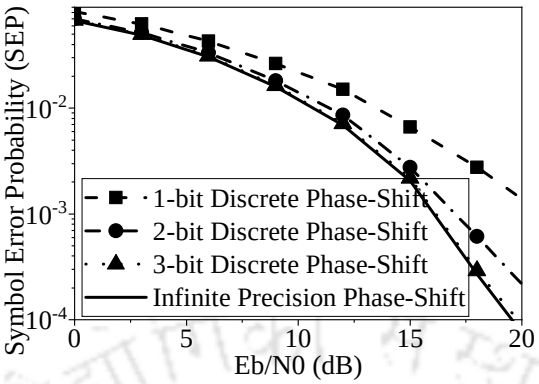


Fig. 4.7: Discrete phase-shift, weak user: $N_1 = 2$, $N_2 = 4$, $T = 100$, $\lambda = 0.7$.

4.3.1.3 Discrete Phase-Shifts

Finally, to address the practical aspect of phase-shift compensation, in Fig. 4.7, we plot the SEP versus Eb/N0 for UE_1 considering 1-bit, 2-bit and 3-bit programmable discrete phase-shifters at the IRS. Note that j-bit discrete phase-shifter can facilitate 2^j levels of phase-shift. It is evident that only 3-bit discrete phase-shifters are sufficient to obtain nearly the same SEP as obtained by phase-shifters with infinite precision at IRS.

4.3.2 Two-User Setup with QAM

Based on the average SEP analysis presented in Theorems 4.3 and 4.4, we present the numerical results in Fig. 4.8. Specifically, we plot SEP vs. Eb/N0 for $T \in \{100, 200\}$, $N_1 = 4$, $N_2 = 16$. The results obtained based on mathematical analysis match well with the simulation results, thus validating the tightness of the derived expressions. We can observe that the SEP performance for both users improves as we increase the number of IRS elements at both users. Furthermore, we can observe that SEP performance of UE_2 (strong user) is worse than that of UE_1 (weak user). This is primarily because of the effect of error propagation while doing SIC at UE_2. Other reasons for this performance difference are the higher constellation size used for UE_2 ($N_2 = 16$) and lower power allocation to UE_2 (5%).

4.3.3 Multi-User Setup: SEP Analysis for Multicast Transmission

Based on the results presented in Corollaries 4.1 and 4.2, we present the numerical results in Fig. 4.9. Specifically, we present the SEP averaged over both small-scale fading and over the number of users for weak users and strong users. That is, in Fig. 4.9 we present the average

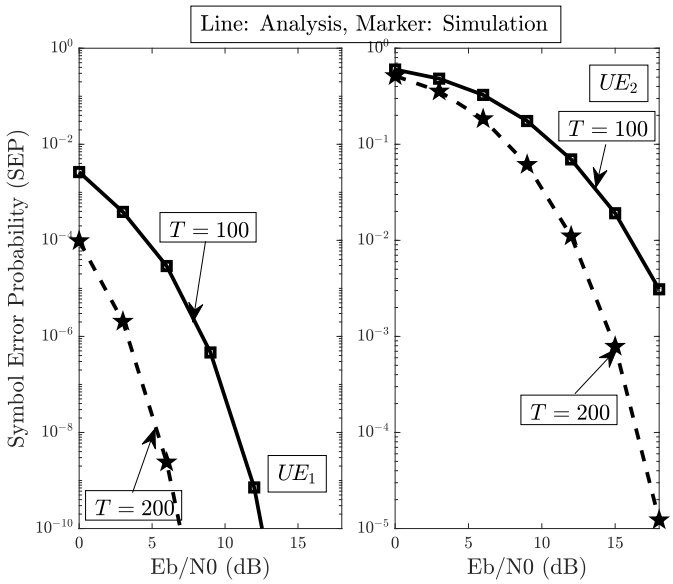


Fig. 4.8: SEP vs. Eb/N0: two-user QAM, $N_2 = 16$, $N_1 = 4$, $\lambda = 0.95$.

SEP given by

$$\mathcal{P}_k^{e,\text{Avg}}(T) = \frac{2}{K} \sum_{k \in \mathcal{K}^{\text{W}}} \mathcal{P}_k^e\left(\frac{2T}{K}\right), \tag{4.117}$$

$$\mathcal{P}_{k'}^{e,\text{Avg}}(T) = \frac{2}{K} \sum_{k' \in \mathcal{K}^{\text{S}}} \mathcal{P}_{k'}^e\left(\frac{2T}{K}\right). \tag{4.118}$$

The results obtained based on mathematical analysis match well with the simulation results, thus validating the tightness of the derived expressions. Here we present the results for $K = 2, 4, 8$ and $T = 256$. As we increase the number of users, the SEP performance worsens for both UE_k and $UE_{k'}$. This is because the number of IRS elements used to beamform the signal to each user reduces with increase in number of users. Alternatively, size of each partition reduces with increase in number of users for a fixed value of T.

4.3.4 Multi-User Setup: PEP Analysis for Unicast Transmission

Finally, based on the derived average PEP expressions presented in Theorem 5, in Fig. 4.10 we present the numerical results for UE_1, i.e., the weakest user in terms of effective channel strength among $K = 4$ users. We plot the average PEP vs. Eb/N0 for BPSK and 4-PAM constellations. The results obtained based on mathematical analysis match well with the simulation results, thus validating the tightness of the derived expressions. The power allocation coeffi-

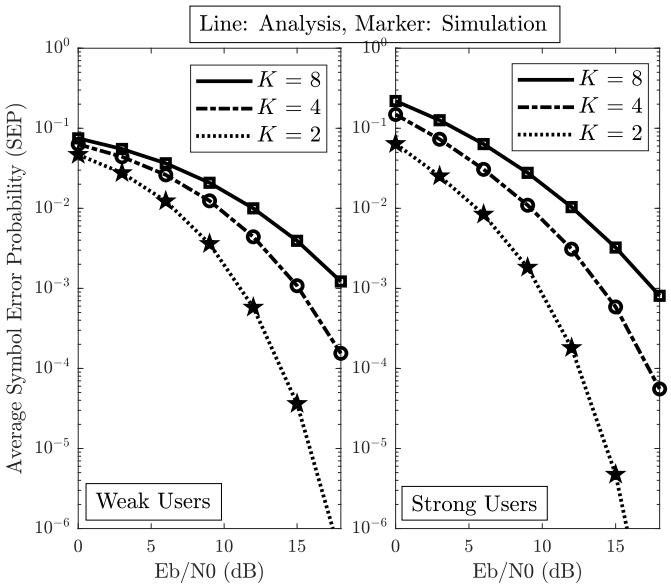


Fig. 4.9: SEP vs. Eb/N0: multi-user multicast transmission with PAM, $T = 256$, $\lambda = 0.7$.

T	4-PAM	BPSK
50	0.0456311	0.0111268
200	0.0222565	0.0009458

Table 4.2: Comparison of Average PEP for PAM and BPSK, SNR = 12 dB.

cients considered are: $P_1 = 0.4P$, $P_2 = 0.3P$, $P_3 = 0.2P$ and $P_4 = 0.1P$, where $P = 10\,W$ is the power budget. We present the results for $T \in \{50, 200\}$. As the number of IRS elements T increases, the number of reflected paths adding constructively at the receiver increases, boosting the received signal strength. This, in turn, improves the average PEP. Thus, then average PEP performance with $T = 200$ is better than that with $T = 50$, for both BPSK and 4-PAM. For example, with SNR = 12 dB, the average PEP of 4-PAM is 75% lesser with $T = 200$ than that with $T = 50$. Similarly, with SNR = 12 dB, the average PEP of BPSK is 95% lesser with $T = 200$ than that with $T = 50$. Furthermore, the average PEP performance of BPSK is better than that of 4-PAM owing to the smaller constellation size, which allows more accurate symbol detection.

4.4 Summary

We derived novel approximate analytical expressions for the average SEP of UE_1 (weak user) and UE_2 (strong user) in an IRS-aided NOMA system employing PAM and QAM constellations. We also analyzed the error performance of multi-user IRS-aided NOMA system in two

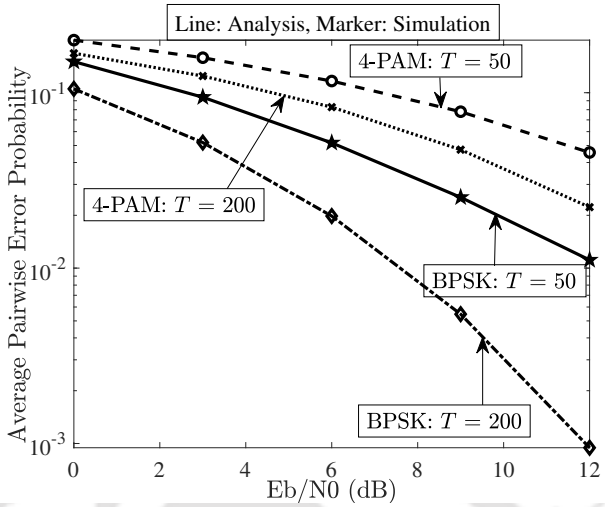


Fig. 4.10: PEP vs. Eb/N0 (dB): Multi-user unicast transmission, $K = 4$, $P_1 = 0.4P$, $P_2 = 0.3P$, $P_3 = 0.2P$, $P_4 = 0.1P$.

ways. Specifically, we derived the average SEP expressions for a multi-user multicast transmission employing PAM constellation. Further, we also derived the average PEP expression for a multi-user unicast transmission employing arbitrary modulation scheme. Using extensive simulations, we verified the tightness of the derived average SEP expressions. We proved that by adding a few extra IRS elements: i) the target average SEP can be obtained at a much lower SNR for both the users in two-user scenario and for all users in multi-user scenario, ii) communication using a higher-order constellation can be supported without significant increase in the average SEP/average PEP, and iii) a smaller fraction of the total power needs to be allocated to the weak user to obtain near-identical average SEP performance as that of the strong user. Also, a 3-bit programmable IRS is sufficient to obtain SEP performance close to that of an IRS with infinite precision phase-shift. Impact of imperfect CSI is also elucidated.

5

Conclusions

Contents

5.1 Research Contributions 129
5.2 Quantification of the Contributions Towards the Objectives of the Thesis 131
5.3 Directions for Future Research 133

We summarize the research contributions of our work in Section 5.1 and conclude this thesis with some of the future directions in Section 5.3.

5.1 Research Contributions

In this thesis, we focused on modeling, optimization and performance analysis of (i) uplink massive MIMO-NOMA systems with the objective of achieving increased connectivity while guaranteeing uniform quality of service (equivalently, performance fairness) in terms of SE with and without channel aging as summarized in Sections 5.1.1 and 5.1.2 below, and (ii) IRS-enabled NOMA systems with the objective of achieving increased reliability of data detection as summarized in Section 5.1.3 below. To be specific, in this thesis, we performed detailed mathematical analysis, extensive numerical simulations and looked at several key design aspects such as developing low overhead channel estimation schemes, understanding the impact of pilot contamination, Doppler induced channel aging, imperfect SIC, discrete phase shifts, user grouping strategies, max-min optimal and proportional fairness power control techniques on system performance and use of prediction to counter the impact of channel aging.

5.1.1 Massive MIMO-NOMA Systems

We first investigated the uplink of a NOMA-based massive MIMO system with a ZF decoder designed using estimated CSI at the BS to completely cancel inter-group interference under quasi-static Rayleigh fading channel. We considered two low-overhead channel estimation schemes, namely, Scheme-I and Scheme-S, and two user grouping strategies, namely, NF grouping and NB grouping. We derived new expressions for achievable SE corresponding to ZF decoder designed using both the estimation schemes. To ensure uniform quality-of-service, we formulated a max-min transmit power allocation problem and obtained corresponding max-min power allocation coefficients using geometric programming and CVX. Through extensive simulations based on derived achievable SE expressions, we showed that, in a substantial portion of the over-loaded regime and in entire under-loaded regime, the designed ZF decoder based on Scheme-S with NF grouping strategy gives the highest max-min achievable SE among all the considered NOMA decoding and grouping strategies. We also showed how scalable NB group-

ing strategy can be used to ensure massive connectivity and uniform service to all users in the network. Therein, we showed that among all the considered scalable decoding strategies, the designed ZF decoder based on Scheme-S with scalable NB grouping gives the highest max-min achievable SE and can simultaneously serve a very large number of users.

5.1.2 Massive MIMO-NOMA Systems under Channel Aging

Next, we investigated a more practical case of time-varying channel, where we identified how the performance of uplink massive MIMO-NOMA system is influenced by channel aging and pilot contamination. Particularly, we considered the ZF and MR decoders for processing at the BS and considered two pilot-sharing-based channel estimation schemes, Scheme-I and Scheme-S. We obtained novel closed-form expressions for the lower bounds on the achievable SE for each decoding scheme. We also obtained an optimal WLP of order p to predict the CSI and derived novel expressions for the lower bound on achievable SE with ZF and MR decoders designed using the predicted CSI.

To ensure fairness in achievable SE performance, we formulated a max-min optimization problem based on the expressions derived for each decoder and converted it to a convex program in order to obtain the corresponding max-min power allocation coefficients. Further, to strike a good balance between fairness and sum SE, we formulated and solved proportional fairness optimization problem based on the expressions derived for each decoder and obtained the corresponding proportional fairness power allocation coefficients using geometric programming. To quantify the impact of channel aging on performance, in addition to max-min and proportional fairness power control, we also considered equal and inversion power control.

Through rigorous derivations and extensive simulations, we showed that the ZF decoder based on Scheme-S performs the best under channel aging among all considered decoders in terms of the per-user achievable SE and the minimum achievable user SE for three practically well-motivated power control strategies, namely, equal, inversion and max-min power control. Moreover, for ZF decoder based on Scheme-S, the inversion power control is more robust to the impact of channel aging when its per-user achievable SE performance is compared to that of equal power control and max-min power control. Finally, for ZF decoder based on Scheme-S,

max-min power control is the most robust to channel aging when its minimum achievable user SE is compared to that under equal and inversion power control.

5.1.3 IRS-aided NOMA Systems

Finally, we investigated the challenges involved in ensuring increased connectivity and reliability in an IRS-aided NOMA system, which is a cost-efficient alternative to a massive MIMO system. Therein, we derived novel approximate analytical expressions for the average SEP in a two-user IRS-aided NOMA system employing PAM and QAM constellations of arbitrary order. We also analyzed the error performance of multi-user IRS-aided NOMA system in two ways. Specifically, we derived the average SEP expressions for a multi-user multicast transmission employing PAM constellation. Further, we also derived the average PEP expression for a multi-user unicast transmission employing arbitrary modulation scheme. Using extensive simulations, we verified the tightness of the derived average SEP/average PEP expressions. We proved that by adding a few extra IRS elements: i) the target average SEP can be obtained at a much lower SNR for both the users in two-user scenario and for all users in multi-user scenario, ii) communication using a higher-order constellation can be supported without significant increase in the average SEP/average PEP, and iii) a smaller fraction of the total power needs to be allocated to the weak user to obtain near-identical average SEP performance as that of the strong user. We also showed that a 3-bit programmable IRS is sufficient to obtain SEP performance close to that of an IRS with infinite precision phase-shift. Impact of imperfect CSI is also elucidated.

5.2 Quantification of the Contributions Towards the Objectives of the Thesis

As we discussed in Chapter 1 of the thesis, the key objectives of the thesis are improving wireless connectivity, ensure fairness in user performance in terms of achievable SE and improving reliability in signal detection. In this direction, we made significant contributions in Chapters 2-4.

Specifically, in Chapters 2 and 3, we considered an annular ring of outer radius of 300 m and inner radius of 100 m and showed that we can serve $L = NK = \min\{MK, \tau_c K\}$, for $N \leq \min\{M, \tau_c\}$, number of users under the assumption of quasi-static fading while guaranteeing

fairness in performance in terms of achievable SE using the proposed decoding strategies and max-min power control. If we scale the numbers for a better picture, with $M = 128$ antennas at the BS and coherence interval of length $\tau_c = 168$ samples, we can provide simultaneous wireless connectivity to 256, 384 and 512 *UE*s with $K = 2$, $K = 3$, and $K = 3$ users per group, respectively, while ensuring the Jain's fairness index equal to 1. Similarly, with $M = 168$ and $\tau_c = 168$, we can provide simultaneous wireless connectivity to 336, 504 and 672 *UE*s with $K = 2$, $K = 3$, and $K = 3$ users per group, respectively, while ensuring the Jain's fairness index equal to 1.

In addition to this, in Chapter 3, we also showed that the we can maintain this increased connectivity even under high mobility, while simultaneously ensuring Jain's fairness index equal to 1. Specifically, among the considered values for the normalized Doppler shift, $f_D T_S = 0.001$ corresponds to the relative velocity of 540 kilo meters per hour with $T_S = 1\mu S$ and $f_c = 2$ GHz. In addition, in Chapter 4 we showed that using IRSs, we can ensure increased reliability in terms of signal detection. Quantitatively, for SNR $\approx$ 15 dB and number of elements at IRS $T = 256$, we can achieve the average SEP of 10^{-5}, and this performance is expected to improve as we increase the value of T based on the observed trends.

