A Comprehensive Study of Multiple Access Techniques in 6G Networks

Mohamed Elhattab

A Thesis
in
The Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Electrical and Computer Engineering) at
Concordia University
Montréal, Québec, Canada

December 2022

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Mohamed Elhattab**

Entitled: **A Comprehensive Study of Multiple Access Techniques in 6G Networks**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Electrical and Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

	Chair
Dr. Anjan Bhowmick	
	External Examiner
Dr. Robert Schober	
	Thesis Supervisor
Dr. Chadi Assi	
	Thesis Co-Supervisor
Dr. Ali Ghrayeb	
	Examiner
Dr. Jun Cai	
	Examiner
Dr. Amr Youssef	
	Examiner
Dr. Reza Soleymani	

Approved by

Dr. Jun Cai, Graduate Program Director

4/28/2022

Dr. Mourad Debbabi, Dean
Gina Cody School of Engineering and Computer Science

Abstract

A Comprehensive Study of Multiple Access Techniques in 6G Networks

Mohamed Elhattab, Ph.D.

Concordia University, 2022

With the proliferation of numerous burgeoning services such as ultra-reliable low-latency communication (URLLC), massive machine type communications (mMTC), enhanced mobile broadband (eMBB), among others, wireless communication systems are expected to face daunting challenges. In order to satisfy these ever-increasing traffic demands, diverse quality-of-services (QoS) requirements, and the massive connectivity accompanied by these new applications, various innovative and promising technologies, and architectures need to be developed. Novel multiple-access techniques are currently being explored in both academia and industry in order to accommodate such unprecedented requirements. Non-orthogonal multiple access (NOMA) has been deemed as one of the vital enabling multiple access techniques for the upcoming six-generation (6G) networks. This is due to its ability to enhance network spectral efficiency (NSE) and support a massive number of connected devices. Owing to its potential benefits, NOMA is recognized as a prominent member of next-generation multiple access (NGMA).

Several emerging techniques such as full-duplex (FD) communication, device-to-device (D2D) communications, reconfigurable intelligent surface (RIS), coordinated multipoint (CoMP), cloud radio access networks, are being gradually developed to address fundamental problems in future wireless networks. In this thesis, and with the goal of converging toward NGMA, we investigate the synergistic integration between NOMA and other evolving physical layer technologies. Specifically, we analyze this integration aiming at improving the performance of cell-edge users (CEUs), mitigating the detrimental effect of inter-cell interference (ICI), designing energy-efficient multiple access toward "green" wireless networks, guarantying reliable communication between NOMA UEs and base stations (BSs)/remote radio heads (RRHs), and maintaining the required QoS in terms of the minimum achievable data rate, especially at CEUs.

Regarding the ICI mitigation in multi-cell NOMA networks and tackling the connectiv-

ity issue in traditional CoMP-based OMA networks, we first investigate the integration between location-aware CoMP transmission and NOMA in downlink heterogeneous C-RAN. In doing so, we design a novel analytical framework using tools from stochastic geometry to analyze the system performance in terms of the average achievable data rate per NOMA UE. Our results reveal that CoMP NOMA can provide a significant gain in terms of network spectral efficiency compared to the traditional CoMP OMA scheme. In addition, with the goal of further improving the performance of CEUs and user fairness, cooperative transmission with the aid of D2D communication and FD or half-duplex (HD) transmission, has been introduced to NOMA, which is commonly known as cooperative NOMA (C-NOMA). As a result, we extend our study to also investigate the potential gains of investigating CoMP and C-NOMA. In such a framework, we exploit the cooperation between the RRHs/BSs and the successive decoding strategy at NOMA UEs that are near the RRHs/BSs. Specifically, we investigate both performance analysis and resource management optimization (power control and user pairing). Our results show that the transmit power at the BS, the transmit power at the relay user, and the self-interference (SI) value at the relay user determine which multiple access technique, CoMP NOMA, CoMP HD C-NOMA, and CoMP FD C-NOMA, should be adopted at the BSs.

Now, to assist in designing energy-efficient multiple access techniques and guaranteeing reliable communication for NOMA UEs, this thesis explores the interplay between FD/HD C-NOMA and RIS. We show that the proposed model has the best performance in terms of network power consumption compared to other multiple access techniques in the literature, which leads to "green" future wireless networks. Moreover, our results show that the network power consumption can be significantly reduced by increasing the number of RIS elements. A more significant finding is that the location of the RIS depends on the adopted multiple access techniques. For example, it is not recommended to deploy the RIS besides the BS if the adopted multiple access is HD C-NOMA. Another insight that has been unveiled is the FD C-NOMA with the assistance of RIS has more resistance to the residual SI effect, due to the FD transmission, and can tolerate high SI values compared to the same scheme without RIS.

Although much work has been conducted to improve the network spectral efficiency of multi-cell NOMA cellular networks, the required QoS by the upcoming 6G applications, in terms of the minimum achievable rate, may not be guaranteed at CEUs. This is due to

their distant locations from their serving BSs, and thus, they experience severe path-loss attenuation and high ICI. This thesis addresses this research gap by studying the synergistic integration between RIS, NOMA, and CoMP in a multi-user multi-cell scenario. Unlike the developed high-complexity optimal solutions or the low-complexity sub-optimal solutions in the literature for the power allocation problem, we derive a low-complexity optimal solution in a such challenging scenario. We also consider the interdependency between the user clustering policies in different coordinated cells, which has been ignored in the literature. Finally, we prove that this integration between RIS, NOMA, and CoMP can attain a high achievable rate for CEUs, ameliorate spectral efficiency compared to existing literature, and can form a novel paradigm for NGMA.

Acknowledgements

First and foremost, all praise and thanks are due to Allah (God) the Almighty, the Most Merciful, and the Most Gracious, for giving me the strength to pursue my dreams. This thesis would not have been possible without the faith I have in Him.

I would like to express my special thanks and appreciation to my advisors Prof. Chadi Assi and Prof. Ali Ghrayeb. I could not have reached this day without your extreme encouragement and constant support. I would like to thank Prof. Assi for being such an excellent advisor and for the support you provided me both personally and technically. My academic achievement as a research scientist would not have been possible without your support and insightful advice. I have benefited tremendously from your guidance. Prof. Assi, thank you for all the ethical and technical values I have learned from you. Also, I would like to thank Prof. Ghrayeb. It has been a pleasure working with you and thank you for providing me with the opportunity to grow as a researcher. Your advice on both career and research has been invaluable. Prof. Ghrayeb, thank you for all the writing and research skills I have gained throughout the time I worked with you. Finally, without your valuable guidance, support, and encouragement, I could not have achieved this.

I would also like to thank my committee members, Prof. Reza Soleymani, Prof. Amr Youssef, Prof. Jun Cai, and Prof. Robert Schober for taking part in my Ph.D. dissertation defense committee, the effort and time spent reading my thesis and providing me with valuable suggestions and comments.

I would like to express my gratitude to many institutes for providing scholarships and financial support. I sincerely thank the Fonds de recherche du Quebec – Nature et technologies (FRQNT) for my doctoral scholarship, and Mitacs Globalink Research Award Abroad for my internship funding at the Network sciEnce, Wireless, and Security laboratory at Virginia Tech (NEWS@VT). Finally, thank you Concordia Graduate Student Support Program (GSSP) for funding my Ph.D. studies.

My gratitude also extends to my colleague, Mohamed Amine Arfaoui. I would like to thank him for his collaboration and friendship and wish him the best of luck in his future

endeavors. I am also deeply grateful to my friends and colleagues with whom I was able to share special moments and make memories that will last a lifetime.

I would like to express my deepest gratitude to my beloved parents. It would not have been possible for me to become who I am today without your endless support, wise advice, and prayers. I am also extremely grateful to my sister, Amira, and my brother, Ahmed, for their continuous support and unconditional love throughout my life.

Contents

List of Figures

List of Tables

List of Acronyms

1G	first-generation.
2G	second-generation.
3G	third-generation.
3GPP	Third Generation Partnership Project.
4G	fourth-generation.
5G	fifth-generation.
6G	sixth-generation.
AF	amplify-and-forward.
AO	alternating optimization.
AWGN	additive white Gaussian noise.
B5G	beyond 5G.
BBU	baseband processing unit.
BS	base station.
CB	coordinated beamforming.
CCU	cell-center user.
CDF	cumulative distribution function.
CDMA	code division multiple access.
CEU	cell-edge user.
CJT	coherent JT.
C-NOMA	cooperative NOMA.
CoMP	coordinated multipoint.
C-RAN	cloud radio access network.
CS	coordinated scheduling.
CSI	channel-state-information.
CT	cooperative transmission.
D2D	device-to-device.
DC	difference-of-convex.
DF	decode-and-forward.

DL	downlink.
DT	direct transmission.
EE	energy efficiency.
eMBB	enhanced mobile broadband.
ETSI	European Telecommunications Standards Institute.
FD	full-duplex.
FDMA	frequency division multiple access.
F-RAN	fog radio access network.
GR	Gaussian randomization.
H-CRAN	heterogeneous cloud radio access network.
HD	half-duplex.
HetNet	heterogeneous network.
ICI	intercell interference.
INUI	intra-NOMA-user interference.
IoE	Internet-of-Everything.
IoT	internet of things.
IRS	intelligent reflecting surface.
ISG	Industry Specification Group.
JT	joint-transmission.
KSD	Kolmogorov-Smirnov distance.
LIS	large intelligent surface.
LoS	line-of-sight.
LT	Laplace transform.
LTE	Long Term Evolution.
LTE-A	LTE-advanced.
MA	multiple access.
MINLP	mixed-integer nonlinear programming.
mMIMO	massive multiple-input multiple-output.
mMTC	massive machine-type communication.
MRC	maximum-ratio-combining.

MRRH	macro remote radio head.
NCJT	non-coherent JT.
NGMA	next-generation multiple access.
NLoS	non line-of-sight.
NOMA	non-orthogonal multiple access.
NR	New Radio.
NSE	network spectral-efficiency.
OFDMA	orthogonal frequency division multiple access.
OMA	orthogonal multiple-access.
PA	power allocation.
PDF	probability density function.
PGFL	probability generating functional.
PPP	Poisson Point Process.
PS	phase-shift.
PSD	Positive semidefinite.
QoS	quality-of-service.
RB	resource block.
RF	radio frequency.
RIS	Reconfigurable intelligent surfaces.
RRH	remote radio head.
RSMA	Rate-Splitting Multiple Access.
RSS	received signal strength.
RV	random variables.
SC	superposition coding.
SCA	successive convex approximation.
SDP	semi-definite programming.
SDR	semidefinite relaxation.
SE	spectral-efficiency.
SI	self-interference.
SIC	successive interference cancellation.
SINR	signal-to-interference-plus-noise-ratio.

SNR	signal-to-noise-ratio.
SRE	smart radio environment.
SRRH	small remote radio head.
TDMA	time division multiple access.
UC	user clustering.
UE	user equipment.
UL	uplink.
URLLC	ultra-reliable low-latency communication.

List of Symbols

$\text{rank}(\boldsymbol{M})$	rank of matrix $\boldsymbol{M}$
$\|\boldsymbol{M}\|_*$	nuclear norm of matrix $\boldsymbol{M}$
$\|\boldsymbol{M}\|_F$	Frobenius norm of matrix $\boldsymbol{M}$
$\|\boldsymbol{M}\|_2$	spectral norm of matrix $\boldsymbol{M}$
$\text{tr}(\boldsymbol{M})$	trace of matrix $\boldsymbol{M}$
$\bigotimes$	generalized Cartesian product
$\succ_k$	preference relation of node k
$\text{sgn}(x)$	sign function
$\dot{\sim}$	approximately following this distribution
$\mathbb{E}\{x\}$	Expectation of random variable x
$\text{Var}\{x\}$	Variance of random variable x
σ^2	noise variance
$\mathbb{I}(\cdot)$	indicator function
$[\![a, b]\!]$	discrete interval of integers between a and b, inclusive
$\langle ., . \rangle$	inner product
$\|\boldsymbol{x}\|$, $\|.\|$	Euclidean norm
$\circledast$	circular convolution
$\odot$	element-wise multiplication
$\text{diag}(\boldsymbol{x})$	diagonal matrix that has $\boldsymbol{x}$ as diagonal.
$Ei(.)$	Exponential integral function
$e^{(\cdot)}$, $\exp(.)$	Exponential function

Chapter 1
Introduction

According to the recent Cisco reports [1], by the end of 2023, it is expected that the number of Internet-of-Things (IoT) devices will be 29.3 billion devices, up from 18.4 billion in 2018. This explosive growth of IoT as well as the multitude of the new emerging use-cases and applications such as augmented-reality, collaborative robots, Holographic Telepresence, space and deep-sea tourism, etc., have created unprecedented demands for heterogeneous mobile traffic, massive access, low-latency services, and highly spectral efficient connectivity, which gradually overwhelms the connectivity and the supported capabilities of fifth-generation (5G) cellular network. In addition, according to that report [1], there is no sign that these increasing demands in terms of both the massive connectivity and the high required data rates will slow down in the next decade.

In order to accommodate this explosive traffic demand and with the current development of 5G New Radio (NR) networks, researchers from both industry and academia have started shifting their attention towards the next-generation wireless network, namely "sixth-generation (6G)" [2–5]. "Connecting the unconnected" is an overriding goal of this next generation cellular network that is driving research to provide seamless and ubiquitous connectivity to every device, given a continued exponential growth of Internet-of-Everything (IoE) in the next decade [2–5]. 6G is expected to significantly enhance existing technologies/architectures and/or develop novel techniques at the infrastructure level, the spectrum level, as well as the protocol/algorithmic level to provide better services and realize full-coverage connectivity [2–5]. It is worth mentioning that substantive research programs are being developed on the global scale to shape and establish the vision of 6G including "6G Hubs" in Germany, "6G Flagship" in Finland, Terahertz (THz) communication studies in the USA, "Broadband Communications and New Networks" in China, etc [6].

In order to support different IoT use cases and the new emerging applications requirements, 6G should accommodate a much higher density of connectivity (estimated to be

10^6 devices per km^2), provide 5-10, and 10-100 times of the spectral efficiency and energy efficiency in comparison with the ones that 5G can achieve, 0.9999999 reliability, and support latency in terms of microseconds [6]. Hence, to achieve the aforementioned strict goals, one of the most fundamental issues is to design sophisticated multiple access techniques for the forthcoming wireless networks denoted as next-generation multiple access (NGMA) [6, 7]. According to [6, 7], non-orthogonal transmission strategies that allow different users to utilize the same radio channels in the same time are promising candidates for the NGMA. In this context, non-orthogonal multiple access (NOMA) is considered as a prominent member of the NGMA family and has been recognized as a key multiple access candidate for the 6G networks. As a result, much effort has been directed and focused on exploiting NOMA to develop NGMA.

One potential way to converge toward NGMA is by synergistically integrating NOMA with other beyond 5G (B5G)/6G technologies and architectures such as device-to-device (D2D) communications, reconfigurable intelligent surfaces (RIS), coordinated multipoint (CoMP), full-duplex (FD) communications, cloud radio access network (C-RAN), to name a few. Specifically, in order to mitigate the effect of the intercell interference (ICI) in multi-cell cellular networks scenario, the combination between NOMA and the Third Generation Partnership Project (3GPP) interference mitigation techniques such as CoMP transmission has been widely investigated to attain higher spectral efficiency, especially at the ICI-prone users. Although there is much work in integrating NOMA with CoMP in multi-cell setup, the performance analysis of such a framework is not well investigated in the literature and should be revisited.

Nevertheless, the network spectral efficiency is significantly improved by invoking NOMA compared to OMA. However, the co-existence of near NOMA user equipments (UEs) with the far NOMA UEs may result in performance degradation to the co-channel interference and the obligation of allocating high power to far NOMA UEs. This is due to the fact that the far NOMA users, which are also can be considered as cell-edge users, mainly suffer from intra-NOMA user interference as well as experience severe path loss attenuation. Therefore, in order to tackle this challenge and by leveraging the successive decoding strategy at the strong NOMA users, the incorporation between cooperative communications and NOMA with the aid of D2D communications and FD and/or half-duplex (HD) technologies has been elaborated, which is known as cooperative NOMA, with the

objective of improving the performance of cell-edge users. Specifically, since the broadcast superimposed signal carries information of each cellular user's message, each user is able to have prior information about the messages intended for other UEs. cooperative NOMA (C-NOMA) exploits the availability of this prior information about the messages intended for other UEs to improve the achievable data rate as well as to enhance the reception reliability, especially at the cell-edge users. However, most of the work investigating the performance of C-NOMA has considered a single-cell set-up, meanwhile, the performance of the C-NOMA in a multi-cell scenario is not studied yet, which needs to be investigated.

Another direction toward this convergence to NGMA is to guarantee reliable communications between NOMA users and the BSs. However, in a realistic sense, the highly uncontrollable and random behavior of wireless environments due to blockages, channel impairments, severe path-loss, etc., impede the communications between cellular users and BSs. A possible method for circumventing the unreliability of wireless links is to sense the propagation environment and to recognize alternative propagation routes through which the information signal can be received at the point-of-interest in a low-cost and power-efficient way. One of the major technological breakthroughs to control wireless propagation, and hence, guarantee reliable communications between source and destination is the reconfigurable meta-surfaces, i.e., Reconfigurable intelligent surfacess (RISs). The amalgamation between RIS and NOMA is gaining momentum to enhance and improve different performance metrics, such as power consumption, network spectral efficiency, network energy efficiency, and user fairness. Despite the crucial role that the RIS plays in NOMA-based cellular networks, most work, in the literature, has only investigated the performance of RIS-aided NOMA under a single-cell set-up, whilst there is a lack of investigation in multi-cell scenarios in the existing literature.

1.1 Motivations

The key objectives of this thesis are to:

1. Investigate and address the potential gains brought by the synergistic integration of NOMA/C-NOMA with CoMP transmission in a multi-cell scenario.

2. Study the role of the RIS in single-cell and multi-cell NOMA/C-NOMA cellular networks to improve the network performance from different perspectives.

3. Provide innovative solutions to improve the performance of the network in terms of spectral efficiency, energy efficiency, and outage probability.

The proposed work certainly addresses a timely and vital topic, i.e., NOMA, which is expected to play an important role toward the convergence toward NGMA. This topic has recently received considerable attention in both academia and industry [6,7]. In fact, much work in the literature has revealed the significant gains of NOMA in realizing the needs of future wireless networks with superior performance to that of the traditional orthogonal multiple-access (OMA) techniques [6, 7]. Specifically, NOMA has been envisioned as a contender multiple access mechanism for the next-generation wireless network [6, 7]. However, there are still several limitations that hinder in unleashing the full potential gains of NOMA in 6G networks such as the effect of serious inter-cell interference and intra-NOMA user interference, the uncontrollable propagation behavior of the wireless channels, and the difficulty of achieving the required quality-of-service (QoS) at the edge of the cell.

In the following, we answer and address one important and fundamental question worth discussing: "*Why we should consider and study the integration of NOMA and other B5G technologies?*" We address this question from a broad perspective by discussing the needs and the features that each technology adds to empower and resolve some limitations and fill some gaps in NOMA-based cellular networks.

First, integrating CoMP with NOMA has two major benefits. On the one hand, in multi-cell NOMA-based cellular systems, far NOMA users experience serious inter-cell interference, which results in performance deterioration. To this end, CoMP transmission can assist in partially/fully canceling and mitigating the ICI for cell-edge users (CEUs) to enhance communication performance. On the other hand, traditional CoMP, which is based on conventional OMA techniques, may lead to low spectral efficiency. This is due to the fact that, in OMA, once a radio channel is occupied by a CEU, it cannot be accessible by other users. The situation will be even worse when the number of CEUs increases. By applying NOMA to CoMP-based cellular networks, each BS is allowed to serve, at least, an additional user, which is usually a CCU having a good channel gain with the BS, concurrently with the CEU. Accordingly, each radio resource is allocated to two UEs, one CCU, and one CEU. According to what we discussed, it can be observed that NOMA and CoMP complement each other in many aspects. It can be shown that NOMA

can resolve the connectivity issue in CoMP-based cellular networks; meanwhile, CoMP can tackle the issue of interference at the far NOMA users. Consequently, this motivates us to investigate the application of CoMP in the downlink transmission of NOMA-based cellular networks in order to enhance the network spectral-efficiency (NSE).

Second, the application of cooperative relaying with the aid of D2D communications and HD and/or FD relaying mode(s) to NOMA scenarios is becoming popular. In the following, we show what are the gaps that each one of these two technologies suffers from, and then, we present how cooperative relaying and NOMA can complement each other in many aspects. On one hand, although cooperative relaying can be deemed as an effective approach to combat path-loss and shadowing, enhance the reception reliability at the destination, and hence, extend the communication coverage, it can potentially sacrifice the NSE due to the two time-slot requirement in the HD relaying mode or due to negative impact of the residual SI in the FD relaying mode. On the other hand, from the NOMA side, the far NOMA users mainly suffer from severe intra-NOMA user interference as well as serious path-loss attenuation, and hence, the BS is forced to provide these users with most of its available transmit power, leaving a small portion of its power budget to users with strong channel gains, which may not be enough to serve their requirements. Consequently, combining cooperative relaying with NOMA promises to tackle and address the aforementioned issues and limitations with a significant enhancement in diversity gains resulting in superior user fairness and reception reliability. This is due to the fact that integrating NOMA with cooperative communications exploits the properties of the successive interference cancellation (SIC) process, on other words, the availability of other users' information at one user to achieve the expected performance by utilizing the duplicate transmission due to the relaying mode adopted by the strong NOMA users to improve the reception of the weak NOMA users. Since the performance of such a framework is strongly affected by the user clustering policy as well as the power control, we provide a novel resource management scheme, which coincides both the power allocation and the user clustering, in multi-cell C-NOMA-based cellular networks.

In other regard, the increasing density of wireless networks has caused the information and communication technology (ICT) community to become a considerable contributor to the overall carbon footprint. According to recent studies, the ICT systems that require energy sources for data transmission and communication increase the global world's CO2

emissions by 5% every year. Moreover, the fast advent of digitization in multiple sectors and the widespread of the IoE services will continue to drive the rapid growth of the energy demand in the ICT sector. In addition, it is expected that the telecom sector over the globe will consume 51% of the global energy by 2030, which causes a huge amount of greenhouse gas (GHG) emissions. As such, energy efficiency (EE) is the paramount consideration to decrease the net present GHG emissions. Therefore, toward a green-energy economy, EE has become one of the main themes of the 6G networks. Although network energy consumption has severe implications on the economic cost and becomes a major issue in the next-generation wireless networks, much work in the literature devoted their attention to improving the network spectral efficiency as they considered the NSE as the key performance indicator for the optimization and design of the wireless systems. In light of the above, it is necessary to develop multiple access techniques that are able to improve the network energy efficiency toward a green wireless environment.

Although the performance gains will increase by integrating CoMP with NOMA, the required QoS, in terms of the minimum achievable data rate, may not be guaranteed at the CEUs. This is because the CEUs still experience severe path-loss attenuation due to their far locations from the serving BSs, which encounter an obstacle for these users to achieve high-data rates that are required by the new emerging 6G applications. It is worth mentioning that guaranteeing the required QoS at the CEUs and tackling the severe path-loss attenuation can be achieved through two main approaches; 1) Increasing the transmit power by the serving BSs, and/or 2) Improving and controlling the channel gains from the BSs to the CEUs. Increasing the transmit power by the serving BSs is neither a cost-efficient nor an energy-efficient solution. On the other hand, improving and controlling the channel gains from the BSs to the CEUs can be realized through the concept of smart radio environment (SRE), i.e., by considering the wireless propagation environment as an optimization variable that can be jointly controlled with the transceiver design [8]. The key enabler technology to realize the vision of making the wireless environment controllable and programmable is the RIS. Due to the induced degrees of freedom, RIS provides a paradigm shift in the performance optimization and evaluation of wireless communication networks. Accordingly, through a proper RIS configuration, the channel gains of the CEUs can be improved and the required QoS can be attained without consuming much transmit power. Consequently, and with the sake of achieving the required QoS at the

CEUs, it is important to study the application of RIS in CoMP-NOMA cellular networks.

1.2 Thesis Contributions

Given the motivating points in the previous section, the main contributions of this thesis can be detailed as follows.

Performance Analysis for CoMP NOMA/C-NOMA cellular Networks. In Chapter 3, we first investigate the integration between CoMP transmission and NOMA in downlink heterogeneous cloud radio access networks (H-CRANs). In H-CRAN, low-power high-density small remote radio heads (SRRHs) are underlaid by high-power low-density macro remote radio heads (MRRHs). However, co-channel deployment of different RRHs gives rise to the problem of ICI that significantly affects the system performance, especially the cell-edge users. Thus, the users are first categorized into Non-CoMP users and CoMP users based on the relation between the useful signal to the dominant interference signal. The Non-CoMP user is the UEs having high signal-to-interference-plus-noise-ratio (SINR), and hence, it associates with only one RRH. On the other hand, the CoMP user, cell-edge user, is the UE that experiences less distinctive received power with the best two RRHs. In the proposed CoMP NOMA framework, each RRH schedules a CoMP-UE and a non-CoMP-UE over the same transmission channel using NOMA. We first design an analytical framework based on tools from stochastic geometry to evaluate the performance of the proposed framework (CoMP NOMA) in terms of the average achievable data rate for each NOMA UE. We then examine the NSE of the proposed CoMP NOMA-based H-CRAN. Simulation results are provided to validate the accuracy of the analytical models and to reveal the superiority of the proposed CoMP NOMA framework compared with the conventional CoMP OMA technique. By reaping the benefits of both CoMP and NOMA, we proved that the proposed framework can successfully deal with the ICI by using CoMP and improving the network's spectral efficiency through the NOMA technique. We also show that, with an appropriate PA coefficient setting at the Non-CoMP-UEs, a fairness performance can be achieved between the CoMP-UEs and the Non-CoMP-UEs.

Second, we investigate the integration of CoMP and C-NOMA aiming to improve the performance of CEUs. Using this framework, we exploit the cooperation between base stations (BSs) to mitigate the ICI and the successive decoding of users that are near the BSs to further enhance the performance of CEUs. In this setting, we derive a closed-

form expression for the outage probability of a CEU along with an analytical expression for its ergodic rate. We validate the derived expressions through various Monte-Carlo simulations, where we show the superiority of the proposed framework compared with other multiple access schemes proposed in the literature. The work done in this chapter leads to two IEEE published journals [9, 10].

Joint Power Allocation and User Pairing in CoMP C-NOMA Cellular Networks. In Chapter 4, we first investigate a dynamic PA problem of a cellular network consisting of two adjacent and coordinating cells. The joint transmission CoMP between the two-cell is introduced to assist users experiencing high ICI, where each cell invokes C-NOMA to serve its associated devices. Both effects of imperfect SIC and imperfect channel estimation are considered within the proposed scheme. Accordingly, a PA framework is formulated as an optimization problem with the objective of maximizing the network sum-rate while guaranteeing a certain QoS for each user. The formulated optimization problem is neither concave nor quasi-concave, which is difficult to be solved directly unless using heuristic methods, which comes with the expense of high computational complexity. To overcome this issue, near-optimal closed-form expressions for the PA are derived. The simulation results show that our purposed scheme achieves an average sum-rate that is 5% less than the one of the optimal power control but it can save up to 99% in computational time. In addition, the superiority of the proposed CoMP C-NOMA scheme is demonstrated when compared to the well-known C-NOMA scheme.

Second, we extend this model to a multi-user scenario. In this part, a framework to jointly optimize the power control and the UEs clustering of the CoMP-assisted FD C-NOMA system is formulated as an optimization problem to maximize the network sum-rate while guaranteeing the required QoS of cellular UEs. The formulated problem is a non-convex mixed-integer nonlinear programming (MINLP) that cannot be solved in a straightforward manner. To tackle this issue, the formulated problem is decomposed into an inner PA problem and an outer user clustering (UC) problem. For the inner problem, a computationally-efficient solution is obtained. Meanwhile, the outer problem is reformulated as a one-to-one three-sided matching game. Then, a low-complexity near-optimal clustering algorithm is proposed. The simulation results demonstrated that 1) the optimality of the power control solution; 2) the CoMP-assisted FD C-NOMA has a superior performance compared to CoMP-assisted HD C-NOMA and CoMP NOMA

schemes for moderate values of SI. It has been also shown that the proposed solution achieves around 96.5% of the average achievable network sum-rate of the optimal solution. The work done in this chapter leads to two IEEE published journals [11, 12].

RIS-assisted C-NOMA Cellular Networks. In Chapter 5, we investigate the downlink transmission of RIS-aided C-NOMA, where both HD and FD relaying modes are considered. The system model consists of one BS, two users and one RIS. The goal is to minimize the total transmit power at both the BS and at the user-cooperating relay for each relaying mode by jointly optimizing the PA coefficients at the BS, the transmit power coefficient at the relay user, and the passive beamforming at the RIS, subject to power budget constraints, the SIC constraint and the minimum required QoS at both cellular users. To address the high-coupled optimization variables, an efficient algorithm is proposed by invoking an alternating optimization (AO) approach that decomposes the original problem into a PA sub-problem and a passive beamforming sub-problem, which are solved alternately. For the PA sub-problem, the optimal closed-form expressions for the PA coefficients are derived. Meanwhile, with the aid of difference-of-convex (DC) rank-one representation and successive convex approximation (SCA), an efficient solution for the passive beamforming is obtained. The simulation results validate the accuracy of the derived power control closed-form expressions and demonstrate the gain in the total transmit power brought by integrating the RIS in C-NOMA networks. The work done in this chapter leads to an IEEE published journal [13].

RIS-Assisted CoMP NOMA Networks: Performance Analysis and Optimization. In Chapter 6, we first study the performance analysis of the integration between RIS with downlink NOMA in a two-cell network assisted by CoMP. The RIS is adopted to construct a strong combined channel gain at the CEU, while CoMP is used to mitigate the effects of ICI. In this proposed framework, we derive first a closed-form expression for the ergodic rate of the CEU, and then we evaluate the NSE. We validate the derived expression through Monte-Carlo simulations, where we demonstrate the efficacy of the proposed framework compared to other multiple access techniques proposed in the literature.

Second, we jointly optimize the PA coefficients at the BSs, the UC policy, and the phase-shift (PS) matrix of the RIS with the objective of maximizing the network sum-rate subject to a target QoS, defined in terms of the minimum required data rate at each

cellular user and the SIC constraints. The formulated problem ends to be a non-convex MINLP that is difficult to be solved in a straightforward manner. To alleviate this issue, and with the aid of AO, the original optimization problem is decomposed into two sub-problems, a joint PA and UC sub-problem and a PS sub-problem, that are solved in an alternating way. For the first sub-problem, we invoke the bi-level optimization approach to decouple the PA sub-problem from the UC sub-problem. For the PA sub-problem, closed-form expressions for the optimal PA coefficients are derived. On the other hand, the UC problem is projected to multiple 2-dimensional assignment problems, each of which is solved using the Hungarian method. Finally, the PS sub-problem is formulated as a DC problem and an efficient solution is obtained using the SCA technique. The numerical results reveal that the network sum-rate of the proposed RIS-assisted CoMP NOMA networks outperforms the conventional CoMP NOMA scheme without the assistance of the RIS, the RIS-assisted CoMP OMA scheme and RIS-assisted NOMA scheme, especially for low transmit power from the BSs. The work done in this chapter leads to two IEEE published journals [14, 15].

1.3 Thesis Organization

The rest of the thesis is organized as follows. In Chapter 2, we present a relevant background for the main and key topics of this thesis. In Chapter 3, we investigate the performance analysis of CoMP NOMA/C-NOMA cellular networks. We evaluate the power allocation and user pairing in CoMP C-NOMA cellular networks in Chapter 4. In Chapter 5, we study the role of the RIS in C-NOMA-based cellular networks to provide energy-efficient communications. We first investigate the performance analysis of RIS-empowered CoMP NOMA cellular networks, and then, we study the joint power allocation, user pairing, and passive beamforming in Chapter 6. Finally, we present a brief summary of our study and investigation, and then, we provide some important conclusions and some potential directions for future research in Chapter 7.

1.4 List of Publications

Journal Articles

[J1] A. Al-Hilo, M. Samir, **M. Elhattab**, C. Assi and S. Sharafeddine, "RIS-Assisted UAV for Timely Data Collection in IoT Networks," in *IEEE Systems Journal*, Nov. 2022.

[J2] **M. Elhattab**, M. Arfaoui, Chadi Assi, Ali Ghrayeb, Marwa Qaraqe "Power Allocation Optimization and Decoding Order Selection in Uplink C-NOMA Networks," *IEEE Communications Letters*, Oct. 2022.

[J3] S. Khisa, M. Almekhlafi, **M. Elhattab** and C. Assi, "Full Duplex Cooperative Rate Splitting Multiple Access for a MISO Broadcast Channel With Two Users," in *IEEE Communications Letters*, vol. 26, no. 8, pp. 1913-1917, Aug. 2022.

[J4] **M. Elhattab**, M. A. Arfaoui and C. Assi, "Joint Clustering and Power Allocation in Coordinated Multipoint Assisted C-NOMA Cellular Networks," in *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 3483-3498, May 2022.

[J5] **M. Elhattab**, Mohamed Amine Arfaoui, Chadi Assi, Ali Ghrayeb "Reconfigurable Intelligent Surface Enabled Half-Duplex/Full-Duplex Cooperative Non Orthogonal Multiple Access, "*IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3349-3364, May 2022.

[J6] **M. Elhattab**, M. A. Arfaoui, C. Assi and A. Ghrayeb, "RIS-Assisted Joint Transmission in a Two-Cell Downlink NOMA Cellular System," in *IEEE Journal on Selected Areas in Communications, Special Issue on Next Generation Multiple Access,* vol. 40, no. 4, pp. 1270-1286, April 2022.

[J7] A. Al-Hilo, M. Samir, **M. Elhattab**, C. Assi and S. Sharafeddine, "Reconfigurable Intelligent Surface Enabled Vehicular Communication: Joint User Scheduling and Passive Beamforming," in *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 2333-2345, March 2022.

[J8] M. Almekhlafi, M. A. Arfaoui, **M. Elhattab**, C. Assi and A. Ghrayeb, "Joint Resource Allocation and Phase Shift Optimization for RIS-Aided eMBB/URLLC

Traffic Multiplexing," in *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 1304-1319, Feb. 2022.

[J9] E. E. Haber, **M. Elhattab**, C. Assi, S. Sharafeddine and K. K. Nguyen, "Latency and Reliability Aware Edge Computation Offloading via an Intelligent Reflecting Surface," in *IEEE Communications Letters*, vol. 25, no. 12, pp. 3947-3951, Dec. 2021.

[J10] **M. Elhattab**, M. Kamel and W. Hamouda, "Edge-Aware Remote Radio Heads Cooperation for Interference Mitigation in Heterogeneous C-RAN," in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 12142-12157, Nov. 2021.

[J11] M. Samir, **M. Elhattab**, C. Assi, S. Sharafeddine and A. Ghrayeb, "Optimizing Age of Information Through Aerial Reconfigurable Intelligent Surfaces: A Deep Reinforcement Learning Approach," in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3978-3983, Apr. 2021.

[J12] **M. Elhattab**, M. A. Arfaoui and C. Assi, "Power Allocation in CoMP-Empowered C-NOMA Networks," in *IEEE Networking Letters*, vol. 3, no. 1, pp. 10-14, Mar. 2021.

[J13] **M. Elhattab**, M. -A. Arfaoui, C. Assi and A. Ghrayeb, "Reconfigurable Intelligent Surface Assisted Coordinated Multipoint in Downlink NOMA Networks," in *IEEE Communications Letters*, vol. 25, no. 2, pp. 632-636, Feb. 2021.

[J14] **M. Elhattab**, M. -A. Arfaoui and C. Assi, "CoMP Transmission in Downlink NOMA-Based Heterogeneous Cloud Radio Access Networks," in *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7779-7794, Dec. 2020.

[J15] **M. Elhattab**, M. A. Arfaoui and C. Assi, "A Joint CoMP C-NOMA for Enhanced Cellular System Performance," in *IEEE Communications Letters*, vol. 24, no. 9, pp. 1919-1923, Sept. 2020.

[J16] **M. Elhattab** and W. Hamouda, "Performance Analysis for H-CRANs Under Constrained Capacity Fronthaul," in *IEEE Networking Letters*, vol. 2, no. 2, pp. 62-66, June 2020.

[J17] **M. Elhattab**, M. M. Elmesalawy, F. M. Salem and I. I. Ibrahim, "Device-Aware Cell Association in Heterogeneous Cellular Networks: A Matching Game Approach," in *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 1, pp. 57-66, Mar. 2019.

[J18] Ali Amhaz, **M. Elhattab**, Sanaa Sharafeddin, and Chadi Assi, "UAV-Assisted Cooperative NOMA: Deployment and Power Allocation,"Oct. 2022. (Submitted)

[J19] **M. Elhattab**, M. A. Arfaoui, Sanaa Sharafeddin, and Chadi Assi, "Sum-Rate Maximization for Network NOMA-Assisted UAVs Communications with Joint Reception,"Oct. 2022. (Submitted)

[J20] S. Khisa, **M. Elhattab**, M. A. Arfaoui, Sanaa Sharafeddin, and Chadi Assi, "Joint Power Allocation and Beamforming for Uplink Rate Splitting Multiple Access with User Cooperation,"Dec. 2022. (Submitted)

[J21] Ali Muhammad, **M. Elhattab**, M. A. Arfaoui, Ahmed Al-Hilo, Chadi Assi "Optimizing Age of Information for Reflective Intelligent Surface-aided wireless network,"Jun. 2022. (Submitted)

[J22] **M. Elhattab**, M. Khabbaz, N. Al Dahabreh, R. Atallah, and Chadi Assi, "Leveraging Real-World Data Sets For QoE Enhancement In Public EV Charging Networks," Sept. 2022. (Submitted).

[J23] Ali Muhammad, **M. Elhattab**, M. A. Arfaoui, Chadi Assi "Optimizing Age of Information in RIS-Empowered Uplink Cooperative NOMA Networks,"Aug. 2022. (Submitted)

[J24] Elie Elhaber, **M. Elhattab**, Sanaa Sharafeddin, and Chadi Assi, "Multi-IRS Aided Mobile Edge Computing for High Reliability and Low Latency Services,"Aug. 2022. (Submitted)

[J25] S. Khisa, **M. Elhattab**, Sanaa Sharafeddin, and Chadi Assi, "Energy Consumption Optimization in RIS-assisted Cooperative RSMA Cellular Networks,"Aug. 2022. (Second round revision, *IEEE Transactions on Communications*)

[J26] Ali Muhammad, **M. Elhattab**, M. A. Arfaoui, Chadi Assi "Optimizing Information Freshness in RIS-assisted NOMA-based IoT Networks," Jun. 2022. (Second round revision, *IEEE Transactions on Network and Service Management*).

Conference Proceedings

[C1] Ali Muhammad, **M. Elhattab**, M. Arfaoui, Chadi Assi "Optimizing Information Freshness Leveraging Multi-RISs in NOMA-based IoT Networks," *IEEE Global Communications Conference (Globecom)*, Dec. 2022.

[C2] E. E. Haber, **M. Elhattab**, C. Assi, S. Sharafeddine and K. K. Nguyen, "Latency and Reliability Aware Edge Computation Offloading in IRS-aided Networks," *ICC 2022 - IEEE International Conference on Communications,* 2022, pp. 5035-5040.

[C3] A. Muhammad, **M. Elhattab**, M. Shokry and C. Assi, "Leveraging Reconfigurable Intelligent Surface to Minimize Age of Information in Wireless Networks," *ICC 2022 - IEEE International Conference on Communications,* 2022, pp. 2525-2530.

[C4] M. Almekhlafi, M. A. Arfaoui, **M. Elhattab**, C. Assi and A. Ghrayeb, "Joint Scheduling of eMBB and URLLC Services in RIS-Aided Downlink Cellular Networks," 2021 *International Conference on Computer Communications and Networks (ICCCN)*, 2021, pp. 1-9.

Chapter 2
Background

In this Chapter, we presented the main enabling network architectures and technologies for fifth-generation (5G)/6G wireless networks that have been considered within this thesis. Specifically, we first presented an overview of NOMA, C-NOMA, and CoMP. Then, we illustrated the centralized-based network architecture C-RAN and showed how this architecture can be considered as an ideal platform to practically implement CoMP. Finally, we described the main characteristics and benefits brought by the RIS to 6G networks.

2.1 Non-Orthogonal Multiple Access

According to whether each UE is assigned orthogonal resources (code/frequency/time) or not, the current multiple access (MA) techniques can be categorized into two classes, namely orthogonal transmission mechanisms and non-orthogonal transmission mechanism [16]. Specifically, given their advantages in terms of both the inter-user interference avoidance and the low complexity in implementation, orthogonal transmission mechanisms have been widely employed in practical cellular communication networks, such as frequency division multiple access (FDMA) in the first-generation (1G) cellular network, time division multiple access (TDMA) in the second-generation (2G) [17], code division multiple access (CDMA) with orthogonal codes in the third-generation (3G) [18], and orthogonal frequency division multiple access (OFDMA) in the fourth-generation (4G) [19]. These OMA techniques were developed with one common theme which is to assign orthogonal resource blocks (RBs) to different UEs so that inter-user interference can be avoided. In particular, in order to orthogonalize the data of different users, the BSs should allocate different resources (time, frequency, space, code) to differentiate between users [20–23] as depicted in Fig. 2.1. However, given the limited amount of available radio spectrum, the ever-increasing number of UEs, traditional OMA techniques have become insufficient due to their low spectral-efficiency (SE) as well as the limited number of UEs that can be

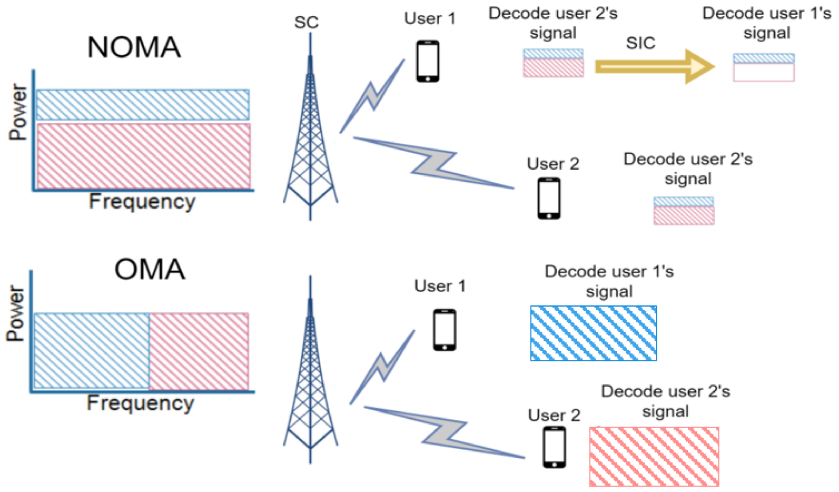


Figure 2.1: Illustration of two-user NOMA compared to OMA (FDMA).

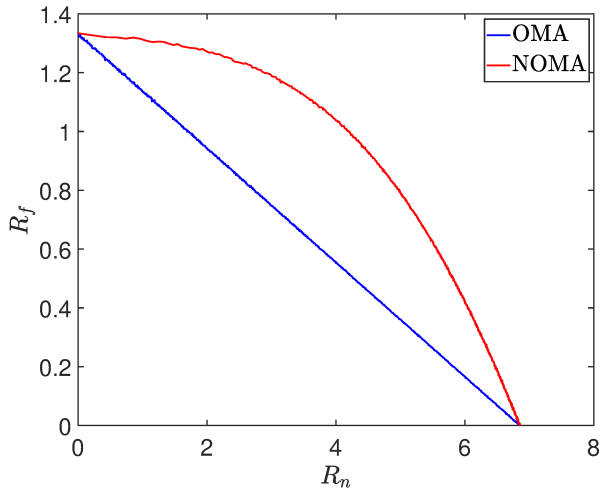


Figure 2.2: Achievable regions for two-user OMA and NOMA with $|h_n| = 10|h_f| = \sqrt{5}$ and $P_{\text{BS}}/\sigma^2 = 40$, where $|h_n|^2, |h_f|^2$ are the channel gains from BS to user 1, which is referred to as UE_1 and user 2, which is denoted as UE_2, respectively.

supported because of the inflexible resource allocation. Consequently, supporting a variety of new emerging applications with dissimilar requirements such as enhanced mobile broadband (eMBB), massive machine-type communication (mMTC), and ultra-reliable low-latency communication (URLLC) in the next-generation wireless networks necessitate novel UE access mechanisms which can potentially serve multiple UE in the same RB [24].

In order to tackle the challenging emerging heterogeneous applications and services (e.g., augmented reality, ultra-high-definition video, and industry 5.0), NOMA has been deemed as one of the vital enabling MA techniques for the upcoming 6G cellular networks and has drawn remarkable research interests in the past few years [20–23]. In contrast to

OMA techniques, NOMA can simultaneously serve multiple UEs using the same transmission block (same frequency channel, time, and/or spatial domain) with the cost of intra-NOMA-user interference (INUI).[1] This can be achieved by using superposition coding (SC) at the transmitter with appropriately PA coefficients and performing SIC at the receiver side to remove the INUI [4]. This allows the UEs with poor channel conditions, which are referred to as "far" NOMA UEs, to be simultaneously served with the UEs that have good channel conditions, which are referred to as "near" NOMA UEs. The operation of power-domain NOMA in comparison with one of the OMA techniques, i.e., FDMA, is presented in Fig. 2.1 for two users. Despite the high complexity brought by NOMA at both the transmitter and receiver, significant benefits can be obtained, such as achieving high SE and EE [25], supporting massive connectivity [26, 27], and guaranteeing UEs fairness [28]. As a result, compared with OMA mechanisms, the aforementioned benefits make NOMA as a promising candidate for the NGMA. In order to emphasize the performance gains of NOMA over OMA, the capacity region for both NOMA and OMA considering single-antenna broadcasting channel is depicted in Fig.2.2.

2.2 Cooperative Non-Orthogonal Multiple Access

With the aid of D2D communication and HD and/or FD technologies, C-NOMA has been proposed to improve the fairness performance, in which near NOMA UE can act as a decode-and-forward (DF) relay to enhance the reception reliability at far NOMA UE [29–32]. C-NOMA can operate in two different modes, i.e., HD relaying mode or FD relaying mode [32]. This is based on whether a relay node, i.e., near NOMA UE, is able to simultaneously transmit and receive in the same time/frequency band. Specifically, the main details for the operation of C-NOMA technique can be explained as follow. In C-NOMA-enabled cellular networks, the data transmission is executed over two different phases as depicted in Fig. 2.3. The first phase is the direct transmission (DT) phase and the second phase is the cooperative transmission (CT) phase that are detailed as follows [32–35].

- DT phase: In this phase, the BS applies SC on the signals of each NOMA pair. Then, the superimposed signals are transmitted simultaneously by the BS. Following the

[1]The signals for other NOMA UEs with higher channel gains than a typical user in the same cluster produce INUI.

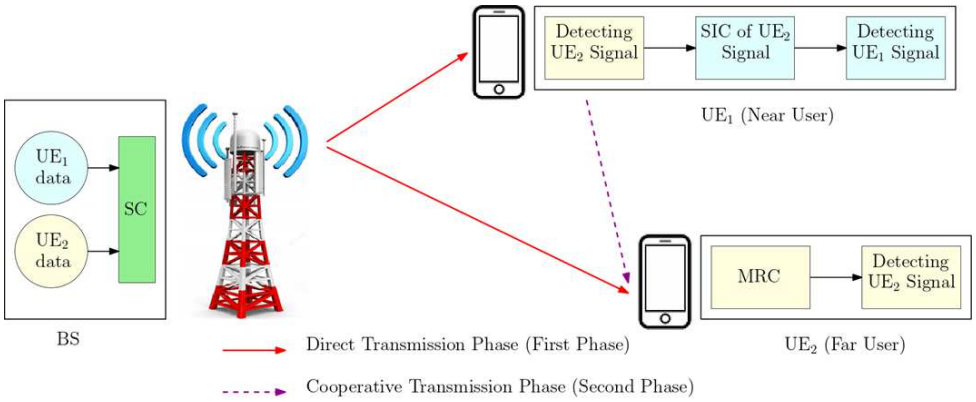


Figure 2.3: Two-user C-NOMA architecture.

NOMA principle, the strong NOMA UE first performs SIC to decode the signal of the weak NOMA UE. Second, each strong NOMA UE removes the decoded signal of weak NOMA UE from its own reception and then decodes its signal from the resulting reception. Meanwhile, the weak NOMA UE treats the signals of the strong NOMA UE as a noise.

- CT phase: In this phase, the strong NOMA UE forwards the weak NOMA UE's decoded message during the SIC process to this weak NOMA UE through a D2D channel.

Here, the strong NOMA UE can adopt two different DF relaying modes, either a HD relaying mode or a FD relaying mode [32–35]. For the case of HD DF relaying mode, the DF phase and the CT phase occur in two consecutive time-slots. In other words, the DT phase occurs in the first time-slot and the CT phase is executed in the second time-slot. However, the existence of the two time-slot transmission mechanism may result in a SE deterioration due to the pre-log penalty. Hence, in order to further enhance the bandwidth usage efficiency of the considered cellular network, FD technology has been proposed as a key promising solution that enables the transmission and the reception of signals within the same time-frequency resource. Consequently, for the case of FD DF relaying mode, the two phases occur in the same time-slot in which the DT phase is followed by the CT phase. Note that, in a general case, because of the imperfect cancellation process or isolation, FD communication may suffer from residual SI that is modeled as a fading channel [32]. With the development of antenna technologies and signal processing, relaying with FD communication is feasible [36].

Finally, the rate region has been characterized for the C-NOMA to emphasize the

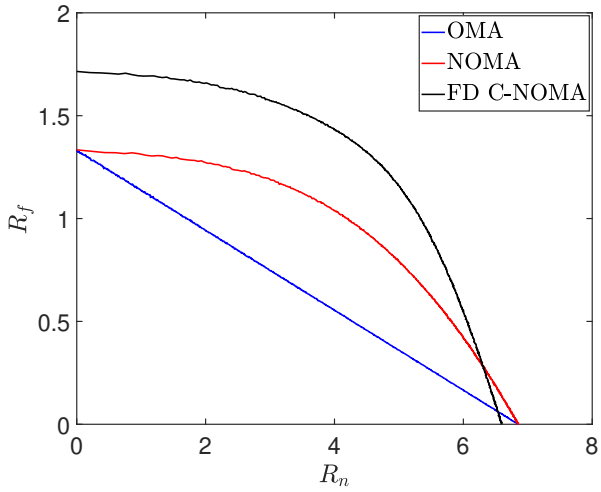


Figure 2.4: Achievable regions for two-user OMA, NOMA, and C-NOMA with $|h_n| = 10|h_f| = \sqrt{5}$, the D2D channel gain $|h_{nf}| = \sqrt{3}$, the SI channel gain $|h_{\mathrm{SI}}| = 0$ dB, and $P_{\mathrm{BS}}/\sigma^2 = 40$.

gain brought by the cooperation between the near NOMA UE and far NOMA UE in comparison with NOMA and OMA as shown in Fig 2.4.

2.3 Coordinated Multipoint Transmission

In order to reduce the additional inter-tier interference (interference between small cells and macro cells) and intra-tier interference (interference between BSs belonging to the same tier) in dense multi-cell networks, the 3GPP introduced the concept of CoMP transmission. CoMP transmission has been considered as one of the main enabling technologies for improving cell-edge users' performance, and hence, enhancing the network coverage for future mobile networks. With CoMP, cooperating BSs can share the channel-state-information (CSI) of UEs when scheduling is performed, which makes joint scheduling possible. Accordingly, three levels of coordination schemes between the BSs are summarized in [37] and are identified by 3GPP for LTE-advanced (LTE-A) dependent on the scheduling complexity and the required backhaul/fronthaul capacity [38]. An illustration of the downlink CoMP types is shown in Fig. 2.5.

2.3.1 Coordinated Scheduling/Coordinated Beamforming

In this coordinated scheduling (CS)/coordinated beamforming (CB) transmission scheme, the beamforming vectors are selected such that the interfering BSs are steered towards the null space of the CEU who experiences high ICI as shown in Fig. 2.5.(a). Note that,

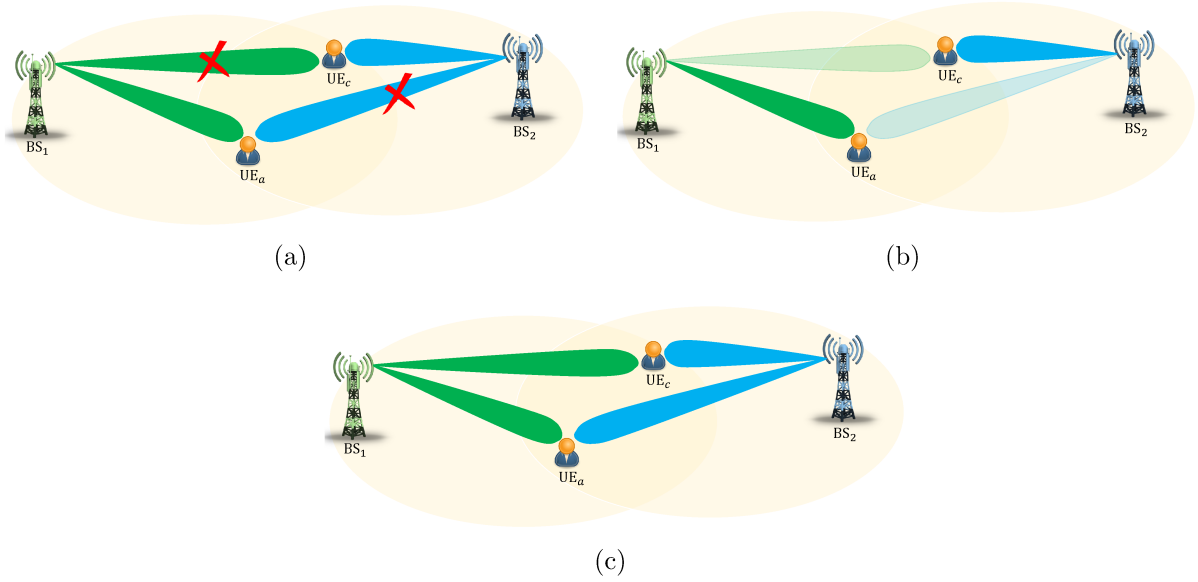


Figure 2.5: Different types of CoMP transmission [37].

in this scheme, only the CSI is required at each BS and the user's data is only available at the serving BS.

2.3.2 Dynamic Point Selection

In this kind of transmission, the serving BS is changed dynamically in each sub-frame based on the channel conditions between the BSs and the UEs as shown in Fig. 2.5.(b). Fading conditions are responsible for selecting the serving BS in each sub-frame. It is worth mentioning that the user data must be available at each BS in a predetermined BSs cluster.

2.3.3 Joint Transmission

In this transmission scheme, a cluster of BSs are coordinated to concurrently transmit the users' data to provide high `SINR` at the CEU as shown in Fig. 2.5.(c). Specifically, joint-transmission (JT)-CoMP technique allows the transmission of the same data from multiple nodes to a UE on the same time-frequency resource unit. This kind of transmission converts the interference signal at the user-of-interest to a useful signal. According to the 3GPP, JT-CoMP can be divided into coherent JT (CJT) and non-coherent JT (NCJT). In CJT, the transmitted signal from the multiple nodes is jointly precoded to attain coherent combining at the receiver. Note that, in this scheme, both CSI and the data of the UEs are shared between the coordinated BSs. Moreover, in this transmission model, small

timing error differences (prior phase alignment) and a strict synchronization between the transmission points are required to realize the full potential gains of this scheme, which limits its applicability [37, 39, 40]. On the other hand, in NCJT scheme, the transmission points cooperate by jointly transmitting the same data to a given UE without a prior phase mismatch alignment or a strict synchronization, and hence, it is a more practical transmission scheme in comparison with the CJT [41–45]. It is worth mentioning that, in this thesis, we apply the NCJT CoMP. As a result, no synchronization is required at the transmission nodes. Note that, in NCJT scheme, the UEs date is only required at each coordinated BS.

2.4 Cloud Radio Access Networks

As mentioned earlier, to reduce the effect of the severe ICI in dense cellular networks, CoMP transmission is adopted as a promising solution for interference mitigation [46]. In the literature, CoMP transmission is reported to achieve significant performance gains in terms of per-cell throughput and coverage probability through the effective mitigation of the ICI in heterogeneous network (HetNet) [47]. However, the coordination between all BSs in the network is a very complex problem [37]. This leads to a practical difficulty in applying CoMP transmission on distributed radio access networks. In addition, this problem is further complicated when the network gets denser. Consequently, to overcome this challenge, the BSs should form a kind of centralized radio access network [48].

Recently, cloud computing has been considered as a popular computing paradigm for improving both energy and spectral efficiencies [49]. Cloud computing has been integrated with the radio access network to provide a large cooperative gain [50]. Due to this integration, a new cellular architecture, named C-RAN is developed [51, 52]. In C-RAN, the radio frequency (RF) function is separated from the baseband processing in which the RF is implemented at the remote radio head (RRH). Meanwhile, the baseband processing units (BBUs) are aggregated and placed in a central location called BBU pool and connected with the densely deployed RRHs via fronthaul links as illustrated in Fig. 2.6. An C-RAN architecture facilitates the cross-cell cooperation and enhances the spectrum and energy efficiencies by densely deploying SRRHs which are empowered by cloud computing to cooperatively support users. Owing to the centralized processing nature of C-RAN, CoMP transmission can be easily applied in this architecture and the CEU's performance

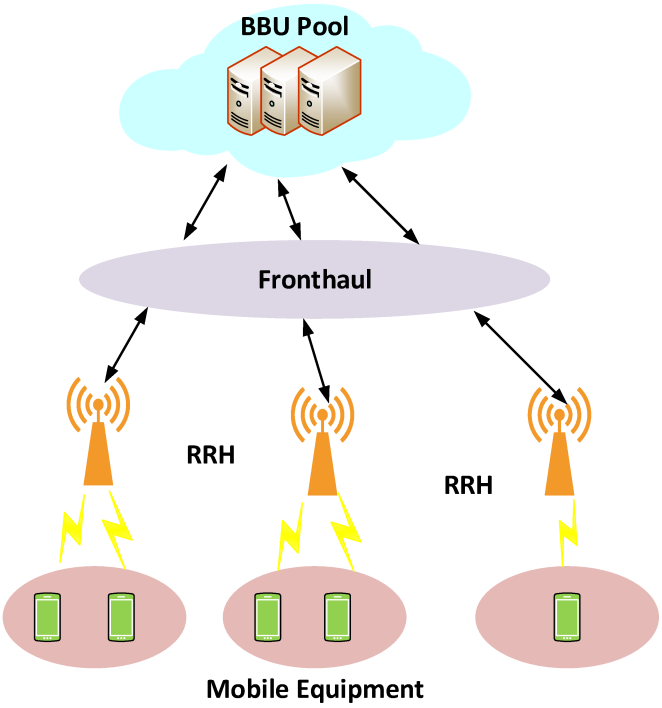


Figure 2.6: Cloud radio access network architecture.

can be boosted to meet the new emerging requirements. By centralizing the baseband processing, the RRHs can cooperatively transmit the users' data concurrently. Then, the received signals at the users are combined, and, hence a high SINR can be achieved at the user side [53]. Note that to reap the main benefits of small cells and CRAN, heterogeneous CRAN (H-CRAN) has been introduced as a promising architecture for 5G cellular networks [51, 54]. H-CRAN inherits the centralized processing and the powerful cloud computing of CRAN and high spectrum reuse of HetNet [54, 55].

2.5 Reconfigurable Intelligent Surface

RIS, also known as intelligent reflecting surface (IRS) or large intelligent surface (LIS), has been introduced as a key promising technology for achieving cost-, energy-, and spectral-efficient communications via smartly reconfiguring and controlling the wireless propagation environment [8, 56–64]. RIS can be thought of as a smart, inexpensive, and thin composite material sheet that is similar to a wallpaper, which may cover parts of buildings, walls, and ceilings. RIS is composed of a large number of passive low-cost elements, each of which is capable of independently tuning the phase-shift of the incident radio waves. For instance, by appropriately configuring the phase-shifts with the aid of the

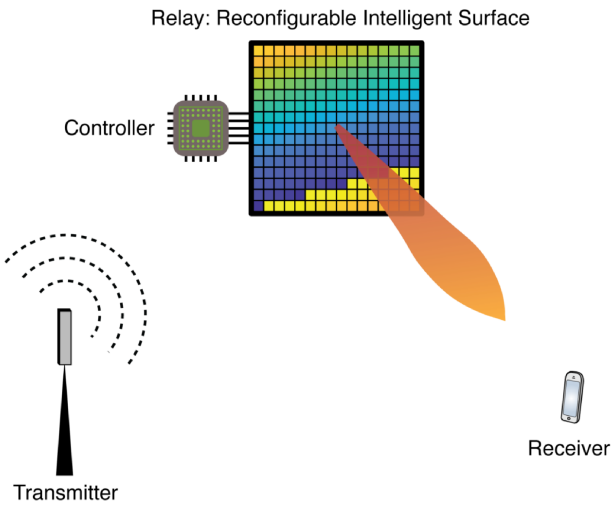


Figure 2.7: Reconfigurable intelligent surface operation [56].

RIS controller, the reflected signals can be constructively added, therefore, enhancing the received signal strength at the point of interest as depicted in Fig. 2.7. Based on the required mission, each RIS element can be configured to do refraction, absorption, reflection, and focusing [8, 56–64].

In contrast to the traditional relays, RISs has the benefits of low-power consumption and cost-efficiency since they passively reflect the incident signals without the need for any RF chains [8, 56–64]. In addition, the radio signal reflected by RIS is free from noise corruption or SI in an inherently FD fashion. Consequently, the deployment of RISs in cellular networks has the potential to remarkably enhance the coverage area of the BS and reduce the blind-spot areas at the higher operating frequencies, where there is no direct link between the user and the BS due to the existence of blockages. In addition, the integration between the RIS and 5G technologies have been also introduced [8, 56–64]. This is because RIS can be also utilized in lower operating frequencies to provide indirect line-of-sight (LoS) for the out-of-coverage UEs and construct a strong combined channel gain for the UEs with weak channel conditions, i.e., CEUs [8, 56–64].

When deployed in wireless systems, RIS has the potential of altering the wireless environment from being highly probabilistic to a controllable and partially deterministic space referred to as SRE or "Wireless 2.0" [8, 56–64]. It is worth mentioning that RIS has attracted significant attention from the research community worldwide, including the USA, Europe, and China. For example, research projects worth over 20 million Euros have been recently funded in Europe to study the impact of RIS on the performance of the next-generation wireless networks [65]. Starting from Sept. 30^{th} 2021 and will run for two years,

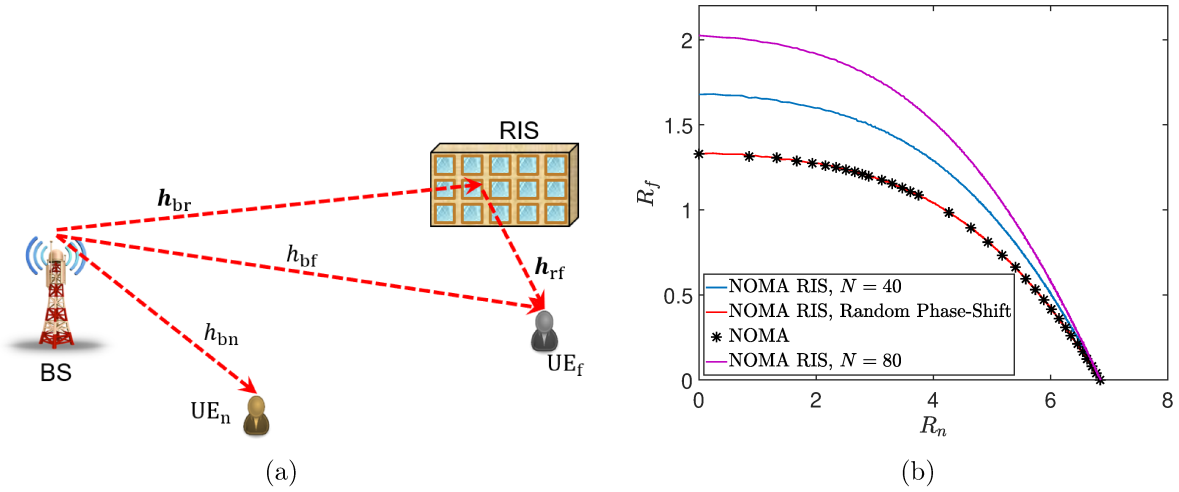


Figure 2.8: Capacity region for RIS-assisted two-UE NOMA cellular system, where N denotes the number of elements at the RIS.

Industry Specification Group (ISG) on RIS within the European Telecommunications Standards Institute (ETSI) will focus on pre-standardization work, including application scenarios, the definition of use cases, and relevant system requirements [65]. To highlight the benefits brought by RIS in two-UE NOMA network, we have evaluated the capacity region as shown in Fig. 2.8. One can see that the capacity region is extended compared to traditional NOMA with a proper configuration at the RIS's meta-atoms. In addition, we can see that the number of elements at RIS plays a crucial role in improving the capacity region.

