Imperial College London
Department of Electrical and Electronic Engineering

Rate-Splitting Multiple Access: SIC-free Implementation, Codeword Segmentation, and Mismatched Decoding

Sibo Zhang

2025

Supervised by Prof. Bruno Clerckx

Submitted in part fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Electronic Engineering of Imperial College London and the Diploma of Imperial College London

Statement of Originality

I declare that the content of this thesis is the result of my own work. Information and ideas derived from the work of others have been acknowledged in the text and a list of references is given in the bibliography. The material of this thesis has not been and will not be submitted for another degree at any other university or institution.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

To all I love.

Abstract

Multiple access techniques have played a pivotal role in advancing wireless communications. Rate-splitting multiple access (RSMA) has been recognized as a promising multiple access technique for future wireless systems. However, practical implementation challenges may restrict its applicability. This thesis addresses such challenges by proposing novel extensions of RSMA with lower implementation complexity and by investigating the performance of different RSMA schemes under practical constraints.

First, we develop alternative receiver designs for RSMA beyond the conventional hard successive interference cancellation (SIC). These designs can be broadly classified into two categories: SIC-based and SIC-free. Using metrics derived from mutual information, we model both categories and evaluate their theoretical limits. Our analysis and link-level simulations demonstrate that RSMA can retain its performance advantages even without SIC at the receiver. Second, we introduce a new RSMA architecture, termed codeword-segmentation RSMA (CS-RSMA). This approach reduces the complexity of encoding/decoding, control signaling, and retransmission, while achieving performance comparable to or better than existing RSMA schemes. These contributions open the door to practical and competitive RSMA strategies for 6G and beyond.

In addition to RSMA, this thesis also investigates the broader challenge of interference in wireless communications. By applying mismatched decoding, we analyze scenarios where interference is treated suboptimally, providing new insights into its impact, proposing novel receiver designs, and introducing a refined framework for performance evaluation that more accurately captures interference effects.

Acknowledgements

My PhD journey has brought challenges from many directions—some shared by most doctoral students, and others unique to my own path. I am grateful to myself for having found the persistence to navigate each of them and to see this work through to the end. I would like to express my deepest gratitude to my parents for their love and continuous support, both emotional and financial, throughout this journey. I am deeply thankful to Ms. Yifei Wang for her unwavering understanding and encouragement.

I would like to sincerely thank Prof. Bruno Clerckx for granting me the opportunity to pursue my PhD under his supervision. His exceptional expertise, insightful guidance, and continuous support have been essential in the completion of this work and in shaping my ability to conduct independent research.

I am profoundly grateful to Prof. Albert Guillen i Fàbregas for hosting me in Cambridge as a visiting scholar. His generous guidance was invaluable in my exploration of the elegance of information theory.

I acknowledge with gratitude Dr. David Vargas and Mr. Andrew Murphy for their remarks and support, through which I gained valuable insight into industrial aspects.

I was fortunate to be accompanied by some wonderful colleagues. Especially, I would like to express my sincere thanks to Dr. Yunnuo Xu for kindly organizing trips and leisure activities; to Dr. Yuanwen Liu for chatting and teaching me philosophy; to Dr. Hongyu Li and Dr. Jiawei Xu for teaching me playing card games and Mahjong; to Dr. Zhenyuan Feng, Dr. Yumeng Zhang, Dr. Onur Dizdar and Dr. Longfei Yin for chatting and generous career advices; and to Mr. Yiyang Peng, Mr. Qiaosen Zhang, Dr. Binggui Zhou, Dr. Zheyu Wu, Dr. Matteo Nerini, along with all the colleagues mentioned above, for the many enjoyable chats we shared.

Sibo Zhang
London, 2025

Contents

Abstract **v**

Acknowledgements **vii**

1 Introduction **1**

1.1 Background 1

1.1.1 Rate-splitting multiple access 1

1.1.2 Mismatched decoding 5

1.2 Contributions 6

1.3 Publications 7

1.4 Notations 9

2 Preliminaries **11**

2.1 Rate-splitting multiple access 11

2.2 Mutual information of finite-alphabet input Gaussian channels 15

2.3 The generalized mutual information 17

2.4 Subgradient 21

3 SIC-free Implementation for RSMA **25**

3.1 Low-complexity precoder designs for RSMA 27

3.1.1 Sum-rate optimization 28
3.1.2 Constellation-constrained sum-rate optimization 29
3.2 Transceiver designs for RSMA 34
3.2.1 Transmitter design 34
3.2.2 Receiver designs 34
3.2.3 Comments on different receivers 38
3.3 Numerical results 41
3.3.1 CC sum-rate comparison 41
3.3.2 Power allocation behavior 43
3.3.3 Link-level simulation 45
3.4 Summary 50

4 Precoder Optimization for SIC-based and SIC-free RSMA 53

4.1 Achievable rate of RSMA under finite constellations 55
4.2 Precoding optimization 59
4.2.1 Weighted sum-rate optimization 60
4.2.2 Max-min fairness optimization 63
4.3 Optimization for large-scale systems 65
4.3.1 User grouping 68
4.3.2 Null space projection and dimensionality reduction 69
4.3.3 Complexity analysis 71
4.4 Numerical results 71
4.4.1 Adaptive modulation 72
4.4.2 Weighted sum-rate performance 72
4.4.3 Max-min fairness performance 75

4.4.4 Performance in large-scale systems 76

4.5 Summary 77

5 Optimal and Suboptimal Decoding under Finite-Alphabet Signals and Interference 81

5.1 Mismatched decoding for finite alphabets 84

5.1.1 Finite-alphabet Gaussian channel under interference 84

5.1.2 Encoding and decoding 85

5.1.3 Matched capacity and a lower bound on mismatched capacity 86

5.1.4 Approximation of mismatched capacity 88

5.1.5 Results on GMI for FAGCI channels 88

5.2 Improvements on decoding metric 94

5.2.1 Generalized Gaussian decoding metric 94

5.2.2 Interference decomposition decoding metric 96

5.2.3 Results on GMI of the proposed decoding metrics 98

5.3 Practical receivers 98

5.4 Signal processing with GMI 101

5.4.1 Achievable rate of MU-MISO incorporating decoding strategy 103

5.4.2 Sum-rate optimization 106

5.4.3 Numerical simulations 107

5.4.4 Future directions 108

5.5 Summary 110

6 Codeword-Segmentation Rate-Splitting Multiple Access 113

6.1 Codeword-segmentation RSMA 116

6.2 Performance evaluation using mismatched decoding 118

6.2.1 Evaluation of RSMA and CS-RSMA under suboptimal decoding 118

6.2.2 Constrained decoding complexity 121

6.3 Precoder optimization 123

6.3.1 Sum-rate optimization 123

6.3.2 Max-min fairness optimization 126

6.4 Numerical results on achievable rate 128

6.4.1 Adaptive modulation 129

6.4.2 Sum-rate evaluation 129

6.4.3 Max-min fairness optimization 130

6.5 Physical-layer transceiver design for CS-RSMA and link-level simulation 134

6.5.1 Transceiver design 134

6.5.2 Link-level simulation 135

6.5.3 Advantages of CS-RSMA 137

6.6 On CS-RSMA with optimal encoding and decoding 139

6.6.1 A superposition perspective 139

6.6.2 Performance evaluation 141

6.6.3 Two other schemes with equivalent asymptotic rate as CS-RSMA with optimal encoding/decoding 143

6.7 Summary 144

7 Conclusion **145**

7.1 Summary of thesis achievements 145

7.2 Future work 147

A Channel Models Involving Fading Correlation **151**

A.1 Ong-ring model 151

A.2 Rician fading model 152

B Appendices for Chapter 3 **153**

B.1 Proof of Proposition 3.1 and 3.2 153

B.2 The equivalence of joint de-mapping and treating interference as noise 154

B.3 Channel realizations 155

B.4 Code rates used in LLS 156

C Appendices for Chapter 4 **159**

C.1 Proof of Proposition 4.1 159

C.2 Proof of Proposition 4.2 160

C.3 Transmission mode dictionaries used for numerical results 161

D Appendices for Chapter 5 **163**

D.1 Proof of Proposition 5.1 163

D.2 Proof of Proposition 5.2 163

D.3 Proof of Proposition 5.4 166

D.4 Proof of Proposition 5.5 167

E Appendix for Chapter 6 **169**

E.1 Proof of the achievability of (6.8) 169

E.2 Proof of Proposition 6.3 172

Bibliography **173**

List of Tables

3.1 Classification and complexity of different receiver designs. 39

4.1 An estimate of the number of FLOPs for computing constellation-constrained rate 66

4.2 An estimate of the number of FLOPs for computing constellation-constrained rate with user grouping and interference nulling 66

6.1 Transmission modes under decoding complexity $\delta = 4$ 129

6.2 Transmission modes under decoding complexity $\delta = 16$ 129

B.1 Code rates used for RSMA in link-level simulations. 157

C.1 Transmission mode dictionary for $K = 2$ and $R_{\max} = 6$ bits. 162

C.2 Transmission mode dictionary for $K = 2$ and $R_{\max} = 8$ bits. 162

C.3 Transmission modes dictionary for $K = 3$ and $R_{\max} = 6$ bits. 162

C.4 Transmission modes dictionary for $K = 3$ and $R_{\max} = 9$ bits. 162

List of Figures

2.1 Single-layer RSMA with SIC receivers. 12

3.1 Comparison of CC sum-rate for RSMA achieved by different methods of searching for $t^\star$ with $N_\mathrm{T} = 4$, $K = 2$. 33

3.2 Transmitter architecture. 35

3.3 Hard CWIC. 35

3.4 Soft CWIC 1. 35

3.5 Soft CWIC 2. 35

3.6 Joint de-mapper. 35

3.7 Soft SLIC. 35

3.8 CC sum-rate comparison for different schemes. 42

3.9 Sum-rate/CC sum-rate as functions of t with $t^\star$ labeled by ★, $K = 2$. 43

3.10 CDF of $t^\star$ under different objective functions with $K = 2$. 44

3.11 BER performance of different RSMA receivers with $K = 2$ and $\Delta = \pi/9$. 46

3.12 Fast fading channel with $K = 2$, $\Delta = \pi/18$ and throughput = 3 bits/symbol. Modulation schemes for RSMA: QPSK for both common and private streams. 47

3.13 Fast fading channel with $K = 2$, $\Delta = \pi/18$ and throughput = 6 bits/symbol. Modulation schemes for RSMA: 16QAM for both common and private streams. 47

3.14 Block fading channel with $K = 2$, $\Delta = \pi/18$ and throughput = 3.6 bits/symbol. Modulation schemes for RSMA: QPSK for both common and private streams. 48

3.15 Block fading channel with $K = 3$, $\Delta = \pi/9$ and throughput = 5.52 bits/symbol. Modulation schemes for RSMA: 16QAM for the common stream and QPSK for the private streams. 48

3.16 Comparison between RSMA and SDMA with $K = 2$, $\Delta = \pi/18$, $\pi/9$ or $\pi/6$. RSMA uses QPSK for both common and private streams. The throughput of different schemes is 3 bits/symbol. 50

3.17 SNR-complexity trade-off with different schemes under fast fading with $K = 2$, $\Delta = \pi/18$ and throughput = 3 bits/symbol. Modulation schemes for RSMA: QPSK for all the streams. 51

4.1 Single-layer RSMA with SIC and SIC-free receivers. 55

4.2 Illustration of user grouping. 66

4.3 Average SR performance. 73

4.4 average common stream power ratio. 74

4.5 Convergence of PSGA and GQMP. 74

4.6 Rate region with $N_{\mathrm{T}} = 2$ and $K = 2$, $\{\theta_1^{\mathrm{az}}, \theta_2^{\mathrm{az}}\} = \{0, \pi/9\}$, $\mathcal{K} = 10$ dB. 75

4.7 Average MMF performance. 76

4.8 Average sum-rate performance in large-scale systems with $N_{\mathrm{T}} = 128$, $K = 64$, and $R_{\max} = 6$ bits per group. 77

4.9 Average MMF performance in large-scale systems with $N_{\mathrm{T}} = 128$, $K = 64$, and $R_{\max} = 6$ bits per group. 78

5.1 Channel model with interference and noise. 84

5.2 GMI and MI with $\sigma_i^2 = 20$dBm, $\sigma_j^2 = 10$dBm and $\sigma_z^2 = 0$dBm. 89

5.3 Visualization of $\mathcal{X} + \mathcal{I} + \mathcal{J}$ at the three marked points in Figure 5.2. 91

5.4 GMI and MI with $\sigma_x^2 = 20$dBm, $\sigma_j^2 = 15$dBm and $\sigma_z^2 = 0$dBm. 92

5.5 GMI and MI with $\sigma_x^2 = 10$, $\sigma_i^2 = 5$ and $\sigma_z^2 = 0.5$. 93

5.6 Probability density comparison of ground truth interference plus noise distribution by a mixture of Gaussian, approximation by the Gaussian distribution and Generalized Gaussian distribution with $\beta = 5.1$. . . . 95

5.7 Probability density comparison of ground truth interference plus noise distribution by a mixture of Gaussian, an approximation by a Gaussian distribution and two approximations by interference decomposition. 97

5.8 GMI and MI with $\sigma_x^2 = 15$dBm and $\sigma_z^2 = 0$dBm. 98

5.9 Transceiver structure. 99

5.10 Comparison of GMI/MI and LLS results with $\mathcal{X}$ being QPSK, $\mathcal{J}$ being 16QAM, $\sigma_j^2 = 20$dBm, $\sigma_z^2 = 0$dBm. 102

5.11 Comparison of GMI/MI and LLS results with $\mathcal{X}$ being 16QAM, $\mathcal{J}$ being 16QAM, $\sigma_j^2 = 20$dBm, $\sigma_z^2 = 0$dBm. 102

5.12 Illustration of a MISO broadcast channel, where solid arrows denote desired signals and dashed arrows denote interfering signals. The three users potentially suffer from multi-user interference. 103

5.13 Sum-rate comparison of MU-MISO under different decoding strategies. 109

5.14 Sum-rate comparison of MU-MISO, with decoders treating interference as a Gaussian RV and precoders by optimizing different objective functions. 109

5.15 Ergodic inter-user interference power received at User 1 from the other two users under different objective functions. 110

5.16 Evolution of the ergodic approximated sum-rate over successive iterations of Algorithm 7. 111

6.1 Codeword-Segmentation RSMA for multi-user MISO. 116

6.2 A visual illustration of codeword segmentation and combination. . . . 117

6.3 Average SR performance of conventional RSMA, CS-RSMA and SDMA. 131

6.4 Average SR performance of different modes of CS-RSMA. 131

6.5 Comparing contribution of common and private stream with $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. 132

6.6 Average MMF performance of conventional RSMA, CS-RSMA and SDMA. 133

6.7 Average MMF performance of different modes of CS-RSMA. 133

6.8 PHY transmitter design for CS-RSMA. 135

6.9 PHY receiver designs for CS-RSMA. 136

6.10 BER comparison between conventional RSMA and CS-RSMA with all streams using QPSK constellation. Code rate for RSMA: 0.85 for common streams, 0.325 for private streams. Code rate for CS-RSMA: 0.5 for all codewords. 137

6.11 An effective codeword construction for user-k in CS-RSMA. 139

6.12 Average SR performance of CS-RSMA with optimal encoder and decoder. 142

6.13 Average MMF performance of CS-RSMA with optimal encoder and decoder. 142

6.14 Illustration of a scheme where $\mathbf{s}_{\mathrm{c}}$ consists of multiple complete codewords, user-k desires $\mathbf{s}_{\mathrm{c},k}$, one codeword in $\mathbf{s}_{\mathrm{c}}$, and $\mathbf{s}_k$, a separately transmitted codeword. 143

A.1 Transmit antenna array and a scattering ring centered at AoD θ with two sided angular spread Δ. 152

Chapter 1

Introduction

Digital communication systems have evolved rapidly since the foundational work in information theory by Shannon [1]. Emerging wireless networks face increasing demands for high spectral efficiency and massive connectivity. Multiple access techniques have played a pivotal role in the advancement of wireless communication systems. Mobile broadband standards, such as 4G and 5G, have converged to the use of Time-Division Multiple Access (TDMA) and Orthogonal Frequency-Division Multiple Access (OFDMA) to split time and frequency resources into resource elements, and linearly precoded Multi-User Multiple-Input Multiple-Output (MU-MIMO) technologies, also called Space-Division Multiple Access (SDMA), to exploit the spatial dimension. The focus of this thesis will be on a novel multiple access technique, namely Rate-Splitting Multiple Access (RSMA) [2, 3][1] and on a study of suboptimal receivers in the presence of interference.

1.1 Background

1.1.1 Rate-splitting multiple access

RSMA was primarily proposed as a spatial domain technique to improve MU-MIMO performance.[2][3] With RSMA, in the simplest form known as 1-layer RSMA, the

[1]The term "Rate-Splitting Multiple Access" was also used in [4] for a rate-splitting technique in single-antenna multiple access channels. This is different from the focus of this thesis.

[2]That said, the idea of RSMA can potentially be applied to other domains. For example, in [5], the idea of RSMA was applied in the frequency domain, instead of the spatial domain.

[3]The term "multiple access" carries different meanings in wireless communications and information theory. In wireless communications, multiple access often refers to techniques that allow

transmitter first splits the unicast messages into common and private parts. The common parts are combined and encoded into a common stream, while the private parts are encoded separately. The transmitter then utilizes linear precoding to send the common and private streams in a superposition manner. Each receiver decodes the common stream first and applies Successive Interference Cancellation (SIC) before decoding the desired private stream. Owing to the presence of both common and private streams, RSMA enables users to partially decode interference and partially treat interference as noise, and hence is recognized as a powerful multiple access strategy.

The idea of rate splitting and superposition coding first emerged in the study of single antenna interference channels [6, 7], where they are also known as the Han–Kobayashi coding scheme. This scheme leads to the best inner bound on the capacity region of two-user Gaussian interference channels for all interference regimes. Specifically, the Han–Kobayashi scheme generalizes treating interference as noise and decoding interference, where the former achieves the sum-capacity under very weak interference, and the latter achieves the capacity region under strong interference. Under weak interference, the Han–Kobayashi scheme achieves higher rates than the two. Notably, [8] proposed a transmission strategy for a single-user slowly fading Gaussian channel without channel feedback to the transmitter. Although not termed rate splitting, this strategy involves message splitting to mitigate channel uncertainty.

Inspired by the power of the Han–Kobayashi coding scheme, [9, 10, 11] initiate the research on RSMA for robust transmission under multi-antenna broadcast channels by showing its superiority with partial Channel State Information at Transmitter (CSIT). Although the fundamental limits of multi-antenna broadcast channels under imperfect CSIT are still unknown, it was shown that the rate-splitting approach achieves the optimal degree-of-freedom/multiplexing gain [9, 10, 12, 13, 14].

It can be noted that treating interference as noise and decoding interference in interference channels are very similar to the two well-known strategies in downlink MU-MIMO systems, namely SDMA and Non-Orthogonal Multiple Access (NOMA) [15]. In light of this, [16, 17] show that even under perfect CSIT, RSMA generalizes and outperforms NOMA, SDMA and physical layer multicasting. Several architectures have been proposed for RSMA. Enabling only one common stream

multiple users to share limited transmission resources, such as time, frequency and space, and can be applied to either uplink or downlink scenarios. In information theory, however, multiple-access channels refer exclusively to the uplink, while broadcast channels correspond to the downlink.

for all the users, 1-layer RSMA [9, 16] is the most attractive form due to its low computation and hardware requirements. However, with more complicated implementations, multiple layers of common streams can be enabled in RSMA. Examples of such designs are the hierarchical RSMA [18] and the generalized RSMA [16]. By transmitting the private streams using dirty paper coding, dirty paper coded RS is proposed in [19]. In this thesis, we will mainly focus on 1-layer RSMA.

Subsequent research has shown that RSMA benefits a wide range of downlink communication scenarios. For scenarios with only statistical CSIT (i.e., only the knowledge on the distribution of the channels is available), the application of RSMA is discussed in [20, 21, 22]. The robustness of RSMA with imperfect CSIT caused by quantized feedback is discussed in [23], where it was shown that RSMA enables overhead reduction compared to SDMA. The superiority of RSMA on energy efficiency is studied in [24]. The common stream in RSMA is naturally suited to multicasting, which motivates joint unicast and multicast using RSMA [25]. Benefits under finite resolution digital-to-analogue and analogue-to-digital are investigated in [26, 27]. Integration of RSMA and cooperative communications has been shown to improve coverage by compensating for performance degradation at cell-edge users [28, 29].

The emerging technologies proposed for 6G and beyond brought new opportunities for RSMA. The emergence of higher frequencies brings challenges in high path loss, blockage, and expensive hardware. RSMA is studied for millimeter wave and hybrid precoding in [21] and terahertz systems in combination with cooperative communications in [30]. For frequency division duplexing (FDD) massive MIMO, [20] shows that RSMA can mitigate the performance degradation from imperfect channel statistics and CSIT. For time division duplexing (TDD) massive MIMO, RSMA benefits under imperfect channel estimation caused by high mobility [31]. For integrated sensing and communications, [32] shows that the challenge of interference caused by sensing can be improved by using the common stream to perform sensing. The interplay between RSMA and reconfigurable intelligent surfaces is studied in [33, 34, 35]. For multibeam satellite communication systems, the challenge of frequency reuse and imperfect/delayed CSIT can be tackled by RSMA [36, 37]. For low-latency scenarios, the performance of RSMA is shown to be stable in [38, 39]. RSMA for the space-air-gound integrated network is discussed in [40]. RSMA with cell-free massive MIMO is discussed in [41]. For wireless caching, the idea of RSMA is applied to the most popular content and the intelligent coded caching policy in [42]. The application of RSMA in the cellular internet of things and massive access is discussed in [43]. [44] investigates RSMA as a multiple access technique for

semantic communications.

Although the majority of the research focuses on RSMA for downlink systems, [4] provides the foundation for RSMA for single-antenna multiple access channels, i.e., uplink systems. For MIMO uplink systems, the benefits have also been studied, for example, in short-packet systems [45] and network slicing [46]

While there is extensive literature on the application of RSMA across various scenarios, comparatively little research has addressed its practical implementation. The low complexity precoder design is involved in [17, 31, 47, 48, 24, 21]. The performance of RSMA under finite-block lengths is studied in [38, 39]. For more convincing performance evaluations, Link-Level Simulations (LLS) of RSMA are presented in [49, 31, 50, 51, 52].

Beyond the aspects mentioned above, other practical limitations remain. A non-exhaustive summary is as follows:

- Alternative receiver designs: Most existing work focuses on the hard-decision SIC receiver, whereas research on alternative RSMA receiver implementations, which potentially pose different complexity and performance, is very limited. In [53], the de-modulation and cancellation modules in the SIC receivers are replaced by deep neural networks to obtain performance improvement over basic SIC receivers under imperfect Channel State Information at Receiver (CSIR). [54, 55] analyze RSMA with joint decoding at the receiver from an information-theoretic perspective, despite the fact that the implementation of joint decoding is very difficult in practice. A more interesting yet unresolved problem of RSMA lies in the necessity of SIC. Although it is convenient to consider SIC or joint decoding for information-theoretic analysis, the requirement of SIC poses a challenge in the practical implementation because of its high complexity and error propagation issues. The effect of imperfect cancellation from SIC is studied in [56]. However, the possibility of resolving the benefit of RSMA without SIC at the receivers is completely undiscovered.

- Finite constellations: The aforementioned research on RSMA assumes Gaussian-distributed input signals for ease of analysis. However, Gaussian signals are impractical to realize, and finite alphabets/constellations (such as Quadrature Amplitude Modulation (QAM) constellations) are commonly used in practice.[4] Very limited work studies the potential of RSMA with finite alphabets. In [57],

[4] In this thesis, both alphabet and constellation refer to the set of possible input signals.

constructive interference precoding with finite constellations is considered for RSMA. In [31, 50, 51, 52], although the LLS performance of RSMA is evaluated under QAM constellations, the system designs assume Gaussian input signals. It is therefore difficult to exploit the potential of RSMA in practice without considering the actual constellations in use during the design process. System designs for RSMA that consider finite-alphabet constellations are desired to reveal the full potential of RSMA in practical applications. Also, finite-alphabet interference is less detrimental than Gaussian-distributed interference, and it is therefore possible to preserve the benefit of superposition coding without successively removing signals during decoding.

- Suboptimal decoders: The majority of the work on RSMA considers performance metrics originating from mutual information, which is proved to be achievable with reliable communications by Shannon's channel coding theorem [1]. Proofs of the channel coding theorem typically assume idealized decoders, which is often difficult to implement in practice. It is therefore interesting to discover the design and evaluation of RSMA systems under suboptimal decoders.

1.1.2 Mismatched decoding

Shannon's channel coding theorem provides limits on reliable communication rates subject to optimal encoding and decoding [1, 58]. The channel coding theorem is often proved by assuming decoding techniques that optimally exploit the channel transition law, also called perfect channel knowledge, such as joint typicality decoding [1] or maximum-likelihood decoding [58]. However, in practice, it is often the case that suboptimal decoding is implemented because, for example, the channel knowledge is not sufficient at the decoder, or there are constraints on the decoding technique in use, possibly regarding computational burden.

Mismatched decoding [59, 60, 61] is a theory that addresses the problem of reliable communications subject to suboptimal decoders. Mismatched capacity, i.e., the supremum of rates that allow reliable communications under mismatched decoding, is often a difficult problem, because this would require the codebook to be optimized according to not only the channel knowledge but also the decoding rule. The common approach is to make certain assumptions on the coding ensemble (i.e., the random process that generates the codewords), which is often suboptimal, and study the resulting achievable rates. This approach leads to lower bounds of

the mismatched capacity. Two well-known lower bounds of mismatched capacity are Generalized Mutual Information (GMI) and LM rate,[5] where the GMI assumes i.i.d. random coding and the LM rate assumes constant-composition random coding [61].

Mismatched decoding has been evaluated under various scenarios. One motivation behind using mismatched decoders is often from concern on the decoder's complexity. For example, [62] addresses nearest neighbour decoding under non-Gaussian channels, as maximum-likelihood decoding for non-Gaussian channels can be computationally complicated depending on the noise distribution. [63] investigates decoders under constraints of finite precision arithmetic. In [64], Bit-Interleaved Coded Modulation (BICM) systems are analyzed using mismatched decoding. [65] considers multiple access channels with simplified decoders. [66] studies multiple access channels with mismatches in successive decoding.

The above examples consider mismatched decoding due to decoder limitations. For wireless systems, fluctuating channels and limited resources for channel estimation bring about the other common cause of mismatched decoding, namely, imperfect channel knowledge. [67] addresses the achievable rate with imperfect side information at the receiver. [68] and [69] consider Multiple-Input-Multiple-Output (MIMO) systems under block fading channels and imperfect Channel State Information at Receiver (CSIR). [70] studies the scenario where imperfect channel knowledge exists at both the transmitter and receiver.

1.2 Contributions

A summary of the following chapters, highlighting the key contributions of this thesis, is presented below.

Chapter 2 provides background knowledge for this thesis, including the fundamentals of RSMA, derivations of the mutual information under finite alphabets, a proof of the GMI and the concept of subgradients.

In Chapter 3, various receiver implementations for RSMA are proposed and classified into SIC and SIC-free types. It further analyzes the achievable rates of RSMA under finite alphabets with and without SIC receivers, and demonstrates that the

[5]The LM rate appears to have originated in [59], implicitly denoting "Lower bound on the Mismatch capacity".

two receiver types correspondingly push the performance of RSMA towards the two achievable rates. The results on achievable rates and LLS demonstrate that RSMA can retain its performance advantages even without SIC at the receivers.

Chapter 4 discovers the theoretical limits of SIC-based and SIC-free RSMA under finite alphabets with fully optimized precoders. It proposes precoder optimization algorithms along with methods to reduce computational complexity in large-scale systems. The results demonstrate that, under the evaluated scenarios, RSMA schemes outperform SDMA with finite alphabets. Moreover, SIC-free RSMA retains most of the performance advantages of SIC-based RSMA, thereby extending the conclusion of Chapter 3 to a wider range of scenarios.

Chapter 5 analyzes the effect of suboptimally treating finite-alphabet interference during decoding. This is achieved by proposing a channel model that incorporates finite-alphabet input and interference and analyzing its achievable rate under various decoders. Building on this analysis, the chapter provides new insights into the impact of interference, introduces novel receiver designs, and establishes a new system design and performance evaluation framework that more accurately captures effects from suboptimally treating interference .

Chapter 6 proposes a novel RSMA architecture, namely codeword-segmentation RSMA (CS-RSMA). Performance of CS-RSMA and conventional RSMA is evaluated under a novel framework based on Chapter 5 to capture the effect of suboptimally treating inter-user interference. A physical layer implementation of CS-RSMA is also proposed. It is demonstrated that CS-RSMA achieves similar performance with the conventional RSMA, while offering significant practical advantages in terms of encoding/decoding complexity, control signaling overhead, and retransmission mechanism design.

Lengthy proofs, some details on channel models and simulation setups, and supplementary discussions are deferred to the appendices for clarity.

1.3 Publications

The content of this thesis builds upon materials from the following journal and conference papers, a industrial contribution, and a patent.

Journal and Conference Papers:

- **S. Zhang**, B. Clerckx, D. Vargas, O. Haffenden and A. Murphy, "Rate-Splitting Multiple Access: Finite Constellations, Receiver Design, and SIC-free Implementation," *IEEE Transactions on Communications*, vol. 72, no. 9, pp. 5319–5333, Sep. 2024. [Chapter 3]

- **S. Zhang**, B. Clerckx, D. Vargas, "SIC-Free Rate-Splitting Multiple Access: Constellation Constrained Sum-Rate Optimization," in *Proceedings of IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Sep. 2024, pp. 903–910. [Chapter 4]

- **S. Zhang**, B. Clerckx, D. Vargas, "SIC-Free Rate-Splitting Multiple Access: Constellation-Constrained Optimization and Application to Large-Scale Systems," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 10669-10683, 2026. [Chapter 4]

- **S. Zhang**, B. Clerckx, "Optimal and Suboptimal Decoders under Finite-Alphabet Interference: A Mismatched Decoding Perspective," *IEEE Journal on Selected Areas in Communications*, early access, Dec. 2025, doi: 10.1109/JSAC.2025.3644829. [Chapter 5]

- **S. Zhang**, B. Clerckx, D. Vargas, "Codeword-Segmentation Rate-Splitting Multiple Access and Evaluation under Suboptimal Decoding," *arXiv preprint* arXiv:2506.17164, Jun. 2025 – submitted to *IEEE Transactions on Communications*. [Chapter 6]

Industrial Contribution:

- 3GPP. Rate Splitting Multiple Access (RSMA) for Multi-User MIMO - Enhancement of Unicast and Joint Unicast/Multicast Delivery. document RP-231938, 3GPP TSG RAN Meeting #101, BBC, Bangalore, India, Sept. 2023. [Chapter 3]

Patent:

- **S. Zhang**, B. Clerckx, D. Vargas, "Codeword-Segmentation Rate-Splitting Multiple Access," *UK Patent Application*, filed, Jun. 2024, patent pending. [Chapter 6]

1.4 Notations

Scalars are denoted by normal letters. Sequences and vectors are denoted by bold lowercase letters. Random variables are denoted by capital letters. Calligraphic letters denote multisets. Bold uppercase letters can denote matrices or a vector of random variables, depending on the context. For example, x is an observation of X, and $\mathbf{x}$ is an observation of $\mathbf{X}$.

$P_{\cdot}(\cdot)$, $P_{\cdot|\cdot}(\cdot \mid \cdot)$, $f_{\cdot}(\cdot)$ and $f_{\cdot|\cdot}(\cdot|\cdot)$ denote probability, conditional probability, probability density, and conditional probability density, respectively. $H(\cdot)$, $H(\cdot|\cdot)$, $I(\cdot;\cdot)$ and $I(\cdot\,;\,\cdot \mid \cdot)$ denote entropy, conditional entropy, mutual information and conditional mutual information, respectively. $\mathbf{I}$ and $\mathbf{0}$ denote the identity matrix and the zero matrix, whose dimensions are given by their superscripts. $[\cdot]_{m,n}$ denotes the (m,n)-th entry of a matrix. $(\cdot)^T$, $(\cdot)^H$, $\|\cdot\|$ and $\|\cdot\|_{\mathrm{F}}$ denote the transpose, conjugate transpose, Euclidean norm and Frobenius norm of the input entity, respectively. $|\cdot|$ denotes the absolute value if the argument is a scalar, or the cardinality if the argument is a multiset. $\times$ denotes the Cartesian product of two sets. $\mathbb{E}_X[\cdot]$ denotes expectation w.r.t. X, and the subscript is omitted when unambiguous. Expectations are taken with respect to the joint distribution of all random variables involved. For conditional expectations, the conditioned random variables are explicitly indicated. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a circular symmetric complex Gaussian distribution for vectors whose mean and covariance matrix are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. $\dim(\cdot)$ denotes the dimension of the argument. $\nabla f(x)$ denotes the gradient of $f(x)$. $\log_2$ represents the base-2 logarithm. log represents the natural logarithm. All information rate expressions are in units of nats for clarity, while the numerical results are presented in bits.

Chapters 3, 4 and 6 reuses notations introduced in Section 2.1 unless otherwise stated.

Chapter 2

Preliminaries

This chapter presents the preliminaries of the thesis, including RSMA, mutual information under finite alphabets, the GMI and subgradient methods.

2.1 Rate-splitting multiple access

This section introduces the principle of RSMA and presents the system model to be used in Chapters 3, 4 and 6.

We consider a downlink multi-user multiple-input single-output (MU-MISO) system where a transmitter equipped with N_{T} transmit antennas serves K users with unicast messages, $\{W_1, ..., W_K\}$. We assume narrowband channels and let n denote the transmission block length. At time instant $t \in \{1, 2, ..., n\}$, the received signal at user-k, $k \in \mathcal{K} = \{1, 2, ..., K\}$, can be written as[1]

$$y_{t,k} = \mathbf{h}_k^H \mathbf{x}_t + z_{t,k}, \tag{2.1}$$

where $\mathbf{x}_t$ is the transmitted signal, $\mathbf{h}_k$ is the channel between the transmitter and user-k and $z_{t,k}$ is the additive white Gaussian noise observed by user-k, model as an observation of $\mathcal{CN}(0, \sigma_z^2)$. We assume that the transmitter applies 1-layer RSMA [9, 16], and the transmit signal can be expressed as

$$\mathbf{x}_t = \mathbf{P}\mathbf{s}_t, \tag{2.2}$$

[1]The time indices will often be omitted in the rest of this thesis for clarity.

where $\mathbf{P} = [\mathbf{p}_{\mathrm{c}}, \mathbf{p}_{\mathrm{p},1}, ..., \mathbf{p}_{\mathrm{p},K}] \in \mathbb{C}^{N_{\mathrm{T}} \times (K+1)}$ represents the precoding matrix and $\mathbf{s}_t = [s_{t,\mathrm{c}}\ s_{t,\mathrm{p},1}, ..., s_{t,\mathrm{p},K}]^T \in \mathbb{C}^{(K+1)\times 1}$ represents the transmitted symbol vector. We assume $\mathbb{E}\{\mathbf{s}_t \mathbf{s}_t^H\} = \mathbf{I}$, and the transmission is subject to a power budget P_{T}, where $P_{\mathrm{T}} = \mathbb{E}\{\mathbf{x}_t^H \mathbf{x}_t\} = \|\mathbf{P}\|_{\mathrm{F}}^2$. A collection of symbol vectors at all the time instances can be written as $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n] \in \mathbb{C}^{(K+1)\times n}$. The rows of $\mathbf{S}$ are $K+1$ symbol blocks (streams), resulting in an alternative expression, $\mathbf{S} = [\mathbf{s}_{\mathrm{c}}, \mathbf{s}_{\mathrm{p},1}, ..., \mathbf{s}_{\mathrm{p},K}]^T$, where $\mathbf{s}_{\mathrm{c}}$ and $\mathbf{s}_{\mathrm{p},k}$, $\forall k \in \mathcal{K}$, are referred to as common and private streams in RSMA.

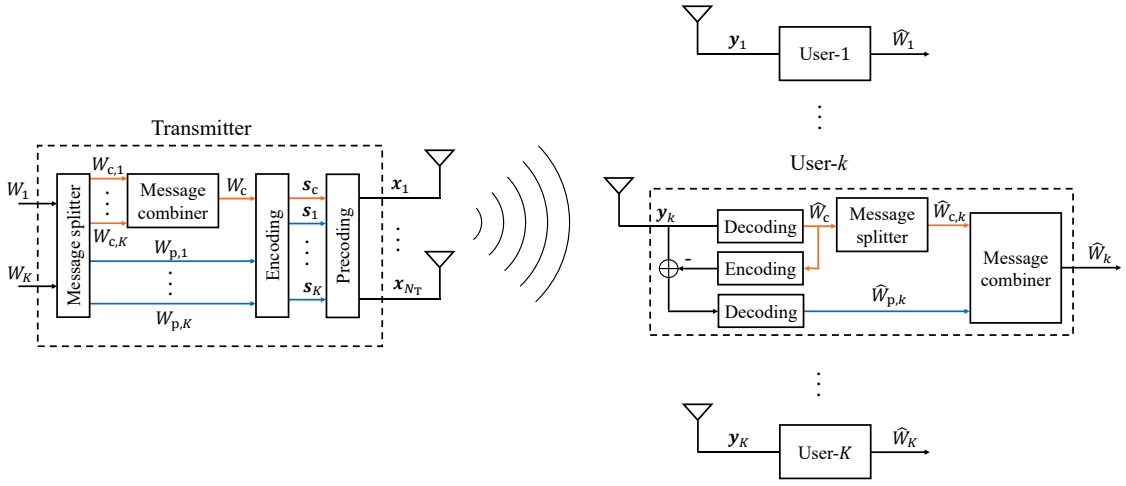


Figure 2.1: Single-layer RSMA with SIC receivers.

The key procedures in transmission and reception are shown in Figure 2.1 and explained as follows. Let $\{W_1, ..., W_K\}$ respectively denote the K unicast messages to be transmitted to the K users. The transmitter first processes the messages by splitting each message into common and private parts, e.g., W_k, $\forall k \in \mathcal{K}$, is split into $W_{\mathrm{c},k}$ and $W_{\mathrm{p},k}$. Next, the K common parts, i.e., $\{W_{\mathrm{c},1}, ..., W_{\mathrm{c},K}\}$ are combined into one common message, denoted by W_{c}, while the private parts, i.e., $\{W_{\mathrm{p},1}, ..., W_{\mathrm{p},K}\}$ are treated as individual messages. The resulting $K+1$ messages, $\{W_{\mathrm{c}}, W_{\mathrm{p},1}, ..., W_{\mathrm{p},K}\}$, are then encoded into $K+1$ common and private symbol streams, denoted by $\{\mathbf{s}_{\mathrm{c}}, \mathbf{s}_1, ..., \mathbf{s}_K\}$. Finally, these symbols are precoded and transmitted.

As for receiver implementation at user-k, $\forall k \in \mathcal{K}$, SIC is conventionally used as follows. The receiver first decodes W_{c} from $\mathbf{y}_k$ by treating $\{\mathbf{s}_1, ..., \mathbf{s}_K\}$ as noise and obtains $\widehat{W}_{\mathrm{c}}$, denoting the resulting estimate of W_{c}. $\widehat{W}_{\mathrm{c}}$ is then re-encoded into $\widehat{\mathbf{s}}_{\mathrm{c}}$, an estimate of the common stream. Next, the receiver decodes $W_{\mathrm{p},k}$ from $y_k - \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} \widehat{\mathbf{s}}_{\mathrm{c}}$, i.e., the received signal with the estimated common stream symbol canceled, by treating $\{\mathbf{s}_{k'} | k' \in \mathcal{K}/k\}$, as noise and obtains $\widehat{W}_{\mathrm{p},k}$, an estimate of the desired private message. Finally, the receiver extracts $\widehat{W}_{\mathrm{c},k}$ from $\widehat{W}_{\mathrm{c}}$ and combines it with $\widehat{W}_{\mathrm{p},k}$ to reconstruct $\widehat{W}_k$ as the resulting estimate of the original unicast message for user-k.

Throughout this thesis, the following three assumptions are made to facilitate the theoretical analysis:

1. channel state information is perfectly known at the transmitter and the receivers;
2. the transmission block is of infinite length;
3. all the messages are encoded at rates that are decodable at the corresponding receivers.

Hence, $\widehat{W_{\mathrm{c}}} = W_{\mathrm{c}}$, i.e., the cancellation of common stream is always perfect.[2]

Let $R_{\mathrm{c},k}$ and $R_{\mathrm{p},k}$ denote the achievable rate of W_{c} and $W_{\mathrm{p},k}$ at user-k. To ensure that the common stream can be decoded by all the users, the achievable common stream rate, denoted by R_{c}, depends on the worst-case user, i.e.,

$$R_{\mathrm{c}} = \min_{k} R_{\mathrm{c},k}. \tag{2.3}$$

Because W_{c} carries common part messages for potentially multiple users, R_{c} is distributed among these users, and this distribution depends on the common part message each user contributed to W_{c}. Let c_k denote the portion of the common stream rate that is allocated to user-k. Inherently, $c_k \geq 0$, $\forall k$, and $\sum_k c_k = 1$. The rate experienced by user-k, denoted by R_k, is given by

$$R_k = c_k R_{\mathrm{c}} + R_{\mathrm{p},k}. \tag{2.4}$$

The expressions for $R_{\mathrm{c},k}$ and $R_{\mathrm{p},k}$ vary throughout this thesis depending on the assumptions regarding the input signals and decoders.

Remark 2.1 (Connection with NOMA and SDMA)**.** *In the case of two users, 1-layer RSMA can be viewed as a generalization of power-domain NOMA [15], since it becomes equivalent to NOMA when a user's private stream is switched off and that user's message fills the entire common stream. For more than two users, NOMA is a special case of generalized RSMA [16]. Regardless of the number of users, Spatial Domain Multiple Access (SDMA), which has been widely used for downlink MU-MIMO, is a special case of 1-layer RSMA, as it reduces to SDMA when* $\|\mathbf{p}_c\|^2 = 0$.

[2]The effect of imperfect cancellation and error propagation is presented in Chapter 3 through LLS.

Remark 2.2 (Capacity region of MU-MIMO)**.** *We note that the capacity region of MU-MIMO under perfect CSIT is given by a non-linear strategy, namely dirty paper coding [71]. However, the practical implementation of dirty paper coding is highly complicated. Although the linear precoding approach is generally suboptimal, it is much more tractable.*

2.2 Mutual information of finite-alphabet input Gaussian channels

Consider a MIMO channel represented by

$$\mathbf{y} = \mathbf{H}^H \mathbf{P}\mathbf{s} + \mathbf{z}. \tag{2.5}$$

where $\mathbf{H}$ is the known channel matrix, $\mathbf{P}$ is the precoding matrix, $\mathbf{s}$ is an observation of a uniformly distributed random variable on finite support given by $\mathcal{X}$, and $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_z^2\mathbf{I})$ models the i.i.d. AWGN noise. Let $\mathbf{s}_m$ denote the m-th element in $\mathcal{X}$ in a certain order. The mutual information in bits between $\mathbf{s}$ and y is given by

$$I(\mathbf{s}; \mathbf{y}) = \log(|\mathcal{X}|) - H(\mathbf{s} \mid \mathbf{y}), \tag{2.6}$$

where $H(\mathbf{s} \mid \mathbf{y})$ is the conditional entropy of $\mathbf{s}$ given $\mathbf{y}$.[3]

Proposition 2.1.

$$H(\mathbf{s} \mid \mathbf{y}) = \frac{1}{|\mathcal{X}|} \sum_{m=1}^{|\mathcal{X}|} \mathbb{E}\Big\{ \log \sum_{l=1}^{|\mathcal{X}|} \exp\Big(- \frac{|\mathbf{H}^H\mathbf{P}(\mathbf{s}_m - \mathbf{s}_l) + \mathbf{z}|^2}{\sigma_z^2} \Big) \Big\} + 1. \tag{2.7}$$

Proof. Let $f(\mathbf{y} \mid \mathbf{s})$ denote the conditional probability density of $\mathbf{y}$ given $\mathbf{s}$.

$$\begin{aligned}
H(\mathbf{s} \mid \mathbf{y}) &= \mathbb{E}\{-\log P(\mathbf{s} \mid \mathbf{y})\} && \text{(2.8a)}\\
&= \mathbb{E}\left\{ -\log \frac{f(\mathbf{y} \mid \mathbf{s}) P(\mathbf{s})}{\sum_{l=1}^{|\mathcal{X}|} f(\mathbf{y} \mid \mathbf{s}_l) P(\mathbf{s}_l)} \right\} && \text{(2.8b)}\\
&= \mathbb{E}\left\{ \log \frac{\sum_{l=1}^{|\mathcal{X}|} f(\mathbf{y} \mid \mathbf{s}_l)}{f(\mathbf{y} \mid \mathbf{s})} \right\} && \text{(2.8c)}\\
&= \mathbb{E}\left\{ \log \frac{\sum_{l=1}^{|\mathcal{X}|} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}^H\mathbf{P}\mathbf{s}_l\|^2}{\sigma_z^2} \right)}{\exp\left(-\frac{\|\mathbf{z}\|^2}{\sigma_z^2} \right)} \right\} && \text{(2.8d)}\\
&= \mathbb{E}\left\{ \log \sum_{l=1}^{|\mathcal{X}|} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}^H\mathbf{P}\mathbf{s}_l\|^2}{\sigma_z^2} \right) \right\} + \log e && \text{(2.8e)}\\
&= \frac{1}{|\mathcal{X}|} \sum_{m=1}^{|\mathcal{X}|} \mathbb{E}\left\{ \log \sum_{l=1}^{|\mathcal{X}|} \exp\left(-\frac{\|\mathbf{H}^H\mathbf{P}(\mathbf{s}_m - \mathbf{s}_l) + \mathbf{z}\|^2}{\sigma_z^2} \right) \right\} + 1, && \text{(2.8f)}
\end{aligned}$$

[3]Throughout this thesis, we assume that the channel is known. The mutual information and conditional entropy are therefore evaluated by considering $\mathbf{H}$ is a known constant.

where

- (2.8a) follows from the definition of conditional entropy.
- (2.8b) follows from Bayes' theorem.
- (2.8c) follows from the assumption that $\mathbf{s}$ is uniformly distributed.
- (2.8d) follows from the definition of the Gaussian distribution.
- (2.8e) follows from splitting the numerator and the denominator in the logarithm and computing the expectation of the denominator part.
- (2.8f) follows from substituting $\mathbf{y}$ by (2.5) and simplify the expectation over $\mathbf{s}$ using $P(\mathbf{s}) = \frac{1}{|\mathcal{X}|}$, $\forall \mathbf{s} \in \mathcal{X}$.

□

The expectation in (2.7) cannot be resolved analytically. Therefore, computing (2.7) requires numerical methods such as Monte Carlo and Gauss-Hermite quadratures.

By applying Jensen's inequality and a constant shift, an approximation of (2.7) can be obtained as [72]

$$H(\mathbf{s} \mid \mathbf{y})^{\text{approx.}} = \frac{1}{|\mathcal{X}|} \sum_{m=1}^{|\mathcal{X}|} \log \sum_{l=1}^{|\mathcal{X}|} \exp\Big(- \frac{|\mathbf{H}^H \mathbf{P}(\mathbf{s}_m - \mathbf{s}_l))|^2}{2\sigma^2} \Big). \tag{2.9}$$

As a consequence of Shannon's channel coding theorem [1], if $\mathbf{H}$ and σ_z^2 are known at the transmitter and receiver, any code rate below $I(\mathbf{s}; \mathbf{y})$ is achievable.[4] Note that the mutual information under finite-alphabet inputs is referred to in the literature by several other names, such as constellation-constrained (CC) rates and the capacity of coded modulation. In Chapter 3, the term "CC achievable rate" is used to distinguish it from the commonly cited achievable rate under Gaussian input signals. In other chapters, "CC" is omitted, as the finite-alphabet assumption applies throughout those discussions.

[4]Shannon's noisy-channel coding theorem establishes that error-free communication is achievable when the code rate is below the mutual information, whose maximum over all input distributions defines the channel capacity. In the literature, mutual information, channel capacity or metrics based on them are conventionally described as "achievable." This thesis follows the same convention.

2.3 The generalized mutual information

In this section, a proof of the generalized mutual information is provided using the method of error exponent, which can be obtained from [73, 61]. More details on the method of error exponent can be found in [58, 74]. Note that the GMI has primal and dual forms. The proof given in this section directly leads to the dual form, which is the form that will be used in this thesis. A proof of the equivalence between the primal and the dual forms is given in [59].

Consider a communication link with block length n characterized by an n-dimensional channel transition probability (density) given by $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x})$. The encoder is represented by the following code:

$$\mathcal{C} = \{\mathbf{x}^{(1)}, \ ..., \ \mathbf{x}^{(M)}\}, \tag{2.10}$$

i.e., $\mathbf{x}^{(m)}$ is transmitted if the message is m. We consider a decoder represented as

$$\hat{m} = \underset{l=1, \ ..., \ M}{\arg\max} \ q^n(\mathbf{x}^{(l)}, \mathbf{y}), \tag{2.11}$$

where $q^n(\mathbf{x}^{(l)}, \mathbf{y})$ is referred as the decoding metric. The decoder in (2.11) chooses the message by maximizing the decoding metric.

We consider memoryless channels, i.e.,

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}) = \prod_{i=1}^{n} f_{Y|X}(y_i \mid x_i), \tag{2.12}$$

and an i.i.d. random coding ensemble on the codewords, i.e.,

$$f_{\mathbf{X}^{(m)}}(\mathbf{x}^{(m)}) = \prod_{i=1}^{n} f_X(x_i^{(l)}), \quad m = 1, \ ..., \ M. \tag{2.13}$$

We also assume that the decoding metric, $q^n(\mathbf{x}, \mathbf{y})$, can be factorized into symbol-wise decoding metrics, $q(x, y_i)$, i.e.,

$$q^n(\mathbf{x}^{(l)}, \mathbf{y}) = \prod_{i=1}^{n} q(x_i^{(l)}, y_i). \tag{2.14}$$

This is because the optimal decoding would be separable in symbols under memoryless channels and the i.i.d. random coding ensemble, and a separable decoding metric, and factorized decoding metrics permit analytical tractability.

Theorem 2.1 (The random coding exponent and the generalized mutual information)**.** *There exists a channel code with $M = e^{nR}$ such that the average error probability of the code, denoted by $p_e(\mathcal{C})$, satisfies*

$$p_e(\mathcal{C}) \leq e^{-n(E_0(\rho,\, s,\, f_X)-\rho R)}. \tag{2.15}$$

where

$$E_0(\rho, s, f_X) = -\log \mathbb{E}\left[\left(\frac{\mathbb{E}\left[q(\bar{X}, Y)^s \mid Y\right]}{q(X,Y)^s}\right)^{\rho}\right] \quad \forall\, 0 \leq \rho \leq 1,\ s > 0 \tag{2.16}$$

with $(X, Y, \bar{X}) \sim f_X(x) f_{Y|X}(y \mid x) f_X(\bar{x})$, and R is achievable if

$$R \leq I_{\text{GMI}} \triangleq \sup_{s>0}\ \mathbb{E}\left[\log\left(\frac{q(X,Y)^s}{\mathbb{E}\left[q(\bar{X}, Y)^s \mid Y\right]}\right)\right] \tag{2.17}$$

Proof. For a randomly generated $\mathcal{C}$ whose codewords are i.i.d., the average error probability over all possible codes is given by

$$\bar{p}_e = \mathbb{E}[p_e(\mathcal{C})] \tag{2.18}$$

$$= \frac{1}{M}\sum_{l=1}^{M}\mathbb{E}\left[P\left[\bigcup_{l'\neq l}\left\{q(\mathbf{X}^{(l')},\mathbf{Y}) \geq q(\mathbf{X}^{(l)},\mathbf{Y}) \mid \mathbf{X}^{(l)},\mathbf{Y}\right\}\right]\right] \tag{2.19}$$

$$\leq \mathbb{E}\left[P\left[\bigcup_{l'=2}^{M}\left\{q(\mathbf{X}^{(l')},\mathbf{Y}) \geq q(\mathbf{X}^{(1)},\mathbf{Y}) \mid \mathbf{X}^{(1)},\mathbf{Y}\right\}\right]\right] \tag{2.20}$$

$$\leq \mathbb{E}\left[\min\left\{1,(M-1)P\left[\left\{q(\bar{\mathbf{X}},\mathbf{Y}) \geq q(\mathbf{X},\mathbf{Y}) \mid \mathbf{X},\mathbf{Y}\right\}\right]\right\}\right] \tag{2.21}$$

$$\leq \mathbb{E}\left[\min\left\{1,(M-1)P\left[\left\{q(\bar{\mathbf{X}},\mathbf{Y})^s \geq q(\mathbf{X},\mathbf{Y})^s \mid \mathbf{X},\mathbf{Y}\right\}\right]\right\}\right] \tag{2.22}$$

$$\leq \mathbb{E}\left[\min\left\{1,(M-1)\frac{\mathbb{E}\left[q(\bar{\mathbf{X}},\mathbf{Y})^s \mid \mathbf{Y}\right]}{q(\mathbf{X},\mathbf{Y})^s}\right\}\right] \tag{2.23}$$

$$\leq \mathbb{E}\left[\left((M-1)\frac{\mathbb{E}\left[q(\bar{\mathbf{X}},\mathbf{Y})^s \mid \mathbf{Y}\right]}{q(\mathbf{X},\mathbf{Y})^s}\right)^{\rho}\right] \tag{2.24}$$

$$\leq M^{\rho}\mathbb{E}\left[\left(\frac{\mathbb{E}\left[q(\bar{\mathbf{X}},\mathbf{Y})^s \mid \mathbf{Y}\right]}{q(\mathbf{X},\mathbf{Y})^s}\right)^{\rho}\right] \tag{2.25}$$

$$= M^{\rho}\prod_{i=1}^{n}\mathbb{E}\left[\left(\frac{\mathbb{E}\left[q(\bar{X}_i,Y_i)^s \mid Y_i\right]}{q(X_i,Y_i)^s}\right)^{\rho}\right] \tag{2.26}$$

$$= M^{\rho}\left(\mathbb{E}\left[\left(\frac{\mathbb{E}\left[q(\bar{X},Y)^s \mid Y\right]}{q(X,Y)^s}\right)^{\rho}\right]\right)^{n} \tag{2.27}$$

where

- (2.18) averages the error probability over all possible codes with slight abuse of the notation;
- (2.19) follows from averaging the error probability of a specific codeword, i.e., the probability that, given $\mathbf{X}^{(l)}$ and $\mathbf{Y}$, there is another codeword, $\mathbf{X}^{(l')}$ yields a larger decoding metric than $\mathbf{X}^{(l)}$;
- (2.20) follows from the fact that every codeword has the same error probability for the random coding ensemble, and we can only consider the error probability of $\mathbf{X}^{(1)}$;
- (2.21) follows from the truncated union bound, i.e., $P(\cup_k \mathcal{A}_k) \leq \min\{1, \sum_k P(\mathcal{A}_k)\}$;
- (2.22) holds for any $s > 0$;
- (2.23) follows from Markov's inequality,

- (2.24) applied the inequality $\min\{1, x\} \leq x^{\rho}$ for $0 \leq \rho \leq 1$;

- (2.25) follows from the assumption that the channel is memoryless and that the symbols of every codeword are i.i.d.;

- (2.27) follows from the fact that each term in the product in (2.25) is the same except the dummy index i.

By defining $R = \frac{\log M}{n}$ and the error exponent

$$E_0(\rho, s, f_X) = -\log \mathbb{E}\left[\left(\frac{\mathbb{E}\left[q(\bar{X}, Y)^s \mid Y\right]}{q(X, Y)^s}\right)^{\rho}\right] \tag{2.28}$$

we obtain

$$\bar{p}_e \leq e^{-n(E_0(\rho, s, f_X) - \rho R)}. \tag{2.29}$$

Since (2.29) gives an upper bound of the average error probability over all the randomly generated codes, there must exist one code in the ensemble whose error probability satisfies this bound.