Now, let us compare the improvements made in connectivity, reliability and fairness. As discussed in Chapter 2, with massive MIMO-NOMA systems, we were able to ensure simultaneous wireless connectivity to $\min\{MK, \tau_c K\}$ users, which is K times of that with massive MIMO-OMA which allows simultaneous wireless connectivity to $\min\{M, \tau_c\}$ users. Further, if we look at the fairness in user performance in terms of achievable SE, the proposed max-min power control provides the Jain's fairness index $J = 1$, which is 23% higher than that of the equal (or greedy) power control ($J = 0.77$) and 14% higher than that of the inversion power control ($J = 0.86$). Further, if we look at the improvement made in the reliability of signal detection, then the proposed IRS-aided NOMA transmission achieves the average SEP equal to 10^{-5}, while the corresponding non-IRS-aided NOMA transmission achieves the average SEP higher than 10^{-2}. Such a huge improvement is achieved only with $T = 256$, and the order to improvement is expected to increase further with increase in T.

5.3 Directions for Future Research

This thesis lays the groundwork for extending its fundamental ideas and mathematical analysis in diverse directions. Below we discuss a few potential problems for future research.

5.3.1 Massive MIMO Systems with RSMA Transmission

In the first part of this thesis, we investigated uplink of a massive MIMO-NOMA systems to derive novel lower bounds on the achievable SE for ZF decoders based on imperfect CSI and for ZF/MR decoders based on imperfect CSI under channel aging. The comprehensive investigations therein can be extended to massive MIMO-RSMA systems, where RSMA is employed instead of NOMA. RSMA principally differs from NOMA in the signal modelling in the sense that, the transmitted signal of each user is separated into a common part and a private part, and at the receiver(s), multi-user detection is employed. It would be crucial to identify the regimes where massive MIMO-RSMA systems outperform massive MIMO-NOMA systems for linear processing at the BS based on imperfect CSI. Also, it would be interesting to investigate the performance of massive MIMO-RSMA systems under channel aging. Below we present probable problem statements for future research.

- **Problem Statement 1:** Investigation of the effects of channel estimation schemes and user grouping strategies on the SE of uplink massive MIMO-RSMA systems, with an emphasis on exploring the interplay between these channel estimation methods and user grouping approaches. Examination of how decoding and power control schemes, designed based on the estimated CSI, influence the SE performance of uplink massive MIMO-RSMA systems.
- **Problem Statement 2:** Investigation of the effects of channel aging on pilot-sharing-based channel estimation schemes in combination with various user grouping strategies, focusing on their impact on the SE of uplink massive MIMO-RSMA systems. Further exploration of how the implementation of different decoding and power control schemes, which utilize outdated estimated CSI, affects the SE performance of these systems.

5.3.2 Error Probability Analysis of NOMA Systems under Imperfect CSI

In this thesis, we performed comprehensive error probability analysis of IRS-enabled NOMA systems assuming that the perfect CSI is available at the BS. While we have quantified the effect of CSI errors on the error probability performance of IRS-enabled NOMA systems using a widely used model. However, the problem of obtaining the analytical closed-form expressions for error probability of NOMA systems with imperfect CSI for PAM/QAM modulation is still open. This study of error probability analysis with imperfect CSI can further be extended to IRS-enabled NOMA systems. Below we state the probable problem statements for future research directions.

- **Problem Statement 1:** Error probability analysis of two-user and multi-user NOMA/RSMA systems when SIC is based on the estimated CSI with PAM/QAM modulation of arbitrary order.
- **Problem Statement 2:** Error probability analysis of two-user and multi-user IRS-enabled NOMA/RSMA systems when IRS phase configurations and SIC are based on the estimated CSI with PAM/QAM modulation of arbitrary order.
- **Problem Statement 3:** Error probability analysis of two-user and multi-user IRS-enabled NOMA/RSMA systems when IRS phase configurations and SIC are based on the outdated estimated CSI (and/or predicted CSI) with PAM/QAM modulation of arbitrary order.

5.3.3 IRS-enabled Massive MIMO-NOMA Systems

The study presented in the later part of the thesis addresses challenges involved in IRS-aided NOMA systems with single-antenna BS. This study can be interestingly extended to IRS-aided massive MIMO systems with limited number of RF chains, wherein it would be crucial to come up with efficient antenna selection algorithm and efficient power allocation algorithms for various system objectives like maximizing the sum SE, maximizing the minimum achievable SE, etc. Below we state probable problem statements for future research directions.

- **Problem Statement 1:** Investigation of the effects of channel estimation schemes and

user grouping strategies on the SE of IRS-aided uplink massive MIMO-NOMA/RSMA systems, with an emphasis on exploring the interplay between these channel estimation methods, user grouping approaches and IRS phase configurations. Examination of how decoding, power control schemes and IRS phase configurations, designed based on the estimated CSI, influence the SE performance of IRS-aided uplink massive MIMO-NOMA/RSMA systems.

- **Problem Statement 2:** Investigation of the effects of channel aging on pilot-sharing-based channel estimation schemes in combination with various user grouping strategies, focusing on their impact on the SE of IRS-aided uplink massive MIMO-RSMA systems. Further exploration of how the implementation of different decoding, power control schemes and IRS phase configuration strategies, which utilize outdated estimated CSI, affects the SE performance of these systems.
- **Problem Statement 3:** Investigation of diversity gains offered by IRS in uplink and downlink massive MIMO-NOMA/RSMA systems under perfect CSI, estimated CSI and outdated CSI.

5.3.4 Multi-IRS-aided Massive MIMO mmWave/THz Systems

It would also be interesting and challenging to develop low-latency beam-training algorithms for multi-IRS-aided massive MIMO systems. Here it is important to note that the beam-training algorithm should jointly take care of active and passive beams at the BS and each IRS, respectively. Tools like matching theory, compressed sensing can be potentially explored to come up with an efficient solution. Below we state the probable problem statement for future research directions.

- **Problem Statement:** Design of computationally efficient active and passive beam-training algorithms for multi-IRS-aided massive MIMO mmWave/THz systems.

A

Appendix

A.1 Proof of Theorem 2.1

The signal model in (2.30) corresponds to the deterministic channel impaired by additive non-Gaussian noise [14]. Thus, lower bound on the achievable SE for UE_{nk} is given by

$$\mathrm{R}_{nk}^{\mathrm{I}} = \delta \log_2\left(1 + \mathrm{SINR}_{nk}^{\mathrm{I}}\right). \tag{A.1}$$

From (2.30), $\mathrm{SINR}_{nk}^{\mathrm{I}}$ is given by

$$\mathrm{SINR}_{nk}^{\mathrm{I}} = \frac{\eta_{nk}\gamma_{nk}}{\sum_{i=k+1}^{K}\eta_{ni}\gamma_{ni} + \mathbb{V}\mathrm{ar}\left\{\left[\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right]_n\right\}}. \tag{A.2}$$

We next compute variance of the effective noise term in closed-form as follows:

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left\{\left[\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right]_n\right\} &= \left[\mathbb{C}\mathrm{ov}\left\{\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right\}\right]_{n,n} \\ &= \left[\mathbb{E}\left\{\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\left(\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right)'\right\}\right]_{n,n} \\ &= \left[\mathbb{E}_{\mathbf{Z}}\left\{\mathbf{A}'\mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\overline{\mathbf{n}}^{\mathrm{I}}\left(\overline{\mathbf{n}}^{\mathrm{I}}\right)'\middle|\mathbf{Z}\right\}\mathbf{A}\right\}\right]_{n,n}.\end{aligned} \tag{A.3}$$

Since $\widetilde{\mathbf{G}}^{\mathrm{I}}$ is independent of $\mathbf{Z}$,

$$\begin{aligned}\mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\overline{\mathbf{n}}^{\mathrm{I}}\left(\overline{\mathbf{n}}^{\mathrm{I}}\right)'\middle|\mathbf{Z}\right\} &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\overline{\mathbf{n}}^{\mathrm{I}}\left(\overline{\mathbf{n}}^{\mathrm{I}}\right)'\right\} \\ &= \sigma^2\mathbf{I}_M + \mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{x}\left(\widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{x}\right)'\right\}. \end{aligned} \tag{A.4}$$

Further,

$$\begin{aligned}\mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{x}\left(\widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{x}\right)'\right\} &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\widetilde{\mathbf{G}}^{\mathrm{I}}\mathbb{E}\left\{\mathbf{x}\mathbf{x}'\right\}\left(\widetilde{\mathbf{G}}^{\mathrm{I}}\right)'\right\} \\ &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\widetilde{\mathbf{G}}^{\mathrm{I}}\mathbf{D}_{\gamma}\left(\widetilde{\mathbf{G}}^{\mathrm{I}}\right)'\right\} \\ &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\sum_{i=1}^{K}\widetilde{\mathbf{G}}_i^{\mathrm{I}}\mathbf{D}_{\gamma_i}\left(\widetilde{\mathbf{G}}_i^{\mathrm{I}}\right)'\right\} \\ &= \sum_{i=1}^{K}\mathbb{E}_{\widetilde{\mathbf{G}}^{\mathrm{I}}}\left\{\sum_{n=1}^{N}\gamma_{ni}\tilde{\mathbf{g}}_{ni}^{\mathrm{I}}\left(\tilde{\mathbf{g}}_{ni}^{\mathrm{I}}\right)'\right\} \\ &= \sum_{n=1}^{N}\sum_{i=1}^{K}\left\{\gamma_{ni}\left(\beta_{ni}-\eta_{ni}\right)\mathbf{I}_M\right\}, \end{aligned} \tag{A.5}$$

where, in the second equality above, $\mathbf{D}_{\gamma} = \mathrm{diag}\left[\gamma_{11},\cdots,\gamma_{N1},\cdots,\gamma_{1K},\cdots,\gamma_{NK}\right]$. Using (A.4) and (A.5), we can rewrite (A.3) as

$$\mathbb{V}\mathrm{ar}\left\{\left[\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right]_n\right\} = \left[\mathbb{E}_{\mathbf{Z}}\left\{\mathbf{A}'\left(\sigma^2\mathbf{I}_M + \lambda^{\mathrm{I}}\mathbf{I}_M\right)\mathbf{A}\right\}\right]_{n,n}, \tag{A.6}$$

where $\lambda^{\mathrm{I}} = \sum_{n=1}^{N}\sum_{i=1}^{K}\left\{\left(\beta_{ni}-\eta_{ni}\right)\gamma_{ni}\right\}$. Using the result that $\mathbb{E}_{\mathbf{Z}}\left\{\left[\mathbf{A}'\mathbf{A}\right]_{i,i}\right\} = (M-N)^{-1}$ [14],

$$\mathbb{V}\mathrm{ar}\left\{\left[\mathbf{A}'\overline{\mathbf{n}}^{\mathrm{I}}\right]_n\right\} = \left(\sigma^2+\lambda^{\mathrm{I}}\right)(M-N)^{-1}. \tag{A.7}$$

Substituting (A.7) in (A.2) gives (2.39).

A.2 Proof of Theorem 2.2

In (2.45), the third term is correlated to the first and second terms. But, given the side information $\Omega = \mathbf{V}$, all three terms in (2.45) are uncorrelated. Thus, we can use the capacity bound for fading channel with additive non-Gaussian noise and side information. Thus, lower bound on the achievable rate for UE_{nk} is given by [14]

$$\mathrm{R}_{nk}^{\mathrm{S}} \geq \overline{\mathrm{R}}_{nk}^{\mathrm{S}} = \mathbb{E}_{\mathbf{V}}\left\{\delta\log_2\left(1+\mathrm{SINR}_{nk}^{\text{S-UB}}\right)\right\}, \tag{A.8}$$

where $\text{SINR}_{nk}^{\text{S-UB}}$ is given by

$$\text{SINR}_{nk}^{\text{S-UB}} = \frac{\eta_n \gamma_{nk}}{\sum_{i=k+1}^{K} \eta_n \gamma_{ni} + \mathbb{V}\text{ar}\left\{\left[\mathbf{B}'\overline{\mathbf{n}}^{\text{S}}\right]_n \middle| \mathbf{V}\right\}}. \tag{A.9}$$

We next compute variance of the effective noise term conditioned on $\Omega = \mathbf{V}$ in the closed-form as follows:

$$\begin{aligned}
\mathbb{V}\text{ar}\left\{\left[\mathbf{B}'\overline{\mathbf{n}}^{\text{S}}\right]_n \middle| \mathbf{V}\right\} &= \left[\mathbb{C}\text{ov}\left\{\mathbf{B}'\overline{\mathbf{n}}^{\text{S}} \middle| \mathbf{V}\right\}\right]_{n,n} \\
&= \left[\mathbf{B}'\mathbb{E}\left\{\overline{\mathbf{n}}^{\text{S}}\left(\overline{\mathbf{n}}^{\text{S}}\right)' \middle| \mathbf{V}\right\}\mathbf{B}\right]_{n,n} \\
&= \left[\mathbf{B}'\mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\overline{\mathbf{n}}^{\text{S}}\left(\overline{\mathbf{n}}^{\text{S}}\right)' \middle| \mathbf{V}\right\}\mathbf{B}\right]_{n,n}.
\end{aligned} \tag{A.10}$$

Since $\widetilde{\mathbf{G}}^{\text{S}}$ and $\mathbf{V}$ are correlated,

$$\mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\overline{\mathbf{n}}^{\text{S}}\left(\overline{\mathbf{n}}^{\text{S}}\right)' \middle| \mathbf{V}\right\} = \sigma^2\mathbf{I}_M + \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\widetilde{\mathbf{G}}^{\text{S}}\mathbf{x}\left(\widetilde{\mathbf{G}}^{\text{S}}\mathbf{x}\right)' \middle| \mathbf{V}\right\}. \tag{A.11}$$

Further,

$$\begin{aligned}
\mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\widetilde{\mathbf{G}}^{\text{S}}\mathbf{x}\left(\widetilde{\mathbf{G}}^{\text{S}}\mathbf{x}\right)' \middle| \mathbf{V}\right\} &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\widetilde{\mathbf{G}}^{\text{S}}\mathbb{E}\left\{\mathbf{x}\mathbf{x}'\right\}\left(\widetilde{\mathbf{G}}^{\text{S}}\right)' \middle| \mathbf{V}\right\} \\
&= \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\widetilde{\mathbf{G}}^{\text{S}}\mathbf{D}_{\gamma}\left(\widetilde{\mathbf{G}}^{\text{S}}\right)' \middle| \mathbf{V}\right\} \\
&= \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\sum_{i=1}^{K}\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}\left(\widetilde{\mathbf{G}}_i^{\text{S}}\right)' \middle| \mathbf{V}\right\} \\
&= \sum_{i=1}^{K}\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}} \middle| \mathbf{V}\right\}.
\end{aligned} \tag{A.12}$$

Therefore, (A.10) can be rewritten as

$$\mathbb{V}\text{ar}\left\{\left[\mathbf{B}'\overline{\mathbf{n}}^{\text{S}}\right]_n \middle| \mathbf{V}\right\} = \left[\mathbf{B}'\left(\sigma^2\mathbf{I}_M + \sum_{i=1}^{K}\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}} \middle| \mathbf{V}\right\}\right)\mathbf{B}\right]_{n,n}. \tag{A.13}$$

Using the result that $\mathbb{C}\text{ov}\{\mathbf{P}|\mathbf{Q}\} = \mathbb{C}\text{ov}\{\mathbf{P}\} - \mathbb{E}\{\mathbf{P}\mathbf{Q}'\}\left(\mathbb{E}\left\{\mathbf{Q}\mathbf{Q}'\right\}\right)^{-1}(\mathbb{E}\{\mathbf{P}\mathbf{Q}'\})'$ [72],

$$\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}} \middle| \mathbf{V}\right\} = \mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\right\} - \mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{V}'\right\}\left(\mathbb{E}\left\{\mathbf{V}\mathbf{V}'\right\}\right)^{-1}\left(\mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{V}'\right\}\right)'. \tag{A.14}$$

Now we calculate each term on the RHS of (A.14) one by one. Firstly,

$$\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\right\} = \mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}\left(\widetilde{\mathbf{G}}_i^{\text{S}}\right)'\right\} = \sum_{j=1}^{N}\gamma_{ji}\mathbb{E}\left\{\tilde{\mathbf{g}}_{ji}^{\text{S}}\left(\tilde{\mathbf{g}}_{ji}^{\text{S}}\right)'\right\}$$
$$= \sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\mathbf{I}_M. \tag{A.15}$$