Chapter 3
Performance Analysis of CoMP-Assisted NOMA/C-NOMA Networks

3.1 Introduction

In order to mitigate the severe effect of the ICI, further enhance the system SE, and support high downlink (DL) data rate requirements for the new emerging B5G applications, the integration of CoMP and NOMA/C-NOMA is expected to be utilized as one of the promising access technique in the physical layer of the downlink H-CRAN [66]. By combining NOMA/C-NOMA with CoMP transmission in H-CRAN, not only the performance of CoMP transmission will be further enhanced, but also a better SE can be achieved especially for far NOMA users [67,68]. We refer to the application of NOMA/C-NOMA technique in CoMP networks as CoMP NOMA/CoMP C-NOMA network. In such a model, CoMP NOMA/CoMP C-NOMA allows multiple BSs/RRHs to serve the weak UEs in NOMA pairs at the same time using the same radio resource to enhance their performance by leveraging the data sharing among cooperating RRHs. Note that, in this thesis, we exploit the concept of "Location-aware" CoMP. In this transmission scheme, the BSs/RRHs and UEs' locations are deemed to decide which UEs should trigger and operate in CoMP and which UEs are served by one RRH/BS.

3.2 State of The Art

Here, we briefly review the works that have been done to analyze and study the performance of DL NOMA/C-NOMA in single-/multi-cell cellular networks.

Specifically, we first discuss the works that evaluate NOMA/C-NOMA performance using tools from stochastic geometry in a *single-cell* cellular network. An analytical framework to evaluate the performance of a NOMA-enabled wireless network in both DL and

uplink (UL) is evaluated in [69]. The Poisson Point Process (PPP) is used to model the locations of both BSs and UEs on a 2-D plane. A network model is considered in [70], where Poisson distributed BSs transmit to N NOMA UEs. The main weakness is that the authors define the largest disk that can fit inside the Voronoi cell and the UEs outside this disk are not considered during their analysis because of having weaker channels in which the authors suggested that these CEUs can be served by OMA technique. A hybrid massive multiple-input multiple-output (mMIMO) aided macro-cells and heterogeneous networks framework with NOMA-enhanced small-cells is investigated in [26]. To do so, analytical expressions of the NOMA-enhanced small-cell networks in terms of SE are derived. Regarding the mMIMO-enabled macro-cells, a tractable analytical lower bound for macro UEs is derived. However, the authors in [26] consider an OMA transmission at the macro BS to reduce the effect of inter-tier interference. A performance analysis framework for the cell association in NOMA-based fog radio access network (F-RAN) is developed. Using stochastic geometry tools, closed-form analytical results are derived in [71]. Then, two cell-association algorithms based on reinforcement learning and evolutionary game, respectively, are presented.

The outage probability and the diversity order achieved by the C-NOMA scheme are analyzed in [29]. In addition, the authors utilize local short-range communication techniques to deliver the data from the strong NOMA UEs to the weak NOMA UEs. An FD D2D-aided C-NOMA scheme to enhance the transmission reliability of the weak NOMA UEs is studied in [30], in which the weak NOMA UE is assisted by the strong NOMA UE that is empowered by FD D2D communications. Then, the outage performance is derived, and a hybrid C-NOMA/OMA is proposed for further enhancing the outage performance. The authors in [72] derive the outage probability and the ergodic sum rate expressions when a pair of near UEs adopt a successive relay scheme to assist a far UE using HD relaying mode. In [72], a two-stage relay selection scheme, in a system model consisting of one BS, two UEs, and a set of relay nodes, is developed to satisfy the required QoS of both near and far NOMA UEs; further, a closed-form expression for the outage probability is derived. In [73], a relay-aided C-NOMA network consisting of one BS and two UEs that communicate via one relay node selected from a group of FD amplify-and-forward (AF) relays is studied. The outage probability at each UE is derived considering imperfect SIC at the strong UE. Note that, all aforementioned works (see [26, 29, 30, 69–74] and the

citations therein) studied the performance of NOMA/C-NOMA under a single-cell set-up.

In the following, we discuss the works that evaluate the NOMA performance using tools from the stochastic geometry in a *multi-cell* scenario. In such a scenario, the ICI inevitably degrades the performance of the CEUs, i.e., far NOMA users, which may lead to low received `SINR` [68, 70]. This is because the far NOMA users, which are considered in reality CEUs, always have worse channel conditions. Moreover, the effect of ICI on the far NOMA users' performance is even more exacerbated in the case of high-density BSs/RRHs [68]. As a result, to resolve the above challenge, NOMA should be combined with the 3GPP interference mitigation techniques such as CoMP transmission to attain higher spectral efficiency, especially for the far NOMA users [67, 68, 70]. Recently, more attention is directed toward studying the application of CoMP in NOMA-based multi-cell downlink transmissions [67, 68, 75].

In [76], Alamouti code is applied for joint DL transmission to a CEU under a CoMP transmission of two BSs. An opportunistic NOMA scheme in a CoMP system is considered in [75], which is outperformed the conventional CoMP in terms of network capacity. The users may associate with one or multiple BSs based on their channel gains [75]. In [66], the outage probability is discussed in DL of the two-user NOMA C-RAN framework, where stochastic geometry is used for modeling the locations of both NOMA users and BSs. The authors assume that all the RRHs are cooperated to serve these NOMA users. The authors in [77] develop a PA scheme for maximizing the EE in DL NOMA networks. They analyze three transmission schemes which are: 1) The BSs coordinated to jointly serve all the users, 2) The CEUs are only served by the coordinated BSs, 3) All the users operate in a single association mode. It is worth mentioning that all the aforementioned works on CoMP NOMA studied the system performance considering homogeneous cellular networks [66, 75–77]. However, these studies cannot be projected to evaluate the system performance in HetNet. This is because the effect of the severe inter-tier interference from the high-power BSs/RRHs on the low-power small cells will negatively impact the performance of the network.

On the other hand, the application of CoMP in NOMA-based HetNet is studied in [67], [78]. A power optimization problem for CoMP NOMA sum-rate maximization is formulated in [67] among the CoMP BSs considering the data rate constraint for each user and the SIC constraint. The authors in [67] demonstrate the gain in the network SE due to

the CoMP NOMA framework in comparison with the conventional CoMP OMA system. Note that, the performance provided in [67] is a simulation-based study that does not include any tractable and/or theoretical analysis. Thus, the system performance analysis is conducted through extensive time-consuming simulations [79–83]. This situation is even more exacerbated in the case of heterogeneous and high-density RRH [81], [82]. In addition, the authors in [67] consider a static network set-up for locations of users and BSs. However, in B5G cellular networks, spatial and topological randomness is a main feature of the BSs, especially for the case of small cells [79–83]. Motivated by this, modeling the heterogeneous RRH by PPP provides a tractable and effective method to evaluate the performance of B5G cellular network using the stochastic geometry theory. Therefore, an approach based on stochastic geometry analysis is required to govern network performance and offer insightful design guidelines.

Meanwhile, in [78], the DL throughput performance in coordinated JT-NOMA-enabled HetNet is investigated by utilizing tools from stochastic geometry. In JT-NOMA, only the void BSs (The BSs that do not serve any users) are coordinated to improve the farthest NOMA user in a K users NOMA-cell. Note that, the authors study the performance of JT-NOMA assuming dense heterogeneous networks in which the small BSs are deployed with high intensity that may not be smaller than the intensity of the users. Therefore, this study is not applicable for a network with the intensity of the BSs less than the intensity of the users. Furthermore, the authors in [78] show that when the density of UEs increases compared to the density of the BSs, the performance of the JT-NOMA scheme is similar to traditional NOMA. In addition, operating all the BSs in a dense network will harm the performance of the system in terms of the network EE [84], [85]. As a result, the performance analysis in terms of the average achievable data rate for CoMP NOMA in a multi-cell network should be carefully addressed. In order to design an energy-efficient network, the RRHs selected to help the ICI-prone users should also serve Non-CoMP-UEs, i.e., CCUs, simultaneously using NOMA transmission.

Based on the above, the main limitations and research gaps can be summarized in two points as follows.

1. To the best of our knowledge, the concept of integrating location-aware CoMP transmission with NOMA network using tools from stochastic geometry has not been investigated in the literature.

2. To the best of our knowledge, the performance of combining CoMP transmission in a multi-cell C-NOMA system still remains unexplored. This motivates us to study the CEU's performance in terms of both the outage performance and the ergodic rate in joint CoMP C-NOMA cellular systems.

3.3 Contributions

First, our contribution towards tackling the first limitation is to investigate the application of CoMP in the DL transmission of NOMA-based H-CRAN to enhance NSE. In doing so, we consider a system model composed of densely deployed low-power SRRHs which are underlaid by low-density high-power MRRHs. The ICI (intra-tier interference and inter-tier interference) is explicitly modeled. The stochastic geometry approach is applied to model the deployment of both UEs and RRHs on a 2-D plane and to evaluate the overall system performance in terms of the SE. Specifically, we use PPP to model the locations of both UEs and RRHs. Against the above background, our main contributions are listed as follows,

- We investigate the amalgamation between location-aware CoMP and NOMA in the DL transmission of H-CRAN architecture by prioritizing the UEs that are vulnerable to the ICI to trigger CoMP. In doing so, we divide the users into two classes, namely, Non-CoMP-UEs, i.e., CCUs, and CoMP-UEs, i.e., CEUs.[1] Under the considered NOMA enhanced CoMP H-CRAN, we allow each CoMP-UE to receive two transmissions from the coordinating RRHs and construct a NOMA cluster with a Non-CoMP-UE from each coordinating RRH.

- We use tools from stochastic geometry to evaluate the performance of CoMP NOMA-based H-CRAN in terms of the average achievable data rate for each UE in each NOMA group and the network SE. However, two challenges are raised in this case. The first challenge is caused by the combination between the CoMP and NOMA, the expression of the corresponding `SINR` at the CEU is different from those of the traditional CoMP OMA scheme. The second challenge is how to consider the impact of the locations of the RRHs and the users on the performance of the proposed CoMP

[1]Throughout the rest of this thesis, the terms "cell-edge UEs" and "CoMP-UEs" will be used interchangeably. In addition, "cell-center UEs" and "Non-CoMP-UEs" will be used interchangeably.

NOMA-based H-CRAN. These two challenges are settled using some mathematical manipulations and rigorous derivations.

- We compare the performance of the proposed scheme with the traditional CoMP OMA scheme. Then, we analyze the performance of the two schemes under various system parameters by varying the RRH densities, cooperation threshold coefficients, operating carrier frequencies, PA coefficient allocated to Non-CoMP-UEs, and SIC decoding error percentage. Then, we highlight the main insights and show the impact of these parameters on the system behavior.

- We validate all the analytical results with Monte-Carlo simulation. Our results show that the proposed CoMP NOMA-enabled H-CRAN outperforms the traditional CoMP OMA in terms of the average achievable data rate and the network SE for different system parameter settings. In addition, we define the PA coefficients for the Non-CoMP-UE to improve fairness performance between the CoMP-UE and Non-CoMP-UE without jeopardizing the performance of the Non-CoMP-UE.

Second, our contribution towards tackling the second limitation is to evaluate the performance of integrating CoMP with C-NOMA network. Specifically, in this part, we consider a system model consisting of two adjacent cells, each equipped with one BS, and three UEs. Two types of users, namely, CoMP-UEs and Non-CoMP-UEs are considered. A Non-CoMP-UE is only served by the nearest BS. On the other hand, in this CoMP C-NOMA cellular network, a CoMP-UE is not only served by the two BSs using CoMP but also receives information from the near UE of each NOMA pair. Thus, each Non-CoMP-UE belongs to one NOMA pair, whereas the CoMP-UE belongs to two NOMA pairs. The main contributions of the second part can be summarized as follows.

- We investigate the integration between the CoMP transmission and C-NOMA technique for enhancing the performance of a CEU.

- We derive a closed-form expression for the outage probability of the CEU along with an analytical expression for its ergodic rate.

- The derived expressions are validated through Monte-Carlo simulations, where the superiority of the proposed framework compared with other multiple access schemes

proposed in the literature, i.e CoMP-assisted HD C-NOMA, CoMP NOMA, and CoMP OMA, is demonstrated.

The rest of this chapter can be organized as follows. To address the two main research gaps introduced above, we divide this chapter into two main parts, which is mapped to sections 3.4 and 3.5, respectively. The first part, which addresses the first research limitation, can be organized as follows. The NOMA-enhanced CoMP H-CRAN architecture, the association model, NOMA pairing scheme, as well as the signal model for the users are introduced in Section 3.4.1. In Section 3.4.2, the association probabilities expressions, the distance(s) distribution(s) for UE to the serving RRH(s), and the Laplace transform (LT) of the interference are discussed. In section 3.4.3, the average achievable data rates for the CoMP-UE and Non-CoMP-UE in the NOMA group as well as the average ergodic rate are obtained. Finally, the numerical results are discussed in Section 3.4.4. On the other hand, the second part of the Chapter can be summarized as follows. The network model, the signal model, the `SINR`, and the data rate analysis are presented in Section 3.5.1. In Section 3.5.2, we derive the outage probability as well as the ergodic rate for the CEU. Then, the numerical results for that model are discussed in Section 3.5.3. Finally, we conclude the Chapter.

3.4 Performance Analysis of CoMP NOMA Networks

3.4.1 System Model

In this section, we first define the two-tier H-CRAN model. Then, the categorization of the users into CoMP-UE and Non-CoMP-UE as well as the user pairing policy are described. Finally, we characterize the `SINR` for CoMP NOMA-based H-CRAN.

3.4.1.1 Network Model

We consider a DL H-CRAN consisting of two tiers of RRHs distinguished by their RRH densities, transmit powers, and path-loss exponents as shown in Fig. 3.1. The RRHs in the k-th tier ($k \in \{m, s\}$ where m refers to the MRRH tier and s refers to the SRRH tier) are modeled as independent PPP Φ_k with intensity λ_k, transmit power P_k, and path-loss exponent α_k. Different MRRHs and SRRHs can utilize the same radio channels in which a co-channel deployment scenario is assumed, i.e., full-frequency reuse [26, 47, 69, 81]. In

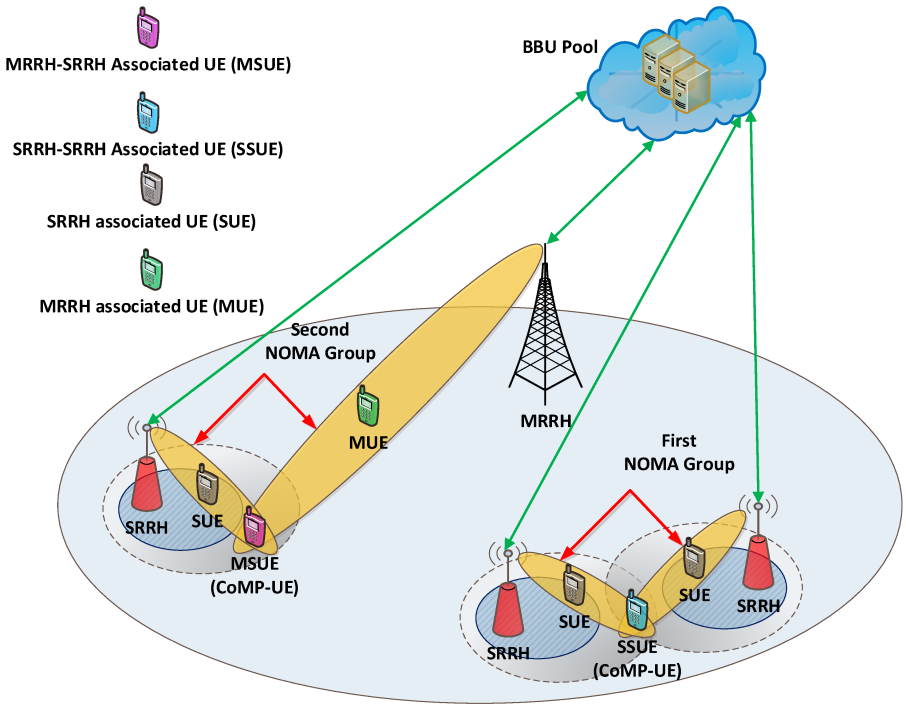


Figure 3.1: CoMP NOMA-based H-CRAN architecture.

this chapter, the NOMA study assumes a group size of two [26], [69].[2] All the UEs are assumed to follow independent PPP Φ_u of intensity λ_u. Moreover, we assume that $\lambda_u > \lambda_s$ such that a sufficient number of users can always be within each cell to form a NOMA group [69].

Due to the properties of the PPP, the system performance can be evaluated by assuming a reference point residing at the origin which is termed as "typical user", denoted by UE_0 [88–90]. Without loss of generality, according to Slivnyak theorem, the typical user can reflect the spatially averaged of the entire system's performance [88–91]. In other words, the typical user is representative of all users in the network [88–91]. In order to model the propagation, the small-scale fading is considered as Rayleigh with average unit power. While the large-scale fading is represented by $K_0\|d_{k,j}\|^{-\alpha_k}$, where $d_{k,j}$ is the distance between a typical user and the jth RRH belonging to the k-th tier and K_0 is a constant. Note that K_0 depends on the carrier frequency, e.g. $K_0 = G_t G_r(\chi/4\pi)^2$, where G_t, G_r, and χ are the transmit antenna gain, the received antenna gain and the wavelength, respectively [92], [93]. For any arbitrary point $z \in \mathbb{R}^2$, we define y_k as the location

[2]Note that a two-user NOMA group configuration is of practical interest and it has been standardized in 3GPP LTE-A [26], [69]. In addition, due to the processing complexity and latency of SIC at the receivers and the practical limitation which causes SIC error propagation, a two-user NOMA group configuration is more practical in realistic scenarios [86], [87].

of the RRH belonging to the k-th tier that provides the maximum average received signal strength (RSS) at this point. This can be described as follows.

$$y_k = \arg\max_{\mathbf{y}\in\Phi_k} \left\{ \frac{K_0 P_k}{\|y-z\|^{\alpha_k}} \right\}, \tag{3.1}$$

where $\|.\|$ denotes the Euclidean norm. Similar to [92], we utilize different path-loss exponents to consider the different propagation characteristics due to the selected carrier frequency. It is empirically observed that the path-loss exponent α_k is generally approximated as a constant between 2 to 5 that depends on the carrier frequency and the propagation environment [93].

3.4.1.2 Classification of Users Into Non-CoMP- And CoMP-UEs

Here, we define the criteria of users' classification into Non-CoMP-UEs and CoMP-UEs. In the proposed model, a CoMP-UE receives its useful signal from two RRHs where the RSS is highest from each and constructs distinct NOMA clusters with Non-CoMP-UEs served by each of these RRHs as seen in Fig. 3.1.[3] Thus, each CoMP-UE is a member in two NOMA clusters. Note that, the best two RRHs may be the nearest RRH from each tier or the two nearest SRRHs, which will be discussed later. On the other hand, a Non-CoMP-UE receives its useful signal only from either the best MRRH or the best SRRH. Therefore, it is a member of one NOMA cluster. The classification is performed based on the relation between the useful received signal from the serving-RRH to the received signal from the dominant interfering RRH.[4]

Before we proceed, let $D_{1s}, D_{1m}, D_{2s}, \text{and}, D_{\text{In}}$ be the distances from a typical user to the nearest RRH from the second tier, which is denoted as SRRH_1, the nearest RRH from the first tier, which is referred to as MRRH_1, the second nearest RRH belongs to the second tier, which is denoted as SRRH_2, and the dominant interfering RRH, respectively. For the Non-CoMP-UEs, these users are associated with the RRHs that result in the maximum RSS regardless of the corresponding tier. Meanwhile, for the CoMP-UE, it should select the best two RRHs that provide the best RSS. We denote by $\mathcal{B}$ the serving

[3]We allow two-RRH instead of $k > 2$ RRHs to serve each CEU due to the following two reasons. The first reason is to enhance the network connectivity by offering more available communication links at the RRHs, which is considered as one of the main requirements in B5G cellular networks by allowing only two communication links to be allocated for each CEU [94]. Second, increasing the number of cooperative RRHs can enhance the achievable data rate for the CEUs at the cost of leading to more fronthaul consumption. However, the practical fronthaul links are often capacity-limited, which will limit the performance gain [95].

[4]Serving-RRH is defined as the RRH that provides the maximum RSS.

RRH(s) of a typical user, which can be expressed as follows.

$$\mathcal{B} \equiv \begin{cases} \{y_m\}, & \text{if} \frac{P_m D_{1m}^{-\alpha_m}}{P_s D_{1s}^{-\alpha_s}} \geq \theta \quad \text{“MUE”} \\ \{y_s\}, & \text{if} \frac{P_s D_{1s}^{-\alpha_s}}{P_{\text{In}} D_{\text{In}}^{-\alpha_{\text{In}}}} \geq \eta \quad \text{“SUE”} \\ \{y_s, \bar{y}_s\}, & \text{if } \frac{P_s D_{1s}^{-\alpha_s}}{P_s D_{2s}^{-\alpha_s}} < \eta, \quad P_s D_{2s}^{-\alpha_s} > P_m D_{1m}^{-\alpha_m} \quad \text{“SSUE”} \\ \{y_m, y_s\}, & \text{if } \frac{1}{\theta} \leq \frac{P_s D_{1s}^{-\alpha_s}}{P_m D_{1m}^{-\alpha_m}} \leq \eta, \quad P_m D_{1m}^{-\alpha_m} > P_s D_{2s}^{-\alpha_s} \quad \text{“MSUE”} \end{cases} \tag{3.2}$$

where $\forall k \in \{m, s\}$, y_k can be given according to (3.1) and $\bar{y}_s$ is the second nearest SRRH to a typical user, which can be also given by (3.1). In addition, θ and η are the cooperation judging coefficients, in which $\theta, \eta > 1$. Note that, θ, MRRH cooperation threshold, controls the cooperation between the RRHs belonging to different tiers. While, η, SRRH cooperation threshold, is a tunable parameter to decide the cooperation between the best two SRRHs or SRRH and MRRH for serving a CEU. As a result, in Fig. 3.1, we define a region around the SRRH in which a UE is served by CoMP transmission, which is controlled by θ and η. To elaborate, if the condition for a single RRH association (case 1 and case 2) is not achieved, the serving RRH should cooperate with the dominant interfering RRH to serve that user, i.e., CoMP-UE. If the serving RRH is SRRH_1 and the interference received from SRRH_2 is larger than the interference power from the MRRH_1, SRRH_1 and SRRH_2 decide to cooperate to serve that user. However, either the interference resulting from the MRRH_1 is greater than the SRRH_2 or the MRRH is the serving RRH, the cooperation occurs between the MRRH_1 and SRRH_1. We define the CoMP set as the set of RRHs that serves the CoMP-UE. For the third case in (3.2), the CoMP set is {SRRH_1 and SRRH_2}; meanwhile for the last case is {MRRH_1 and SRRH_1}. As a conclusion, from (3.2), it is clear that the users are split into four disjoint groups: Non-CoMP MRRH user (MUE), Non-CoMP SRRH user (SUE), CoMP user associated with the nearest RRH from each tier (MSUE), and CoMP user associated with the two nearest SRRHs (SSUE) as shown in Fig. 3.1.

Note that this user-type classification can be done in practical and real cellular systems as follows. Each UE first associates with the RRH that provides the maximum RSS, which represents the conventional association scheme in Long Term Evolution (LTE) cellular networks [67, 96]. After that, each UE will decode the reference signal from the serving RRH and the dominant interfering RRH, i.e., the two RRHs that provide the

two highest RSS. Then, each UE reports these estimated RSS measures to the serving RRH [67]. If the measured RSS values are less distinctive (according to the adopted values of the RRH cooperation thresholds, i.e., θ and η), this UE will be served by the two RRHs and is considered as a CEU. Otherwise, if the user is experiencing good channel conditions from the serving RRH compared to the interfering RRH, i.e., the ratio between RSS measures from the serving RRH and the dominant interfering RRH is higher than the adopted cooperation threshold, then this UE is considered as a CCU. Thanks to the powerful computational processing resources available at the BBU pool, these measurements reported by the users can be processed in an efficient and low-latency way [55, 97].

In a NOMA network, user pairing plays an important role in the performance of the system. In the proposed system model, the CoMP-UE is a CEU with respect to its serving RRH and experiences less distinctive received power from the serving RRH and the coordinating RRH (The dominant interfering RRH). On the contrary, the Non-CoMP-UE is a CCU and has a good received `SINR` from the serving RRH. Consequently, the Non-CoMP-UE has a good channel gain and is considered as the strong user or the near user. Meanwhile, the CoMP-UE has low channel gain and is deemed as the weak user or the far user in our model. Therefore, we assume that the SIC process is occurring at the near user (Non-CoMP-UE), and therefore, it can cancel the INUI from the CoMP-UEs before decoding their own signals [67], [68], [26]. However, each CoMP-UE cannot cancel INUI from the associated Non-CoMP-UE and treat it as a noise. Based on that, the four kinds of users construct two different NOMA groups. As shown in Fig. 3.1, for the first NOMA group, one SSUE is considered along with one SUE from each RRH that serves the CoMP-UE. Meanwhile, for the other NOMA group, one MSUE is considered while one MUE and one SUE are considered from each cell that serves this MSUE.

Specifically, we consider a two-user NOMA group configuration where one CoMP-UE and one Non-CoMP-UE are clustered into one group. However, since multiple UEs may be admitted by each RRH, a promising technique to support the admitted UEs is to construct a hybrid MA technique, in which the NOMA technique is combined with traditional orthogonal MA techniques (i.e., FDMA, TDMA, etc) [68], [69], [98]. In particular, in our proposed model, we have two categories of UEs, namely, CoMP-UEs and Non-CoMP-UEs. Each member belonging to the first category is grouped with one member from

the second category to construct multiple groups, where the NOMA technique is applied within each group. Meanwhile, different NOMA groups are served through orthogonal radio resources in order to cancel the intra-cell interference [68], [69], [98].

3.4.1.3 DL NOMA Signal-to-Interference-Plus-Noise-Ratio Model

We define the PA coefficients allocated to the near user and the far user from the serving RRH j that belongs to k-th tier as $a^n_{k,j}$ and $a^f_{k,j}$, $\forall k \in \{m, s\}$, respectively, for which we have $a^n_{k,j} + a^f_{k,j} = 1$ and $a^f_{k,j} > a^n_{k,j}$ [69]. Note that, throughout the analysis, we denote the near user and the far user in each NOMA cluster as n and f, respectively. In addition, in the first NOMA group, for CoMP-UE, the serving SRRH provides a useful signal which is always greater than the coordinated RRH (the RRH that provides the second largest RSS at the user). This means that the CoMP-UE's performance is more sensitive to the PA parameters from the serving RRH than the coordinated RRH. As a result, we define a different PA factor for the near user and far user attached to the coordinated RRH as $b^n_{k,j}$ and $b^f_{k,j}$, respectively. Furthermore, for the second NOMA group, $a^n_{m,j}$ represents the PA coefficient assigned to the near user from the MRRH, and $b^n_{s,j}$ denotes the fraction of power assigned to the near user from the SRRH. In the considered model, the received interference at a user consists of three components, namely the INUI, the interference from all the MRRHs except the serving MRRH_1, and the interference coming from all the other SRRHs except the serving SRRH(s). It is worth mentioning that the Non-CoMP-UE has a good channel condition; hence it can cancel the INUI from the CoMP-UE in the same NOMA group [68]. However, SIC decoding at the Non-CoMP-UE may not be successful due to some practical limitations. Consequently, the imperfect SIC process leads to an error that propagates to the next level of decoding when the Non-CoMP-UE decodes its own message [69]. This is called the SIC error propagation, which will be considered in our analysis. In the following, we define first the `SINR` for each NOMA group.

1) For the first NOMA group: Considering the imperfect SIC process during the decoding process of the message of the CoMP-UE, the `SINR` for the Non-CoMP-UE in the jth SRRH is expressed as

$$\gamma_s{}^{\text{NC}} = \frac{c^n_{s,j} P_s h_{j,0} \|d_{s,j}\|^{-\alpha_s}}{\mu\ c^f_{s,j} P_s h_{j,0} \|d_{s,j}\|^{-\alpha_s} + I_{ms} + I_{ss} + \hat{\sigma}^2}, \tag{3.3}$$

where $\hat{\sigma}^2 = \sigma^2 / K_0$ in which σ^2 is the additive white Gaussian noise (AWGN) power, $d_{s,j}$ and $h_{j,0}$ denote the distance and the small-scale fading between the user and RRH

j belonging to the second tier, respectively. Moreover, $c^n_{s,j}$ is equal to $a^n_{s,j}$ or $b^n_{s,j}$ according to if the SRRH is the serving RRH or the coordinating RRH, respectively. $I_{ms} = P_m \sum_{i\in\Phi_m} g_{i,0}\|d_{m,i}\|^{-\alpha_m}$ and $I_{ss} = P_s \sum_{i\in\Phi_s\setminus\mathcal{B}} g_{i,0}\|d_{s,i}\|^{-\alpha_s}$ represent the interference from all the MRRHs and the SRRHs except the serving RRH $\mathcal{B}$, respectively. $\mu \in [0,1]$ is the portion of the NOMA interference due to the imperfect SIC decoding [69], where $\mu = 0$ means that the SIC is perfect, whereas $\mu = 1$ means that the Non-CoMP-UE completely fails in decoding the CoMP-UE message. On the other hand, the CoMP user (SSUE) can decode its signal directly by considering the signals coming from the Non-CoMP-UEs at both the serving RRH and the coordinated RRH as a noise. Thus, the SINR at that user can be defined as follows [44],

$$\gamma_s{}^{\mathrm{C}} = \frac{\sum_{j\in\mathcal{B}} c^f_{s,j} P_s h_{j,0}\|d_{s,j}\|^{-\alpha_s}}{\bar{I}_m + \bar{I}_s + \sum_{j\in\mathcal{B}} c^n_{s,j} P_s g_{j,0}\|d_{s,j}\|^{-\alpha_s} + \hat{\sigma}^2}, \tag{3.4}$$

where $c^f_{s,j} \triangleq a^f_{s,j}$ for the SRRH y_s and $c^f_{s,j} \triangleq b^f_{s,j}$ for the SRRH $\bar{y}_s$, and $\bar{I}_m = P_m \sum_{i\in\Phi_m} g_{i,0}\|d_{m,i}\|^{-\alpha_m}$ and $\bar{I}_s = P_s \sum_{i\in\Phi_s\setminus\mathcal{B}} g_{i,0}\|d_{s,i}\|^{-\alpha_s}$.

2) *For the second NOMA group:* Similar to the first group, the SINR for the Non-CoMP-UE attached to the MRRH_1 is

$$\gamma_m{}^{\mathrm{NC}} = \frac{a^n_{m,j} P_m h_{j,0}\|d_{m,j}\|^{-\alpha_m}}{\mu\ a^f_{m,j} P_m h_{j,0}\|d_{m,j}\|^{-\alpha_m} + I_{mm} + I_{sm} + \hat{\sigma}^2}, \tag{3.5}$$

where $I_{mm} = P_m \sum_{i\in\Phi_m\setminus\mathcal{B}} g_{i,0}\|d_{m,i}\|^{-\alpha_m}$ and $I_{sm} = P_s \sum_{i\in\Phi_s} g_{i,0}\|d_{s,i}\|^{-\alpha_s}$. Moreover, the SINR for MSUE can be written as follows [44],

$$\gamma_{ms}{}^{\mathrm{C}} = \frac{\sum_{k\in\{m,s\}} \sum_{j\in\mathcal{B}} c^f_{k,j} P_k h_{j,0}\|d_{k,j}\|^{-\alpha_k}}{\tilde{I}_m + \tilde{I}_s + \sum_{k\in\{m,s\}} \sum_{j\in\mathcal{B}} c^n_{k,j} P_k g_{j,0}\|d_{k,j}\|^{-\alpha_k} + \hat{\sigma}^2}, \tag{3.6}$$

where $c^f_{m,j}$ and $c^f_{s,j}$ are equal to $a^f_{m,j}$ and $b^f_{s,j}$ for the MRRH and the SRRH, respectively. $\tilde{I}_m = P_m \sum_{i\in\Phi_m\setminus y_m} g_{i,0}\|d_{m,i}\|^{-\alpha_m}$ and $\tilde{I}_s = P_s \sum_{i\in\Phi_s\setminus y_s} g_{i,0}\|d_{s,i}\|^{-\alpha_s}$. Finally, the SINR for the Non-CoMP-UE served by the SRRH can be calculated as in (3.3).

3.4.2 Association Model, Distance Model, and Laplace transform of The Interference

In this section, we derive the association probabilities for the four kinds of users. Then, the probability density function (PDF) of the distance(s) between the serving RRH(s) and a typical user is derived. Finally, the LT of the ICI at each user type is defined.

3.4.2.1 Association Model

As described in Section II-B, the users are classified into four disjoint groups. Two groups associated with only one RRH, i.e., MRRH-associated UEs and SRRH-associated UEs. While, the other two groups, CoMP-UEs, are the MSUEs and SSUEs. Let $\mathcal{Q}_m, \mathcal{Q}_s, \mathcal{Q}_{ms}$, and $\mathcal{Q}_{ss}$ denote the association probabilities that a typical user associates with the MRRH, the SRRH, the nearest RRH from each tier, and the best two SRRHs, respectively. For simplicity, we denote $P_{ij} = \frac{P_i}{P_j}, \alpha_{ij} = \frac{\alpha_i}{\alpha_j}$, and $\zeta_i = \frac{2}{\alpha_i}$ in the following parts of the work.

Lemma 3.1. *The cell association probabilities that a typical user associates with the $MRRH_1$, $SRRH_1$, $SRRH_1$ and $SRRH_2$, and $MRRH_1$ and $SRRH_1$ are respectively given by (3.7)-(3.10).*

$$\mathcal{Q}_m = 2\pi\lambda_m \int_{\mathbb{R}_+} r \exp\left[-\pi\left(\lambda_m r^2 + \lambda_s \left(\theta P_{sm}\right)^{\zeta_s} r^{2\alpha_{ms}}\right)\right] dr, \tag{3.7}$$

$$\mathcal{Q}_s = \frac{2\pi\lambda_s}{\eta^{\zeta_s}} \int_{\mathbb{R}_+} r \exp\left[-\pi\left(\lambda_s \eta^{\zeta_s} r^2 + \lambda_m \left(\eta P_{ms}\right)^{\zeta_m} r^{2\alpha_{sm}}\right)\right] dr, \tag{3.8}$$

$$\mathcal{Q}_{ss} = 2(\pi\lambda_s)^2 \left(1 - \frac{1}{\eta^{\zeta_s}}\right) \int_{\mathbb{R}_+} r^3 exp\left[-\pi\left[\lambda_m \left(P_{ms}\right)^{\zeta_m} r^{2\alpha_{sm}} + \lambda_s r^2\right]\right] \; dr, \tag{3.9}$$

$$\mathcal{Q}_{ms} = 1 - \mathcal{Q}_m - \mathcal{Q}_s - \mathcal{Q}_{ss}. \tag{3.10}$$

Proof. See appendix A1. ■

It can be observed that the judging cooperation thresholds, i.e., η and θ, affect the association probabilities. Increasing the judging cooperation thresholds reduces the single-RRH association probability ($\mathcal{Q}_m$ and $\mathcal{Q}_s$), which reflects on increasing the probability that a typical user triggers CoMP mode.

3.4.2.2 Distance Model

Here, we derive the PDFs for the distance(s) between a typical user and its serving RRH(s) in the predefined four different cases. We define $f_U(u)$ as the PDF of the distance between the MUE and its serving MRRH, $f_V(v)$ as the PDF of the distance between the SUE and its serving SRRH, $f_{U,V}(u, v)$ as the joint PDF of the distances between MSUE and its two strongest RRHs, i.e., the nearest RRH from each tier, and $f_{V,W}(v, w)$ is the joint PDF between SSUE and the two nearest SRRHs, which receives the strongest power from each.

Lemma 3.2. *The PDF of the distance(s) between a typical user and its connected RRH(s) are respectively given by (3.11)-(3.14).*

$$f_U(u) = \frac{2\pi\lambda_m u}{\mathcal{Q}_m} \exp\left[-\pi\left[\lambda_m u^2 + \lambda_s \left(\theta P_{sm}\right)^{\zeta_s} u^{2\alpha_{ms}}\right]\right], \tag{3.11}$$

$$f_V(v) = \frac{2\pi\lambda_s v}{\mathcal{Q}_s} \exp\left[-\pi\left(\lambda_s(\eta)^{\zeta_s} v^2 + \lambda_m \left(\eta P_{ms}\right)^{\zeta_m} v^{2\alpha_{sm}}\right)\right], \tag{3.12}$$

$$f_{U,V}(u,v) = \begin{cases} \frac{(2\pi)^2\lambda_m\lambda_s uv}{\mathcal{Q}_{ms}} \exp\left[-\pi\left[\lambda_m u^2 + \lambda_s v^2\right]\right], & (u,v) \in \Delta_1, \\ \frac{(2\pi)^2\lambda_m\lambda_s uv}{\mathcal{Q}_{ms}} \exp\left[-\pi\left[\lambda_s (P_{sm})^{\zeta_s} u^{2\alpha_{ms}} + \lambda_m u^2\right]\right], & (u,v) \in \Delta_2, \end{cases} \tag{3.13}$$

$$f_{V,W}(v,w) = \frac{(2\pi\lambda_s)^2 vw}{\mathcal{Q}_{ss}} \exp\left[-\pi\left[\lambda_s w^2 + \lambda_m (P_{ms})^{\zeta_m} w^{2\alpha_{sm}}\right]\right], \quad (v,w) \in \Theta, \tag{3.14}$$

where

$$\Delta_1 = \{(u,v) : u \geq 0 \text{ and } (P_{sm})^{\frac{1}{\alpha_s}} u^{\alpha_{ms}} \leq v \leq (\theta P_{sm})^{\frac{1}{\alpha_s}} u^{\alpha_{ms}}\}, \tag{3.15}$$

$$\Delta_2 = \{(u,v) : u \geq 0 \text{ and } \left(\frac{P_{sm}}{\eta}\right)^{\frac{1}{\alpha_s}} u^{\alpha_{ms}} \leq v \leq (P_{sm})^{\frac{1}{\alpha_s}} u^{\alpha_{ms}}\}, \tag{3.16}$$

$$\Theta = \{(v,w) : w \geq 0 \text{ and } \left(\frac{1}{\eta}\right)^{\frac{1}{\alpha_s}} w \leq v \leq w\}. \tag{3.17}$$

Proof. Proof: The proof is given in appendix A2. ■

It can be observed that the joint PDF for the distances between a typical user and the nearest RRH from each tier has two different expressions as indicated in (3.13) based on the value of u and v. This is because in the first region Δ_1 regardless of the distance to the second nearest SRRH a typical user associates with the MRRH_1 and SRRH_1. This is because the serving RRH is the MRRH. As a result, cooperation must be triggered between the MRRH_1 and SRRH_1. Meanwhile, for $(u,v) \in \Delta_2$, in this case, the serving RRH is the SRRH and the coordinated RRH is the MRRH. As long as the distance from the MRRH to a typical user is bounded by region Δ_2, the SRRH_1 is the serving RRH and the MRRH_1 is the coordinated RRH. This is because the received signal from the MRRH_1 is greater than the received power from the SRRH_2.

3.4.2.3 Laplace Transform of The Interference

Here, we analyze the LT of the ICI for the predefined four kinds of users. We denote $I^{\text{tot}} = I_m + I_s$ as the overall interference to a typical user from the two tiers. Note that, the LT of I^{tot} is $\mathcal{L}_{I^{\text{tot}}}(s) = \mathcal{L}_{I_m}(s)\mathcal{L}_{I_s}(s)$. Before calculating the average achievable data rate for each defined user type, the LT for the different inter-tier interference and intra-tier interference in the different association modes need to be evaluated. We first evaluate the LT of the interference at the Non-CoMP-UE in the following Lemma.

Lemma 3.3. *The LT of the interference from the SRRHs and MRRHs to the Non-CoMP-UE (SUE and MUE) can be expressed as*

$$\mathcal{L}_{I_{ij}}(s) = \exp\left[-\pi\lambda_i(sP_i)^{\zeta_i}\mathcal{Z}\left(\left(\frac{\varphi_{ij}}{sP_j}\right)^{\zeta_i}\vartheta_j^{2\alpha_{ji}}, \alpha_i\right)\right], \tag{3.18}$$

where

$$\varphi_{ij} = \begin{cases} 1, & i = m \quad and \quad j = m, \\ \theta, & i = s \quad and \quad j = m, \\ \eta, & otherwise, \end{cases}$$

$$\vartheta_j = \begin{cases} u, & j = m, \\ v, & j = s, \end{cases} \qquad \mathcal{Z}(\psi, \alpha) = \int_{\psi}^{\infty} \frac{1}{1 + u^{\frac{\alpha}{2}}} du. \tag{3.19}$$

Proof. See appendix A3. ■

Lemma 3.4. *The LT of the interference from the SRRHs and MRRHs to the CoMP-UE associated with the two nearest SRRHs (SSUE) can be expressed as*

$$\mathcal{L}_{\tilde{I}_k} = \exp\left[-\pi\lambda_k(sP_k)^{\zeta_k}\mathcal{Z}\left(\left(\frac{1}{sP_s}\right)^{\zeta_k} w^{2\alpha_{sk}}, \alpha_k\right)\right]. \tag{3.20}$$

Proof. See appendix A4. ■

Lemma 3.5. *The LT of the interference from the SRRHs and MRRHs to the CoMP-UE associated with the best RRH from each tier depends on which RRH (MRRH or SRRH) is the serving RRH. If the MRRH is the serving RRH and the SRRH is the coordinated RRH, the LT for the interference can be expressed as*

$$\mathcal{L}_{\tilde{I}_k}(s) = \exp\left[-\pi\lambda_k(sP_k)^{\zeta_k}\mathcal{Z}\left(\left(\frac{1}{sP_k}\right)^{\zeta_k}\vartheta_k^2, \alpha_k\right)\right]. \tag{3.21}$$

If the SRRH is the serving RRH and the MRRH is the coordinated RRH, the LT of the interference is expressed as follows

$$\mathcal{L}_{\tilde{I}_m}(s) = \mathcal{L}_{\tilde{I}_k}(s)\big|_{k=m}$$

$$\mathcal{L}_{\tilde{I}_s}(s) = \exp\left[-\pi\lambda_s(sP_s)^{\zeta_s}\mathcal{Z}\left(\left(\frac{1}{sP_m}\right)^{\zeta_s} u^{2\alpha_{ms}}, \alpha_s\right)\right]. \tag{3.22}$$

Proof. See appendix A5. ■

3.4.3 Performance Analysis

In this section, we focus on analyzing the system's performance. In order to evaluate the performance of the proposed CoMP transmission for the NOMA-based H-CRAN framework, we calculate the average achievable rate for each NOMA user. Then, we evaluate the *average ergodic rate* for the proposed framework.

3.4.3.1 Average Achievable Rate for Non-CoMP- And CoMP-UEs

In this subsection, we find an expression for the average achievable rate for each user (Non-CoMP user and CoMP user) in the two different NOMA groups which represents the user's spectral efficiency. Note that, if the CoMP user can decode its message, the Non-CoMP user can definitely decode the CoMP user's message [68], [26], [69], [99]. This is because the Non-CoMP user has a better channel condition than the CoMP user [68]. According to [26], [69], we define the average achievable data rate at a typical user when served by RRH(s) for the proposed scheme as

$$\mathcal{R}_s \triangleq \mathbb{E}_v \left[\mathbb{E}_{\gamma_s^{\mathrm{NC}}} \left[\log(1+\gamma_s^{\mathrm{NC}}(v) \right] \right], \tag{3.23}$$

$$\mathcal{R}_{ss} \triangleq \mathbb{E}_{v,w} \left[\mathbb{E}_{\gamma_s^{\mathrm{C}}} \left[\log(1+\gamma_s^{\mathrm{C}}(v,w) \right] \right], \tag{3.24}$$

$$\mathcal{R}_m \triangleq \mathbb{E}_u \left[\mathbb{E}_{\gamma_m^{\mathrm{NC}}} \left[\log(1+\gamma_m^{\mathrm{NC}}(u) \right] \right], \tag{3.25}$$

$$\mathcal{R}_{ms} \triangleq \mathbb{E}_{u,v} \left[\mathbb{E}_{\gamma_{ms}^{\mathrm{C}}} \left[\log(1+\gamma_{ms}^{\mathrm{C}}(u,v) \right] \right], \tag{3.26}$$

where $\mathbb{E}[.]$ is the statistical average operator. This ergodic rate can measure the long-term achievable data rate averaged over all channels and interference. The average achievable data rate is calculated by considering a typical user at a given distance(s) from its serving RRH(s).[5] Then, this rate is averaged over these distance(s). In the following theorems, the expressions for the average achievable rate of the two paired UEs in each NOMA group can be obtained.

Theorem 3.1. *The average achievable data rate of the Non-CoMP-UE and the CoMP-UE for the first NOMA group are, respectively, given by*

$$\mathcal{R}_s = \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \left[\Psi(c_{s,j}^n + \mu\ c_{s,j}^f) - \Psi(\mu\ c_{s,j}^f) \right] f_V(v) dv dt, \tag{3.27}$$

$$\mathcal{R}_{ss} = \int_{\mathbb{R}_+} \int_{r_{1s}} \int_{r_{2s}} \left[\Phi(P_s, P_s) - \Phi(a_{s,j}^n P_s, b_{s,j^{'}}^n P_s) \right] f_{V,W}(r_{1s}, r_{2s}) dr_{1s} dr_{2s} dt, \tag{3.28}$$

where,

$$\Psi(\omega) = \exp\left[-\frac{\beta(\omega) v^{\alpha_s} \hat{\sigma}^2}{P_s} \right] \prod_{k \in \{m,s\}} \mathcal{L}_{I_{sk}} \left(\frac{\beta(\omega) v^{\alpha_s}}{P_s} \right), \tag{3.29}$$

$$\Phi(\Omega_1, \Omega_2) = \sum_{b=\{1,2\}} \frac{\exp\left[-\frac{\delta(t)\hat{\sigma}^2}{\Omega_b r_{bs}^{-\alpha_s}} \right] \mathbb{E}_{\bar{I}_k} \left[\exp\left[-\frac{\delta(t)(\bar{I}_m + \bar{I}_s)}{\Omega_b r_{bs}^{-\alpha_s}} \right] \right]}{\prod_{j \in \{1,2\}, j \neq b} \left(1 - \frac{\Omega_j r_{js}^{-\alpha_s}}{\Omega_b r_{bs}^{-\alpha_s}} \right)}, \tag{3.30}$$

and $\delta(t) = 2^t - 1$ *and* $\beta(\omega) = \frac{\delta(t)}{\omega}, r_{1s} \triangleq v$, *and* $r_{2s} \triangleq w$.

Proof. The average achievable data rate for the Non-CoMP-UE in the first NOMA group

[5] Note that the average achievable data rate is measured by bits/sec/Hz [26, 47, 69, 90]. According to the definition in [26, 47, 69, 90], this average achievable data rate is independent of the user intensity λ_u.

can be evaluated as follows

$$\begin{aligned}\mathcal{R}_s &= \mathbb{E}\left[\log_2\left[1+\gamma_s^{\mathrm{NC}}(v)\right]\right] \\ &= \mathbb{E}\left[\log_2\left[1+\frac{c_{s,j}^n P_s h_{j,0}\|d_{s,j}\|^{-\alpha_s}}{\mu\ c_{s,j}^f P_s h_{j,0}\|d_{s,j}\|^{-\alpha_s}+I_{ms}+I_{ss}+\hat{\sigma}^2}\right]\right] \\ &\overset{(a)}{=} \underbrace{\mathbb{E}\left[\log_2\left[1+\frac{(c_{s,j}^n+\mu\ c_{s,j}^f)P_s h_{j,0}v^{-\alpha_s}}{I_{ms}+I_{ss}+\hat{\sigma}^2}\right]\right]}_{\mathbf{P1}} - \underbrace{\mathbb{E}\left[\log_2\left[1+\frac{\mu\ c_{s,j}^f P_s h_{j,0}v^{-\alpha_s}}{I_{ms}+I_{ss}+\hat{\sigma}^2}\right]\right]}_{\mathbf{P2}},\end{aligned} \tag{3.31}$$

where (a) obtains after some mathematical manipulations and substituting $\|d_{s,j}\| = v$. By evaluating **P1** and **P2**, we can calculate the average achievable data rate for the Non-CoMP-UE. This can be achieved by finding an expression for this term,

$$\begin{aligned}\Psi(\omega) &= \mathbb{E}\left[\log_2\left[1+\frac{\omega\ c_{s,j}^f P_s h_{j,0}v^{-\alpha_s}}{I_{ms}+I_{ss}+\hat{\sigma}^2}\right]\right], \\ &= \int_{\mathbb{R}_+}\left[\int_{\mathbb{R}_+}\mathbb{P}\left[\log_2\left(1+\frac{\omega P_s h_{j,0}v^{-\alpha_s}}{I_{ms}+I_{ss}+\hat{\sigma}^2}\right)>t)\right]f_V(v)dv\right]dt, \\ &= \int_{\mathbb{R}_+}\left[\int_{\mathbb{R}_+}\underbrace{\mathbb{P}\left[\frac{\omega P_s h_{j,0}v^{-\alpha_s}}{I_{ms}+I_{ss}+\hat{\sigma}^2}>\delta(t)\right]}_{\mathbf{P3}}f_V(v)dv\right]dt,\end{aligned} \tag{3.32}$$

where, **P3** can be calculated as

$$\begin{aligned}\mathbf{P3} =& \mathbb{P}\left[\frac{P_s h_{j,0}v^{-\alpha_s}}{I_{ms}+I_{ss}+\hat{\sigma}^2}\geq\beta(\omega)\right], \\ =& \mathbb{P}\left[h_{j,0}\geq\beta(\omega)\frac{I_{ms}+I_{ss}+\hat{\sigma}^2}{P_s v^{-\alpha_s}}\right], \\ \overset{(a)}{=}& \exp\left[-\frac{\beta(\omega)v^{\alpha_s}\hat{\sigma}^2}{P_s}\right]\mathbb{E}_{I_{ms}}\left[-\frac{\beta(\omega)v^{\alpha_s}}{P_s}i_{ms}\right]\mathbb{E}_{I_{ss}}\left[-\frac{\beta(\omega)v^{\alpha_s}}{P_s}i_{ss}\right], \\ =& \exp\left[-\frac{\beta(\omega)v^{\alpha_s}\hat{\sigma}^2}{P_s}\right]\mathcal{L}_{I_{ms}}\left(\frac{\beta(\omega)v^{\alpha_s}}{P_s}\right)\mathcal{L}_{I_{ss}}\left(\frac{\beta(\omega)v^{\alpha_s}}{P_s}\right),\end{aligned} \tag{3.33}$$

where (a) follows from the fact that $h\sim\exp(1)$ and $\mathcal{L}_{I_{ss}}, \mathcal{L}_{I_{ms}}$ are the LT of the interference, which can be calculated through *Lemma 3*. Then, by plugging (3.18) into (3.33), we can obtain the average achievable data rate for that user. On the other hand, the average achievable data rate for the CoMP-UE (This user is associated with the two nearest SRRHs), which is denoted by $\mathcal{R}_{ss}$, can be calculated as follows,

$$\mathcal{R}_{ss} = \mathbb{E}\left[\log\left(1+\frac{a_{s,j}^f P_s h_{j,0}v^{-\alpha_s}+b_{s,j}^f P_s h_{j,0}w^{-\alpha_s}}{\bar{I}_m+\bar{I}_s+a_{s,j}^n P_s h_{j,0}v^{-\alpha_s}+b_{s,j}^n P_s h_{j,0}w^{-\alpha_s}+\hat{\sigma}^2}\right)\right]. \tag{3.34}$$

Note that for any $x, y, z > 0$, we have the following property,

$$\log\left(1+\frac{x}{y+z}\right) = \log\left(1+\frac{x+y}{z}\right) - \log\left(1+\frac{y}{z}\right). \tag{3.35}$$

Consider $y = a^n_{s,j} P_s h_{j,0} v^{-\alpha_s} + b^n_{s,j} P_s h_{j^{`},0} w^{-\alpha_s}$ and $z = \bar{I}_m + \bar{I}_s + \hat{\sigma}^2$, (3.34) can be rewritten as follows,

$$\mathcal{R}_{ss} = \mathbb{E}\left[\log\left(1 + \frac{P_s h_{j,0} v^{-\alpha_s} + P_s h_{j^{`},0} w^{-\alpha_s}}{\bar{I}_m + \bar{I}_s + \hat{\sigma}^2}\right) - \log\left(1 + \frac{a^n_{s,j} P_s h_{j,0} v^{-\alpha_s} + b^n_{s,j} P_s h_{j,0} w^{-\alpha_s}}{\bar{I}_m + \bar{I}_s + \hat{\sigma}^2}\right)\right]. \tag{3.36}$$

In order to have an expression for the user average achievable rate, we need to first find an expression for this term,

$$\begin{aligned}\Phi(\Omega_1, \Omega_2) &= \mathbb{E}\left[\log\left(1 + \frac{\Omega_1 h_{j,0} v^{-\alpha_s} + \Omega_2 h_{j,0} w^{-\alpha_s}}{\bar{I}_m + \bar{I}_s + \hat{\sigma}^2}\right)\right] \\ &= \int_{\mathbb{R}_+}\int_w \int_v \underbrace{\mathbb{P}\left[\frac{\Omega_1 h_{j,0} v^{-\alpha_s} + \Omega_2 h_{j,0} w^{-\alpha_s}}{\bar{I}_m + \bar{I}_s + \hat{\sigma}^2} \geq \delta(t)\right]}_{\mathbf{P2}} f_{V,W}(v,w) dv dw dt, \end{aligned} \tag{3.37}$$

where **P2** can be calculated as follows,

$$\begin{aligned}\mathbf{P2} =& \mathbb{P}\left[\frac{\Omega_1 h_{j,0} v^{-\alpha_s} + \Omega_2 h_{j^{`},0} w^{-\alpha_s}}{\bar{I}_m + \bar{I}_s + \hat{\sigma}^2} \geq \delta(t)\right], \\ \overset{(a)}{=}& \mathbb{E}_{I_k}\left[\sum_{b=\{1,2\}} \frac{\exp\left[-\frac{\delta(t)(\bar{I}_m + \bar{I}_s + \hat{\sigma}^2)}{\Omega_b r_b^{-\alpha_s}}\right]}{\prod_{j\in\{1,2\}, j\neq b}\left(1 - \frac{\Omega_j r_j^{-\alpha_s}}{\Omega_b r_b^{-\alpha_s}}\right)}\right], \\ \overset{(b)}{=}& \sum_{b=\{1,2\}} \frac{\exp\left[-\frac{\delta(t)\hat{\sigma}^2}{\Omega_b r_b^{-\alpha_s}}\right] \mathbb{E}_{\bar{I}_k}\left[\exp\left[-\frac{\delta(t)(\bar{I}_m + \bar{I}_s)}{\Omega_b r_b^{-\alpha_s}}\right]\right]}{\prod_{j\in\{1,2\}, j\neq b}\left(1 - \frac{\Omega_j r_j^{-\alpha_s}}{\Omega_b r_b^{-\alpha_s}}\right)}, \end{aligned} \tag{3.38}$$

where (a) is obtained from the distribution of the sum of Exponential random variabless (RVs) [100]. The LT of the interference can be calculated through *Lemma 4*. Thus, by substituting (3.20) into (3.38), the average achievable rate for the CoMP-UE can be obtained. ■

Theorem 3.2. *The average achievable data rate for the Non-CoMP-UE (MUE) and the CoMP-UE (MSUE) for the second NOMA group are respectively given by (3.39)-(3.40), where $r_{1m} \triangleq u$ and $r_{1s} \triangleq v$. In addition, $\Psi(.)$ and $\Phi(.,.)$ can be obtained from Theorem 1.*

$$\mathcal{R}_m = \int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \left[\Psi(c^n_{m,j} + \mu\ c^f_{m,j}) - \Psi(\mu\ c^f_{m,j})\right] f_U(u) du dt, \tag{3.39}$$

$$\mathcal{R}_{ms} = \int_{\mathbb{R}_+}\int_{r_{1m}}\int_{r_{1s}} \left[\Phi(P_m, P_s) - \Phi(a^n_{m,j} P_m, b^n_{s,j} P_s)\right] f_{U,V}(r_{1m}, r_{1s}) dr_{1s} dr_{1m} dt. \tag{3.40}$$

Proof. Following the same procedures that are given in *Theorem 1*, we can obtain the results in *Theorem 2*. Note that the derivations of the rate expressions for the two different NOMA groups are dissimilar in two particular points. The first point is the parameters associated with each tier where the MRRHs and SRRHs are different in the transmit power, the RRHs density, and the path-loss exponent. The second point is that the calculation of the aggregated interference generated by all the RRHs except the serving

Table 3.1: Simulation Parameters

Parameter	Symbol	Value
MRRH transmit power	P_m	46 dBm
SRRH transmit power	P_s	30 dBm
MRRH path-loss exponent	α_m	3.5
SRRH path-loss exponent	α_s	4
MRRH cooperation coefficient	θ	4 dB
SRRH cooperation coefficient	η	3 dB
MRRH intensity	λ_m	$(500^2\pi)^{-1}$
SRRH intensity	λ_s	$50 * (500^2\pi)^{-1}$
Noise power spectral density	σ^2	-174 dBm/Hz
Monte Carlo simulation trials	T	10^6

RRH(s) and the distance distribution to the serving RRH(s) are different. ■

3.4.3.2 Average Ergodic Rate for The Proposed CoMP NOMA-Based H-CRAN

In this subsection, we define the *average ergodic rate* to evaluate the system performance in terms of the spectral efficiency for the proposed CoMP transmission for NOMA-based H-CRAN [26, 47, 81]. This can be expressed as,

$$\mathcal{R} = \mathcal{Q}_m\mathcal{R}_m + \mathcal{Q}_s\mathcal{R}_s + \mathcal{Q}_{ms}\mathcal{R}_{ms} + \mathcal{Q}_{ss}\mathcal{R}_{ss}, \tag{3.41}$$

where $\mathcal{Q}_m$, $\mathcal{Q}_s$, $\mathcal{Q}_{ms}$, and $\mathcal{Q}_{ss}$ are the association probabilities defined in *Lemma 1.* $\mathcal{R}_s$ and $\mathcal{R}_{ss}$ are the average achievable data rates of the Non-CoMP-UE and CoMP-UE for the first NOMA group, which are determined using *Theorem 1.* On the other hand, $\mathcal{R}_m$ and $\mathcal{R}_{ms}$ are the average achievable data rates of the Non-CoMP-UE and CoMP-UE belonging to the second NOMA group calculated in *Theorem 2.*

3.4.4 Simulation And Discussion Results

In this section, we numerically evaluate the system performance based on both simulations and analytical models. As comparisons to the proposed CoMP NOMA scheme, CoMP OMA results from simulations are also obtained. The bandwidth of one resource unit is normalized to one. For a fair comparison to the proposed NOMA, OMA allocates half a unit of resource for each of the two users in the NOMA group [69]. For the DL OMA technique, each RRH transmits with half of its power budget to each user on its owned resource [69]. The system parameters for the simulations are listed in Table 3.1 [47,81]. For simplicity, we denote the PA coefficient of NOMA user in each tier as $a^n_{k,j} = a^n_k, b^n_{k,j} = b^n_k$,

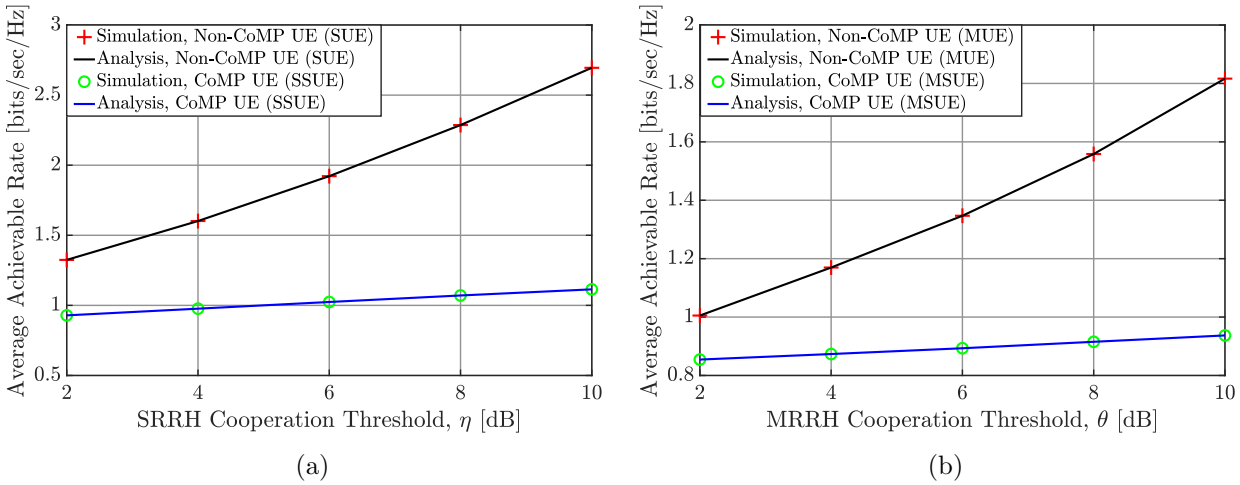


Figure 3.2: Average achievable data rate for the CoMP-UE and Non-CoMP-UE in the different NOMA group.

and $c^n_{k,j} = c^n_k$,. In addition, we assume that the PA coefficient of NOMA users from the serving RRH and coordinated RRH are equal. Finally, μ is set to be 0, i.e., perfect SIC, otherwise it will be mentioned.

3.4.4.1 Accuracy of the Analysis

In order to validate our analysis, the average achievable data rate for each user in each NOMA group is obtained from both the simulation and the analysis. Monte Carlo simulations, which is obtained via MATLAB, are used where we consider a simulation area of 10km x 10km. In each Monte-Carlo trial, the RRH locations, user locations, and channel gains are independently generated. It can be seen that from Fig. 3.2 the analytical results perfectly match the simulation results, which validates the accuracy of the obtained analytical results. As a result, from now onwards, the analytical expressions are used to evaluate the system's performance. Moreover, it seems that the variations of the two cooperation thresholds have a significant impact on the performance of the Non-CoMP-UE than the CoMP-UE. The main purpose of these cooperation thresholds, i.e. η or θ, is to determine which users associated with a single RRH (SRRH only or MRRH only) should be offloaded to the CoMP transmission with the two best SRRHs or with the nearest RRH from each tier. It can also define the coverage area of the RRH. Consequently, when the cooperation threshold increases, the coverage area of the RRHs gets reduced, and therefore, more users with poor `SINR` are offloaded to the CoMP transmission mode and the users with the best channel gains are operated in single association mode. As a result,

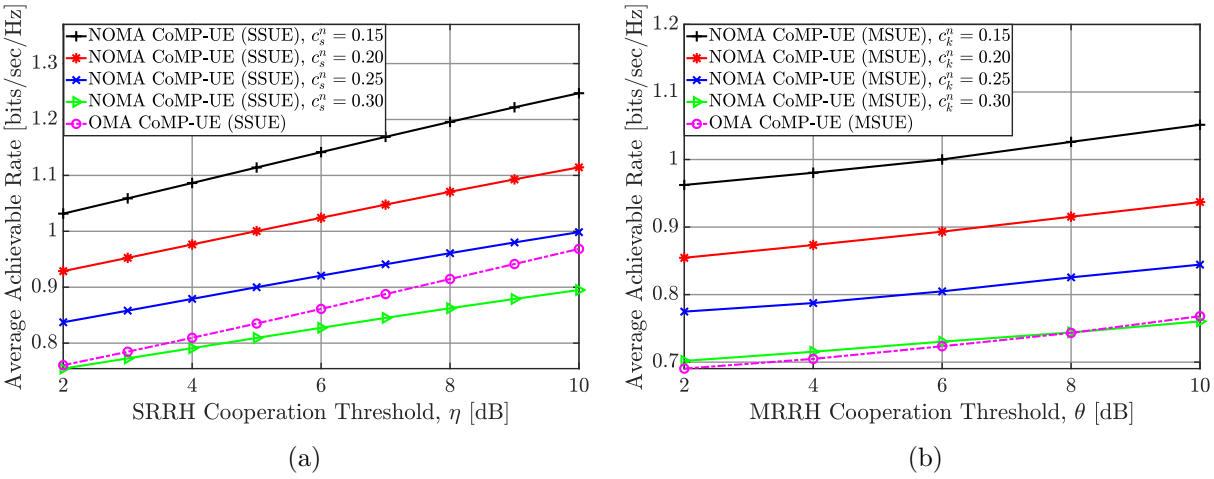


Figure 3.3: (a) and (b) compare the SSUE's and MSUE's average achievable rate for the proposed CoMP NOMA and CoMP OMA with varying the value of η and θ, respectively, for different PA setting.

increasing the cooperation thresholds results in shrinking the RRH's coverage area and so offloading most of the Non-CoMP-UE with poor SINR to CoMP transmission. This explains why the improvement for the average achievable data rate for the Non-CoMP-UE is higher than the one of the CoMP-UE.

