

# Non-Orthogonal Multiple Access for MIMO Wireless Communications

*Sara Norouzi*



Department of Electrical & Computer Engineering  
McGill University  
Montréal, Canada

March 2023

---

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

©2023 Sara Norouzi

For

**Women, Life, Freedom**

---

## Abstract

Through spatial diversity, multiplexing or beamforming gain, the multiple-input multiple-output (MIMO) techniques can offer significant performance improvements in terms of user capacity, spectral efficiency, and peak data rates. Recently, the application of MIMO techniques along with non-orthogonal multiple access (NOMA) has aroused great interest as an enabling technology to meet the exacting demands of fifth generation (5G) and beyond 5G (B5G) wireless networks. In effect, by allowing multiple users to access overlapping time and frequency resources in the same spatial layer, NOMA has the potential to provide higher system throughput and solve the massive connectivity needed for future wireless networks. The primary objective of this thesis is to develop new approaches for multi-user MIMO NOMA systems from the perspectives of spectral and energy efficiency.

First, the joint design of user clustering, downlink beamforming and power allocation is formulated as a mixed-integer non-linear programming (MINLP) model for a MIMO NOMA system. In this problem the aim is to minimize the total transmission power while satisfying quality-of-service (QoS) and power constraints. To tackle this challenging problem, we reformulate it into a more tractable form and conceive two algorithms based on the branch-and-bound and penalty dual decomposition techniques for its solution. The performance of the proposed joint design algorithms for MIMO NOMA is validated by means of simulations over millimeter-wave (mmWave) channels. The results show the advantages of the proposed algorithms in terms of total transmit power and spectral efficiency over competing multiple access schemes.

Then, we study the application of spatial user clustering along with downlink beamforming for MIMO sparse code multiple access (SCMA) in a cloud radio access

---

network (C-RAN). A user clustering algorithm based on a constrained  $K$ -means method is proposed to limit the number of users in each cluster. Subsequently, two iterative algorithms for beamforming design are developed by minimizing the total transmission power under QoS and fronthaul capacity constraints. The performance of the proposed user clustering and downlink beamforming approaches in MIMO SCMA systems is evaluated through simulations. The results provide useful insights into the advantages of the proposed schemes over benchmark approaches, in terms of transmit power and spectral efficiency.

Finally, we propose a novel SCMA decoder based on deep residual neural network (ResNet), wherein the decoder is trained to predict the transmit codewords. In our approach, batch normalization is utilized to enhance the stability and robustness of the decoder, while residual blocks are employed to tackle the problems with deep learning based decoder such as accuracy saturation and vanishing gradients. The performance of the proposed ResNet decoder for SCMA is validated by means of simulations over AWGN and Rayleigh fading channels. The results show that besides a much reduced complexity, the proposed decoder leads to improvements in term of bit error rate (BER) over competing deep neural network (DNN) based decoders.

## Sommaire

Grâce à la diversité spatiale, au multiplexage ou au gain de formation de faisceaux, les techniques multiple-input multiple-output (MIMO) peuvent offrir des améliorations significatives des performances en termes de capacité utilisateur, d'efficacité spectrale et de débits de données de pointe. Récemment, l'application des techniques MIMO ainsi que l'accès multiple non orthogonal (NOMA) ont suscité un grand intérêt en tant que technologie habilitante pour répondre aux exigences rigoureuses des réseaux sans fil de cinquième génération (5G) et au-delà de la 5G (B5G). En effet, en permettant à plusieurs utilisateurs d'accéder à des ressources de temps et de fréquence qui se chevauchent dans la même couche spatiale, NOMA a le potentiel de fournir un débit système plus élevé et de résoudre la connectivité massive nécessaire pour les futurs réseaux sans fil. L'objectif principal de cette thèse est de développer de nouvelles approches pour les systèmes MIMO NOMA multi-utilisateurs du point de vue de l'efficacité spectrale et énergétique.

Tout d'abord, la conception conjointe du regroupement d'utilisateurs, de la formation de faisceaux sur la liaison descendante et de l'allocation de puissance est formulée sous la forme d'un modèle de programmation non linéaire à nombre entier mixte (MINLP) pour un système MIMO NOMA. Dans ce problème, le but est de minimiser la puissance de transmission totale tout en satisfaisant les contraintes de qualité de service (QoS) et de puissance. Pour résoudre ce problème difficile, nous le reformulons sous une forme plus simple et concevons deux algorithmes basés sur les techniques de décomposition par branches-limites et par pénalité. Les performances des algorithmes de conception conjointe proposés pour MIMO NOMA sont validées au moyen de simulations sur des canaux à ondes millimétriques (mmWave). Les résultats montrent les avantages des algorithmes proposés en termes de puissance d'émission totale et d'efficacité spectrale par rapport aux schémas d'accès multiples concurrents.

Ensuite, nous étudions l'application du regroupement spatial d'utilisateurs avec la formation de faisceaux en liaison descendante pour l'accès multiple à code creux MIMO (SCMA) dans un réseau d'accès radio cloud (C-RAN). Un algorithme de regroupement d'utilisateurs basé sur une méthode  $K$ -means contrainte est proposé pour limiter le nombre d'utilisateurs dans chaque groupe. Par la suite, deux algorithmes itératifs pour la conception de formation de faisceaux sont développés en minimisant la puissance de transmission totale sous les contraintes de QoS et de capacité de transmission. Les performances des approches proposées de regroupement d'utilisateurs et de formation de faisceaux de liaison descendante dans les systèmes SCMA MIMO sont évaluées au moyen de simulations. Les résultats fournissent des informations utiles sur les avantages des schémas proposés en termes de puissance d'émission et d'efficacité spectrale par rapport aux approches de référence.

Enfin, nous proposons un nouveau décodeur SCMA basé sur un réseau neuronal résiduel profond (ResNet), dans lequel le décodeur est entraîné pour prédire les mots de code transmis. Dans notre approche, la normalisation par lots est utilisée pour améliorer la stabilité et la robustesse du décodeur, tandis que les blocs résiduels sont utilisés pour résoudre les problèmes liés au décodeur basé sur l'apprentissage en profondeur, tels que la saturation de la précision et les gradients de fuite. Les performances du décodeur ResNet proposé pour SCMA sont validées au moyen de simulations sur les canaux d'évanouissement AWGN et Rayleigh. Les résultats montrent qu'en plus d'une complexité très réduite, le décodeur proposé conduit à des améliorations en termes de taux d'erreur sur les bits (BER) par rapport aux décodeurs concurrents basés sur le réseau neuronal profond (DNN).

## Acknowledgements

Working towards a Ph.D. degree is a long journey which takes commitment, dedication, and perseverance. However, one might never make it to the destination without the help and advice from others along the way. For this, I feel fortunate and privileged to have Professor Benoit Champagne as my supervisor. None of what I have accomplished would have been possible without his excellent guidance and encouragement throughout my doctoral study. In addition, I would like to express my gratitude towards my Ph.D. committee members, Professor Ioannis Psaromiligkos and Professor Tim Hoheisel for their time and efforts to evaluate my work and provide valuable feedback. I would also like to thank all the jury members who have spent time reviewing and critiquing this dissertation, the presentation of which has significantly benefited from their insightful comments.

I am deeply grateful for the generous funding support from the McGill Engineering Doctoral Awards (MEDA), the Natural Sciences and Engineering Research Council (NSERC) of Canada and InterDigital Canada through the NSERC CRD grant program. I would like to thank the technical staff of InterDigital, especially Dr. Afshin Haghighat, for the constructive suggestions and inputs at our progress meetings. I will always cherish the fond memories of being surrounded by a group of talented and affable colleagues: Ali, Alireza, Seyed Saleh, Toluwaleke, Hanwook, Lu, Ryan, Farnood.

I am forever grateful to my family for their unconditional love, understanding, and support. A debt of gratitude is owed to Azadeh, Mahbod, Siavash, and Arash, whose love and support helped me overcome various difficulties despite the distance from my hometown. Last but not least, my heartfelt love is devoted to my husband, Abolfazl, for his patience, continuous support and encouragement over the course of my Ph.D. studies. I hope that this accomplishment makes all of you proud.

# Contents

<b>Contents</b>	<b>vii</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xiii</b>
<b>Acronyms</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Future Wireless Networks . . . . .	1
1.2 NOMA: Concept, Challenges, and Opportunities . . . . .	5
1.3 Thesis Objective and Contributions . . . . .	7
1.4 Organization . . . . .	11
<b>2 Literature Review</b>	<b>13</b>
2.1 PD-NOMA . . . . .	13
2.2 CD-NOMA . . . . .	18
2.2.1 SCMA Encoder Design . . . . .	20
2.2.2 SCMA Decoder Design . . . . .	22

---

2.3	Comparison of NOMA Schemes . . . . .	27
2.4	C-RAN Architecture . . . . .	28
2.5	Concluding Remarks . . . . .	30
<b>3</b>	<b>Joint User Clustering, Beamforming, and Power Allocation for NOMA</b>	<b>32</b>
3.1	Introduction . . . . .	32
3.2	System Model . . . . .	35
3.2.1	Channel Model . . . . .	35
3.2.2	Signal Model . . . . .	37
3.2.3	SIC Procedure . . . . .	37
3.3	Problem Formulation . . . . .	39
3.4	Proposed BB-Based Algorithm . . . . .	42
3.4.1	Convex Relaxation . . . . .	44
3.4.2	Proposed Algorithm . . . . .	47
3.4.3	Convergence and Complexity Analysis . . . . .	50
3.5	Proposed PDD-Based Algorithm . . . . .	52
3.5.1	Problem Reformulation . . . . .	52
3.5.2	Proposed Algorithm . . . . .	54
3.5.3	Convergence and Complexity Analysis . . . . .	57
3.6	Simulation Results . . . . .	59
3.6.1	Methodology . . . . .	59
3.6.2	Results and Discussion . . . . .	60
3.7	Concluding Remarks . . . . .	69
<b>4</b>	<b>User Clustering and Beamforming for MIMO SCMA in C-RAN</b>	<b>70</b>

---

4.1	Introduction . . . . .	70
4.2	System Model and Problem Description . . . . .	73
4.2.1	SCMA Encoder . . . . .	74
4.2.2	Channel Model . . . . .	75
4.2.3	Signal Model . . . . .	77
4.2.4	Problem Description . . . . .	79
4.3	User Clustering . . . . .	80
4.3.1	Constrained $K$ -means Clustering . . . . .	80
4.3.2	Number of Clusters . . . . .	84
4.3.3	Complexity Analysis . . . . .	85
4.4	Downlink Beamforming . . . . .	86
4.4.1	Beamforming Problem . . . . .	87
4.4.2	MI-SOCP Beamforming Approach . . . . .	89
4.4.3	Two-Stage Beamforming Approach . . . . .	93
4.4.4	Convergence and Complexity Analysis . . . . .	98
4.5	Simulation Results . . . . .	99
4.5.1	Methodology . . . . .	99
4.5.2	Results and Discussion . . . . .	103
4.6	Concluding Remarks . . . . .	109
<b>5</b>	<b>Deep Residual Neural Network Decoder for SCMA</b>	<b>112</b>
5.1	Introduction . . . . .	112
5.2	System Model . . . . .	114
5.2.1	SCMA Encoder . . . . .	114
5.2.2	Deep Neural Networks (DNNs) . . . . .	115

---

5.3	Proposed ResNet Decoder . . . . .	117
5.3.1	Batch Normalization . . . . .	117
5.3.2	Residual Blocks . . . . .	119
5.3.3	Structure of Proposed SCMA Decoder . . . . .	120
5.4	Simulation Results . . . . .	122
5.5	Concluding Remarks . . . . .	127
<b>6</b>	<b>Conclusion</b>	<b>128</b>
6.1	Summary . . . . .	128
6.2	Potential Future Works . . . . .	130
	<b>Appendices</b>	<b>133</b>
<b>A</b>		<b>133</b>
A.1	Proof of Theorem 3.1 . . . . .	133
A.2	Proof of Theorem 3.2 . . . . .	134
<b>B</b>		<b>136</b>
B.1	Proof of Proposition 4.2 . . . . .	136
	<b>Bibliography</b>	<b>140</b>

---

# List of Figures

1.1	Illustration of basic PD-NOMA with a SIC receiver (adapted from [17]). . . . .	6
2.1	The multi-user MIMO NOMA system model. . . . .	14
2.2	SCMA encoder model ( $C = 2, N = 4$ ). . . . .	20
2.3	The SCMA factor graph. . . . .	21
2.4	The MPA process model. . . . .	25
2.5	The C-RAN architecture. . . . .	29
3.1	The multi-user MIMO NOMA system model. . . . .	36
3.2	An illustration of the BB process (adapted from [74]). . . . .	43
3.3	An illustration for Proposition 3.1. . . . .	46
3.4	Convergence behavior of the BB-based algorithm. . . . .	61
3.5	Convergence behavior of the PDD-based algorithm. . . . .	63
3.6	Achievable sum rate versus total transmit power. . . . .	64
3.7	Total transmit power versus target SINR. . . . .	65
3.8	Total transmit power versus target SINR. . . . .	66
3.9	Total transmit power versus number of users. . . . .	67
3.10	Sum rate versus total transmit power for perfect and imperfect CSI. . . . .	68

---

4.1	The MIMO SCMA system model under C-RAN. . . . .	73
4.2	The impact of the number of clusters. . . . .	101
4.3	The convergence of the proposed algorithms. . . . .	102
4.4	Transmit power versus sum rate for different clustering and beamforming approaches. . . . .	105
4.5	Transmit power versus sum rate for different transmission schemes. . . . .	106
4.6	Transmit power versus sum rate for perfect and imperfect CSI. . . . .	108
4.7	Average number of associated RRHs versus fronthaul link capacity for $M = 20$ . . . . .	109
4.8	Transmit power versus number of antennas. . . . .	110
4.9	Transmit power versus total number of users. . . . .	110
5.1	DNN structure. . . . .	115
5.2	(a) Batch normalization. (b) Residual block. . . . .	119
5.3	Internal structure of ResNet decoder. . . . .	121
5.4	BER performance of SCMA decoders under AWGN. . . . .	124
5.5	BER performance of SCMA decoders under Rayleigh fading channel. . . . .	125
5.6	Computation time comparison among SCMA decoders. . . . .	126
B.1	The MCF equivalent directed graph structure. . . . .	138

# List of Tables

2.1	Comparison between different NOMA schemes. . . . .	28
3.1	Summary of parameters. . . . .	60
4.1	Simulation setting parameters . . . . .	100
4.2	Run-time of different beamforming approaches . . . . .	104
5.1	Complexity comparison. . . . .	125

# Acronyms

**3GPP** Third-Generation Partnership Project.

**5G** Fifth Generation.

**AL** Augmented Lagrangian.

**AO** Alternating Algorithm.

**BB** Branch-and-Bound.

**BD** Block Diagonalization.

**BER** Bit Error Rate.

**BS** Base Station.

**CCCP** Concave-Convex Procedure.

**C-RAN** Cloud Radio Access Network.

**CDMA** Code Division Multiple Access.

**CSI** Channel State Information.

**DC** Differences of Convex Functions.

<b>DL</b>	Deep Learning.
<b>DNN</b>	Deep Neural Network.
<b>IoT</b>	Internet of Things.
<b>LTE</b>	Long-Term Evolution.
<b>MI-SOCP</b>	Mixed-Integer Second Order Cone Program.
<b>MIMO</b>	Multiple-Input Multiple-Output.
<b>MINLP</b>	Mixed-Integer Non-Linear Program.
<b>mmWave</b>	Millimeter Wave.
<b>MPA</b>	Message Passing Algorithm.
<b>MSE</b>	Mean Squared Error.
<b>NOMA</b>	Non-Orthogonal Multiple Access.
<b>OFDMA</b>	Orthogonal Frequency Division Multiple Access.
<b>OMA</b>	Orthogonal Multiple Access.
<b>PDD</b>	Penalty Dual-Decomposition.
<b>PDMA</b>	Pattern Division Multiple Access.
<b>QoS</b>	Quality of Service.

<b>RAT</b>	Radio Access Technology.
<b>ResNet</b>	Residual Neural Network.
<b>RRH</b>	Remote Radio Heads.
<b>SCMA</b>	Sparse Code Multiple Access.
<b>SIC</b>	Successive Interference Cancellation.
<b>SINR</b>	Signal-to-Interference-plus-Noise Ratio.
<b>SSE</b>	Sum of Squared Euclidean.
<b>TDMA</b>	Time Division Multiple Access.

# Chapter 1

## Introduction

In this Chapter, we first discuss potential candidate technologies for emerging and future wireless networks. We then present the concept and challenges of different non-orthogonal multiple access schemes. Finally, we state the objectives of the research and summarize the main contributions of the thesis.

### 1.1 Future Wireless Networks

In modern wireless networks, the design of radio access technology (RAT) is of critical importance for improving system performance in a cost-effective manner. RATs are typically characterized by multiple access schemes which play key roles in optimizing system performance through efficient allocation of resources. The conventional orthogonal multiple access (OMA) schemes, such as time division multiple access (TDMA), code division multiple access (CDMA) and orthogonal frequency division multiple access (OFDMA), can completely remove inter-user interference under ideal conditions, which help improve system performance. However, this comes at the cost of limiting the number

of supported users, and consequently, overall network capacity [1].

During the last few decades, we have witnessed more than a 1,000-fold capacity increase in wireless communications systems, with the main catalysts being the improvement in spectrum efficiency and the acquisition of new spectrum. Such a capacity increase has fostered the rapid growth of the mobile Internet accompanied by various new applications and services. Looking into the next decades of wireless communications evolution, it is expected that the continuing growth of mobile Internet applications and services will trigger a huge growth, in the range of 10-100 times for mobile traffic [2]. Considering that the capacity of currently deployed systems such as long-term evolution (LTE) Release 8/9 is already quite high, further capacity growth seems to be a very challenging task at first glance. As a matter of fact, in order to handle such data traffic, improving the spectral efficiency remains one of the key challenges. Moreover, the rapid development of the Internet of Things (IoT) calls for the massive connectivity of users and/or devices, and future generations of wireless networks will hence need to meet the demand for low-latency, low-cost devices, and diverse service types. Hence, further enhancement in technologies is a must to satisfy these requirements for future wireless networks [3].

Orthogonal multiple access is an appropriate choice for good system-level throughput performance with a simple receiver design. However, more advanced receiver designs are required in order to mitigate intra-cell and inter-cell interference, and boost the spectrum efficiency in the future [4]. So far, some of potential candidates that have been proposed to address the challenges of the fifth-generation (5G) and beyond are millimeter wave communications (mmWave), massive multiple-input multiple-output (MIMO) systems, ultra-dense networks, non-orthogonal multiple access (NOMA), and cloud radio access networks (C-RAN).

Traditionally, the microwave (sub-6 GHz) band has been the most widely-used spectrum for the consumer wireless systems owing to its favorable propagation behavior. Over time, the frequency spectrum under 6 GHz became densely occupied by TV and radio signals and hence, the mmWave band (30 GHz to 300 GHz) is now attracting considerable attention for prospective wireless networks [5, 6]. On the one hand, operating in mmWave band brings its own challenges since the electromagnetic waves become highly sensitive to path loss and blockage [7]. In effect, the use of shorter wavelength in mmWave band (and hence larger transmission bandwidth) can help transfer data at a faster rate, although the distance over which the data transfer can be accomplished is significantly reduced compared to the microwave band. As a result, mmWave communications would be restricted to short-range, line-of-sight applications unless further enhancements are made at the physical-layer.

On the other hand, the mmWave frequency band offers several benefits such as the possibility of integrating a large-scale antenna array into a much smaller area compared to its microwave counterparts. The form factor of an antenna array is proportional to the wavelength. For instance, at a carrier frequency of 60 GHz, the form factor of an  $8 \times 8$  planar antenna array is only  $4 \text{ cm}^2$ . Consequently, the application of massive MIMO beamforming along with mmWave can efficiently address the issues of severe path loss and blockage cellular networks [8]. Moreover, the use of massive MIMO techniques can lead to significant performance improvements in terms of user capacity, spectral efficiency, and peak data rates, by taking advantage of spatial diversity, multiplexing or beamforming gains.

In order to make network capacity improvements in terms of spectral efficiency within a given geographical area, we can admit as many users as possible over an allocated set of time-frequency resources. we can approach this issue by densifying the deployment of base stations (BSs) with universal frequency reuse, which reduces the competition for resources

within a cell among a large number of users [9]. Due to reduced cell radius obtained in ultra-dense networks, lower transmit power is required which results in increased energy efficiency. Moreover, thanks to load sharing, the backhaul links can accommodate more data traffic from each user in service. However, by increasing the density of BSs and user devices, the problem of inter-cell interference may become severe in multi-cell networks.

In recent years, so-called NOMA schemes have shown great promises in meeting the exacting demands of emerging and future generations of wireless networks. In effect, NOMA has the potential to provide higher spectral efficiency and enable massive connectivity by allowing multiple users and devices to access overlapping time and frequency resource elements in the same spatial layer [10]. As such, NOMA is particularly well-suited as a prospective multiple access technology for dense heterogeneous networks, as envisaged for machine-to-machine (M2M) communications and IoT applications [11]. Nevertheless, NOMA requires effective interference cancellation and management approaches. Moreover, low-complexity resource allocation algorithms is another challenge that needs to be addressed in NOMA systems.

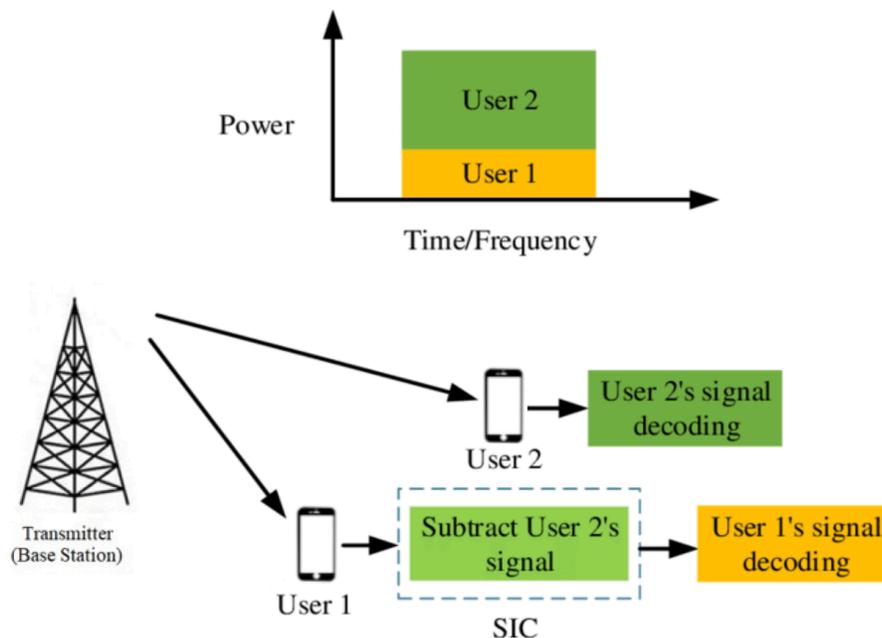
As an emerging network architecture for 5G and beyond, C-RAN offers several benefits, e.g., improved energy efficiency, better ability to handle interference on a larger scale and increased network capacity. In the C-RAN architecture, the transceivers are connected to a central processor via fronthaul links [12]. The central processor jointly encodes the user messages using linear precoding or beamforming techniques for interference mitigation purposes in the downlink. This separation of the central processor and transceivers functionalities reduces the power consumption, operating expenses and complexity of the transceivers. However, the implementation of C-RAN architecture raises several challenges such as high fronthaul capacities needed, central processor cooperation, resource allocation

mechanism, cell clustering, etc.

## 1.2 NOMA: Concept, Challenges, and Opportunities

NOMA has the potential to be integrated into the existing and future wireless systems because of its compatibility with other communication technologies. For example, NOMA has been shown to be compatible with conventional OMA, such as TDMA and OFDMA. Because of this, NOMA has also been proposed for inclusion in the 3rd generation partnership project (3GPP) long-term evolution advanced (LTE-A) standard, where NOMA is referred to as multiuser superposition transmission (MUST) [1]. Particularly, without requiring any changes to the LTE resource blocks (i.e., OFDMA subcarriers), the use of the NOMA principle ensures that multiple users can be simultaneously served on the same OFDMA subcarriers.

Basically, NOMA techniques can be classified into three main categories, namely: code domain, power domain, and multiple domain [13]. At the transmitter side, power domain NOMA (PD-NOMA) allocates different power levels to the users according to their channel condition. At the receiver side, successive interference cancellation (SIC) is employed, whereby the stronger signals are decoded first and subtracted from the received signal sequentially. Code domain NOMA (CD-NOMA) techniques, such as sparse code multiple access (SCMA) and multi-user shared access (MUSA), assign and apply different codes to the data streams of different users. In this way, the transformed data is spread over multiple resource elements in a sparse manner that allows the control of interference [14]. In multiple domain NOMA, such as pattern division multiple access (PDMA) [15] and lattice partition multiple access (LPMA) [16], the signals from multiple users are



**Figure 1.1:** Illustration of basic PD-NOMA with a SIC receiver (adapted from [17]).

superimposed in multiple domains, such as power, code and spatial domains.

PD-NOMA attempts to fully exploit multiplexing by overlapping multiple user signals in the time, code and frequency domains. According to the PD-NOMA principle, the users with lower channel gain are considered as weak users, and should therefore be allocated a larger fraction of the total available power at the transmitter, as illustrated in Fig. 1.1. PD-NOMA is capable of enhancing signal quality, network capacity, and cell-edge throughput. However, its implementation raises several research challenges. Firstly, the efficient allocation of available resources to multiple users with different data rates and quality of service (QoS) requirements poses a formidable optimization problem. Secondly, individual decoding of the superimposed user signals in the PD-NOMA receiver calls for advanced interference cancellation algorithms whose implementation significantly increases the complexity of the equipment and processing delay.

In CD-NOMA, the data streams of the users are modulated by using user-specific spreading sequences, having sparse, low-density and low inter-correlation properties. One of the advantages of these techniques is the grant-free access provided by the non-orthogonality feature, although this comes at the cost of introducing users interference. This kind of technique is inspired from the well-known CDMA technology extensively studied in the 90s. We note that for certain CD-NOMA schemes, such as MUSA, the receiver does not know the spreading sequences in advance. Hence, the decoder has to detect and estimate the transmitted data streams in a blind manner.

SCMA is a popularized CD-NOMA technique. While CDMA extends each information symbol (taken, e.g., from a quadrature amplitude modulation (QAM) constellation) into a finite sequence of complex symbols by using orthogonal or near orthogonal spreading codes, SCMA directly maps each group of bits into a sequence of complex symbols by merging together the symbol mapper and the CDMA spreader [18]. The overall process can be interpreted as a coding procedure from the binary domain to a multidimensional complex domain, which in turn raises new problems in terms of codebook and decoder designs [14].

### 1.3 Thesis Objective and Contributions

The primary objective of this thesis is to develop and investigate improved transceiver design approaches for the application of NOMA within the emerging framework of multi-user MIMO communications. To achieve this goal, we address the aforementioned research challenges by proposing new designs that can provide both energy and spectral efficiency under practical constraints. In this regard, the main research contributions and findings of this thesis are summarized as follows.

- We first investigate the application of downlink beamforming along with PD-NOMA in a multi-user MIMO system operating at mmWave frequencies. The joint design of user clustering, downlink beamforming and power allocation scheme is formulated as a novel mixed-integer non-linear program (MINLP), where the aim is to minimize the total transmission power while satisfying QoS, user clustering and power constraints. Owing to the non-convexity and combinatorial nature of the problem, obtaining an optimal solution is challenging. To tackle this issue, we first develop an algorithm based on branch-and-bound (BB), whereby the feasible space is successively partitioned and searched by means of lower and upper bounds on the objective function. While this algorithm is shown to return an  $\epsilon$ -optimal solution within a finite number of iterations, it entails high computational complexity. Considering this limitation, we then reformulate the original problem into a more tractable form and conceive a low-complexity algorithm for its solution based on the penalty dual-decomposition technique. The proposed joint design algorithms for MIMO NOMA are evaluated by means of simulations over mmWave channels. Results show significant improvements in terms of total transmit power and spectral efficiency compared to benchmark approaches.
- We then address the key problems of user clustering and downlink beamforming for MIMO SCMA in a C-RAN. Using channel state information available at the central processor, an efficient user clustering algorithm based on the constrained K-means method is proposed. Subsequently, two iterative algorithms for beamforming design are developed by minimizing the total transmission power under QoS and fronthaul capacity constraints. In the first approach, we approximate the continuous non-convex constraints by convex conic ones using first-order Taylor expansion and

iteratively solve a sequence of mixed-integer second order cone programs (MI-SOCPs) to achieve a high-quality solution, but with higher complexity. In the second approach, a two-stage low-complexity solution is developed in which beamforming matrices obtained from each stage are combined to form a single beamformer for each user. In the first stage, cluster beamformers are designed by taking advantage of block diagonalization, while in the second stage, user-specific beamformers are determined by minimizing transmission power. The performance of the proposed user clustering and downlink beamforming approaches for MIMO SCMA in C-RAN is validated through simulations over mmWave channels. Compared to benchmark approaches, the results show significant improvements in terms of transmit power and spectral efficiency.

- Finally, we focus on the design of SCMA decoders and propose a novel solution to this problem based on a deep residual neural network (ResNet). To tackle the problem of accuracy saturation and vanishing gradients, we employ residual blocks, while batch normalization is utilized to enhance the stability and performance of the decoder. Under the assumption that the channel state information (CSI) is available at the receiver side, the decoder is trained to predict the transmitted codewords by the users. The received signal and CSI are fed into the ResNet decoder as input, while the output consists of multiple branches, one for each user, wherein the transmit codewords are predicted. Through simulations, it is demonstrated that the proposed SCMA scheme with ResNet decoder can notably reduce bit error rate (BER) compared to DNN-based benchmark approaches, yet with much lower complexity.

This thesis has resulted into the following published contributions:

- **Journal papers:**

- [J1] S. Norouzi, Y. Cai, and B. Champagne. “Energy-efficient user clustering and downlink beamforming for MIMO-SCMA in C-RAN,” *IEEE Access*, vol. 9, pp. 115175-115191, Aug. 2021.
- [J2] S. Norouzi, B. Champagne, and Y. Cai. “Joint Optimization Framework for User Clustering, Downlink Beamforming, and Power Allocation in MIMO NOMA Systems,” *IEEE Transactions on Communications*, Nov. 2022. (accepted for publication)

- **Conference papers:**

- [C1] S. Norouzi, A. Morsali, and B. Champagne. “Optimizing Transmission Rate in NOMA via Block Diagonalization Beamforming and Power Allocation,” *IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing (PACRIM)*, pp. 1-5, Victoria, Canada, Aug. 2019.
- [C2] S. Norouzi, Y. Cai, and B. Champagne. “Constrained  $K$ -means user clustering and downlink beamforming in MIMO-SCMA systems,” *IEEE 32nd Annual International Symposium on PIMRC*, pp. 1091-1096, Oct. 2021.
- [C3] S. Norouzi, Y. Cai, and B. Champagne. “Joint design of user clustering, beamforming, and power allocation for NOMA,” *IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct. 2022.
- [C4] S. Norouzi and B. Champagne. “Deep residual neural network decoder for sparse code multiple access,” *IEEE Wireless Communications and Networking Conf. (WCNC)*, Scotland, UK, Mar. 2023.

In all of these publications, Ms. Sara Norouzi, as the first author, developed and/or proposed the idea, formulated the problem, developed the algorithms, implemented simulations, and prepared the draft of the manuscripts. Mr. Alireza Morsali, Ph.D. student, assisted in reviewing [C1]. Prof. Yunlong Cai collaborated in [J1-J2], [C2-C3], to review the works and to assist in editing and writing. Prof. Benoit Champagne supervised and reviewed the works, and assisted in the editing and writing of the manuscripts at different stages.

## 1.4 Organization

The rest of the thesis is organized as follows. Chapter 2 provides a comprehensive literature survey of prior contributions on the application of NOMA in MIMO systems from various perspectives. Chapter 3 investigates the joint design of user clustering, downlink beamforming and power allocation scheme in a MIMO NOMA system operating at mmWave frequencies. Chapter 4 focuses on the problems of user clustering and downlink beamforming for MIMO SCMA in a C-RAN. Chapter 5 considers the design of SCMA decoders by employing deep learning based methods. Chapter 6 gives a summary and provides suggestions for future investigations. Certain mathematical proofs and derivations are relegated to the Appendices.

*Notations:* The following notations are used throughout the thesis, unless otherwise noted. Scalars, vectors and matrices are respectively denoted by lower case, boldface lower case and boldface upper case letters. For a matrix  $\mathbf{A}$ ,  $a_{i,j}$  and  $[\mathbf{A}]_{i,j}$  denote its  $(i, j)$ th entry, while  $\mathbf{A}^T$  and  $\mathbf{A}^H$  denote its transpose and conjugate transpose, respectively. The operators  $\|\cdot\|_2$ ,  $\|\cdot\|_0$  and  $\|\cdot\|_\infty$  denote the Euclidean, zero and infinity norms of a vector, respectively.

---

For a set  $\mathcal{A}$ ,  $|\mathcal{A}|$  denotes its cardinality.  $\Re(z)$  and  $\Im(z)$  denotes the real and imaginary parts of complex number  $z$ , respectively.  $\mathbb{C}^{m \times n}$  ( $\mathbb{R}^{m \times n}$ ) denotes the space of  $m \times n$  complex (real) matrices.  $\mathbb{B}^{m \times n}$  denotes binary matrices of size  $m \times n$  where the set  $\mathbb{B} = \{0, 1\}$ . Given matrices  $\mathbf{A}$  and  $\mathbf{B} \in \mathbb{R}^{m \times n}$ , we define  $[\mathbf{A}, \mathbf{B}] = \{\mathbf{X} \in \mathbb{R}^{m \times n} | a_{i,j} \leq x_{i,j} \leq b_{i,j}, \forall i, j\}$  and refer to this set as a box. We use  $\mathcal{CN}(\mu, \sigma^2)$  to denote a complex circular Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$ .

# Chapter 2

## Literature Review

Before we embark on the study of NOMA for MIMO wireless systems, we give a brief review of the related background and recent developments in this chapter, which serve as the basis for our proposed research in the subsequent chapters. First, we focus on MIMO PD-NOMA where the application of multiple antenna techniques and user clustering provide further performance improvements. We then review related works on SCMA, specifically the design of the codebook and decoder. Afterwards, a comparison between major NOMA techniques is presented. Finally, MIMO beamforming in CRAN architecture is reviewed followed by some concluding remarks.

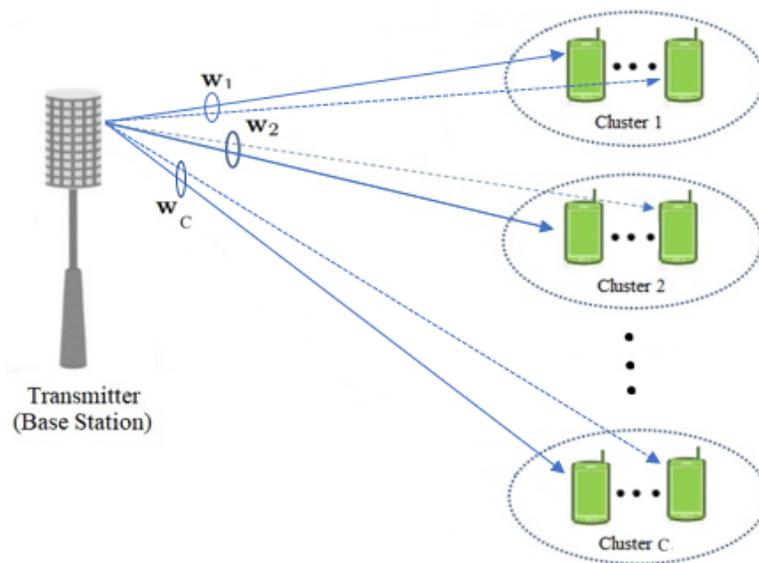
### 2.1 PD-NOMA

The basic principles of PD-NOMA rely on the employment of superposition coding (SC) at the transmitter side and SIC techniques at the receiver side. Although the application of PD-NOMA in cellular networks is relatively new, related concepts have been studied in communications and information theory for a long time. For instance, the key components

of PD-NOMA i.e., SC and SIC, have already been invented more than two decades ago [13,19,20]. Nevertheless, the principle of removing orthogonality for multiple access has not been used in the previous generations of wireless cellular networks. Since the main focus of this section is on PD-NOMA, in the sequel, we refer to the latter as NOMA for simplicity.

From the NOMA perspective, power-domain multiplexing means that different users are allocated different power levels according to their channel conditions to obtain the maximum gain in system performance. The weaker users with lower channel gains will be apportioned a larger fraction of the total available transmit power. Such power allocation can be exploited to separate different users, using SIC at the receiver to cancel the interference. Specifically, SIC will be employed at the receiver whereby the strongest signal is decoded first and subtracted from the received signal, after which the second strongest signal is extracted from the residual, and so on in a sequential manner.

As illustrated in Fig. 2.1, let us consider the downlink of a multi-user MIMO NOMA



**Figure 2.1:** The multi-user MIMO NOMA system model.

system with one BS. The transmitter is equipped with  $N$  antennas, serving  $K$  users indexed by  $k \in \mathcal{K} = \{1, \dots, K\}$ , where each user is equipped with a single-antenna receiver<sup>1</sup>. The users are partitioned into  $C$  non-overlapping clusters, indexed by  $c \in \mathcal{C} = \{1, \dots, C\}$ . The users in a given cluster jointly share a common beamforming vector denoted as  $\mathbf{w}_c \in \mathbb{C}^{N \times 1}$ . Hence, the BS utilizes the NOMA superposition of signals and beamforming simultaneously.

Let  $\mathcal{U}_c$  denote the index set of users in the  $c$ th cluster<sup>2</sup> and  $\mathbf{h}_i \in \mathbb{C}^{N \times 1}$  denote the channel vector from BS to the  $i$ th user in the  $c$ th cluster, i.e.,  $i \in \mathcal{U}_c$ , and  $c \in \mathcal{C}$ . Without loss of generality, we assume that the  $K$  users in the system are sorted in ascending order according to their channel gains, i.e.,  $\|\mathbf{h}_i\| > \|\mathbf{h}_j\|$  for  $i > j$ . The received signal at the  $i$ th user in the  $c$ th cluster is given by,

$$r_i = \mathbf{h}_i^H \mathbf{z}_c + \sum_{c'=1, c' \neq c}^C \mathbf{h}_i^H \mathbf{z}_{c'} + n_i \quad (2.1)$$

where  $\mathbf{z}_c \in \mathbb{C}^{N \times 1}$  is the transmitted signal by the BS towards the  $c$ th cluster, and  $n_i \sim \mathcal{CN}(0, \sigma_i^2)$  is an additive white Gaussian noise term.