Then, using $\tilde{\mathbf{g}}_{nk}^{\text{S}} = \left(\sqrt{\tau_p\rho_{ul}}\mu_n - 1\right)\mathbf{g}_{nk} + \sqrt{\tau_p\rho_{ul}}\mu_n\sum_{j=1,j\neq k}^{K}\mathbf{g}_{nj} + \mu_n\overline{\mathbf{w}}_{pn}$ and $\hat{\mathbf{c}}_n = \sqrt{\eta_n}\mathbf{v}_n$, we get

$$\mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\mathbf{V}'\right\} = \sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\mathbb{E}\left\{\tilde{\mathbf{g}}_{ji}^{\text{S}}\mathbf{v}_j'\right\} = \sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\mathbf{I}_M, \tag{A.16}$$

where $\epsilon_{ji} = \left(\sqrt{\tau_p\rho_{ul}}\mu_j - 1\right)\left(\sqrt{\tau_p\rho_{ul}}\mu_j\right)\beta_{ji} + \left(\sqrt{\tau_p\rho_{ul}}\mu_j\right)^2\sum_{l=1,l\neq i}^{K}\beta_{jl} + \mu_j^2\sigma^2$. Next,

$$\left(\mathbb{E}\left\{\mathbf{V}\mathbf{V}'\right\}\right)^{-1} = \left(\sum_{j=1}^{N}\mathbb{E}\left\{\mathbf{v}_j\mathbf{v}_j'\right\}\right)^{-1} = \left(\sum_{j=1}^{N}\mathbf{I}_M\right)^{-1} = \frac{1}{N}\mathbf{I}_M. \tag{A.17}$$

Substituting (A.15), (A.16) and (A.17) in (A.14), we get,

$$\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}\mathbf{D}_{\gamma_i}^{\frac{1}{2}}\middle|\mathbf{V}\right\} = \left[\left(\sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\right) - \frac{1}{N}\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right]\mathbf{I}_M. \tag{A.18}$$

Substituting (A.18) in (A.13),

$$\mathbb{V}\text{ar}\left\{\left[\mathbf{B}'\overline{\mathbf{n}}^{\text{S}}\right]_n\middle|\mathbf{V}\right\} = \left[\mathbf{B}'\left(\sigma^2\mathbf{I}_M + \lambda^{\text{S}}\mathbf{I}_M\right)\mathbf{B}\right]_{n,n}$$
$$= \left(\sigma^2 + \lambda^{\text{S}}\right)\left[\mathbf{B}'\mathbf{B}\right]_{n,n} \tag{A.19}$$

where $\lambda^{\text{S}} = \sum_{i=1}^{K}\left[\left(\sum_{j=1}^{N}\gamma_{ji}\alpha_{ji}\right) - \frac{1}{N}\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right]$. Therefore, (A.9) can written as

$$\text{SINR}_{nk}^{\text{S-UB}} = \frac{\eta_n\gamma_{nk}}{\sum_{i=k+1}^{K}\eta_n\gamma_{ni} + \left(\sigma^2 + \lambda^{\text{S}}\right)\left[\mathbf{B}'\mathbf{B}\right]_{n,n}}. \tag{A.20}$$

Now, using Jensen's inequality, for any random variable z,

$$\mathbb{E}\left\{\log_2\left\{1+z\right\}\right\} \geq \log_2\left\{1 + \frac{1}{\mathbb{E}\left\{\frac{1}{z}\right\}}\right\}. \tag{A.21}$$

Therefore,

$$\overline{\mathrm{R}}_{nk}^{\mathrm{S}} \geq \delta \log_2 \left\{ 1 + \frac{1}{\mathbb{E}\left\{ \frac{1}{\mathrm{SINR}_{nk}^{\text{S-UB}}} \right\}} \right\}. \tag{A.22}$$

Using the result $\mathbb{E}_{\mathbf{V}}\{[\mathbf{B}'\mathbf{B}]_{i,i}\} = (M-N)^{-1}$ [14],

$$\begin{aligned}
\mathbb{E}\left\{ \frac{1}{\mathrm{SINR}_{nk}^{\text{S-UB}}} \right\} &= \mathbb{E}\left\{ \frac{\sum_{i=k+1}^{K} \eta_n \gamma_{ni}}{\eta_n \gamma_{nk}} + \frac{\left(\sigma^2 + \lambda^{\mathrm{S}}\right)\left[\mathbf{B}'\mathbf{B}\right]_{n,n}}{\eta_n \gamma_{nk}} \right\} \\
&= \frac{\sum_{i=k+1}^{K} \eta_n \gamma_{ni}}{\eta_n \gamma_{nk}} + \frac{\left(\sigma^2 + \lambda^{\mathrm{S}}\right)\mathbb{E}\left\{\left[\mathbf{B}'\mathbf{B}\right]_{n,n}\right\}}{\eta_n \gamma_{nk}} \\
&= \frac{\sum_{i=k+1}^{K} \eta_n \gamma_{ni}}{\eta_n \gamma_{nk}} + \frac{\left(\sigma^2 + \lambda^{\mathrm{S}}\right)}{(M-N)\eta_n \gamma_{nk}} \\
&= \frac{\sigma^2 + \lambda^{\mathrm{S}} + \sum_{i=k+1}^{K} (M-N)\,\eta_n \gamma_{ni}}{(M-N)\ \eta_n\ \gamma_{nk}}.
\end{aligned} \tag{A.23}$$

Therefore, (A.22) can be written as

$$\overline{\mathrm{R}}_{nk}^{\mathrm{S}} \geq \delta \log_2 \left\{ 1 + \frac{(M-N)\ \eta_n\ \gamma_{nk}}{\sigma^2 + \lambda^{\mathrm{S}} + \sum_{k'=k+1}^{K} (M-N)\,\eta_n \gamma_{nk'}} \right\}. \tag{A.24}$$

Using (A.24), and (A.8), we get (2.46).

B

Appendix

B.1 Proof of Theorem 3.1

As (3.26) represents the signal with deterministic channel corrupted by additive non-Gaussian noise [14], the corresponding SE at $(t+u)^{\text{th}}$ symbol instant for UE_{nk} is lower bounded by

$$\mathrm{R}_{nk}^{\text{ZF-I}}[t+u] = \log_2\left(1 + \mathrm{SINR}_{nk}^{\text{ZF-I}}[t+u]\right). \tag{B.1}$$

Since all terms in (3.26) are uncorrelated, $\mathrm{SINR}_{nk}^{\text{ZF-I}}[t+u]$ can be written as

$$\mathrm{SINR}_{nk}^{\text{ZF-I}}[t+u] = \frac{\mathcal{N}_{nk}^{\text{ZF-I}}[t+u]}{\mathcal{D}_{nk}^{\text{ZF-I}}[t+u]}, \tag{B.2}$$

where

$$\mathcal{N}_{nk}^{\text{ZF-I}}[t+u] = \left(\alpha(u)^2 \eta_{nk}\gamma_{nk}[t+u]\right), \tag{B.3}$$

$$\mathcal{D}_{nk}^{\text{ZF-I}}[t+u] = \sum_{i=k+1}^{K} \alpha(u)^2 \eta_{ni}\gamma_{ni}[t+u] + \mathbb{V}\mathrm{ar}\left\{\left[\left(\mathbf{A}^{\text{ZF}}[t+u]\right)' \overline{\mathbf{n}}^{\text{I}}[t+u]\right]_n\right\}. \tag{B.4}$$

To compute $\mathrm{SINR}_{nk}^{\text{ZF-I}}[t+u]$, now we calculate the variance of the third term in (3.26) as

$$\mathbb{V}\mathrm{ar}\left\{\left[\left(\mathbf{A}^{\text{ZF}}[t+u]\right)' \overline{\mathbf{n}}^{\text{I}}[t+u]\right]_n\right\} = \left[\mathbb{C}\mathrm{ov}\left\{\left(\mathbf{A}^{\text{ZF}}[t+u]\right)' \overline{\mathbf{n}}^{\text{I}}[t+u]\right\}\right]_{n,n} \tag{B.5}$$

$$= \left[\mathbb{E}_{\mathbf{Z}[t]} \left\{ \left(\mathbf{A}^{\text{ZF}}[t+u] \right)' \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{I}}[t]} \left\{ \overline{\mathbf{n}}^{\text{I}}[t+u] \left(\overline{\mathbf{n}}^{\text{I}}[t+u] \right)' \middle| \mathbf{Z}[t] \right\} \mathbf{A}^{\text{ZF}}[t+u] \right\} \right]_{n,n}. \tag{B.6}$$

Since $\widetilde{\mathbf{G}}^{\text{I}}[t]$ is independent of $\mathbf{Z}[t]$,

$$\begin{aligned} \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{I}}[t]} \left\{ \overline{\mathbf{n}}^{\text{I}}[t+u] \left(\overline{\mathbf{n}}^{\text{I}}[t+u] \right)' \middle| \mathbf{Z}[t] \right\} &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{I}}[t]} \left\{ \overline{\mathbf{n}}^{\text{I}}[t+u] \left(\overline{\mathbf{n}}^{\text{I}}[t+u] \right)' \right\} \\ &= \sigma^2 \mathbf{I}_M + \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{I}}[t]} \left\{ \alpha(u)^2 \widetilde{\mathbf{G}}^{\text{I}}[t] \mathbf{x}[t+u] \left(\widetilde{\mathbf{G}}^{\text{I}}[t] \mathbf{x}[t+u] \right)' \right\} \\ &\quad + \mathbb{E} \left\{ \mathbf{E}[t+u] \mathbf{x}[t+u] \left(\mathbf{E}[t+u] \mathbf{x}[t+u] \right)' \right\}. \end{aligned} \tag{B.7}$$

Further,

$$\begin{aligned} \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{I}}[t]} \left\{ \widetilde{\mathbf{G}}^{\text{I}}[t] \mathbf{x}[t+u] \left(\widetilde{\mathbf{G}}^{\text{I}}[t] \mathbf{x}[t+u] \right)' \right\} &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{I}}[t]} \left\{ \sum_{i=1}^{K} \widetilde{\mathbf{G}}_i^{\text{I}}[t] \mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u] \left(\widetilde{\mathbf{G}}_i^{\text{I}}[t] \mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u] \right)' \right\} \\ &= \sum_{n=1}^{N} \sum_{i=1}^{K} \left\{ \gamma_{ni}[t+u] \left(\beta_{ni} - \eta_{ni} \right) \mathbf{I}_M \right\}. \end{aligned} \tag{B.8}$$

Similarly,

$$\mathbb{E} \left\{ \mathbf{E}[t+u] \mathbf{x}[t+u] \left(\mathbf{E}[t+u] \mathbf{x}[t+u] \right)' \right\} = \sum_{n=1}^{N} \sum_{i=1}^{K} \left\{ \gamma_{ni}[t+u] \left(1 - \alpha(u)^2 \right) \beta_{ni} \mathbf{I}_M \right\}. \tag{B.9}$$

Substituting (B.7), (B.8) and (B.9) in (B.6), then substituting (B.6) in (B.2) and simplifying using the result that $\mathbb{E}_{\mathbf{Z}[t]} \left\{ \left[\left(\mathbf{A}^{\text{ZF}}[t+u] \right)' \mathbf{A}^{\text{ZF}}[t+u] \right]_{i,i} \right\} = (M-N)^{-1}$ [14], we get (3.28), where $\lambda^{\text{ZF-I}}[t+u] = \sum_{n=1}^{N} \sum_{i=1}^{K} \left\{ \left(\beta_{ni} - \alpha(u)^2 \eta_{ni} \right) \gamma_{ni}[t+u] \right\}$.

B.2 Proof of Theorem 3.2

The terms in (3.39) are correlated. However, if the side information $\Omega[t+u] = \mathbf{V}[t]$ is available, then all three terms in (3.39) are uncorrelated. Thus, using the corresponding result from [14], the SE for UE_{nk} is lower bounded by

$$\widetilde{\text{R}}_{nk}^{\text{ZF-S}}[t+u] = \mathbb{E}_{\mathbf{V}[t]} \left\{ \log_2 \left(1 + \text{SINR}_{nk}^{\text{ZF-S-UB}}[t+u] \right) \right\}, \tag{B.10}$$

where $\text{SINR}_{nk}^{\text{ZF-S-UB}}[t+u]$ can be written as

$$\text{SINR}_{nk}^{\text{ZF-S-UB}}[t+u] = \frac{\mathcal{N}_{nk}^{\text{ZF-S-UB}}[t+u]}{\mathcal{D}_{nk}^{\text{ZF-S-UB}}[t+u]}, \tag{B.11}$$

where

$$\mathcal{N}_{nk}^{\text{ZF-S-UB}}[t+u] = \alpha(u)^2\eta_n\gamma_{nk}[t+u], \tag{B.12}$$

$$\mathcal{D}_{nk}^{\text{ZF-S-UB}}[t+u] = \sum_{i=k+1}^{K}\alpha(u)^2\eta_n\gamma_{ni}[t+u] + \mathbb{V}\text{ar}\left\{\left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'\overline{\mathbf{n}}^{\text{S}}[t+u]\right]_n\middle|\mathbf{V}[t]\right\}. \tag{B.13}$$

Given $\Omega[t+u] = \mathbf{V}[t]$, we now derive the variance of the effective noise term as

$$\begin{aligned}\mathbb{V}\text{ar}\left\{\left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'\overline{\mathbf{n}}^{\text{S}}[t+u]\middle|\mathbf{V}[t]\right]_n\right\} &= \left[\mathbb{C}\text{ov}\left\{\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'\overline{\mathbf{n}}^{\text{S}}[t+u]\middle|\mathbf{V}[t]\right\}\right]_{n,n}\\ &= \left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'\mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}}\left\{\overline{\mathbf{n}}^{\text{S}}[t+u]\left(\overline{\mathbf{n}}^{\text{S}}[t+u]\right)'\middle|\mathbf{V}[t]\right\}\mathbf{B}^{\text{ZF}}[t+u]\right]_{n,n}.\end{aligned} \tag{B.14}$$

Since $\widetilde{\mathbf{G}}^{\text{S}}[t]$ and $\mathbf{V}[t]$ are correlated,

$$\begin{aligned}\mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}[t]}\left\{\overline{\mathbf{n}}^{\text{S}}[t+u]\left(\overline{\mathbf{n}}^{\text{S}}[t+u]\right)'\middle|\mathbf{V}[t]\right\} &= \sigma^2\mathbf{I}_M + \mathbb{E}\left\{\mathbf{E}[t+u]\mathbf{x}[t+u]\left(\mathbf{E}[t+u]\mathbf{x}[t+u]\right)'\right\}\\ &\quad + \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}[t]}\left\{\alpha(u)^2\widetilde{\mathbf{G}}^{\text{S}}[t]\mathbf{x}[t+u]\left(\widetilde{\mathbf{G}}^{\text{S}}[t]\mathbf{x}[t+u]\right)'\middle|\mathbf{V}[t]\right\}.\end{aligned} \tag{B.15}$$

Further,

$$\begin{aligned}\mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}[t]}\left\{\widetilde{\mathbf{G}}^{\text{S}}[t]\mathbf{x}[t+u]\left(\widetilde{\mathbf{G}}^{\text{S}}[t]\mathbf{x}[t+u]\right)'\middle|\mathbf{V}[t]\right\} &= \mathbb{E}_{\widetilde{\mathbf{G}}^{\text{S}}[t]}\left\{\sum_{i=1}^{K}\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}[t+u]\left(\widetilde{\mathbf{G}}_i^{\text{S}}[t]\right)'\middle|\mathbf{V}[t]\right\}\\ &= \sum_{i=1}^{K}\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\middle|\mathbf{V}[t]\right\}.\end{aligned} \tag{B.16}$$

Using the result that $\mathbb{C}\text{ov}\{\mathbf{X}|\mathbf{Y}\} = \mathbb{C}\text{ov}\{\mathbf{X}\} - \mathbb{E}\{\mathbf{X}\mathbf{Y}'\}\left(\mathbb{E}\left\{\mathbf{Y}\mathbf{Y}'\right\}\right)^{-1}(\mathbb{E}\{\mathbf{X}\mathbf{Y}'\})'$ [72],

$$\begin{aligned}\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\middle|\mathbf{V}[t]\right\} &= \mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\right\} - \mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\left(\mathbf{V}[t]\right)'\right\}\\ &\quad \left(\mathbb{E}\left\{\mathbf{V}[t]\left(\mathbf{V}[t]\right)'\right\}\right)^{-1}\left(\mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\left(\mathbf{V}[t]\right)'\right\}\right)'.\end{aligned} \tag{B.17}$$

Now, we need to compute the terms in (B.17). Firstly,

$$\begin{aligned}\mathbb{C}\text{ov}\left\{\widetilde{\mathbf{G}}_i^{\text{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\right\} &= \sum_{j=1}^{N}\gamma_{ji}[t+u]\mathbb{E}\left\{\widetilde{\mathbf{g}}_{ji}^{\text{S}}[t]\left(\widetilde{\mathbf{g}}_{ji}^{\text{S}}[t]\right)'\right\}\\ &= \sum_{j=1}^{N}\gamma_{ji}[t+u]\iota_{ji}\mathbf{I}_M.\end{aligned} \tag{B.18}$$