It can be shown that, for $\rho \geq 0$,

$$E_0(\rho, s, f_X) \geq 0 \tag{2.30}$$

$$\frac{\partial}{\partial \rho} E_0(\rho, s, f_X) > 0 \tag{2.31}$$

and

$$\frac{\partial^2}{\partial \rho^2} E_0(\rho, s, f_X) \leq 0, \tag{2.32}$$

From (2.29), for fixed s, $\lim_{n\to\infty} \bar{p}_e = 0$ only if

$$R < \frac{\partial}{\partial \rho} E_0(\rho, s, f_X)\Big|_{\rho=0} = \mathbb{E}\left[\log\left(\frac{q(X, Y)^s}{\mathbb{E}\left[q(\bar{X}, Y)^s \mid Y\right]}\right)\right]. \tag{2.33}$$

By optimizing s, we conclude that any $R < I_{\mathrm{GMI}}$, where

$$I_{\mathrm{GMI}} = \sup_{s>0} \mathbb{E}\left[\log\left(\frac{q(X, Y)^s}{\mathbb{E}\left[q(\bar{X}, Y)^s \mid Y\right]}\right)\right] \tag{2.34}$$

is achievable. □

Note that

$$\frac{d^2}{ds^2}\mathbb{E}\left[\log\left(\frac{q(X,Y)^s}{\mathbb{E}\left[q(\bar{X},Y)^s \mid Y\right]}\right)\right]$$
$$= -\mathbb{E}\left[\frac{\mathbb{E}\left[q(\bar{X},Y)^s\left(\log q(\bar{X},Y)\right)^2 \mid Y\right]\mathbb{E}\left[q(\bar{X},Y)^s \mid Y\right] - \mathbb{E}\left[q(\bar{X},Y)^s \log q(\bar{X},Y) \mid Y\right]^2}{\mathbb{E}\left[q(\bar{X},Y)^s \mid Y\right]^2}\right]$$

$$\leq 0, \tag{2.35}$$

where the inequality is given by Cauchy's inequality ($\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$) as follows.

$$\begin{aligned}
&\mathbb{E}\left[q(\bar{X},Y)^s \log q(\bar{X},Y) \mid Y\right]^2 \\
&\leq \mathbb{E}\left[\left(q(\bar{X},Y)^{s/2}\left(\log q(\bar{X},Y)\right)\right)^2 \mid Y\right]\mathbb{E}\left[\left(q(\bar{X},Y)^{s/2}\right)^2 \mid Y\right] \\
&\leq \mathbb{E}\left[q(\bar{X},Y)^s\left(\log q(\bar{X},Y)\right)^2 \mid Y\right]\mathbb{E}\left[q(\bar{X},Y)^s \mid Y\right]
\end{aligned} \tag{2.36}$$

Therefore, optimization of $s > 0$ in (2.34) is a convex problem.

More properties of the GMI can be found in [61], for example, the expurgation and the ensemble tightness of the GMI, i.e., it is the highest achievable rate under the i.i.d. random coding ensemble.

2.4 Subgradient

Gradient descent, as a widely used optimization algorithm, requires the objective function to be differentiable. To address non-differentiable functions, algorithms based on subgradients can be employed. In this section, we introduce the concept of the subgradient.

Given a convex and differentiable function f of real vectors, the first order criterion states that

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f). \tag{2.37}$$

The first order criterion defines the tightest affine function that lower-bounds f and equals it when $\mathbf{x} = \mathbf{y}$. Subgradients serve the same purpose for non-differentiable functions.

Definition 2.1 (Subgradient)**.** $\mathbf{g}$ *is a subgradient of a convex function* f *at* $\mathbf{x}$ *if*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}), \quad \forall \mathbf{y} \in dom(f). \tag{2.38}$$

Subgradients have the following properties:

- Subgradients always exist in the relative interior of the domain of f.
- If f is differentiable at $\mathbf{x}$, then $\mathbf{g} = \nabla f(\mathbf{x})$ uniquely.
- There can be an infinite number of subgradients.

For example, with $f(x) = |x|$, for $x \neq 0$, the subgradient is given by $g = \text{sign}(x)$ uniquely; for $x = 0$, the subgradient is any element of [-1,1].

The collection of all subgradients of f at $\mathbf{x}$ is called the subdifferential.

Definition 2.2 (subdifferential)**.** *Given a convex function* f*, the subdifferential at* $\mathbf{x} \in dom(f)$ *is*

$$\partial f(\mathbf{x}) = \{\mathbf{g} \mid f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}),\ \forall \mathbf{y} \in dom(f)\}. \tag{2.39}$$

For convex non-differential optimization, one can apply subgradient descent methods, which take one subgradient as the search direction. [75, 76, 77] are recommended for further readings on subgradients, subdifferential and subgradient descent.

For non-convex functions, subgradients satisfying (2.38) may or may not exist. To enable the analysis of non-convex and non-smooth functions through convex analysis, a natural generalization of subdifferential, namely Clarke Subdifferential [78], was introduced.

Definition 2.3 (Clarke Subdifferential)**.** *Given a function* f *of real vectors, the Clarke differential at* $\mathbf{x} \in dom(f)$*, denoted by* $\partial_C f(\mathbf{x})$*, is defined as*

$$\partial_C f(\mathbf{x}) = \text{Co}\left\{\mathbf{s} \in \mathbb{R}^n : \exists \mathbf{x}' \to \mathbf{x},\ \nabla f(\mathbf{x}') \text{ exists, and } \nabla f(\mathbf{x}') \to \mathbf{s}\right\}, \tag{2.40}$$

where $\text{Co}\{\cdot\}$ *represents the convex hull of the argument set.*

Every element of $\partial_C f(\mathbf{x})$ is a subgradient of $f(\mathbf{x})$, denoted by $\nabla_C f(\mathbf{x})$, and can be used as a search direction for non-convex non-differential optimization problems.[5] [78] is recommended for additional readings on Clarke's subdifferential.

[5] [79] shows that, if $f(\mathbf{x})$ is a locally Lipschitz function, for every $\mathbf{x}$, there exists a point $\mathbf{y}$ for which $\nabla f(\mathbf{y})$ exists.

Chapter 3

SIC-free Implementation for RSMA

In this chapter, we jointly address three practical but fundamental questions on RSMA: (1) How to exploit the benefit of RSMA under finite constellations? (2) What are the potential and promising ways to implement RSMA receivers? (3) Can RSMA retain its superiority in the absence of successive interference cancellers (SIC)? To address these concerns, we first propose low-complexity precoder designs incorporating finite constellations and show that the potential of RSMA is better achieved with such designs than those assuming Gaussian input signals. We then consider some practical receiver designs that can be applied to RSMA. We notice that these receiver designs follow one of two principles: (1) SIC: cancelling upper layer signals before decoding the lower layer, and (2) SIC-free: treating upper layer signals as noise when decoding the lower layer. In light of this, we propose to alter the system design according to the receiver category. Through link-level simulations, the effectiveness of the proposed power allocation and receiver designs is verified. More importantly, we show that it is possible to preserve the superiority of RSMA over Spatial Domain Multiple Access (SDMA), including SDMA with advanced receivers, even without SIC in the receivers. Those results therefore open the door to competitive implementable RSMA strategies for 6G and beyond communications.

The key contributions of this chapter are listed as follows.

1. With the distinction between receivers with and without SIC clarified, we derive the CC sum-rate expression and its approximation for RSMA with or without SIC, with some assumptions on the precoders made such that the

expressions are more tractable. Subject to actual constraints on the constellations in use, these rate expressions provide more accurate metrics for RSMA than those commonly used for Gaussian input signals.

2. We simplify the problem of precoder design by fixing precoder directions and power distribution among private streams. Based on the derived approximations of CC sum-rate expressions, we propose low-complexity methods to choose the power allocation of common and private streams. Closed-form power allocation for optimizing the sum-rate assuming Gaussian input signals is also obtained as a benchmark.

3. Despite being rarely studied in the literature, variations or substitutes of SIC are widely studied in standardization bodies [80, 81, 82]. In light of this, apart from the basic SIC receiver as in [49, 31, 50, 51, 52], we consider several receiver designs that can be applied with RSMA in real applications to provide more choices in terms of complexity and performance. Unlike works such as [53] where some components of a basic SIC receiver are improved, we focus on the overall architecture of the receiver and do not limit the design of any building blocks in it; that is, most of the de-mapper designs (such as maximum-likelihood de-mapper, max-log de-mapper and list sphere de-mapper) and decoder designs (such as belief propagation decoding and successive-cancellation list decoding for polar codes) can be adopted into the proposed receiver designs. The complexity of different receivers is quantified, such that a complexity-performance trade-off with different receiver designs is revealed. We further notice that the receiver designs can be classified into SIC-type receivers and SIC-free-type receivers based on their principles, and therefore propose to alter the system design according to this classification. This also indicates that these receivers respectively push RSMA towards the two different information theoretic limits, which are characterized by the CC sum-rate expressions assuming whether SIC is adopted or not.

4. We evaluate RSMA with proposed power allocation methods and transceiver designs using LLS. We show that CC sum-rate oriented system designs lead to a significant performance improvement in practice in comparison with designs that assume Gaussian signalling and SDMA. We also show that, although receivers with SIC lead to some performance loss compared to SIC-type receivers, they still retain most of the superiority of RSMA over SDMA. Thus, we reveal an important fact that it is possible to preserve the superiority of RSMA over SDMA, including SDMA with advanced receivers, even without

SIC at the receivers.

Organization: The remainder of this chapter is organized as follows: In Section 3.1, low-complexity precoder designs for RSMA are proposed that incorporate different optimization objectives. Section 3.2 summarizes the practical receiver implementations for RSMA. In Section 3.3, the numerical results are presented and analyzed. Finally, Section 3.4 summarizes this paper.

Unless otherwise stated, the system model presented in Section 2.1 is assumed throughout this chapter.

3.1 Low-complexity precoder designs for RSMA

We follow the system model for RSMA introduced in Section 2.1. To obtain simple system designs for practical uses, we assume that the transmit power is uniformly distributed among the private streams, and hence the transmit signal can be expressed (2.2) as

$$\mathbf{x} = \sqrt{P_{\mathrm{T}} t}\bar{\mathbf{p}}_{\mathrm{c}} s_{\mathrm{c}} + \sqrt{\frac{P_{\mathrm{T}}(1-t)}{K}} \sum_{k=1}^{K} \bar{\mathbf{p}}_{\mathrm{p},k} s_{\mathrm{p},k}, \tag{3.1}$$

where P_{T} is the total transmit power budget, $\{\bar{\mathbf{p}}_{\mathrm{c}}, \bar{\mathbf{p}}_{\mathrm{p},1}, ..., \bar{\mathbf{p}}_{\mathrm{p},K}\}$ are unit-norm vectors in $\mathbb{C}^{N_{\mathrm{T}}\times 1}$ indicating the direction of the precoders, and $t \in [0,1]$ is the fraction of the power budget allocated to the common stream. $\{\bar{\mathbf{p}}_{\mathrm{p},1}, ..., \bar{\mathbf{p}}_{\mathrm{p},K}\}$ are designed using zero-forcing (ZF) precoding. Specifically, we define $\bar{\mathbf{H}} \triangleq [\bar{\mathbf{h}}_1, \bar{\mathbf{h}}_2, ..., \bar{\mathbf{h}}_K] = [\frac{\mathbf{h}_1}{\|\mathbf{h}_1\|}, \frac{\mathbf{h}_2}{\|\mathbf{h}_2\|}, ..., \frac{\mathbf{h}_K}{\|\mathbf{h}_K\|}]$. With ZF, $[\bar{\mathbf{p}}_{\mathrm{p},1}, ..., \bar{\mathbf{p}}_{\mathrm{p},K}]$ are obtained by column-wise normalization of $\mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1}$. For $\bar{\mathbf{p}}_{\mathrm{c}}$, given its multicast nature, we heuristically hope to employ solutions for the following multicast problem.[1]

$$\begin{aligned} \max_{\bar{\mathbf{p}}_{\mathrm{c}}} \ & \min_{k} |\bar{\mathbf{h}}_k^H \bar{\mathbf{p}}_{\mathrm{c}}|^2 \\ \text{s.t.} \ & \|\bar{\mathbf{p}}_{\mathrm{c}}\|^2 = 1 \end{aligned} \tag{3.2}$$

The solution to (3.2) for general cases has been addressed in [83, 84, 85]. Specifically, for two-user cases, [85] provides the globally optimal $\bar{\mathbf{p}}_{\mathrm{c}}$ in closed form assuming that $|\bar{\mathbf{h}}_1^H \bar{\mathbf{p}}_{\mathrm{c}}|^2 = |\bar{\mathbf{h}}_2^H \bar{\mathbf{p}}_{\mathrm{c}}|^2$ is necessary for the solution. In this chapter, we design $\bar{\mathbf{p}}_{\mathrm{c}}$ using the solution from [85] for $K = 2$, and the dominant left singular vector of $\bar{\mathbf{H}}$ for $K \geq 3$ for their low complexity. Other designs, for example, those in [83] and [84], are also valid.

[1] Any other choice of $\bar{\mathbf{p}}_{\mathrm{c}}$ is valid here without affecting any other parts of the rest of the design.

With $\{\bar{\mathbf{p}}_{\mathrm{c}}, \bar{\mathbf{p}}_{\mathrm{p},1}, ..., \bar{\mathbf{p}}_{\mathrm{p},K}\}$ decided, t becomes the only remaining variable. In what follows, two methods for determining t, namely optimizing sum-rate and optimizing CC sum-rate,[2] are introduced.

3.1.1 Sum-rate optimization

The power allocation coefficient, t, is obtained by solving a sum-rate maximization problem. With the above choice of $\{\bar{\mathbf{p}}_{\mathrm{c}}, \bar{\mathbf{p}}_1, ..., \bar{\mathbf{p}}_K\}$ and the assumption of Gaussian input signals, the sum-rate of RSMA with SIC and Gaussian inputs can be written as

$$R_{\mathrm{sum}} = \min_k \log\Big(1 + \frac{P_{\mathrm{T}} t |\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{c}}|^2}{\frac{P_{\mathrm{T}}(1-t)}{K}|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},k}|^2 + \sigma_z^2}\Big) + \sum_{k=1}^{K} \log\Big(1 + \frac{P_{\mathrm{T}}(1-t)}{K\sigma_z^2}|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},k}|^2\Big), \tag{3.3}$$

where we have used the fact that $|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},i}|^2 = 0, \forall i \neq k$. The first term in (3.3) represents the achievable rate of the common stream such that it can be decoded by all the users. The second term represents the sum of the achievable rates of all the private streams.

If $|\bar{\mathbf{h}}_1^H \bar{\mathbf{p}}_{\mathrm{c}}| = |\bar{\mathbf{h}}_2^H \bar{\mathbf{p}}_{\mathrm{c}}| = ... = |\bar{\mathbf{h}}_K^H \bar{\mathbf{p}}_{\mathrm{c}}|$, or $P_{\mathrm{T}} \to +\infty$, it is trivial to see that

$$\begin{aligned} k' &\triangleq \arg\min_k \log\Big(1 + \frac{P_{\mathrm{T}} t |\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{c}}|^2}{\frac{P_{\mathrm{T}}(1-t)}{K}|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},k}|^2 + \sigma_z^2}\Big) \\ &= \arg\min_k \frac{|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{c}}|^2}{|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},k}|^2}, \end{aligned} \tag{3.4}$$

and k' depends on $\{\bar{\mathbf{p}}_c, \bar{\mathbf{p}}_{\mathrm{p},k}, ..., \bar{\mathbf{p}}_{\mathrm{p},k}\}$ but not t and therefore can be pre-determined.

We can then further simplify (3.3) into

$$\begin{aligned} R_{\mathrm{sum}} =& \log\Big(1 + \frac{P_{\mathrm{T}} t}{\sigma_z^2}|\mathbf{h}_{k'}^H \bar{\mathbf{p}}_{\mathrm{c}}|^2 + \frac{P_{\mathrm{T}}(1-t)}{K\sigma_z^2}|\mathbf{h}_{k'}^H \bar{\mathbf{p}}_{\mathrm{p},k'}|^2\Big) \\ &+ \sum_{k=1, k\neq k'}^{K} \log\Big(1 + \frac{P_{\mathrm{T}}(1-t)}{K\sigma_z^2}|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},k}|^2\Big). \end{aligned} \tag{3.5}$$

(3.5) is a concave function of t. Let $t^\star$ be the globally optimal t that maximizes (3.5). Using the first-order derivative, $t^\star$ is given by

[2] To clarify, in this chapter, sum-rate implies that all the streams are transmitted generated from Gaussian distributed symbols, while CC sum-rate implies that the input signal is constrained to be a discrete and finite set.

$$t^\star = \begin{cases} 0 & \text{if } -\sum_i \frac{a_i}{1+a_i} + \frac{b+a_{k'}}{1+a'_k} \leq 0 \\ 1 & \text{if } -\sum_i \frac{a_i}{1+a_i} + \frac{b+a_{k'}}{1+a'_k} > 0 \text{ and} \\ & \quad -\sum_{i\neq k} a_i + \frac{b}{1+a_{k'}+b} \geq 0 \\ t_0 & \text{otherwise,} \end{cases} \tag{3.6}$$

where $a_k = \frac{P_\mathrm{T}}{K\sigma_z^2}|\mathbf{h}_k^H \bar{\mathbf{p}}_{\mathrm{p},k}|^2,\ k \in \{1, ..., K\}$, $b = P_\mathrm{T}\left(|\mathbf{h}_{k'}^H \bar{\mathbf{p}}_\mathrm{c}|^2 - \frac{1}{K\sigma_z^2}|\mathbf{h}_{k'}^H \bar{\mathbf{p}}_{\mathrm{p},k'}|^2\right)$, and t_0 is the only real root in $[0, 1]$ of the following polynomial of t.

$$\sum_{i=1,i\neq k'}^{K} a_i \prod_{j=1,j\neq i,k'}^{K} (ta_j - a_j - 1)(tb + a_{k'} + 1) + b \prod_{i=1,i\neq k'}^{K} (-ta_i + a_i + 1) \tag{3.7}$$

Although the assumption of Gaussian input signals eases the analysis of system performance, practical systems utilize only finite constellations. Next, two designs considering finite constellations are introduced.

3.1.2 Constellation-constrained sum-rate optimization

In what follows, designing t to optimize the achievable sum rate under finite constellations, i.e., CC sum-rate, is addressed. We consider two cases:

1. all the users apply SIC receivers: each user can perfectly remove common stream signals before decoding the desired private stream, as long as the common stream is decodable by all the users[3];
2. none of the users applies SIC receivers: each user decodes the private stream without any prior knowledge of the common message, i.e., treating the common stream signal as noise when decoding the desired private stream.

Proposition 3.1. *The CC sum-rate of RSMA with SIC receivers is expressed as (3.8) and can be well approximated by (3.9), where* $\mathbf{p}_\mathrm{c} = \sqrt{P_\mathrm{T} t}\bar{\mathbf{p}}_\mathrm{c}$, $\mathbf{p}_{\mathrm{p},k} = \sqrt{\frac{P_\mathrm{T}(1-t)}{K}}\bar{\mathbf{p}}_{\mathrm{p},k}$, *$\mathcal{X}_\mathrm{c}$ is the finite size constellation set used for the common stream, $\mathcal{X}_k$ is the constellation set used for the private stream intended for the k-th user, $s_{k,m}$ is the m-th element in $\mathcal{X}_k$ and $\mathbf{s}_{k,m} \in \mathbb{C}^{2\times 1}$ contains two entries drawn from the m-th element in $\mathcal{X}_\mathrm{c} \times \mathcal{X}_k$.*

Proof. See Appendix B.1. □

[3]This is guaranteed by the assumptions made in Section 2.1.

Proposition 3.2. *The CC sum-rate of RSMA without SIC receivers is expressed as (3.10) and can be well approximated by (3.11).*

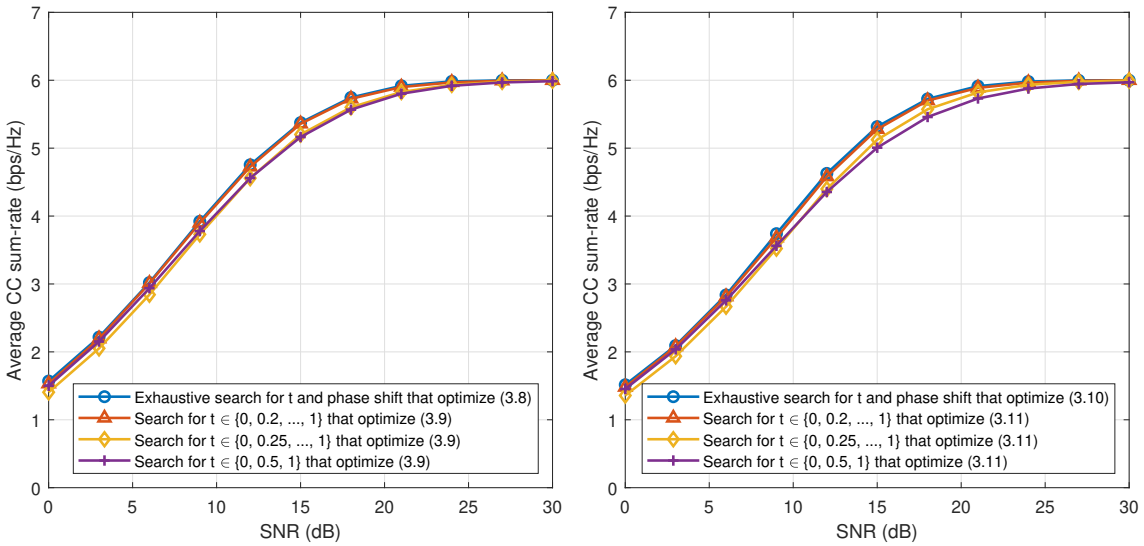
Proof. See Appendix B.1. □

$$
\begin{aligned}
R_{\mathrm{CC,sum}}^{\mathrm{SIC}} =& \log|\mathcal{X}_{\mathrm{c}}| - K + \min_{k\in\{1,\ldots,K\}} \Bigg(-\frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_k|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H[\mathbf{p}_{\mathrm{c}}\ \mathbf{p}_{\mathrm{p},k}](\mathbf{s}_{k,m}-\mathbf{s}_{k,l})+z|^2}{\sigma_z^2}\Big)\Big\}\Bigg) \\
&+\frac{1}{|\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_k|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H\mathbf{p}_{\mathrm{p},k}(s_{k,m}-s_{k,l})+z|^2}{\sigma_z^2}\Big)\Big\} \\
&+\sum_{k'=1}^{K}\log|\mathcal{X}_{k'}| - \frac{1}{|\mathcal{X}_{k'}|}\sum_{m=1}^{|\mathcal{X}_{k'}|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_{k'}|}\exp\Big(-\frac{|\mathbf{h}_{k'}^H\mathbf{p}_{\mathrm{p},k'}(s_{k',m}-s_{k',l})+z|^2}{\sigma_z^2}\Big)\Big\}
\end{aligned} \tag{3.8}
$$

$$
\begin{aligned}
R_{\mathrm{CC,sum}}^{\mathrm{SIC}} \approx& \log|\mathcal{X}_{\mathrm{c}}| + \min_{k\in\{1,\ldots,K\}} \Bigg(-\frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_k|}\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H[\mathbf{p}_{\mathrm{c}}\ \mathbf{p}_{\mathrm{p},k}](\mathbf{s}_{k,m}-\mathbf{s}_{k,l})|^2}{2\sigma_z^2}\Big)\Bigg) \\
&+\frac{1}{|\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_k|}\log\sum_{l=1}^{|\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H\mathbf{p}_{\mathrm{p},k}(s_{k,m}-s_{k,l})|^2}{2\sigma_z^2}\Big)\Bigg) \\
&+\sum_{k'=1}^{K}\log|\mathcal{X}_{k'}| - \frac{1}{|\mathcal{X}_{k'}|}\sum_{m=1}^{|\mathcal{X}_{k'}|}\log\sum_{l=1}^{|\mathcal{X}_{k'}|}\exp\Big(-\frac{|\mathbf{h}_{k'}^H\mathbf{p}_{\mathrm{p},k'}(s_{k',m}-s_{k',l})|^2}{2\sigma_z^2}\Big)
\end{aligned} \tag{3.9}
$$

$$
\begin{aligned}
R_{\text{CC,sum}}^{\text{SIC-free}} = \log|\mathcal{X}_{\text{c}}| + \min_{k\in\{1,\ldots,K\}} \Bigg(& -\frac{1}{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H[\mathbf{p}_{\text{c}}\ \mathbf{p}_{\text{p},k}](\mathbf{s}_{k,m}-\mathbf{s}_{k,l})+z|^2}{\sigma_z^2}\Big)\Big\} \\
& + \frac{1}{|\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_k|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H\mathbf{p}_{\text{p},k}(s_{k,m}-s_{k,l})+z|^2}{\sigma_z^2}\Big)\Big\}\Bigg) \\
& + \sum_{k'=1}^{K}\log|\mathcal{X}_{k'}| - \frac{1}{|\mathcal{X}_{\text{c}}\times\mathcal{X}_{k'}|}\sum_{m=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_{k'}|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_{k'}|}\exp\Big(-\frac{|\mathbf{h}_{k'}^H[\mathbf{p}_{\text{c}}\ \mathbf{p}_{\text{p},k'}](\mathbf{s}_{k',m}-\mathbf{s}_{k',l})+z|^2}{\sigma_z^2}\Big)\Big\} \\
& + \frac{1}{|\mathcal{X}_{\text{c}}|}\sum_{m=1}^{|\mathcal{X}_{\text{c}}|}\mathbb{E}_z\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\text{c}}|}\exp\Big(-\frac{|\mathbf{h}_{k'}^H\mathbf{p}_{\text{c}}(s_{\text{c},m}-s_{\text{c},l})+z|^2}{\sigma_z^2}\Big)\Big\}
\end{aligned} \tag{3.10}
$$

$$
\begin{aligned}
R_{\text{CC,sum}}^{\text{SIC-free}} \approx \log|\mathcal{X}_{\text{c}}| + \min_{k\in\{1,\ldots,K\}} \Bigg(& -\frac{1}{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k|}\log\sum_{l=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H[\mathbf{p}_{\text{c}}\ \mathbf{p}_{\text{p},k}](\mathbf{s}_{k,m}-\mathbf{s}_{k,l})|^2}{2\sigma_z^2}\Big) \\
& + \frac{1}{|\mathcal{X}_k|}\sum_{m=1}^{|\mathcal{X}_k|}\log\sum_{l=1}^{|\mathcal{X}_k|}\exp\Big(-\frac{|\mathbf{h}_k^H\mathbf{p}_{\text{p},k}(s_{k,m}-s_{k,l})|^2}{2\sigma_z^2}\Big)\Bigg) \\
& + \sum_{k'=1}^{K}\log|\mathcal{X}_{k'}| - \frac{1}{|\mathcal{X}_{\text{c}}\times\mathcal{X}_{k'}|}\sum_{m=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k'|}\log\sum_{l=1}^{|\mathcal{X}_{\text{c}}\times\mathcal{X}_k'|}\exp\Big(-\frac{|\mathbf{h}_{k'}^H[\mathbf{p}_{\text{c}}\ \mathbf{p}_{\text{p},k'}](\mathbf{s}_{k',m}-\mathbf{s}_{k',l})|^2}{2\sigma_z^2}\Big) \\
& + \frac{1}{|\mathcal{X}_{\text{c}}|}\sum_{m=1}^{|\mathcal{X}_{\text{c}}|}\log\sum_{l=1}^{|\mathcal{X}_{\text{c}}|}\exp\Big(-\frac{|\mathbf{h}_{k'}^H\mathbf{p}_{\text{c}}(s_{\text{c},m}-s_{\text{c},l})|^2}{2\sigma_z^2}\Big)
\end{aligned} \tag{3.11}
$$



(a) RSMA with SIC (b) RSMA without SIC

Figure 3.1: Comparison of CC sum-rate for RSMA achieved by different methods of searching for $t^\star$ with $N_\mathrm{T} = 4$, $K = 2$.

$t^\star$ is obtained by searching for the point in $[0, 1]$ that maximizes either $R^{\mathrm{SIC}}_{\mathrm{CC,sum}}$ or $R^{\text{SIC-free}}_{\mathrm{CC,sum}}$, for scenarios where users apply or not apply SIC respectively.

Remark 3.1 (Phase shift and search space)**.** *Because of its circular symmetry, a Gaussian distributed signal remains statistically unchanged under arbitrary phase shifts, optimizing the power of symbol streams is sufficient when a Gaussian input is assumed. However, a finite constellation generally changes under a phase shift. After superposition, the sum constellation, i.e., the effective constellation of the sum of signals observed by each user, depends on the phase shifts applied to the component signals. Different sum constellations lead to different performance. Hence, initially, both the power and the phase shift applied to each stream need to be optimized in order to maximize the performance under finite constellations. Figure 3.1.2 and 3.1.2 evaluate the CC sum-rate achieved by two searching strategies:*

1. *To exhaustively search for $t^\star$ and the optimal phase shift between the common and private streams that optimize (3.8) or (3.10), depending on the availability of SIC.*

2. *To search among only a small number of samples in $[0, 1]$ for one that maximizes (3.9) or (3.11), depending on the availability of SIC, without optimizing the phase shifts applied to the streams.*

It can be observed that searching $t^\star$ among $\{0, 0.2, 0.4, ..., 1\}$ without any control

on the phase shifts leads to a near-optimal performance. We conclude that, given enough freedom in designing t*, optimizing the phase provides very little performance improvement, and hence we neglect the phase shifts in the proposed design. Also, taking (3.9) and (3.11) as objective functions, instead of (3.8) and (3.10), not only significantly reduces the computational complexity of determining* $t^\star$*, since there is no need to compute the expectations in (3.8) and (3.10), but also poses very little performance loss. As a consequence, the second strategy with the search space* $\{0, 0.2, 0.4, ..., 1\}$ *was applied to generate numerical results in Section. 3.3.*

3.2 Transceiver designs for RSMA

This section introduces potential transceiver designs for RSMA employing a bit-interleaved coded modulation (BICM) approach.

3.2.1 Transmitter design

Figure 3.2 depicts the proposed RSMA transmitter architecture, where FEC, Π and M represent forward error correction coding, interleaving and mapping (modulator) correspondingly. Message splitting and combining in RSMA can be achieved by de-multiplexing and multiplexing the information bits/uncoded bits. The common message, which is generated by multiplexing $W_{\mathrm{c},1}, W_{\mathrm{c},2}, ..., W_{\mathrm{c},K}$, and the private messages, $W_{\mathrm{p},1}, W_{\mathrm{p},2}, ..., W_{\mathrm{p},K}$, are encoded, interleaved and mapped into symbols separately to generate blocks of common and private stream symbols, $\{s_{\mathrm{c}}, s_1, ..., s_K\}$. The precoders then mix the common and private streams before transmission.

3.2.2 Receiver designs

We introduce five RSMA receiver designs, offering flexibility in terms of complexity and performance in practical applications. The receivers are classified into SIC and SIC-free type receivers, and hence they also represent specific implementations that push practical systems towards the information theoretic limits characterized by (3.8) and (3.10) respectively.

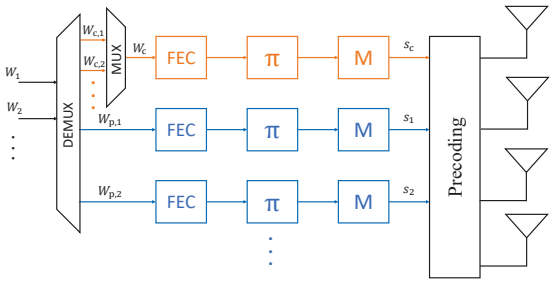


Figure 3.2: Transmitter architecture.

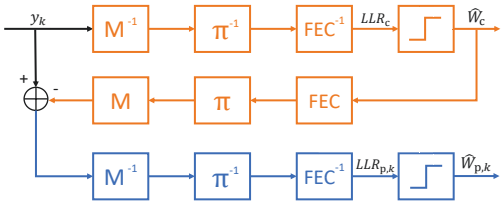


Figure 3.3: Hard CWIC.

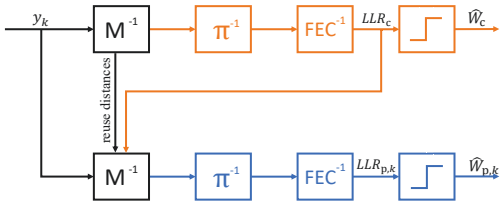


Figure 3.4: Soft CWIC 1.

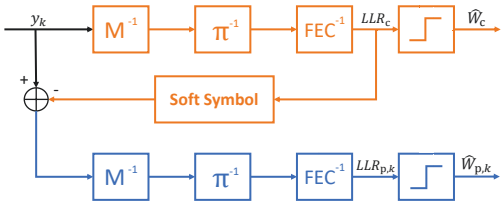


Figure 3.5: Soft CWIC 2.

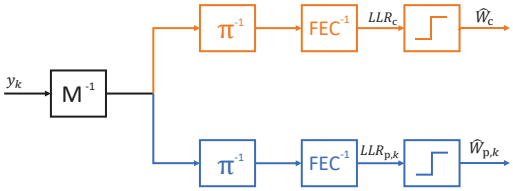


Figure 3.6: Joint de-mapper.

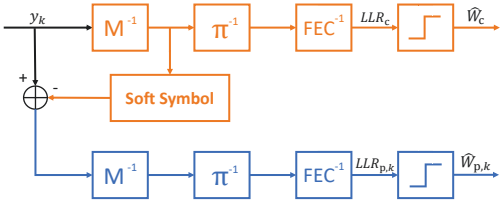


Figure 3.7: Soft SLIC.

Hard CWIC

Based on the definition of SIC from information theory, hard codeword-level SIC (CWIC) [81, 82] is the most common way of implementing SIC in practice and is also the most widely studied RSMA receiver in the literature, for example, in [49, 31, 50, 51, 52]. The structure of a hard CWIC is depicted in Figure 3.3, where M^{-1}, Π^{-1} and FEC^{-1} represent de-mapper (demodulator), de-interleaver and decoder respectively.[4] The receiver first de-maps the common stream symbols from y_k by treating the private symbols as Gaussian noise and obtains the log-likelihood ratios (LLRs) of the coded common stream bits. Next, the LLRs are de-interleaved and decoded to obtain the LLRs of the uncoded common stream bits. Then, hard decisions are applied to the LLRs to obtain estimates on W_c, denoted by $\tilde{W}_\mathrm{c}$. This

[4]In the receiver diagrams, M^{-1} in orange or blue represents marginal de-mapper for the common or private stream respectively, i.e., it marginally de-maps one stream by considering the rest of the signals as Gaussian noise. M^{-1} in black represents a joint de-mapper which jointly de-maps all the received signals and obtains soft-estimates for the coded bits carried by either the common stream or the private stream or both.

is followed by re-encoding, re-interleaving and re-mapping $\tilde{W}_{\mathrm{c}}$ to obtain estimates on a block of common stream symbols, $\tilde{s}_{\mathrm{c}}$. After that,, a block of $\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\tilde{s}_{\mathrm{c}}$ is subtracted from the received signal block to cancel the effect of the common stream. Finally, the remaining signals, $y_k - \mathbf{h}^H\mathbf{p}_{\mathrm{c}}\tilde{s}_{\mathrm{c}}$, pass through de-mapping, de-interleaving and decoding to obtain the uncoded bits carried by the private stream.

Hard CWIC has a known issue with error propagation, i.e., if the decoding of the common stream fails, errors occur in the cancellation process and severely harm the decoding for the private stream. To mitigate this issue, two extended CWIC designs exploiting, instead of the hard information, the soft information from the common stream decoding to perform cancellation, are introduced as follows.

Soft CWIC 1

The principle of soft CWIC 1 is depicted in Figure 3.4. The receiver first jointly de-maps the common stream and the desired private stream but only computes the common stream coded-bit-LLRs.[5] As with hard CWIC, these LLRs are de-interleaved and decoded to obtain uncoded-bit-LLRs of the common stream. Assuming the decoder also updates the coded-bit-LLRs,[6] the probability distribution of a block of common stream symbols, $P(s_{\mathrm{c}})$, is then computed. Note that this estimate on $P(s_{\mathrm{c}})$ is more accurate than the one obtained from the de-mapper, as knowledge of the error correcting code is utilized. To de-map the private stream symbols, the receiver performs another joint de-mapping but also uses $P(s_{\mathrm{c}})$ as prior information. The computation results on distances from the previous de-mapping can be reused at this stage, hence the computational complexity of the second de-mapping is negligible. Following the de-mapping, the private stream is then de-interleaved and decoded. Readers are referred to [80] for more details.

[5]Here, joint de-mapping with marginal outputs is used instead of marginal de-mapping for two reasons: 1) Joint de-mapping is the optimum way of detecting signals in the presence of interference, while marginal de-mapping treats the discretely distributed interference as Gaussian noise and hence is sub-optimal; 2) Applying prior knowledge of the interference distribution in the LLR computation, which in our context is required for de-mapping the private stream, does not introduce additional complexity to joint de-mapping [80].

[6]Many decoder designs used in practice are soft-input-soft-out, e.g., the belief propagation decoders.

Soft CWIC 2

With soft CWIC 1, joint de-mapping for the common and private symbols potentially poses high computational complexity. To reduce this computational burden, soft CWIC 2 is introduced and depicted in Figure 3.5 [86]. In this design, the receiver first de-maps the common stream by treating the private stream symbols as Gaussian noise. $P(s_c)$ is obtained after de-interleaving and decoding in a similar way as with soft CWIC 1. Unlike soft CWIC 1, $P(s_c)$ is used to compute the mean (also called the soft symbol estimates), $\tilde{s}_c$, and the variance, $\sigma^2_{s_c}$, of the common symbols. Subsequently, de-mapping for the private symbols is performed by treating the common stream symbols plus noise as an effective noise following the distribution $\mathcal{CN}(\mathbf{h}_k^H \mathbf{p}_c \tilde{s}_c, \sigma_z^2 + \sigma^2_{s_c})$. To do so, the receiver removes $\mathbf{h}_k^H \mathbf{p}_c \tilde{s}_c$ from y_k then calculates the LLRs by taking $\sigma_z^2 + \sigma^2_{s_c}$ as the noise variance. In comparison to soft CWIC 1, soft CWIC 2 avoids any joint de-mapping and utilizes $P(s_c)$ in a simpler way and hence is more computationally efficient.

Joint De-mapper

Figure 3.6 depicts a joint de-mapper [80] based receiver design for RSMA. The receiver first jointly de-maps symbols from common and private streams. The resulting LLRs are then de-multiplexed and separately de-interleaved into coded-bit-LLRs of common and private streams. Finally, these LLRs are fed into decoders to obtain uncoded bits of the common and private streams.

Soft SLIC

To avoid the potentially high complexity brought by the joint de-mapper, a soft symbol level interference cancellation (SLIC) design [87] is introduced as a substitute, as depicted in Figure 3.7. The process is similar to the soft CWIC 2 design mentioned above. The only difference is that the prior knowledge of the common stream symbols is not from the decoder, but from the de-mapper for the common stream. Another way to interpret soft SLIC is to view several steps, $M^{-1} \rightarrow$ Soft Symbol $\rightarrow$ Cancellation $\rightarrow M^{-1}$, as a simplified joint de-mapping process. Indeed, it de-map the received signals into soft estimates on the common and private coded bits. Instead of jointly de-mapping the two streams, using two marginal de-mapping blocks reduces the computational burden, albeit at the cost of accuracy.

3.2.3 Comments on different receivers

Complexity

We summarize the complexity of different receiver designs under three criteria, namely computational complexity, additional buffer size and additional delay. We recognize that the joint de-mapper leads to the simplest structure and the shortest delay among others, and hence measure the buffer size and delay in terms of the extra requirement in comparison with the joint de-mapper. The three criteria are explained as follows:

- Computational complexity: The majority of the computation burden comes from de-mapping and decoding. For comparative purposes, we focus on the de-mapping complexity and neglect the computational burden of decoding, since the decoding process is identical for all the proposed receivers, i.e., all the receivers decode the common stream and the desired private streams.
- Additional buffer size: This refers to the additional buffers required for cancellation, compared with the joint de-mapper. For CWIC receivers, the decoded common stream message needs to be stored before performing cancellation. The required buffer size depends on the number of coded bits during one transmission and the type of information to be stored, i.e., hard or soft.
- Additional delay: This refers to extra time required to perform cancellation in comparison to the joint de-mapper. Apart from all the processes of joint de-mapper, CWIC receivers need extra time to perform re-interleaving and re-modulation (or more generally, converting bit level information into symbol-level information) before cancellation. Furthermore, the time consumption on joint de-mapping and soft SLIC can be shortened by parallelized decoding of common and private messages, but this is not valid for CWIC receivers due to the existence of the cancellation process.

Let $\mathcal{X}_\mathrm{c}$ and $\mathcal{X}_\mathrm{p}$ be the constellations for common and private streams, F be the number of bits required to store soft information, i.e., the LLRs, and B be the blocklength in terms of symbols. A quantitative evaluation of receiver complexity is given in Table 3.1. It can be concluded that the soft CWIC 1 and the soft SLIC are of the highest and lowest complexity, respectively.

Table 3.1: Classification and complexity of different receiver designs.

	Receiver	Computational complexity	Additional buffer size	Additional delay
SIC Type	Hard CWIC	$\mathcal{O}\Big(\lvert\mathcal{X}_{\mathrm{c}}\rvert + \lvert\mathcal{X}_{\mathrm{p}}\rvert\Big)$	$B \log \lvert\mathcal{X}_{\mathrm{c}}\rvert$	Interleaving & Mapping
	Soft CWIC 1	$\mathcal{O}\Big(\lvert\mathcal{X}_{\mathrm{c}} \times \mathcal{X}_{\mathrm{p}}\rvert\Big)$	$FB \log \lvert\mathcal{X}_{\mathrm{c}}\rvert$	Interleaving & Mapping
	Soft CWIC 2	$\mathcal{O}\Big(\lvert\mathcal{X}_{\mathrm{c}}\rvert + \lvert\mathcal{X}_{\mathrm{p}}\rvert\Big)$	$FB \log \lvert\mathcal{X}_{\mathrm{c}}\rvert$	Interleaving & Mapping
SIC-free Type	Joint De-mapper	$\mathcal{O}\Big(\lvert\mathcal{X}_{\mathrm{c}} \times \mathcal{X}_{\mathrm{p}}\rvert\Big)$	0	-
	Soft SLIC	$\mathcal{O}\Big(\lvert\mathcal{X}_{\mathrm{c}}\rvert + \lvert\mathcal{X}_{\mathrm{p}}\rvert\Big)$	0	-

SIC and SIC-free receivers

The above-mentioned receiver designs can be classified into two groups according to their underlying principle. In particular, hard CWIC, soft CWIC 1 and soft CWIC 2 utilize the common stream decoder outputs in the process of decoding the private stream and hence are classified as SIC-type receivers. Joint de-mapping and soft SLIC decode the private stream without using the output of the common stream decoder by treating the common stream as noise. Hence, they are classified as SIC-free-type receivers. This classification is also illustrated in Table3.1.

Jointly de-mapping the desired signal and interference is equivalent to treating interference as noise. Two supporting evidences are given:

1. From an information theoretic perspective, approaches to handling interference can be classified into: 1) decoding interference and 2) treating interference as noise. From Figure 3.6, it is easy to see that joint de-mapping does not perform SIC or joint decoding, hence it does not belong to the category of decoding interference.

2. From a detection problem perspective, optimally detecting a discrete signal in the presence of discrete interference and Gaussian noise is equivalent to jointly de-mapping the desired signal and interference. A proof is given in Appendix B.2.

A common confusion exists between two ideas: 1) treating interference as noise and 2) treating interference as Gaussian noise. With finite constellations, interference is not Gaussian, and hence the sum of interference and noise is not Gaussian. The optimal way of treating interference as noise is to use its exact probability distribution when de-mapping as shown in Appendix B.2. This leads to joint de-mapping. However, in practice, it is computationally more efficient to pretend that the interference is Gaussian, and hence the sum of interference and noise is also Gaussian. This leads to receivers treating the sum of interference and noise as an effective Gaussian noise (such as receivers for downlink MU-MIMO), which could result in a huge performance loss in comparison to joint de-mapping because the wrong probability distribution for the interference is used.[7] For RSMA, jointly de-mapping the common and the private streams is equivalent to treating the common stream as noise when decoding the private stream.

[7] An information-theoretic study of this performance loss will be the focus of Chapter 5.

In light of the classification of SIC-type and SIC-free-type receivers, it is evident that, depending on the type of receivers in use, the system should be designed to maximize the corresponding CC sum-rate (with or without SIC) in order to optimize the performance.

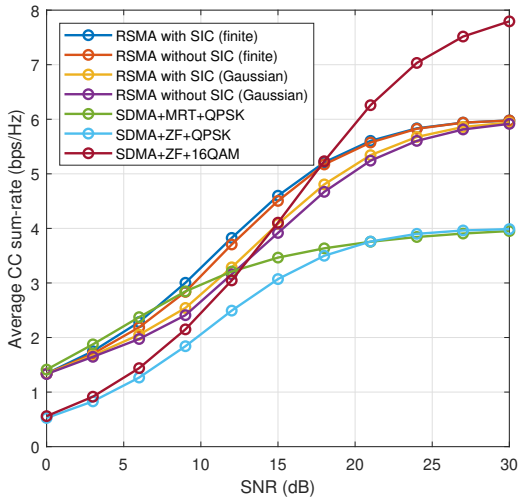
3.3 Numerical results

We consider the one-ring scattering model (See Appendix A for details). For simplicity, we assume $\theta_1 = \theta_2 = ... = \theta_K = \theta$ and $\Delta_1 = \Delta_2 = ... = \Delta_K = \Delta$, which implies that users with similar channel statistics are served at the same time. With this setting, we intend to model scenarios with a certain user density. A small Δ indicates a high-density demand area. Throughout this section, we assume $\theta = \pi/3$, $N_{\mathrm{T}} = 4$ for two-user cases, and $N_{\mathrm{T}} = 6$ for three-user cases.

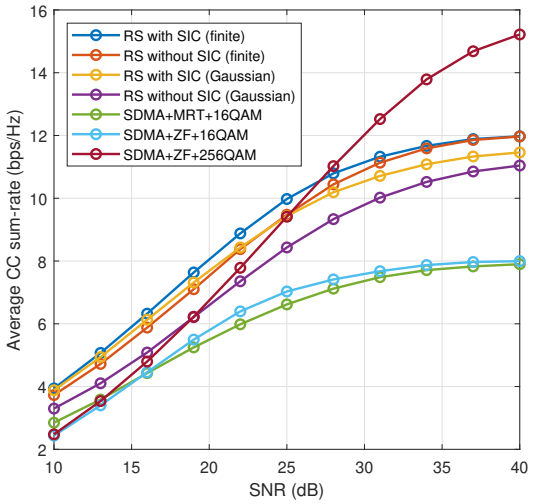
3.3.1 CC sum-rate comparison

For rate evaluation, we consider a block-fading model where the channel remains static and known within each codeword. For each realization, we design the precoder based on the instantaneous channel state and compute the resulting achievable rate. The results are obtained by averaging these rates over a sufficiently large number of independent channel realizations. This approach is particularly suited for slow-fading environments where the channel is approximately constant over a single codeword but varies across different transmission blocks.

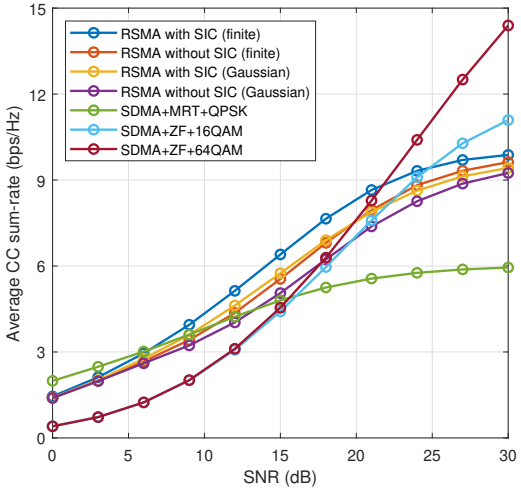
We verify that the proposed low-complexity precoder design that optimizes (3.9) and (3.11) actually leads to a higher CC sum-rate than other schemes. In Figure 3.8, we compare the average CC sum-rate of three schemes, namely RSMA with precoder optimizing the sum-rate assuming Gaussian input signals (with "Gaussian" in the legend), RSMA with precoder optimizing the approximation of CC sum-rate (with "finite" in the legend) and SDMA with either ZF or MRT precoding under different scenarios. Constellations used for different schemes are given in the legend and caption of the figure. The constellations were chosen such that different schemes require similar decoding complexity. It can be observed that RSMA with precoders optimizing the sum-rate does not essentially optimize the CC sum-rate and often underperforms SDMA. Especially when an SIC-free receiver is assumed, RSMA by



(a) $K = 2$, $\Delta = \pi/18$, QPSK for all the streams in RSMA.



(b) $K = 2$, $\Delta = \pi/18$, 16QAM for all the streams in RSMA.



(c) $K = 3$, $\Delta = \pi/9$, 16QAM for common streams, QPSK for private streams.

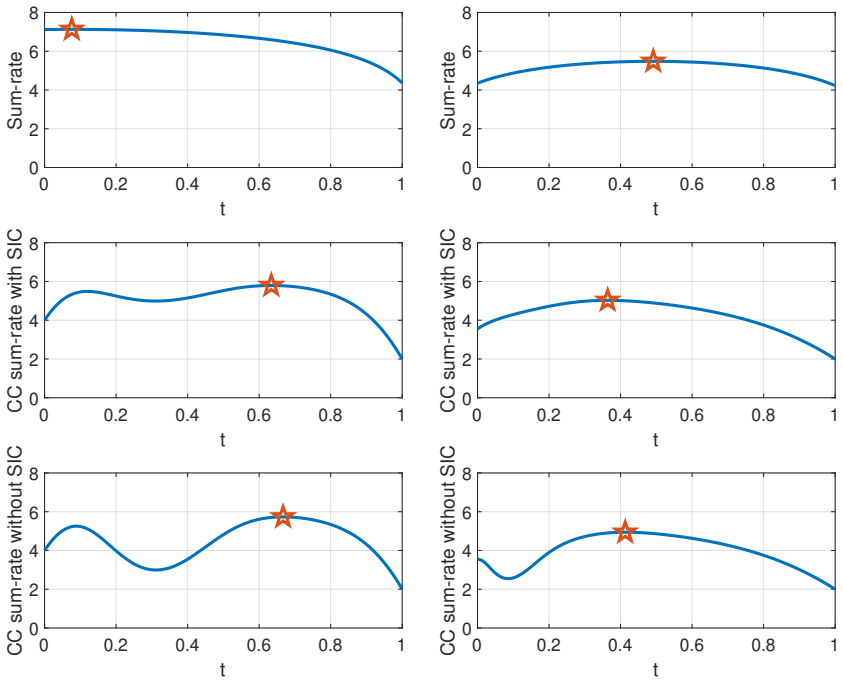
Figure 3.8: CC sum-rate comparison for different schemes.

optimizing the sum-rate always underperforms SDMA. On the other hand, the designs that optimize (3.9) and (3.11) lead to nearly the best performance among others at low and medium SNRs. At a very low SNR regime, SDMA with MRT precoding slightly outperforms RSMA. However, this performance requires more complicated receivers in practice. The reason is that CC sum-rate provides the performance upper-bound and, to approach this, the receiver should at least jointly de-maps the desired signal and interference rather than simply treating interference as Gaussian noise. To do so, advanced receivers and extra knowledge on interference are necessary. At high SNR regime, SDMA with high-order constellations outperforms RSMA because the latter saturates at a lower rate. For example, RSMA with

QPSK used for both common and private streams transmits at most 6 bits/symbol, while SDMA with two 16QAM streams transmits at most 8 bits/symbol.

The above results lead to two important implications. First, the potential of RSMA can be better discovered when finite constellations are considered during the system design process. Second, the superiority of RSMA over SDMA remains even under constraints on similar decoding complexity.

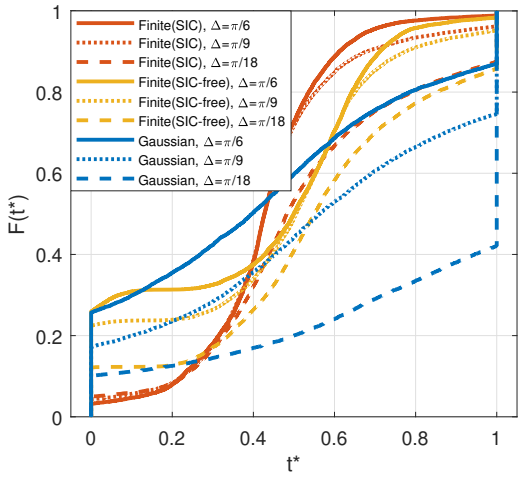
3.3.2 Power allocation behavior



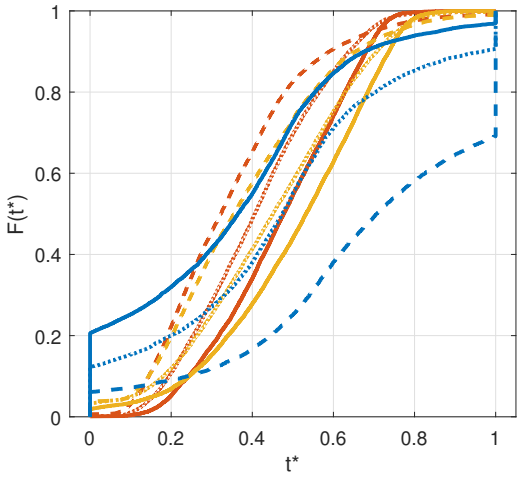
(a) Channel 1. (b) Channel 2.

Figure 3.9: Sum-rate/CC sum-rate as functions of t with $t^\star$ labeled by ★, $K = 2$.

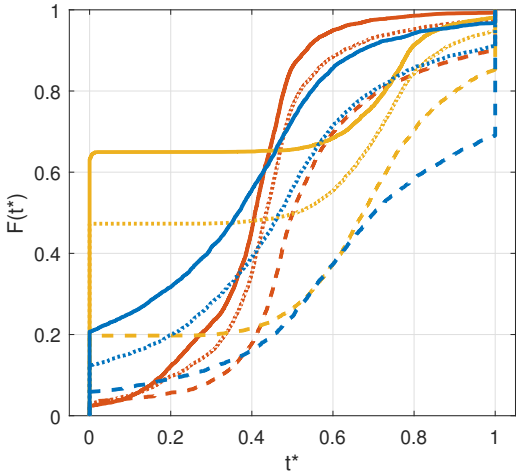
In the following, we show that taking sum-rate or CC sum-rate as the objective function leads to distinct choices on the power allocation factor, t. We first consider two channel realizations given in Appendix B.3. Figure 3.9 shows, under the two realizations, the sum-rate, approximated CC sum-rate with SIC receivers and approximated CC sum-rate with SIC-free receivers, i.e., (3.3), (3.9) and (3.11) respectively, as a function of t, where $t^\star$ for (3.9) and (3.11) are obtained from exhaustive search. It can be observed that $t^\star$ under different objective functions is indeed different. Also, it is can be illustrated that (3.3) is convex over t while (3.9) and (3.11) are not.



(a) SNR=5dB, QPSK assumed for CC sum-rates.



(b) SNR=10dB, QPSK assumed for CC sum-rates.



(c) SNR=10dB, 16QAM assumed for CC sum-rates.

Figure 3.10: CDF of $t^\star$ under different objective functions with $K = 2$.

Figure 3.10 depicts the CDF of $t^\star$, denoted by $F(t^\star)$, under different objective functions and channel conditions. It shows that the statistics of $t^\star$ can be very different under different objective functions. One interesting observation can be made from the variation of CDF with different Δ. As Δ decreases, the system with Gaussian signalling always allocates more power to the common stream, as this leads to less interference. However, with finite constellations, this is not always the case. For example, in Figure 3.10(b), the system allocates less power to the common stream with a decreasing Δ. This is because the rate of finite constellations is upper-bounded and saturates after a certain SNR. With a smaller Δ, the system can achieve the saturating point for the common stream with a smaller amount of power, and it is more beneficial to allocate the rest of the power to the private streams as they have

not saturated yet. In Figure 3.10(c), the system again allocates more power to the common streams as Δ drops because 16QAM requires more power than QPSK to saturate in CC rates.

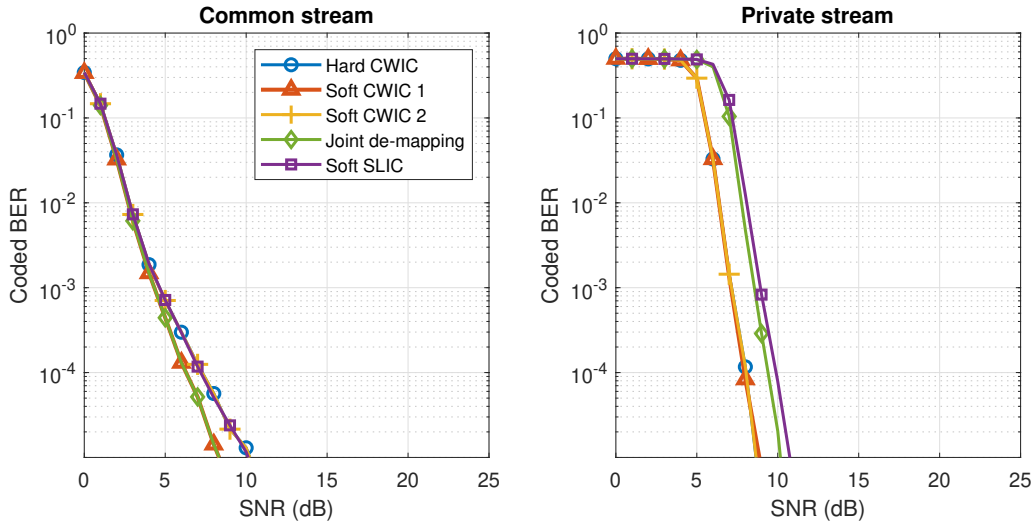
3.3.3 Link-level simulation

Link-level simulation was performed to provide a more comprehensive evaluation of different precoders and receivers. In the following, fast fading and block fading channels[8] with perfect CSIT are considered. For the error correction code, polar code was chosen for its good error-correcting capability. Belief propagation decoding was used as the decoding technique to obtain soft outputs. All the results were obtained by averaging the Bit Error Rate (BER) performance of 50000 blocks. Symbol block length is 512 for fast fading channels and 4096 for block fading channels. When block fading is considered, it is assumed that the channel is coherent for 32 symbols. Max-log approximation was used for de-mapping as it is common in real applications. The target BER was 10^{-4}.