It is assumed that the BS applies power domain NOMA within each cluster. That is, it transmits the superposition of the individual data symbols with different power levels to all users in a cluster simultaneously with the same radio resources, such as time slot and frequency channel. Hence, we have,

$$\mathbf{z}_c = \mathbf{w}_c \left( \sum_{i \in \mathcal{U}_c} \sqrt{\alpha_i} s_i \right) \quad (2.2)$$

---

<sup>1</sup>The system model formulation presented in this section can be extended to the case where users are equipped with multiple antennas. We refer the interested reader to [21, 22] for additional details.

<sup>2</sup>For instance, five users could be distributed into 2 non-overlapping clusters as  $U_1 = \{1, 3\}$  and  $U_2 = \{2, 4, 5\}$ .

where  $s_i \in \mathbb{C}$  is the data symbol intended for the  $i$ th user and  $\alpha_i$  denotes the fraction of the total power available for the  $c$ th cluster that is allocated to the  $i$ th user, i.e.,  $\alpha_i \in [0, 1]$  and  $\sum_{i \in \mathcal{U}_c} \alpha_i = 1$ . We assume that the data symbols of the different users are independent with zero mean and unit variance, i.e.  $E[|s_i|^2] = 1$ . Upon substitution of (2.2) into (2.1), we can express the received signal at the  $i$ th user in the  $c$ th cluster as a sum of the desired signal, the interference from the other user in that cluster (intra-cluster interference), the inter-cluster interference and the noise term, i.e.,

$$r_i = \mathbf{h}_i^H \mathbf{w}_c \left( \sum_{i' \in \mathcal{U}_c} \sqrt{\alpha_{i'}} s_{i'} \right) + \underbrace{\sum_{c' \neq c} \mathbf{h}_i^H \mathbf{w}_{c'} \left( \sum_{k \in \mathcal{U}_{c'}} \sqrt{\alpha_k} s_k \right)}_{\text{Inter-cluster Interference}} + n_i \quad (2.3)$$

According to the NOMA principle, each user in a given cluster employs SIC to mitigate the intra-cluster interference, by decoding and removing the message of the weaker users in that cluster. Assume that both  $i$ th and  $j$ th users belong to the  $c$ th cluster and  $i > j$ . For the SIC operation at the  $i$ th user, the signal-to-interference-plus-noise ratio (SINR) of the  $j$ th user signal after perfect interference cancellation of weaker user signals (i.e., those with index  $i' < j$ ) is given by,

$$\text{SINR}_{j,i}^c = \frac{|\mathbf{h}_i^H \mathbf{w}_c|^2 \alpha_j}{\sum_{i' > j, i' \in \mathcal{U}_c} |\mathbf{h}_i^H \mathbf{w}_c|^2 \alpha_{i'} + \sum_{c' \neq c} |\mathbf{h}_i^H \mathbf{w}_{c'}|^2 + \sigma_i^2} \quad (2.4)$$

where the first term in the denominator of (2.4) represents the residual intra-cluster interference and the second term represents the inter-cluster interference.

The application of MIMO NOMA along with mmWave is of great interest for new generations of wireless networks (i.e. 5G, Beyond 5G and ensuing 6G); accordingly, it has been the focus of significant research efforts in recent years. In [23], a random beamforming

approach is applied to a MIMO NOMA system to reduce overhead in a dense network with a large number of users. In [24], a transmission scheme relying on user clustering, random beamforming and power allocation is proposed for a MIMO NOMA system. In [25], a multi-beam NOMA framework is developed for mmWave systems, such that a limited number of radio frequency (RF) chains may be used to accommodate multiple users with various angles of departures (AoDs). The design of beamforming and power allocation for a MIMO NOMA system is addressed in [26], where user clustering is handled through the  $K$ -means algorithm. The authors in [27] evaluate the effect of beam misalignment on rate performance in a MIMO NOMA system with hybrid beamforming and propose a design scheme for the digital and analog precoders and power allocation based on sum-rate maximization. The performance of a MIMO NOMA system operating at mmWave frequencies with hybrid beamforming is investigated in [28], where the clustering of users is based on the correlation among their channel vectors.

Some recent works on NOMA investigate the *joint* design of beamforming and power allocation. In [29], such a joint design is proposed for simultaneous wireless information and power transfer (SWIPT) in a two-user NOMA system, such that the data rate of the strong user is maximized under a quality of service (QoS) constraint for the weak user. The joint design of power allocation and beamforming using decomposed optimization (DO) is proposed in [30] for a two-user MIMO NOMA system at mmWave frequencies. In [31], the joint design of robust beamforming and power splitting ratio is addressed to maximize the data rate of the cell-center user which adopts a SWIPT technique. The authors in [32] investigate a joint design for artificial noise aided beamforming and power allocation, such that the transmission security and reliability are increased in the presence of untrusted near users and external eavesdroppers.

It is worth-mentioning that [29–32] focused on the joint design of beamforming and power allocation in two-user scenarios. However, this joint design problem for multi-user NOMA (with more than 2 users) is quite challenging and has been the focus of several recent studies, as summarized below. In [33], a joint design is addressed by maximizing the achievable rate at the destination in a multi-user NOMA-based amplify-and-forward (AF) relay network where all users are grouped into one cluster. In [34], joint design of beamforming and power allocation is proposed for a multi-cell multi-user MIMO NOMA network in which the users are divided into two groups according to their QoS requirements, rather than their channel quality. The performance of a NOMA-based satellite-terrestrial integrated network with a joint design of beamforming and power allocation is investigated in [35] where a novel user clustering based on the channel gain and correlation is proposed. In [36], an alternating optimization (AO) algorithm is proposed for the joint design of a NOMA system with intelligent reflecting surface, where the transmit power is minimized. Specifically, it is assumed that user clustering is given and the original problem is divided into two subproblems which are solved iteratively. In [37], users are clustered based on matching algorithm and then an AO algorithm is developed to transform the joint design into multiple subproblems, wherein relevant variables are optimized while keeping the remaining variables fixed. In [38], an algorithm based on deep reinforcement learning is proposed for the joint design in a single-cell MIMO NOMA system, where users are partitioned into two groups.

## 2.2 CD-NOMA

SCMA is a CD-NOMA scheme inspired from the well-known CDMA technique. In SCMA, groups of user data bits are directly mapped to sparse  $N$ -dimensional codewords. The latter

are selected from a codebook specified for each user and then transmitted over  $N$  radio resources, e.g., OFDMA subcarriers [39,40]. Hence, the SCMA encoder for the  $i$ th user can be defined as a one-to-one mapping from the set of  $u$ -bit tuples to a codebook  $\mathcal{X}_i \subset \mathbb{C}^N$ , using a function  $f_i : \mathbb{B}^u \rightarrow \mathcal{X}_i$ , where we define  $\mathbb{B} = \{0, 1\}$ . The codebook contains  $N$ -dimensional codewords, with cardinality  $|\mathcal{X}_i| = 2^u = U$ . Specifically, for  $\mathbf{b}_i = [b_{i,1}, \dots, b_{i,u}] \in \mathbb{B}^u$ , the corresponding codeword is obtained as,

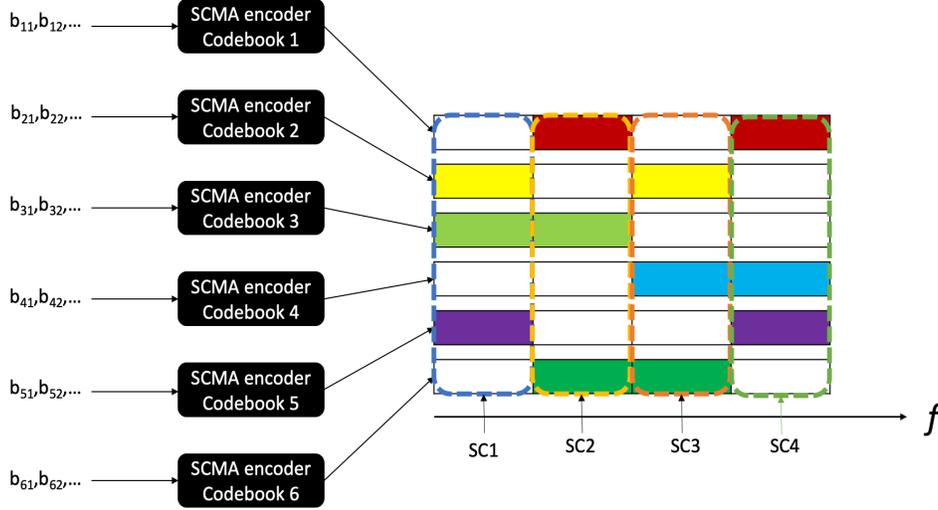
$$\mathbf{x}_i = f_i(\mathbf{b}_i) = [x_{i,1}, \dots, x_{i,N}] \quad (2.5)$$

where  $\mathbf{x}$  is a sparse vector with  $C < N$  non-zero elements. Each user is assigned  $C$  subcarriers such that no two users occupy the same subset of subcarriers. Hence, only  $q$  users can be supported by SCMA, as given by [14],

$$q = \binom{N}{C} = \frac{N!}{C!(N-C)!}. \quad (2.6)$$

Fig. 2.2 illustrates the operation of an SCMA encoder for  $C = 2$  non-zero elements and  $N = 4$  subcarriers (SCs), where a different color is employed for each one of the  $q = 6$  users.

The SCMA encoder can be expressed as  $f_i(\mathbf{b}_i) = \mathbf{S}_i g_i(\mathbf{b}_i)$  where  $g_i : \mathbb{B}^u \rightarrow \mathbb{C}^C$  is a mapping from the set of  $u$ -bit tuples to a  $C$ -dimensional constellation point with non-zero elements, while matrix  $\mathbf{S}_i$  maps this latter point into an  $N$ -dimensional codeword. It is worth noting that  $\mathbf{S}_i$  contains  $N - C$  all-zero rows and an identity matrix of order  $C$  is obtained by removing them. Hence, all the codewords in  $\mathcal{X}_i$  contain 0 in the same  $N - C$  positions. The codebook can be conveniently represented as a matrix, i.e.,  $\mathbf{X}_i \in \mathbb{C}^{N \times U}$  for the  $i$ th user, where the different columns, each with  $C$  non-zero elements at the same positions, correspond to the possible codewords.

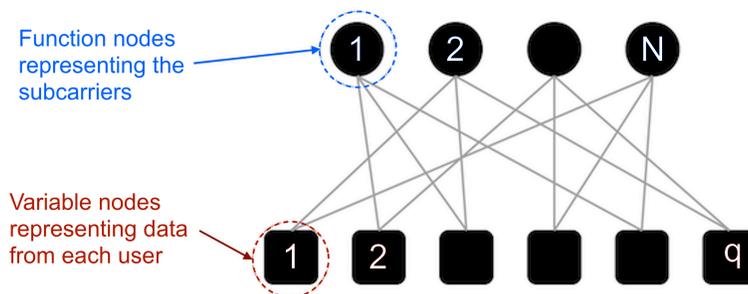


**Figure 2.2:** SCMA encoder model ( $C = 2$ ,  $N = 4$ ).

The positions (or indices) of the non-zero elements of the binary indicator vector obtained by  $\mathbf{f}_i = \text{diag}(\mathbf{S}_i \mathbf{S}_i^T) \in \mathbb{B}^{N \times 1}$  determine the set of subcarriers occupied by user  $i$ . In effect, the complete SCMA encoder structure for  $q$  users and  $N$  subcarriers can be represented by a factor graph with associated matrix  $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_q] \in \mathbb{B}^{N \times q}$ , as shown in Fig. 2.3. In this interpretation, subcarrier node  $n$  and user node  $i$  are connected if and only if the corresponding element of matrix  $\mathbf{F}$  is equal to 1, i.e.,  $[\mathbf{F}]_{n,i} = 1$ . For later use, we define  $\mathcal{A}_n$  and  $\mathcal{B}_i$  as the subsets of indices corresponding to the non-zero locations in the  $n$ th row and the  $i$ th column of matrix  $\mathbf{F}$ , respectively.

### 2.2.1 SCMA Encoder Design

There have been extensive studies devoted to the design of multidimensional constellations for downlink and uplink SCMA systems. In [41], the performance of a systematic sub-optimal design for the mother constellation (from which the individual user codebooks



**Figure 2.3:** The SCMA factor graph.

are derived) is investigated and a unified metric is proposed to obtain the optimum codebooks using a specific mother constellation. The authors in [42] evaluate the average BER performance of SCMA systems in which codebooks are based on star-QAM signaling constellations. Multidimensional constellations with a low number of projections are designed in [43] based on the extrinsic information transfer (EXIT) chart using a multistage optimization. Subsequently, an appropriate labeling method based on the EXIT chart is optimized for the resulting constellation. In [44], the design of SCMA codebooks based on star-QAM constellations is addressed and an analytical approach to obtain the theoretical BER performance over Rayleigh fading channels is proposed. The design of an efficient suboptimal SCMA codebook is proposed in [45] for a large scale scenario with growing number of resources and users.

The codebook design for an uplink SCMA system is formulated as an optimization problem aiming to maximize the average mutual information (AMI) in [46], wherein a message passing algorithm (MPA) with non-equiprobable distribution is introduced as the multiuser detection algorithm. In [47], a multi-stage optimization approach is proposed for the uplink SCMA codebook design such that the multi-dimensional mother constellation obtained from the first stage is combined with the user-specific mapping matrix from the

second stage to generate the final user codebooks. The problem of SCMA codebook design is addressed in [48], where an extended mother codebook is divided into several mother codebooks according to the modulation order. The authors in [49] propose a low-complexity codebook design based on uniquely decomposable constellation group for a SCMA system over Gaussian and Rayleigh fading channels. In [50], SCMA codebook design is addressed by maximizing the minimum Euclidean distance (MED) of superimposed codewords under power constraints, wherein the alternating maximization algorithm with exact penalty is utilized to obtain the solution.

### 2.2.2 SCMA Decoder Design

As mentioned before, for a non-orthogonal system like SCMA, more than one user's symbol are superposed on each subcarrier. Thus, joint multi-user detection algorithms such as MPA are needed. The MPA relies on the corresponding factor graph, as exemplified in Fig. 2.3, where each subcarrier is represented by a function node (FN) and data from each user by a variable node (VN). The FN degree, defined as the common number of connected neighboring VNs to each FN, and the VN degree, defined as the common number of connected neighboring FNs to each VN, are two useful parameters to design the structure of the codebook and will greatly impact the complexity in MPA decoding. Below, we provide additional details about the log-MPA [14] in the case of superimposed signal transmission from  $q$  users to a common access point (AP), where the users and AP are equipped with single-antenna for simplicity.

Let  $\mathbf{h}_n = [h_{n,1}, \dots, h_{n,q}] \in \mathbb{C}^q$  denote the channel vector between the  $q$  users and the AP for the  $n$ th subcarrier,  $n \in \mathcal{N} = \{1, \dots, N\}$ . Let us assume that user  $j \in \{1, \dots, q\}$  transmit codeword with index  $m_j \in \mathcal{U} = \{1, \dots, U\}$ . The received signal at the AP over the  $n$ th

subcarrier can then be expressed as,

$$y_n = \sum_{j \in \mathcal{A}_n} h_{n,j} [\mathbf{X}_j]_{n,m_j} + z_n \quad (2.7)$$

where  $z_n \sim \mathcal{CN}(0, \sigma^2)$  is additive Gaussian noise term with variance  $\sigma^2$ , assumed to be the same for each subcarrier.

The inputs of the Log-MPA are the received signal  $y_n$ , the channel vector  $\mathbf{h}_n \in \mathbb{C}^q$  on the  $n$ th subcarrier, and the codebook matrix of each user  $\mathbf{X}_i$ . The output of the algorithm are the log likelihood ratio (LLR) for the coded bits calculated from the probability estimations of each codeword, which are then served as input to the turbo decoder. The overall Log-MPA consists of the following steps:

**Step 1:** Initial calculation of the log conditional probabilities that is,

$$\phi(y_n | \{m_j\}, \mathbf{h}_n) = -\frac{1}{2\sigma^2} |y_n - \sum_{j \in \mathcal{A}_n} h_{n,j} [\mathbf{X}_j]_{n,m_j}|^2 \quad (2.8)$$

for all  $n \in \mathcal{N}$ ,  $m_j \in \mathcal{U}$ , and  $j \in \mathcal{A}_n$ .

**Step 2:** Iterative message passing along edges, which involves the update of the mutual information between FN and VN nodes

1. **Step 2a:** Update the mutual information for FN node  $g$  passing the extrinsic information to its neighboring VN nodes  $v_q$

$$\mathbf{I}_{g \rightarrow v_q}^{\log}(m_q) = \max_{\{m_j\}} \{ \phi(y_n | \{m_j\}, \mathbf{h}_n) + \sum_{j \in \mathcal{A}_n} \mathbf{I}_{v_j \rightarrow g}^{\log}(m_j) \} - \mathbf{I}_{v_q \rightarrow g}^{\log}(m_q) \quad (2.9)$$

for all  $m_q \in \mathcal{U}$  and  $q \in \mathcal{A}_g$ .

2. **Step 2b:** Update the mutual information for VN node  $v$  passing the extrinsic

information to its neighboring FN nodes  $g_{q_1}$  and  $g_{q_2}$

$$\mathbf{I}_{v \rightarrow g_{q_2}}^{log}(m) = \mathbf{I}_{g_{q_1} \rightarrow v}^{log}(m), \quad \mathbf{I}_{v \rightarrow g_{q_1}}^{log}(m) = \mathbf{I}_{g_{q_2} \rightarrow v}^{log}(m) \quad (2.10)$$

for all  $m \in \mathcal{U}$  and  $q_1, q_2 \in \mathcal{B}_v$ .

**Step 3:** Following  $N_{iter}$  iteration of step 2, calculate the LLR output at each VN node

$$\mathbf{I}_v^{log}(m) = \sum_{n \in \mathcal{B}_v} \mathbf{I}_{g_n \rightarrow v}^{log}(m), \quad (2.11)$$

$$LLR_v(b_i) = \max_{k \in \{b_i=0\}} (\mathbf{I}_v^{log}(m_k)) - \max_{k \in \{b_i=1\}} (\mathbf{I}_v^{log}(m_k)) \quad (2.12)$$

where  $m_k \in \mathcal{U}$  and  $i \in \{1, \dots, u\}$ .

The main steps of the Log-MPA are illustrated in Fig 2.4.

Although MPA can achieve near optimum performance, it entails high computational complexity. In recent years, many significant research efforts have been made to achieve a better trade-off between the performance and complexity of the decoder. In [51], a novel framework for MPA is proposed in which a belief threshold is applied to control the algorithm process. Tikhonov regularization is used to propose a low-complexity optimal modified sphere decoding (MSD) detection scheme for SCMA systems in [52], where the original rank-deficient detection problem is formulated as an equivalent full-rank detection problem. The design of a message-passing receiver for uplink grant-free SCMA systems is studied in [53], where the proposed receiver performs joint channel estimation, data decoding, and active user detection semi-blindly. In order to improve the system performance, a joint channel estimation and decoding scheme is proposed for polar-coded SCMA in [54]. The application of Gaussian-approximated message passing detection of SCMA along with the

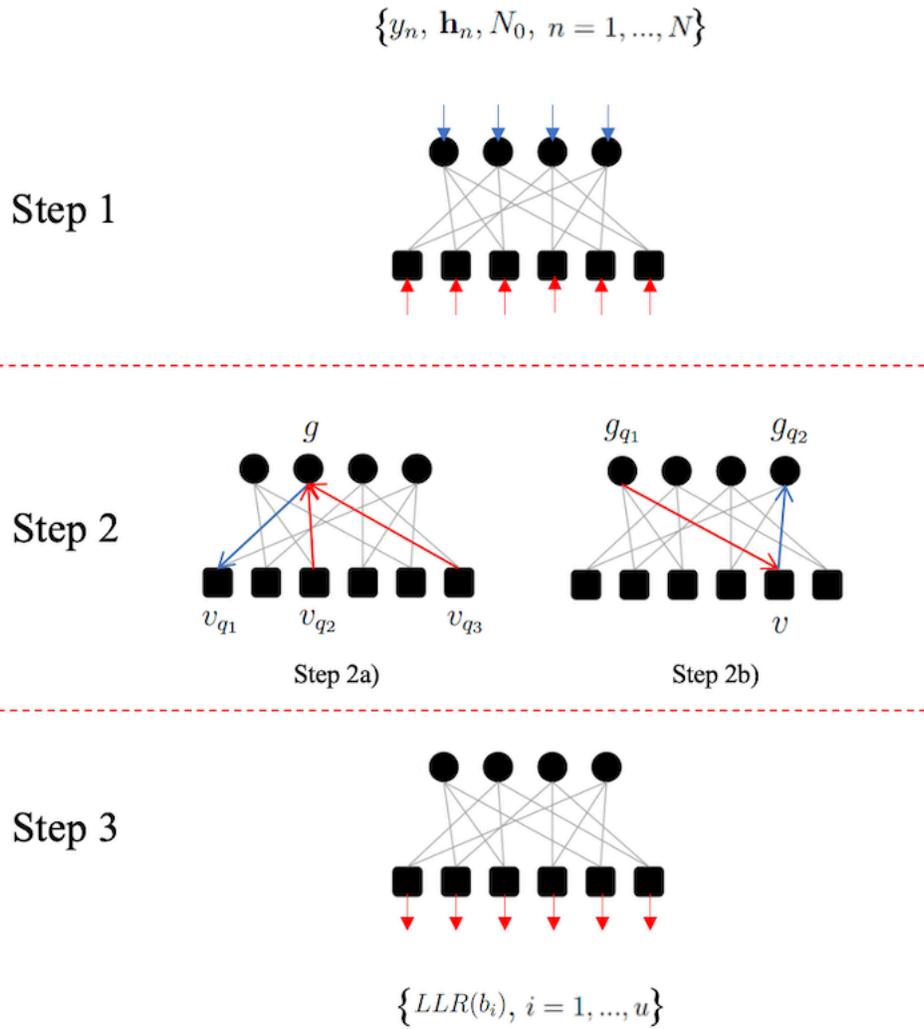


Figure 2.4: The MPA process model.

soft list decoding (SLD) of polar codes is investigated for polar-coded SCMA in [55].

The use of multiple antennas along with MIMO techniques in SCMA systems can lead to significant performance improvements in terms of network capacity and spectral efficiency. In [56], a joint sparse graph is constructed for a MIMO SCMA system model, and the corresponding virtual SCMA codebooks are designed for the detector, wherein the MPA is employed to reconstruct the transmitted data bits. In [57], a joint decoding algorithm is proposed for MIMO SCMA systems based on space frequency block codes (SFBC), which exhibits lower computational complexity than MPA and yet achieves a similar block error rate (BLER). A novel downlink MIMO mixed-SCMA scheme is proposed in [58], such that the transmitted codewords for each user over different antennas come from different codebooks. The authors in [59] propose near-optimal low-complexity iterative receivers based on factor graph for a downlink MIMO SCMA system over frequency selective fading channels.

Recently, applying deep learning (DL) methods to wireless communication problems has aroused great interest as it can bring significant performance improvements [60]. In particular, deep neural networks (DNNs), can deal with multidimensional and nonlinear characteristics of raw input and extract higher level features by using multiple processing layers. Motivated by this consideration, several works have investigated the application of DL techniques to enable autonomous derivation of an efficient algorithm for SCMA encoder and decoder. A DL-aided SCMA scheme is proposed in [61], where the constructed adaptive codebook and learned decoding strategy aim to minimize the BER. In [62], a DL-based decoding approach for SCMA is proposed in which a deep neural network is utilized to unfold the procedure of MPA. In [63], DL methods are utilized to conceive a SCMA auto-encoder which generates and then decodes codewords under AWGN channel. A deep learning decoder (DLD) is proposed in [64] for SCMA systems under AWGN and Rayleigh fading channel.

In [65], a deep neural network aided MPA (DNN-MPA) is developed to speed up the decoding procedure and reduce the computational complexity.

### 2.3 Comparison of NOMA Schemes

Table 2.1 highlights the main features, advantages, and disadvantages of major NOMA techniques. By comparing different schemes, it can be seen that SCMA outperforms other techniques in terms of spectral efficiency. However, a codebook design at the transmitter side and MPA at the receiver side are needed in SCMA systems. Hence, the performance improvement comes at the cost of high computational complexity. While each user is assigned a fixed number of resource blocks in SCMA (such that no two users occupy the same set of resource blocks), the number of resource blocks assigned to users in PDMA can be flexibly adjusted to match desired levels of overload. The MUSA scheme can be used to support massive connectivity with low cost and low power consumption; however, large number of spreading codes with relatively low cross-correlation are required.

PD-NOMA can achieve a useful balance between the sum rate of all multiplexed users and the throughput fairness among individual users through power allocation. Moreover, in PD-NOMA, more users can be multiplexed over a given number of resources which in turn, improves system throughput at the cost of transmit power. Due to its multiple benefits along with low complexity, PD-NOMA is a celebrated multiple access radio technology now being widely considered for emerging and future generations of wireless networks. As discussed in Section 2.1, this method has indeed been studied in several papers with different aims, such as the optimization of spectral or energy efficiency; however, many important questions remain to be answered [66].

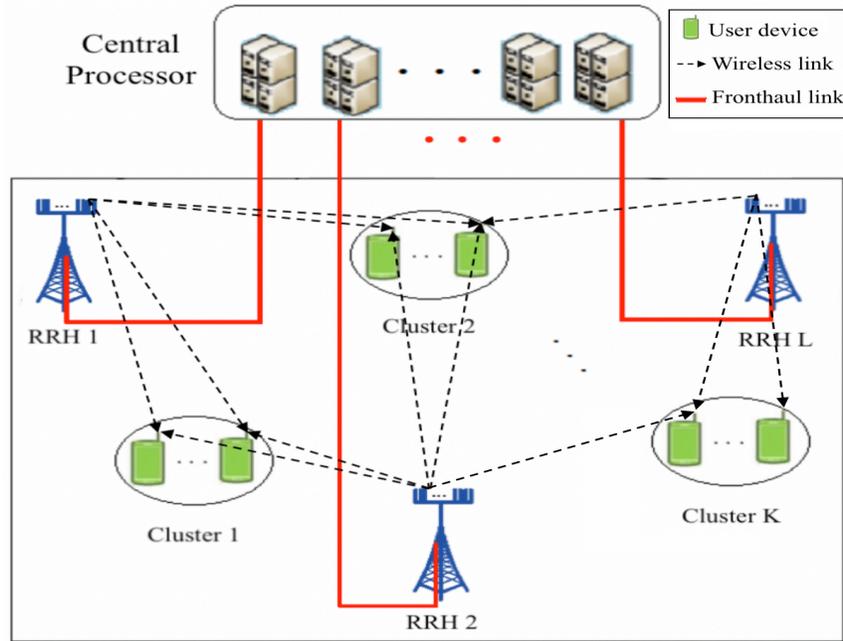
	PD-NOMA	SCMA	MUSA	PDMA
<b>Multiplexing domain</b>	Power	Code	Code	Power, Code, and/or Spatial
<b>Technique</b>	(i) Takes advantage of different channel conditions (ii) Allocates different power levels to the users	(i) Utilizes pre-designed user-specific codebooks (ii) Directly maps user data bits to sparse multi-dimensional complex codewords	(i) Utilizes low-correlation spreading sequences (ii) Users randomly pick one from pool of the sequences	(i) Utilizes nonorthogonal patterns (ii) Patterns are designed to maximize the diversity and minimize the overlaps among users
<b>Receiver</b>	SIC	MPA	SIC	SIC or MPA
<b>Advantage</b>	(i) Low complexity (ii) High downlink throughput	(i) High spectral efficiency (ii) Low bit-error-rate	(i) Low block error rate (ii) Massive connectivity	(i) High spatial diversity (ii) High uplink system capacity
<b>Disadvantage</b>	(i) High transmit power (ii) Error propagation in case of imperfect CSI	(i) High complexity of receiver (ii) Codebook design	(i) Increased inter-user interference (ii) Spread symbol design	(i) Increased inter-user interference (ii) Pattern optimization and design

**Table 2.1:** Comparison between different NOMA schemes.

## 2.4 C-RAN Architecture

C-RAN is considered as an emerging network architecture that is particularly suitable for handling interference allocating resources in a centralized fashion. As illustrated in Fig. 2.5, the C-RAN architecture consists of three main components, namely: the central processor, the remote radio heads (RRH) and the fronthaul links. The central processor, which is located in one or more data centers within the cloud, is responsible for all the baseband processing. The RRHs connect wireless devices to the network, alike base stations in current cellular networks. The fronthaul link provides connectivity (e.g., via dedicated optical fiber or microwave links) between the central processor and the RRHs. The C-RAN architecture concentrates the baseband processing in the central processor and coordinates the operation of the RRHs. This separation of the central processor and RRHs functionalities reduces the power consumption and complexity of the RRHs, since the latter only need to perform basic transceiver operations. Moreover, it reduces both the network capital expenditure and operating expenses.

Recently, the C-RAN architecture has aroused great interest for the implementation of



**Figure 2.5:** The C-RAN architecture.

MIMO NOMA transmission schemes. In [67], a novel framework for C-RAN is proposed in which two users are scheduled over the same resources according to power domain NOMA, while the performance of cell-edge users is further enhanced by means of coordinated beamforming. Stochastic geometry is used to analyze the outage probability of NOMA under C-RAN in [68], where power domain multiplexing along with SIC are employed to increase downlink system capacity. The application of beamforming along with power domain NOMA is investigated for cache-enabled C-RAN in [69]. The design of robust radio resource allocation and beamforming approaches for MIMO SCMA systems under C-RAN is studied in [70], where the aim is to maximize the total sum rate of users subject to a minimum required rate for each slice.

## 2.5 Concluding Remarks

In this chapter, we have presented a discussion of important design approaches for PD-NOMA in MIMO wireless systems from the perspectives of user clustering, beamforming and power allocation. We have also provided a literature review on SCMA encoder and decoder design. We have then presented a comparison between various NOMA schemes. Finally, we have surveyed recent works on the NOMA transmission schemes within C-RAN framework. The literature review in this chapter provides motivations and serves as basis for the research contributions presented in subsequent chapters of this thesis.

For MIMO NOMA systems, previous works have shown that the joint design of beamforming and power allocation can lead to significant improvements in terms of transmit power and spectral efficiency. Moreover, the application of spatial user clustering along with beamforming and power allocation techniques in MIMO NOMA systems has the potential to reduce the interference among users and hence improve spectral efficiency. To the best of our knowledge, in most of the existing works, the user clustering problem for MIMO NOMA systems is addressed separately from the beamforming and power allocation. While this approach considerably simplifies the design, it is inherently suboptimal and can therefore lead to performance degradation. Hence, a novel approach for the joint design of user clustering, beamforming and power allocation that can lead to superior performance is desired in a downlink multi-user MIMO NOMA system.

Previous works related to PD-NOMA have shown that the joint application of spatial user clustering along with beamforming techniques can improve spectral efficiency and reduce the total transmit power. Nonetheless, an approach to derive the user clustering and beamforming solutions has not yet been attempted in MIMO SCMA systems. Additionally, the C-RAN architecture has the potential to increase the number of

supported users in the network by using a common codebook for users in different clusters, while the effect of inter-cluster interference can be eliminated by centralized beamformer design and coordinated RRH operation. Hence, the design of energy-efficient user clustering and downlink beamforming approaches for MIMO SCMA in C-RAN remains of considerable interest.

In recent years, the application of DL-based methods to SCMA encoder and decoder design has aroused great interest as it can bring significant performance improvements. However, using DNNs bring its own plethora of challenges. For instance, by increasing the number of layers, the accuracy of detection will saturate at one point and may eventually degrade, notwithstanding the additional complexity in training. Moreover, DNNs suffer from problems such as vanishing gradients and curse of dimensionality. Even using a sufficiently large number of layers, DNN may not be able to learn simple functions (like an identity function, because it is trying to approximate a linear function with a nonlinear function). Hence, novel design approaches that overcome such disadvantages while enjoying comparable BER performance and computational complexity are required in the decoding stage.

## Chapter 3

# Joint User Clustering, Beamforming, and Power Allocation for NOMA

In this Chapter, we investigate the application of downlink beamforming along with PD-NOMA in a MIMO system. The joint design of user clustering, power allocation, and beamforming is formulated for a multi-user MIMO NOMA system operating at mmWave frequencies. To find the solution of the problem, two algorithms are developed whose advantages in terms of total transmit power and spectral efficiency are shown through simulation results.

### 3.1 Introduction

Since the main focus of this chapter is on PD-NOMA, in the sequel, we refer to the latter as NOMA for simplicity. As seen in [35–37], the application of user clustering along with

---

Parts of the material in this chapter have been presented at the 2022 Asilomar Conference on Signals, Systems, and Computers in Pacific Grove, CA, USA [71], and accepted for publication in the IEEE Transactions on Communications [72].

beamforming and power allocation techniques in MIMO NOMA systems has the potential to improve sum-rate and reduce the total transmit power. However, in the existing literature on MIMO NOMA, the user clustering problem is either neglected or addressed separately from the beamforming and power allocation design. While this approach brings important simplification in the design phase, it is inherently suboptimal and can lead to a degradation of system performance. As discussed before, high transmit power is a limiting factor in practical implementation of NOMA-based systems. Nevertheless, minimizing the total transmit power of NOMA under quality-of-service constraints for the user has only received limited attention. In [36], the authors explore this avenue by enforcing signal-to-interference-plus noise constraints on the users, but assuming that the user clustering is already available. In this paper, motivated by such considerations, we present and investigate a novel joint optimization framework for user clustering, beamforming, and power allocation, in a downlink MIMO NOMA system operating at mmWave frequencies. Our main contributions in addressing the above challenges are summarized as follows:

- Under the assumption that CSI is available at the BS, we formulate the joint design as a MINLP optimization problem where the objective is to minimize the total transmit power under practical constraints on the SINR, power allocation, and clustering, so as to ensure adequate SIC performance at the user terminals. In effect, this MINLP problem is non-convex and NP-hard.
- Different from [30, 33, 36], and [37] which propose suboptimal algorithms based on decomposed optimization (DO) or alternating optimization (AO), we develop a first algorithm based on branch-and-bound (BB) to obtain the optimal solution of the problem. BB is a systematic algorithm which lower bounds the objective function by relaxing the problem and partitions the feasible space successively to find the global

optimum solution, with guaranteed convergence to an  $\epsilon$ -optimal solution within a finite number of steps. Although the BB-based algorithm exhibits high computational complexity, it can serve as a benchmark for evaluating the performance of suboptimal algorithms for the same problem.

- To address the complexity issue, we then reformulate the original problem as an equivalent difference of convex functions (DC) program. To solve the resulting optimization problem, we develop a second algorithm based on the penalty dual-decomposition (PDD), which is guaranteed to converge to a local stationary solution of the transformed problem. The obtained solution is feasible for the original MINLP problem and meets the necessary conditions of optimal solutions.
- We evaluate the performance of the proposed algorithms for the joint design of user clustering, beamforming and power allocation, using in-depth simulations in a multi-user MIMO NOMA system with mmWave downlink transmissions and different parameter configurations. The results provide valuable insights into the advantages of user clustering and superiority of the proposed approaches over competing ones from the literature. Specifically, the results show that PDD-based algorithm outperforms DO and AO algorithms and exhibits a comparable performance to the optimal solution obtained by the BB-based algorithm in a small-scale network, yet with much lower complexity. The results also illustrate the benefits of user clustering and significant performance gains obtained in terms of transmit power when a joint design is considered compared to benchmark approaches such as  $K$ -means and random clustering. The convergence behavior of the new algorithms and the effect of various parameters on the system performance are also presented in the simulation results.

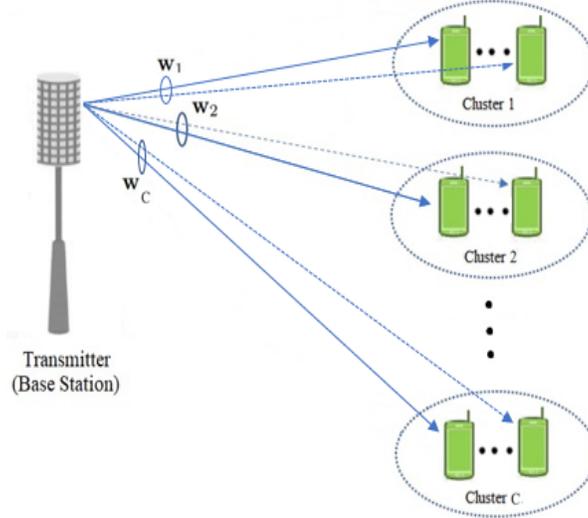
The rest of the chapter is organized as follows: Section 3.2 introduces the system model under consideration. The joint design of user clustering, beamforming and power allocation is formulated as an optimization problem in Section 3.3. In Section 3.4, the proposed BB-based algorithm for obtaining its solution is developed. The problem reformulation as a DC program along with the proposed PDD-based algorithm are presented in Section 3.5. The simulation results are presented in Section 3.6, followed by concluding remarks in Section 3.7.

## 3.2 System Model

As illustrated in Fig. 3.1, we consider downlink transmission in a multi-user MIMO NOMA system operating at mmWave frequencies. The system consists of a BS transmitter equipped with  $N$  antennas, serving  $K$  users indexed by  $k \in \mathcal{K} = \{1, \dots, K\}$ , where each user is equipped with a single-antenna receiver. The users are partitioned into  $C$  non-overlapping clusters indexed by  $c \in \mathcal{C} = \{1, \dots, C\}$ . The  $c$ th cluster comprises  $q_c$  users, which jointly share a common beamforming vector denoted as  $\mathbf{w}_c \in \mathbb{C}^N$  but can be allocated different power levels depending on their respective channel conditions.

### 3.2.1 Channel Model

The application of MIMO NOMA communications in the mmWave band is more challenging than in conventional low frequency scenarios, due to strenuous radio propagation characteristics. Herein, we focus on narrowband flat-fading transmissions, and assume without loss in generality that the BS transmitter employs a uniform linear array (ULA) with half-wavelength antenna spacing. Hence, the mmWave-based channel vector



**Figure 3.1:** The multi-user MIMO NOMA system model.

$\mathbf{h}_k \in \mathbb{C}^{N \times 1}$  from the BS to the  $k$ th user can be expressed as [87],

$$\mathbf{h}_k = \sqrt{N} \frac{a_{k,0} \mathbf{a}(\phi_k^0)}{\sqrt{1 + d_k^{\alpha_{\text{LOS}}}}} + \sqrt{N} \sum_{l=1}^L \frac{a_{k,l} \mathbf{a}(\phi_k^l)}{\sqrt{1 + d_k^{\alpha_{\text{NLOS}}}}} \quad (3.1)$$

where  $d_k$  denotes the distance between the BS and the  $k$ th user;  $\alpha_{\text{LOS}}$  and  $\alpha_{\text{NLOS}}$  are the path loss exponents of the line-of-sight (LOS) and non-LOS (NLOS) paths, respectively, and  $a_{k,l}$  denotes the complex gain for the  $l$ th path which follows a complex circular Gaussian distribution, i.e.,  $a_{k,l} \sim \mathcal{CN}(0, 1)$ . In the case of an ULA, the antenna array steering vector  $\mathbf{a}(\phi_k^l) \in \mathbb{C}^{N \times 1}$  is given by,

$$\mathbf{a}(\phi_k^l) = \frac{1}{\sqrt{N}} [1, e^{-j\pi \sin(\phi_k^l)}, \dots, e^{-j\pi(N-1) \sin(\phi_k^l)}]^T \quad (3.2)$$

where  $\phi_k^l \in [0, 2\pi]$  is the angle of departure of the  $l$ th path. In this work, we assume that near perfect CSI is available at the BS for joint processing, i.e. the channel vectors  $\mathbf{h}_k$  are

known with sufficient accuracy.