Using $\widetilde{\mathbf{g}}_{nk}^{\mathrm{S}}[t]$ and $\hat{\mathbf{c}}_n[t]$ as given in Section 3.1.1.2, we get

$$\mathbb{E}\left\{\widetilde{\mathbf{G}}_i^{\mathrm{S}}[t]\mathbf{D}_{\gamma_i}^{\frac{1}{2}}[t+u]\left(\mathbf{V}[t]\right)'\right\} = \sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}[t+u]\mathbb{E}\left\{\widetilde{\mathbf{g}}_{ji}^{\mathrm{S}}[t]\left(\mathbf{v}_j[t]\right)'\right\}$$
$$= \sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}[t+u]\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\mathbf{I}_M, \tag{B.19}$$

where $\epsilon_{ji} = \left(\sqrt{\tau_p\rho_{ul}}\mu_j - 1\right)\left(\sqrt{\tau_p\rho_{ul}}\mu_j\right)\beta_{ji} + \left(\sqrt{\tau_p\rho_{ul}}\mu_j\right)^2\sum_{l=1,l\neq i}^{K}\beta_{jl} + \mu_j^2\sigma^2$. Next,

$$\left(\mathbb{E}\left\{\mathbf{V}[t]\left(\mathbf{V}[t]\right)'\right\}\right)^{-1} = \left(\sum_{j=1}^{N}\mathbb{E}\left\{\mathbf{v}_j[t]\left(\mathbf{v}_j[t]\right)'\right\}\right)^{-1}$$
$$= \left(\sum_{j=1}^{N}\mathbf{I}_M\right)^{-1} = \frac{1}{N}\mathbf{I}_M. \tag{B.20}$$

Substituting (B.18), (B.19) and (B.20) in (B.17), then substituting (B.17) in (B.16), substituting (B.16) in (B.15), substituting (B.15) in (B.14) and finally substituting (B.14) in (B.13), we get

$$\mathcal{D}_{nk}^{\text{ZF-S-UB}}[t+u] = \sum_{i=k+1}^{K}\alpha(u)^2\eta_n\gamma_{ni}[t+u]$$
$$+\left(\sigma^2 + \lambda^{\text{ZF-S}}[t+u] + \omega^{\text{ZF-S}}[t+u]\right)\left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'\mathbf{B}^{\text{ZF}}[t+u]\right]_{n,n}, \tag{B.21}$$

with

$$\lambda^{\text{ZF-S}}[t+u] = \sum_{i=1}^{K}\alpha(u)^2\left[\left(\sum_{j=1}^{N}\gamma_{ji}[t+u]\iota_{ji}\right) - \frac{1}{N}\left(\sum_{j=1}^{N}\gamma_{ji}^{\frac{1}{2}}[t+u]\frac{\epsilon_{ji}}{\sqrt{\eta_j}}\right)^2\right], \tag{B.22}$$

$$\omega^{\text{ZF-S}}[t+u] = \sum_{n=1}^{N}\sum_{i=1}^{K}\left\{\gamma_{ni}[t+u]\left(1-\alpha(u)^2\right)\beta_{ni}\right\}. \tag{B.23}$$

Now, using the Jensen's inequality and using a result that $\mathbb{E}_{\mathbf{Z}[t]}\left\{\left[\left(\mathbf{B}^{\text{ZF}}[t+u]\right)'\mathbf{B}^{\text{ZF}}[t+u]\right]_{i,i}\right\} = (M-N)^{-1}$ [14], we get

$$\widetilde{\mathrm{R}}_{nk}^{\text{ZF-S}}[t+u] \geq \log_2\left\{1 + \left(\frac{1}{\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_{nk}^{\text{ZF-S-UB}}[t+u]}\right\}}\right)\right\}, \tag{B.24}$$

where

$$\mathbb{E}\left\{\frac{1}{\mathrm{SINR}_{nk}^{\text{ZF-S-UB}}[t+u]}\right\} = \frac{\sigma^2 + \lambda^{\text{ZF-S}}[t+u] + \omega^{\text{ZF-S}}[t+u] + \sum_{i=k+1}^{K}(M-N)\eta_n\gamma_{ni[t+u]}}{(M-N)\ \eta_n\ \gamma_{nk}[t+u]}. \tag{B.25}$$

Therefore, (B.10) can be written as

$$\widetilde{\mathrm{R}}_{nk}^{\text{ZF-S}}[t+u] \geq \log_2\left\{1+\widetilde{\mathrm{SINR}}_{nk}^{\text{ZF-S}}[t+u]\right\}, \tag{B.26}$$

where

$$\widetilde{\mathrm{SINR}}_{nk}^{\text{ZF-S}}[t+u] = \frac{\mathcal{N}_{nk}^{\text{ZF-S}}[t+u]}{\mathcal{D}_{nk}^{\text{ZF-S}}[t+u]}, \tag{B.27}$$

where

$$\mathcal{N}_{nk}^{\text{ZF-S}}[t+u] = (M-N)\,\alpha(u)^2\,\eta_n\,\gamma_{nk}[t+u], \tag{B.28}$$

$$\mathcal{D}_{nk}^{\text{ZF-S}}[t+u] = \sigma^2 + \lambda^{\text{ZF-S}}[t+u] + \omega^{\text{ZF-S}}[t+u] + \sum_{k'=k+1}^{K}(M-N)\,\eta_n\gamma_{nk'}[t+u]. \tag{B.29}$$

B.3 Proof of Theorem 3.4

The post-SIC received signal corresponding to UE_{nk} is given in (3.55). Note that all the terms in (3.55) are uncorrelated to each other. Thus, the SE for UE_{nk} is lower bounded by

$$\begin{aligned}\mathrm{R}_{nk}^{\text{MR-I}}[t+u] &= \log_2\left(1+\mathrm{SINR}_{nk}^{\text{MR-I}}[t+u]\right)\\ &= \log_2\left(1+\frac{\mathcal{V}_1}{\sum_{i=2}^{6}\mathcal{V}_i}\right),\end{aligned} \tag{B.30}$$

where $\mathcal{V}_i$ is the variance of the i^{th} term in (3.55), for $i \in \{1,\cdots,6\}$. Now, we calculate the variance of each term in (3.55) one-by-one.

To begin with, the first term in (3.55) is

$$\mathcal{T}_1 = \sqrt{\gamma_{nk}[t+u]}\,\mathbb{E}\left\{\widehat{\mathbf{g}}_{nk}^{\prime}[t]\mathbf{g}_{nk}[t+u]\right\} q_{nk}[t+u]. \tag{B.31}$$

Here,

$$\mathbb{E}\left\{\widehat{\mathbf{g}}_{nk}^{\prime}[t]\mathbf{g}_{nk}[t+u]\right\} = \mathbb{E}\left\{\widehat{\mathbf{g}}_{nk}^{\prime}[t]\left(\alpha(u)\widehat{\mathbf{g}}_{nk}[t] - \alpha(u)\widetilde{\mathbf{g}}_{nk}^{\text{I}}[t] - \mathbf{e}_{nk}[t+u]\right)\right\} = \alpha(u)M\eta_{nk}. \tag{B.32}$$

Thus,

$$\mathcal{V}_1 = \mathbb{V}\mathrm{ar}\left\{\mathcal{T}_1\right\} = \gamma_{nk}[t+u]\alpha(u)^2M^2\eta_{nk}^2. \tag{B.33}$$

Second term in (3.55) is

$$\mathcal{T}_2 = \sqrt{\gamma_{nk}[t+u]}\ \left(\widehat{\mathbf{g}}'_{nk}[t]\mathbf{g}_{nk}[t+u] - \mathbb{E}\left\{\widehat{\mathbf{g}}'_{nk}[t]\mathbf{g}_{nk}[t+u]\right\}\right) q_{nk}[t+u]. \tag{B.34}$$

Thus,

$$\mathcal{V}_2 = \gamma_{nk}[t+u]\mathbb{V}\mathrm{ar}\left\{\widehat{\mathbf{g}}'_{nk}[t]\mathbf{g}_{nk}[t+u]\right\} = \mathcal{V}_{21} - \mathcal{V}_{22}, \tag{B.35}$$

where

$$\mathcal{V}_{21} = \gamma_{nk}[t+u]\mathbb{E}\left\{\left|\widehat{\mathbf{g}}'_{nk}[t]\mathbf{g}_{nk}[t+u]\right|^2\right\} \tag{B.36}$$

$$\mathcal{V}_{22} = \gamma_{nk}[t+u]\left(\mathbb{E}\left\{\widehat{\mathbf{g}}'_{nk}[t]\mathbf{g}_{nk}[t+u]\right\}\right)^2. \tag{B.37}$$

Now, substituting for $\mathbf{g}_{nk}[t+u]$, we get $\mathcal{V}_{21}$ as

$$\begin{aligned}
\mathcal{V}_{21} &= \gamma_{nk}[t+u]\mathbb{E}\left\{\left\|\widehat{\mathbf{g}}'_{nk}[t]\left(\alpha(u)\widehat{\mathbf{g}}_{nk}[t] - \alpha(u)\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t] - \mathbf{e}_{nk}[t+u]\right)\right\|^2\right\} \\
&= \gamma_{nk}[t+u]\mathbb{E}\left\{\alpha(u)^2\left\|\widehat{\mathbf{g}}_{nk}[t]\right\|^4 + \alpha(u)^2\left\|\widehat{\mathbf{g}}'_{nk}[t]\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk}[t]\right\|^2 + \left\|\widehat{\mathbf{g}}'_{nk}[t]\mathbf{e}_{nk}[t+u]\right\|^2\right\} \\
&= \gamma_{nk}[t+u]\left(\beta_{nk} + \alpha(u)^2 M\eta_{nk}\right) M\eta_{nk}.
\end{aligned} \tag{B.38}$$

From (B.32), we know that $\mathcal{V}_{22} = \gamma_{nk}[t+u]\alpha(u)^2M^2\eta_{nk}^2$. Therefore, $\mathcal{V}_2 = \mathcal{V}_{21} - \mathcal{V}_{22}$ can be simplified to get

$$\mathcal{V}_2 = \gamma_{nk}[t+u]\beta_{nk}M\eta_{nk}. \tag{B.39}$$

Next, the third term in (3.55) is given by

$$\mathcal{T}_3 = \sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\widehat{\mathbf{g}}'_{nk}[t]\left(\mathbf{g}_{nk'}[t+u] - \alpha(u)\widehat{\mathbf{g}}_{nk'}[t]\right) q_{nk'}[t+u]. \tag{B.40}$$

We have

$$\mathbf{g}_{nk'}[t+u] - \alpha(u)\widehat{\mathbf{g}}_{nk'}[t] = -\alpha(u)\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk'}[t] - \mathbf{e}_{nk'}[t+u]. \tag{B.41}$$

Thus, $\mathcal{V}_3$ can be computed as

$$\mathcal{V}_3 = \sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\left(\alpha(u)^2\mathbb{E}\left\{\widehat{\mathbf{g}}'_{nk}[t]\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk'}[t]\left(\widetilde{\mathbf{g}}^{\mathrm{I}}_{nk'}[t]\right)'\widehat{\mathbf{g}}_{nk}[t]\right\} + \mathbb{E}\left\{\widehat{\mathbf{g}}'_{nk}[t]\mathbf{e}_{nk'}[t]\mathbf{e}'_{nk'}[t]\widehat{\mathbf{g}}_{nk}[t]\right\}\right)$$

$$= \sum_{k'=1}^{k-1} \gamma_{nk'}[t+u]\left(\beta_{nk'} - \alpha(u)^2\eta_{nk'}\right) M\eta_{nk}. \tag{B.42}$$

Further, the fourth term in (3.55) is

$$\mathcal{T}_4 = \sum_{k''=k+1}^{K} \sqrt{\gamma_{nk''}[t+u]}\widehat{\mathbf{g}}_{nk}^{\prime}[t]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u]. \tag{B.43}$$

Thus, substituting for $\mathbf{g}_{nk}[t+u]$, $\mathcal{V}_4$ can be computed as

$$\mathcal{V}_4 = \sum_{k''=k+1}^{K}\gamma_{nk''}[t+u]\left(\mathbb{E}\left\{\alpha(u)^2\left\|\widehat{\mathbf{g}}_{nk}^{\prime}[t]\widehat{\mathbf{g}}_{nk''}[t]\right\|^2 + \alpha(u)^2\left\|\widehat{\mathbf{g}}_{nk}^{\prime}[t]\widehat{\mathbf{g}}_{nk''}^{\prime}[t]\right\|^2 + \left\|\widehat{\mathbf{g}}_{nk}^{\prime}[t]\mathbf{e}_{nk''}[t]\right\|^2\right\}\right)$$

$$= \sum_{k''=k+1}^{K}\gamma_{nk''}[t+u]\left(\alpha(u)^2 M\eta_{nk''} + \beta_{nk''}\right) M\eta_{nk}. \tag{B.44}$$

Now, the 5$^{\text{th}}$ term in (3.55) is

$$\mathcal{T}_5 = \sum_{n'\neq n, n'=1}^{N}\sum_{k'''=1}^{K} \sqrt{\gamma_{n'k'''}[t+u]}\widehat{\mathbf{g}}_{nk}^{\prime}[t]\mathbf{g}_{n'k'''}[t+u]q_{n'k'''}[t+u]. \tag{B.45}$$

Therefore,

$$\mathcal{V}_5 = \sum_{n'\neq n, n'=1}^{N}\sum_{k'''=1}^{K}\gamma_{n'k'''}[t+u]\beta_{n'k'''}M\eta_{nk}. \tag{B.46}$$

Similarly, $\mathcal{V}_6 = \sigma^2 M\eta_{nk}$.

B.4 Proof of Theorem 3.5

The post-SIC received signal corresponding to UE_{nk} is given in (3.65). Note that all six terms on the RHS of (3.65) are uncorrelated to each other. Therefore, the SE for UE_{nk} at $(t+u)^{\text{th}}$ symbol instant is lower bounded by

$$\begin{aligned}\mathrm{R}_{nk}^{\text{MR-S}}[t+u] &= \log_2\left(1 + \mathrm{SINR}_{nk}^{\text{MR-S}}[t+u]\right)\\ &= \log_2\left(1 + \frac{\mathcal{X}_1}{\sum_{i=2}^{6}\mathcal{X}_i}\right),\end{aligned} \tag{B.47}$$

where $\mathcal{X}_i$ is the variance of the i^{th} term in (3.65), for $i \in \{1, \cdots, 6\}$. Now, we calculate the variance of each term in (3.65) one-by-one.

The desired signal term in (3.65) is

$$\mathcal{P}_1 = \sqrt{\gamma_{nk}[t+u]}\,\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\} q_{nk}[t+u]. \tag{B.48}$$

Therefore,

$$\mathcal{X}_1 = \gamma_{nk}[t+u]\left|\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\}\right|^2. \tag{B.49}$$

Here,

$$\begin{aligned}\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\} &= \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\left(\alpha(u)\widehat{\mathbf{c}}_n[t] - \alpha(u)\widetilde{\mathbf{g}}_{nk}^{\mathrm{S}}[t] - \mathbf{e}_{nk}[t+u]\right)\right\}\\ &= \alpha(u)M\eta_n - \alpha(u)\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widetilde{\mathbf{g}}_{nk}^{\mathrm{S}}[t]\right\}.\end{aligned} \tag{B.50}$$

Now, $\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widetilde{\mathbf{g}}_{nk}^{\mathrm{S}}[t]\right\}$ is computed as

$$\begin{aligned}\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widetilde{\mathbf{g}}_{nk}^{\mathrm{S}}[t]\right\} &= \mathbb{E}\left\{\left(\bar{b}\sum_{i=1}^{K}\mathbf{g}_{ni}'[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}'[t]\right)\left(\bar{a}\mathbf{g}_{nk}[t] + \bar{b}\sum_{j\neq k}^{K}\mathbf{g}_{nj}[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}[t]\right)\right\}\\ &= \mathbb{E}\left\{\bar{a}\bar{b}\mathbf{g}_{nk}'[t]\mathbf{g}_{nk}[t] + \bar{b}^2\sum_{j\neq k}^{K}\mathbf{g}_{nj}'[t]\mathbf{g}_{nj}[t] + \bar{c}^2\,\overline{\mathbf{w}}_{pn}'[t]\overline{\mathbf{w}}_{pn}[t]\right\} = \bar{\epsilon}_{nk}M,\end{aligned} \tag{B.51}$$

where $\bar{a} = \sqrt{\tau_p\rho_{ul}}\mu_n - 1$, $\bar{b} = \sqrt{\tau_p\rho_{ul}}\mu_n$, $\bar{c} = \mu_n$ and $\bar{\epsilon}_{nk} = \left(\bar{a}\bar{b}\beta_{nk} + \bar{b}^2\sum_{j\neq k}^{K}\beta_{nj} + \bar{c}^2\sigma^2\right)$. Substituting (B.51) in (B.50), and then substituting (B.50) in (B.49), we get

$$\mathcal{X}_1 = \gamma_{nk}[t+u]\alpha(u)^2M^2\left|\eta_n - \bar{\epsilon}_{nk}\right|^2. \tag{B.52}$$

Now, the second term in (3.65) is

$$\mathcal{P}_2 = \sqrt{\gamma_{nk}[t+u]}\left(\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u] - \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\}\right) q_{nk}[t+u]. \tag{B.53}$$

Thus,

$$\mathcal{X}_2 = \gamma_{nk}[t+u]\left[\mathbb{E}\left\{\left|\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right|^2\right\} - \left(\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right\}\right)^2\right]. \tag{B.54}$$