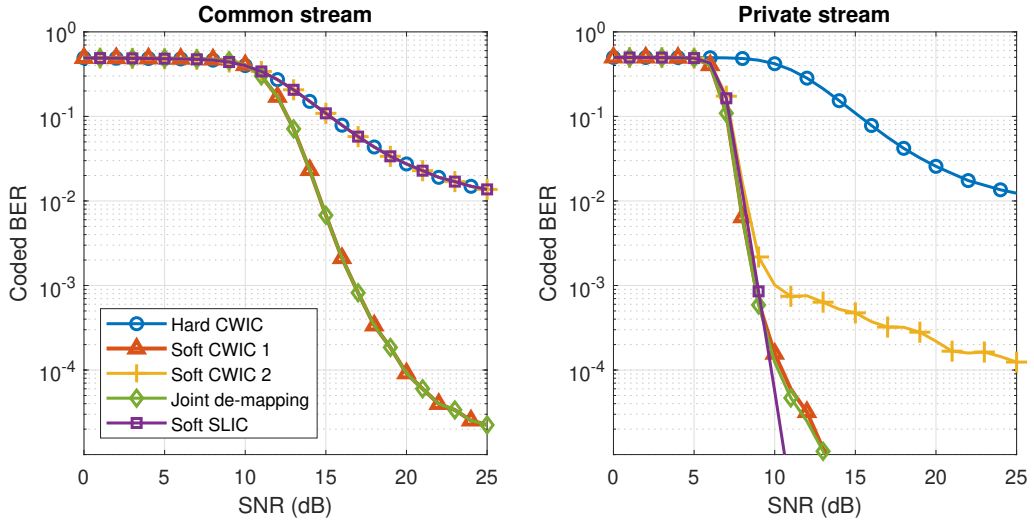
We first make a comparison of different receiver designs mentioned in Section 3.2 based on Figure 3.11(a) and Figure 3.11(b). Here, the precoders were designed to maximize the sum-rate. Details on the modulation and coding schemes (MCS) in use are shown in the caption of the figures in the format of "modulation/code rate". We first set the common stream code rate to a medium value such that common and private streams reach 10^{-4} BER with similar SNR. This simulates cases where a good operating point is found. In Figure 3.11(a), it can be concluded that the performance of different receivers follows the order: soft CWIC 1 > soft CWIC 2 > hard CWIC > joint de-mapping > soft SLIC.

We now consider cases where the common stream rate is not properly set, and hence errors may exist when decoding the common stream. In these cases, error propagation effects with different receivers become important. Figure 3.11(b) shows the BER performance with different receiver designs when a high error rate occurs in the common stream. The code rate for the common stream was set to be high, such that a high BER (around 0.4) occurs when the SNR is 10 dB. When hard CWIC is applied, it can be observed that a similar BER occurs in the private stream due to the error propagation. In contrast, with other receivers, the BER degradation caused by error propagation is very weak or not visible.

[8]Fast fading refers to i.i.d. channel realizations for each symbol. Block fading refers to channel realizations that are i.i.d. from block to block and identical within the same block.



(a) MCS: QPSK/0.54 for common stream and QPSK/0.3 for private streams.



(b) MCS: QPSK/0.6 (hard CWIC, soft CWIC 2 and soft SLIC) or QPSK/0.74 (soft CWIC 1 and joint de-mapping) for common stream and QPSK/0.3 for private streams.

Figure 3.11: BER performance of different RSMA receivers with $K = 2$ and $\Delta = \pi/9$.

Next, we compare the performance of the sum-rate oriented precoder design and the CC sum-rate oriented precoder designs through LLS. For clarity, only soft CWIC 1 and joint de-mapping are considered in the following as representatives of SIC-type receivers and SIC-free-type receivers. RSMA schemes with power allocation methods considering sum-rate or CC sum-rate, as proposed in Section III, are indicated by "(Gaussian)" and "(finite)" in the legend. We also include several SDMA-based schemes as benchmarks:

- SDMA+ZF+SD+QPSK/16QAM/64QAM/256QAM: The precoders are de-

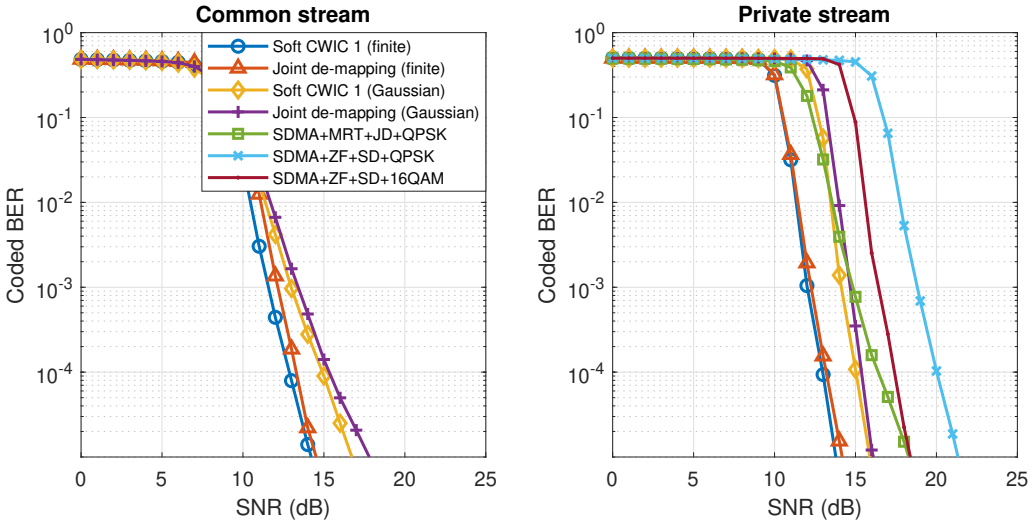


Figure 3.12: Fast fading channel with $K = 2$, $\Delta = \pi/18$ and throughput $= 3$ bits/symbol. Modulation schemes for RSMA: QPSK for both common and private streams.

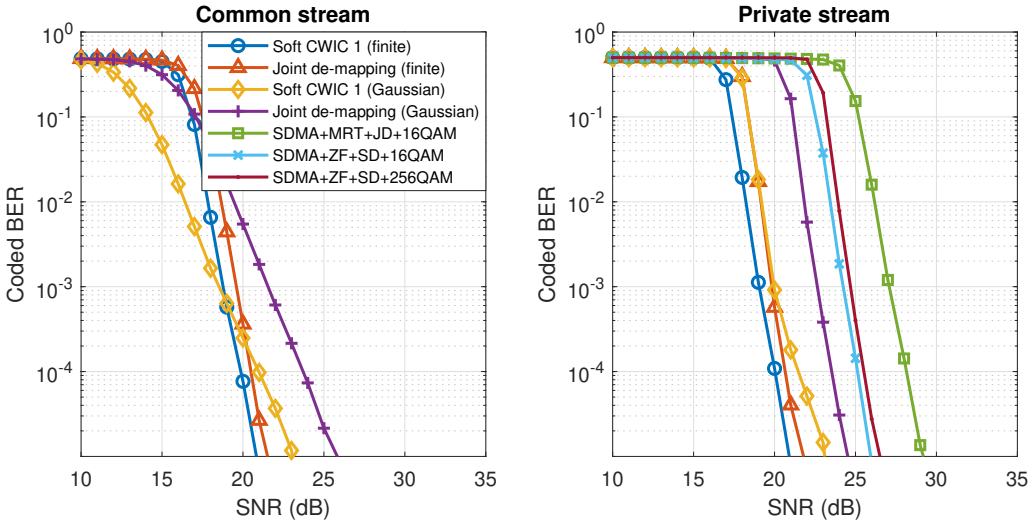


Figure 3.13: Fast fading channel with $K = 2$, $\Delta = \pi/18$ and throughput $= 6$ bits/symbol. Modulation schemes for RSMA: 16QAM for both common and private streams.

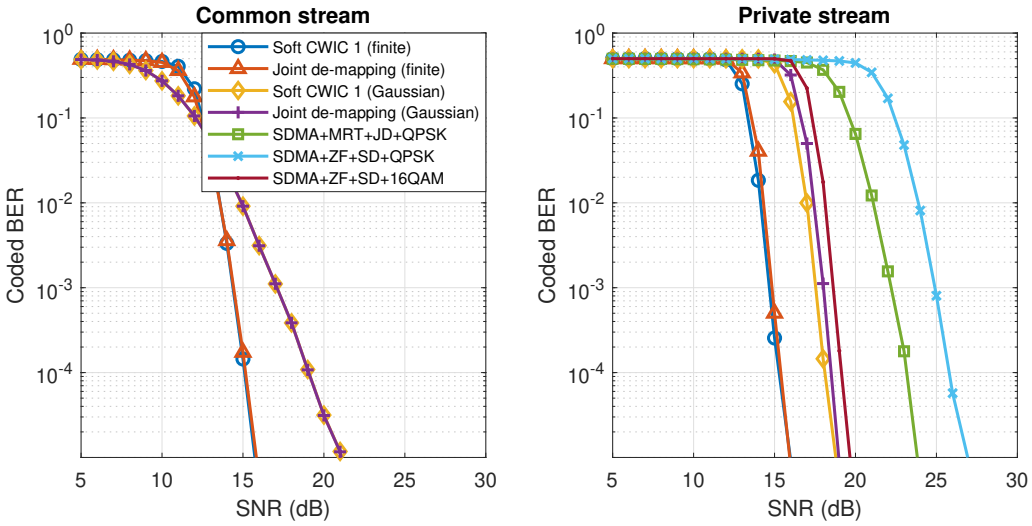


Figure 3.14: Block fading channel with $K = 2$, $\Delta = \pi/18$ and throughput = 3.6 bits/symbol. Modulation schemes for RSMA: QPSK for both common and private streams.

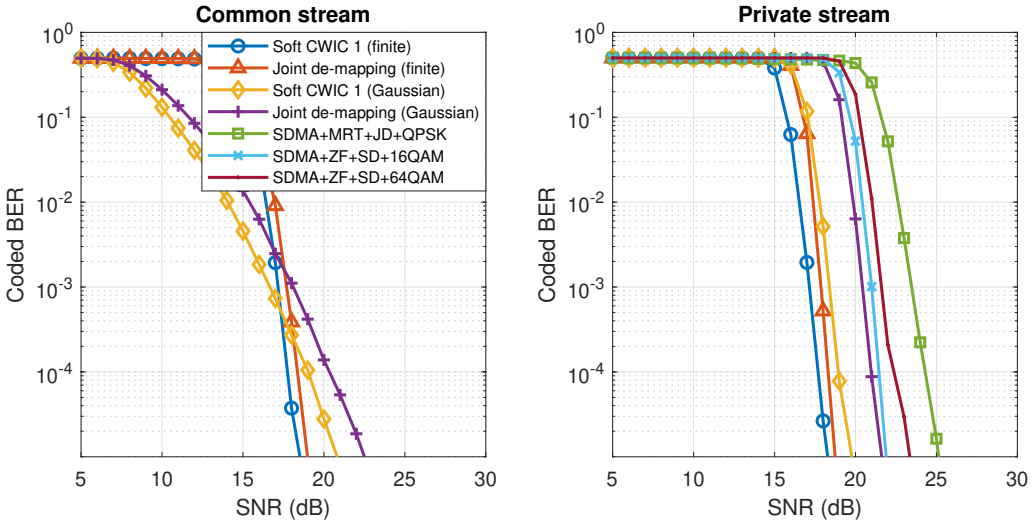


Figure 3.15: Block fading channel with $K = 3$, $\Delta = \pi/9$ and throughput = 5.52 bits/symbol. Modulation schemes for RSMA: 16QAM for the common stream and QPSK for the private streams.

signed using ZF with uniform power allocation. The receivers apply a single-user detector (marginally de-map the desired signal). Both streams are transmitted in QSPK, 16QAM or 256QAM signals.

- SDMA+MRT+JD+QPSK/16QAM: The precoders are designed using Maximum Ratio Transmission (MRT) with uniform power allocation. The receivers jointly de-map the desired stream and interference. This receiver design is studied as an advanced receiver for MU-MIMO in 3GPP [88].[9] Both streams are transmitted in QPSK or 16QAM signals.

The code rates used in these simulations are given in Appendix B.4.

The BER performance of RSMA with different precoders and receiver designs, as well as three SDMA benchmarks, is depicted in Figure 3.12-3.15, with maximum throughput and modulation schemes in use for RSMA given in the caption of the figures. In particular, Figure 3.12 and Figure 3.13 examine scenarios with fast fading and two users. Figure 3.14 assumes block fading with two users, while Figure 3.15 assumes block fading with three users. In each figure, different schemes achieve the same throughput as indicated in the captions. For SDMA schemes, the code rates are uniform across users. For RSMA schemes, the code rates for common and private streams are selected such that both streams achieve the target BER at similar SNRs. It can be observed that RSMA schemes generally outperform SDMA schemes. Among different RSMA schemes, the performance depends on the precoder and receiver designs. In particular, it can be observed that the CC sum-rate oriented RSMA (with "(finite)" in the legends) significantly outperforms the sum-rate oriented RSMA (with "(Gaussian)" in the legends), with either SIC-type or SIC-free-type receivers. For example, in Figure 3.12, with either SIC-type or SIC-free-type receivers, CC sum-rate oriented precoder design achieves roughly 2.5 dB gain over sum-rate–oriented designs. This again reveals the important fact that the practical capability of RSMA is not well exploited by the sum-rate oriented designs but is better discovered by the CC sum-rate oriented designs.

Figure 3.16 depicts how the angle spread, Δ, of the scattering ring affects the superiority of RSMA over SDMA. It can be concluded that RSMA offers greater advantages under small-angle spreads.

In Figure 3.17, the trade-offs between performance and complexity with different schemes and receivers are depicted based on the LLS results and Table I. SDMA

[9] As discussed in Section IV C, joint de-mapping is the optimal detector in the presence of finite-constellation interference.

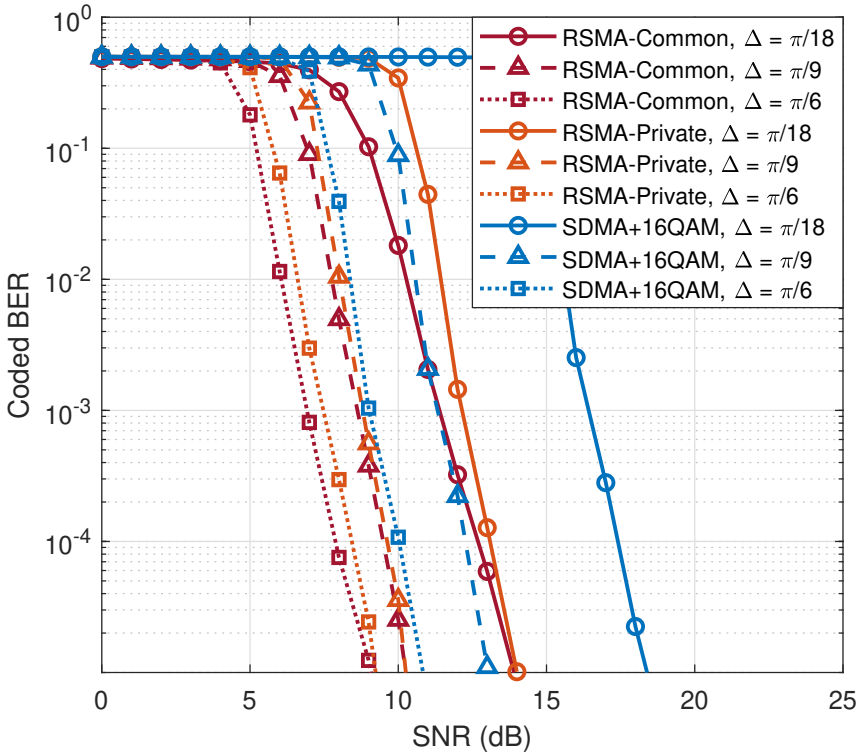


Figure 3.16: Comparison between RSMA and SDMA with $K = 2$, $\Delta = \pi/18$, $\pi/9$ or $\pi/6$. RSMA uses QPSK for both common and private streams. The throughput of different schemes is 3 bits/symbol.

with QPSK constellation was chosen as the reference. SNR reduction was measured in the amount of SNR saved by a scheme when a BER of 10^{-4} is achieved. Relative complexity increase was measured in the additional complex distances to be computed per symbol during de-mapping. Without considering the cost of the buffering memory and delay, CC sum-rate oriented RSMA with soft CWIC 2 leads to almost optimal performance with low complexity. To save buffer and delay, CC-rate-oriented RSMA with soft SLIC and joint de-mapper becomes a promising candidate. We again observe that RSMA outperforms SDMA, even when the decoding complexities are comparable.

3.4 Summary

In this chapter, we considered the practical implementation of RSMA and addressed several problems. We first introduced three low-complexity power allocation designs that respectively optimize sum-rate, CC sum-rate with SIC receiver and CC sum-rate without SIC receiver. We then introduced several transceiver implementations

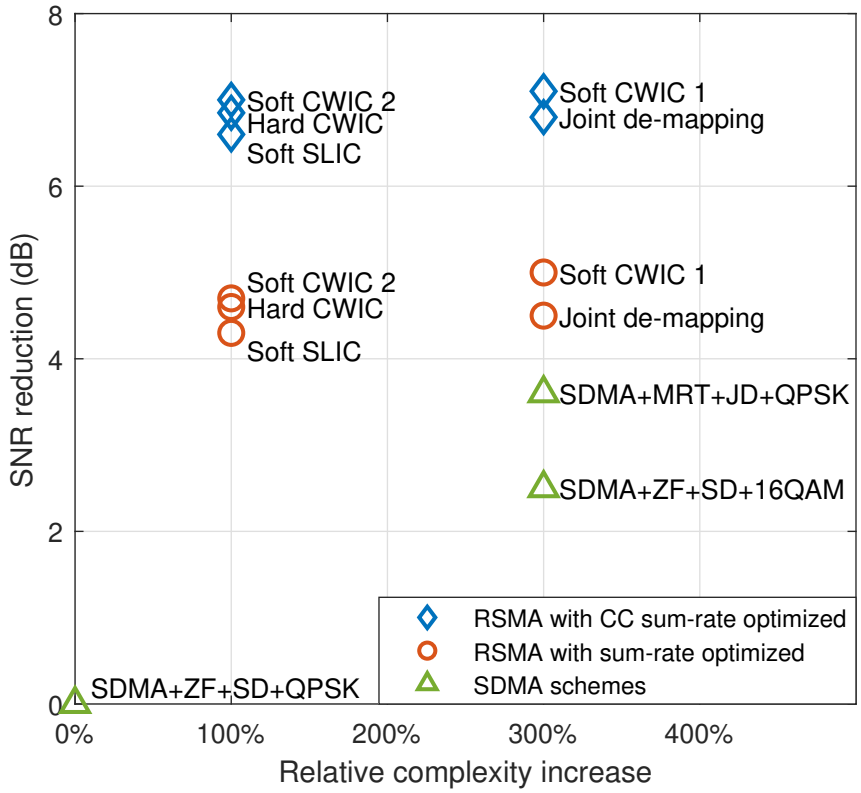


Figure 3.17: SNR-complexity trade-off with different schemes under fast fading with $K = 2$, $\Delta = \pi/18$ and throughput = 3 bits/symbol. Modulation schemes for RSMA: QPSK for all the streams.

that can be applied with RSMA. We noticed that these designs can be classified into either SIC-type receivers or SIC-free-type receivers based on their principle and proposed to choose the power allocation design in accordance with this classification. Through numerical simulations, we reveal the fact that, under practical constellation constraints, sum-rate oriented designs do not always bring the benefit of RSMA, while CC sum-rate oriented designs better exploit the potential of RSMA. It was also observed that, with CC sum-rate oriented precoder design, RSMA often preserves its superiority over SDMA even without SIC receivers. LLS was performed to verify this conclusion. Different practical receiver designs were also compared using LLS. We show that the most common SIC-type receiver design, the hard CWIC, is not promising due to its weak immunity to error propagation. Soft CWICs suffer very little from error propagation and hence provide better performance. Joint demapping and soft SLIC, which incur less complexity on the user side with acceptable performance loss, are also promising candidates.

Chapter 4

Precoder Optimization for SIC-based and SIC-free RSMA

The last chapter demonstrates that RSMA can maintain its superiority without relying on SIC receivers. In practical systems, SIC-free receivers are more attractive than SIC receivers because of their low complexity and latency. This chapter evaluates the theoretical limits of RSMA with and without SIC receivers under finite constellations. We first derive the constellation-constrained rate expressions for RSMA with an arbitrary precoding matrix. We then design algorithms based on a projected subgradient method to optimize the precoders and maximize the weighted sum-rate (WSR) and max-min fairness (MMF) (or minimum rate per user) among users. To apply the proposed optimization algorithms to large-scale systems, such as massive MIMO, one challenge lies in the exponentially increasing computational complexity brought about by the constellation-constrained rate expressions. In light of this, we propose methods to avoid such a computational burden. Numerical results show that, under optimized precoders, SIC-free RSMA leads to minor losses in both weighted sum-rate and max-min fairness in comparison to RSMA with SIC receivers, making it a viable option for future implementations.

The key contributions of this chapter are listed as follows.

1. We provide constellation-constrained rate expressions and their approximation for SIC-based and SIC-free RSMA. Note that the RSMA rate expressions derived in Chapter 3 are based on the assumption that zero-forcing precoders are used for private streams. As a result of this assumption, inter-user interference was removed during the derivations, and the rate expressions were simplified.

The rate expressions given in this chapter do not make any assumption on the precoders and are therefore more general than those in Chapter 3.

2. We propose precoder optimization methods to maximize the WSR and MMF performance for SIC-based and SIC-free RSMA under finite constellation constraints. To enable this, we developed global optimal solutions to common stream rate allocation for both WSR maximization and MMF problems. These solutions hold for not only finite-constellation inputs but also any other input signal distribution, including Gaussian inputs.

3. We observe that the computational complexity of the proposed algorithms grows exponentially as the system scales up in terms of the number of users. In light of this, we propose methods to scale up optimization algorithms to handle large-scale systems, such as massive MIMO, and avoid prohibitive computational cost.

4. We show through simulations that both SIC-based and SIC-free RSMA outperform SDMA under different evaluation metrics and dimensions of the system. Moreover, SIC-free RSMA often preserves most of the superiority of SIC-based RSMA. The results also show that SIC-based and SIC-free RSMA require very different precoders, despite the fact that they achieve similar performance in the end. This indicates the need for dedicated precoder designs for SIC-free RSMA. Therefore, we extend the conclusion in Chapter 3 from low-complexity precoders to fully optimized precoders, from the sum-rate objective to the WSR and MMF objectives, and from small-scale to large-scale systems.

Organization: The rest of this chapter is organized as follows. Section 4.1 provides achievable rate expressions for RSMA under finite constellations. In Section 4.2, precoder optimization algorithms are proposed for WSR and MMF optimization. Section 4.3 addresses optimization in large-scale systems. Section 4.4 presents and analyzes numerical results. Finally, Section 4.5 summarizes this chapter.

Unless otherwise stated, the system model presented in Section 2.1 is assumed throughout this chapter.

4.1 Achievable rate of RSMA under finite constellations

Chapter 3 extends the system model of RSMA in the literature by introducing the possibility of SIC-free receivers. Figure 4.1 is a schematic of one-layer RSMA involving either SIC or SIC-free receivers.

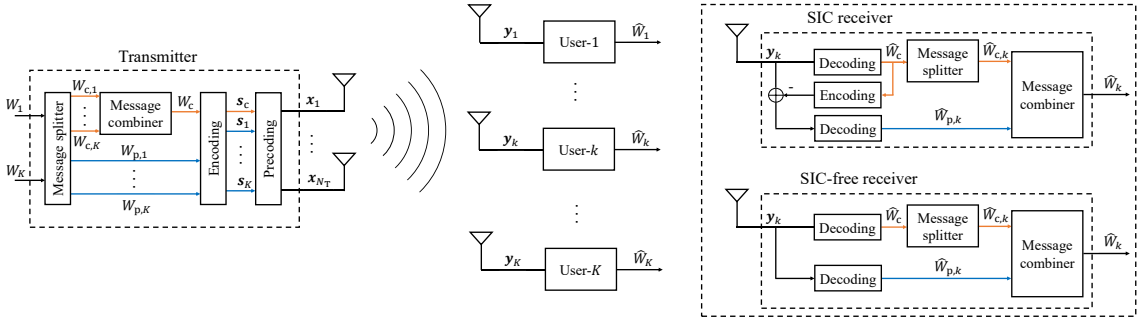


Figure 4.1: Single-layer RSMA with SIC and SIC-free receivers.

The procedures in the transmitter and SIC receivers are described in Section 2.1. For theoretical analysis, SIC-free receivers can be characterized by the following steps: The receiver decodes W_{c} from y_k by treating $\{\mathbf{s}_{\mathrm{p},1}, ..., \mathbf{s}_{\mathrm{p},K}\}$ as noise, and also decodes $W_{\mathrm{p},k}$ by treating $\mathbf{s}_{\mathrm{c}}$ and $\mathbf{s}_{\mathrm{p},k'}$, $\forall k' \in \mathcal{K}/k$, as noise. After this, the receiver extracts $\widehat{W}_{\mathrm{c},k}$ from $\widehat{W}_{\mathrm{c}}$ and combines it with $\widehat{W}_{\mathrm{p},k}$ to reconstruct $\widehat{W}_k$. The main difference from the SIC implementation is that $\widehat{\mathbf{s}}_{\mathrm{c}}$ remains in the received signal when the receiver decodes $W_{\mathrm{p},k}$ due to the absence of the cancellation step.

For RSMA with finite-alphabet (constellation) inputs, the achievable rate expressions of the common stream at user-k, denoted by $R_{\mathrm{c},k}$, and the private stream intended for user-k, denoted by $R_{\mathrm{p},k}^{\mathrm{SIC}}$ and $R_{\mathrm{p},k}^{\text{SIC-free}}$ under SIC and SIC-free receivers respectively, are given in (4.4)-(4.6), where $\mathcal{X}_{\mathrm{c}}$ is the constellation set used for the common stream, $\mathcal{X}_{\mathrm{p},k}$ is the constellation set used for the private stream intended for the k-th user, $\mathcal{X}_{\mathcal{M}} \triangleq \prod_{k\in\mathcal{M}} \mathcal{X}_{\mathrm{p},k}$ for certain set $\mathcal{M}$, $\mathbf{d}_{m,l}(\mathcal{X}) = \mathbf{s}_m - \mathbf{s}_l$, and $\mathbf{s}_m$ denotes the l-th element of $\mathcal{X}$ in a certain order. Their approximations are given in (4.7)-(4.9).

The derivation of (4.4)-(4.9) can be obtained by applying the mutual information with finite alphabet inputs (See Section 2.2) to

$$R_{\mathrm{c},k} = I(s_{\mathrm{c}}; y_k) = I\Big(\{s_{\mathrm{c}}, s_{\mathrm{p},1}, ..., s_{\mathrm{p},K}\}; y_k\Big) - I\Big(\{s_1, ..., s_K\}; y_k \mid s_{\mathrm{c}}\Big), \tag{4.1}$$

$$R_{\mathrm{p},k}^{\mathrm{SIC}} = I(s_{\mathrm{p},k}; y_k | s_{\mathrm{c}}), \tag{4.2}$$

and

$$R_{\mathrm{p},k}^{\text{SIC-free}} = I(s_{\mathrm{p},k}; y_k). \tag{4.3}$$

This extends the proof in Appendix B.1 by accounting for interference from private streams intended for users other than user-k and is therefore omitted. Note that the difference between (4.2) and (4.3) lies in the availability of knowledge on s_{c}.

Note that the "CC" subscript used in Chapter 3 is omitted in the notation, as the remainder of this thesis focuses exclusively on finite-alphabet inputs.

$$R_{\mathrm{c},k} = \log|\mathcal{X}_{\mathrm{c}}| - \frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\mathbb{E}_{z_k}\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}+\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_{\mathrm{p},a}\mathcal{X}_{\mathrm{p},a})+z_k|^2}{\sigma_z^2}\Big)\Big\}$$
$$+\frac{1}{|\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathcal{K}}|}\mathbb{E}_{z_k}\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_{\mathrm{p},a}\mathcal{X}_{\mathrm{p},a})+z_k|^2}{\sigma_z^2}\Big)\Big\} \tag{4.4}$$

$$R_{\mathrm{p},k}^{\mathrm{SIC}} = \log|\mathcal{X}_k| - \frac{1}{|\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathcal{K}}|}\mathbb{E}_{z_k}\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_{\mathrm{p},a}\mathcal{X}_{\mathrm{p},a})+z_k|^2}{\sigma_z^2}\Big)\Big\}$$
$$+\frac{1}{|\mathcal{X}_{\mathcal{K}/k}|}\sum_{m=1}^{|\mathcal{X}_{\mathcal{K}/k}|}\mathbb{E}_{z_k}\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathcal{K}/k}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\sum_{a\in\mathcal{K}/k}\mathbf{h}^H\mathbf{p}_{\mathrm{p},a}\mathcal{X}_{\mathrm{p},a})+z_k|^2}{\sigma_z^2}\Big)\Big\} \tag{4.5}$$

$$R_{\mathrm{p},k}^{\text{SIC-free}} = \log|\mathcal{X}_k| - \frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\mathbb{E}_{z_k}\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}+\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_{\mathrm{p},a}\mathcal{X}_{\mathrm{p},a})+z_k|^2}{\sigma_z^2}\Big)\Big\}$$
$$+\frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}/k}|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}/k}|}\mathbb{E}_{z_k}\Big\{\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}/k}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}+\sum_{a\in\mathcal{K}/k}\mathbf{h}^H\mathbf{p}_{\mathrm{p},a}\mathcal{X}_{\mathrm{p},a})+z_k|^2}{\sigma_z^2}\Big)\Big\} \tag{4.6}$$

$$R_{\mathrm{c},k}^{\text{approx.}} = \log|\mathcal{X}_{\mathrm{c}}| - \frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}+\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_a\mathcal{X}_{\mathrm{p},a})|^2}{2\sigma_z^2}\Big)$$
$$+\frac{1}{|\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathcal{K}}|}\log\sum_{l=1}^{|\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_a\mathcal{X}_{\mathrm{p},a})|^2}{2\sigma_z^2}\Big) \tag{4.7}$$

$$R_{\mathrm{p},k}^{\text{SIC, approx.}} = \log|\mathcal{X}_k| - \frac{1}{|\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathcal{K}}|}\log\sum_{l=1}^{|\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_a\mathcal{X}_{\mathrm{p},a})|^2}{2\sigma_z^2}\Big)$$
$$+\frac{1}{|\mathcal{X}_{\mathcal{K}/k}|}\sum_{m=1}^{|\mathcal{X}_{\mathcal{K}/k}|}\log\sum_{l=1}^{|\mathcal{X}_{\mathcal{K}/k}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\sum_{a\in\mathcal{K}/k}\mathbf{h}^H\mathbf{p}_a\mathcal{X}_{\mathrm{p},a})|^2}{2\sigma_z^2}\Big) \tag{4.8}$$

$$R_{\mathrm{p},k}^{\text{SIC-free, approx.}} = \log|\mathcal{X}_k| - \frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}+\sum_{a\in\mathcal{K}}\mathbf{h}^H\mathbf{p}_a\mathcal{X}_{\mathrm{p},a})|^2}{2\sigma_z^2}\Big)$$
$$+\frac{1}{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}/k}|}\sum_{m=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}/k}|}\log\sum_{l=1}^{|\mathcal{X}_{\mathrm{c}}\times\mathcal{X}_{\mathcal{K}/k}|}\exp\Big(-\frac{|\mathbf{d}_{m,l}(\mathbf{h}^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}+\sum_{a\in\mathcal{K}/k}\mathbf{h}^H\mathbf{p}_a\mathcal{X}_{\mathrm{p},a})|^2}{2\sigma_z^2}\Big) \tag{4.9}$$

Recall from Section 2.1 that c_k denotes the portion of the common stream rate that is allocated to user-k. Following (2.4), the rate experienced by user-k can be expressed as

$$R_k^{\text{SIC/SIC-free}} = c_k R_{\text{c}} + R_{\text{p},k}^{\text{SIC/SIC-free}}, \tag{4.10}$$

where the superscript "SIC/SIC-free" indicates that the expression depends on whether SIC receivers or SIC-free receivers are used.

The performance loss incurred by removing the SIC at user-k can be characterized as

$$\begin{aligned} &R_k^{\text{SIC}} - R_k^{\text{SIC-free}} \\ =&I(s_{\text{p},k}; y_k) - I(s_{\text{p},k}; y_k | s_{\text{c}}) \\ =&I(s_{\text{p},k}; \sum_{k' \in \mathcal{K}} \mathbf{h}_k^H \mathbf{p}_{\text{p},k'} s_{\text{p},k'} + \mathbf{h}_k^H \mathbf{p}_{\text{c}} s_{\text{c}} + z_k) - I(s_{\text{p},k}; \sum_{k' \in \mathcal{K}} \mathbf{h}_k^H \mathbf{p}_{\text{p},k'} s_{\text{p},k'} + z_k). \end{aligned} \tag{4.11}$$

Understanding this difference requires addressing the following question: Consider a channel that takes $\mathbf{h}_k^H \mathbf{p}_{\text{p},k} s_{\text{p},k}$ as input and is disturbed by interference and noise denoted by $\sum_{k' \in \mathcal{K}/k} \mathbf{h}_k^H \mathbf{p}_{\text{p},k'} s_{\text{p},k'} + z_k$. How detrimental is it to the achievable rate if an additional interference, $\mathbf{h}_k^H \mathbf{p}_{\text{c}} s_{\text{c}}$, is added? It is intuitive that s_{c} is less detrimental to the achievable rate if it is from a finite alphabet rather than Gaussian distributed, which motivated our analysis of SIC-free RSMA under finite alphabets. The impact of finite alphabet interference will be further studied in Chapter 5, where we will see that this impact is non-monotonic w.r.t. the interference power and can be low even if the interference is strong. In particular, it can be shown that

$$\lim_{\text{E}[|\mathbf{h}_k^H \mathbf{p}_{\text{c}}|^2] \to +\infty} I(s_{\text{p},k}; \sum_{k' \in \mathcal{K}} \mathbf{h}_k^H \mathbf{p}_{\text{p},k'} s_{\text{p},k'} + \mathbf{h}_k^H \mathbf{p}_{\text{c}} s_{\text{c}} + z_k) = I(s_{\text{p},k}; \sum_{k' \in \mathcal{K}} \mathbf{h}_k^H \mathbf{p}_{\text{p},k'} s_{\text{p},k'} + z_k). \tag{4.12}$$

This indicates that the achievable rate of SIC-free RSMA can be made similar to SIC-based RSMA through proper precoding design.

4.2 Precoding optimization

In this section, we consider WSR and MMF maximization problems for RSMA with SIC and SIC-free receivers, and propose precoding optimization algorithms to address them.

4.2.1 Weighted sum-rate optimization

The WSR problem can be expressed as

$$\begin{aligned} \max_{\mathbf{P},\mathbf{c}} \ & \sum_{k\in\mathcal{K}} u_k R_k^{\text{SIC/SIC-free}} \\ \text{s.t.} \ & \|\mathbf{P}\|_{\text{F}}^2 \leq P_{\text{T}} \\ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}^{1\times K}, \end{aligned} \tag{4.13}$$

where u_k, $k \in \mathcal{K}$, denotes the weight assigned to user-k, P_{T} denotes the transmit power budget, and $\mathbf{c} = [c_1, c_2, ..., c_K]$.

Proposition 4.1. *Let $i \in \arg\max_{k\in\mathcal{K}} u_k$. If $(\mathbf{P}^\star, \mathbf{c}^\star)$ is a solution of 4.13, $(\mathbf{P}^\star, \mathbf{c}^{\star'})$ is another solution where $\mathbf{c}^{\star'} = [c_1^{\star'}, c_2^{\star'}, ..., c_K^{\star'}]$ satisfies $c_i^{\star'} = 1$ and $c_j^{\star'} = 0, \forall j \in \mathcal{K}\backslash i$.*

Proof. See Appendix C.1. □

Remark 4.1. *The intuition behind Proposition 4.1 is of significant value for the practical implementation of RSMA, as it implies that when the objective is to maximize the WSR, only splitting the message of the most weighted users (or one of the most weighted users if multiple users are most weighted) is no worse than splitting multiple user messages. In practice, the splitting of messages requires additional implementation and control signaling, making splitting only one message a more favorable option.*

Remark 4.2. *The proof of Proposition 4.1 does not assume any input distribution, which means that Proposition 4.1 holds for arbitrary input distributions, including Gaussian input, and any other distribution for discrete or continuous space.*

As a direct consequence of Proposition 4.1, one solution of $\mathbf{c}$ can already be determined given u_k, $\forall k \in \mathcal{K}$. As a result, $\mathbf{c}$ can be removed from the design variables, and 4.13 is simplified into

$$\begin{aligned} \max_{\mathbf{P}} \ & u_{k'} R_{\text{c}} + \sum_{k\in\mathcal{K}} u_k R_{\text{p},k}^{\text{SIC/SIC-free}} \\ \text{s.t.} \ & \|\mathbf{P}\|_{\text{F}}^2 \leq P_{\text{T}}, \end{aligned} \tag{4.14}$$

where $k' \in \arg\max_{k\in\mathcal{K}} u_k$.

One challenge in solving (4.14) lies in the fact that the expectation in (2.7) cannot be reduced to an analytical form. One way to address it is to estimate (2.7) empirically and optimize that estimate, but this can introduce a significant computational burden to algorithms. We adopt an alternative approach by using the approximation given by (4.7)-(4.9) during the optimization process. The effectiveness of this approach has been illustrated, for example, in [72, 89, 90]. To tackle the non-convexity and complicated structure of (2.9) and the discontinuity brought about by the point-wise minimum in (2.3), a subgradient method, as introduced in Section 2.4, is chosen to solve (4.14).

With a vector channel $\mathbf{h}$, the gradient of (2.9) is given by

$$\nabla_{\mathbf{P}} H(\mathbf{s}|\mathbf{y})^{\text{approx.}} = \frac{1}{|\mathcal{X}|} \sum_{m=1}^{|\mathcal{X}|} \frac{1}{\sum_{l=1}^{|\mathcal{X}|} u(\mathbf{P})} \sum_{l=1}^{|\mathcal{X}|} u(\mathbf{P}) \left(-\frac{\mathbf{h}\mathbf{h}^H \mathbf{P} \mathbf{d}_{m,l}(\mathcal{X}) \mathbf{d}_{m,l}(\mathcal{X})^H}{2\sigma_z^2} \right), \tag{4.15}$$

where

$$u(\mathbf{P}) = \sum_{l=1}^{|\mathcal{X}|} \exp\Big(- \frac{|\mathbf{h}^H \mathbf{P} \mathbf{d}_{m,l}(\mathcal{X})|^2}{2\sigma_z^2} \Big). \tag{4.16}$$

Based on (4.15), $\nabla_{\mathbf{P}} R_{\mathrm{c},k}$, $\nabla_{\mathbf{P}} R_k^{\text{SIC}}$ and $\nabla_{\mathbf{P}} R_k^{\text{SIC-free}}$ can be obtained trivially. One subgradient[1] of the WSR can be obtained as

$$g(R_{\text{WSR}}^{\text{SIC/SIC-free}}) = u_{k'} \nabla_{\mathbf{P}} R_{\mathrm{c},k^\circ} + \sum_{k \in \mathcal{K}} u_k \nabla_{\mathbf{P}} R_k^{\text{SIC/SIC-free}}, \tag{4.17}$$

where $k^\circ \in \arg\min_{k \in \mathcal{K}} R_{\mathrm{c},k}$.[2]

It is easy to see that any solution of (4.14), denoted by $\mathbf{P}^\star$, must satisfy $\|\mathbf{P}^\star\|_{\mathrm{F}}^2 = P_{\mathrm{T}}$. We therefore use the projection of $\mathbf{P}$ in $\{\mathbf{P} \mid \|\mathbf{P}\|_{\mathrm{F}}^2 = P_{\mathrm{T}}\}$ as the updated solution after every iteration. The proposed Projected SubGradient Ascent (PSGA) algorithm is summarized in Algorithm 1.[3] For choosing the step length, t, we use a modified backtracking line search specific for PSGA as described in Algorithm 2. Compared to the classical backtracking for gradient descent/ascent, the proposed backtracking method incorporates projection when evaluating the predicted objective value after updating, ensuring a non-decreasing sequence of objectives. The convergence of general projected gradient methods for convex problems is discussed in [91], although the convergence rate of such methods for non-convex problems

[1] We use the Clarcke subgradient introduced in Section 2.4.

[2] When $\arg\min_{k \in \mathcal{K}} R_{\mathrm{c},k}$ has multiple values, any of them can be used as k°, as they all lead to valid subgradients.

[3] In Algorithm 1, depending on the assumption about receivers, $R_{\text{WSR}} = R_{\text{WSR}}^{\text{SIC}}$ or $R_{\text{WSR}}^{\text{SIC-free}}$.

Algorithm 1 Projected Subgradient Ascent for Weight Sum-rate Maximization

Input: P_{T}, $\mathcal{X}_{\mathrm{c}}$, $\mathcal{X}_k$, $\mathbf{h}_k$, $k \in \mathcal{K}$, ϵ, $v_{\max}$.
Output: $\mathbf{P}^\star$

Initialize $\mathbf{P}^0$.
$v := 0$.
Update. $R_{\mathrm{WSR}}^v := R_{\mathrm{WSR}}(\mathbf{P}^v)$.
repeat
  $\Delta\mathbf{P} := g(R_{\mathrm{WSR}})$ according to (4.17).
  Choose step size t via Algorithm 2.
  Update. $\mathbf{P}^{v+1} := \mathbf{P}^v + t\Delta\mathbf{P}$.
  Projection. $\mathbf{P}^{v+1} := \frac{\sqrt{P_{\mathrm{T}}}}{\|\mathbf{P}^{v+1}\|_{\mathrm{F}}}\mathbf{P}^{v+1}$.
  Update. $R_{\mathrm{WSR}}^{v+1} := R_{\mathrm{WSR}}(\mathbf{P}^{v+1})$.
  $v := v + 1$.
until $\mid R_{\mathrm{WSR}}^v - R_{\mathrm{WSR}}^{v-1} \mid < \epsilon$ or $v \geq v_{\max}$.
return $\mathbf{P}^\star := \mathbf{P}^v$.

Algorithm 2 Backtracking Line Search for PSGA

Input: $\Delta\mathbf{P}$, $f(\cdot)$,[4] P_{T}, $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$ and $t_{\min}$.
Output: t

$t := 1$.
while $f(\frac{\sqrt{P_{\mathrm{T}}}}{\|\mathbf{P}+t\Delta\mathbf{P}\|_{\mathrm{F}}}(\mathbf{P} + t\Delta\mathbf{P})) \leq f(\mathbf{P}) + \alpha t\|\Delta\mathbf{P}\|_{\mathrm{F}}^2$ **and** $t > t_{\min}$, **do**
  $t := \beta t$.
end while
if $t \leq t_{\min}$, **then**
  $t := 0$.
end if
return t

is generally difficult to analyze. Note that each iteration in Algorithm 2 leads to a non-decreasing variation of the objective function, and the objective function of (4.13) is upper-bounded. Therefore, Algorithm 1 is guaranteed to converge. The convergence is also numerically evaluated in Figure 4.5, where a monotonic increase in the objective function and the convergence can be observed.

[4] $f(\cdot)$ refers to the objective function of interest. It is kept general here, as Algorithm 2 will be reused later to solve the MMF problem.

4.2.2 Max-min fairness optimization

We now consider an MMF problem as follows:

$$\begin{aligned} \max_{\mathbf{P},\mathbf{c}} \ & \min_{k\in\mathcal{K}} \ R_k^{\text{SIC/SIC-free}} \\ \text{s.t.} \ & \|\mathbf{P}\|_{\text{F}}^2 \leq P_{\text{T}} \\ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}^{1\times K}. \end{aligned} \tag{4.18}$$

Note that (4.18) is also a non-smooth non-convex problem and hence is difficult to solve. To develop efficient algorithms for 4.18, we decouple it into the following two sub-problems,

$$\begin{aligned} \max_{\mathbf{c}} \ & \min_{k\in\mathcal{K}} \ R_k(\mathbf{c};\mathbf{P}) \\ \text{s.t.} \ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}^{1\times K}, \end{aligned} \tag{4.19}$$

and

$$\begin{aligned} \max_{\mathbf{P}} \ & \min_{k\in\mathcal{K}} \ R_k(\mathbf{P};\mathbf{c}) \\ \text{s.t.} \ & \|\mathbf{P}\|_{\text{F}}^2 \leq P_{\text{T}}, \end{aligned} \tag{4.20}$$

where $R_k(\mathbf{c};\mathbf{P})$ indicates the case where $R_k^{\text{SIC/SIC-free}}$ is determined with $\mathbf{P}$ as a given parameter and $\mathbf{c}$ as a variable, whereas $R_k(\mathbf{P};\mathbf{c})$ indicates the case in which $\mathbf{c}$ is given and $\mathbf{P}$ is the variable. Hence, (4.19) optimizes $\mathbf{c}$ with a fixed $\mathbf{P}$ while (4.20) optimizes $\mathbf{P}$ with a given $\mathbf{c}$.

Note that (4.19) is linear programming; therefore, its global optimal solution can be obtained efficiently. In particular, we developed a closed-form solution as in Proposition 4.2.

Proposition 4.2. *The global optimal solution of (4.19) is given in closed form by Algorithm 3.*

Proof. See Appendix C.2. □

Algorithm 3 Global Optimal Common Stream Allocation

Input: R_{c}, $R_{\mathrm{p},k}$, $k \in \mathcal{K}$.
Output: $\mathbf{c}^{\star}$

$k' := K + 1$.
repeat
  $k' := k' - 1$.
  $\mathbf{d} := [R_{\mathrm{p},1}, ..., R_{\mathrm{p},k'}]$.
  $\eta := \frac{R_{\mathrm{c}} + \mathbf{d}\mathbf{1}^{k' \times 1}}{k'}$.
  $\mathbf{c}' := (\eta \mathbf{1}^{k' \times 1} - \mathbf{d})/R_{\mathrm{c}}$.
until $\mathbf{c}' \succcurlyeq 0$.
return $\mathbf{c}^{\star} := [\mathbf{c}',\ \mathbf{0}^{1 \times (K-k')}]$.

To remove the discontinuity caused by the point-wise minimum in (4.20), we apply a log-sum-exp approximation to its objective function and consider the following problem.

$$\begin{aligned} \max_{\mathbf{P}} \ & \frac{1}{\gamma} \log \sum_{k \in \mathcal{K}} \exp\left(\gamma R_k(\mathbf{P}; \mathbf{c})\right) \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{T}}. \end{aligned} \tag{4.21}$$

The gap between the point-wise minimum and its log-sum-exp approximation is characterized by the following [92].

$$0 \leq \min_{1 \leq i \leq L} f_i - \frac{1}{\gamma} \log \sum_{i=1}^{L} \exp(\gamma f_i) \leq \frac{1}{\gamma} \log L, \quad \gamma < 0 \tag{4.22}$$

Let $R_{\text{max-min}}$ denote the objective function in (4.21). One subgradient of $R_{\text{max-min}}$ w.r.t. $\mathbf{P}$ is derived to be

$$g(R_{\text{max-min}}) = \frac{\sum_{k \in \mathcal{K}} \exp\left(\gamma R_k(\mathbf{P}; \mathbf{c})\right) g\big(R_k(\mathbf{P}; \mathbf{c})\big)}{\gamma \sum_{k \in \mathcal{K}} \exp\left(\gamma R_k(\mathbf{P}; \mathbf{c})\right)}, \tag{4.23}$$

where

$$g\left(R_k(\mathbf{P}; \mathbf{c})\right) = c_k \nabla_{\mathbf{P}} R_{\mathrm{c},k^{\circ}} + \nabla_{\mathbf{P}} R_{\mathrm{p},k}, \tag{4.24}$$

and $k^{\circ} = \arg\min_k R_{\mathrm{c},k}$. Similarly to (4.14), (4.21) can also be solved using PSGA. The solution of (4.18) is obtained by updating the solution for (4.19) and (4.21) alternatively, and the procedures are summarized in Algorithm 4. Note that, in Algorithm 4, the updates on $\mathbf{P}$ do not decrease the objective function due to the step sizes selected by Algorithm 2, and the updates on $\mathbf{c}$ also do not decrease the objective function, because Algorithm 3 leads to the global optimum. Hence, every

Algorithm 4 Subgradient Ascent with Projection for MMF Optimization

Input: P_{T}, $\mathcal{X}_{\mathrm{c}}$, $\mathcal{X}_k$, $\mathbf{h}_k$, $k \in \mathcal{K}$, γ, ϵ, $v_{\max}$.
Output: $\mathbf{P}^\star$ and $\mathbf{c}^\star$.

Initialize $\mathbf{P}^0$.
$v := 0$.
Update. $R^0_{\text{max-min}} := R_{\text{max-min}}(\mathbf{P}^0)$.
repeat
  Update R_{c}, $R_{\mathrm{p},k}$ based on $\mathbf{P}^v$.
  Update $\mathbf{c}$ using Algorithm 3.
  $\Delta\mathbf{P} := g(R_{\text{max-min}})$ according to (4.23) and (4.24).
  Choose step size t via Algorithm 2.
  Update. $\mathbf{P}^{v+1} := \mathbf{P}^v + t\Delta\mathbf{P}$.
  Projection. $\mathbf{P}^{v+1} := \frac{\sqrt{P_{\mathrm{T}}}}{\|\mathbf{P}^{v+1}\|_{\mathrm{F}}}\mathbf{P}^{v+1}$.
  Update. $R^{v+1}_{\text{max-min}} := R_{\text{max-min}}(\mathbf{P}^{v+1})$.
  $v := v + 1$.
until $| R^v_{\text{max-min}} - R^{v-1}_{\text{max-min}} | < \epsilon$ or $v \geq v_{\max}$.
return $\mathbf{P}^\star := \mathbf{P}^v$, $\mathbf{c}^\star = \mathbf{c}$.

iteration in Algorithm 4 leads to a non-decreasing variation of the objective function. It is trivial to see that the objective function of 4.18 is upper-bounded. Therefore, Algorithm 4 is guaranteed to converge.

4.3 Optimization for large-scale systems

In this section, we propose methods to extend the precoding designs introduced in Section 4.3 to systems with a large number of transmit antennas and a large number of co-scheduled users.

To solve linear precoding design problems for large-scale systems, the inherent challenge is the growing size of the design variable, i.e., the precoding matrix, which leads to growing computational complexity. For problems considering finite constellations, the even more fatal issue is the exponential growth in the cardinality of observed constellations at each user. It can be observed from (2.7) and (2.9) that the computational complexity grows exponentially with respect to the number of received signal sources. An estimate of the computational complexity of the sum-rate of a MU-MIMO system with different numbers of users and constellation cardinality is given in Table 4.1 in terms of floating point operations (FLOPs). In practice, evaluating scenarios with more than 12 bits transmitted across all streams becomes difficult.

Table 4.1: An estimate of the number of FLOPs for computing constellation-constrained rate

K	QPSK	16QAM	64QAM
2	1.18×10^{8}	2.77×10^{10}	7.05×10^{12}
6	2.46×10^{13}	3.56×10^{20}	5.95×10^{27}
10	2.45×10^{18}	2.55×10^{30}	2.79×10^{42}
50	1.79×10^{67}	2.72×10^{127}	4.36×10^{187}

Table 4.2: An estimate of the number of FLOPs for computing constellation-constrained rate with user grouping and interference nulling

K	QPSK	16QAM	64QAM
2	1.18×10^{8}	2.77×10^{10}	7.05×10^{12}
6	3.55×10^{8}	8.31×10^{10}	2.11×10^{13}
10	5.91×10^{8}	1.38×10^{11}	3.52×10^{13}
50	2.96×10^{9}	6.92×10^{11}	1.76×10^{14}

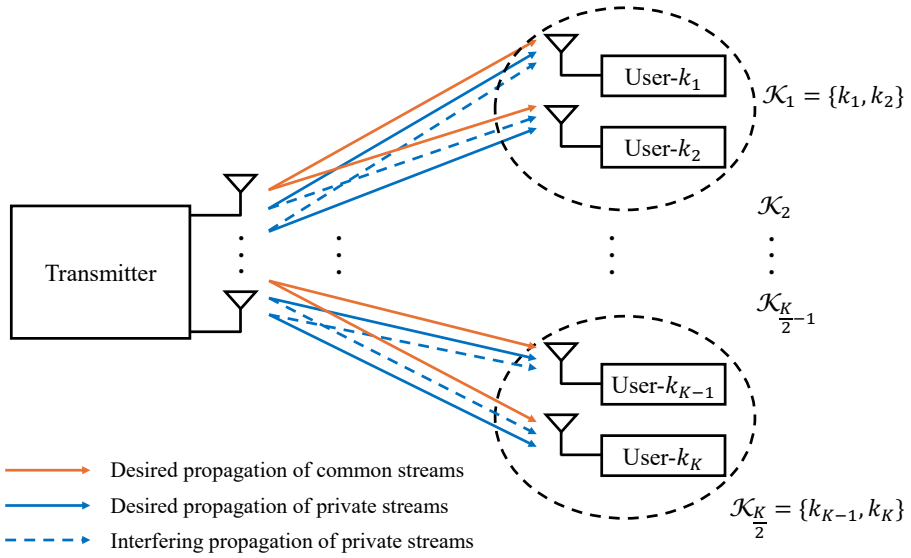


Figure 4.2: Illustration of user grouping.

To enable precoding designs and evaluation for large-scale systems (e.g., massive MIMO) with practical computational burden and acceptable reduction in performance, we introduce user grouping and null space projection techniques. The idea is to partition the co-scheduled users into groups, enable one common stream for each group, and constrain the precoders such that inter-group interference is eliminated. Figure 4.2 illustrates a system in which K users (assuming that K is even) are partitioned into $\frac{K}{2}$ groups, and each group contains two users. The elimination of inter-group interference is illustrated since signals carrying the common and private streams intended for a particular user do not propagate to other groups, although intra-group interference, i.e., interference caused by the private stream intended for a co-group user, potentially exists. In contrast to Table 4.1, an estimate of computational complexity with user grouping and interference nulling is given by Table 4.2, where the growth in FLOPs is roughly linear w.r.t. the number of users.

Problem formulation involving user grouping and interference nulling can be written as follows:

$$
\begin{aligned}
&\max_{\mathbf{P},\mathbf{c},\mathcal{G}} \ \Omega(\mathbf{P},\mathbf{c},\mathcal{G}) \\
&\text{s.t.} \ \ \|\mathbf{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{T}} && \text{(4.25a)} \\
&\quad |h_m^H \mathbf{p}_k|^2 = 0, \forall m \in \mathcal{K}_a, k \in \mathcal{K}_b, \mathcal{K}_a \in \mathcal{G}, \mathcal{K}_b \in \mathcal{G}, \mathcal{K}_a \neq \mathcal{K}_b && \text{(4.25b)} \\
&\quad |h_m^H \mathbf{p}_{\mathcal{K}_1}|^2 = 0, \forall m \notin \mathcal{K}_a \in \mathcal{G} && \text{(4.25c)} \\
&\quad \sum_{k \in \mathcal{K}_a} c_k = 1, \forall \mathcal{K}_a \in \mathcal{G} && \text{(4.25d)} \\
&\quad \mathbf{c} \succcurlyeq \mathbf{0}^{1\times K}, && \text{(4.25e)}
\end{aligned}
$$

where, for the WSR problem,

$$
\Omega(\mathbf{P},\mathbf{c},\mathcal{G}) = \sum_{k\in\mathcal{K}} u_k R_k^{\text{SIC/SIC-free}}, \tag{4.26}
$$

for the MMF problem,

$$
\Omega(\mathbf{P},\mathbf{c},\mathcal{G}) = \min_{k\in\mathcal{K}} R_k^{\text{SIC/SIC-free}}, \tag{4.27}
$$

and

$$
R_k^{\text{SIC/SIC-free}} = c_k R_{\mathcal{K}_a} + R_{\text{p},k}^{\text{SIC/SIC-free}}, \forall \mathcal{K}_a \in \mathcal{G} \text{ and } \forall k \in \mathcal{K}_a. \tag{4.28}
$$

$\mathcal{G} = \{\mathcal{K}_1, \mathcal{K}_2, ...\}$ denotes the grouping, whose elements are sets of user indices allocated to different groups. $\mathbf{P} = [\mathbf{p}_{\mathcal{K}_1}, \mathbf{p}_{\mathcal{K}_2}, ..., \mathbf{p}_{\text{p},1}, \mathbf{p}_{\text{p},2}, ..., \mathbf{p}_{\text{p},K}]$, where $\mathbf{p}_{\mathcal{K}_i}$, $\forall \mathcal{K}_i \in$

$\mathcal{G}$, is the common stream precoder intended for users in group $\mathcal{K}_i$. $R_{\mathcal{K}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$, denotes the achievable rate of the common stream intended for group $\mathcal{K}_i$.