### 3.2.2 Signal Model

As previously explained, within each cluster of users, the BS utilizes NOMA superposition and beamforming simultaneously. That is, a common beamforming vector is shared by the users who are allocated different power levels on the basis of their respective channel conditions. Let  $\alpha_{c,k} \in [0, 1]$  denote the fraction of the total power available for the  $c$ th cluster that is allocated to the  $k$ th user. Specifically,  $\alpha_{c,k} = 0$  if the  $k$ th user does not belong to the  $c$ th cluster, and for each cluster  $\sum_k \alpha_{c,k} = 1$ . Hence, assuming that the  $m$ th user belongs to the  $c$ th cluster, its received signal can be expressed as,

$$r_m = \mathbf{h}_m^H \mathbf{w}_c \left( \sum_{k=1}^K \sqrt{\alpha_{c,k}} s_k \right) + \underbrace{\sum_{c' \neq c} \mathbf{h}_m^H \mathbf{w}_{c'} \left( \sum_{k=1}^K \sqrt{\alpha_{c',k}} s_k \right)}_{\text{Inter-cluster Interference}} + n_m \quad (3.3)$$

where  $s_k \in \mathbb{C}$  denotes the data symbol intended for the  $k$ th user and  $n_m \sim \mathcal{CN}(0, \sigma_m)$  is an additive noise term. The data symbols of different users are statistically independent with zero mean and unit variance, i.e.  $E[|s_k|^2] = 1$ .

### 3.2.3 SIC Procedure

At the receiver side, each user in a given cluster employs SIC to mitigate the intra-cluster interference, by decoding and removing the message of the weaker users in that cluster. Hence, user decoding order plays a key role in MIMO NOMA systems. For SIC procedure, let  $\pi_c(k)$  denote the order of decoding for the  $k$ th user in the  $c$ th cluster. Assume that both  $k$ th and  $m$ th users belong to the  $c$ th cluster and that  $\pi_c(k) < \pi_c(m)$ . For the SIC operation

at the  $m$ th user, the SINR of the  $k$ th user signal after perfect interference cancellation of the  $i$ th users,  $\pi_c(i) < \pi_c(k)$ , is given by,

$$\text{SINR}_{k,m}^c = \frac{|\mathbf{h}_m^H \mathbf{w}_c|^2 \alpha_{c,k}}{\sum_{\pi_c(k') > \pi_c(k)} |\mathbf{h}_m^H \mathbf{w}_c|^2 \alpha_{c,k'} + \sum_{c' \neq c} |\mathbf{h}_m^H \mathbf{w}_{c'}|^2 + \sigma_m^2} \quad (3.4)$$

where the first and second terms in the denominator of (3.4) represent the residual intra-cluster and inter-cluster interference, respectively. In the case  $k = m$ , (3.4) simply represents the available SINR of the  $m$ th user signal after perfect cancellation of signals from weaker users. To ensure successful SIC, the SINR of the  $k$ th user signal decoded at the  $m$ th user should be no smaller than the SINR of the  $k$ th user signal at the  $k$ th user, i.e.,  $\text{SINR}_{k,m}^c \geq \text{SINR}_{k,k}^c$ .

The optimal decoding order is determined by the effective channel gains which depend on the user clustering, beamforming, and channel gains [37]. For a given cluster partition and corresponding beamforming, the decoding order of the  $k$ th and  $m$ th users in the  $c$ th cluster is  $\pi_c(k) < \pi_c(m)$ , if the effective channel gain satisfies  $h_{c,k}^{\text{eff}} < h_{c,m}^{\text{eff}}$  [87], where,

$$h_{c,m}^{\text{eff}} = \frac{|\mathbf{h}_m^H \mathbf{w}_c|^2}{\sum_{c' \neq c} |\mathbf{h}_m^H \mathbf{w}_{c'}|^2 + \sigma_m^2}. \quad (3.5)$$

Under the assumption of the optimal decoding order, it is sufficient to ensure that the SINR of the  $k$ th user signal after removal of the other user signals is no smaller than a suitable SINR threshold, i.e.,  $\text{SINR}_{k,k}^c \geq \gamma_{\min}$  [87, Proposition 2].

In our work, the clustering and beamforming are not given and hence finding the optimal decoding order is not trivial. In this regard, without loss of generality, we assume that  $\|\mathbf{h}_1\| \leq \dots \leq \|\mathbf{h}_K\|$ . According to the NOMA principle, the *weaker* users (here, users with lower channel gains, i.e., smaller  $\|\mathbf{h}_k\|$ ) should be allocated a *larger* fraction of the

total available power at the transmitter and are decoded first at the receiver. To guarantee successful performance of SIC, for any  $k$ th and  $m$ th users in the  $c$ th cluster such that  $\forall k \leq m$ , the condition  $\text{SINR}_{k,m}^c \geq \gamma_{min}$  should be satisfied.

### 3.3 Problem Formulation

In this work, our goal is to jointly design the beamforming vectors, user clustering strategy and power allocation in the above MIMO NOMA system, so as to minimize the total transmit power at the BS, subject to relevant SINR, power and clustering constraints. Let the binary variable  $\iota_{c,k} = 1$  indicate that the  $k$ th user belongs to the  $c$ th cluster and  $\iota_{c,k} = 0$  otherwise. Then, it should be noted that  $\iota_{c,k} = 0$  must coerce into  $\alpha_{c,k} = 0$ . Moreover, to guarantee successful performance of SIC, the condition  $\text{SINR}_{k,m}^c \geq \gamma_{min}$  should be satisfied only if the  $k$ th and  $m$ th users are both in the  $c$ th cluster, i.e.,  $\iota_{c,k} = \iota_{c,m} = 1$ , which can be reformulated as,

$$\text{SINR}_{k,m}^c \geq \gamma_{min} \iota_{c,k} \iota_{c,m}, \quad \forall k \leq m \in \mathcal{K}, c \in \mathcal{C} \quad (3.6)$$

where  $\gamma_{min}$  represents a suitable SINR threshold<sup>1</sup>. If either of the  $k$ th or  $m$ th users are not in the  $c$ th cluster, the condition would be  $\text{SINR}_{k,m}^c \geq 0$  which is always true. Hence, the joint design problem for the MIMO NOMA system can be formulated as,

$$\text{P0} : \min_{\mathbf{Z}_0} \sum_{c \in \mathcal{C}} \|\mathbf{w}_c\|_2^2 \quad (3.7a)$$

$$\text{s.t.} \quad \text{SINR}_{k,m}^c \geq \gamma_{min} \iota_{c,k} \iota_{c,m}, \quad \forall k \leq m \in \mathcal{K}, c \in \mathcal{C} \quad (3.7b)$$

---

<sup>1</sup>Different values of the SNR threshold  $\gamma_{min}$  will be considered during simulations to explore the effects of this parameter on the performance of the system.

$$\sum_c \|\mathbf{w}_c\|_2^2 \leq P_{\max} \quad (3.7c)$$

$$0 \leq \alpha_{c,k} \leq \iota_{c,k}, \quad \forall c \in \mathcal{C}, k \in \mathcal{K} \quad (3.7d)$$

$$\sum_{k \in \mathcal{K}} \alpha_{c,k} = 1, \quad \forall c \in \mathcal{C} \quad (3.7e)$$

$$\sum_{k \in \mathcal{K}} \iota_{c,k} = q_c, \quad \forall c \in \mathcal{C} \quad (3.7f)$$

$$\sum_{c \in \mathcal{C}} \iota_{c,k} = 1, \quad \forall k \in \mathcal{K} \quad (3.7g)$$

$$\iota_{c,k} \in \{0, 1\}, \quad \forall c \in \mathcal{C}, k \in \mathcal{K} \quad (3.7h)$$

where to ease notation, we let  $\mathcal{Z}_0 \triangleq \{\mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\iota}\}$  with  $\mathbf{W} \triangleq \{\mathbf{w}_c | c \in \mathcal{C}\}$ ,  $\boldsymbol{\alpha} \triangleq \{\alpha_{c,k} | c \in \mathcal{C}, k \in \mathcal{K}\}$ , and  $\boldsymbol{\iota} \triangleq \{\iota_{c,k} | c \in \mathcal{C}, k \in \mathcal{K}\}$ . The objective function in (3.7a) represents the total transmit power at the BS under the assumptions made in Section 3.2.2 for the data symbol  $s_k$ . Constraint (3.7b) corresponds to (3.6). The quantity  $P_{\max}$  in constraint (3.7c) denotes the maximum available transmit power<sup>2</sup>. Constraint (3.7d) compels the power allocation to be zero if the  $k$ th user is not in the  $c$ th cluster. Constraints (3.7f) and (3.7g) reflect the facts that the  $c$ th cluster must contain  $q_c$  users, and each user must be scheduled in only one cluster, respectively.

The SINR constraint (3.7b) is difficult to handle due to the binary variables on the right-hand side of the inequality and the fractional form (3.4) of the SINR on the left-hand side. By introducing auxiliary variables,  $x_{k,m}^c$ , the SINR constraint can be rewritten as follows,

$$\text{SINR}_{k,m}^c \geq x_{k,m}^c \quad (3.8)$$

---

<sup>2</sup> $P_{\max}$  can be any large number so as to impose no constraint on the beamformer during transmission. However, it should be noted that for the proposed BB-base algorithm, all variables must be bounded below and/or above to guarantee global optimality, as explained later.

$$\gamma_{min} \iota_{c,k} \iota_{c,m} \leq x_{k,m}^c \quad (3.9)$$

which are assumed to apply  $\forall k \leq m \in \mathcal{K}, c \in \mathcal{C}$ . To tackle the difficulty posed by the fractional SINR form, we introduce the auxiliary variable  $y_{k,m}^c$  as the upper bound of the denominator, and then equivalently express (3.8) as the following pair of constraints,

$$|\mathbf{h}_m^H \mathbf{w}_c|^2 \alpha_{c,k} \geq x_{k,m}^c y_{k,m}^c \quad (3.10)$$

$$\sum_{k'=k+1}^K |\mathbf{h}_m^H \mathbf{w}_c|^2 \alpha_{c,k'} + \sum_{c' \neq c} |\mathbf{h}_m^H \mathbf{w}_{c'}|^2 + \sigma_m^2 \leq y_{k,m}^c \quad (3.11)$$

Without loss of optimality, using the factorable programming technique [74], a tight relaxation for the binary bilinear terms in (3.9) can be derived as,

$$x_{k,m}^c \geq \gamma_{min} (\iota_{c,k}^u \iota_{c,k} + \iota_{c,m}^u \iota_{c,m} - \iota_{c,k}^u \iota_{c,m}^u) \quad (3.12)$$

where  $\iota_{c,k}^u$  denotes the upper bound of  $\iota_{c,k}$  which is equal to one in this case, i.e.,  $\iota_{c,k}^u = 1$ . In case that either of  $\iota_{c,k}$  or  $\iota_{c,m}$  are equal to 1, both (3.9) and (3.12) would be the same. If  $\iota_{c,k}$  and  $\iota_{c,m}$  are equal to 0, (3.9) gives  $x_{k,m}^c \geq 0$ , while (3.12) gives  $x_{k,m}^c \geq -\gamma_{min}$ . In this case, these inequalities exert no influence on the final solution of the problem, since  $x_{k,m}^c$  is the lower bound of  $\text{SINR}_{k,m}^c$  and always greater than zero.

From the above discussion, we can reformulate problem P0 into an equivalent problem as given below

$$\text{P1 : } \min_{\mathcal{Z}_1} \sum_c \|\mathbf{w}_c\|_2^2 \quad (3.13a)$$

$$\text{s.t. (3.7c)-(3.7h), (3.10), (3.11), (3.12)} \quad (3.13b)$$

where we introduce  $\mathcal{Z}_1 \triangleq \{\mathbf{W}, \mathbf{X}, \mathbf{Y}, \boldsymbol{\alpha}, \boldsymbol{\iota}\}$  with  $\mathbf{X} \triangleq \{x_{k,m}^c | k \leq m \in \mathcal{K}, c \in \mathcal{C}\}$ , and  $\mathbf{Y} \triangleq \{y_{k,m}^c | k \leq m \in \mathcal{K}, c \in \mathcal{C}\}$ . We emphasize that the presence of non-convex constraints in a MINLP can make the problem significantly more difficult to solve. In fact, MINLP problems are already known to be NP-hard and the addition of non-convex constraints only adds to the difficulty of the problem [73]. Hence, problem P1 which is a non-convex MINLP, is NP-hard. Accordingly, obtaining the optimal solution is challenging due to the non-convexity of the constraints and the combinatorial nature of the user clustering variables,  $\iota_{c,k}$ . In the following sections, we develop two algorithms to find the near-optimal solutions of P1.

### 3.4 Proposed BB-Based Algorithm

MINLP represents one of the most challenging classes of mathematical programming, typically requiring either simulation-based approaches or special decomposition methods for the solution of very large scale problems. In particular, the application of the BB algorithms to the MINLPs has shown promising results [74]. BB is a systematic method for solving non-convex optimization problems. In the present context, BB can be applied to P1 by constructing and solving its relaxation and branching the feasible space successfully, where *relaxation* herein refers to transformation into a convex form. A lower bound on the optimal objective function value of P1 is obtained by solving the relaxation problem, while an upper bound is derived by employing upper bounding heuristics. The procedure quits if the upper and lower bounds are sufficiently close, as they delimit the global minimum value of the objective. Otherwise, the feasible space is partitioned into convex sets and explored further to locate an optimal solution and verify its globality. As illustrated in Fig. 3.2, the BB process is typically depicted as a tree where the nodes and branches correspond

to bounding and partitioning respectively.

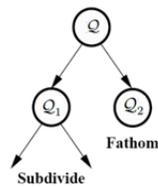
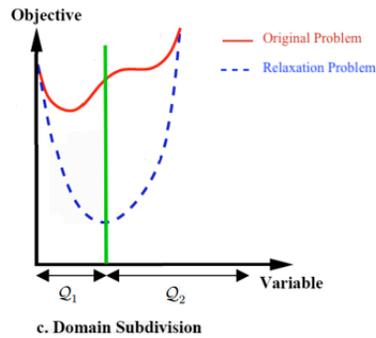
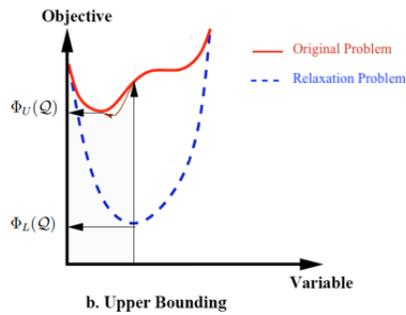
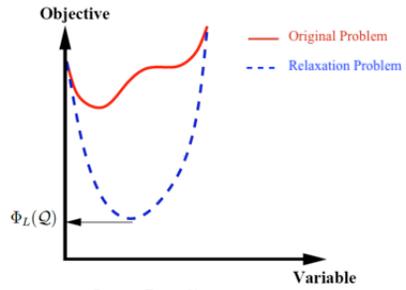


Figure 3.2: An illustration of the BB process (adapted from [74]).

### 3.4.1 Convex Relaxation

The relaxation problem is constructed by enlarging the feasible region and/or underestimating the objective function of P1. Hence, the solution of the relaxation problem provides a lower bound on the optimal objective function value of P1. In this subsection, we introduce some effective convex relaxations for the non-convex constraints of P1.

We introduce new variables  $\beta_{c,m}$ ,  $u_{c,m}$ , and  $l_{c,m}$ ,  $\forall m \in \mathcal{K}$  and  $c \in \mathcal{C}$ , such that,

$$\beta_{c,m} = \mathbf{h}_m^H \mathbf{w}_c \in \mathbb{C}, \quad (3.14)$$

$$|\beta_{c,m}|^2 \leq u_{c,m} \quad (3.15)$$

$$|\beta_{c,m}| \geq l_{c,m} \quad (3.16)$$

Without loss of optimality, we can then reformulate problem P1 into a more tractable form as given below,

$$\text{P2 : } \min_{\mathcal{Z}_2} \sum_c \|\mathbf{w}_c\|_2^2 \quad (3.17a)$$

$$\text{s.t. (3.7c)-(3.7h), (5.12)-(3.16)} \quad (3.17b)$$

$$\sum_{k'=k+1}^K u_{c,m} \alpha_{c,k'} + \sum_{c' \neq c} u_{c',m} + \sigma^2 \leq y_{k,m}^c \quad (3.17c)$$

$$l_{c,m}^2 \alpha_{c,k} \geq x_{k,m}^c y_{k,m}^c \quad (3.17d)$$

$$x_{k,m}^c \geq \gamma \min(\iota_{c,k} + \iota_{c,m} - 1) \quad (3.17e)$$

where we let  $\mathcal{Z}_2 \triangleq \{\mathbf{U}, \mathbf{L}, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \boldsymbol{\alpha}, \boldsymbol{\iota}, \boldsymbol{\beta}\}$  with  $\mathbf{U} \triangleq \{u_{c,m} | c \in \mathcal{C}, m \in \mathcal{K}\}$ ,  $\mathbf{L} \triangleq \{l_{c,m} | c \in \mathcal{C}, m \in \mathcal{K}\}$ , and  $\boldsymbol{\beta} \triangleq \{\beta_{c,m} | c \in \mathcal{C}, m \in \mathcal{K}\}$ . The feasible region of P2 is non-convex due

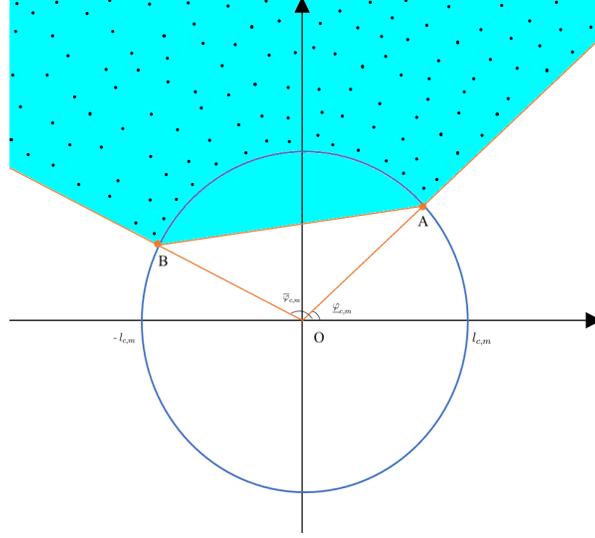
to constraints (3.7h), (3.16), (3.17c), and (3.17d), while the objective function is convex. We construct a convex relaxation of P2 by outer-approximating its feasible space with a convex set. To the end, we can simply drop the binary constraints and treat the variables as continuous ones in the range  $[0, 1]$ . Next, we construct convex relaxations for the non-convex constraint (3.16) by applying the following proposition.

**Proposition 3.1:** Let  $\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m})$  denote the subset of complex numbers  $\beta_{c,m} = \rho_{c,m}e^{j\varphi_{c,m}}$ , with amplitude and phase respectively satisfying the inequalities  $\rho_{c,m} \geq l_{c,m}$  and  $\underline{\varphi}_{c,m} \leq \varphi_{c,m} \leq \overline{\varphi}_{c,m}$ , where  $l_{c,m} \geq 0$  and  $0 \leq \underline{\varphi}_{c,m} \leq \overline{\varphi}_{c,m} \leq 2\pi$ . Suppose that  $\overline{\varphi}_{c,m} - \underline{\varphi}_{c,m} \leq \pi$ , then the convex envelop of  $\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m})$  is given by [75]

$$\begin{aligned} \text{Conv}(\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m})) &= \{\beta_{c,m} \in \mathbb{C} \mid \sin(\underline{\varphi}_{c,m})\Re(\beta_{c,m}) - \cos(\underline{\varphi}_{c,m})\Im(\beta_{c,m}) \leq 0, \\ &\quad \sin(\overline{\varphi}_{c,m})\Re(\beta_{c,m}) - \cos(\overline{\varphi}_{c,m})\Im(\beta_{c,m}) \geq 0, \\ &\quad f_{c,m}\Re(\beta_{c,m}) + g_{c,m}\Im(\beta_{c,m}) \geq (f_{c,m}^2 + g_{c,m}^2)l_{c,m}\} \end{aligned} \quad (3.18)$$

where  $f_{c,m} = (\cos(\underline{\varphi}_{c,m}) + \cos(\overline{\varphi}_{c,m}))/2$  and  $g_{c,m} = (\sin(\underline{\varphi}_{c,m}) + \sin(\overline{\varphi}_{c,m}))/2$ .

*Proof.* Fig. 3.3 illustrates the relationship between sets  $\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m})$  and  $\text{Conv}(\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m}))$ . In Fig. 3.3, the area outside the blue circle corresponds to  $|\beta_{c,m}| \geq l_{c,m}$ , the set  $\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m})$  corresponds to the light blue polka dot region outside the arc AB which is obviously nonconvex, and its convex envelope,  $\text{Conv}(\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m}))$ , corresponds to the light blue region determined by three lines,  $\overline{OA}$ ,  $\overline{OB}$ , and  $\overline{AB}$ . It is simple to see that the two extreme points of set  $\text{Conv}(\mathcal{D}_{[\underline{\varphi}_{c,m}, \overline{\varphi}_{c,m}]}(l_{c,m}))$  are  $A = (\cos(\underline{\varphi}_{c,m}), \sin(\underline{\varphi}_{c,m}))$  and  $B = (\cos(\overline{\varphi}_{c,m}), \sin(\overline{\varphi}_{c,m}))$ . To prove the proposition, it is sufficient to determine the three lines  $\overline{OA}$ ,  $\overline{OB}$ , and  $\overline{AB}$  which can be determined by equations  $\sin(\underline{\varphi}_{c,m})x - \cos(\underline{\varphi}_{c,m})y = 0$ ,  $\sin(\overline{\varphi}_{c,m})x - \cos(\overline{\varphi}_{c,m})y = 0$ , and



**Figure 3.3:** An illustration for Proposition 3.1.

$f_{c,m}x + g_{c,m}y = (f_{c,m}^2 + g_{c,m}^2)l_{c,m}$ , respectively.  $\square$

It can be verified that as  $\bar{\varphi}_{c,m} - \varphi_{c,m}$  goes to zero, the convex envelope becomes tight, i.e.,  $\text{Conv}(\mathcal{D}_{[\varphi_{c,m}, \bar{\varphi}_{c,m}]}(l_{c,m}))$  is equal to  $\mathcal{D}_{[\varphi_{c,m}, \bar{\varphi}_{c,m}]}(l_{c,m})$ . Note that the convex envelope does not take effect when  $\bar{\varphi}_{c,m} - \varphi_{c,m} > \pi$ . According to Proposition 3.1, we can now replace constraint (3.16) by  $\beta_{c,m} \in \text{Conv}(\mathcal{D}_{[\varphi_{c,m}, \bar{\varphi}_{c,m}]}(l_{c,m}))$ .

By defining  $\mu_{k,m}^c = u_{c,m}\alpha_{c,k}$ ,  $\nu_{k,m}^c = l_{c,m}\alpha_{c,k}$ ,  $\zeta_{k,m}^c = l_{c,m}\nu_{k,m}^c$  and  $\eta_{k,m}^c = y_{k,m}^c x_{k,m}^c$ , we can reformulate (3.17c) and (3.17d) as below,

$$\sum_{k'=k+1}^K \mu_{k',m}^c + \sum_{c' \neq c} u_{c',m} + \sigma^2 - y_{k,m}^c \leq 0 \quad (3.19)$$

$$\zeta_{k,m}^c \geq \eta_{k,m}^c. \quad (3.20)$$

We address the relaxation of the bilinear terms, i.e.,  $\mu_{k,m}^c$ ,  $\nu_{k,m}^c$ ,  $\zeta_{k,m}^c$ , and  $\eta_{k,m}^c$ , using their convex and concave envelopes. For two variables  $y_i$  and  $y_j$ , the bilinear term defined as

$y_{i,j} = y_i y_j$  is relaxed using the convex and concave envelopes as [74],

$$y_{i,j} \geq y_i^u y_j + y_j^u y_i + y_i^u y_j^u \quad (3.21)$$

$$y_{i,j} \geq y_i^l y_j + y_j^l y_i - y_i^l y_j^l \quad (3.22)$$

$$y_{i,j} \leq y_i^u y_j + y_j^l y_i - y_i^u y_j^l \quad (3.23)$$

$$y_{i,j} \leq y_i^l y_j + y_j^u y_i - y_i^l y_j^u \quad (3.24)$$

where the superscripts  $u$  and  $l$  stand for upper bound and lower bound of the corresponding variable, respectively. Since the bilinear terms are convex or concave in both variables, the convex and concave envelopes then provide tight approximations.

### 3.4.2 Proposed Algorithm

For ease of notation, let  $\Gamma = [\boldsymbol{\nu}, \boldsymbol{\alpha}, \boldsymbol{\varphi}] \in \mathbb{R}^{C \times 3K}$  be the variable matrix of interest. Initially, this matrix belongs to the box  $\mathcal{Q}_{\text{init}} = [\underline{\Gamma}, \overline{\Gamma}]$  where,

$$\underline{\Gamma} \triangleq \mathbf{0}_{C \times 3K}, \quad \overline{\Gamma} \triangleq [\mathbf{1}_{C \times 2K}, 2\pi \mathbf{1}_{C \times K}]. \quad (3.25)$$

The proposed BB-based algorithm involves a sequence of iterations indexed by integer  $t \in \mathbb{N}$ . Within each iteration,  $\mathcal{R}^t$ ,  $\Phi_U^t$ , and  $\Phi_L^t$  denote the box list, the upper bound, and the lower bound of the optimal objective function value of problem P2, respectively. Let  $\Phi_U(\mathcal{Q})$  and  $\Phi_L(\mathcal{Q})$  represent the upper bound and the lower bound of the objective function value over a given box  $\mathcal{Q}$ . Each iteration consists of two main parts, i.e. Branch and Bound, as further explained below.

1) *Branch*: At the  $t$ th iteration, we select a box in  $\mathcal{R}^t$  and split it into two smaller ones. An effective method for selecting the candidate box is to choose the one with the least lower bound, i.e.,  $\mathcal{Q}^* = \arg \min_{\mathcal{Q} \in \mathcal{R}^t} \Phi_L(\mathcal{Q})$ . The selected box  $\mathcal{Q}^* = [\mathbf{A}, \mathbf{B}]$  is then split along the longest edge, i.e.,  $(i^*, j^*) = \arg \max_{i,j} \{b_{i,j} - a_{i,j}\}$ , to create two boxes with equal size, that is,

$$\mathcal{Q}_1^* = \begin{cases} [\mathbf{A}, \mathbf{B} - \mathbf{J}_{i^*,j^*}], & \text{if } j^* \leq K \\ [\mathbf{A}, \mathbf{B} - \frac{1}{2}(b_{i^*,j^*} - a_{i^*,j^*})\mathbf{J}_{i^*,j^*}], & \text{if } j^* > K \end{cases} \quad (3.26)$$

$$\mathcal{Q}_2^* = \begin{cases} [\mathbf{A} + \mathbf{J}_{i^*,j^*}, \mathbf{B}], & \text{if } j^* \leq K \\ [\mathbf{A} + \frac{1}{2}(b_{i^*,j^*} - a_{i^*,j^*})\mathbf{J}_{i^*,j^*}, \mathbf{B}], & \text{if } j^* > K \end{cases} \quad (3.27)$$

where  $\mathbf{J}_{i^*,j^*}$  is a  $C \times 3K$  matrix with  $(i^*, j^*)$ th entry equal to 1 and all other entries equal to 0.

2) *Bound*: The bounding operation consists in computing the upper and lower bounds over the newly added box  $\mathcal{Q} \in \{\mathcal{Q}_1^*, \mathcal{Q}_2^*\}$ , and update the lower bound  $\Phi_L^t$  and the upper bound  $\Phi_U^t$ .

*Lower bound*: The lower bound  $\Phi_L(\mathcal{Q})$  is obtained by solving the convex relaxation of problem P2 developed in Section 3.4.1, where the variables  $\iota, \alpha$  and  $\phi$  are searched over the box  $\mathcal{Q}$ . Since the relaxation problem is convex, it can be solved via any general-purpose solver using interior-point methods [76]. Note that if the relaxation problem is infeasible, the box  $\mathcal{Q}$  does not contain the optimal solution. In this case, we simply set  $\Phi_L(\mathcal{Q})$  and  $\Phi_U(\mathcal{Q})$  as  $+\infty$ .

After obtaining the lower bounds  $\Phi_L(\mathcal{Q})$ , for  $\mathcal{Q} \in \{\mathcal{Q}_1^*, \mathcal{Q}_2^*\}$ , we can obtain the updated box list  $\mathcal{R}^{t+1}$  by removing  $\mathcal{Q}^*$  from  $\mathcal{R}^t$  and adding  $\mathcal{Q}_1^*$  and  $\mathcal{Q}_2^*$  if their lower bounds are less

than or equal to the current best upper bound  $\Phi_U^t$ , i.e.,  $\mathcal{R}^{t+1} = (\mathcal{R}^t - \{\mathcal{Q}^*\}) \cup \{\mathcal{Q}_i^* | \Phi_L(\mathcal{Q}_i^*) \leq \Phi_U^t, i = 1, 2\}$ . The lower bound of the optimal objective value of the original problem is updated as  $\Phi_L^{t+1} = \min_{\mathcal{Q} \in \mathcal{R}^{t+1}} \Phi_L(\mathcal{Q})$ .

*Upper bound:* To obtain an upper bound, we need to find a feasible solution of problem P1, which can be done by utilizing the solution of the relaxation problem. Let  $\{\mathbf{W}^*, \boldsymbol{\alpha}^*, \boldsymbol{\iota}^*, \boldsymbol{\beta}^*\}$  denote the optimal solution of the latter problem. If  $\boldsymbol{\iota}^* \notin \mathbb{B}^{C \times K}$ , we set  $\Phi_U(\mathcal{Q})$  as  $+\infty$ . If  $\boldsymbol{\iota}^* \in \mathbb{B}^{C \times K}$  and constraint (3.7b) is satisfied, then the solution of the relaxation problem is a feasible solution of the original problem. If constraint (3.7b) is not satisfied, we can scale  $\mathbf{W}^*$  to be feasible. Therefore, a feasible solution of problem P2 is given by  $\{\tilde{\mathbf{W}}, \tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\iota}}\}$  where  $\tilde{\boldsymbol{\iota}} = \boldsymbol{\iota}^*$ ,  $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha}^*$ , and,

$$\tilde{\mathbf{W}}_c = \max\left\{1, \max_{\{k \leq m | \iota_{c,k}^* = \iota_{c,m}^* = 1\}} \left\{ \frac{\sqrt{a_{k,m}^c}}{|\beta_{c,m}^*|} \right\}\right\} \mathbf{W}_c^*, \quad (3.28)$$

$$a_{k,m}^c = \frac{(C-1) \|\mathbf{h}_m\|^2 P_{\max} \gamma_{\min} + \sigma^2 \gamma_{\min}}{\alpha_{c,k} - \sum_{k'=k+1}^K \alpha_{c,k'} \gamma_{\min}}, \quad (3.29)$$

$$\Phi_U(\mathcal{Q}) = \sum_c \|\tilde{\mathbf{w}}_c\|_2^2. \quad (3.30)$$

To derive these equations, we make the assumption that the optimal beamforming solution, i.e.,  $\mathbf{W}_c^*$ , is multiplied by a scaling factor to meet the requirements of constraint (3.7b). Next, a lower bound for this factor is calculated for each user in the cluster. Then, the maximum lower bound among all users in a cluster is chosen to obtain the feasible solution  $\tilde{\mathbf{W}}_c$ . Finally, a better upper bound of the optimal objective value of P2 is obtained by applying the update  $\Phi_U^{t+1} = \min_{\mathcal{Q} \in \mathcal{R}^{t+1}} \Phi_U(\mathcal{Q})$ . The overall BB-based algorithm for solving problem P1 is summarized in Algorithm 3.1.

---

**Algorithm 3.1:** The BB-based algorithm.

---

**Initialization:** Initialize  $\mathcal{R}^t \leftarrow \mathcal{Q}_{\text{init}}$ , Find the lower bound  $\Phi_L(\mathcal{Q}_{\text{init}})$  by solving the relaxation of problem (3.17), and the upper bound  $\Phi_U(\mathcal{Q}_{\text{init}})$  according to (3.30). Set  $t = 0$ ,  $\Phi_L^0 = \Phi_L(\mathcal{Q}_{\text{init}})$ ,  $\Phi_U^0 = \Phi_U(\mathcal{Q}_{\text{init}})$ , and the tolerance  $\epsilon > 0$ .

**While:**  $(\Phi_U^t - \Phi_L^t)/\Phi_L^t > \epsilon$

1) **Branch:** Select the box  $\mathcal{Q}^*$  in  $\mathcal{R}^t$  with the least lower bound and split it into two boxes  $\mathcal{Q}_1^*$  and  $\mathcal{Q}_2^*$  according to the splitting rules (3.26)-(3.27).

2) **Bound:** For each box  $\mathcal{Q}_i^*$  ( $i = 1, 2$ ), find its lower bound  $\Phi_L(\mathcal{Q}_i^*)$  by solving the relaxation of problem (3.17) and its upper bound  $\Phi_U(\mathcal{Q}_i^*)$  according to (3.30).

3) Update  $\mathcal{R}^{t+1} = (\mathcal{R}^t - \{\mathcal{Q}^*\}) \cup \{\mathcal{Q}_i^* | \Phi_L(\mathcal{Q}_i^*) \leq \Phi_U^t, i = 1, 2\}$ .

4) Update  $\Phi_L^{t+1} = \min_{\mathcal{Q} \in \mathcal{R}^{t+1}} \Phi_L(\mathcal{Q})$ .

5) Update  $\Phi_U^{t+1} = \min_{\mathcal{Q} \in \mathcal{R}^{t+1}} \Phi_U(\mathcal{Q})$ .

6) Set  $t \leftarrow t + 1$ .

**End**

---

### 3.4.3 Convergence and Complexity Analysis

Let  $\text{size}(\mathcal{Q})$  denote the maximum half-length of the sides of box  $\mathcal{Q}$ . The following Theorem 1 shows that the upper and lower bounds over a box region become tight as the box shrinks to a point. That is, as  $\text{size}(\mathcal{Q})$  goes to zero, the difference between upper and lower bounds converges to zero.

**Theorem 3.1:** For any  $\epsilon > 0$ , let  $\delta \in (0, \frac{\pi}{2})$  be defined as,

$$\delta = \arccos\left(\frac{1}{\sqrt{1 + \epsilon}}\right). \quad (3.31)$$

For any given  $\mathcal{Q} \subset \mathcal{Q}_{\text{init}}$ , if  $\text{size}(\mathcal{Q}) \leq 2\delta$ , we can obtain an  $\epsilon$ -optimal solution of problem P1 by applying Algorithm 3.1, i.e.,  $\frac{\Phi_U^t - \Phi_L^t}{\Phi_L^t} \leq \epsilon$  for some integer  $t$ .

*Proof.* See Appendix A.1. □

By adopting the splitting rule in Section 3.4.2, at least one box in the partition has

size not exceeding  $\delta$  if  $t$  is sufficiently large. It follows from Theorem 3.1 that when the corresponding box is added to the list at the  $t$ th iteration, the algorithm should terminate and return  $\epsilon$ -optimal solution. We refer the interested reader to [77] for additional details on the convergence properties of the branch and bound method.

The following Theorem proves that the number of iterations of the BB-based algorithm for obtaining the solution is finite.

**Theorem 3.2:** For any given  $\epsilon > 0$ , the proposed BB-based algorithm returns an  $\epsilon$ -optimal solution of the given problem within at most,

$$T_B = \left\lceil \left( \frac{4\pi}{\delta^2} \right)^{CK} \right\rceil + 1 \quad (3.32)$$

iterations, where  $\delta$  is given in (3.31).

*Proof.* See Appendix A.2. □

At each iteration, the complexity of the proposed BB-based algorithm is dominated by calculating the lower bounds in Step 2. Obtaining the lower bound requires solving a convex quadratic problem via a general-purpose solver, e.g., SeDuMi in CVX [78] with a complexity of  $\mathcal{O}((KCN)^{3.5})$ . Assuming that the BB-based algorithm converges after  $T_B$  iterations, the worst-case computational complexity can be expressed as  $\mathcal{O}(T_B(KCN)^{3.5})$ . Theorem 3.2 shows that  $T_B$  can be very large if the tolerance  $\epsilon$  is small. Nevertheless, the proposed BB-based algorithm can be used as the system performance benchmark.

### 3.5 Proposed PDD-Based Algorithm

To circumvent the complexity issue of the BB-based algorithm, we herein conceive an alternative algorithm based on the penalty dual-decomposition (PDD) that can obtain a suboptimal solution of problem P1 with reduced computational cost. Specifically, the proposed PDD-based algorithm consists of two embedded loops, where the inner loop seeks to approximately solve an augmented Lagrangian (AL) problem (see e.g., [79–81]) via the concave-convex procedure (CCCP) [82], while the outer loop updates either the dual variables or the penalty parameter, depending on a constraint violation status. Below, we first transform the problem P1 into a more tractable form, which is then used as a basis to develop our proposed PDD-based algorithm. Finally, the convergence and computational complexity of the new algorithm are discussed.

#### 3.5.1 Problem Reformulation

The binary constraints in (3.7h) are generally difficult to handle. To address this issue, we introduce the auxiliary variables  $\tilde{l}_{c,k}$  such that,

$$\tilde{l}_{c,k} = \iota_{c,k}, \quad \forall k \in \mathcal{K}, c \in \mathcal{C}. \quad (3.33)$$

Hence, we can equivalently replace the binary constraints in (3.7h) by,

$$\iota_{c,k}(1 - \tilde{l}_{c,k}) = 0, \quad \forall k \in \mathcal{K}, c \in \mathcal{C}. \quad (3.34)$$

It can be seen that the equality constraints hold only when  $\tilde{l}_{c,k} = \iota_{c,k} \in \{0, 1\}$ . Moreover, the introduction of equality constraints (3.33)-(3.34) does not change the actual feasible set

of the solution.