3.4.4.2 Average Achievable Data Rate versus Interference Judging Coefficients

Fig. 3.3(a) depicts the average achievable data rate for the CoMP user (SSUE) in the first NOMA group versus the SRRH cooperation coefficient, i.e., η, for different PA coefficients assigned to Non-CoMP user belonging to the same NOMA group. Meanwhile, Fig. 3.3(b) illustrates the average achievable data rate for the CoMP user (MSUE) in the second NOMA group versus the MRRH cooperation coefficient, i.e., η, for different PA coefficients allocated Non-CoMP user in the same group. It can be observed that the performance of the CoMP NOMA scheme outperforms the performance of the conventional CoMP OMA when the PA coefficient is properly selected. For instance, in Fig. 3.3(a) when $\eta = 4$ dB and the power control coefficient assigned to the Non-CoMP-UE in the same NOMA group is reduced from 0.25 to 0.15, the proposed CoMP NOMA can achieve about 8.6% to 34.2% better than the traditional CoMP OMA. Meanwhile, in Fig. 3.3(b) when $\theta = 4$ dB and the power control coefficient assigned to the Non-CoMP-UE in the same NOMA group is reduced from 0.25 to 0.15, the proposed CoMP NOMA can achieve about 11.7% to 39% better than the traditional CoMP OMA. We also observe that the average achievable

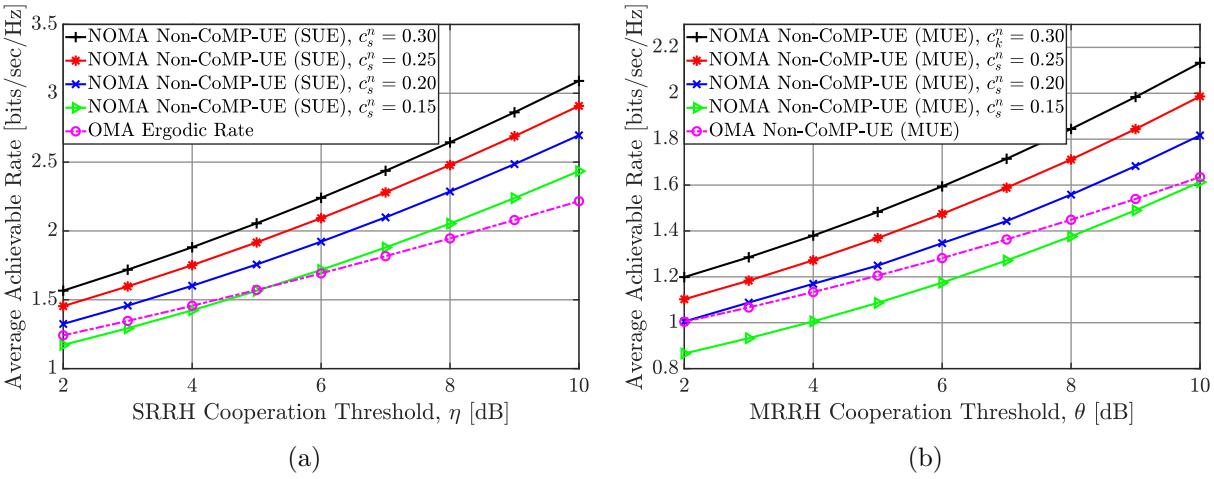


Figure 3.4: (a) and (b) compare the SUE's and MUE's average achievable rate for the proposed CoMP NOMA and CoMP OMA with varying the value of the SRRH cooperation threshold η and the MRRH cooperation threshold θ, respectively, for different PA settings.

rate increases when the interference judging coefficient (η and θ) increases. This behavior can be explained as follows: higher interference judging coefficient values encourage more users to trigger CoMP transmission, and hence, improve the useful received signal power along with reducing the received interference power.

On the other hand, Fig. 3.4(a) and Fig. 3.4(b) describe the average achievable data rate for the Non-CoMP-UE in both NOMA group versus the interference judging coefficients (η and θ) for different PA coefficients. It can be seen that the average achievable rate improves when the interference judging coefficient values increase. This is because increasing η and θ lead to only the strongest users will be Non-CoMP-UEs, and hence, Non-CoMP-UEs will achieve high improvement in terms of the average data rate. In addition, we can see from Fig.3.4(a) and Fig. 3.4(b) that when $\eta = 4$ dB ($\theta = 4$dB) and the PA coefficients allocated to the Non-CoMP-UE increases from 0.2 to 0.3, the proposed CoMP NOMA technique provides a gain of 10% to 29% for the SUE (a gain of 3.2% to 21.7% for the MUE) over the conventional CoMP OMA scheme, respectively.

3.4.4.3 Achievable data Rate Versus The Power Allocation Coefficient

In this subsection, we show how the average achievable data rate for the CoMP-UE and Non-CoMP-UE in the first NOMA group changes with increasing the PA coefficient of the Non-CoMP user. As shown in Fig. 3.5, as the power control coefficient increases, the performance of the Non-CoMP-UE improves and the performance of the CoMP-UE

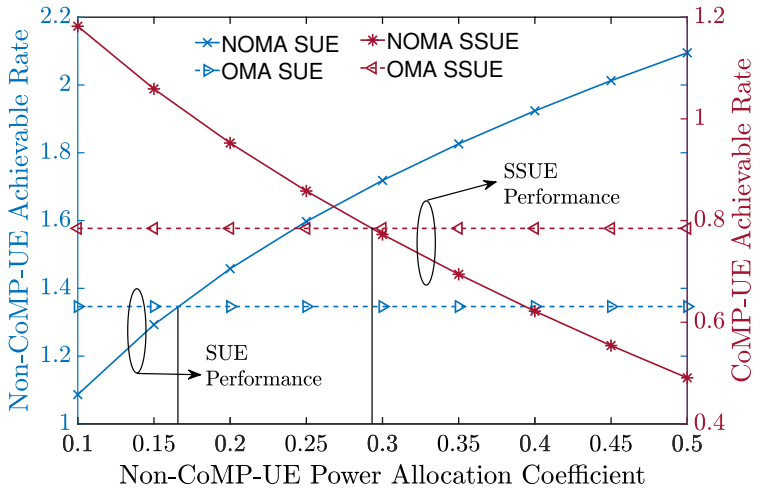
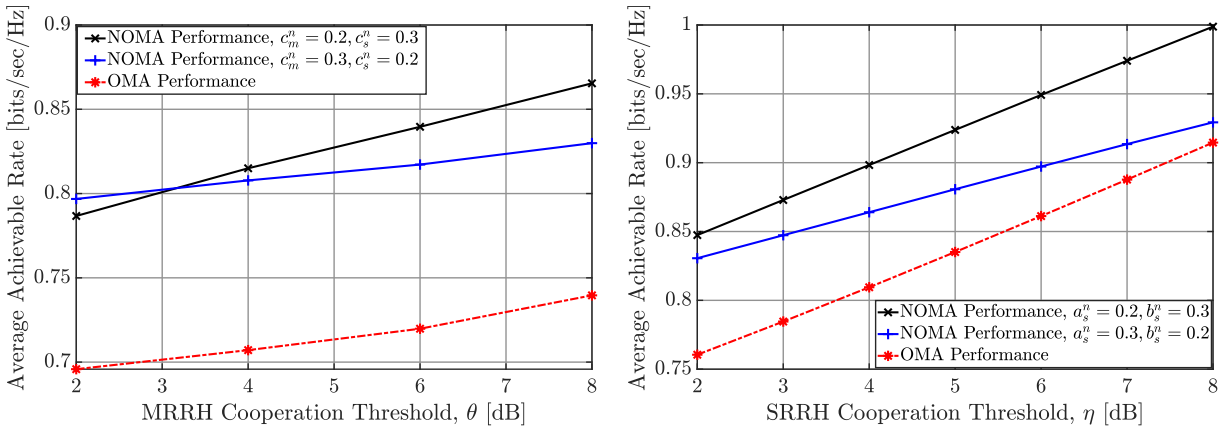


Figure 3.5: Illustrates the performance of the Non-CoMP-UE and the CoMP-UE with varying the PA coefficient for the near user.



(a) (b)

Figure 3.6: (a) and (b) show the effect of the different interference judging coefficient values from the MRRH and the SRRH or from the serving SRRH and the coordinated SRRH on the performance of the CoMP-UE, respectively.

degrades. However, to achieve better performance for the two-user NOMA compared to the OMA scheme, the PA factor should approximately take values between 0.15 to 0.30. In addition, in order to achieve fairness among the Non-CoMP-UE and CoMP-UE, we should assign the lower valid value for PA factor, i.e., 0.15, to guarantee that the Non-CoMP-UE performance does not become lower than its performance when the OMA scheme is adopted. At the same time, for lower PA factor value, the CoMP-UE has the best average achievable data rate. For instant, for $a_s^n = b_s^n = 0.15$, $\mathcal{R}_{ss} = 1.0589$ bits/sec/Hz and $\mathcal{R}_s = 1.2929$ bits/sec/Hz.

In Fig. 3.6(a) and Fig. 3.6(b), we show the effect of the fraction of the PA coefficients

assigned by the cooperating RRHs on the performance of the CoMP-UE in terms of the average achievable data rate. In Fig. 3.6a, it can be seen that the fraction of the power assigned by the SRRH has a better impact on the CoMP-UE (MSUE) than the MRRH for lower values of θ. This is because for lower values of θ the extend region around the SRRH is low. Besides, the high density of SRRHs leads to the SRRHs be closer to the UEs than the MRRH, and hence, provides better performance than the MRRH for lower values of θ. However, when the value of θ increases, the extended region around the SRRH increases. Consequently, the probability that the MRRH provides a stronger power than the SRRH in this region is high. This means that the effect of the PA factor from the MRRH is more dominant than that from the SRRH. For instance, when $\theta = 6$, the propose NOMA scheme can achieve about 16.6% than the OMA scheme when $a_m^n = 0.2$ and $b_s^n = 0.3$; meanwhile achieves about 13.5% when $a_m^n = 0.3$ and $b_s^n = 0.2$. Fig. 3.6(b) indicates the effect of the serving SRRH compared to the coordinated SRRHs in the performance of the SSUE. It can be observed that the coordinated RRH has a closer impact in comparison with the serving SRRH for lower values of the SRRH cooperation coefficient η on the performance of the SSUE. This is because lower values of η mean that the extended region around the serving SRRH is small and hence the coordinated SRRH is near to the SSUE. For instance, for $\eta = 2, a_s^n = 0.2$, and $b_s^n = 0.3$, the NOMA scheme achieves about 11.4% than the OMA scheme; meanwhile for $a_s^n = 0.3$ and $b_s^n = 0.2$, the performance is 9.2% better compared to OMA scheme. However, the effect of the coordinated SRRH on the performance of the SSUE decreases when the η increases. The reason behind this is that high values of η lead to an increase in the extended region around the serving SRRH, which leads to the probability of the SSUE being far from the coordinated SRRH being high. For instance, for $\eta = 6, a_s^n = 0.2$, and $b_s^n = 0.3$, the NOMA scheme achieves about 10.2% than the OMA scheme; meanwhile for $a_s^n = 0.3$ and $b_s^n = 0.2$, the performance is 4.2% better compared to the OMA.

3.4.4.4 Spectral Efficiency

Fig. 3.7.(a) plots the SE of the proposed CoMP NOMA and the traditional CoMP OMA techniques versus the ratio between the densities of the SRRHs and the MRRHs, i.e., λ_s/λ_m, for different PA coefficients allocated to the near user (Non-CoMP-UE), i.e. c_k^n. We can observe that, with appropriate power control, the system performance of the

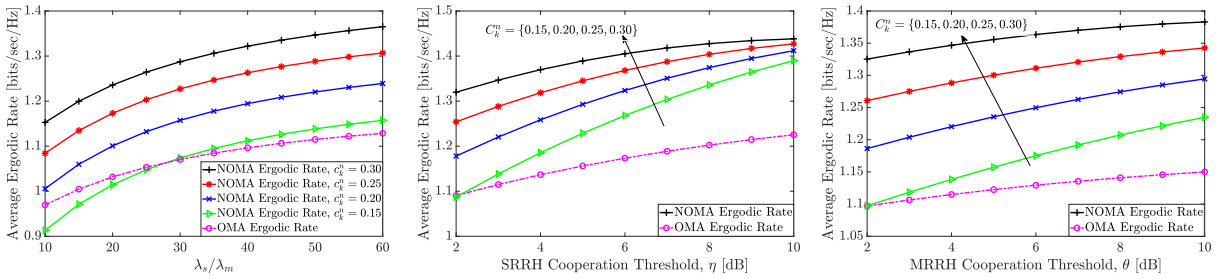


Figure 3.7: (a), (b), and (c) illustrates the *average ergodic rate* of the proposed CoMP NOMA technique versus the ratio of SRRHs density to the MRRHs density, SRRH cooperation threshold, and SRRH cooperation threshold for different power control coefficients.

proposed CoMP NOMA technique outperforms the counterpart CoMP OMA technique in terms of the SE. In addition, it can be also shown that the spectral efficiency improves as the SRRHs density increases. The reason behind this is that when the density of the SRRHs increases, two benefits are gained. First, the CCU becomes near to the serving SRRH, which leads to improving its achievable rate. Second, the probability of users triggering CoMP is high. This is because densifying the network also increases interference. Therefore, the CEUs can exploit the CoMP technique by transforming the received signal from the dominant interfering SRRH into a useful signal and, therefore, enhance their performance.

Fig. 3.7.(b) and Fig. 3.7.(c) depict the comparison between the proposed CoMP NOMA and conventional CoMP OMA techniques in terms of the SE versus the SRRH and the MRRH cooperation thresholds η and θ, respectively, for different PA coefficient assigned to the Non-CoMP-UE. It can be seen that for all the shown PA settings, the proposed CoMP NOMA outperforms the CoMP OMA, which validates the performance of the integration between CoMP and NOMA. As a result, to guarantee the superiority of the proposed CoMP NOMA technique over the CoMP OMA scheme, the PA coefficient for the Non-CoMP-UE should be greater than 0.15. Moreover, assigning more power to the Non-CoMP-UE results in enhancing the network SE. As depicted in Fig. 3.3 and Fig. 3.4, the performance of the SSUE and SUE is better than the performance of the MSUE and MUE. In addition, the rate of improvement for the SSUE and SUE performance when increasing the SRRH cooperation threshold is better than the rate of improvement for the MSUE and MUE when the MRRH cooperation threshold is increasing. This is because the density of SRRHs is much higher than the density of the MRRHs. This explains why the rate of improvement in the network SE is better when increasing the SRRH

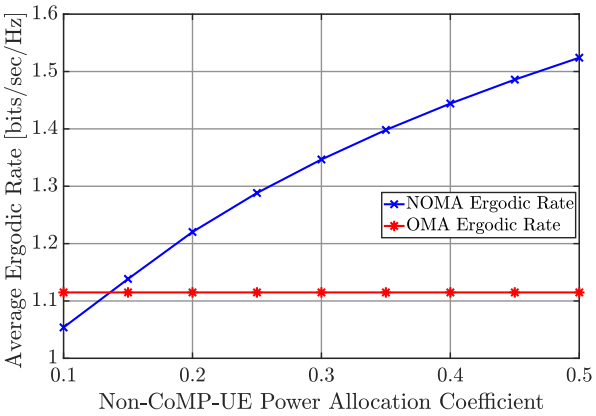


Figure 3.8: Shows the *average ergodic rate* versus PA coefficients allocated to the Non-CoMP-UE.

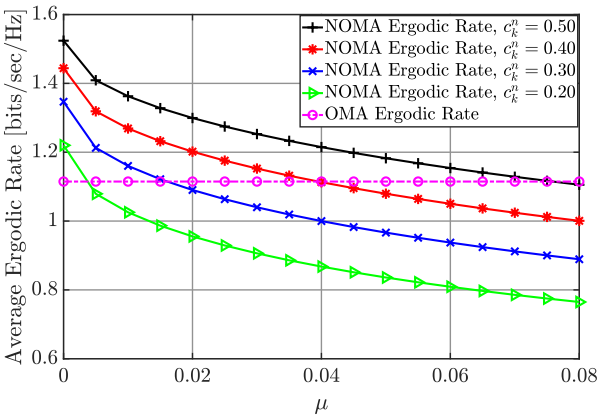


Figure 3.9: *Average ergodic rate* versus μ caused by the imperfect SIC decoding.

cooperation threshold than when increasing the MRRH cooperation threshold.

Fig. 3.8 depicts the SE of the proposed CoMP NOMA and the conventional CoMP OMA techniques versus the PA coefficient allocated to the near user (Non-CoMP-UE), i.e. c_k^n. It can be observed that with an appropriate PA setting, the performance of the proposed framework outperforms the CoMP OMA, especially at higher values of the PA coefficients, i.e. c_k^n. Furthermore, higher values of c_k^n mean more power is assigned to the Non-CoMP-UE, and hence, a higher gain for CoMP NOMA over CoMP OMA is achieved.

3.4.5 Effect of Imperfect SIC and Different Carrier Frequencies on the System Performance

Fig. 3.9 presents the *average ergodic rate* for the proposed system versus the portion of the NOMA interference caused by the imperfect SIC decoding μ. This figure shows that

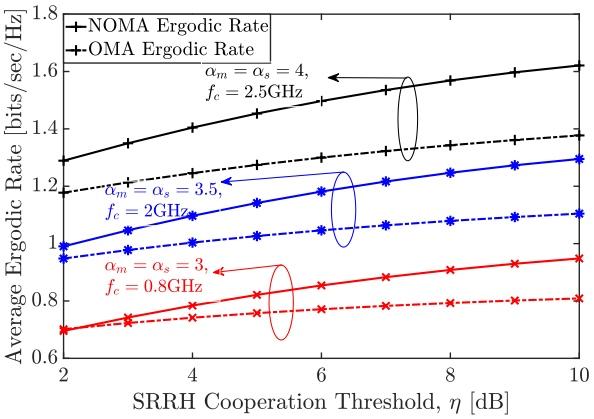


Figure 3.10: *Average ergodic rate* versus η for different operating carrier frequencies.

the imperfect SIC has a great impact on the system performance in terms of the *average ergodic rate*. When $\mu = 0.075$, i.e., 7.5% INUI fails to be canceled at the Non-CoMP-UE, the gain of CoMP NOMA completely vanishes in all the possible PA settings (Since $c_k^n \leq 0.5$). Moreover, higher values of c_k^n mean more power is allocated to the Non-CoMP-UE. Thus, with a greater c_k^n, the gain of CoMP NOMA over CoMP OMA is higher. So a greater c_k^n makes the Non-CoMP-UE more robust against the impact of imperfect SIC. Finally, one can observe that, with perfect SIC and for all shown PA settings for the Non-CoMP-UE, the performance of CoMP NOMA outperforms the one of CoMP OMA as discussed above.

Fig. 3.10 presents the *average ergodic rate* of the system versus the SRRH cooperation threshold η, for different path-loss exponents, each associated to a given carrier frequency. When the operating carrier frequency is high, the RRHs experience high path-loss as discussed in [92] and the *average ergodic rate* improves. Intuitively, a higher path-loss value reduces the effect of the ICI between the different RRHs either in the same tier or in different tiers. Consequently, the different RRHs are more isolated from each other. On the other hand, for lower carrier frequencies, the CoMP NOMA technique does not achieve much gain compared to CoMP OMA technique. This is because the INUI will have severe impacts on the far user.

After investigating the performance of CoMP NOMA networks, in the following, we investigate the potential of integrating cooperative communication with CoMP NOMA to enhance the performance of the CEU in terms of both outage probability and ergodic rate.

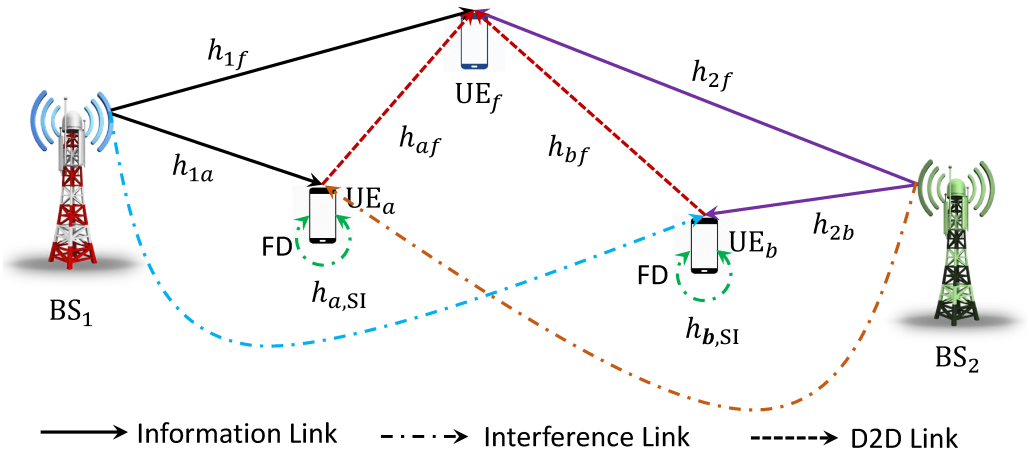


Figure 3.11: CoMP enabled C-NOMA cellular network.

3.5 Performance Analysis of CoMP C-NOMA Networks

3.5.1 System Model

In this section, we present first the network model, and then we describe the signal model and the corresponding SINR.

3.5.1.1 Network Model

We consider a two-cell DL CoMP-enabled C-NOMA system, which consists of two BSs, two Non-CoMP-UEs, and one CoMP-UE as shown in Fig. 3.11. We assume that each Non-CoMP-UE is associated with one BS, where we denote the Non-CoMP-UEs associated with BS_1 and BS_2 as UE_a and UE_b, respectively. This is because UE_a and UE_b have good SINR with their serving BSs. Meanwhile, the CoMP-UE, which we denote by UE_f, is considered as a CEU that experiences less distinctive channel gain from the two BSs. Hence, we consider this user as a CoMP-UE that is associated with both BSs. Each BS forms a NOMA pair that includes its relative Non-CoMP-UE with the CoMP-UE. Therefore, the CoMP-UE belongs to two NOMA pair, while each Non-CoMP UE is a part of only one NOMA pair. All UEs and BSs have one transmit antenna and one receive antenna.

Both UE_a and UE_b relay the signal for UE_f in FD DF relaying mode. There are six wireless communication links namely $BS_1 \rightarrow UE_a$, $BS_1 \rightarrow UE_f$, $UE_a \rightarrow UE_f$, $BS_2 \rightarrow UE_b$, $BS_2 \rightarrow UE_f$ and $UE_b \rightarrow UE_f$, whose channel coefficients are denoted, respectively,

as $h_{1a}, h_{1f}, h_{af}, h_{1b}, h_{2f}$ and h_{bf}, and two interference links namely $\text{BS}_1 \rightarrow \text{UE}_b$ and $\text{BS}_2 \rightarrow \text{UE}_a$, whose channel coefficients are h_{1b} and h_{2a}, respectively. In addition, there are two SI links namely $\text{UE}_a \rightarrow \text{UE}_a$ and $\text{UE}_b \rightarrow \text{UE}_b$, whose channel coefficients are denoted, respectively, as $h_{a,\text{SI}}$ and $h_{b,\text{SI}}$. All channels are independent and follow Rayleigh distribution. As a result, the channel gains for predefined channels follow Exponential distribution with parameters λ_{1a}, λ_{1f}, λ_{af}, λ_{2b}, λ_{2f}, and λ_{bf}, respectively, for the wireless links, λ_{1b}, λ_{2a}, for the interference links and λ_{SI} for the SI links. We assume that the average channel gains of the considered links are determined by the path-loss, that is, $\lambda_{ij} = d_{ij}^{-\alpha}, \forall i \in \{1,2\}, j \in \{a,b,f\}$, where d is the distance between BS_i and UE_j, and $\lambda_{kf} = d_{kf}^{-\alpha}, \forall k \in \{a,b\}$, for the D2D communication [33]. Finally, we denote by P_s and P_r the transmit power from the BSs and relay nodes, respectively, in which $P_r = P_s/2$ [72].

3.5.1.2 Signal Model and SINR Analysis

The communication between each BS and the cellular users consists of two different phases, namely, the DT phase and the CT phase. In the DT phase, each BS (BS_1 and BS_2) transmits the superimposed signal, which consists of the messages of both its corresponding Non-CoMP-UE and the CoMP-UE. Using SIC, the Non-CoMP-UE first decodes the CoMP-UE's message, subtracts it from its own reception, and then decodes its own message. In the CT phase, the Non-CoMP-UE forwards the message of the CoMP-UE using FD and D2D communication and then the CoMP-UE combines the signals from the cooperated BSs and the Non-CoMP-UEs. Note that, since each Non-CoMP-UE is operating under FD mode, the direct and the cooperative transmission phases occur simultaneously.

We start first by discussing the signal model for the Non-CoMP-UEs, i.e., UE_a and UE_b. Then, we define the signal model for the CoMP-UE, i.e., UE_f. In the kth time slot, $k = 1, 2, 3, \ldots$, BS_1 transmits the superposition signal $x_a[k]$ and $x_f[k]$ to UE_a and UE_f simultaneously. Also, BS_2 transmits the superimposed signal to its associated UEs. Thus, the received signal at UE_a can be written as

$$y_a[k] = h_{1a}[k](\sqrt{\beta_1^a P_s}x_a[k] + \sqrt{\beta_1^f P_s}x_f[k]) + h_{a,\text{SI}}[k]\sqrt{P_r}s_a[k] + \sqrt{P_s}h_{2a}x_2[k] + n_1[k], \tag{3.42}$$

where β_1^a and β_1^f are the PA coefficients allocated by BS_1 to the near and far UE, respectively, $s_a[k]$ is the transmit signal from the Non-CoMP-UE to the CoMP-UE that causes

SI at UE_a with zero mean and unit variance $\sim \mathcal{CN}(0,1)$, $x_2[k]$ is the superimposed signal transmitted by BS_2, and $n_1[k]$ is an AWGN that is $\sim \mathcal{CN}(0,\sigma^2)$ distributed. According to NOMA principle, since $|h_{1f}| \leq |h_{1a}|$, BS_1 assigns less power to Non-CoMP-UE than the CoMP-UE, i.e., $\beta_1^f > \beta_1^a$ and $\beta_1^a + \beta_1^f \leq 1$. The Non-CoMP-UE, i.e. UE_a, first decodes the CoMP-UE's message $x_f[k]$ and then cancels it from its received signal to decode its own message. Hence, the received `SINR` at the UE_a to decode the message of UE_f is given by

$$\gamma_{a,f} = \frac{\beta_1^f P_s |h_{1a}|^2}{\beta_1^a P_s |h_{1a}|^2 + |h_{a,\text{SI}}|^2 P_r + P_s |h_{2a}|^2 + \sigma^2}, \tag{3.43}$$

Similarly, the `SINR` at the Non-CoMP-UE UE_b that is associated with BS_2 to detect the UE_f's message can be given as

$$\gamma_{b,f} = \frac{\beta_2^f P_s |h_{2b}|^2}{\beta_2^b P_s |h_{2b}|^2 + P_r |h_{b,\text{SI}}|^2 + P_s |h_{1b}|^2 + \sigma^2}, \tag{3.44}$$

where β_2^b and β_2^f are the PA coefficients allocated by BS_2 to the near and far UE, respectively. Thus, the achievable data rate at UE_a and UE_b to decode the message of UE_f can be calculated as

$$\mathcal{R}_{a,f} = \log_2(1 + \gamma_{a,f}), \tag{3.45}$$

$$\mathcal{R}_{b,f} = \log_2(1 + \gamma_{b,f}). \tag{3.46}$$

Both UE_a and UE_b decode and forward the message to the CoMP-UE UE_f. As a result, the UE_f receives four versions of its own signal, two copies through CoMP and the other two copies from the D2D communication links by UE_a and UE_b. Therefore, the received signal at the CoMP-UE can be expressed as

$$\begin{aligned} y_2[k] = &(\sqrt{\beta_1^f P_s} h_{1f}[k] + \sqrt{\beta_2^f P_s} h_{2f}[k]) x_f[k] + (h_{af}[k]\sqrt{P_r} + h_{bf}[k]\sqrt{P_r}) x_f[k-t] \\ &+ \sqrt{\beta_1^a P_s} h_{1f}[k] x_a[k] + \sqrt{\beta_2^b P_s} h_{2f}[k] x_b[k] + n_2[k], \end{aligned} \tag{3.47}$$

where $x_b[k]$ is the message of UE_b, $n_2[k]$ is an AWGN that is $\sim \mathcal{CN}(0,\sigma^2)$ distributed and t is the delay resulting from the SIC processing at each Non-CoMP-UE. We assume that UE_a and UE_b have the same processing capability and perfect SIC [72]. The `SINR` of the CoMP-UE at its own end is expressed as[6]

$$\gamma_{f,f}(\omega_a, \omega_b) = \frac{P_s \left(\beta_1^f |h_{1f}|^2 + \beta_2^f |h_{2f}|^2\right) + P_r \left(\omega_a |h_{af}|^2 + \omega_b |h_{bf}|^2\right)}{P_s \left(\beta_1^a |h_{1f}|^2 + \beta_2^b |h_{2f}|^2\right) + \sigma^2}, \tag{3.48}$$

where, for $k \in \{a, b\}$, ω_k is a binary indicator such that $\omega_k = 1$ when the near user k can decode the message of the far user in the same NOMA group and 0 otherwise. One can see that the `SINR` is improved by invoking CoMP in conventional C-NOMA. The reason

[6] A detailed proof is given in Appendix B2

behind this is that adding another BS with the serving BS to help the ICI-prone user will not only suppress the ICI interference and enhance the received useful signal, but also will add one extra stream from the Non-CoMP-UE in the same NOMA pair. Note that, in the decode and forward relaying channel, the achievable data rate is limited by the weakest link [101]. As a result, the overall achievable data rate for the CoMP-UE UE_f is expressed as [101],

$$\mathcal{R}_f = \min\{\mathcal{R}_{a,f}, \mathcal{R}_{f,f}, \mathcal{R}_{b,f}\}, \tag{3.49}$$

where $\mathcal{R}_{f,f} = \log_2(1 + \gamma_{f,f}(\omega_a, \omega_b))$. The overall achievable data rate at CoMP-UE is the minimum between three terms. The first term is the rate at the UE_a to decode UE_f's message, the second term is the rate achieved by combining the `SINR` from the coordinated BSs and the relay nodes and the last term is the rate at UE_b to detect the UE_f's message.

3.5.2 Performance Analysis

In this section, a detailed analysis of the network performance for the proposed CoMP-assisted FD C-NOMA is presented. Both outage performance and ergodic rate for the CoMP-UE are investigated.

3.5.2.1 Outage Probability

In order to guarantee the required QoS at the CoMP-UE, its achievable data rate must be greater than a predefined threshold R_f^{th}. The outage event occurs when the achievable data rate of the CoMP-UE is smaller than its required threshold rate. The `SINR` threshold below which the CoMP-UE is in outage is $\gamma_f^{\mathrm{th}} = 2^{R_f^{\mathrm{th}}} - 1$. We first denote the outage event as $E_{j,f} = \{\gamma_{j,f} < \gamma_f^{\mathrm{th}}\}$, for $j \in \{a, b, f\}$. One can see that, when $j \neq f$, the outage occurs when UE_j cannot decode the message of the CoMP-UE and when $j = f$ the outage occurs when the CoMP-UE cannot decode its own message.

The outage probability of the CoMP-UE can be decomposed into three independent events. The first event is when both UE_a and UE_b can decode the message of UE_f but the received `SINR` after combining the four different streams from UE_a, UE_b, BS_1, and BS_2 is below the predefined threshold. The second event occurs when only one Non-CoMP-UE can decode the message x_f of UE_f and the combined received `SINR` at UE_f is below the target threshold. The last event is that neither UE_a nor UE_b can detect x_f as well as UE_f cannot decode its own message received from the CoMP transmission. Therefore,

the outage probability at UE_f can be expressed as:

$$\begin{aligned}
P_{\text{out},f} &= \mathbb{P}(E_{f,f}(\omega_a=1,\omega_b=1)\cap\bar{E}_{a,f}\cap\bar{E}_{b,f})+\mathbb{P}(E_{f,f}(\omega_a=1,\omega_b=0)\cap\bar{E}_{a,f}\cap E_{b,f})+\\
&\quad\mathbb{P}(E_{f,f}(\omega_a=0,\omega_b=1)\cap E_{a,f}\cap\bar{E}_{b,f})+\mathbb{P}(E_{f,f}(\omega_a=0,\omega_b=0)\cap E_{a,f}\cap E_{b,f}),\\
&=\mathbb{P}(E_{f,f}(1,1))(1-\mathbb{P}(E_{a,f}))(1-\mathbb{P}(E_{b,f}))+\mathbb{P}(E_{f,f}(1,0))(1-\mathbb{P}(E_{a,f}))\mathbb{P}(E_{b,f})+\\
&\quad\mathbb{P}(E_{f,f}(0,1))\mathbb{P}(E_{a,f})(1-\mathbb{P}(E_{b,f}))+\mathbb{P}(E_{f,f}(0,0))\mathbb{P}(E_{a,f})\mathbb{P}(E_{b,f}),
\end{aligned}\tag{3.50}$$

where $\bar{E}_{j,f}$ is complementary event of $E_{j,f}$ and $E_{f,f}(\omega_a=\theta_a,\omega_b=\theta_b)=\{\gamma_{f,f}(\omega_a=\theta_a,\omega_b=\theta_b)<\gamma_f^{\text{th}}\}$. In order to find the outage probability expression for the CoMP-UE, we need to calculate these probabilities $\mathbb{P}(E_{f,f}(\theta_a,\theta_b)),\mathbb{P}(E_{a,f})$, and $\mathbb{P}(E_{b,f})$. The probability of the event $E_{f,f}(\theta_a,\theta_b)$ can be calculated as follows:

$$\begin{aligned}
\mathbb{P}(E_{f,f}(\theta_a,\theta_b)) &= \mathbb{P}\left(\gamma_{f,f}(\omega_a=\theta_a,\omega_b=\theta_b)<\gamma_f^{\text{th}}\right),\\
&=\mathbb{P}\left(\frac{P_s\left(\beta_1^f|h_{1f}|^2+\beta_2^f|h_{2f}|^2\right)+P_r\left(\theta_a|h_{af}|^2+\theta_b|h_{bf}|^2\right)}{P_s\left(\beta_1^a|h_{1f}|^2+\beta_2^b|h_{2f}|^2\right)+\sigma^2}<\gamma_f^{\text{th}}\right),\\
&=\mathbb{P}\left(\Omega_1P_s|h_{1f}|^2+\Omega_2P_s|h_{2f}|^2+\theta_aP_r|h_{af}|^2+\theta_bP_r|h_{bf}|^2<\sigma^2\gamma_f^{\text{th}}\right),
\end{aligned}\tag{3.51}$$

where $\Omega_1=\beta_1^f-\gamma_f^{\text{th}}\beta_1^a$ and $\Omega_2=\beta_2^f-\gamma_f^{\text{th}}\beta_2^b$. Since the channel gain is modeled as an Exponential random variable, it holds that $\Omega_1\rho_s|h_{1f}|^2,\Omega_2\rho_s|h_{2f}|^2,\theta_a\rho_r|h_{af}|^2$, and $\theta_b\rho_r|h_{bf}|^2$ are Exponential random variables with parameters $\phi_1=\Omega_1\rho_s\lambda_{1f},\phi_2=\Omega_2\rho_s\lambda_{2f},\phi_3=\theta_a\rho_r\lambda_{af}$, and $\phi_4=\theta_b\rho_r\lambda_{bf}$, respectively, where $\rho_s\triangleq\frac{P_s}{\sigma^2}$ and $\rho_r\triangleq\frac{P_r}{\sigma^2}$ are the signal-to-noise-ratio (SNR) at the BS and the relay nodes, respectively. Thus, the probability of the event $E_{f,f}(\theta_a,\theta_b)$ can be obtained using the distribution of the sum of independent Exponential random variables [102]:

$$\mathbb{P}(E_{f,f}(\theta_a,\theta_b))=1-\sum_{i=1}^{2m}\frac{\exp\left[-\gamma_f^{\text{th}}/\phi_i\right]}{\prod_{j=1,j\neq i}^{2m}(1-\frac{\phi_j}{\phi_i})},\tag{3.52}$$

where m is the number of the coordinated BSs, and in this case $m=2$. On the other hand, the probability of the second event $E_{a,f}$ can be calculated as follows.

$$\begin{aligned}
\mathbb{P}(E_{a,f}) &= \mathbb{P}(\gamma_{a,f}<\gamma_f^{\text{th}}),\\
&=\mathbb{P}(|h_{1a}|^2<\tau_1(\rho_r|h_{a,\text{SI}}|^2+\rho_s|h_{2a}|^2+1)),\\
&=\int_0^\infty\int_0^\infty\int_0^{\tau_1(\rho_rv+\rho_sw+1)}f_{|h_{1a}|^2}(u)f_{|h_{a,\text{SI}}|^2}(v)f_{|h_{2a}|^2}(w)dudvdw,\\
&=1-\frac{(\lambda_{1a})^2}{(\lambda_{1a}+\tau_1\rho_r\lambda_{\text{SI}})(\lambda_{1a}+\tau_1\rho_s\lambda_{2a})}\exp^{-\frac{\tau_1}{\lambda_{1a}}},
\end{aligned}\tag{3.53}$$

where $\tau_1=\frac{\gamma_f^{\text{th}}}{\Omega_1\rho_s}$. Note that $\mathbb{P}(E_{a,f})=1$ when $\beta_1^f-\gamma_f^{\text{th}}\beta_1^a<0$. Following the same steps in (3.53), the probability of the event $E_{b,f}$ can be evaluated. By substituting (3.52), (3.53)

in (3.50), the outage probability at the CoMP-UE can be obtained.

Remark 1: The proposed scheme can be relaxed to C-NOMA [33] by considering the parameter of the channel gain between the coordinated BS, i.e., BS_2 and UE_f tends to infinity. In addition, it may be also relaxed to the CoMP NOMA [68] by considering $\omega_a = \omega_b = 0$ with zero SI.

Remark 2: For further enhancement in terms of the outage probability performance, the proposed scheme can be generalized for $m > 2$ cooperated BSs for serving the CEUs. In this case, the outage probability can be expressed as follows:

$$P_{\text{out}}^{g,f} = \sum_{i \in \mathcal{M}} \mathbb{P}(E_{f,f}(\mathbf{W}_i)) \left[\prod_{k \in \mathcal{K}_i} \mathbb{P}(E_{k,f}) \prod_{u \in \mathcal{U}_i} (1 - \mathbb{P}(E_{u,f})) \right], \tag{3.54}$$

where $\mathcal{M} = 2^m$ is the total number of combinations of D2D links and for all $i = 1, 2, ..., m$, $\mathbf{W}_i$ is an $m \times 1$ binary vector such that, for all $j = 1, 2, ...m$, the element $W_{i,j} = 1$ if the near user of the jth D2D link has correctly decoded the message of the far user, and $W_{i,j} = 0$ otherwise, and $\mathcal{K}_i$ and $\mathcal{U}_i$ are the sets that contain the indices of the ones and zeros in $\mathbf{W}_i$, respectively.

3.5.2.2 CoMP-UE Ergodic Rate

In this section, we analyze the performance of the proposed CoMP-enabled C-NOMA in terms of the ergodic rate of the far UE. The ergodic rate is a key metric to evaluate the system performance when the rate of the CoMP-UE is determined by its channel condition [72]. Assuming that the message x_f from the CoMP transmission and relaying links can be decoded at the CoMP-UE as well as at the UE_a and UE_b using SIC. Thus, the achievable data rate for CoMP-UE in the proposed model can be given:

$$\begin{aligned}
\bar{\mathcal{R}}_f &= \mathbb{E}\left[\log(1 + \min(\gamma_{a,f}, \gamma_{f,f}, \gamma_{b,f})\right], \\
&\overset{(a)}{=} \frac{1}{\ln 2} \int_0^{\vartheta} \frac{\bar{F}_{\gamma_{\min}}(z)}{1+z} dz, \\
&= \frac{1}{\ln 2} \int_0^{\vartheta} \frac{1}{1+z} \left[\mathbb{P}(\bar{E}_{a,f}) \mathbb{P}(\bar{E}_{f,f}(1,1)) \mathbb{P}(\bar{E}_{b,f})\right] \Big|_{\gamma_f^{\text{th}} = z} dz,
\end{aligned} \tag{3.55}$$

where $\gamma_{\min} = \min(\gamma_{a,f}, \gamma_{f,f}, \gamma_{b,f})$ and $\bar{F}_{\gamma_{\min}}(z)$ is its complementary cumulative distribution function, and $\vartheta = \min\left(\frac{\beta_1^f}{\beta_1^a}, \frac{\beta_2^f}{\beta_2^b}\right)$. The integration limits in (a) come from the NOMA principle. A simplified expression can be obtained at a high SNR approximation. In fact, at high SNR, $\mathbb{P}(\bar{E}_{a,f}), \mathbb{P}(\bar{E}_{f,f}(1,1))$, and $\mathbb{P}(\bar{E}_{b,f})$, can be rewritten as follows.

$$\mathbb{P}(\bar{E}_{a,f}) = (\lambda_{1a}\Gamma_1(z))^2 / ((\lambda_{1a}\Gamma_1(z) + \lambda_{\text{SI}} z/2)(\lambda_{1a}\Gamma_1(z) + \lambda_{2a} z)),$$

Table 3.2: Simulation Parameters [33, 74]

Parameter	Symbol	Value
Distance between BS_1 and UE_a and between BS_2 and UE_b	d_{1a}, d_{2b}	0.3
Distance between BS_1 and UE_f	d_{1f}	1
Distance between between BS_2 and UE_f	d_{2f}	1.4
Distance between UE_a and UE_f	d_{af}	d_{1f} - d_{1a}
Distance between UE_b and UE_f	d_{bf}	d_{2f} - d_{2b}
Distance between BS_1 and UE_b and between BS_2 and UE_a	d_{1b}, d_{2a}	2
SI channel gain	λ_{SI}	-10 dB
Pathloss exponent	α	4
PA coefficients for UE_a and UE_b	β_1^a, β_2^b	0.2
PA coefficients for UE_f	β_1^f, β_2^f	0.8
Rate threshold	γ_{th}^f	0.5 bits/sec/Hz

$$\mathbb{P}(\bar{E}_{f,f}(1,1)) = \sum_{i=1}^{2m} \frac{\phi_i}{\prod_{\substack{j=1,\\ j\neq i}}(\phi_i - \phi_j)},$$

$$\mathbb{P}(\bar{E}_{b,f}) = (\lambda_{2b}\Gamma_2(z))^2/((\lambda_{2b}\Gamma_2(z) + \lambda_{\text{SI}}z/2)(\lambda_{2b}\Gamma_2(z) + \lambda_{1b}z)), \tag{3.56}$$

where $\Gamma_1(z) = \beta_1^f - \beta_1^a z$ and $\Gamma_2(z) = \beta_2^f - \beta_2^b z$. By substituting in (3.55), the ergodic achievable rate for the CoMP-UE can be calculated after evaluating the single integration numerically.

3.5.3 Results And Discussion

In this section, the simulation results are presented to validate the performance of the proposed CoMP-enabled C-NOMA scheme. We assume that the power control schemes employed by the two BSs are the same, i.e., $(\beta_1^a, \beta_1^f) = (\beta_2^b, \beta_2^f)$. The simulation results are obtained through 10^9 independent Monte-Carlo trials and the main simulation parameters are shown in Table 3.2. We compare the proposed scheme with four other schemes, CoMP with HD C-NOMA, CoMP NOMA, CoMP OMA, and C-NOMA schemes. These schemes can be explained as follows.

1. CoMP with HD C-NOMA: this is similar to the proposed scheme but with HD communication instead of FD, which can be derived from the proposed scheme by adopting the SI parameter to zero and the `SINR` thresholds for the CoMP-UE is $\gamma_f^{\text{HD,th}} = 2^{2\times R_f^{\text{th}}} - 1$.

2. CoMP NOMA scheme [68]: this is similar to the proposed scheme without any D2D communication, which can be obtained by calculating only this term $E_{f,f}(0,0)$.

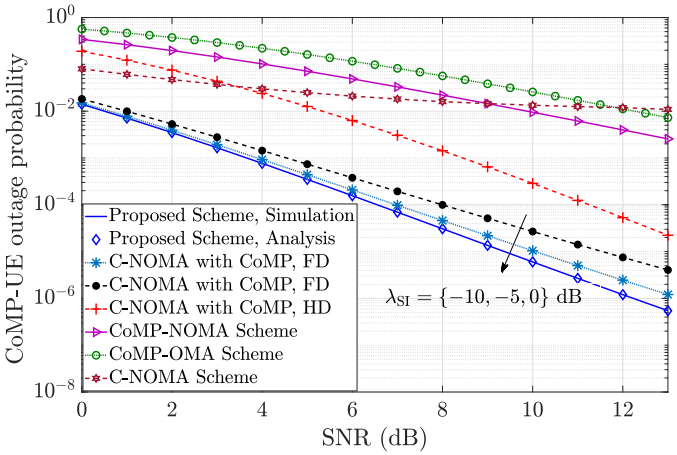


Figure 3.12: Outage probability performance versus the transmit SNR for the proposed scheme and the other four comparable schemes.

3. CoMP OMA scheme [68]: this scheme allows the Non-CoMP-UE to allocate orthogonal resources from the CoMP-UE by applying the OMA technique and the SINR threshold is similar to the first scheme.

4. C-NOMA scheme [33]: this is the traditional scheme that is applied in the literature considering only a single cell.

3.5.3.1 Outage Probability Performance versus Transmit SNR

Fig. 3.12 describes the outage probability of the CoMP-UE versus the SNR. It is important to note that the analytical derivations are in excellent agreement with the simulation results. In addition, it can be observed that the outage performance of the proposed scenario exceeds the other schemes. This is because our proposed scheme allows the CEU to exploit both the coordination between the BSs to mitigate the ICI by triggering CoMP and the coordination between the UEs belonging to the same NOMA group by exploiting the successive detection, D2D, and FD communication. It can also be seen that the C-NOMA scheme performs worse when the SNR increases. This is because the CEU experiences two types of interference, i.e., the INUI and the ICI interference, which seriously affects its performance.

3.5.3.2 Outage Probability Performance versus Distance from BS_2 to Far NOMA UE

Fig. 3.13 compares the outage performance of C-NOMA, CoMP NOMA, and the proposed scheme when the distance from the CEU to the coordinated BS increases. The lower

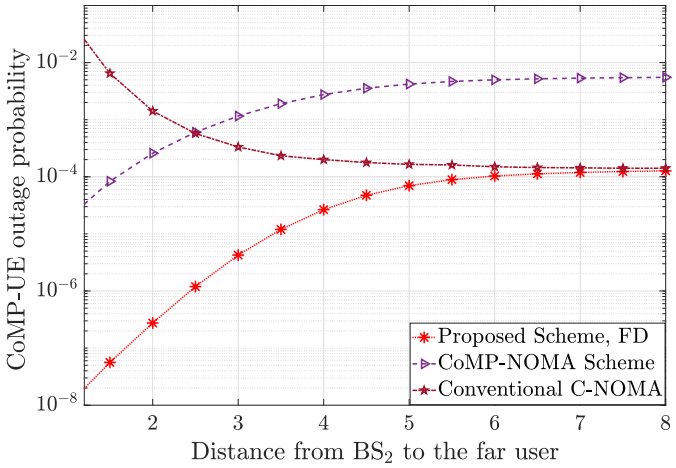


Figure 3.13: Outage probability performance versus the distance from the BS_2 to the far UE.

distance from BS_2 to UE_f means a high density of BSs. Thus, for lower values of d_{2f}, the ICI is high and the CoMP NOMA achieves better performance than the C-NOMA. Since the proposed scheme exploits the benefits of both C-NOMA and CoMP transmission, it achieves the best performance among the comparable schemes. However, for high values of d_{2f}, the ICI effect is low and the CoMP technique will not provide any gain.

3.5.3.3 Ergodic Rate Performance versus Transmit SNR

Fig. 3.14 presents the comparison between the proposed scheme and the CoMP NOMA scheme versus the transmit SNR in terms of the ergodic rate of the CEU. Markers and lines represent the analytical and simulation results, respectively, which show a perfect match between them. It can be seen that the proposed scheme has a better performance than the CoMP NOMA scheme at low SNR values. This demonstrates the superiority of the proposed scheme compared to the CoMP NOMA scheme and it is due to the contribution of the near users and the coordinated BSs to enhance the SINR at the far user. However, as we see from the rate equation of the far user, the achievable ergodic rate is constrained by the weakest link amongst the three available signals, i.e., the combined SINR and the rates at the near users to decode the far user message. Therefore, the achievable ergodic rate at the far user is saturated at higher SNR values. The reason behind this is that when the transmit SNR increases the rate at the near user to decode the far user is almost constant. In the meantime, the combined signal at the far user is further enhanced. Once the combined SINR exceeds the SINR at the near user to decode the message of the far user, the rate will be constant. Furthermore, when the near user gets closer to the BS,

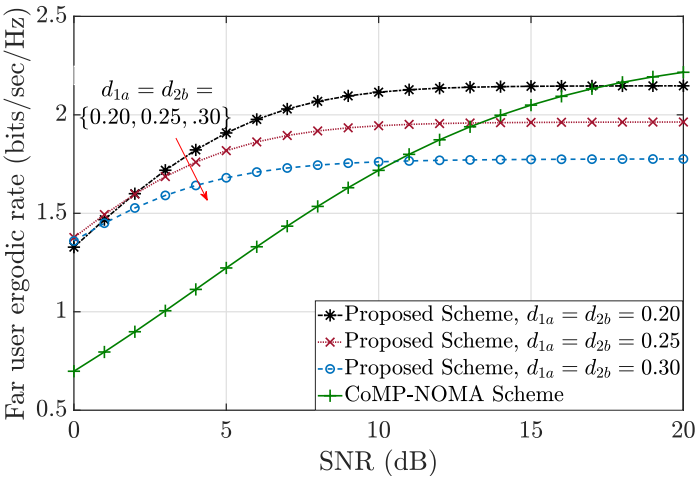


Figure 3.14: Ergodic rate performance versus SNR.

the achieved ergodic rate at the far user increases. This is because the rate at the near user to decode the far user message is enhanced.

3.6 Summary

With the goal of enhancing the DL SE, we investigate the integration between CoMP transmission and DL NOMA-based cellular networks. In addition, in the quest of improving network connectivity, the CoMP technique is primarily triggered to help UEs that is vulnerable to severe ICI with the assumption of only the two best RRHs are selected to serve such users. Our analytical results are validated using Monte Carlo simulation. Then, we compare our results with the conventional CoMP OMA transmission. We show that the proposed CoMP NOMA outperforms the CoMP OMA in terms of the average achievable data rate and the network SE for most of the PA settings where the combination of CoMP NOMA techniques has a great potential to be considered as the key enabling access technique in the physical layer of the DL H-CRAN. In this work, we numerically show the effect of the PA on the system performance in terms of the average achievable rate for each user and the NSE. However, it is worth analytically optimizing the PA coefficients so as to maximize the NSE considering the OMA achievable rate as the rate constraint for the NOMA users. This can be considered as a future research direction.

In addition, we also investigated the application of CoMP in C-NOMA system to further improve the CEU's performance. Two BSs with their associated strong users

are cooperating to serve the weak NOMA user belonging to the same NOMA pair. A closed-form expression for the outage probability of the far user is derived for the two BSs model. Then, we derive the outage probability for m coordinated BSs. Finally, an analytical expression for the ergodic rate of the far user is derived. We showed that the proposed CoMP C-NOMA scheme achieves better outage performance than the existing schemes. In addition, the ergodic rate of the proposed scheme is also superior compared to the CoMP NOMA scheme, especially at lower SNR.

Chapter 4
User Clustering and Power Allocation in CoMP C-NOMA Networks

4.1 Introduction

With the goal of enhancing the performance of multi-cell networks and as mentioned in the last chapter, more attention was recently directed towards the integration between C-NOMA and CoMP in which the authors in [10, 103] demonstrated the effectiveness of CoMP C-NOMA on the performance of the CEU in terms of both the ergodic rate and the outage probability. Note that, these works studied only the performance analysis of the system and a fixed non-optimal PA scheme was considered. However, we need also to study the optimization of the PA coefficients at the BSs as well as the relay node due to the following reasons.

1. Due to the FD operation, the higher values of the SI will discourage the CCUs to transmit with high power for avoiding harming themselves. Consequently, there is a high chance that the required QoS at the CEUs cannot be met. Thus, in order to overcome this issue, we should play with the PA coefficients assigned to the CEUss at the BSs to achieve their required QoS.

2. The SIC constraints and the QoS constraints at the CCUs are functions of the PA coefficients at the BSs and the power transmit from the CCUs. Therefore, in order to guarantee a successful SIC and to achieve the required QoS at the CCUs, the PA coefficients at the BSs should be jointly optimized with the power transmit at the CCUs.

3. When the required rate threshold at the CEUs increases if the PA coefficients at the BSs are not optimized, then the CCUs should increase their transmit power regardless the damages brought by to the SI. This may deteriorate the performance

of the network in terms of sum-rate and UEs admission.

4. Finally, since the CCUs operate in DF relaying mode, the performance of the CEUs is limited by the weakest link, which depends on the PA coefficients at the BSs.

4.2 State of The Art

The performance analysis of C-NOMA networks was studied in [33, 72] using tools from stochastic geometry. On the other hand, in order to harness the full potential of C-NOMA, it is necessary to study its PA problem, where a variety of such studies were developed in [32, 104, 105]. In [32], the objective is to maximize the minimum achievable rate, whereas in [104], the objective is to minimize the system outage probability. For the case when there is no direct link between the BS and the far NOMA UE, the energy efficiency-delay tradeoff was studied in [105]. In [106], a one-to-one two-sided matching game is proposed to pair each far NOMA UE with one near NOMA UE to maximize the system ergodic rate given the randomness of the UEs' channel gains. In [107] a joint UC and PA scheme is optimized with the objective of maximizing the network sum-rate. Note that, to the best of our knowledge, the existing research contributions in [32, 33, 72, 104–107] (and references therein) studied the PA problem in the context of single-cell C-NOMA networks. However, the more challenging multi-cell scenario was not addressed in the literature. With the goal of enhancing the performance of multi-cell networks, more attention was recently directed towards the integration between NOMA and CoMP [9, 66, 67, 108, 109]. In spite of tremendous works in CoMP NOMA, the works [9, 66, 67, 108, 109] (and references therein) only combine NOMA with CoMP without exploiting the cooperative communication between the cell-center users and the cell-edge users. Based on the above discussion, the main limitations and research gaps can be summarized as follows.

1. To the best of our knowledge, the joint UC and PA problem of CoMP-assisted FD/HD C-NOMA enabled cellular networks has not been studied in the literature and the performance of combining CoMP transmission in multi-user multi-cell C-NOMA systems remains unexplored.

2. The effect of the imperfection of both the CSI and SIC on the performance of CoMP-assisted C-NOMA cellular networks has not been investigated in the literature.

In order to proceed with tackling the above research gaps, it is worth mentioning that the main contributions of this chapter are composed of two key parts that can be read as follows.

1. We propose a low-complexity sub-optimal PA scheme for one coordinated C-NOMA cluster consisting of two CCUs and one CEU considering imperfect CSI as well as imperfect SIC.

2. We extend the previous study to consider multiple coordinated C-NOMA clusters (multi-user set-up) while providing low-complexity and high-quality solutions for both the PA and the UC policies.

4.3 Contributions

First, towards addressing the first limitation, we consider one coordinated C-NOMA cluster, according to that and against the above background and the aforementioned observations, the main contributions of this part can be summarized as follows.

- The power optimization framework for a coordinated C-NOMA cluster is studied, where the effects of both the imperfect SIC and the imperfect channel estimation are considered. This framework is formulated as an optimization problem that maximizes the cluster sum-rate while satisfying the minimum required rate for each UE.

- The formulated problem ends up being neither concave nor quasi-concave, which is difficult to be solved directly. To overcome this challenge, a near-optimal power control scheme is derived in a closed-form expression.

- The simulation results show that our purposed scheme achieves an average sum-rate that is 5% less than one of the optimal power control but it can save up to 99% in computational time. In addition, the superiority of the proposed scheme is demonstrated when compared to the C-NOMA scheme, which is a widely used access technique.

Second, towards resolving the second research limitation, we extend the above-proposed model to a multi-user scenario and we provide not only an efficient UC solution but also

a near-optimal PA solution that reaches 99.99% from the optimal solution. The main contributions of this part can be summarized as follows.

- We investigate the joint UC and the PA framework in the DL of a two-cell CoMP-assisted FD C-NOMA network. This framework is formulated as an optimization problem that maximizes the network sum-rate while guaranteeing the required data rate QoS for the UEs, the SIC constraints, and the power budget at the CCUs.

- The formulated problem ends up being a mixed-integer non-convex program that includes not only binary indicators but also non-smooth utility functions. Therefore, derivative-based convexification approaches cannot be applied directly. With the aid of the bi-level optimization, we decompose the original optimization problem into two sub-problems: a PA sub-problem for a single cluster (inner problem) and a UC sub-problem (outer problem).

- For the inner problem, we derive the feasibility conditions as relations between the QoS requirements, SIC constraints, and the power budget of the nodes. Then, for each member of the coordinated C-NOMA cluster, a computational-efficient solution is provided for the PA. Different from [11], it is important to mention here that no approximations are considered in this paper when conducting the derivations of the proposed PA solution.

- After obtaining the power control solution for each possible UEs cluster, the outer problem ends up being a linear assignment problem. Therefore, we develop a one-to-one three-sided matching scheme that can match one CCU from each cell with one CEU to construct a near-optimal coordinated C-NOMA cluster. Finally, the overall computational complexity is analyzed.

We perform extensive performance evaluations of the proposed CoMP-assisted FD C-NOMA network through various simulations. To demonstrate its effectiveness, we compare it with both the CoMP-assisted HD C-NOMA scheme and the CoMP NOMA scheme.[1] Numerical results validate the efficacy of the proposed framework in comparison

[1]To the best of our knowledge, investigating the potential gains of the integration between the CoMP transmission and HD C-NOMA has not been studied in the literature. It has been considered as a benchmark scheme to provide some insights when it is compared with the proposed CoMP-assisted FD C-NOMA. The performed analysis for the CoMP-assisted HD C-NOMA is described in detail in Appendix B1.

with the two benchmark schemes in terms of the network sum-rate and the connectivity for low and moderate values of SI. In addition, it is demonstrated that with lower total transmit power by the BSs, our proposed CoMP-assisted FD C-NOMA scheme provides a significant gain compared to the CoMP NOMA scheme.

The rest of this chapter can be organized as follows. To address the two main research gaps introduced above, we divide this chapter into two main parts, which are mapped into two principle Sections 4.4 and 4.5, respectively. The first part, which addresses the second research limitation, can be organized as follows. Section 4.4.1 presents the proposed system model, `SINR`, and rate analysis. In Section 4.4.2, we introduce the problem formulation with the proposed solution approach. Then, we present the main simulation results in Section 4.4.3. Meanwhile, the second part, which addresses the first research limitation, can be organized as follows. Section 4.5.1 presents the proposed system model, `SINR`, and rate analysis. Section 4.5.2 presents the formulated optimization problem and the solution approach. Section 4.5.3 presents the proposed power control scheme. Section 4.5.4 presents the proposed one-to-one three-sided matching game for solving the formulated UEs clustering problem. The simulation results are presented in Sections 4.5.5. Finally, we summarize the Chapter and provide the main insights in Section 4.6.

4.4 Power Allocation in CoMP C-NOMA Networks

4.4.1 System Model

In this section, we first describe the network model, and then, we discuss the signal model, the corresponding `SINR` for the near and far users, and the data rate expressions.

4.4.1.1 Network Model

As shown in Fig. 4.1, we consider a CoMP C-NOMA system that consists of two BSs, denoted by BS_1 and BS_2, two CCUs, denoted by UE_a and UE_b, and one CEU denoted by UE_f. The CCUs are the UEs that receive high `SINR` from the nearest BS. Meanwhile, the CEU is defined as the UE that has less distinctive channel gains from BS_1 and BS_2. Consequently, this UE is simultaneously served by the two BSs using CoMP. We define a coordinated C-NOMA cluster with two CCUs associated with two different BSs and one

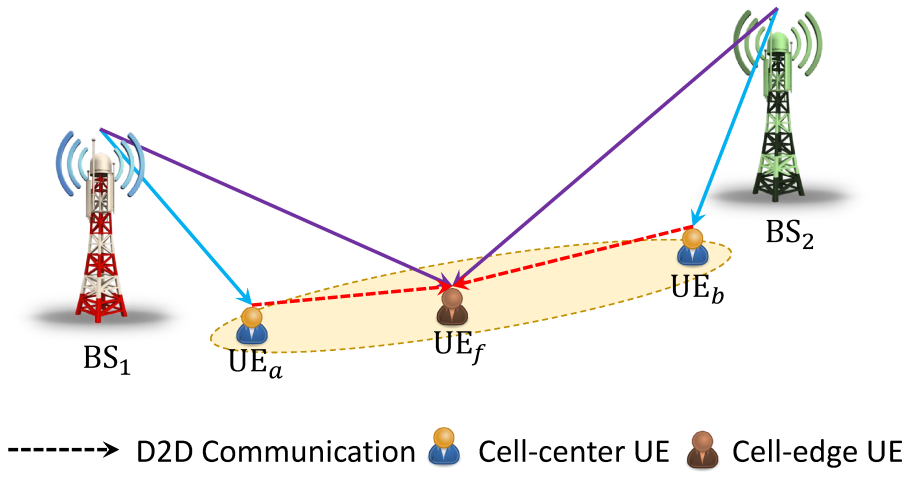


Figure 4.1: CoMP enabled three-UE two-BS C-NOMA architecture.

CEU served simultaneously by the two BSs. We assume that the UEs and the BSs have each one transmit antenna and one received antenna. Each BS serves its own CCU along with the CEU, i.e., UE_f, using power domain NOMA. Finally, each CCU, i.e. UE_a and UE_b, relays the message for the CEU in a DF FD relaying mode.

For notation convenience, we define $\mathcal{I} = \{1, 2\}, \mathcal{J} = \{a, b, f\}$ and $\mathcal{K} = \{a, b\}$ as the set of BSs, the set of all UEs, and the set of CCUs in the coordinated C-NOMA cluster, respectively. Moreover, for all $i \in \mathcal{I}$, $j \in \mathcal{J}$ and $k \in \mathcal{K}$, $h_{i,j} \sim \mathcal{CN}(0, \Omega_{ij})$, $h_{k,f} \sim \mathcal{CN}(0, \Omega_{kf})$, and $h_{k,\mathrm{SI}} \sim \mathcal{CN}(0, \Omega_{\mathrm{SI}})$ account for the Rayleigh fading channel coefficients from BS_i to UE_j, from UE_k to UE_f, and from UE_k to itself (SI due to the FD relaying mode), respectively. In practice, however, not only is it challenging but also there is no guarantee to obtain the perfect channel state information (CSI). Accordingly, we assume imperfect CSI in our model. By using the minimum mean square error (MMSE) channel estimation technique [110], we can model the Rayleigh fading coefficient between node v and node w as $h_{v,w} = \hat{h}_{v,w} + e_{v,w}$, where $\hat{h}_{v,w} \sim \mathcal{CN}(0, (1 - \eta_e)\Omega_{vw})$ denotes the estimated channel gain and $e_{v,w} \sim \mathcal{CN}(0, \eta_e \Omega_{vw})$ denotes the residual estimation error, which is uncorrelated with $\hat{h}_{v,w}$ [110]. In addition, we denote the transmit power from the BS_i and from UE_k by P_i and P_k, respectively.

4.4.1.2 Signal Model And SINR Analysis

We start by presenting the signal model for the CCUs, i.e. UE_a and UE_b, and then we present the signal model at the CEU, i.e. UE_f. During the lth time slot, BS_1 broadcasts

a superimposed mixture of the messages intended to UE_a and UE_f as

$$\mathcal{S}_1[l] = \sqrt{(1-\alpha_1)\,P_1}x_a[l] + \sqrt{\alpha_1 P_1}x_f[l], \tag{4.1}$$

where x_a and x_f are the signals intended to UE_a and UE_f, respectively, and $\alpha_1 \in [0,1]$ is the PA coefficient assigned by BS_1 to UE_f. The received signal at UE_a is expressed as

$$y_a[l] = \left(\hat{h}_{1,a} + e_{1,a}\right)\mathcal{S}_1[l] + \sqrt{P_a}h_{a,\mathrm{SI}}Z_f[l] + \left(\hat{h}_{2,a} + e_{2,a}\right)\mathcal{S}_2[l] + w_{1a}, \tag{4.2}$$

where $Z_f[l]$ is the message of UE_f decoded at UE_a, which causes SI at UE_a [32], $\mathcal{S}_2$ is the superimposed signal transmitted by BS_2, and w_{1a} is an AWGN at UE_a, which is $\mathcal{CN}(0,\sigma^2)$ distributed. At this point, UE_a decodes first the data of UE_f and then cancels it from its received signal to decode its own message. Therefore, the received `SINR` at UE_a, to decode the message of UE_f is expressed as

$$\delta_{a\to f} = \frac{\alpha_1 P_1 \gamma_{1,a}}{(1-\alpha_1)\,P_1\gamma_{1,a} + \gamma_{a,\mathrm{SI}}P_a + 1}, \tag{4.3}$$

where $\gamma_{1,a} = |\hat{h}_{1,a}|^2/\Theta_a$, $\gamma_{a,\mathrm{SI}} = |h_{a,\mathrm{SI}}|^2/\Theta_a$ and $\Theta_a \triangleq P_2|\hat{h}_{2,a}|^2 + P_1\eta_e\Omega_{1,a} + P_2\eta_e\Omega_{2,a} + \sigma^2$. However, the SIC decoding at the CCU may be imperfect due to some practical limitations. Therefore, this imperfect SIC process results in an error that propagates to the next level of data detection where the CCU decodes its own message. Based on this, the `SINR` for UE_a is expressed as

$$\delta_{a\to a} = \frac{(1-\alpha_1)\,P_1\gamma_{1,a}}{\zeta\alpha_1 P_1\gamma_{1,a} + P_a\gamma_{a,\mathrm{SI}} + 1}, \tag{4.4}$$

where $\zeta \in [0,1]$ denotes the fraction of the residual INUI from the SIC process, such that $\zeta = 0$ means a perfect SIC process, whereas $\zeta = 1$ means that UE_a completely fails in decoding the message of the CEU.