In addition to the constraints in (4.13) and 4.18, (4.25b) and (4.25c) are added to (4.25) to enforce inter-group interference nulling. As a result of (4.25b) and (4.25c), computation of the achievable rate of user-k, $\forall k \in \mathcal{K}$, can be simplified by only considering signals intended for users in the same group as user-k. Specifically, $\forall k \in \mathcal{K}_i$, $\forall \mathcal{K}_i \in \mathcal{G}$, the computation of (4.4-4.6) and (4.7-4.9) can be reduced by substituting $\mathcal{K}_i$ for $\mathcal{K}$ without affecting the results.

(4.25) involves designing $\mathcal{G}$, i.e., the user grouping, and therefore is a combinatorial optimization problem. The joint optimization of $\mathbf{P}$, $\mathbf{c}$ and $\mathcal{G}$ is difficult. We therefore design $\mathcal{G}$ through a heuristic approach and ensure (4.25b) and (4.25c) through multi-stage precoding, followed by optimizing $\mathbf{P}$ and $\mathbf{c}$. To address the growth in the dimension of $\mathbf{P}$, we also utilize a technique to reduce the dimensionality of the design variables to further accelerate optimization algorithms.

4.3.1 User grouping

It is intuitive that, with an optimal precoding matrix, a user would suffer less interference from streams intended for co-scheduled users that have dissimilar channels. Hence, we heuristically group users with similar channels to limit the potential performance degradation. For the purpose of limiting the computational burden, the number of users within each group should be limited. We consider a special case where each group contains two users. The proposed user grouping algorithm is described in Algorithm 5[5], where $q(\mathbf{h}_m, \mathbf{h}_n)$ is a predefined channel similarity metric. We use the cosine similarity of channels for our analysis, i.e.,

$$q(\mathbf{h}_m, \mathbf{h}_n) = \frac{|\mathbf{h}_m^H \mathbf{h}_n|}{\|\mathbf{h}_m\| \|\mathbf{h}_n\|}, \tag{4.29}$$

although alternative metrics might provide superior results or serve different purposes. It is evident that Algorithm 5 can be extended to handle cases with more than two users per group.

[5]We assume K to be even for clarity, while extension to odd numbers is straightforward. For example, one can assume that one group contains only one user, while all other groups contain two users.

Algorithm 5 User Grouping Algorithm

Input: $\mathbf{h}_k$, $k \in \mathcal{K}$.
Output: $\mathcal{G}$

1: Initialize $\mathcal{G} = \varnothing$
2: Compute $\mathbf{Q} \in \mathbb{C}^{K \times K}$, where
$[\mathbf{Q}]_{m,n} = q(\mathbf{h}_m, \mathbf{h}_n)$, $\forall m < n, m \in \mathcal{K}, n \in \mathcal{K}$, and
$[\mathbf{Q}]_{m,n} = -1$, $\forall m \geq n, m \in \mathcal{K}, n \in \mathcal{K}$.
3: **for** $l \leftarrow 1$ **to** $K/2$ **do**
4: $(a, b) \leftarrow \arg\max_{m,n}[\mathbf{Q}]_{m,n}$.
5: Add $\{a, b\}$ to $\mathcal{G}$.
6: $[\mathbf{Q}]_{m,n} := -1$, $\forall (m, n) \in$
$\{(m, n) | m = a \ \vee \ m = b \ \vee \ n = a \ \vee \ n = b\}$.
7: **end for**
8: **return** $\mathcal{G}$

4.3.2 Null space projection and dimensionality reduction

We propose three-stage precoding by decomposing each precoder into a product of two matrices and a vector, and design them sequentially. Suppose that the user grouping algorithm suggests $k \in \mathcal{K}_i$, where $\mathcal{K}_i$ denotes one of the groups. The common stream precoder for group $\mathcal{K}_i$ is denoted and decomposed as

$$\mathbf{p}_{\mathcal{K}_i} = \mathbf{F}_{\mathcal{K}_i} \mathbf{G}_{\mathcal{K}_i} \mathbf{v}_{\mathcal{K}_i}, \tag{4.30}$$

and the private stream precoder for user-k is decomposed as

$$\mathbf{p}_{\mathrm{p},k} = \mathbf{F}_{\mathcal{K}_i} \mathbf{G}_{\mathcal{K}_i} \mathbf{v}_k, \tag{4.31}$$

where $\mathbf{F}_{\mathcal{K}_i}$ projects the subsequent vector onto the null space of the channels of users outside $\mathcal{K}_i$, such that inter-group interference is eliminated. We denote and express the inter-group channels to users in $\mathcal{K}_i$ as

$$\mathbf{H}_{\bar{\mathcal{K}}_i} = [\mathbf{H}_{\mathcal{K}_1}, ..., \mathbf{H}_{\mathcal{K}_{i-1}}, \mathbf{H}_{\mathcal{K}_{i+1}}, ...], \tag{4.32}$$

where the columns of $\mathbf{H}_{\mathcal{K}_j}$ are the channel vectors of users in group $\mathcal{K}_j$. With eigenvalue decomposition (EVD), it is readily seen that inter-group interference nulling can be achieved by having

$$\mathbf{F}_{\mathcal{K}_i} = \tilde{\mathbf{V}}_{\bar{\mathcal{K}}_i}, \tag{4.33}$$

where $\tilde{\mathbf{V}}_{\bar{\mathcal{K}}_i}$ denotes the eigenvectors corresponding to the zero eigenvalues of $\mathbf{H}_{\bar{\mathcal{K}}_i} \mathbf{H}_{\bar{\mathcal{K}}_i}^H$.

Apart from interference nulling, an additional benefit of $\mathbf{F}_{\mathcal{K}_i}$ is that the dimensionality of the subsequent design variable is reduced from N_{T} to $N_{\mathrm{T}} - K + 2$, i.e., the dimensionality of the null space of inter-group channels. We introduce $\mathbf{G}_{\mathcal{K}_i}$ to further reduce the dimensionality of the design variable by extending the low-dimensional subspace property (LSP) introduced by [93] to finite constellation scenarios as follows.

Proposition 4.3 (Low-Dimensional Subspace Property)**.** *Any non-trivial stationary point* $\mathbf{P}^\star$ *of (4.13) and* 4.18 *is in the range of* $[\mathbf{h}_1, ..., \mathbf{h}_K]$.

This is trivial to see as any components of $\mathbf{P}$ that lie outside the range of $[\mathbf{h}_1, ..., \mathbf{h}_K]$ do not affect the objective function of (4.13) and 4.18, hence the proof of Proposition 4.3 is omitted.

With inter-group interference nulling, the LSP implies that optimal $\mathbf{G}_{\mathcal{K}_i}\mathbf{v}_{\mathcal{K}_i}$ and $\mathbf{G}_{\mathcal{K}_i}\mathbf{v}_k$ lies in the column space of $\mathbf{F}_{\mathcal{K}_i}^H\mathbf{H}_{\mathcal{K}_j}$. This allows a reduction in the dimensionality of the design variable from $N_{\mathrm{T}} - K + 2$ to 2 while ensuring that the optimal solution can still be achieved. Let

$$\mathbf{F}_{\mathcal{K}_i}^H\mathbf{H}_{\mathcal{K}_i} = \mathbf{U}_{\mathcal{K}_i}\sigma_{\mathbf{z}}\mathbf{V}_{\mathcal{K}_i}^H \tag{4.34}$$

be compact singular value decomposition (SVD) of $\mathbf{F}_{\mathcal{K}_i}^H\mathbf{H}_{\mathcal{K}_j}$. We restrict $\mathbf{G}_{\mathcal{K}_i}\mathbf{v}_k$ to be in the desired subspace by choosing[6]

$$\mathbf{G}_{\mathcal{K}_i} = \mathbf{U}_{\mathcal{K}_i}\mathbf{I}\mathbf{V}_{\mathcal{K}_i}^H. \tag{4.35}$$

The procedures of interference nulling and dimensionality reduction are summarized in Algorithm 6. Given solutions for the first and second stages of precoding, the third stage, i.e., $\mathbf{v}_k$ and $\mathbf{v}_{\mathcal{K}_i}$, is designed in a manner similar to that in Section 4.2. Note that, as a result of inter-group interference nulling, inter-group interference does not need to be considered when computing the achievable rates and their gradients for (4.13) and 4.18. For the MMF problem, as each group posses a common stream, the design of common rate allocation, i.e., Algorithm 3, is carried out for each group separately.

[6]To achieve the same purpose, one can also have $\mathbf{G}_{\mathcal{K}_i} = \mathbf{F}_{\mathcal{K}_i}^H\mathbf{H}_{\mathcal{K}_j}$. However, it was observed numerically that the proposed $\mathbf{G}_{\mathcal{K}_i}$ leads to better convergence when optimizing $\mathbf{v}_k$. One potential reason is that the proposed $\mathbf{G}_{\mathcal{K}_i}$ guarantees that $\|\mathbf{p}_{\mathrm{p},k}\|^2 = \|\mathbf{v}_k\|^2$ and $\|\mathbf{p}_{\mathcal{K}_i,\mathrm{c}}\|^2 = \|\mathbf{v}_{\mathcal{K}_i}\|^2$.

Algorithm 6 Interference Nulling and Dimensionality Reduction

Input: $\mathcal{G}$, $\mathbf{h}_k$, $k \in \mathcal{K}$.
Output: $\mathbf{F}_{\mathcal{K}_i}$ and $\mathbf{G}_{\mathcal{K}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$.
1: Construct $\mathbf{H}_{\bar{\mathcal{K}}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$.
2: Perform EVD to $\mathbf{H}_{\bar{\mathcal{K}}_i}\mathbf{H}_{\bar{\mathcal{K}}_i}^H$ to obtain $\tilde{\mathbf{V}}_{\bar{\mathcal{K}}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$.
3: $\mathbf{F}_{\mathcal{K}_i} := \tilde{\mathbf{V}}_{\bar{\mathcal{K}}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$.
4: Perform SVD to $\mathbf{F}_{\mathcal{K}_i}^H\mathbf{H}_{\mathcal{K}_i}$ to obtain $\mathbf{U}_{\mathcal{K}_i}$ and $\mathbf{V}_{\mathcal{K}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$.
5: $\mathbf{G}_{\mathcal{K}_i} := \mathbf{U}_{\mathcal{K}_i}\mathbf{I}\mathbf{V}_{\mathcal{K}_i}^H$.
6: **return** $\mathbf{F}_{\mathcal{K}_i}$ and $\mathbf{G}_{\mathcal{K}_i}$, $\forall \mathcal{K}_i \in \mathcal{G}$.

4.3.3 Complexity analysis

For clarity, we assume that all private streams apply the same alphabet, denoted by $\mathcal{X}$. For the general approaches proposed in Section 4.2, computing (4.7), (4.9) and their gradients w.r.t. $\mathbf{P}$ requires $\mathcal{O}(|\mathcal{X}_\mathrm{c}|^2|\mathcal{X}|^{2K}KN_\mathrm{T})$ operations. Computing (4.8) and its gradient requires $\mathcal{O}(|\mathcal{X}|^{2K}KN_\mathrm{T})$. Therefore, the complexity of Algorithms 1 and 4 is $\mathcal{O}(I_\mathrm{it}|\mathcal{X}_\mathrm{c}|^2|\mathcal{X}|^{2K}K^2N_\mathrm{T})$, where I_it is the number of iterations.

With user grouping, interference nulling and dimensionality reduction, $\mathcal{O}(|\mathcal{X}_\mathrm{c}|^2|\mathcal{X}|^2)$ is required for (4.7), (4.9) and their gradients. $\mathcal{O}(|\mathcal{X}|^2)$ is required for (4.8) and its gradient. For user grouping, Algorithm 5 with the proposed similarity metric requires $\mathcal{O}(K^2N_\mathrm{T})$. The complexity of EVD for an $n \times n$ matrix is $\mathcal{O}(n^3)$, and that of SVD for an $m \times n$ matrix is $\mathcal{O}(\min(mn^2, m^2n))$. Hence, Algorithm 6 requires $\mathcal{O}(N_\mathrm{T}^3 + N_\mathrm{T}^2K + \min(N_\mathrm{T}K^2, N_\mathrm{T}^2K))$. The overall complexity of user grouping, interference nulling, dimensionality reduction, and WSR/MMF optimization proposed in this section requires $\mathcal{O}(I_\mathrm{it}|\mathcal{X}_\mathrm{c}|^2|\mathcal{X}|^2K + N_\mathrm{T}^3 + N_\mathrm{T}^2K + N_\mathrm{T}K^2)$. Compared to $\mathcal{O}(I_\mathrm{it}|\mathcal{X}_\mathrm{c}|^2|\mathcal{X}|^{2K}KN_\mathrm{T})$ for the general algorithm, the prohibitive exponent K disappears, although there is an increase in the power of N_T and K due to EVD and SVD.

4.4 Numerical results

This section introduces the simulation setup in use and presents numerical results on the performance of SIC-based and SIC-free RSMA.

The Racian channel model is used (See Appendix A. We aim to model channels in practical scenarios where users are in high-density demand areas, therefore simulating users in similar directions. For small-scale systems, a uniform linear array is

assumed along the y axis with $\theta_k^{\text{el}} = 0, \forall k \in \mathcal{K}$ since elevation does not affect the channels. For large-scale systems, a uniform rectangular array along the y-z plane is assumed. It is assumed that $\kappa_1 = ... = \kappa_K = \kappa$.

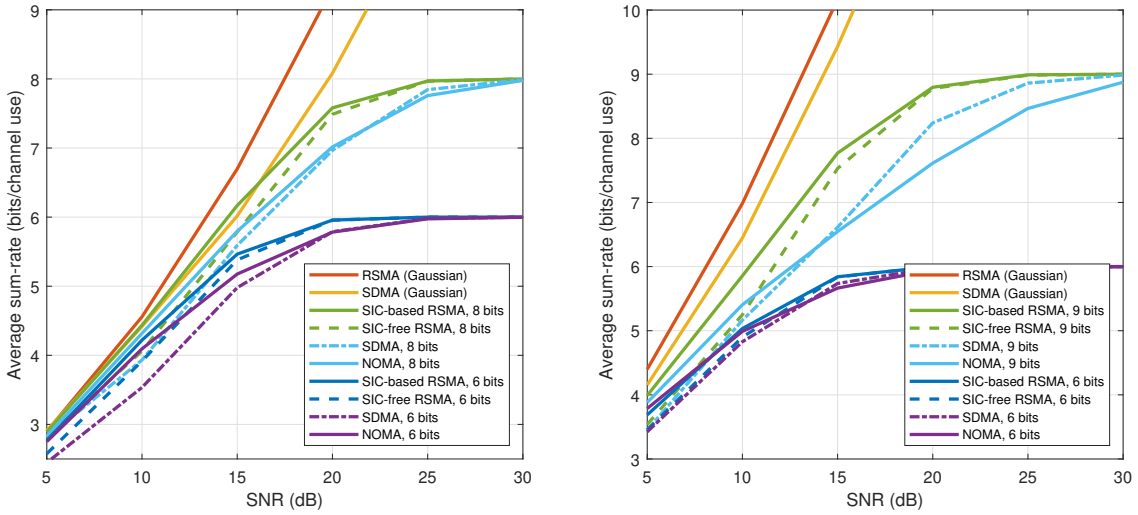
4.4.1 Adaptive modulation

The advantage of RSMA lies in its flexibility in adjusting the power and rates between common and private streams. To fully exploit this advantage under finite constellations, we allow RSMA to dynamically adjust the constellations. In particular, we enable RSMA to maximize the objective function by exhaustively searching for constellations from a pre-defined dictionary of transmission modes. To reduce the searching complexity, private streams are limited to using the same constellation in each transmission mode. To ensure fairness, different modes within the same dictionary are designed to produce the same maximum achievable rate, denoted by $R_{\max}$. According to the above design principles, the four transmission mode dictionaries used in the simulations are documented in Table C.1 - C.4 in Appendix C.3. Note that with Mode 1 in all the examples, RSMA is equivalent to SDMA, since the precoding optimization will allocate zero power to the common stream when the common stream constellation is empty. For large-scale systems, we assume that all the groups use the same transmission mode, which is searched from one of the pre-defined dictionaries. In the numerical results, where necessary, K and $R_{\max}$ are provided in the legends or captions to identify the transmission mode dictionary in use.

4.4.2 Weighted sum-rate performance

We first evaluate the sum-rate performance by setting $[u_1, ..., u_k] = \mathbf{1}^{1\times K}$. Figure 4.3 depicts the average sum-rate of SIC-based and SIC-free RSMA under finite constellations. The performance of SDMA under finite constellations is also depicted by fixing the transmission mode to be the first one in each dictionary. For two users, we also included power-domain Non-Orthogonal Multiple Access (NOMA) [15] as a benchmark, where both users apply the same constellation as in SDMA, and the user ordering is determined by the channel strength.[7] Furthermore, the performance

[7]For example, if $\|\mathbf{h}_1\|^2 \geq \|\mathbf{h}_2\|^2$, user-1 decodes user-2's message, applies SIC and decodes its own message, whereas user-2 decodes its own message directly.



(a) $N_T = 2$, $K = 2$, $\{\theta_1^{\text{az}}, \theta_2^{\text{az}}\} = \{0, \pi/18\}$, $\kappa = 10$ dB.

(b) $N_T = 4$, $K = 3$, $\{\theta_1^{\text{az}}, \theta_2^{\text{az}}, \theta_3^{\text{az}}\} = \{-\pi/36, 0, \pi/36\}$, $\kappa = 10$ dB.

Figure 4.3: Average SR performance.

of RSMA and SDMA under Gaussian distributed input signals, following [16], is included as a reference. It can be observed that, in all cases, SIC-free RSMA preserves most of the superiority of SIC-based RSMA over SDMA and NOMA, which is consistent with the conclusion in Chapter 3.

Figure 4.4 depicts the ratio of the average common stream power to the transmit power budget with SIC-based and SIC-free RSMA, i.e., $\mathbb{E}\left[\|\mathbf{p}_{\text{c}}\|^2/P_{\text{T}}\right]$, under the same setup as in Figure 4.3 (a). This result visualizes the fact that, despite SIC-based and SIC-free achieving similar performance, they demand very different precoder designs and power allocation strategies. Notably, SIC-free RSMA tends to allocate less power to the common stream in the low-SNR regime. The intuition behind this is as follows: at low SNR, there is limited flexibility in power allocation, and the algorithm may fail to find a solution where the common stream is allocated sufficient power while remaining nearly non-detrimental to the private streams. In such cases, the solution naturally assigns less power to the common stream to minimize its interference with the private streams. It can be observed that the algorithm tends to turn off the common stream at extremely high and low SNRs. It can also be observed that as $R_{\max}$ increases, the power allocation curves shift to the right-hand side, because a higher achievable rate needs to be supported by a higher SNR.

We numerically evaluate the convergence of the proposed PSGA algorithm for solving the sum-rate maximization problem of SIC-free RSMA. Additionally, another algorithm that serves the same purpose using Generalized Quadratic Matrix Pro-

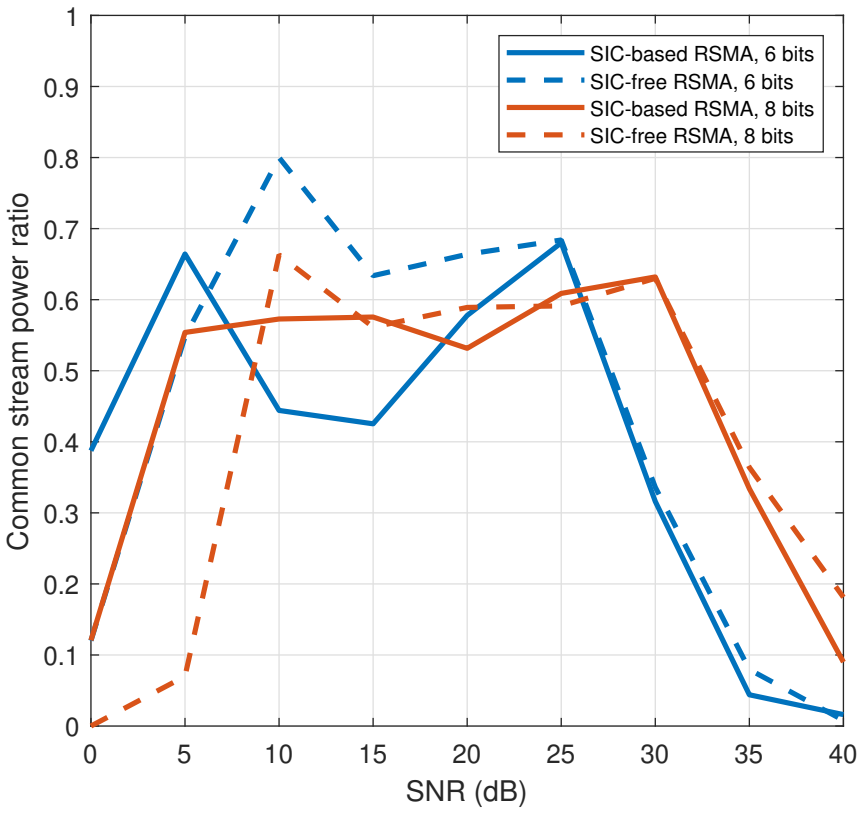


Figure 4.4: average common stream power ratio.

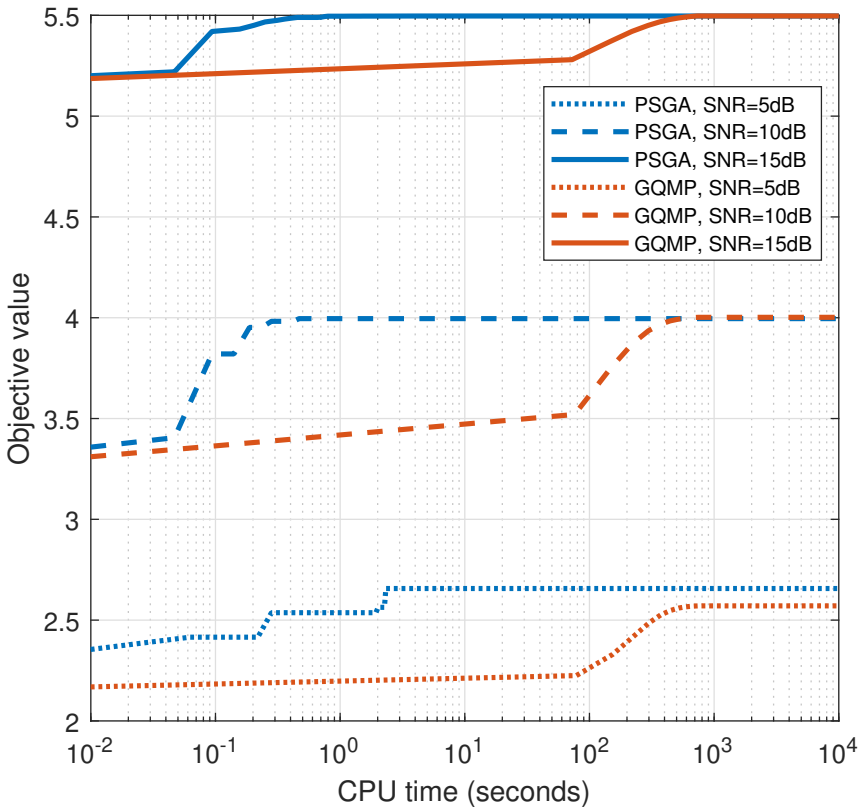


Figure 4.5: Convergence of PSGA and GQMP.

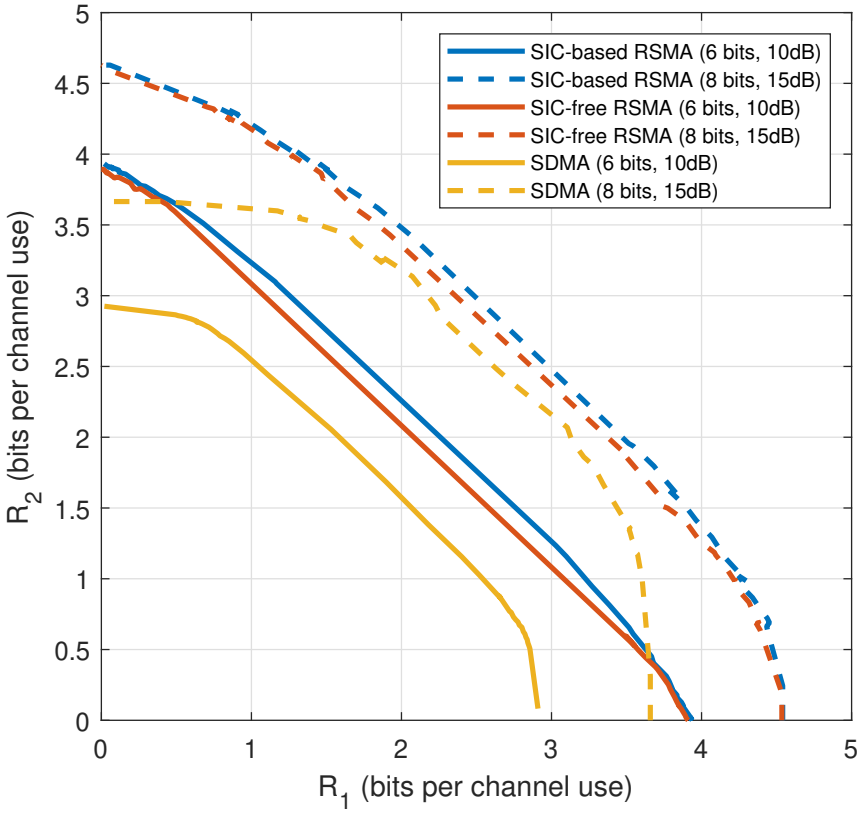


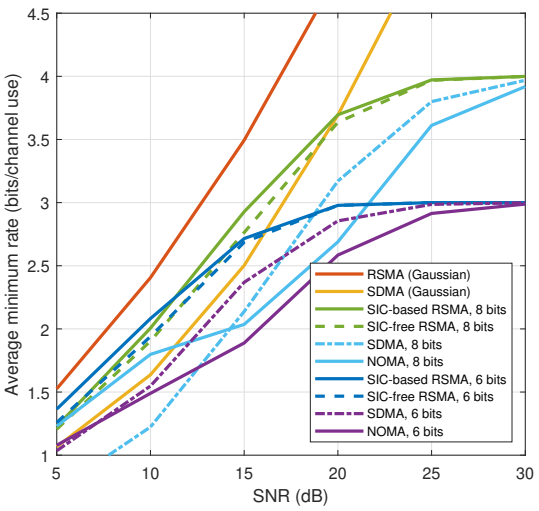
Figure 4.6: Rate region with $N_{\mathrm{T}} = 2$ and $K = 2$, $\{\theta_1^{\mathrm{az}}, \theta_2^{\mathrm{az}}\} = \{0, \pi/9\}$, $\mathcal{K} = 10$ dB.

gramming (GQMP) [90] is incorporated for comparison. The simulation is under $N_T = 2$, $K = 2$ and QPSK as the constellation for all the streams. In Figure 4.5, it is evident that both algorithms eventually converge, but PSGA always converges with a much shorter computational time than GQMP, and sometimes converges at a higher objective value.

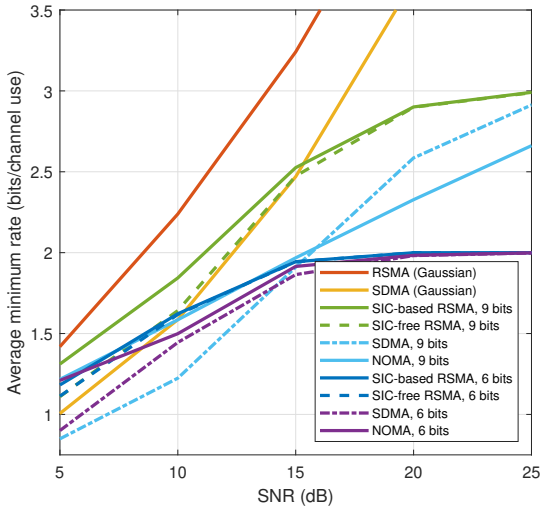
By applying various weights to the users, the rate region can be obtained. Figure 4.6 depicts the average rate region with $N_{\mathrm{T}} = 2$ and $K = 2$ and indicates that RSMA achieves better rate regions than SDMA under finite constellations, similar to Gaussian distributed inputs [16]. It also suggests that SIC-free RSMA performs similarly to SIC-based RSMA under different weights; hence, the two achieve similar rate regions. This result extends the conclusion in Chapter 3 from sum-rate to rate regions.

4.4.3 Max-min fairness performance

Figure 4.7 depicts the MMF performance of SIC-based and SIC-free RSMA under finite constellations. The performance of SDMA under finite constellations is also



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\{\theta_1^{\mathrm{az}}, \theta_2^{\mathrm{az}}\} = \{0, \pi/18\}$, $\kappa = 10$ dB.



(b) $N_{\mathrm{T}} = 4$, $K = 3$, $\{\theta_1^{\mathrm{az}}, \theta_2^{\mathrm{az}}, \theta_3^{\mathrm{az}}\} = \{-\pi/36, 0, \pi/36\}$, $\kappa = 10$ dB.

Figure 4.7: Average MMF performance.

depicted by fixing the transmission mode as in the WSR evaluations. For two users, we also included power-domain NOMA as a benchmark, where both users apply the same constellation as in SDMA, and the user ordering is determined by the channel strength[7]. In addition, the MMF performance of RSMA and SDMA under Gaussian distributed input signals, following the work in [94], is included as a reference. It can be observed that RSMA schemes achieve better MMF performance than SDMA and NOMA under finite constellations; meanwhile, SIC-free RSMA preserves most of the MMF performance of SIC-based RSMA. This result extends the conclusion in Chapter 3 from sum-rate to MMF evaluations.

4.4.4 Performance in large-scale systems

We consider a large-scale rectangular antenna array with $N_{\mathrm{T}} = 128$, with 16 antennas in each row along the y-axis and 8 antennas in each column along the z-axis, serving 64 users. Figures 4.8 and 4.9 depict the sum-rate and MMF performance of SDMA, SIC-based and SIC-free RSMA for large-scale systems, which is enabled by techniques including user grouping, null space projection, and dimensionality reduction introduced in Section 4.3. We also evaluate the impact of user grouping by comparing the performance under the proposed grouping method in Algorithm 5 and random grouping, respectively labeled as "(ordered)" and "(unordered)" in the legends. For the channel model, θ_k^{az} and θ_k^{el}, $\forall k \in \mathcal{K}$, are randomly drawn from the range of $[-\pi/4, \pi/4]$ and $[-\pi/6, \pi/6]$ with uniform distribution. κ is set to 20dB

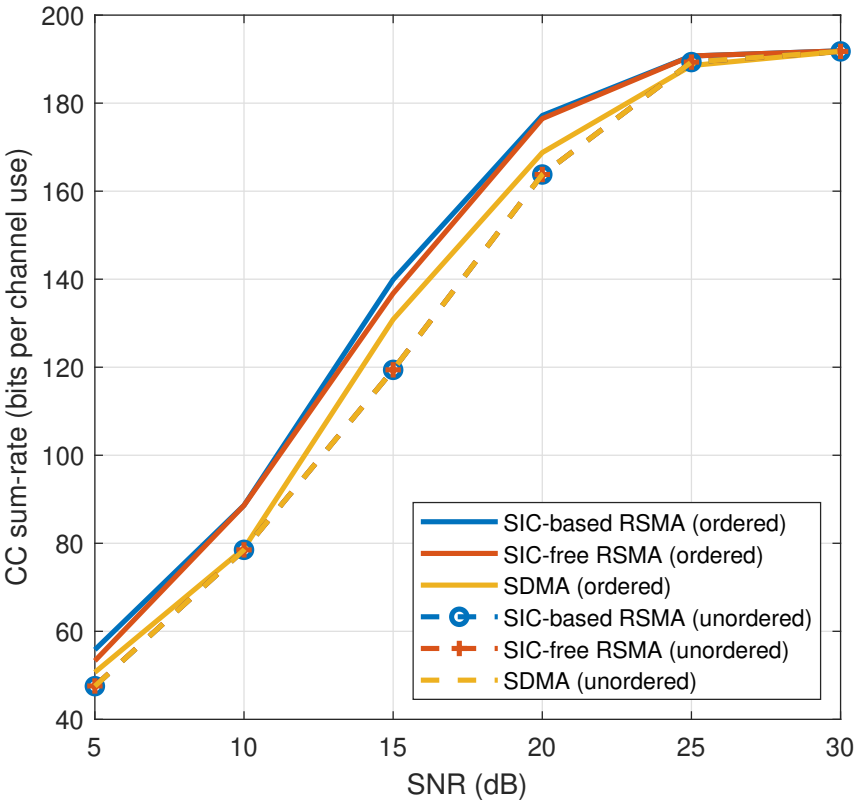


Figure 4.8: Average sum-rate performance in large-scale systems with $N_{\mathrm{T}} = 128$, $K = 64$, and $R_{\max} = 6$ bits per group.

to reflect the sparse nature of channels in typical large-scale antenna arrays. It can be observed that the conclusions drawn from small-scale systems as discussed previously, i.e., RSMA schemes outperform SDMA in both sum-rate and MMF, and SIC-free RSMA preserves most of the superiority of SIC-based RSMA, still hold in large-scale systems. Moreover, the proposed user grouping strategy generally enhances the performance of both SDMA and RSMA, with a more pronounced benefit for RSMA due to the high channel similarity within each user group.

4.5 Summary

In this chapter, we aimed to investigate the fundamental limits of SIC-based and SIC-free RSMA under finite constellations. We first derived the rate expressions for SIC-based and SIC-free RSMA. We then proposed precoding optimization algorithms for the WSR and MMF optimization problems. To enable these optimization algorithms in large-scale systems with a feasible computational burden, we introduced user grouping, null space projection and dimensionality reduction techniques.

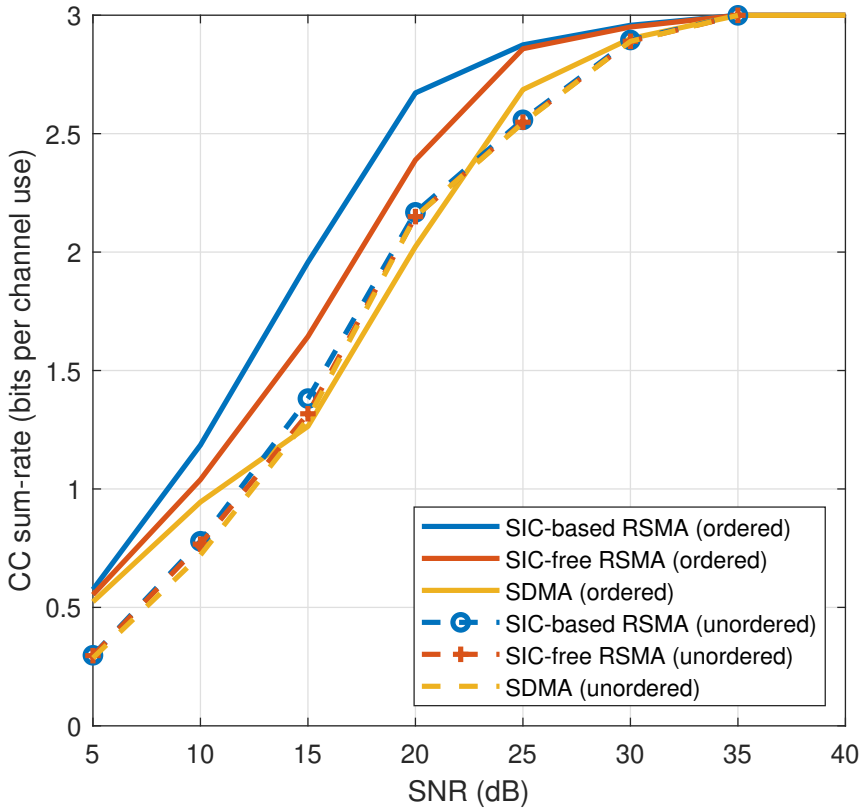


Figure 4.9: Average MMF performance in large-scale systems with $N_{\mathrm{T}} = 128$, $K = 64$, and $R_{\max} = 6$ bits per group.

The numerical results show that RSMA schemes outperform SDMA under finite constellations and in both WSR and MMF evaluation, similar to the Gaussian distributed inputs. Furthermore, SIC-free RSMA preserves most of the superiority of SIC-based RSMA, which extends the conclusion in Chapter 3 from low-complexity precoders to fully optimized precoders, from sum-rate objective to WSR and MMF objectives, and from small-scale systems to large-scale systems.

Before concluding, it is useful to summarize the trajectory of the preceding chapters: In Chapter 3, we examined various receiver implementations and categorized them into SIC-based and SIC-free types. In both Chapters 3 and 4, we analyzed the performance of these two categories using a physical layer abstraction based on mutual information, under the assumption that SIC is either implemented or not. However, mutual information alone cannot distinguish among receivers within the same category. Such a limitation arises not only in our study of RSMA but also throughout the broader literature on wireless communications. We observe that a key difference among receivers lies in the assumed interference distribution at the demodulator. In the next chapter, we shift our focus to mismatched decoding and develop tools that capture this distinction, applicable not only to RSMA receivers

but also to a broader class of receivers operating under interference.

Chapter 5

Optimal and Suboptimal Decoding under Finite-Alphabet Signals and Interference

Interference is a common challenge in wireless communication systems and is addressed in various aspects of the system design. For cellular networks, in 3GPP Release 18, multiple work items address interference from different perspectives [95]. Among others, advanced MIMO receiver designs that suppress intra- and inter-cell interference are studied [96]. [97] provides viewpoints on potential 6G RAN technologies, where the impact of interference is highlighted in various aspects. To understand the impact of interference, the vast majority of work in the literature models both the input signal and interference using Gaussian distributions, although most systems apply finite-alphabet signals. Such approaches lead to acceptable approximations of system performance under the impact of interference and benefit from the simple analytical achievable rate expressions. More precise but computationally more complicated evaluation of achievable rates can be obtained using constellation-constrained Mutual Information (MI) [98]. However, both approaches assume optimal decoders, which can be difficult in practice due to limited interference knowledge and the complexity issue. The impact of suboptimal decoders when dealing with finite-alphabet interference remains to be explored.

In this chapter, we investigate the effect of finite-alphabet interference under optimal and suboptimal decoders. The contribution is as follows:

1. We define a special channel model, namely Finite-Alphabet Gaussian Chan-

nel under Interference (FAGCI), which is a Single-Input-Single-Output (SISO) channel with finite-alphabet input, finite-alphabet interference and Gaussian noise. This channel model is to reflect practical scenarios where the communication channel suffers from not only noise but also interference from co-scheduled users or other communication links. Maximum-likelihood decoding under FAGCI is given, but can be computationally complicated once the interference alphabet has a high cardinality. A more practical decoding approach can be to treat part of the interference as Gaussian random variables (RV). We derive the MI and GMI of FAGCI, which correspond to the achievable rates under optimal and suboptimal decoders. Through analytical properties of the GMI/MI and numerical simulations, insights into the impact of finite-alphabet interference under different decoders are obtained.

2. Noting the difference in computational complexity between the matched decoding metric and mismatched decoding metrics, we propose two novel decoding metrics, namely the generalized Gaussian decoding metric and the interference decomposition decoding metric, whose computational complexity remains low. Numerical results show that they lead to improved achievable rates compared to decoders that treat interference as a Gaussian RV.

3. We show that the notion of decoding metric in the FAGCI model is related to BICM-demodulators, and propose low-complexity demodulators based on the proposed decoding metrics. Through Link-Level Simulations (LLS), it is shown that the proposed demodulators lead to improved throughput. We also show a strong correlation between the GMI and throughput when the decoding metric and demodulator apply the same principle for treating interference. Since computing the GMI is much more time-efficient than conducting LLS, such a correlation indicates that GMI can be an important performance indicator for BICM systems.

4. Motivated by the precise performance evaluation given the GMI and the correlation between the GMI and throughput, we further consider signal processing based on the GMI. In particular, we extend the GMI results for FAGCI to multiple-antenna cases and provide an example of a Multi-User Multiple-Input-Single-Output (MU-MISO) system, where each user potentially suffers from inter-user interference. We propose optimizing the precoders by taking the GMI as the objective function. Therefore, such an optimization problem incorporates decoding strategies on the user side. An optimization algorithm is proposed, and the numerical results reflect the impact of suboptimal de-

coders on the achievable rate and the necessity of using GMI as the objective function. The imperfection in receivers on treating finite-alphabet interference widely exists in practice, yet means of incorporating such imperfection are mostly overlooked in signal processing. The proposed method pioneers a novel system design framework that addresses this problem.

The following is a brief summary of the new insights and implications of finite-alphabet interference derived from this work.

1. The achievable rate under a finite-alphabet input signal and interference is in general non-monotonic w.r.t. input power.
2. Different approaches to addressing interference during decoding result in significant variations in the achievable rate.
3. Strong interference is not always detrimental: under proper interference treatment, the achievable rate under very strong interference is as if that interference were not present.
4. Decoding metrics and BICM-demodulators are strongly related, and innovations in the former can inspire improvements in the latter.
5. The GMI/MI can be a computationally efficient tool for evaluating the performance of BICM-demodulators.
6. The GMI/MI is a precise physical layer abstraction for signal processing and for system-level simulations (SLS).
7. Discrepancy between the decoding metric and the objective function in signal processing can be harmful to system performance.

Organization: The rest of this chapter is organized as follows. In Section 5.1, we propose the FAGCI model, introduce assumptions on the encoder and decoder, the resulting GMI and its approximation, and present numerical results from the GMI along with some new insights. In Section 5.2, two improved decoding metrics for the FAGCI model are proposed, namely the generalized Gaussian decoding metric and the interference decomposition decoding metric, followed by numerical results on the GMI led by them. In Section 5.3, we show the connection between the decoding metric in the FAGCI model and the BICM-demodulators, propose low-complexity demodulators based on the proposed decoding metrics, and provide numerical results

comparing the GMI and the throughput performance of the demodulators. Section 5.4 provides an example of precoder optimization in MU-MISO systems utilizing the GMI, which pioneers a novel system design framework that captures receiver imperfection in treating finite-alphabet interference. Finally, Section 5.5 summarizes this chapter.

5.1 Mismatched decoding for finite alphabets

5.1.1 Finite-alphabet Gaussian channel under interference

We consider coded communications with finite-alphabet inputs under additive interference and additive Gaussian noise, or FAGCI. Without loss of generality, we assume that there are two interference sources. Thus, at time instance t, the received signal can be expressed as

$$y_t = x_t + i_t + j_t + z_t, \tag{5.1}$$

where x_t and y_t are the complex-valued channel input and output, i_t and j_t are the two complex-valued interference observations, and z_t is the noise observation. x_t, i_t and j_t are uniformly drawn from alphabets $\mathcal{X}$, $\mathcal{I}$ and $\mathcal{J}$ correspondingly,[1] while z_t are samples of $Z_t \sim \mathcal{CN}(0,\ \sigma_z^2)$. Unless any of $\mathcal{X}$, $\mathcal{I}$ and $\mathcal{J}$ is $\{0\}$, $\mathbb{E}\left[|X_t|^2\right] = \sigma_x^2$, $\mathbb{E}\left[|I_t|^2\right] = \sigma_i^2$ and $\mathbb{E}\left[|J_t|^2\right] = \sigma_j^2$. X_t, I_t, J_t and Z_t and are mutually independent are i.i.d. over t, leading to a memoryless channel, i.e.,

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}) = \prod_{t=1}^{n} f_{Y|X}(y_t \mid x_t), \tag{5.2}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the input and output sequences of length n. Henceforth, the time index is omitted if no ambiguity arises.

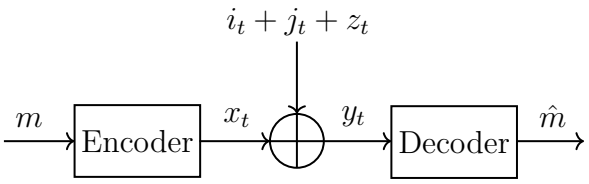


Figure 5.1: Channel model with interference and noise.

[1] This work focuses on uniformly distributed inputs and interference for clarity. The extension to non-uniform distributions is straightforward and is briefly addressed at the end of Appendix D.2.

Remark 5.1. *Note that (5.1) can represent channels under an arbitrary number of interference sources. Channels under one interference source can be represented by having either $\mathcal{I}$ or $\mathcal{J}$ equal to $\{0\}$. For more than two interference sources, one can have*

$$i_t = \sum_p i_t^p \ , \ j_t = \sum_q j_t^q \ \textit{with} \ \mathcal{I} = \sum_p \mathcal{I}^p \ \textit{and} \ \mathcal{J} = \sum_q \mathcal{J}^q, \tag{5.3}$$

where, for arbitrary range of p and q, i_t^p and j_t^q represents the elementary interference observations drawn from $\mathcal{I}^p$ and $\mathcal{J}^q$. We keep two interference sources rather than one because, in the following, we are interested in scenarios where the two interference sources are treated differently during decoding.

5.1.2 Encoding and decoding

We aim at reliably transmitting a message, m, over n channel uses, where m is drawn randomly and uniformly from $\{1, ..., M\}$. The encoding scheme is given by a codebook $\mathcal{C} = \{\mathbf{x}^{(1)}, ..., \mathbf{x}^{(M)}\} \subset \mathcal{X}^n$, i.e., to transmit the message m, the corresponding codeword $\mathbf{x}^{(m)}$ is transmitted. We further assume i.i.d. random coding ensemble, hence

$$P_{\mathbf{X}}(\mathbf{x}^{(m)}) = \prod_{t=1}^{n} P_X(x_t^{(m)}), \tag{5.4}$$

where $x_t^{(m)}$ is the t-th symbol of $\mathbf{x}^{(m)}$, and n is the length of codewords.

The decoder is represented by

$$\hat{m} = \arg \max_{l=1,...,M} q^n(\mathbf{x}^{(l)}, \ \mathbf{y}), \ \text{with} \ q^n(\mathbf{x}^{(l)}, \ \mathbf{y}) = \prod_{t=1}^{n} q(x_t^{(l)}, \ y_t), \tag{5.5}$$

where $q(x, \ y)$ is a non-negative decoding metric.

For the FAGCI channel model described above, the optimal (matched) decoder maximizes the likelihood, which leads to the following optimal decoding metric:

$$q^\star(x, \ y) = \sum_{\bar{i} \in \mathcal{I}} \sum_{\bar{j} \in \mathcal{J}} \exp\left(-\frac{|y - \bar{i} - \bar{j} - x|^2}{\sigma_z^2} \right). \tag{5.6}$$

We are also interested in suboptimal (mismatched) but simpler decoding metrics. Primarily, the following two are considered:

1. Partially treat interference as Gaussian RV by treating I optimally and treating J as a Gaussian RV[2]:

$$q(x,\ y) = \sum_{\bar{i}\in\mathcal{I}} \exp\left(-\frac{|y-\bar{i}-x|^2}{\sigma_j^2+\sigma_z^2}\right). \tag{5.7}$$

2. Fully treat interference as a Gaussian RV:

$$q(x,\ y) = \exp\left(-\frac{|y-x|^2}{\sigma_i^2+\sigma_j^2+\sigma_z^2}\right). \tag{5.8}$$

Indeed, if $J \sim \mathcal{CN}(0,\ \sigma_j^2)$ while I remains to be from finite alphabet, (5.7) is the optimal decoding metric for (5.1); if $I \sim \mathcal{CN}(0,\ \sigma_i^2)$ and $J \sim \mathcal{CN}(0,\ \sigma_j^2)$, (5.8) is the optimal decoding metric. Note that (5.8) is a special case of (5.7) as (5.7) becomes (5.8) when viewing i as 0, and j as $i+j$.

5.1.3 Matched capacity and a lower bound on mismatched capacity

Proposition 5.1 (Matched capacity)**.** *With optimal decoding, i.e., using (5.6) as the decoding metric, the capacity of (5.1) is given by the constellation-constrained MI as in (5.9).*

Proof. See Appendix D.1. □

Proposition 5.2 (A lower bound on mismatched capacity)**.** *With* ***any*** *suboptimal decoding metric denoted by $q(x,\ y)$, the mismatched capacity of (5.1) is lower bounded by the GMI as in (5.10), which is a function of $q(x,\ y)$.*

Proof. See Appendix D.2. □

[2]Another interpretation of (5.7) is to approximate $P_{I+J}(i+j)$ with a Gaussian mixture model.

$$I_{\mathrm{MI}}(\mathcal{X},\,\mathcal{I},\,\mathcal{J},\,\sigma_z^2) = \log|\mathcal{X}| - \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{x\in\mathcal{X},\,i\in\mathcal{I},\,j\in\mathcal{J}} \mathbb{E}\left[\log\sum_{\bar{x}\in\mathcal{X},\,\bar{i}\in\mathcal{I},\,\bar{j}\in\mathcal{J}} \exp\left(-\frac{|x+i+j+Z-\bar{x}-\bar{i}-\bar{j}|^2}{\sigma_z^2}\right)\right] + \frac{1}{|\mathcal{I}||\mathcal{J}|}\sum_{i\in\mathcal{I},\,j\in\mathcal{J}} \mathbb{E}\left[\log\sum_{\bar{i}\in\mathcal{I},\,\bar{j}\in\mathcal{J}} \exp\left(-\frac{|i+j+Z-\bar{i}-\bar{j}|^2}{\sigma_z^2}\right)\right] \tag{5.9}$$

$$I_{\mathrm{GMI}}(\mathcal{X},\,\mathcal{I},\,\mathcal{J},\,\sigma_z^2) = \sup_{s\geq 0}\;\log|\mathcal{X}| - \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{x\in\mathcal{X},\,i\in\mathcal{I},\,j\in\mathcal{J}} \mathbb{E}\left[\log\sum_{\bar{x}\in\mathcal{X}} q(\bar{x},\,x+i+j+Z)^s - s\log\;q(x,\,x+i+j+Z)\right] \tag{5.10}$$

$$I^{\text{approx.}}_{\mathrm{GMI,partial}}(\mathcal{X},\,\mathcal{I},\,\mathcal{J},\,\sigma_z^2) = \log|\mathcal{X}| - \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{x\in\mathcal{X},\,i\in\mathcal{I},\,j\in\mathcal{J}} \log\sum_{\bar{x}\in\mathcal{X},\,\bar{i}\in\mathcal{I}} \exp\left(-\frac{|x+i+j-\bar{x}-\bar{i}|^2}{\sigma_j^2+2\sigma_z^2}\right) + \frac{1}{|\mathcal{I}||\mathcal{J}|}\sum_{i\in\mathcal{I},\,j\in\mathcal{J}} \log\sum_{\bar{i}\in\mathcal{I}} \exp\left(-\frac{|i+j-\bar{i}|^2}{\sigma_j^2+2\sigma_z^2}\right) \tag{5.11}$$

$$I^{\text{approx.}}_{\mathrm{GMI,full}}(\mathcal{X},\,\mathcal{I},\,\mathcal{J},\,\sigma_z^2) = \log|\mathcal{X}| - \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{x\in\mathcal{X},\,i\in\mathcal{I},\,j\in\mathcal{J}} \log\sum_{\bar{x}\in\mathcal{X}} \exp\left(-\frac{|x+i+j-\bar{x}|^2}{\sigma_i^2+\sigma_j^2+2\sigma_z^2}\right) + \frac{1}{|\mathcal{I}||\mathcal{J}|}\sum_{i\in\mathcal{I},\,j\in\mathcal{J}} \left(-\frac{|i+j|^2}{\sigma_i^2+\sigma_j^2+2\sigma_z^2}\right) \tag{5.12}$$

Remark 5.2. *In (5.10), the optimization over s is convex [61].*

Remark 5.3. *The GMI given by (5.10) is the highest achievable rate of (5.1) under the assumption of the i.i.d. random coding ensemble given by (5.4). Breaking such an assumption may lead to higher achievable rates, for example, by using a constant-composition random coding ensemble[61].*

Remark 5.4. *In (5.10), $I_{GMI}(\mathcal{X},\ \mathcal{I}+\mathcal{J},\ \{0\},\ Z)$ with $s=1$ and $q(x,\ y)$ from (5.7) equals $I_{MI}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ Z)$, as optimal decoding is a special case of mismatched decoding.*

5.1.4 Approximation of mismatched capacity

One potential challenge of utilizing (5.10) stems from the fact that it is difficult to compute the expectation term analytically. One may estimate (9) empirically at the cost of a high computational burden. For the sake of computation time, we provide a closed-form approximation of (5.10).

Proposition 5.3 (Approximation of the GMI)**.** *With $q(x,y)$ given by (5.7) or (5.8), (5.10) can be approximated by (5.11) or (5.12) respectively.*

Proof. By setting $s=1$ and applying Jensen's inequality to both expectations in (5.10), the expectations can be moved inside the $\log_2$ operators. With $q(x,\ y)$ given by (5.7) or (5.8), the result is further simplified using Gaussian integral into (5.11) or (5.12) respectively. □

5.1.5 Results on GMI for FAGCI channels

We first investigate the behavior of GMI and MI in the SISO FAGCI model as expressed by (5.1) under different scenarios. In the following results, the two aforementioned mismatched decoding metrics are considered:

- Treating only J as a Gaussian RV, as given by (5.7): this is denoted by "GMI (partial)" in the following results.
- Treating both I and J as Gaussian RVs, as given by (5.8): this is denoted by "GMI (full)" in the following results.

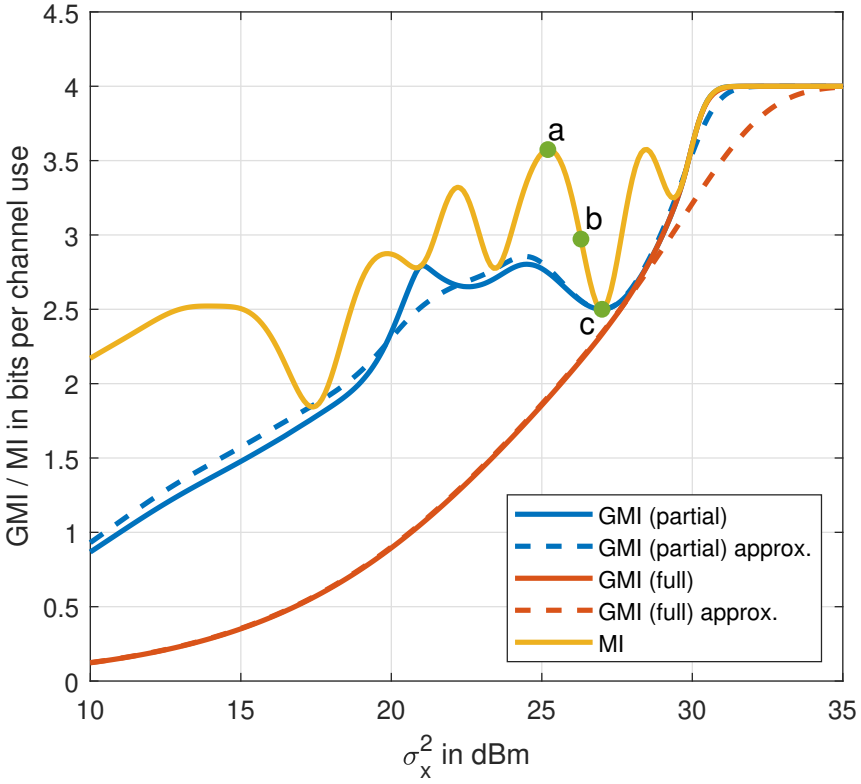


Figure 5.2: GMI and MI with $\sigma_i^2 = 20$dBm, $\sigma_j^2 = 10$dBm and $\sigma_z^2 = 0$dBm.