Let  $u_{c,m}$  and  $l_{c,m}$  denote the upper and lower bounds of  $|\mathbf{h}_m^H \mathbf{w}_c|^2$  respectively. Then, we can rewrite (3.10) and (3.11) as follows,

$$l_{c,m} \alpha_{c,k} \geq x_{k,m}^c y_{k,m}^c \quad (3.35)$$

$$\sum_{k'=k+1}^K u_{c,m} \alpha_{c,k'} + \sum_{c' \neq c} u_{c',m} + \sigma_m^2 \leq y_{k,m}^c. \quad (3.36)$$

Moreover, for any two variables  $x$  and  $y$ , we have  $2xy = (x+y)^2 - x^2 - y^2$ . Hence, (3.35) and (3.36) can be equivalently rewritten as,

$$[(x_{k,m}^c + y_{k,m}^c)^2 + l_{c,m}^2 + \alpha_{c,k}^2] - [(l_{c,m} + \alpha_{c,k})^2 + (x_{k,m}^c)^2 + (y_{k,m}^c)^2] \leq 0 \quad (3.37)$$

$$\left[ \frac{1}{2} \sum_{k'=k+1}^K (u_{c,m} + \alpha_{c,k'})^2 + \sum_{c' \neq c} u_{c',m} + \sigma_m^2 \right] - \left[ \frac{1}{2} \sum_{k'=k+1}^K (u_{c,m}^2 + \alpha_{c,k'}^2) + y_{k,m}^c \right] \leq 0 \quad (3.38)$$

which are expressed as differences of convex functions (DC).

From the above discussion, we can reformulate P1 into an equivalent problem as follows,

$$\text{P3 : } \min_{\mathcal{Z}_3} \sum_c \|\mathbf{w}_c\|_2^2 \quad (3.39a)$$

$$\text{s.t. (3.7c)-(3.7g), (3.12), (3.33), (3.34), (3.37), (3.38)} \quad (3.39b)$$

$$|\mathbf{h}_m^H \mathbf{w}_c|^2 \leq y_{c,m} \quad (3.39c)$$

$$l_{c,m} \leq |\mathbf{h}_m^H \mathbf{w}_c|^2 \quad (3.39d)$$

$$0 \leq l_{c,k} \leq 1. \quad (3.39e)$$

To ease the notation, we let  $\mathcal{Z}_3 \triangleq \{\mathbf{U}, \mathbf{L}, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \boldsymbol{\alpha}, \boldsymbol{\iota}, \tilde{\boldsymbol{\iota}}\}$  with  $\tilde{\boldsymbol{\iota}} \triangleq \{\tilde{\iota}_{c,m} | c \in \mathcal{C}, m \in \mathcal{K}\}$ . We can observe that: the objective function is a scalar continuously differentiable function, the functions appearing in the inequality constraints are all continuously differentiable, and the functions in the equality constraints are continuously differentiable. Hence, we can apply the PDD method to solve problem (3.39).

### 3.5.2 Proposed Algorithm

In this subsection, we conceive an efficient PDD-based algorithm to solve problem (3.39).

To tackle the equality constraints, we first formulate the AL of problem (3.39) as follows,

$$\text{AL} : \min_{\mathcal{Z}_3} \sum_c \|\mathbf{w}_c\|_2^2 + P_\rho \tag{3.40a}$$

$$\text{s.t. (3.7c), (3.7d), (3.12), (3.37), (3.38), (3.39c)-(3.39e)} \tag{3.40b}$$

where the penalty term  $P_\rho$  is given by (3.41), and  $\{\lambda_k\}$ ,  $\{\lambda_c^{(1)}\}$ ,  $\{\lambda_c^{(2)}\}$ ,  $\{\lambda_{c,k}^{(1)}\}$ , and  $\{\lambda_{c,k}^{(2)}\}$  denote the dual variables corresponding to equality constraints (3.7g), (3.7e), (3.7f), (3.33), and (3.34) respectively. The coefficient  $\rho$  is used to control the size of the penalty such that decreasing  $\rho$  increases the penalty.

$$P_\rho \triangleq \frac{1}{2\rho} \left[ \sum_k \left| \sum_c \iota_{c,k} - 1 + \rho\lambda_k \right|^2 + \sum_c \left( \left| \sum_k \alpha_{c,k} - 1 + \rho\lambda_c^{(1)} \right|^2 + \left| \sum_k \iota_{c,k} - q_c + \rho\lambda_c^{(2)} \right|^2 \right) + \sum_c \sum_k \left( \left| \tilde{\iota}_{c,k} - \iota_{c,k} + \rho\lambda_{c,k}^{(1)} \right|^2 + \left| \iota_{c,k}(1 - \tilde{\iota}_{c,k}) + \rho\lambda_{c,k}^{(2)} \right|^2 \right) \right] \tag{3.41}$$

Our proposed algorithm has a double-loop structure. In the inner loop, we employ the CCCP method to iteratively optimize the variables  $\mathcal{Z}$  while keeping the dual variables and

penalty parameter fixed. In the outer loop, we update either the dual variables or the penalty parameter. In the following, we first develop the CCCP method in details, then present the update of the dual variables and penalty parameter, and finally summarize the overall PDD-based algorithm.

*Inner loop:* In the AL problem, there exist non-convex constraints in DC forms. Hence, the former can be efficiently solved using the iterative CCCP. Specifically, by linearizing the non-convex part of the constraints using first-order Taylor expansion, we obtain convex sub-problems. A sub-optimal solution to the AL problem for the current outer loop can be efficiently found by iteratively solving a sequence of convex sub-problems. For example, by applying first-order Taylor expansion, (3.37), (3.38) and (3.39d) can be respectively transformed to the convex constraint expressed in (3.42), (3.43), and (3.44), where a variable with circumflex denotes the current value of the variable at the current iteration of the inner loop.

$$\begin{aligned} (x_{k,m}^c + y_{k,m}^c)^2 + l_{c,m}^2 + \alpha_{c,k}^2 + (\hat{l}_{c,m} + \hat{\alpha}_{c,k})^2 - 2(\hat{l}_{c,m} + \hat{\alpha}_{c,k})(l_{c,m} + \alpha_{c,k}) \\ + (\hat{x}_{k,m}^c)^2 - 2\hat{x}_{k,m}^c x_{k,m}^c + (\hat{y}_{k,m}^c)^2 - 2\hat{y}_{k,m}^c y_{k,m}^c \leq 0 \end{aligned} \quad (3.42)$$

$$\frac{1}{2} \sum_{k'=k+1}^K \left[ (u_{c,m} + \alpha_{c,k'})^2 + \hat{u}_{c,m}^2 - 2\hat{u}_{c,m} u_{c,m} + \hat{\alpha}_{c,k'}^2 - 2\hat{\alpha}_{c,k'} \alpha_{c,k'} \right] + \sum_{c' \neq c} u_{c',m} + \sigma_m^2 - y_{k,m}^c \leq 0 \quad (3.43)$$

$$l_{c,m} + |\mathbf{h}_m^H \hat{\mathbf{w}}_c|^2 - 2\Re\{\hat{\mathbf{w}}_c^H \mathbf{h}_m \mathbf{h}_m^H \mathbf{w}_c\} \leq 0 \quad (3.44)$$

From the discussion above, the AL problem in (3.40) can be approximated as,

$$\min_{\mathcal{Z}_3} \sum_c \|\mathbf{w}_c\|_2^2 + P_\rho \quad (3.45a)$$

$$\text{s.t. (3.7c), (3.7d), (3.12), (3.39c), (3.39e), (3.42)-(3.44)} \quad (3.45b)$$

We can solve subproblem (3.45) at each iteration of the inner loop by employing the CCCP. To this end, we divide the variables  $\mathcal{Z}$  into two blocks and update each block successively. The first block contains variables  $\tilde{\mathbf{i}}$ , which only exist in the objective function. Therefore, by fixing the variables in the second block, we can easily obtain the optimal value of  $\tilde{\mathbf{i}}$  in closed-form as follows,

$$\tilde{\iota}_{c,k} = \frac{\iota_{c,k} + \iota_{c,k}^2 + \rho\lambda_{c,k}^{(1)} + \rho\iota_{c,k}\lambda_{c,k}^{(2)}}{1 + \iota_{c,k}^2}. \quad (3.46)$$

In order to update the variables in the second block  $\bar{\mathcal{Z}} \triangleq \mathcal{Z} - \tilde{\mathbf{i}}$ , we fix  $\tilde{\mathbf{i}}$  and solve problem (3.45). Since the latter is now convex, this can be achieved by employing any general-purpose solver using interior-point methods [76].

*Outer loop:* At each iteration of the outer loop, the dual variables  $\{\lambda_k\}$ ,  $\{\lambda_c^{(1)}\}$ ,  $\{\lambda_c^{(2)}\}$ ,  $\{\lambda_{c,k}^{(1)}\}$ ,  $\{\lambda_{c,k}^{(2)}\}$ , or the penalty parameter  $\rho$  are updated as follows,

$$\lambda_k \leftarrow \lambda_k + \frac{1}{\rho}(\sum_c \iota_{c,k} - 1) \quad (3.47)$$

$$\lambda_c^{(1)} \leftarrow \lambda_c^{(1)} + \frac{1}{\rho}(\sum_k \alpha_{c,k} - 1) \quad (3.48)$$

$$\lambda_c^{(2)} \leftarrow \lambda_c^{(2)} + \frac{1}{\rho}(\sum_k \iota_{c,k} - q_c) \quad (3.49)$$

$$\lambda_{c,k}^{(1)} \leftarrow \lambda_{c,k}^{(1)} + \frac{1}{\rho}(\tilde{\iota}_{c,k} - \iota_{c,k}) \quad (3.50)$$

$$\lambda_{c,k}^{(2)} \leftarrow \lambda_{c,k}^{(2)} + \frac{1}{\rho}(\iota_{c,k}(1 - \tilde{\iota}_{c,k})) \quad (3.51)$$

$$\rho \leftarrow \tau\rho \tag{3.52}$$

where the symbol  $\leftarrow$  denotes an overwrite operation. A control parameter  $0 < \tau < 1$  is used to increase the value of the penalty term  $P_\rho$  during each outer iteration.

The overall PDD-based algorithm for the joint design is summarized in Algorithm 3.2, where  $\mathbf{e}(\mathcal{Z})$  is a vector that combines all functions in the equality constraints of problem (3.39) and  $\|\mathbf{e}(\mathcal{Z})\|_\infty$  denotes the maximum constraint violation. Moreover,  $\varpi_t(\mathcal{Z})$  denotes the value of the objective function of problem (3.45) at the  $t$ th iteration of the inner loop and  $\kappa_{t'}$  denotes the constraint violation parameter at the  $t'$ th iteration of the outer loop.

*Initialization:* Choosing a feasible point for initialization of Algorithm 2 is essential. For this purpose, we randomly initialize  $\mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\nu}$ , and  $\tilde{\mathbf{l}}$  such that (3.7d)-(3.7h), (3.33) are satisfied. Subsequently, we initialize  $\mathbf{U}$  and  $\mathbf{L}$  by setting  $u_{c,m} = l_{c,m} = |\mathbf{h}_m^H \mathbf{w}_c|^2, \forall m \in \mathcal{K}, c \in \mathcal{C}$ . We then set  $x_{k,m}^c = \gamma_{\min} \nu_{c,k} l_{c,m}$  and  $y_{k,m}^c = \sum_{k'=k+1}^K u_{c,m} \alpha_{c,k'} + \sum_{c' \neq c} u_{c',m} + \sigma_m^2, \forall k, m \in \mathcal{K}, c \in \mathcal{C}$ .

### 3.5.3 Convergence and Complexity Analysis

With a feasible initial point, the sequence generated by the PDD method is guaranteed to converge to a stationary solution of the problem. We omit the proof here for brevity and refer the interested reader to [83] and [84] for a rigorous proof of the convergence of the PDD methods. It is worth noting that there is no approximation or relaxation during the transformation from the original problem P1 to problem (3.39), hence they share the same solutions. Moreover, since the obtained stationary solution meets the necessary conditions for the optimal solution of problem (3.39), it also meets the necessary conditions for the optimal solution of problem P1.

---

**Algorithm 3.2:** The PDD-based algorithm.

---

**Initialization:** Define the tolerance of accuracy  $\epsilon_I$  and  $\epsilon_O$ . Initialize the algorithm with a feasible point  $\mathcal{Z}$ . Set the iteration numbers  $t = 0$  and  $t' = 0$ .

**Repeat:**

**Repeat:**

- Update  $\tilde{\mathbf{t}}$  based on (3.46).
- Update  $\tilde{\mathcal{Z}}$  by solving problem (3.45).
- Set  $t \leftarrow t + 1$ .

**Until:**  $\frac{|\varpi_{t+1}(\mathcal{Z}) - \varpi_t(\mathcal{Z})|}{|\varpi_t(\mathcal{Z})|} \leq \epsilon_I$  or  $t \leq N_{max}$

**if**  $\|\mathbf{e}(\mathcal{Z})\|_\infty \leq \kappa_{t'}$

- Update the dual variables based on (3.47)-(3.51).

**else**

- Update the penalty parameter based on (3.52).

**end**

Set  $t' \leftarrow t' + 1$ .

**Until:**  $\|\mathbf{e}(\mathcal{Z})\|_\infty \leq \epsilon_O$

---

For the PDD-based algorithm, the overall complexity is dominated by solving the convex problem (3.45) which can be approximated by a sequence of SOCPs via the successive approximation method. Each SOCP can then be solved via a general-purpose software, e.g., SeDuMi in CVX [78] with a complexity of  $\mathcal{O}((KCN)^{3.5})$ . The worst-case computational complexity is therefore given by,

$$C_P = \mathcal{O}(T_I T_O (KCN)^{3.5}) \quad (3.53)$$

where  $T_I$  and  $T_O$  denote the numbers of the inner and outer iterations, respectively. Owing to the replacement of the binary constraints and decomposition of the problem, the proposed PDD-based algorithm can converge much faster with lower complexity compared to the

proposed BB-based algorithm in Section 3.4.

## 3.6 Simulation Results

In this section, numerical experiments are carried out to illustrate the performance of the proposed algorithms for joint design of user clustering, downlink beamforming, and power allocation in a MIMO NOMA system.

### 3.6.1 Methodology

We consider downlink transmission in a MIMO NOMA system, wherein a BS equipped with  $N$  antennas serves  $K = 6$  users, each equipped with a single antenna. The BS is located at the center of a circular cell with radius 200m, over which the users are randomly distributed. The users are grouped into  $C$  non-overlapping clusters, each allocated a different beamforming vector. For simplicity, unless otherwise stated, we assume that all clusters comprise equal number of users, i.e.,  $q_c = q, \forall c \in \mathcal{C}$ . We consider the channel model (3.1) described in Section 3.2.1 with bandwidth of  $W = 2$  GHz and carrier frequency of  $f_c = 28$  GHz. The AoDs of the different channel paths follow a uniform distribution in  $[0, 2\pi]$ . Throughout the experiments, it is assumed that the noise variance is the same for all users, i.e.,  $\sigma_k^2 = \sigma^2, \forall k \in \mathcal{K}$ . The pathloss exponents of the LOS and NLOS paths are  $\alpha_{\text{LOS}} = 2$  and  $\alpha_{\text{NLOS}} = 3$ , respectively.

For the PDD-based algorithm, we set the initial penalty parameter and control parameter as  $\rho = 4$  and  $\tau = 0.8$ , respectively and the equality constraint violation tolerance parameter as  $\kappa_j = 0.3^j$  at the  $j$ th outer iteration. Table 3.1 summarizes the key simulation parameters.

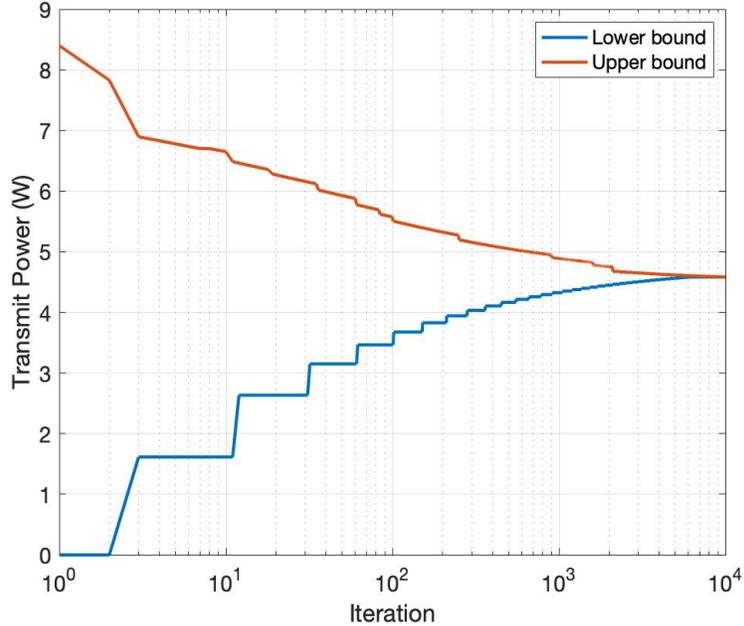
Description	Varibales	Value
Number of users	$K$	6
Number of BS antennas	$N$	2 or 8
Noise power (mW)	$\sigma^2$	10
Bandwidth (GHz)	$W$	2
Carrier frequency (GHz)	$f_c$	28
Path loss exponent of LOS	$\alpha_{\text{LOS}}$	2
Path loss exponent of NLOS	$\alpha_{\text{NLOS}}$	3
Number of the NLOS paths	$L$	0 or 3
Tolerance for the algorithms	$\epsilon, \epsilon_I, \epsilon_O$	$10^{-6}$
Control parameter	$\tau$	0.8
Initial penalty parameter	$\rho$	4
Maximum number of inner iterations	$N_{max}$	10
Maximum available transmit power (W)	$P_{max}$	50

**Table 3.1:** Summary of parameters.

We use Monte Carlo experiments to evaluate the performance of the proposed BB and PDD-based algorithms for joint user clustering, downlink beamforming, and power allocation. The total transmit power and sum rate are measured for different parameter configurations and the results are compared with benchmark approaches from the literature, as will be specified below.

### 3.6.2 Results and Discussion

Fig. 3.4 illustrates the convergence behavior of the proposed BB-based algorithm for a particular scenario where  $K = 6$  users are grouped into  $C = 3$  clusters with  $q = 2$  users in each. The target SINR and number of NLOS paths are set to  $\gamma_{min} = 0.1$  and  $L = 0$ , respectively. We consider  $N = 2$  transmit antennas for the BS to limit the computational



**Figure 3.4:** Convergence behavior of the BB-based algorithm.

complexity of the algorithm<sup>3</sup>. It can be seen that the upper bound and the lower bound are non-increasing and non-decreasing, respectively, while the gap between them becomes smaller as the iteration index increases and infeasible subregions are removed. The algorithm performs a branching process for each dimension to approach the  $\epsilon$ -optimal solution. Hence, the number of iterations can be very large if the tolerance  $\epsilon$  is small, which is not practical due to the high complexity. However, the achieved results can still be used as the network performance benchmark.

Figs. 3.5a and 3.5b present the convergence behavior of the proposed PDD-based algorithm with three different initial points. For comparison purposes, the same parameter configuration as in Fig. 3.4 is considered. Fig. 3.5a shows the maximum constraint

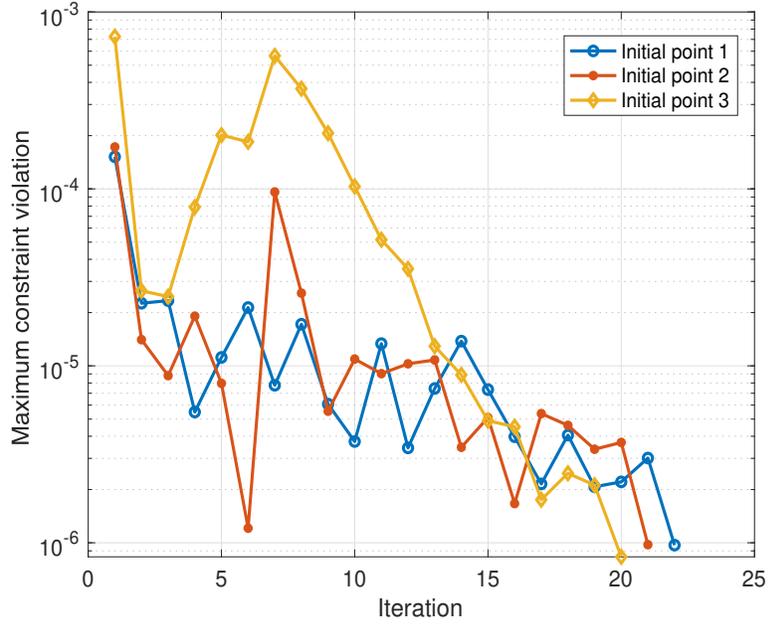
<sup>3</sup>With  $N = 2$  antennas and  $C = 3$  clusters, the beamforming gain will be limited since there is not enough degrees of freedom to design directional beams in the clusters' direction; that is, the design problem is overdetermined. We refer the interested reader to [85,86] for discussion on classical beamformer design.

violation, i.e.,  $\|\mathbf{e}(\mathcal{Z})\|_\infty$ , versus the number of iterations. We observe that the constraint violation decrease below  $10^{-6}$  within 20 iterations. Fig. 3.5b shows the objective value of (3.45), i.e.,  $\varpi_t(\mathcal{Z})$ , achieved by the PDD-based algorithm at the end of each outer iteration. It can be seen that the objective value converges in less than 15 iterations to the same optimum value obtained by the BB-based algorithm for all initial points.

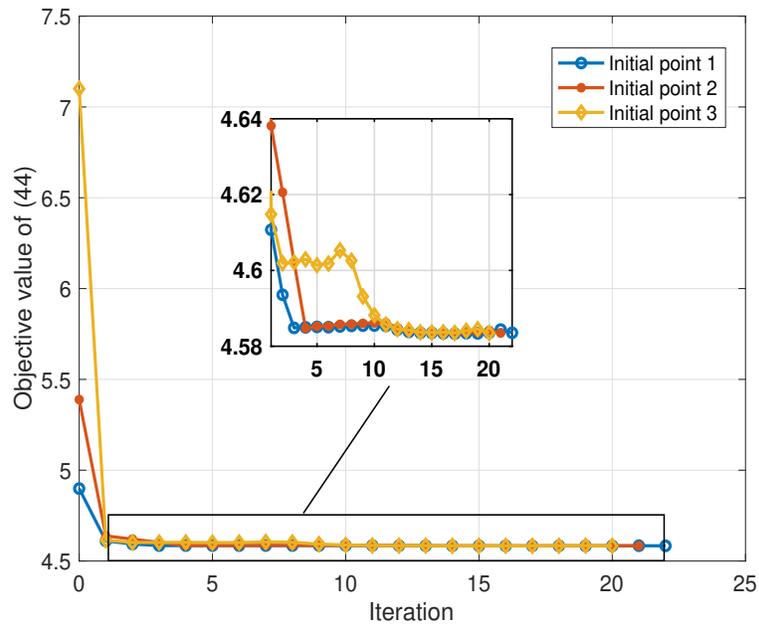
Fig. 3.6 compares the achievable sum rate versus total transmit power among the proposed BB and PDD-based algorithms, the AO [37], the DO [30] and the orthogonal multiple access (OMA) method. In this scenario, all  $K = 6$  users are grouped into  $C = 1$  cluster and the BS is equipped with  $N = 2$  antennas. As shown in the figure, the average achievable rate increases as the total transmit power increases. We can also see that the proposed BB and PDD-based algorithms outperform the AO, DO and OMA methods. For the BB algorithm, this performance improvement is not surprising since it finds the global optimum solution for the power allocation and beamforming, while for the PDD algorithms, the results suggest that it achieves near optimal performance. We note that the sum rate for BB slightly exceeds that for PDD when the transmit power is small; however, the performance gap between PDD and BB goes to zero as the transmit power increases. While it is not taken into account in this figure, user clustering can further improve system performance, will be illustrated shortly.

Fig. 3.7 shows the total transmit power versus target SINR,  $\gamma_{min}$ , for the case with  $C = 3$  clusters and  $q = 2$  users per cluster. The BS is equipped with  $N = 2$  antennas and  $L = 0$  NLOS paths are considered for the mmWave-based channel vectors. Different clustering, beamforming, and power allocation schemes are considered for comparison as described below:

- RCL: Random clustering scheme.

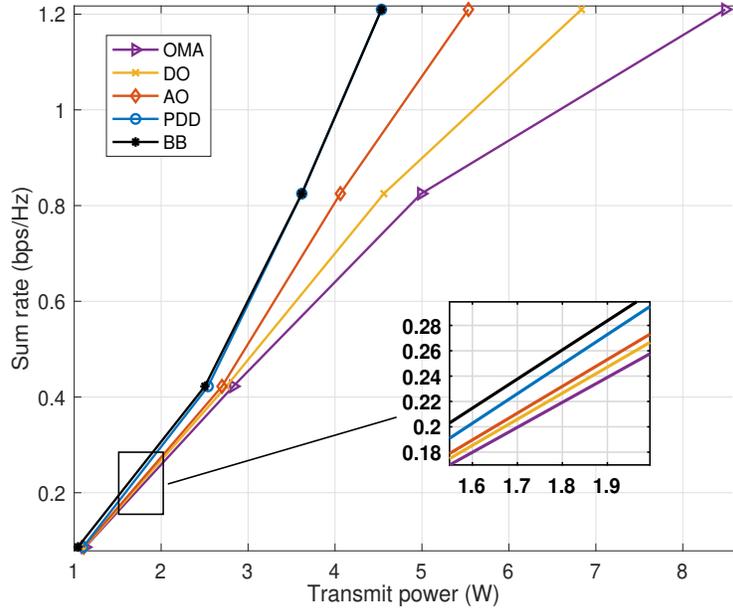


(a) Maximum constraint violation,  $\|e(\mathcal{Z})\|_\infty$ .



(b) Objective value of (44),  $\varpi_t(\mathcal{Z})$ .

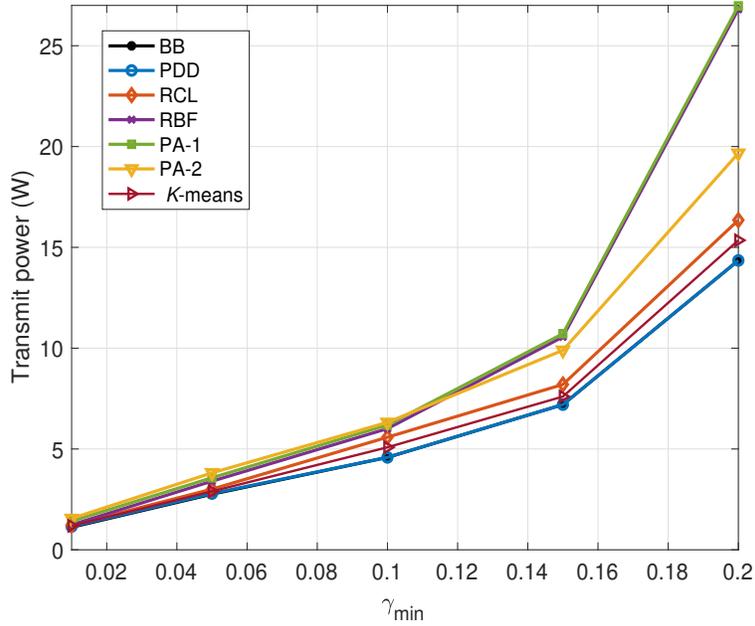
Figure 3.5: Convergence behavior of the PDD-based algorithm.



**Figure 3.6:** Achievable sum rate versus total transmit power.

- RBF: Random beamforming scheme [87].
- PA-1: Fixed power allocation with equal power for the users in one cluster, i.e.,  $\alpha_{c,k} = \alpha_{c,m} = 0.5$  when the  $k$ th and  $m$ th users are both in the  $c$ th cluster.
- PA-2: Fixed power allocation with unequal power, i.e.,  $\alpha_{c,k} = 0.8$  and  $\alpha_{c,m} = 0.2$  when the  $k$ th and  $m$ th users are both in the  $c$ th cluster and  $k < m$ .
- $K$ -means: Machine-learning based user clustering algorithm constrained to limit the number of users in each cluster [92].

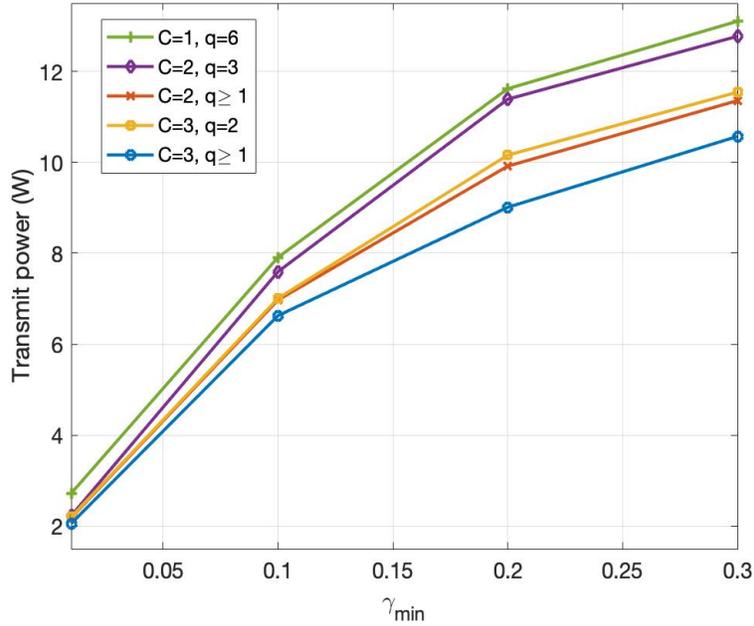
We observe that the proposed BB-based and PDD-based algorithms outperform other schemes and and that once again, the latter nearly achieve the optimal performance of the former. For RBF and PA-1, the total transmit power for small  $\gamma_{min}$  is less than PA-2.



**Figure 3.7:** Total transmit power versus target SINR.

However, as  $\gamma_{min}$  increases, a noticeable increase in total transmit power is observed for RBF and PA-1, while the unbalanced power allocation, i.e., PA-2 obtains lower transmit power.

To better appreciate the benefits of user clustering, we examine the total transmit power versus target SINR,  $\gamma_{min}$ , for different clustering schemes in Fig. 3.8. In this scenario,  $K = 6$  users are grouped into  $C = 1, 2,$  and  $3$  clusters with  $6, 3, 2$  users in each, respectively. We also consider grouping into  $C = 2$  and  $3$  clusters, but with no limit on the number of users per cluster, as indicated by, e.g. “ $C = 2, q \geq 1$ ”. The BS is equipped with  $N = 8$  antennas and  $L = 3$  NLOS paths are considered for the channel vectors. We observe that when all users are grouped into one cluster, the total transmit power is greater than for other cases. It can also be seen that the results for the unconstrained cases outperform other cases with fixed  $q$  and the performance gap increasing with  $\gamma_{min}$ . This means that the algorithm

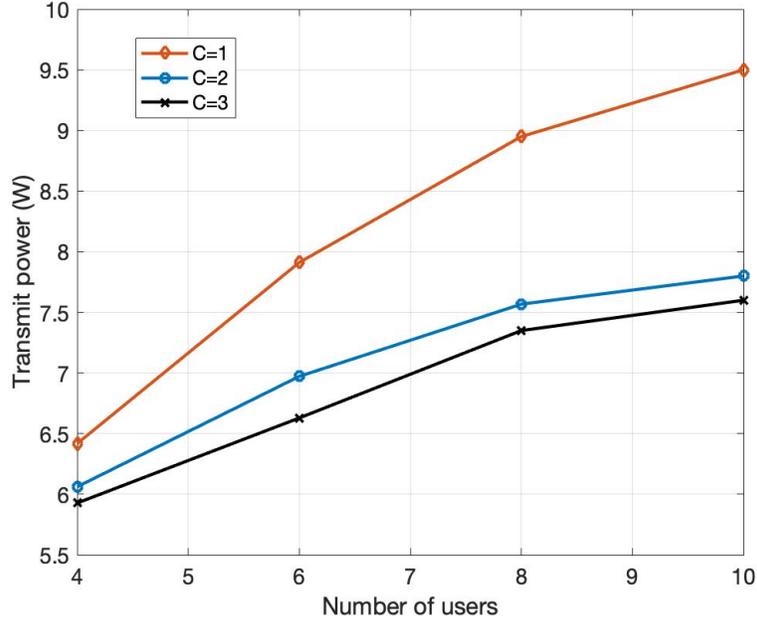


**Figure 3.8:** Total transmit power versus target SINR.

can find better user clustering with lower total transmit power when there is no constraint on the number users per cluster. Finally, by increasing the number of antennas, the total transmission power can be reduced, which is a consequence of narrower beamforming.

Fig. 3.9 shows the total transmit power versus number of users  $K$  with target SINR  $\gamma_{\min} = 0.1$ . Here,, users are grouped into  $C = 1, 2,$  and  $3$  clusters while no limit on the number of users per cluster is enforced. Moreover, the BS is equipped with  $N = 8$  antennas and  $L = 3$  NLOS paths are considered for the channel vectors. Obviously, the total transmit power increases with the number of users. When all users are grouped into one cluster, a rapid increase in the total transmit power as a function of  $K$  is observed. It can also be seen that by increasing the number of clusters transmit power can be significantly reduced. Hence, user clustering has a great impact on the performance of the system.

To investigate the impact of imperfect CSI on the proposed joint design of user

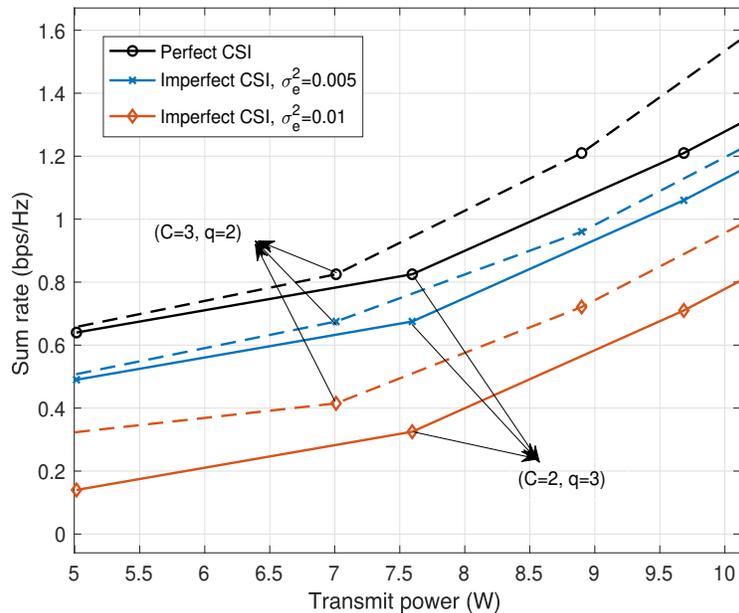


**Figure 3.9:** Total transmit power versus number of users.

clustering, power allocation and downlink beamforming, we implement the proposed PDD-based algorithm with an estimated channel vector model as follows

$$\hat{\mathbf{h}}_k = \mathbf{h}_k + \Delta_k \quad (3.54)$$

where  $\mathbf{h}_k$  is the actual channel vector and  $\Delta_k$  is the CSI error, which follows a complex Gaussian distribution with i.i.d. entries, i.e.,  $\Delta_k \sim \mathcal{CN}(\mathbf{0}, \sigma_e^2 \mathbf{I})$ . Fig. 3.10 shows sum rate versus total transmit power, where the perfect and imperfect CSI with different value of  $\sigma_e$  are considered. Here, two scenarios for user clustering is considered, i.e.,  $C = 2, q = 3$  shown by solid lines and  $C = 3, q = 2$  shown by dashed lines. We observe that the system performance is sensitive to the CSI accuracy. This is due to the fact that the proposed PDD-based algorithm largely depends on the obtained CSI at the transmitter. Moreover,



**Figure 3.10:** Sum rate versus total transmit power for perfect and imperfect CSI.

SIC does not work well in the presence of imperfect CSI and can not remove intra-cluster interference totally. It can also be seen that the performance degradation is worse when the number of users in a cluster increases. Once an error occurs in SIC (due to e.g., imperfect CSI), signal of the corresponding user will not be completely removed, leaving some residual signals as interference. Consequently, the message of all remaining users in the corresponding cluster will likely be decoded erroneously. Although robust beamforming can potentially compensate the impact of imperfect CSI and error propagation [88–90], this approach can add more complexity to the existing problem. This aspect falls beyond the scope of this contribution, and nevertheless remains an interesting avenue for future work.

## 3.7 Concluding Remarks

In this work, the joint optimization framework for user clustering, power allocation and beamforming was investigated in a multi-user MIMO NOMA system operating at mmWave frequencies. In the proposed scheme, users are partitioned into non-overlapping clusters. In any given clusters, users share a common beamforming vector but are distinguished with different power allocation. The joint design of user clustering, power allocation, and beamforming was formulated as an optimization problem, with the aim to minimize the total transmit power under the SINR, power and clustering constraints. We first developed a BB-based algorithm to find the global optimum of the problem. We then proposed a low-complexity algorithm using PDD method to obtain the suboptimal solution. Through simulations, it was shown that applying the proposed design to the multi-user MIMO NOMA system can effectively decrease total transmit power and improve spectral efficiency compared to the benchmark approaches.

## Chapter 4

# User Clustering and Beamforming for MIMO SCMA in C-RAN

In this chapter, we investigate the key problems of user clustering and downlink beamforming for MIMO SCMA in a C-RAN. The proposed approaches are evaluated by means of simulations over mmWave channels. Results show significant improvements in terms of total transmit power and spectral efficiency compared to benchmark approaches.

### 4.1 Introduction

MIMO SCMA combines MIMO techniques, which increase capacity by transmitting different signals over multiple antennas, and SCMA which improves spectral efficiency and device connectivity by transmitting multiple user signals over the same radio resources. As seen in works related to power domain NOMA [93,94], the joint application of spatial user clustering

---

Parts of the material in this chapter have been presented at the 2021 IEEE 32nd Annual International Symposium on PIMRC [91], and published in the IEEE Access [92].

along with beamforming techniques in MIMO SCMA systems has the potential to improve spectral efficiency and reduce the total transmit power. Additionally, when considered within a C-RAN architecture, this approach makes it possible to increase the number of supported users in the network by using a common codebook for users in different clusters, while the effect of inter-cluster interference can be eliminated by centralized beamformer design and coordinated RRH operation. In spite of its importance, the joint problem of user clustering and beamforming has not received considerable attention in the literature on MIMO SCMA, let alone C-RAN.