Both UE_a and UE_b will decode and forward UE_f's message to it. Consequently, the received signal at the CEU can be expressed as

$$R_{a\to f} = \log\left(1+\delta_{a\to f}\right) \quad \text{and} \quad R_a = \log\left(1+\delta_{a\to a}\right). \tag{4.5}$$

Similarly, the data rate at UE_b to decode UE_f's message and to decode its own message can be obtained. Both UE_a and UE_b will decode and forward the UE_f's message to the CEU. Consequently, the UE_f receives four streams for its own signal, two copies from UE_a and UE_b through the D2D communication and two copies from the two BSs through CoMP. The received signal at the CEU can be expressed as

$$\begin{aligned} y_f[l] = &\left(\sqrt{\alpha_1 P_1}\left(\hat{h}_{1,f} + e_{1,f}\right) + \sqrt{\alpha_2 P_2}\left(\hat{h}_{2,f} + e_{2,f}\right)\right)x_f[l] \\ &+ \left(\left(\hat{h}_{a,f} + e_{a,f}\right)\sqrt{P_a} + \left(\hat{h}_{b,f} + e_{b,f}\right)\sqrt{P_b}\right)x_f[l-\tau] \\ &+ \sqrt{(1-\alpha_1)\,P_1}\left(\hat{h}_{1,f} + e_{1,f}\right)x_a[l] + \sqrt{(1-\alpha_2)\,P_2}\left(\hat{h}_{2,f} + e_{2,f}\right)x_b[l] + w_f[l], \end{aligned} \tag{4.6}$$

where τ denotes the processing delay at the CCUs, where we assume that UE_a and UE_b

have the same processing capability [10]. Similar to [10, 111], the SINR of the CEU at its own end is expressed as

$$\delta_{f\to f} = \frac{\alpha_1 P_1 \gamma_{1,f} + \alpha_2 P_2 \gamma_{2,f} + P_a \gamma_{a,f} + P_b \gamma_{b,f}}{(1-\alpha_1) P_1 \gamma_{1,f} + (1-\alpha_2) P_2 \gamma_{2,f} + P_a \mu_1 + P_b \mu_2 + 1}, \tag{4.7}$$

where $\forall i \in \mathcal{I},\ f \in \mathcal{F},\ k \in \mathcal{K},\ \gamma_{i,f} = |\hat{h}_{i,f}|^2/\Theta_f,\ \gamma_{k,f} = |\hat{h}_{k,f}|^2/\Theta_f,\ \mu_1 = \eta_e \Omega_{a,f}/\Theta_f,$ $\mu_2 = \eta_e \Omega_{b,f}/\Theta_f$ and $\Theta_f \triangleq P_1 \eta_e \Omega_{1,f} + P_2 \eta_e \Omega_{2,f} + \sigma^2$. Then, the overall achievable rate for the CEU, i.e., UE_f is formulated as [101],

$$R_f = \min\{R_{a\to f}, R_{f,\text{CoMP}}, R_{b\to f}\}, \tag{4.8}$$

where $R_{f,\text{CoMP}} = \log(1+\delta_{f\to f})$.

4.4.2 Proposed Power Control Scheme for CoMP C-NOMA Network

4.4.2.1 Problem Formulation

The main objective of this paper is to derive a power control scheme with the goal of maximizing the sum-rate of the CoMP C-NOMA cellular system while achieving the SIC constraints (4.9b), considering the power budget constraints at each CCU (4.9c), and guaranteeing a certain QoS for each UE (4.9d). This objective can be obtained by solving the following optimization problem.

$$\mathcal{P}: \quad R^*_{\text{sum}} = \max_{\alpha_1,\alpha_2,P_a,P_b} R_a + R_b + R_f, \tag{4.9a}$$

$$\text{s.t.} \qquad \frac{1}{2} \le \alpha_1, \alpha_2 \le 1, \tag{4.9b}$$

$$0 \le P_a, P_b \le P_{\text{d}}, \tag{4.9c}$$

$$R_j, R_{k\to f} \ge R_{\text{th}}, \qquad j \in \mathcal{J}, k \in \mathcal{K}. \tag{4.9d}$$

The formulated optimization problem is neither concave nor quasi-concave. The optimal solution to this optimization problem is challenging to obtain in practice. Hence, to overcome this issue, we propose a near-optimal closed-form expression for power control, which is detailed in the following subsection.

4.4.2.2 Proposed Approach

As it can be seen from their expressions, the rates R_a and R_b of the two CCUs are increasing when the power splitting factors from the BSs to the CEU, i.e., α_1 and α_2, and the transmit power from the two CCUs, i.e. P_a and P_b, decrease. Since, the CCUs have better channel conditions than the CEU, it is reasonable to allocate most of the

power of each BS to its associated CCU, with the consideration of the QoS at the CEU. In addition, allocating more transmit power at the CCU will harm its performance due to the SI effect. Therefore, our approach consists of determining the lowest possible power coefficients (α_1, α_2) and the lowest possible D2D powers at the CCUs (P_a, P_b) that achieve the QoS constraint at the CEU as well as a successful SIC at CCUs. Recall that the rate of the CEU should satisfy $R_{\text{th}} \leq R_f$. However, our approach consists of forcing the different terms of R_f in (4.8) to be equal to R_{th}. Based on this, assuming $R_{a\to f} = R_{\text{th}}$ and $R_{b\to f} = R_{\text{th}}$ implies that

$$P_{k_i} = \frac{\gamma_{i,k_i} P_i (1+\beta)}{\beta \gamma_{k_i,\text{SI}}} \alpha_i - \frac{\gamma_{i,k_i} P_i + 1}{\gamma_{k_i,\text{SI}}}, \tag{4.10}$$

where $\beta = \exp(R_{\text{th}}) - 1$ and $k_i = a$ when $i = 1$ and $k_i = b$ when $i = 2$. Therefore, since $0 \leq P_{k_i} \leq P_{\text{d}}$, then for all $i \in \mathcal{I}, \alpha_i$ should satisfy

$$\frac{\beta(\gamma_{i,k_i} P_i + 1)}{\gamma_{i,k_i} P_i (\beta+1)} \leq \alpha_i \leq \frac{\beta \gamma_{k_i,\text{SI}}}{\gamma_{i,k_i} P_i (\beta+1)} P_{\text{d}} + \frac{\beta(\gamma_{i,k_i} P_i + 1)}{\gamma_{i,k_i} P_i (\beta+1)}. \tag{4.11}$$

By replacing the expressions of P_{k_i} in (4.10) into (4.5) we obtain

$$R_{k_i} = \log\left[1 + \frac{(1-\alpha_i)\gamma_{i,k_i} P_i \beta}{\gamma_{i,k_i} P_i (1+\beta+\zeta\beta)\alpha_i - \gamma_{i,k_i} P_i \beta}\right]. \tag{4.12}$$

On the other hand, since $R_{k_i} \geq R_{\text{th}}$ and using the D2D power expressions in (4.10), then α_i should satisfy

$$\alpha_i \leq \frac{1+\beta}{2+\beta+\zeta\beta}. \tag{4.13}$$

Thus, by considering the inequalities in (4.11) and (4.13), and by taking into consideration (9b), α_i, for $i \in \mathcal{I}$, should satisfy $\alpha_{i,\min} \leq \alpha_1 \leq \alpha_{i,\max}$, where

$$\begin{cases} \alpha_{i,\min} = \max\left(\frac{1}{2}, \frac{\beta(\gamma_{i,k_i} P_i + 1)}{\gamma_{i,k_i} P_i (\beta+1)}\right), \\ \alpha_{i,\max} = \min\left(\frac{1+\beta}{2+\beta+\zeta\beta}, \frac{\beta\gamma_{k_i,\text{SI}}}{\gamma_{i,k_i} P_i (\beta+1)} P_{\text{d}} + \frac{\beta(\gamma_{i,k_i} P_i + 1)}{\gamma_{i,k_i} P_i (\beta+1)}\right), \end{cases} \tag{4.14}$$

After replacing the expressions of P_a and P_b in (4.10) into the rate $R_{f,\text{CoMP}}$ and making $R_{f,\text{CoMP}} = R_{\text{th}}$, we obtain the following equality on α_1 and α_2.

$$\alpha_2 = \frac{-C_1}{C_2}\alpha_1 + \frac{C_3}{C_2}, \tag{4.15}$$

where

$$\begin{cases} C_i = (1+\beta) P_i \left(\gamma_{i,f} + \frac{\gamma_{i,k_i}(\gamma_{k_i,f} - \beta\mu_i)}{\beta\gamma_{k_i,\text{SI}}}\right), \quad \forall i \in \{1,2\}, \\ C_3 = \beta(\gamma_{1,f} P_1 + \gamma_{2,f} P_2 + 1) + \frac{(\gamma_{a,f} - \beta\mu_1)(\gamma_{1,a} P_1 + 1)}{\gamma_{a,\text{SI}}} \\ \qquad + \frac{(\gamma_{b,f} - \beta\mu_2)(\gamma_{2,b} P_2 + 1)}{\gamma_{b,\text{SI}}}. \end{cases} \tag{4.16}$$

Consequently, substituting α_2 by its expression on R_b, we obtain

$$R_b = \log\left[1 + \frac{\frac{C_1}{C_2}\gamma_{2,b} P_2 \beta \alpha_1 + \left(1 - \frac{C_3}{C_2}\right)\gamma_{2,b} P_2 \beta}{-\frac{C_1}{C_2}\gamma_{2,b} P_2 (1+\beta+\zeta\beta)\alpha_1 + \gamma_{2,b} P_2 \left(\frac{C_3}{C_2}(1+\beta+\zeta\beta) - \beta\right)}\right]. \tag{4.17}$$

Table 4.1: Optimal α_1

	Solutions of f_1	$C_1 \leq C_2$	$C_1 \geq C_2$
$\Delta < 0$	$\emptyset$	$\alpha'_{1,\max}$	$\alpha'_{1,\min}$
$\Delta = 0$	$\alpha_0 \leq \alpha'_{1,\min}$	$\alpha'_{1,\max}$	$\alpha'_{1,\min}$
	$\alpha'_{1,\min} \leq \alpha_0$ $\alpha_0 \leq \alpha'_{1,\max}$	$\operatorname{argmax}(y_1, y_2)$	α_0
	$\alpha'_{1,\max} \leq \alpha_0$	$\alpha'_{1,\min}$	$\alpha'_{1,\max}$
$\Delta > 0$	$\alpha_0^1, \alpha_0^2 \leq \alpha'_{1,\min}$	$\alpha'_{1,\max}$	$\alpha'_{1,\min}$
	$\alpha_0^1 \leq \alpha'_{1,\min}$ $\alpha'_{1,\min} \leq \alpha_0^2 \leq \alpha'_{1,\max}$	$\operatorname{argmax}(y_1, y_2)$	α_0^2
	$\alpha_0^1 \leq \alpha'_{1,\min}$ $\alpha'_{1,\max} \leq \alpha_0^2$	$\alpha'_{1,\min}$	$\alpha'_{1,\max}$
	$\alpha'_{1,\min} \leq \alpha_0^1 \leq \alpha'_{1,\max}$ $\alpha'_{1,\min} \leq \alpha_0^2 \leq \alpha'_{1,\max}$	$\operatorname{argmax}(f(\alpha_0^1), y_2)$	$\operatorname{argmax}(y_1, f(\alpha_0^2))$
	$\alpha'_{1,\min} \leq \alpha_0^1 \leq \alpha'_{1,\max}$ $\alpha'_{1,\max} \leq \alpha_0^2$	α_0^1	$\operatorname{argmax}(y_1, y_2)$
	$\alpha'_{1,\max} \leq \alpha_0^1, \alpha_0^2$	$\alpha'_{1,\max}$	$\alpha'_{1,\min}$

In addition, since $\alpha_{1,\min} \leq \alpha_1 \leq \alpha_{1,\max}$, $\alpha_{2,\min} \leq \alpha_2 \leq \alpha_{2,\max}$ and $\alpha_2 = \frac{-C_1}{C_2}\alpha_1 + \frac{C_3}{C_2}$, we deduce that

$$\alpha'_{1,\min} \leq \alpha_1 \leq \alpha'_{1,\max}, \tag{4.18}$$

where

$$\begin{cases} \alpha'_{1,\min} = \max\left(\alpha_{1,\min}, \dfrac{C_3 - C_2\alpha_{2,\max}}{C_1}\right), \\ \alpha'_{1,\max} = \min\left(\alpha_{1,\max}, \dfrac{C_3 - C_2\alpha_{2,\min}}{C_1}\right). \end{cases} \tag{4.19}$$

Based on this, solving problem $\mathcal{P}$ is reduced to solving the following optimization problem

$$\mathcal{P}' : \quad \alpha_1^* = \operatorname{argmax}_{[\alpha'_{1,\min}, \alpha'_{1,\max}]} f(\alpha_1), \tag{4.20}$$

where the function f is expressed as

$$f(\alpha_1) = \log\left[1 + \frac{-\alpha_1 a_1 + a_1}{\alpha_1 a_2 - a_1}\right] + \log\left[1 + \frac{\alpha_1 b_1 + b_2}{-\alpha_1 b_3 + b_4}\right], \tag{4.21}$$

such that

$$\begin{cases} a_1 = \gamma_{1,a} P_1 \beta, \quad a_2 = \gamma_{1,a} P_1 (1 + \beta + \zeta\beta), \\ b_1 = \dfrac{C_1}{C_2}\gamma_{2,b} P_2 \beta, \quad b_2 = \left(1 - \dfrac{C_3}{C_2}\right)\gamma_{2,b} P_2 \beta, \\ b_3 = \dfrac{C_1}{C_2}\gamma_{2,b} P_2 (1 + \beta + \zeta\beta), \\ b_4 = \gamma_{2,b} P_2 \left(\dfrac{C_3}{C_2}(1 + \beta + \zeta\beta) - \beta\right). \end{cases} \tag{4.22}$$

Consequently, in order to solve problem $\mathcal{P}'$, one just needs to determine the extrema of

the function f within the interval $[\alpha'_{1,\min}, \alpha'_{1,\max}]$. The derivative of the function f with respect to α_1 is given by

$$\frac{\partial f}{\partial \alpha_1}(\alpha_1) = \frac{f_1(\alpha_1)}{f_2(\alpha_1)}, \tag{4.23}$$

where

$$\begin{cases} f_1(\alpha_1) = d_1\alpha_1^2 + d_2\alpha_1 + d_3, \\ f_2(\alpha_1) = A\alpha_1\left(\alpha_1 - x_1\right)\left(\alpha_1 - x_2\right)\left(\alpha_1 - x_3\right), \end{cases} \tag{4.24}$$

such that

$$\begin{cases} d_1 = \dfrac{C_1}{C_2}\left(\dfrac{C_1}{C_2} - 1\right)(1 + \beta + \zeta\beta), \\ d_2 = -2\dfrac{C_1}{C_2}\left(\dfrac{C_3}{C_2}(1 + \beta + \zeta\beta) - \beta\right), \\ d_3 = \dfrac{C_3}{C_2}\left(\dfrac{C_3}{C_2}(1 + \beta + \zeta\beta) - \beta\right), \ A = -\dfrac{C_1^2(1 + \beta + \zeta\beta)^2}{C_2^2\beta}, \\ x_1 = \dfrac{\beta}{1 + \beta + \zeta\beta}, \ x_2 = \dfrac{C_3}{C_1} - \dfrac{C_2\beta}{C_1\left(1 + \beta + \zeta\beta\right)}, \ x_3 = \dfrac{C_3}{C_1}. \end{cases} \tag{4.25}$$

Now, note that $x_1 = \frac{\beta}{1+\beta+\zeta\beta} < \frac{\beta}{1+\beta} < \frac{\beta(P_1\gamma_{1,a}+1)}{P_1\gamma_{1,a}(1+\beta)} \leq \alpha_{1,\min} \leq \alpha'_{1,\min}$. In addition, since $\alpha'_{1,\max} \leq \frac{C_3}{C_1} - \frac{C_2}{C_1}\alpha_{2,\min}$ and $\alpha_{2,\min} \geq \frac{\beta=\left(P_2\gamma_{2,b}+1\right)}{P_2\gamma_{2,b}(\beta+1)} \geq \frac{\beta}{1+\beta} > \frac{\beta}{1+\beta+\zeta\beta}$, then $\alpha'_{1,\max} \leq \frac{C_3}{C_1} - \frac{C_2}{C_1}\frac{\beta}{1+\beta+\zeta\beta} = x_2$. Moreover, since $x_2 < x_3$, then $\alpha'_{1,\max} \leq x_3$. Therefore, we conclude that $x_1 \leq \alpha'_{1,\min} \leq \alpha'_{1,\max} \leq \min\left(x_2, x_3\right)$. Moreover, since $A < 0$, we conclude that the function f_2 is strictly negative in the interval $\left[\alpha'_{1,\min}, \alpha'_{1,\max}\right]$. On the other hand, the sign of the function f_1 depends mainly on the sign of its discriminant $\Delta = d_2^2 - 4d_1d_3$ and the sign of d_1, which is the same as the sign of $C1 - C2$.

Based on the above analysis, the optimal solution for α_1 is summarized in Table 4.1, where $y_1 = f\left(\alpha'_{1,\min}\right)$, $y_2 = f\left(\alpha'_{1,\max}\right)$, $\alpha_0 = \frac{-d_2}{2d_1}$ is the root of f_2 when $\Delta = 0$ and $\alpha_0^1, \alpha_0^2 = \frac{-d_2 \pm \sqrt{d_2^2 - 4d_1d_2}}{2d_1}$ are the roots of f_2 when $\Delta > 0$, such that $\alpha_0^1 < \alpha_0^2$. Hence, we conclude from Table 4.1 that the optimal solution α_1^* is either $\alpha'_{1,\min}$ or $\alpha'_{1,\max}$ or one of the roots of f_2 if they existed and they are within the interval $\left[\alpha'_{1,\min}, \alpha'_{1,\max}\right]$. Finally, the proposed power control scheme of the CoMP-empowered C-NOMA system is given in the **Algorithm 1**, where $\mathbb{I}(\cdot)$ denotes the indicator function within the interval $[\alpha'_{1,\min}, \alpha'_{1,\max}]$.

4.4.3 Simulation Results

In this section, the simulation results are presented to validate the performance of the proposed CoMP-assisted C-NOMA scheme. The simulation results are obtained through 10^5 independent Monte-Carlo trials. We consider the same simulation setup in [32]. In addition, unless otherwise stated, the main simulation setting is presented in Table 4.2.

Algorithm 1: Proposed Power Control Scheme

Initialization:

i) Estimate the environment parameters;

ii) Express f as shown in (21);

iii) Calculate d_1, d_2 and d_3 as shown in (25) and $\Delta = d_2^2 - 4d_1d_3$;

iv) Calculate $y_1 = f\left(\alpha'_{1,\min}\right)$ and $y_2 = f\left(\alpha'_{1,\max}\right)$;

if $\Delta < 0$ **then**

$\alpha_1^* = \text{argmax}\left(y_1, y_2\right)$;

end if

else if $\Delta = 0$ **then**

i) $\alpha_0 = -\frac{d_2}{2d_1}$;

ii) $\alpha_1^* = \text{argmax}\left(y_1, y_2, f\left(\alpha_0\right)\mathbb{I}(\alpha_0)\right)$;

end if

else

i) $\alpha_0^1 = \min\left(\frac{-d_2-\sqrt{d_2^2-4d_1d_2}}{2d_1}, \frac{-d_2+\sqrt{d_2^2-4d_1d_2}}{2d_1}\right)$;

ii) $\alpha_0^2 = \max\left(\frac{-d_2-\sqrt{d_2^2-4d_1d_2}}{2d_1}, \frac{-d_2+\sqrt{d_2^2-4d_1d_2}}{2d_1}\right)$;

iii) $\alpha_1^* = \text{argmax}\left(y_1, y_2, f\left(\alpha_0^1\right)\mathbb{I}(\alpha_0^1), f\left(\alpha_0^2\right)\mathbb{I}(\alpha_0^2)\right)$;

end if

$\alpha_2^* = \frac{-C_1}{C_2}\alpha_1^* + \frac{C_3}{C_2}$;

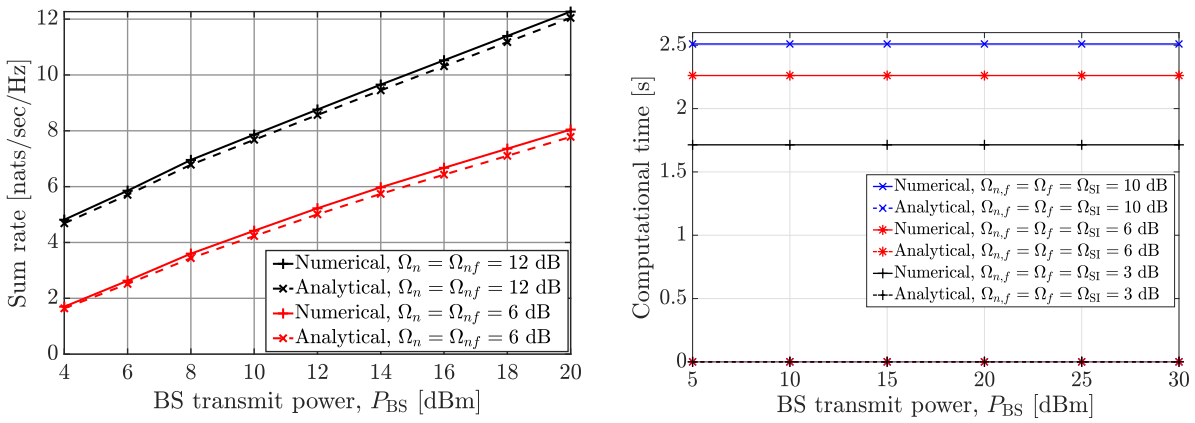
$P_a^* = \frac{\gamma_{1,a}P_1(1+\beta)}{\beta\gamma_{a,\text{SI}}}\alpha_1^* - \frac{\gamma_{1,a}P_1+1}{\gamma_{a,\text{SI}}}$;

$P_b^* = \frac{\gamma_{2,b}P_2(1+\beta)}{\beta\gamma_{b,\text{SI}}}\alpha_2^* - \frac{\gamma_{2,b}P_2+1}{\gamma_{b,\text{SI}}}$;

Table 4.2: Simulation Parameters

Parameter	Symbol	Value
Channel gain parameter between BS_1 and UE_a and between BS_2 and UE_b	$\Omega_{1a} = \Omega_{2b} = \Omega_n$	12 dB
Channel gain parameter between BS_1 and UE_f and between BS_2 and UE_f	$\Omega_{1f} = \Omega_{2f} = \Omega_f$	3 dB
Channel gain parameter between UE_a and UE_f and between UE_b and UE_f	$\Omega_{af} = \Omega_{bf} = \Omega_{nf}$	12 dB
SI channel gain	Ω_{SI}	3 dB
Power budget of the BS_1 and BS_2	$P_1 = P_2 = P_{\text{BS}}$	20 dBm
Power budget for UE_a and UE_b	P_d	10 dBm
Rate threshold	R_{th}	1 bits/sec/Hz
Noise variance	σ^2	1

Fig. 4.2 compares the performance of the optimal power control scheme with one of the proposed schemes in terms of accuracy and computational time. Fig. 4.2(a) presents the



(a) (b)

Figure 4.2: (a) sum-rate versus the power budget at the BS when $\Omega_{\mathrm{SI}} = 3$ dB and (b) Computational time versus the power budget at the BS.

analytical and numerical average sum-rate versus the power budget at the BSs P_{BS}. The analytical results are obtained by running **Algorithm 1** whereas the numerical results are obtained by solving the main problem $\mathcal{P}$ in (4.9) using an off-the-shelf optimization solver, where we generate 100 initial points and then we select the best result. It can be seen from Fig. 4.2(a) that the near-optimal closed form PA gets around 95% of the sum-rate of the optimal one obtained through an expensive numerical method. It is worth noting that the gap between the proposed and the optimal power control scheme is due to the fact that the proposed approach consists of forcing the different terms of R_f in (4.8) to be equal to R_{th}, and hence, loosing in terms of optimality. In addition, Fig. 4.2(b) shows that the computational time of the proposed scheme is negligible compared to the optimal one, which validates the potential of the proposed power control scheme. In order to show the effectiveness of integrating CoMP with C-NOMA under two-cell scenario, we compare the performance of the proposed scheme with the conventional C-NOMA scheme proposed in [32]. For a fair comparison, we sum the achievable rate at one CCU along with the CEU's achievable rate. Meanwhile, for the C-NOMA, the signal received at the CEU from BS_2 is treated as noise. Fig. 4.3(a) compares the performance of the joint CoMP C-NOMA scheme with the conventional C-NOMA scheme when the SI channel gain increases. As depicted in this figure, the proposed scheme provides a significant performance improvement over the conventional C-NOMA scheme. This is resulting from the consideration of the ICI mitigation technique in our model. In addition, increasing the SI channel gain forces the CCU to transmit with low power in order to avoid the

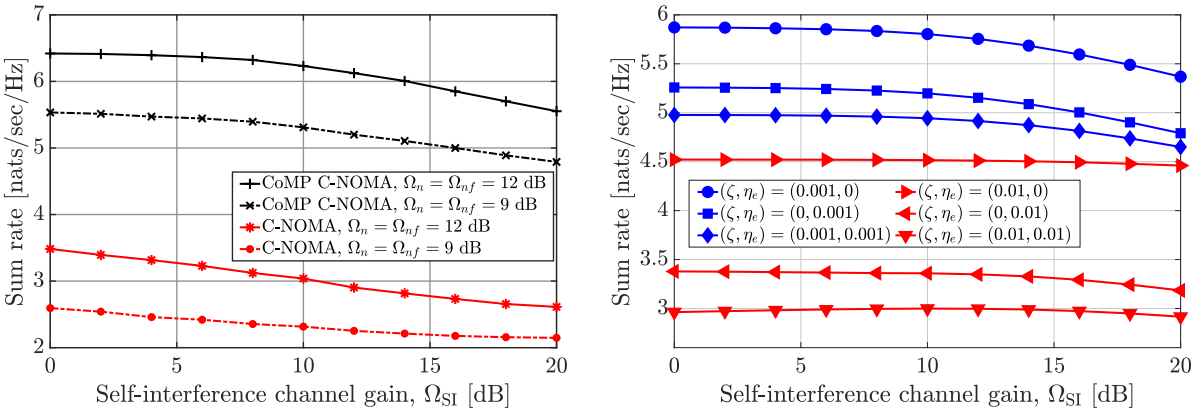


Figure 4.3: (a) shows sum-rate versus the SI parameter and (b) Network sum-rate of the coordinated C-NOMA cluster versus the SI channel gain Ω_{SI}, for different values of the channel estimation error and the imperfect SIC process factor, when $\Omega_n = \Omega_{nf} = 12$ dB.

SI harmful effect. This means that the BSs should allocate more power to the CEU to achieve its required QoS, which degrades the sum-rate performance.

Fig. 4.3(b) presents the sum-rate of the coordinated C-NOMA cluster versus the SI channel gain Ω_{SI}, for different values of the channel estimation error and of the imperfect SIC process factor. As observed from Fig. 4.3(b), the sum-rate decreases when the errors from both imperfect SIC and CSI estimation increase. It can be also seen that for high values of both errors, the interference generated from both types of errors dominates the performance in comparison with the SI. In addition, the channel estimation error is seen more effects on the sum-rate performance than that the effect of the imperfect SIC. After discussing the PA in one coordinated C-NOMA cluster, we move on to the investigation of the performance of the network when a multi-user scenario is considered.

4.5 Joint UE Clustering and Power Allocation in CoMP C-NOMA Networks

4.5.1 System Model

We consider a network model consisting of a DL transmission of two adjacent cells as illustrated in Fig. 4.4, where each cell is equipped with one BS similar to part I of this chapter.[2] Two types of UEs, namely, CCUs and CEUs (the UEs located in the overlapped

[2]Our system model can be easily extended to a multi-cell scenario. This can be achieved by allowing any two-adjacent cells in a multi-cell scenario to coordinate together in serving their associated UEs,

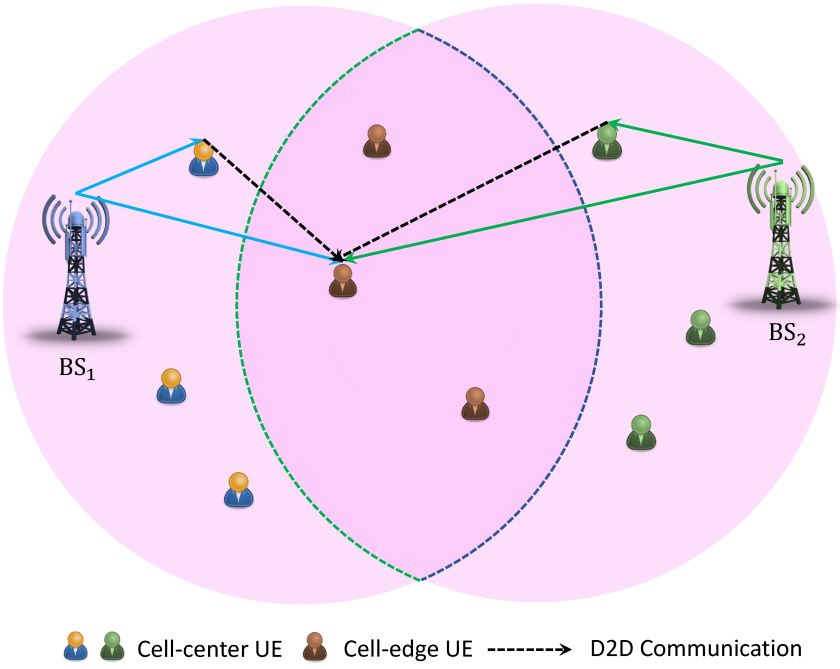


Figure 4.4: An illustration of the proposed CoMP C-NOMA model.

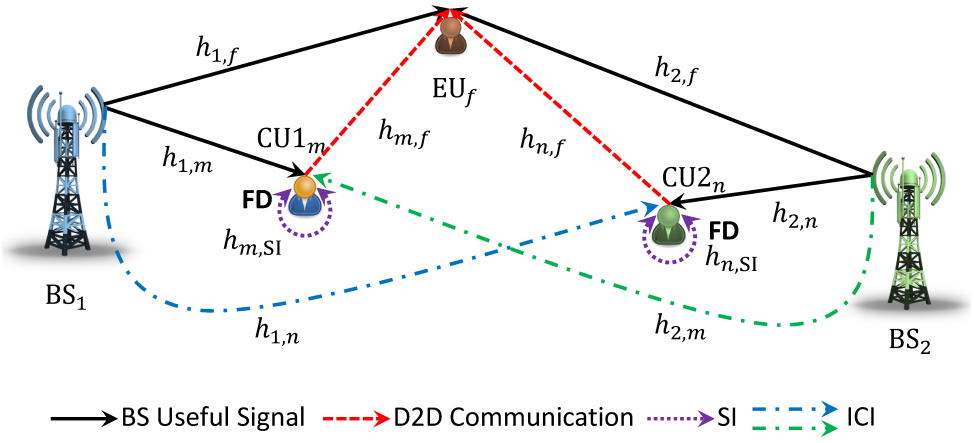


Figure 4.5: Details explanation of one coordinated C-NOMA cluster.

area as shown in Fig. 4.4) are considered within each cell. Within this setup, each cell invokes the C-NOMA technique to serve its associated UEs, while the CoMP technique between the two BSs is introduced to assist the CEUs. Consequently, each BS forms C-NOMA clusters, each of a size of two UEs by including one CCU and one CEU. In addition, each CCU relays the message of the CEU belonging to the same C-NOMA cluster in an FD DF relaying mode. We assume that the UEs and the BSs have each one transmit antenna and one received antenna [67, 109, 113].

Notably, based on this set-up, the CEUs do not suffer from ICI; however, they only experience INUI due to the transmissions of the messages of the CCUs in the same time-

which is similar to the proposed network architecture in [10, 11, 76, 108, 109, 112].

Table 4.3: Table of Notations

System Parameters	
P_i	Transmit power of BS_i
$P_{\max}$	Power budget of $\text{CCU}^i_{k_i}$
$\alpha_{i,k_i,f}$	PA coefficient by BS_i to CEU_f paired with $\text{CCU}^i_{k_i}$
$\beta_{i,k_i,f}$	PA coefficient allocated by $\text{CCU}^i_{k_i}$ associated with BS_i to CEU_f
$x_{m,f,n}$	UC indicator
$\mathcal{M},\mathcal{N},\mathcal{F},\mathcal{I},\mathcal{U},\mathcal{K}$	Set of CCU^1, CCU^2, CEU, BSs, UEs, near UEs
$\mathcal{S}$	Set of preference profiles
Δ	Matching triple
U_Φ	Social welfare of the network
$R_{m,f,n}$	Achievable rate of one C-NOMA cluster (CCU^1_m, CEU_f, CCU^2_n)
Φ	Matching function
R_{th}	Minimum required rate

frequency resource. On the other hand, a CCU receives its signal from the adjacent BS only.[3] Hence, this UE belongs to only one C-NOMA cluster and experiences two types of interference:

- *The inter-cell interference*: Since each CCU is associated with a single BS, the transmission from the other BS is considered as interference at this UE. For instance, the signal transmitted by BS_2 is treated as interference at CCU^1_m (the dashed green line in Fig. 4.5.).

- *The Self-interference*: Since we consider a FD DF relaying between the CCU and the CEU, the CCU suffers from a SI resulting from receiving data from the BS and transmitting data to the CEU simultaneously within the same resource (the dotted purple line in Fig. 4.5.).

In this part, different from the definition of a C-NOMA cluster reported in the literature [72, 104, 105, 107], we define a coordinated C-NOMA cluster as a group of three UEs, where two of them are two CCUs, each of them is served by its adjacent BS, and one CEU served simultaneously by the two BSs. Since multiple coordinated C-NOMA clusters are constructed by the two BSs, a promising technique to avoid the inter-cluster interference

[3]In order to achieve a higher performance gain compared to orthogonal multiple access (OMA) techniques, the disparity in the channel gains between the far NOMA user, which is the CEU, and the near NOMA UE, which is the CCU, should be large. As a result, we limit the distribution of the CCUs to be near the BSs; meanwhile, the CEUs are distributed on the edge of the cells. Accordingly, the interference between the two CCUs is neglected due to the large distance between them, which is validated in the simulation section. Note that, this interference is due to the relaying of the CEUs' data by the CCUs in the CT phase.

within each cell is to utilize a hybrid multiple access technique, in which C-NOMA is combined with a conventional OMA technique such as FDMA or TDMA [69, 98, 107, 114]. In this paper, we consider an FDMA technique to avoid inter-cluster interference [69,98,107]. Specifically, the overall bandwidth is equally divided into a fixed number of equally-sized radio channels that have equal transmit powers and each coordinated C-NOMA cluster is allocated one radio channel [69, 98, 107].

For notation convenience, $\mathcal{I} \triangleq \{1, 2\}, \mathcal{M} \triangleq \{1, 2, \ldots, M\}, \mathcal{F} \triangleq \{1, 2, \ldots, F\}$ and $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$ define the set of BSs, the set of the CCUs associated with BS_1 (denoted as CCU^1), the set of CEUs, and the set of CCUs associated with BS_2 (denoted as CCU^2), respectively. Moreover, M, F and N denote the cardinality of $\mathcal{M}, \mathcal{F}$, and $\mathcal{N}$, respectively. The set of all CCUs in both cells is denoted by $\mathcal{K} \triangleq \{\mathcal{M} \cup \mathcal{N}\}$ and the set of all UEs in the network is denoted by $\mathcal{U} \triangleq \{\mathcal{M} \cup \mathcal{F} \cup \mathcal{N}\}$. Moreover, $\forall i \in \mathcal{I}, u \in \mathcal{U}$, $k \in \mathcal{K}$ and $f \in \mathcal{F}$, $h_{i,u}$, $h_{k,f}$ and $h_{k,\text{SI}}$, denote the channel gains from BS_i to UE u, from the CCU_k to CEU_f, and from the CCU_k to itself (SI caused by the FD relaying mode), respectively.[4] Finally, $\forall i \in \mathcal{I}, P_i$ is the allocated power by BS_i for each coordinated C-NOMA cluster, where $\mathcal{I} = \{1, 2\}$.

4.5.1.1 Transmission Model

In C-NOMA-enabled cellular network, the data transmission is executed over two different phases [13, 29, 30, 32, 33, 107]. The first phase is the DT phase and the second phase is the CT phase, which are detailed as follows [13, 29, 30, 32, 33, 107].

- **DT phase**: In this phase, the two BSs, i.e. BS_1 and BS_2, apply superposition coding on the signals of their associated NOMA pairs, which are (CCU^1_m, EU_f) for BS_1 and (CCU^2_n, EU_f) for BS_2. Then, the superimposed signals are transmitted simultaneously by the BSs to their associated NOMA pairs. Following the NOMA principle, each CCU first performs SIC to decode the signal of the CEU_f. Second, each CCU removes the decoded signal of CEU_f from its own reception and then decodes its signal from the resulting reception. Meanwhile, the CEU_f treats the signals of the CCUs as noise.

[4] Through passive SI suppression techniques, such as antennae separation, wavetraps, etc., followed by the active SI suppression techniques, i.e., analog and digital cancellation, the SI in a SISO set-up can be effectively mitigated [30]. Although the effectiveness of these SI suppression techniques was demonstrated in the literature, the SI effect cannot be totally canceled. Hence, the effect of the residual SI is considered in our model.

- **CT phase**: In this phase, each CCU, i.e., CCU_m^1 and CCU_n^2, forwards the CEU's decoded message during the SIC process to the CEU_f through a D2D channel.

Since the CCUs adopt FD relaying mode, the two phases occur within the same time-slot, with the penalty of inducing SI at the CCUs (CCU_m^1 and CCU_n^2). Based on this, we evaluate in the following the received `SINR` and the corresponding achievable rates at both CCUs and CEUs.

4.5.1.2 Signal Model and `SINR` Analysis

We consider a single coordinated C-NOMA cluster as shown in Fig. 4.5. We start by describing the signal model for the CCUs, and then, we present the signal model for the CEU. In particular, $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, assuming that CCU_m^1, CEU_f and CCU_n^2 form a coordinated C-NOMA cluster as shown in Fig. 4.5, BS_1 broadcasts a superimposed mixture of the messages intended to CCU_m^1 and CEU_f as $z_{1,m,f} = \sqrt{(1-\alpha_{1,m,f})\,P_1} z_{1,m} + \sqrt{\alpha_{1,m,f} P_1} z_f$, where $z_{1,m}$ and z_f are the signals intended to CCU_m^1 and CEU_f, respectively, and $\alpha_{1,m,f}$ is the PA coefficient assigned by BS_1 to CEU_f. According to NOMA principle, since CCU_m^1 should be able to decode the message of CEU_f, the PA coefficient $\alpha_{1,m,f}$ should satisfy $0.5 < \alpha_{1,m,f} < 1$ [115, 116]. Based on this and due to the consideration of DF FD relaying mode, CCU_m^1 first decodes the message of CEU_f and then transmits the decoded message to this CEU. Consequently, the received signal at CCU_m^1 is expressed as

$$y_{1,m} = h_{1,m} z_{1,m,f} + \sqrt{\beta_{1,m,f} P_{\max}} h_{m,\text{SI}} z_f + h_{2,m} z_{2,n,f} + w_{1,m}, \tag{4.26}$$

where $\beta_{1,m,f} \in [0,1]$ is the fraction of the allocated power by CCU_m^1 to CEU_f, $P_{\max}$ is the maximum allowable transmit power at each CCU, $\sqrt{\beta_{1,m,f} P_{\max}} h_{m,\text{SI}} z_f$ is the SI signal resulting from the transmission of the decoded message of CEU_f, i.e. z_f, by CCU_m^1, $z_{2,n,f}$ is the broadcast signal by BS_2, which causes an ICI at that UE, and $w_{1,m}$ is an AWGN experienced at CCU_m^1, which is $\sim \mathcal{CN}(0, \sigma^2)$ distributed. Hence, the received `SINR` at CCU_m^1 to decode the message of CEU_f is given by

$$\delta_{1,m\to f} = \frac{\alpha_{1,m,f}\gamma_{1,m}}{(1-\alpha_{1,m,f})\gamma_{1,m} + \beta_{1,m,f}\gamma_{m,\text{SI}} + 1}, \tag{4.27}$$

where $\gamma_{1,m} = P_1 |h_{1,m}|^2 / \mathcal{T}$, $\gamma_{m,\text{SI}} = P_{\max} |h_{m,\text{SI}}|^2 / \mathcal{T}$, and $\mathcal{T} \triangleq P_2 |h_{2,m}|^2 + \sigma^2$. Moreover, the received `SINR` at CCU_m^1 to decode its own message after canceling the CEU_f's signal is expressed as

$$\delta_{1,m\to m} = \frac{(1-\alpha_{1,m,f})\gamma_{1,m}}{\beta_{1,m,f}\gamma_{m,\text{SI}} + 1}. \tag{4.28}$$

Thus, the achievable rate at CCU_m^1 to decode the message of CEU_f and to decode its own message are expressed, respectively, as

$$\mathcal{R}_{1,m\to f} = \log(1+\delta_{1,m\to f}), \tag{4.29}$$

$$\mathcal{R}_{1,m} = \log(1+\delta_{1,m\to m}). \tag{4.30}$$

By following the same steps as in (4.26)-(4.30), the received SINR and the achievable rate at CCU_n^2 to decode the message of CEU_f, i.e $\mathcal{R}_{2,n\to f}$, and to decode its own message, i.e., $\mathcal{R}_{2,n}$, can be obtained. Note that both CCU_m^1 and CCU_n^2, which belong to the same coordinated C-NOMA cluster along with CEU_f, will decode and forward the CEU_f's message. Consequently, the CEU_f receives four copies of its own message, two streams from CCU_m^1 and CCU_n^2 through the D2D relaying and two streams from the two BSs through CoMP. Therefore, the received signal at CEU_f can be expressed as

$$y_f = \underbrace{\left(\overbrace{\sqrt{\alpha_{1,m,f}}z_f}^{\text{useful signal}} + \overbrace{\sqrt{(1-\alpha_{1,m,f})z_{1,m}}}^{\text{INUI}}\right)\sqrt{P_1}h_{1,f} +}_{\text{received signal from BS}_1}$$

$$\underbrace{\left(\overbrace{\sqrt{\alpha_{2,n,f}}z_f}^{\text{useful signal}} + \overbrace{\sqrt{(1-\alpha_{2,n,f})z_{2,n}}}^{\text{INUI}}\right)\sqrt{P_2}h_{2,f} +}_{\text{received signal from BS}_2}$$

$$\underbrace{\sqrt{\beta_{1,m,f}P_{\max}}h_{m,f}z_f}_{\text{received signal from CCU}_m^1} + \underbrace{\sqrt{\beta_{2,n,f}P_{\max}}h_{n,f}z_f}_{\text{received signal from CCU}_n^2} + \underbrace{w_f}_{\text{noise at CEU}}, \tag{6}$$

where w_f is an AWGN experienced at CEU_f, which is $\sim \mathcal{CN}(0,\sigma^2)$ distributed. We assume that CCU_m^1 and CCU_n^2 have perfect SIC [10]. In our proposed model, four nodes, which are BS_1, BS_2, CCU_m^1, and CCU_n^2, cooperate to jointly serve the CEU_f. Consequently, the SINR expression at the CEU_f can be expressed as follows.

$$\delta_{f\to f} = \frac{\alpha_{1,m,f}\gamma_{1,f} + \alpha_{2,n,f}\gamma_{2,f} + \beta_{1,m,f}\gamma_{m,f} + \beta_{2,n,f}\gamma_{n,f}}{(1-\alpha_{1,m,f})\gamma_{1,f} + (1-\alpha_{2,n,f})\gamma_{2,f} + 1}, \tag{4.31}$$

where $\forall i \in \mathcal{I}, \gamma_{i,f} = \frac{P_i|h_{i,f}|^2}{\sigma^2}$ and $\gamma_{k,f} = \frac{P_{\max}|h_{k,f}|^2}{\sigma^2}$. The derivation of the SINR $\delta_{f\to f}$ in (4.31) is provided in Appendix B2. One can observe that the SINR is improved with the amalgamation between CoMP and C-NOMA. This is motivated by the fact that allowing two cells to serve the ICI-prone UE will not only mitigate the ICI effects but also improve the signal reception through the D2D communication from the two CCUs. However, since CCU_m^1 and CCU_n^2 operate in DF relaying mode, the achievable rate of CEU_f is limited

by the weakest link and it is expressed as [11]

$$\mathcal{R}_f = \min\{\mathcal{R}_{m\to f}, \mathcal{R}_{f,\text{CoMP}}, \mathcal{R}_{n\to f}\}, \tag{4.32}$$

where

$$\mathcal{R}_{f,\text{CoMP}} = \log(1+\delta_{f\to f}). \tag{4.33}$$

Based on the above analysis, the achievable sum rate of the coordinated C-NOMA cluster (CCU_m^1, CEU_f, CCU_n^2) is expressed as

$$\mathcal{R}_{m,f,n} = \mathcal{R}_{1,m} + \mathcal{R}_f + \mathcal{R}_{2,n}. \tag{4.34}$$

Now, let us define $\mathcal{X} = \{x_{m,f,n} \in \{0,1\} | \, m \in \mathcal{M}, f \in \mathcal{F}, \;\; n \in \mathcal{N}\}$, such that, $x_{m,f,n}$ is a binary UC indicator in which $x_{m,f,n} = 1$ when CCU_m^1, CEU_f, and CCU_n^2 form one coordinated C-NOMA cluster and $x_{m,f,n} = 0$ otherwise, $\boldsymbol{\alpha} = \{(\alpha_{1,m,f}, \;\; \alpha_{2,n,f}) \in [0.5,1]^2 | \; m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$ is the collection of the BSs PA coefficients, and $\boldsymbol{\beta} = \{(\beta_{1,m,f}, \; \beta_{2,n,f}) \in [0,1]^2 | \;\; m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$ is the collection of the PA fractions from the CCUs to the CEUs. Consequently, the network sum-rate, which is defined as the sum of the achievable sum rates of all the constructed coordinated C-NOMA clusters, can be expressed as

$$\mathcal{R}^{\text{sum}}(\mathcal{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{m\in\mathcal{M}} \sum_{f\in\mathcal{F}} \sum_{n\in\mathcal{N}} x_{m,f,n} \mathcal{R}_{m,f,n}. \tag{4.35}$$

4.5.2 Problem Formulation and Solution Methodology

With the goal of enhancing the performance of the proposed CoMP-assisted FD C-NOMA scheme, a framework that jointly optimizes the PA coefficients at the BSs, the PA fractions at the CCUs in each coordinated C-NOMA cluster, and the UC is proposed. The UC policy defines the members of each coordinated C-NOMA cluster, i.e., which two CCUs should be grouped with each CEU. The proposed framework is formulated as an optimization problem aiming at maximizing the network sum-rate while guaranteeing the required QoS at the UEs and the SIC constraints at the CCUs, which can be written as follows.

$$\mathcal{P}: \mathcal{R}^{\text{sum}}(\mathcal{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \tag{4.36a}$$

$$\text{s.t. } 0.5 x_{k_1,f,k_2} \le \alpha_{i,k_i,f} \le x_{k_1,f,k_2}, \qquad \forall i \in \mathcal{I}, k_i \in \mathcal{K}_i, f \in \mathcal{F}, \tag{4.36b}$$

$$0 \le \beta_{i,k_i,f} \le x_{k_1,f,k_2}, \qquad \forall i \in \mathcal{I}, k_i \in \mathcal{K}_i, f \in \mathcal{F}, \tag{4.36c}$$

$$\mathcal{R}_{i,k_i} \ge x_{k_1,f,k_2} R_{\text{th}}, \qquad \forall i \in \mathcal{I}, k_i \in \mathcal{K}_i, f \in \mathcal{F}, \tag{4.36d}$$

$$\mathcal{R}_f \ge x_{k_1,f,k_2} R_{\text{th}}, \qquad \forall f \in \mathcal{F}, \tag{4.36e}$$

$$\sum_{n\in\mathcal{N}}\sum_{f\in\mathcal{F}} x_{m,f,n} \leq 1, \qquad \forall m \in \mathcal{M}, \tag{4.36f}$$

$$\sum_{f\in\mathcal{F}}\sum_{m\in\mathcal{M}} x_{m,f,n} \leq 1, \qquad \forall n \in \mathcal{N}, \tag{4.36g}$$

$$\sum_{n\in\mathcal{N}}\sum_{m\in\mathcal{M}} x_{m,f,n} \leq 1, \qquad \forall f \in \mathcal{F}, \tag{4.36h}$$

$$x_{m,f,n} \in \{0,1\}, \qquad \forall m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}, \tag{4.36i}$$

where k_i and $\mathcal{K}_i$ are defined for all $i \in \mathcal{I}$, as

$$k_i = \begin{cases} m, & i = 1, \\ n, & i = 2\ , \end{cases} \qquad \mathcal{K}_i = \begin{cases} \mathcal{M}, & i = 1, \\ \mathcal{N}, & i = 2. \end{cases} \tag{4.37}$$

Looking deep into problem $\mathcal{P}$, constraint (4.36b) assures that more power is allocated to each CEU than the CCUs to achieve the NOMA principle, constraint (4.36c) indicates that the PA fractions from the CCUs to the CEU in the same coordinated C-NOMA cluster should be less than one. Constraints (4.36d) and (4.36e) guarantee the required QoS in terms of the data rate threshold R_{th} at the CCUs and CEU, respectively. In addition, constraint (4.36e) guarantees that each CCU can decode the message of the associated CEU, which ensures the SIC criteria at the CCU. Constraints (4.36f), (4.36g), and (4.36h) guarantee that each UE can be a member in at most one coordinated C-NOMA cluster. Finally, constraint (4.36i) guarantees that the value of the UEs clustering indicator is either 0 or 1.

One key component in the optimization problem $\mathcal{P}$ is the availability of the full CSI of all invoked wireless links in order to perform the optimization in a centralized manner. At this stage, it is important to mention first that the CSI between the BSs and their associated UEs can be obtained using one of the CSI acquisition techniques for multi-user wireless communication systems. For instance, the UL sounding reference signal transmitted by all UEs (CCUs and CEUs) in LTE is used to estimate the required channel gains for the communication links between the BSs and the cellular UEs in real time as in LTE-A standard. Nevertheless, considering the D2D and SI CSI, one can apply the same approach used in [117], where a method to obtain the full CSI in a source-relay-destination topology with FD relaying was proposed. Specifically, based on [117], we can assume in our system model that there is a control channel tunneling from the source (BS), the relays (CCUs), and the destinations (CEUs) for channel estimation. Based on this discussion, a perfect CSI is assumed to be available. On the other hand, it is worth

noting that the computation of the resource allocation usually takes place in a server located at the core network or in an edge server located near to one of the BSs.

To maximize the network sum-rate, the UC, the D2D PA, and the BSs PA coefficients should be jointly optimized. It can be seen that the formulated optimization problem $\mathcal{P}$ is an MINLP problem, which is hard to be solved in a straightforward manner. In order to overcome this issue, we solve the optimization problem $\mathcal{P}$ using the concept of bi-level optimization as detailed in the following [118]. In problem $\mathcal{P}$, it can be seen that the achievable sum rate expression $\mathcal{R}_{m,f,n}$ of a single coordinated C-NOMA cluster is only a function of the PA policy and is independent of the UC indicator $x_{m,f,n}$. Specifically, one can observe that, $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, if the optimal UEs cluster indicator $x^*_{m,f,n} = 0$, then the optimal power control policy $\left(\alpha^*_{1,m,f}, \beta^*_{1,m,f}, \alpha^*_{2,n,f}, \beta^*_{2,n,f}\right) = \mathbf{0}_{4\times 1}$. However, if $x^*_{m,f,n} = 1$, then $\left(\alpha^*_{1,m,f}, \beta^*_{1,m,f}, \alpha^*_{2,n,f}, \beta^*_{2,n,f}\right)$ should be the optimal solutions of the PA scheme of the coordinated C-NOMA cluster (CCU^1_m,CEU_f,CCU^2_n). In other words, if we assume that the UEs CCU^1_m, CEU_f and CCU^2_n are clustered together and that we can obtain their optimal PA scheme $\left(\alpha^*_{1,m,f}, \beta^*_{1,m,f}, \alpha^*_{2,n,f}, \beta^*_{2,n,f}\right)$, problem $\mathcal{P}$ becomes a linear assignment problem and hence the remaining part is to determine the optimal UEs clustering policy $x^*_{m,f,n}$. Thus, since we aim to obtain the optimal power control policy such that the network sum-rate of the CoMP-assisted FD C-NOMA system can be maximized, we can reduce the feasible set of the PA of problem $\mathcal{P}$ to the set of PA coefficients that maximize the achievable sum rate for each coordinated C-NOMA cluster. Therefore, the formulated optimization problem $\mathcal{P}$ is decoupled into two sub-problems based on the bi-level optimization approach [118], where the first is the PA sub-problem (inner problem) and the second is the UC sub-problem (outer problem) as shown below.

4.5.2.1 Power Control: CCUs and BSs Power Allocation Coefficients

In this part, we study the PA policy for each coordinated C-NOMA cluster (CCU^1_m,CEU_f, CCU^2_n), $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, by defining both the PA coefficients at the BSs and the PA fractions from CCU^1_m and CCU^2_n to CEU_f. As a result, the PA solution is obtained by maximizing the achievable sum rate for a given coordinated C-NOMA cluster (CCU^1_m,CEU_f,CCU^2_n), which can be formulated as follow.

$$\mathcal{P}_{m,f,n}: \quad \max_{\mathbf{c}_{m,f,n}} \ \mathcal{R}_{m,f,n}, \tag{4.38a}$$

$$\text{s.t.} \quad 0.5 \leq \alpha_{1,m,f}, \ \alpha_{2,n,f} \leq 1, \tag{4.38b}$$

$$0 \leq \beta_{1,m,f},\ \beta_{2,n,f} \leq 1, \tag{4.38c}$$

$$\mathcal{R}_{1,m}, \mathcal{R}_f, \mathcal{R}_{2,n} \geq R_{\text{th}}, \tag{4.38d}$$

where $\mathbf{c}_{m,f,n} = (\alpha_{1,m,f}, \beta_{1,m,f}, \alpha_{2,n,f}, \beta_{2,n,f})$. The formulated PA problem is neither concave nor quasi-concave and the optimal solution is challenging to obtain in practice. Nevertheless, since $\mathcal{P}_{m,f,n}$ must be solved for all possible combinations of (CCU_m^1,CEU_f,CCU_n^2), a computationally efficient approach for solving this problem needs to be investigated, which is introduced in Section 4.5.3.

4.5.2.2 UEs Clustering: Defining The Coordinated C-NOMA Members

The two coordinating BSs jointly perform the UC by grouping one CCU from each cell with one CEU shared between the two cells. Specifically, we need to define the optimal UEs to be grouped in each coordinated C-NOMA cluster given the optimal power control policy obtained from solving problem $\mathcal{P}_{m,f,n}$ for all $m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$. Consequently, the optimal UC solution $\mathcal{X}^*$ can be obtained by solving the following optimization problem

$$\mathbf{OPT-UC}: \max_{\mathcal{X}} \quad \mathcal{R}^{\text{sum}}(\mathcal{X}, \mathbf{c}^*_{m,f,n}), \tag{4.39a}$$

$$\text{s.t.} \qquad (4.36f) - (4.36i), \tag{4.39b}$$

where $\mathbf{c}^*_{m,f,n}$ is the solution of problem $\mathcal{P}_{m,f,n}$. Note that, the optimal solution for the UC in problem (4.39) requires an exhaustive search, which is impractical due to its extremely high complexity. In order to overcome this issue, problem (4.39) will be formulated as a one-to-one three-sided matching problem [119]. Then, a low-complexity matching algorithm is developed to obtain a near-optimal solution, which will be discussed in Section 4.5.4.

4.5.3 Power Allocation for Each Coordinated C-NOMA Cluster

In this section, our objective is to solve problem $\mathcal{P}_{m,f,n}$ for a given UEs cluster (CCU_m^1, CEU_f,CCU_n^2). First, we define the feasibility conditions for problem $\mathcal{P}_{m,f,n}$. Then, we determine $\mathbf{c}^*_{m,f,n}$ within the feasibility region of problem $\mathcal{P}_{m,f,n}$.

4.5.3.1 Feasibility Conditions

The feasibility conditions of problem $\mathcal{P}_{m,f,n}$ define the conditions under which at least one feasible solution for this problem does exist. The optimization problem $\mathcal{P}_{m,f,n}$ in (4.38)

can be rewritten as

$$\mathcal{P}_{m,f,n}: \max_{\mathbf{c}_{m,f,n}} \ \mathcal{R}_{m,f,n}, \tag{4.40a}$$

$$\text{s.t.} \quad 0.5 \leq \alpha_{i,k_i,f} \leq 1, \qquad \forall i \in \mathcal{I}, \tag{4.40b}$$

$$0 \leq \beta_{i,k_i,f} \leq 1, \qquad \forall i \in \mathcal{I}, \tag{4.40c}$$

$$\alpha_{i,k_i,f} \leq \mathrm{UB}_{i,k_i,f}\left(\beta_{i,k_i,f}\right), \qquad \forall i \in \mathcal{I}, \tag{4.40d}$$

$$\alpha_{i,k_i,f} = \mathrm{OB}_{i,k_i,f}\left(\beta_{i,k_i,f}, Y_{i,k_i,f}\right), \qquad \forall i \in \mathcal{I}, \tag{4.40e}$$

$$\alpha_{i,k_i,f} \geq \mathrm{LB}_{i,k_i,f}\left(\beta_{i,k_i,f}\right), \qquad \forall i \in \mathcal{I}, \tag{4.40f}$$

$$Y_f \leq \sum\nolimits_{i=1}^{|\mathcal{I}|} Y_{i,k_i,f}, \tag{4.40g}$$

where $t = \exp\left(R_{\mathrm{th}}\right) - 1$, $Y_f = t(\gamma_{1,f} + \gamma_{2,f} + 1)$ and for all $i \in \mathcal{I}$, $Y_{i,k_i,f}$, $\mathrm{UB}_{i,k_i,f}$, $\mathrm{OB}_{i,k_i,f}$ and $\mathrm{LB}_{i,k_i,f}$ are expressed, respectively, as

$$\begin{cases} Y_{i,k_i,f} = \alpha_{i,k_i,f}\gamma_{i,f}(1+t) + \beta_{i,k_i,f}\gamma_{k_i,f}, \\ \mathrm{UB}_{i,k_i,f}\left(\beta_{i,k_i,f}\right) = \dfrac{-\gamma_{k_i,\mathrm{SI}} t \beta_{i,k_i,f} + \gamma_{i,k_i} - t}{\gamma_{i,k_i}}, \\ \mathrm{OB}_{i,k_i,f}\left(\beta_{i,k_i,f}, Y_{i,k_i,f}\right) = \dfrac{-\gamma_{i,k_i,f}\beta_{i,k_i,f} + Y_{i,k_i,f}}{\gamma_{i,f}(1+t)}, \\ \mathrm{LB}_{i,k_i,f}\left(\beta_{i,k_i,f}\right) = \dfrac{\gamma_{k_i,\mathrm{SI}} t \beta_{i,k_i,f} + t(1+\gamma_{i,k_i})}{\gamma_{i,k_i}(1+t)}. \end{cases} \tag{4.41}$$

Constraints (4.40d) and (4.40f) result from the facts that $R_{\mathrm{th}} \leq R_{i,k_i}$ and $R_{\mathrm{th}} \leq R_{i,k_i \to f}$, whereas constraints (4.40e) and (4.40g) result from the fact that $R_{\mathrm{th}} \leq R_f$. Based on this, the feasibility conditions of problem $\mathcal{P}_{m,f,n}$ can be obtained through the following theorem.

Theorem 4.1. *Problem $\mathcal{P}_{m,f,n}$ is feasible iff the following conditions hold.*

$$\textit{Condition 1: } t \leq -1 + \sqrt{1 + \min_{i \in \mathcal{I}} \gamma_{i,k_i}}, \tag{4.42a}$$

$$\textit{Condition 2: } Y_f \leq \sum_{i=1}^{|\mathcal{I}|} Y_{i,k_i,f}^{\max}, \tag{4.42b}$$

where for all $i \in \mathcal{I}$, $Y_{i,k_i,f}^{\max}$ is expressed as shown in (4.43)

$$Y_{i,k_i,f}^{\max} = \gamma_{i,f}(1+t)\left(\frac{1}{2}\left(1 + \mathrm{sgn}\left(\frac{\gamma_{i,f}}{\gamma_{i,f}(1+t)} - \frac{\gamma_{k_i\mathrm{SI}}}{\gamma_{i,k_i}t}\right)\right)\left(\frac{\gamma_{i,f}}{\gamma_{i,f}(1+t)} - \frac{\gamma_{k_i,\mathrm{SI}}}{\gamma_{i,k_i}t}\right)\beta_{i,k_i,f}^{\max} + \frac{\gamma_{i,k_i} - t}{\gamma_{i,k_i}}\right), \tag{4.43}$$

in which

$$\beta_{i,k_i,f}^{\max} = \min\left(1, \frac{\gamma_{i,k_i} - 2t - t^2}{\gamma_{k_i,\mathrm{SI}} t(2+t)}\right). \tag{4.44}$$

Proof. See Appendix B3. ■

4.5.3.2 Proposed Power Control Scheme

In this subsection, our objective is to solve the optimization problem $\mathcal{P}_{m,f,n}$ under the assumption that it is feasible. First, we start with the following corollary.

Corollary 1. *For all $i \in \mathcal{I}$, the range of $Y_{i,k_i,f}$ is given by*

$$Y_{i,k_i,f} \in \left[Y^{\min}_{i,k_i,f}, Y^{\max}_{i,k_i,f}\right], \tag{4.45}$$

where $Y^{\max}_{i,k_i,f}$ is expressed as shown in (4.43) *and*

$$Y^{\min}_{i,k_i,f} \max\left(\frac{\gamma_{i,f}\,(1+t)}{2}, \frac{t\gamma_{i,f}(1+\gamma_{i,k_i})}{\gamma_{i,k_i}}\right). \tag{4.46}$$

Proof. See Appendix B4. ■

Now, let us consider constraint (4.40e). For all $i \in \mathcal{I}$, since $\alpha_{i,k_i,f} = \mathrm{OB}_{i,k_i,f}\,(\beta_{i,k_i,f}, Y_{i,k_i,f})$, then we have $Y_{i,k_i,f} = \gamma_{i,f}(1+t)\alpha_{i,k_i,f} + \gamma_{i,k_i,f}\beta_{i,k_i,f}$. In addition, the variables $(Y_{1,m,f}, Y_{2,n,f})$ are linked through the inequality constraint in (4.40g). If we transform this inequality constraint into an equality constraint, then we obtain $Y_{2,n,f} = Y_f - Y_{1,m,f}$. Therefore, using the results of corollary 1, since $Y_{2,n,f} \in \left[Y^{\min}_{2,n,f}, Y^{\max}_{2,n,f}\right]$, then we obtain $Y_{1,m,f} \in \left[Y^{\min'}_{1,m,f}, Y^{\max'}_{1,m,f}\right]$, where $Y^{\min'}_{1,m,f} = \max\left(Y^{\min}_{1,m,f}, Y_f - Y^{\max}_{2,n,f}\right)$ and $Y^{\max'}_{1,m,f} = \min\left(Y^{\max}_{1,m,f}, Y_f - Y^{\min}_{2,n,f}\right)$. Our approach to solving problem $\mathcal{P}_{m,f,n}$ is detailed as follows. First, we solve problem $\mathcal{P}_{m,f,n}$ for fixed values of $(Y_{1,m,f}, Y_{2,n,f})$ and assuming that constraint (4.40g) is satisfied. Then, we obtain the optimal solutions of problem $\mathcal{P}_{m,f,n}$ through an optimal design of the parameters $(Y_{1,m,f}, Y_{2,n,f})$. Based on this, for fixed values of $(Y_{1,m,f}, Y_{2,n,f})$, and assuming that (4.40g) is satisfied, problem $\mathcal{P}_{m,f,n}$ is reduced to the following optimization problem.

$$\mathcal{P}'_{m,f,n} : \max_{\mathbf{c}_{m,f,n}} \ \mathcal{R}_{m,n,f}, \tag{4.47a}$$

$$\text{s.t.} \qquad (4.40b) - (4.40f). \tag{4.47b}$$

Let $\mathcal{C}\left(Y_{i,k_i,f}, i \in \mathcal{I}\right) \triangleq \bigtimes_{i=1}^{|\mathcal{I}|} \mathcal{C}_i\left(Y_{i,k_i,f}\right)$, where $\forall i \in \mathcal{I}$, the set $\mathcal{C}_i\left(Y_{i,k_i,f}\right) \subset \left[\frac{1}{2}, 1\right] \times [0,1]$ contains the intersection points $(\alpha_{i,k_i,f}, \beta_{i,k_i,f})$ between the bound $\mathrm{OB}_{i,k_i,f}$ and the boundaries of the feasibility region of problem $\mathcal{P}_{m,f,n}$ except the upper bound $\mathrm{UB}_{i,k_i,f}$. Based on this, and $\forall i \in \mathcal{I}$, the set $\mathcal{C}_i\left(Y_{i,k_i,f}\right)$ is characterized as shown in the following theorem.