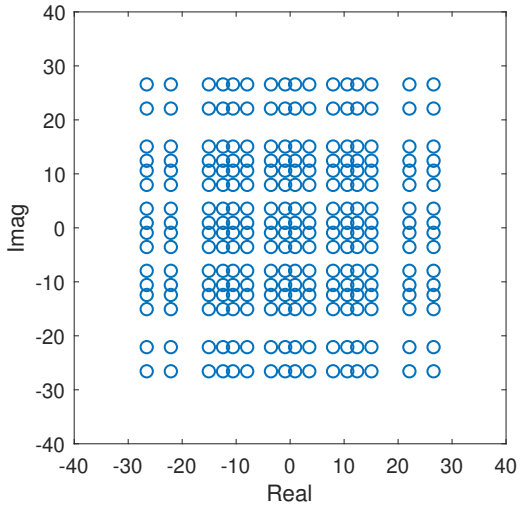
We also plot the approximations of GMI, i.e., (5.11) and (5.12), which are denoted as "GMI (partial) approx." and "GMI (full) approx." respectively. In the following, we assume that x is generated using 16-QAM, while i and j are generated using QPSK.[3]

In Figure 5.2, we consider a scenario with fixed interference and noise powers and characterize GMI and MI under different input power, i.e., σ_x^2. It can be observed that all the considered decoding metrics generally lead to improved achievable rates as the signal power increases, and all saturate at 4 bits at high input power, which is limited by the cardinality of 16-QAM. Hence, with very high input signal power, different decoding metrics under consideration lead to the same performance. With limited σ_x^2, it can be observed that "MI" upper-bounds "GMI (partial)", and "GMI (partial)" upper-bounds "GMI (full)". This aligns with expectations because matched decoding uses the optimal decoding metric, which utilizes full knowledge of interference. With mismatched decoding, only part of the interference knowledge is used. (5.7) uses full knowledge of I but only the variance of J by approximating it as a Gaussian RV. (5.8) uses only the variance of I and J by approximating both I and J as Gaussian RVs, hence uses even less knowledge on interference. In most

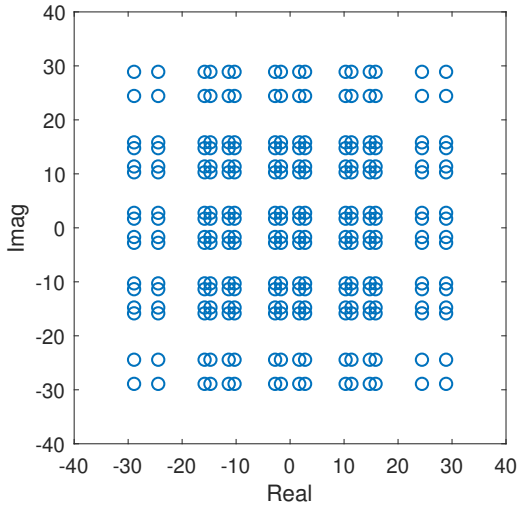
[3]The observations generally hold under different constellation setups.

cases, the more interference knowledge is utilized, the more accurate the decoding is. Hence, the order is almost always "MI" $\geq$ "GMI (partial)" $\geq$ "GMI (full)". When σ_x^2 is between 28dBm and 29.5dBm, the two GMI metrics become similar, and both are lower than MI. This is because the impact of suboptimally treating J becomes negligible, whereas the effect of suboptimally treating I remains significant, as I is stronger than J in the setup. When σ_x^2 is higher than 29.5dBm, the three rates become similar, because the impact of suboptimally treating both I and J becomes negligible. Interestingly, MI does not increase monotonically as σ_x^2 grows but fluctuates. This is because, with finite constellations, the achievable rate also depends on the sum constellations, $\mathcal{X}+\mathcal{I}+\mathcal{J}$ and $\mathcal{I}+\mathcal{J}$. Note that in (5.9), the second and third terms are conditional entropy, $H(X,\ I,\ J|Y)$ and $H(I,\ J|Y,\ X)$, respectively (this can also be seen from (D.1)). Such conditional entropy depends on the distinguishability of different points in the sum constellations. Since σ_i^2 and σ_j^2 are fixed, $H(I,\ J|Y,\ X)$ is unchanged (computed to be 0.0066 bits). The MI only depends on $H(X,\ I,\ J|Y)$, and therefore the sum constellation, $\mathcal{X}+\mathcal{I}+\mathcal{J}$. Figure 5.3 visualizes $\mathcal{X}+\mathcal{I}+\mathcal{J}$ at the three marked points in Figure 5.2, with the corresponding $H(X,\ I,\ J|Y)$ and $I(X;Y)$ given in the captions. It can be observed that, in the three examples, as σ_x^2 increases, the sum constellation becomes less distinguishable, leading to an increase in $H(X,\ I,\ J|Y)$, and therefore a drop in $I(X;Y)$.

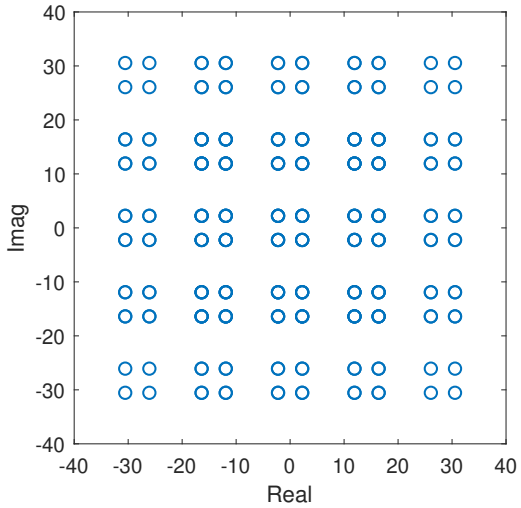
By fixing σ_x^2 and σ_j^2, and adjusting σ_i^2, Figure 5.4 can be obtained. It can be observed that there is always a gap between the GMI and MI, indicating that the two mismatched decoders always underperform the optimal decoding metric. This gap exists because J with non-negligible power is approximated as a Gaussian RV by the mismatched decoders, while optimally treated by the matched decoder. It can also be observed that, as σ_i^2 grows, "GMI (partial)" first drops, then grows and saturates in the end, while "GMI (full)" only decreases until zero. The gap between the two GMI curves is due to the difference in treating I, and such a gap increases as σ_i^2 increases. As for the MI, the fluctuation at the medium σ_i^2 regime is due to the non-monotonic change in the distinguishability of the sum constellations as analyzed above. Notably, both the MI and "GMI (partial)" saturate to their respective values, which remain the same as $\sigma_i^2 \to 0$ or $+\infty$. This trend is formally stated by Proposition 5.4.



(a) $\sigma_x^2 = 25.2\text{dBm}$, $H(X, I, J|Y) = 0.4324\text{bits}$, $I_{\text{MI}} = 3.5742\text{bits}$.



(b) $\sigma_x^2 = 26.3\text{dBm}$, $H(X, I, J|Y) = 1.0344\text{bits}$, $I_{\text{MI}} = 2.9722\text{bits}$.



(c) $\sigma_x^2 = 27\text{dBm}$, $H(X, I, J|Y) = 1.5064\text{bits}$, $I_{\text{MI}} = 2.5002\text{bits}$.

Figure 5.3: Visualization of $\mathcal{X} + \mathcal{I} + \mathcal{J}$ at the three marked points in Figure 5.2.

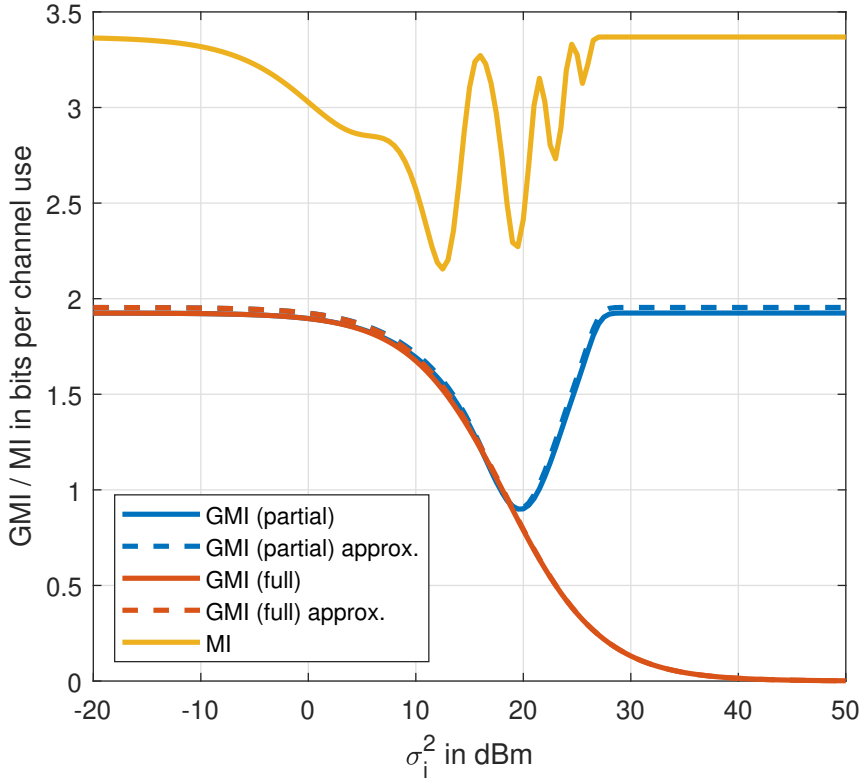


Figure 5.4: GMI and MI with $\sigma_x^2 = 20\text{dBm}$, $\sigma_j^2 = 15\text{dBm}$ and $\sigma_z^2 = 0\text{dBm}$.

Proposition 5.4.

$$\begin{aligned}\lim_{\sigma_i^2 \to 0} I_{MI}(\mathcal{X},\, \mathcal{I},\, \mathcal{J},\, \sigma_z^2) &= \lim_{\sigma_i^2 \to +\infty} I_{MI}(\mathcal{X},\, \mathcal{I},\, \mathcal{J},\, \sigma_z^2) \\ &= I_{MI}(\mathcal{X},\, \{0\},\, \mathcal{J},\, \sigma_z^2),\end{aligned} \tag{5.13}$$

and

$$\begin{aligned}\lim_{\sigma_i^2 \to 0} I_{GMI}(\mathcal{X},\, \mathcal{I},\, \mathcal{J},\, \sigma_z^2) &= \lim_{\sigma_i^2 \to +\infty} I_{GMI}(\mathcal{X},\, \mathcal{I},\, \mathcal{J},\, \sigma_z^2) \\ &= I_{GMI}(\mathcal{X},\, \{0\},\, \mathcal{J},\, \sigma_z^2)\end{aligned} \tag{5.14}$$

with $q(x,\, y)$ given by (5.7).

Proof. See Appendix D.3. □

We also consider a case with fixed σ_x^2 and σ_i^2, and with varying σ_j^2, which leads to Figure. 5.5. It can be observed that, with low σ_j^2, there is a gap between "GMI (partial)" and "GMI (full)" due to the difference in treating I from the corresponding mismatched decoders, while "GMI (partial)" is very close to "MI" because $q(x,\, y)$ given by (5.7) is almost $q^\star(x,\, y)$ when σ_j^2 is negligible. As σ_j^2 grows, the gap between

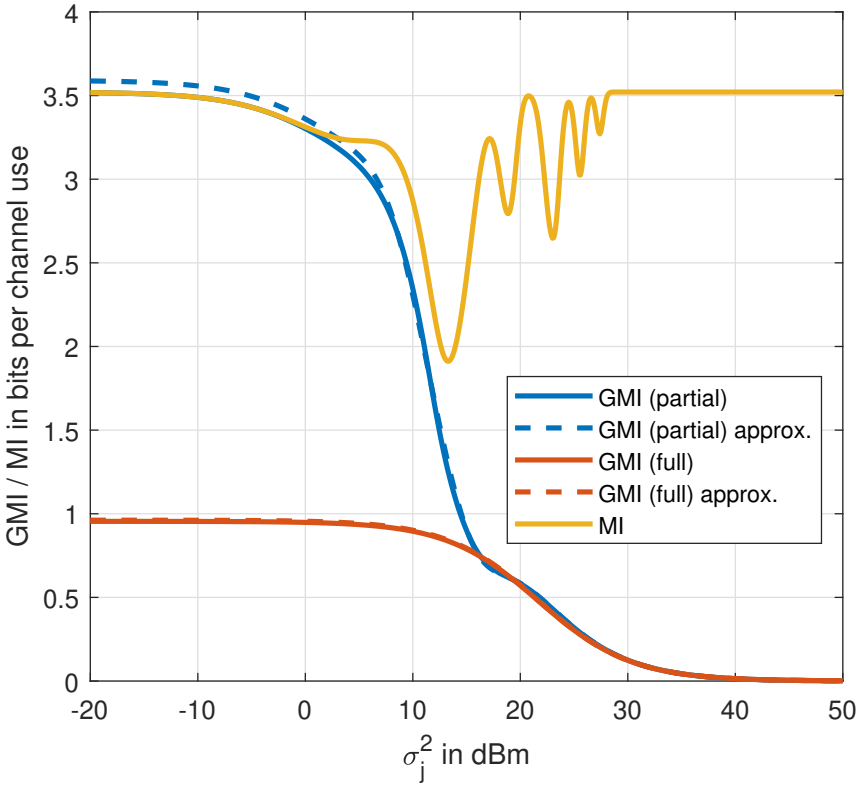


Figure 5.5: GMI and MI with $\sigma_x^2 = 10$, $\sigma_i^2 = 5$ and $\sigma_z^2 = 0.5$.

"GMI (partial)" and "GMI (full)" shrinks and eventually vanishes. This is because, when interference is dominated by J, optimally treating I in decoding leads to limited benefits. As $\sigma_j^2 \to +\infty$, both "GMI (partial)" and "GMI (full)" tend to be zero, whereas "MI" saturates at a high level as stated by Proposition 5.4 with interchanged i and j. Figure 5.5 indicates that treating j as a Gaussian RV is nearly optimal when σ_j^2 is small but clearly suboptimal with large σ_j^2.In Proposition 5, we formally show that the GMI indeed vanishes as σ_j^2 grows.

Proposition 5.5. *With $q(x,\ y)$ given by (5.7),*

$$\lim_{\sigma_j^2 \to +\infty} I_{GMI}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2) = 0. \tag{5.15}$$

Proof. See Appendix D.4. □

Another interesting observation from Figure 5.5 is that "GMI (partial)" underperforms "GMI (full)" in the vicinity of $\sigma_j^2 = 17.5$dBm, indicating that the mismatched decoder given by (5.7) does not always outperform the one given by (5.8) "GMI (full)", though very often.

Figures 5.2-5.5 also depict that (5.11) and (5.12) provide an adequate approximation of (5.10) under $q(x,\ y)$ given by (5.7) and (5.8).

5.2 Improvements on decoding metric

In the spirit of decoding under finite-alphabet interference with reduced computational complexity in comparison to maximum-likelihood decoders, we propose two decoding matrices and evaluate their performance.

5.2.1 Generalized Gaussian decoding metric

In light of improving the similarity between the ground truth interference distribution and the one utilized by the decoding metric, we propose a decoding metric assuming that the real and imaginary parts of $J+Z$ follow i.i.d. generalized Gaussian distributions with their total variance equal to $\sigma_j^2 + \sigma_z^2$. This leads to

$$q(x,\ y) = \sum_{\bar{i}\in\mathcal{I}} \exp\left(-\frac{|\Re\{y-\bar{i}-x\}|^\beta + |\Im\{y-\bar{i}-x\}|^\beta}{\alpha^\beta}\right), \tag{5.16}$$

where β is the shape parameter to be chosen, α is the scale parameter and can be determined based on the variance as $\alpha = \sqrt{\frac{(\sigma_j^2+\sigma_z^2)\Gamma(1/\beta)}{2\Gamma(3/\beta)}}$, with $\Gamma(\cdot)$ representing the Gamma function.[4][5]

Figure 5.6 provides a graphical comparison among a possible ground truth interference plus noise distribution, an approximation given by a Gaussian distribution whose variance is equal to $\sigma_j^2 + \sigma_z^2$, and an approximation given by a generalized Gaussian distribution with the same variance and $\beta = 5.1$. For ease of illustration, the signals are restricted to being real. For clarity, x is drawn from uniform 4-PAM with variance of 15dB, $\mathcal{I} = \{0\}$, j is drawn from uniform 4-PAM with variance of 8dBm, and $\sigma_z^2 = 0$dBm. For the generalized Gaussian distribution, $\beta = 5.1$ is de-

[4]We consider both real and imaginary parts following the same generalized Gaussian distribution. This is sensible if j, or a phase-shifted version of j, has symmetric distributions in real and imaginary parts, for example, QAMs. For asymmetric distributed j (for example, a BPSK signal), it is straightforward to extend (5.16) by assuming asymmetric generalized Gaussian distributions in real and imaginary parts. For example, in Figure 5.6, only the real part is considered to approximate the 4-PAM signal.

[5]The normalization constant of the generalized Gaussian distribution has been ignored in the decoding metric, as it does not affect decoding results.

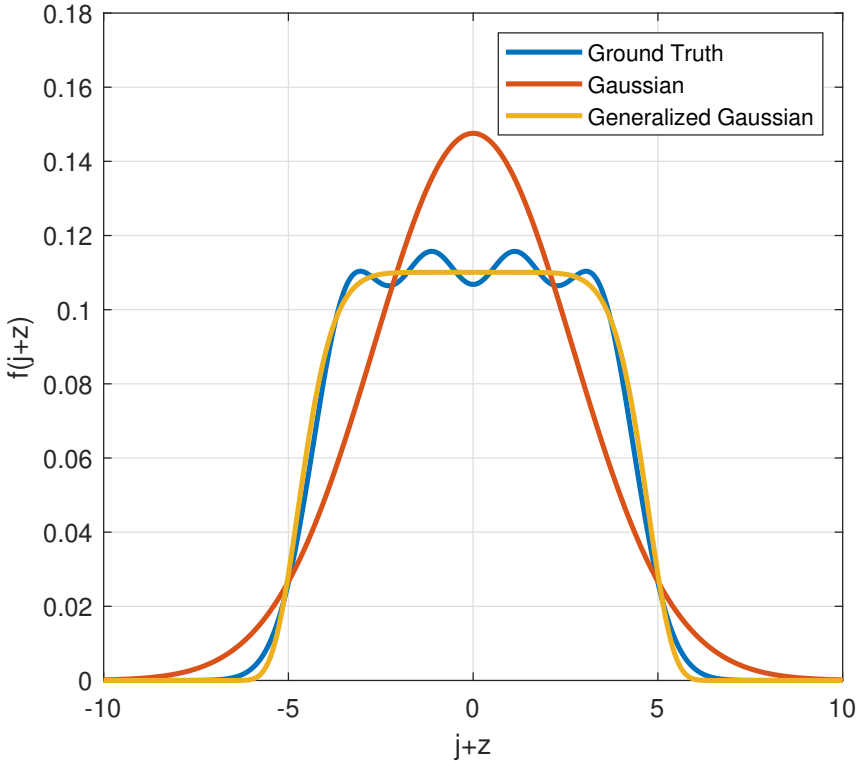


Figure 5.6: Probability density comparison of ground truth interference plus noise distribution by a mixture of Gaussian, approximation by the Gaussian distribution and Generalized Gaussian distribution with $\beta = 5.1$.

termined by optimizing the GMI under the generalized Gaussian decoding metric.[6] It can be observed that, compared to the Gaussian distribution, the generalized Gaussian distribution provides a better approximation of the ground truth distribution, indicating the potential of utilizing the generalized Gaussian distribution in decoding.

5.2.2 Interference decomposition decoding metric

Some constellations can be decomposed into the Minkowski sum of constellations of smaller orders. We propose an interference decomposition decoding metric, which treats one summand constellation optimally and the rest as Gaussian-distributed RVs. Formally, if there exists one decomposition of $\mathcal{J}$ as the following,

$$\mathcal{J} = \mathcal{J}^+ + \mathcal{J}^-, \tag{5.17}$$

one interference decomposition decoding metric can be

$$q(x,\ y) = \sum_{\bar{i}\in\mathcal{I}} \sum_{j^+\in\mathcal{J}^+} \exp\left(-\frac{|y-\bar{i}-j^+-x|^2}{\sigma_{j^-}^2+\sigma_z^2}\right). \tag{5.18}$$

Note that $|\mathcal{J}^+| < |\mathcal{J}|$ if $|\mathcal{J}^-| \geq 2$, therefore, the computational complexity of (5.18) is always lower than (5.6).

Figure 5.7 provides a graphical comparison among a possible ground truth interference plus noise distribution, a Gaussian approximation, and two approximations given by interference decomposition. The assumptions for interference and noise are the same as in Figure 5.6. For interference decomposition, a uniform 4-PAM can be decomposed into two BPSK constellations, where the powers of the two summand constellations are 0.8 and 0.2 of the power of the 4-PAM, respectively. By optimally treating either one of the two summand constellations, two different underlying assumptions on interference-plus-noise distributions are obtained and depicted in Figure 5.7, where "Int. Summand 1" representing $\mathcal{J}^+$ being a BPSK with $\sigma_{j^+}^2 = 0.8\sigma_j^2$, and "Int. Summand 2" representing $\sigma_{j^+}^2 = 0.2\sigma_j^2$. It can be observed that optimally treating the stronger summand leads to improved approximation on the interference plus noise distribution in comparison with Gaussian approximation, while optimally treating the weaker summand leads to minor changes in the resulting

[6] In the simulations, β is determined by exhaustive search to maximize the GMI. In practice, a lookup table for β may also be constructed.

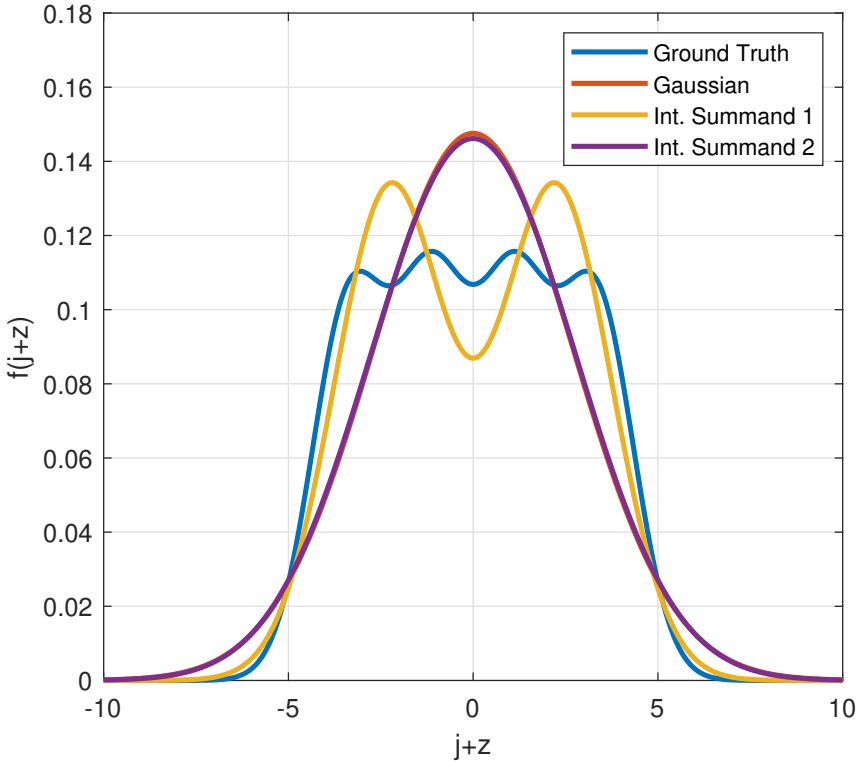


Figure 5.7: Probability density comparison of ground truth interference plus noise distribution by a mixture of Gaussian, an approximation by a Gaussian distribution and two approximations by interference decomposition.

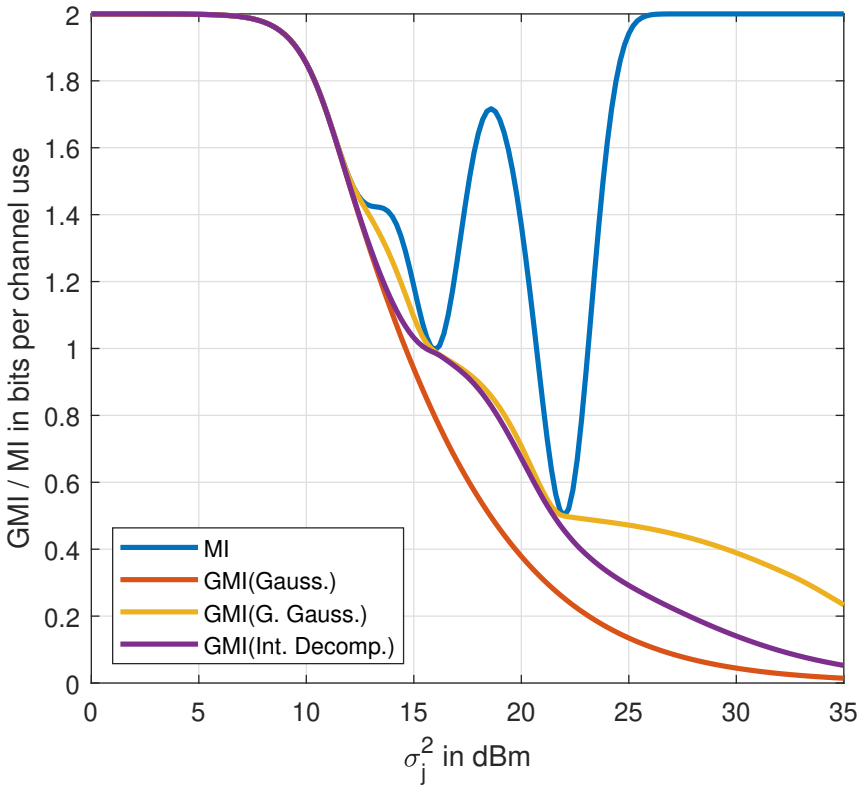


Figure 5.8: GMI and MI with $\sigma_x^2 = 15$dBm and $\sigma_z^2 = 0$dBm.

distribution.

5.2.3 Results on GMI of the proposed decoding metrics

Consider x and j are generated by QPSK and 16QAM respectively, whereas $\mathcal{I} = \{0\}$ for clarity. From Figure 5.8, it can be observed that both the generalized Gaussian decoding metric and the interference decomposition decoding metric lead to improved achievable rates compared to treating interference as a Gaussian RV. Often, the generalized Gaussian decoding metric outperforms the interference decomposition decoding metric. However, such superiority comes with an additional computational cost in determining the optimal shape parameter, β.

5.3 Practical receivers

To shine more light on the designs of the physical layer receivers, the decoding metrics discussed above can be converted into demodulators in BICM-type systems.

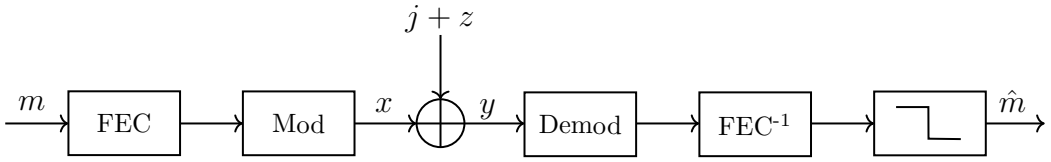


Figure 5.9: Transceiver structure.

Consider the physical-layer transceiver architecture shown in Figure 5.9, where, for clarity, only one interference source, j, is considered. In Figure 5.9, "FEC", "Mod", "Demod" and "FEC^{-1}" denote channel encoder,[7] modulator, demodulator and channel decoder, respectively. The demodulator takes the received signal, y, as input and estimates the posterior probability distribution of the channel input signals, $\hat{P}_{X|Y}(x \mid y)$, $x \in \mathcal{X}$, based on its knowledge of the input signal, interference and noise. The estimated posterior input distribution can then be converted to log-likelihood ratios of coded bits, which feed the channel decoder. The decoding metrics mentioned above for the FAGCI model make different assumptions about the interference distribution. The same principle can be applied to the design of demodulators.

The optimal decoding metric (5.6), which applies the true interference distribution, corresponds to the optimal demodulator, which jointly demodulates x and j as follows.[8]

$$\hat{P}_{X|Y}(x \mid y) = \frac{1}{2\pi\sigma_z^2|\mathcal{J}|} \sum_{\bar{j}\in\mathcal{J}} \exp\left(-\frac{|y-\bar{j}-x|^2}{\sigma_z^2}\right) \tag{5.19}$$

The suboptimal decoding metric (5.8), which treats interference as Gaussian random variables, corresponds to a demodulator that treats interference as Gaussian noise as follows.

$$\hat{P}_{X|Y}(x \mid y) = \frac{1}{2\pi(\sigma_j^2+\sigma_z^2)} \exp\left(-\frac{|y-x|^2}{\sigma_j^2+\sigma_z^2}\right) \tag{5.20}$$

The generalized Gaussian decoding metric given in (5.16) can be linked to a demodulator that treats the sum of interference and noise as a generalized Gaussian-distributed random variable, which can be written as

$$\hat{P}_{X|Y}(x \mid y) = \left(\frac{\beta}{2\alpha\Gamma(1/\beta)}\right)^2 \exp\left(-\frac{|\Re\{y-x\}|^\beta + |\Im\{y-x\}|^\beta}{\alpha^\beta}\right), \tag{5.21}$$

with $\alpha = \sqrt{\frac{(\sigma_j^2+\sigma_z^2)\Gamma(1/\beta)}{2\Gamma(3/\beta)}}$. Finally, inspired by the interference decomposition decoder

[7] This potentially includes interleaving and scrambling.

[8] The constant factor $P_X(x)/P_Y(y)$ has been neglected from the approximated posterior distributions, making them also approximated likelihood functions.

in (5.18), an interference decomposition demodulator can be designed to jointly demodulate the input signal and a summand of the interference constellation, while treating the other summand of the interference as a Gaussian RV. This can be written as

$$\hat{P}_{X|Y}(x \mid y) = \frac{1}{2\pi(\sigma_{j^-}^2 + \sigma_z^2)|\mathcal{J}^+|} \sum_{j^+ \in \mathcal{J}^+} \exp\left(-\frac{|y - j^+ - x|^2}{\sigma_{j^-}^2 + \sigma_z^2}\right). \tag{5.22}$$

The similarity between the channel model in Section 5.1 and Figure 5.9 also implies that the derived GMI in Section 5.1 may resemble the throughput of practical transceivers with the corresponding demodulator.

We examine the throughput of different decoding metrics when they are used as demodulators and compare it with the GMI led by different decoding metrics in Figures 5.10 and 5.11. The legends are explained as follows:

- "JD" (MI) and (LLS): optimal decoding, i.e., (5.6) evaluated by MI, and demodulator jointly demodulating x and j, i.e., (5.19), evaluated by LLS.
- "ID" (GMI) and (LLS): interference decomposition decoder, i.e., (5.18) evaluated by GMI, and interference decomposition demodulator, i.e., (5.22) evaluated by LLS.
- "IN" (GMI) and (LLS): decoder that treats interference as a Gaussian RV, i.e., (5.8), evaluated by GMI, and demodulator that treats interference as a Gaussian RV, i.e., (5.20), evaluated by LLS.
- "G Gauss" (GMI) and (LLS): the generalized Gaussian decoding metric, i.e., (5.16), evaluated by GMI, and the generalized Gaussian demodulator, i.e., (5.21), evaluated by LLS.

The channel code used in LLS is the LDPC code specified in [99],[9] with the block length being 4000 symbols. The throughput is determined by searching for the highest code rate that leads to a block error rate lower than 10^{-2}. In Figure 5.10 and 5.11, it can be observed that the proposed demodulators, i.e., (5.21) and (5.22), lead to improved throughput compared to that given by (5.20). In addition, GMI/MI provides adequate prediction on the physical layer throughput when the demodulator

[9] [99] specifies the LDPC code for code rate between 0.2 and 0.92. Hence, in Figure 5.10 and 5.11, the code rate suddenly increases from 0 to 0.2 with medium σ_x^2, and is capped at 0.92 at high σ_x^2.

used in the LLS corresponds to the decoding metric assumed by GMI/MI. The gap between GMI/MI and throughput is due to the non-ideal encoding and decoding.

These results lead to two key implications:

1. GMI/MI provides alternative tools to LLS for predicting demodulator performance.[10] The computational complexity for computing GMI/MI is much lower than that of running LLS, making the former promising. The approximations by GMI/MI also provide adequate prediction with even less computational burden.

2. GMI/MI can be important for accurate physical layer abstraction in SLS. Optimizing GMI/MI indirectly optimizes the physical layer throughput under the corresponding demodulators.

The second implication motivates the next section, where an example of GMI optimization for a multi-user MISO system is explored.

5.4 Signal processing with GMI

In multi-user communication systems, it is common that multi-user interference is not optimally treated due to reasons such as limited receiver complexity and lack of knowledge on interference. The channel model, the mismatched decoder, the corresponding GMI and its approximation from Section 5.1 provide novel tools for analyzing such scenarios and designing signal processing methods. In this section, we extend the results from the SISO FAGCI model to multi-user MISO systems to enable novel signal processing algorithms. In particular, we provide an example on precoder optimization with the objective of maximizing sum-rates under different decoding strategies adapted by users.

[10]Note that another approach for demodulator evaluation is proposed in [87] where interference is not considered, and the resulting expression is not in closed form, hence not viable for system design.

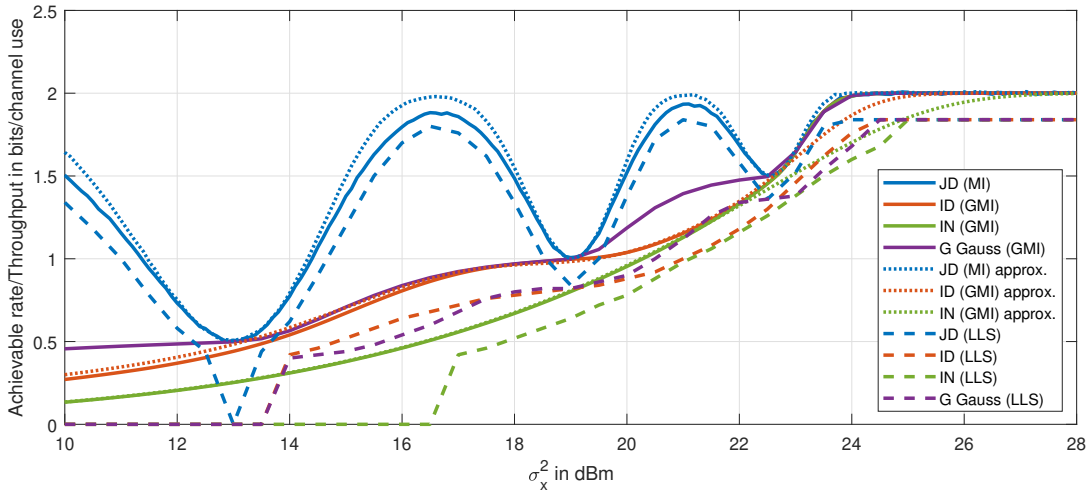


Figure 5.10: Comparison of GMI/MI and LLS results with $\mathcal{X}$ being QPSK, $\mathcal{J}$ being 16QAM, $\sigma_j^2 = 20$dBm, $\sigma_z^2 = 0$dBm.

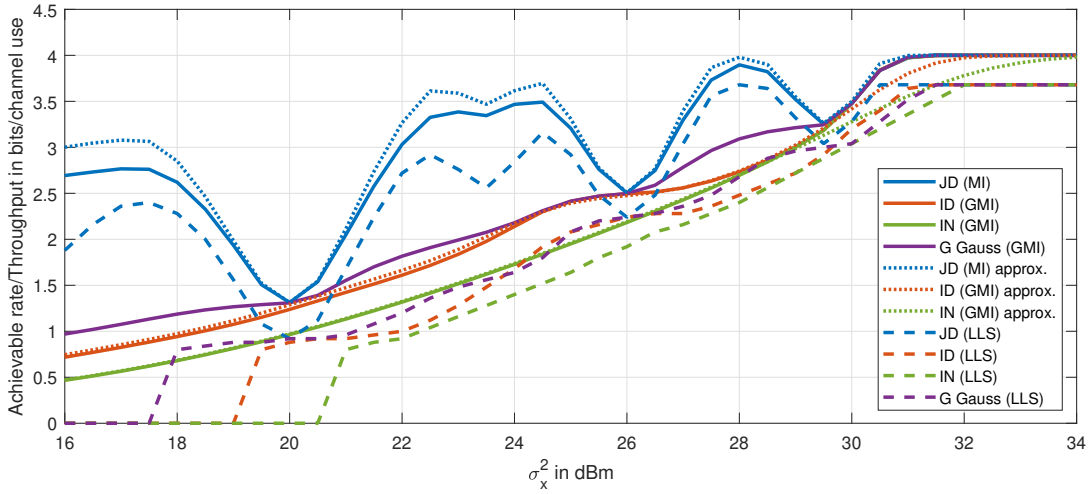


Figure 5.11: Comparison of GMI/MI and LLS results with $\mathcal{X}$ being 16QAM, $\mathcal{J}$ being 16QAM, $\sigma_j^2 = 20$dBm, $\sigma_z^2 = 0$dBm.

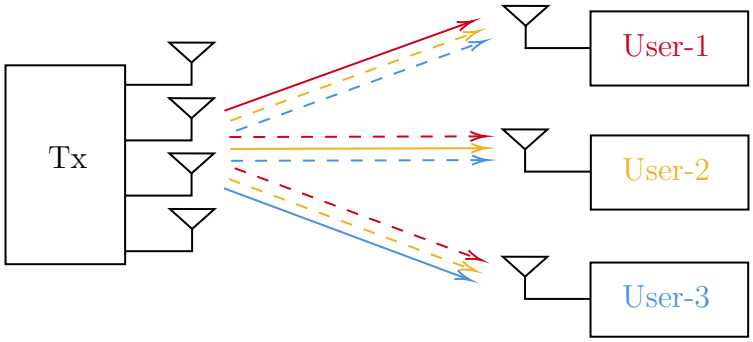


Figure 5.12: Illustration of a MISO broadcast channel, where solid arrows denote desired signals and dashed arrows denote interfering signals. The three users potentially suffer from multi-user interference.

5.4.1 Achievable rate of MU-MISO incorporating decoding strategy

Consider a downlink MU-MISO system with N_{T} antennas at the transmitter and K single-antenna users as illustrated in Figure 5.12, where the transmitted signal can be written as

$$\mathbf{x} = \mathbf{P}\mathbf{s}, \tag{5.23}$$

where $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_K] \in \mathcal{C}^{N_{\mathrm{T}} \times K}$ is the precoding matrix, and $\mathbf{s} = [s_1,\ s_2,\ ...,\ s_K] \in \mathcal{C}^{K \times 1}$ is the symbol vector to be transmitted. User-k desires s_k and generally suffers from multi-user interference and noise. For simplicity, we assume that all users adopt the same constellation, i.e., $s_k \in \mathcal{X}$, $\forall k \in \{1, ..., k\}$.

Given a decoding strategy, i.e., deciding which interfering streams to be treated optimally and as Gaussian RV at each user, the achievable rate at each user is given by extending the GMI in (5.10). Consider a generic expression for the received signal at user-k, $\forall k \in \{1, ..., K\}$, as follows:

$$y_k = \mathbf{h}_k^H(\mathbf{P}_x\mathbf{x} + \mathbf{P}_i\mathbf{i} + \mathbf{P}_j\mathbf{j}) + z_k, \tag{5.24}$$

where $\mathbf{h}_k$ is the channel vector at user-k, and $\mathbf{x}$, $\mathbf{i}$ and $\mathbf{j}$ are respectively the desired symbol vector for user-k, interfering symbol vector to be treated optimally and interfering symbol vector to be treated as Gaussian RVs by user-k, respectively drawn from $\mathcal{X}$, $\mathcal{I}$ and $\mathcal{J}$. $\mathbf{P}_x$, $\mathbf{P}_i$ and $\mathbf{P}_j$ are their precoders. z_k is additive noise generated from i.i.d. $\mathcal{CN}(0, \sigma_z^2)$. Without loss of generality, we assume that $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}^{\dim(\mathbf{x})}$, $\mathbb{E}[\mathbf{i}\mathbf{i}^H] = \mathbf{I}^{\dim(\mathbf{i})}$ and $\mathbb{E}[\mathbf{j}\mathbf{j}^H] = \mathbf{I}^{\dim(\mathbf{j})}$ unless any of the constellations is $\{0\}$. WLOG,

user-k utilizes a decoding metric given by

$$q(\mathbf{x},\ y_k) = \sum_{\bar{\mathbf{i}} \in \mathcal{I}} \exp\left(- \frac{|y_k - \mathbf{h}_k^H \mathbf{P}_i \bar{\mathbf{i}} - \mathbf{h}_k^H \mathbf{P}_x \mathbf{x}|^2}{\|\mathbf{h}_k^H \mathbf{P}_j\|^2 + \sigma_z^2} \right). \tag{5.25}$$

$$
\begin{aligned}
&I_{\mathrm{GMI},k}\left(\mathbf{h}_k^H\mathbf{P}_x\mathcal{X},\ \mathbf{h}_k^H\mathbf{P}_i\mathcal{I},\ \mathbf{h}_k^H\mathbf{P}_j\mathcal{J},\ \sigma_z^2\right)\\
=&\sup_{s\geq 0}\ \log|\mathcal{X}| - \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{\mathbf{x}\in\mathcal{X},\,\mathbf{i}\in\mathcal{I},\,\mathbf{j}\in\mathcal{J}}\mathbb{E}_Z\left[\log\sum_{\bar{\mathbf{x}}\in\mathcal{X}}\left(\sum_{\bar{\mathbf{i}}\in\mathcal{I}}\exp\left(-\frac{|\mathbf{h}_k^H\mathbf{P}_x(\mathbf{x}-\bar{\mathbf{x}})+\mathbf{h}_k^H\mathbf{P}_\mathbf{I}(\mathbf{i}-\bar{\mathbf{i}})+\mathbf{h}_k^H\mathbf{P}_j\mathbf{j}+Z|^2}{\|\mathbf{h}_k^H\mathbf{P}_j\|^2+\sigma_z^2}\right)\right)^s\right]\\
&+\frac{s}{|\mathcal{I}||\mathcal{J}|}\sum_{\mathbf{i}\in\mathcal{I},\,\mathbf{j}\in\mathcal{J}}\mathbb{E}_Z\left[\log\left(\sum_{\bar{i}\in\mathcal{I}}\exp\left(-\frac{|\mathbf{h}_k^H\mathbf{P}_i(\mathbf{i}-\bar{\mathbf{i}})+\mathbf{h}_k^H\mathbf{P}_j\mathbf{j}+Z|^2}{\|\mathbf{h}_k^H\mathbf{P}_j\|^2+\sigma_z^2}\right)\right)\right]
\end{aligned}
\tag{5.26}
$$

$$
\begin{aligned}
&I^{\mathrm{approx}}_{\mathrm{GMI},k}\left(\mathbf{h}_k^H\mathbf{P}_x\mathcal{X},\ \mathbf{h}_k^H\mathbf{P}_i\mathcal{I},\ \mathbf{h}_k^H\mathbf{P}_j\mathcal{J},\ \sigma_z^2\right)\\
=&\log|\mathcal{X}| - \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{\mathbf{x}\in\mathcal{X},\,\mathbf{i}\in\mathcal{I},\,\mathbf{j}\in\mathcal{J}}\log\sum_{\bar{\mathbf{x}}\in\mathcal{X},\,\bar{\mathbf{i}}\in\mathcal{I}}\exp\left(-\frac{|\mathbf{h}_k^H\mathbf{P}_x(\mathbf{x}-\bar{\mathbf{x}})+\mathbf{h}_k^H\mathbf{P}_i(\mathbf{i}-\bar{\mathbf{i}})+\mathbf{h}_k^H\mathbf{P}_j\mathbf{j}|^2}{\|\mathbf{h}_k^H\mathbf{P}_j\|^2+2\sigma_z^2}\right)\\
&+\frac{1}{|\mathcal{I}||\mathcal{J}|}\sum_{\mathbf{i}\in\mathcal{I},\,\mathbf{j}\in\mathcal{J}}\log\sum_{\bar{\mathbf{i}}\in\mathcal{I}}\exp\left(-\frac{|\mathbf{h}_k^H\mathbf{P}_i(\mathbf{i}-\bar{\mathbf{i}})+\mathbf{h}_k^H\mathbf{P}_j\mathbf{j}|^2}{\|\mathbf{h}_k^H\mathbf{P}_j\|^2+2\sigma_z^2}\right)
\end{aligned}
\tag{5.27}
$$

$$
\begin{aligned}
&\nabla_{\widetilde{\mathbf{P}}_k} I^{\mathrm{approx}}_{\mathrm{GMI},k}\left(\mathbf{h}_k^H\mathbf{P}_x\mathcal{X},\ \mathbf{h}_k^H\mathbf{P}_i\mathcal{I},\ \mathbf{h}_k^H\mathbf{P}_j\mathcal{J},\ \sigma_z^2\right)\\
=&-\frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{\mathbf{x}\in\mathcal{X},\,\mathbf{i}\in\mathcal{I},\,\mathbf{j}\in\mathcal{J}}\frac{1}{\sum_{\bar{\mathbf{x}}\in\mathcal{X},\,\bar{\mathbf{i}}\in\mathcal{I}}g(\widetilde{\mathbf{P}}_k,\mathbf{v})}\sum_{\bar{\mathbf{x}}\in\mathcal{X},\,\bar{\mathbf{i}}\in\mathcal{I}}g_1(\widetilde{\mathbf{P}}_k,\mathbf{v})+\frac{1}{|\mathcal{I}||\mathcal{J}|}\sum_{\mathbf{i}\in\mathcal{I},\,\mathbf{j}\in\mathcal{J}}\frac{1}{\sum_{\bar{\mathbf{i}}\in\mathcal{I}}h(\widetilde{\mathbf{P}}_k,\mathbf{w})}\sum_{\bar{\mathbf{i}}\in\mathcal{I}}h_1(\widetilde{\mathbf{P}}_k,\mathbf{w})
\end{aligned}
\tag{5.28}
$$

$$
h_1(\widetilde{\mathbf{P}}_k,\mathbf{w}) = h(\widetilde{\mathbf{P}}_k,\mathbf{w})\left(\frac{-\left(\|\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\,\mathbf{I}_j\|^2+2\sigma_z^2\right)\mathbf{h}_k\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\mathbf{w}\mathbf{w}^H+(\mathbf{w}^H\widetilde{\mathbf{P}}_k^H\mathbf{h}_k\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\mathbf{w})\mathbf{h}_k\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\,\mathbf{I}_j\,\mathbf{I}_j^H}{\left(\|\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\,\mathbf{I}_j\|^2+2\sigma_z^2\right)^2}\right)
\tag{5.29}
$$

$$
h_1(\widetilde{\mathbf{P}}_k,\mathbf{w}) = h(\widetilde{\mathbf{P}}_k,\mathbf{w})\left(\frac{-\left(\|\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\,\mathbf{I}_j\|^2+2\sigma_z^2\right)\mathbf{h}_k\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\mathbf{w}\mathbf{w}^H+(\mathbf{w}^H\widetilde{\mathbf{P}}_k^H\mathbf{h}_k\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\mathbf{w})\mathbf{h}_k\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\,\mathbf{I}_j\,\mathbf{I}_j^H}{\left(\|\mathbf{h}_k^H\widetilde{\mathbf{P}}_k\,\mathbf{I}_j\|^2+2\sigma_z^2\right)^2}\right)
\tag{5.30}
$$

$$g(\widetilde{\mathbf{P}}_k, \mathbf{v}) = \exp\left(-\frac{|\mathbf{h}_k^H \widetilde{\mathbf{P}}_k \mathbf{v}|^2}{\|\mathbf{h}_k^H \widetilde{\mathbf{P}}_k \, \mathbf{I}_j\|^2 + 2\sigma_z^2}\right) \tag{5.31}$$

$$h(\widetilde{\mathbf{P}}_k, \mathbf{w}) = \exp\left(-\frac{|\mathbf{h}_k^H \widetilde{\mathbf{P}}_k \mathbf{w}|^2}{\|\mathbf{h}_k^H \widetilde{\mathbf{P}}_k \, \mathbf{I}_j\|^2 + 2\sigma_z^2}\right) \tag{5.32}$$

By extending (5.10) and (5.11), the GMI at user-k can be expressed as (5.26) and approximated by (5.27). To reduce the computational complexity, we focus on (5.27) during signal processing, but the results will be evaluated by (5.26). To enable optimization of GMI, the gradient of GMI w.r.t. precoders is given by (5.28), where $g(\widetilde{\mathbf{P}}_k, \mathbf{v})$, $h(\widetilde{\mathbf{P}}_k, \mathbf{v})$, $g_1(\widetilde{\mathbf{P}}_k, \mathbf{v})$ and $h_1(\widetilde{\mathbf{P}}_k, \mathbf{v})$ are given by (5.29-5.31), and

$$\widetilde{\mathbf{P}}_k = [\mathbf{P}_x \; \mathbf{P}_i \; \mathbf{P}_j], \tag{5.33}$$

$$\mathbf{v} = [\mathbf{x} - \bar{\mathbf{x}};\; \mathbf{i} - \bar{\mathbf{i}};\; \mathbf{j}], \tag{5.34}$$

$$\mathbf{w} = [\mathbf{0}^{\dim(\mathcal{X})\times 1};\; \mathbf{i} - \bar{\mathbf{i}};\; \mathbf{j}], \tag{5.35}$$

$$\mathbf{I}_j = [\mathbf{0}^{\dim(\mathcal{X})\times\dim(\mathcal{J})};\; \mathbf{0}^{\dim(\mathcal{I})\times\dim(\mathcal{J})};\; \mathbf{I}^{\dim(\mathcal{J})}]. \tag{5.36}$$

Note that $\widetilde{\mathbf{P}}_k$ is a column-wise permutation of the precoding matrix at the transmitter, $\mathbf{P}$, based on user-k's decoding strategy.

5.4.2 Sum-rate optimization

We consider a sum-rate maximization problem under fixed transmit power budget as follows:

$$\begin{aligned} \max_{\mathbf{P}} \;& \sum_{k\in\mathcal{K}} I_{\mathrm{GMI},k}^{\mathrm{approx}} \\ \text{s.t.} \;& \|\mathbf{P}\|_{\mathrm{F}}^2 \le P_{\mathrm{T}}, \end{aligned} \tag{5.37}$$

where P_{T} is the transmit power budget. Note that the approximated GMI by (5.27) is used to reduce computational complexity. The logarithmic barrier [100] is used to approximate (5.37) into the following unconstrained optimization problem

$$\max_{\mathbf{P}} \; \Omega(\mathbf{P}) \triangleq \tau \sum_{k\in\mathcal{K}} I_{\mathrm{GMI},k}^{\mathrm{approx}} + \log\left(P_{\mathrm{T}} - \|\mathbf{P}\|_{\mathrm{F}}^2\right), \tag{5.38}$$

where τ is a positive parameter that sets the accuracy of the approximation. (5.38) can be solved by gradient ascent as given in Algorithm 7, where $\nabla_{\mathbf{P}} I_{\mathrm{sum}}^{\mathrm{approx}}(\mathbf{P})$ is

Algorithm 7 Gradient Ascent with Barrier Method for Sum-Rate Maximization

Input: P_{T}, $\mathcal{X}_k$, $\mathbf{h}_k$, $k \in \mathcal{K}$, β, ϵ, $v_{\max}$, $\tau_{\max}$.
Output: $\mathbf{P}^\star$

Initialize $\mathbf{P}^0$ randomly and $\tau := 1$.
repeat
  $v := 0$.
  repeat
    $\Delta\mathbf{P} := \tau \nabla_{\mathbf{P}} I_{\text{sum}}^{\text{approx}}(\mathbf{P}^v) + \frac{\mathbf{P}^v}{\|\mathbf{P}^v\|_{\mathrm{F}}^2 - P_{\mathrm{T}}}$.
    Choose step size α via backtracking line search.
    Update. $\mathbf{P}^{v+1} := \mathbf{P}^v + \alpha\Delta\mathbf{P}$.
    Update. $\Omega^{v+1} := \Omega(\mathbf{P}^{v+1})$.
    $v := v + 1$.
  until $|\Omega^v - \Omega^{v-1}| < \epsilon$ or $v \geq v_{\max}$.
  Update $\tau := \beta\tau$.
until $\tau \geq \tau_{\max}$.
return $\mathbf{P}^\star := \mathbf{P}^v$.

computed by adding the GMI gradient at all users given by (5.28). The convergence of Algorithm 7 is ensured by employing a backtracking line search to determine the step sizes.

5.4.3 Numerical simulations

We consider the geometrical one-ring scattering model (See Appendix A) to model correlated user channels where interference may hinder performance. The settings are $\theta_k = \pi/3$ and $\Delta_k = \pi/9$, $\forall k \in \{1, ..., K\}$.

With $N_{\mathrm{T}} = 4$, $K = 3$ and $\mathcal{X}$ given by QPSK, we evaluate the optimal sum-rate of the system under different decoding strategies deployed by the users. The results are depicted in Figure 5.13, where the legends and their corresponding decoding strategies are as follows.

1. "MI" indicates that the optimization objective is the mutual information, whose underlying assumption is that all users apply the optimal decoding metric, which treats the other two users' interference optimally;

2. "GMI (partial)" implies optimizing a GMI which assumes that each user treats interference from one of the other users optimally and the interference from the third as a Gaussian RV. In particular, we assume that user-1 treats user-2's interference optimally, user-2 treats user-3's interference optimally, and user-3 treats user-1's interference optimally;

3. "GMI (full)" indicates that the objective function is the GMI expression, which assumes that each user treats all multi-user interference as a Gaussian RV.

From Figure 5.13, it can be observed that different decoding metrics lead to different sum-rates in general. The order is always "MI" > "GMI (partial)" > "GMI (full)", which aligns with the intuition that utilizing more knowledge on interference leads to better decoding, and therefore better performance. Figure 5.14 evaluates the "GMI (full)" but with precoders obtained by optimizing different objective functions as in Figure 5.13. It shows that not only the GMI optimization solution for one decoding metric is different from other decoding metrics, but also the mismatch in the objective function can even be harmful, noticing the non-monotonic change in sum-rate under discrepancy between the decoding metric and the objective function. This indicates that mismatched receivers require dedicated signal processing approaches based on the corresponding GMI.

To gain further insight into how the precoders are designed differently under various decoding metrics, Figure 5.15 depicts the ergodic inter-user interference power received by user-1 from the other two users for different decoding metrics. The results show that the optimized solutions allow interference that is optimally treated to remain strong while keeping suboptimally treated interference weak. This observation is consistent with the results in Section 5.1, which indicate that strong interference is not detrimental when optimally treated, but becomes harmful when treated suboptimally.

Figure 5.16 depicts the convergence of the proposed algorithm through the approximated sum-rate obtained at each inner-loop iteration in Algorithm 7.

5.4.4 Future directions

The above example pioneers a novel system design framework that considers the receiver's imperfection in treating finite-alphabet interference, which can be applied to various communication scenarios where interference is critical. Examples of such scenarios include, but are not limited to:

- MIMO receivers implemented by linear combiners and single stream decoders, where inter-stream interference may exist during decoding;

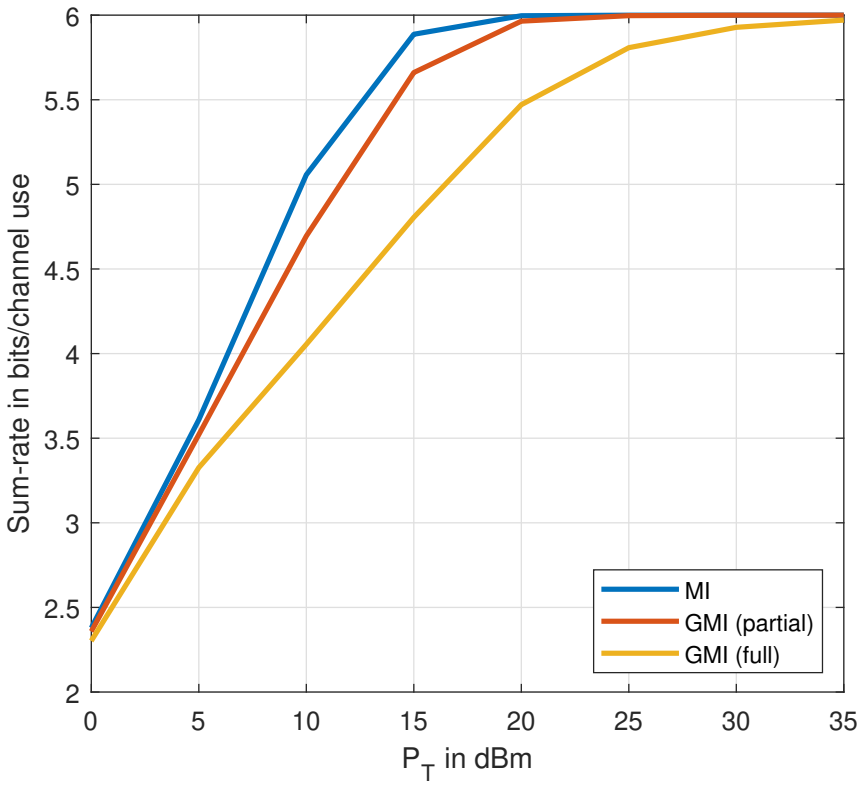


Figure 5.13: Sum-rate comparison of MU-MISO under different decoding strategies.

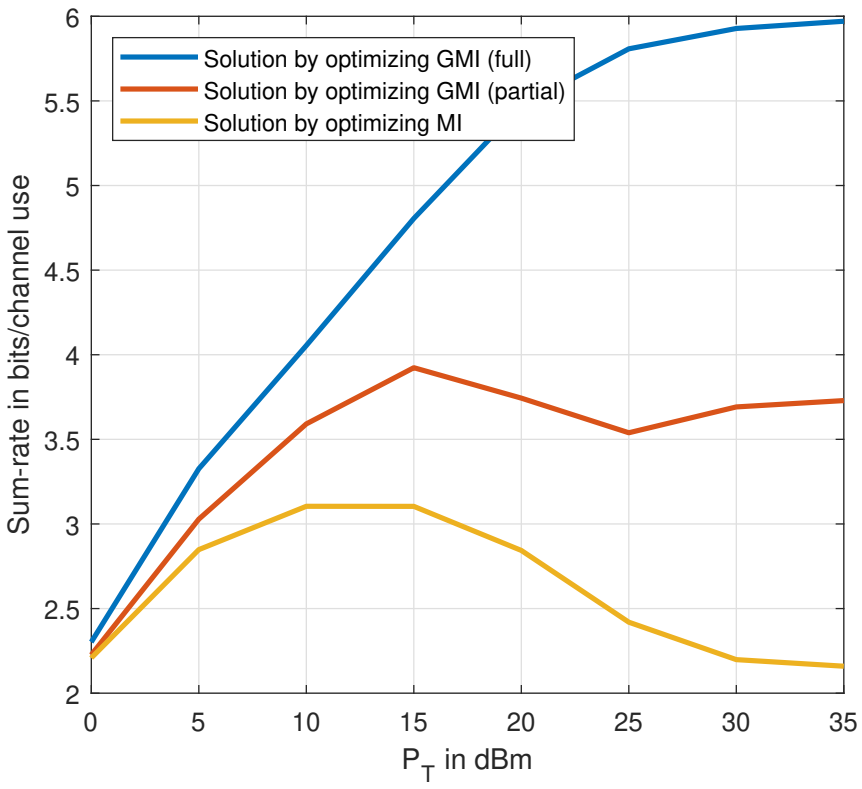


Figure 5.14: Sum-rate comparison of MU-MISO, with decoders treating interference as a Gaussian RV and precoders by optimizing different objective functions.