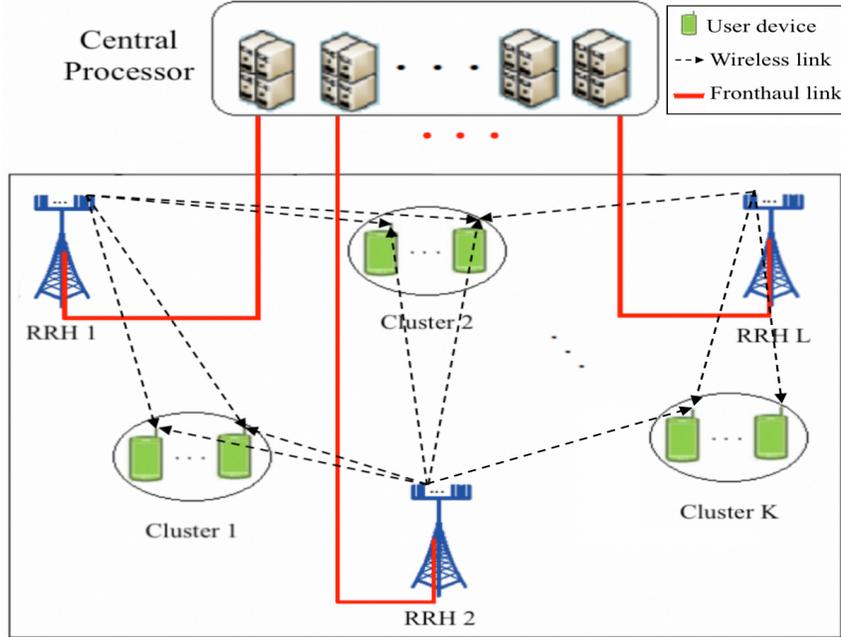
Motivated by the above considerations, we propose energy-efficient user clustering and downlink beamforming approaches for MIMO SCMA in C-RAN. Our main contributions in addressing the above challenges are summarized as follows:

- We approach the user clustering problem by modifying the widely-used  $K$ -means method from the field of machine learning, in order to limit the number of users in each cluster. Specifically, the proposed *constrained*  $K$ -means algorithm uses the Euclidian metric to characterize the similarities between the user channel vectors and the cluster centers, and seeks to group users with channel vectors exhibiting large correlation. The elbow method is utilized to find the optimum number of clusters for the network.
- We formulate the beamforming design and RRH selection as a non-convex mixed-integer nonlinear programming (MINLP) optimization problem, aiming to minimize the total transmit power while satisfying the signal-to-interference-plus-noise ratio (SINR) and fronthaul capacity constraints. We then propose transformations to reformulate the problem as a difference of convex functions (DC) program and derive two algorithms for solving the problem. In the

first algorithm, we iteratively approximate the continuous non-convex constraints by convex ones using first-order Taylor expansion and solve a sequence of mixed-integer second-order cone programming (MI-SOCP) using dedicated solvers. This algorithm entails high computational complexity, yet it can achieve high quality solution.

- The second algorithm is based on a two-stage low-complexity beamforming approach wherein the beamforming matrices obtained from each stage are multiplied to form the final beamformer. In the first stage, specifically, a block diagonalization (BD) technique is adopted to design the cluster beamformers (one for each cluster), which remove the inter-cluster interference and thus enhance the QoS for intra-cluster users. In the second stage, the user-specific beamformers are designed along with RRH selection by employing a smoothed  $\ell_0$ -norm approximation. The resulting optimization problem is solved via the convex-concave procedure (CCCP) with guaranteed convergence [95].
- We evaluate the performance of the proposed algorithms for user clustering and downlink beamforming using in-depth simulations of MIMO SCMA in C-RAN with mmWave channel models and different parameter configurations. The results illustrate the convergence behavior of the new algorithms and the effect of various parameters on the system performance, while providing useful insights into the advantages of the proposed approaches over competing ones from the literature.

The rest of the chapter is organized as follows: Section 4.2 introduces the MIMO SCMA system model under C-RAN and describes the problem under consideration. The proposed constrained  $K$ -means algorithm for user clustering is introduced in Section 4.3. The two-stage energy-efficient beamforming approach for eliminating inter-cluster interference and minimizing total transmit power is developed In Section 4.4. The results of our simulation



**Figure 4.1:** The MIMO SCMA system model under C-RAN.

experiments are presented in Section 4.5, followed by the concluding remarks in Section 4.6.

## 4.2 System Model and Problem Description

We consider downlink transmission in a MIMO SCMA system under C-RAN, as illustrated in Fig. 4.1. The system consists of  $L$  RRHs, each equipped with  $M$  antennas, and  $J$  single-antenna users. The RRHs indexed by  $l \in \mathcal{L} \triangleq \{1, \dots, L\}$ , are connected to the central processor via limited-capacity fronthaul links. Due to the fronthaul constraint, each user is cooperatively served by a specific subset of RRHs through joint beamforming. Moreover, the users are partitioned into  $K$  non-overlapping clusters, indexed by  $k \in \mathcal{K} \triangleq \{1, \dots, K\}$  with the  $k$ th cluster comprising  $J_k$  users such that  $J = \sum_{k=1}^K J_k$ . Below, we provide further details on the SCMA encoder, mmWave channel, received signal model, and problem description.

### 4.2.1 SCMA Encoder

In SCMA, contiguous groups of data bits from each user are directly mapped to sparse  $N$ -dimensional codewords selected from a predefined codebook and then transmitted over  $N$  radio resources, e.g., orthogonal frequency division multiple access (OFDMA) subcarriers. The SCMA encoder for the  $i$ th user can be defined as  $f_i : \mathbb{B}^U \rightarrow \mathcal{X}_i$  which is a one-to-one mapping from the set of  $U$ -bit tuples to a codebook  $\mathcal{X}_i \subset \mathbb{C}^N$  of  $N$ -dimensional codewords, with cardinality  $|\mathcal{X}_i| = 2^U$ . Specifically, for  $\mathbf{b} = [b_1, \dots, b_U] \in \mathbb{B}^U$ , the corresponding codeword is obtained as,

$$\mathbf{x} = f_i(\mathbf{b}) = [x(1), \dots, x(N)] \quad (4.1)$$

where  $\mathbf{x}$  is a sparse vector with  $C < N$  non-zero elements.

Each user is assigned  $C$  subcarriers such that no two users occupy the same set of subcarriers. Hence, only  $q$  users can be supported by SCMA, as given by [14],

$$q = \binom{N}{C} = \frac{N!}{C!(N-C)!}. \quad (4.2)$$

In this work, we group users into  $K$  clusters of size  $J_k \leq q$  and remove inter-cluster interference so that the users in different clusters can use common codebooks.

Referring to (4.1), we can associate to each codeword  $\mathbf{x}$  a vector  $\mathbf{y}$  containing its  $C$  non-zero elements in the same order, i.e.,  $\mathbf{y}$  is obtained from  $\mathbf{x}$  by removing its zero elements. For convenience, we represent this operation by the function  $\phi : \mathbb{C}^N \rightarrow \mathbb{C}^C$ , so that  $\mathbf{y} = \phi(\mathbf{x}) = [y(1), \dots, y(C)]$ . Through this operation, the original codebook  $\mathcal{X}_i \subset \mathbb{C}^N$  is transformed into a constellation of  $C$ -dimensional codewords, i.e.,  $\mathcal{Y}_i \subset \mathbb{C}^C$ , where  $\mathcal{Y}_i = \{\phi(\mathbf{x}) : \mathbf{x} \in \mathcal{X}_i\}$ . We also let  $g_i = \phi \circ f_i : \mathbb{B}^U \rightarrow \mathcal{Y}_i$  denote the composite mapping of  $f_i$  and  $\phi$ , so that for any

$\mathbf{b} \in \mathbb{B}^U$ , and  $\mathbf{x} = f_i(\mathbf{b})$ , we have,

$$\mathbf{y} = \phi(\mathbf{x}) = g_i(\mathbf{b}). \quad (4.3)$$

From this perspective, the SCMA encoder can be redefined as  $f_i(\mathbf{b}) = \mathbf{S}_i g_i(\mathbf{b})$ , where matrix  $\mathbf{S}_i \in \mathbb{B}^{N \times C}$  maps a  $C$ -dimensional constellation point to an  $N$ -dimensional codeword. Note that  $\mathbf{S}_i$  contains  $N - C$  all-zero rows and hence, all the codewords in codebook  $\mathcal{X}_i$  contain 0 in the same  $N - C$  positions. Moreover, an identity matrix of order  $C$  is obtained by removing the all-zero rows from  $\mathbf{S}_i$ .

The set of resources occupied by user  $i$  is determined by the positions (or indices) of the non-zero elements of the binary indicator vector  $\mathbf{f}_i = \text{diag}(\mathbf{S}_i \mathbf{S}_i^T) \in \mathbb{B}^{N \times 1}$ . In effect, the complete SCMA encoder structure for  $q$  users and  $N$  subcarriers can be represented by a factor graph, with associated matrix  $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_q] \in \mathbb{B}^{N \times q}$ . In this interpretation, subcarrier node  $n$  and user node  $i$  are connected if and only if the corresponding element of matrix  $\mathbf{F}$  is equal to 1, i.e.,  $[\mathbf{F}]_{n,i} = 1$ .

## 4.2.2 Channel Model

Due to the propagation characteristics at such high frequencies, the application of MIMO SCMA communication in the mmWave band is more challenging than in a conventional low-frequency scenario. The mmWave-based channel vector  $\mathbf{h}_{jk}^{(l)}(n) \in \mathbb{C}^{1 \times M}$  from the  $l$ th RRH to the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier can be expressed as the discrete sum of a line-of-sight (LOS) and  $P$  non line-of-sight (NLOS) components [87, 96], i.e.,

$$\mathbf{h}_{jk}^{(l)}(n) = \sum_{p=0}^P \frac{\sqrt{M} a_{jk}^{(lp)}(n) \mathbf{a}(\theta_{jk}^{(lp)})}{\sqrt{P+1}(1 + (d_{jk}^{(l)})^{\alpha^{(p)}})} \quad (4.4)$$

where:  $p$  is the path index, with  $p = 0$  corresponding to LOS and  $p \geq 1$  to NLOS paths;  $d_{jk}^{(l)}$  is the distance between the RRH and the user;  $\alpha^{(p)}$  is the path loss exponent;  $a_{jk}^{(lp)}(n)$  denotes the complex gain for the  $p$ th path which follows a complex circular Gaussian distribution, i.e.,  $a_{jk}^{(lp)}(n) \sim \mathcal{CN}(0, 1)$ ; and  $\mathbf{a}(\theta_{jk}^{(lp)}) \in \mathbb{C}^{1 \times M}$  is the antenna array steering vector. In the case of a uniform linear antenna array, the steering vector is given by,

$$\mathbf{a}(\theta_{jk}^{(lp)}) = \frac{1}{\sqrt{M}} [1, e^{-j\pi\theta_{jk}^{(lp)}}, \dots, e^{-j\pi(M-1)\theta_{jk}^{(lp)}}] \quad (4.5)$$

where  $\theta_{jk}^{(lp)}$  is the normalized direction of the  $p$ th path. The latter can be expressed as,

$$\theta_{jk}^{(lp)} = \frac{2d}{\lambda} \sin(\phi_{jk}^{(lp)}) \quad (4.6)$$

where  $\phi_{jk}^{(lp)} \in [0, 2\pi]$  is the angle of departure (AoD) of the  $p$ th path,  $d$  is the inter-antenna element spacing, and  $\lambda$  is the wavelength at the operating frequency.

In MIMO systems operating at mmWave frequencies, a single-path model is often adopted for the channel vectors by retaining only one dominant path in (4.4) [97]. In most cases, the latter will be the LOS path, whose gain can be as much as 20dB stronger than that of NLOS paths [98]. However, when there is no LOS path due to blockage, the dominant NLOS path can be considered instead. Hence, the mmWave channel model can be simplified to,

$$\mathbf{h}_{jk}^{(l)}(n) = \frac{\sqrt{M}a_{jk}^{(l)}(n)}{(1 + (d_{jk}^{(l)})^\alpha)} \mathbf{a}(\theta_{jk}^{(l)}) \quad (4.7)$$

where, for simplicity of notation, the superscript  $p$  for the path index has been removed.

### 4.2.3 Signal Model

Let  $x_{jk}(n) \in \mathbb{C}$  denote the codeword element intended for the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier. Due to the sparsity of the SCMA encoder,  $x_{jk}(n)$  can be either 0, or a non-zero element with normalized power, i.e.,  $E\{|x_{jk}(n)|^2\} = 1$ . Codeword element  $x_{jk}(n)$  is transmitted from the  $M$  antennas of the  $l$ th RRH by employing the beamforming vector  $\mathbf{w}_{jk}^{(l)}(n) \in \mathbb{C}^{M \times 1}$ . Hence, the transmit signal of the  $l$ th RRH over the  $n$ th subcarrier can be expressed as,

$$\mathbf{z}^{(l)}(n) = \sum_{k=1}^K \sum_{j \in \mathcal{U}_{n,k}} \mathbf{w}_{jk}^{(l)}(n) x_{jk}(n) \quad (4.8)$$

where  $\mathcal{U}_{n,k}$  denotes the set of users in the  $k$ th cluster occupying the  $n$ th subcarrier. Owing to the limited-capacity fronthaul link, only a selected group of RRHs serve a specific user cooperatively. The process of RRH selection for transmission can be performed through beamforming. That is,  $\|\mathbf{w}_{jk}^{(l)}(n)\|_2 = 0$  implies that the  $l$ th RRH does not participate in the transmission for that user over its assigned subcarrier. Hence, the corresponding network-wide beamforming vector,  $\mathbf{w}_{jk}(n) = [\mathbf{w}_{jk}^{(1)}(n)^T, \dots, \mathbf{w}_{jk}^{(L)}(n)^T]^T \in \mathbb{C}^{LM \times 1}$  may be sparse.

Let  $\mathbf{h}_{jk}(n) = [\mathbf{h}_{jk}^{(1)}(n), \dots, \mathbf{h}_{jk}^{(L)}(n)] \in \mathbb{C}^{1 \times LM}$  denote the network-wide channel vector for the  $j$ th user in the  $k$ th cluster and  $\mathbf{z}(n) = [\mathbf{z}^{(1)}(n)^T, \dots, \mathbf{z}^{(L)}(n)^T]^T \in \mathbb{C}^{LM \times 1}$  denote the network-wide transmit signal over the  $n$ th subcarrier. The received signal at the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier is given by,

$$r_{jk}(n) = \mathbf{h}_{jk}(n)\mathbf{z}(n) + n_{jk} \quad (4.9)$$

where  $n_{jk} \sim \mathcal{CN}(0, \sigma_{jk}^2)$  is an additive noise term. We can express the received signal of

this user as a sum of the desired signal, the interference from the other users in that cluster (intra-cluster interference), the inter-cluster interference and the noise, i.e.,

$$\begin{aligned}
r_{jk}(n) = & \mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)x_{jk}(n) + \underbrace{\sum_{j' \neq j, j' \in \mathcal{U}_{n,k}} \mathbf{h}_{jk}(n)\mathbf{w}_{j'k}(n)x_{j'k}(n)}_{\text{Intra-cluster Interference}} \\
& + \underbrace{\sum_{k' \neq k} \sum_{j \in \mathcal{U}_{n,k'}} \mathbf{h}_{jk}(n)\mathbf{w}_{jk'}(n)x_{jk'}(n)}_{\text{Inter-cluster Interference}} + n_{jk}.
\end{aligned} \tag{4.10}$$

The SINR of the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier with non-zero codeword element is given by,

$$\gamma_{jk}(n) = \frac{|\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)|^2}{I_{jk}^{(1)}(n) + I_{jk}^{(2)}(n) + \sigma_{jk}^2} \tag{4.11}$$

where the first term in the denominator represents the intra-cluster interference and the second term represents the inter-cluster interference, i.e.,

$$I_{jk}^{(1)}(n) = \sum_{j' \neq j, j' \in \mathcal{U}_{n,k}} |\mathbf{h}_{jk}(n)\mathbf{w}_{j'k}(n)|^2 \tag{4.12}$$

$$I_{jk}^{(2)}(n) = \sum_{k' \neq k} \sum_{j \in \mathcal{U}_{n,k'}} |\mathbf{h}_{jk}(n)\mathbf{w}_{jk'}(n)|^2. \tag{4.13}$$

The total transmit power for the whole network over  $N$  subcarriers is given by,

$$P_T = \sum_{n=1}^N E\{\mathbf{z}(n)^H \mathbf{z}(n)\} = \sum_{n=1}^N \sum_{l=1}^L E\{\mathbf{z}^{(L)}(n)^H \mathbf{z}^{(L)}(n)\}. \tag{4.14}$$

Upon substitution of (4.8) into (4.14) and assuming that the transmitted codewords  $x_{jk}(n)$  from different sources are uncorrelated and have zero mean and unit variance, we can write

the total transmit power as,

$$P_T = \sum_n \sum_l \sum_k \sum_j \|\mathbf{w}_{jk}^{(l)}\|^2 = \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}\|^2, \quad (4.15)$$

where the last equality follows from the definition of the network-wide beamforming vector.

#### 4.2.4 Problem Description

In this work, our objective is to group users into non-overlapping clusters and design beamformers such that the total transmit power is minimized while constraining the inter-cluster interference, the user SINRs and the fronthaul capacity. Indeed, removing inter-cluster interference not only enhances the SINR at the user terminal, but also allows the transmitter to use a common SCMA codebook to serve users in different clusters, which in turn boosts network capacity. To further satisfy the requirements imposed by the limited-capacity fronthaul links of C-RAN, dynamic RRH selection is taken into consideration in our formulation.

In order to address the above challenges and obtain the desire solution, we conceive efficient algorithms for user clustering and beamforming design with low complexity. Specifically, we propose an efficient user clustering algorithm based on the constrained  $K$ -means method in Section 4.3. Then, the beamformer design is addressed in Section 4.4 by means of a two-stage energy-efficient approach wherein the inter-cluster interference is removed using a BD technique in the first stage and the total transmit power is optimized under SINR and fronthaul capacity constraints in the second stage.

## 4.3 User Clustering

In this section, we first introduce the proposed constrained  $K$ -means algorithm for user clustering. We then apply the elbow method to determine the number of clusters. Finally, we evaluate the computational complexity of the proposed algorithm.

### 4.3.1 Constrained $K$ -means Clustering

$K$ -means is a celebrated method for grouping inharmonious multi-dimensional data points into  $K$  clusters such that a similarity criterion within clusters is maximized [99, 100]. In effect,  $K$ -means attempts to group  $J$  data points (or vectors)  $\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_J\}$  into  $K$  clusters by finding cluster centers  $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_K\}$  such that similarities between the points in the same group are high while similarities between the points in different groups are low. Two key factors in the  $K$ -means method are the number of clusters  $K$ , which is pre-determined, and the similarity metric [101].

In the current MIMO SCMA application, high correlation between the channel vectors of the users in a cluster can provide a better beamforming performance. Indeed, if users in a cluster have highly correlated channels, more degrees of freedom will be left for the inter-cluster interference cancellation (as explained in Section IV). In this work, we utilize the Euclidian distance as a similarity metric to measure the correlation between a user's channel vector and the cluster centers. Moreover, to account for variations of channel gains due to fading and other propagation effects, the channel vectors are normalized, averaged over subcarriers, and treated as the data points in the application of the  $K$ -means method,

i.e.,

$$\mathbf{d}_j = \frac{1}{N} \sum_{n=1}^N \frac{\mathbf{h}_j(n)}{\|\mathbf{h}_j(n)\|_2} \quad (4.16)$$

where  $\mathbf{h}_j(n) \in \mathbb{C}^{1 \times LM}$  for  $j \in \mathcal{J} \triangleq \{1, \dots, J\}$  are the known network-wide channel vectors of all users prior to clustering.

The  $K$ -means method can be presented as an optimization problem for finding the  $K$  best centers such that the sum of squared Euclidean (SSE) distance between the data points and their nearest cluster centers is minimized. Specifically, this optimization problem can be expressed as follows,

$$\min_{\mathbf{C}} \sum_{j=1}^J \min_{k \in \mathcal{K}} \|\mathbf{d}_j - \mathbf{c}_k\|_2^2 \quad (4.17)$$

where  $\mathbf{C} \triangleq \{\mathbf{c}_k | k \in \mathcal{K}\}$ .

**Proposition 4.1** Given  $\mathbf{d}_j$  and  $\mathbf{c}_k \in \mathbb{C}^{1 \times LM}$ , we have,

$$\min_{k \in \mathcal{K}} \|\mathbf{d}_j - \mathbf{c}_k\|_2^2 = \min_{\{\iota_{j,k} | k \in \mathcal{K}\}} \sum_{k=1}^K \iota_{j,k} \|\mathbf{d}_j - \mathbf{c}_k\|_2^2 \quad (4.18a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \iota_{j,k} = 1 \quad (4.18b)$$

$$\iota_{j,k} \geq 0, \quad \forall k \in \mathcal{K}. \quad (4.18c)$$

*Proof.* The result follows directly from the linear programming duality theory [102].  $\square$

By introducing selection variables  $\boldsymbol{\iota} \triangleq \{\iota_{j,k} | j \in \mathcal{J}, k \in \mathcal{K}\}$  and using Proposition 4.1, we

can reformulate problem (4.15) as the following problem,

$$\min_{\boldsymbol{\iota}, \mathbf{C}} \sum_{j=1}^J \sum_{k=1}^K \iota_{j,k} \|\mathbf{d}_j - \mathbf{c}_k\|_2^2 \quad (4.19a)$$

$$\text{s.t.} \quad \sum_k \iota_{j,k} = 1, \quad \forall j \in \mathcal{J} \quad (4.19b)$$

$$\iota_{j,k} \geq 0, \quad \forall j \in \mathcal{J}, k \in \mathcal{K} \quad (4.19c)$$

where  $\iota_{j,k} = 1$  if the  $j$ th data point is closest to the  $k$ th cluster center, i.e., belongs to the  $k$ th cluster, and  $\iota_{j,k} = 0$  otherwise.

While the  $K$ -means method does not involve *a priori* constraint on the number of users in each cluster [103], the SCMA encoder in the current application can support at most  $q$  users over  $N$  subcarriers. To avoid solutions with more than  $q$  data points in a cluster, we propose adding explicit constraints to problem (4.19) so that each cluster contains at most  $q$  data points, i.e.,

$$\min_{\boldsymbol{\iota}, \mathbf{C}} \sum_j \sum_k \iota_{j,k} \|\mathbf{d}_j - \mathbf{c}_k\|_2^2 \quad (4.20a)$$

$$\text{s.t.} \quad (4.19b), (4.19c) \quad (4.20b)$$

$$\sum_i \iota_{i,k} \leq q, \quad \forall k \in \mathcal{K}. \quad (4.20c)$$

The constrained  $K$ -means algorithm solves problem (4.20) iteratively by uncoupling cluster center and selection variables. Specifically, in each iteration, this algorithm alternates between solving a linear program for variable  $\boldsymbol{\iota}$  with fixed  $\mathbf{c}$  and solving a problem for  $\mathbf{c}$  with fixed  $\boldsymbol{\iota}$ . The overall constrained  $K$ -means algorithm for solving problem (4.20) is summarized in Algorithm 4.1, where the superscript  $t$  denotes the iteration index.

---

**Algorithm 4.1:** The proposed constrained  $K$ -means algorithm for user clustering.

---

**Initialization:** Initialize cluster centers  $\mathbf{c}^{(0)} = \{\mathbf{c}_1^{(0)}, \mathbf{c}_2^{(0)}, \dots, \mathbf{c}_K^{(0)}\}$  by selecting  $K$  data points from the dataset randomly. Set  $t = 0$ .

**Repeat:**

1) **Cluster assignment:** Solve the following linear program with fixed  $\mathbf{c}^{(t)}$ .

$$\begin{aligned} \boldsymbol{\iota}^{(t)} = \arg \min_{\boldsymbol{\iota}} \quad & \sum_j \sum_k \iota_{j,k} \|\mathbf{d}_j - \mathbf{c}_k^{(t)}\|_2^2 \\ \text{s.t.} \quad & (4.19\text{b}), (4.19\text{c}), (4.20\text{c}). \end{aligned}$$

2) **Cluster update:** Update the cluster centers as,

$$\mathbf{c}_k^{(t+1)} = \frac{\sum_j \iota_{j,k}^{(t)} \mathbf{d}_j}{\sum_j \iota_{j,k}^{(t)}}, \quad \forall k \in \mathcal{K}.$$

3) Set  $t \leftarrow t + 1$ .

**Until:**  $\mathbf{c}_k^{(t)} = \mathbf{c}_k^{(t-1)}, \forall k \in \mathcal{K}$ .

---

**Proposition 4.2** There exists an optimal solution for the cluster assignment subproblem in Algorithm 4.1 such that  $\iota_{j,k} \in \{0, 1\}$ .

*Proof.* See Appendix B.1. □

According to Proposition 4.2 and Appendix B.1, we can use the network simplex algorithm which is faster than mixed integer solvers for tackling the cluster assignment subproblem.

**Proposition 4.3** The constrained  $K$ -means algorithm terminates in a finite number of iterations at a cluster assignment that is locally optimal. That is, the limit point of the iterates generated by the constrained  $K$ -means algorithm is a stationary point that satisfies the Karush-Kuhn-Tucker (KKT) conditions for problem (4.20).

*Proof.* At each iteration, the cluster assignment step cannot increase the objective function of (4.20). The cluster update step will either strictly decrease the value of the objective function of (4.20) or the algorithm will terminate since,

$$\mathbf{c}^{(t+1)} = \arg \min_{\mathbf{c}} \sum_j \sum_k \iota_{j,k}^{(t)} \|\mathbf{d}_j - \mathbf{c}_k\|_2^2 \quad (4.21)$$

is a strictly convex optimization problem with a unique global solution (as shown in the cluster update step in Algorithm 1). Thus, the objective of (4.20) is strictly non-increasing and bounded below by zero. Moreover, there are a finite number of ways to assign  $J$  points to  $K$  clusters such that each cluster has at most  $q$  points and Algorithm 4.1 does not permit repeated assignments. Consequently, the algorithm must terminate at some cluster assignment that is locally optimal.  $\square$

### 4.3.2 Number of Clusters

The choice of the number of clusters  $K$  plays a key role in the performance of  $K$ -means clustering [104]. An appropriate number of clusters can accurately reflect specific distribution characteristics of users in the network. While the number of clusters cannot exceed the number of users, it should also satisfy the constraint on the maximum number of users in each cluster. However, finding the optimal  $K$  is a major challenge in clustering analysis, and there is no definitive solution. To address this problem, a number of approaches have been proposed such as the elbow [105], silhouette [106], and gap statistic [107] methods. Among these, the elbow method is possibly the most well-known and utilized, as it entails the lowest computational complexity while providing very good performance.

Herein, we employ the elbow method to determine the number of clusters. The elbow

method is a heuristic method which involves running the clustering algorithm on the dataset and evaluating a clustering criterion for different values of  $K$ . The plot of the clustering criterion versus the number of clusters resembles an arm in which the elbow point (the point of discontinuity in the slope of the curve) determines the appropriate number of clusters for the dataset. The sum of the normalized within-cluster SSE distance is a common clustering criterion for applying the elbow method along with  $K$ -means.

In a given cluster  $\mathcal{C}_k$ , the within-cluster SSE distance between the data points is given by,

$$D_k = \frac{1}{2} \sum_{\mathbf{d}_i \in \mathcal{C}_k} \sum_{\mathbf{d}_{i'} \in \mathcal{C}_k} \|\mathbf{d}_i - \mathbf{d}_{i'}\|_2^2. \quad (4.22)$$

Hence, the sum of the normalized within-cluster SSE distances can be expressed as,

$$W_K = \sum_{k=1}^K \frac{1}{|\mathcal{C}_k|} D_k \quad (4.23)$$

where  $|\mathcal{C}_k|$  shows the cardinality of the cluster  $\mathcal{C}_k$ . It should be noted that although the sum of the normalized within-cluster SSE distance can give a proper measure of the compactness of the clustering, we may encounter cases with more than one elbow point or no elbow point. In such cases, other reliable methods mentioned before can be used to find the best  $K$ .

### 4.3.3 Complexity Analysis

In this subsection, we analyze the computational complexity of the proposed constrained  $K$ -means algorithm by considering the number of required operations (e.g. complex addition and multiplication) in each step and in each iteration of the algorithm. Specifically, we divide

the operations for each iteration into three steps:

- Calculation of Euclidean distances: The complexity of calculating the Euclidean distance between the data points and the cluster centers is  $O(JKLM)$ .
- Cluster assignment: The complexity of solving cluster assignment subproblem via network simplex algorithm is  $O(J^3K^2(\log(J))^2)$  (See Appendix A).
- Cluster update: The complexity of updating the cluster centers is  $O(JKLM)$ .

Assuming that the algorithm converges after  $T_K$  iterations. The overall complexity of Algorithm 4.1 can be expressed as,

$$C_C \triangleq O(T_K J^3 K^2 (\log(J))^2 + T_K JKLM). \quad (4.24)$$

## 4.4 Downlink Beamforming

In this section, we first formulate the beamforming design as a non-convex mixed-integer nonlinear programming (MINLP) optimization problem, aiming to minimize the total transmit power while satisfying the QoS and fronthaul capacity constraints. We then propose transformations and convex approximation techniques to derive two iterative algorithms for solving the problem. In the first algorithm, we approximate the continuous non-convex constraints by convex ones using first-order Taylor expansion. Hence, we are able to arrive at a sequence of mixed-integer second-order cone programming (MI-SOCP), for which dedicated solvers are available. Although the MI-SOCP algorithm entails high computational complexity, it is shown that it can achieve high quality solution [108]. Hence, in this work, we use MI-SOCP algorithm as a benchmark. A simplified suboptimal

approach is also proposed which designs the beamformers in two stages to achieve lower complexity. In the first stage, the cluster beamformers are determined by taking advantage of BD to remove intercluster interference. In the second stage, we obtain the user-specific beamformers with the aid of CCCP method to minimize the total transmit power. Finally, the convergence and computational complexity of the proposed algorithms are discussed.

#### 4.4.1 Beamforming Problem

Our objective is to optimize the total transmit power through joint design of the dynamic RRH selection scheme and the beamforming vectors subject to the QoS and fronthaul capacity constraints on each individual RRH. Let the binary variable  $s_{j,k}^{(l)}(n) = 1$  indicate that the  $l$ th RRH participates in transmission for the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier and  $s_{j,k}^{(l)}(n) = 0$  otherwise. Hence, our optimization problem can be mathematically formulated as,

$$\min_{\mathbf{w}_{jk}(n), s_{j,k}^{(l)}(n)} \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}(n)\|^2 \quad (4.25a)$$

$$\text{s.t. } \mathbf{h}_{j'k'}(n)\mathbf{w}_{jk}(n) = \mathbf{0}, \quad \forall k' \neq k \quad (4.25b)$$

$$\gamma_{jk}(n) \geq \gamma_{\min}, \quad \forall n \in \mathcal{N}_{jk} \quad (4.25c)$$

$$\sum_k \sum_j s_{j,k}^{(l)}(n) R_{jk}(n) \leq C_{\max}, \quad \forall l, n \quad (4.25d)$$

$$\|\mathbf{w}_{jk}^l(n)\|_2 \leq s_{j,k}^{(l)}(n) P_{\max}, \quad \forall l, n \quad (4.25e)$$

$$\sum_l s_{j,k}^{(l)}(n) \geq 1, \quad \forall n \in \mathcal{N}_{jk} \quad (4.25f)$$

$$s_{j,k}^{(l)}(n) \in \{0, 1\} \quad (4.25g)$$

where  $R_{jk}(n) = \log_2(1 + \gamma_{jk}(n))$  denotes the transmission rate,  $\mathcal{N}_{jk}$  shows the set of subcarriers occupied by the  $j$ th user in the  $k$ th cluster,  $\gamma_{\min}$ ,  $C_{\max}$ , and  $P_{\max}$  are the minimum required SINR for the user over the subcarrier, the maximum capacity constraint for each RRH over the subcarrier, the maximum available total transmit power, respectively. Constraints (4.25b) and (4.25c) guarantee QoS by removing the inter-cluster interference and satisfying SINR requirements, respectively. The constraint (4.25d) shows that the sum-rate of the users served by the  $l$ th RRH over the  $n$ th subcarrier should be smaller than the maximum fronthaul capacity  $C_{\max}$ . Constraint (4.25e) utilizes the so-called Big M method which indicates that the beamformer  $\|\mathbf{w}_{jk}^l(n)\|_2 = 0$  if the  $l$ th RRH does not participate in transmission for the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier, i.e.,  $s_{j,k}^{(l)}(n) = 0$ , but leaves the beamformer "open" otherwise. Therefore,  $P_{\max}$  can be any large number. Constraint (4.25f) guarantees that each user is served by at least one RRH. Although constraint (4.25f) appears to be redundant, it is added to reduce the size of the feasible set of the associated problem which in turn improves the convergence time of the solver. We refer the interested reader to [109] for additional details.

Problem (4.25) is a non-convex MINLP problem, which can be considered as an NP-hard problem in general and is one of the most challenging class of mathematical optimization problems [110]. Obtaining its optimal solution is challenging due to the non-convexity of the SINR constraints, the combinatorial nature of the RRH selection variable  $s_{j,k}^{(l)}(n)$ , and the coupling between the variables  $s_{j,k}^{(l)}(n)$  and  $R_{jk}(n)$  in the fronthaul constraint. Even when the RRH selection scheme  $s_{j,k}^{(l)}(n)$  is given, problem (4.25) is still non-convex and computationally difficult. In the following subsections, we develop two beamforming approaches to find a suboptimal solution.

### 4.4.2 MI-SOCP Beamforming Approach

In this section, we first reformulate the problem (4.25) into a more tractable form. We then solve the resulting optimization problem via a CCCP-based algorithm with guaranteed convergence to a local stationary solution of the transformed problem.

Without loss of optimality, SINR constraint (4.25c) can be rewritten as the following second-order cone (SOC) constraint,

$$\sqrt{I_{jk}^{(1)}(n) + I_{jk}^{(2)}(n) + \sigma_{jk}^2} \leq \frac{\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)}{\sqrt{\gamma_{\min}}} \quad (4.26)$$

where  $I_{jk}^{(1)}(n)$  and  $I_{jk}^{(2)}(n)$  are the intra- and inter-cluster interference as expressed in (4.12) and (4.13) respectively. We have restricted  $\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)$  to be positive real, which incurs no loss of optimality since we can always phase-rotate the vector  $\mathbf{w}_{jk}(n)$  such that  $\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)$  is positive real without affecting the cost function or the constraints.

Let us introduce the auxiliary variables  $u_{j,k}(n)$  and  $v_{j,k}(n)$  as the upper bounds on the SINR and transmission rate for the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier. Hence, constraint (4.25d) can be rewritten as follows,

$$\gamma_{jk}(n) \leq u_{jk}(n) \quad (4.27)$$

$$\log_2(1 + u_{jk}(n)) \leq v_{jk}(n) \quad (4.28)$$

$$\sum_k \sum_j s_{j,k}^{(l)}(n)v_{jk}(n) \leq C_{\max}. \quad (4.29)$$

Since the expression of  $\gamma_{jk}(n)$  is in fractional form, the constraint in (4.27) is difficult to handle. Therefore, we introduce the auxiliary variables  $l_{j,k}(n)$  as the lower bound of the

denominator, and then equivalently transform (4.27) as the following two constraints,

$$|\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)|^2 \leq l_{jk}(n)u_{jk}(n), \quad (4.30)$$

$$l_{jk}(n) \leq I_{jk}^{(1)}(n) + I_{jk}^{(2)}(n) + \sigma_{jk}^2. \quad (4.31)$$

From the above discussion, we can finally reformulate the problem (4.25) into an equivalent problem as given below,

$$\min \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}(n)\|^2 \quad (4.32a)$$

$$\text{s.t. (4.25b),(4.25e)-(4.25g),(4.26)} \quad (4.32b)$$

$$\sqrt{4|\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)|^2 + (l_{jk}(n) - u_{jk}(n))^2} \leq l_{jk}(n) + u_{jk}(n) \quad (4.32c)$$

$$l_{jk}(n) \leq I_{jk}^{(1)}(n) + I_{jk}^{(2)}(n) + \sigma_{jk}^2 \quad (4.32d)$$

$$1 + u_{jk}(n) \leq 2^{v_{jk}(n)} \quad (4.32e)$$

$$\sum_k \sum_j (s_{j,k}^{(l)}(n) + v_{jk}(n))^2 - 4C_{\max} \leq \sum_k \sum_j (s_{j,k}^{(l)}(n) - v_{jk}(n))^2. \quad (4.32f)$$

where the identity  $4xy = (x + y)^2 - (x - y)^2$  is used to obtain (4.32f). We note that even by continuous relaxation of binary variables  $s_{j,k}^{(l)}(n)$ , optimization problem (4.32) is still non-convex due to constraints (4.32d)-(4.32f). However, the latter can be expressed as differences of two convex functions. Thus, the obtained optimization problem can be efficiently solved using the iterative CCCP.

Basically, CCCP iteratively solves a sequence of convex subproblems, each of which is constructed by linearizing the concave part of the DC constraints using their first-order Taylor expansions [95]. Specifically, the first-order Taylor expansion of the right side of

constraint (4.32d) around the current point  $\hat{\mathbf{w}}_{jk}(n)$  is expressed as,

$$\begin{aligned}
F(\mathbf{w}_{j'k}(n); \hat{\mathbf{w}}_{j'k}(n)) &= \sum_{j' \neq j, j' \in \mathcal{U}_{n,k}} [-|\mathbf{h}_{jk}(n)\hat{\mathbf{w}}_{j'k}(n)|^2 + 2\Re\{\hat{\mathbf{w}}_{j'k}^H(n)\mathbf{h}_{jk}^H(n)\mathbf{h}_{jk}(n)\mathbf{w}_{j'k}(n)\}] \\
&+ \sum_{k' \neq k} \sum_{j' \in \mathcal{U}_{n,k'}} [-|\mathbf{h}_{jk}(n)\hat{\mathbf{w}}_{j'k'}(n)|^2 + 2\Re\{\hat{\mathbf{w}}_{j'k'}^H(n)\mathbf{h}_{jk}^H(n)\mathbf{h}_{jk}(n)\mathbf{w}_{j'k'}(n)\}]
\end{aligned} \tag{4.33}$$

where  $\Re\{\cdot\}$  denotes the real part of its argument. In the same way, we convexify the right side of constraints (4.32d) and (4.32e) by using the first-order Taylor expansions around the current points  $\hat{v}_{jk}(n)$ ,  $\hat{s}_{j,k}^{(l)}(n)$ , and  $\hat{v}_{jk}(n)$  as,

$$\Gamma(v_{jk}(n); \hat{v}_{jk}(n)) = 2^{\hat{v}_{jk}(n)} + (\ln 2)2^{\hat{v}_{jk}(n)}(v_{jk}(n) - \hat{v}_{jk}(n)), \tag{4.34}$$

$$\Omega(s_{j,k}^{(l)}(n), v_{jk}(n); \hat{s}_{j,k}^{(l)}(n), \hat{v}_{jk}(n)) = -(\hat{s}_{j,k}^{(l)}(n) - \hat{v}_{jk}(n))^2 + 2(\hat{s}_{j,k}^{(l)}(n) - \hat{v}_{jk}(n))(s_{j,k}^{(l)}(n) - v_{jk}(n)). \tag{4.35}$$

By applying the above approximations to the non-convex constraints (4.32d)-(4.32f), we can formulate the convex approximation of problem (4.32) as shown below,

$$\min \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}(n)\|^2 \tag{4.36a}$$

$$\text{s.t. (4.25b), (4.25e)-(4.25g), (4.26), (4.32c)} \tag{4.36b}$$

$$l_{jk}(n) \leq F(\mathbf{w}_{j'k}(n); \hat{\mathbf{w}}_{j'k}(n)) + \sigma_{jk}^2 \tag{4.36c}$$

$$1 + u_{jk}(n) \leq \Gamma(v_{jk}(n); \hat{v}_{jk}(n)) \tag{4.36d}$$

$$\sum_k \sum_j (s_{j,k}^{(l)}(n) + v_{jk}(n))^2 - 4C_{\max} \leq \Omega(s_{j,k}^{(l)}(n), v_{jk}(n); \hat{s}_{j,k}^{(l)}(n), \hat{v}_{jk}(n)). \tag{4.36e}$$

---

**Algorithm 4.2:** MI-SOCP beamforming algorithm.

---

Initialize the algorithm with feasible points  $\hat{\mathbf{w}}_{jk}(n)$ ,  $\hat{s}_{j,k}^{(l)}(n)$ , and  $\hat{v}_{jk}(n)$ . Set iteration index  $t = 0$  and termination threshold  $\epsilon > 0$ .