Theorem 4.2. *For all $i \in \mathcal{I}$, the set $\mathcal{C}_i\left(Y_{i,k_i,f}\right)$ is expressed as follows. If $\frac{\gamma_{i,k_i}-t}{\gamma_{i,k_i}} \leq \frac{Y_{i,k_i,f}}{\gamma_{i,f}(1+t)}$, then $\mathcal{C}_i$ contains only one critical point that is expressed as*

$$(\alpha_{i,k_i,f}, \beta_{i,k_i,f})_1 = \left(f_{i,k_i,f}\left(\beta^c_{i,k_i,f}, Y_{i,k_i,f}\right), \beta^c_{i,k_i,f}\right), \tag{4.48}$$

and if $\frac{\gamma_{i,k_i}-t}{\gamma_{i,k_i}} \geq \frac{Y_{i,k_i,f}}{\gamma_{i,k_i}(1+t)}$, then $\mathcal{C}_i$ contains two critical points that are expressed, as

$$\begin{aligned}(\alpha_{i,k_i,f}, \beta_{i,k_i,f})_1 &= \left(f_{i,k_i,f}\left(\beta^c_{i,k_i,f}, Y_{i,k_i,f}\right), \beta^c_{i,k_i,f}\right),\\ (\alpha_{i,k_i,f}, \beta_{i,k_i,f})_2 &= \left(f_{i,k_i,f}\left(0, Y_{i,k_i,f}\right), 0\right),\end{aligned} \tag{4.49}$$

where the function $f_{i,k_i,f}(\cdot,\cdot)$ *is expressed as*

$$f_{i,k_i,f}(\beta,x)=\max\left(\frac{1}{2},-\frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}\beta+\frac{x}{\gamma_{i,f}(1+t)}\right), \tag{4.50}$$

and $\beta^{c}_{i,k_i,f}$ *is expressed as*

$$\beta^{c}_{i,k_i,f}=\min\left(\beta^{\max}_{i,k_i,f},\beta^{\text{int}}_{i,k_i,f}\right), \tag{4.51}$$

in which

$$\beta^{\text{int}}_{i,k_i,f}=\frac{\gamma_{i,k_i}Y_{i,k_i,f}-t\gamma_{i,f}(\gamma_{i,k_i}+1)}{\gamma_{k_i,\text{SI}}\gamma_{i,f}+\gamma_{i,k_i}\gamma_{i,k_i,f}}. \tag{4.52}$$

Proof. See Appendix B5. ■

The critical points are the possible candidates to be optimal solutions for problem $\mathcal{P}'_{m,f,n}$, which are functions of $(Y_{1,m,f},Y_{2,n,f})$. Consequently, the proposed solutions of problem $\mathcal{P}'_{m,f,n}$ are expressed as

$$\mathbf{c}^{*}_{m,f,n}\left(Y_{1,m,f},Y_{2,n,f}\right)=\underset{\mathbf{c}_{m,f,n}}{\text{argmax}}\ R_{m,f,n}\left(\mathbf{c}_{m,f,n},Y_{1,m,f},Y_{2,n,f}\right). \tag{4.53}$$

Based on the above, the original problem $\mathcal{P}_{m,f,n}$ is reduced to a linear search problem. Specifically, using the linear relation $Y_{2,n,f}=Y_f-Y_{1,m,f}$, the proposed solutions of problem $\mathcal{P}_{m,f,n}$ are fully characterized by obtaining the optimal value of the variable $Y_{1,m,f}$. Therefore, let us define the L-point discrete uniform range for $Y_{1,m,f}$ that is given by $\mathcal{Y}_{1,m,f}=\left\{\left(Y^{\max'}_{1,m,f}-Y^{\min'}_{1,m,f}\right)\frac{l-1}{L-1}+Y^{\min'}_{1,m,f},l\in[\![1,L]\!]\right\}$, where $L\in\mathbb{N}\setminus\{0\}$. Therefore, for each value of $Y_{1,m,f}\in\mathcal{Y}_{1,m,f}$, we calculate $Y_{2,n,f}=Y_f-Y_{1,m,f}$. Then, we obtain the value of $\mathbf{c}^{*}_{m,f,n}(Y_{1,m,f},Y_f-Y_{1,m,f})$ from equation (4.53) and the corresponding optimal cluster rate $R^{*}_{m,f,n}(Y_{1,m,f},Y_f-Y_{1,m,f})$. Consequently, we select the power control vector $\mathbf{c}^{*}_{m,f,n}$ that achieves the highest cluster rate $R^{*}_{m,f,n}(Y_{1,m,f},Y_f-Y_{1,m,f})$ over all values of $Y_{1,m,f}\in\mathcal{Y}_{1,m,f}$, i.e., $\mathbf{c}^{*}_{m,f,n}=\underset{Y_{1,m,f}\in\mathcal{Y}_{1,m,f}}{\text{argmax}}\ R^{*}_{m,f,n}(Y_{1,m,f},Y_f-Y_{1,m,f})$. Finally, based on the above analysis, the algorithm of the proposed power control scheme is presented in **Algorithm 2**.

4.5.4 UEs Clustering: Three-Sided Matching Game Formulation

To efficiently solve the UEs clustering problem, we reformulate the optimization problem in (4.39) as a one-to-one three-sided matching game. To better describe the matching game between the two CCUs and the CEU, we first introduce some necessary definitions and notations for the matching game.

Algorithm 2: Proposed power control scheme

Initialization: $\mathbf{c}^*_{m,f,n} = \mathbf{0}_{4\times 1}$, $R^*_{m,f,n} = 10^5$, $\mathbf{c}_{m,f,n} = \mathbf{0}_{4\times L}$ and $\mathbf{r}_{m,n,f} = \mathbf{0}_L$;
for l *from* 1 *to* L **do**

i) Compute $Y_{1,m,f} = \left(Y^{\max'}_{1,m,f} - Y^{\min'}_{1,m,f}\right)\frac{l-1}{L-1} + Y^{\min'}_{1,m,f}$ and then $\mathcal{C}_1\left(Y_{1,m,f}\right)$ as shown in **Theorem 2**;

ii) Compute $Y_{2,n,f} = Y_f - Y_{1,m,f}$ and then $\mathcal{C}_2\left(Y_{2,n,f}\right)$ as shown in **Theorem 2**;

iii) Compute $\mathcal{C}\left(Y_{1,m,f}, Y_{2,n,f}\right) = \mathcal{C}_1\left(Y_{1,m,f}\right) \times \mathcal{C}_2\left(Y_{2,n,f}\right)$;

iv) $\mathbf{c}_{m,f,n}\left(:,l\right) = \underset{c_{m,n,f}\in\mathcal{C}\left(Y_{1,m,f},Y_{2,n,f}\right)}{\operatorname{argmax}} R_{m,f,n}\left(\mathbf{c}_{m,f,n}\right)$ and then $\mathbf{r}_{m,n,f}\left(l\right) = R_{m,f,n}\left(\mathbf{c}_{m,f,n}\left(:,l\right)\right)$;

v) **if** $\boldsymbol{R}_{m,n,f}\left(l\right) < R^*_{m,f,n}$ **then**
$R^*_{m,f,n} = \mathbf{r}_{m,n,f}\left(l\right)$;
$\mathbf{c}^*_{m,f,n} = \mathbf{c}_{m,f,n}\left(:,l\right)$;
end if

end for

4.5.4.1 Preliminaries

Definition 4.1. *When a CCU^1_m, and another CCU^2_n, jointly serve a CEU_f, along with the two BSs, we say that CCU^1_m, CEU_f, and CCU^2_n are matched with each other, and they form a matching triple. This matching triple is denoted by $\Delta \triangleq (CCU^1_m, CEU_f, CCU^2_n)$. A player is defined as one UE in the cellular system, i.e., each UE from CCU^1_m, CEU_f and CCU^2_n represents a player. Additionally, any two of these three players compose a player pair, i.e. (CCU^1_m, CEU_f), (CCU^1_m, CCU^2_n) and (CEU_f, CCU^2_n).*

We denote the preference utility of any triple (CCU^1_m,CEU_f,CCU^2_n) by the sum rate of the players inside the potential matching triple, i.e., $U_{m,f,n} \triangleq \mathcal{R}_{1,m} + \mathcal{R}_f + \mathcal{R}_{2,n}$, i.e., $U_{m,f,n} \triangleq \mathcal{R}_{m,f,n}$. The intuition behind selecting such preference utility comes from the objective of $\mathbf{OPT-UC}$, where each matching triple Δ constructs a coordinated C-NOMA cluster such that its achievable sum rate is maximized. Each triple has unique channel gains with the BSs and among its players (D2D channel gains from the CCUs to the CEU) and therefore results in a distinctive triple utility. Thus, when a player is matched with different player pairs and composes different triples, this player should only select one triple to be a member of it and, in the meantime, each agent in that triple should accept that player (agent and player are used interchangeably throughout the paper). To better describe the decision of each agent and the competitive behavior among the players, we define

a *preference utility* for each player over the player pairs (CCU_m^1,CEU_f), (CCU_m^1,CCU_n^2), and (CEU_f,CCU_n^2). For example, if CEU_f prefers player pair (CCU_m^1,CCU_n^2) over the player pair (CCU_a^1,CCU_c^2), the triple utility (CCU_m^1,CEU_f,CCU_n^2) should yield higher utility than the triple (CCU_a^1,CEU_f,CCU_c^2) does, which can be presented as (CCU_m^1,CCU_n^2) $\succ_{\text{CEU}_f}$ (CCU_a^1,CCU_c^2) iff $U_{m,f,n} > U_{a,f,c}$. By evaluating the utilities of the different triples, each player can build up a descending-ordered preference profile. Using this preference profile, each player can identify its preference over all player pairs from the other two player sets. We denote the set of the preference profiles of all players as $\boldsymbol{S} \triangleq \{\boldsymbol{S}(\text{CCU}_1^1), \ldots, \boldsymbol{S}(\text{CCU}_M^1), \boldsymbol{S}(\text{CEU}_1), \ldots, \boldsymbol{S}(\text{CEU}_F), \boldsymbol{S}(\text{CCU}_1^2), \ldots, \boldsymbol{S}(\text{CCU}_N^2)\}$, where, for all $m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, $\boldsymbol{S}(\text{CCU}_m^1), \boldsymbol{S}(\text{CEU}_f)$, and $\boldsymbol{S}(\text{CCU}_n^2)$ are the preference profiles of CCU_m^1, CEU_f, and CCU_n^2, which have at most FN (CEU_f,CCU_n^2) pairs, MN (CCU_m^1,CCU_n^2) pairs, and MF (CCU_m^1, CEU_f) pairs, respectively. Next, and based on the above definitions and notations, we will reformulate the UEs clustering optimization problem $\mathbf{OPT-UC}$ as a three-sided matching problem.

Definition 4.2. *Given a set $\mathcal{M}$ of CCU^1, a set $\mathcal{F}$ of CEUs, and a set $\mathcal{N}$ of CCU^2, a one-to-one three-sided matching Φ is a matching function defined as $\mathcal{M} \longmapsto \mathcal{F} \times \mathcal{N}, \mathcal{F} \longmapsto \mathcal{M} \times \mathcal{N}$, and $\mathcal{N} \longmapsto \mathcal{M} \times \mathcal{F}$. Therefore, under the matching Φ, we have at most $Q = \min\{M, F, N\}$ disjoint triples, which are denoted by $\{\Delta_1, \Delta_2, \ldots \Delta_Q\}$ such that:*

1. $\Phi(CCU_m^1) = (CEU_f, CCU_n^2) \iff \Phi(CEU_f) = (CCU_m^1, CCU_n^2) \iff \Phi(CCU_n^2) = (CCU_m^1, CEU_f)$,
2. $\Phi(CCU_m^1) = \varnothing$, *if* CCU_m^1 *is not in any matching triple,* $\Phi(CEU_f) = \varnothing$, *if* CEU_f *is not in any matching triple, and* $\Phi(CCU_n^2) = \varnothing$, *if* CCU_n^2 *is not in any matching triple,*
3. $\Delta_i \cap \Delta_j = \varnothing$, $\forall i, j \in \{1, 2, \ldots Q\}$,

where condition 1) guarantees that CCU_m^1, CEU_f, and CCU_n^2 are matched with each other and construct a matching triple. Since it is not necessary that the sizes of the three sets, $\mathcal{M}, \mathcal{F}$, and $\mathcal{N}$ are the same, it may be that some agents will not be matched, which is represented by conditions 2). Condition 3) indicates that each agent must belong to at most one matching triple.

Definition 4.3. *The social welfare of the whole system can be defined as the total utility of all the matching triples, which can be expressed as*

$$U_\Phi = \sum_{m \in \mathcal{M}} \sum_{f \in \mathcal{F}} \sum_{n \in \mathcal{N}} x_{m,f,n} U_{m,f,n}. \tag{4.54}$$

Corresponding to the optimization problem $\mathbf{OPT-UC}$, the one-to-one three-sided CCU^1-CEU-CCU^2 matching problem can be described as follows. We aim to obtain a one-to-one three-sided matching Φ from a set of CCU^{1}s, a set of CEUs, and a set of CCU^{2}s that can maximize the *social welfare* U_Φ in (4.54) subject to constraints (4.36f)-(4.36i) with decision parameters $\{x_{m,f,n}, \forall m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$. In order to solve this matching game, we explain the proposed matching algorithm in the following subsection.

4.5.4.2 Proposed Matching Algorithm

In the considered three-sided CCU^1-CEU-CCU^2 matching game, each CEU$_f$ is an agent which tries to be matched with its most preferable $(\text{CCU}_m^1, \text{CCU}_n^2)$ pair and the CCU^{1}s and CCU^{2}s tend to choose the CEUs that can maximize the total utility (4.54). First, the CEUs propose themselves to the most favourable $(\text{CCU}^1, \text{CCU}^2)$ pairs, i.e., for all $m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, each CEU$_f$ makes an offer to its favourite $(\text{CCU}_m^1, \text{CCU}_n^2)$ pair. Then, when CCU^{1}s and CCU^{2}s receive the offers from the CEUs, they have the right to decide whether to reject or accept the offers. In the proposing stage of the CEUs, it is likely that one CCU^1 or one CCU^2 receives more than one offer from different CEUs, which results in a conflict between the offers. To discuss how the CCU^1 and CCU^2 choose among various offers, the concept of blocking triples is introduced, which can be defined as follows.

Definition 4.4. *Given a matching Φ and a player triple Δ_B having at least one matched player and one unmatched player under Φ, we have:*

1. *If only one player of Δ_B, denoted by q_1, is matched under Φ, then Δ_B is a blocking triple under Φ when $U_{\Delta_B} > U_{q_1,\Phi(q_1)}$.*
2. *If two players of Δ_B, denoted by q_1 and q_2, are matched under Φ, then we have two cases*
   (a) *If $q_1 \in \Phi(q_2)$, then Δ_B is a blocking triple under Φ when $U_{\Delta_B} > U_{q_1,\Phi(q_1)}$;*
   (b) *If $q_1 \notin \Phi(q_2)$, then Δ_B is a blocking triple under Φ when $U_{\Delta_B} > U_{q_1,\Phi(q_1)} + U_{q_2,\Phi(q_2)}$.*

Following up on ***Definition 4.4***, denote q_1, q_2 and q_3 as the three players of triple Δ_B, where at least one among q_1 and q_2 is matched under Φ and q_3 is an unmatched player. Item 1) indicates the case when only one agent is matched under Φ (say q_1 is the matched

agent). In this case, if agent q_1 can achieve a higher utility by constructing a matching triple with agents q_2 and q_3 instead of its original partners under Φ, then q_1, q_2, and q_3 is a blocking triple together [119]. In general, a player in a blocking triple Δ_B can switch its current matching triple under Φ as long as it can achieve a better utility. As a result, q_1 must reject its current partners and, therefore, one original matching triple under Φ will be broken up. On the other hand, if (q_1, q_2) are partners in other matching triples under Φ, we have two sub-cases: 2a) when q_1 and q_2 belong to the same matching triple under Φ or 2b) when they have been matched to two different matching triples under Φ. Case 2a) is similar to case 1) because only one matching triple under Φ must be broken up. Meanwhile, in case 2b), the two agents are matched with two different matching triples, which leads to breaking up two matching triples under Φ. However, changing the partners of q_1 and q_2 may affect the number of matching triples, which results in dissatisfaction with the other players and reduces the *social welfare* of the network. Consequently, these players should change their partners and break up two matching triples only when a remarkable enhancement can be obtained by these switching. Therefore, Δ_B is considered as a blocking pair, if the utility of Δ_B is higher than the profit that can be achieved from the sum of the triples utilities of q_1 and q_2 under Φ, which should be broken-up [119].

The details of the proposed $\underline{\mathbf{C}}\mathrm{CU}^1$-C$\underline{\mathbf{E}}$U-$\underline{\mathbf{C}}\mathrm{CU}^2$ $\underline{\mathbf{M}}$atching (CECM) algorithm are presented in **Algorithm 3**, which consists of three main stages: *Stage I:* PA phase, *Stage II:* Initialization phase and *Stage III:* Matching phase. In *Stage I*, the PA for each coordinated C-NOMA cluster (CCU^1_m,CEU_f,CCU^2_n) is evaluated according to **Algorithm 3**, and the corresponding achievable sum rate of each coordinated C-NOMA cluster is computed. In *Stage II*, the preference profile of each CEU_f is obtained by sorting the sum rate of the coordinated C-NOMA cluster for every potential triple $(\mathrm{CCU}^1_m, \mathrm{CEU}_f, \mathrm{CCU}^2_n)$. In addition, three unmatched players sets, $\mathcal{V}, \mathcal{W}$, and $\mathcal{Z}$ are built to identify whether or not each $\mathrm{CCU}^1, \mathrm{CEU}$, and CCU^2, respectively, are matched.

In *Stage III*, the matching process requires multiple rounds. In each round r, each unmatched CEU_f proposes itself to the most preferred ($\mathrm{CCU}^1_{m^*}$, $\mathrm{CCU}^2_{n^*}$) pair, which is the first ranked in CEU_f's preference profile, which corresponds to steps 5-6 in **Algorithm 3**. It should be noted that the pair ($\mathrm{CCU}^1_{m^*}$, $\mathrm{CCU}^2_{n^*}$) has never rejected CEU_f request in the previous rounds. If the two agents $\mathrm{CCU}^1_{m^*}$ and $\mathrm{CCU}^2_{n^*}$ have never received any requests in the previous rounds and only receive a request from CEU_f in round r, the

Algorithm 3: Proposed CECM Algorithm

Input: Set of CCU^1s $\mathcal{N}$; Set of CEUs $\mathcal{F}$; Set of CCU^2s $\mathcal{M}$

Phase I: Power Allocation

Phase II: Initialization of Matching Algorithm

Calculate the utility $\mathcal{R}_{m,f,n}, \forall m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}$;

Form $M \times N$ $(\text{CCU}_m^1, \text{CCU}_n^2)$ pairs from $\mathcal{M} \times \mathcal{N}$;

Construct the preference profiles of all CEUs in $\mathcal{F}$ overall $(\text{CCU}_m^1, \text{CCU}_n^2)$ pairs: $\{\mathcal{S}(\text{CEU}_1), \ldots, \mathcal{S}(\text{CEU}_F)\}$;

Initialize the sets of unmatched CCU^1s $\mathcal{V} \triangleq \mathcal{M}$, unmatched CEUs $\mathcal{W} \triangleq \mathcal{F}$, and unmatched CCU^2s $\mathcal{Z} \triangleq \mathcal{N}$;

Phase III: Matching

while $\mathcal{W}$ *is not empty and the preference profile* $\mathcal{S}(\text{CEU}_f)$ *of each* CEU_f *is not empty* **do**

- CEU_f proposes itself to its most favourable $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$, the first pair in its preference profile $\mathcal{S}(\text{CEU}_f)$;
- **if** $\text{CCU}_{m^*}^1 \in \mathcal{V}$ *and* $\text{CCU}_{n^*}^2 \in \mathcal{Z}$ **then**
  - Match CEU_f with $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ and record this triple;
  - Remove $\text{CCU}_{m^*}^1$, EU_f, and $\text{CCU}_{n^*}^2$ from $\mathcal{V}, \mathcal{W}$, and $\mathcal{Z}$, respectively;
- **end if**
- **else if** $(\text{CCU}_{m^*}^1, \text{CEU}_f, \text{CCU}_{n^*}^2)$ *is a blocking triple* **then**
  - Remove CEU_f from $\mathcal{W}$
  - Each member of $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ kicks out its current partners;
  - Match CEU_f with $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$, and record this triple matching;
  - Add the rejected players in the corresponding unmatched player sets;
- **end if**
- **else**
  - CEU_f is rejected by $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$;
- **end if**
- Remove $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ from $\mathcal{S}(\text{CEU}_f)$;

end while

pair $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ will accept temporarily the request from CEU_f. Then, $(\text{CCU}_{m^*}^1, \text{CEU}_f, \text{CCU}_{n^*}^2)$ will be matched and form a matching triple, which corresponds to steps 7-9 in **Algorithm 3**. However, if either $\text{CCU}_{m^*}^1$ and/or $\text{CCU}_{n^*}^2$ have/has already been matched in earlier rounds or have/has received requests from other CEUs in the current round, $\text{CCU}_{m^*}^1$ and/or $\text{CCU}_{n^*}^2$ should check whether $(\text{CCU}_{m^*}^1, EU_f, \text{CCU}_{n^*}^2)$ can form a blocking triple under the current matching, which corresponds to step 10 in **Algorithm 3**. If it can, then the pair $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ will accept the request of CEU_f and kicks out the previously matched CEU, which corresponds to steps 11-14 in **Algorithm 3**. Otherwise $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ rejects CEU_f and keeps the previously matched partners, which corresponds to steps 15-16 in **Algorithm 3**. It should be noted that the current

accepted CEU_f by $(\text{CCU}^1_{m^*}, \text{CCU}^2_{n^*})$ pair might be kicked out in the next coming rounds because this pair still have the right to accept another CEU if it can offer a better request in terms of the total utility than the current matching. The rounds stop when no CEU_f proposes itself to the $(\text{CCU}^1_m, \text{CCU}^2_n)$ pairs, which means that either a CEU_f is matched to an $(\text{CCU}^1_m, \text{CCU}^2_n)$ pair or it has been rejected by all $(\text{CCU}^1_m, \text{CCU}^2_n)$ pairs. Moreover, the output of Φ in **Algorithm 3** can be mapped to a feasible UEs clustering set $\mathcal{X}$ of the original problem $\mathbf{OPT-UC}$, where if CCU^1_m, EU_f, and CCU^2_n are matched, we set $x_{m,f,n} = 1$, and $x_{m,f,n} = 0$ otherwise. Note that the UEs that are not members of any coordinated C-NOMA cluster can be served through OMA technique. Based on that, after the coordinated C-NOMA clusters are constructed, one possible solution to serve those remaining UEs is to solve another optimization problem to maximize the sum-rate while considering the OMA technique. In this paper, we only focus on analyzing the performance of the UEs which are members of the coordinated C-NOMA clusters. Thus, the problem is formulated to guarantee the required QoS for admitted UEs.

4.5.4.3 Stability And Convergence Analysis of The Proposed CECM Algorithm

Definition 4.5. *A matching Φ is said to be stable if there are no blocking triples other than the matching triples under Φ.*

Lemma 4.1. *If **Algorithm 3** converges to Φ^*, then Φ^* is a stable matching.*

Proof. See Appendix B6. ■

Theorem 4.3. *The CECM algorithm converges to a stable matching after a finite number of rounds.*

Proof. See Appendix B7. ■

4.5.4.4 Computational Complexity Analysis

In this section, we evaluate the computational complexity of the CoMP-assisted FD C-NOMA, through both the optimal and the proposed solutions of problem $\mathcal{P}$, and of the conventional CoMP NOMA scheme. Note that, since we have obtained the optimal PA coefficients at the BSs and the optimal transmit power at the CCUs, the main reason of the sub-optimality of the proposed approach is the solution of the linear assignment problem of the UC. In order to find the optimal solution of the optimization problem $\mathcal{P}$

in (4.38), we need to optimally solve the UC policy through an exhaustive search, which refers to as optimal UC (OUC). In order to evaluate the computational complexity of these three schemes, the computational complexity of the PA optimization for all the potential coordinated C-NOMA clusters, the UC optimization using the CECM algorithm, and the OUC scheme need to be evaluated, which is the focus of the following paragraphs.

1) ***Complexity Analysis of the PA Solution****:* **Algorithm 1** obtains the optimal solution for one coordinated C-NOMA cluster with an efficient computational complexity of $\mathcal{O}(L)$. Thus, the computational complexity of obtaining the optimal sum rate for all the possible coordinated C-NOMA cluster configurations is $\mathcal{O}(LMFN)$.

2) ***Complexity Analysis of the OUC Scheme****:* In the OUC scheme, all the possible combinations of (CCU_m^1,CEU_f,CCU_n^2), for all $m \in \mathcal{M}$, $f \in \mathcal{F}$, $n \in \mathcal{N}$, are tested through an exhaustive search and then the optimal solution is selected. Since the number of CCU^1s, CEUs, and CCU^2s are not necessarily equal, we assume that $Q \triangleq F$. This indicates that each of the F CEUs can be matched with one CCU^1 and one CCU^2. By choosing F CCU^1s from M and F CCU^2 from N, there are $\binom{M}{F}\binom{N}{F}$ patterns in total. For every selected pattern, we have $(F!)^2$ matching cases between the selected players. As a result, the total number of triple-matching combinations is $M!N!/[(M-F)!(N-F)!]$. In order to find the best matching triples that can maximize the network sum-rate, we need to perform an exhaustive search over all possible matching combinations. Thus, the computational complexity of the exhaustive search is $\mathcal{O}((MN)^F)$, which is an exponential complexity. The same steps can be applied if $Q \triangleq N$ or $Q \triangleq M$.

3) ***Complexity Analysis of the Proposed CECM Algorithm****:* to quantify the complexity of the proposed CECM algorithm, it is necessary to analyze the computational complexity of forming the preference profile for each CEU (Phase II in **Algorithm 3**) and the running complexity of the proposed CECM algorithm (Phase III in **Algorithm 3**). First, each CEU needs a sorting process to construct its preference profile, which has a complexity of $\mathcal{O}(MN\log(MN))$. Consequently, the overall computational complexity of Phase II is $\mathcal{O}(FMN\log(MN))$ [120]. Meanwhile, in Phase III, under the worst case, it can be observed that the computational complexity of the proposed CECM algorithm is linear with respect to the size of the preference profiles, and it is given by $\mathcal{O}(MFN)$. Therefore, the computational complexity of the proposed CECM algorithm is $\mathcal{O}(FMN\log(MN))$, which is a polynomial complexity, and therefore, more efficient than the OUC scheme.

As a result, the computational complexity of the optimal and proposed solutions of the CoMP-assisted FD C-NOMA are $\mathcal{O}((MN)^F + LMNF)$ and $\mathcal{O}(FMN(L+\log(MN)))$, respectively, which is exactly the same as the one of the CoMP-assisted HD C-NOMA scheme that is considered as a benchmark in the simulation results section. Finally, it is worth mentioning that the adopted solution approach in the proposed CoMP-assisted FD C-NOMA scheme is applied for solving the CoMP NOMA as explained in section VI. However, for the PA optimization, the line search is not needed after doing the required substitutions. As a result, the computational complexity for the CoMP NOMA scheme is $\mathcal{O}(FMN(\log(MN)))$.

4.5.5 Results and Discussion

In this section, the efficacy of the proposed model and the solution approach is validated through the following three points:

- The validation of the optimal low-complexity solution for the PA problem.
- The assessment of the proposed UEs clustering, i.e. the CECM algorithm, in comparison with the OUC scheme.
- The evaluation of the performance of the proposed CoMP-assisted FD C-NOMA compared to two benchmark schemes, which are CoMP-assisted HD C-NOMA and CoMP NOMA, under various system parameters by varying the power budget at the CUs, the rate threshold at the UEs, and the power allocated for each C-NOMA cluster.

The baseline schemes can be explained as follows.

- ***CoMP-assisted HD C-NOMA***: In this scheme, the CCUs cooperate with the CEUs in HD DF relaying mode. The UEs clustering is performed by applying the proposed CECM algorithm. In addition, a computational-efficient solution for the PA sub-problem is obtained following the same approach used in the CoMP-assisted FD C-NOMA.
- ***CoMP NOMA*** [9, 100]: It is important to remind here that the conventional CoMP NOMA scheme is the same as the proposed CoMP FD C-NOMA scheme, but without cooperation between the CCUs and the CEUs. In the CoMP NOMA

Table 4.4: Simulation Parameters

Parameter	Symbol	Value
Distance between BS_1 and CCU_m^1 and between BS_2 and CCU_n^2	$d_{1m} = d_{2n}$	$[0.2, 0.4]$
Distance between BS_1 and CEU_f	d_{1f}	$[0.8, 1]$
Distance between BS_2 and CEU_f	d_{2f}	$(1.8 - d_{1f})$
Distance between CCU_m^1 and UE_f	d_{mf}	$(d_{1f} - d_{1m})$
Distance between CCU_n^2 and CEU_f	d_{nf}	$(d_{2f} - d_{2n})$
Pathloss exponent	α	4
Noise variance	σ^2	1
Rate threshold	γ_{th}^f	1 bits/sec/Hz
Number of CCU^1s and Number of CCU^2s	M, N	6
Number of CEUs	F	4

scheme, the UEs clustering is performed by applying the proposed CECM algorithm. In addition, closed-form expressions for the PA coefficients adopted by the BSs to the CCUs and CEUs are obtained by following the same proposed procedures for the proposed power control considering $\Omega_{\text{SI}} = 0$ and $P_{\max} = 0$.

4.5.5.1 Simulation Settings

Considering a homogeneous cellular network, we assume that the power allocated by the BSs for each C-NOMA cluster is equal in which $P_1 = P_2 = P_b$ [10, 109, 113]. Apart from SI channel, both the small-scale fading and the large-scale fading are considered for each wireless communication link. The large-scale fading, i.e., the distance-dependent path-loss, is modeled as $PL(d_{tr}) = d_{tr}^{-\eta}$, where d_{tr} is the distance between node t and node r and η is the path-loss exponent. Meanwhile, the small-scale fading, which is denoted as g_{tr}, is modeled as Rayleigh fading with zero mean and unit variance [10, 33, 109]. Consequently, the corresponding channel coefficients can be expressed as $h_{tr} = g_{tr}\sqrt{PL(d_{tr})}$. Finally, the SI channel coefficient follows a complex symmetric Gaussian random variable with a zero mean and Ω_{SI} standard deviation, i.e., $\mathcal{CN}(0, \Omega_{\text{SI}}^2)$ [10, 33, 109]. Moreover, we assume that the cell radius is normalized to unity [33]. The distance from each BS (BS_1 and BS_2) to its CCUs, i.e., $d_{1,m}$ and $d_{2,n}$, is randomly selected from 0.2 to 0.4, while the distance from BS_1 to the CEUs is randomly chosen from 0.8 to 1. The main simulation parameters are shown in Table 4.4 and are similar to the ones adopted in the literature [10, 33]. Unless otherwise stated, we assume that the number of CCUs at each cell is equal to 6, i.e., $M = N = 6$. Meanwhile, the number of CEUs is equal to 4, i.e., $F = 4$. Note that, Monte Carlo simulations are employed over 10^6 independent channel and distance

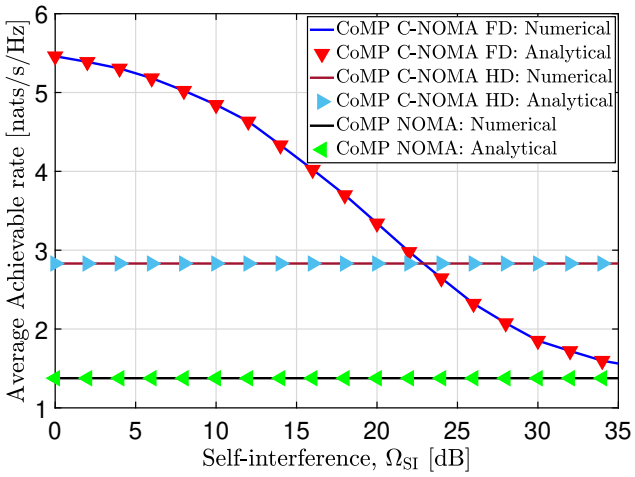


Figure 4.6: Analytical and numerical average sum-rate versus SI channel gain, i.e., Ω_{SI} ($P_b = 27$ dBm and $P_{\max} = 25$ dBm).

realizations. Finally, the adopted carrier frequency is 2.5 GHz [9].

4.5.5.2 Validation of The Optimality of The Proposed PA Scheme

Fig. 3 depicts the numerical and analytical average achievable sum rate for one coordinated C-NOMA cluster for the proposed CoMP-assisted FD C-NOMA scheme and the two benchmark schemes versus the average SI channel gain Ω_{SI}. The analytical results are obtained following the procedures in **Algorithm 2**, whereas the numerical results are obtained by solving problem $\mathcal{P}_{m,f,n}$ using an off-the-shelf optimization solver.[5] It can be seen from Fig. 4.6 that the analytical results match perfectly the numerical results, which validate the optimality of the PA solution obtained by **Algorithm 1**. For most of the SI values, Fig. 4.6 shows that the proposed scheme remarkably outperforms the benchmarked schemes. This is because, compared to the CoMP NOMA scheme, the CCUs are encouraged to assist the CEUs, which reflects on reducing the power allocated from the BSs to these CEUs and the BSs allocate most of the power to the CCUs. Meanwhile, compared to CoMP-assisted HD C-NOMA, the proposed scheme makes benefits from the one time-slot required for transmission. This results in a significant gain compared to the other benchmarked schemes. Furthermore, Fig. 4.6 shows that when the SI channel gain increases, the average achievable cluster sum rate decreases until nearly reaching the same performance as the CoMP NOMA scheme. This is because when SI increases, each CCU prefers to transmit with a low power value to avoid harming itself and thus maintain a

[5]The adopted solver is fmincon, which is a predefined MATLAB solver [107].

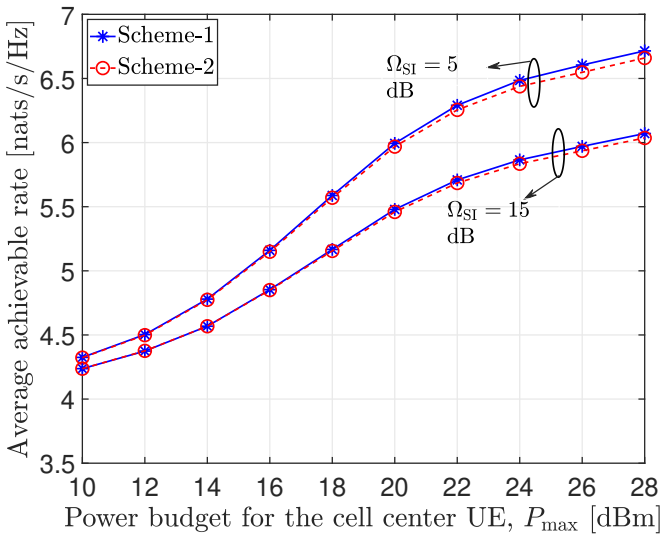


Figure 4.7: Average achievable sum rate per cluster versus the power budget at the CCUs $P_{\max}$, where $P_b = 30$ dBm.

better achievable rate. Thus, more power should be allocated from the BS to the CEUs to achieve its required QoS. Consequently, the allocated power from BS to CCUs decreases, and a lower average cluster rate performance is obtained. One can see that at a certain value of the SI, the performance of the CoMP-assisted HD C-NOMA achieves a better performance than the proposed scheme. This is due to the SI that limits the transmit power at CCUs in the FD mode. In addition, at very high SI values, the CCUs do not contribute to the transmission to the CEUs, and hence the proposed CoMP C-NOMA achieves the same performance as the CoMP NOMA scheme.

4.5.5.3 Validation of The Ignorance of Inter-CCUs Interference

In order to validate this assumption, we compare the performance of the proposed model, i.e., without considering the interference between the CCUs belonging to different cells, with the case when this interference is considered. In doing so, we have considered one coordinated C-NOMA cluster and we have optimally solved problem $\mathcal{P}_{m,f,n}$ for the proposed model without/with considering the interference between the CCUs belonging to different cells. For simplicity, let us denote the case of neglecting the interference between the two CCUs as "*Scheme-1*" and the case of considering the interference between the two CCUs as "*Scheme-2*". For *Scheme-1*, we solve problem $\mathcal{P}_{m,f,n}$ following the procedures in **Algorithm 2**, whereas for *Scheme-2*, we first derive the feasibility conditions. Then, we use an off-the-shelf optimization solver. Specifically, the adopted solver is fmincon, which

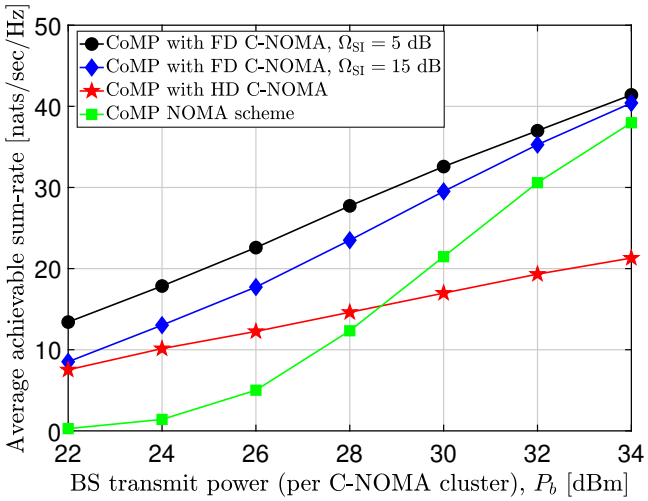


Figure 4.8: Average network sum-rate versus P_b with different SI channel gains ($R_{\text{th}} = 1$ nats/sec/Hz and $P_{\max} = 25$ dBm).

is a predefined MATLAB solver. Moreover, 100 different initial points were generated so as to guarantee convergence to the optimal solution.

Fig. 4.7 depicts the effect of increasing the transmit power budget $P_{\max}$ at the CCUs on the performance of *Scheme-1* and *Scheme-2* for different values of the SI channel gains. It can be seen that for most of the values of $P_{\max}$, *Scheme-1* and *Scheme-2* have the same performance. In addition, by considering the worst-case scenario when the SI channel gain is low and the power budget of the CCUs is high, which encourages the CCUs to transmit with high power, the impact of interference between the CCUs belonging to different cells is less than 1% on the performance of the coordinated C-NOMA cluster. For instance, when $\Omega_{\text{SI}} = 5$ dB and $P_{\max} = 28$ dBm, *Scheme-2* achieves 99.15% of the performance of *Scheme-1*, which validates the assumption of neglecting the effect of the interference between the two cell-center users.

4.5.5.4 Network Sum-Rate Performance

Fig. 4.8 presents the average network sum-rate achieved by the three schemes versus the power budget per C-NOMA cluster at the BS P_b for different values of SI channel gains. It can be observed that for most of the values of P_b, the proposed scheme outperforms the benchmarked schemes. In order to maximize the sum-rate of the network, the performance of the CCUs should be maximized. This can be approximately achieved by assigning the maximum allowable power control coefficients to the CCUs while considering the SIC constraints. Consequently, the BSs will allocate the minimum power control coefficients

to the CEUs in order to guarantee a successful SIC process at the CCUs. Then, the CCUs will be able to assist the transmission from the BSs to the CEUs in order to achieve their QoS. Thus, with low transmit power from the BSs, the CCUs should assist the CEUs to achieve their required QoS, which explains the performance gap between the CoMP C-NOMA in FD and HD relaying mode over the CoMP NOMA. For instance, when $P_b = 26$ dBm, the proposed scheme can achieve about 350% and 252% gain over the CoMP NOMA scheme for $\Omega_{\text{SI}} = 5$ dB and 15 dB, respectively; and the CoMP with HD C-NOMA achieves around 150% gain over CoMP NOMA. On the other hand, when the transmit power from the BSs increases, the gap between the proposed scheme and the CoMP NOMA scheme decreases. This is because with a high transmit power from the BSs, the power coefficients allocated by the BSs to the CEUs are able to guarantee the SIC constraint at the CCUs and are approximately enough to achieve the rate constraints at the CEUs. Thus, the required contributions from the CCUs will be reduced, and hence the performance of the CoMP with the FD C-NOMA scheme will converge to the one of the CoMP NOMA scheme.

As a conclusion, the transmit power from the BSs play an important role to determine whether the CCUs are required to assist the transmission from the BSs to the CEUs in the proposed scheme. However, for the CoMP-assisted HD C-NOMA, the gain resulting in the BS transmit power on the CCUs performance is only leveraged in one time-slot. Hence, when the power per C-NOMA cluster is sufficiently large compared to the power budget at the CCUs, the CoMP NOMA scheme achieves better performance than the CoMP with HD C-NOMA.

Fig. 4.9 compares the performance of the proposed CoMP-assisted FD C-NOMA scheme with the two benchmarked schemes when varying the required data rate threshold for different values of the SI. It can be seen that the proposed scheme has a significant gain compared to the CoMP NOMA scheme, especially at high data rate threshold values. This observation is explained as follows. Increasing the data rate threshold forces the BSs to allocate more power to the CEUs in order to guarantee the SIC process at the CCUs, which results in decreasing the achievable rates of that CCUs and, consequently, the network sum-rate. Before we proceed in further detail, we have to explain first how the average network sum-rate is calculated. For each coordinated C-NOMA/NOMA cluster, we first check the feasibility conditions defined in ***Theorem 4.2***. If these conditions

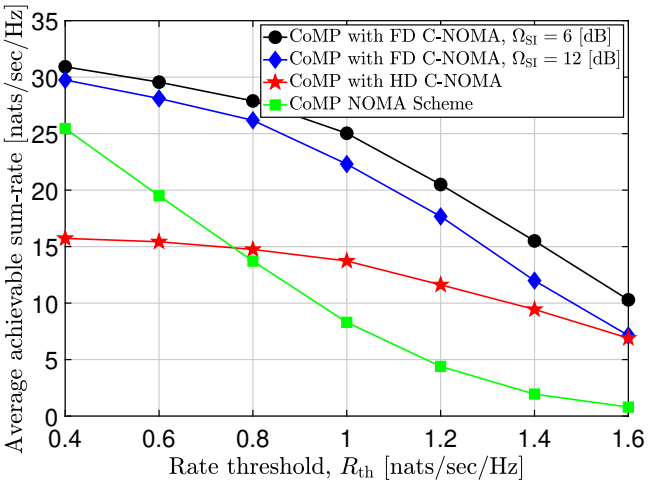


Figure 4.9: Average network sum-rate versus R_{th} with different SI channel gains (P_b = 27 dBm and $P_{\max}$ = 25 dBm).

are not satisfied, the achievable sum rate for this cluster is set to zero. Consequently, when the required QoS R_{th} increases, the BSs allocate more power to the CEUs in order to satisfy the SIC constraints. However, for certain channel realizations, increasing the power allocated to the CEUs will break the satisfaction of the rate constraint at the CCUs. Based on this, by increasing the rate threshold R_{th}, the probability that the feasibility conditions to be satisfied decreases. Therefore, the majority of the possible coordinated C-NOMA/NOMA clusters are in an outage, and hence, the average network sum-rate decreases.

One can see that CoMP-assisted HD C-NOMA has the worst performance at lower values of the data rate threshold due to the pre-log penalty in the DF HD relaying mode. Specifically, at low values of R_{th}, the PA coefficients at the BSs to achieve the SIC constraints is enough to achieve the required QoS at the CEUs. Thus, the benefits brought by the cooperation between CCUs and CEUs in the second time-slot go for only the CEUs. However, when the R_{th} increases, the D2D cooperation is required to meet the required QoS at the CEUs and to allocate sufficient power to the CCUs for maximizing the sum-rate, which makes this scheme outperform the CoMP NOMA scheme.

Fig. 4.10 depicts the effect of increasing the transmit power budget $P_{\max}$ at the CCUs on the performance of the three schemes for different values of SI channel gains. For the proposed scheme, when the power budget at the CCUs increases, the gain resulting from the contribution of the CCUs through the D2D links to the CEUs increases. In fact, for lower values of $P_{\max}$, each BS tries to allocate PA coefficients in a way that

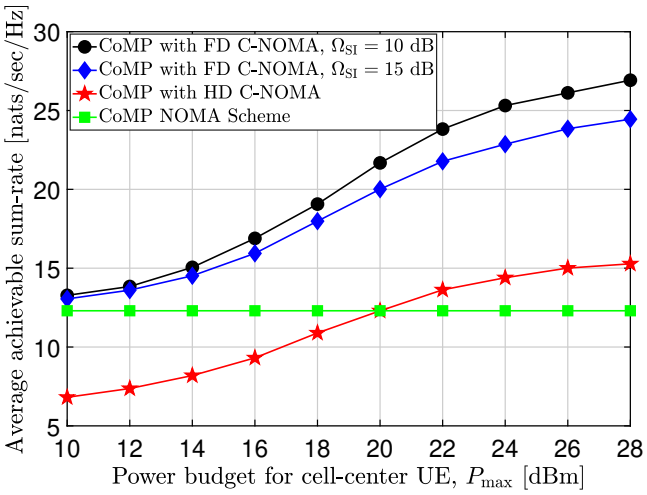


Figure 4.10: Average network sum-rate versus $P_{\max}$ with different SI channel gains ($R_{\text{th}} = 1$ nats/sec/Hz and $P_b = 28$ dBm).

not only guarantees the SIC process at the CCUs, but also achieves the QoS at the CEUs. This is because the contributions from the D2D relaying links are weak. Thus, the gain from the proposed scheme on the network sum-rate compared to the conventional CoMP NOMA scheme will be low. However, when $P_{\max}$ increases, each BS tries only to allocate the PA coefficients such that the SIC constraints are achieved. Meanwhile, for the QoS constraints, the CCUs will assist the CEUs to preserve the required rate threshold. Therefore, most of the transmit power for the BSs goes to the CCUs, which in turn results in an improvement in the network sum-rate. On the other hand, the CoMP-assisted HD C-NOMA only outperforms the CoMP NOMA when the power budget at the CCUs is sufficiently large such that a gain can be obtained from the D2D cooperation. In addition, due to the pre-log penalty in the HD mode, the proposed scheme always outperforms the HD C-NOMA scheme.

4.5.5.5 Network Sum-Rate Performance for Different Clustering Schemes

Fig. 4.11 shows the average network sum-rate versus P_b for different numbers of CEUs F in the network. We compare the performance of the proposed CECM algorithm with the ones of the OUC scheme and the random matching (RM) scheme. Through the RM scheme, the power control policy of each C-NOMA cluster is determined according to **Algorithm 2** and the UEs clustering is randomly performed. Moreover, the PA in the OUC scheme is also determined through **Algorithm 2**. Fig. 4.11 shows that the performance of the CECM algorithm is very close to the one of the OUC scheme.

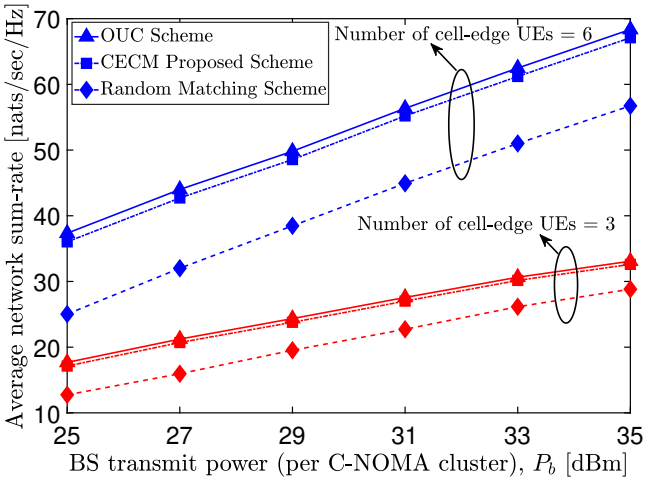


Figure 4.11: Average network sum-rate versus P_b for OUC, CECM, and random matching schemes for different number of CEUs ($R_{\text{th}} = 1$ nats/sec/Hz and $P_{\max} = 28$ dBm).

For instance, when $F = 6$ and $P_b = 29$ dBm, the CECM proposed scheme gets around 96.5% of the network sum-rate achieved by the OUC scheme, which demonstrates the high performance of the proposed CECM algorithm. Furthermore, the proposed UEs clustering scheme significantly outperforms the RM scheme. These observations demonstrate the high performance-complexity efficiency of the CECM algorithm compared to the ones of the OUC and RM schemes.

4.6 Summary

With the quest of enhancing the network connectivity and meeting the high data rate requirements of the B5G networks, we investigated the amalgamation between CoMP and FD C-NOMA. First, we provided a low-complexity sub-optimal power allocation solution for one coordinated C-NOMA cluster considering imperfect CSI and imperfect SIC. Afterward, we studied the joint UEs clustering and power control problem with the objective of maximizing the network sum-rate while guaranteeing the QoS requirement for the UEs, the SIC constraints, and the UEs' power budget constraints. We compare the performance of the proposed scheme with two benchmark schemes, which are CoMP-assisted HD C-NOMA and CoMP NOMA. We show that the proposed CoMP with FD C-NOMA outperforms the other benchmark schemes when the SI channel gains is not sufficiently large regardless of the value of the transmit powers at the BS and the near UEs. Meanwhile, when the SI value is sufficiently large, it is recommended to use either CoMP with HD C-NOMA or CoMP NOMA as an access technique depending on the rate

requirements, the transmit power at the BSs, and the transmit power at the near NOMA UEs. When the power transmit at the BS is sufficiently larger than the power transmit at the near UEs or when the rate requirement is sufficiently low, it is recommended to consider CoMP NOMA as a multiple access technique, otherwise, the CoMP-assisted HD C-NOMA should be adopted. One extension of the proposed scheme is to investigate the beamforming problem when deploying multiple antennae at each BS. This would provide additional performance gains besides the gain brought by exploiting the cooperation between the CCUs and the CEUs.

Chapter 5
RIS-Enabled HD/FD C-NOMA Cellular Networks

5.1 Introduction

Recently, C-NOMA has been deemed as one of the vital enabling MA techniques for the upcoming 6G cellular networks [9, 121]. On the other hand, RIS has recently been recognized as a key promising technology for the 6G wireless network to provide cost-, energy-, and spectral-efficient communications [122, 123]. Motivated by the aforementioned benefits of these two technologies, the potential performance enhancement brought by effectively integrating RIS technology with C-NOMA-based cellular networks is investigated in this Chapter. This combination of RIS and C-NOMA technologies can provide a promising paradigm for upcoming 6G networks by providing additional paths that can jointly construct a stronger combined channel gain for user of interest by leveraging RIS technology and by improving network connectivity by adopting C-NOMA technique.

5.2 State of The Art

The research on RIS-enabled NOMA-based cellular networks is gaining momentum to enhance and improve different performance metrics, such as power consumption [124–126], network spectral-efficiency [127–129], network energy-efficiency [130], and user fairness [131]. The authors in [124] evaluate the minimum power consumption in a two-UE NOMA network for three different multiple access schemes to achieve the same required data rate threshold and have proved that NOMA can achieve superior performance compared to the counterpart OMA schemes in most of the system settings. The transmit beamforming at the BS, phase-shift matrix at the RIS, and the channel gains ordering for a multi-UE NOMA cellular network are jointly optimized to minimize the total transmit power

in [126]. With the goal of maximizing the sum-rate in a two-UE and multi-UE NOMA cellular network, joint optimization of the PA at the BS and the passive beamforming at the RIS were investigated in [127, 128], respectively. Aiming at improving the spectral efficiency of a multi-cell RIS-NOMA network, the problem of joint user association, sub-channel allocation, power control, and passive beamforming is formulated in [129]. By jointly optimizing the active beamforming at the BS and the passive beamforming at the RIS, an efficient algorithm is developed in [130] to maximize the network energy efficiency of a two-user RIS-NOMA network. Joint active beamforming at the BS and passive beamforming at the RIS is investigated in [131] to maximize the minimum achievable data rate for ensuring user fairness.

Going deep into the investigation of RIS-NOMA cellular networks, the performance analysis using tools from stochastic geometry is considered in [14, 132–134]. Specifically, the authors in [14] analyzed the network spectral efficiency when both the RIS technology and CoMP transmission are jointly integrated in NOMA cellular networks. The network spectral efficiency, energy efficiency, and outage probability are derived in [132]. Closed-form expressions for the coverage probability and the ergodic rate are derived in a RIS-assisted two-UE NOMA network without considering the direct link between far NOMA user and BS [133]. Meanwhile, in [134], the rate coverage probability is derived in RIS-assisted NOMA in HetNet considering the effect of the direct link.

All the aforementioned research works [14, 124–134] have only investigated the performance of integrating RIS with NOMA without considering the user-relaying cooperation, whilst there is a lack of investigations in the existing literature on the performance of C-NOMA when the RIS is considered in the network. Recently, the authors in [135] studied the performance of a RIS-assisted HD C-NOMA system by jointly optimizing the active beamforming at the BS, the user relaying power, and the phase-shifts at the RIS. However, this work has not comprehensively analyzed the performance of RIS-aided HD C-NOMA compared to the FD C-NOMA, either with or without the RIS technology.

5.3 Contributions

Against the above background, and to the best of our knowledge, integrating RIS with C-NOMA in HD relaying mode has not been well studied in the literature and remains still unexplored. Moreover, this chapter is one of the early attempts to explore the per-

formance of FD C-NOMA systems with the assistance of RIS. These facts motivate us to study the joint PA at the BS, transmit relaying power at the near UE, and passive beamforming at the RIS that minimizes the total transmit power. Driven by the aforementioned observations, the main contributions of this chapter can be summarized as follows.

- A DL RIS-enabled HD/FD C-NOMA framework consisting of one BS, one near NOMA user, one far NOMA user, and one RIS is considered, where the near NOMA user can relay the message of the far NOMA user in either an HD or an FD relaying mode. For each relaying mode, this framework is formulated as an optimization problem with the objective of minimizing the total transmit power while guaranteeing the data rate QoS requirements for the users, the power budget at both BS and near NOMA user, and the SIC constraint.

- Due to the high coupling between the PA coefficients at both the BS and near NOMA user from one side and the passive beamforming at the RIS from the other side, the formulated total transmit power minimization problem is neither linear nor convex, and hence, is difficult to be directly solved. In order to overcome this challenge, the AO approach is adopted, in which the original optimization problem is decomposed into two sub-problems, namely, a PA sub-problem and a PS optimization sub-problem, which are solved in an alternating manner.

- For the PA sub-problem, and for given phase-shift matrices at the RIS, the feasibility conditions as relations between the QoS requirements, the power budget of the active nodes (BS and near NOMA user), and SIC constraints are derived. Then, the optimal solution of the PA sub-problem is determined in closed-form expressions.

- For given values of the PA coefficients at the BS and fixed transmit relaying power at the near NOMA user, the passive beamforming sub-problem is reformulated as a rank-one constrained optimization problem via matrix lifting and change of variables. Then, a DC representation is formulated for the rank-one constraint. Finally, an efficient SCA algorithm is proposed to solve the resulting problem and to obtain a feasible rank-one solution. In addition, we analyze the convergence and the computational complexity of the overall proposed algorithm.

Extensive simulations were carried out to evaluate the performance of the proposed RIS-enabled HD/FD C-NOMA cellular network. In order to validate the effectiveness of the proposed framework, we compare the performance of the proposed scheme with the other four baseline schemes. Then, we evaluate the performance of these schemes with respect to different system parameters, such as the SI channel gain, the number of RIS elements, the required QoS at the UEs, the RIS location, and the far NOMA user location. In fact, the numerical results unveil that

- The RIS-enabled FD C-NOMA scheme provides a significant gain in terms of the total transmit power compared to the RIS-enabled HD C-NOMA, the FD C-NOMA without RIS, and the RIS-assisted NOMA schemes, despite the existence of high residual SI at near NOMA user.

- The FD C-NOMA with the assistance of RIS has more resistance to the residual SI effect and can tolerate high SI values compared to the same scheme without RIS, which demonstrated the potential of integrating RIS with FD C-NOMA.

- The RIS-enabled HD C-NOMA scheme can beat the FD C-NOMA without the RIS scheme despite the pre-log penalty in the HD relaying mode. This performance gain depends on the number of RIS reflecting elements, the SI channel gain at the near NOMA user, and the required QoS at the NOMA users.

- The location of the RIS depends on the adopted access technique. For instance, in the RIS-assisted HD C-NOMA scheme, the RIS cannot be deployed at the BS and it should be located beside either the near NOMA user or the far NOMA user.

The rest of the Chapter is organized as follows. Section 5.4 presents the system model of the proposed RIS-enabled C-NOMA system. Section 5.5 presents the `SINR` and the rate analysis for both HD and FD relaying modes. Section 5.6 and 5.7 present the formulated optimization problem and the solution approach for the RIS-enabled HD C-NOMA system and the RIS-enabled FD C-NOMA system, respectively. The simulation results and the conclusion are presented in Sections 5.8 and 5.9, respectively. Note that, The notations and symbols adopted throughout this Chapter are summarized in Table 5.1.

Table 5.1: Table of Notations

System Parameters	
P_{BS}, P_{n}	Power budget of BS and $\mathrm{UE_n}$, respectively.
$R_{\mathrm{n}}^{\mathrm{th}}$ and $R_{\mathrm{f}}^{\mathrm{th}}$	Minimum required data rate for $\mathrm{UE_n}$ and $\mathrm{UE_f}$, respectively.
$\beta^{\mathrm{HD}}, \beta^{\mathrm{FD}}$	Power fraction coefficients at $\mathrm{UE_n}$ in HD and FD relaying mode, respectively.
$\alpha_{\mathrm{n}}^{\mathrm{HD}}, \alpha_{\mathrm{f}}^{\mathrm{HD}}$	Power allocation coefficients at BS for $\mathrm{UE_n}$ and $\mathrm{UE_f}$ in HD case, respectively.
$\alpha_{\mathrm{n}}^{\mathrm{FD}}, \alpha_{\mathrm{f}}^{\mathrm{FD}}$	Power allocation coefficients at BS for $\mathrm{UE_n}$ and $\mathrm{UE_f}$ in FD case, respectively.
$\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}$, $\boldsymbol{\Theta}_{(2)}^{\mathrm{HD}}$	RIS phase-shift matrix in the first and second time slots, respectively, in HD case.
$\boldsymbol{\Theta}^{\mathrm{FD}}$	RIS phase-shift matrix in FD case.
$\boldsymbol{h}_{\mathrm{br}}, \boldsymbol{h}_{\mathrm{rn}}, \boldsymbol{h}_{\mathrm{rf}}, \boldsymbol{h}_{\mathrm{nr}}$	Channel gains for BS-RIS, RIS-$\mathrm{UE_n}$, RIS-$\mathrm{UE_f}$, and $\mathrm{UE_n}$-RIS links, respectively.
$h_{\mathrm{bn}}, h_{\mathrm{bf}}, h_{\mathrm{nf}}$	Channel gains for BS-$\mathrm{UE_n}$, BS-$\mathrm{UE_f}$, $\mathrm{UE_n}$-$\mathrm{UE_f}$ links, respectively.
$\hat{\boldsymbol{h}}_{\mathrm{rf}}$	Channel gain for RIS-$\mathrm{UE_f}$ link in the second time-slot for HD case.
η_x	Path-loss exponent for the communication link x.

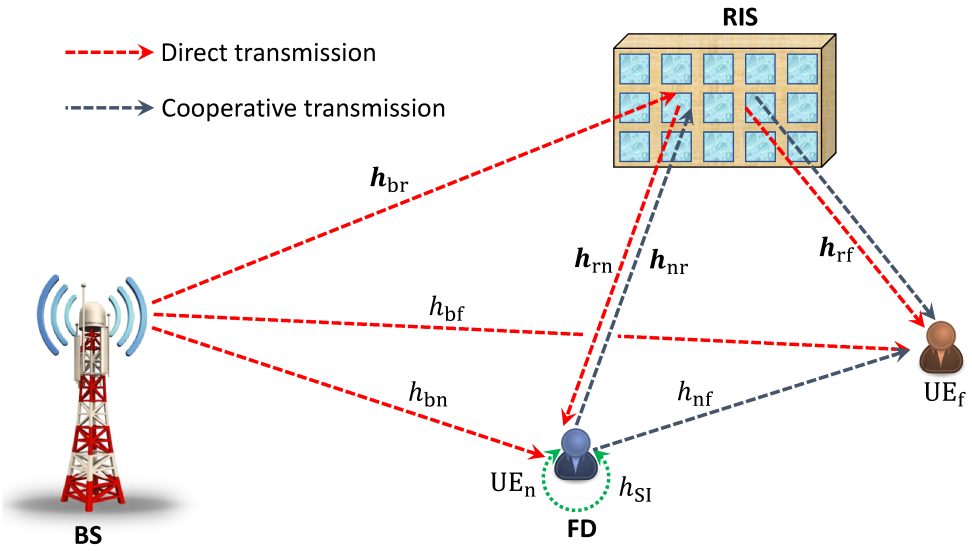


Figure 5.1: RIS-enabled HD/FD C-NOMA cellular network.

5.4 System Model

5.4.1 Network Model

We consider a DL transmission in a RIS-enabled two-UE C-NOMA cellular system, which consists of one BS, one near UE denoted by $\mathrm{UE_n}$, and one far UE denoted by $\mathrm{UE_f}$ as shown in Fig. 5.1. To enhance the performance of C-NOMA-enabled wireless networks, a RIS equipped with M reflecting elements is deployed for improving the wireless service from BS to both $\mathrm{UE_n}$ and $\mathrm{UE_f}$ as well as for boosting the performance of the D2D communication from the $\mathrm{UE_n}$ to $\mathrm{UE_f}$. We assume that the UEs and the BS both have one transmit antenna and one receive antenna. With the assistance of the RIS, the BS serves $\mathrm{UE_n}$ and $\mathrm{UE_f}$ simultaneously using NOMA, while $\mathrm{UE_n}$ relays the signal for $\mathrm{UE_f}$ either in an HD DF relaying mode or in an FD DF relaying mode. Let $h_{\mathrm{bn}} \in \mathbb{C}, h_{\mathrm{bf}} \in \mathbb{C}, \boldsymbol{h}_{\mathrm{br}} \in$

$\mathbb{C}^{M\times 1}, \boldsymbol{h}_{\mathrm{rn}} \in \mathbb{C}^{M\times 1}, \boldsymbol{h}_{\mathrm{nr}} \in \mathbb{C}^{M\times 1}, \boldsymbol{h}_{\mathrm{rf}} \in \mathbb{C}^{M\times 1}, h_{\mathrm{nf}} \in \mathbb{C}$, and h_{SI} be the channel coefficients of the communication links from BS $\longrightarrow$ $\mathrm{UE_n}$, BS $\longrightarrow$ $\mathrm{UE_f}$, BS $\longrightarrow$ RIS, RIS $\longrightarrow$ $\mathrm{UE_n}$, $\mathrm{UE_n}$ $\longrightarrow$ RIS, $\mathrm{UE_n}$ $\longrightarrow$ $\mathrm{UE_f}$, and of the SI link, respectively. Note that, we assume that the CSI of all the communication links is perfectly known at the BS [14, 124–133].[1] Since $\mathrm{UE_n}$ can adopt two different relaying modes, i.e., either HD or FD, two different transmission schemes, i.e., RIS-enabled HD C-NOMA and RIS-enabled FD C-NOMA, can be adopted, which are adequately defined in the next subsection.

5.4.2 Transmission Model

The transmission model for the two-UE C-NOMA consists of DT and CT phases that are detailed as follows.

- Direct transmission: the BS applies superposition coding on the signals intended to $\mathrm{UE_n}$ and $\mathrm{UE_f}$. Then, it transmits the superimposed signal to both of them. The transmitted signal by the BS will hit the RIS and will be then reflected back to $\mathrm{UE_n}$ and $\mathrm{UE_f}$ to improve their signals reception diversity. Following the NOMA principle, $\mathrm{UE_n}$ first performs SIC to decode the intended signal for $\mathrm{UE_f}$. Second, it cancels the decoded signal of the $\mathrm{UE_f}$ from its own reception. Afterward, $\mathrm{UE_n}$ decodes its own signal from the resulting reception; while $\mathrm{UE_f}$ treats the signal of $\mathrm{UE_n}$ as a noise.

- Cooperative transmission: $\mathrm{UE_n}$ relays the decoded signal of $\mathrm{UE_f}$ through a D2D channel. Consequently, two signals will be received at the $\mathrm{UE_f}$'s side. The first is resulting from the transmission of $\mathrm{UE_n}$ and the second signal comes from the reflection by the RIS. Finally, $\mathrm{UE_f}$ combines the received signals coming from the BS, RIS, and $\mathrm{UE_n}$ and then decodes its own signal.

For the case of HD relaying mode, DT and CT occur in two consecutive time-slots. However, for the case of FD relaying mode, they occur in the same time-slot with the cost

[1] In order to characterize the theoretical performance gain and to provide useful insights from the interplay between RIS and C-NOMA, we assume that the channel state information of all the communication links is perfectly known at the BS. Although it is generally difficult to have perfect CSI, various efforts in the literature have proposed recently efficient channel estimation methods for RIS-assisted wireless networks that can be adopted in our system model to provide accurate CSI [136, 137]. For instance, the authors in [136] developed an alternating least square method based on the parallel factor framework that can continuously estimate all the channels with low complexity. Two channel estimation approaches based on compressive sensing and deep learning were proposed in RIS-assisted wireless systems [137].

of inducing SI at $\mathrm{UE_n}$ (the green line in Fig. 5.1). After presenting the main operations of RIS-enabled C-NOMA, we direct our attention toward calculating the received SINR and the corresponding achievable data rates at $\mathrm{UE_n}$ and $\mathrm{UE_f}$ in the two different operation modes (FD and HD) as shown in the next section.