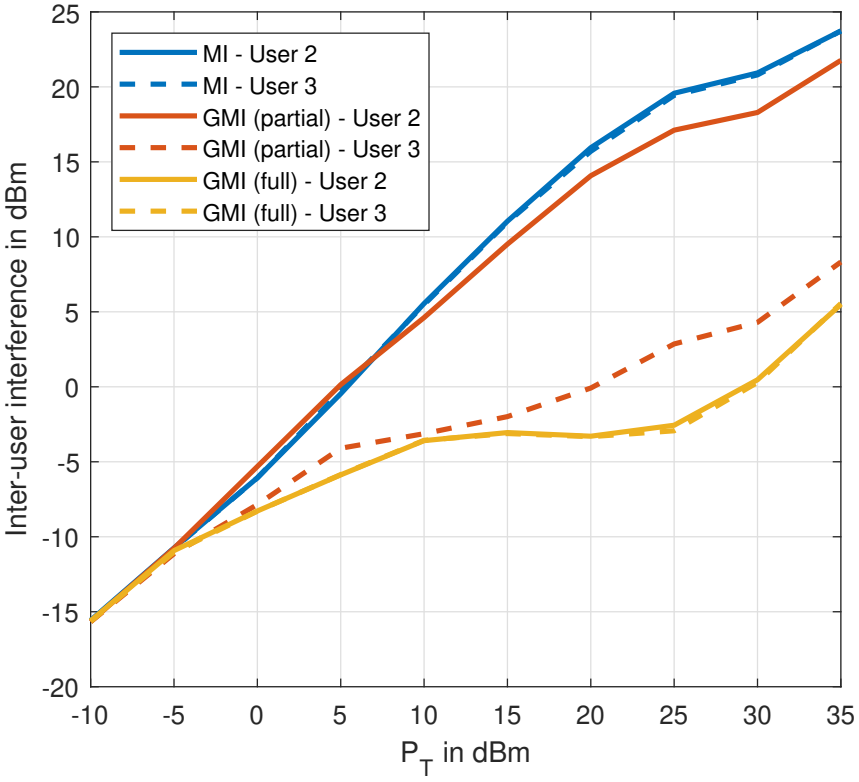


Figure 5.15: Ergodic inter-user interference power received at User 1 from the other two users under different objective functions.

- Massive random access channels, for example, massive (machine-type) communications in 6G and internet of things, where collision in transmission may happen.
- Superposition/non-orthorgonal transmission schemes, where the receivers need to decode under interference brought by the nature of the schemes.
- Support for Unmanned Aerial Vehicles (UAV) from cellular networks, where uplink/downklink interference between base stations and UAVs needs to be resolved [95].

5.5 Summary

In this chapter, we aim at understanding the impact of mismatched/suboptimal decoders under finite-alphabet interference. We first defined the FAGCI model, which takes a finite-alphabet input and adds finite-alphabet interference and Gaussian noise during transmission. We initially considered the optimal decoding metric and two suboptimal decoding metrics, where the suboptimal decoding metrics treat

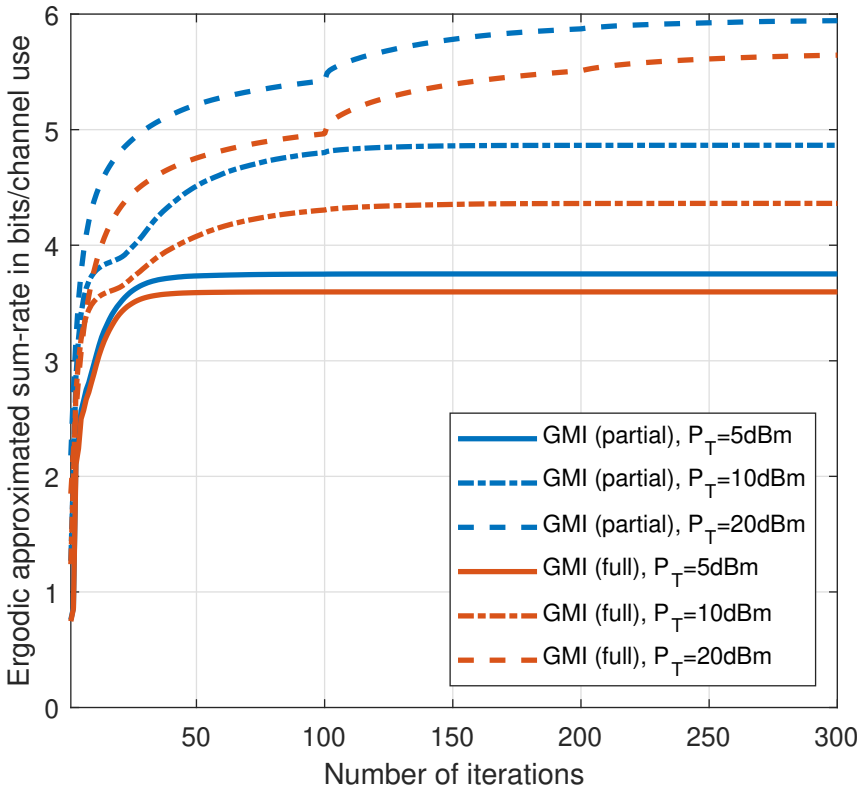


Figure 5.16: Evolution of the ergodic approximated sum-rate over successive iterations of Algorithm 7.

part or all of the interference as a Gaussian RV. Achievable rates under different decoding metrics were given by the derived MI and GMI. To improve the achievable rates of FAGCI channel while maintaining low decoding complexity, we further proposed two novel decoding metrics, namely the generalized Gaussian decoding metric and the interference decomposition decoding metric, whose superiority was validated through numerical evaluation on their corresponding GMI. Noting the connection between decoding metrics and demodulators in BICM systems, we also applied the two proposed decoding metrics as demodulators. LLS show that the two novel demodulators also lead to improved throughput. More interestingly, it was also observed that there is a strong correlation between GMI/MI and throughput when the decoding metric corresponds to the demodulator. This indicates that GMI/MI can be used as an efficient tool for evaluating demodulators, and optimizing GMI/MI improves the physical layer throughput under the corresponding demodulator. Lastly, we extended the FAGCI model and the GMI to a multiple-antenna case, and provided an example on MU-MISO precoder optimization, where the objective function is given by MI or GMI, depending on the decoding strategy deployed by the users. The numerical results showed that different decoding strategies led to different throughput and that suboptimal decoders required a dedicated optimization

algorithm. Overall, this chapter provides novel theoretical tools, receiver designs, a signal processing framework and insights that can potentially enhance interference management in future wireless networks.

Chapter 6

Codeword-Segmentation Rate-Splitting Multiple Access

SDMA is currently the predominant spatial domain technique in use. Despite the numerous benefits of RSMA from theoretical analysis, replacing SDMA with RSMA in practical systems will introduce several challenges. Although the SIC-free receiver designs proposed in Chapter 3 addressed the increase in receiver complexity, other challenges remain. These include:

- Increased encoding/decoding complexity: with RSMA, more codewords are transmitted (especially with multi-layer schemes such as the hierarchical RSMA in [18] and the general framework of multi-layer RSMA in [16]), which leads to more complexity in encoding and decoding.

- Control signaling overhead: The increased number of streams requires an increased control signaling overhead to assist in decoding at receivers.

- Retransmission: As the common stream is intended for multiple users, it requires retransmission mechanisms that are different from SDMA, such as in [101]. The design of an efficient retransmission process for RSMA schemes poses a challenge.

In this chapter, a novel downlink RSMA architecture is proposed, namely Codeword-Segmentation RSMA (CS-RSMA), to reduce the implementation challenges of RSMA listed above.

Another direction that deserves exploration is to evaluate and design RSMA schemes using performance metrics that account for more practical limitations. The benefit of RSMA has been well studied under performance metrics that originate from the mutual information, but is less considered under suboptimal receivers. In Chapter 5, we studied communication in finite-alphabet input channels under Gaussian noise and finite-alphabet interference, namely FAGCI, where the decoders fully or partially approximate the interference as Gaussian random variables to reduce computational complexity. This model reflects a common issue that interference is not properly detected in receivers due to complexity reasons. In this chapter, we will consider a novel performance analysis framework using the GMI developed in Chapter 5 to evaluate conventional RSMA and the proposed CS-RSMA.

The contribution of this chapter is summarized as follows:

- We propose a novel downlink architecture for RSMA, namely CS-RSMA. Unlike conventional RSMA, which splits users' messages into common and private parts, then combines the common parts into a common message before encoding, CS-RSMA encodes users' messages directly, segments the resulting codewords into common and private parts, and combines the common parts into a common stream. Details on encoding and decoding procedures for both conventional RSMA and CS-RSMA are provided.

- We apply the GMI derived in Chapter 5 to analyze the achievable rate of CS-RSMA and conventional RSMA to capture receiver imperfection in practice. This leads to a different analysis from Chapter 4, where the achievable rate is measured by mutual information, which is often achieved by decoders that optimally exploit channel knowledge, e.g., the maximum likelihood decoder and typical-set decoder. With the GMI, conventional RSMA and CS-RSMA can be evaluated under the assumption that the decoders treat undesired private streams as Gaussian random variables, despite the fact that they are from finite alphabets. Such suboptimal decoders reveal the performance of commonly used low-complexity receivers that treat interference as Gaussian noise instead of demodulating it, as demonstrated in Chapter 5. To ensure a fair comparison, a decoding complexity constraint is defined.

- To reveal the theoretical performance of conventional RSMA and CS-RSMA, we develop precoder optimization methods to maximize SR and MMF among users, under the assumption of finite alphabets and suboptimal decoders. To exploit the flexibility of RSMA, we allow adaptive alphabet selection under

a decoding complexity constraint. Numerical results from the proposed optimization algorithms are provided, which show that CS-RSMA outperforms conventional RSMA with and without SIC in terms of SR and performs similarly to the two in terms of MMF.

- While the achievable rate given by GMI measures theoretical limits and provides closed-form expressions for system design, they are based on idealized signal-space encoding and decoding, which remain a challenge in practice. Although the achievable rate of a proper channel model reflects the throughput in practical systems, a more accurate evaluation can be obtained through LLS, which involves practical designs, i.e., those incorporating binary encoding/decoding and modulation. For this purpose, as well as improving the practicality of CS-RSMA, we propose PHY transceiver designs for CS-RSMA and apply LLS to measure the BER performance of these designs. Similarly to the achievable rate results, the LLS results show that CS-RSMA slightly outperforms conventional RSMA. From the perspective of practical applications, the benefits of CS-RSMA over conventional RSMA are also discussed in terms of encoding/decoding complexity, control signalling overhead, and retransmission design.

Organization: The rest of this chapter is organized as follows. In Section 6.1, we introduce the proposed CS-RSMA scheme. In Section 6.2, results from Section 5.4 on multi-user communications with suboptimal decoders are used as the performance metric of conventional RSMA and CS-RSMA. Decoding complexity is also defined. In 6.3, to evaluate the theoretical performance of conventional RSMA and CS-RSMA, SR and MMF optimization methods that capture the suboptimal decoders are proposed. Section 6.4 summarizes the numerical results on the theoretical performance. In Section 6.5, we propose the PHY implementation of CS-RSMA and evaluate it using LLS, and discuss the implementation benefits of CS-RSMA over conventional RSMA. 6.6 discusses the potential of CS-RSMA with optimal encoding and decoding. Finally, Section 6.7 summarizes this chapter.

Unless otherwise stated, the system model presented in Section 2.1 is assumed throughout this chapter.

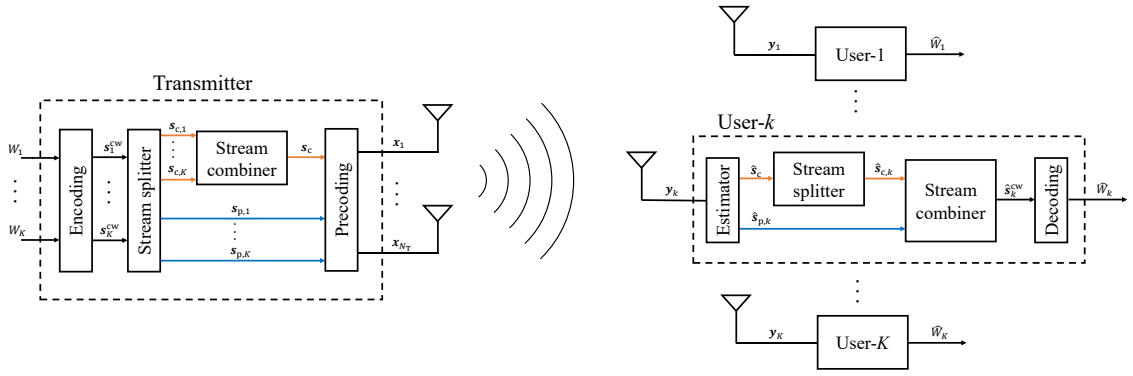


Figure 6.1: Codeword-Segmentation RSMA for multi-user MISO.

6.1 Codeword-segmentation RSMA

We continue to use expressions (2.1) and (2.2) for the transmitted and received signals and keep the associated notations unchanged. We assume that the symbols in $\mathbf{s}_{\mathrm{c}}$ and $\mathbf{s}_{\mathrm{p},k}$, $\forall k \in \mathcal{K}$, are uniformly drawn from their corresponding alphabets, denoted by $\mathcal{X}_{\mathrm{c}}$ and $\mathcal{X}_{\mathrm{p},k}$. CS-RSMA is defined by its encoding of $\{W_1, ..., W_K\}$ into $\{\mathbf{s}_{\mathrm{c}},\ \mathbf{s}_{\mathrm{p},1},\ ...,\ \mathbf{s}_{\mathrm{p},K}\}$ along with its distinct decoding procedure. Both the encoding and decoding differentiate it from conventional RSMA. The general procedures of CS-RSMA are depicted in Figure 6.1 and described as follows.

In CS-RSMA, $\{W_1, ..., W_K\}$ are first encoded into K codewords, $\{\mathbf{s}_1^{\mathrm{cw}},\ ...,\ \mathbf{s}_K^{\mathrm{cw}}\}$, where $\mathbf{s}_k^{\mathrm{cw}} \in \mathbb{C}^{n_k \times 1}$, $\forall k \in \mathcal{K}$, and n_k is the codeword length for user-k satisfying

$$n_k \geq n \tag{6.1}$$

and

$$\sum_{k=1}^{K} n_k = (K+1)n. \tag{6.2}$$

$\{\mathbf{s}_1^{\mathrm{cw}},\ ...,\ \mathbf{s}_K^{\mathrm{cw}}\}$ are then segmented into private and common parts. The private parts have a uniform length n and are denoted by $\{\mathbf{s}_{\mathrm{p},1},\ ...,\ \mathbf{s}_{\mathrm{p},K}\} \subset \mathbb{C}^{n \times 1}$. The common parts are denoted by $\{\mathbf{s}_{\mathrm{c},1},\ ...,\ \mathbf{s}_{\mathrm{c},K}\}$, where $\mathbf{s}_{\mathrm{c},k} \in \mathbb{C}^{(n_k - n) \times 1}$, $\forall k \in \mathcal{K}$. The common parts are then combined (e.g., concatenated or interleaved) into one symbol stream, denoted by $\mathbf{s}_{\mathrm{c}}$, and the private parts remain as separate private streams. As a consequence of (6.2), $\mathbf{s}_{\mathrm{c}} \in \mathbb{C}^{n \times 1}$. An example of codeword segmentation and combination is illustrated in Figure 6.2. Upon reception, user-k estimates $\mathbf{s}_{\mathrm{c}}$ and $\mathbf{s}_{\mathrm{p},k}$ from y_k and output $\widehat{\mathbf{s}}_{\mathrm{c}}$ and $\widehat{\mathbf{s}}_{\mathrm{p},k}$. Then it extracts an estimate of its desired codeword segment from $\widehat{\mathbf{s}}_{\mathrm{c}}$, denoted by $\widehat{\mathbf{s}}_{\mathrm{c},k}$. $\widehat{\mathbf{s}}_{\mathrm{c},k}$ and $\widehat{\mathbf{s}}_{\mathrm{p},k}$ are combined into $\widehat{\mathbf{s}}_k^{\mathrm{cw}}$, an

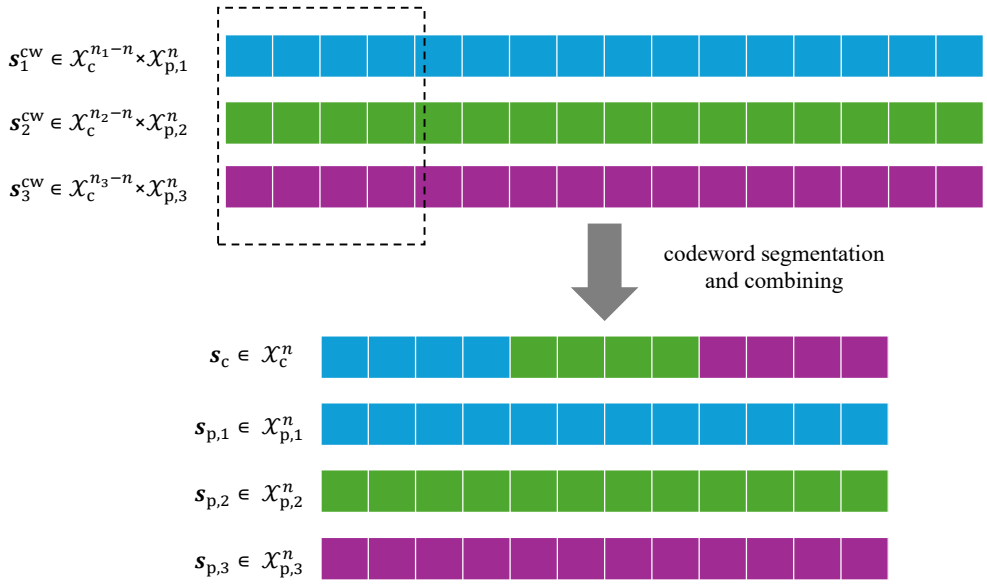


Figure 6.2: A visual illustration of codeword segmentation and combination.

estimate of $\mathbf{s}_k^{\text{cw}}$. Finally, based on $\widehat{\mathbf{s}}_k^{\text{cw}}$, user-$k$ decodes its desired message, $\widehat{W}_k$.

The main difference between conventional RSMA and CS-RSMA lies in the order of encoding and splitting: at the transmitter, conventional RSMA splits the messages before encoding them, whereas CS-RSMA directly encodes the messages and segments the codewords; at the receivers, the conventional RSMA decodes the common and private streams before reconstructing the message, whereas CS-RSMA first reconstructs the desired codeword from the codeword segments, and then decodes the reconstructed codeword into the desired message.

Regarding receiver implementation, SIC can be implemented with conventional RSMA. With SIC receivers, each user can first decode the common message, re-encode it into an estimate of the common stream, remove/cancel the estimated common stream from the received signal, and decode the desired private message. Meanwhile, SIC-free receivers provide a low-complexity option for conventional RSMA as discussed in Chapters 3 and 4, where re-encoding and cancellation are not performed. With CS-RSMA, SIC cannot be implemented. This is because the common stream is not a codeword, but a combination of segmented codewords from different users, and therefore cannot be decoded directly. SIC-free receivers are still valid for CS-RSMA. Fortunately, Chapter 4 shows that SIC-free receivers only incur a minor performance loss in RSMA compared to SIC receivers. Therefore, the absence of SIC in the CS-RSMA is not a critical limitation.

The fact that the common stream consists of codeword segments not only leads to the drawback of not employing SIC, but also introduces a benefit: each user is only required to decode its own codeword segment within the $\mathbf{s}_c$, rather than the entire $\mathbf{s}_c$, unless a single user occupies the whole stream.

Remark 6.1 (Relationship between CS-RSMA and rate splitting)**.** *Conventional rate-splitting approaches, such as [6, 7, 4, 8], split messages and encode them at different rates. In this sense, CS-RSMA does not strictly fall under the class of rate-splitting schemes. Nevertheless, we refer to it as CS-RSMA to acknowledge its conceptual connection to RSMA and because, as shown later, it provides an alternative implementation of SIC-free RSMA that preserves its performance while offering several practical advantages.*

6.2 Performance evaluation using mismatched decoding

We evaluate the achievable rates of RSMA and CS-RSMA when employing suboptimal decoders that treat part of the interference as Gaussian noise. The resulting decoding complexity is also quantified.

6.2.1 Evaluation of RSMA and CS-RSMA under suboptimal decoding

In Chapters 3 and 4, the performance of SIC-based and SIC-free receivers is studied considering constellation-constrained mutual information. However, mutual information has been shown to be achievable only under the assumption that the decoder optimally uses channel knowledge (e.g., maximum-likelihood decoding in [58]). Such decoders can be computationally prohibitive when the number of signals observed by the receivers increases. In this section, to address scenarios where complexity constraints are imposed on receivers, we apply the GMI results for FAGCI models with MU-MISO derived in Chapter 5 to analyze and evaluate the performance of conventional RSMA and CS-RSMA.

We assume that each user applies a suboptimal decoding strategy by treating the desired signals optimally, but the undesired signals as Gaussian random variables

with the same variance.[1] For example, when user-k decodes s_{c}, it treats $\mathbf{h}_k^H \mathbf{p}_k s_{\mathrm{p},k}$ optimally, but treats $\{\mathbf{h}_k^H \mathbf{p}_{k'} s_{k'}\}$ as $\mathcal{CN}(0, |\mathbf{h}_k^H \mathbf{p}_{k'}|^2)$, $\forall k' \in \mathcal{K}/k$. This decoding strategy is based on the assumption that desired signals are typically stronger than unwanted signals. An insight from Chapter 5 is that treating strong interference as Gaussian random variables can lead to significant losses, whereas the impact of treating weak interference as Gaussian random variables can be negligible. Hence, we consider decoders that treat the desired signals optimally to enable high performance under limited allowance for decoding complexity.

Recalling (5.26), the contribution to the achievable rate of user-k from the common stream and the desired private stream can be characterized using the GMI as follows:

- For the common stream:

$$I_{\mathrm{c},k} = I_{\mathrm{GMI}} \left(\mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} \mathcal{X}_{\mathrm{c}}, \mathbf{h}_k^H \mathbf{p}_k \mathcal{X}_{\mathrm{p},k}, \sum_{k' \neq k} \mathbf{h}_k^H \mathbf{p}_{k'} \mathcal{X}_{\mathrm{p},k'}, \sigma_z^2 \right). \tag{6.3}$$

- For the private stream, assuming that SIC is applied to remove the common stream signal:

$$I_{\mathrm{p},k}^{\mathrm{SIC}} = I_{\mathrm{GMI}} \left(\mathbf{h}_k^H \mathbf{p}_k \mathcal{X}_{\mathrm{p},k}, \{0\}, \sum_{k' \neq k} \mathbf{h}_k^H \mathbf{p}_{k'} \mathcal{X}_{\mathrm{p},k'}, \sigma_z^2 \right). \tag{6.4}$$

- For the private stream, assuming SIC-free receivers, i.e., the decoding of the private stream is under interference from the common stream:

$$I_{\mathrm{p},k}^{\text{non-SIC}} = I_{\mathrm{GMI}} \left(\mathbf{h}_k^H \mathbf{p}_k \mathcal{X}_{\mathrm{p},k}, \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} \mathcal{X}_{\mathrm{c}}, \sum_{k' \neq k} \mathbf{h}_k^H \mathbf{p}_{k'} \mathcal{X}_{\mathrm{p},k'}, \sigma_z^2 \right). \tag{6.5}$$

Note that, with CS-RSMA, each stream carries only a segment of a codeword. Therefore, the GMI of each stream quantifies only its contribution to the overall rate and does not represent an achievable rate on its own.

We assign c_k, $\forall k \in \mathcal{K}$, two roles: we let c_k denote the portion of W_{c} occupied by $W_{\mathrm{c},k}$ in conventional RSMA (the same as in Section 2.1), and the portion of $\mathbf{s}_{\mathrm{c}}$

[1] This decoding strategy is based on implications on soft codeword level interference cancellation for SIC receivers, and joint-demapping receivers for non-SIC receivers proposed in Chapter 3. The joint-demapping receiver is classified as a non-SIC receiver in Chapter 3, which effectively treats the desired signals optimally when demodulating a stream. This is not the only possible decoding strategy.

allocated to $\mathbf{s}_{\mathrm{c},k}$ in CS-RSMA (hence, $c_k = (n_k - n)/n$ for CS-RSMA). Therefore, the conditions $c_k \geq 0$ and $\sum_{k\in\mathcal{K}} c_k = 1$ hold under both interpretations.

The achievable rate for user-k under conventional RSMA schemes and CS-RSMA is expressed as follows:

- Conventional RSMA with SIC:

$$R_k = c_k \min_{k'\in\mathcal{K}} I_{\mathrm{c},k'} + I_{\mathrm{p},k}^{\mathrm{SIC}}. \tag{6.6}$$

- Conventional RSMA with SIC-free receivers:

$$R_k = c_k \min_{k'\in\mathcal{K}} I_{\mathrm{c},k'} + I_{\mathrm{p},k}^{\text{non-SIC}}. \tag{6.7}$$

- CS-RSMA:

$$R_k = c_k I_{\mathrm{c},k} + I_{\mathrm{p},k}^{\text{non-SIC}}. \tag{6.8}$$

Note that (6.8) assumes that the user-k ignores the dependence between the decoding of $\mathbf{s}_{\mathrm{c},k}$ and $\mathbf{s}_{\mathrm{p},k}$. The achievability of (6.8) under such separable decoders is proved in Appendix E.1.

Remark 6.2 (On the dependence between the decoding of $\mathbf{s}_{\mathrm{c},k}$ and $\mathbf{s}_{\mathrm{p},k}$)**.** *In the above discussion, we assumed that the decoding of* $\mathbf{s}_{\mathrm{c},k}$ *and* $\mathbf{s}_{\mathrm{p},k}$ *is separable for two reasons:*

- *In practical implementations,* $\mathbf{s}_{\mathrm{c},k}$ *and* $\mathbf{s}_{\mathrm{p},k}$ *are typically encoded and decoded without incorporating their mutual dependency during transmission. By ignoring this dependency in the rate expression, we capture this limitation.*
- *Ignoring this dependency allows us to maintain the notion of common and private streams, thereby enabling a meaningful comparison with conventional RSMA.*

A higher achievable rate of CS-RSMA will be addressed in Section 6.6, where the dependency is properly incorporated.

Remarkably, "$\min_{k'\in\mathcal{K}} I_{\mathrm{c},k'}$" in (6.7) is replaced by "$I_{\mathrm{c},k}$" in (6.8). Since

$$\min_{k'\in\mathcal{K}} I_{\mathrm{c},k'} \leq I_{\mathrm{c},k}, \tag{6.9}$$

the rate of CS-RSMA is lower bounded by SIC-free RSMA.

6.2.2 Constrained decoding complexity

Comparing communication schemes in the finite alphabet regime is more involved than in the Gaussian signaling regime. On the one hand, the achievable rate ultimately depends on the cardinality of alphabets under sufficient SNR. On the other hand, the decoding complexity is proportional to the number of observed signals and the cardinality of their alphabets, which can be observed from (5.6) and (5.7). Therefore, we introduce the concept of decoding complexity constraint to ensure the fairness of comparison.

Definition 6.1 (Decoding Complexity)**.** *A decoder of complexity-δ can compute "*$\exp(|\cdot|^2)$*" for at most n times for each received symbol.*

For example, for the channel model given by (5.1), the matched decoding, as in (5.6), is only valid if $\delta \geq |\mathcal{X}||\mathcal{I}||\mathcal{J}|$. The mismatched decoder given by (5.7) is valid if $\delta \geq |\mathcal{X}||\mathcal{I}|$. Therefore, it requires a lower decoding complexity than the matched decoding.

Proposition 6.1 (Decoding Complexity of RSMA)**.** *For conventional RSMA (with or without SIC) and CS-RSMA, if user-k treats all the undesired private streams as Gaussian noise, the decoding is feasible with at least complexity-$|\mathcal{X}_c||\mathcal{X}_{p,k}|$ at user-k, $\forall k \in \mathcal{K}$.*

Proof. For all the RSMA schemes considered, we show that, if user-k treats all the undesired private streams as Gaussian random variables, it needs to compute

$$d(s_{\mathrm{c}}, s_{\mathrm{p},k}, y_{t,k}) = \exp\left(-\frac{|y_{t,k} - \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}} - \mathbf{h}_k^H \mathbf{p}_k s_{\mathrm{p},k}|^2}{\sum_{k' \neq k} |\mathbf{h}_k^H \mathbf{p}_{k'}|^2 + \sigma_z^2}\right), \tag{6.10}$$

$\forall s_{\mathrm{c}} \in \mathcal{X}_{\mathrm{c}}$, $\forall s_{\mathrm{p},k} \in \mathcal{X}_{\mathrm{p},k}$ and for each time instance t. The rest of the decoding is achieved by only summing and multiplying $d(s_{\mathrm{c}}, s_{\mathrm{p},k}, y_{t,k})$ with different arguments. The exact decoders in terms of $d(s_{\mathrm{c}},\ s_{\mathrm{p},k},\ y_{t,k})$ are listed as follows:

- For conventional RSMA, Let $\{\mathbf{s}_{\mathrm{c}}^{(1)}, ..., \mathbf{s}_{\mathrm{c}}^{(M_{\mathrm{c}})}\}$ denote the codebook for common stream and $\{\mathbf{s}_{\mathrm{p},k}^{(1)}, ..., \mathbf{s}_{\mathrm{p},k}^{(M_{\mathrm{p},k})}\}$ denote the codebook for the private stream intended for user-k, $\forall k \in \mathcal{K}$, with M_{c} and $M_{\mathrm{p},k}$ being the corresponding codebook sizes. Let $s_{\mathrm{c},t}^{(q)}$ and $s_{\mathrm{p},t,k}^{(q)}$ respectively denote the symbols of $\mathbf{s}_{\mathrm{c}}^{(q)}$ and $\mathbf{s}_{\mathrm{p},k}^{(q)}$

at time instance t. The decoder for the common stream can be represented as

$$\begin{aligned}\widehat{W}_{\mathrm{c}} &= \arg \max_{q\in 1,\ldots,M_{\mathrm{c}}} q^n(\mathbf{s}_{\mathrm{c}}^{(q)}, \mathbf{y}_k) \\ &= \arg \max_{q\in 1,\ldots,M_{\mathrm{c}}} \prod_{t=1}^{n} \sum_{s_{\mathrm{p},k}\in\mathcal{X}_{\mathrm{p},k}} d(s_{\mathrm{c},t}^{(q)}, s_{\mathrm{p},k}, y_{t,k}).\end{aligned} \tag{6.11}$$

If SIC is applied, the decoder for the private stream is based on $\widehat{W}_{\mathrm{c}}$ and can be represented as

$$\begin{aligned}\widehat{W}_k &= \arg \max_{q\in 1,\ldots,M_{\mathrm{p},k}} q^n(\mathbf{s}_{\mathrm{p},k,t}^{(q)}, \mathbf{y}_k) \\ &= \arg \max_{q\in 1,\ldots,M_{\mathrm{p},k}} \prod_{t=1}^{n} d(s_{\mathrm{c},t}^{(\widehat{W}_{\mathrm{c}})}, s_{\mathrm{p},t,k}^{(q)}, y_{t,k}).\end{aligned} \tag{6.12}$$

If SIC is not applied,

$$\begin{aligned}\widehat{W}_k &= \arg \max_{q\in 1,\ldots,M_k} q^n(\mathbf{s}_{\mathrm{p},k}^{(q)}, \mathbf{y}_k) \\ &= \arg \max_{q\in 1,\ldots,M_k} \prod_{t=1}^{n} \sum_{s_{\mathrm{c}}\in\mathcal{X}_{\mathrm{c}}} d(s_{\mathrm{c}}, s_{\mathrm{p},t,k}^{(q)}, y_{t,k}).\end{aligned} \tag{6.13}$$

- For CS-RSMA, let $\{\mathbf{s}_k^{(1)}, ..., \mathbf{s}_k^{(M_k)}\}$ represent the codebook for user-k with M_k being the codebook size, $\forall k \in \mathcal{K}$, and $s_{t,k}^{(q)}$ represent the symbol of q-th codeword at time instance t. The decoder can be represented as

$$\begin{aligned}\widehat{W}_k &= \arg \max_{q\in 1,\ldots,M_k} q^n(\mathbf{s}_k^{(q)}, \mathbf{y}_k) \\ &= \arg \max_{q\in 1,\ldots,M_k} \prod_{t'\in\mathcal{T}_k} \sum_{s_k\in\mathcal{X}_{\mathrm{p},k}} d(s_{\tau_{\mathrm{c}}(t'),k}^{(q)}, s_k,\ y_{t,k}) \\ &\qquad \prod_{t=1}^{n} \sum_{s_{\mathrm{c}}\in\mathcal{X}_{\mathrm{c}}} d(s_{\mathrm{c}}, s_{\tau_{\mathrm{p},k}(t),k}^{(q)}, y_{t,k}),\end{aligned} \tag{6.14}$$

where $\mathcal{T}_k$ represents the time indices where the symbols of $\mathbf{s}_{\mathrm{c}}$ are generated from $\mathbf{s}_{\mathrm{c},k}$, and $\tau_{\mathrm{c}}(\cdot)$ and $\tau_{\mathrm{p},k}(\cdot)$ respectively maps the time index into the corresponding symbol index of user-k's codeword.

Hence, computing $d(s_{\mathrm{c}}, s_{\mathrm{p},k}, y_{t,k})$, $\forall s_{\mathrm{c}} \in \mathcal{X}_{\mathrm{c}}$ and $\forall s_{\mathrm{p},k} \in \mathcal{X}_{\mathrm{p},k}$, covers all "$\exp(|\cdot|^2)$" operations required for (6.11)-(6.13). This shows that complexity-δ is sufficient to decode conventional RSMA and CS-RSMA. It is also necessary to compute $d(s_{\mathrm{c}}, s_{\mathrm{p},k}, y_{t,k})$, $\forall s_{\mathrm{c}} \in \mathcal{X}_{\mathrm{c}}$ and $\forall s_{\mathrm{p},k} \in \mathcal{X}_{\mathrm{p},k}$, as long as, at each time instance and

for each stream, the probability of any symbol in the alphabet occurring is non-zero. □

6.3 Precoder optimization

We propose precoder optimization algorithms for SR and MMF optimization under finite alphabets and suboptimal decoders. The alphabets are assumed to be fixed in this section, but will be allowed to adjust adaptively in Section 6.4. To reduce computational complexity, the approximation given by (5.27) is utilized in the optimization, but the performance evaluation in Section 6.4 will be based on the exact GMI given by (5.26).

6.3.1 Sum-rate optimization

We consider a sum-rate maximization problem as follows:

$$\begin{aligned} \max_{\mathbf{P},\mathbf{c}} \ & \sum_{k\in\mathcal{K}} R_k \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \le P_{\mathrm{T}} \\ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}, \end{aligned} \tag{6.15}$$

where $\mathbf{c} = [c_1, c_2, ..., c_K]$.

Proposition 6.2 (Sum-rate maximization reformulation)**.** *Under different RSMA schemes, (6.15) can be simplified into the following formulations:*

- *For conventional RSMA with or without SIC,*

$$\begin{aligned} \max_{\mathbf{P}} \ & \min_{k\in\mathcal{K}} \ I_{\mathrm{c},k} + \sum_{k\in\mathcal{K}} I_{\mathrm{p},k}^{\text{SIC/non-SIC}} \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \le P_{\mathrm{T}}. \end{aligned} \tag{6.16}$$

- *For CS-RSMA,*

$$\begin{aligned} \max_{\mathbf{P}} \ & \max_{k\in\mathcal{K}} \ I_{\mathrm{c},k} + \sum_{k\in\mathcal{K}} I_{\mathrm{p},k}^{\text{non-SIC}} \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \le P_{\mathrm{T}}. \end{aligned} \tag{6.17}$$

Proof. (6.16) is readily seen by substituting (6.6) and (6.7) into (6.15). For (6.17), substituting (6.8) into (6.15) leads to the following objective function:

$$\sum_{k\in\mathcal{K}} c_k I_{\mathrm{c},k} + \sum_{k\in\mathcal{K}} I_{\mathrm{p},k}^{\text{non-SIC}}. \tag{6.18}$$

It is trivial to see that the optimal $\mathbf{c}$ is always given by having $c_k = 1$ for $k = \arg\max_k I_{\mathrm{c},k}$, and $c_{k'} = 0$ for $k' \neq k$.[2] Using this optimal $\mathbf{c}$ results in (6.17). □

Remark 6.3. *An interesting difference between (6.16) and (6.17) is that the "*min*" operator in 6.16) is replaced by "*max*" in (6.17). This indicates that CS-RSMA is expected to perform at least as well as conventional RSMA without SIC in SR, since its problem formulation only differs from the latter by changing the "*min*" to a "*max*". The comparison between CS-RSMA and conventional RSMA with SIC is difficult to reveal analytically because, although CS-RSMA is disadvantaged in the omission of SIC, its common stream is not restricted to being decodable by all users.*

Remark 6.4. *To achieve the maximum SR of CS-RSMA, if the common stream is activated, the number of codewords to split is strictly one. In contrast, in conventional RSMA, the number of messages to split can be any number between 1 and K, as discussed in Chapter 4.*

We further reformulate the problems in Proposition 6.2 using the logarithmic barrier [100] to approximate 6.16) and (6.17) as the following unconstrained optimization problem:

$$\max_{\mathbf{P},\mathbf{c}} \ \tau I_{\text{sum}} + \log\left(P_{\mathrm{T}} - \|\mathbf{P}\|_{\mathrm{F}}^2\right), \tag{6.19}$$

where I_{sum} represents the objective function of 6.16) and (6.17) depending on the RSMA scheme in use, and τ is a parameter that sets the accuracy of the approximation.

For RSMA with SIC, one subgradient of I_{sum} w.r.t. $\mathbf{P}$ is given by

$$g(I_{\text{sum}}) = \nabla_{\mathbf{P}} I_{\mathrm{c},k'} + \sum_{k\in\mathcal{K}} \nabla_{\mathbf{P}} I_{\mathrm{p},k}^{\text{SIC}}, \tag{6.20}$$

[2] In the cases where $k \in \arg\max_k I_{\mathrm{c},k}$ has more than one element, we simply select one element as k.

Algorithm 8 Subgradient Ascent with Barrier Method for Sum-Rate Maximization

Input: P_{T}, $\mathcal{X}_{\mathrm{c}}$, $\mathcal{X}_{\mathrm{p},k}$, $\mathbf{h}_k$, $k \in \mathcal{K}$, β, ϵ, $v_{\max}$, $\tau_{\max}$.
Output: $\mathbf{P}^\star$

Initialize $\mathbf{P}^0$ and τ.
repeat
  $v := 0$.
  repeat
    $\Delta\mathbf{P} := \tau\, g(I_{\mathrm{sum}}) + \frac{\mathbf{P}^v}{\|\mathbf{P}^v\|_{\mathrm{F}}^2 - P_{\mathrm{T}}}$.
    Choose step size α via backtracking line search.
    Update. $\mathbf{P}^{v+1} := \mathbf{P}^v + \alpha\Delta\mathbf{P}$.
    Update. $\Omega_{\mathrm{SR}}^{v+1} := \Omega_{\mathrm{SR}}(\mathbf{P}^{v+1})$.
    $v := v + 1$.
  until $\mid \Omega_{\mathrm{SR}}^{v} - \Omega_{\mathrm{SR}}^{v-1} \mid < \epsilon$ or $v \geq v_{\max}$.
  Update $\tau := \beta\tau$.
until $\tau \geq \tau_{\max}$.
return $\mathbf{P}^\star := \mathbf{P}^v$.

where $k' = \arg\min_k I_{\mathrm{c},k}$. For RSMA without SIC and CS-RSMA,

$$g(I_{\mathrm{sum}}) = \nabla_{\mathbf{P}} I_{\mathrm{c},k'} + \sum_{k\in\mathcal{K}} \nabla_{\mathbf{P}} I_{\mathrm{p},k}^{\text{non-SIC}}, \tag{6.21}$$

where, for RSMA without SIC, $k' = \arg\min_k I_{\mathrm{c},k}$ and, for CS-RSMA, $k' = \arg\max_k I_{\mathrm{c},k}$. The gradients in (6.20) and (6.21) can be computed by referring to (6.3)-(6.5) and (5.28).

It can be noticed that the problems described by (6.15)-(6.19) are non-convex due to the non-convexity of (5.26) and (5.27) w.r.t. the precoders, and non-smooth due to the "min/max" operators over k. Hence, it is difficult to obtain the global optimal solutions of them. We use the subgradient descent method to optimize (6.19). The complete algorithm is summarized in Algorithm 8, with $\Omega_{\mathrm{SR}}(\mathbf{P})$ representing the objective function in (6.19). Algorithm 8 is guaranteed to converge because each update on $\mathbf{P}$ leads to a non-decreasing update on the objective function, and the objective function is clearly upper-bounded.

6.3.2 Max-min fairness optimization

We now consider a max-min fairness problem as follows:

$$\begin{aligned} \max_{\mathbf{P},\mathbf{c}} \ & \min_{k\in\mathcal{K}} \ I_{\mathrm{GMI},k} \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{T}} \\ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}. \end{aligned} \tag{6.22}$$

(6.22) is a non-convex problem. To solve it efficiently, we decouple it into the following two sub-problems,

$$\begin{aligned} \max_{\mathbf{c}} \ & \min_{k\in\mathcal{K}} \ I_{\mathrm{GMI},k}(\mathbf{c};\ \mathbf{P}) \\ \text{s.t.} \ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}, \end{aligned} \tag{6.23}$$

and

$$\begin{aligned} \max_{\mathbf{P}} \ & \min_{k\in\mathcal{K}} \ I_{\mathrm{GMI},k}(\mathbf{P};\ \mathbf{c}) \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{T}}. \end{aligned} \tag{6.24}$$

(6.23) optimizes $\mathbf{c}$ with a fixed $\mathbf{P}$, whereas (6.24) optimizes $\mathbf{P}$ with a given $\mathbf{c}$. Note that (6.23) is a linear programming problem, and hence its global optimal solution can be efficiently obtained. In particular, we developed a closed-form solution as in Proposition 6.3.

Proposition 6.3. *The global optimal solution of (6.23) is given in closed form by Algorithm 9.*

Proof. See Appendix E.2. □

To remove the discontinuity caused by the point-wise minimum in (6.24), we apply a log-sum-exp approximation to its objective function and consider the following:

$$\begin{aligned} \max_{\mathbf{P}} \ & \frac{1}{\gamma} \log \sum_{k\in\mathcal{K}} \exp \ \left(\beta I_{\mathrm{GMI},k}(\mathbf{P};\ \mathbf{c})\right) \\ \text{s.t.} \ & \|\mathbf{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{T}}. \end{aligned} \tag{6.25}$$

Algorithm 9 Global Optimal Common Stream Allocation

Input: $I_{\mathrm{c},k},\ I_{\mathrm{p},k},\ k \in \mathcal{K}$.
Output: $\mathbf{c}^\star$

$I'_{\mathrm{c},k} := \min_k\ I_{\mathrm{c},k}$ for conventional RSMA and
$I'_{\mathrm{c},k} := I_{\mathrm{c},k}$ for CS-RSMA.
$k' := K$.
repeat
  $\mathbf{A} := \mathrm{diag}\{[I'_{\mathrm{c},1},\ ...,\ I'_{\mathrm{c},k'}]\}$, $\mathbf{b} := [I_{\mathrm{p},1},\ ...,\ I_{\mathrm{p},k'}]^T$.
  $\xi := \frac{1+\mathbf{b}^T\mathbf{A}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{1}\mathbf{A}^{-1}}$.
  $\mathbf{c}' := \xi\mathbf{A}^{-1}\mathbf{1} - \mathbf{b}\mathbf{A}^{-1}$.
  $k' := k' - 1$.
until $\mathbf{c}' \succcurlyeq 0$.
return $\mathbf{c}^\star := [\mathbf{c}',\ \mathbf{0}^{1\times(K-k'-1)}]$.

The gap between point-wise minimum and its log-sum-exp approximation is characterized by (4.22).

A subgradient of the objective function in (6.25) w.r.t. $\mathbf{P}$ is derived to be

$$g(I_{\text{max-min}}) = \frac{1}{\gamma}\frac{\sum_{k\in\mathcal{K}} \beta \exp\ (\beta I_{\mathrm{GMI},k}(\mathbf{P};\mathbf{c}))\, g(I_{\mathrm{GMI},k}(\mathbf{P};\mathbf{c}))}{\sum_{k\in\mathcal{K}} \exp\ (\beta I_{\mathrm{GMI},k}(\mathbf{P};\mathbf{c}))}, \tag{6.26}$$

with

$$g(I_{\mathrm{GMI},k}(\mathbf{P};\mathbf{c})) = c_k g(I_{\mathrm{c}}) + \nabla_{\mathbf{P}} I_{\mathrm{p},k}, \tag{6.27}$$

where $g(I_{\mathrm{c}}) = \nabla_{\mathbf{P}} I_{\mathrm{c},k'}^{\text{SIC/non-SIC}}$ with $k' = \arg\min_k I_{\mathrm{c},k'}^{\text{SIC/non-SIC}}$ for RSMA with/without SIC, and $g(I_{\mathrm{c}}) = \nabla_{\mathbf{P}} I_{\mathrm{c},k}^{\text{non-SIC}}$ for CS-RSMA.

Similarly to (6.19), the inequality constraint in (6.25)can be integrated into the objective function through the logarithmic barrier. This leads to

$$\max_{\mathbf{P}}\ \frac{\tau}{\gamma}\log\sum_{k\in\mathcal{K}}\exp\ (\gamma I_{\mathrm{GMI},k}(\mathbf{P};\mathbf{c})) + \log\left(P_{\mathrm{T}} - \parallel \mathbf{P} \parallel_{\mathrm{F}}^2\right). \tag{6.28}$$

(6.28) can be solved using subgradient ascent. The solution of (6.22) is obtained by solving (6.23) and (6.28) alternatively and the steps are summarized in Algorithm 10, with $\Omega_{\mathrm{MMF}}(\mathbf{P})$ representing the objective function in (6.28). Algorithm 10 is guaranteed to converge because each update on $\mathbf{P}$ and $\mathbf{c}$ leads to a non-decreasing update on the objective function, and the objective function is clearly upper-bounded.

Algorithm 10 Subgradient Ascent with Barrier Method for Max-Min Fairness

Input: P_{T}, $\mathcal{X}_{\mathrm{c}}$, $\mathcal{X}_{\mathrm{p},k}$, $\mathbf{h}_k$, $k \in \mathcal{K}$, β, γ, ϵ, $v_{\max}$, $\tau_{\max}$.
Output: $\mathbf{P}^\star$ and $\mathbf{c}^\star$

Initialize $\mathbf{P}^0$ and τ.
repeat
  $v := 0$.
  repeat
    Update $I_{\mathrm{c},k}$, $I_{\mathrm{p},k}$ based on $\mathbf{P}^v$.
    Update $\mathbf{c}$ using Algorithm 9.
    $\Delta\mathbf{P} := \tau\, g(I_{\text{max-min}}) + \frac{\mathbf{P}^v}{\|\mathbf{P}^v\|_{\mathrm{F}}^2 - P_{\mathrm{T}}}$.
    Choose step size α via backtracking line search.
    Update. $\mathbf{P}^{v+1} := \mathbf{P}^v + \alpha\Delta\mathbf{P}$.
    Update. $\Omega_{\mathrm{MMF}}^{v+1} := \Omega_{\mathrm{MMF}}(\mathbf{P}^{v+1})$.
    $v := v + 1$.
  until $\mid \Omega_{\mathrm{MMF}}^{v} - \Omega_{\mathrm{MMF}}^{v-1} \mid < \epsilon$ or $v \geq v_{\max}$.
  Update $\tau := \beta\tau$.
until $\tau \geq \tau_{\max}$.
return $\mathbf{P}^\star := \mathbf{P}^v$ and $\mathbf{c}^\star := \mathbf{c}$.

6.4 Numerical results on achievable rate

In this section, we investigate the theoretical performance of CS-RSMA and conventional RSMA schemes under mismatched decoding through numerical simulations. All the results in this section are average performance measured by averaging 100 channel realizations. The precoder initialization in Algorithm 8 and 10 involves the low-complexity precoders proposed in Chapter 3 for constellation-constrained sum-rate optimization and some randomly generated precoders.

We follow the one-ring scattering model (see Appendix A for details) with $\theta_k = \pi/3$ and $\Delta_k = \Delta$, $\forall k \in \mathcal{K}$, which implies that users with similar channel statistics are served at the same time. The purpose of this setting is to model communication scenarios with a certain user density.

Prior to the results, we introduce the adaptive modulation scheme in use, which differs from the one in Chapter 4.

Table 6.1: Transmission modes under decoding complexity $\delta = 4$

Mode	1	2	3
Private Stream	QPSK	BPSK	{0}
Common Stream	{0}	BPSK	QPSK

Table 6.2: Transmission modes under decoding complexity $\delta = 16$

Mode	1	2	3	4	5
Private Stream	16QAM	8QAM	QPSK	BPSK	{0}
Common Stream	{0}	BPSK	QPSK	8QAM	16QAM

6.4.1 Adaptive modulation

We assume that all the users are subject to a decoding complexity $\delta = 4$ or 16.[3] The alphabets applied to the streams are chosen from BPSK and QAM constellations, subject to the decoding complexity constraints. This leads to a few transmission modes in terms of alphabets for each complexity constraint, as summarized in Tables 6.1 and 6.2. The simulations allow adaptive alphabet selection under certain decoding complexity.

Remark 6.5. *With Mode 1 in both Tables 6.1 and 6.2, RSMA becomes SDMA; with Mode 3 in Table 6.1 and Mode 5 in Table 6.2, RSMA is equivalent to multicasting all messages [17].*

Remark 6.6. *Achievable rates under finite alphabets are upper-limited by the cardinality of the alphabet. Different transmission modes lead to different maximum achievable rates. It can be observed from Tables 6.1 and 6.2 that the sum rate of Mode 1, i.e., SDMA, is upper-bounded by* $K \log_2 \delta$ *bits and is the highest among all the modes, whereas the sum rate of multicasting, i.e., the last mode in both tables, is upper-bounded by* $\log_2 \delta$ *bits and is the lowest among all the modes. For example, for* $K = 2$ *and* $\delta = 16$*, Mode 1 can bring at most 8 bits per transmission in SR, which cannot be achieved by any other mode.*

6.4.2 Sum-rate evaluation

Figure 6.3 depicts the average SR performance of conventional RSMA schemes, CS-RSMA and SDMA. It can be observed that RSMA schemes, including CS-RSMA,

[3]It is assumed that δ is the same for all the users for simplicity and clarity. The problem and solution introduced in Section 6.3 can be trivially generalized to ones with non-uniform complexity constraints among users.

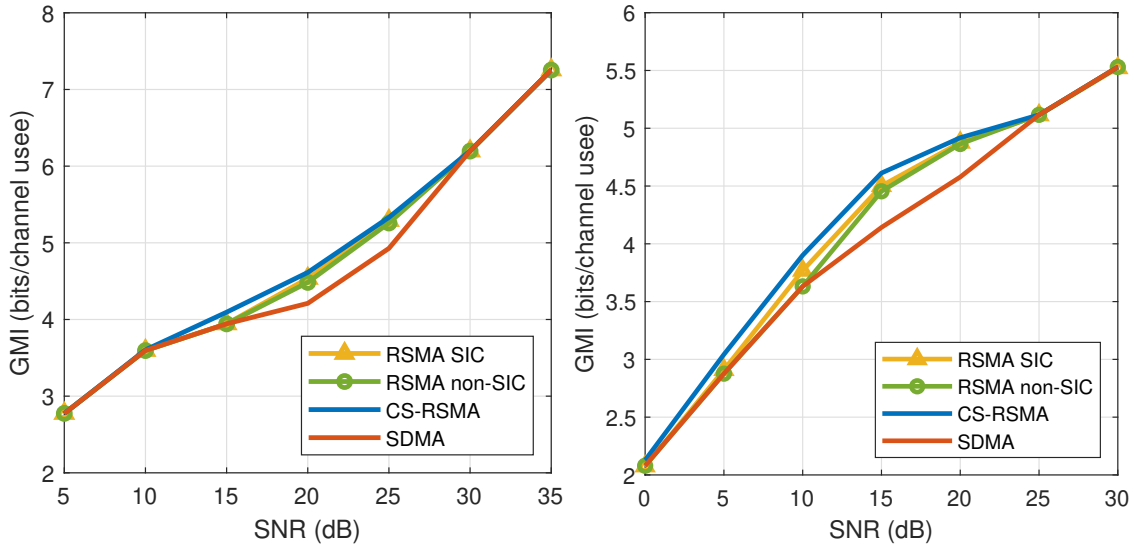
bring benefits in the medium SNR regime compared to SDMA. Furthermore, CS-RSMA slightly outperforms conventional RSMA schemes. The average SR led by different transmission modes under CS-RSMA is depicted in Figure 6.4, together with the performance of CS-RSMA given by adaptive alphabet/mode selection as a reference. It can be observed that the transmission modes that enable the common stream achieve higher SR than Mode 1, i.e., SDMA, at medium SNR. At high SNR, Mode 1 dominates as other modes saturate at lower SR, as suggested by Remark 6.6.

Figure 6.5 depicts the average contributions of the common and private streams to the user rates for CS-RSMA and conventional RSMA. It can be observed that CS-RSMA tends to convey more information in the common stream than conventional RSMA. This is a consequence of having "max" on the common rates in the SR problem formulation for CS-RSMA instead of the "min" in the one for conventional RSMA, as shown in (6.16) and (6.17).

6.4.3 Max-min fairness optimization

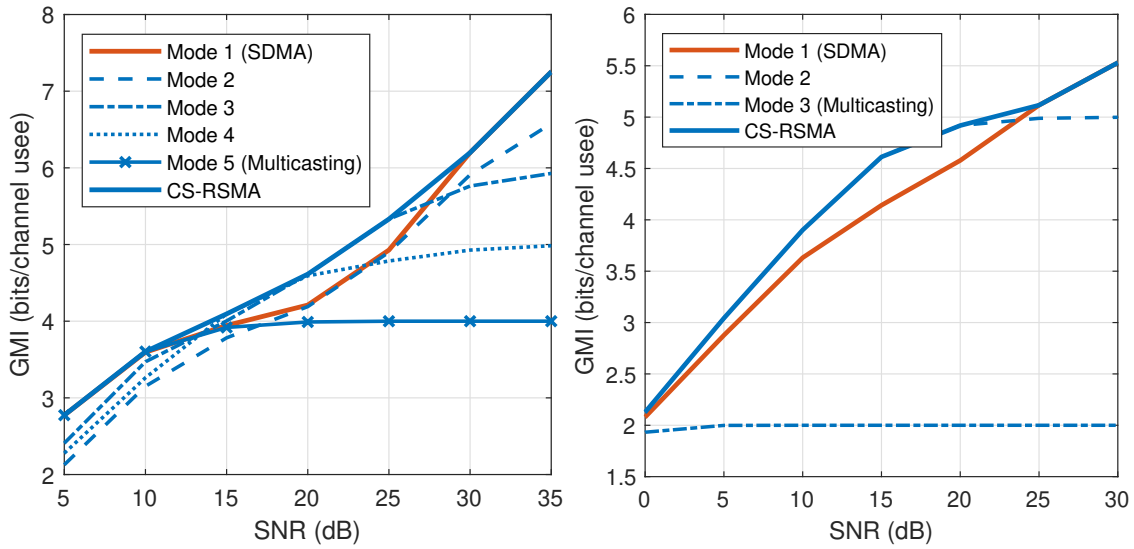
Figure 6.6 depicts the average MMF performance of conventional RSMA schemes, CS-RSMA and SDMA. Similar to SR performance, RSMA offers benefits at medium SNR by activating the common stream, but converges to SDMA at high SNR. CS-RSMA performs very similar to conventional RSMA schemes.[4] The performance of different modes of CS-RSMA is depicted in Figure 6.7, along with the performance of CS-RSMA given by adaptive alphabet/mode selection as a reference. It shows that activating the common stream brings higher minimum rates among users under medium SNR than Mode 1, i.e., SDMA, whereas Mode 1 dominates under high SNR, because the modes enabling the common stream saturate at lower rates, as suggested by Remark 6.6.