**Repeat**

- 1) Update  $\hat{\mathbf{w}}_{jk}(n)$ ,  $\hat{s}_{j,k}^{(l)}(n)$ , and  $\hat{v}_{jk}(n)$  by solving problem (4.36).
- 2) Set  $t = t + 1$ .

**Until:** Termination criterion is met:  $\Delta P_T < \epsilon$ .

---

Hence, based on CCCP, we solve subproblem (4.36) at each iteration. Problem (4.36) is a MI-SOCP which can be solved via modern solvers such as MOSEK [111] or GUROBI [112]. The proposed iterative algorithm is summarized in Algorithm 4.2. The algorithm terminates if the variation of the total transmit power, i.e.,  $\Delta P_T$ , is less than a preset threshold  $\epsilon$ .

*Initialization:* Choosing a feasible point for initialization of Algorithm 4.2 is essential. For this purpose, we simply set  $\hat{v}_{jk}(n) = \log_2(1 + \gamma_{\min})$  and then solve the following feasibility problem  $P = \text{find}\{s_{j,k}^{(l)}(n) | (25\text{f}), (25\text{g}), \sum_k \sum_j s_{j,k}^{(l)}(n) \hat{v}_{jk}(n) \leq C_{\max}\}$  which is a mixed-integer linear program which can be solved optimally by off-the-shelf solvers such as MOSEK or GUROBI. Subsequently, we solve the following quadratic program with fixed  $\hat{s}_{j,k}^{(l)}(n)$  via any general-purpose solver using interior-point method,

$$\begin{aligned} \hat{\mathbf{w}}_{jk}(n) = \arg \min_{\mathbf{w}_{jk}(n)} & \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}(n)\|^2 \\ \text{s.t.} & (4.25\text{b}), (4.26), \\ & \|\mathbf{w}_{jk}^l(n)\|_2 \leq \hat{s}_{j,k}^{(l)}(n) P_{\max}. \end{aligned} \tag{4.37}$$

### 4.4.3 Two-Stage Beamforming Approach

In order to reduce the computational complexity, we propose a two-stage energy-efficient beamforming approach such that,

$$\mathbf{w}_{jk}(n) = \mathbf{B}_k(n)\mathbf{v}_{jk}(n) \quad (4.38)$$

where  $\mathbf{B}_k(n) \in \mathbb{C}^{LM \times a}$  is the  $k$ th cluster beamformer obtained in the first stage which should eliminate the inter-cluster interference and  $\mathbf{v}_{jk}(n) \in \mathbb{C}^{a \times 1}$  is the user-specific beamformer for the  $j$ th user in the  $k$ th cluster optimized in the second stage.

Using channel state information (CSI) available at the central processor, BD beamforming can be adopted in a MIMO SCMA system to remove the inter-cluster interference and enhance the QoS for intra-cluster users [113]. Hence, the users in different clusters can share codebooks. Although BD algorithm does not work well in the presence of imperfect CSI, we considered a second stage for beamforming in which the QoS can be guaranteed. Specifically, the BD beamforming projects the transmitted signal onto the null-space of the interfering channels and hence eliminates the inter-cluster interference.

To find the corresponding null-space, let us define,

$$\mathbf{H}_k(n) = [\mathbf{h}_{1k}(n)^T \dots \mathbf{h}_{J_k}(n)^T] \in \mathbb{C}^{LM \times J_k} \quad (4.39)$$

$$\mathbf{H}_{-k}(n) = [\mathbf{H}_1(n) \dots \mathbf{H}_{k-1}(n) \mathbf{H}_{k+1}(n) \dots \mathbf{H}_K(n)] \quad (4.40)$$

where  $k \in \mathcal{K}$  and  $\mathbf{H}_{-k}(n) \in \mathbb{C}^{LM \times (J - J_k)}$  is the matrix containing all interfering channels for the  $k$ th cluster. We seek  $\mathbf{B}_k(n)$  orthogonal to the column span of  $\mathbf{H}_{-k}(n)$ , i.e.,  $\mathbf{H}_{-k}(n)^T \mathbf{B}_k(n) = \mathbf{0}$ . Here, it is assumed that the total number of antennas  $LM$  is larger

than the total number of users  $J$ .

The singular value decomposition (SVD) can be employed to calculate the cluster beamformers. Applying the SVD to  $\mathbf{H}_{-k}(n)$  yields,

$$\mathbf{H}_{-k}(n) = \mathbf{U}_k(n)\mathbf{\Sigma}_k(n)\mathbf{V}_k(n)^H \quad (4.41)$$

where  $\mathbf{U}_k(n) \in \mathbb{C}^{LM \times LM}$  and  $\mathbf{V}_k(n) \in \mathbb{C}^{(J-J_k) \times (J-J_k)}$  are unitary matrices and  $\mathbf{\Sigma}_k(n) \in \mathbb{R}^{LM \times (J-J_k)}$  is the rectangular diagonal matrix of singular values. Let  $r$  denote the rank of matrix  $\mathbf{H}_{-k}(n)$ , which corresponds to the number of non-zero diagonal entries in  $\mathbf{\Sigma}_k(n)$ . The null-space of the interfering channel matrix  $\mathbf{H}_{-k}(n)$  is spanned by the left singular vectors (i.e. columns of matrix  $\mathbf{U}_k(n)$ ) associated to the zero singular values of  $\mathbf{H}_{-k}(n)$ . We can express the  $k$ th cluster beamformer as,

$$\mathbf{B}_k(n) = [\mathbf{u}_{r+1,k}(n) \ \mathbf{u}_{r+2,k}(n) \ \dots \ \mathbf{u}_{LM,k}(n)] \quad (4.42)$$

where  $\mathbf{u}_{i,k}(n)$  denotes the  $i$ th column of  $\mathbf{U}_k(n)$ .

As mentioned before, constraint (4.25e) implies that  $\|\mathbf{w}_{jk}^{(l)}(n)\|_2 = 0$  if  $s_{j,k}^{(l)}(n) = 0$ . Without loss of optimality, the binary RRH selection variable  $s_{j,k}^{(l)}(n)$  can be replaced by  $\|\|\mathbf{w}_{jk}^{(l)}(n)\|_2\|_0$ , as in [114], [115]. Therefore, upon substitution of (4.38) and  $\ell_0$ -norm, problem (4.25) can be rewritten as,

$$\min_{\mathbf{v}_{j,k}(n)} \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}(n)\|^2 \quad (4.43a)$$

$$\text{s.t. } \mathbf{w}_{jk}(n) = \mathbf{B}_k(n)\mathbf{v}_{jk}(n), \quad \forall j, k, n \quad (4.43b)$$

$$\gamma_{jk}(n) \geq \gamma_{\min}, \quad \forall n \in \mathcal{N}_{jk} \quad (4.43c)$$

$$\sum_k \sum_j \|\mathbf{w}_{jk}^{(l)}(n)\|_2^2 \|_0 R_{jk}(n) \leq C_{\max}, \quad \forall l, n \quad (4.43d)$$

It should be noted that the fronthaul capacity constraint (4.43d) which is expressed in the form of  $\ell_0$ -norm, indicates the inherently dynamic RRH selection. That is, owing to this fronthaul constraint, the network-wide beamforming vectors  $\mathbf{w}_{jk}(n)$  may have a sparse structure. Although the number of constraints is reduced and the binary RRH selection variable is removed, problem (4.43) is still non-convex due to constraints (4.43c) and (4.43d).

As mentioned before, using cluster beamformer  $\mathbf{B}_k(n)$  obtained from BD can remove inter-cluster interference. Hence, the SINR of the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier can be expressed as,

$$\gamma_{jk}(n) = \frac{|\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)|^2}{\sum_{j' \neq j}^{J_k} |\mathbf{h}_{jk}(n)\mathbf{w}_{j'k}(n)|^2 + \sigma_{jk}^2} \quad (4.44)$$

where the inter-cluster interference term in the denominator is removed. Consequently, SINR constraint (4.43c) can be rewritten as follows,

$$\sqrt{\sum_{j' \neq j}^{J_k} |\mathbf{h}_{jk}(n)\mathbf{w}_{j'k}(n)|^2 + \sigma_{jk}^2} \leq \frac{|\mathbf{h}_{jk}(n)\mathbf{w}_{jk}(n)|}{\sqrt{\gamma_{\min}}} \quad (4.45)$$

which is a SOC constraint.

To address the non-convexity of constraint (4.43d), we first introduce the auxiliary variables  $u_{j,k}(n)$ ,  $v_{j,k}(n)$ , and  $t_{j,k}^{(l)}(n)$  as the upper bounds of the SINR, transmission rate, and  $\ell_0$ -norm for the  $j$ th user in the  $k$ th cluster over the  $n$ th subcarrier. Hence, constraint

(4.43d) can be rewritten as follows,

$$\gamma_{jk}(n) \leq u_{jk}(n), \quad (4.46)$$

$$\log_2(1 + u_{jk}(n)) \leq v_{jk}(n), \quad (4.47)$$

$$\| \|\mathbf{w}_{jk}^{(l)}(n)\|_2^2 \|_0 \leq t_{j,k}^{(l)}(n), \quad (4.48)$$

$$\sum_k \sum_j t_{j,k}^{(l)}(n) v_{jk}(n) \leq C_{\max}. \quad (4.49)$$

We then propose to approximate the non-convex  $\ell_0$ -norm by a reweighted  $\ell_1$ -norm as follows [116],

$$\| \|\mathbf{w}_{jk}^{(l)}(n)\|_2^2 \|_0 \approx \beta_{jk}^{(l)}(n) \|\mathbf{w}_{jk}^{(l)}(n)\|_2^2. \quad (4.50)$$

$\beta_{jk}^{(l)}(n)$  is a constant weight which is updated in each iteration according to,

$$\beta_{jk}^{(l)}(n) = \frac{1}{\|\hat{\mathbf{w}}_{jk}^{(l)}(n)\|_2^2 + \tau} \quad (4.51)$$

where  $\hat{\mathbf{w}}_{jk}^{(l)}(n)$  is obtained from previous iteration and  $\tau$  is a small constant regularization factor controlling the smoothness of the approximation. Based on the updating rule (4.51),  $\beta_{jk}^{(l)}(n)$  is inversely proportional to the transmit power level  $\|\hat{\mathbf{w}}_{jk}^{(l)}(n)\|_2^2$ . Hence, the RRHs with lower transmit power for the  $j$ th user in the  $k$ th cluster would have higher weights and hence would be forced to further reduce its transmit power and eventually be dropped out of the group of participating RRHs for that user.

We can employ the approach mentioned in IV.B to deal with the non-convexity of the

---

**Algorithm 4.3:** Proposed CCCP-based iterative algorithm for beamforming.

---

**Initialization:** Initialize  $\hat{\mathbf{v}}_{jk}(n)$  randomly. Calculate  $\hat{\mathbf{w}}_{jk}(n)$ ,  $\hat{t}_{j,k}^{(l)}(n)$  and  $\beta_{jk}^{(l)}(n)$ .  
Set iteration index  $t = 0$  and termination threshold  $\epsilon > 0$ .

**Repeat**

- 1) Update  $\hat{\mathbf{w}}_{jk}(n)$ ,  $\hat{t}_{j,k}^{(l)}(n)$ , and  $\hat{v}_{jk}(n)$  via solving problem (4.52).
- 2) Calculate  $\beta_{jk}^{(l)}(n)$ .
- 3) Set  $t = t + 1$ .

**Until** Termination criterion is met:  $\Delta P_T < \epsilon$ .

---

constraints and use CCCP to solve the optimization problem. Hence, based on CCCP, we solve the following subproblem at each iteration,

$$\min_{\mathbf{v}_{jk}(n)} \sum_n \sum_k \sum_j \|\mathbf{w}_{jk}(n)\|^2 \quad (4.52a)$$

$$\text{s.t. (4.32c),(4.36c),(4.36d),(4.38),(4.43b),(4.45)} \quad (4.52b)$$

$$\beta_{jk}^{(l)}(n) \|\mathbf{w}_{jk}^{(l)}(n)\|_2^2 \leq t_{j,k}^{(l)}(n) \quad (4.52c)$$

$$\sum_k \sum_j (t_{j,k}^{(l)}(n) + v_{jk}(n))^2 - 4C_{\max} \leq \Omega(t_{j,k}^{(l)}(n), v_{jk}(n); \hat{t}_{j,k}^{(l)}(n), \hat{v}_{jk}(n)). \quad (4.52d)$$

Problem (4.52) is convex and can be solved via any general-purpose solver using interior-point methods [76]. The proposed CCCP-based iterative algorithm is summarized in Algorithm 4.3.

*Initialization:* In this case, an initial point for Algorithm 4.3 is obtained by generating  $\hat{\mathbf{v}}_{jk}(n)$  randomly. Then,  $\hat{\mathbf{w}}_{jk}(n)$  and  $\beta_{jk}^{(l)}(n)$  are calculated as in (4.38) and (4.51) respectively.  $\hat{t}_{j,k}^{(l)}(n)$  is set to  $\|\hat{\mathbf{w}}_{jk}^{(l)}(n)\|_2^2$ , and  $\hat{v}_{jk}(n)$  is set to the transmission rate calculated using  $\hat{\mathbf{w}}_{jk}(n)$ .

#### 4.4.4 Convergence and Complexity Analysis

With a feasible initial point, repeated application of the CCCP iteration is guaranteed to converge to a stationary solution of the problem with DC constraints. It can be seen that the optimal solution obtained from the previous iteration, i.e.,  $\hat{\mathbf{w}}_{jk}(n)$ , is feasible for the convex subproblem at the next iteration for both algorithms. The achieved objective at the current iteration cannot be greater than the one at the previous iteration. Since, the objective function is non-increasing and bounded below by zero, it follows that both algorithms will converge to a point that according to [82] is locally optimal. We refer the interested reader to [82] for a rigorous proof of the convergence.

For Algorithm 4.2, the overall complexity is dominated by solving the MI-SOCP problem in (4.36). In particular, there are  $JLN$  binary variables  $s_{j,k}^{(l)}(n)$ , resulting in  $2^{JLN}$  combinations for all the binary variables. Thus, assuming that MI-SOCP algorithm terminates after  $T_M$  iterations, the worst-case complexity can be written as

$$C_M \triangleq \mathcal{O}(T_M 2^{JLN} (JCLM)^3). \quad (4.53)$$

At each iteration, the CCCP-based algorithm solves the convex subproblem (4.52) which can be approximated by a sequence of SOCPs via the successive approximation method. Each SOCP can then be solved via a general-purpose solver, e.g., SDPT3 in CVX [78] with a complexity of  $\mathcal{O}((JCLM)^3)$ . Assuming that the CCCP terminates after  $T_C$  iterations, the worst-case computational complexity is therefore given by,

$$C_B \triangleq \mathcal{O}(T_C (JCLM)^3). \quad (4.54)$$

## 4.5 Simulation Results

In this section, numerical experiments are carried out to illustrate the performance of the proposed energy-efficient user clustering and downlink beamforming for MIMO SCMA in C-RAN.

### 4.5.1 Methodology

In our simulations, unless otherwise specified, we consider a network with  $L = 3$  RRHs, each equipped with  $M = 5$  antennas and serving  $J = 12$  single-antenna users. The RRHs and the users are independently distributed in a square area  $[-50, 50] \times [-50, 50]$  meters. The RRHs are connected to the central processor via a limited-capacity fronthaul link with maximum capacity  $C_{\max} = 50$  bps/Hz. The maximum available total transmit power is  $P_{\max} = 50$  dBm.

We consider the channel model as described in Section 4.2.1 with bandwidth of  $W = 2$  GHz and carrier frequency of 28 GHz. The AoDs are assumed to follow a uniform distribution in  $[0, 2\pi]$ . The inter-antenna spacing is  $d = \lambda/2$  to reduce the effect of mutual coupling and correlation among neighbouring antenna elements. The noise figure is  $N_f = 40$  dBm, hence, the noise power is  $\sigma_{jk}^2 = -174 + 10 \log_{10}(W) + N_f$  dBm [87]. The pathloss exponent of the LOS and NLOS paths in (4.4) are  $\alpha^{(0)} = 2$  and  $\alpha^{(p)} = 3$ , respectively. For SCMA encoder, the number of subcarriers is  $N = 4$ , and the number of non-zero elements for each codeword

Notation	Description	Value
$L$	Number of RRHs	3
$M$	Number of antennas per RRH	5
$J$	Total number of users	12
$N$	Number of SCs	5
$C$	Number of non-zero elements for a codeword	2
$\alpha^{(0)}$	path loss exponent of the LOS path	2
$\alpha^{(p)}$	path loss exponent of the NLOS path	3
$C_{\max}$	Maximum capacity constraint for each RRH	50 bps/Hz
$P_{\max}$	Maximum available total transmit power	50 dBm
$\epsilon$	Termination threshold	$10^{-6}$

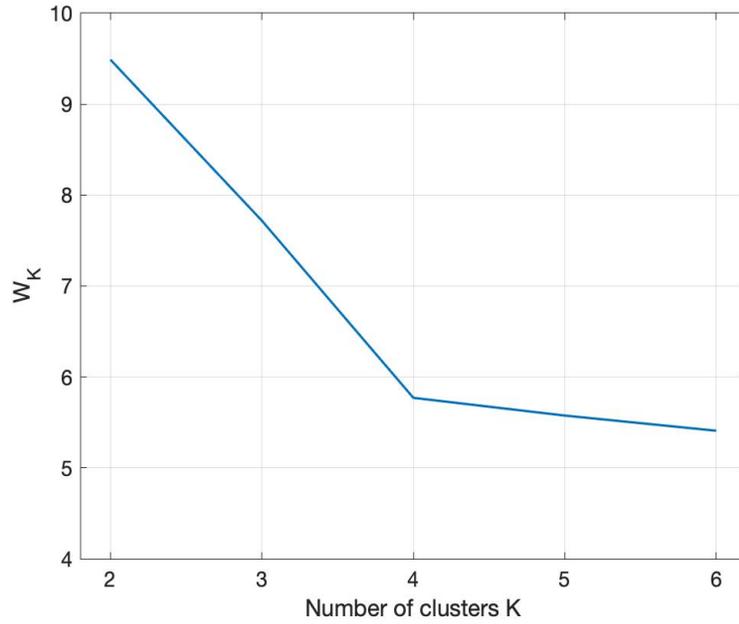
**Table 4.1:** Simulation setting parameters

is  $C = 2$ . The corresponding factor graph matrix is,

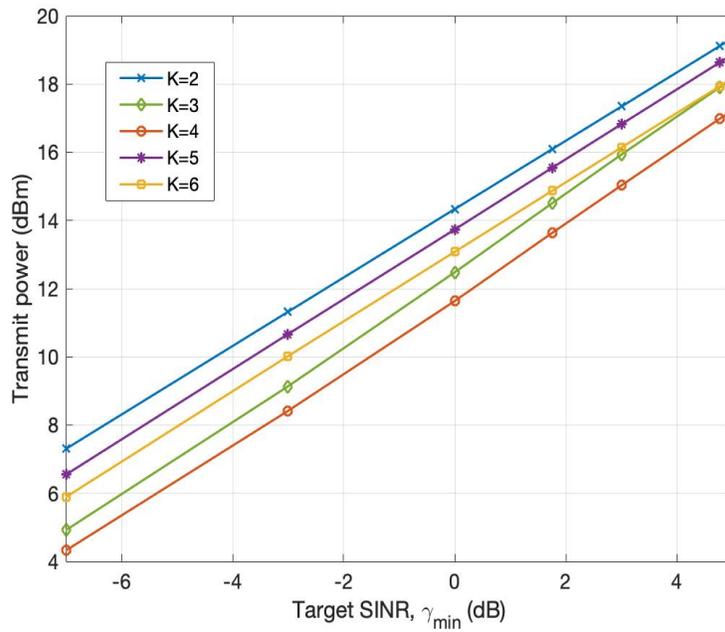
$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \quad (4.55)$$

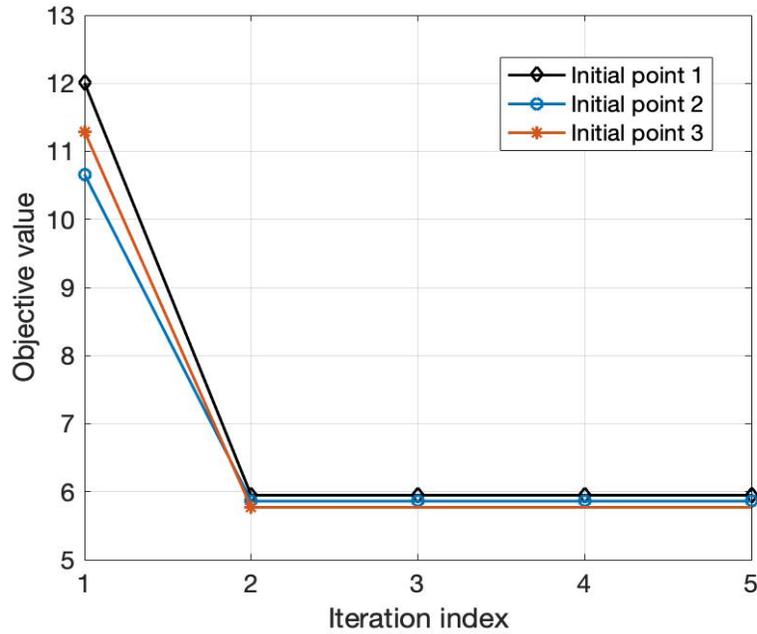
It should be noted that the structure of the factor graph matrix with fixed  $N$  and  $C$  does not affect system performance significantly. Table 4.1 summarizes the simulation setting parameters.

We use Monte Carlo experiments with 100 runs to evaluate the performance of the proposed algorithms for user clustering and downlink beamforming. The total transmit power and sum rate are measured for different parameter configurations and the results are compared with benchmark approaches in the literature.

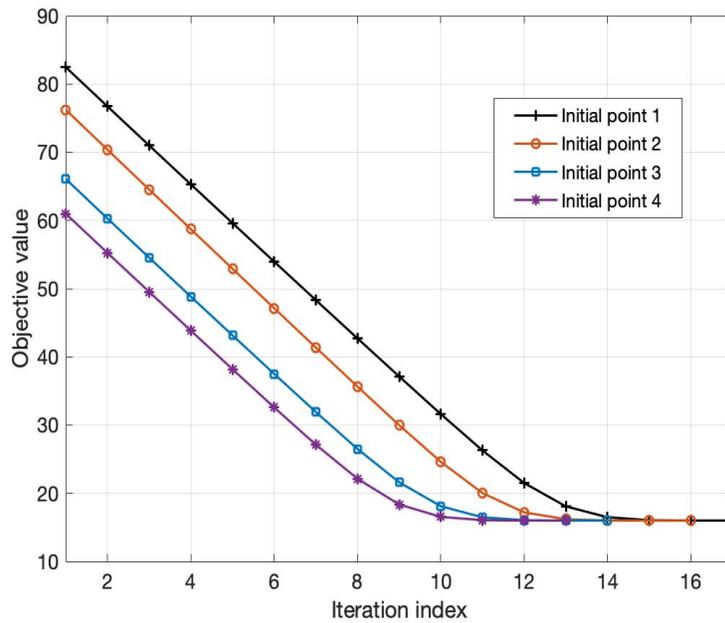


(a) Elbow method.

(b) Transmit power versus target SINR  $\gamma_{\min}$ .**Figure 4.2:** The impact of the number of clusters.



(a) The constrained K-means algorithm.



(b) The CCCP-based iterative algorithm.

**Figure 4.3:** The convergence of the proposed algorithms.

### 4.5.2 Results and Discussion

Figs. 4.2a and 4.2b find the optimal number of clusters  $K$  and evaluate its impact on the performance of the proposed scheme. In Fig. 4.2a, we plot the sum of the normalized within-cluster SSE distance which serves as clustering criterion in the elbow method described in Section III.B. It can be seen  $W_K$  decreases when  $K$  increases and the elbow point can be found at  $K = 4$ <sup>1</sup>. To gain further insight into the impact of the number of clusters, we investigate the transmit power performance versus target SINR,  $\gamma_{min}$  in Fig. 4.2b, where the number of clusters increases from 2 to 6. It is observed that the total transmit power increases monotonically as  $\gamma_{min}$  increases. Moreover, the best performance is achieved when the number of clusters is  $K = 4$ . On one hand, for  $K < 4$ , an increase in the number of users in a cluster results in larger intra-cluster interference which results in higher transmit power. On the other hand, increasing the number of clusters intensifies inter-cluster interference which increases power consumption in the first stage beamforming for interference cancellation. We thereby observe that better user clustering with lower total transmit power can be found at  $K = 4$  and the elbow method can efficiently find the optimal number of clusters in this case.

Figs. 4.3a and 4.3b present the convergence behaviour of the proposed constrained  $K$ -means and CCCP-based algorithms. Fig. 4.3a shows the objective value achieved by the constrained  $K$ -means algorithm with three different initial points. In this regards,  $J = 12$  users are grouped into  $K = 4$  non-overlapping clusters of size less than 6, i.e.,  $q = 6$ . We observe that the algorithm converges rapidly in a few steps and the gap between final results of different initial points is small. In Fig. 4.3b, the convergence performance of the CCCP-

---

<sup>1</sup>We omit the results on the Silhouette and gap statistic methods here for brevity. However, it should be noted that using each of these methods for determining the optimal number clusters gives the same result while elbow method entails lower computational complexity.

	$\gamma_{min}$			
	0	2	4	6
Two-Stage approach	3.6108	4.3852	5.7032	7.0414
MI-SOCP approach	182.3750	238.3348	323.9511	397.0151

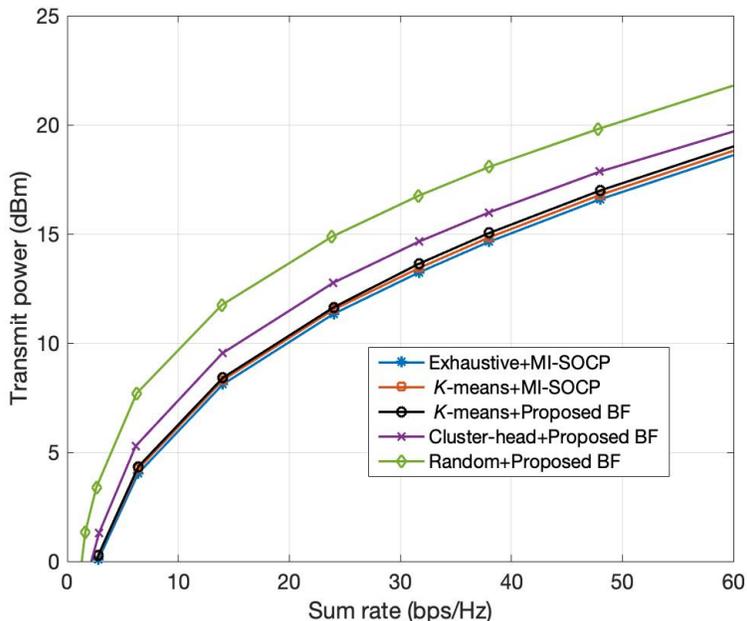
**Table 4.2:** Run-time of different beamforming approaches

based algorithms is investigated for the case of  $\gamma_{min} = 3$  dB. It can be seen that the algorithm converges in less than 15 iterations monotonically to a same value for different initial points.

In Table 4.2, we present the run-time comparison between the proposed two-stage approach and the MI-SOCP beamforming for different values of  $\gamma_{min}$ . The results<sup>2</sup>, show that the complexity of the proposed two-stage beamforming algorithm is much less than that of the MI-SOCP approaches, owing to the use of the smoothed  $\ell_0$ -norm approximation.

In Fig. 4.4, we compare the transmit power versus sum rate among different clustering and beamforming algorithms. The cluster-head approach proposed in [117] is used as a benchmark for user clustering which selects the  $K$  users with the highest channel gains as the cluster centers. The cluster assignment is then used to group users into clusters. We also consider the performance of the MIMO SCMA system with exhaustive and random clustering. The combination of exhaustive search for user clustering and MI-SOCP for beamforming is shown to attain the best performance among all the algorithms. However, this comes at the cost of high computational complexity. As it can be seen from Fig. 4.4, the proposed constrained  $K$ -means clustering algorithm exhibits better performance compared to random search and cluster-head approach and can partition users more efficiently. Regarding the beamforming algorithms, it is shown in Fig. 4.4 that the suboptimal solution achieved

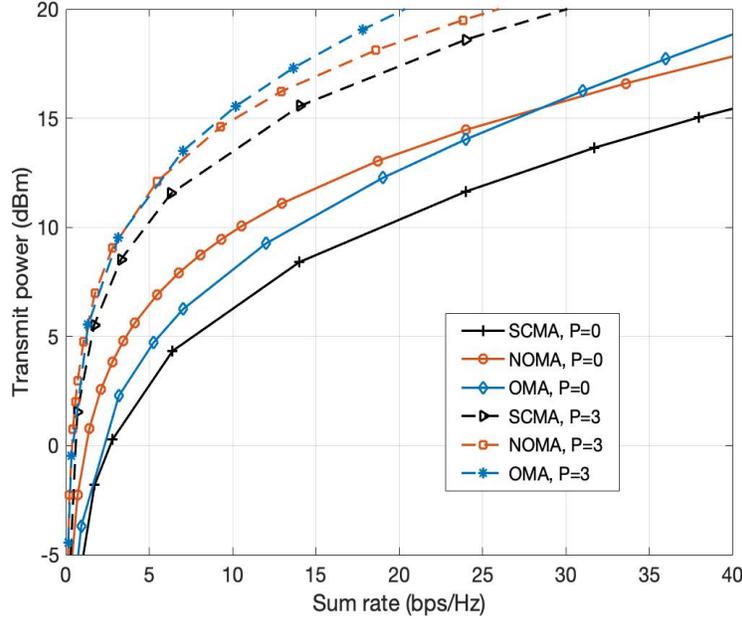
<sup>2</sup>Based on the use of a desktop computer equipped with 8th Generation Intel i7-8700 6-core processor (12M Cache, 4.6 GHz) and 32GB RAM.



**Figure 4.4:** Transmit power versus sum rate for different clustering and beamforming approaches.

by the proposed two-stage beamforming is very close to the high-quality solution obtained by MI-SOCP.

To better appreciate the benefits of the proposed MIMO SCMA scheme in terms of spectral efficiency, we examine the achievable sum rate of the users within the network. We consider orthogonal multiple access (OMA) and power domain NOMA as benchmarks with similar parameters except the number of multiplexed signals over each subcarrier. Specifically, in OMA, each user is assigned only one subcarrier such that no interference occurs with other user signals. Hence, the maximum number of users in each cluster is equal to the number of subcarriers, i.e.,  $q = N$ . In power domain NOMA, all users have access to all the subcarriers and no constraint is applied to the maximum number of users in a cluster. In this section, we refer to power domain NOMA simply by NOMA.



**Figure 4.5:** Transmit power versus sum rate for different transmission schemes.

Fig. 4.5 compares the total transmit power versus achievable sum rate among the proposed SCMA, power domain NOMA and OMA schemes. Two different channel models are considered for this purpose, one with no NLOS path, i.e.,  $P = 0$ , and the other with  $P = 3$  NLOS components. It is observed that in both cases, the results of the proposed SCMA scheme outperforms other schemes in terms of sum rate and the performance gap gets larger as the transmit power increases. Moreover, we observe that the transmit power for NOMA is more than that of OMA. However, as the sum rate increases, the results for OMA exhibits a noticeable increase in transmit power compared to NOMA.

To investigate the impact of imperfect CSI on the proposed user clustering and downlink beamforming, we model the estimated channel vector as follows,

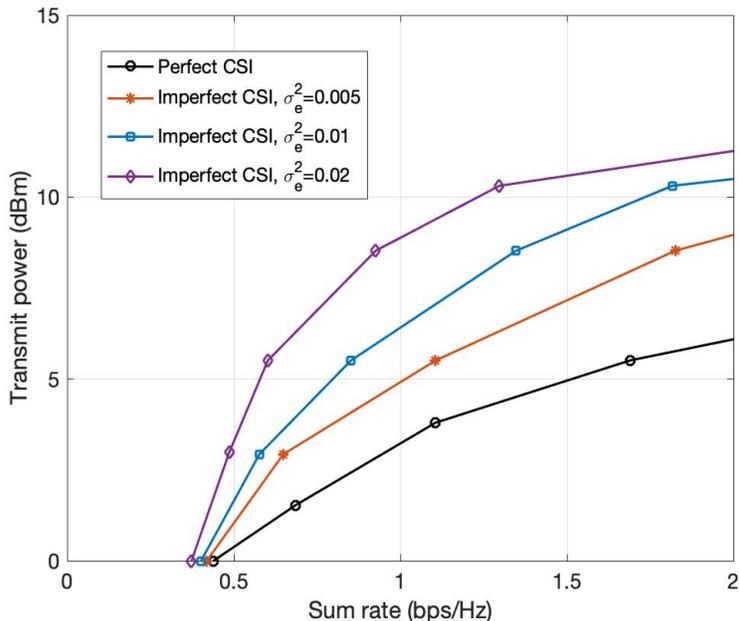
$$\hat{\mathbf{h}}_{jk}(n) = \mathbf{h}_{jk}(n) + \Delta_{jk}(n) \quad (4.56)$$

where  $\mathbf{h}_{jk}(n)$  is the actual channel vector and  $\Delta_{jk}(n)$  is CSI error with i.i.d. entries following a complex Gaussian distribution, i.e.,  $\Delta_{jk}(n) \sim \mathcal{CN}(\mathbf{0}, \sigma_e^2 \mathbf{I})$ .

Fig. 4.6 shows the transmit power comparison for the channel model with  $P = 3$  NLOS paths, where the perfect and imperfect CSI with different  $\sigma_e$  scenarios are considered. It can be seen that the system performance is sensitive to the CSI accuracy. This is due to the fact that the channel correlation is used as the similarity metric for the proposed constrained  $K$ -means algorithm, which largely depends on the obtained CSI at the central processor. Moreover, the BD algorithm does not work well in the presence of imperfect CSI and can not remove inter-cluster interference totally. In order to enhance the performance of the proposed user clustering and downlink beamforming in the presence of imperfect CSI, one can use a more sophisticated similarity metric in the clustering algorithm or robust beamforming in the second stage of the proposed approach [118–120]. However, these considerations are beyond of the scope of this work.

Fig. 4.7 presents the average number of associated RRHs per user versus the fronthaul link capacity for different target SINRs,  $\gamma_{min}$ . It can be seen that due to the limitation on the capacity of the fronthaul links, each user can be only served by a small group of RRHs. For fixed  $\gamma_{min}$ , the number of RRHs associated with each user will increase as the fronthaul link capacity grows. Moreover, for a fixed fronthaul link capacity, the group of associated RRHs will increase as  $\gamma_{min}$  gets smaller. In fact, the data rate of each user will become smaller for a lower  $\gamma_{min}$ . Thus, each RRH can serve more users with lower data rate.

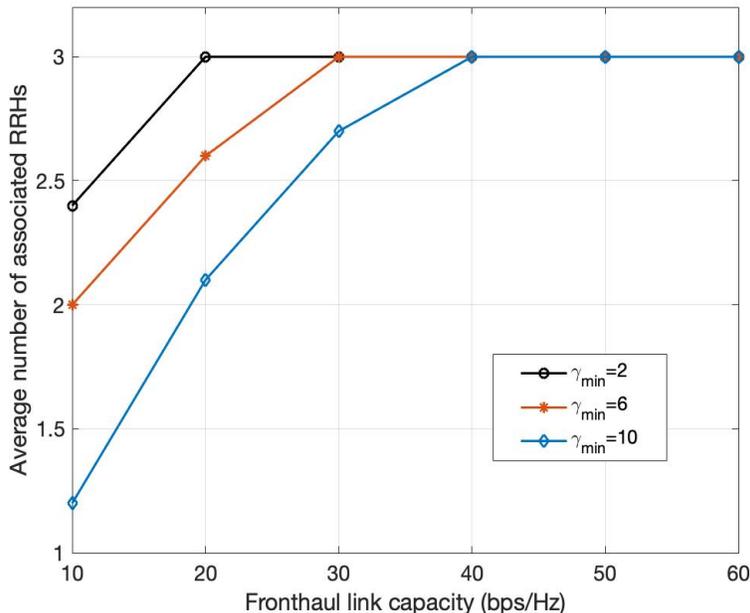
Fig. 4.8 shows the total transmit power versus number of antennas  $M$  for different target SINRs,  $\gamma_{min}$ . In this regard, the channel model with  $P = 3$  NLOS paths is considered and  $J = 18$  users are grouped into  $K = 4$  clusters. It is worth noting that the number of clusters are determined through the elbow method for this case. For larger  $M$ , better beamforming



**Figure 4.6:** Transmit power versus sum rate for perfect and imperfect CSI.

results are expected as more degrees of freedom will be left for inter-cluster interference cancellation. We can observe that the total transmission power decreases as the number of antennas increases which is a consequence of narrower beamforming.

Fig. 4.9 shows the total transmit power versus total number of users  $J$  for different target SINRs,  $\gamma_{min}$ . In this regard, the channel model with  $P = 3$  NLOS paths is considered and the users are grouped into  $K = 4$  clusters. As the results indicate, the total transmission power depends on the number of users and  $\gamma_{min}$ . As the number of users or the target SINR increases, larger total transmission power is needed to satisfy QoS and fronthaul capacity requirements.