5.5 Downlink SINRs Model and Achievable Rates Analysis

In this section, we present the analysis of both the SINRs and the achievable rates for a C-NOMA-based cellular system assisted by RIS for both HD and FD relaying modes. For the two different relaying modes and at each channel use, the superimposed mixture of the signals intended to $\mathrm{UE_n}$ and $\mathrm{UE_f}$ at the BS are expressed, respectively, as

$$\mathcal{S}^{\mathrm{HD}} = \sqrt{\alpha_{\mathrm{n}}^{\mathrm{HD}} P_{\mathrm{BS}}}\mathcal{S}_{\mathrm{n}} + \sqrt{\alpha_{\mathrm{f}}^{\mathrm{HD}} P_{\mathrm{BS}}}\mathcal{S}_{\mathrm{f}}, \quad \text{and} \quad \mathcal{S}^{\mathrm{FD}} = \sqrt{\alpha_{\mathrm{n}}^{\mathrm{FD}} P_{\mathrm{BS}}}\mathcal{S}_{\mathrm{n}} + \sqrt{\alpha_{\mathrm{f}}^{\mathrm{FD}} P_{\mathrm{BS}}}\mathcal{S}_{\mathrm{f}}, \quad (5.1)$$

where $\mathcal{S}_{\mathrm{n}}$ and $\mathcal{S}_{\mathrm{f}}$ represent the intended signals of $\mathrm{UE_n}$ and $\mathrm{UE_f}$, respectively, such that $\mathbb{E}[|\mathcal{S}_{\mathrm{n}}|^2] = 1$ and $\mathbb{E}[|\mathcal{S}_{\mathrm{f}}|^2] = 1$, P_{BS} represents the power budget of the BS, and for $\mathrm{R} \in \{\mathrm{HD}, \mathrm{FD}\}$, $\alpha_{\mathrm{n}}^{\mathrm{R}}$ and $\alpha_{\mathrm{f}}^{\mathrm{R}}$ represent the power control coefficients allocated by the BS to $\mathrm{UE_n}$ and $\mathrm{UE_f}$ in the relaying mode R, respectively. Note that, $\mathrm{UE_n}$ is located near the BS and has a better channel gain compared to $\mathrm{UE_f}$, which is located far from the BS. Hence, following the principle of NOMA, the power allocation coefficients $\alpha_{\mathrm{n}}^{\mathrm{R}}$ and $\alpha_{\mathrm{f}}^{\mathrm{R}}$ of $\mathrm{UE_n}$ and $\mathrm{UE_f}$ should satisfy the conditions $\alpha_{\mathrm{n}}^{\mathrm{R}} < \alpha_{\mathrm{f}}^{\mathrm{R}}$ and $0 \leq \alpha_{\mathrm{n}}^{\mathrm{R}} + \alpha_{\mathrm{f}}^{\mathrm{R}} \leq 1$ in order to perform downlink NOMA and maintain the power budget at the BS, respectively [69, 107, 138].

5.5.1 RIS-Enabled HD C-NOMA

For the case of RIS-enabled HD C-NOMA, the received signal at $\mathrm{UE_n}$ can be expressed as

$$\mathcal{Y}_{\mathrm{n}}^{\mathrm{HD}} = \left(h_{\mathrm{bn}} + \boldsymbol{h}_{\mathrm{rn}}^{H} \boldsymbol{\Theta}_{(1)}^{\mathrm{HD}} \boldsymbol{h}_{\mathrm{br}}\right) \mathcal{S}^{\mathrm{HD}} + \mathcal{W}_{\mathrm{n}}, \quad (5.2)$$

where $\mathcal{W}_{\mathrm{n}} \sim \mathcal{CN}(0, \sigma_{\mathrm{n}}^2)$ is the additive white Gaussian noise (AWGN) with zero mean and variance σ_{n}^2 and $\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}} = \mathrm{diag}\{\phi_1^{(1)}, \phi_2^{(1)}, \ldots, \phi_M^{(1)}\}$ is the phase-shift matrix of the RIS in the DT, where $\forall m \in [\![1, M]\!]$, $\phi_m^{(1)} = e^{j\theta_m^{(1)}}$, such that $\theta_m^{(1)} \in [0, 2\pi]$ represents the phase-shift of the mth reflecting element of the RIS in the first time slot. According to the NOMA principle, SIC is adopted at $\mathrm{UE_n}$ to decode the message $\mathcal{S}_{\mathrm{f}}$ of $\mathrm{UE_f}$. Consequently, the

received SINR at $\mathrm{UE_n}$ to decode the message of $\mathrm{UE_f}$ can be expressed as

$$\mathtt{SINR}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{f}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\boldsymbol{\theta}_{(1)}\right)=\frac{\alpha^{\mathrm{HD}}_{\mathrm{f}}P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(1)}\boldsymbol{h}_{\mathrm{br}}|^2}{\alpha^{\mathrm{HD}}_{\mathrm{n}}P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(1)}\boldsymbol{h}_{\mathrm{br}}|^2+\sigma^2_{\mathrm{n}}},\tag{5.3}$$

where $\boldsymbol{\theta}_{(1)}=[\theta^{(1)}_1,\theta^{(1)}_2,\ldots,\theta^{(1)}_M]$. Thus, after cancelling $\mathrm{UE_f}$'s message, the SINR at $\mathrm{UE_n}$ to decode its own message can be expressed as

$$\mathtt{SINR}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{n}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\boldsymbol{\theta}_{(1)}\right)=\frac{\alpha^{\mathrm{HD}}_{\mathrm{n}}P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(1)}\boldsymbol{h}_{\mathrm{br}}|^2}{\sigma^2_{\mathrm{n}}}.\tag{5.4}$$

Thus, the achievable data rate at $\mathrm{UE_n}$ to decode the message of $\mathrm{UE_f}$ and to decode its own message can be expressed, respectively, as

$$\mathcal{R}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{f}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\boldsymbol{\theta}_{(1)}\right)=\frac{1}{2}\log_2\left[1+\mathtt{SINR}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{f}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\boldsymbol{\theta}_{(1)}\right)\right],\tag{5.5}$$

$$\mathcal{R}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{n}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\boldsymbol{\theta}_{(1)}\right)=\frac{1}{2}\log_2\left[1+\mathtt{SINR}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{n}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\boldsymbol{\theta}_{(1)}\right)\right].\tag{5.6}$$

On the other hand, the received signal at $\mathrm{UE_f}$ in the DT is expressed as

$$\mathcal{Y}^{(1)}_{\mathrm{f}}=(h_{\mathrm{bf}}+\boldsymbol{h}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(1)}\boldsymbol{h}_{\mathrm{br}})\mathcal{S}^{\mathrm{HD}}+\mathcal{W}^{(1)}_{\mathrm{f}},\tag{5.7}$$

where $\mathcal{W}^{(1)}_{\mathrm{f}}\sim\mathcal{CN}(0,\sigma^2_{\mathrm{f}})$ is the AWGN at $\mathrm{UE_f}$ in the DT. Therefore, the SINR at $\mathrm{UE_f}$ is given by

$$\mathtt{SINR}^{\mathrm{HD}}_{(1)}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\boldsymbol{\theta}_{(1)}\right)=\frac{\alpha^{\mathrm{HD}}_{\mathrm{f}}P_{\mathrm{BS}}|h_{\mathrm{bf}}+\boldsymbol{h}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(1)}\boldsymbol{h}_{\mathrm{br}}|^2}{\alpha^{\mathrm{HD}}_{\mathrm{n}}P_{\mathrm{BS}}|h_{\mathrm{bf}}+\boldsymbol{h}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(1)}\boldsymbol{h}_{\mathrm{br}}|^2+\sigma^2_{\mathrm{f}}}.\tag{5.8}$$

Meanwhile, in the CT, $\mathrm{UE_n}$ forwards the decoded message $\mathcal{S}_{\mathrm{f}}$ to $\mathrm{UE_f}$. Consequently, the observation at $\mathrm{UE_f}$ resulting from the transmission of $\mathrm{UE_n}$ and the reflections by the RIS can be given by

$$\mathcal{Y}^{(2)}_{\mathrm{f}}=\beta^{\mathrm{HD}}P_{\mathrm{n}}\left(h_{\mathrm{nf}}+\hat{\boldsymbol{h}}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(2)}\boldsymbol{h}_{\mathrm{nr}}\right)\mathcal{S}_{\mathrm{f}}+\mathcal{W}^{(2)}_{\mathrm{f}},\tag{5.9}$$

where $\beta^{\mathrm{HD}}\in[0,1]$ is the fraction of the allocated power by $\mathrm{UE_n}$, P_{n} is the power budget at $\mathrm{UE_n}$, $\hat{\boldsymbol{h}}^{H}_{\mathrm{rf}}$ is the channel gain between the RIS and $\mathrm{UE_f}$ in the CT and $\mathcal{W}^{(2)}_{\mathrm{f}}\sim\mathcal{CN}(0,\sigma^2_{\mathrm{f}})$ is the AWGN at $\mathrm{UE_f}$ in the CT. In addition, $\boldsymbol{\Theta}^{\mathrm{HD}}_{(2)}=\mathrm{diag}\{\phi^{(2)}_1,\phi^{(2)}_2,\ldots,\phi^{(2)}_M\}$ is the phase-shift matrix of the RIS in the CT, where $\forall m\in[\![1,M]\!]$, $\phi^{(2)}_m=e^{j\theta^{(2)}_m}$, such that $\theta^{(2)}_m\in[0,2\pi]$ represents the phase-shift of the mth reflecting element of the RIS in the second time slot. Therefore, the received SINR at $\mathrm{UE_f}$ can be expressed as

$$\mathtt{SINR}^{\mathrm{HD}}_{(2)}\left(\beta^{\mathrm{HD}},\boldsymbol{\theta}_{(2)}\right)=\frac{\beta^{\mathrm{HD}}P_{\mathrm{n}}|h_{\mathrm{nf}}+\hat{\boldsymbol{h}}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{HD}}_{(2)}\boldsymbol{h}_{\mathrm{nr}}|^2}{\sigma^2_{\mathrm{f}}},\tag{5.10}$$

where $\boldsymbol{\theta}_{(2)}=[\theta^{(2)}_1,\theta^{(2)}_2,\ldots,\theta^{(2)}_M]$. It is clear that $\mathrm{UE_f}$ receives its own signal from both the DT and CT. As a result, at the end of the CT, $\mathrm{UE_f}$ employs the maximum-ratio-combining (MRC) technique to combine these receptions in both the DT and the CT in order to decode its own signal [35, 107].

$$\mathcal{R}^{\mathrm{HD}}_{\mathrm{MRC}}\left(\mathbf{p}^{\mathrm{HD}},\boldsymbol{\theta}^{\mathrm{HD}}\right)=\frac{1}{2}\log_2\left[1+\mathtt{SINR}^{\mathrm{HD}}_{(1)}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\boldsymbol{\theta}_{(1)}\right)+\mathtt{SINR}^{\mathrm{HD}}_{(2)}\left(\beta^{\mathrm{HD}},\boldsymbol{\theta}_{(2)}\right)\right],\tag{5.11}$$

where $\mathbf{p}^{\mathrm{HD}}=\{\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\beta^{\mathrm{HD}}\}$ and $\boldsymbol{\theta}^{\mathrm{HD}}=\{\boldsymbol{\theta}_{(1)},\boldsymbol{\theta}_{(2)}\}$. However, the rate $\mathcal{R}_{\mathrm{MRC}}$ can be

achievable if and only if $\mathrm{UE_n}$ has the ability to decode the message $\mathcal{S}_\mathrm{f}$ of $\mathrm{UE_f}$. Thus, the data rate achieved at $\mathrm{UE_f}$ is bounded by the data rate of $\mathrm{UE_n}$ to decode the message of $\mathrm{UE_f}$, i.e. $\mathcal{R}_{\mathrm{n}\longrightarrow\mathrm{f}}$ [107]. Therefore, the achievable data rate of $\mathrm{UE_f}$ to decode its own message can be given as

$$\mathcal{R}^{\mathrm{HD}}_{\mathrm{f}\longrightarrow\mathrm{f}}\left(\mathbf{p}^{\mathrm{HD}},\boldsymbol{\theta}^{\mathrm{HD}}\right)=\min\left(\mathcal{R}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{f}}\left(\alpha^{\mathrm{HD}}_{\mathrm{n}},\alpha^{\mathrm{HD}}_{\mathrm{f}},\boldsymbol{\theta}_{(1)}\right),\mathcal{R}^{\mathrm{HD}}_{\mathrm{MRC}}\left(\mathbf{p}^{\mathrm{HD}},\boldsymbol{\theta}^{\mathrm{HD}}\right)\right). \tag{5.12}$$

5.5.2 RIS-Enabled FD C-NOMA

For the case of RIS-enabled FD C-NOMA, both the DT and CT are executed simultaneously using the same radio channel. Due to this, $\mathrm{UE_n}$ suffers from a SI resulting from receiving data from the BS and transmitting data to the $\mathrm{UE_f}$ simultaneously within the same co-channel [32]. Hence, the received signal at $\mathrm{UE_n}$ is given by

$$\mathcal{Y}^{\mathrm{FD}}_{\mathrm{n}}=\left(h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{FD}}\boldsymbol{h}_{\mathrm{br}}\right)\mathcal{S}^{\mathrm{FD}}+h_{\mathrm{SI}}\tilde{\mathcal{S}}_{\mathrm{f}}+\mathcal{W}_{\mathrm{n}}, \tag{5.13}$$

where h_{SI} represents the SI channel coefficient at $\mathrm{UE_n}$ and it is assumed to follow a complex symmetric Gaussian random variable with a zero mean and Ω_{SI} standard deviation, i.e., $\mathcal{CN}(0,\Omega^2_{\mathrm{SI}})$ [107, 111]. Due to the processing time resulting from decoding the message $\mathcal{S}_\mathrm{f}$ of $\mathrm{UE_f}$ at $\mathrm{UE_n}$, the message $\tilde{\mathcal{S}}_\mathrm{f}$ relayed from $\mathrm{UE_n}$ to $\mathrm{UE_f}$ is a delayed version of the original message $\mathcal{S}_\mathrm{f}$ intended to $\mathrm{UE_f}$. In other words, by denoting τ the decoding time of the message $\mathcal{S}_\mathrm{f}$ of $\mathrm{UE_f}$ at $\mathrm{UE_n}$, the message $\tilde{\mathcal{S}}_\mathrm{f}$ satisfies $\tilde{\mathcal{S}}_\mathrm{f}(i)=\mathcal{S}_\mathrm{f}(i-\tau)$, where i denotes the ith time slot [107, 111]. This transmission of $\tilde{\mathcal{S}}_\mathrm{f}$ causes an interference at $\mathrm{UE_n}$. In addition,$\boldsymbol{\Theta}^{\mathrm{FD}}=\mathrm{diag}\{\phi_1,\phi_2,\ldots,\phi_M\}$ is the phase-shift matrix of the RIS, where $\forall m\in[\![1,M]\!]$, $\phi_m=e^{j\theta_m}$ such that $\theta_m\in[0,2\pi]$ represents the phase-shift of the mth reflecting element of the RIS. Therefore, the data rate of $\mathrm{UE_n}$ to decode the message of $\mathrm{UE_f}$ can be expressed as

$$\mathcal{R}^{\mathrm{FD}}_{\mathrm{n}\longrightarrow\mathrm{f}}(\mathbf{p}^{\mathrm{FD}},\boldsymbol{\theta}^{\mathrm{FD}})=\log_2\left[1+\frac{\alpha^{\mathrm{FD}}_{\mathrm{f}}P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{FD}}\boldsymbol{h}_{\mathrm{br}}|^2}{\alpha^{\mathrm{FD}}_{\mathrm{n}}P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{FD}}\boldsymbol{h}_{\mathrm{br}}|^2+\beta^{\mathrm{FD}}P_{\mathrm{n}}\gamma_{\mathrm{SI}}+\sigma^2_{\mathrm{n}}}\right], \tag{5.14}$$

where $\gamma_{\mathrm{SI}}=|h_{\mathrm{SI}}|^2$, $\boldsymbol{\theta}^{\mathrm{FD}}=[\theta_1,\theta_2,\ldots,\theta_M]$, and $\mathbf{p}^{\mathrm{FD}}=(\alpha^{\mathrm{FD}}_{\mathrm{n}},\alpha^{\mathrm{FD}}_{\mathrm{f}},\beta^{\mathrm{FD}})$. Then, after the successive decoding and canceling of the signal of $\mathrm{UE_f}$, $\mathrm{UE_n}$ decodes its own signal $\mathcal{S}_\mathrm{n}$. Consequently, the achievable data rate of $\mathrm{UE_n}$ to decode its own signal can be given as

$$\mathcal{R}^{\mathrm{FD}}_{\mathrm{n}\longrightarrow\mathrm{n}}\left(\alpha^{\mathrm{FD}}_{\mathrm{n}},\beta^{\mathrm{FD}},\boldsymbol{\theta}^{\mathrm{FD}}\right)=\log_2\left[1+\frac{\alpha^{\mathrm{FD}}_{\mathrm{n}}P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}^{H}_{\mathrm{rn}}\boldsymbol{\Theta}^{\mathrm{FD}}\boldsymbol{h}_{\mathrm{br}}|^2}{\beta^{\mathrm{FD}}P_{\mathrm{n}}\gamma_{\mathrm{SI}}+\sigma^2_{\mathrm{n}}}\right]. \tag{5.15}$$

Afterward, $\mathrm{UE_n}$ forwards the signal $\tilde{\mathcal{S}}_\mathrm{f}$ to $\mathrm{UE_f}$. Thus, the received signal at $\mathrm{UE_f}$ is given by

$$\mathcal{Y}^{\mathrm{FD}}=\left(h_{\mathrm{bf}}+\boldsymbol{h}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{FD}}\boldsymbol{h}_{\mathrm{br}}\right)\mathcal{S}^{\mathrm{FD}}+\beta^{\mathrm{FD}}P_{\mathrm{n}}\left(h_{\mathrm{nf}}+\boldsymbol{h}^{H}_{\mathrm{rf}}\boldsymbol{\Theta}^{\mathrm{FD}}\boldsymbol{h}_{\mathrm{nr}}\right)\tilde{\mathcal{S}}_{\mathrm{f}}+\mathcal{W}_{\mathrm{f}}, \tag{5.16}$$

where the first term in (5.16) is due to the transmission of the BS, whereas the second term results from the transmission of the near user over the D2D communication link. Note that, in FD scenario, $\boldsymbol{h}_{\rm nr} \triangleq \boldsymbol{h}_{\rm rn}$. Note that, the processing delay τ is very small compared to the duration of one time slot, i.e., $\tau < t_{i+1} - t_i$, where t_{i+1} and t_i are the time of the $(i+1)$th and ith time slots, respectively [107, 111]. Thus, $\mathrm{UE_f}$ receives its own message from the BS and from $\mathrm{UE_n}$ at approximately the same channel use. Based on [10, 107, 111], the two signals from the BS and $\mathrm{UE_n}$ are fully resolvable at $\mathrm{UE_f}$, and hence, they can be appropriately co-phased and combined using MRC. Thus, based on the above discussion and on the results of [10, 111], the achievable data rate of $\mathrm{UE_f}$ can be expressed as

$$\mathcal{R}^{\rm FD}_{\rm MRC}(\mathbf{p}^{\rm FD}, \boldsymbol{\theta}^{\rm FD}) = \log_2 \left[1 + \frac{\alpha^{\rm FD}_{\rm f} P_{\rm BS} |h_{\rm bf} + \boldsymbol{h}^H_{\rm rf} \boldsymbol{\Theta}^{\rm FD} \boldsymbol{h}_{\rm br}|^2 + \beta^{\rm FD} P_{\rm n} |h_{\rm nf} + \boldsymbol{h}^H_{\rm rf} \boldsymbol{\Theta}^{\rm FD} \boldsymbol{h}_{\rm nr}|^2}{\alpha^{\rm FD}_{\rm n} P_{\rm BS} |h_{\rm bf} + \boldsymbol{h}^H_{\rm rf} \boldsymbol{\Theta}^{\rm FD} \boldsymbol{h}_{\rm br}|^2 + \sigma^2_{\rm f}} \right]. \tag{5.17}$$

Based on this analysis and according to what was explained in the HD case, the achievable data rate of $\mathrm{UE_f}$ to decode its own message can be given by

$$\mathcal{R}^{\rm FD}_{\rm f \longrightarrow f} \left(\mathbf{p}^{\rm FD}, \boldsymbol{\theta}^{\rm FD} \right) = \min \left(\mathcal{R}^{\rm FD}_{\rm n \longrightarrow f} \left(\mathbf{p}^{\rm FD}, \boldsymbol{\theta}^{\rm FD} \right), \mathcal{R}^{\rm FD}_{\rm MRC} \left(\mathbf{p}^{\rm FD}, \boldsymbol{\theta}^{\rm FD} \right) \right). \tag{5.18}$$

5.6 RIS-Enabled HD C-NOMA: Problem Formulation and Solution Approach

5.6.1 Problem Formulation

With the quest of improving the performance of the proposed RIS-enabled HD C-NOMA, an optimization problem is formulated with the objective of minimizing the total transmit power by the BS and the near user. By optimizing the power allocation coefficients at the BS $\left(\alpha^{\rm HD}_{\rm n}, \alpha^{\rm HD}_{\rm f}\right)$, the power fraction coefficient at the near user $\beta^{\rm HD}$, and the phase-shift matrices for the RIS $\left(\boldsymbol{\theta}_{(1)}, \boldsymbol{\theta}_{(2)}\right)$, the total transmit power minimization problem for the proposed RIS-enabled HD C-NOMA framework can be formulated as follows.

$$\mathrm{OPT-HD}: \min_{\substack{\boldsymbol{\theta}_{(1)}, \boldsymbol{\theta}_{(2)}, \\ \alpha^{\rm HD}_{\rm n}, \alpha^{\rm HD}_{\rm f}, \beta^{\rm HD}}} (\alpha^{\rm HD}_{\rm n} + \alpha^{\rm HD}_{\rm f}) P_{\rm BS} + \beta^{\rm HD} P_{\rm n}, \tag{5.19a}$$

$$\text{s.t. } 0 \leq \alpha^{\rm HD}_{\rm n} \leq \alpha^{\rm HD}_{\rm f}, \tag{5.19b}$$

$$0 \leq \alpha^{\rm HD}_{\rm n} + \alpha^{\rm HD}_{\rm f} \leq 1, \tag{5.19c}$$

$$0 \leq \beta^{\rm HD} \leq 1, \tag{5.19d}$$

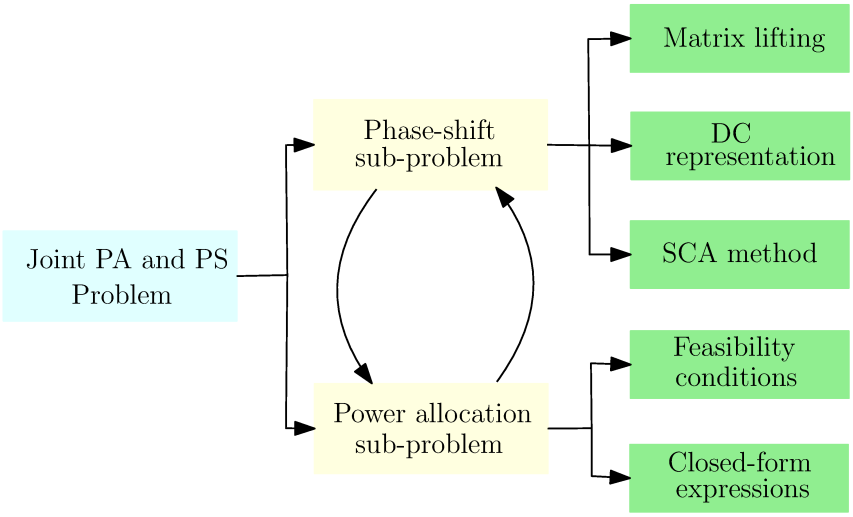


Figure 5.2: Problem formulation and solution roadmap.

$$\mathcal{R}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{n}}(\alpha^{\mathrm{HD}}_{\mathrm{n}}, \boldsymbol{\theta}_{(1)}) \geq R^{\mathrm{th}}_{\mathrm{n}}, \tag{5.19e}$$

$$\min\left(\mathcal{R}^{\mathrm{HD}}_{\mathrm{MRC}}(\mathbf{p}^{\mathrm{HD}}, \boldsymbol{\theta}^{\mathrm{HD}}), \mathcal{R}^{\mathrm{HD}}_{\mathrm{n}\longrightarrow\mathrm{f}}(\alpha^{\mathrm{HD}}_{\mathrm{n}}, \alpha^{\mathrm{HD}}_{\mathrm{f}}, \boldsymbol{\theta}_{(1)})\right) \geq R^{\mathrm{th}}_{\mathrm{f}}, \tag{5.19f}$$

$$|\phi^{(1)}_m|, |\phi^{(2)}_m| \ = 1, \qquad \forall\ m \in [\![1, M+1]\!], \tag{5.19g}$$

where constraint (5.19b) represents the SIC constraint, and constraints (5.19c) and (5.19d) guarantee that the total transmit power by the BS and $\mathrm{UE_n}$ do not exceed their power budget, respectively. Constraints (5.19e)-(5.19f) represent the QoS constraints for $\mathrm{UE_n}$ and $\mathrm{UE_f}$, where $R^{\mathrm{th}}_{\mathrm{n}}$ and $R^{\mathrm{th}}_{\mathrm{f}}$ represent the required QoS in terms of minimum data rate for $\mathrm{UE_n}$ and $\mathrm{UE_f}$, respectively. It can be seen that it is challenging to solve $\mathrm{OPT}-\mathrm{HD}$ directly due to the high coupling between the power allocation coefficients $\left(\alpha^{\mathrm{HD}}_{\mathrm{n}}, \alpha^{\mathrm{HD}}_{\mathrm{f}}, \beta^{\mathrm{HD}}\right)$ and the RIS phase-shift coefficients $\left(\boldsymbol{\theta}_{(1)}, \boldsymbol{\theta}_{(2)}\right)$ as well as the non-convexity of constraints (5.19e)-(5.19f). Therefore, problem $\mathrm{OPT}-\mathrm{HD}$ is hard to be solved by common standard optimization techniques. Thus, it is necessary to transform problem $\mathrm{OPT}-\mathrm{HD}$ into tractable sub-problems that can be solved alternatively. Toward this end, the AO approach is utilized to solve $\mathrm{OPT}-\mathrm{HD}$ in an efficient manner. The solution roadmap is discussed as shown in Fig. 5.2.

Accordingly, unlike [135], which decomposes the main problem into the DT sub-problem and CT sub-problem, we divide problem $\mathrm{OPT}-\mathrm{HD}$ into two sub-problems, i.e. power control optimization sub-problem and RIS passive beamforming (phase-shift coefficients) optimization sub-problem. This is because we seek for deriving optimal closed-form expressions for the power control coefficients $\left(\alpha^{\mathrm{HD}}_{\mathrm{n}}, \alpha^{\mathrm{HD}}_{\mathrm{f}}, \beta^{\mathrm{FD}}\right)$ as a function of the phase-shift coefficients $\left(\boldsymbol{\theta}_{(1)}, \boldsymbol{\theta}_{(2)}\right)$. In particular, we start by optimizing the power allocation

coefficients at the BS and the power fraction coefficient at the near UE given the phase-shift coefficients $\boldsymbol{\theta}_{(1)}$ and $\boldsymbol{\theta}_{(2)}$ at the RIS in the DT and CT. Then, once the optimal power allocation coefficients at the BS and at $\mathrm{UE_n}$ are obtained, we direct our attention to the optimal phase-shift coefficients of the RIS at the DT and CT. Based on the above discussion, the power control optimization problem can be written as

$$\mathrm{PC-HD}: \min_{\alpha_{\mathrm{n}}^{\mathrm{HD}},\alpha_{\mathrm{f}}^{\mathrm{HD}},\beta^{\mathrm{HD}}} \left(\alpha_{\mathrm{n}}^{\mathrm{HD}}+\alpha_{\mathrm{f}}^{\mathrm{HD}}\right) P_{\mathrm{BS}}+\beta^{\mathrm{HD}} P_{\mathrm{n}}, \tag{5.20a}$$

$$\text{s.t. (5.19b)} - \text{(5.19f)}, \tag{5.20b}$$

whereas the passive beamforming optimization problem can be presented as

$$\mathrm{PS-HD}: \text{Find } \ \boldsymbol{\theta}_{(1)}, \boldsymbol{\theta}_{(2)}, \tag{5.21a}$$

$$\text{s.t. (5.19e)} - \text{(5.19g)}. \tag{5.21b}$$

5.6.2 RIS-enabled HD C-NOMA: Power Control Optimization

In this part, we assume that the phase-shift matrices $\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}$ and $\boldsymbol{\Theta}_{(2)}^{\mathrm{HD}}$ are fixed. Based on this, we denote by

$$\begin{aligned}
\gamma_{\mathrm{bn}} &\triangleq \frac{P_{\mathrm{BS}}|h_{\mathrm{bn}}+\boldsymbol{h}_{\mathrm{rn}}^{H}\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}\boldsymbol{h}_{\mathrm{br}}|^2}{\sigma_{\mathrm{n}}^2},\\
\gamma_{\mathrm{bf}} &\triangleq \frac{P_{\mathrm{BS}}|h_{\mathrm{bf}}+\boldsymbol{h}_{\mathrm{rf}}^{H}\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}\boldsymbol{h}_{\mathrm{br}}|^2}{\sigma_{\mathrm{f}}^2},\\
\gamma_{\mathrm{d}} &\triangleq \frac{P_{\mathrm{n}}|h_{\mathrm{nf}}+\hat{\boldsymbol{h}}_{\mathrm{rf}}^{H}\boldsymbol{\Theta}_{(2)}^{\mathrm{HD}}\boldsymbol{h}_{\mathrm{nr}}|^2}{\sigma_{\mathrm{f}}^2}.
\end{aligned} \tag{5.22}$$

Before deriving the optimal power control of problem $\mathrm{PC-HD}$, one needs to specify its feasibility conditions. The feasibility conditions of problem $\mathrm{PC-HD}$ define the conditions under which at least one feasible solution for this problem does exist. In addition, a feasible solution for problem $\mathrm{PC-HD}$ defines a solution that satisfies the constraint of this problem. In this context, the feasibility conditions of problem $\mathrm{PC-HD}$ are presented in the following theorem.

Theorem 5.1. *Problem* $\mathrm{PC-HD}$ *is feasible if and only if the following conditions hold.*

$$\textit{Condition 1: } \alpha_{\min}^{\mathrm{HD}} \le \alpha_{\max}^{\mathrm{HD}}, \tag{5.23a}$$

$$\textit{Condition 2: } \beta_{\min}^{\mathrm{HD}} \le \beta_{\max}^{\mathrm{HD}}, \tag{5.23b}$$

where $\alpha_{\min}^{\mathrm{HD}}$, $\alpha_{\max}^{\mathrm{HD}}$, $\beta_{\min}^{\mathrm{HD}}$ *and* $\beta_{\max}^{\mathrm{HD}}$ *are expressed, respectively, as*

$$\begin{aligned}
\alpha_{\min}^{\mathrm{HD}} &= \frac{t_{\mathrm{n}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}},\\
\alpha_{\max}^{\mathrm{HD}} &= \min\left(0, \frac{\gamma_{\mathrm{bn}}-t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}\left(t_{\mathrm{f}}^{\mathrm{HD}}+1\right)}\right),
\end{aligned}$$

$$\beta_{\min}^{\mathrm{HD}} = \max\left(0, \frac{1}{\gamma_{\mathrm{d}}}\left(t_{\mathrm{f}}^{\mathrm{HD}} - \frac{\gamma_{\mathrm{bn}} - t_{\mathrm{n}}^{\mathrm{HD}}}{t_{\mathrm{n}}^{\mathrm{HD}} + \frac{\gamma_{\mathrm{bn}}}{\gamma_{\mathrm{bf}}}}\right)\right),$$

$$\beta_{\max}^{\mathrm{HD}} = 1, \tag{5.24}$$

such that $t_{\mathrm{n}}^{\mathrm{HD}} = 2^{2R_{\mathrm{n}}^{\mathrm{th}}} - 1$ *and* $t_{\mathrm{f}}^{\mathrm{HD}} = 2^{2R_{\mathrm{n}}^{\mathrm{th}}} - 1$.

Proof. See Appendix C1. ■

Afterward, assuming that $\mathrm{PC}-\mathrm{HD}$ is feasible, its optimal solution is given in the following theorem.

Theorem 5.2. *Assuming that* $\mathrm{PC}-\mathrm{HD}$ *is feasible, i.e., conditions* (5.23a) *and* (5.23b) *hold, its optimal solution is expressed as follows. Let* $\mathbf{p}_1^{\mathrm{HD}}$ *and* $\mathbf{p}_2^{\mathrm{HD}}$ *denote the power control scheme expressed, respectively, as*

$$\mathbf{p}_1^{\mathrm{HD}} =$$

$$\left(\alpha_{\mathrm{n},1}^{\mathrm{HD}}, \alpha_{\mathrm{f},1}^{\mathrm{HD}}, \beta_1^{\mathrm{HD}}\right) = \left(\alpha_{\min}^{\mathrm{HD}}, \max\left(\alpha_{\min}^{\mathrm{HD}}, \alpha_{\min}^{\mathrm{HD}} t_{\mathrm{f}}^{\mathrm{HD}} + \frac{t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{n}}}\right), \frac{1}{\gamma_{\mathrm{d}}}\left(t_{\mathrm{f}}^{\mathrm{HD}} - \frac{\alpha_{\mathrm{f},1}^{\mathrm{HD}}\gamma_{\mathrm{f}}}{\alpha_{\mathrm{n},1}^{\mathrm{HD}}\gamma_{\mathrm{f}} + 1}\right)\right), \tag{5.25a}$$

$$\mathbf{p}_2^{\mathrm{HD}} = \left(\alpha_{\mathrm{n},2}^{\mathrm{HD}}, \alpha_{\mathrm{f},2}^{\mathrm{HD}}, \beta_2^{\mathrm{HD}}\right) = \left(\alpha_{\min}^{\mathrm{HD}}, \frac{\left(\alpha_{\min}^{\mathrm{HD}}\gamma_{\mathrm{bn}} + 1\right) t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}}, 0\right). \tag{5.25b}$$

Then, the optimal power control scheme is expressed as

$$\mathbf{p}^{\mathrm{HD}^*} = \left(\alpha_{\mathrm{n}}^{\mathrm{HD}^*}, \alpha_{\mathrm{f}}^{\mathrm{HD}^*}, \beta^{\mathrm{HD}^*}\right) = \underset{\mathbf{p}^{\mathrm{HD}} \in \left\{\mathbf{p}_1^{\mathrm{HD}}, \mathbf{p}_2^{\mathrm{HD}}\right\}}{\arg\min} f\left(\mathbf{p}^{\mathrm{HD}}\right), \tag{5.26}$$

where the function f *is the objective function of problem* $\mathrm{PC}-\mathrm{HD}$*, i.e.,*

$$f\left(\mathbf{p}^{\mathrm{HD}}\right) = f\left(\alpha_{\mathrm{n}}^{\mathrm{HD}}, \alpha_{\mathrm{f}}^{\mathrm{HD}}, \beta^{\mathrm{HD}}\right) = \left(\alpha_{\mathrm{n}}^{\mathrm{HD}} + \alpha_{\mathrm{f}}^{\mathrm{HD}}\right) P_{\mathrm{BS}} + \beta^{\mathrm{HD}} P_{\mathrm{n}}. \tag{5.27}$$

Proof. See Appendix C2. ■

5.6.3 RIS-enabled HD C-NOMA: Phase-Shift Coefficients Optimization

In this subsection, the PS coefficients in both DT and CT are optimized with given values of $\left(\alpha_{\mathrm{n}}^{\mathrm{HD}}, \alpha_{\mathrm{f}}^{\mathrm{HD}}, \beta^{\mathrm{HD}}\right)$. One can see that $\mathrm{PS}-\mathrm{HD}$ is a feasibility check problem (finding the phase-shift coefficient for each element such that the QoS constraints are satisfied). Note that, the RIS provides additional paths to construct a stronger combined channel gain at the intended receiver. Therefore, the best channel gain in the second time-slot can be achieved when the reflected signals from the RIS can be constructively added at $\mathrm{UE_f}$ and be co-phased with the direct D2D link from $\mathrm{UE_n}$ to $\mathrm{UE_f}$. As a result, optimal PS coefficients in CT can be obtained as follows [14, 139].

$$\theta_m^{(2)} = \arg\left(h_{\mathrm{nf}}\right) - \arg\left([\boldsymbol{h}_{\mathrm{nr}}]_m [\hat{\boldsymbol{h}}_{\mathrm{rf}}]_m\right). \tag{5.28}$$

Consequently, $\texttt{SINR}_{(2)}^{\mathrm{HD}}$ can be expressed as

$$\texttt{SINR}_{(2)}^{\mathrm{HD}}\left(\beta^{\mathrm{HD}}\right)=\frac{\beta^{\mathrm{HD}}P_{\mathrm{n}}\left(|h_{\mathrm{nf}}|+\sum_{m=1}^{M}|[\hat{\boldsymbol{h}}_{\mathrm{rf}}]_m[\boldsymbol{h}_{\mathrm{nr}}]_m|\right)^2}{\sigma_{\mathrm{f}}^2}. \tag{5.29}$$

After obtaining the optimal value of $\theta_{(2)}^{\mathrm{HD}}$, it can be seen that the optimization problem PS − HD is non-convex due to the unit modulus constraint $|\phi_m^{(1)}|=1$. In order to overcome this issue and to find the phase-shift in the first time slot, i.e., $\theta_{(1)}$, we first reformulate PS − HD into a rank-one constrained optimization problem via matrix lifting and change of variables. Then, a DC representation is formulated for the rank-one constraint. Finally, an efficient SCA algorithm is proposed to solve the resulting problem and to obtain a feasible rank-one solution.

1) *Rank-one constrained optimization problem*: Let us start by defining $\boldsymbol{v}\triangleq[\phi_1^{(1)},\ldots,\phi_M^{(1)}]$. By applying the change of variables $\boldsymbol{h}_{\mathrm{rn}}^H\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}\boldsymbol{h}_{\mathrm{br}}=\boldsymbol{v}^H\boldsymbol{\Phi}$, where $\boldsymbol{\Phi}=\mathrm{diag}(\boldsymbol{h}_{\mathrm{rn}}^H)\boldsymbol{h}_{\mathrm{br}}\in\mathbb{C}^{M\times 1}$ and $\boldsymbol{h}_{\mathrm{rf}}^H\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{br}}=\boldsymbol{v}^H\boldsymbol{\Psi}$, where $\boldsymbol{\Psi}=\mathrm{diag}(\boldsymbol{h}_{\mathrm{rf}}^H)\boldsymbol{h}_{\mathrm{br}}\in\mathbb{C}^{M\times 1}$, we have $|h_{\mathrm{bn}}+\boldsymbol{h}_{\mathrm{rn}}^H\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}\boldsymbol{h}_{\mathrm{br}}|^2$ $=|h_{\mathrm{bn}}+\boldsymbol{v}^H\boldsymbol{\Phi}|^2$ and $|h_{\mathrm{bf}}+\boldsymbol{h}_{\mathrm{rf}}^H\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}}\boldsymbol{h}_{\mathrm{br}}|^2=|h_{\mathrm{bf}}+\boldsymbol{v}^H\boldsymbol{\Psi}|^2$. By introducing an auxiliary variable t, the PS − HD can be written as

$$\mathcal{P}:\text{Find}\quad\boldsymbol{\theta}_{(1)}^{\mathrm{HD}}, \tag{5.30a}$$

$$\text{s.t. }\alpha_{\mathrm{n}}^{\mathrm{HD}}P_{\mathrm{BS}}\left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bn}}\boldsymbol{V}\right)+|h_{\mathrm{bn}}|^2\right)\geq t_{\mathrm{n}}^{\mathrm{HD}}\sigma_{\mathrm{n}}^2, \tag{5.30b}$$

$$\alpha_{\mathrm{f}}^{\mathrm{HD}}P_{\mathrm{BS}}\left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bf}}\boldsymbol{V}\right)+|h_{\mathrm{bf}}|^2\right)\geq\left(t_{\mathrm{f}}^{\mathrm{HD}}-\texttt{SINR}_{(2)}^{\mathrm{HD}}\right)\left(\alpha_{\mathrm{n}}^{\mathrm{HD}}P_{\mathrm{BS}}\left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bf}}\boldsymbol{V}\right)+|h_{\mathrm{bf}}|^2\right)+\sigma_{\mathrm{f}}^2\right), \tag{5.30c}$$

$$\alpha_{\mathrm{f}}^{\mathrm{HD}}P_{\mathrm{BS}}\left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bn}}\boldsymbol{V}\right)+|h_{\mathrm{bn}}|^2\right)\geq t_{\mathrm{f}}^{\mathrm{HD}}\left(\alpha_{\mathrm{n}}^{\mathrm{HD}}P_{\mathrm{BS}}\left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bn}}\boldsymbol{V}\right)+|h_{\mathrm{bn}}|^2\right)+\sigma_{\mathrm{n}}^2\right), \tag{5.30d}$$

$$\boldsymbol{V}\succeq 0, \tag{5.30e}$$

$$[\boldsymbol{V}]_{m,m}=1,\qquad\forall\, m\in[\![1,M+1]\!], \tag{5.30f}$$

$$\mathrm{rank}(\boldsymbol{V})=1, \tag{5.30g}$$

where

$$\boldsymbol{Q}_{\mathrm{bn}}=\begin{bmatrix}\boldsymbol{\Phi}\boldsymbol{\Phi}^H & \boldsymbol{\Phi}h_{\mathrm{bn}}^H\\ h_{\mathrm{bn}}\boldsymbol{\Phi}^H & 0\end{bmatrix},\quad\boldsymbol{Q}_{\mathrm{bf}}=\begin{bmatrix}\boldsymbol{\Psi}\boldsymbol{\Psi}^H & \boldsymbol{\Psi}h_{\mathrm{bf}}^H\\ h_{\mathrm{bf}}\boldsymbol{\Psi}^H & 0\end{bmatrix}\quad\text{and}\quad\bar{\boldsymbol{v}}=\begin{bmatrix}\boldsymbol{v}\\ t\end{bmatrix}, \tag{5.31}$$

where $\boldsymbol{V}=\bar{\boldsymbol{v}}\bar{\boldsymbol{v}}^H$, which should achieve rank-one constraint. Note that $\bar{\boldsymbol{v}}^H\boldsymbol{Q}_z\bar{\boldsymbol{v}}=\mathrm{tr}(\boldsymbol{Q}_z\bar{\boldsymbol{v}}\bar{\boldsymbol{v}}^H)$ for all $z\in\{\mathrm{bn},\mathrm{bf}\}$. One common approach adopted in the literature to tackle the rank-one constraint is to apply semidefinite relaxation (SDR). By dropping the rank-one constraint, the relaxed optimization problem ends up with convex semidefinite programming that can be efficiently solved by existing convex optimization solvers such as CVX [131]. By solving the relaxed optimization problem, one can obtain a global optimal

solution of the original problem if the solution of the relaxed problem achieves rank-one. However, after carrying out deep investigations in the literature and according to recent studies, SDR may not be the best choice to tackle rank-one constrained optimization problems [140–144]. This is because in the case when the dimension of the optimization variable is large, the SDR method has a low probability of returning a rank-one solution and it often fails to reach it [140–142]. Hence, the Gaussian randomization (GR) should be applied after the SDR technique to obtain a rank-one sub-optimal solution. However, this sub-optimal solution may not be feasible to meet the QoS and the SIC constraints, and hence, this causes early stopping in the AO approach. In addition, due to the application of the GR method in solving the feasibility check of the passive beamforming problem, the convergence is not always guaranteed [143–145].

2) *DC representation for the rank-one constraint*: To tackle the limitations brought by directly dropping the rank-one constraint and with the goal of improving the performance loss of the SDR, we design a DC representation for the rank-one constraint. This representation guarantees to have a solution that satisfies the non-convex rank-one constraint if the phase-shift optimization problem is feasible. It is worth mentioning that for a Positive semidefinite (PSD) matrix $\boldsymbol{V} \in \mathbb{C}^{M+1\times M+1}$, the rank-one constraint means that $\sigma_1(\boldsymbol{V}) > 0$ and $\sigma_m(\boldsymbol{V}) = 0, \forall m = 2, \ldots, M+1$, where $\sigma_m(\boldsymbol{V})$ is the mth largest singular value of $\boldsymbol{V}$. Consequently, the rank-one constraint can be expressed through the following proposition.

***Proposition* 1.** *For a PSD* $\boldsymbol{V} \in \mathbb{C}^{M+1\times M+1}$ *and* $\mathrm{tr}(\boldsymbol{V}) > 0$*, we can have*

$$\mathrm{rank}(\boldsymbol{V}) = 1 \Leftrightarrow \|\boldsymbol{V}\|_* - \|\boldsymbol{V}\|_2 = 0, \tag{5.32}$$

where $\|\boldsymbol{V}\|_2 = \sigma_1(\boldsymbol{V})$ *and* $\|\boldsymbol{V}\|_* = \sum_{m=1}^{M+1} \sigma_m(\boldsymbol{V})$ *are the spectral norm and the nuclear norm of* $\boldsymbol{V}$*, respectively.*

Proof: The complete proof can be found in [142, Proposition 3].

Since the norm is a convex function, one can deduce that the function $\boldsymbol{V} \mapsto \|\boldsymbol{V}\|_* - \|\boldsymbol{V}\|_2$ is a DC function. Therefore, by applying this DC representation to the phase-shift matrix feasibility check problem, we can reformulate the original rank-one constraint as the difference between the nuclear norm and the spectral norm. As a result, problem $\mathcal{P}$ can be reformulated as

$$\begin{aligned} \mathcal{P}1 : \min_{\boldsymbol{V}} \quad & \|\boldsymbol{V}\|_* - \|\boldsymbol{V}\|_2 , \\ \text{s.t.} \quad & (5.30\text{b}) - (5.30\text{f}). \end{aligned} \tag{5.33}$$

One can see that when the objective function of $\mathcal{P}1$ reaches zero, a rank-one feasible solution can be obtained, which is denoted as $\boldsymbol{V}^*$. Therefore, a feasible solution for problem $\mathcal{P}$ can be obtained by applying Cholesky decomposition $\boldsymbol{V}^* = \bar{v}\bar{v}^H$. As stated above, $\mathcal{P}1$ is a DC problem. Thus, in the following step, we design an SCA algorithm for efficiently solving the DC optimization problem $\mathcal{P}1$.

Before we proceed in the solution procedures, we represent first the objective function of $\mathcal{P}1$ as a difference of strongly convex functions by adding quadratic terms to the convex functions while their difference remains unchanged. The main reason behind that is to guarantee the convergence for the SCA algorithm [142]. Particularly, the objective function of $\mathcal{P}1$ can be rewritten as follows,

$$\min_{\boldsymbol{V}} \quad \mathcal{F}(\boldsymbol{V}) = \|\boldsymbol{V}\|_* - \|\boldsymbol{V}\|_2 + \Delta_{\mathcal{C}}(\boldsymbol{V}), \tag{5.34}$$

where $\mathcal{C}$ is a convex cone that achieves the constraints of $\mathcal{P}1$ and $\Delta_{\mathcal{C}}(\boldsymbol{V})$ is an indicator function which is equal to zero when $\boldsymbol{V} \in \mathcal{C}$ and $+\infty$, otherwise. Then, in order to establish a convergence result for the DC program, $\mathcal{F}(\boldsymbol{V})$ can be rewritten as a difference of strongly convex functions, i.e., $\mathcal{F}(\boldsymbol{V}) = g(\boldsymbol{V}) - q(\boldsymbol{V})$, where $g(\boldsymbol{V}) = \|\boldsymbol{V}\|_* + \Delta_{\mathcal{C}}(\boldsymbol{V}) + \frac{\mu}{2}\|\boldsymbol{V}\|_F^2$ and $q(\boldsymbol{V}) = \|\boldsymbol{V}\|_2 + \frac{\mu}{2}\|\boldsymbol{V}\|_F^2$ [142]. Due to the added quadratic terms, $g(\boldsymbol{V})$ and $q(\boldsymbol{V})$ are μ-strongly convex functions. Hence, problem (5.34) ends up with minimizing the difference of strongly convex functions

$$\min_{\boldsymbol{V}} \quad \mathcal{F}(\boldsymbol{V}) = g(\boldsymbol{V}) - q(\boldsymbol{V}). \tag{5.35}$$

For complex domain $\boldsymbol{V}$, Wirtinger calculus should be applied for algorithm design [146]. In order to solve this non-convex DC program, a sequence of candidates is iteratively constructed to the primal solutions and dual solutions via applying SCA [147].

3) *SCA Algorithm for Rank-One Constrained Optimization Problem* (5.35)*:* According to Fenchel's duality [146], the dual problem of (5.35) can be given by

$$\min_{\boldsymbol{Z}} \quad \mathcal{F}(\boldsymbol{Z}) = q^*(\boldsymbol{Z}) - g^*(\boldsymbol{Z}), \tag{5.36}$$

where $q^*(\boldsymbol{Z})$ and $g^*(\boldsymbol{Z})$ are the conjugate functions of $q(\boldsymbol{Z})$ and $g(\boldsymbol{Z})$, respectively. The conjugate function $g^*(\boldsymbol{Z})$ can be defined as $g^*(\boldsymbol{Z}) = \sup_{\boldsymbol{V}} \langle \boldsymbol{V}, \boldsymbol{Z}\rangle - g(\boldsymbol{V})$ in which the inner product can be given as $\langle \boldsymbol{V}, \boldsymbol{Z}\rangle = \Re(\mathrm{tr}(\boldsymbol{V}^H\boldsymbol{Z}))$. In the rth iteration, the convex approximation of the primal and dual problems should be iteratively solved by linearizing the concave part as follow.

$$\begin{aligned}
\boldsymbol{Z}^{[r]} &= \arg\inf_{\boldsymbol{Z}} \; q^*(\boldsymbol{Z}) - [g^*(\boldsymbol{Z}^{[r-1]}) + \langle \boldsymbol{Z} - \boldsymbol{Z}^{[r-1]}, \boldsymbol{V}^{[r]}\rangle], \\
\boldsymbol{V}^{[r+1]} &= \arg\inf_{\boldsymbol{V}} \; g(\boldsymbol{V}) - [q(\boldsymbol{V}^{[r]}) + \langle \boldsymbol{V} - \boldsymbol{V}^{[r]}, \boldsymbol{Z}^{[r]}\rangle].
\end{aligned} \tag{5.37}$$

According to Fenchel bi-conjugation theorem [146], the solution of (5.37) can be rewritten as $\boldsymbol{Z}^{[r]} \in \partial_{\boldsymbol{V}^{[r]}} q$, where $\partial_{\boldsymbol{V}^{[r]}} q$ is the subgradient of q with respect to $\boldsymbol{V}$ at $\boldsymbol{V}^{[r]}$. Thus, $\boldsymbol{V}^{[r]}$ at the rth iteration can be obtained by solving the following optimization problem

$$\min_{\boldsymbol{V}} \; g(\boldsymbol{V}) - \langle \boldsymbol{V}, \partial_{\boldsymbol{V}^{[r-1]}} q(\boldsymbol{V}) \rangle,$$
$$\text{s.t. (5.30b)} - \text{(5.30f)}, \tag{5.38}$$

where $\boldsymbol{V}^{[r-1]}$ is the solution obtained at iteration $r-1$ and $\partial_{\boldsymbol{V}^{[r-1]}} q(\boldsymbol{V}) = \partial_{\boldsymbol{V}^{[r-1]}} \|\boldsymbol{V}\|_2 + \mu \boldsymbol{V}^{[r-1]}$. It is worth mentioning that $\partial_{\boldsymbol{V}^{[r-1]}} \|\boldsymbol{V}\|_2$ can be efficiently evaluated as $\boldsymbol{a}_1 \boldsymbol{a}_1^H$, where $\boldsymbol{a}_1$ is the eigenvector corresponding to the largest singular value $\sigma_1(\boldsymbol{V})$ [142]. Given an initial value of $\boldsymbol{V}^{[0]}$ and by iteratively solving (5.38) until the objective function $\mathcal{F}(\boldsymbol{V})$ in (5.35) becomes zero, we get an exact rank-one solution based on **Proposition 1**. A practical stopping criterion can also be adopted as $\|\boldsymbol{V}\|_* - \|\boldsymbol{V}\|_2 < \epsilon_{\text{DC}}$, where $\epsilon_{\text{DC}} > 0$ is a sufficiently small constant [142]. Note that, the convergence characteristic of the solution obtained by iteratively solving problem (5.38), i.e. steps (4)-(7) in **Algorithm 4**, can be presented in the following proposition.

***Proposition* 2.** *For any $r = 0, 1, 2, \ldots,$ the sequence $\{\boldsymbol{V}^{[r]}\}$ generated by iteratively solving problem* (5.38) *has the following properties.*

(i) The generated sequence $\{\boldsymbol{V}^{[r]}\}$ converges to a critical point of $\mathcal{F}(\boldsymbol{V})$ in (5.34) *from an arbitrary initial point $\boldsymbol{V}^{[0]}$ and the sequence $\{\mathcal{F}(\boldsymbol{V}^{[r]})\}$ is strictly decreasing and convergent.*

(ii) For any $r = 0, 1, 2 \ldots,$ we have

$$\text{Avg}\left[\{\left\|\boldsymbol{V}^{[i]} - \boldsymbol{V}^{[i+1]}\right\|_F^2\}_{i=0}^{r}\right] \leq \frac{\mathcal{F}(\boldsymbol{V}^{[0]}) - \mathcal{F}(\boldsymbol{V}^{\star})}{\mu(r+1)}, \tag{5.39}$$

where Avg[.] *represents the average value and $\mathcal{F}(\boldsymbol{V}^{\star})$ denotes the global minimum of $\mathcal{F}$.*

Proof: The complete proof can be found in [142, Proposition 5], which is based on [148, Proposition 1].

Hence, the proposed solution based on SCA can always reach a feasible $\boldsymbol{V}$ to problem $\text{PS} - \text{HD}$ that guarantees that the objective value of problem (5.34) converges to zero. This means that a feasible rank-one solution is obtained [142]. Finally, the steps of the proposed optimization scheme for RIS-enabled HD C-NOMA, which is referred to as "HD: AO-based SCA algorithm", are presented in detail in Algorithm 1. Specifically, Algorithm 1 optimizes the power allocation coefficients at both the BSs and the near NOMA user and the phase-shift at the RIS in an alternative manner. Based on Proposition 2, the convergence of the proposed AO-based SCA algorithm is guaranteed according to the

Algorithm 4: AO Algorithm for RIS-enabled HD C-NOMA

Input: $P_{\mathrm{BS}}, P_{\mathrm{n}}, \boldsymbol{Q}_{\mathrm{bn}}, \boldsymbol{Q}_{\mathrm{bf}}, \sigma_{\mathrm{n}}^2, \sigma_{\mathrm{f}}^2$ and $\mathtt{SINR}_{(2)}^{\mathrm{HD}}$

Initialize the phase-shift $\boldsymbol{\theta}_{(1)}^{\mathrm{HD},[0]}$ and set the iteration number $t = 1$.

repeat

Using the closed-form expressions in ***Theorem 5.2*** to find the optimal power allocation coefficients at the BS $(\alpha_{\mathrm{n}}^{\mathrm{HD},[t]}, \alpha_{\mathrm{f}}^{\mathrm{HD},[t]})$ and the optimal power fraction coefficient at $\mathrm{UE_n}$, i.e. $\beta^{\mathrm{HD},[t]}$ for given $\boldsymbol{\theta}_{(1)}^{\mathrm{HD},[t]}$.

Set r = 1.

while *The objective value of* $\mathcal{P}1 \geq \epsilon_{\mathrm{DC}}$ *and* $r \leq N_1$ **do**

Compute the subgradient $\partial_{\boldsymbol{V}^{[r-1]}} \|\boldsymbol{V}\|_2$ and obtain a solution $\boldsymbol{V}^{[r]}$ by solving (5.38).

Update $r = r + 1$.

end while

Obtain $\bar{\boldsymbol{v}}^{[t]}$ using Cholesky decomposition, where $\boldsymbol{V}^{[r]} = \bar{\boldsymbol{v}}^{[t]} \left(\bar{\boldsymbol{v}}^{[t]}\right)^H$.

Obtain phase-shift matrix $\boldsymbol{\Theta}_{(1)}^{\mathrm{HD}} = \mathrm{diag}\left(\left(\boldsymbol{v}^{[t]}\right)^H\right)$, where $\boldsymbol{v}^{[t]} = [\bar{\boldsymbol{v}}]_{1:M} / [\bar{\boldsymbol{v}}]_{M+1}$.

Update $t = t + 1$.

Until The decrease of the objective value in (5.19) is below a threshold $\epsilon > 0$ or the maximum number of iterations N_2 is reached.

following proposition.

***Proposition* 3.** *By applying the proposed AO-based SCA algorithm, the objective value of* $\mathrm{OPT} - \mathrm{HD}$ *decreases when the number of iterations increases until convergence.*

Proof: See Appendix C3.

5.6.4 Complexity Analysis

To quantify the computational complexity of the proposed AO-based SCA algorithm, i.e., Algorithm 1, the computational complexity of the power allocation optimization sub-problem, i.e., $\mathrm{PA} - \mathrm{HD}$, and the phase-shift matrix optimization sub-problem, i.e., $\mathrm{PS} - \mathrm{HD}$ need to be analyzed. First, for the power allocation solution, since we obtained closed-form expressions in ***Theorem 5.2***, the computational complexity of having the optimal total transmit power is approximately $\mathcal{O}(1)$. Regarding the phase-shift matrix sub-problem, it is a semi-definite programming (SDP) that can be solved using interior point methods [149]. Based on [149, Theorem 3.12], the order of the computational complexity of the SDP problem with k SDP constraints that includes an $a \times a$ PSD matrix is given by $\mathcal{O}\left(\sqrt{a}\log\left(1/\zeta\right)\left(ka^3 + k^2a^2 + k^3\right)\right)$, where $\zeta > 0$ is the solution accuracy. For the phase-shift optimization, we have $a = M + 1$ and $k = M + 4$. Hence, the approximate

complexity can be obtained as $\mathcal{O}\left(N_1 \log\left(1/\zeta\right) M^{4.5}\right)$, where N_1 is the maximum number of required iterations until the decrease in the objective value in (5.35) is below the adopted threshold. As a result, the approximate overall complexity for solving $\mathrm{OPT}-\mathrm{HD}$ is $\mathcal{O}\left(N_2 N_1 \log\left(1/\zeta\right) M^{4.5}\right)$, where N_2 is the maximum number of iterations until convergence.

5.7 RIS-Enabled FD C-NOMA: Problem Formulation and Solution Approach

5.7.1 Problem Formulation

In this subsection, we investigate the power minimization problem for RIS-enabled FD C-NOMA systems. In contrast to the HD scenario that requires two time-slot and adjusts the RIS's configuration in each one of them, only one RIS's configuration is required in the FD case. By optimizing the power allocation coefficients at the BS $(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \alpha_{\mathrm{f}}^{\mathrm{FD}})$, the power fraction coefficient at UE_{n}, β^{FD}, and the phase-shift coefficients for the RIS $\boldsymbol{\theta}^{\mathrm{FD}}$, the total transmit power minimization problem for the proposed RIS-enabled FD C-NOMA framework can be formulated as follows.

$$\mathrm{OPT}-\mathrm{FD}: \min_{\substack{\boldsymbol{\theta}^{\mathrm{FD}}, \alpha_{\mathrm{n}}^{\mathrm{FD}}, \\ \alpha_{\mathrm{f}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}}} \left(\alpha_{\mathrm{n}}^{\mathrm{FD}} + \alpha_{\mathrm{f}}^{\mathrm{FD}}\right) P_{\mathrm{BS}} + \beta^{\mathrm{FD}} P_{\mathrm{n}}, \tag{5.40a}$$

$$\text{s.t. } 0 \leq \alpha_{\mathrm{n}}^{\mathrm{FD}} \leq \alpha_{\mathrm{f}}^{\mathrm{FD}}, \tag{5.40b}$$

$$0 \leq \alpha_{\mathrm{n}}^{\mathrm{FD}} + \alpha_{\mathrm{f}}^{\mathrm{FD}} \leq 1, \tag{5.40c}$$

$$0 \leq \beta^{\mathrm{FD}} \leq 1, \tag{5.40d}$$

$$\mathcal{R}_{\mathrm{n}\longrightarrow\mathrm{n}}^{\mathrm{FD}}\left(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}, \boldsymbol{\theta}^{\mathrm{FD}}\right) \geq R_{\mathrm{n}}^{\mathrm{th}}, \tag{5.40e}$$

$$\min\left(\mathcal{R}_{\mathrm{MRC}}^{\mathrm{FD}}\left(\mathbf{p}^{\mathrm{FD}}, \boldsymbol{\theta}^{\mathrm{FD}}\right), \mathcal{R}_{\mathrm{n}\longrightarrow\mathrm{f}}^{\mathrm{FD}}\left(\mathbf{p}^{\mathrm{FD}}, \boldsymbol{\theta}^{\mathrm{FD}}\right)\right) \geq R_{\mathrm{f}}^{\mathrm{th}}, \tag{5.40f}$$

$$|\phi_m| = 1, \qquad \forall\, m \in [\![1, M+1]\!]. \tag{5.40g}$$

Similar to $\mathrm{OPT}-\mathrm{HD}$, $\mathrm{OPT}-\mathrm{FD}$ is hard to be solved by common standard optimization techniques. Therefore, we resort to the AO technique similar to the RIS-enabled HD C-NOMA case in solving problem $\mathrm{OPT}-\mathrm{FD}$. Thus, the power control optimization problem with a fixed $\boldsymbol{\theta}^{\mathrm{FD}}$ can be written as

$$\mathrm{PC}-\mathrm{FD}: \min_{\alpha_{\mathrm{n}}^{\mathrm{FD}}, \alpha_{\mathrm{f}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}} \left(\alpha_{\mathrm{n}}^{\mathrm{FD}} + \alpha_{\mathrm{f}}^{\mathrm{FD}}\right) P_{\mathrm{BS}} + \beta^{\mathrm{FD}} P_{\mathrm{n}}, \tag{5.41a}$$

$$\text{s.t. (5.40b)} - \text{(5.40f)}, \tag{5.41b}$$

whereas the passive beamforming optimization problem with given power allocation coefficients, i.e., $\alpha_{\rm n}^{\rm FD}, \alpha_{\rm f}^{\rm FD}$ and $\beta^{\rm FD}$ can be presented as

$$\mathrm{PS-FD: Find}\ \ \boldsymbol{\theta}^{\rm FD}, \tag{5.42a}$$

$$\text{s.t. (5.40e)} - \text{(5.40g)}. \tag{5.42b}$$

5.7.2 RIS-enabled FD C-NOMA: Power Allocation Optimization

In this part, we start by determining the feasibility conditions of problem $\mathrm{PC-FD}$. Assuming that the phase-shift matrix $\boldsymbol{\Theta}^{\rm FD}$ is fixed, we denote by $\gamma_{\rm SI} \triangleq \frac{P_{\rm n}\gamma_{\rm SI}}{\sigma_{\rm n}^2}$ and by

$$\begin{aligned}
\gamma_{\rm bn} &\triangleq \frac{P_{\rm BS}|h_{\rm bn} + \boldsymbol{h}_{\rm rn}^H \boldsymbol{\Theta}^{\rm FD} \boldsymbol{h}_{\rm br}|^2}{\sigma_{\rm n}^2}, \\
\gamma_{\rm bf} &\triangleq \frac{P_{\rm BS}|h_{\rm bf} + \boldsymbol{h}_{\rm rf}^H \boldsymbol{\Theta}^{\rm FD} \boldsymbol{h}_{\rm br}|^2}{\sigma_{\rm f}^2}, \\
\gamma_{\rm d} &\triangleq \frac{P_{\rm n}|h_{\rm nf} + \boldsymbol{h}_{\rm rf}^H \boldsymbol{\Theta}^{\rm FD} \boldsymbol{h}_{\rm nr}|^2}{\sigma_{\rm f}^2}.
\end{aligned} \tag{5.43}$$

Based on this, the feasibility conditions of problem $\mathrm{PC-FD}$ are presented in the following theorem.