[4]For both SR and MMF evaluation, if the input signals are not from finite alphabets but are Gaussian distributed, the gap between RSMA with SIC and SDMA exists in both medium and high SNR regime, as shown in [9, 102], because the common stream rate will not saturate at high SNR. CS-RSMA, SIC-free RSMA and SDMA will perform similarly under Gaussian input, because the interference from the common stream when decoding the private stream is more detrimental than with finite-alphabet inputs.



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. (b) $N_{\mathrm{T}} = 4$, $K = 4$, $\Delta = \pi/12$, $\delta = 4$.

Figure 6.3: Average SR performance of conventional RSMA, CS-RSMA and SDMA.



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. (b) $N_{\mathrm{T}} = 4$, $K = 4$, $\Delta = \pi/12$, $\delta = 4$.

Figure 6.4: Average SR performance of different modes of CS-RSMA.

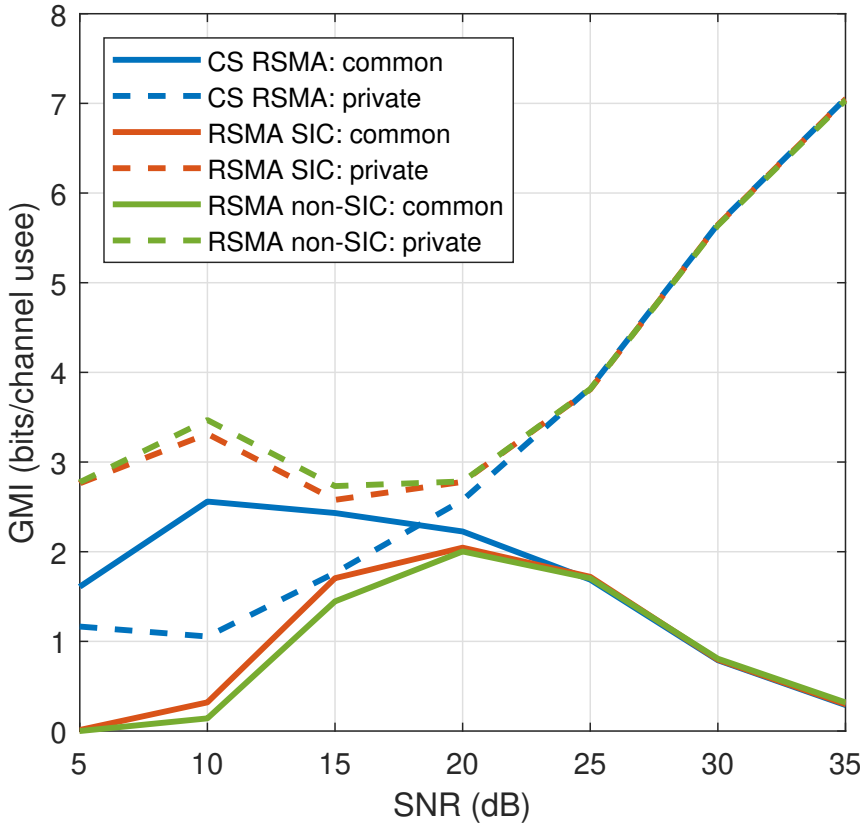
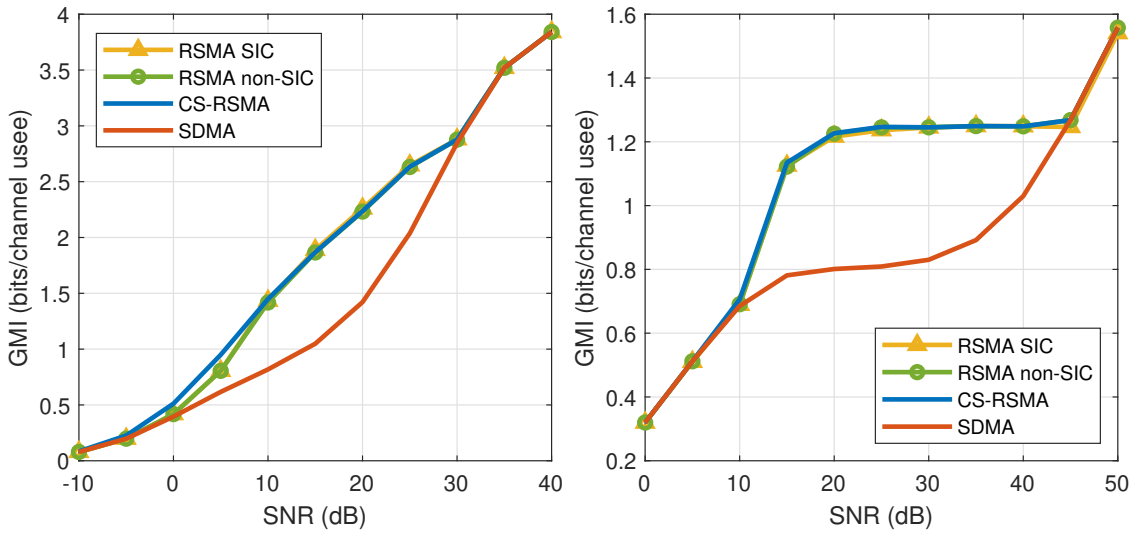
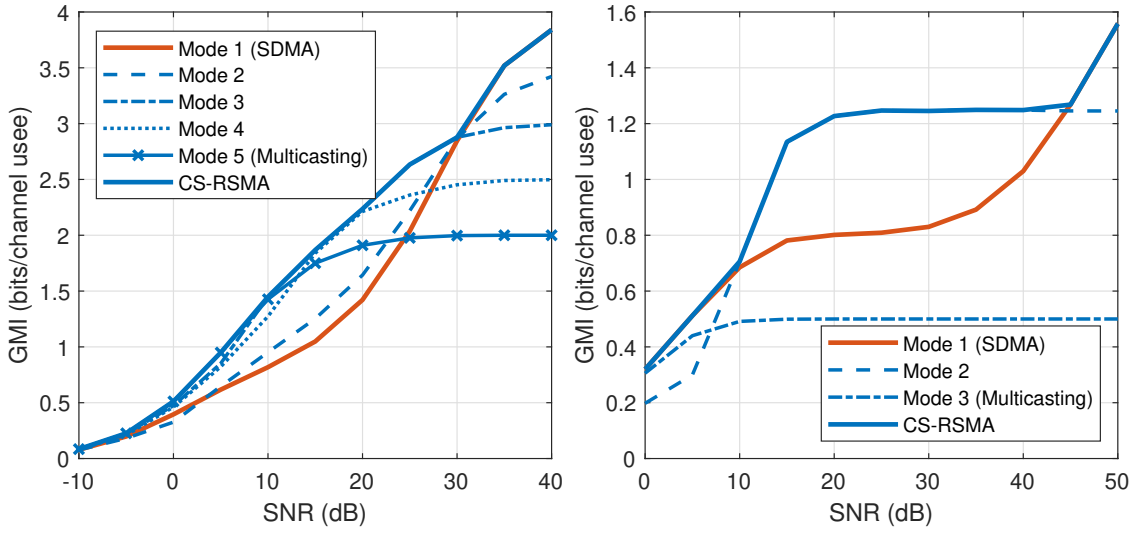


Figure 6.5: Comparing contribution of common and private stream with $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$.



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. (b) $N_{\mathrm{T}} = 4$, $K = 4$, $\Delta = \pi/9$, $\delta = 4$.

Figure 6.6: Average MMF performance of conventional RSMA, CS-RSMA and SDMA.



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. (b) $N_{\mathrm{T}} = 4$, $K = 4$, $\Delta = \pi/9$, $\delta = 4$.

Figure 6.7: Average MMF performance of different modes of CS-RSMA.

6.5 Physical-layer transceiver design for CS-RSMA and link-level simulation

In this section, we propose a transceiver architecture for CS-RSMA and evaluate its performance using link-level simulations.

6.5.1 Transceiver design

Figure 6.8 depicts the proposed CS-RSMA transmitter architecture, where "FEC" and "Mod" represent forward error correction coding[5] and modulation (mapping) respectively. For clarity, the time index is neglected in the notation. In Figure 6.8, the messages, W_1, W_2, ..., W_K, are first encoded into binary codewords, denoted by X_1, X_2, ..., X_K. Codeword segmentation and combining can be achieved by de-multiplexing and multiplexing the encoded bits into X_{c}, $X_{\mathrm{p},1}$, $X_{\mathrm{p},2}$,..., $X_{\mathrm{p},K}$. The common and private streams are obtained by modulating the resulting $K+1$ bit streams into symbol streams. The precoders then mix the symbol streams into a stream of vector symbols to feed the multiple transmit antennas. The resource mapper allocates different instances of the vector symbol stream to different physical layer resource elements.[6]

Figure 6.9 depicts two proposed receiver architectures for CS-RSMA, where "FEC^{-1}" and "Demod" represent channel decoder and demodulator (demapper), respectively. They are based on the joint de-mapping receiver and soft Symbol Level Interference Cancellation (SLIC) receiver for conventional RSMA proposed in Section 3.2. The two receiver architectures are explained as follows.

1. Joint de-mapping: The receiver first jointly demodulates the common stream and the desired private stream from the receiver signal. Assuming a soft-output demodulator is in use, the demodulator produces Log-Likelihood Ratios (LLRs) for X_{c} and $X_{\mathrm{p},k}$, denoted as $\mathrm{LLR}_{\mathrm{c}}$ and $\mathrm{LLR}_{\mathrm{p},k}$. The receiver then de-multiplexes (extracts) the LLRs for $X_{\mathrm{c},k}$ from $\mathrm{LLR}_{\mathrm{c}}$, denoted as $\mathrm{LLR}_{\mathrm{c},k}$. $\mathrm{LLR}_{\mathrm{c},k}$ and $\mathrm{LLR}_{\mathrm{p},k}$ are then multiplexed (combined) into LLRs of the desired unicast message, X_k, denoted as LLR_k. Finally, the receiver decodes LLR_k into an estimate of the desired message $\hat{W}_k$.

[5] This potentially includes interleaving and scrambling.

[6] These can be created, for example, by Orthogonal Frequency-Division Multiplexing (OFDM) and Time Domain Multiplexing (TDM).

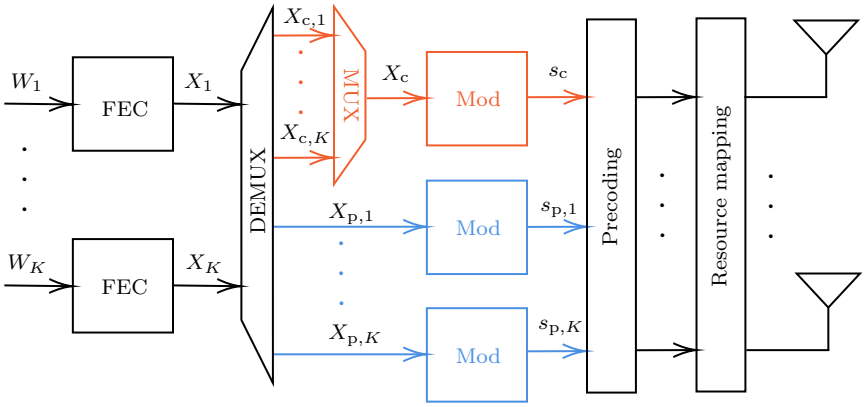
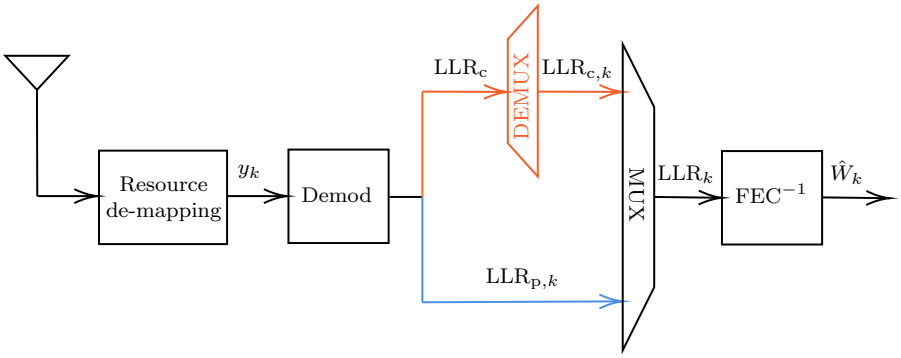


Figure 6.8: PHY transmitter design for CS-RSMA.

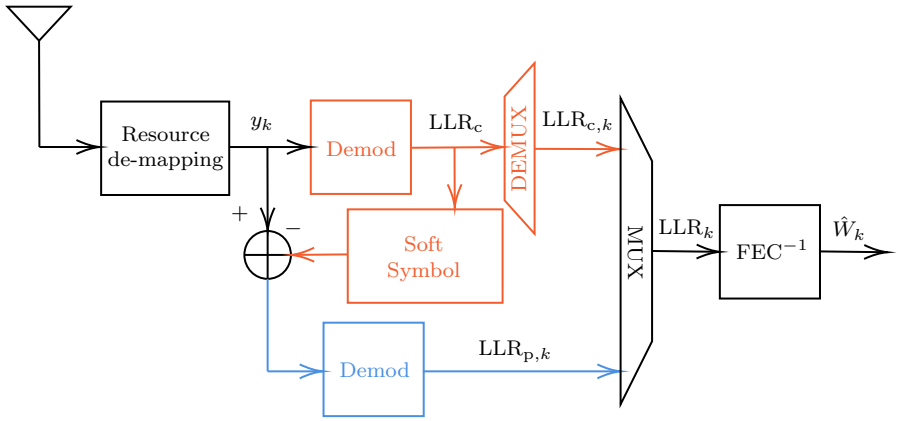
2. Soft SLIC: Jointly demodulating X_c and $X_{p,k}$ may introduce a high computational burden. Therefore, soft SLIC is proposed as a low-complexity alternative to joint demapping. With soft SLIC, the receiver first demodulates s_c by treating all private streams as Gaussian noise and obtains LLR_c. Based on LLR_c, the receiver then computes the posterior probability distribution of the common stream symbols, denoted by $P(s_c|y)$. $P(s_c|y)$ is used to compute the soft symbol estimates, i.e., $\mathbb{E}[s_c|y]$, denoted by $\tilde{s}$, and the variance of the residue noise, $\mathbb{E}[|s_c - \tilde{s}|^2|y]$, denoted as $\sigma^2_{s_c}$. Next, the receiver applies soft cancellation of s_c by subtracting $\mathbf{h}_k^H \mathbf{p}_c \tilde{s}_c$ from y_k and demodulates $s_{p,k}$ from the remaining signal into $\text{LLR}_{p,k}$. The remaining steps are the same as in the joint de-mapping implementation: the receiver de-multiplexes (extracts) $\text{LLR}_{c,k}$ from LLR_c, multiplexes $\text{LLR}_{c,k}$ and $\text{LLR}_{p,k}$ into LLR_k, and decodes $\hat{W}_k$ based on LLR_k.

6.5.2 Link-level simulation

Link-level simulations were performed to evaluate the proposed CS-RSMA transceiver designs. In the following, we assume fast fading channels, where individual channel observations are generated using the one-ring channel model introduced in Section V, with $N_T = 4$, $K = 2$, and $\Delta = \pi/18$. For channel encoding and decoding, the LDPC code specified in 5G NR [99] with belief propagation decoding is chosen for its popularity and efficiency for long code lengths. All the BER results were obtained by averaging the results from 50000 blocks. The symbol block length is set to 512. For de-mapping, the max-log approximation was applied for its popularity in practice. To reduce simulation time, the low-complexity precoder design proposed



(a) Joint de-mapping



(b) Soft SLIC

Figure 6.9: PHY receiver designs for CS-RSMA.

in Section 3.2 for RSMA without SIC is used for CS-RSMA.

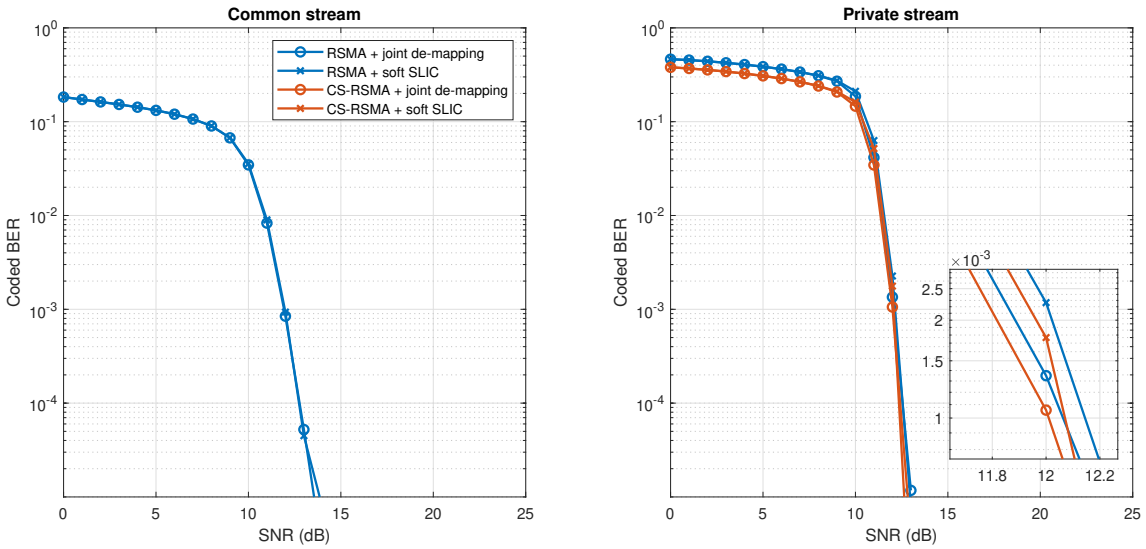


Figure 6.10: BER comparison between conventional RSMA and CS-RSMA with all streams using QPSK constellation. Code rate for RSMA: 0.85 for common streams, 0.325 for private streams. Code rate for CS-RSMA: 0.5 for all codewords.

Figure 6.10 depicts BER curves of conventional RSMA and CS-RSMA with both joint de-mapping and soft SLIC receivers. It can be observed that, with the same receiver implementation, CS-RSMA performs slightly better than conventional RSMA. Taking 10^{-4} as the target BER, the SNR gap is around 0.05dB. In addition, a soft SLIC receiver only leads to a minor performance loss compared to joint mapping, which is consistent with the observation in Chapter 3.

6.5.3 Advantages of CS-RSMA

Although the numerical results in Section V and this section show that the achievable rate and BER performance of CS-RSMA are not significantly different from conventional RSMA, the introduction of CS-RSMA is of significant importance for practical applications. We briefly discuss these benefits of implementing CS-RSMA compared to conventional RSMA in the following aspects.

1. Reduced encoding/decoding complexity: With CS-RSMA, the transmitter encodes K messages rather than $K+1$ messages as in conventional RSMA; the receiver decodes one codeword rather than two as in conventional RSMA. The reduced number of codewords for each user leads not only to reduced encoding/decoding complexity (in the sense that fewer codewords are needed) but

also to simplification of error detection (e.g., CRC check), interleaving, and scrambling.

2. Reduced signaling overhead: With CS-RSMA, each receiver only needs information on the encoding scheme for one codeword, rather than two (for common and private streams) as with conventional RSMA. This leads to fewer requirements on control signaling from CS-RSMA.

3. Simplified retransmission design: With conventional RSMA, an additional retransmission process, e.g., Hybrid Automatic Repeat Request (HARQ), needs to be implemented for the common stream. Due to the fact that the common stream is encoded with messages for potentially multiple users, designing the retransmission scheme for the common stream can be a complicated problem (some examples are given in [101]). With CS-RSMA, the retransmission mechanism only needs to be implemented for the K encoded unicast messages, exactly the same as the widely used MU-MIMO. This means that the existence of CS-RSMA can be transparent to the retransmission process for conventional MU-MIMO. For example, to implement RSMA in 5G NR PDSCH channels, CS-RSMA is naturally compatible with the existing HARQ process, whereas conventional RSMA requires redesigning the HARQ process for the common stream.

Although this work only presents the 1-layer CS-RSMA, i.e., with one common stream, the idea of codeword segmentation and creating common streams using part of segments can be trivially extended to multi-layer RSMA schemes, such as the hierarchical RSMA in [18] and the generalized multi-layer RSMA in [16], where more common streams need to be implemented. In these cases, the above three benefits will be significantly magnified as the number of common streams increases.

Overall, CS-RSMA brings significant convenience in terms of complexity in implementation and compatibility with existing systems and performs slightly better than conventional RSMA without SIC.

6.6 On CS-RSMA with optimal encoding and decoding

In Section 6.2, it is assumed that the dependence between $\mathbf{s}_{\mathrm{c},k}$ and $\mathbf{s}_{\mathrm{p},k}$ during transmission is ignored in both encoding and decoding, which directly leads to (6.8). The achievable rate of CS-RSMA was derived based on this assumption. While this assumption has the benefit of capturing practical limitations, as discussed in Remark 6.2, it also reduces the achievable rate of CS-RSMA. This section investigates the achievable rate of CS-RSMA when this constraint is removed.

6.6.1 A superposition perspective

We recall the block structure shown in Figure 6.2 and focus on the codeword for user-k, i.e., $\mathbf{s}_k^{\mathrm{cw}}$. During encoding, $\mathbf{s}_k^{\mathrm{cw}}$ is segmented into $\mathbf{s}_{\mathrm{c},k}$ and $\mathbf{s}_{\mathrm{p},k}$. The transmission involves precoding $\mathbf{s}_{\mathrm{c},k}$ (as part of $\mathbf{s}_{\mathrm{c}}$) and $\mathbf{s}_{\mathrm{p},k}$. At the receiver side, user-k observes a linear combination of $\mathbf{s}_{\mathrm{c},k}$ and $\mathbf{s}_{\mathrm{p},k}$, where the coefficients depend on the precoders and channel at user-k. The communication channel between the transmitter and user-k can therefore be viewed equivalently as a point-to-point channel without any fading. The inputs of this channel are drawn from alphabets determined by the precoders and channel at user-k, the transmission is interrupted by finite alphabet interference and Gaussian noise. This is illustrated in Figure 6.11.

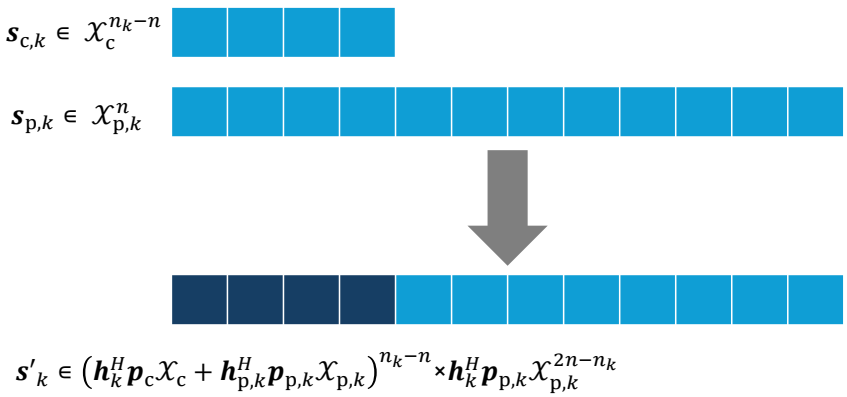


Figure 6.11: An effective codeword construction for user-k in CS-RSMA.

Formally, the received signal at user-k at time instance t can be written as

$$y_{t,k} = s'_{t,k} + \sum_{k' \in \mathcal{K}/k} s'_{t,k'} + z_{t,k'}. \tag{6.29}$$

The alphabet of $s'_{t,k}$ and $s'_{t,k}$ depends on the time index. Let $\mathcal{T} = \{1, ..., n\}$ and $\mathcal{T}_k$ represent the set of time indices when the common stream symbol, $s_{t,\mathrm{c}}$, is from $\mathbf{s}_{\mathrm{c},k}$, $\forall k$. Therefore, $\{\mathcal{T}_1, ..., \mathcal{T}_K\}$ is a partition of $\mathcal{T}$. For $t \in \mathcal{T}_k$,

$$s'_{t,k} \in \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} \mathcal{X}_{\mathrm{c}} + \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k} \mathcal{X}_{\mathrm{p},k},$$

$$s'_{t,k'} \in \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k'} \mathcal{X}_{\mathrm{p},k'},$$

and the instantaneous rate is denoted by $I_k^{t \in \mathcal{T}_k}$,

For $t \notin \mathcal{T}_k$,

$$s'_{t,k} \in \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k} \mathcal{X}_{\mathrm{p},k}$$

$$s'_{t,k'} \in \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} \mathcal{X}_{\mathrm{c}} + \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k'} \mathcal{X}_{\mathrm{p},k'}.$$

and the instantaneous rate is denoted by $I_k^{t \notin \mathcal{T}_k}$.

The assumption on the uniform distribution of s_{c} and $s_{\mathrm{p,k}}$, $\forall k \in \mathcal{K}$, from Chapter 6.1 implies that $s'_{t,k}$ and $s'_{t,k'}$ are also uniformly distributed.[7]

The overall rate at user-k is given by

$$\begin{aligned} R_k =& \frac{n_k - n}{n} I_k^{t \in \mathcal{T}_k} + \frac{2n - n_k}{n} I_k^{t \notin \mathcal{T}_k} \\ =& c_k I_k^{t \in \mathcal{T}_k} + (1 - c_k) I_k^{t \notin \mathcal{T}_k}, \end{aligned} \tag{6.30}$$

where we used $c_k = (n_k - n)/n$ from the definition of c_k for CS-RSMA. (6.30) can be rewritten as

$$R_k = c_k (I_k^{t \in \mathcal{T}_k} - I_k^{t \notin \mathcal{T}_k}) + I_k^{t \notin \mathcal{T}_k}. \tag{6.31}$$

By comparing (6.31) with (6.8), we observe that the two expressions resemble each other, where $I_k^{t \in \mathcal{T}_k} - I_k^{t \notin \mathcal{T}_k}$ plays the role of $I_{\mathrm{c},k}$, meanwhile $I_{\mathrm{p},k}^{\text{non-SIC}}$ and $I_k^{t \notin \mathcal{T}_k}$ are essentially the same.

Under the mismatched decoding conditions given in Section 6.2, i.e., user-k treats $s_{t,k'}$, $\forall k' \in \mathcal{K}/k$, as Gaussian random variables during decoding, we have

[7] We ignore the possibility that some points merge in the superimposed constellation.

$$I_k^{t\in\mathcal{T}_k} = I_{\mathrm{GMI}}\left(\mathbf{h}_k^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}} + \mathbf{h}_k^H\mathbf{p}_{\mathrm{p},k}\mathcal{X}_{\mathrm{p},k}, \{0\}, \sum_{k'\neq k}\mathbf{h}_k^H\mathbf{p}_{\mathrm{p},k'}\mathcal{X}_{\mathrm{p},k'}, \sigma_z^2\right) \tag{6.32}$$

and

$$I_k^{t\notin\mathcal{T}_k} = I_{\mathrm{GMI}}\left(\mathbf{h}_k^H\mathbf{p}_{\mathrm{p},k}\mathcal{X}_{\mathrm{p},k}, \mathbf{h}_k^H\mathbf{p}_{\mathrm{c}}\mathcal{X}_{\mathrm{c}}, \sum_{k'\neq k}\mathbf{h}_k^H\mathbf{p}_{\mathrm{p},k'}\mathcal{X}_{\mathrm{p},k'}, \sigma_z^2\right). \tag{6.33}$$

Note that, to achieve (6.32), the encoder needs to have knowledge of the distribution of $s'_{t,k'}$, which includes precoding and channel information.

6.6.2 Performance evaluation

Since (6.31) closely resembles (6.8), all optimization problem formulations, reformulations, and algorithms introduced in Section 6.3 can be directly applied to precoder optimization in CS-RSMA with optimal encoding and decoding, by replacing the corresponding rate expressions. By omitting the intermediate steps, the task reduces to solving the following two problems. For SR optimization, we need to solve

$$\max_{\mathbf{P}}\ \tau\left(\max_{k\in\mathcal{K}}\left(I_k^{t\in\mathcal{T}_k} - I_k^{t\notin\mathcal{T}_k}\right) + \sum_{k\in\mathcal{K}} I_k^{t\notin\mathcal{T}_k}\right) + \log\left(P_{\mathrm{T}} - \|\mathbf{P}\|_{\mathrm{F}}^2\right), \tag{6.34}$$

and, for MMF optimization, we need to solve

$$\begin{aligned} &\max_{\mathbf{c}}\ \min_{k\in\mathcal{K}}\ R_k(\mathbf{c};\ \mathbf{P}) \\ &\text{s.t.}\ \ \mathbf{c}^T\mathbf{1} = 1 \\ &\qquad\ \mathbf{c} \succcurlyeq \mathbf{0}, \end{aligned} \tag{6.35}$$

and

$$\max_{\mathbf{P}}\ \frac{\tau}{\gamma}\log\sum_{k\in\mathcal{K}}\exp\ (\gamma R_k(\mathbf{P};\mathbf{c})) + \log\left(P_{\mathrm{T}} - \|\mathbf{P}\|_{\mathrm{F}}^2\right). \tag{6.36}$$

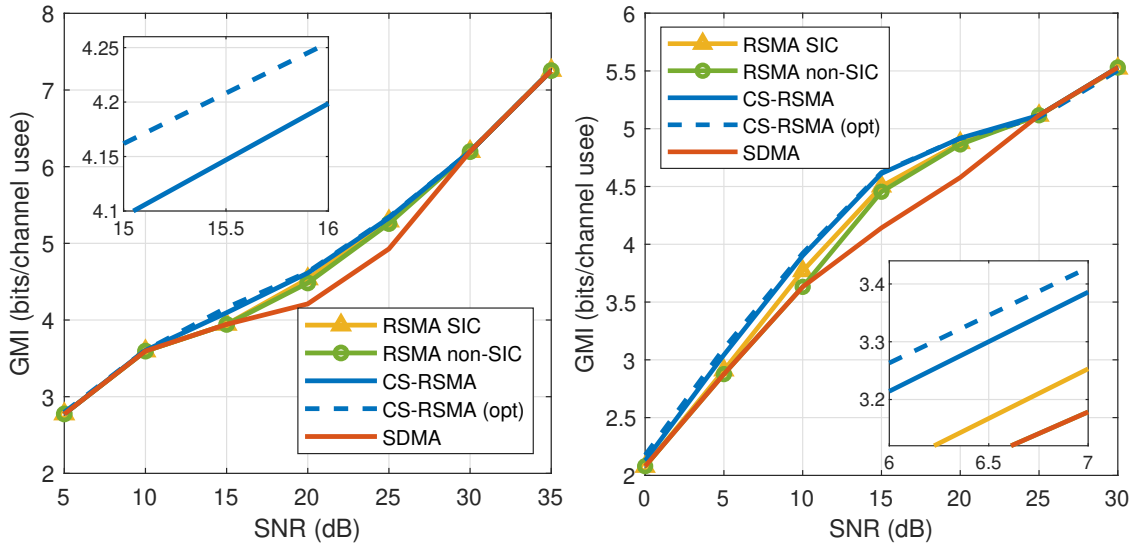
iteratively. (6.34) and (6.36) can be solved using subgradient descent, and (6.35) can be solved using Algorithm 9 with

$$I_{\mathrm{c},k} = I_k^{t\in\mathcal{T}_k} - I_k^{t\notin\mathcal{T}_k} \tag{6.37}$$

and

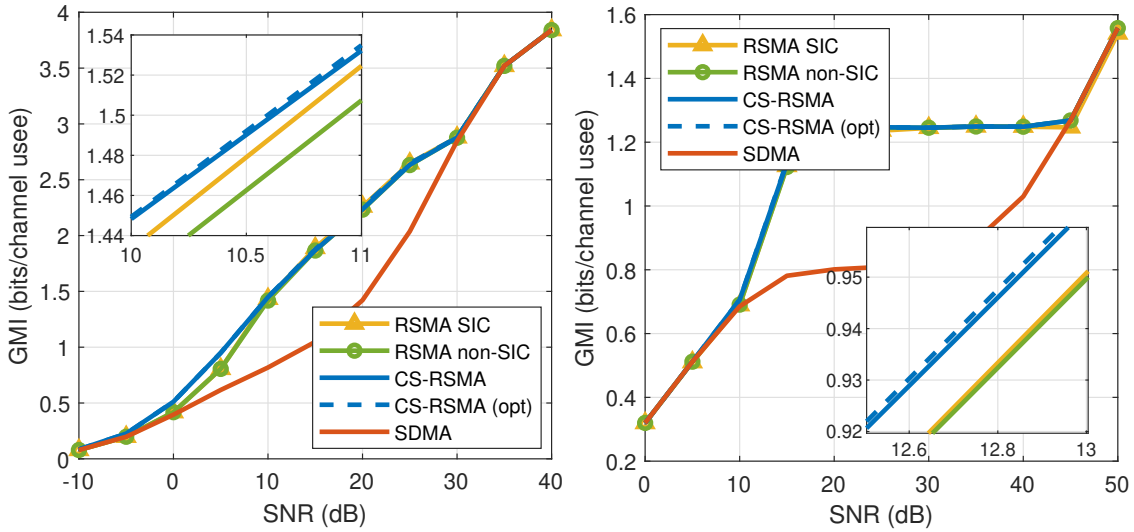
$$I_{\mathrm{p},k} = I_k^{t\notin\mathcal{T}_k} \tag{6.38}$$

as the inputs.



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. (b) $N_{\mathrm{T}} = 4$, $K = 4$, $\Delta = \pi/9$, $\delta = 4$.

Figure 6.12: Average SR performance of CS-RSMA with optimal encoder and decoder.



(a) $N_{\mathrm{T}} = 2$, $K = 2$, $\Delta = \pi/18$, $\delta = 16$. (b) $N_{\mathrm{T}} = 4$, $K = 4$, $\Delta = \pi/9$, $\delta = 4$.

Figure 6.13: Average MMF performance of CS-RSMA with optimal encoder and decoder.

By applying the same simulation setup as in Figures 6.3 and 6.6, the average SR and MMF performance of CS-RSMA with optimal encoder and decoders in comparison to other schemes is depicted in Figures 6.12 and 6.13. It can be observed that the optimal encoders and decoders only yield marginal improvements.

6.6.3 Two other schemes with equivalent asymptotic rate as CS-RSMA with optimal encoding/decoding

Apart from CS-RSMA, one can also interpret (6.30) as SDMA with geometrically shaped constellations. In such a scheme, user-k's transmission block of length n is divided into two parts: one part of block length $n_k - n$ uses a geometrically shaped constellation, which is designed through the linear combination of $\mathcal{X}_{\mathrm{c}}$ and $\mathcal{X}_{\mathrm{p},k}$; the other part of block length $2n - n_k$ uses a constellation given by $\mathcal{X}_{\mathrm{p},k}$.

In (6.31), if we compute $I_k^{t\in\mathcal{T}_k}$ and $I_k^{t\notin\mathcal{T}_k}$ using mutual information, instead of the generalized mutual information,

$$\begin{aligned} R_k =& c_k I(s_{\mathrm{c}}, s_{\mathrm{p},k}; y_k) + (1-c_k) I(s_{\mathrm{p},k}; y_k) \\ =& c_k (I(s_{\mathrm{c}}; y_k \mid s_{\mathrm{p},k}) + I(s_{\mathrm{p},k}; y_k)) + (1-c_k) I(s_{\mathrm{p},k}; y_k) \\ =& c_k I(s_{\mathrm{c}}; y_k \mid s_{\mathrm{p},k}) + I(s_{\mathrm{p},k}; y_k), \end{aligned} \tag{6.39}$$

where the chain rule of mutual information is applied. The last line of (6.39) can be achieved by another scheme: each user's message is encoded into two codewords, $\mathbf{s}_{\mathrm{p},k} \in \mathbb{C}^{n\times 1}$ and $\mathbf{s}_{\mathrm{c},k} \in \mathbb{C}^{(n_k-n)\times 1}$. $\{\mathbf{s}_{\mathrm{p},k} \mid k \in \mathcal{K}\}$ are separately precoded and transmitted, while $\{\mathbf{s}_{\mathrm{c},k} \mid k \in \mathcal{K}\}$ are combined (e.g., concatenated or multiplexed) into one block, denoted by $\mathbf{s}_{\mathrm{c}}$, which is precoded and transmitted along with $\{\mathbf{s}_{\mathrm{p},k} \mid k \in \mathcal{K}\}$. The codewords involved in this scheme are visually illustrated in Figure 6.14.

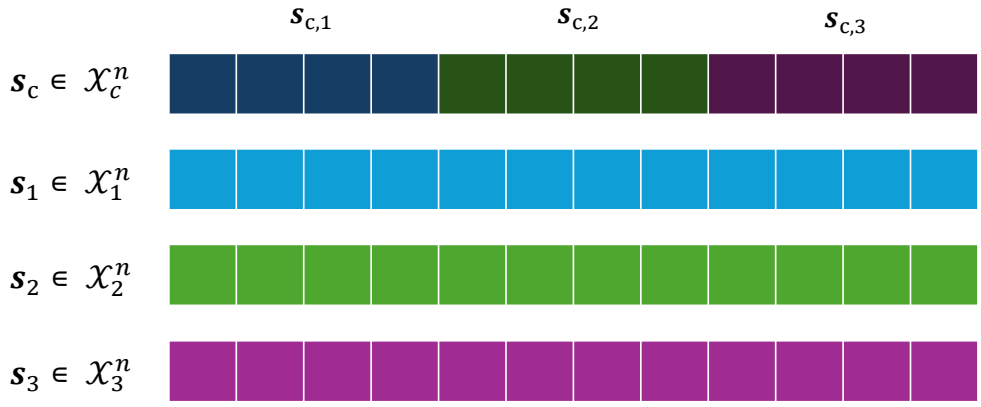


Figure 6.14: Illustration of a scheme where $\mathbf{s}_{\mathrm{c}}$ consists of multiple complete codewords, user-k desires $\mathbf{s}_{\mathrm{c},k}$, one codeword in $\mathbf{s}_{\mathrm{c}}$, and $\mathbf{s}_k$, a separately transmitted codeword.

Note that although this scheme achieves the same rate as CS-RSMA at infinite block

length, it has a higher probability of error, as each user's message is encoded into short blocks, rather than one long block as in CS-RSMA.

6.7 Summary

In this chapter, we proposed a novel architecture for RSMA, namely CS-RSMA. To capture practical constraints on receiver complexity, we evaluated the performance of conventional RSMA and CS-RSMA using the GMI for FAGCI channels, which assumes that users implement suboptimal decoders that treat undesired signals as Gaussian random variables for lower complexity. We further formulated SR and MMF optimization problems and developed optimization algorithms for precoder design. The numerical results in the SR and MMF evaluations reveal that CS-RSMA slightly outperforms conventional RSMA schemes in SR, while it performs very similarly to them in MMF. Furthermore, PHY transceiver designs were proposed for CS-RSMA, and LLS was utilized to further evaluate CS-RSMA under the proposed transceiver implementations. Overall, we concluded that CS-RSMA performs very similarly to conventional RSMA, and occasionally outperforms it, both under information-theoretical metrics and LLS. We also discussed the significant practical benefits of CS-RSMA on encoding/decoding complexity, control signalling overhead, and retransmission mechanism design in comparison to conventional RSMA. Finally, potentially higher rates from CS-RSMA, by implementing optimal encoding and decoding, were discussed.

Chapter 7

Conclusion

This thesis concludes with a summary of its key contributions and a discussion of future research directions.

7.1 Summary of thesis achievements

To advance both theory and practice, this thesis develops two extensions of RSMA: SIC-free RSMA and CS-RSMA. In addition, it presents an analysis of the impact of suboptimal interference treatment.

Chapter 3 started by considering low-complexity precoder designs that respectively optimize two information-theoretic metrics: the sum-rate (under Gaussian inputs) and constellation-constrained sum-rate of RSMA. It also investigated several receiver designs that can be applied with RSMA to provide more choices in terms of complexity and performance. The complexity of different receivers are quantified such that a complexity-performance trade-off with different receiver designs is revealed. It was further noticed that receiver designs can be classified into SIC-type receivers and SIC-free-type receivers based on their principles, and propose to alter the system design according to this classification. This also indicates that these receivers respectively push RSMA towards the two information-theoretic limits. Through theoretical results and link-level simulations, it was shown that, although non-SIC-type receivers lead to some performance loss compared to SIC-type receivers, they still preserve most of the superiority of RSMA over SDMA. Chapter 3 revealed an important fact that it is possible to preserve the superiority of RSMA over SDMA, including SDMA with advanced receivers, even without SIC in the receivers.

In Chapter 4, the fundamental limits of SIC-based RSMA and SIC-free RSMA under finite constellations were discovered by allowing the precoders to be fully designed under various scenarios. Precoder optimization methods were proposed to maximize the WSR and MMF performance for SIC-based and SIC-free RSMA under finite constellation constraints. Chapter 4 also proposed methods to scale up optimization algorithms to handle large-scale systems and avoid prohibitive computational cost. Numerical results show that both SIC-based and SIC-free RSMA outperform SDMA under different evaluation metrics and dimensions of the system. Moreover, SIC-free RSMA often preserved most of the superiority of SIC-based RSMA. Therefore, Chapter 4 extended the conclusion in Chapter 3 from low-complexity precoders to fully optimized precoders, from the sum-rate objective to the WSR and MMF objectives, and from small-scale to large-scale systems.

In Chapter 5, the focus was shifted to the effect of finite-alphabet interference under optimal and suboptimal decoders. Mismatched decoding was used to analyze communications with decoders that treat interference suboptimally, aiming at more accurate performance metrics. Specifically, it defines the FAGCI model, a finite-alphabet input Gaussian channel under interference, where the decoder can be matched (optimal) or mismatched (suboptimal) to the channel. The matched capacity is derived using mutual information, while a lower bound on the mismatched capacity under various decoding metrics is derived using the generalized mutual information. It was shown that the decoding metric in the proposed channel model is closely related to the behavior of the demodulator in BICM systems. Simulations illustrated that GMI/MI accurately predicts the throughput performance of BICM-type systems. Finally, the channel model and the GMI are extended to multiple-antenna cases, with an example of a multi-user multiple-input-single-output (MU-MISO) precoder optimization problem considering GMI under different decoding strategies. Overall, Chapter 5 discovered new insights about the impact of interference, proposed novel receivers, and introduced a new design and performance evaluation framework that more accurately captures the effect of interference.

Chapter 6 proposed a novel architecture for downlink RSMA, namely Codeword-Segmentation RSMA (CS-RSMA). Different from conventional RSMA, which splits users' messages into common and private parts before encoding, CS-RSMA encodes the users' messages directly, segments the codewords into common and private parts, and transmits the codeword segments using common and private streams. A novel performance analysis framework based on the derived GMI from Chapter 5 was also proposed. The numerical results revealed the theoretical performance of con-

ventional RSMA and CS-RSMA. It was observed that CS-RSMA leads to better performance than conventional RSMA in SR, and similar performance in MMF. Furthermore, a physical-layer implementation of CS-RSMA was proposed and evaluated through link-level simulations. In addition to performance benefits, Chapter 6 demonstrated that CS-RSMA brings significant benefits on the encoding/decoding, control signaling, and retransmission process compared to conventional RSMA.

7.2 Future work

We outline several possible avenues for future research, inspired by the strengths and shortcomings identified in this thesis.

- **RSMA/SIC-free RSMA/CS-RSMA under various channel conditions:** A significant part of this thesis has focused on precoder design for RSMA based on a single channel realization. In practice, this assumption is reasonable when the entire transmission block experiences the same channel conditions (which is less realistic in wireless applications), or when different channel realizations occur but the precoder can be adapted for each realization.[1] It is more often the case that the precoders are determined for a group of channels. For example, in 5G NR, the same precoders are applied to a group of subcarriers and time indices. It is therefore desirable to investigate the performance of the proposed RSMA schemes under precoders designed per-group of channels. Such studies can also reveal performance under imperfect CSIT, as in [9]. As shown in Section 5.1, performance in finite constellations can be very sensitive to the ratio between the desired signal strength and interference strength. The design of RSMA (including SIC-free RSMA and CS-RSMA) under finite alphabets needs to be more sophisticated in producing the desired signal-to-interference power ratio at the receivers. Therefore, designs for a group of channels may not be as trivial as in [9], and may require restrictions on the size of group/variance of channels within the groups.

- **Reference signals:** One shortcoming of RSMA that has often been overlooked is its requirement for additional channel estimation at receivers. In downlink systems, each user must estimate the precoded channels corresponding to all its intended streams. Since RSMA introduces common streams, users

[1] With a large number of channel realizations involved in one transmission, the ergodic rate is achievable for every transmission [103].

are also required to estimate the channels associated with these additional streams, which incurs additional estimation overhead and resource consumption. For example, in 5G NR, demodulation reference signals (DM-RS) are used for downlink channel estimation. Future research is therefore needed to investigate: (i) whether RSMA can maintain its performance advantages when constrained to a fair allocation of downlink channel estimation resources; and (ii) whether novel reference signal designs can be developed specifically for RSMA, and more generally, for non-orthogonal transmission schemes.

- **CS-RSMA under finite block lengths:** All the analyses in this thesis are based on the ideal assumption of infinite block lengths, which is expected to yield similar conclusions under finite but sufficiently long block lengths. Yet, at least one benefit of CS-RSMA has remained hidden under this assumption. In conventional RSMA and SIC-free RSMA, each user's message is encoded into two separate codewords, each of length n. In contrast, in CS-RSMA, each user's message is encoded into a single codeword of length n_k, where $n_k \geq n$. The longer block length of CS-RSMA provides an inherent advantage in terms of error probability compared to conventional RSMA and SIC-free RSMA. A detailed comparison of the resulting differences in error probability and achievable rates under finite block-length and error-probability constraints should therefore be investigated in future work. Due to the SIC-free nature of CS-RSMA, such an analysis needs to be based on finite constellations.

- **Novel receiver design for SIC-free RSMA and CS-RSMA:** The physical layer designs of SIC-free RSMA and CS-RSMA proposed in this thesis consider joint de-mapping and soft SLIC at the receivers. The soft SLIC technique was originally introduced to mitigate unwanted interference rather than to process desired signals as in RSMA. It would therefore be interesting to investigate novel receiver architectures specifically designed for SIC-free RSMA and CS-RSMA, which preserve their SIC-free nature while providing additional benefits in terms of complexity reduction and throughput improvement. The mismatched decoding framework introduced in Section 5 can serve as a foundation for inspiring new design approaches and assist in the analytical evaluation.

- **Analysis of the FAGCI model:** Further analysis of the FAGCI model and the corresponding suboptimal decoders can be pursued. For example, the LM rate [59] can be investigated. As the LM rate is an upper bound to the GMI, the discovery of the LM rate may provide new insights into designing

systems with suboptimal interference treatment. The LM rate is the achievable rate under the constant-composition random code ensemble subject to the optimal composition. Therefore, the challenge in determining the LM rate lies in finding the optimal composition.

- **Extensions to broader scenarios:** While conventional RSMA has demonstrated advantages in various wireless scenarios in the literature as outlined in Section 1.1.1, it is of interest to examine whether similar benefits can be achieved with SIC-free RSMA and CS-RSMA. For the FAGCI model, as discussed at the end of Section 5.4, the performance metrics based on GMI/MI, together with the system design framework, can also be applied to a wide range of communication scenarios where interference is expected.

- **Reducing computational complexity:** MI and GMI under the finite alphabet assumption are used throughout this thesis, yet they have high computational complexity due to the nature of finite alphabets. Methods for reducing such complexity can be extremely valuable.

- **Prototyping:** The prototyping of SIC-free RSMA and CS-RSMA is a promising direction. Such efforts would not only validate some of the conclusions in this thesis but also help to identify potential gaps between these schemes and their practical implementations and standardization. The novel demodulator designs proposed in Section 5.4 can also be implemented and examined through prototyping.

Appendix A

Channel Models Involving Fading Correlation

This appendix presents detailed descriptions of the channel models employed in this thesis for generating the numerical results.

A.1 Ong-ring model

The spatially correlated Rayleigh-fading channel is assumed, i.e. $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}^{N_t \times 1},\ \mathbf{R}_k)$ and $\mathbf{R}_k \in \mathbb{C}^{N_t \times N_t}$ is the channel covariance matrix for user-k. Applying eigendecomposision to $\mathbf{R}_k$, one can have

$$\mathbf{R}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H, \tag{A.1}$$

where $\mathbf{\Lambda}_k$ is an $r_k \times r_k$ diagonal matrix of the nonzero eigenvalues of $\mathbf{R}_k$ and $\mathbf{U}_k \in \mathbb{C}^{N_t \times r_k}$ is a matrix containing the eigenvectors of $\mathbf{R}_k$ corresponding to the nonzero eigenvalues. Using Karhunen-Loeve representation, $\mathbf{h}_k$ can be expressed in the form

$$\mathbf{h}_k = \mathbf{U}_k \mathbf{\Lambda}_k^{\frac{1}{2}} \mathbf{w}_k, \tag{A.2}$$

where $\mathbf{w}_k \in \mathbb{C}^{r_k \times 1} \sim \mathcal{CN}(\mathbf{0},\ \mathbf{I})$.

We further consider the one-ring scattering model [104], where the underlying assumption is that, between the transmitter and user-k, there is no line-of-sight path and the angle-of-departure (AoD) of all the scatters is uniformly distributed around

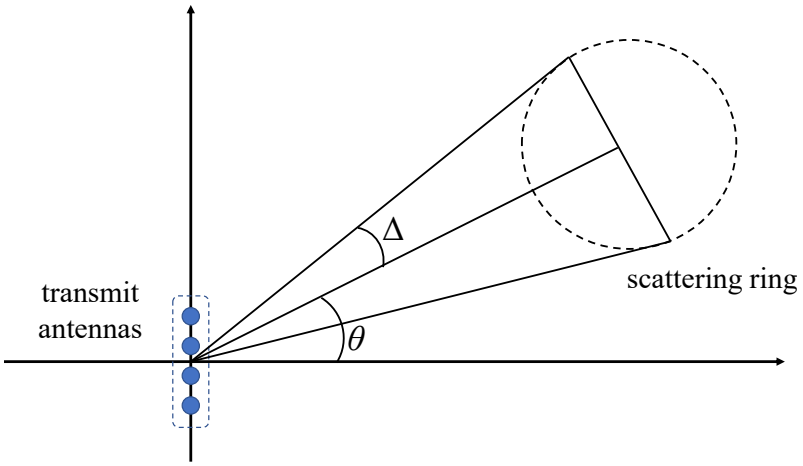


Figure A.1: Transmit antenna array and a scattering ring centered at AoD θ with two sided angular spread Δ.

a centre AoD, θ_k, with an angular spread of Δ_k, as depicted in Fig. A.1. Assuming a uniform and linear antenna array centered at the origin and along the y-axis with the spacing of the antenna elements to be $\lambda/2$, as depicted in Fig. A.1, the (m,n)-th element of $\mathbf{R}_k$ is given by

$$[\mathbf{R}_k]_{m,n} = \frac{1}{2\Delta_k} \int_{\theta_k-\Delta_k}^{\theta_k+\Delta_k} e^{-j\pi(m-n)\sin\alpha} d\alpha. \tag{A.3}$$

A.2 Rician fading model

The Rician fading model assumes that there is a dominant line-of-sight (LOS) component in addition to scattered non-LOS (NLOS) components, i.e.,

$$\mathbf{h}_k = \sqrt{\frac{\kappa_k}{\kappa_k+1}}\mathbf{h}_k^{\mathrm{LoS}}(\theta_k^{\mathrm{az}}, \theta_k^{\mathrm{el}}) + \sqrt{\frac{1}{\kappa_k+1}}\mathbf{h}_k^{\mathrm{NLoS}}, \forall k \in \mathcal{K}, \tag{A.4}$$

where $\mathbf{h}_k^{\mathrm{LoS}}(\theta_k^{\mathrm{az}}, \theta_k^{\mathrm{el}})$, $\mathbf{h}_k^{\mathrm{NLoS}}$ and κ_k denote the Line-of-Sight (LoS) channel component, non-LoS channel component and the Rician factor. $\mathbf{h}_k^{\mathrm{LoS}}$ is characterized by the array geometry and the direction of departure in azimuth and elevation, denoted by θ_k^{az} and θ_k^{el} respectively. $\mathbf{h}_k^{\mathrm{NLoS}}$ is modeled by Rayleigh fading with all the entries following $\mathcal{CN}(0,1)$.

Appendix B

Appendices for Chapter 3

B.1 Proof of Proposition 3.1 and 3.2

The constellation-constrained achievable rate of the common stream at user-k is expressed as (B.1), where (B.1a) follows the chain rule of mutual information, (B.1b) follows the relationship between mutual information and conditional entropy, and (B.1c) utilizes the assumption that the directions of $\{\mathbf{p}_{\mathrm{p},1}, ..., \mathbf{p}_{\mathrm{p},K}\}$ are designed using ZF precoding.

$$
\begin{aligned}
R_{\mathrm{CC,c},k} =& I(s_{\mathrm{c}}; y_k \mid \mathbf{h}_k, \mathbf{P}) \\
=& I\Big(\{s_{\mathrm{c}}, s_{\mathrm{p},1}, ..., s_{\mathrm{p},K}\}; y_k \;\Big|\; \mathbf{h}_k, \mathbf{P}\Big) - I\Big(\{s_{\mathrm{p},1}, ..., s_{\mathrm{p},K}\}; y_k \;\Big|\; \mathbf{h}_k, \mathbf{P}, s_{\mathrm{c}}\Big) && \text{(B.1a)} \\
=& H(s_{\mathrm{c}}, s_{\mathrm{p},1}, ..., s_{\mathrm{p},K}) - H\Big(s_{\mathrm{c}}, s_{\mathrm{p},1}, ..., s_{\mathrm{p},K} \;\Big|\; y_k, \mathbf{h}_k, \mathbf{P}\Big) \\
& - H(s_{\mathrm{p},1}, ..., s_{\mathrm{p},K}) + H\Big(s_{\mathrm{p},1}, ..., s_{\mathrm{p},K} \;\Big|\; y_k, \mathbf{h}_k, \mathbf{P}, s_{\mathrm{c}}\Big) && \text{(B.1b)} \\
=& H(s_{\mathrm{c}}, s_{\mathrm{p},k}) - H\Big(s_{\mathrm{c}}, s_{\mathrm{p},k} \;\Big|\; y_k, \mathbf{h}_k, [\mathbf{p}_{\mathrm{c}}, \mathbf{p}_{\mathrm{p},k}]\Big) \\
& - H(s_{\mathrm{p},k}) + H\Big(s_{\mathrm{p},k} \;\Big|\; y_k, \mathbf{h}_k, [\mathbf{p}_{\mathrm{c}}, \mathbf{p}_{\mathrm{p},k}], s_{\mathrm{c}}\Big) && \text{(B.1c)}
\end{aligned}
$$

It is assumed that the probability distributions of the constellation points are uniform, hence

$$H(s_{\mathrm{c}}, s_{\mathrm{p},k}) = \log |\mathcal{X}_{\mathrm{c}} \times \mathcal{X}_k| \tag{B.2}$$

and

$$H(s_{\mathrm{p},k}) = \log |\mathcal{X}_k|. \tag{B.3}$$

Assuming user-k successfully decodes s_{c} and applies SIC before decoding s_k, the constellation-constrained achievable rate of s_k is expressed as

$$\begin{aligned} R_{\mathrm{CC},k}^{\mathrm{SIC}} &= I(s_{\mathrm{p},k}; y_k \mid s_{\mathrm{c}}) \\ &= H(s_{\mathrm{p},k}) - H(s_{\mathrm{p},k} \mid y_k - \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}), \end{aligned} \tag{B.4}$$

where $|\mathbf{h}_k^H \mathbf{p}_{\mathrm{p},i}|^2 = 0, \forall i \neq k$, has been used. Alternatively, if user-k does not apply SIC but treats s_{c} as noise when decoding $s_{\mathrm{p},k}$,

$$\begin{aligned} R_{\mathrm{CC},k}^{\text{non-SIC}} =& I(s_{\mathrm{p},k}; y_k) \\ =& I\Big(\{s_{\mathrm{c}}, s_{\mathrm{p},k}\}; y_k\Big) - I(s_{\mathrm{c}}; y_k \mid s_{\mathrm{p},k}) \\ =& H(s_{\mathrm{c}}, s_{\mathrm{p},k}) - H(s_{\mathrm{c}}, s_{\mathrm{p},k} \mid y_k) - H(s_{\mathrm{c}}) + H(s_{\mathrm{c}} \mid y_k, s_{\mathrm{p},k}) \end{aligned} \tag{B.5}$$

The CC sum-rate of RSMA is expressed as

$$R_{\mathrm{CC,sum}}^{\mathrm{SIC}} = \min_{k \in \mathcal{K}} R_{\mathrm{CC,c},k} + \sum_{k=1}^{K} R_{\mathrm{CC},k}^{\mathrm{SIC}} \tag{B.6}$$

when SIC is applied at the receivers and

$$R_{\mathrm{CC,c},k} + \sum_{k=1}^{K} R_{\mathrm{CC},k}^{\text{non-SIC}} \tag{B.7}$$

when the receivers do not use SIC.

By combining (2.7), (2.9), (B.1b-B.4) and (B.6), Proposition 3.1 is proved. By combining (2.7), (2.9), (B.1b-B.3), (B.5) and (B.7), Proposition 3.2 is proved.