Let $\mathcal{A}_{nk} = \mathbb{E}\left\{\left|\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right|^2\right\}$. Therefore, $\mathcal{A}_{nk}$ is computed as

$$\begin{aligned}\mathcal{A}_{nk} &= \mathbb{E}\left\{\left|\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\right|^2\right\}\\ &= \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk}[t+u]\mathbf{g}_{nk}'[t+u]\widehat{\mathbf{c}}_n[t]\right\}\end{aligned}$$

$$= \alpha(u)^2\mathbb{E}\left\{\left\|\widehat{\mathbf{c}}_n[t]\right\|^4\right\} - 2\alpha(u)^2\Re\left[\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widehat{\mathbf{c}}_n[t]\left(\widetilde{\mathbf{g}}_{nk}^{S}[t]\right)'\widehat{\mathbf{c}}_n[t]\right\}\right]$$
$$+ \alpha(u)^2\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widetilde{\mathbf{g}}_{nk}^{S}[t]\left(\widetilde{\mathbf{g}}_{nk}^{S}[t]\right)'\widehat{\mathbf{c}}_n[t]\right\} + \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{e}_{nk}[t+u]\mathbf{e}_{nk}'[t+u]\widehat{\mathbf{c}}_n[t]\right\}$$
$$= \mathcal{A}_{nk}^1 - \mathcal{A}_{nk}^2 + \mathcal{A}_{nk}^3 + \mathcal{A}_{nk}^4, \tag{B.55}$$

where $\mathcal{A}_{nk}^1$, $\mathcal{A}_{nk}^2$, $\mathcal{A}_{nk}^3$ and $\mathcal{A}_{nk}^4$ are given by

$$\mathcal{A}_{nk}^1 = \alpha(u)^2\mathbb{E}\left\{\left\|\widehat{\mathbf{c}}_n[t]\right\|^4\right\} = \alpha(u)^2 M(M+1)\eta_n^2, \tag{B.56}$$
$$\mathcal{A}_{nk}^2 = 2\alpha(u)^2\Re\left[\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widehat{\mathbf{c}}_n[t]\left(\widetilde{\mathbf{g}}_{nk}^{S}[t]\right)'\widehat{\mathbf{c}}_n[t]\right\}\right], \tag{B.57}$$
$$\mathcal{A}_{nk}^3 = \alpha(u)^2\mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\widetilde{\mathbf{g}}_{nk}^{S}[t]\left(\widetilde{\mathbf{g}}_{nk}^{S}[t]\right)'\widehat{\mathbf{c}}_n[t]\right\}, \tag{B.58}$$
$$\mathcal{A}_{nk}^4 = \mathbb{E}\left\{\widehat{\mathbf{c}}_n'[t]\mathbf{e}_{nk}[t+u]\mathbf{e}_{nk}'[t+u]\widehat{\mathbf{c}}_n[t]\right\} = \left(1-\alpha(u)^2\right)\eta_n M\beta_{nk}. \tag{B.59}$$

Further, $\mathcal{A}_{nk}^2$ and $\mathcal{A}_{nk}^3$ can be computed as

$$\mathcal{A}_{nk}^2 = 2\alpha(u)^2\Re\left[\mathbb{E}\left\{\left(\bar{b}\sum_{i=1}^{K}\mathbf{g}_{ni}'[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}'[t]\right)\left(\bar{b}\sum_{i=1}^{K}\mathbf{g}_{ni}[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}[t]\right)\right.\right.$$
$$\left.\left.\left(\bar{a}\mathbf{g}_{nk}'[t] + \bar{b}\sum_{j\neq k}^{K}\mathbf{g}_{nj}'[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}'[t]\right)\left(\bar{b}\sum_{i=1}^{K}\mathbf{g}_{ni}[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}[t]\right)\right\}\right]$$
$$= 2\alpha(u)^2\Re\left\{\left[\bar{a}\bar{b}^3\beta_{nk}^2 + \bar{b}^4\sum_{j\neq k}^{K}\beta_{nj}^2 + \bar{c}^4\sigma^4\right]M(M+1) + \left[\bar{a}\bar{b}^3\sum_{j\neq k}^{K}\beta_{nj}\beta_{nk} + \bar{b}^2\bar{c}^2\sum_{i=1}^{K}\beta_{ni}\sigma^2\right.\right.$$
$$\left. + \bar{a}\bar{b}\bar{c}^2\beta_{nk}\sigma^2 + \bar{b}^2\bar{c}^2\sum_{j\neq k}^{K}\beta_{nj}\sigma^2 + \bar{b}^4\sum_{j\neq k}^{K}\beta_{nj}\beta_{nk} + \bar{b}^4\sum_{i\neq j,k}^{K}\sum_{j\neq i,k}^{K}\beta_{ni}\beta_{nj}\right]M^2$$
$$\left. + \left[\bar{a}\bar{b}\bar{c}^2\beta_{nk}\sigma^2 + \bar{b}^2\bar{c}^2\sum_{j\neq k}^{K}\beta_{nj}\sigma^2 + \bar{b}^2\bar{c}^2\sum_{i=1}^{K}\beta_{ni}\sigma^2\right]M\right\}, \tag{B.60}$$

and

$$\mathcal{A}_{nk}^3 = \alpha(u)^2\mathbb{E}\left\{\left(\bar{b}\sum_{i=1}^{K}\mathbf{g}_{ni}'[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}'[t]\right)\left(\bar{a}\mathbf{g}_{nk}[t] + \bar{b}\sum_{j\neq k}^{K}\mathbf{g}_{nj}[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}[t]\right)\right.$$
$$\left.\left(\bar{a}\mathbf{g}_{nk}'[t] + \bar{b}\sum_{j\neq k}^{K}\mathbf{g}_{nj}'[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}'[t]\right)\left(\bar{b}\sum_{i=1}^{K}\mathbf{g}_{ni}[t] + \bar{c}\,\overline{\mathbf{w}}_{pn}[t]\right)\right\}$$
$$= \alpha(u)^2\Re\left\{\left[\bar{a}^2\bar{b}^2\beta_{nk}^2 + \bar{b}^4\sum_{j\neq k}^{K}\beta_{nj}^2 + \bar{c}^4\sigma^4\right]M(M+1) + \left[2\bar{a}\bar{b}^3\sum_{j\neq k}^{K}\beta_{nj}\beta_{nk} + 2\bar{a}\bar{b}\bar{c}^2\beta_{nk}\sigma^2\right.\right.$$

$$+2\bar{b}^2\bar{c}^2\sum_{j\neq k}^{K}\beta_{nj}\sigma^2+\bar{b}^4\sum_{i\neq j,k}^{K}\sum_{j\neq i,k}^{K}\beta_{ni}\beta_{nj}\Bigg]M^2+\Bigg[\bar{a}^2\bar{b}^2\sum_{j\neq k}^{K}\beta_{nj}\beta_{nk}+\bar{b}^4\sum_{i\neq j,k}^{K}\sum_{j\neq i,k}^{K}\beta_{ni}\beta_{nj}$$
$$+\bar{b}^2\bar{c}^2\sum_{i=1}^{K}\beta_{ni}\sigma^2+\bar{a}^2\bar{c}^2\beta_{nk}\sigma^2+\bar{b}^2\bar{c}^2\sum_{j\neq k}^{K}\beta_{nj}\sigma^2+\bar{b}^4\sum_{j\neq k}^{K}\beta_{nj}\beta_{nk}\Bigg]M\Bigg\}. \quad \text{(B.61)}$$

Substituting $\mathcal{A}^1_{nk}$, $\mathcal{A}^2_{nk}$, $\mathcal{A}^3_{nk}$ and $\mathcal{A}^4_{nk}$ in (B.55) and then substituting (B.55) and $\mathcal{X}_1$ in (B.54), we get $\mathcal{X}_2$.

Now, the third term in (3.65) is given by

$$\mathcal{P}_3=\sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\widehat{\mathbf{c}}'_n[t]\left(\mathbf{g}_{nk'}[t+u]-\alpha(u)\widehat{\mathbf{c}}_n[t]\right)q_{nk'}[t+u]. \quad \text{(B.62)}$$

Here,

$$\mathbf{g}_{nk'}[t+u]-\alpha(u)\widehat{\mathbf{c}}_n[t]=\alpha(u)\widehat{\mathbf{c}}_n[t]-\alpha(u)\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]-\mathbf{e}_{nk'}[t+u]-\alpha(u)\widehat{\mathbf{c}}_n[t]$$
$$=-\alpha(u)\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]-\mathbf{e}_{nk'}[t+u]. \quad \text{(B.63)}$$

Thus,

$$\mathcal{P}_3=\sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+u]}\widehat{\mathbf{c}}'_n[t]\left(-\alpha(u)\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]-\mathbf{e}_{nk'}[t+u]\right)q_{nk'}[t+u]. \quad \text{(B.64)}$$

Thus, $\mathcal{X}_3$ can be computed as

$$\mathcal{X}_3=\sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\mathbb{E}\left\{\widehat{\mathbf{c}}'_n[t]\left(-\alpha(u)\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]-\mathbf{e}_{nk'}[t+u]\right)\left(-\alpha(u)\left(\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]\right)'-\mathbf{e}'_{nk'}[t+u]\right)\widehat{\mathbf{c}}_n[t]\right\}$$
$$=\sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\left[\alpha(u)^2\mathbb{E}\left\{\widehat{\mathbf{c}}'_n[t]\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]\left(\widehat{\mathbf{c}}'_n[t]\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]\right)'\right\}+\mathbb{E}\left\{\widehat{\mathbf{c}}'_n[t]\mathbf{e}_{nk'}[t+u]\left(\widehat{\mathbf{c}}'_n[t]\mathbf{e}_{nk'}[t+u]\right)'\right\}\right]. \quad \text{(B.65)}$$

where $\alpha(u)^2\mathbb{E}\left\{\widehat{\mathbf{c}}'_n[t]\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]\left(\widetilde{\mathbf{g}}^{\mathrm{S}}_{nk'}[t]\right)'\widehat{\mathbf{c}}_n[t]\right\}$ can be obtained by replacing k with k' in $\mathcal{A}^3_{nk}$ given in (B.58), and we denote it as $\mathcal{A}^3_{nk'}$. Moreover,

$$\mathbb{E}\left\{\widehat{\mathbf{c}}'_n[t]\mathbf{e}_{nk'}[t+u]\mathbf{e}'_{nk'}[t+u]\widehat{\mathbf{c}}_n[t]\right\}=\left(1-\alpha(u)^2\right)\eta_n M\beta_{nk'}. \quad \text{(B.66)}$$

Thus,

$$\mathcal{X}_3=\sum_{k'=1}^{k-1}\gamma_{nk'}[t+u]\left(\mathcal{A}^3_{nk'}+\left(1-\alpha(u)^2\right)\eta_n M\beta_{nk'}\right). \quad \text{(B.67)}$$

Next, the fourth and fifth terms in (3.65) are

$$\mathcal{P}_4 = \sum_{k''=k+1}^{K} \sqrt{\gamma_{nk''}[t+u]}\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk''}[t+u]q_{nk''}[t+u], \tag{B.68}$$

$$\mathcal{P}_5 = \sum_{n'\neq n,n'=1}^{N}\sum_{k'''=1}^{K} \sqrt{\gamma_{n'k'''}[t+u]}\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{n'k'''}[t+u]q_{n'k'''}[t+u]. \tag{B.69}$$

Thus, $\mathcal{X}_4$, $\mathcal{X}_5$ can be obtained as

$$\mathcal{X}_4 = \sum_{k''=k+1}^{K} \gamma_{nk''}[t+u]\mathcal{A}_{nk''}, \tag{B.70}$$

$$\mathcal{X}_5 = \sum_{n'\neq n,n'=1}^{N}\sum_{k'''=1}^{K} \gamma_{n'k'''}[t+u]\eta_n M\beta_{n'k'''}, \tag{B.71}$$

where $\mathcal{A}_{nk''} = \mathbb{E}\left\{\left|\widehat{\mathbf{c}}_n'[t]\mathbf{g}_{nk''}[t+u]\right|^2\right\}$ and it is obtained by replacing k with k'' in $\mathcal{A}_{nk}$ given in (B.55).

Finally, the sixth term in (3.65) is $\mathcal{P}_6 = \widehat{\mathbf{c}}_n'[t]\mathbf{w}[t+u]$. Therefore, $\mathcal{X}_6 = \eta_n M\sigma^2$.

B.5 Proof of Lemma 3.1

We know that the observation vector is orthogonal to the error in prediction. Using this orthogonality property of linear prediction, we get

$$\mathbb{E}\left\{\left(\mathbf{g}_{nk}[t+1] - \mathcal{P}_{nk}\widetilde{\mathbf{y}}_{pn}[t]\right)\widetilde{\mathbf{y}}_{pn}'[t]\right\} = \mathbf{0}. \tag{B.72}$$

After expansion, we get

$$\mathbb{E}\left\{\mathbf{g}_{nk}[t+1]\widetilde{\mathbf{y}}_{pn}'[t]\right\} = \mathcal{P}_{nk}\mathbb{E}\left\{\widetilde{\mathbf{y}}_{pn}[t]\widetilde{\mathbf{y}}_{pn}'[t]\right\}. \tag{B.73}$$

Therefore,

$$\mathcal{P}_{nk} = \mathbf{R}_{\mathbf{g}\widetilde{\mathbf{Y}}}[1]\mathbf{R}_{\widetilde{\mathbf{Y}}}^{-1}[0], \tag{B.74}$$

where

$$\mathbf{R}_{\mathbf{g}\widetilde{\mathbf{Y}}}[1] \triangleq \mathbb{E}\left\{\mathbf{g}_{nk}[t+1]\widetilde{\mathbf{y}}_{pn}'[t]\right\}, \tag{B.75}$$

$$\mathbf{R}_{\widetilde{\mathbf{Y}}}[0] \triangleq \mathbb{E}\left\{\widetilde{\mathbf{y}}_{pn}[t]\widetilde{\mathbf{y}}_{pn}'[t]\right\}. \tag{B.76}$$

Here,

$$\begin{aligned}\mathbf{R}_{\mathbf{g}\widetilde{\mathbf{y}}}[k+1] &\triangleq \mathbb{E}\left\{\mathbf{g}_{nk}[t+1]\overline{\mathbf{y}}'_{pn}[t-k]\right\}\\ &= \alpha^{|k+1|}\mathbf{I}_M. \end{aligned} \tag{B.77}$$

Therefore,

$$\mathbf{R}_{\mathbf{g}\widetilde{\mathbf{Y}}}[1] = \alpha\left[\delta(\alpha,p)\otimes\mathbf{I}_M\right]. \tag{B.78}$$

Next,

$$\mathbf{R}_{\overline{\mathbf{y}}}[k-q] \triangleq \mathbf{E}\left\{\overline{\mathbf{y}}_{pn}[t-q]\overline{\mathbf{y}}'_{pn}[t-k]\right\}. \tag{B.79}$$

Therefore,

$$\begin{aligned}\mathbf{R}_{\widetilde{\mathbf{Y}}}[0] &= \mathbb{E}\left\{\widetilde{\mathbf{y}}_{pn}[t]\widetilde{\mathbf{y}}'_{pn}[t]\right\}\\ &= \begin{bmatrix} \mathbf{R}_{\overline{\mathbf{y}}}[0] & \mathbf{R}_{\overline{\mathbf{y}}}[1] & \cdots & \mathbf{R}_{\overline{\mathbf{y}}}[p] \\ \mathbf{R}_{\overline{\mathbf{y}}}[1] & \mathbf{R}_{\overline{\mathbf{y}}}[0] & \cdots & \mathbf{R}_{\overline{\mathbf{y}}}[p-1] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{R}_{\overline{\mathbf{y}}}[p] & \mathbf{R}_{\overline{\mathbf{y}}}[p-1] & \cdots & \mathbf{R}_{\overline{\mathbf{y}}}[0] \end{bmatrix}\\ &= \mathbf{T}_{nk}(\alpha,p)\end{aligned} \tag{B.80}$$

Finally, substituting (B.80) and (B.78) in (B.74), we get (3.79).