**Figure 4.7:** Average number of associated RRHs versus fronthaul link capacity for  $M = 20$ .

## 4.6 Concluding Remarks

In this work, the design of user clustering and beamforming approach was investigated for MIMO SCMA in C-RAN. We proposed a constrained  $K$ -means algorithm for user clustering. By taking advantage of CSI available at the central processor, this algorithm was applied to partition users into non-overlapping clusters based on the correlation between channel vectors. The beamforming design was formulated as an optimization problem, with the aim to minimize the total transmit power under the SINR and fronthaul capacity constraints, and two iterative algorithms were proposed for its solution. In the first approach, the high-quality solution was achieved by solving a MI-SOCP in each iteration via dedicated solvers. In the second approach, a two-stage low-complexity beamforming design was proposed where in the first stage, the BD was employed to obtain the cluster beamformers, while in the second stage,

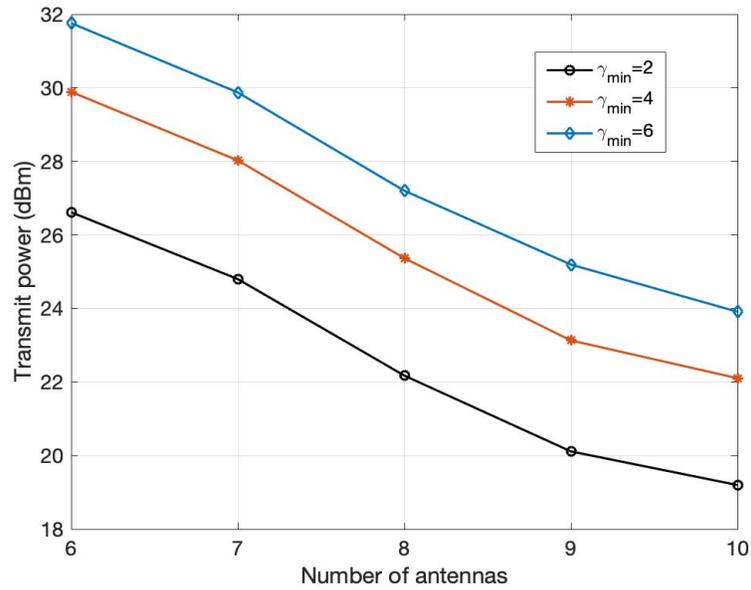


Figure 4.8: Transmit power versus number of antennas.

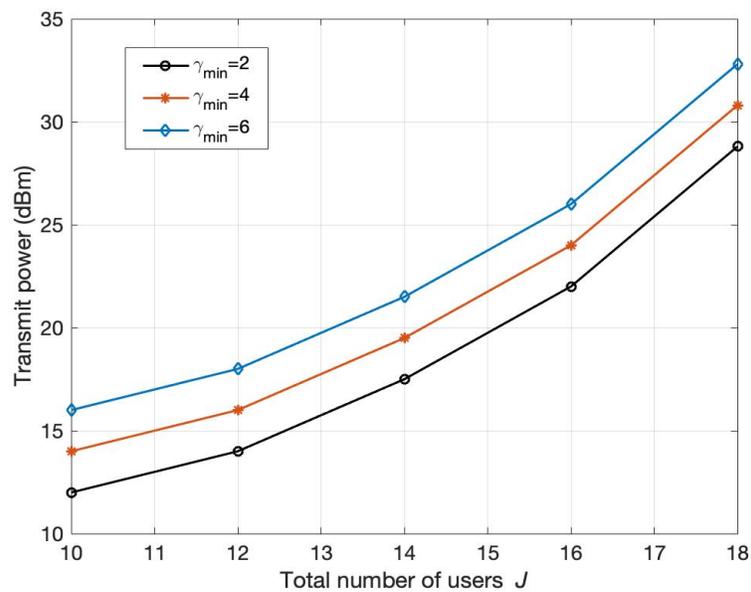


Figure 4.9: Transmit power versus total number of users.

the design of user-specific beamformers was formulated as an optimization problem. Through simulations, it was shown that the proposed user clustering and beamforming approaches for MIMO SCMA systems can effectively decrease total transmit power, eliminate inter-cluster interference, and improve spectral efficiency compared to the benchmark approaches.

## Chapter 5

# Deep Residual Neural Network Decoder for SCMA

In this chapter, we focus on the design of SCMA decoders and develop a novel decoding algorithm based on deep residual neural network (ResNet). The proposed decoder is validated by means of simulations over AWGN and Rayleigh fading channels.

### 5.1 Introduction

Applying deep learning (DL) methods to wireless communications problems is of great interest as it can bring significant performance improvements. DL methods can solve challenging problems by using multiple processing layers to progressively extract higher level features from the raw input. On the one hand, deep neural networks (DNNs), as a class of DL methods, can deal with multidimensional and nonlinear characteristics of raw

---

Parts of the materials in this chapter have been submitted to the IEEE Wireless Communications and Networking Conference 2023 (WCNC'23) in Scotland, UK [122]

input. On the other hand, MPA entails high computational complexity, although it can achieve near optimum performance as an SCMA decoder. Hence, in recent years, DNNs have been applied to enable autonomous derivation of efficient algorithms for SCMA decoder. In [61], the DNN-based decoder reconstructs the input data symbol, given the received signal. In this reference only AWGN channel is considered and a fully connected network is utilized for the DNN-based decoder. In [63], considering AWGN channels, a fully connected DNN is used to reconstruct the transmitted bits of all users given the channel and received signal. A DNN-based multi-output classification is proposed in [64], where the transmitted codewords are predicted given received signal in the presence of noise and Rayleigh fading channels. In spite of their good performance and low complexity, these DNN-based decoders suffer from several drawbacks such as accuracy saturation, vanishing gradients and instability.

In our proposed DNN structure, residual blocks are utilized to tackle the above mentioned problems with existing DNN approaches, while batch normalization is employed to improve the robustness of the decoder to different initialization and learning rates. Under the assumption that the channel state information (CSI) is available at the receiver side, the decoder is trained to predict the transmitted codewords by users. In order to predict the transmitted user bits, a non-mutually exclusive classification problem should be considered in which the activation function in the output layer of the decoder would be the sigmoid function. However, the sigmoid function suffers from the problem of vanishing gradients. Hence, our aim in this work is to predict the user codewords using the softmax function in the output layer. Specifically, the received signal and CSI are fed into the ResNet decoder as input, while the output consists of multiple branches, one for each user, wherein the transmit codewords are predicted. Through simulations, it is demonstrated

that the proposed SCMA scheme with ResNet decoder can notably reduce bit error rate (BER) compared to DL-based benchmark approaches with lower complexity.

The rest of the chapter is organized as follows: Section 5.2 introduces the system model under consideration. In Section 5.3, the proposed ResNet decoder for SCMA is developed. The simulation results are presented in Section 5.4, followed by conclusion in Section 5.

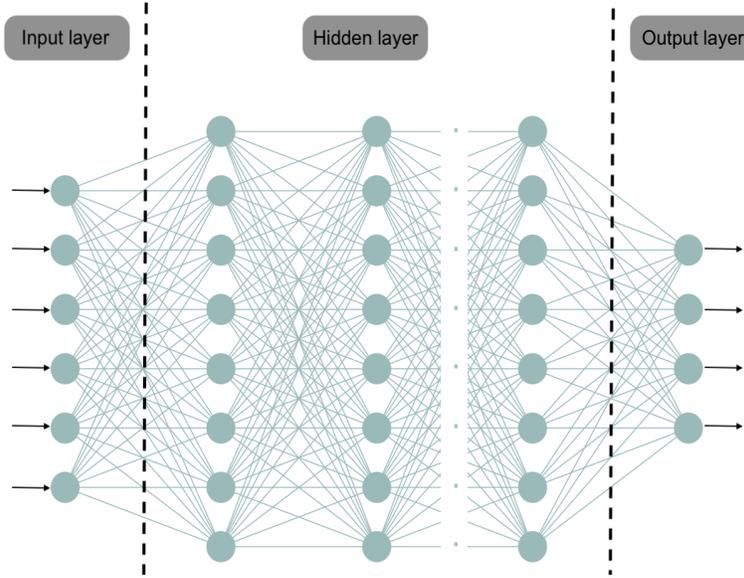
## 5.2 System Model

We consider the uplink of a SCMA system with  $J$  independent users, indexed by  $j \in \mathcal{J} = \{1, \dots, J\}$ , transmitting data over  $K$  shared resources, e.g., orthogonal frequency division multiple access (OFDMA) subcarriers or multiple-input multiple-output (MIMO) spatial layers. It is assumed that the number of resources is less than the number of users, i.e.,  $K < J$ . Below, we provide further details on the SCMA encoder and then introduce relevant DNN framework and notations as needed for further development of the proposed decoder.

### 5.2.1 SCMA Encoder

SCMA directly maps groups of user data bits to  $K$ -dimensional complex codewords selected from a user-specified codebook. Hence, the SCMA encoder for the  $j$ th user can be defined as an injective function, i.e.,  $f_j : \mathbb{B}^m \rightarrow \mathcal{S}_j$  where  $\mathbb{B}^m$  denotes the set of  $m$ -bit tuples and  $\mathcal{S}_j \subset \mathbb{C}^K$  denotes the user-specified codebook. It can be inferred that each user codebook contains  $M = 2^m$  multi-dimensional complex codewords, i.e.,  $|\mathcal{S}_j| = M$ . Specifically, the corresponding codeword for input bits  $\mathbf{b} \in \mathbb{B}^m$  is obtained as,

$$\mathbf{s}_j = f_j(\mathbf{b}) = [s_{j,1}, \dots, s_{j,K}]^T. \quad (5.1)$$



**Figure 5.1:** DNN structure.

The complex codeword  $\mathbf{s}_j$  is sparse with  $N < K$  non-zero elements. Each user is assigned  $N$  subcarriers such that no two users occupy the same set of subcarriers.

Let  $\mathbf{h}_j = [h_{j1}, \dots, h_{jK}] \in \mathbb{C}^{1 \times K}$  denote the channel vector for the  $j$ th user. The received signal for the uplink scenario can be expressed as,

$$\mathbf{r} = \sum_{j=1}^J \text{diag}(\mathbf{h}_j) \mathbf{s}_j + \mathbf{n} = \sum_{j=1}^J \text{diag}(\mathbf{h}_j) f_j(\mathbf{b}) + \mathbf{n} \quad (5.2)$$

where  $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$  is an additive noise term. In this work, we assume that the channel vectors  $\mathbf{h}_j, \forall j \in \mathcal{J}$ , are known at the receiver side.

### 5.2.2 Deep Neural Networks (DNNs)

As shown in Fig. 5.1, a fully connected DNN is composed of three main components, namely: input layer, multiple hidden layers, and output layer. The input layer passes on

the features to the hidden layer without performing any computation. Each hidden layer consists of multiple nodes that progressively extract higher-level features from the raw input. Let  $\mathbf{x}_l \in \mathbb{R}^{N_l \times 1}$  denote the input of the  $l$ th layer. In each layer, the following computation is accomplished,

$$\mathbf{x}_{l+1} = \varphi_l(\mathbf{W}_l^T \mathbf{x}_l + \mathbf{b}_l) \quad (5.3)$$

where  $\mathbf{W}_l \in \mathbb{R}^{N_l \times N_{l+1}}$  is the weight matrix, and  $\mathbf{b}_l \in \mathbb{R}^{N_{l+1} \times 1}$  is the bias vector, and  $\varphi_l(\cdot)$  is a non-linear activation function operating element-wise on its vector input. The activation function adds non-linearity to the DNN and determines whether a node should be activated or not. The output layer delivers the resulting information learned through the hidden layers as the final layer. The activation function utilized in the output layer is different from that in the hidden layers and depends on the prediction type and DNN model.

A loss function  $\mathcal{L}(\cdot)$  is required to evaluate the difference between the network output and its expected output. Examples of loss functions include the mean squared error (MSE) for regression tasks, and the cross entropy for binary classification tasks. The network has to be trained such that the loss function is minimized. Accordingly, the weights and biases are updated using back propagation and gradient descent as follows,

$$\begin{aligned} [\mathbf{W}'_l]_{i,j} &= [\mathbf{W}_l]_{i,j} - \alpha \frac{\partial \mathcal{L}(\mathbf{W}, \mathbf{B})}{\partial [\mathbf{W}_l]_{i,j}} \\ [\mathbf{b}'_l]_i &= [\mathbf{b}_l]_i - \alpha \frac{\partial \mathcal{L}(\mathbf{W}, \mathbf{B})}{\partial [\mathbf{b}_l]_i} \end{aligned} \quad (5.4)$$

where  $\mathbf{W}$  and  $\mathbf{B}$  denote the concatenation of weight matrices and bias vectors of all layers in the network, respectively, and  $\alpha$  is the learning rate. Several techniques such as stochastic gradient descent [123], momentum [124], and adaptive moment estimation (Adam) [125]

have been proposed to improve the training performance and decrease the computational complexity of the gradient descent method.

### 5.3 Proposed ResNet Decoder

DNNs are universal function approximators which can model complex non-linear relationships. However, using deep networks creates problems [126]. For instance, by increasing the number of layers, the accuracy will saturate at one point and eventually degrade. Moreover, DNNs suffer from problems such as vanishing gradients and instability. Herein, we conceive a novel DNN-base decoder to tackle the aforementioned problems. Batch normalization is used to improve the speed, performance and stability of the DNN-base decoder. Moreover, residual blocks are employed to avoid the problem of vanishing gradients, and to mitigate the accuracy saturation problem. Below, we provide further details on the batch normalization and residual blocks. Then, we introduce the structure of the proposed DNN-based SCMA decoder.

#### 5.3.1 Batch Normalization

Batch normalization is a technique that can stabilize the learning process and significantly reduce the number of training epochs in deep networks by standardizing the inputs to a layer [127]. In effect, the aim is to keep the input of each layer in the active region of the activation function using linear transformations. In this regard, the whole training data set is divided into training min-batches. Let us define  $\mathbf{z}_l \in \mathbb{R}^{N_{l+1} \times 1}$  such that  $\mathbf{z}_l = \mathbf{W}_l^T \mathbf{x}_l + \mathbf{b}_l$  and denote the min-batch of size  $N_b$  by  $\mathcal{B} = \{\mathbf{z}_l^{(1)}, \dots, \mathbf{z}_l^{(N_b)}\}$ . The batch normalization is implemented through the following steps:

- Mean calculation:

$$\mu_{l,k} = \frac{1}{N_b} \sum_{i=1}^{N_b} [\mathbf{z}_l^{(i)}]_k. \quad (5.5)$$

- Variance calculation:

$$\sigma_{l,k}^2 = \frac{1}{N_b} \sum_{i=1}^{N_b} ([\mathbf{z}_l^{(i)}]_k - \mu_{l,k})^2. \quad (5.6)$$

- Normalization:

$$[\hat{\mathbf{z}}_l^{(i)}]_k = \frac{[\mathbf{z}_l^{(i)}]_k - \mu_{l,k}}{\sqrt{\sigma_{l,k}^2 + \epsilon}}. \quad (5.7)$$

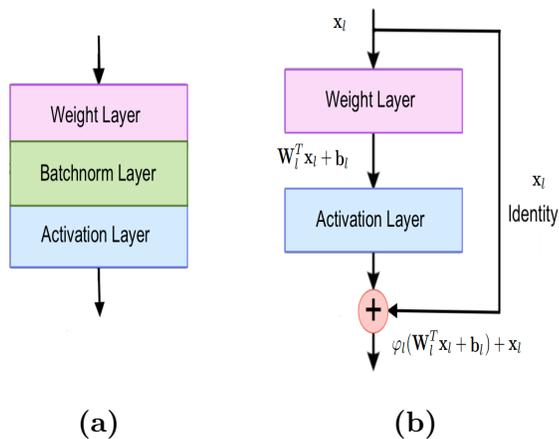
where  $\epsilon > 0$  is a constant added to the batch variance for numerical stability.

- Scaling and shifting:

$$[\boldsymbol{\alpha}_l^{(i)}]_k = \gamma_{l,k} [\hat{\mathbf{z}}_l^{(i)}]_k + \beta_{l,k} \quad (5.8)$$

where  $\gamma_{l,k}$  and  $\beta_{l,k}$  denote trainable parameters that are subsequently learned as part of the optimization process.

It should be noted that adjusting trainable parameters allows the model to choose the optimum distribution for each hidden layers. Specifically,  $\gamma_{l,k}$  allows to adjust the standard deviation of the  $\boldsymbol{\alpha}_l^{(i)}$ , while  $\beta_{l,k}$  allows to adjust their mean. As illustrated in Fig. 5.2a, batch normalization layer is positioned before the activation layer, i.e.,  $\boldsymbol{\alpha}_l^{(i)}, \forall i \in \{1, \dots, N_b\}$ , are fed into the activation functions to produce the output of the layer. By utilizing a batch



**Figure 5.2:** (a) Batch normalization. (b) Residual block.

normalization layer, the network becomes more robust to different initialization and learning rates.

### 5.3.2 Residual Blocks

A network with residual blocks is called residual network (ResNet). In the residual blocks, each layer feeds into the next layer and directly into the one of the layers about 2–3 hops away, as illustrated in Fig. 5.2b. The skip connections (or identity shortcuts) may jump over activation, weight, and batch normalization layers in between. Consequently, the network can skip the training of a few layers using skip connections.

Utilizing residual blocks in a DNN offers several benefits. For instance, ResNet is capable of learning simple functions such as an identity function, which is not possible using a fully connected deep networks. Moreover, skip connections propagate larger gradients to initial layers and avoid the problem of vanishing gradients. In a DNN with residual blocks, stacking layers would not degrade the network performance and hence, the accuracy saturation problem is mitigated.

### 5.3.3 Structure of Proposed SCMA Decoder

In this work, our goal is to predict the transmitted codewords,  $\mathbf{s}_j, \forall j \in \mathcal{J}$ , from the received signal,  $\mathbf{r} \in \mathbb{C}^K$ , in the presence of noise and fading channels, as represented by model equation (5.2). Hence, we approach this multi-output classification problem by implementing a ResNet decoder which predicts multiple outputs simultaneously. Specifically, we utilize the batch normalization and residual blocks to enhance the performance and stability of the decoder.

Fig. 5.3 illustrates the complete structure of the proposed ResNet SCMA decoder, which consists of a shared pre-processor and  $J$  classifiers for each user. The pre-processor can separate the superimposed signals while the classifiers can predict the transmitted codewords. Increasing the number of hidden layers in the classifiers can improve the prediction of the decoder, but this will significantly raise the complexity of the decoder. By increasing the number of hidden layer in the pre-processor, we can reach a good trade-off between the complexity of the network and its performance.

In this structure, the combination of weight, batchnorm, and activation layers are called neural network (NN) blocks. In the input layer, the decoder accepts the concatenation of the real and imaginary parts of the received signal,  $\mathbf{r}$ , and channel vectors,  $\mathbf{h}_j, \forall j \in \mathcal{J}$ . The rectified linear unit (ReLU) is used for the activation function of the hidden layers which can be expressed as,

$$\varphi(z) = \max\{0, z\}. \quad (5.9)$$

while in the output layer, softmax is used for the activation function. Hence, the decoder returns  $J$  vectors  $\mathbf{p}_j = [p_{j,1}, \dots, p_{j,M}], \forall j \in \mathcal{J}$ , such that the element  $p_{j,m} \in [0, 1]$  denotes the softmax probability of transmission of the  $m$ th codeword in the  $j$ th user codebook. For

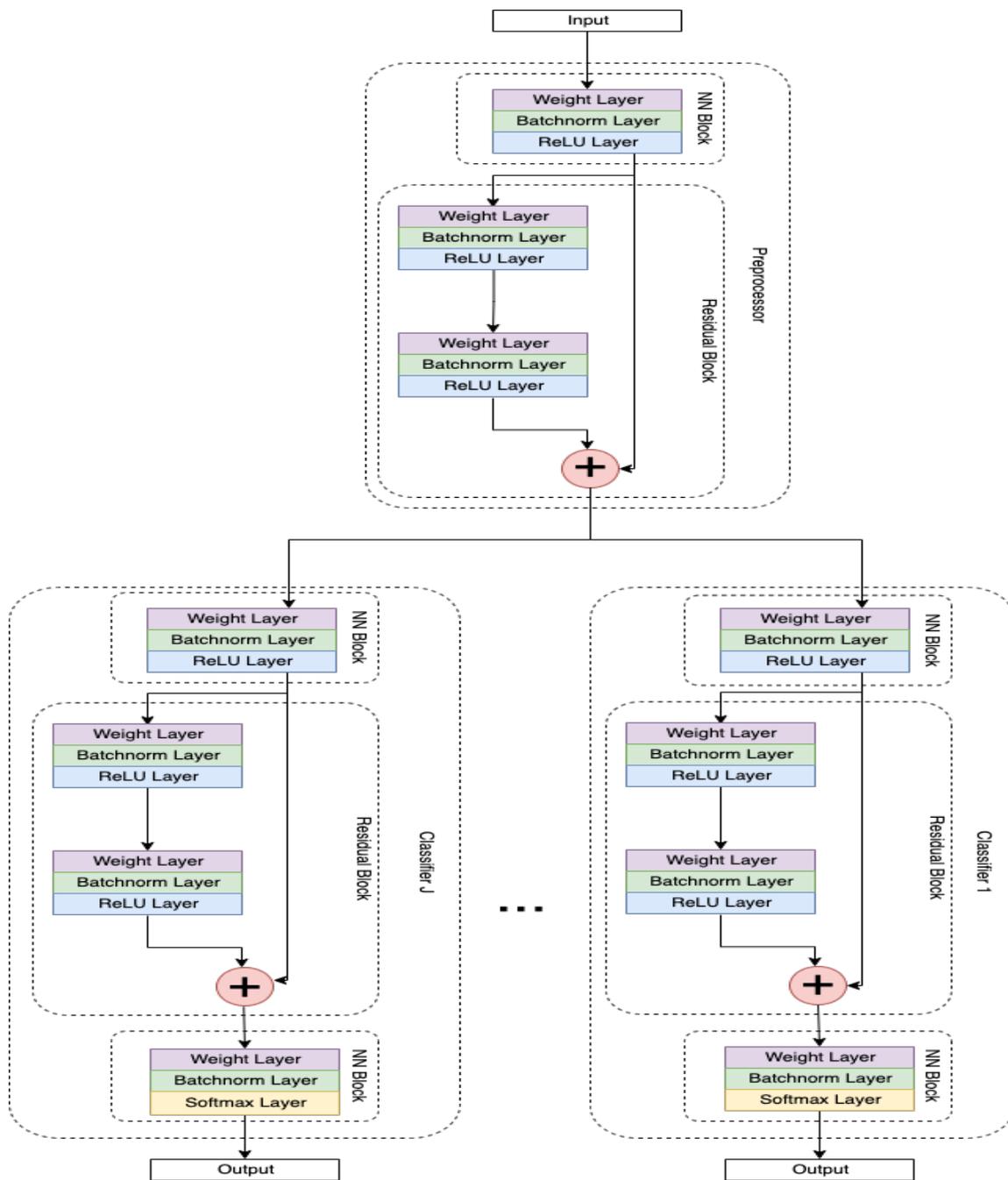


Figure 5.3: Internal structure of ResNet decoder.

an input  $\mathbf{z} \in \mathbb{R}^M$ , the softmax function returns a vector of the same dimension whose  $i$ th entry is calculated as below,

$$[\varphi(\mathbf{z})]_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}. \quad (5.10)$$

During the training phase, the aim is to find the optimum solution for the set of network weights and biases, represented by  $\mathbf{W}_l$  and  $\mathbf{b}_l$  such that the loss function between the original transmitted bits from all users and the decoder outputs is minimized. In this work, the cross entropy is used as the loss function which can be written as

$$\mathcal{L} \equiv \mathcal{L}(\{\mathbf{t}_j, \mathbf{p}_j\}; \mathbf{W}, \mathbf{B}) = - \sum_{j=1}^J \sum_{m=1}^M t_{j,m} \log(p_{j,m}) \quad (5.11)$$

where  $t_{j,m}$  and  $p_{j,m}$  denote the  $m$ th entry of the truth label of the transmitted codeword and the output probability of the  $m$ th codeword for the  $j$ th user, respectively. Specifically, the truth label of the  $j$ th user, i.e.,  $\mathbf{t}_j$ , is a binary vector with only one non-zero element whose index is determined by the index of the transmitted codeword in the  $j$ th user codebook.

## 5.4 Simulation Results

This section presents simulation results to evaluate the performance of the proposed ResNet decoder for SCMA systems over AWGN and Rayleigh fading channels. For this purpose, we consider  $J = 6$  independent users, which are transmitting data over  $K = 4$  shared resources through the SCMA encoder described in Section 5.2.1. The codebook design given in [128] is utilized in the SCMA encoder. Throughout the experiments, it is assumed that the channels are the same for all users, i.e.,  $\mathbf{h}_j = \mathbf{h}, \forall j \in \mathcal{J}$ . For Rayleigh fading, we consider

block fading where the channel coefficients are modeled as complex Rayleigh fading random variables, with zero-mean and unit variance. For AWGN, we model the channels as constant all-one vectors.

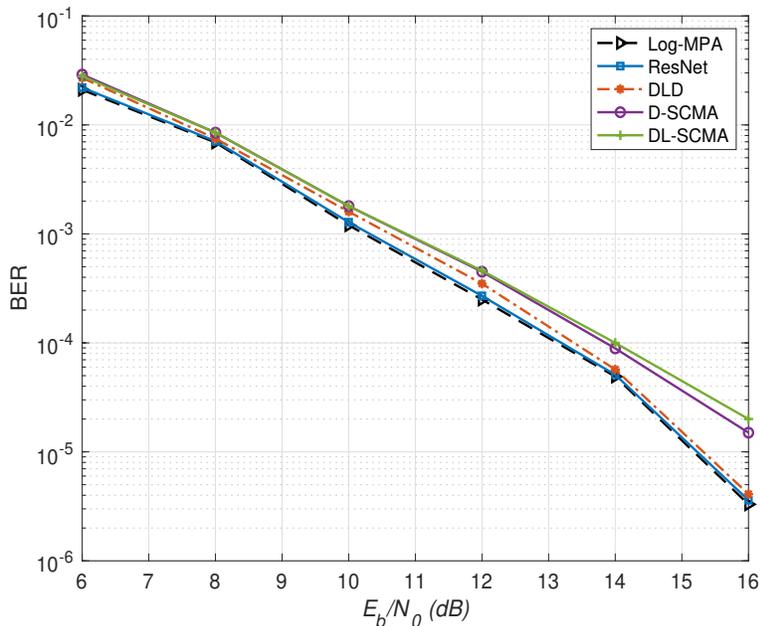
During the simulations, the user data bits are generated randomly and passed through the encoder and channel to produce the received signal<sup>1</sup>. The training dataset consists of five groups, each containing  $4 \times 10^6$  samples with a different value of SNR per bit, i.e.,  $E_b/N_0 \in \{8, 10, 12, 14, 16\}$  in dB. The learning rate in (5.4) is set to  $\alpha = 0.0001$ . The weights and biases of the network are updated using Adam. The batch sizes for the training and test phase are set to 500 and 1000, respectively. All simulations are run on a DELL OptiPlex 7040 desktop computer with an Intel Core i7-6700 CPU @3.4GHz (4 cores and 8 threads), 32GB RAM and an 4GB Nvidia GM107 (GeForce GTX 745) GPU. Software environment includes Ubuntu 18.04.6 LTS operating system, PyTorch 1.9.0 and Python 3.8.

Fig. 5.4 compares the BER performance versus  $E_b/N_0$  among conventional Log-MPA, D-SCMA [61], DL-SCMA [63], DLD [64], and the proposed ResNet decoder under AWGN channels. In this regard, the numbers of hidden nodes in each NN block are set to  $\{32, 32, 32, 16, 16, 16, 8\}$ . We observe that the the proposed ResNet SCMA decoder has notably lower BER than D-SCMA and DL-SCMA, while it shows slight performance improvements compared to DLD decoder. The outstanding performance of Log-MPA is not surprising since it achieves the near optimum solution. However, this improvement comes at the cost of high computational complexity. Meanwhile, the BER performance of the proposed ResNet decoder is very close to that of Log-MPA, but is obtained with much lower complexity (as illustrated later).

Fig. 5.5 shows the BER performance of SCMA decoders under Rayleigh fading

---

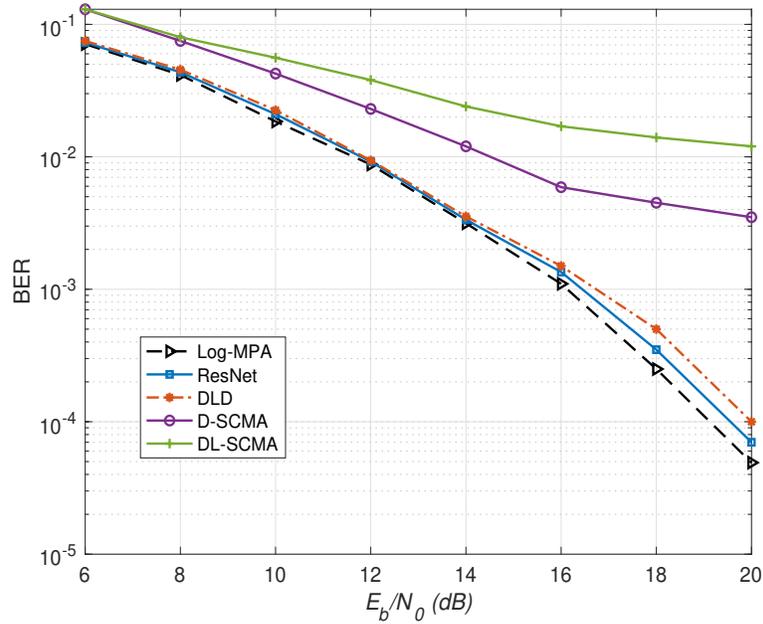
<sup>1</sup>During the training phase, it is assumed that all users are transmitting signals. However, if one or more users are not transmitting data, the corresponding classifier can be deactivated during the test phase.



**Figure 5.4:** BER performance of SCMA decoders under AWGN.

channels. Since decoding in presence of Rayleigh fading is harder than AWGN, we use the same overall DNN structure but with larger numbers of hidden nodes, which are set to  $\{256, 256, 256, 128, 128, 128, 64\}$ . It can be seen that the BER performance of all decoders degrades compared to previous results as Rayleigh fading channels impose more challenges on the decoders. It can also be noted D-SCMA and DL-SCMA exhibit inferior performance compared to the other decoders and cannot cope with Rayleigh fading channels. It is shown in Fig. 5.5 that the sub-optimal solution achieved by the proposed ResNet decoder is very close to the near-optimal solution obtained by Log-MPA.

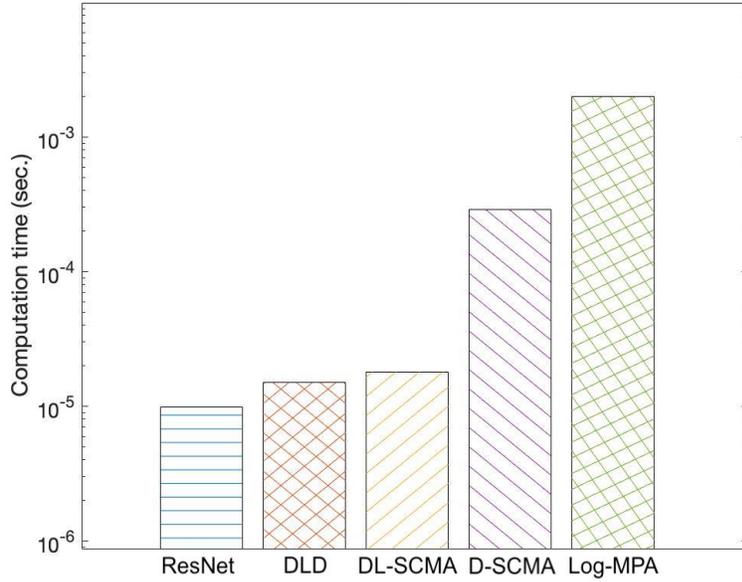
Let  $N_{p,i}$  and  $N_{c,i}$  denote the number of nodes in the  $i$ th NN block of the pre-processor and classifier, respectively. The computational complexity of the proposed network is measured



**Figure 5.5:** BER performance of SCMA decoders under Rayleigh fading channel.

	ResNet	DLD	DL-SCMA	D-SCMA	Log-MPA
Add.	568	424	252	1,560	20,520
Mul.	10,880	12,224	14,784	54,672	9,456
Log/exp	0	0	0	0	5,713

**Table 5.1:** Complexity comparison.



**Figure 5.6:** Computation time comparison among SCMA decoders.

by the number of multiply-accumulate (MAC) operations which can be expressed as,

$$\begin{aligned}
 & N_{p,1}(N_{\text{in}} + 2N_{p,2} + 3) + JN_{c,1}(N_{p,1} + 2N_{c,2} + N_{\text{out}} + 3) \\
 & + N_{p,2} + JN_{c,2} + JN_{\text{out}}
 \end{aligned} \tag{5.12}$$

where  $N_{\text{in}} = N_{p,0} = 4K$  and  $N_{\text{out}} = N_{c,3} = M$ . Regarding simulation configurations for AWGN channels, Table 5.1 compares the complexity of the proposed and benchmark decoders for passing the data through the network by calculating the number of multiplication, addition, log/exp operations. It should be noted that log/exp will cost much more time, as shown in the next results. It can be observed that the proposed ResNet decoder reduces the complexity compared to DLD, DL-SCMA, D-SCMA, and Log-MPA, respectively.

Finally, Fig. 5.6 shows the computation time of different approaches for passing the data through the network considering AWGN channels. It should be noted that all DNN-based decoders are simulated in PyTorch which provides Tensor computing with strong acceleration via GPU. It can be seen that the computation time of the proposed ResNet decoder is 1.5, 1.8, 2.93, and 200.2 times lower than that of DLD, DL-SCMA, D-SCMA, and Log-MPA, respectively.

## 5.5 Concluding Remarks

Recently, DNN has been used to address the problems of codebook design and decoding strategy for emerging SCMA systems. In this work, We have investigated the application of ResNets for SCMA decoder under AWGN and Rayleigh fading channels. Assuming that CSI is known at the receiver side, we developed a multi-output classifier to predict the transmitted user codewords. In the proposed structure, batch normalization and residual blocks are utilized to improve the performance and stability the decoder. Through simulations, it was shown that the proposed ResNet decoder outperforms competing DNN-based benchmarks in terms of BER while exhibiting notably lower complexity.

# Chapter 6

## Conclusion

In this Chapter, we review our main research contributions and discuss possible avenues for future work.

### 6.1 Summary

Through spatial diversity, multiplexing or beamforming gain, the MIMO techniques can offer significant performance improvements in terms of user capacity, spectral efficiency, and peak data rates. In addition, multiple access radio technologies play a crucial role in improving system performance in cellular mobile networks. Recently, NOMA has shown great promises in meeting higher system throughput and solving the massive connectivity issue in future radio access technologies. This thesis focused on the application of MIMO techniques along with NOMA to meet the exacting demands of 5G and beyond 5G (B5G) wireless networks. Specifically, the main contributions of the thesis are summarized as follows.

In Chapter 3, we addressed the joint design of user clustering, downlink beamforming and power allocation in a MIMO NOMA system operating at mmWave frequencies. In the

proposed scheme, users were partitioned into non-overlapping clusters to improve resource utilization: within each cluster, a common beamforming vector was shared by the users who were distinguished on the basis of allocated power. The joint optimization framework was formulated as a MINLP model, aiming to minimize the total transmission power while satisfying QoS, user clustering and power constraints. To tackle the challenges imposed by the non-convexity and combinatorial nature of the problem, we first developed an algorithm based on BB to find the global optimum within a finite number of iterations. In the BB-based algorithm, the feasible space is successively partitioned and searched by means of lower and upper bounds on the objective function to return an  $\epsilon$ -optimal solution. Considering the computational complexity of BB, we then reformulated the original problem into a more tractable form and conceived a low-complexity algorithm for its solution based on the PDD technique. Through simulations, it was shown that applying the proposed design algorithms to the multi-user MIMO NOMA system can effectively decrease total transmit power and improve spectral efficiency compared to the benchmark approaches.

In Chapter 4, we considered the problem of user clustering and downlink beamforming for MIMO SCMA in a C-RAN, assuming that CSI is available at the BS. The constrained  $K$ -means algorithm was proposed and applied to spatially partition users into non-overlapping clusters based on the correlation between channel vectors. Subsequently, two iterative algorithms for beamforming design were developed, aiming to minimize the total transmission power under QoS and fronthaul capacity constraints. In the first approach, the continuous non-convex constraints were approximated using first-order Taylor expansion and the high-quality solution was achieved by solving a MI-SOCP in each iteration via dedicated solvers. In the second two-stage beamforming approach, the cluster and user-specific beamformers were designed to remove the inter-cluster interference and

reduce total transmit power, respectively. In the first stage, the BD beamforming approach was employed to design the cluster beamformers. In the second stage, the design of user-specific beamformers was determined by minimizing the total transmit power under the SINR constraints. Simulation results showed that applying the proposed design to MIMO SCMA systems under C-RAN environment can effectively decrease total transmit power and improve spectral efficiency compared to the benchmark approaches.

In Chapter 5, we designed a ResNet decoder for SCMA to address the computational complexity limitation of current MPA decoders. In our approach, residual blocks are employed to tackle the problems of accuracy saturation and vanishing gradients with deep learning based decoder, while batch normalization is utilized to enhance the stability and performance of the decoder. The performance of the proposed ResNet decoder for SCMA is validated by means of simulations over AWGN and Rayleigh fading channels. The results show that besides a much reduced complexity compared to the MPA detector, the proposed decoder leads to improvements in terms of BER over competing DL based decoders in literature.

## 6.2 Potential Future Works

In this section, we discuss some potential future works, which are closely related to the contributions presented in this thesis.

In Chapter 3, the proposed joint design of user clustering, downlink beamforming, and power allocation for MIMO NOMA systems assumes that near perfect CSI is available at the BS for joint processing and perfect interference cancellation is performed during SIC procedure. However, once an error occurs in SIC (due to e.g., noise or imperfect CSI), signal

of the corresponding user will not be completely removed, leaving some residual signals as interference. Consequently, the message of all remaining users in the corresponding cluster will likely be decoded erroneously. Indeed, the performance degradation is worse when the number of users in a cluster increases. Robust beamforming and nonlinear detection techniques can be considered to enhance the system performance and suppress the error propagation. Using robust beamforming, we can extend the formulation of the problem to the case of imperfect CSI which leads to more general results. The solution of the robust design problem with probabilistic constraints is far from trivial and may require a totally different solution approach. While this aspect falls beyond the scope of this thesis, it nevertheless remains an interesting avenue for future work.

In Chapter 4, the proposed constrained  $K$ -means user clustering and beamforming approach for MIMO SCMA in C-RAN was developed under the assumption of perfect CSI. The CSI collected at the central processor in a C-RAN is subject to various sources of imperfections such as the estimation errors at the receiver side, the quantization errors due to the finite-rate feedback and the delay of feedback. Therefore, it is desirable to take into the consideration the effects of CSI error in the design of user clustering and beamforming. As shown in the simulation results, the proposed design is sensitive to the CSI accuracy owing to the similarity metric for the proposed constrained  $K$ -means algorithm and the BD algorithm for beamforming. In order to enhance the performance of the proposed user clustering and downlink beamforming in the presence of imperfect CSI, one can use a more sophisticated similarity metric in the clustering algorithm. Another possible extension of the works in Chapter 4 may be the consideration of robust beamforming with different CSI error models, e.g., statistical (Gaussian), norm-bounded, and random vector quantization.