B.2 The equivalence of joint de-mapping and treating interference as noise

Consider a receiver that observes a symbol x, which is corrupted by interference u, and Gaussian noise n. x is associated with a binary label sequence, $\{c_1, ..., c_L\}$, representing the code bits. The received signal is expressed as

$$y = x + u + n, \tag{B.8}$$

where x and u are from finite constellation sets, $\mathcal{X}$ and $\mathcal{U}$, following certain prior probability distributions, $P(x)$ and $P(u)$. For clarity, the channel effects on the desired symbol and the interference are considered implicitly in x and u.

The receiver detects x by treating u as part of the noise. To do so, we treat the sum of interference and noise as an effective noise, $n' \triangleq u + n$, following a probability density function given by a Gaussian mixture,

$$f_{N'}(n') = \sum_{u \in \mathcal{U}} P(u) f_N(n' - u), \tag{B.9}$$

where $f_N(n)$ is the probability density function of n. The optimum soft-output maximum a posteriori (MAP) de-mapper for x in the presence of n' can use the log of the ratio of a posteriori of the l-th coded bit carried by x, denoted by c_l, as follows

$$\begin{aligned}
\Lambda_l &= \log \frac{P(c_l = 0|y)}{P(c_l = 1|y)} \\
&= \log \frac{\sum_{x \in \mathcal{X}_l^0} P(y, x)}{\sum_{x \in \mathcal{X}_l^1} P(y, x)} \\
&= \log \frac{\sum_{x \in \mathcal{X}_l^0} P(y|x) P(x)}{\sum_{x \in \mathcal{X}_l^1} P(y|x) P(x)} \\
&= \log \frac{\sum_{x \in \mathcal{X}_l^0} f_{N'}(y - x) P(x)}{\sum_{x \in \mathcal{X}_l^1} f_{N'}(y - x) P(x)} \\
&= \log \frac{\sum_{x \in \mathcal{X}_l^0} \sum_{u \in \mathcal{U}} f_N(y - x - u) P(x) P(u)}{\sum_{x \in \mathcal{X}_l^1} \sum_{u \in \mathcal{U}} f_N(y - x - u) P(x) P(u)},
\end{aligned} \tag{B.10}$$

where $\mathcal{X}_l^0$ and $\mathcal{X}_l^1$ are the set of symbols with the c_l-th bit being 0 or 1.

Note that the last expression in (B.10) is exactly the same expression to use if the receiver jointly de-maps x and u. Hence, we conclude that jointly de-mapping the desired signal and interference is equivalent to the optimal detector for signals corrupted by finite-constellation interference and noise when interference is treated as noise.

B.3 Channel realizations

The channel realizations used to produce Figure 12 are as follows.

Channel 1:

$$[\mathbf{h}_1\ \mathbf{h}_2] = \begin{bmatrix} -0.3010 - 1.0804j & -0.6889 - 0.3357j \\ -0.3562 + 0.8451j & 0.7694 + 0.0572j \\ 0.5462 - 0.2465j & -1.0358 - 0.1752j \\ -0.1919 - 0.1582j & 1.0553 + 0.5879j \end{bmatrix} \tag{B.11}$$

Channel 2:

$$[\mathbf{h}_1\ \mathbf{h}_2] = \begin{bmatrix} 0.9101 + 0.3530j & -0.0353 + 0.8155j \\ -1.0081 - 0.7742j & 0.3403 - 0.6657j \\ 0.7555 + 1.2008j & -0.3133 + 0.4655j \\ -0.2739 - 1.3710j & 0.1513 - 0.5489j \end{bmatrix} \tag{B.12}$$

B.4 Code rates used in LLS

This appendix provides code rates used in the LLSs to generate Figure 3.12-3.17. The code rates for SDMA are uniform across all the users and can be computed based on the throughput given in the captions. For example, in Figure 3.12, the throughput of 3 bits/symbol and QPSK leads to a code rate of 0.75 for both users. The code rates for RSMA schemes are selected to reach good operating points, i.e., both common and private streams achieve 10^{-4} BER at similar SNRs. These code rates are summarized in Table. II in the format of "common stream code rate + private stream code rate" and, in all cases, all the private streams share the same code rate. In Figure 3.17, SIC-type and non-SIC-type receivers use the same code rates as soft CWIC 1 and joint de-mapping respectively in Figure 3.12.

Table B.1: Code rates used for RSMA in link-level simulations.

	Fig. 3.12	Fig. 3.13	Fig. 3.14	Fig. 3.15	Fig. 3.16, $\Delta = \pi/18, \pi/9, \pi/6$		
Soft CWIC 1 (finite)	$0.81 + 0.345$	$0.75 + 0.375$	$0.88 + 0.46$	$0.78 + 0.4$	$0.8 + 0.35$	$0.7 + 0.4$	$0.6 + 0.45$
Joint de-mapping (finite)	$0.85 + 0.325$	$0.78 + 0.36$	$0.94 + 0.43$	$0.84 + 0.36$	-	-	-
Soft CWIC 1 (Gaussian)	$0.815 + 0.3425$	$0.7 + 0.4$	$0.8 + 0.5$	$0.66 + 0.48$	-	-	-
Joint de-mapping (Gaussian)	$0.82 + 0.34$	$0.75 + 0.375$	$0.8 + 0.5$	$0.69 + 0.46$	-	-	-

Appendix C

Appendices for Chapter 4

C.1 Proof of Proposition 4.1

If $\{u_1, ..., u_K\}$ are all of the same value, it is trivial to see that $\mathbf{c}$ does not affect the objective function of (4.13). Hence, $c_i = 1$ and $c_j = 0$, $\forall j \in \mathcal{K} \setminus i$, is one solution for $\mathbf{c}$ in (4.13). If $\{u_1, ..., u_K\}$ are of different values, we need the following lemma.

Lemma C.1. *If $u_i > u_j$, $\forall j \in \mathcal{K} \setminus i$, for a given $\mathbf{P}$, the optimal $\mathbf{c}$ of (4.13) satisfies $c_i^\star = 1$ and $c_j^\star = 0, \forall j \in \mathcal{K} \setminus \{i\}$.*

Proof of Lemma C.1. Suppose $\mathbf{P}$ is given and the private rates, $R_{\mathrm{p},k}, \forall k \in \mathcal{K}$, are not functions of $\mathbf{c}$, (4.13) is equivalent to

$$\begin{aligned} \max_{\mathbf{c}} \ & R_{\mathrm{c}} \sum_{k \in \mathcal{K}} u_k c_k \\ \text{s.t.} \ & \sum_{k \in \mathcal{K}} c_k \leq 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}^{1 \times K}, \end{aligned} \tag{C.1}$$

which is a simple linear programming, and it is easy to see that Lemma C.1 is true. □

Suppose $(\mathbf{P}^\star, \mathbf{c}^\star)$ is optimal in (4.13), and $\mathbf{c}^\star$ does not satisfy $c_i = 1$ and $c_j = 0, \forall j \in \mathcal{K} \setminus i$. Then, according to Lemma C.1, we can always find another $\mathbf{c}$ that gives a higher WSR than $\mathbf{c}^\star$ given that $\mathbf{P} = \mathbf{P}^\star$ in (4.13). By contradiction, $\mathbf{c}^\star$ must satisfy $c_i = 1$ and $c_j = 0$, $\forall j \in \mathcal{K} \setminus i$.

In the cases where $\arg\max_{k\in\mathcal{K}} u_k$ is a set of more than one element, a trivial extension of the above reasoning implies that any $\mathbf{c}$ allocating zero rate to users who are not the most weighted is an optimal solution. This includes $c_i^\star = 1$ and $c_j^\star = 0, \forall j \in \mathcal{K}\backslash i$, where $i \in \arg\max_{k\in\mathcal{K}} u_k$.

C.2 Proof of Proposition 4.2

(4.19) can be written as follows.

$$\begin{aligned} \max_{\mathbf{c}} \quad & \min_k c_k R_\mathrm{c} + R_{\mathrm{p},k} \\ \text{s.t.} \quad & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}^{1\times K} \end{aligned} \tag{C.2}$$

We introduce a slack variable, η, and convert (C.2) into

$$\begin{aligned} \max_{\mathbf{c},t} \quad & \eta \\ \text{s.t.} \quad & \sum_{k\in\mathcal{K}} c_k = 1 \\ & \mathbf{c} \succcurlyeq \mathbf{0}^{1\times K} \\ & c_k R_\mathrm{c} + R_{\mathrm{p},k} \geq \eta, \forall k. \end{aligned} \tag{C.3}$$

Let $(\mathbf{c}^\star, \eta^\star)$ denote the solution of (C.3). Without loss of generality, we assume $R_{\mathrm{p},1} < R_{\mathrm{p},2} < ... < R_{\mathrm{p},K}$, and the solution of (C.3) leads to $c_k^\star > 0$, $\forall k \in \{1, ..., k'\}$ with $1 \leq k' \leq K$, and $c_k^\star = 0$, $\forall k \in \{k'+1, ..., K\}$. We have the following lemma.

Lemma C.2. *$c_k^\star R_c + R_{p,k} = \eta^\star$, $\forall k \in \{1, ..., k'\}$, and with some value for $\eta^\star$.*

Proof. It is trivial to see that $R_{\mathrm{p},k} < \eta^\star$, $\forall k \in \{1, ..., k'\}$. Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be a partition of $\{1, ..., k'\}$. Suppose $(\mathbf{c}^\star, \eta^\star)$ leads to $c_k^\star R_\mathrm{c} + R_{\mathrm{p},k} > \eta^\star$ for $k \in \mathcal{K}_1$, and $c_k^\star R_\mathrm{c} + R_{\mathrm{p},k} = \eta^\star$ for $k \in \mathcal{K}_2$. Then, another choice of $\mathbf{c}$, denoted by $\mathbf{c}^\diamond$, can be found by $\mathbf{c}^\diamond = \alpha \mathbf{c}^\circ$, where $c_k^\circ R_\mathrm{c} = (c_k^\star R_\mathrm{c} - R_{\mathrm{p},k} + \eta^\star)/2$ for $k \in \mathcal{K}_1$, $c_k^\circ = c_k^\star$ for $k \in \mathcal{K}_2$, and $\alpha = 1/\sum_{k\in\mathcal{K}} c_k^\circ$. It is trivial to see that $\alpha > 1$. Therefore, $R_\mathrm{c} \min_{k\in\{1,...,k'\}} c_k^\circ + R_{\mathrm{p},k} > R_\mathrm{c} \min_{k\in\{1,...,k'\}} c_k^\star + R_{\mathrm{p},k}$, which indicates that $\mathbf{c}^\star$ is not the solution of (C.3). By contradiction, Lemma C.2 is proved. □

Lemma C.2 can be re-written as

$$R_{\mathrm{c}}\mathbf{c}' + \mathbf{d} = \eta \mathbf{1}^{1\times k'}, \tag{C.4}$$

where $\mathbf{d} = [R_{\mathrm{p},1}, ..., R_{\mathrm{p},k'}]$, and $\mathbf{c}' = [c_1, ..., c_{k'}]$. Combining (C.4) and the constraint on $\sum_{k\in\mathcal{K}} c_k = 1$, the global optimal solution of $\mathbf{c}'$ and η, denoted by $\mathbf{c}'^{\star}$ and $\eta^{\star}$, can then be obtained by

$$\mathbf{c}'^{\star} = \frac{\eta^{\star}\mathbf{1}^{1\times k'} - \mathbf{d}}{R_{\mathrm{c}}}, \tag{C.5}$$

and

$$\eta^{\star} = \frac{R_{\mathrm{c}} + \mathbf{d}\mathbf{1}^{k'\times 1}}{k'}. \tag{C.6}$$

Although k', i.e., the number of users strictly served by the common stream, is still not known, we can test it using (C.5) and (C.6). This is because allocating the common stream rate based on (C.5) and (C.6) to more users than k' will lead to $\mathbf{c}'$ containing negative entries, which is infeasible. Hence, by iteratively guessing the number of users to be served by the common stream, Algorithm 3 is obtained.

C.3 Transmission mode dictionaries used for numerical results

The transmission mode dictionaries used to produce the numerical results in Section 4.4 are depicted in Tables C.1-C.4.

Table C.1: Transmission mode dictionary for $K = 2$ and $R_{\max} = 6$ bits.

Mode	1	2	3	4
s_{c}	-	QPSK	16QAM	64QAM
$s_{\mathrm{p},k}$	8QAM	QPSK	BPSK	-

Table C.2: Transmission mode dictionary for $K = 2$ and $R_{\max} = 8$ bits.

Mode	1	2	3	4	5
s_{c}	-	QPSK	16QAM	64QAM	256QAM
$s_{\mathrm{p},k}$	16QAM	8QAM	QPSK	BPSK	-

Table C.3: Transmission modes dictionary for $K = 3$ and $R_{\max} = 6$ bits.

Mode	1	2	3
s_{c}	-	8QAM	64QAM
$s_{\mathrm{p},k}$	QPSK	BPSK	-

Table C.4: Transmission modes dictionary for $K = 3$ and $R_{\max} = 9$ bits.

Mode	1	2	3	4
s_{c}	-	8QAM	64QAM	512QAM
$s_{\mathrm{p},k}$	8QAM	QPSK	BPSK	-

Appendix D

Appendices for Chapter 5

D.1 Proof of Proposition 5.1

Since uniformly distributed input is assumed, the capacity is given by the mutual information under uniform distribution, i.e., $I_{\mathrm{MI}}(\mathcal{X}, \mathcal{I}, \mathcal{J}, \sigma_z^2) = I(X;Y)$. The latter can be expanded using the chain rule of mutual information into

$$\begin{aligned} I(X;Y) =& I(X,I,J;Y) - I(I,J;Y \mid X) \\ =& H(X,I,J) - H(X,I,J \mid Y) - H(I,J \mid X) + H(I,J \mid X,Y) \\ =& \log(|\mathcal{X}|) - H(X,I,J \mid Y) + H(I,J \mid Y-X). \end{aligned} \tag{D.1}$$

The last two conditional entropy terms can be computed by definition, resulting in (5.9).

D.2 Proof of Proposition 5.2

For any mismatched memoryless channel with channel transition probability $P_{Y|X}(y|x)$, decoding metric $q(x,y)$ under i.i.d. random coding with input distribution $P_X(x)$, the GMI is given by [61] (a proof is also given in Section 2.3)

$$\begin{aligned} I_{\mathrm{GMI}} =& \sup_{s \geq 0} \mathbb{E}_{X,Y} \left[\log \frac{q(X,Y)^s}{\mathbb{E}_{\bar{X}} \left[q(\bar{X},Y)^s \mid Y \right]} \right] \\ =& \sup_{s \geq 0} s \mathbb{E}_{X,Y} \left[\log q(X,Y) \right] - \mathbb{E}_Y \left[\log \mathbb{E}_{\bar{X}} \left[q(\bar{X},Y)^s \mid Y \right] \right], \end{aligned} \tag{D.2}$$

where $(X, Y, \bar{X}) \sim P_X(\bar{x})P_{Y|X}(y|x)P_X(x)$.

Under the assumptions given in Section 5.1.1, one can obtain

$$\mathbb{E}_{X,Y}\left[\log q(X,Y)\right] = \frac{1}{|\mathcal{X}||\mathcal{I}||\mathcal{J}|}\sum_{x\in\mathcal{X}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\mathbb{E}_Z\left[\log q(x, x+i+j+Z)\right], \tag{D.3}$$

and

$$\mathbb{E}_{\bar{X}}\left[q(\bar{X},Y)^s \mid Y\right] = \frac{1}{|\mathcal{X}|}\sum_{\bar{x}\in\mathcal{X}} q(\bar{x},Y)^s. \tag{D.4}$$

Based on (D.4), (D.5) can be obtained. By substituting (D.3) and (D.5) into (D.2), Proposition 5.1 is proved. To derive the GMI under non-uniformly distributed input signals and/or interference, one only needs to apply the corresponding distribution when evaluating the expectations in (D.3), (D.4), and (D.5).

$$\mathbb{E}_Y\left[\log\ \mathbb{E}_{\bar{X}}\left[q(\bar{X},\ Y)^s \mid Y\right]\right] = -\log \mid \mathcal{X} \mid + \frac{1}{\mid \mathcal{X} \mid\mid \mathcal{I} \mid\mid \mathcal{J} \mid} \sum_{x\in\mathcal{X},\ i\in\mathcal{I},\ j\in\mathcal{J}} \mathbb{E}\left[\log \sum_{\bar{x}\in\mathcal{X}} q(\bar{x},\ x+i+j+Z)^s\right] \tag{D.5}$$

$$\begin{aligned} I_{\mathrm{MI}}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2) =& \log|\mathcal{X}| - \mathbb{E}\left[\log\left(\sum_{\bar{i}\in\mathcal{I}/I,\ \bar{x}\in\mathcal{X},\ \bar{j}\in\mathcal{J}} \exp\left(\frac{-|X+I+J+Z-\bar{x}-\bar{i}-\bar{j}|^2}{\sigma_z^2}\right) + \sum_{\bar{x}\in\mathcal{X},\ \bar{j}\in\mathcal{J}} \exp\left(\frac{-|X+J+Z-\bar{x}-\bar{j}|^2}{\sigma_z^2}\right)\right)\right] \\ &+ \mathbb{E}\left[\log\left(\sum_{\bar{i}\in\mathcal{I}/I,\ \bar{j}\in\mathcal{J}} \exp\left(\frac{-|I+J+Z-\bar{i}-\bar{j}|^2}{\sigma_z^2}\right) + \sum_{\bar{j}\in\mathcal{J}} \exp\left(\frac{-|J+Z-\bar{j}|^2}{\sigma_z^2}\right)\right)\right] \end{aligned} \tag{D.6}$$

$$\begin{aligned} \lim_{\sigma_i^2\to 0} I_{\mathrm{MI}}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2) =& \log|\mathcal{X}| - \mathbb{E}\left[\log\left(|\mathcal{I}| \sum_{\bar{x}\in\mathcal{X},\ \bar{j}\in\mathcal{J}} \exp\left(\frac{-|X+J+Z-\bar{x}-\bar{j}|^2}{\sigma_z^2}\right)\right)\right] + \mathbb{E}\left[\log\left(|\mathcal{I}| \sum_{\bar{j}\in\mathcal{J}} \exp\left(\frac{-|J+Z-\bar{j}|^2}{\sigma_z^2}\right)\right)\right] \\ =& \log|\mathcal{X}| - \mathbb{E}\left[\log \sum_{\bar{x}\in\mathcal{X},\ \bar{j}\in\mathcal{J}} \exp\left(\frac{-|X+J+Z-\bar{x}-\bar{j}|^2}{\sigma_z^2}\right)\right] + \mathbb{E}\left[\log \sum_{\bar{j}\in\mathcal{J}} \exp\left(\frac{-|J+Z-\bar{j}|^2}{\sigma_z^2}\right)\right] \\ =& I_{\mathrm{MI}}(\mathcal{X},\ \{0\},\ \mathcal{J},\ \sigma_z^2) \end{aligned} \tag{D.7}$$

$$\begin{aligned} \lim_{\sigma_i^2\to+\infty} I_{\mathrm{MI}}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2) =& \log|\mathcal{X}| - \mathbb{E}\left[\log \sum_{\bar{x}\in\mathcal{X},\ \bar{j}\in\mathcal{J}} \exp\left(\frac{-|X+J+Z-\bar{x}-\bar{j}|^2}{\sigma_z^2}\right)\right] + \mathbb{E}\left[\log \sum_{\bar{j}\in\mathcal{J}} \exp\left(\frac{-|J+Z-\bar{j}|^2}{\sigma_z^2}\right)\right] \\ =& I_{\mathrm{MI}}(\mathcal{X},\ \{0\},\ \mathcal{J},\ \sigma_z^2) \end{aligned} \tag{D.8}$$

$$I_{\mathrm{GMI}}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2) = \sup_{s\geq 0}\ \log|\mathcal{X}| - \mathbb{E}\left[\log \sum_{\bar{x}\in\mathcal{X}} \left(\sum_{\bar{i}\in\mathcal{I}} \exp\left(-\frac{|X+I+J+Z-\bar{i}-\bar{x}|^2}{\sigma_j^2+\sigma_z^2}\right)\right)^s\right] + s\mathbb{E}\left[\log \sum_{\bar{i}\in\mathcal{I}} \exp\left(-\frac{|I+J+Z-\bar{i}|^2}{\sigma_j^2+\sigma_z^2}\right)\right] \tag{D.9}$$

D.3 Proof of Proposition 5.4

(5.9) can be expanded by separating the exponential terms according to whether $I = \bar{i}$ or not, resulting in (D.6). If $I \neq \bar{i}$, $\lim_{\sigma_i^2 \to 0} I - \bar{i} = 0$. Therefore, for terms with $I \neq \bar{i}$,

$$\lim_{\sigma_i^2 \to 0} \exp\left(\frac{-|X + I + J + Z - \bar{x} - \bar{i} - \bar{j}|^2}{\sigma_z^2}\right) = \exp\left(\frac{-|X + J + Z - \bar{x} - \bar{j}|^2}{\sigma_z^2}\right), \tag{D.10}$$

and

$$\lim_{\sigma_i^2 \to 0} \exp\left(\frac{-|I + J + Z - \bar{i} - \bar{j}|^2}{\sigma_z^2}\right) = \exp\left(\frac{-|J + Z - \bar{j}|^2}{\sigma_z^2}\right), \tag{D.11}$$

leading to (D.7).

From the reverse triangle inequality,

$$\exp\left(\frac{-|X + I + J + Z - \bar{x} - \bar{i} - \bar{j}|^2}{\sigma_z^2}\right) \leq \exp\left(-\frac{|I - \bar{i}|^2}{\sigma_z^2}\right) \exp\left(\frac{|X + J + Z - \bar{x} - \bar{j}|^2}{\sigma_z^2}\right). \tag{D.12}$$

If $I \neq \bar{i}$,

$$\lim_{\sigma_i^2 \to +\infty} \exp\left(-\frac{|I - \bar{i}|^2}{\sigma_z^2}\right) = 0. \tag{D.13}$$

Since the exponential is nonnegative, the squeeze theorem yields

$$\lim_{\sigma_i^2 \to +\infty} \exp\left(\frac{-|X + I + J + Z - \bar{x} - \bar{i} - \bar{j}|^2}{\sigma_z^2}\right) = 0. \tag{D.14}$$

Following a similar approach, it can be proved that

$$\lim_{\sigma_i^2 \to +\infty} \exp\left(\frac{-|I + J + Z - \bar{i} - \bar{j}|^2}{\sigma_z^2}\right) = 0, \tag{D.15}$$

if $I \neq \bar{i}$. Applying (D.14) and (D.15) to (D.6) results in (D.8).

The proof of

$$\begin{aligned} \lim_{\sigma_i^2 \to 0} I_{\mathrm{GMI}}(\mathcal{X},\, \mathcal{I},\, \mathcal{J},\, \sigma_z^2) &= \lim_{\sigma_i^2 \to +\infty} I_{\mathrm{GMI}}(\mathcal{X},\, \mathcal{I},\, \mathcal{J},\, \sigma_z^2) \\ &= I_{\mathrm{GMI}}(\mathcal{X},\, \{0\},\, \mathcal{J},\, \sigma_z^2) \end{aligned} \tag{D.16}$$

with $q(x, y)$ given by (5.7), it follows similar steps and is therefore skipped for clarity.

D.4 Proof of Proposition 5.5

With $q(x,\ y)$ given by (5.7), (5.10) is written as (D.9). From the reverse triangle inequality,

$$\exp\left(-\frac{|X+I+J+Z-\bar{i}-\bar{x}|^2}{\sigma_j^2+\sigma_z^2}\right) \leq \exp\left(-\frac{|J|^2}{\sigma_j^2+\sigma_z^2}\right)\exp\left(\frac{|X+I+Z-\bar{i}-\bar{x}|^2}{\sigma_j^2+\sigma_z^2}\right), \tag{D.17}$$

and

$$\lim_{\sigma_j^2\to+\infty}\exp\left(-\frac{|J|^2}{\sigma_j^2+\sigma_z^2}\right)\exp\left(\frac{|X+I+Z-\bar{i}-\bar{x}|^2}{\sigma_j^2+\sigma_z^2}\right) = \exp\left(-\frac{|J|^2}{\sigma_j^2}\right). \tag{D.18}$$

Similarly,

$$\exp\left(-\frac{|I+J+Z-\bar{i}|^2}{\sigma_j^2+\sigma_z^2}\right) \leq \exp\left(-\frac{|J|^2}{\sigma_j^2+\sigma_z^2}\right)\exp\left(\frac{|I+Z-\bar{i}|^2}{\sigma_j^2+\sigma_z^2}\right), \tag{D.19}$$

and

$$\lim_{\sigma_j^2\to+\infty}\exp\left(-\frac{|J|^2}{\sigma_j^2+\sigma_z^2}\right)\exp\left(\frac{|I+Z-\bar{i}|^2}{\sigma_j^2+\sigma_z^2}\right) = \exp\left(-\frac{|J|^2}{\sigma_j^2}\right). \tag{D.20}$$

Therefore, an upper bound on (D.9) follows from (D.17) and (D.19) and yields

$$\begin{aligned}
&\lim_{\sigma_j^2\to+\infty} I_{\mathrm{GMI}}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2)\\
&\leq \sup_{s\geq 0}\ \log|\mathcal{X}| - \mathbb{E}\left[\log\left(|\mathcal{X}|\left(|\mathcal{I}|\exp\left(-\frac{|J|^2}{\sigma_j^2}\right)\right)^s\right)\right]\\
&\qquad + s\mathbb{E}\left[\log\left(|\mathcal{I}|\exp\left(-\frac{|J|^2}{\sigma_j^2}\right)\right)\right]\\
&= \sup_{s\geq 0}\ -s\left(\log|\mathcal{I}| + \mathbb{E}\left[-\frac{|J|^2}{\sigma_j^2}\right]\right) + s\left(\log|\mathcal{I}| + \mathbb{E}\left[-\frac{|J|^2}{\sigma_j^2}\right]\right)\\
&= 0.
\end{aligned} \tag{D.21}$$

Since the rate is nonnegative, we conclude that

$$\lim_{\sigma_j^2\to+\infty} I_{\mathrm{GMI}}(\mathcal{X},\ \mathcal{I},\ \mathcal{J},\ \sigma_z^2) = 0. \tag{D.22}$$

Appendix E

Appendix for Chapter 6

E.1 Proof of the achievability of (6.8)

We consider a CS-RSMA decoder at user-k that takes $\mathbf{y}_k$ as the input and outputs an estimate of the message, which is written in the following form

$$\hat{m} = \underset{m=1,...,M}{\arg\max}\ q(\mathbf{s}_{\mathrm{c},k}^{m}, \mathbf{s}_{\mathrm{p},k}^{m}, \mathbf{y}_k), \tag{E.1}$$

where $[\mathbf{s}_{\mathrm{c},k}^{m};\ \mathbf{s}_{\mathrm{p},k}^{m}] = \mathbf{s}_{k}^{\mathrm{cw},m}$ is the codeword for user-k corresponding to message m. We assume without loss of generality that $\mathbf{s}_{\mathrm{c},k}$ occupies $\mathbf{s}_{\mathrm{c}}$ at $t = 1, ..., n_k - n$. (2.1) can be then rewritten as

$$y_{t,k} = \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{t,\mathrm{c},k} + \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k} s_{t,\mathrm{p},k} + \tilde{z}_{t,k},\ \ \text{for}\ \ t = 1, ..., n_k - n \tag{E.2}$$

and

$$y_{t,k} = \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k} s_{t,\mathrm{p},k} + \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{t,\mathrm{c},k'} + \tilde{z}_{t,k},$$
$$\text{for}\ \ t = n_k - n + 1, ..., n\ \ \text{and}\ k' \in \mathcal{K}/k, \tag{E.3}$$

where $s_{t,\mathrm{p},k}$ is the private stream symbol intended for user-k at time instance t, $s_{t,\mathrm{c},k}$ is the common stream symbol at time instance t and is from $\mathbf{s}_{\mathrm{c},k}$, and $\tilde{z}_{t,k}$ contains interference from the private streams intended for other users and the AWGN noise, i.e.,

$$\tilde{z}_{t,k} = \sum_{k^\circ \in \mathcal{K}/k} \mathbf{h}_k^H \mathbf{p}_{\mathrm{p},k^\circ} s_{t,\mathrm{p},k^\circ} + z_{t,k}. \tag{E.4}$$

Since the channel is memoryless across different time indices, we consider a decoding metric that can be decomposed as follows.

$$\begin{aligned}\hat{m} &= \arg\max_{m=1,\ldots,M} q(\mathbf{s}_{\mathrm{c},k}^m, \mathbf{s}_{\mathrm{p},k}^m, \mathbf{y}_k) \\ &= \arg\max_{m=1,\ldots,M} \prod_{t=1}^{n_k-n} q_a(s_{t,\mathrm{c},k}^m, s_{t,\mathrm{p},k}^m, y_{t,k}) \prod_{t=n_k-n+1}^{n_k} q_b(s_{t,\mathrm{p},k}^m, y_{t,k})^{\rho_b}\end{aligned} \tag{E.5}$$

We further consider suboptimal decoding metrics that are separable in $s_{t,\mathrm{c},k}$ and $s_{t,\mathrm{p},k}$, i.e.,

$$q_a(s_{t,\mathrm{c},k}, s_{t,\mathrm{p},k}, y_{t,k}) = q_{a,\mathrm{c}}(s_{t,\mathrm{c},k}, y_{t,k})^{\rho_{a,\mathrm{c}}} q_{a,\mathrm{p}}(s_{t,\mathrm{p},k}, y_{t,k})^{\rho_{a,\mathrm{p}}}. \tag{E.6}$$

$\rho_{a,\mathrm{c}}$, $\rho_{a,\mathrm{p}}$ and ρ_b are positive parameters. We evaluate the achievable rate using the GMI. With capital letters representing the random variables and the time indices omitted when unnecessary, the GMI for user-k can be derived as in (E.7).

$$
\begin{aligned}
I_{\mathrm{GMI}} &= \sup_{s>0,\rho_{a,\mathrm{c}}>0,\rho_{a,\mathrm{p}}>0,\rho_b>0} \frac{1}{n}\mathbb{E}\left[\log \frac{q(\mathbf{S}_{\mathrm{c},k},\mathbf{S}_{\mathrm{p},k},\mathbf{Y}_k)^s}{\mathbb{E}\left[q(\bar{\mathbf{S}}_{\mathrm{c},k},\bar{\mathbf{S}}_{\mathrm{p},k},\mathbf{Y}_k)^s \mid \mathbf{Y}_k\right]}\right] \\
&= \sup_{s>0,\rho_{a,\mathrm{c}}>0,\rho_{a,\mathrm{p}}>0,\rho_b>0} \frac{1}{n}\mathbb{E}\left[\log \frac{\prod_{t=1}^{n_k-n} q_{a,\mathrm{c}}(S_{t,\mathrm{c},k},Y_{t,k})^{\rho_{a,\mathrm{c}}s} q_{a,\mathrm{p}}(S_{t,\mathrm{p},k},Y_{t,k})^{\rho_{a,\mathrm{p}}s} \prod_{t=n_k-n+1}^{n_k} q_b(S_{t,\mathrm{p},k},Y_{t,k})^{\rho_b s}}{\mathbb{E}\left[\prod_{t=1}^{n_k-n} q_{a,\mathrm{c}}(\bar{S}_{t,\mathrm{c},k},Y_{t,k})^{\rho_{a,\mathrm{c}}s} q_{a,\mathrm{p}}(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^{\rho_{a,\mathrm{p}}s} \prod_{t=n_k-n+1}^{n_k} q_b(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^{\rho_b s} \mid \mathbf{Y}_k\right)\right]} \right] \\
&= \sup_{s_1>0,s_2>0,s_3>0} \\
&\quad \frac{1}{n}\mathbb{E}\left[\log \frac{\prod_{t=1}^{n_k-n} q_{a,\mathrm{c}}(S_{t,\mathrm{c},k},Y_{t,k})^{s_1} q_{a,\mathrm{p}}(S_{t,\mathrm{p},k},Y_{t,k})^{s_2} \prod_{t=n_k-n+1}^{n_k} q_b(S_{t,\mathrm{p},k},Y_{t,k})^{s_3}}{\mathbb{E}\left[\prod_{t=1}^{n_k-n} q_{a,\mathrm{c}}(\bar{S}_{t,\mathrm{c},k},Y_{t,k})^{s_1} \mid \mathbf{Y}_k\right]\mathbb{E}\left[\prod_{t=1}^{n_k-n} q_{a,\mathrm{p}}(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^{s_2} \mid \mathbf{Y}_k\right]\mathbb{E}\left[\prod_{t=n_k-n+1}^{n_k} q_b(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^{s_3} \mid \mathbf{Y}_k\right]}\right] \\
&= \sup_{s>0} \frac{1}{n}\mathbb{E}\left[\log \frac{\prod_{t=1}^{n_k-n} q_{a,\mathrm{c}}(S_{t,\mathrm{c},k},Y_{t,k})^s}{\mathbb{E}\left[\prod_{t=1}^{n_k-n} q_{a,\mathrm{c}}(\bar{S}_{t,\mathrm{c},k},Y_{t,k})^s \mid \mathbf{Y}_k\right]}\right] + \sup_{s>0} \frac{1}{n}\mathbb{E}\left[\log \frac{\prod_{t=1}^{n_k-n} q_{a,\mathrm{p}}(S_{t,\mathrm{p},k},Y_{t,k})^s}{\mathbb{E}\left[\prod_{t=1}^{n_k-n} q_{a,\mathrm{p}}(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^s \mid \mathbf{Y}_k\right]}\right] \\
&\quad + \sup_{s>0} \frac{1}{n}\mathbb{E}\left[\log \frac{\prod_{t=n_k-n+1}^{n_k} q_b(S_{t,\mathrm{p},k},Y_{t,k})^s}{\mathbb{E}\left[\prod_{t=n_k-n+1}^{n_k} q_b(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^s \mid \mathbf{Y}_k\right]}\right] \\
&= \sup_{s>0} \frac{n_k-n}{n}\mathbb{E}\left[\log \frac{q_{a,\mathrm{c}}(S_{t,\mathrm{c},k},Y_{t,k})^s}{\mathbb{E}\left[q_{a,\mathrm{c}}(\bar{S}_{t,\mathrm{c},k},Y_{t,k})^s \mid Y_k\right]}\right] + \sup_{s>0} \frac{n_k-n}{n}\mathbb{E}\left[\log \frac{q_{a,\mathrm{p}}(S_{t,\mathrm{p},k},Y_{t,k})^s}{\mathbb{E}\left[q_{a,\mathrm{p}}(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^s \mid Y_k\right]}\right] \\
&\quad + \sup_{s>0} \frac{2n-n_k}{n}\mathbb{E}\left[\log \frac{q_b(S_{t,\mathrm{p},k},Y_{t,k})^s}{\mathbb{E}\left[q_b(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^s \mid Y_k\right]}\right] \\
&= \sup_{s>0} c_n\mathbb{E}\left[\log \frac{q_{a,\mathrm{c}}(S_{t,\mathrm{c},k},Y_{t,k})^s}{\mathbb{E}\left[q_{a,\mathrm{c}}(\bar{S}_{t,\mathrm{c},k},Y_{t,k})^s \mid Y_k\right]}\right] + \sup_{s>0} c_n\mathbb{E}\left[\log \frac{q_{a,\mathrm{p}}(S_{t,\mathrm{p},k},Y_{t,k})^s}{\mathbb{E}\left[q_{a,\mathrm{p}}(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^s \mid Y_k\right]}\right] \\
&\quad + \sup_{s>0} (1-c_n)\mathbb{E}\left[\log \frac{q_b(S_{t,\mathrm{p},k},Y_{t,k})^s}{\mathbb{E}\left[q_b(\bar{S}_{t,\mathrm{p},k},Y_{t,k})^s \mid Y_k\right]}\right]
\end{aligned}
\tag{E.7}
$$

Under the channel model in (E.2) and (E.3) and the suboptimal decoding setup given in Section 6.2.1, i.e., treating undesired private streams as Gaussian random variables, we have

$$\sup_{s>0} \ \mathbb{E}\left[\log \ \frac{q_{a,\mathrm{c}}(S_{t,\mathrm{c},k}, Y_{t,k})^s}{\mathbb{E}\left[q_{a,\mathrm{c}}(\bar{S}_{t,\mathrm{c},k}, Y_{t,k})^s \mid Y_k\right)]}\right] = I_{\mathrm{c},k} \tag{E.8}$$

and

$$\begin{aligned} &\sup_{s>0} \ \mathbb{E}\left[\log \ \frac{q_{a,\mathrm{p}}(S_{t,\mathrm{p},k}, Y_{t,k})^s}{\mathbb{E}\left[q_{a,\mathrm{p}}(\bar{S}_{t,\mathrm{p},k}, Y_{t,k})^s \mid Y_k\right)]}\right] \\ =&\sup_{s>0} \ \mathbb{E}\left[\log \ \frac{q_{b}(S_{t,\mathrm{p},k}, Y_{t,k})^s}{\mathbb{E}\left[q_{b}(\bar{S}_{t,\mathrm{p},k}, Y_{t,k})^s \mid Y_k\right)]}\right] \\ =&I_{\mathrm{p},k}^{\text{non-SIC}}. \end{aligned} \tag{E.9}$$

Substituting (E.8) and (E.9) into (E.7), we obtain (6.8).

Remark E.1. *This proof is strongly related to the proof in Section IV.D of [64], which proves that for BICM systems that apply a decoding metric that is separable in bits, the GMI equals the rate of the virtual parallel channels that carry the bits separately.*

E.2 Proof of Proposition 6.3

(6.24) can be written as follows:

$$\begin{aligned} \max_{\mathbf{c}} \ \ &\min_{k} \ \ c_k I'_{\mathrm{c},k} + I_{\mathrm{p},k} \\ \text{s.t.} \ \ &\mathbf{c} \succcurlyeq \mathbf{0} \\ &\sum_{k\in\mathcal{K}} c_k = 1. \end{aligned} \tag{E.10}$$

For conventional RSMA (either with or without SIC), $I'_{\mathrm{c},k} = \min_k I_{\mathrm{c},k}$, $\forall k \in \mathcal{K}$. For CS-RSMA, $I'_{\mathrm{c},k} = I_{\mathrm{c},k}$, , $\forall k \in \mathcal{K}$.

By introducing a slack variable, ξ, (E.10) can be converted into

$$\begin{aligned} \max_{\mathbf{c}} \quad & \xi \\ \text{s.t.} \quad & \mathbf{c} \succcurlyeq \mathbf{0} \\ & \sum_{k\in\mathcal{K}} c_k = 1 \\ & c_k I'_{\mathrm{c},k} + I_{\mathrm{p},k} \geq \xi, \ \forall k. \end{aligned} \tag{E.11}$$

W.L.O.G., we assume $I_{\mathrm{p},1} < I_{\mathrm{p},2} < ... < I_{\mathrm{p},K}$. Assume that the solution of (E.11) leads to $c_k > 0$ for $k = 1, \ ..., \ k'$. It is easy to see that $c_k I'_{\mathrm{c},k} + I_{\mathrm{p},k} = \xi$ for $k = 1, ..., k'$. Equivalently,

$$\mathbf{c}'\mathbf{A} + \mathbf{b} = \xi\mathbf{1}, \tag{E.12}$$

where $\mathbf{A} = \mathrm{diag}\{[I'_{\mathrm{c},1}, \ ..., \ I'_{\mathrm{c},k'}]\}$, $\mathbf{b} = [I_{\mathrm{p},1}, \ ..., \ I_{\mathrm{p},k'}]^T$, $\mathbf{c}' = [c_1, \ ..., \ c_{k'}]$, and $\mathbf{1}$ is a column vector of length k' whose elements are all ones. The global optimal solution of $\mathbf{c}'$ and ξ, denoted by $\mathbf{c}'^{\star}$ and $\xi^{\star}$, can then be obtained by

$$\mathbf{c}'^{\star} = \xi\mathbf{1}\mathbf{A}^{-1} - \mathbf{b}\mathbf{A}^{-1} \tag{E.13}$$

and

$$\xi^{\star} = \frac{1 + \mathbf{b}^T\mathbf{A}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{A}^{-1}\mathbf{1}}. \tag{E.14}$$

Note that (E.13) is directly from (C.4), and (E.14) is from (C.5) combined the condition of $\sum_{k\in\mathcal{K}} c_k = 1$.

Although k', i.e., the number of users strictly served by the common stream, is still unknown, we can test it using (E.13) and (E.14). This is because allocating the common stream resource to more users than k' will lead to $\mathbf{c}'$ containing negative entries, which is infeasible. Hence, by iteratively guessing the number of users to be served by the common stream, Algorithm 9 is obtained.

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[2] B. Clerckx, Y. Mao, E. A. Jorswieck, J. Yuan, D. J. Love, E. Erkip, and D. Niyato, "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1265–1308, May 2023.

[3] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-splitting multiple access: Fundamentals, survey, and future research trends," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2073–2126, Fourthquarter 2022.

[4] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 364–375, Mar. 1996.

[5] M. M. Şahin, O. Dizdar, B. Clerckx, and H. Arslan, "OFDM-RSMA: Robust transmission under inter-carrier interference," *IEEE Transactions on Communications*, vol. 73, no. 7, pp. 4602–4615, Jul. 2025.

[6] A. Carleial, "Interference channels," *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[7] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[8] S. Shamai, "A broadcast strategy for the gaussian slowly fading channel," in *Proceedings of IEEE International Symposium on Information Theory*, Jun. 1997, pp. 150–.

[9] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4847–4861, Nov. 2016.

[10] E. Piovano and B. Clerckx, "Optimal DoF region of the K -user MISO BC with partial CSIT," *IEEE Communications Letters*, vol. 21, no. 11, pp. 2368–2371, Nov. 2017.

[11] Y. Mao and B. Clerckx, "Beyond dirty paper coding for multi-antenna broadcast channel with partial CSIT: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6775–6791, Nov. 2020.

[12] H. Joudeh and B. Clerckx, "DoF region of the MISO BC with partial CSIT: Proof by inductive fourier-motzkin elimination," in *Proceedings of 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2019, pp. 1–5.

[13] A. Gholami Davoodi, B. Yuan, and S. A. Jafar, "GDoF region of the MISO BC: Bridging the gap between finite precision and perfect CSIT," *IEEE Transactions on Information Theory*, vol. 64, no. 11, pp. 7208–7217, Nov. 2018.

[14] A. G. Davoodi and S. A. Jafar, "Degrees of freedom region of the (M, N1, N2) mimo broadcast channel with partial CSIT: An application of sum-set inequalities," in *Proceedings of 2019 IEEE International Symposium on Information Theory (ISIT)*, Jul. 2019, pp. 1637–1641.

[15] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[16] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming SDMA and NOMA," *EURASIP Journal on Wireless Communications and Networking*, no. 1, p. 133, May 2018.

[17] B. Clerckx, Y. Mao, R. Schober, and H. V. Poor, "Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: A simple two-user rate analysis," *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 349–353, Mar. 2020.

[18] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A rate splitting strategy for massive mimo with imperfect CSIT," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4611–4624, Jul. 2016.

[19] Y. Mao and B. Clerckx, "Beyond dirty paper coding for multi-antenna broadcast channel with partial CSIT: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6775–6791, Nov. 2020.

[20] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A rate splitting strategy for massive MIMO with imperfect CSIT," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4611–4624, Jul. 2016.

[21] M. Dai and B. Clerckx, "Multiuser millimeter wave beamforming strategies with quantized and statistical CSIT," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7025–7038, Nov. 2017.

[22] L. Yin, B. Clerckx, and Y. Mao, "Rate-splitting multiple access for multi-antenna broadcast channels with statistical CSIT," in *Proceedings of 2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Mar. 2021, pp. 1–6.

[23] C. Hao, Y. Wu, and B. Clerckx, "Rate analysis of two-receiver MISO broadcast channel with finite rate feedback: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3232–3246, Sep. 2015.

[24] G. Zhou, Y. Mao, and B. Clerckx, "Rate-splitting multiple access for multi-antenna downlink communication systems: Spectral and energy efficiency tradeoff," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 4816–4828, Jul. 2022.

[25] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting for multi-antenna non-orthogonal unicast and multicast transmission: Spectral and energy efficiency analysis," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8754–8770, Dec. 2019.

[26] O. Dizdar, A. Kaushik, B. Clerckx, and C. Masouros, "Energy efficient dual-functional radar-communication: Rate-splitting multiple access, low-resolution DACs, and RF chain selection," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 986–1006, Jun. 2022.

[27] S. Park, J. Choi, J. Park, W. Shin, and B. Clerckx, "Rate-splitting multiple access for quantized multiuser MIMO communications," *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 7696–7711, Nov. 2023.

[28] J. Zhang, B. Clerckx, J. Ge, and Y. Mao, "Cooperative rate splitting for MISO broadcast channel with user relaying, and performance benefits over cooperative NOMA," *IEEE Signal Processing Letters*, vol. 26, no. 11, pp. 1678–1682, Nov. 2019.

[29] Y. Mao, B. Clerckx, J. Zhang, V. O. K. Li, and M. A. Arafah, "Max-min fairness of K-user cooperative rate-splitting in MISO broadcast channel with user relaying," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6362–6376, Oct. 2020.

[30] H. Cho, B. Ko, B. Clerckx, and J. Choi, "Coverage increase at THz frequencies: A cooperative rate-splitting approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9821–9834, Dec. 2023.

[31] O. Dizdar, Y. Mao, and B. Clerckx, "Rate-splitting multiple access to mitigate the curse of mobility in (massive) MIMO networks," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6765–6780, Oct. 2021.

[32] C. Xu, B. Clerckx, S. Chen, Y. Mao, and J. Zhang, "Rate-splitting multiple access for multi-antenna joint radar and communications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1332–1347, Nov. 2021.

[33] H. Li, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for 6G—part III: Interplay with reconfigurable intelligent surfaces," *IEEE Communications Letters*, vol. 26, no. 10, pp. 2242–2246, Oct. 2022.

[34] H. Li, S. Shen, and B. Clerckx, "Synergizing beyond diagonal reconfigurable intelligent surface and rate-splitting multiple access," *IEEE Transactions Wireless Communications*, vol. 23, no. 8, pp. 8717–8729, Aug. 2024.

[35] A. S. de Sena, P. H. J. Nardelli, D. B. da Costa, P. Popovski, and C. B. Papadias, "Rate-splitting multiple access and its interplay with intelligent reflecting surfaces," *IEEE Communications Magazine*, vol. 60, no. 7, pp. 52–57, Jul. 2022.

[36] L. Yin and B. Clerckx, "Rate-splitting multiple access for multigroup multicast and multibeam satellite systems," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 976–990, February 2021.

[37] Y. Xu, L. Yin, Y. Mao, W. Shin, and B. Clerckx, "Distributed rate-splitting multiple access for multilayer satellite communications," *IEEE Transactions on Communications*, vol. 72, no. 10, pp. 6131–6144, Oct. 2024.

[38] Y. Xu, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access with finite blocklength for short-packet and low-latency downlink communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 12 333–12 337, Nov. 2022.

[39] ——, "Max-min fairness of rate-splitting multiple access with finite blocklength communications," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 5, pp. 6816–6821, February 2023.

[40] L. Yin and B. Clerckx, "Rate-splitting multiple access for satellite-terrestrial integrated networks: Benefits of coordination and cooperation," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 317–332, Jan. 2023.

[41] Y. Zhang, H. Zhao, Y. Mao, W. Xia, W. Lu, and H. Zhu, "Rate-splitting multiple access in cell-free massive MIMO-URLLC systems: Achievable rate analysis and optimization," *IEEE Transactions on Communications*, vol. 72, no. 11, pp. 6752–6767, Nov. 2024.

[42] E. Piovano, H. Joudeh, and B. Clerckx, "Generalized degrees of freedom of the symmetric cache-aided MISO broadcast channel with partial CSIT," *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5799–5815, Sep. 2019.

[43] Y. Mao, E. Piovano, and B. Clerckx, "Rate-splitting multiple access for overloaded cellular internet of things," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4504–4519, Jul. 2021.

[44] Y. Liu and B. Clerckx, "Rate-splitting multiple access for coexistence of semantic and bit communications," *arXiv preprint arXiv:2409.10314*, sep 2024.

[45] J. Xu, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for short-packet uplink communications: A finite blocklength analysis," *IEEE Communications Letters*, vol. 27, no. 2, pp. 517–521, February 2023.

[46] Y. Liu, B. Clerckx, and P. Popovski, "Network slicing for eMBB, URLLC, and mMTC: An uplink rate-splitting multiple access approach," *IEEE Transactions Wireless Communications*, vol. 23, no. 3, pp. 2140–2152, Mar. 2024.

[47] O. Dizdar, A. Sattarzadeh, Y. X. Yap, and S. Wang, "RSMA for overloaded mimo networks: Low-complexity design for max–min fairness," *IEEE Transactions Wireless Communications*, vol. 23, no. 6, pp. 6156–6173, Jun. 2024.

[48] A. R. Flores, R. C. de Lamare, and B. Clerckx, "Linear precoding and stream combining for rate splitting in multiuser MIMO systems," *IEEE Communications Letters*, vol. 24, no. 4, pp. 890–894, Apr. 2020.

[49] O. Dizdar, Y. Mao, W. Han, and B. Clerckx, "Rate-splitting multiple access for downlink multi-antenna communications: Physical layer design and link-level simulations," in *Proceedings of 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, Aug. 2020, pp. 1–6.

[50] H. Chen, D. Mi, Z. Chu, P. Xiao, Y. Xu, and D. He, "Link-level performance of rate-splitting based downlink multiuser MISO systems," in *Proceedings of 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, Aug. 2020, pp. 1–5.

[51] L. Yin, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for multi-group multicast cellular and satellite communications: PHY layer design and link-level simulations," in *Proceedings of 2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.

[52] A. Mishra, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for downlink multiuser MIMO: Precoder optimization and PHY-layer design," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 874–890, February 2022.

[53] R. Cerna Loli, O. Dizdar, B. Clerckx, and C. Ling, "Model-based deep learning receiver design for rate-splitting multiple access," *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 8352–8365, Nov. 2023.

[54] Z. Li, C. Ye, Y. Cui, S. Yang, and S. Shamai, "Rate splitting for multi-antenna downlink: Precoder design and practical implementation," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1910–1924, Aug. 2020.

[55] Z. Li, S. Yang, and S. Shamai, "On linearly precoded rate splitting for gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 67, no. 7, pp. 4693–4709, Jul. 2021.

[56] A. S. de Sena, P. H. J. Nardelli, D. B. da Costa, P. Popovski, C. B. Papadias, and M. Debbah, "Dual-polarized massive MIMO-RSMA networks: Tackling imperfect SIC," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3194–3215, May 2023.

[57] A. Salem, C. Masouros, and B. Clerckx, "Rate splitting with finite constellations: The benefits of interference exploitation vs suppression," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1541–1557, Jun. 2021.

[58] R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Transactions on Information Theory*, vol. 11, no. 1, pp. 3–18, Jan. 1965.

[59] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai Shitz, "On information rates for mismatched decoders," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[60] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.

[61] J. Scarlett, A. G. i Fàbregas, A. Somekh-Baruch, and A. Martinez, "Information-Theoretic Foundations of Mismatched Decoding," *Foundations and Trends® in Communications and Information Theory*, vol. 17, no. 2–3, pp. 149–401, 2020. [Online]. Available: http://dx.doi.org/10.1561/0100000101

[62] A. Lapidoth, "Nearest neighbor decoding for additive non-gaussian noise channels," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1520–1529, Sep. 1996.

[63] N. Binshtok and S. Shamai, "Integer metrics for binary input symmetric output memoryless channels," *IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1636–1645, Nov. 1999.

[64] A. Martinez, A. Guillen i Fabregas, G. Caire, and F. M. J. Willems, "Bit-interleaved coded modulation revisited: A mismatched decoding perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2756–2765, Jun. 2009.

[65] A. Lapidoth, "Mismatched decoding and the multiple-access channel," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1439–1452, Sep. 1996.

[66] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Mismatched multi-letter successive decoding for the multiple-access channel," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2253–2266, Apr. 2018.

[67] A. Lapidoth and S. Shamai, "Fading channels: How perfect need "perfect side information" be?" *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[68] H. Weingarten, Y. Steinberg, and S. Shamai, "Gaussian codes and weighted nearest neighbor decoding in fading multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1665–1686, Aug. 2004.

[69] A. T. Asyhari and A. Guillén i Fàbregas, "Nearest neighbor decoding in MIMO block-fading channels with imperfect CSIR," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1483–1517, Mar. 2012.

[70] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading mimo channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[71] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in *Proceedings of International Symposium onInformation Theory*, Jun. 2004, p. 174.

[72] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," *IEEE Wireless Communications Letters*, vol. 1, no. 1, pp. 38–41, February 2012.

[73] J. Scarlett, A. Martinez, and A. G. i. Fabregas, "Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2647–2666, May 2014.

[74] R. Gallager, *Information Theory and Reliable Communications*. John Wiley & Sons, 1968.

[75] R. Tibshirani, "Lecture notes in Convex Optimization," 2019.

[76] S. Boyd, "Lecture notes in EE364b - Convex Optimization II," 2022.

[77] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton University Press, 2015.

[78] F. H. Clarke, *Optimization and Nonsmooth Analysis*. SIAM, 1990.

[79] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer Science & Business Media, 2009.

[80] DVB, "Study mission reports: Commercial & technical reviews of wib," DVB, Tech. Rep., Jun. 2018, study mission reports.

[81] 3GPP, "Study on network-assisted interference cancellation and suppression (naic) for LTE (release 12)," 3rd Generation Partnership Project (3GPP), Technical Report TR 36.866, Mar. 2014.

[82] ——, "Study on downlink multiuser superposition transmission (MUST) for LTE (release 13)," 3rd Generation Partnership Project (3GPP), Technical Report TR 36.859, Dec. 2015.

[83] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[84] K. Son and W. Choi, "Single group multicast beamformer design using active constraints," in *Proceedings of 2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Oct. 2020, pp. 1364–1366.

[85] C.-L. Hsiao, J.-C. Guey, W.-H. Sheen, and R.-J. Chen, "A two-user approximation-based transmit beamforming for physical-layer multicasting in mobile cellular downlink systems," *Journal of the Chinese Institute of Engineers*, vol. 38, no. 6, pp. 742–750, Mar. 2015. [Online]. Available: https://doi.org/10.1080/02533839.2015.1016879

[86] W.-J. Choi, K.-W. Cheong, and J. Cioffi, "Iterative soft interference cancellation for multiple antenna systems," in *Proceedings of IEEE Wireless Communications and Networking Conference*, vol. 1, Sep. 2000, pp. 304–309 vol.1.

[87] P. Fertl, J. Jalden, and G. Matz, "Performance assessment of MIMO-BICM demodulators based on mutual information," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1366–1382, Mar. 2012.

[88] 3GPP, "NR demodulation performance evolution (Release 18)," 3rd Generation Partnership Project (3GPP), Technical Report TR 38.878, Sep. 2023.

[89] P. Yang and H. Yang, "Low-complexity weighted sum-rate maximization of MIMO MAC with finite-alphabet inputs," in *Proceedings of IEEE Global Communications Conference (GLOBECOM)*, Dec. 2020, pp. 1–5.

[90] J. Jin, Y. R. Zheng, W. Chen, and C. Xiao, "Generalized quadratic matrix programming: A unified framework for linear precoding with arbitrary input distributions," *IEEE Transactions on Signal Processing*, vol. 65, no. 18, pp. 4887–4901, Sep. 2017.

[91] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 5th ed. New York, NY, USA: Springer, 2021.

[92] M. Chen, S. C. Liew, Z. Shao, and C. Kai, "Markov approximation for combinatorial network optimization," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6301–6327, Oct. 2013.

[93] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, "Rethinking WMMSE: Can its complexity scale linearly with the number of BS antennas?" *IEEE Transactions on Signal Processing*, vol. 71, pp. 433–446, February 2023.

[94] Y. Mao, B. Clerckx, J. Zhang, V. O. K. Li, and M. A. Arafah, "Max-min fairness of k-user cooperative rate-splitting in MISO broadcast channel with user relaying," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6362–6376, Oct. 2020.

[95] 3GPP, "Technical Specification Group Services and System Aspects; Release 18 Description; Summary of Rel-18 Work Items (Release 18)," 3GPP, Tech. Rep. TR 21.918 V2.2.0, Mar. 2025.

[96] ——, "Technical Specification Group Radio Access Networks; NR demodulation performance evolution (Release 18)," 3GPP, Tech. Rep. TR 38.878 V18.4.0, Sep. 2024.

[97] M. Shafi, E. G. Larsson, X. Lin, D. Panaitopol, S. Parkvall, F. Ronteix-Jacquet, and A. Toskala, "Industrial viewpoints on ran technologies for 6G," *arXiv preprint arXiv:2508.08225*, 2025.

[98] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "A survey on MIMO transmission with finite input signals: Technical challenges, advances, and future trends," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1779–1833, Oct. 2018.

[99] 3GPP, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Multiplexing and channel coding (Release 17)," 3GPP, Tech. Rep. TS 38.212 V18.6.0, Mar. 2025.

[100] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[101] R. Cerna Loli, O. Dizdar, B. Clerckx, and P. Popovski, "Hybrid automatic repeat request for downlink rate-splitting multiple access," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 15 261–15 275, Oct. 2024.

[102] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser MISO systems: A rate-splitting approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016.

[103] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[104] D.-S. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, Sep. 2000.