B.6 Proof of Theorem 3.6

As (3.88) represents the signal with deterministic channel corrupted by non-Gaussian additive white noise, the corresponding SE at $(t+1)^{\text{th}}$ symbol instant for UE_{nk} is lower bounded by [14]

$$\mathrm{R}_{nk}^{\text{ZF-I-P}}[t+1] = \log_2\left(1+\mathrm{SINR}_{nk}^{\text{ZF-I-P}}[t+1]\right). \tag{B.81}$$

Since all terms in (3.88) are uncorrelated, $\mathrm{SINR}_{nk}^{\text{ZF-I-P}}[t+1]$ can be written as

$$\mathrm{SINR}_{nk}^{\text{ZF-I-P}}[t+1] = \frac{\theta_{nk}(\alpha,p)\gamma_{nk}[t+u]}{\sum_{i=k+1}^{K}\theta_{ni}(\alpha,p)\gamma_{ni}[t+1] + \mathbb{V}\mathrm{ar}\left\{\left[\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\overline{\mathbf{n}}^{\text{P}}[t+1]\right]_n\right\}}. \tag{B.82}$$

To compute $\text{SINR}_{nk}^{\text{ZF-I-P}}[t+1]$, now we calculate the variance of the third term in (3.88) as:

$$\begin{aligned}\mathbb{V}\text{ar}\left\{\left[\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\overline{\mathbf{n}}^{\text{P}}[t+1]\right]_n\right\} &= \left[\mathbb{C}\text{ov}\left\{\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\overline{\mathbf{n}}^{\text{P}}[t+1]\right\}\right]_{n,n}\\ &= \left[\mathbb{E}_{\mathbf{H}[t+1]}\left\{\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\mathbb{E}_{\check{\mathbf{E}}[t+1]}\left\{\overline{\mathbf{n}}^{\text{P}}[t+1]\left(\overline{\mathbf{n}}^{\text{P}}[t+1]\right)'\middle|\mathbf{H}[t+1]\right\}\mathbf{F}^{\text{ZF-I-P}}[t+1]\right\}\right]_{n,n}. \quad \text{(B.83)}\end{aligned}$$

Since $\check{\mathbf{E}}[t+1]$ is independent of $\mathbf{H}[t+1]$,

$$\begin{aligned}\mathbb{E}_{\check{\mathbf{E}}[t+1]}\left\{\overline{\mathbf{n}}^{\text{P}}[t+1]\left(\overline{\mathbf{n}}^{\text{P}}[t+1]\right)'\middle|\mathbf{H}[t+1]\right\} &= \mathbb{E}_{\check{\mathbf{E}}[t+1]}\left\{\overline{\mathbf{n}}^{\text{P}}[t+1]\left(\overline{\mathbf{n}}^{\text{P}}[t+1]\right)'\right\}\\ &= \sigma^2\mathbf{I}_M + \mathbb{E}_{\check{\mathbf{E}}[t+1]}\left\{\check{\mathbf{E}}[t+1]\mathbf{x}[t+1]\left(\check{\mathbf{E}}[t+1]\mathbf{x}[t+1]\right)'\right\}\\ &= \sigma^2\mathbf{I}_M + \sum_{n=1}^{N}\sum_{i=1}^{K}\left\{\gamma_{ni}[t+u]\left(\beta_{ni}-\theta_{ni}(\alpha,p)\right)\mathbf{I}_M\right\}. \quad \text{(B.84)}\end{aligned}$$

Substituting (B.84) in (B.83), then substituting (B.83) in (B.82) and simplifying using the results that

$$\mathbb{E}_{\mathbf{H}[t+1]}\left\{\left[\left(\mathbf{F}^{\text{ZF-I-P}}[t+1]\right)'\mathbf{F}^{\text{ZF-I-P}}[t+1]\right]_{i,i}\right\} = (M-N)^{-1},$$

we get (3.90), where $\lambda^{\text{ZF-I-P}}[t+1] = \sum_{n=1}^{N}\sum_{i=1}^{K}\{(\beta_{ni}-\theta_{ni}(\alpha,p))\,\gamma_{ni}[t+1]\}$.

B.7 Proof of Theorem 3.7

The post-SIC received signal corresponding to UE_{nk} is given in (3.96). Note that all the terms in (3.96) are uncorrelated to each other. Thus, the SE for UE_{nk} is lower bounded by

$$\begin{aligned}\text{R}_{nk}^{\text{MR-I-P}}[t+1] &= \log_2\left(1+\text{SINR}_{nk}^{\text{MR-I-P}}[t+1]\right)\\ &= \log_2\left(1+\frac{\mathcal{W}_1}{\sum_{i=2}^{6}\mathcal{W}_i}\right), \quad \text{(B.85)}\end{aligned}$$

where $\mathcal{W}_i$ is the variance of the i^{th} term in (3.96), for $i \in \{1,\cdots,6\}$. Now, we calculate the variance of each term in (3.96) one-by-one.

To begin with, the first term in (3.96) is

$$\overline{\mathcal{T}}_1 = \sqrt{\gamma_{nk}[t+1]}\,\mathbb{E}\left\{\overline{\mathbf{g}}_{nk}'[t+1]\mathbf{g}_{nk}[t+1]\right\}q_{nk}[t+1]. \quad \text{(B.86)}$$

Here,

$$\mathbb{E}\left\{\overline{\mathbf{g}}'_{nk}[t+1]\mathbf{g}_{nk}[t+1]\right\} = \mathbb{E}\left\{\overline{\mathbf{g}}'_{nk}[t+1]\left(\mathbf{g}_{nk}[t+1] + \check{\mathbf{e}}_{nk}[t+1]\right)\right\} = M\theta_{nk}. \tag{B.87}$$

Thus,

$$\mathcal{W}_1 = \mathbb{V}\mathrm{ar}\left\{\overline{\mathcal{T}}1\right\} = \gamma_{nk}[t+1]M^2\theta_{nk}^2. \tag{B.88}$$

Second term in (3.96) is

$$\overline{\mathcal{T}}_2 = \sqrt{\gamma_{nk}[t+1]}\left(\overline{\mathbf{g}}'_{nk}[t+1]\mathbf{g}_{nk}[t+1] - \mathbb{E}\left\{\overline{\mathbf{g}}'_{nk}[t+1]\mathbf{g}_{nk}[t+1]\right\}\right)q_{nk}[t+1]. \tag{B.89}$$

Thus,

$$\mathcal{W}_2 = \gamma_{nk}[t+1]\mathbb{V}\mathrm{ar}\left\{\overline{\mathbf{g}}'_{nk}[t+1]\mathbf{g}_{nk}[t+1]\right\} = \mathcal{W}_{21} - \mathcal{W}_{22}, \tag{B.90}$$

where

$$\mathcal{W}_{21} = \gamma_{nk}[t+1]\mathbb{E}\left\{\left|\overline{\mathbf{g}}'_{nk}[t+1]\mathbf{g}_{nk}[t+1]\right|^2\right\} \tag{B.91}$$

$$\mathcal{W}_{22} = \gamma_{nk}[t+1]\left(\mathbb{E}\left\{\overline{\mathbf{g}}'_{nk}[t+1]\mathbf{g}_{nk}[t+1]\right\}\right)^2. \tag{B.92}$$

Now, substituting for $\mathbf{g}_{nk}[t+1]$, we get $\mathcal{W}_{21}$ as

$$\begin{aligned}\mathcal{W}_{21} &= \gamma_{nk}[t+1]\mathbb{E}\left\{\left\|\overline{\mathbf{g}}'_{nk}[t+1]\left(\overline{\mathbf{g}}_{nk}[t+1] + \check{\mathbf{e}}_{nk}[t+1]\right)\right\|^2\right\} \\ &= \gamma_{nk}[t+1]\mathbb{E}\left\{\left\|\overline{\mathbf{g}}_{nk}[t+1]\right\|^4 + \left\|\widehat{\mathbf{g}}'_{nk}[t+1]\right\|^2\right\} \\ &= \gamma_{nk}[t+1]\left(\beta_{nk} + M\theta_{nk}\right)M\theta_{nk}.\end{aligned} \tag{B.93}$$

From (B.87), we know that $\mathcal{W}_{22} = \gamma_{nk}[t+1]M^2\theta_{nk}^2$. Therefore, $\mathcal{W}_2 = \mathcal{W}_{21} - \mathcal{W}_{22}$ can be simplified to get

$$\mathcal{W}_2 = \gamma_{nk}[t+1]\beta_{nk}M\theta_{nk}. \tag{B.94}$$

Next, the third term in (3.96) is given by

$$\overline{\mathcal{T}}_3 = \sum_{k'=1}^{k-1}\sqrt{\gamma_{nk'}[t+1]}\overline{\mathbf{g}}'_{nk}[t+1]\left(\mathbf{g}_{nk'}[t+1] - \overline{\mathbf{g}}_{nk'}[t+1]\right)q_{nk'}[t+1]. \tag{B.95}$$

We have

$$\mathbf{g}_{nk'}[t+u] - \overline{\mathbf{g}}_{nk'}[t] = -\check{\mathbf{e}}_{nk}[t+1]. \tag{B.96}$$

Thus, $\mathcal{W}_3$ can be computed as

$$\begin{aligned}\mathcal{W}_3 &= \sum_{k'=1}^{k-1} \gamma_{nk'}[t+1] \left(\mathbb{E}\left\{ \overline{\mathbf{g}}'_{nk}[t+1] \check{\mathbf{e}}_{nk}[t+1] \left(\check{\mathbf{e}}_{nk}[t+1]\right)' \overline{\mathbf{g}}_{nk}[t+1] \right\} \right) \\ &= \sum_{k'=1}^{k-1} \gamma_{nk'}[t+u] \left(\beta_{nk'} - \theta_{nk'}\right) M\theta_{nk}. \end{aligned} \tag{B.97}$$

Further, the fourth term in (3.96) is

$$\overline{\mathcal{T}}_4 = \sum_{k''=k+1}^{K} \sqrt{\gamma_{nk''}[t+u]}\, \overline{\mathbf{g}}'_{nk}[t+1] \mathbf{g}_{nk''}[t+1] q_{nk''}[t+1]. \tag{B.98}$$

Thus, substituting for $\mathbf{g}_{nk''}[t+1]$, $\mathcal{W}_4$ can be computed as

$$\begin{aligned}\mathcal{W}_4 &= \sum_{k''=k+1}^{K} \gamma_{nk''}[t+1] \left(\mathbb{E}\left\{ \left\| \overline{\mathbf{g}}'_{nk}[t+1] \overline{\mathbf{g}}_{nk''}[t+1] \right\|^2 + \left\| \overline{\mathbf{g}}'_{nk}[t+1] \check{\mathbf{e}}_{nk''}[t+1] \right\|^2 \right\} \right) \\ &= \sum_{k''=k+1}^{K} \gamma_{nk''}[t+u] \left(M\theta_{nk''} + \beta_{nk''} \right) M\theta_{nk}. \end{aligned} \tag{B.99}$$

Now, the 5^{th} term in (3.96) is

$$\overline{\mathcal{T}}_5 = \sum_{n' \neq n, n'=1}^{N} \sum_{k'''=1}^{K} \sqrt{\gamma_{n'k'''}[t+u]}\, \overline{\mathbf{g}}'_{nk}[t+1] \mathbf{g}_{n'k'''}[t+1] q_{n'k'''}[t+1]. \tag{B.100}$$

Therefore,

$$\mathcal{W}_5 = \sum_{n' \neq n, n'=1}^{N} \sum_{k'''=1}^{K} \gamma_{n'k'''}[t+1] \beta_{n'k'''} M\theta_{nk}. \tag{B.101}$$

Similarly, $\mathcal{W}_6 = \sigma^2 M\theta_{nk}$.

C Appendix

C.1 Proof of Lemma 4.1

As explained in detail in 4.1, the effective channel gain is given by

$$|z_k| = u_k + v_k, \tag{C.1}$$

where $u_k \triangleq |t_k|$ and $v_k \triangleq \sum_{i=1}^{T} \left|g_{k,i}\right| \left|h_{k,i}\right|$. Since u_k and v_k are independent random variables, the PDF of their sum is given by

$$f_{|z_k|}(z) = \frac{\partial \mathcal{F}_{|z_k|}(z)}{\partial z}, \tag{C.2}$$

where $\mathcal{F}_{|z_k|}(z)$ is the CDF of $|z_k|$. The CDF $\mathcal{F}_{|z_k|}(z)$ is given by

$$\mathcal{F}_{|z_k|}(z) = \int_{x=0}^{z} \int_{y=0}^{z-x} f_{u_k v_k}(x, y)\, dydx, \tag{C.3}$$

where $f_{u_k v_k}(x, y)$ is the joint PDF of u_k and v_k. Note that the joint PDF of two independent random variables is equal to the product of their individual PDFs. We know that u_k is Rayleigh distributed random variable with scale parameter σ_{u_k}. Therefore, the PDF of u_k is given by

$$f_{u_k}(x) = \frac{x}{\sigma_{u_k}} \exp\left(-\frac{x^2}{2\sigma_{u_k}^2}\right). \tag{C.4}$$

Moreover, as stated above, v_k is the sum of T i.i.d. random variables. Therefore, invoking CLT, the distribution of v_k approaches complex Gaussian distribution with mean μ_{v_k} and variance $\sigma^2_{v_k}$. Therefore, the PDF of v_k is given by

$$f_{v_k}(y) = \bar{q_k}^{-1} \frac{1}{\sqrt{2\pi\sigma^2_{v_k}}} \exp\left(-\frac{(y-\mu_{v_k})^2}{2\sigma^2_{v_k}}\right). \tag{C.5}$$

The scaling by $\bar{q_k}^{-1} = \left\{\frac{1}{2}\left(1 + \text{Erf}\left[\frac{\mu_{v_k}}{\sqrt{2\sigma^2_{v_k}}}\right]\right)\right\}^{-1}$ ensures that $\int_{y=0}^{\infty} f_{v_k}(y)\, dy = 1$. Substituting the PDF of u_k from (C.4) and the PDF of v_k from (C.5) to simplify (C.3), taking the derivative with respect to z as in (C.2) and simplifying further, we get the PDF of the effective channel gain $|z_k|$ between the BS and UE_k as

$$f_{|z_k|}(x) = f_{1,k}(x) - [f_{2,k,1}(x) + f_{2,k,2}(x)] - f_{3,k}(x)\,, \tag{C.6}$$

where $f_{1,k}(x)$, $f_{2,k,1}(x)$, $f_{2,k,2}(x)$, and $f_{3,k}(x)$ are given in (4.4), (4.5), (4.6) and (4.7) with $a_k = \sigma^2_{u_k}$, $b_k = \sigma^2_{v_k}$, $c_k = \mu_{v_k}$, and $\bar{q_k} = \left(1 + \text{Erf}\left[\frac{c_k}{\sqrt{2b_k}}\right]\right)$.

C.2 Proof of Theorem 4.1

We know that, the average SEP for UE_1 is given by

$$\mathcal{P}_1^e(T) = \mathbb{E}_{|z_1|}\left\{\mathcal{P}_1(e|\,|z_1|)\right\} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C}}(e) f_{|z_1|}(x) dx\,, \tag{C.7}$$

where $f_{|z_1|}(x)$ is given in (4.3), and $\mathcal{P}_1^{\text{C}}(e)$ is given in (4.10). Expanding (C.7) using (4.3) and (4.10), average SEP at UE_1 is

$$\mathcal{P}_1^e(T) = \mathcal{I}_{11} - (\mathcal{I}_{12,1} + \mathcal{I}_{12,2}) - \mathcal{I}_{13}\,, \tag{C.8}$$

where

$$\mathcal{I}_{11} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C}}(e) f_{1,1}(x) dx\,, \tag{C.9}$$

$$\mathcal{I}_{12,1} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C}}(e) f_{2,1,1}(x) dx\,, \tag{C.10}$$

$$\mathcal{I}_{12,2} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C}}(e) f_{2,1,2}(x) dx,\,, \tag{C.11}$$

$$\mathcal{I}_{13} = \int_{x=0}^{\infty} \mathcal{P}_1^{\mathrm{C}}(e) f_{3,1}(x) dx. \tag{C.12}$$

Using the following approximation of Q-function [95]

$$\mathbf{Q}\{j\} \approx (1/12)e^{-\frac{j^2}{2}} + (1/4)e^{-\frac{2j^2}{3}} \triangleq \widetilde{\mathbf{Q}}\{j\}, \ \ j > 0 \tag{C.13}$$

to expand the integrand of $\mathcal{I}_{11}$ and $\mathcal{I}_{12,2}$, we get

$$\mathcal{I}_{11} \approx \sum_{m=1}^{N_2/2} \int_{x=0}^{\infty} \mathcal{K}_{I_{11}} e^{-\frac{(c_1-x)^2}{2b_1}} \sum_{j=1}^{2} \widetilde{\mathbf{Q}}\left\{L_j(m)\, x\right\} dx, \tag{C.14}$$

$$\mathcal{I}_{12,2} \approx \sum_{m=1}^{N_2/2} \int_{x=0}^{\infty} \mathcal{K}_{I_{12,2}} e^{-\frac{c_1^2 a_1 + b_1 x^2}{2a_1 b_1}} \sum_{j=1}^{2} \widetilde{\mathbf{Q}}\left\{L_j(m)\, x\right\} dx, \tag{C.15}$$

where $\mathcal{K}_{I_{11}} = \mathcal{K}_1/\sqrt{b_1}$, $\mathcal{K}_{I_{12,2}} = (\mathcal{K}_1\sqrt{b_1})/(a_1 + b_1)$ with $\mathcal{K}_1 = (2(N_1 - 1)\sqrt{2/\pi})/(N_2 N_1 \bar{q}_1)$. Using the identity given in [96, Eqn. (2.33.1)],

$$\int_{x=0}^{\infty} e^{-[\mathcal{A}x^2 + 2\mathcal{B}x + C]} dx = 0.5\left(\sqrt{\pi/\mathcal{A}}\right) e^{\frac{\mathcal{B}^2 - \mathcal{A}C}{\mathcal{A}}} \left(1 - \mathrm{Erf}\left[\mathcal{B}/\sqrt{\mathcal{A}}\right]\right), \tag{C.16}$$

we get $\mathcal{I}_{11}$ and $\mathcal{I}_{12,2}$ as given in (4.16) and (4.18), respectively.