Theorem 5.3. *Problem* $\mathrm{PC-FD}$ *is feasible if and only if the following conditions hold.*

$$\textit{Condition 1: } \beta_{\min}^{\rm FD} \le \beta_{\max}^{\rm FD}, \tag{5.44a}$$

$$\textit{Condition 2: } \frac{\gamma_{\rm SI} t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}} \beta_{\min}^{\rm FD} + \frac{t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}} \le \frac{1}{2}, \tag{5.44b}$$

where $\beta_{\min}^{\rm FD}$ *and* $\beta_{\max}^{\rm FD}$ *are expressed, respectively, as*

$$\begin{aligned}
\beta_{\min}^{\rm FD} &= \max\left(0, \frac{\gamma_{\rm bf} - t_{\rm f}^{\rm FD} - \frac{\gamma_{\rm bf}\left(1+t_{\rm f}^{\rm FD}\right)t_{\rm n}^{\rm FD}}{\gamma_{\rm n}}}{c_1 - c_2}\right), \\
\beta_{\max}^{\rm FD} &= \min\left(1, \frac{\gamma_{\rm n} - t_{\rm f}^{\rm FD} - t_{\rm n}^{\rm FD}\left(1 + t_{\rm f}^{\rm FD}\right)}{\gamma_{\rm SI} t_{\rm n}^{\rm FD}\left(1 + t_{\rm f}^{\rm FD}\right) + t_{\rm f}^{\rm FD}}\right),
\end{aligned} \tag{5.45}$$

if $c_1 < c_2$ *and*

$$\beta_{\min}^{\rm FD} = 0, \quad \textit{and} \quad \beta_{\max}^{\rm FD} = \min\left(1, \frac{\gamma_{\rm n} - t_{\rm f}^{\rm FD} - t_{\rm n}^{\rm FD}\left(1 + t_{\rm f}^{\rm FD}\right)}{\gamma_{\rm SI} t_{\rm n}^{\rm FD}\left(1 + t_{\rm f}^{\rm FD}\right) + t_{\rm f}^{\rm FD}}, \frac{\gamma_{\rm bf} - t_{\rm f}^{\rm FD} - \frac{\gamma_{\rm bf}\left(1+t_{\rm f}^{\rm FD}\right)t_{\rm n}^{\rm FD}}{\gamma_{\rm n}}}{c_1 - c_2}\right), \tag{5.46}$$

if $c_1 > c_2$, *such that* $c_1 = \frac{\gamma_{\rm bf}}{\gamma_{\rm n}}\left(1 + t_{\rm f}^{\rm FD}\right) t_{\rm n}^{\rm FD} \gamma_{\rm SI}$, $c_2 = \gamma_{\rm d}$, $t_{\rm n}^{\rm FD} = 2^{R_{\rm n}^{\rm th}} - 1$ *and* $t_{\rm f}^{\rm FD} = 2^{R_{\rm n}^{\rm th}} - 1$.

Proof. See Appendix C4. ■

Before continuing with the derivation of optimal power control, let us define the following quantities. Let $\beta_c^{\rm FD}$, $\alpha_c^{\rm FD}$, $\alpha_{\min}^{\rm FD}$ and $\alpha_{\max}^{\rm FD}$ be defined, respectively, as

$$\beta_{\rm c}^{\rm FD} \triangleq \frac{t_{\rm f}^{\rm FD}\left(1-\frac{\gamma_{\rm bn}}{\gamma_{\rm bf}}\right)}{\frac{\gamma_{\rm bn}}{\gamma_{\rm bf}}\gamma_{\rm SI}t_{\rm f}^{\rm FD}+\gamma_d},$$
$$\alpha_{\rm c}^{\rm FD} \triangleq \frac{\gamma_{\rm SI}t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}}\beta_{\rm c}^{\rm FD}+\frac{t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}},$$
$$\alpha_{\rm min}^{\rm FD} \triangleq \frac{\gamma_{\rm SI}t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}}\beta_{\rm min}^{\rm FD}+\frac{t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}},$$
$$\alpha_{\rm max}^{\rm FD} \triangleq \frac{\gamma_{\rm SI}t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}}\beta_{\rm max}^{\rm FD}+\frac{t_{\rm n}^{\rm FD}}{\gamma_{\rm bn}}, \tag{5.47}$$

and let $\alpha_0^{\rm FD}$ and $\beta_0^{\rm FD}$ be the quantities defined, respectively, as

$$\left(\alpha_0^{\rm FD},\beta_0^{\rm FD}\right)=\begin{cases}\left(\alpha_{\rm min}^{\rm FD},\beta_{\rm min}^{\rm FD}\right), & \text{if } \beta_c^{\rm FD}<\beta_{\rm min}^{\rm FD},\\ \left(\alpha_{\rm c}^{\rm FD},\beta_{\rm c}^{\rm FD}\right), & \text{if } \beta_c^{\rm FD}\in\left[\beta_{\rm min}^{\rm FD},\beta_{\rm max}^{\rm FD}\right],\\ \left(\alpha_{\rm max}^{\rm FD},\beta_{\rm max}^{\rm FD}\right), & \text{otherwise.}\end{cases} \tag{5.48}$$

Based on the above definitions, let $\mathcal{X}$ be the set defined as

$$\mathcal{X}=\begin{cases}\left\{\left(\alpha_{\rm min}^{\rm FD},\beta_{\rm min}^{\rm FD}\right),\left(\alpha_0^{\rm FD},\beta_0^{\rm FD}\right),\left(\alpha_{\rm max}^{\rm FD},\beta_{\rm max}^{\rm FD}\right)\right\}, & \text{if } \alpha_0^{\rm FD}\le\frac{1}{2} \text{ and } \alpha_{\rm max}^{\rm FD}\le\frac{1}{2},\\ \left\{\left(\alpha_{\rm min}^{\rm FD},\beta_{\rm min}^{\rm FD}\right),\left(\alpha_{\rm max}^{\rm FD},\beta_{\rm max}^{\rm FD}\right)\right\}, & \text{if } \alpha_0^{\rm FD}\le\frac{1}{2} \text{ and } \alpha_{\rm max}^{\rm FD}\ge\frac{1}{2},\\ \left\{\left(\alpha_{\rm min}^{\rm FD},\beta_{\rm min}^{\rm FD}\right),\left(\alpha_{\rm max}^{\rm FD},\beta_{\rm max}^{\rm FD}\right)\right\}, & \text{if } \alpha_0^{\rm FD}\ge\frac{1}{2} \text{ and } \alpha_{\rm max}^{\rm FD}\le\frac{1}{2},\\ \left\{\left(\alpha_{\rm min}^{\rm FD},\beta_{\rm min}^{\rm FD}\right)\right\}, & \text{if } \alpha_0^{\rm FD}\ge\frac{1}{2} \text{ and } \alpha_{\rm max}^{\rm FD}\ge\frac{1}{2}.\end{cases} \tag{5.49}$$

Let $L\in\{1,2,3\}$ be the number of elements of $\mathcal{X}$ (the cardinal of $\mathcal{X}$). In addition, let $\boldsymbol{\alpha}_{\rm n}^{\rm FD}$ and $\boldsymbol{\beta}_{\rm n}^{\rm FD}$ be the two $3\times L$ vectors, such that, for all $i\in[\![1,L]\!]$, $\left([\boldsymbol{\alpha}_{\rm n}^{\rm FD}]_i,[\boldsymbol{\beta}^{\rm FD}]_i\right)$, is the ith element of $\mathcal{X}$. Based on this, let $\boldsymbol{\alpha}_{\rm f}^{\rm FD}$ be the $1\times L$ vector defined as

$$[\boldsymbol{\alpha}_{\rm f}^{\rm FD}]_i=\max\left(t_{\rm f}^{\rm FD}[\boldsymbol{\alpha}_{\rm n}^{\rm FD}]_i+\frac{\gamma_{\rm SI}t_{\rm f}^{\rm FD}}{\gamma_{\rm bn}}[\boldsymbol{\beta}^{\rm FD}]_i+\frac{t_{\rm f}^{\rm FD}}{\gamma_{\rm bn}},t_{\rm f}^{\rm FD}[\boldsymbol{\alpha}_{\rm n}^{\rm FD}]_i-\frac{\gamma_{\rm d}}{\gamma_{\rm bf}}[\boldsymbol{\beta}^{\rm FD}]_i+\frac{t_{\rm f}^{\rm FD}}{\gamma_{\rm bf}}\right),$$
$$\forall\ i\in[\![1,L]\!]. \tag{5.50}$$

Afterwards, let $\mathcal{P}^{\rm FD}\triangleq\left\{\left([\boldsymbol{\alpha}_{\rm n}^{\rm FD}]_i,[\boldsymbol{\alpha}_{\rm f}^{\rm FD}]_i,[\boldsymbol{\beta}^{\rm FD}]_i\right)\middle|\, i\in[\![1,L]\!]\right\}$. To this end, based on the above and assuming that problem PC − FD is feasible, its optimal solution is given in the following theorem.

Theorem 5.4. *Assuming that problem* PC − FD *is feasible, i.e., conditions* (5.44a) *and* (5.44b) *hold, its optimal solution is given by*

$$\left(\alpha_{\rm n}^{\rm FD^*},\alpha_{\rm f}^{\rm FD^*},\beta^{\rm FD^*}\right)=\underset{\mathbf{p}^{\rm FD}\in\mathcal{P}^{\rm FD}}{\arg\min}\, f\left(\mathbf{p}^{\rm FD}\right). \tag{5.51}$$

Proof. See Appendix C5. ■

One can see that problem (5.51) is a brute force search over a finite set with a size of at most $L=3$ elements, which can be easily solved.

5.7.3 RIS-enabled FD C-NOMA: Phase-Shift Coefficients Optimization

Similar to the solution of PS − HD, we apply the same three procedures to find the RIS phase-shift matrix. We first reformulate PS − FD into a rank-one constrained optimization problem as follows: by defining $\boldsymbol{v} \triangleq [\phi_1, \ldots, \phi_M]$ and applying the change of variables $\boldsymbol{h}_{\mathrm{rn}}^H \boldsymbol{\Theta}^{\mathrm{FD}} \boldsymbol{h}_{\mathrm{br}} = \boldsymbol{v}^H \boldsymbol{\Phi}$, where $\boldsymbol{\Phi} = \mathrm{diag}(\boldsymbol{h}_{\mathrm{rn}}^H)\boldsymbol{h}_{\mathrm{br}} \in \mathbb{C}^{M\times 1}$, $\boldsymbol{h}_{\mathrm{rf}}^H \boldsymbol{\Theta}^{\mathrm{FD}} \boldsymbol{h}_{\mathrm{br}} = \boldsymbol{v}^H \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = \mathrm{diag}(\boldsymbol{h}_{\mathrm{rf}}^H)\boldsymbol{h}_{\mathrm{br}} \in \mathbb{C}^{M\times 1}$, and $\boldsymbol{h}_{\mathrm{rf}}^H \boldsymbol{\Theta}^{\mathrm{FD}} \boldsymbol{h}_{\mathrm{nr}} = \boldsymbol{v}^H \boldsymbol{\Xi}$, where $\boldsymbol{\Xi} = \mathrm{diag}(\boldsymbol{h}_{\mathrm{rf}}^H)\boldsymbol{h}_{\mathrm{nr}} \in \mathbb{C}^{M\times 1}$, we have $|h_{\mathrm{bn}} + \boldsymbol{h}_{\mathrm{rn}}^H \boldsymbol{\Theta}^{\mathrm{FD}} \boldsymbol{h}_{\mathrm{br}}|^2 = |h_{\mathrm{bn}} + \boldsymbol{v}^H \boldsymbol{\Phi}|^2$, $|h_{\mathrm{bf}} + \boldsymbol{h}_{\mathrm{rf}}^H \boldsymbol{\Theta}^{\mathrm{FD}} \boldsymbol{h}_{\mathrm{br}}|^2 = |h_{\mathrm{bf}} + \boldsymbol{v}^H \boldsymbol{\Psi}|^2$, and $|h_{\mathrm{nf}} + \boldsymbol{h}_{\mathrm{rf}}^H \boldsymbol{\Theta}^{\mathrm{FD}} \boldsymbol{h}_{\mathrm{nr}}|^2 = |h_{\mathrm{nf}} + \boldsymbol{v}^H \boldsymbol{\Xi}|^2$. Then, PS − FD can be transformed into

$$\hat{\mathcal{P}}: \text{Find } \boldsymbol{\theta}^{\mathrm{FD}}, \tag{5.52a}$$

$$\text{s.t. } \alpha_{\mathrm{n}}^{\mathrm{FD}} P_{\mathrm{BS}} \left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bn}} \boldsymbol{V}\right) + |h_{\mathrm{bn}}|^2\right) \geq t_{\mathrm{n}}^{\mathrm{FD}} \left(\beta^{\mathrm{FD}} P_{\mathrm{n}} \gamma_{\mathrm{SI}} + \sigma_{\mathrm{n}}^2\right), \tag{5.52b}$$

$$\mathrm{tr}\left(\boldsymbol{Q}\boldsymbol{V}\right) + \alpha_{\mathrm{f}}^{\mathrm{FD}} P_{\mathrm{BS}} |h_{\mathrm{bf}}|^2 + \beta^{\mathrm{FD}} P_{\mathrm{n}} |h_{\mathrm{nf}}|^2 \geq t_{\mathrm{f}}^{\mathrm{FD}} \left(\alpha_{\mathrm{n}}^{\mathrm{FD}} P_{\mathrm{BS}} \left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bf}} \boldsymbol{V}\right) + |h_{\mathrm{bf}}|^2\right) + \sigma_{\mathrm{f}}^2\right), \tag{5.52c}$$

$$\alpha_{\mathrm{f}}^{\mathrm{FD}} P_{\mathrm{BS}} \left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bn}} \boldsymbol{V}\right) + |h_{\mathrm{bn}}|^2\right) \geq t_{\mathrm{f}}^{\mathrm{FD}} \left(\alpha_{\mathrm{n}}^{\mathrm{FD}} P_{\mathrm{BS}} \left(\mathrm{tr}\left(\boldsymbol{Q}_{\mathrm{bn}} \boldsymbol{V}\right) + |h_{\mathrm{bn}}|^2\right) + \beta^{\mathrm{FD}} P_{\mathrm{n}} \gamma_{\mathrm{SI}} + \sigma_{\mathrm{f}}^2\right), \tag{5.52d}$$

$$\boldsymbol{V} \succeq 0, \tag{5.52e}$$

$$[\boldsymbol{V}]_{m,m} = 1, \qquad \forall\, m \in [\![1, M+1]\!], \tag{5.52f}$$

$$\mathrm{rank}(\boldsymbol{V}) = 1, \tag{5.52g}$$

where

$$\boldsymbol{Q}_{\mathrm{bn}} = \begin{bmatrix} \boldsymbol{\Phi}\boldsymbol{\Phi}^H & \boldsymbol{\Phi} h_{\mathrm{bn}}^H \\ h_{\mathrm{bn}} \boldsymbol{\Phi}^H & 0 \end{bmatrix}, \; \boldsymbol{Q}_{\mathrm{bf}} = \begin{bmatrix} \boldsymbol{\Psi}\boldsymbol{\Psi}^H & \boldsymbol{\Psi} h_{\mathrm{bf}}^H \\ h_{\mathrm{bf}} \boldsymbol{\Psi}^H & 0 \end{bmatrix}, \; \boldsymbol{Q}_{\mathrm{nf}} = \begin{bmatrix} \boldsymbol{\Xi}\boldsymbol{\Xi}^H & \boldsymbol{\Xi} h_{\mathrm{nf}}^H \\ h_{\mathrm{nf}} \boldsymbol{\Xi}^H & 0 \end{bmatrix}, \; \bar{\boldsymbol{v}} = \begin{bmatrix} \boldsymbol{v} \\ t \end{bmatrix}, \tag{5.53}$$

where $\boldsymbol{V} = \bar{\boldsymbol{v}}\bar{\boldsymbol{v}}^H$ and $\boldsymbol{Q} = \alpha_{\mathrm{f}}^{\mathrm{FD}} P_{\mathrm{BS}} \boldsymbol{Q}_{\mathrm{bf}} + \beta^{\mathrm{FD}} P_{\mathrm{n}} \boldsymbol{Q}_{\mathrm{nf}}$. Note that $\bar{\boldsymbol{v}}^H \boldsymbol{Q}_z \bar{\boldsymbol{v}} = \mathrm{tr}(\boldsymbol{Q}_z \bar{\boldsymbol{v}}\bar{\boldsymbol{v}}^H)$ for all $z \in \{\mathrm{bn}, \mathrm{bf}, \mathrm{nf}\}$. Then, we apply the proposed DC representation for rank-one constrained which is followed by the proposed SCA procedures to find the phase-shift matrix in the FD relaying mode. Note that the same procedures in **Algorithm 4** are used in solving problem OPT − FD, but by substituting ***Theorem 5.2*** by ***Theorem 5.4*** to find the optimal power allocation coefficients at the BS and the optimal power fraction coefficient at near NOMA user for the FD relaying mode. In addition, problem $\hat{\mathcal{P}}$ is solved instead of problem $\mathcal{P}$ in order to find the RIS phase-shift matrix. We denote the proposed scheme for RIS-enabled FD C-NOMA as "FD: AO-based SCA". Finally, it is

worth mentioning that the computational complexity as well as the convergence analysis of the RIS-enabled FD C-NOMA can be analyzed in the same way as the RIS-enabled HD C-NOMA scheme.

5.8 Results and Discussion

In this section, several numerical examples and simulation results are presented to examine the performance of the proposed schemes RIS-enabled FD C-NOMA and RIS-enabled HD C-NOMA networks. In order to validate the effectiveness of the proposed AO-based SCA algorithm in both HD and FD relaying modes, we compare its performance with the following baselines schemes

- *AO-based SDR Algorithm*: In this scheme, the SDR is applied on the phase-shift optimization problem by simply dropping the rank-one constraint. If the obtained solution of the relaxed problem is not rank-one, the GR method is used to construct a rank-one solution. The power allocation coefficients are obtained through the derived closed-form expressions. We denote this scheme as "HD: AO-based SDR" in the HD relaying mode and as "FD: AO-based SDR" in the FD relaying mode. The complete analysis of SDR can be found in Appendix C6.

- *FD C-NOMA without RIS* [69, 107]: In this scheme, the BS communicates with $\mathrm{UE_n}$ and $\mathrm{UE_f}$ and $\mathrm{UE_n}$ communication with $\mathrm{UE_f}$ in the same time-slot without any assistance from the RIS. Hence, with the same target objective (power consumption minimization), we only need to use the power allocation coefficients at the BS and $\mathrm{UE_n}$ to minimize the total transmit power, which are derived in ***Theorem 5.4*** considering only the direct channel gains (BS $\longrightarrow$ $\mathrm{UE_n}$, BS $\longrightarrow$ $\mathrm{UE_f}$, and $\mathrm{UE_n}$ $\longrightarrow$ $\mathrm{UE_f}$) in the derived closed-form expressions.

- *Random Phase-shift*: In this method, the phase-shifts of the reflecting element are independent and generated uniformly within $[0, 2\pi]$. In addition, the power allocation coefficients are obtained through the derived closed-form expressions.

- *RIS with NOMA* [131]: The RIS is deployed to assist the transmission from the BS to two NOMA users without any cooperation between the near and the far

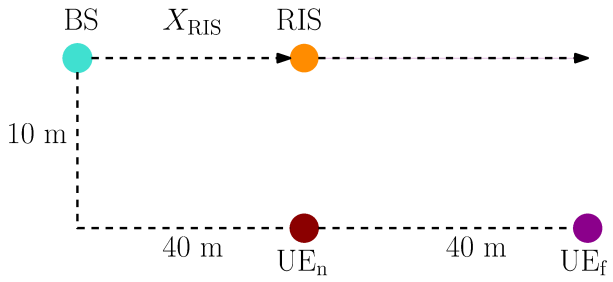


Figure 5.3: Simulation environment.

NOMA users. We adopt the same approach proposed in [131], but with the goal of minimizing the total transmit power from the BS.

5.8.1 Simulation Settings

The simulation environment consists of one BS, one RIS, one UE_n, and one UE_f which are located in a 3-dimensional Cartesian coordinates systems (X, Y, Z) at $(0\,\text{m}, 10\,\text{m}, 0\,\text{m})$, $(80\,\text{m}, 10\,\text{m}, 0\,\text{m})$, $(40\,\text{m}, 0\,\text{m}, 0\,\text{m})$ and $(80\,\text{m}, 0\,\text{m}, 0\,\text{m})$, respectively, as shown in Fig. 5.3. The small-scale fading and the large-scale fading are considered both for each communication link. The large-scale fading, i.e., distance-dependent path-loss, is modeled as $PL(d_x) = \rho_0 \left(\frac{d_x}{d_0}\right)^{-\eta_x}$, where ρ_0 is the path-loss at a reference distance d_0, η_x is the path-loss exponent, and d_x is the distance between the end-to-end node in the xth link. Considering the communication links between the BS and the RIS and between the RIS and UE_f, LoS components are assumed to exist [127]. Thus, these communication links experience small-scale fading that are modeled as Rician fading. The corresponding channel coefficients for these communication links can be expressed as, $\boldsymbol{h}_x = \sqrt{PL(d_x)}\left(\sqrt{\frac{1}{1+\kappa_x}}\mathbf{g}_x + \sqrt{\frac{\kappa_x}{1+\kappa_x}}\hat{\mathbf{g}}_x\right), \forall\, x \in \{\text{br}, \text{rf}\}$, where κ_x is the Rician factor, $\hat{\mathbf{g}}_x$ denotes the deterministic LoS component and $\mathbf{g}_x$ represents the non line-of-sight (NLoS) component, which follows a Rayleigh distribution with mean zero and variance one. On the other hand, $\forall y \in \{\text{bn}, \text{bf}, \text{nf}\}$ and $z \in \{\text{rn}, \text{nr}\}$, the small scale fading of h_y and $\boldsymbol{h}_z$ are modeled as Rayleigh fading. Consequently, the corresponding channel coefficients can be expressed as $h_y = g_y\sqrt{PL(d_y)}, \forall\, y \in \{\text{bn}, \text{bf}, \text{nf}\}$, and $\boldsymbol{h}_z = \mathbf{g}_z\sqrt{PL(d_z)}, \forall\, z \in \{\text{rn}, \text{nr}\}$, respectively, where $g_y \in \mathbb{C}$ and $\mathbf{g}_z \in \mathbb{C}^{M\times 1}$ are the small-scale Rayleigh fading with zero mean and unit variance, and $PL(d_y)$ and $PL(d_z)$ are the large-scale path-losses. Note that, Monte-Carlo simulations are employed over 10^3 independent channel realizations. The system parameters for the simulations are listed in Table 5.2 [120, 127, 135].

Table 5.2: Simulation Parameters

Parameter	Symbol	Value
Rician factors of the BS-RIS and the RIS-UE_f links	$\kappa_{\text{br}}, \kappa_{\text{rf}}$	3 dB
Path-loss at reference distance of 1 m	ρ_0	-30 dB
Path-loss exponents for the BS-RIS and for the RIS-UE_f links	$\eta_{\text{br}}, \eta_{\text{rf}}$	2.2
Path-loss exponents for the BS-UE_f and for the UE_n-UE_f links	$\eta_{\text{bf}}, \eta_{\text{nf}}$	4
Path-loss exponent for the BS-UE_n link	η_{bn}	3.5
Path-loss exponent for the UE_n-RIS link	η_{nr}	3
Power budget at the BS	P_{BS}	46 dBm
Power budget at UE_n	P_{n}	23 dBm
Noise power at UE_n and UE_f	$\sigma_{\text{n}}^2, \sigma_{\text{f}}^2$	-90 dBm
Minimum rate QoS requirement for UE_n	R_{n}^{th}	1 bits/sec/Hz

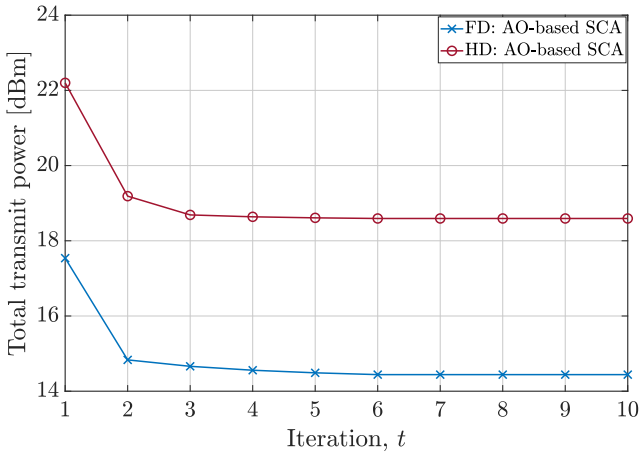


Figure 5.4: Algorithm convergence.

Fig. 5.4 describes the convergence behavior of the proposed RIS-enabled FD C-NOMA and RIS-enabled HD C-NOMA algorithms versus the iteration number with the number of RIS elements $M = 30$, SI parameter $\Omega_{\text{SI}} = -100$ dB, and $R_{\text{f}}^{\text{th}} = 2$ bits/sec/Hz. It can be observed that the proposed RIS-enabled FD C-NOMA and RIS-enabled HD C-NOMA algorithms converge in about 4 to 6 iterations. In addition, Fig. 5.5 depicts the computational time for both "FD: AO-based SCA" and "FD: AO-based SDR".[2] The run time of FD: AO-based SCA is higher than the one with SDR, i.e., FD: AO-based SDR. This is due to the iterations required by the DC algorithm to reach zero objective. Given the fact that there is a compromise between computational complexity and system performance, the AO-based SCA algorithm can achieve a better solution at the cost of a higher run time as we will see later. However, it is worth mentioning that owing to the advances of cloud computing with the introduction of C-RAN, the run time can be

[2]The algorithm was implemented in MATLAB using a machine with the following characteristics: System Type: x64-based PC, Processor: Intel(R) i7-4510U, CPU @2GHz, 8 Gigabyte RAM.

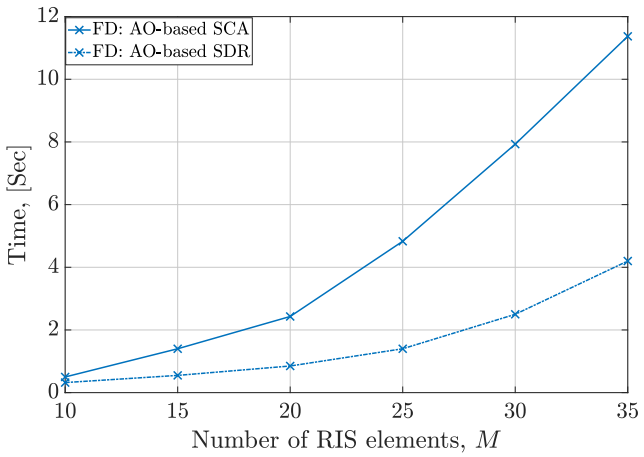


Figure 5.5: Computation time.

further decreased and hence the computational complexity can be easily tolerated [97].

5.8.2 Validation of The Closed-form Expressions for The Power Allocation Coefficients

It can be seen that the closed-form expressions in ***Theorem 5.2*** and ***Theorem 5.4*** are derived for a given phase-shift matrix. As a result, in order to validate the closed-form expression for the power allocation coefficients, we consider FD C-NOMA without RIS and HD- C-NOMA without RIS as the schemes that validate the analytical results. Fig. 5.6 depicts the analytical and numerical total transmit power for FD C-NOMA without RIS and HD C-NOMA without RIS schemes versus the minimum required rate for $\mathrm{UE_f}$, $R_{\mathrm{f}}^{\mathrm{th}}$ with $\Omega_{\mathrm{SI}} = -90$ dB. The analytical results are obtained based on the closed-from expressions derived in ***Theorem 5.2*** and ***Theorem 5.4***, while the numerical results are obtained by solving problem $\mathrm{PC}-\mathrm{HD}$ and $\mathrm{PC}-\mathrm{FD}$ using an off-the-shelf optimization solver. [3] It can be seen from Fig. 5.6 that the analytical results match perfectly the numerical results which validate the optimality of the closed-form expressions of the power allocation coefficients obtained by ***Theorem 2*** and ***Theorem 4***.

5.8.3 Effect of The number of RIS elements

Fig. 5.7 presents the total transmit power for the proposed schemes and the four baseline schemes versus the number of RIS reflecting elements. First, it can be seen that the total

[3]The adopted solver is fmincon that is a predefined MATLAB solver [11, 107, 150]. Moreover, 10^3 different initial points are generated to guarantee the convergence of the solver to the optimal solution.

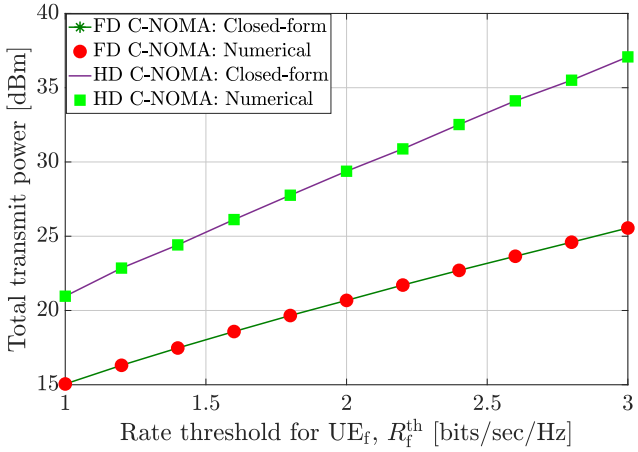


Figure 5.6: Analytical vs. numerical Results.

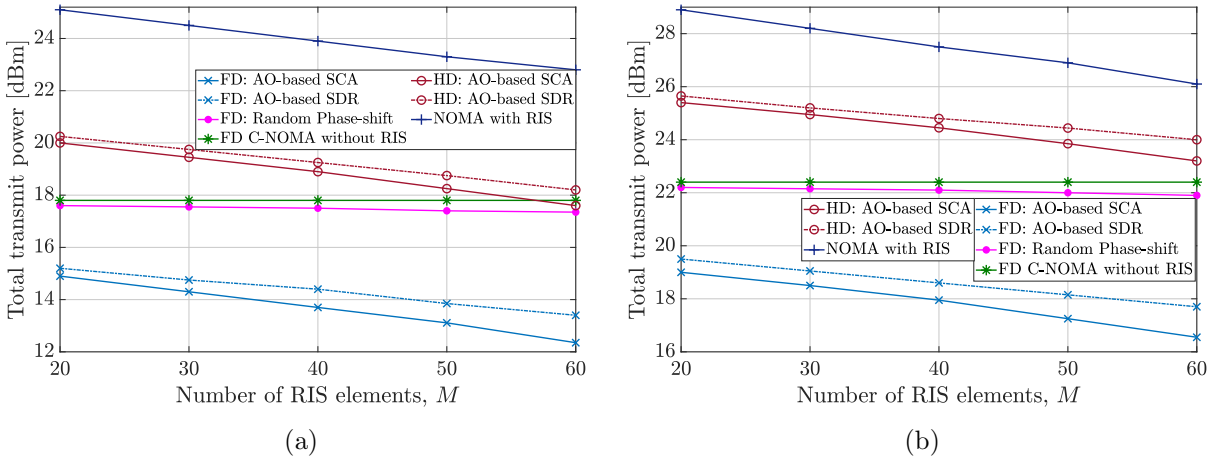


Figure 5.7: Total transmit power versus the number of RIS elements, when $\Omega_{SI} = -100$ dB, where (a) for the case when $R_f^{th} = 2$ bits/s/Hz and (b) for the case when $R_f^{th} = 3$ bits/s/Hz.

transmit power obtained by the proposed AO-based SCA approach is lower than the AO-based SDR approach. This is due to dropping rank-one constraint causing performance degradation; meanwhile, the proposed AO-based SCA approach can obtain an exact rank-one solution. In addition, when the number of RIS elements increases, the probability that the AO-based SDR approach fails to obtain a feasible solution is high and hence it early terminates the AO procedure. This explains why the proposed AO-based SCA approach achieves a better performance than the AO-based SDR approach when the number of elements increases. Second, it is observed that the total transmit power that is required by the RIS-based schemes decreases when the number of meta-atoms of the RIS increases while the total transmit power for FD C-NOMA without the RIS scheme remains unchanged. This is because a larger number of RIS meta-atoms leads to higher

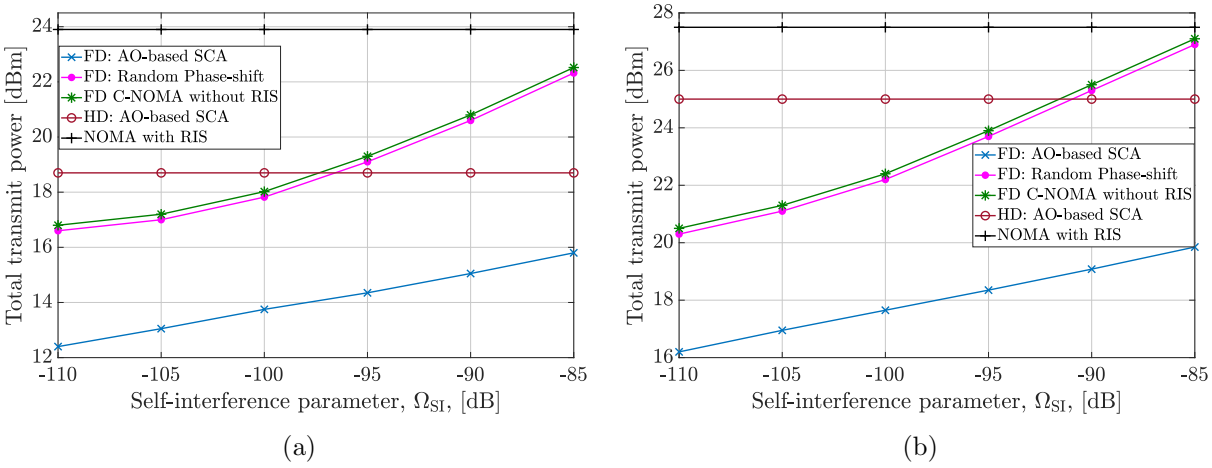
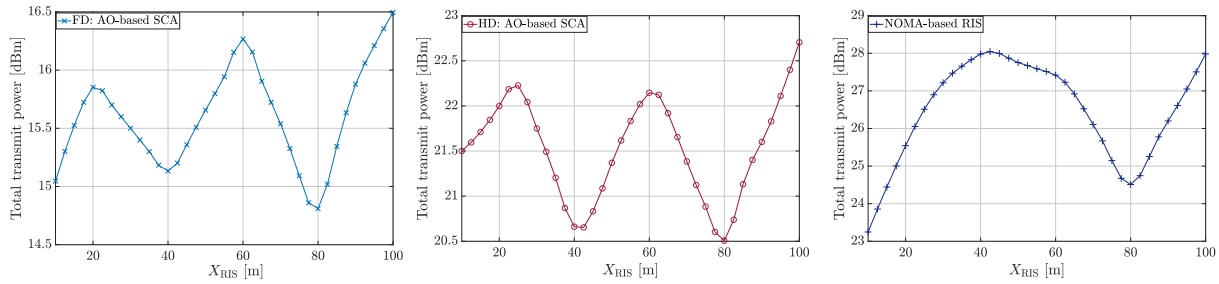


Figure 5.8: Total transmit power versus the SI parameter at UE_n, when the number of reflecting elements at the RIS is $M = 40$, where (a) for the case when $R_f^{th} = 2$ bits/s/Hz and (b) for the case when $R_f^{th} = 3$ bits/s/Hz.

combined channel gains and hence higher passive array gains. Third, it can be seen that the number of meta-atoms required at the RIS to allow the HD C-NOMA with RIS to beat the traditional FD C-NOMA depends on the required QoS at UE_f. This is because a high QoS requirement at UE_f needs a high passive array gain to overcome the pre-log penalty in the HD mode. Finally, the RIS-enabled FD C-NOMA network significantly outperforms the other schemes, which reveals the potential of integrating RIS in FD C-NOMA networks in enhancing the network power efficiency.

5.8.4 Effect of The SI channel

Fig. 5.8.a and Fig. 5.8.b depict the total transmit power versus the SI values at the near NOMA user. First, it can be observed that the proposed FD C-NOMA with RIS scheme gives a significant performance enhancement than both the HD C-NOMA with RIS and RIS-assisted NOMA schemes when Ω_{SI} is relatively small. However, as Ω_{SI} increases, the performance gain between them decreases. This is because increasing the SI value restricts the power transmission at UE_n and hence the BS should increase its transmit power in order to meet the QoS constraint at UE_f. Second, the HD C-NOMA with RIS scheme can achieve a significant performance compared to FD C-NOMA without RIS when Ω_{SI} increases. This is because increasing Ω_{SI} leads to deteriorating the performance of the FD mode and hence the passive array gain at the RIS can make the HD C-NOMA with RIS be a favorable scheme compared to the FD C-NOMA without RIS scheme. Third, it



(a) (b) (c)

Figure 5.9: (a), (b), and (c) illustrates the impact of the RIS location on the total power transmit for RIS-assisted FD C-NOMA, RIS-assisted HD C-NOMA, and RIS-assisted NOMA respectively, when $M = 30, \Omega_{\text{SI}} = -100\text{dB}, R_{\text{f}} = 2\text{bits/sec/Hz}$.

can be also observed that the total transmit power in the network without RIS is sharply increasing in comparison with the rate of increase in the network with RIS. This means that the system with RIS can tolerate high values of SI, which validates the effectiveness of the amalgamation between FD C-NOMA and RIS.

5.8.5 Effect of The RIS Location

Fig. 5.9.a, Fig. 5.9.b, and Fig. 5.9.c present the total transmit power versus the location of the RIS X_{RIS} for the RIS-assisted FD C-NOMA, the RIS-assisted HD C-NOMA, and the RIS-assisted NOMA, respectively. First, we set the coordinate of the RIS as $(X_{\text{RIS}}\text{ m}, 10\text{ m}, 0\text{ m})$. It has been shown in the literature that the best location for the RIS is either beside the BS or the UE of interest, which is the far NOMA user, in this case, to achieve a strong combined channel gain. Different from the literature, in the RIS-assisted FD C-NOMA, the RIS may be located as well beside the near NOMA user as shown in Fig. 5.9.a. This is basically due to two points: 1) the combined channel gain BS-RIS-UE_{n} at UE_{n} will be improved, and 2) the combined channel gain UE_{n}-RIS-UE_{f} will be also enhanced. Moreover, the best location in the FD relaying mode is near the far NOMA user UE_{f}. This is because the received `SINR` at the UE_{f} will be boosted with the assistance of the RIS. Meanwhile, in the HD C-NOMA, in order to make benefit from the RIS in the second slot, the RIS should be located at either the near NOMA user UE_{n} or the far NOMA user UE_{f}. In the case of the RIS-assisted NOMA scheme, it can be shown from Fig. 5.9.c that, in order to minimize the transmit power, the RIS should be located beside the BS so that both NOMA users get benefits from the RIS.

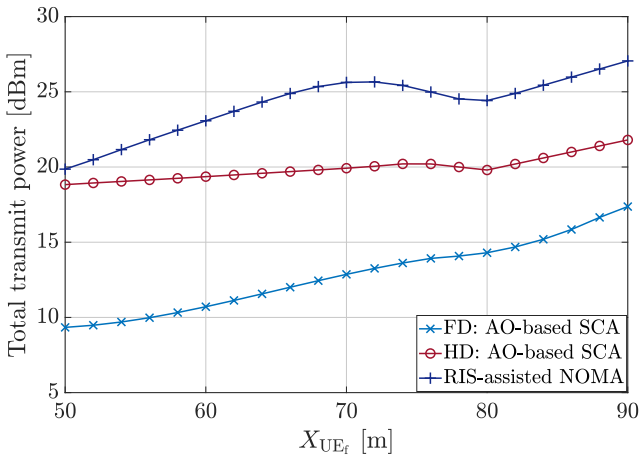


Figure 5.10: Illustrates the impact of the $\mathrm{UE_f}$ location on the total power transmit, when $M = 30, \Omega_{\mathrm{SI}} = -100\mathrm{dB}, R_{\mathrm{f}} = 2\mathrm{bits/sec/Hz}$.

5.8.6 Effect of The Location of The Far NOMA UE

Fig. 5.10 depicts the total transmit power versus the location $X_{\mathrm{UE_f}}$ of the far NOMA user $\mathrm{UE_f}$. We set the coordinate of the far NOMA user as $(X_{\mathrm{UE_f}}\ \mathrm{m}, 0\ \mathrm{m}, 0\ \mathrm{m})$ and of the RIS as $(80\ \mathrm{m}, 10\ \mathrm{m}, 0\ \mathrm{m})$. It can be seen that when the $X_{\mathrm{UE_f}}$ increases, the total transmit power increases, since the far NOMA user is moving away from the BS, and hence, its channel conditions get worse. Moreover, after a certain location of $\mathrm{UE_f}$, one can notice that the transmit power decreases for both the RIS-assisted HD C-NOMA and the RIS-assisted NOMA schemes and the rate of increase in the total power transmit for the RIS-assisted FD C-NOMA reduces. This is resulting from the improvement of the combined channel gain when the $\mathrm{UE_f}$ is moving towards the RIS.

5.8.7 Effect of The Required Rate QoS Threshold

Fig. 5.11 shows the effect of increasing the required data rate threshold for $\mathrm{UE_f}$ on the total transmit power. First, it can be seen that FD C-NOMA with RIS scheme has a significant gain compared to HD C-NOMA with RIS, RIS-assisted NOMA, and FD C-NOMA without RIS. Second, due to the pre-log penalty in the HD scenario, the gap between the FD C-NOMA with RIS and the HD C-NOMA with RIS increases when the required data rate threshold increases. Finally, it can be seen that the HD C-NOMA with RIS has the ability to beat the FD C-NOMA without RIS proposed in [30, 111] in the low data rate requirements at $\mathrm{UE_f}$.

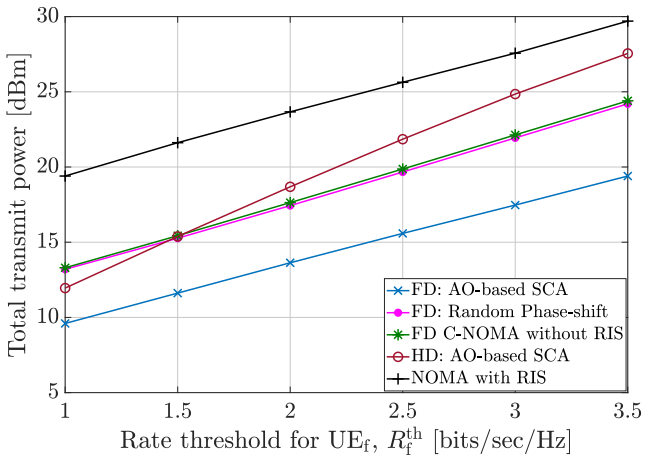


Figure 5.11: Total transmit power versus the rate threshold $R_{\mathrm{f}}^{\mathrm{th}}$ of $\mathrm{UE_f}$ when $M = 40$.

5.9 Summary

In this chapter, we investigated the RIS-empowered HD/FD C-NOMA downlink transmission scheme in order to minimize the total transmit power at the BS and at the relay user by jointly optimizing the power allocation coefficients at the BS and the transmit relaying power at the near NOMA user, along with the phase-shift coefficients at the RIS. By invoking the alternating optimization technique, the non-convex power minimization optimization problem is decomposed into two sub-problems, the power allocation optimization sub-problem and phase-shift optimization sub-problem, which are solved in an alternate manner. By leveraging the DC representation of rank-one constraint and the SCA approach, the RIS phase-shift coefficients are obtained. Meanwhile, for the power allocation sub-problem, we derived closed-form expressions for the optimal power allocation coefficients at the BS and the optimal transmit-relaying power from the near NOMA user. The Simulation results show that the proposed RIS-enabled FD C-NOMA scheme significantly outperforms both the FD C-NOMA without the assistance of the RIS and the RIS-assisted NOMA. In addition, in spite of the pre-log penalty in the HD relaying mode, and according to the required QoS at the far NOMA user, the number of reflecting elements, and the SI value, the proposed RIS-enabled HD C-NOMA can outperform the FD C-NOMA without RIS. Finally, the location of the RIS depends on the adopted access technique.

Chapter 6
RIS-Assisted CoMP NOMA Networks: Performance Analysis and Optimization

6.1 Introduction

Nevertheless, the gain of integrating RIS in wireless networks has been widely studied in the literature. For instance, the RIS has been utilized to significantly enhance the sum-rate for the CEUs in a multi-RIS wireless network [151], the network latency and reliability in a mobile edge computing system [152,153], the power consumption in a cooperative NOMA network [13], data collection for internet of things (IoT) networks [154], the max-min data rate in vehicular communications [155] and the information freshness in a wireless network [156,157]. Motivated by the benefits of RIS, CoMP and NOMA, we investigate, in this chapter, the potential enhancement brought by effectively integrating RIS with CoMP NOMA networks. Such a combination of RIS, NOMA, and CoMP technologies provides a promising paradigm for the upcoming 6G networks. This is due to the extra paths that can jointly construct a strong combined channel gain at the users-of-interests by leveraging the RIS, the improvement of the network connectivity and the spectral efficiency by invoking NOMA, and the mitigation of the ICI effects at the CEUs by adopting CoMP transmission.

6.2 State of The Art

6.2.1 NOMA-enabled CoMP transmission

With the quest of improving the performance of multi-cell networks, much attention was recently directed towards the integration between CoMP and NOMA [9,11,67,78,158,159].

The performance analysis of CoMP NOMA networks in terms of coverage probability and ergodic rate were studied in [9, 78] using tools from stochastic geometry. In doing so, the authors in [9, 78] assumed fixed PA between different NOMA clusters, which is not an optimal PA strategy. Meanwhile, in [67, 158], a power optimization problem was formulated with the goal of maximizing the network sum-rate while maintaining the required QoS at the users. However, these works considered only one cluster per cell. Moreover, the work in [159] investigated the joint resource and power allocation in CoMP NOMA network. The UC problem has been inherently considered by allowing the users that allocated the same radio resource to construct a NOMA cluster. However, the intertwined between the UC policies in different cells has not been considered.

6.2.2 RIS-assisted CoMP OMA networks

The authors in [160] studied RIS-enabled multi-cell network, in which a RIS was deployed to assist the CoMP transmission to multiple CEUs. Then, joint optimization of the active and passive beamforming to maximize the minimum user data rate was studied. On the other hand, joint optimization of the coordinated transmit beamforming at the BSs and the passive beamforming at the RIS to maximize the minimum weighted received SINR at the cellular users was investigated in [144]. The authors in [161] investigate a resource allocation design for RIS-assisted CoMP in the DL transmission cellular networks with underlying D2D communications. Specifically, the authors formulated an optimization problem with the goal of maximizing the network sum-rate by jointly designing user association, the active beamforming at the BSs, the passive beamforming at the RIS, and the transmit power of each D2D node while considering the QoS requirement of UEs, the power budget for both cellular UE and D2D pairs, and the backhaul capacity. Finally, a CoMP-assisted RIS framework is proposed in [162]. The main objective of this work is to maximize the network EE by jointly optimizing BS clustering, user association, radio resource assignment, PA, and passive beamforming while guaranteeing the users' QoS requirements.

6.2.3 RIS-assisted NOMA networks

In [163], the authors investigated the joint PA, PS matrix of the RIS, i.e., passive beamforming, and hybrid beamforming problem to maximize the sum-rate in RIS-assisted mil-

limeter wave in NOMA system. The joint active and passive beamforming for RIS-aided multiple two-user NOMA clusters was studied in [164] to minimize the total transmit power. A joint active and passive beamforming was investigated in [131] to maximize the minimum achievable rate for ensuring user fairness. However, the works [131, 163, 164] considered a RIS-assisted single-cell NOMA scenario in which the effect of the ICI has not been investigated. On the other hand, the work in [129] considered a multi-cell setup. There was not, however, any coordination between the BSs. Moreover, the performance analysis for the RIS-assisted NOMA system was investigated using tools from stochastic geometry in [132, 165, 166] in which a fixed and a non-optimal PA was considered. Note that, all the aforementioned works in [129, 131, 132, 144, 160–166] have adopted either the NOMA scheme or the CoMP scheme in RIS-assisted cellular networks.

6.2.4 RIS-assisted CoMP NOMA networks

The authors in [112] investigated the potential gain for integrating CoMP, NOMA, and RIS. Specifically, in [112], a joint optimization problem of the PA and passive beamforming was formulated to minimize the uplink power consumption in a two-cell RIS-aided NOMA system with joint detection CoMP. However, the works in [112] considered a simple setup consisting of only one CCU at each cell and one CEU in the overlapped area between the two cells. In fact, typical next-generations cellular systems are expected to serve large numbers of users. Nevertheless, in multi-user scenarios, the UC policy for the CoMP NOMA scheme has a great impact on the system performance, which should be carefully investigated.

Based on the above, the main research limitations and gaps can be presented as follows,

1. To the best of our knowledge, the performance analysis and the potential gains of RIS-assisted downlink CoMP NOMA cellular network have not been investigated in the literature.

2. Most of the works investigated the performance optimization of RIS-assisted NOMA networks assuming that each cell supports either a single NOMA cluster [126–128, 130,131,167] or pre-defined multi-cluster [138,163,164,168–170]. As a result, the UC policy is not investigated in these works [126–128, 130, 131, 138, 163, 164, 167–170].

3. In [129, 171], the authors studied the joint resource allocation, PA, and passive

beamforming for RIS-assisted NOMA system in which the users that allocated the same radio resource were grouped in one NOMA cluster. However, the work in [171] considered a single-cell setup; hence the effect of the ICI is not considered. On the other hand, the work in [129] assumed a multi-cell setup. However, there was not any coordination between the BSs. As a result, these works cannot be extended to the RIS-assisted CoMP NOMA network. This is due to the fact that the PA and the UC policy at the two BSs are intertwined.

4. The authors in [112] studied the performance optimization of the RIS-assisted CoMP NOMA network. However, this work considered a two-cell setup consisting of only one CCU at each cell and one CEU in the overlapped area between the two cells, which is not practical. In fact, typical next generations cellular systems are expected to serve large numbers of users. Nevertheless, in multi-user multi-user scenarios, the UC policy for the CoMP NOMA scheme has a great impact on the system performance, which should be carefully investigated.

5. To the best of our knowledge, the joint optimization of PA, UC policy, and PS matrix optimization has not been studied in the literature in the context of RIS-assisted CoMP NOMA cellular network, which is the main focus of the second part of this chapter.

6.3 Contributions

First, towards tackling the first point in the limitation and research gaps, we consider a system model, which is composed of two adjacent cells, each equipped with one BS, two RISs, and three UEs. Two types of UEs exist in this model: A CCU and a CEU. With the goal of enhancing the performance of the CEU, the PS matrix for each RIS can be tuned such that the reflected signals can be added constructively at the CEU. Against the above background, the performance analysis of RIS-assisted NOMA in multi-cell systems remains unexplored yet. The main contributions of this part can be summarized as follows,

- We investigate the integration between RIS and CoMP in a two-cell NOMA-based network with the goal of improving the performance of a CEU without affecting the performance of CCUs. While the CoMP is adopted to mitigate the ICI effects, the RIS is utilized to construct a stronger combined channel gain at the CEU.

- We derive first a closed-form expression for the ergodic rate of the CEU, and then we evaluate the overall NSE.

- We finally demonstrate the effectiveness of the proposed algorithm by comparing its performance with other multiple access schemes proposed in the literature.

Second, we extend the above-proposed model to a multi-user scenario with centralized-based RIS architecture. Against the above background and the aforementioned observations, the main contributions of this paper are summarized as follows.

- We investigate the joint problem of PA, UC, and RIS PS matrix, i.e. passive beamforming, in the downlink transmission of multi-user two-cell RIS-assisted CoMP in NOMA networks. This framework is formulated as an optimization problem with the goal of maximizing the network sum-rate while guaranteeing the required QoS for each user and the SIC constraints at the CCUs.

- The formulated problem turns out to be an MINLP problem, which is difficult to solve. To overcome this issue, we invoke the AO approach in which the original optimization problem is decomposed into two sub-problems, a joint PA and UC sub-problem and a PS sub-problem, that are solved in an alternating way. To the best of our knowledge, most of the works studying the performance of the CoMP NOMA network assumed that there was only one NOMA cluster per cell [11, 67, 158, 172]. Then, they developed dynamic PA schemes to achieve their objectives. In addition, the intertwined between the UC policies in different cells has not been considered in the literature for such a model. Different from [11,67,158,159,172], we propose and discuss a joint design of intertwined PA and UC. In doing so, and for a given PS matrix, the joint optimization problem of the PA and the UC is decomposed into a PA optimization sub-problem for a single cluster of users, referred to as the inner problem, and a UC optimization sub-problem, referred to as an outer problem, with the aid of the bi-level optimization.

- For the inner sub-problem, for each possible cluster, and as opposed to the sub-optimal PA solutions in [11, 67], the heuristic optimal PA solution in [158], and the high-complexity optimal PA solution in [67], the optimal PA coefficients are derived in closed-form expressions, which have a computational complexity of $\mathcal{O}(1)$.

Once obtaining the optimal PA for all possible cluster configurations, the outer sub-problem boils down to a *3-dimensional matching problem*. Due to the required QoS at the CEUs, the pairing of the CEU with the CCU from the first cell is intertwined with the CCU from the second cell that should be paired with that CEU, which is different from the traditional two-user paring in a single-cell NOMA network. To solve this problem, we project the outer sub-problem into three *2-dimensional matching problems*. Then, an *iterative Hungarian method* is proposed to find an efficient and low-complexity UC policy.

- For given PA coefficients and a UC policy, the RIS PS matrix optimization sub-problem is reformulated as a rank-one constrained optimization problem through change-of-variables and matrix lifting. Then, a DC representation for the rank-one constraint is designed. Finally, an efficient solution based on the SCA technique is proposed to obtain a feasible PS matrix for the RIS. Finally, the computational complexity for the overall proposed AO algorithm is analyzed.

We perform a thorough performance evaluation of the proposed scheme through various simulations. We also compare its performance to that of the CoMP NOMA scheme [9,67], which does not incorporate the RIS, the RIS-assisted NOMA network without considering the coordination between the BSs, and to that of the RIS-assisted CoMP OMA scheme [144, 160]. Numerical results demonstrate the efficacy of the proposed framework compared to these three benchmark schemes in terms of sum-rate performance.

To address the aforementioned research gaps introduced above, we divide this chapter into two main parts, which are mapped into two principal Sections 6.4 and 6.5, respectively. The first part, which addresses the performance analysis study for RIS-assisted CoMP NOMA networks, can be organized as follows. The network model, the signal model, the `SINR`, and the data rate analysis are presented in Section 6.4.1. In Section 6.4.2, we derive the ergodic rate for both the CCU and the CEU. Then, the numerical results for that model are discussed in Section 6.4.3. On the other hand, the second part of this Chapter addresses the resource management in RIS-assisted CoMP NOMA network is organized as follows. Section 6.5.1 presents the system model and rate analysis. Section 6.5.2 presents the formulated optimization problem and the proposed solution roadmap. Section 6.5.3 presents the proposed PA scheme and UC policy. Section 6.5.4 presents the PS optimization for the RIS. The simulation results are presented in Sections 6.5.7.

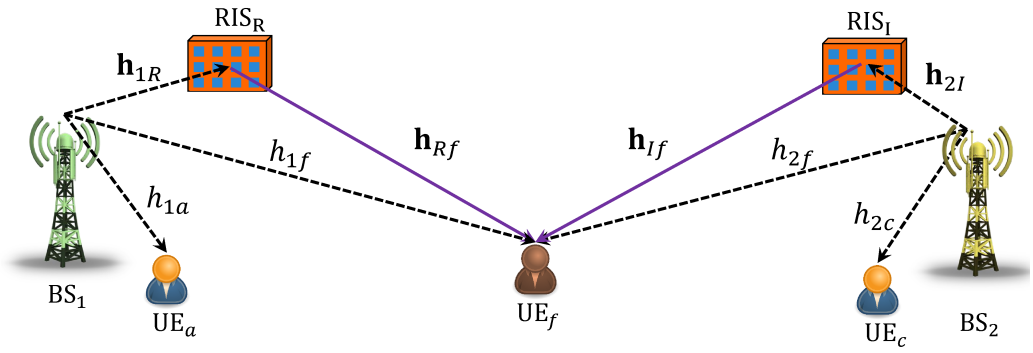


Figure 6.1: RIS-empowered CoMP NOMA cellular network.

Finally, we conclude the Chapter in Section 6.6.

6.4 RIS-assisted CoMP NOMA Network: Performance Analysis

6.4.1 System Model

6.4.1.1 Network Model

We consider a DL NOMA transmission in a two-cell system in which the RIS technology and the CoMP are integrated. As shown in Fig. 6.1, each cell consists of one BS and one CCU. In addition, one CEU is located at the overlapping area of the two cells. Due to the good channel gain with its nearby BS, we assume that each CCU is associated with one BS. We denote the CCUs attached to BS$_1$ and BS$_2$ as UE$_a$ and UE$_c$, respectively. On the other hand, the CEU, denoted by UE$_f$, is far from its serving BS, either BS$_1$ or BS$_2$, and experiences high ICI from the other BS. Thus, both BSs are cooperating together to serve this UE. In this model, we assume a two-user NOMA group within each cell, where each BS pairs its CCU with the CEU in one group. Thus, the CEU is a part of the two NOMA groups. Furthermore, as shown in Fig. 6.1, we assume that a RIS is located in the vicinity of each BS, where we denote by RIS$_R$ and RIS$_I$ the RIS associated with BS$_1$ and BS$_2$, respectively. The phase shifts for each RIS can be reshaped such that the reflected signals are constructively added at the CEU.

We assume that each BS and each UE is equipped with a single antenna, while each RIS is equipped with L reflecting elements [10, 132, 173]. In total, there are eight communication links in the considered system, namely BS$_1$ $\rightarrow$ UE$_a$, BS$_1$ $\rightarrow$ UE$_f$, BS$_1$ $\rightarrow$

RIS_R, $\text{RIS}_R \rightarrow \text{UE}_f$, $\text{BS}_2 \rightarrow \text{UE}_c$, $\text{BS}_2 \rightarrow \text{UE}_f$, $\text{BS}_2 \rightarrow \text{RIS}_I$ and $\text{RIS}_I \rightarrow \text{UE}_f$, whose channel coefficients are denoted as $h_{1a} \in \mathbb{C}, h_{1f} \in \mathbb{C}, \mathbf{h}_{1R} \in \mathbb{C}^{L\times 1}, \mathbf{h}_{Rf} \in \mathbb{C}^{L\times 1}, h_{2c} \in \mathbb{C}, h_{2f} \in \mathbb{C}, \mathbf{h}_{2I} \in \mathbb{C}^{L\times 1}$ and $\mathbf{h}_{If} \in \mathbb{C}^{L\times 1}$, respectively. We assume that all the channel amplitudes are independent and follow distinct Rayleigh distributions. Therefore, for all $q \in \{1a, 1f, 1R, Rf, 2c, 2f, 2I, If\}$, the channel gain $|h_q|^2$ follows an Exponential distribution with parameter λ_q. In addition, we assume that the average channel gain of each link is determined by the path-loss, that is, $\lambda_q = d_q^{-\eta}$, where d_q is the distance between two involved nodes and α is the path-loss exponent [33]. Finally, we denote by $\mathbf{\Phi}_s = \text{diag}\{e^{j\phi_{s,1}}, e^{j\phi_{s,2}}, \ldots, e^{j\phi_{s,L}}\} \in \mathbb{C}^{L\times L}$ the phase-shift matrix for RIS_s $s \in \{R, I\}$, where $\phi_{s,l}$ is the phase-shift for the lth reflecting element.

6.4.1.2 Signal Model and SINR Analysis

We first present the signal model at the CCUs, and then we discuss the signal model at the CEU. Following the NOMA principle, BS_1 and BS_2 broadcast superimposed signals to the UEs, that are given, respectively, as

$$\mathcal{Y}_1 = \sqrt{\alpha_1^a P_1} x_a + \sqrt{\alpha_1^f P_1} x_f, \tag{6.1a}$$

$$\mathcal{Y}_2 = \sqrt{\alpha_2^c P_2} x_c + \sqrt{\alpha_2^f P_2} x_f, \tag{6.1b}$$

where P_1 and P_2 are the transmit powers by BS_1 and BS_2, respectively, x_a, x_c and x_f are the signals intended to UE_a, UE_c and UE_f, respectively, α_1^a and α_2^c are the power control coefficients associated to UE_a and UE_c, respectively, and α_1^f and α_2^f are the power control coefficients of UE_f allocated by BS_1 and BS_2, respectively, such that $\alpha_1^a + \alpha_1^f \leq 1$ and $\alpha_2^c + \alpha_2^f \leq 1$. Let us consider first the downlink NOMA transmission from BS_1. According to the NOMA principle, since $|h_{1f}| \leq |h_{1a}|$, BS_1 assigns less power to the CCU than the CEU, i.e., $\alpha_1^f > \alpha_1^a$. We assume that the distance between the two RISs is sufficiently large so that the signal transmitted from one BS and reflected by the RIS of the other BS is negligible at the CEU and at its CCU due to the high path-loss [174]. Therefore, the received signal at UE_a is expressed as

$$\mathcal{Z}_a = (h_{1a} + \mathbf{h}_{1R}^T \mathbf{\Phi}_R \mathbf{h}_{Ra}) \mathcal{Y}_1 + w_a, \tag{6.2}$$

where w_a is the noise experienced at UE_a, which is $\mathcal{CN}(0, \sigma^2)$ distributed, such that σ^2 denotes the noise power. As explained in [173], the reflected signal from RIS_R towards UE_a is significantly weaker than the received signal from the direct path, and hence, it can be ignored at UE_a [173]. Based on this, after UE_a decodes the UE_f's message $x_f[l]$ and then cancels it from its reception, the received SINR of UE_a to decode its own message

is given by

$$\gamma_a = \alpha_1^a \rho_1 |h_{1a}|^2, \tag{6.3}$$

where $\rho_1 = P_1/\sigma^2$ is the transmit SNR from BS_1. Following the same analysis, the SINR at UE_c, denoted by γ_c, can be obtained. Due to their good channel gains from their serving BSs, we assume that the ICI at each CCU (the interference from BS_2 to UE_a or from BS_1 to UE_c) is negligible and thus can be ignored [108].

Considering UE_f, its received signal is expressed as

$$\mathcal{Z}_f = (h_{1f} + \mathbf{h}_{1R}^T \mathbf{\Phi}_R \mathbf{h}_{Rf})\mathcal{Y}_1 + (h_{2f} + \mathbf{h}_{2I}^T \mathbf{\Phi}_I \mathbf{h}_{If})\mathcal{Y}_2 + w_f, \tag{6.4}$$

where w_f is the noise at UE_f, which is $\mathcal{CN}(0, \sigma^2)$ distributed. Thus, the received SINR at UE_f to decode its own message can be written as

$$\gamma_f = \frac{\alpha_1^f \rho_1 (h_{1f} + \mathbf{h}_{1R}^T \mathbf{\Phi}_R \mathbf{h}_{Rf})^2 + \alpha_2^f \rho_2 (h_{2f} + \mathbf{h}_{2I}^T \mathbf{\Phi}_I \mathbf{h}_{If})^2}{\alpha_1^a \rho_1 (h_{1f} + \mathbf{h}_{1R}^T \mathbf{\Phi}_R \mathbf{h}_{Rf})^2 + \alpha_2^c \rho_2 (h_{2f} + \mathbf{h}_{2I}^T \mathbf{\Phi}_I \mathbf{h}_{If})^2 + 1}. \tag{6.5}$$

Similar to [132,138,175], the CSI related to the direct link and the cascaded link is assumed to be available at the RIS controller. Consequently, the maximum SINR at the CEU can be achieved when the phase-shift is selected as $\phi_{R,n} = \arg(h_{1f}) - \arg([\mathbf{h}_{1R}]_n [\mathbf{h}_{Rf})]_n)$ and $\phi_{I,n} = \arg(h_{2f}) - \arg([\mathbf{h}_{2I}]_n [\mathbf{h}_{If}]_n)$ for RIS_R and RIS_I, respectively. Therefore, the SINR at UE_f can be rewritten as,

$$\gamma_f = \frac{\alpha_1^f \rho_1 \Omega_1 + \alpha_2^f \rho_2 \Omega_2}{\alpha_1^a \rho_1 \Omega_1 + \alpha_2^c \rho_2 \Omega_2 + 1}, \tag{6.6}$$

where

$$\begin{aligned} \Omega_1 &\triangleq \left(|h_{1f}| + \sum\nolimits_{l=1}^{L} |[\mathbf{h}_{1R}]_l [\mathbf{h}_{Rf}]_l| \right)^2, \\ \Omega_2 &\triangleq \left(|h_{2f}| + \sum\nolimits_{l=1}^{L} |[\mathbf{h}_{2I}]_l [\mathbf{h}_{If}]_l| \right)^2. \end{aligned} \tag{6.7}$$

6.4.2 Spectral Efficiency Analysis

The NSE for the proposed RIS-empowered CoMP NOMA scheme is expressed as[1]

$$\mathcal{R} = \mathcal{R}_a + \mathcal{R}_c + \mathcal{R}_f, \tag{6.8}$$

where for $u \in \{a, c, f\}$, $\mathcal{R}_u$ denotes the ergodic rate of UE_u. In the following, we derive a closed-form expression of the ergodic rate for each UE.