In Chapter 5, we focused on the design of SCMA decoders and developed a novel decoding

algorithm based on ResNet. It may be desirable to also consider the design of an encoder, which may bring additional challenges. Appropriate codebooks, e.g., sparse codes with only a small number of non-zero elements or low-correlation spreading codes, need to be designed to efficiently reduce the inter-user interference when all the codewords are multiplexed over the same wideband. It may be useful to exploit the possibility of applying deep learning based approaches for encoder design to reduce decoding complexity while improving robustness of user's signal against various types of interference.

# Appendix A

Below, we provide the proofs of key theorems stated in Chapter 3.

## A.1 Proof of Theorem 3.1

We first provide the following lemma which is a basic extension of [77, Proposition 2].

Lemma A.1: Given any interval  $[\underline{\varphi}_{c,m}, \bar{\varphi}_{c,m}]$ , with  $\bar{\varphi}_{c,m} - \underline{\varphi}_{c,m} \leq \pi$ . For all  $\beta_{c,m} \in \text{Conv}(\mathcal{D}_{[\underline{\varphi}_{c,m}, \bar{\varphi}_{c,m}]})$ , we have,

$$\frac{|\beta_{c,m}|^2}{a_{k,m}^c} \geq \cos^2\left(\frac{\bar{\varphi}_{c,m} - \underline{\varphi}_{c,m}}{2}\right) \quad (\text{A.1})$$

We omit its proof here for brevity. For any  $\mathcal{Q} \subseteq \mathcal{Q}_{\text{init}}$ , assume that  $\mathcal{Q} = [\mathbf{A}, \mathbf{B}]$ . Since  $\text{size}(\mathcal{Q}) \leq 2\delta$ , we have  $\max_{\forall i,j} b_{i,j} - a_{i,j} \leq 2\delta$ , where  $a_{i,j}$  and  $b_{i,j}$  denote  $(i, j)$ th entries of  $\mathbf{A}$  and  $\mathbf{B}$  respectively. Hence, we can use (3.31) to obtain,

$$\frac{|\beta_{c,m}|^2}{a_{k,m}^c} \geq \frac{1}{1 + \epsilon}. \quad (\text{A.2})$$

Then, it follows from (3.28) and (3.36) that the scaled feasible solution  $\tilde{\mathbf{W}}_c$  satisfies,

$$\|\tilde{\mathbf{W}}_c\|_2^2 \leq \|\mathbf{W}^*\|_2^2(1 + \epsilon). \quad (\text{A.3})$$

Moreover, since  $\Phi_U^t$  is the objective value at the best known feasible solution at the  $t$ th iteration, we get,

$$\Phi_U^t \leq \sum_c \|\mathbf{W}^*\|_2^2(1 + \epsilon). \quad (\text{A.4})$$

Using the fact that  $\Phi_L^t = \sum_c \|\mathbf{W}^*\|_2^2$ , we have,

$$\frac{\Phi_U^t - \Phi_L^t}{\Phi_L^t} \leq \frac{\sum_c \|\mathbf{W}^*\|_2^2(1 + \epsilon) - \sum_c \|\mathbf{W}^*\|_2^2}{\sum_c \|\mathbf{W}^*\|_2^2} \leq \epsilon. \quad (\text{A.5})$$

The proof is completed.

## A.2 Proof of Theorem 3.2

We prove Theorem 3.2 based on the contradiction principle. Suppose that Algorithm 3.1 does not terminate within  $T_B$  iterations. Then, according to Theorem 3.1, we conclude that the selected box at the  $t$ th iteration satisfies  $\text{size}(\mathcal{Q}^*) \geq 2\delta$  for all  $t = 1, 2, \dots, T_B$ . If the longest edge chosen to be split satisfies  $j^* > K$ , then, after the splitting, the width of the longest edge of the two boxes  $\mathcal{Q}_2^*$  and  $\mathcal{Q}_2^*$  is greater than  $\delta$ . Similarly, for each box  $\mathcal{Q}$  partitioned from the original box  $\mathcal{Q}_{\text{init}}$ , there holds  $b_{i,j} - a_{i,j} \geq \delta$  for all  $j > K$ . Hence, the volume of each box  $\mathcal{Q}$  is not less than  $\delta^{2CK}$ . Note that due to the binary nature of the variable  $\boldsymbol{\nu}$ , the volume of a box  $\mathcal{Q}$  is calculated without taking the variable  $\boldsymbol{\nu}$  into account. If the longest edge satisfies

---

$j^* \leq K$ , we get two boxes with the same volume after the splitting. At the  $T_B$  iteration, we have  $T_B$  boxes such that the total volume of all boxes is not less than  $T_B \delta^{2CK}$ . Obviously, the volume of  $\mathcal{Q}_{\text{init}}$  is  $(4\pi)^{CK}$ . By the choice of  $T_B$ , we get  $T_B \delta^{2CK} > (4\pi)^{CK}$ , which implies that the total volume of all  $T_B$  boxes is greater than that of the original box  $\mathcal{Q}_{\text{init}}$ . This is in contradiction of splitting rules (3.26)-(3.27). Hence, the algorithm will terminate within at most  $T_B$  iterations.

# Appendix B

Below, we provide the proof of the proposition stated in Chapter 4.

## B.1 Proof of Proposition 4.2

In order to prove proposition 4.2, we first transform the cluster assignment subproblem in Algorithm 4.1 into its equivalent form as a minimum cost flow (MCF) linear network optimization problem. We then show that the optimal selection variable  $\iota_{j,k}$  is binary, which can be found using fast network simplex algorithms instead of complex mixed integer linear programming [121].

In general, a MCF problem has an underlying directed graph structure  $G = (\mathcal{V}, \mathcal{E})$  defined by a set of vertices (nodes),  $\mathcal{V}$ , and a set of edges (arcs),  $\mathcal{E}$ . For each node  $\nu \in \mathcal{V}$ , we associate a value  $b(\nu)$  indicating whether it is a supply node ( $b(\nu) > 0$ ), a demand node ( $b(\nu) < 0$ ), or a transshipment node ( $b(\nu) = 0$ ). For each edge  $(\nu, \omega) \in \mathcal{E}$ , we associate a flow of  $f(\nu, \omega)$  on the edge with cost of  $c(\nu, \omega)$  per unit flow. The optimization model for the MCF problem can be formulated as,

$$\min \sum_{(\nu, \omega) \in \mathcal{E}} f(\nu, \omega) c(\nu, \omega) \tag{B.1a}$$

$$\text{s.t. } \sum_{\omega} f(\nu, \omega) - \sum_{\nu} f(\omega, \nu) = b(\nu), \quad \forall \nu \in \mathcal{V} \quad (\text{B.1b})$$

$$0 \leq f(\nu, \omega) \leq u(\nu, \omega), \quad \forall (\nu, \omega) \in \mathcal{E} \quad (\text{B.1c})$$

where  $u(\nu, \omega)$  is the maximum capacity of flow on the edge  $(\nu, \omega) \in \mathcal{E}$ . The problem is feasible if the sum of the supplies equals the sum of the demands, i.e.,

$$\sum_{\nu \in \mathcal{V}} b(\nu) = 0. \quad (\text{B.2})$$

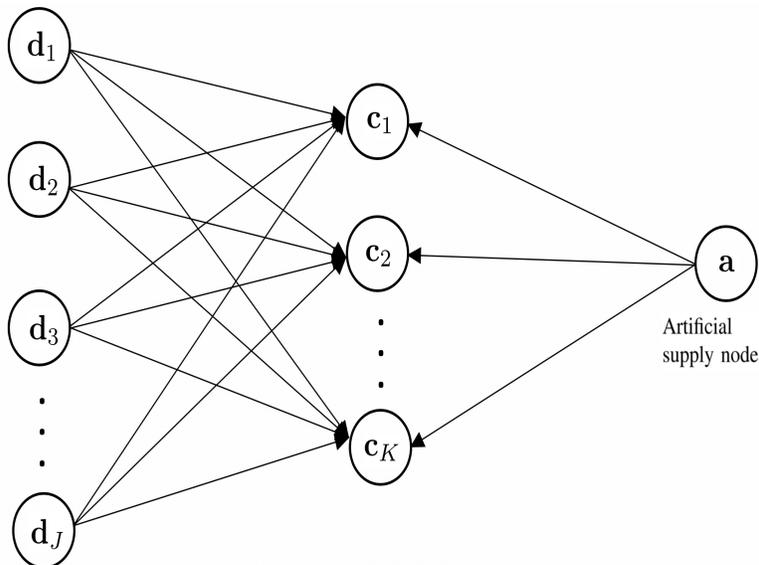
Let each data point  $\mathbf{d}_j$  correspond to a supply node with  $b(\mathbf{d}_j) = 1$  and each cluster center  $\mathbf{c}_k$  correspond to a demand node with  $b(\mathbf{c}_k) = -q$ . The cost of the edge  $(\mathbf{d}_j, \mathbf{c}_k)$  can be expressed as,

$$c(\mathbf{d}_j, \mathbf{c}_k) = \|\mathbf{d}_j - \mathbf{c}_k\|_2^2. \quad (\text{B.3})$$

To satisfy the feasibility constraint of the problem, we consider an artificial supply node,  $\mathbf{a}$ , such that,

$$b(\mathbf{a}) = -J + Kq. \quad (\text{B.4})$$

This artificial node has no edge to or from data points, while the cost of edge from node  $\mathbf{a}$  to cluster center  $\mathbf{c}_k$  is zero, i.e.  $c(\mathbf{a}, \mathbf{c}_k) = 0 \forall k \in \mathcal{K}$ . These identifications establish the equivalence between the MCF and the cluster assignment subproblem in Algorithm 4.1 in which the selection variable  $\iota_{j,k}$  corresponds to flow  $f(\mathbf{d}_j, \mathbf{c}_k)$ . The MCF equivalent directed graph structure is shown in Fig. B.1.



**Figure B.1:** The MCF equivalent directed graph structure.

According to [121, Proposition 5.4], since  $b(\mathbf{d}_j)$ ,  $b(\mathbf{c}_k)$ , and  $b(\mathbf{a})$  are all integers, the optimal flow solution is integer-valued. Since the selection variable  $\boldsymbol{\nu}_{j,k}$  corresponds to flow  $f(\mathbf{d}_j, \mathbf{c}_k)$ , and since  $\sum_k f(\mathbf{d}_j, \mathbf{c}_k) = 1$ , the optimal  $\boldsymbol{\nu}_{j,k}$  is integer with maximum value equal to 1, i.e.  $\boldsymbol{\nu}_{j,k} \in \{0, 1\}$ .

The MCF formulation allows one to solve the cluster assignment subproblem via network simplex algorithm which is faster than general linear programming codes. Specifically, the complexity of solving cluster assignment subproblem via network simplex algorithm is given by [121],

$$O(|\mathcal{V}||\mathcal{E}|^2(\log(|\mathcal{V}|))^2) \quad (\text{B.5})$$

where the number of vertices  $|\mathcal{V}|$ , and number of edges  $|\mathcal{E}|$  in our case are,

$$|\mathcal{V}| = J + K + 1, \tag{B.6}$$

$$|\mathcal{E}| = JK + K. \tag{B.7}$$

It is of interest to investigate the asymptotic complexity of the algorithms when  $J$  and  $K$  are large, i.e., when we let  $J > K \rightarrow \infty$ . Under this condition, we can obtain the asymptotic complexity as,

$$C \triangleq O(J^3 K^2 (\log(J))^2). \tag{B.8}$$

# Bibliography

- [1] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang. “Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends,” *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sept. 2015.
- [2] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi. “Non-orthogonal multiple access (NOMA) for cellular future radio access,” in *IEEE Vehicular Technology Conf. (VTC Spring)*, Dresden, Germany, June 2013, pp. 1–5.
- [3] W. U. Khan, J. Liu, F. Jameel, V. Sharma, R. Jäntti, and Z. Han. “Spectral efficiency optimization for next generation NOMA-enabled IoT networks,” *IEEE Trans. on Vehicular Technology*, vol. 69, no. 12, pp. 15284–15297, Nov. 2020.
- [4] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann. “Modulation and multiple access for 5G networks,” *IEEE Commun. Surveys and Tutorials*, vol. 20, no. 1, pp. 629–646, 2018.
- [5] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. “What will 5G be?,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

- 
- [6] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed. “An overview of signal processing techniques for millimeter wave MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [7] Z. Pi and F. Khan. “An introduction to millimeter-wave mobile broadband systems,” *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
- [8] W. Roh, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar. “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.
- [9] M. Kamel, W. Hamouda, and A. Youssef. “Ultra-dense networks: A survey,” *IEEE Commun. Surveys and Tutorials*, vol. 18, no. 4, 2522–2545, Fourth Quarter 2016.
- [10] B. Makki, K. Chitti, A. Behravan, and M. S. Alouini. “A survey of NOMA: Current status and open research challenges,” *IEEE Open J. of the Commun. Society*, pp. 179–189, Jan. 2020.
- [11] Y. Yuan, S. Wang, Y. Wu, H. V. Poor, Z. Ding, X. You, L. Hanzo. “NOMA for next-generation massive IoT: Performance potential and technology directions,” *IEEE Communications Magazine*, vol. 59, no. 7, pp. 115–121, Apr. 2021.
- [12] M. Zhao, Y. Cai, M. Zhao, B. Champagne, and T. Tsiftsis. “Improving caching efficiency in content aware CRAN-based cooperative beamforming: A joint design approach,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4125–4140, Jun. 2020.

- [13] Y. Liu, Z. Qin, M. ElKashlan, Z. Ding, A. Nallanathan, and L. Hanzo. “Nonorthogonal multiple access for 5G and beyond,” *Proc. of the IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.
- [14] M. Vameghestahbanati, I. Marsland, R. H. Gohary, and H. Yanikomeroglu. “Multidimensional constellations for uplink SCMA systems—A comparative study,” *IEEE Commun. Surveys and Tutorials*, vol. 21, no. 3, pp. 2169-2194, April 2019.
- [15] X. Dai, Z. Zhang, B. Bai, S. Chen, and S. Sun. “Pattern division multiple access: A new multiple access technology for 5G,” *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 54-60, April 2018.
- [16] M. Qiu, Y. C. Huang, S. L. Shieh, and J. Yuan. “A lattice-partition framework of downlink non-orthogonal multiple access without SIC,” *IEEE Trans. on Commun.*, vol. 66, no. 6, pp. 2532-2546, Feb. 2018.
- [17] S. M. R. Islam, M. Zeng, and O. A. Dobre. “NOMA in 5G systems: Exciting possibilities for enhancing spectral efficiency,” *IEEE 5G Tech Focus*, vol. 1, no. 2, pp. 1-6, Jun. 2017.
- [18] H. Nikopour and H. Baligh. “Sparse code multiple access,” in *Personal, Indoor, and Mobile Radio Communications (PIMRC)*, London, UK, Sept. 2013, pp. 332–336.
- [19] P. Patel and J. Holtzman. “Analysis of a simple successive interference cancellation scheme in a ds/cdma system,” *IEEE J. Sel. Areas Commun.*, vol. 12, no. 5, pp. 796–807, June 1994.
- [20] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi. “Superposition coding strategies: Design and experimental evaluation,” *IEEE Trans. on Commun.*, vol. 11, no. 7, pp. 2628–2639, Jul. 2012.

- [21] A. Krishnamoorthy and R. Schober. “Uplink and downlink MIMO-NOMA with simultaneous triangularization,” *IEEE Trans. Wireless Commun.*, early access, Jan. 2021, doi: 10.1109/TWC.2021.3049594.
- [22] A. Krishnamoorthy, Z. Ding, and R. Schober. “Precoder design and statistical power allocation for MIMO-NOMA via user-assisted simultaneous diagonalization,” *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 929–945, Feb. 2021.
- [23] Z. Ding, P. Fan, and H. V. Poor. “Random beamforming in millimeter-wave NOMA networks,” *IEEE Access*, vol. 5, pp. 7667–7681, Feb. 2017.
- [24] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan. “Optimal user scheduling and power allocation for millimeter wave NOMA systems,” *IEEE Trans. on Wireless Commun.*, vol. 17, no. 3, pp. 1502–1517, Dec. 2017.
- [25] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan. “A multi-beam NOMA framework for hybrid mmWave systems,” *IEEE Int. Conf. on Commun. (ICC)*, pp. 1–7, Kansas City, USA, May 2018.
- [26] J. Cui, Z. Ding, and P. Fan. “The application of machine learning in mmWave-NOMA systems,” *IEEE Vehicular Technology Conf. (VTC Spring)*, pp. 1–6, Porto, Portugal, June 2018.
- [27] M. A. Almasi, M. Vaezi, and H. Mehrpouyan. “Impact of beam misalignment on hybrid beamforming NOMA for mmWave communications,” *IEEE Trans. on Commun.*, vol. 67, no. 6, pp. 4505–4518, March 2019.

- [28] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D.O. Wu, and X.G. Xia. "Millimeter-Wave NOMA with user grouping, power allocation and hybrid beamforming," *IEEE Trans. on Wireless Commun.*, vol. 18, no. 11, pp. 5065-5079, Aug. 2019.
- [29] Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong. "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. on Signal Processing*, vol. 65, no. 18, pp. 4874-4886, June 2017.
- [30] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.G. Xia. "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 5, pp. 2961-2974, Feb. 2018.
- [31] Y. Yuan, P. Xu, Z. Yang, Z. Ding, and Q. Chen. "Joint robust beamforming and power-splitting ratio design in SWIPT-based cooperative NOMA systems with CSI uncertainty," *IEEE Trans. Vehicular Technology*, vol. 68, no. 3, pp. 2386-2400, Jan. 2019.
- [32] K. Cao, B. Wang, H. Ding, T. Li, J. Tian, F. Gong. "Secure transmission designs for NOMA systems against internal and external eavesdropping," *IEEE Trans. on Information Forensics and Security*, vol. 15, pp. 2930-2943, Mar. 2020.
- [33] C. Xue, Q. Zhang, Q. Li, and J. Qin. "Joint power allocation and relay beamforming in nonorthogonal multiple access amplify-and-forward relay networks," *IEEE Trans. on Vehicular Technology*, vol. 66, no. 8, pp. 7558-7562, Jan. 2017.
- [34] X. Sun, N. Yang, S. Yan, Z. Ding, D. W. K. Ng, C. Shen, and Z. Zhong. "Joint beamforming and power allocation in downlink NOMA multiuser MIMO networks," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 8, pp. 5367-5381, June 2018.

- [35] Z. Lin, M. Lin, J. B. Wang, T. De Cola, and J Wang. “Joint beamforming and power allocation for satellite-terrestrial integrated networks with non-orthogonal multiple access,” *IEEE J. of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 657-670, Feb. 2019.
- [36] X. Ding, Y. Chen, and Y. Lu. “Joint user ordering, beamforming and power allocation for downlink MIMO-NOMA systems,” *Int. Symposium on Commun. Systems, Networks and Digital Signal Processing (CSNDSP)*, pp. 1-6, Porto, Portugal, July 2020.
- [37] Y. Xiu, J. Zhao, W. Sun, M. Di Renzo, G. Gui, Z. Zhang, and N. Wei. “Reconfigurable intelligent surfaces aided mmwave NOMA: Joint power allocation, phase shifts, and hybrid beamforming optimization,” *IEEE Trans. on Wireless Commun.*, vol. 20, no. 12, pp. 8393-8409, Jul. 2021.
- [38] T. Lu, H. Zhang, and K. Long. “Joint beamforming and power control for MIMO-NOMA with deep reinforcement learning,” *IEEE Int. Conf. on Commun. (ICC)*, pp. 1-5, Montreal, Canada, June 2021.
- [39] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh. “SCMA codebook design,” *IEEE Vehicular Technology Conf. (VTC-Fall)*, pp. 1-5, Vancouver, Canada, Sept. 2014.
- [40] H. Nikopour, E. Yi, A. Bayesteh, K. Au, M. Hawryluck, H. Baligh, and J. Ma. “SCMA for downlink multiple access of 5G wireless networks,” *Proc. IEEE Global Commun. Conf. (Globecom)*, pp. 3940-394, Austin, USA, Dec. 2014.
- [41] M. Alam, and Q. Zhang. “Designing optimum mother constellation and codebooks for SCMA,” *IEEE Int. Conf. on Commun. (ICC)*, pp. 1-6, Paris, France, May 2017.

- [42] L. Yu, P. Fan, X. Lei, and P. T. Mathiopoulos. “BER analysis of SCMA systems with codebooks based on star-QAM signaling constellations,” *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 1925–1928, Sept. 2017.
- [43] J. Bao, Z. Ma, M. Xiao, T. A. Tsiftsis, and Z. Zhu. “Bit-interleaved coded SCMA with iterative multiuser detection: Multidimensional constellations design,” *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5292-5304, Dec. 2017.
- [44] L. Yu, P. Fan, D. Cai, Z. Ma. “Design and analysis of SCMA codebook based on star-QAM signaling constellations,” *IEEE Trans. Vehicular Technology*, vol. 67, no. 11, pp. 10543-10553, Aug. 2018.
- [45] Q. Wang, T. Li, R. Feng, and C. Yang. “An efficient large resource-user scale SCMA codebook design method,” *IEEE Commun. Lett.*, vol. 23, no. 10, pp. 1787-1790, Jul. 2019.
- [46] C. Jiang, and Y. Wang. “An uplink SCMA codebook design combining probabilistic shaping and geometric shaping,” *IEEE Access*, vol. 8, pp. 76726-76736, Apr. 2020.
- [47] M. Gao, W. Ge, P. Zhang, and Y. Zhang. “An efficient codebook design for uplink SCMA,” *IEEE Access*, vol. 8, pp. 11665-211675, Nov. 2020.
- [48] Z. Hou, Z. Xiang, P. Ren, and B. Cao. “SCMA codebook design based on divided extended mother codebook,” *IEEE Access*, vol. 9, pp. 71563-71576, May 2021.
- [49] X. Zhang, D. Zhang, L. Yang, G. Han, H. H. Chen, and D. i. Zhang. “SCMA codebook design based on uniquely decomposable constellation groups,” *IEEE Trans. on Wireless Commun.*, vol. 20, no. 8, pp. 4828-4842, Mar. 2021.

- [50] C. Huang, B. Su, T. Lin, and Y. Huang. “Downlink SCMA codebook design with low error rate by maximizing minimum Euclidean distance of superimposed codewords,” *IEEE Trans. Vehicular Technology*, vol. 71, no. 5, pp. 5231-5245, Mar. 2022.
- [51] L. Yang, Y. Liu, and Y. Siu. “Low complexity message passing algorithm for SCMA system,” *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2466-2469, Sept. 2016.
- [52] M. Vameghestahbanati, E. Bedeer, I. Marsland, R. H. Gohary, and H. Yanikomeroglu. “Enabling sphere decoding for SCMA,” *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2750-2753, Aug. 2017.
- [53] F. Wei, W. Chen, Y. Wu, J. Ma, and T. A. Tsiftsis. “Message-passing receiver design for joint channel estimation and data decoding in uplink grant-free SCMA systems,” *IEEE Trans. on Wireless Commun.*, vol. 18, no. 1, pp. 167-181, Nov. 2018.
- [54] J. Jiao, K. Liang, B. Feng, Y. Wang, S. Wu, and Q. Zhang. “Joint channel estimation and decoding for polar coded SCMA system over fading channels,” *IEEE Trans. on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 210-221, Apr. 2020.
- [55] L. Xiang, Y. Liu, C. Xu, R. G. Maunder, L. L. Yang, and L. Hanzo. “Iterative receiver design for polar-coded SCMA systems,” *IEEE Trans. on Commun.*, vol. 69, no. 7, pp. 4235-4246, Mar. 2021.
- [56] Y. Du, B. Dong, Z. Chen, P. Gao, and J. Fang. “Joint sparse graph-detector design for downlink MIMO-SCMA systems,” *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 14-17, Nov. 2016.

- [57] Z. Wu, C. Zhang, X. Shen, and H. Jiao. "Low complexity uplink SFBC-based MIMO-SCMA joint decoding algorithm," *IEEE Int. Conf. on Computer and Commun. (ICCC)*, pp. 968-972, Chengdu, China, Dec. 2017.
- [58] C. Yan, N. Zhang, and G. Kang. "Downlink multiple input multiple output mixed sparse code multiple access for 5G system," *IEEE Access*, vol. 6, pp. 20837-20847, Apr. 2018.
- [59] W. Yuan, N. Wu, Q. Guo, Y. Li, C. Xing, and J. Kuang. "Iterative receivers for downlink MIMO-SCMA: Message passing and distributed cooperative detection," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3444-3458, June 2018.
- [60] H. Huang, S. Guo, G. Gui, Z. Yang, J. Zhang, H. Sari, and F. Adachi. "Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 214-222, Aug. 2019.
- [61] M. Kim, N. Kim, W. Lee, and D. Cho. "Deep learning-aided SCMA." *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 720-723, Apr. 2018.
- [62] C. Lu, W. Xu, H. Shen, H. Zhang, and X. You. "An enhanced SCMA detector enabled by deep neural network," *IEEE Int. Conf. Commun. in China (ICCC)*, pp. 835-839, Beijing, China, Aug. 2018.
- [63] J. Lin, S. Feng, Z. Yang, Y. Zhang, and Y. Zhang. "A novel deep neural network based approach for sparse code multiple access." *Neurocomputing*, vol. 382, pp. 52-63, Mar. 2019.
- [64] C. P. Wei, H. Yang, C. P. Li, and Y. M. Chen. "SCMA decoding via deep learning." *IEEE Commun. Lett.*, vol. 10, no. 4, pp. 878-881, Dec. 2020.

- [65] H. Cheng, Y. Xia, Y. Huang, Z. Lu, L. Yang . “Deep neural network aided low-complexity MPA receivers for uplink SCMA systems.” *IEEE Trans. on Vehicular Technology*, vol. 70, no. 9, pp. 9050-9062, Jul. 2021.
- [66] B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, G. Caire. “Is NOMA efficient in multi-antenna networks? A critical look at next generation multiple access techniques,” *IEEE Open J. of the Commun. Society*, vol. 10, no. 2, pp. 1310-1343, Jun. 2021.
- [67] F.J. Martin-Vega, Y. Liu, G. Gomez, M.C. Aguayo-Torres, and M. ElKashlan. “Modeling and analysis of NOMA enabled CRAN with cluster point process,” *Proc. IEEE Global Commun. Conf. (Globecom)*, pp. 1-6, Singapore, Singapore, Dec. 2017.
- [68] X. Gu, X. Ji, Z. Ding, W. Wu, and M. Peng. “Outage probability analysis of nonorthogonal multiple access in cloud radio access networks,” *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 149-152, Jan. 2018.
- [69] J. Zhao, Y. Liu, T. Mahmoodi, K. K. Chai, Y. Chen, and Z. Han. “Resource allocation in cache-enabled CRAN with non-orthogonal multiple access,” *IEEE Int. Conf. Commun. (ICC)*, pp. 1 - 6, Kansas City, USA, May 2018.
- [70] M. Moltafet, S. Parsaeefard, M. R. Javan, and N. Mokari. “Robust radio resource allocation in MISO-SCMA assisted C-RAN in 5G networks,” *IEEE Trans. on Vehicular Technology*, vol. 68, no. 6, pp. 5758 - 5768, Apr. 2019.
- [71] S. Norouzi, Y. Cai and B. Champagne. “Joint design of user clustering, beamforming, and power allocation for NOMA,” *IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct. 2022.

- [72] S. Norouzi, B. Champagne, and Y. Cai. “Joint optimization framework for user clustering, downlink beamforming, and power allocation in MIMO NOMA systems,” *IEEE Transactions on Communications*, Nov. 2022. (accepted for publication)
- [73] S. Burer, and A. N. Letchford. “Non-convex mixed-integer nonlinear programming: A survey,” *Surveys in Operations Research and Management Science*, vol. 17, no. 2, pp. 97-106, 2012.
- [74] M. Tawarmalani, N. V. Sahinidis, and N. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*, Springer Science and Business Media, 2002.
- [75] C. Lu and Y. F. Liu. “An efficient global algorithm for single-group multicast beamforming,” *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3761–3774, Jul. 2017.
- [76] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [77] S. Boyd and J. Mattingley. “*Branch and Bounds Methods*.” EE364b course notes, Stanford University, May 2011.
- [78] M. Grant and S. Boyd. “CVX: Matlab software for disciplined convex programming, Version 2.1,” Mar. 2014.
- [79] M. R. Hestenes. “Multiplier and gradient methods,” *J. Optim. Theory Appl.*, vol. 4, no. 5, pp. 303–320, Nov. 1969.
- [80] M. J. D. Powells. “A method for nonlinear constraints in minimization problems,” in *Optimization.*, R. Fletcher, Ed. New York, NY, USA: Academic, 1972, pp. 283–298.

- [81] M. M. Zhao, Q. Shi, Y. Cai, M. J. Zhao, and Q. Yu. “Decoding binary linear codes using penalty dual decomposition method,” *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 958–962, Jun. 2019.
- [82] G. R. Lanckriet, and B. K. Sriperumbudur. “On the convergence of the concave-convex procedure,” *Advances in Neural Information Processing Systems*, 2009, pp. 1759-1767.
- [83] Q. Shi and M. Hong. “Penalty dual decomposition method with application in signal processing,” *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, New Orleans, LA, USA, Mar. 2017, pp. 4059–4063.
- [84] Q. Shi and M. Hong. “Penalty dual decomposition method for nonsmooth nonconvex optimization—Part I: Algorithms and convergence analysis,” *IEEE Trans. on Signal Processing*, vol. 68, pp. 4108-4122, Jun. 2020.
- [85] B. D. Van Veen and K. M. Buckley. “Beamforming: a versatile approach to spatial filtering,” *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4-24, Apr. 1988.
- [86] H. L. Van Trees. “*Optimum array processing: Part IV of detection, estimation, and modulation theory*,” John Wiley & Sons; Apr. 2002.
- [87] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir. “Unsupervised machine learning-based user clustering in millimeter-wave-NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7425-7440, Sept. 2018.
- [88] F. Cui, Z. Qin, Y. Cai, M. Zhao, and G. Y. Li. “Rethinking outage constraints for resource management of NOMA networks,” *IEEE J. on Selected Topics in Signal Processing.*, vol. 13, no. 3, pp. 423-435, Jun. 2019.

- [89] Y. Yuan, P. Xu, Z. Yang, Z. Ding, and Q. Chen. “Joint robust beamforming and power-splitting ratio design in SWIPT-based cooperative NOMA systems with CSI uncertainty,” *IEEE Trans. on Vehicular Technology.*, vol. 68, no. 3, pp. 2386-240, Jan. 2019.
- [90] Z. Wang, H. Kong, S. Huang, M. Cheng, W. P. Zhu. “Beamforming and power allocation in NOMA-based multibeam satellite systems with outage constrain,” *IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1-6, Madrid, Spain, Dec 2021.
- [91] S. Norouzi, Y. Cai and B. Champagne. “Constrained  $K$ -means user clustering and downlink beamforming in MIMO-SCMA systems,” *IEEE 32nd Annual International Symposium on PIMRC*, pp. 1091-1096, Oct. 2021.
- [92] S. Norouzi, Y. Cai, and B. Champagne. “Energy-efficient user clustering and downlink beamforming for MIMO-SCMA in C-RAN,” *IEEE Access*, vol. 9, pp. 115175-115191, Aug. 2021.
- [93] J. Choi. “Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems,” *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, March 2015.
- [94] S. Norouzi, A. Morsali, and B. Champagne. “Optimizing transmission rate in NOMA via block diagonalization beamforming and power allocation,” *IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing (PACRIM)*, Victoria, Canada, Aug. 2019, pp. 1-5.
- [95] A.L. Yuille, and A. Rangarajan. “The concave-convex procedure,” *Neural Computation*, vol. 15, no. 4, pp. 915-936, April 2003.

- [96] J. Cui, Z. Ding, and P. Fan. “The application of machine learning in mmWave-NOMA systems,” *IEEE 87th Vehicular Technology Conf. (VTC Spring)*, pp. 1 - 6, Porto, Portugal, June 2018.
- [97] G. Lee, Y. Sung, and J. Seo. “Randomly-directional beamforming in millimeter-wave multiuser MISO downlink,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1086-1100, Sept. 2015.
- [98] A. Alkhateeb, G. Leus, and R. W. Heath. “Limited feedback hybrid precoding for multi-user millimeter wave systems,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481-6494, Jul. 2015.
- [99] T. Niknam, E. T. Fard, N. Pourjafarian, and A. Roustaei. “An efficient hybrid algorithm based on modified imperialist competitive algorithm and  $K$ -means for data clustering,” *Engineering Applications of Artificial Intelligence*, vol. 24, no. 2, pp. 306-317, Mar. 2011.
- [100] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai. “Structured massive access for scalable cell-free massive MIMO systems,” *IEEE J. on Selected Areas in Communications*, vol. 39, no. 4, pp. 1086-1100, Aug. 2020.
- [101] A. K. Jain. “Data clustering: 50 years beyond  $K$ -means,” *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651-666, Jun. 2010.
- [102] P. S. Bradley, O. L. Mangasarian, and W. N. Street. “Clustering via concave minimization,” *Advances in Neural Information Processing Systems*, pp. 368-374, 1997.
- [103] P. S. Bradley, K. P. Bennett, A. Demiriz. “Constrained  $K$ -means clustering,” *Technical Report Microsoft Research Redmond, WA, USA*, May 2000.

- [104] M. Inaba, N. Katoh, and H. Imai. “Applications of weighted Voronoi diagrams and randomization to variance-based  $K$ -clustering: (extended abstract),” *Proc. of the Tenth Annual Symposium on Computational Geometry*, New York, USA, 1994.
- [105] N. G. Andrew. “*Clustering with K-means algorithm*,” Machine Learning, 2012.
- [106] P. J. Rousseeuw. “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis,” *J. of Computational and Applied Mathematics*, vol. 20, no. 1, pp. 53-65, Jan. 1987.
- [107] R. Tibshirani, G. Walther, and T. Hastie. “Stimating the number of clusters in a data set via the gap statistic,” *J. of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 411–42, 2001.
- [108] P. Luong, F. Gagnon, C. Despins, and L. N. Tran. “Optimal joint remote radio head selection and beamforming design for limited fronthaul C-RAN,” *IEEE Trans. Signal Processing*, vol. 65, no. 21, pp. 5605-5620, Nov. 2017.
- [109] Y. Cheng, M. Pesavento, and A. Philipp. “Joint network optimization and downlink beamforming for CoMP transmission using mixed integer conic programming,” *IEEE Trans. Signal Processing*, vol. 61, no. 16, pp. 3972–3987, Aug. 2013.
- [110] M. Tawarmalani, and N.V. Sahinidis. “Global optimization of mixed-integer nonlinear programs: A theoretical and computational study,” *Mathematical Programming*, vol. 99, no. 3, pp. 563-591, Apr. 2004.
- [111] MOSEK ApS, *The MOSEK Optimization Toolbox for MATLAB Manual*. Version 9.0., 2019. Available: <http://docs.mosek.com/9.0/toolbox/index.html>

- [112] Gurobi Optimization, LLC, *Gurobi Optimizer Reference Manual*. 2021. Available: <https://www.gurobi.com>
- [113] X. Zhu, Z. Wang, C. Qian, L. Dai, J.i Chen, S. Chen, and L. Hanzo. “Soft pilot reuse and multicell block diagonalization precoding for massive MIMO systems,” *IEEE Trans. Vehicular Technology*. vol. 65, no. 5, pp. 3285-3298, May 2016.
- [114] B. Dai, W. Yu. “Sparse beamforming and user-centric clustering for downlink cloud radio access network,” *IEEE Access*. vol. 2, pp. 1326-1339, Oct. 2014.
- [115] E. Chen, M. Tao, and Y. F. Liu. “Joint base station clustering and beamforming for non-orthogonal multicast and unicast transmission with backhaul constraints,” *IEEE Trans. Wireless Commun.* vol. 17, no. 9, pp. 6265-6279, Jul. 2018.
- [116] B. Dai, and W. Yu. “Energy efficiency of downlink transmission strategies for cloud radio access networks,” *IEEE J. on Selected Areas in Communications*. vol. 34, no. 4, pp. 1037-1050, Mar. 2016.
- [117] S. Ali, E. Hossain, and D. I. Kim. “Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: User clustering, beamforming, and power allocation,” *IEEE Access*. vol. 5, pp. 565–577, Dec. 2016.
- [118] Y. Teng and W. Zhao. “Robust group sparse beamforming for dense C-RANs with probabilistic SINR constraints,” *IEEE Wireless Communications and Networking Conf. (WCNC)*, pp. 1-6, San Francisco, USA, Mar. 2017.
- [119] Y. Wang, L. Ma, and Y. Xu. “Robust beamforming and base station activation for energy efficient downlink C-RAN,” *IEEE Vehicular Technology Conf. (VTC-Fall)*, pp. 1-5, Toronto, Canada, Sept. 2017.

- [120] D. Yan, R. Wang, E. Liu, and Q. Hou. “Robust beamforming optimization for downlink cloud radio access networks,” *IEEE Global Communications Conf. (GLOBECOM)*, pp. 1-6, Abu Dhabi, United Arab Emirates, Dec. 2018.
- [121] D. P. Bertsekas. “*Network Optimization: Continuous and Discrete Models.*” Belmont, MA: Athena Scientific, 1998.
- [122] S. Norouzi and B. Champagne. “Deep residual neural network decoder for sparse code multiple access,” *IEEE Wireless Communications and Networking Conf. (WCNC)*, Scotland, UK, Mar. 2023.
- [123] L. Bottou. “Stochastic gradient descent tricks.” in *Neural networks: Tricks of the trade*, Berlin, Heidelberg, 2012.
- [124] I. Sutskever, J. Martens, G. Dahl, and G. Hinton. “On the importance of initialization and momentum in deep learning.” in *Proc. Int. Conf. Mach. Learn. Res. (PMLR)*, vol. 23, no. 3, pp. 1139-1147, May 2013.
- [125] D. P. Kingma and J. L. Ba. “Adam: A method for stochastic optimization.” in *Proc. Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015.
- [126] K. He, X. Zhang, S. Ren, and J. Sun. “Deep residual learning for image recognition,” *Proc. IEEE conf. on computer vision and pattern recognition (CVPR)*, pp. 770-778, 2016.
- [127] S. Ioffe, and C. Szegedy. “Batch normalization: Accelerating deep network training by reducing internal covariate shift.” in *Proc. Int. Conf. Mach. Learn. Res. (PMLR)*, vol. 37, pp. 448-456, June 2015.
- [128] Altera Innovate Asia. “*The 1st 5G Algorithm Innovation Competition SCMA.*” [Online]. Available: <http://www.innovateasia.com/5g/en/gp2.html>