Using (4.4), (4.5) and $\mathcal{I}_{11}$,

$$\mathcal{I}_{12,1} = \frac{a_1}{(a_1 + b_1)} \mathcal{I}_{11}. \tag{C.17}$$

Since $\mathcal{I}_{13}$ involves the product of error function and exponential function, the integration is mathematically intractable and we use numerical integration just to evaluate this term.

C.3 Proof of Theorem 4.2

We know that the average SEP of UE_2 is given by

$$\mathcal{P}_2^e(T) = \mathbb{E}_{|z_2|}\left\{\mathcal{P}_2(e\,||z_2|)\right\} = \int_{x=0}^{\infty} \mathcal{P}_2^{\mathrm{C}}(e) f_{|z_2|}(x) dx, \tag{C.18}$$

where $f_{|z_2|}(x)$ is given in (4.3) and $\mathcal{P}_2^{\mathrm{C}}(e) = \mathcal{P}_{2,1}^{\mathrm{C}}(e) - \mathcal{P}_{2,2}^{\mathrm{C}}(e)$ is given in (4.12). Expanding (C.18) using (4.3) and (4.12), we get

$$\mathcal{P}_2^e(T) = \sum_{j=1}^{2} (-1)^{(j+1)} \left[\mathcal{J}_2^{j1} - \left(\mathcal{J}_2^{j2,1} + \mathcal{J}_2^{j2,2}\right) - \mathcal{J}_2^{j3}\right], \tag{C.19}$$

where, for $\bar{\alpha}, \bar{\rho}, \bar{\delta} \in \{1, 2\}$ and $\bar{\zeta} \in \{1, 3\}$,

$$\mathcal{J}_2^{\bar{\alpha}\bar{\zeta}} = \int_{x=0}^{\infty} \mathcal{P}_{2,\bar{\alpha}}^{\mathrm{C}}(e) f_{\bar{\zeta},2}(x) dx\,, \tag{C.20}$$

$$\mathcal{J}_2^{\bar{\rho}2,\bar{\delta}} = \int_{x=0}^{\infty} \mathcal{P}_{2,\bar{\rho}}^{\mathrm{C}}(e) f_{2,2,\bar{\delta}}(x) dx. \tag{C.21}$$

Expanding RHS of $\mathcal{J}_2^{11}$, $\mathcal{J}_2^{12,2}$, $\mathcal{J}_2^{21}$ and $\mathcal{J}_2^{22,2}$ using (C.62), we get

$$\mathcal{J}_2^{11} \approx \int_{x=0}^{\infty} \mathcal{K}_{\mathcal{J}_2^{11}} e^{-\frac{(c_2-x)^2}{2b_2}} \widetilde{\mathbf{Q}}\{L_3\, x\}\, dx\,, \tag{C.22}$$

$$\mathcal{J}_2^{12,2} \approx \int_{x=0}^{\infty} \mathcal{K}_{\mathcal{J}_2^{12,2}} e^{-\frac{c_2^2 a_2 + b_2 x^2}{2a_2 b_2}} \widetilde{\mathbf{Q}}\{L_3\, x\}\, dx\,, \tag{C.23}$$

$$\mathcal{J}_2^{21} \approx \sum_{n=1}^{N_1-1} \int_{x=0}^{\infty} \mathcal{K}_{\mathcal{J}_2^{21}} e^{-\frac{(c_2-x)^2}{2b_2}} \Upsilon dx\,, \tag{C.24}$$

$$\mathcal{J}_2^{22,2} \approx \sum_{n=1}^{N_1-1} \int_{x=0}^{\infty} \mathcal{K}_{\mathcal{J}_2^{22,2}} e^{-\frac{c_2^2 a_2 + b_2 x^2}{2a_2 b_2}} \Upsilon dx\,, \tag{C.25}$$

where $\Upsilon = \sum_{j=4}^{5}(-1)^j (N_2 - 1)\widetilde{\mathbf{Q}}\left\{L_j(n)\, x\right\} + \sum_{j=6}^{7}(-1)^j \widetilde{\mathbf{Q}}\left\{L_j(n)\, x\right\}$ and

$$\mathcal{K}_{\mathcal{J}_2^{11}} = \mathcal{K}_2 / \sqrt{b_2}\,, \tag{C.26}$$

$$\mathcal{K}_{\mathcal{J}_2^{12,2}} = (\mathcal{K}_2 \sqrt{b_2})/(a_2 + b_2)\,, \tag{C.27}$$

$$\mathcal{K}_2 = 2(N_2 - 1)/(N_2 \bar{q}_2 \sqrt{\pi/2})\,, \tag{C.28}$$

$$\mathcal{K}_{\mathcal{J}_2^{21}} = \mathcal{K}_3 / \sqrt{b_2}\,, \tag{C.29}$$

$$\mathcal{K}_{\mathcal{J}_2^{22,2}} = (\mathcal{K}_3 \sqrt{b_2})/(a_2 + b_2)\,, \tag{C.30}$$

$$\mathcal{K}_3 = 2(N_1 - n)/(N_2 N_1 \bar{q}_2 \sqrt{\pi/2}). \tag{C.31}$$

Solving $\mathcal{J}_2^{11}$, $\mathcal{J}_2^{12,2}$, $\mathcal{J}_2^{21}$ and $\mathcal{J}_2^{22,2}$ using the identity given in [96, Eqn. (2.33.1)], we get $\mathcal{J}_2^{11}$, $\mathcal{J}_2^{12,2}$, $\mathcal{J}_2^{21}$ and $\mathcal{J}_2^{22,2}$ as in (4.24), (4.25), (4.26), and (4.27), respectively.

Using (4.4), (4.5) and (C.21),

$$\mathcal{J}_2^{12,1} = \frac{a_2}{(a_2 + b_2)} \mathcal{J}_2^{11}\,, \tag{C.32}$$

$$\mathcal{J}_2^{22,1} = \frac{a_2}{(a_2 + b_2)} \mathcal{J}_2^{21}. \tag{C.33}$$

Next, since $\mathcal{J}_2^{13}$ and $\mathcal{J}_2^{23}$ involve the product of error function and exponential function, the

integration is intractable and we use numerical integration just to evaluate these terms.

C.4 Proof of Theorem 4.3

We know that, for $N_1 = 4$ and $N_2 = 16$, the conditional SEP for UE_1 is given by

$$\mathcal{P}_1^{\text{C-QAM}}(e) = 2\mathcal{A}_1 - (\mathcal{A}_1)^2, \tag{C.34}$$

where $\mathcal{A}_1 = \frac{1}{4}\sum_{m=1}^{2}\left(\mathbf{Q}\left\{\widetilde{L}_1\,|z_1|\right\} + \mathbf{Q}\left\{\widetilde{L}_2\,|z_1|\right\}\right)$. Therefore, the average SEP of UE_1 is given by

$$\mathcal{P}_1^{e,\text{QAM}}(T) = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C-QAM}}(e) f_{|z_1|}(x)\, dx. \tag{C.35}$$

Expanding (C.35) using the PDF of the effective channel gain $|z_1|$, the average SEP at UE_1 is given by

$$\mathcal{P}_1^{e,\text{QAM}}(T) = \mathcal{I}_{11}^{\text{QAM}} - \left(\mathcal{I}_{12,1}^{\text{QAM}} + \mathcal{I}_{12,2}^{\text{QAM}}\right) - \mathcal{I}_{13}^{\text{QAM}}, \tag{C.36}$$

where

$$\mathcal{I}_{11}^{\text{QAM}} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C-QAM}}(e) f_{1,1}(x) dx\,, \tag{C.37}$$

$$\mathcal{I}_{12,1}^{\text{QAM}} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C-QAM}}(e)(e) f_{2,1,1}(x) dx\,, \tag{C.38}$$

$$\mathcal{I}_{12,2}^{\text{QAM}} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C-QAM}}(e) f_{2,1,2}(x) dx\,, \tag{C.39}$$

$$\mathcal{I}_{13}^{\text{QAM}} = \int_{x=0}^{\infty} \mathcal{P}_1^{\text{C-QAM}}(e) f_{3,1}(x) dx. \tag{C.40}$$

Expanding the integrands of $\mathcal{I}_{11}^{\text{QAM}}$ and $\mathcal{I}_{12,2}^{\text{QAM}}$ using the following approximation of Q-function [95]

$$\mathbf{Q}\{j\} \approx (1/12)e^{-\frac{j^2}{2}} + (1/4)e^{-\frac{2j^2}{3}} \triangleq \widetilde{\mathbf{Q}}\{j\}, \;\; j > 0 \tag{C.41}$$

and using the identity given in [96, Eqn. (2.33.1)],

$$\int_{x=0}^{\infty} e^{-[\mathcal{A}x^2+2\mathcal{B}x+C]} dx = 0.5\left(\sqrt{\pi/\mathcal{A}}\right) e^{\frac{\mathcal{B}^2-\mathcal{A}C}{\mathcal{A}}}\left(1 - \text{Erf}\left[\mathcal{B}/\sqrt{\mathcal{A}}\right]\right), \tag{C.42}$$

we get $\mathcal{I}_{11}^{\text{QAM}}$ and $\mathcal{I}_{12,2}^{\text{QAM}}$ as given in (4.44) and (4.45), respectively.

Using the PDF of the effective channel gain and $\mathcal{I}_{11}^{\text{QAM}}$,

$$\mathcal{I}_{12,1}^{\text{QAM}} = \frac{a_1}{(a_1+b_1)}\mathcal{I}_{11}^{\text{QAM}}. \tag{C.43}$$

Since the integrand of $\mathcal{I}_{13}^{\text{QAM}}$ involves the product of error function and exponential function, the integration is mathematically intractable and we use numerical integration just to evaluate this term.

C.5 Proof of Theorem 4.4

We know that, for $N_1 = 4$ and $N_2 = 16$, the conditional SEP for UE_2 is given by

$$\mathcal{P}_2^{\text{C-QAM}}(e) = 2\mathcal{A}_2 - 2\mathcal{B}_2 + 2\mathcal{A}_2\mathcal{B}_2 - (\mathcal{A}_2)^2 - (\mathcal{B}_2)^2, \tag{C.44}$$

where

$$\mathcal{A}_2 = \frac{3}{2}\mathbf{Q}\left\{\widetilde{L}_3\,|z_2|\right\}, \tag{C.45}$$

$$\mathcal{B}_2 = \frac{1}{4}\left(2\sum_{i=4}^{5}(-1)^i\mathbf{Q}\left\{\widetilde{L}_i\,|z_2|\right\} + \sum_{j=6}^{7}\mathbf{Q}\left\{\widetilde{L}_j\,|z_2|\right\}\right). \tag{C.46}$$

Therefore, the average SEP of UE_2 is given by

$$\mathcal{P}_2^{e,\text{QAM}}(T) = \int_{x=0}^{\infty}\mathcal{P}_2^{\text{C-QAM}}(e)f_{|z_2|}(x)\,dx. \tag{C.47}$$

Expanding (C.47) using the PDF of the effective channel gain $|z_2|$, the average SEP at UE_2 is

$$\mathcal{P}_2^{e,\text{QAM}}(T) = \mathcal{J}_{11}^{\text{QAM}} - \left(\mathcal{J}_{12,1}^{\text{QAM}} + \mathcal{J}_{12,2}^{\text{QAM}}\right) - \mathcal{J}_{13}^{\text{QAM}}, \tag{C.48}$$

where

$$\mathcal{J}_{11}^{\text{QAM}} = \int_{x=0}^{\infty}\mathcal{P}_2^{\text{C-QAM}}(e)f_{1,2}(x)dx, \tag{C.49}$$

$$\mathcal{J}_{12,1}^{\text{QAM}} = \int_{x=0}^{\infty}\mathcal{P}_2^{\text{C-QAM}}(e)(e)f_{2,2,1}(x)dx, \tag{C.50}$$

$$\mathcal{J}_{12,2}^{\text{QAM}} = \int_{x=0}^{\infty}\mathcal{P}_2^{\text{C-QAM}}(e)f_{2,2,2}(x)dx, \tag{C.51}$$

$$\mathcal{J}_{13}^{\text{QAM}} = \int_{x=0}^{\infty}\mathcal{P}_2^{\text{C-QAM}}(e)f_{3,2}(x)dx. \tag{C.52}$$

Expanding the integrands of $\mathcal{J}_{11}^{\text{QAM}}$ and $\mathcal{J}_{12,2}^{\text{QAM}}$ using the following approximation of Q-function [95]

$$\mathbf{Q}\{j\} \approx (1/12)e^{-\frac{j^2}{2}} + (1/4)e^{-\frac{2j^2}{3}} \triangleq \widetilde{\mathbf{Q}}\{j\}, \;\; j > 0, \tag{C.53}$$

and using the identity given in [96, Eqn. (2.33.1)],

$$\int_{x=0}^{\infty} e^{-[\mathcal{A}x^2+2\mathcal{B}x+C]}dx = 0.5\left(\sqrt{\pi/\mathcal{A}}\right)e^{\frac{\mathcal{B}^2-\mathcal{A}C}{\mathcal{A}}}\left(1-\text{Erf}\left[\mathcal{B}/\sqrt{\mathcal{A}}\right]\right), \tag{C.54}$$

we get $\mathcal{J}_{11}^{\text{QAM}}$ and $\mathcal{J}_{12,2}^{\text{QAM}}$ as given in (4.65) and (4.66), respectively.

Using the PDF of the effective channel gain and $\mathcal{J}_{11}^{\text{QAM}}$,

$$\mathcal{J}_{12,1}^{\text{QAM}} = \frac{a_2}{(a_2+b_2)}\mathcal{J}_{11}^{\text{QAM}}. \tag{C.55}$$

Since the integrand of $\mathcal{J}_{13}^{\text{QAM}}$ involves the product of error function and exponential function, the integration is mathematically intractable and we use numerical integration just to evaluate this term.

C.6 Proof of Theorem 4.5

Using the conditional PEP given in (4.108) and the PDF given in Lemma 4.1, the average PEP of UE_k is given by

$$\mathcal{P}^{\text{PEP}}(s_k \to \bar{s}_k) = \int_{x=0}^{\infty} \mathcal{P}(s_k \to \bar{s}_k | |z_k|)\, f_{|z_k|}(x)\, dx. \tag{C.56}$$

Expanding (C.56) using Lemma 4.1 and (4.108), the average PEP at UE_k is given by

$$\mathcal{P}^{\text{PEP}}(s_k \to \bar{s}_k) = \mathcal{P}_{k,1}^{\text{PEP}} - \left(\mathcal{P}_{k,21}^{\text{PEP}} + \mathcal{P}_{k,22}^{\text{PEP}}\right) - \mathcal{P}_{k,3}^{\text{PEP}}, \tag{C.57}$$

where

$$\mathcal{P}_{k,1}^{\text{PEP}} = \int_{x=0}^{\infty} \mathcal{P}(s_k \to \bar{s}_k | |z_k|)\, f_{1,k}(x)\, dx, \tag{C.58}$$

$$\mathcal{P}_{k,21}^{\text{PEP}} = \int_{x=0}^{\infty} \mathcal{P}(s_k \to \bar{s}_k | |z_k|)\, f_{2,k,1}(x) dx, \tag{C.59}$$

$$\mathcal{P}_{k,22}^{\text{PEP}} = \int_{x=0}^{\infty} \mathcal{P}(s_k \to \bar{s}_k | |z_k|)\, f_{2,k,2}(x) dx, \tag{C.60}$$

$$\mathcal{P}_{k,3}^{\text{PEP}} = \int_{x=0}^{\infty} \mathcal{P}\left(s_k \rightarrow \bar{s}_k | |z_k|\right) f_{3,k}(x) dx. \tag{C.61}$$

Using the following approximation of Q-function [95]

$$\mathbf{Q}\{j\} \approx (1/12)e^{-\frac{j^2}{2}} + (1/4)e^{-\frac{2j^2}{3}} \triangleq \widetilde{\mathbf{Q}}\{j\}, \ \ j > 0 \tag{C.62}$$

to expand the integrand of $\mathcal{P}_{k,1}^{\text{PEP}}$ and $\mathcal{P}_{k,21}^{\text{PEP}}$, we get

$$\mathcal{P}_{k,1}^{\text{PEP}} \approx \sqrt{2/(\pi\, b_k)}(1/\bar{q}_k) \int_{x=0}^{\infty} e^{-\frac{(c_k - x)^2}{2b_k}} \widetilde{\mathbf{Q}}\left\{\frac{x\upsilon_k}{\iota_k}\right\} dx, \tag{C.63}$$

$$\mathcal{P}_{k,22}^{\text{PEP}} \approx \frac{\sqrt{(2b_k)/\pi}}{(a_k + b_k)\,\bar{q}_k} \int_{x=0}^{\infty} e^{-\frac{\left(a_k c_k^2 + b_k x^2\right)}{2a_k b_k}} \widetilde{\mathbf{Q}}\left\{\frac{x\upsilon_k}{\iota_k}\right\} dx, \tag{C.64}$$

Using the identity given in [96, Eqn. (2.33.1)],

$$\int_{x=0}^{\infty} e^{-[\mathcal{A}x^2 + 2\mathcal{B}x + C]} dx = 0.5\left(\sqrt{\pi/\mathcal{A}}\right) e^{\frac{\mathcal{B}^2 - \mathcal{A}C}{\mathcal{A}}} \left(1 - \text{Erf}\left[\mathcal{B}/\sqrt{\mathcal{A}}\right]\right), \tag{C.65}$$

we get $\mathcal{P}_{k,1}^{\text{PEP}}$ and $\mathcal{P}_{k,22}^{\text{PEP}}$ as given in (4.112) and (4.113), respectively. Using (4.4), (4.5) and $\mathcal{P}_{k,1}^{\text{PEP}}$,

$$\mathcal{P}_{k,21}^{\text{PEP}} = \frac{a_k}{(a_k + b_k)} \mathcal{P}_{k,1}^{\text{PEP}}. \tag{C.66}$$

Since the integrand of $\mathcal{P}_{k,3}^{\text{PEP}}$ involves the product of error function and exponential function, the integration is mathematically intractable and cannot be obtained in closed-form. So we use numerical integration just to evaluate this term.

Bibliography

[1] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.

[2] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G wireless communications: Vision and potential techniques," *IEEE Network*, vol. 33, no. 4, pp. 70–75, Aug. 2019.

[3] N. Rajatheva, I. Atzeni, E. Bjornson, A. Bourdoux, S. Buzzi, J.-B. Dore, S. Erkucuk, M. Fuentes, K. Guan, Y. Hu, X. Huang, J. Hulkkonen, J. M. Jornet, M. Katz, R. Nilsson, E. Panayirci, K. Rabie, N. Rajapaksha, M. Salehi, H. Sarieddeen, T. Svensson, O. Tervo, A. Tolli, Q. Wu, and W. Xu, "White paper on broadband connectivity in 6G," *arXiv, eprint 2004.14247*, Apr. 2020.

[4] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Techno. Mag.*, vol. 14, no. 3, pp. 28–41, Jul. 2019.

[5] "IMT Traffic Estimates for the Years 2020 to 2030," *document ITU-R M.2370-0*, ITU-R Radiocommunication Sector of ITU, 2015.

[6] B. Makki and A. Behravan, "Fairness and complexity constrained uplink transmission using coordinated multipoint (COMP) non-orthogonal multiple access (NOMA)," *United States Patent*, Mar. 2024.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. USA: Cambridge University Press, 2005.

[8] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. Commun. and Net.*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

[9] M. Vaezi, R. Schober, Z. Ding, and H. V. Poor, "Non-orthogonal multiple access: Common myths and critical questions," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 174–180, Oct. 2019.

[10] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 629–633, May 2016.

[11] S. C. R. Gaddam, D. Kudathanthirige, and G. Amarasuriya, "Achievable rate analysis for NOMA-aided massive MIMO uplink," in *Proc. IEEE Int. Conf. Commun. (ICC) 2019*, May 2019, pp. 1–7, May 2019.

[12] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[13] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 176–183, Oct. 2017.

[14] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[15] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017. [Online]. Available: http://dx.doi.org/10.1561/2000000093.

[16] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[17] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. on Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[18] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, May 2020.

[19] X. Liu and X. Wang, "Efficient antenna selection and user scheduling in 5G massive MIMO-NOMA system," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Jul. 2016, pp. 1–5.

[20] D. Zhang, K. Yu, Z. Wen, and T. Sato, "Outage probability analysis of NOMA within massive MIMO systems," in *2016 IEEE 83rd Veh. Techno. Conf. (VTC Spring)*, Jul. 2016, pp. 1–5.

[21] D. Kudathanthirige and G. Amarasuriya, "Massive MIMO NOMA downlink," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018, pp. 1–7.

[22] S. Kusaladharma, G. Amarasuriya, W.-P. Zhu, and W. Ajib, "Rate analysis for NOMA in massive MIMO based stochastic cellular networks with pilot contamination," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018, pp. 1–6.

[23] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and A. Nallanathan, "Non-orthogonal multiple access in massive MIMO aided heterogeneous networks," in *2016 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[24] H. V. Cheng, E. Björnson, and E. G. Larsson, "Noma in multiuser MIMO systems with imperfect CSI," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2017, pp. 1–5.

[25] J. Ma, C. Liang, C. Xu, and L. Ping, "On orthogonal and superimposed pilot schemes in massive MIMO NOMA systems," *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 12, pp. 2696–2707, Dec. 2017.

[26] X. Chen, F.-K. Gong, G. Li, H. Zhang, and P. Song, "User pairing and pair scheduling in massive MIMO-NOMA systems," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 788–791, Apr. 2018.

[27] W. A. Al-Hussaibi and F. H. Ali, "Efficient user clustering, receive antenna selection, and power allocation algorithms for massive MIMO-NOMA systems," *IEEE Access*, vol. 7, pp. 31 865–31 882, Feb. 2019.

[28] M. Vaezi, G. A. Aruma Baduge, Y. Liu, A. Arafa, F. Fang, and Z. Ding, "Interplay between NOMA and other emerging technologies: A survey," *IEEE Trans. Cognitive Commun. and Net.*, vol. 5, no. 4, pp. 900–919, Dec. 2019.

[29] K. Senel, H. V. Cheng, E. Björnson, and E. G. Larsson, "What role can NOMA play in massive MIMO?" *IEEE J. Sel. Topics in Signal Process.*, vol. 13, no. 3, pp. 597–611, Jun. 2019.

[30] A. S. de Sena, F. R. M. Lima, D. B. da Costa, Z. Ding, P. H. J. Nardelli, U. S. Dias, and C. B. Papadias, "Massive MIMO-NOMA networks with imperfect SIC: Design and fairness enhancement," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6100–6115, Sep. 2020.

[31] M. R. G. Aghdam, R. Abdolee, F. A. Azhiri, and B. M. Tazehkand, "Random user pairing in massive-MIMO-NOMA transmission systems based on mmWave," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Aug. 2018, pp. 1–6.

[32] C. Goztepe, B. Ozbek, and G. K. Kurt, "Design and implementation of spatial correlation-based clustering for multiuser MISO-NOMA systems," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 254–258, Jan. 2021.

[33] Y. Fu, M. Zhang, L. Salaün, C. W. Sung, and C. S. Chen, "Zero-forcing oriented power minimization for multi-cell MISO-NOMA systems: A joint user grouping, beamforming, and power control perspective," *IEEE J. Sel. Areas in Commun.*, vol. 38, no. 8, pp. 1925–1940, Aug. 2020.

[34] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, "Energy-efficient power allocation in uplink mmwave massive MIMO with NOMA," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3000–3004, Mar. 2019.

[35] M. Hua, W. Ni, H. Tian, and G. Nie, "Energy-efficient uplink power control in NOMA enhanced cell-free massive MIMO networks," in *2021 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, Jul. 2021, pp. 7–12.

[36] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez, and T. Sato, "Capacity analysis of NOMA with mmWave massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1606–1618, Apr. 2017.

[37] A. Krishnamoorthy and R. Schober, "Precoder design for two-user uplink MIMO-NOMA with simultaneous triangularization," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[38] X. Chen, Z. Zhang, C. Zhong, R. Jia, and D. W. K. Ng, "Fully non-orthogonal communication for massive access," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1717–1731, Apr. 2018.

[39] D. Kudathanthirige and G. A. A. Baduge, "NOMA-Aided Multicell Downlink Massive MIMO," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 612–627, Jun. 2019.

[40] H. V. Cheng, E. Björnson, and E. G. Larsson, "Performance analysis of NOMA in training-based multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 372–385, Jan. 2018.

[41] Z. Ding, L. Dai, R. Schober, and H. Vincent Poor, "NOMA meets finite resolution analog beamforming in massive MIMO and millimeter-wave networks," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1879–1882, Aug. 2017.

[42] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.

[43] G. Amarasuriya and H. V. Poor, "Impact of channel aging in multi-way relay networks with massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1951–1957.

[44] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Techno.*, vol. 66, no. 2, pp. 1428–1442, May 2017.

[45] R. Chopra, C. R. Murthy, and H. A. Suraweera, "On the throughput of large MIMO beamforming systems with channel aging," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1523–1527, Aug. 2016.

[46] H. Lee, S. Park, and S. Bahk, "Enhancing spectral efficiency using aged CSI in massive MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[47] T. Younas, J. Li, J. Arshad, H. M. Munir, M. M. Tulu, and O. Younas, "Performance of ZF and RZF in massive MIMO with channel aging," in *Proc. 3rd IEEE Int. Conf. on Comp. and Commun. (ICCC)*, Dec. 2017, pp. 930–934.

[48] A. K. Papazafeiropoulos, H. Q. Ngo, and T. Ratnarajah, "Performance of Massive MIMO Uplink with Zero-Forcing Receivers under Delayed Channels," *IEEE Trans. Veh. Techno.*, vol. 66, no. 4, pp. 3158–3169, Apr. 2017.

[49] R. Chopra, C. R. Murthy, H. A. Suraweera, and E. G. Larsson, "Performance Analysis of FDD Massive MIMO Systems Under Channel Aging," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1094–1108, Feb. 2018.

[50] R. Deng, Z. Jiang, S. Zhou, and Z. Niu, "How often should CSI be updated for massive MIMO systems with massive connectivity?" in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[51] Z. Tao and S. Wang, "How often do we need to estimate wireless channels in massive MIMO with channel aging?" in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[52] R. Deng, Z. Jiang, S. Zhou, and Z. Niu, "Intermittent CSI update for massive MIMO systems with heterogeneous user mobility," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4811–4824, Apr. 2019.

[53] M. M. Taygur and T. F. Eibert, "Analyzing the channel aging effects on massive MIMO downlink by ray-tracing," in *Proc. IEEE 29th Annual Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1–5.

[54] R. Chopra, C. R. Murthy, H. A. Suraweera, and E. G. Larsson, "Analysis of nonorthogonal training in massive MIMO under channel aging with SIC receivers," *IEEE Signal Process. Lett.*, vol. 26, no. 2, pp. 282–286, Feb. 2019.

[55] J. Zuo, Y. Liu, Z. Qin, and C. Shen, "The application of intelligent reflecting surface in downlink NOMA systems," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.

[56] Y. Guo, Z. Qin, Y. Liu, and N. Al-Dhahir, "Intelligent reflecting surface assisted NOMA over fading channels," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Dec. 2020, pp. 1–6.

[57] M. Zeng, X. Li, G. Li, W. Hao, and O. A. Dobre, "Sum rate maximization for IRS-assisted uplink NOMA," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 234–238, Jan. 2021.

[58] F. Fang, Y. Xu, Q.-V. Pham, and Z. Ding, "Energy-efficient design of IRS-NOMA networks," *IEEE Trans. Veh. Techno.*, vol. 69, no. 11, pp. 14 088–14 092, Nov. 2020.

[59] M. Zhang, M. Chen, Z. Yang, H. Asgari, and M. Shikh-Bahaei, "Joint user clustering and passive beamforming for downlink NOMA system with reconfigurable intelligent surface," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, Aug. 2020, pp. 1–6.

[60] Y. Xu, M. Chen, Z. Yang, Y. Liu, H. Long, and M. Shikh-Bahaei, "Fair non-orthogonal multiple access communication systems with reconfigurable intelligent surface," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, Aug. 2020, pp. 1–6.

[61] G. Yang, X. Xu, and Y.-C. Liang, "Intelligent reflecting surface assisted non-orthogonal multiple access," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020, pp. 1–6.

[62] A. Hemanth, K. Umamaheswari, A. C. Pogaku, D.-T. Do, and B. M. Lee, "Outage performance analysis of reconfigurable intelligent surfaces-aided NOMA under presence of hardware impairment," *IEEE Access*, vol. 8, pp. 212 156–212 165, Nov. 2020.

[63] Y. Cheng, K. H. Li, Y. Liu, and K. C. Teh, "Outage performance of downlink IRS-assisted NOMA systems," in *GLOBECOM 2020 - 2020 IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[64] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, Nov. 2020.

[65] Z. Ding and H. Vincent Poor, "A simple design of IRS-NOMA transmission," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1119–1123, May 2020.

[66] C. Zhang, W. Yi, Y. Liu, Z. Qin, and K. K. Chai, "Downlink analysis for reconfigurable intelligent surfaces aided NOMA networks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Dec. 2020, pp. 1–6.

[67] B. Zheng, Q. Wu, and R. Zhang, "Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA?" *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 753–757, Apr. 2020.

[68] V. C. Thirumavalavan and T. S. Jayaraman, "BER analysis of reconfigurable intelligent surface assisted downlink power domain NOMA system," in *COMSNETS*, Jan. 2020, pp. 519–522.

[69] L. Bariah, S. Muhaidat, P. C. Sofotasios, F. E. Bouanani, O. A. Dobre, and W. Hamouda, "Large intelligent surface-assisted non-orthogonal multiple access for 6G networks: Performance analysis," *IEEE Internet of Things J.*, vol. 8, no. 7, pp. 5129–5140, Apr. 2021.

[70] S. Mohandass and P. Reba, "BER analysis of intelligent reflecting surface and relay assisted NOMA," in *2022 International Conference on Intelligent Innovations in Engineering and Technology (ICIIET)*, Sep. 2022, pp. 164–170.

[71] Y. Xie, Z. Ding, and X. Dai, "Integrated multicast and unicast transmission in IRS-NOMA networks," *IEEE Trans. Veh. Techno.*, vol. 72, no. 4, pp. 5473–5478, Apr. 2023.

[72] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[73] Z. Wei, L. Yang, D. W. K. Ng, J. Yuan, and L. Hanzo, "On the performance gain of NOMA over OMA in uplink communication systems," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 536–568, Jan. 2020.

[74] N. Zhang, J. Wang, G. Kang, and Y. Liu, "Uplink nonorthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 458–461, Mar. 2016.

[75] B. Clerckx and C. Oestges, "MIMO Wireless Networks." Oxford: Academic Press, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123850553000122

[76] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[77] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optimization and Engineering*, vol. 8, no. 1, p. 67, Apr 2007.

[78] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Techno.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[79] "Further advancements for E-UTRA physical layer aspects (Release 9)," *3GPP TS 36.814*, Mar. 2010.

[80] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Massive MIMO with spatially correlated rician fading channels," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3234–3250, May 2019.

[81] Z. Ding, R. Schober, and H. V. Poor, "A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4438–4454, Jun. 2016.

[82] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *ArXiv*, vol. cs.NI/9809099, 1998.

[83] A. B. Sediq, R. H. Gohary, R. Schoenen, and H. Yanikomeroglu, "Optimal tradeoff between sum-rate efficiency and Jain's fairness index in resource allocation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3496–3509, Jul. 2013.

[84] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," in *50th Asilomar Conf. on Sig., Sys. and Comp.*, Mar. 2017, pp. 974–978.

[85] K. T. Truong, A. Lozano, and R. W. Heath, "Optimal training in continuous flat-fading massive MIMO systems," in *European Wireless 2014; 20th European Wireless Conf.*, Jun. 2014, pp. 1–6.

[86] A. R. Pawar, S. Kashyap, and S. Chouhan, "Impact of max-min power control, channel estimation and user grouping strategies on uplink massive MIMO-NOMA systems," *IEEE Trans. Veh. Techno.*, vol. 70, no. 8, pp. 7858–7869, Jun. 2021.

[87] A. Ghazanfari, H. V. Cheng, E. Björnson, and E. G. Larsson, "A fair and scalable power control scheme in multi-cell massive MIMO," in *2019 IEEE Int. Conf. on Acoustics, Speech and Sig. Process. (ICASSP)*, May 2019, pp. 4499–4503.

[88] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. Boston: McGraw Hill, 2002. [Online]. Available: http://www.worldcat.org/search?qt=worldcat_org_all&q=0071226613

[89] Q. He, Y. Hu, and A. Schmeink, "Closed-form symbol error rate expressions for non-orthogonal multiple access systems," *IEEE Trans. Veh. Techno.*, vol. 68, no. 7, pp. 6775–6789, Jul. 2019.

[90] M. Sadeghi, E. Björnson, E. G. Larsson, C. Yuen, and T. Marzetta, "Joint unicast and multi-group multicast transmission in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6375–6388, Oct. 2018.

[91] V. K. Shrivastava, S. Baek, and Y. Baek, "5G evolution for multicast and broadcast services in 3GPP release 17," *IEEE Commun. Standards Mag.*, vol. 6, no. 3, pp. 70–76, Sep. 2022.

[92] E. Björnson, O. Özdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

[93] *Technical Specification Group Radio Access Network; Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations*, document 3GPP TR 25.996 V14.0.0, 3rd Generation Partnership Project, Mar. 2017.

[94] Z. Abdullah, G. Chen, S. Lambotharan, and J. Chambers, "Low-complexity antenna selection and discrete phase-shifts design in IRS-assisted multiuser massive MIMO networks," *IEEE Trans. Veh. Techno.*, pp. 1–1, Feb. 2022.

[95] M. Chiani, D. Dardari, and M. Simon, "New Exponential Bounds and Approximations for the Computation of Error Probability in Fading Channels," *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 840–845, Jul. 2003.

[96] D. Zwillinger and V. Moll, *Table of Integrals, Series, and Products*, 8th ed. Boston: Academic Press, 2015.