The ergodic rate of a UE can be expressed as $\mathcal{R}_u = \mathbb{E}\left[\log_2(1 + \gamma_u)\right]$, for all $u \in \{a, c, f\}$ [33]. Since the performance of a CCU is not affected by the existence of RISs, its ergodic rate corresponds to the one of traditional NOMA. Thus, the ergodic rate for UE_a is given

[1]We focus in this part on studying the NSE. One may also use the NSE statistics provided in this part to consider other performance metrics such as reliability, coverage probability, and energy efficiency, which have been omitted for space limitation

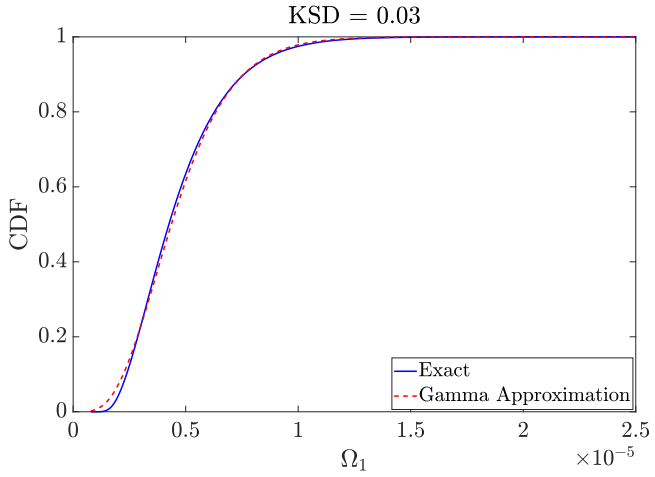


Figure 6.2: Comparison between the simulation and approximation results of the CDF of Ω_1.

by [33]

$$\mathcal{R}_a = \frac{-1}{\ln(2)} \exp\left(\frac{1}{\alpha_1^a \rho_1 \lambda_{1a}}\right) \mathrm{Ei}\left(\frac{-1}{\alpha_1^a \rho_1 \lambda_{1a}}\right), \tag{6.9}$$

where $\mathrm{Ei}(\cdot)$ denotes *Exponential integral function*. The ergodic rate of UE_c can be obtained similarly. Note that, the CCU is not forced to assist the CEU as in C-NOMA, and hence its performance is not affected by either the SI in FD C-NOMA or the pre-log penalty in HD C-NOMA due to the required two-time slots for transmission.

On the other hand, in order to evaluate the ergodic rate of UE_f, we first need to determine the distributions of the $\mathcal{RV}$s Ω_1 and Ω_2. In the following, we start by deriving the distribution of Ω_1. For $l \in [\![1, L]\!]$, the channel amplitude of the signal transmitted from BS_1, incident on the nth element of the RIS_R, and then reflected towards UE_f is subject to a double-Rayleigh fading. Let $\mathbf{h}_{1,R,f} \triangleq \mathbf{h}_{1R} \odot \mathbf{h}_{Rf}$, where $\odot$ denotes the element-wise multiplication. Therefore, for $l \in [\![1, L]\!]$, the mean and the variance of $\mathcal{RV}$ $[\mathbf{h}_{1,R,f}]_l$ can be expressed as $\mathbb{E}\{|[\mathbf{h}_{1,R,f}]_l|\} = \frac{\pi}{4}\sqrt{\lambda_{1R}\lambda_{Rf}}$ and $\mathrm{Var}\{|[\mathbf{h}_{1,R,f}]_l|\} = (1 - \frac{\pi^2}{16})\lambda_{1R}\lambda_{Rf}$ [122]. For a sufficiently large number of reflecting elements, i.e., $L >> 1$, and according to the central limit theorem (CLT), $h_{1,R,f} = \sum_{n=1}^{L} [\mathbf{h}_{1,R,f}]_l$ can be approximated with a Gaussian distributed $\mathcal{RV}$ with mean and variance [122]

$$\mathbb{E}\{h_{1,R,f}\} = \mu_{R,f} = \frac{L\pi}{4}\sqrt{\lambda_{1R}\lambda_{Rf}}, \tag{6.10}$$

$$\mathrm{Var}\{h_{1,R,f}\} = \sigma_{R,f}^2 = L(1 - \frac{\pi^2}{16})\lambda_{1R}\lambda_{Rf}. \tag{6.11}$$

Based on the above, it can be seen that $\sqrt{\Omega_1}$ is a sum of a Gaussian and a Rayleigh distributed $\mathcal{RV}$s, and therefore, obtaining its exact PDF is not straightforward. To overcome this issue, we determine alternatively an approximated expression for the PDF of Ω_1. We approximate the distribution of Ω_1 by a Gamma distribution, i.e., $\Omega_1 \dot{\sim} \Gamma(k_1, \theta_1)$.

In order to evaluate and validate the accuracy and the good fitness of this approximation, we have used the well-known Kolmogorov-Smirnov distance (KSD) [176]. Specifically, the KSD measures the absolute distance between two different cumulative distribution functions (CDFs) F_1 and F_2 [176], i.e.,

$$\text{KSD} = \max_{\Omega_1} |F_1(\Omega_1) - F_2(\Omega_1)|. \tag{6.12}$$

Obviously, smaller values of KSD correspond to more similarity between these distributions [176]. In our case, $F_1(\cdot)$ and $F_2(\cdot)$ represent the exact CDF obtained empirically from the simulations and the approximated CDF resulting from the Gamma distribution, respectively. For the considered simulation settings in Section IV with $L = 150$, the obtained KSD of the approximation problem in hand is 0.03, which validates the good approximation obtained by the use of the Gamma distribution. In addition, Fig. 6.2 presents the exact distribution of Ω_1 obtained empirically from the simulations and the approximated CDF resulting from the use of the Gamma distribution. Fig. 6.2 shows that the Gamma distribution matches perfectly the exact CDF of Ω_1, which demonstrates the good fitness of the proposed approximation. Hereafter, the shape and scale parameters k_1 and θ_1 can be obtained using the moment-matching technique, which is one of the most popular techniques for distribution approximation is [177]. Specifically, the exact first and second moments of Ω_1 are given, respectively, as

$$\begin{aligned}
\mathbb{E}\{\Omega_1\} &= \mathbb{E}\Big\{ \left(|h_{1f}| + |h_{1,R,f}|\right)^2 \Big\} \\
&= \left[\frac{L^2\pi^2}{16} + L\left(1 - \frac{\pi^2}{16}\right)\right] \lambda_{1R}\lambda_{Rf} + \frac{L\pi}{4}\sqrt{\pi\lambda_{1f}\lambda_{1R}\lambda_{Rf}} + \lambda_{1f}, \qquad (6.13) \\
\mathbb{E}\{\Omega_1^2\} &= \mathbb{E}\Big\{ \left(|h_{1f}| + |h_{1,R,f}|\right)^4 \Big\}, \\
&= \sum_{i=0}^{4} \binom{4}{i} \mathbb{E}\left\{|h_{1f}|^i\right\} \mathbb{E}\left\{|h_{1,R,f}|^{4-i}\right\}, \qquad (6.14)
\end{aligned}$$

where

$$\begin{aligned}
&\mathbb{E}\{|h_{1f}|\} = \sqrt{\pi\lambda_{1f}}/2, && \mathbb{E}\{|h_{1f}|^2\} = \lambda_{1f}, \\
&\mathbb{E}\{|h_{1f}|^3\} = 3\sqrt{\pi\lambda_{1f}^3}/4, && \mathbb{E}\{|h_{1f}|^4\} = 2\lambda_{1f}^2, \\
&\mathbb{E}\{|h_{1,R,f}|^2\} = \sigma_{R,f}^2 + \mu_{R,f}^2, && \qquad (6.15) \\
&\mathbb{E}\{|h_{1,R,f}|^3\} = \mu_{R,f}^3 + 3\mu_{R,f}\sigma_{R,f}^2, \\
&\mathbb{E}\{|h_{1,R,f}|^4\} = \mu_{R,f}^4 + 6\mu_{R,f}^2\sigma_{R,f}^2 + 3\sigma_{R,f}^4.
\end{aligned}$$

Consequently, using the moment-matching technique, we obtain

$$k_1 = \frac{\mathbb{E}\{\Omega_1\}^2}{\mathbb{E}\{\Omega_1^2\} - \mathbb{E}\{\Omega_1\}^2},$$

$$\theta_1 = \frac{\mathbb{E}\{\Omega_1^2\} - \mathbb{E}\{\Omega_1\}^2}{\mathbb{E}\{\Omega_1\}}. \tag{6.16}$$

Following the same procedures (6.10)-(6.15), Ω_2 can be also approximated with a Gamma distribution, i.e., $\Omega_2 \dot{\sim} \Gamma(k_2, \theta_2)$, where k_2 and θ_2 can be also obtained. At this point, since the distributions of Ω_1 and Ω_2 are determined, we can derive now the closed-from expression of the ergodic rate of the CEU. However, before we proceed, we present a related Fact and some necessary *lemmas* for the derivations.

***Fact* 1.** *For any $Y \sim \Gamma(k, \theta)$ and any scalar $\delta > 0$, $\delta Y \sim \Gamma(k, \delta\theta)$ [178].*

Lemma 6.1. *If $\{Y_i\}$ are independent Gamma distributed $\mathcal{RV}$s with shape k_i and scale θ_i, i.e., $Y_i \sim \Gamma(k_i, \theta_i)$, then $\sum_i Y_i$ can be approximated with a Gamma distributed $\mathcal{RV}$ with shape and scale parameters given, respectively, as*

$$k_y = \frac{\left(\sum_i k_i\theta_i\right)^2}{\sum_i k_i\theta_i^2} \quad \textit{and} \quad \theta_y = \frac{\sum_i k_i\theta_i^2}{\sum_i k_i\theta_i}. \tag{6.17}$$

Lemma 6.2. *If X is a $\mathcal{RV}$ with mean $\mathbb{E}\{X\}$ and variance $\mathrm{Var}\{X\}$, then $\mathbb{E}\{\ln(1+X)\}$ can be approximated as*

$$\mathbb{E}\{\ln(1+X)\} \approx \ln(1+\mathbb{E}\{X\}) - \frac{\mathrm{Var}\{X\}}{2(1+\mathbb{E}\{X\})^2}. \tag{6.18}$$

Proof. By using the Taylor approximation of the function $x \mapsto \ln(1+x)$ around a point $x_0 \in]-1, \infty[$, we get

$$\ln(1+x) \approx \ln(1+x_0) + \frac{x-x_0}{1+x_0} - \frac{(x-x_0)^2}{2(1+x_0)^2}. \tag{6.19}$$

Hence, by letting $x_0 = \mathbb{E}\{X\}$ and then applying the expectation for both sides of (6.19), we obtain

$$\mathbb{E}\{\ln(1+X)\} \approx \ln(1+\mathbb{E}\{X\}) - \frac{\mathrm{Var}\{X\}}{2(1+\mathbb{E}\{X\})^2}, \tag{6.20}$$

which completes the proof. ■

Theorem 6.1. *The ergodic rate of UE_f can be approximated as*

$$\mathcal{R}_f \approx \frac{1}{\ln 2}\left(\ln\left(\frac{1+k_{\Delta_1}\theta_{\Delta_1}}{1+k_{\Delta_2}\theta_{\Delta_2}}\right) - \frac{k_{\Delta_1}\theta_{\Delta_1}^2}{2(1+k_{\Delta_1}\theta_{\Delta_1})^2} + \frac{k_{\Delta_2}\theta_{\Delta_2}^2}{2(1+k_{\Delta_2}\theta_{\Delta_2})^2}\right), \tag{6.21}$$

where for all $j \in \{1, 2\}$,

$$k_{\Delta_j}(\omega_j) = \frac{\left(\sum\limits_{i=1}^{2} k_i\omega_j\theta_i\right)^2}{\sum\limits_{i=1}^{2} k_i(\omega_j\theta_i)^2}, \quad \textit{and} \quad \theta_{\Delta_j}(\omega_j) = \frac{\sum\limits_{i=1}^{2} k_i(\omega_j\theta_i)^2}{\sum\limits_{i=1}^{2} k_i\omega_j\theta_i}, \tag{6.22}$$

in which $\omega_1 = \rho_i$ and $\omega_2 = \beta_i^m\rho_i$.

Proof. The ergodic rate of UE_f can be obtained as

$$\begin{aligned}\mathcal{R}_f &= \mathbb{E}\{\log_2(1+\gamma_f)\},\\ &= \mathbb{E}\left\{\log_2\left(1 + \frac{\alpha_1^f\rho_1\Omega_1 + \alpha_2^f\rho_2\Omega_2}{\alpha_1^a\rho_1\Omega_1 + \alpha_2^c\rho_2\Omega_2 + 1}\right)\right\}.\end{aligned} \tag{6.23}$$

Table 6.1: Simulation Parameters [33, 175]

Parameter	Symbol	Value
Distance between BS_1 and UE_a and between BS_2 and UE_c	d_{1a}, d_{2c}	20m
Distance between BS_1 and UE_f and between BS_2 and UE_f	d_{1f}, d_{2f}	100m, 110m
Path-loss exponent for the wireless links	η	3
Power allocation coefficients assigned by BS_1 and BS_2 to UE_f	α_1^f, α_2^f	0.8
Power allocation coefficients assigned by BS_1 and BS_2 to UEa and UE_c	α_1^a, α_2^c	0.2
Distance between BS_1 and RIS_R and between BS_2 and RIS_I	d_{1R}, d_{2I}	20m
Distance between RIS_R and UE_f	d_{Rf}	$(d_{1f} - d_{1R})$
Distance between RIS_I and UE_f	d_{If}	$(d_{2f} - d_{2I})$

Now, note that for any $x, y, z > 0$, we have,

$$\log\left(1 + \frac{x}{y+z}\right) = \log\left(1 + \frac{x+y}{z}\right) - \log\left(1 + \frac{y}{z}\right). \tag{6.24}$$

Hence, by considering $x = \alpha_1^f \rho_1 \Omega_1 + \alpha_2^f \rho_2 \Omega_2$, $y = \alpha_1^a \rho_1 \Omega_1 + \alpha_2^c \rho_2 \Omega_2$ and $z = 1$, then $\mathcal{R}_f$ can be rewritten as

$$\mathcal{R}_f = \underbrace{\mathbb{E}\{\log_2(1 + \Delta_1)\}}_{\mathbf{P1}} - \underbrace{\mathbb{E}\{\log_2(1 + \Delta_2)\}}_{\mathbf{P2}}, \tag{6.25}$$

where $\Delta_1 = \rho_1 \Omega_1 + \rho_2 \Omega_2$ and $\Delta_2 = \alpha_1^a \rho_1 \Omega_1 + \alpha_2^c \rho_2 \Omega_2$. Moreover, based on Fact 1 and *Lemma 6.1*, we can see that for $j \in \{1, 2\}$, $\Delta_j \sim \Gamma(k_{\Delta_j}, \theta_{\Delta_j})$. Consequently, by applying the results of *Lemma 6.2* to the expressions **P1** and **P2**, we obtain the ergodic rate of UE_f, which completes the proof. ■

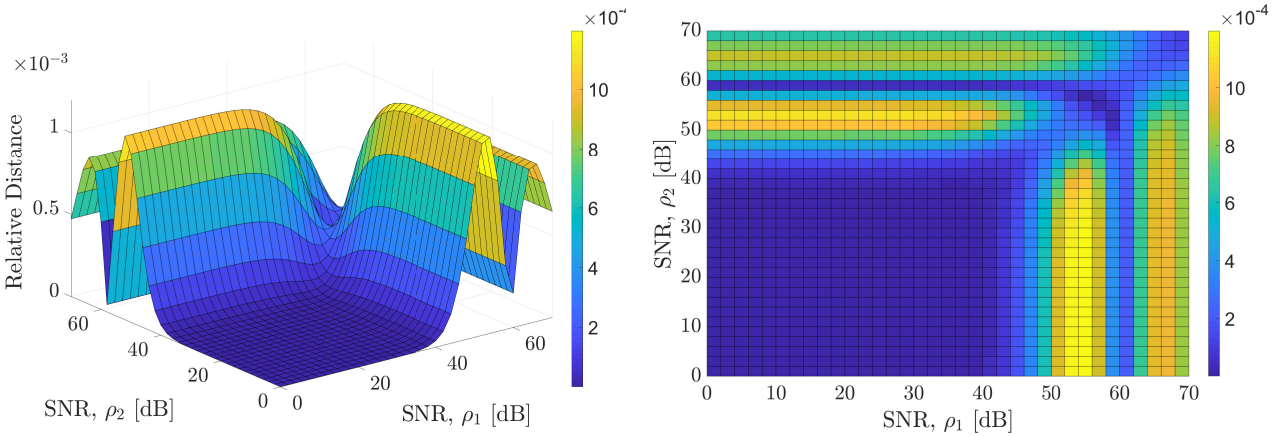
6.4.3 Results And Discussion

In this section, our objective is to validate the performance of the proposed RIS-empowered CoMP NOMA. The simulation results are obtained by generating 10^7 independent Monte-Carlo trials. The main simulation parameters are shown in Table 6.1. The proposed scheme is compared with the three following baselines.

1. CoMP NOMA [68]: this scheme is the same as the proposed one but without any assistance from the RISs.

2. RIS-NOMA without CoMP [132, 138, 173, 179]: this is the conventional scheme that is applied in the literature of RIS-assisted NOMA cellular networks.

3. CoMP C-NOMA [10]: this is similar to the proposed scheme but the CCUs act as active relays. That is, the two CCUs assist the transmission between the two BSs and UE_f instead of using the two RISs.



(a) 3D map. (b) 2D map.

Figure 6.3: Approximation accuracy.

6.4.3.1 Validation of The Analytical Expressions

In order to validate the accuracy of the provided expression for the ergodic rate of UE_f for any values of ρ_1 and ρ_2, we define the relative distance metric as follows.

$$\text{Relative distance} = \frac{|\mathcal{R}_{\text{Sim}} - \mathcal{R}_f|}{\mathcal{R}_{\text{Sim}}}, \tag{6.26}$$

where $\mathcal{R}_{\text{Sim}}$ is the exact ergodic rate of UE_f obtained from simulations. It can be seen from Fig. 6.3 that for most values of the transmit SNRs ρ_1 andρ_2, the analytical ergodic rate perfectly matches the exact ergodic rate. In fact, it can be seen that the highest value of the relative distance over all the transmit SNRs values is equal to 1×10^{-3}, which justifies the accuracy of the proposed approximation. Based on this, and since a homogeneous cellular architecture is considered, we assume that the transmit SNRs from the BSs are equal, i.e., $\rho_1 = \rho_2 = \rho$, for the rest of the results. In addition, from now onwards, the analytical expressions are used to evaluate the system's performance.

6.4.3.2 Ergodic Rate and Network Spectral Efficiency versus Transmit SNR

Fig. 6.4(a) presents the ergodic rate of the CEU vs the SNR, ρ. It can be seen that the ergodic rate of the proposed system outperforms that of the CoMP NOMA and the RIS-NOMA without CoMP baselines. This is due to two main facts. First, our proposed framework allows the two BSs to exploit both RISs in serving the CEU by appropriately adjusting the phase-shift of each meta-atom. Therefore, a stronger combined channel gain at the cell edge UE can be achieved. Second, our proposed framework allows cooperation between the BSs through JT-CoMP, which can mitigate the effect of the ICI. For instance,

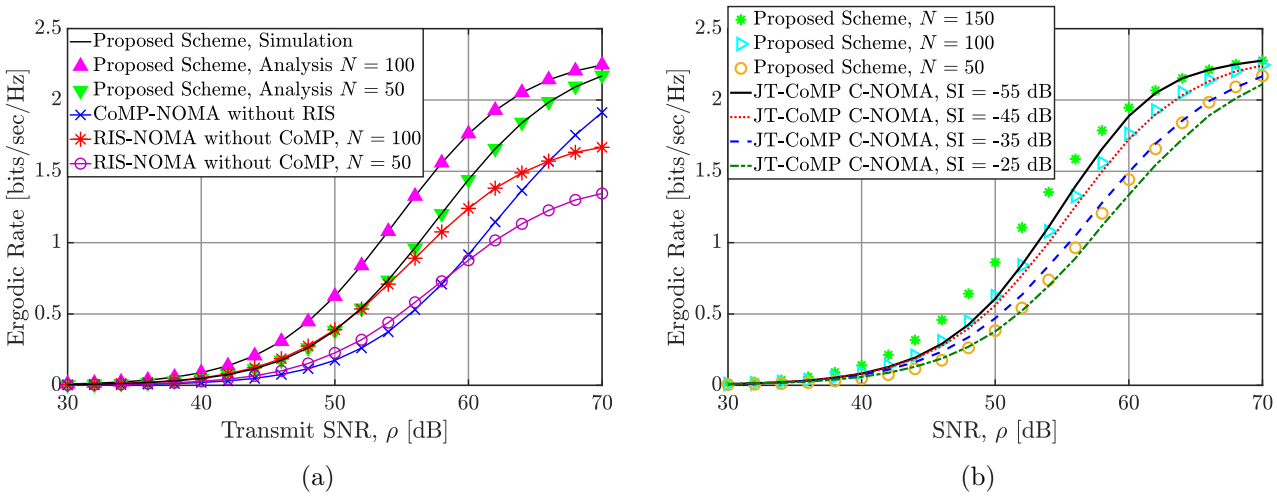


Figure 6.4: The ergodic rate for the CEU vs SNR.

when $\rho = 60$ dB, the proposed scheme achieves a gain of 64% and 57% when $N = 50$ and a gain of 43% and 92% when $N = 100$ against RIS-NOMA without CoMP and CoMP NOMA, respectively, which demonstrates the superiority of the proposed scheme. Furthermore, it can be also observed from Fig. 6.4(a) that, for lower values of ρ, the RIS-NOMA without CoMP scheme achieves a better performance than the CoMP NOMA scheme, especially when the number of elements per RIS increases. This is resulting from the strong combined channel gain that can be obtained at the CEU, which makes it more robust against intra-NOMA interference. However, for high values of SNR, the RIS-NOMA without CoMP has the worst performance. This is because the cell edge UE experiences in this scheme severe ICI interference, which strongly affects its performance.

In Fig. 6.4(b), we compare the performance of the proposed framework with the CoMP C-NOMA scheme with FD relaying proposed in [10] in terms of the ergodic rate of the CEU and the NSE. Aiming to be consistent with [139], and in order to have a fair comparison between the two systems, we consider the constraint $P_m^{\mathrm{D}} + P_b^{\mathrm{BS}} = P_b$ in the CoMP C-NOMA system, where P_m^{D} is the transmit power from the CCU $m \in \{a, c\}$ to the cell edge UE and P_b^{BS} is the transmit power from BS $b \in \{1, 2\}$. These constraints ensure that the proposed scheme and the CoMP C-NOMA scheme utilize the same budget of power. Note that, in the CoMP C-NOMA scheme, the SI limits the gain of the system and the whole system performance depends on the FD performance at the CCUs. However, in the proposed scheme, the number of elements per RIS controls the gain that we can achieve. It can be seen from Fig. 6.4 that, when $N = 150$, the proposed scheme outperforms CoMP

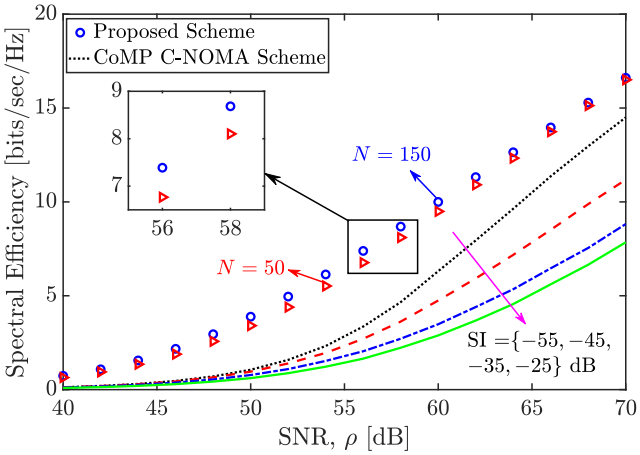


Figure 6.5: The network spectral efficiency vs SNR for various values of SI and RIS elements.

C-NOMA with lower values of SI. This means that we can achieve better performance by using the RIS instead of allowing the CCUs to help the CEU. A similar observation can be seen in Fig. 6.5. In fact, due to the effects of the SI on the performance of the CCUs, the proposed framework has better spectral efficiency than the CoMP C-NOMA scheme.

6.5 Resource Management for RIS-assisted CoMP NOMA Networks

6.5.1 System Model

6.5.1.1 Network Model

We consider a downlink transmission in a RIS-assisted CoMP NOMA cellular network that consists of two adjacent cells as illustrated in Fig. 6.6, where each cell is equipped with one BS. Two classes of users, namely, CCUs and CEUs, are considered within this framework. The CCUs are the users that are near one of the two BSs, whereas the CEUs are the users residing in the overlapped area of the two cells as shown in Fig. 6.6.[2] Such classification can be performed through the relative received signal strength (RSS) between each cellular user and its serving BS using the same approach in **Chapter**

[2] Our system model can be extended to a multi-cell set-up. This can be achieved by allowing any two adjacent cells in a multi-cell set-up to coordinate together to serve the ICI-prone users. In order to tackle the complex coordination between them, the BSs may construct a type of centralized-based radio access network. This can be realized by the introduction of the cloud radio access networks (C-RAN), which is an ideal network architecture to realize CoMP transmission in cellular networks [180].

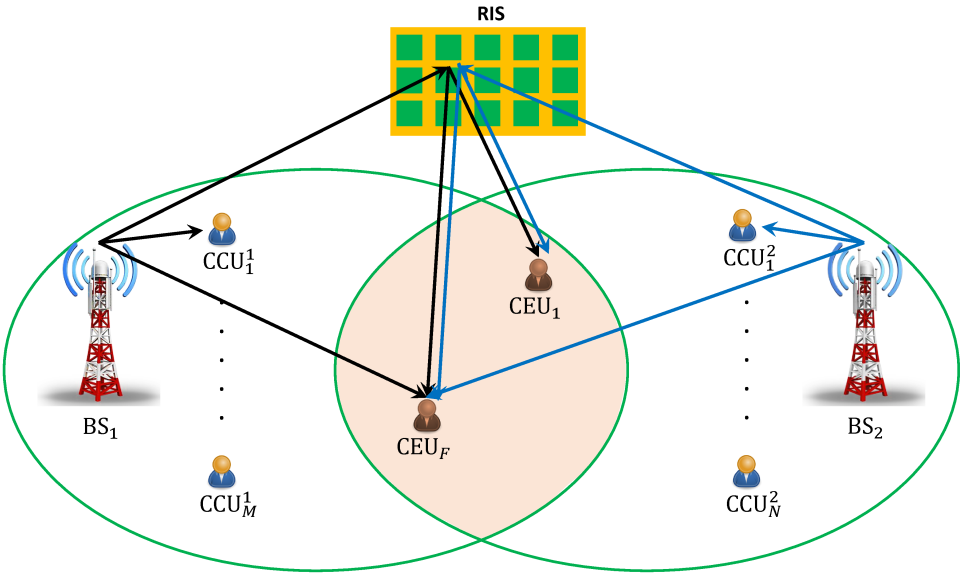


Figure 6.6: An RIS-enabled CoMP NOMA cellular network.

4. In the considered model, each cell invokes two-user NOMA to serve its associated cellular users, where one CCU is paired with one CEU, for which the CoMP is adopted between the two BSs as discussed above. In this context, each CCU from each cell belongs to one NOMA pair within each associated cell, whereas each CEU belongs to two NOMA pairs from the two cells, i.e., each CEU is paired with two CCUs, each from a different cell. Moreover, we assume that the BSs and the cellular users have each one transmit antenna and one received antenna [67, 113, 129, 166]. We also assume that the BSs are connected to a central processing unit through high-speed fronthaul links for exchanging the scheduling/association decisions, signaling information, and the data of CEUs. Although having high-capacity fronthaul links may be a challenge, there are some state-of-the-art fronthaul technologies that have been investigated to release the potential gains of CoMP transmission [181]. In addition, there are transmission techniques such as estimate-compress-forward, compress-forward-estimate, and estimate-multiply-compress-forward that are deemed as promising methods to circumvent the issue of limited-capacity fronthaul links [160]. In this paper, and similar to [67, 158, 160], we assume that the available fronthaul capacity is sufficient for exchanging signaling and data among BSs.

In order to further enhance the performance of the CEUs, a RIS equipped with L reflecting elements is deployed at the overlapped edge of the two cells to boost the power of the useful signal by carefully controlling the PS of the RIS's meta-atoms. Hence, as shown in Fig. 6.6, each CEU receives not only its desired signals from the two BSs, but

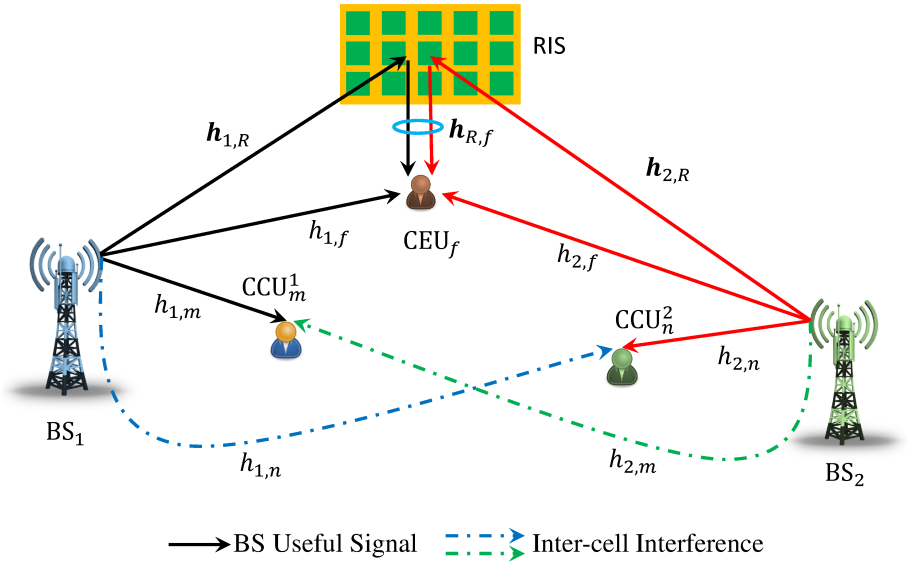


Figure 6.7: Communication links of one *coordinated NOMA cluster*.

also the reflected signals from the RIS. Note that, we assume that the received signals at the CCUs due to the reflection of the RIS can be neglected due to the unavailability of the communication links resulting from the long distance and blockages [14, 173]. For ease of illustration, let BS_1 and BS_2 denote the BSs relative to the first and second cell, respectively. In addition, let CCU^1 and CCU^2 denote the CCUs associated with the first and second cell, respectively. In this context, we define a *coordinated NOMA cluster* as a cluster of cellular users that consists of one CCU^1, one CEU, and one CCU^2. Since in practical cellular systems, each cell is serving multiple coexisting users simultaneously, different *coordinated NOMA clusters* may coexist within the considered cellular system. Hence, in order to cancel the inter-cluster interference, i.e., the interference between the different clusters of users, the resulting distinct *coordinated NOMA clusters* are simultaneously served over orthogonal and equally divided frequency channels [69, 98].

Let $\mathcal{I} = \{1, 2\}, \mathcal{M} = \{1, 2, \ldots, M\}, \mathcal{F} = \{1, 2, \ldots, F\}$ and $\mathcal{N} = \{1, 2, \ldots, N\}$ be the sets of indices of the BSs, the CCUs^1, the CEUs, and the CCUs^2, respectively, where, M, F and N are the cardinality of $\mathcal{M}, \mathcal{F}$, and $\mathcal{N}$, respectively. In addition, let $\mathcal{K} = \{\mathcal{M} \cup \mathcal{N}\}$ be the set of all CCUs in both cells and let $\mathcal{U} = \{\mathcal{M} \cup \mathcal{F} \cup \mathcal{N}\}$ be the set of all cellular users. Regarding the channel models between the different communication nodes, both the large-scale fading and the small-scale fading are considered for each communication link. Specifically, for all $i \in \mathcal{I}$, $u \in \mathcal{U}$, and $f \in \mathcal{F}$, $h_{i,u} \in \mathbb{C}$, $\boldsymbol{h}_{i,R} \in \mathbb{C}^{L\times 1}$ and $\boldsymbol{h}_{R,f} \in \mathbb{C}^{L\times 1}$ denote the channel coefficients of the wireless links from $\text{BS}_i \longrightarrow$ user u, $\text{BS}_i \longrightarrow$ RIS, and RIS $\longrightarrow \text{CEU}_f$, respectively. In the proposed model, we assume that the wireless

links between the BSs and all cellular users, i.e., the CCUs and the CEUs, experience each a small-scale fading that follows a Rayleigh distribution. Therefore, $\forall i \in \mathcal{I}$, $u \in \mathcal{U}$, the channel coefficient $h_{i,u} = g_{i,u}\sqrt{PL(d_{i,u})}$, where $g_{i,u} \in \mathbb{C}$ is the small-scale Rayleigh fading with a zero mean and a unit variance, and $PL(d_{i,u})$ is the large-scale path-loss, which is modeled as $PL(d_{i,u}) = \rho_0(\frac{d}{d_0})^{-\eta_{i,u}}$, such that ρ_0 is the path-loss at a reference distance d_0, $\eta_{i,u}$ is the path-loss exponent and $d_{i,u}$ is the distance between BS_i and user u. On the other hand, regarding the communication links between each BS and the RIS and also between the RIS and each CEU, the line-of-sight (LoS) components are assumed to exist [127]. Therefore, these wireless links experience a small-scale fading that follows each Rician distribution. Hence, $\forall i \in \mathcal{I}$, the corresponding channel coefficients for the links between BS_i and the RIS elements can be expressed as,

$$\boldsymbol{h}_{i,R} = \sqrt{PL(d_{i,R})}\left(\sqrt{\frac{1}{1+\kappa_{i,R}}}\mathbf{g}_{i,R} + \sqrt{\frac{\kappa_{i,R}}{1+\kappa_{i,R}}}\hat{\mathbf{g}}_{i,R}\right), \tag{6.27}$$

where $\kappa_{i,R}$ represents the Rician factor, $\hat{\mathbf{g}}_{i,R}$ represents the deterministic LoS components, and $\mathbf{g}_{i,R}$ represents the non-line-of-sight (NLoS) components, which follows a Rayleigh distribution with a mean zero and unit variance. The channel model for the wireless links between the RIS elements and the CEUs can be obtained similarly. Moreover, the channel state information (CSI) of the considered wireless links is assumed to be perfectly estimated at the BS [129, 131, 132, 165, 166]. Despite the difficulty of obtaining perfect CSI, various efforts in the literature have recently designed efficient channel estimation techniques for RIS-enabled wireless networks that can be adopted in our system model to provide accurate CSI [136, 137, 182]. For instance, the authors in [136] designed an alternating least square approach dependent on the parallel factor framework which is able to continuously estimate all the wireless links with low complexity. Two channel estimation methods based on deep learning and compressive sensing were discussed in RIS-assisted wireless systems [137]. In [182], the RIS-to-user and the BS-to-RIS channels are estimated by using matrix completion and sparse matrix factorization, respectively.

6.5.1.2 Signal Model and Rates Analysis

To precisely describe the signal model at each CCU and CEU, a single *coordinated NOMA cluster* is considered in this part as shown in Fig. 6.7. In particular, $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, we assume that the tuple (CCU_m^1, CEU_f, CCU_n^2) forms one *coordinated NOMA cluster* as illustrated in Fig. 6.7. We start by defining the main operations

occurring in the first cell. At the beginning of a time block, BS_1 broadcasts a superimposed signal of the messages intended to CCU_m^1 and CEU_f, which is expressed as $y_{1,m,f} = \sqrt{(1-\alpha_{1,m,f})P_1}y_{1,m} + \sqrt{\alpha_{1,m,f}P_1}y_f$, where $y_{1,m}$ and y_f denote the normalized signals intended to CCU_m^1 and CEU_f, respectively, i.e., $\mathbb{E}[|y_{1,m}|^2] = 1$ and $\mathbb{E}[|y_f|^2] = 1$, P_1 is the budget of power allocated by BS_1 to the NOMA pair (CCU_m^1, CEU_f), and $\alpha_{1,m,f}$ is the PA coefficient assigned by BS_1 to CEU_f. Following the principle of NOMA, since CCU_m^1 is the strong NOMA user within the NOMA pair (CCU_m^1, CEU_f), it should be able to detect the message of CEU_f. Therefore, the PA coefficient $\alpha_{1,m,f}$ should satisfy the constraint $0.5 < \alpha_{1,m,f} < 1$ [115, 183]. Consequently, the received baseband signal at CCU_m^1 can be expressed as

$$z_{1,m} = h_{1,m}y_{1,m,f} + h_{2,m}y_{2,n,f} + w_{1,m}, \tag{6.28}$$

where $y_{2,n,f}$ is the message broadcasted by BS_2 and $w_{1,m}$ is an additive white Gaussian noise experienced at CCU_m^1, which is $\mathcal{CN}(0,\sigma^2)$ distributed. The term $h_{2,m}y_{2,n,f}$ represents the ICI signal broadcast from BS_2 and experienced at CCU_m^1. Afterwards, using the SIC technique, CCU_m^1 decodes first the message of CEU_f, i.e., y_f, and then removes it from its received signal in order to decode its own message, i.e., $y_{1,m}$. Based on this, the achievable rate at CCU_m^1 to decode the message of CEU_f is expressed as by

$$\mathcal{R}_{1,m\to f} = \log\left(1 + \frac{\alpha_{1,m,f}\gamma_{1,m}}{(1-\alpha_{1,m,f})\gamma_{1,m}+1}\right), \tag{6.29}$$

where $\gamma_{1,m} = P_1|h_{1,m}|^2/\mathcal{T}$ and $\mathcal{T} \triangleq P_2|h_{2,m}|^2 + \sigma^2$. In addition, the achievable data rate at CCU_m^1 to decode its own message is expressed, as

$$\mathcal{R}_{1,m} = \log\left(1 + (1-\alpha_{1,m,f})\gamma_{1,m}\right). \tag{6.30}$$

By following similar steps as in (6.28)-(6.30), the achievable rate at CCU_n^2 to decode the message of CEU_f, which is denoted by $\mathcal{R}_{2,n\to f}$, and the achievable rate at CCU_n^2 to decode its own message, i.e., which is denoted by $\mathcal{R}_{2,n}$ can be obtained. On the other hand, CEU_f receives its own signal through the transmissions from the two BSs, i.e., BS_1 and BS_2, and from the reflection by the RIS. Therefore, the received signal at CEU_f can be expressed as

$$z_f = (h_{1,f} + \boldsymbol{h}_{R,f}^H\boldsymbol{\Theta}\boldsymbol{h}_{1,R})y_{1,m,f} + (h_{2,f} + \boldsymbol{h}_{R,f}^H\boldsymbol{\Theta}\boldsymbol{h}_{2,R})y_{2,n,f} + w_f, \tag{6.31}$$

where w_f is an additive white Gaussian noise experienced at CEU_f, which is $\mathcal{CN}(0,\sigma^2)$ distributed, and $\boldsymbol{\Theta} = \mathrm{diag}\{\phi_1(\theta_1), \phi_2(\theta_2), \ldots, \phi_L(\theta_L)\}$ is the PS matrix of the RIS, where $\forall l \in [\![1, L]\!]$, $\phi_l(\theta_l) = e^{j\theta_l}$, such that $\theta_l \in [0, 2\pi]$ represents the PS of the lth reflecting element. Due to the consideration of non-coherent JT-CoMP, the achievable rate at

CEU_f to decode its message can be expressed as [41]

$$\mathcal{R}_f = \log\left(1 + \frac{\alpha_{1,m,f}\gamma_{1,f} + \alpha_{2,n,f}\gamma_{2,f}}{(1-\alpha_{1,m,f})\gamma_{1,f} + (1-\alpha_{2,n,f})\gamma_{1,f} + 1}\right), \tag{6.32}$$

where $\gamma_{1,f} = P_1|h_{1,f} + \boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{1,R}|^2/\sigma^2$ and $\gamma_{2,f} = P_2|h_{2,f} + \boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{2,R}|^2/\sigma^2$. Based on the above analysis, the achievable sum-rate of the *coordinated NOMA cluster* defined by the tuple (CCU_m^1, CEU_f, CCU_n^2) can be expressed as

$$\mathcal{R}_{m,f,n} = \mathcal{R}_{1,m} + \mathcal{R}_f + \mathcal{R}_{2,n}. \tag{6.33}$$

6.5.2 Problem Formulation and Solution Approach

6.5.2.1 Network Sum-Rate Problem Formulation

With the quest of enhancing the performance of the proposed RIS-assisted CoMP NOMA network, a framework that jointly optimizes the PA coefficients at the BS for each *coordinated NOMA cluster*, the UC policy, and the PS matrix is proposed. The UC policy determines the members of each *coordinated NOMA cluster*, i.e., which CCU from each cell should be paired with which CEU. This framework is formulated as an optimization problem with the goal of maximizing the network sum-rate while guaranteeing a successful SIC process at CCUs and the required QoS for all users, which can be written as follows.

$$\text{OPT}: \quad \max_{\mathcal{A},\mathcal{X},\boldsymbol{\theta}} \sum_{m\in\mathcal{M}} \sum_{f\in\mathcal{F}} \sum_{n\in\mathcal{N}} x_{m,f,n}\mathcal{R}_{m,f,n}, \tag{6.34a}$$

$$\text{s.t. } 0.5x_{k_1,f,k_2} \le \alpha_{i,k_i,f} \le x_{k_1,f,k_2}, \qquad \forall i \in \mathcal{I}, k_i \in \mathcal{K}_i, f \in \mathcal{F}, \tag{6.34b}$$

$$\mathcal{R}_{i,k_i} \ge x_{k_1,f,k_2} R_{k_i}^{\text{th}}, \qquad \forall i \in \mathcal{I}, k_i \in \mathcal{K}_i, f \in \mathcal{F}, \tag{6.34c}$$

$$\mathcal{R}_f \ge x_{k_1,f,k_2} R_f^{\text{th}}, \qquad \forall f \in \mathcal{F}, \tag{6.34d}$$

$$\mathcal{R}_{i,k_i \longrightarrow f} \ge x_{k_1,f,k_2} R_f^{\text{th}}, \qquad \forall i \in \mathcal{I}, f \in \mathcal{F}, \tag{6.34e}$$

$$\sum_{n\in\mathcal{N}} \sum_{f\in\mathcal{F}} x_{m,f,n} \le 1, \qquad \forall m \in \mathcal{M}, \tag{6.34f}$$

$$\sum_{f\in\mathcal{F}} \sum_{m\in\mathcal{M}} x_{m,f,n} \le 1, \qquad \forall n \in \mathcal{N}, \tag{6.34g}$$

$$\sum_{n\in\mathcal{N}} \sum_{m\in\mathcal{M}} x_{m,f,n} \le 1, \qquad \forall f \in \mathcal{F}, \tag{6.34h}$$

$$x_{m,f,n} \in \{0,1\}, \qquad \forall m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}, \tag{6.34i}$$

$$|\phi_l(\theta_l)| = 1, \qquad \forall l \in [\![1, L]\!], \tag{6.34j}$$

where $\mathcal{X} = \{ x_{m,f,n} \in \{0,1\} | \, m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$ represents the set of UC indicators, such that $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, $x_{m,f,n}$ is the UC indicator of the tuple (CCU_m^1, CEU_f, CCU_n^2), $\mathcal{A} = \{(\alpha_{1,m,f}, \alpha_{2,n,f}) \in [0,1]^2 \, | m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$ is the collection of the BSs PA coefficients, and $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_L]$ is the $L \times 1$ vector of the PS for the RIS's elements, and k_i and $\mathcal{K}_i$ are defined $\forall i \in \mathcal{I}$, as $k_i = m$ and $\mathcal{K}_i = \mathcal{M}$ if $i = 1$, and $k_i = n$ and $\mathcal{K}_i = \mathcal{N}$ if $i = 2$.

Looking in depth into problem OPT, constraint (6.34b) represents the SIC constraints at the CCUs within each *coordinated NOMA cluster*, constraints (6.34c) and (6.34d) guarantee the required QoS at the users, where $\forall i \in \mathcal{I}$ and $f \in \mathcal{F}$, $R_{k_i}^{\text{th}}$ and R_f^{th} represent the minimum required rates at the CCUs and at the CEU, respectively. Furthermore, constraint (6.34e) assures that each CCU can decode the signal of the paired CEU to satisfy the SIC constraint, constraints (6.34f), (6.34g), and (6.34h) guarantee that each user, either a CCU or a CEU, is a member of at most one *coordinated NOMA cluster*. Finally, constraint (6.34i) shows that the value of the UC indicator is either binary, where $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, $x_{m,f,n} = 1$ indicates that the tuple (CCU_m^1, CEU_f, CCU_n^2) construct a *coordinated NOMA cluster* and $x_{m,f,n} = 0$ otherwise.

To maximize the network sum-rate, the PA at the two BSs, the UC policy, and the PS matrix at the RIS should be jointly optimized. Due to the coexistence and the high coupling of the binary variables $\mathcal{X}$ and the continuous variables $\mathcal{A}$, $\boldsymbol{\theta}$, the objective function and the constraints of the optimization problem OPT are non-convex. In fact, it can be easily seen that the formulated optimization problem OPT is an MINLP problem, which is difficult to solve. As a result, it is necessary to transform problem OPT into some tractable sub-problems that can be solved separately and alternatively over multiple iterations. Towards this end, we invoke the AO method to solve the original problem OPT in an efficient manner, which will be detailed in the following part.

6.5.2.2 Solution Roadmap

The decomposition of problem OPT and the proposed solution approach for each obtained sub-problem are presented in Fig. 6.8. Due to its intractability, problem OPT is decomposed into two sub-problems, namely, a joint PA and UC optimization sub-problem, which is an MINLP problem, and a PS optimization sub-problem, which is a non-convex. With the aid of the AO approach, the two sub-problems are solved in an alternating

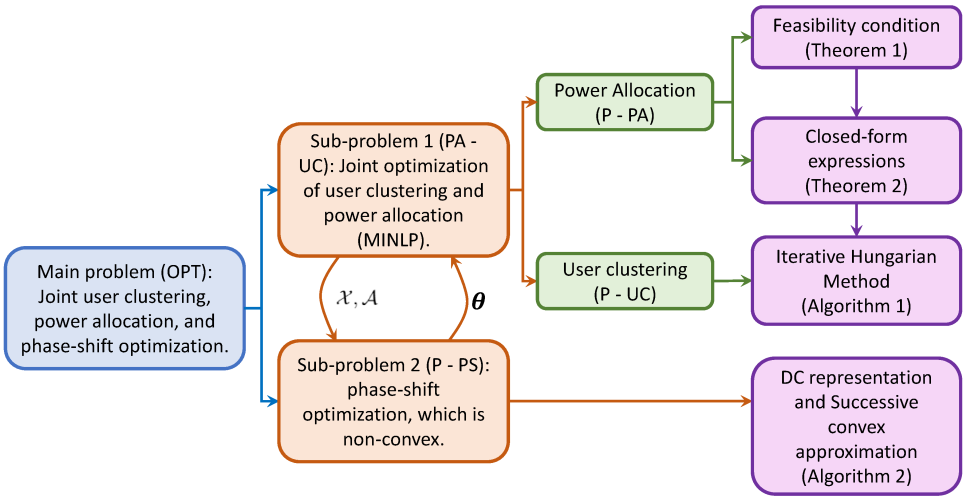


Figure 6.8: Problem decomposition and solution roadmap.

way [127, 129, 131]. Precisely, at the first iteration, a random RIS PS matrix is generated and inputted to the joint PA and UC optimization sub-problem. Then, the resulting problem is solved and the obtained PA scheme and UC policy are inputted to the PS optimization sub-problem. Afterwards, the derived PS optimization sub-problem is solved and an RIS PS matrix is obtained, which will be then inputted to the joint PA and UC optimization sub-problem at the second iteration, and the same steps of the first iteration will be repeated. This procedure, which is known as the AO procedure, will keep running until convergence.

Within the invoked AO approach, efficient solution approaches for the joint PA and UC optimization sub-problem and the PS optimization sub-problem are presented in section 6.5.3 and section 6.5.4, respectively. With the aid of the bi-level optimization, the joint PA and UC optimization sub-problem is divided into two sub-problems: a PA sub-problem for a single *coordinated NOMA cluster*, which is defined as the inner problem and a UC sub-problem, which is denoted as outer problem [107]. For the inner problem, by analyzing the objective function and all the constraints, we derive the feasibility conditions as in presented in ***Theorem 6.2***. Then, for all possible *coordinated NOMA cluster*, closed-form expressions of the optimal PA coefficients are derived as shown in ***Theorem 6.3***. After finding the optimal PA scheme in a closed-form for all possible *coordinated NOMA cluster*, the outer problem, i.e., the UC policy, ends up to be a linear assignment problem, or precisely, a *3-dimensional matching problem*, as shown in (6.42). Such matching problem is an *NP-complete* problem [184]. Due to this, we resolve this challenge by introducing an iterative technique that projects the resulting *3-dimensional matching problem* into three

2-dimensional matching problems, which can be solved using the *Hungarian method* [184]. Based on this, the proposed UC policy is obtained as shown in **Algorithm** 5. Finally, for the PS sub-problem, we resort to DC representation for rank-one constraint and we then invoke the SCA to solve the PS matrix optimization sub-problem, and the proposed approach is summarized in **Algorithm** 6.

Note that, for the case of multiple BSs association scenario, the proposed methods of designing the UC policy and the RIS PS configuration are primordial tools to investigate the more general multi-cell set-up. However, there will be two major challenges that should be considered: 1) Selecting the group of BSs that cooperate to serve a given CEU should be addressed, which is denoted as BSs clustering, and 2) Obtaining an optimal and a low-complexity PA scheme will be more challenging while considering multiple BSs to serve a CEU and, therefore, it should be revisited.

6.5.3 Joint Power Allocation and User Clustering Optimization

Given an RIS's PS matrix, we seek in this section to solve the joint optimization of the PA coefficients and the UC policy, which is expressed as

$$\mathrm{PA}-\mathrm{UC}: \max_{\mathcal{A},\mathcal{X}} \sum_{m\in\mathcal{M}}\sum_{f\in\mathcal{F}}\sum_{n\in\mathcal{N}} x_{m,f,n}\mathcal{R}_{m,f,n}, \tag{6.35a}$$

$$\text{s.t.} \qquad (6.34\text{b})-(6.34\text{i}). \tag{6.35b}$$

In problem $\mathrm{PA}-\mathrm{UC}$, it can be seen that, $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, the achievable sum-rate $\mathcal{R}_{m,f,n}$ of a single *coordinated NOMA cluster*, which is shown in (6.33), is independent of the UC binary indicator $x_{m,f,n}$ and it is only a function of the PA coefficients. Particularly, one can see that $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, if the optimal UC indicator $x^*_{m,f,n} = 0$, then the optimal PA policy $\left(\alpha^*_{1,m,f}, \alpha^*_{2,n,f}\right) = \mathbf{0}_{2\times 1}$. However, if $x^*_{m,f,n} = 1$, then $\left(\alpha^*_{1,m,f}, \alpha^*_{2,n,f}\right)$ should be the optimal solutions of the PA coefficients of the *coordinated NOMA cluster* (CCU^1_m, CEU_f, CCU^2_n). This means that, $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, if we assume that CCU^1_m, CEU_f and CCU^2_n construct a *coordinated NOMA cluster* and that we can determine their optimal PA coefficients $\left(\alpha^*_{1,m,f}, \alpha^*_{2,n,f}\right)$, then problem $\mathrm{PA}-\mathrm{UC}$ transforms into a linear assignment problem, which aims at obtaining the optimal UC policy that maximizes the network sum-rate. Thus, with the aid of the bi-level optimization approach, the formulated optimization problem can be decoupled into two sub-problems [107], one inner sub-problem that consists of a PA sub-problem

and one outer sub-problem that consists of a UC sub-problem. In the next subsections, we present the two sub-problems along with their proposed solution approaches.

6.5.3.1 PA for Each Coordinated NOMA Cluster

In this subsection, and $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$, we investigate the PA scheme for the *coordinated NOMA cluster* (CCU_m^1, CEU_f, CCU_n^2) by determining the PA coefficients at both BS_1 and BS_2. Consequently, the optimal PA solution can be obtained by maximizing the achievable sum-rate of a given *coordinated NOMA cluster* (CCU_m^1, CEU_f, CCU_n^2), which can be formulated as follows.

$$\text{P}-\text{PA}: \quad \max_{\alpha_{1,m,f},\alpha_{2,n,f}} \quad \mathcal{R}_{m,f,n}, \tag{6.36a}$$

$$\text{s.t.} \quad 0.5 \leq \alpha_{1,m,f},\ \alpha_{2,n,f} \leq 1, \tag{6.36b}$$

$$\mathcal{R}_{1,m\to f}, \mathcal{R}_f, \mathcal{R}_{2,n\to f} \geq R_f^{\text{th}}, \tag{6.36c}$$

$$\mathcal{R}_{1,m} \geq R_m^{\text{th}}, \tag{6.36d}$$

$$\mathcal{R}_{2,n} \geq R_n^{\text{th}}. \tag{6.36e}$$

Problem $\text{P}-\text{PA}$ is a non-convex optimization problem and, hence, it is challenging to obtain its optimal solution. Nevertheless, since $\text{P}-\text{PA}$ should be solved for all possible *coordinated NOMA clusters*, a computational-time efficient approach needs to be explored. In doing so, we investigate the objective function and the constraints of problem $\text{P}-\text{PA}$ so that closed-form expressions for the PA coefficients can be obtained. In this context, due to the existence of both SIC and QoS constraints, problem $\text{P}-\text{PA}$ might be infeasible for some channel realizations of the cellular users. Due to this, we check first the conditions under which problem $\text{P}-\text{PA}$ is feasible, which are given in the following theorem.

Theorem 6.2. *Problem* $\text{P}-\text{PA}$ *is feasible if and only if the following conditions hold.*

$$\textit{Condition 1: } t_m \leq \min\left(\frac{\gamma_{1,m}}{2}, -1+\sqrt{1+\gamma_{1,m}}\right), \tag{6.37a}$$

$$\textit{Condition 2: } t_n \leq \min\left(\frac{\gamma_{2,n}}{2}, -1+\sqrt{1+\gamma_{2,n}}\right), \tag{6.37b}$$

$$\textit{Condition 3: } Y_f \leq Y_{1,m,f} + Y_{2,n,f}, \tag{6.37c}$$

where $\forall k \in \{m, f, n\}, t_k = e^{R_k^{\text{th}}} - 1$ *and*

$$Y_{1,m,f} = \frac{(\gamma_{1,m} - t_m)}{\gamma_{1,m}}(1+t_f)\gamma_{1,f},$$

$$Y_{2,n,f} = \frac{(\gamma_{2,n} - t_n)}{\gamma_{2,n}}(1+t_f)\gamma_{2,f},$$

$$Y_f = t_f(\gamma_{1,f} + \gamma_{2,f} + 1). \tag{6.38a}$$

Proof. See Appendix D1. ■

Based on this, assuming that problem P − PA is feasible, the corresponding optimal PA coefficients are presented in the following theorem.

Theorem 6.3. *The optimal PA coefficients* $\alpha^*_{1,m,f}$ *and* $\alpha^*_{2,n,f}$ *of problem* P − PA *are expressed as*

$$\alpha^*_{1,m,f} = \alpha^{\min}_{m,f,n} U\left(\alpha^{\min}_{m,f,n} - x_0\right) + x_0 U\left(x_0 - \alpha^{\min}_{m,f,n}\right) U\left(\alpha^{\max}_{m,f,n} - x_0\right) + \alpha^{\max}_{m,f,n} U\left(x_0 - \alpha^{\max}_{m,f,n}\right), \tag{6.39a}$$

$$\alpha^*_{2,n,f} = -\frac{\gamma_{1,f}}{\gamma_{2,f}} \alpha^*_{1,m,f} + \frac{t_f\left(\gamma_{1,f} + \gamma_{2,f} + 1\right)}{\left(1 + t_f\right)\gamma_{2,f}}, \tag{6.39b}$$

where $U(\cdot)$ *denotes the unit step function and*

$$\alpha^{\min}_{m,f,n} = \max\left(\alpha^{\min}_{1,m,f}, -\frac{\gamma_{2,f}}{\gamma_{1,f}} \alpha^{\max}_{2,n,f} + \frac{t_f\left(\gamma_{1,f} + \gamma_{2,f} + 1\right)}{\left(1 + t_f\right)\gamma_{1,f}}\right), \tag{6.40a}$$

$$\alpha^{\max}_{m,f,n} = \min\left(\alpha^{\max}_{1,m,f}, -\frac{\gamma_{2,f}}{\gamma_{1,f}} \alpha^{\min}_{2,n,f} + \frac{t_f\left(\gamma_{1,f} + \gamma_{2,f} + 1\right)}{\left(1 + t_f\right)\gamma_{1,f}}\right), \tag{6.40b}$$

such that, for all $i \in \mathcal{I}$ *and* $k_i \in \mathcal{K}_i$

$$\alpha^{\min}_{i,k_i,f} = \max\left(\frac{1}{2}, \frac{\left(\gamma_{i,k_i} + 1\right) t_{k_i}}{\left(t_{k_i} + 1\right)\gamma_{i,k_i}}\right), \quad \alpha^{\max}_{i,k_i,f} = \frac{\gamma_{i,k_i} - t_{k_i}}{\gamma_{i,k_i}}. \tag{6.41}$$

Proof. See Appendix D2. ■

For all $m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}$, let us define the optimal PA scheme for the *coordinated NOMA cluster* (CCU^1_m, CEU_f, CCU^2_n) as $\mathcal{P}^*_{m,f,n} \triangleq (\alpha^*_{1,m,f}, \alpha^*_{2,n,f})$. In addition, we denote by $\mathcal{P}^* \triangleq \{\mathcal{P}^*_{m,f,n} | m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$ the optimal PA scheme of all possible clusters. Based on this, a near-optimal UC policy will be obtained in the next sub-section.

6.5.3.2 UC: An Iterative Hungarian Method

It is important to note that achieving the required QoS at each CEU depends on the PA coefficients assigned by both BS_1 and BS_2, which in turn depends on the two CCUs that are members with the considered CEU in the same *coordinated NOMA cluster*. As a result, the two BSs should jointly perform the UC policy by grouping one CCU from each cell along with one CEU. In particular, we need to get the best users that should be grouped within each *coordinated NOMA cluster* given the optimal PA coefficients in $\mathcal{P}^*$ that was obtained from solving problem P − PA $\forall m \in \mathcal{M}$, $f \in \mathcal{F}$ and $n \in \mathcal{N}$. Consequently, the optimal UC policy $\mathcal{X}^\star$ is obtained by solving the following optimization problem.

$$\text{P} - \text{UC}: \max_{\mathcal{X}} \sum_{m \in \mathcal{M}} \sum_{f \in \mathcal{F}} \sum_{n \in \mathcal{N}} x_{m,f,n} \mathcal{R}_{m,f,n}(\mathcal{P}^*_{m,f,n}), \tag{6.42a}$$

$$\text{s.t.} \qquad (6.34\text{f}) - (6.34\text{i}). \tag{6.42b}$$

Problem $\mathrm{P}-\mathrm{UC}$ is a *3-dimensional matching problem* and its optimal solution requires an exhaustive search over all possible *coordinated NOMA clusters*, which is impractical due to its immensely high computational complexity especially in large-scale networks. In order to resolve this issue, we propose an iterative technique that projects the *3-dimensional matching problem* in (6.42) into iterative *2-dimensional matching problems* that can be easily solved. The main idea of this iterative method is to give the opportunity for each user to choose the best other two members to form a potential *coordinated NOMA cluster* such that the network sum-rate is maximized.

Consider an arbitrary initial UC policy $\mathcal{X}=\mathcal{X}^{0^\star}$ that is feasible to problem $\mathrm{P}-\mathrm{UC}$ in (6.42), i.e., satisfying the constraints (6.34f)-(6.34i). Then, define the set of indices of CCUs^1 and CEUs that are clustered together through the UC policy $\mathcal{X}^{0^\star}$ as $\Omega_0=\big\{(m,f)|\,x^{0^\star}_{m,f,n}=1,\,\forall\,(m,f,n)\in\mathcal{M}\times\mathcal{F}\times\mathcal{N}\big\}$. Then, define the binary set $\mathcal{X}^1=\big\{x^1_{m,f,n}\in\{0,1\}|\,\forall\,(m,f,n)\in\Omega_0\times\mathcal{N}\big\}$. Consequently, the problem of designing the optimal assignment of each CCU^2 to a predefined pair in Ω_0 can be expressed as follows.

$$\mathrm{P}.1-\mathrm{UC}:\quad \mathcal{X}^{1^\star}=\arg\max_{\mathcal{X}^1}\sum_{(m,f)\in\Omega_0}\sum_{n\in\mathcal{N}}x^1_{m,f,n}\mathcal{R}^\star_{m,f,n} \tag{6.43a}$$

$$\text{s.t.}\quad \sum_{(m,f)\in\Omega_0}x^1_{m,f,n}=1,\forall n\in\mathcal{N}, \tag{6.43b}$$

where constraint (6.43b) indicates that, for all $n\in\mathcal{N}$, only one CCU^2_n can be uniquely assigned to a given predefined pair (CCU^1_m, CEU_f), for all $(m,f)\in\Omega_0$. Since the set Ω_0 is already defined and fixed, it can be seen that problem $\mathrm{P}.1-\mathrm{UC}$ is now a 2-dimensional matching problem. As a result, the *Hungarian method* can be employed to optimally solve $\mathrm{P}.1-\mathrm{UC}$, which yields to the UC policy $\mathcal{X}^{1^\star}$ [185].

Next, we give the opportunity to each CEU to select the optimal CCU^1 and CCU^2 from the first and second cell, respectively, in order to maximize the network utility in terms of the network sum-rate. Let $\mathcal{M}^1$ and $\mathcal{N}^1$ denote the set of indices of CCUs^1 and CCUs^2 resulting from the clustering policy $\mathcal{X}^{1^\star}$, i.e., $\mathcal{M}^1=\big\{m\,\big|x^1_{m,n,f}=1,\,\forall\,(m,f,n)\in\mathcal{M}\times\mathcal{F}\times\mathcal{N}\big\}$ and $\mathcal{N}^1=\big\{n\,\big|x^1_{m,n,f}=1,\,\forall\,(m,f,n)\in\mathcal{M}\times\mathcal{F}\times\ \ \mathcal{N}\big\}$. Then, define the set of indices of CCUs^1 and CCUs^2 that are clustered together through the UC policy $\mathcal{X}^{1^\star}$ as $\Omega_1=\big\{(m,n)|\,x^{1^*}_{m,f,n}=1,\,\forall\,(m,f,n)\in\mathcal{M}\times\mathcal{F}\times\mathcal{N}\big\}$. Afterwards, define the binary set $\mathcal{X}^2=\big\{x^2_{m,f,n}\in\{0,1\}\big|\,\forall\,(m,f,n)\in\mathcal{M}^1\times\mathcal{F}\times\mathcal{N}^1\big\}$. Consequently, the problem of designing the optimal assignment of each CEU to a predefined pair in Ω_1 can be expressed

as follows.

$$\text{P.2}-\text{UC}: \quad \mathcal{X}^{2^\star} = \arg\max_{\mathcal{X}^2} \sum_{(m,n)\in\Omega_1} \sum_{f\in\mathcal{F}} x^2_{m,f,n} \mathcal{R}^\star_{m,f,n} \tag{6.44a}$$

$$\text{s.t.} \quad \sum_{(m,n)\in\Omega_1} x^2_{m,f,n} = 1, \forall f \in \mathcal{F}. \tag{6.44b}$$

It can be seen that problem $\text{P.2}-\text{UC}$ has the same structure as $\text{P.1}-\text{UC}$ and the *Hungarian method* can be applied to obtain the UC policy $\mathcal{X}^{2^\star}$.

Finally, it is the turn for each CCU^1 to select the optimal CEU and the optimal CCU^2, in order to maximize the network sum-rate. First, define the set of indices of CEUs and CCUs^2 that are clustered together through the UC policy $\mathcal{X}^{2^\star}$ as $\Omega_2 = \left\{ (f,n) | \, x^{2^*}_{m,f,n} = 1, \, \forall \, (m,f,n) \in \mathcal{M} \times \mathcal{F} \times \mathcal{N} \right\}$. Then, define the binary set $\mathcal{X}^3 = \left\{ x^3_{m,f,n} \in \{0,1\} | \, \forall \, (m,f,n) \in \mathcal{M} \times \Omega_2 \right\}$. Consequently, the problem of designing the optimal assignment of each CCU^1 to a predefined pair in Ω_1 can be expressed as follows.

$$\text{P.3}-\text{UC}: \quad \mathcal{X}^{3^\star} = \arg\max_{\mathcal{X}^3} \sum_{(f,n)\in\Omega_2} \sum_{m\in\mathcal{M}} x^3_{m,f,n} \mathcal{R}^\star_{m,f,n} \tag{6.45a}$$

$$\text{s.t.} \quad \sum_{(f,n)\in\Omega_2} x^3_{m,f,n} = 1, \forall m \in \mathcal{M}. \tag{6.45b}$$

The *Hungarian method* can be utilized to get the policy $\mathcal{X}^{3^\star}$. The iterative approach continues by solving $\text{P.1}-\text{UC}$ with the updated set $\Omega_3 = \left\{ (m,f) | \, x^{3^*}_{m,f,n} = 1, \, \forall \, (m,f,n) \in \mathcal{M} \times \mathcal{F} \times \mathcal{N} \right\}$. We develop the subsequent iteration $\mathcal{X}^{0^\star} \longrightarrow \mathcal{X}^{1^\star} \longrightarrow \mathcal{X}^{2^\star} \longrightarrow \mathcal{X}^{3^\star} \longrightarrow \ldots$, which is denoted as the *iterative Hungarian method*. The iteration process continues until there is no change in the UC policy. In fact, the output sequence of the *iterative Hungarian method* is non-decreasing and converges when $\mathcal{X}^{(t+1)^\star} = \mathcal{X}^{(t+2)^\star} = \mathcal{X}^{(t+3)^\star}$, for a $t \geq 0$ [184, Theorem 1]. Finally, the details of the *iterative Hungarian method* are described in **Algorithm** 5.

6.5.4 Passive Beamforming Optimization and Overall Complexity Analysis

6.5.4.1 Passive Beamforming Optimization

For a given PA scheme $\mathcal{P}^*$ and a UC policy $\mathcal{X}^\star$, the optimal PS matrix for the RIS can be obtained by solving the following optimization problem.

$$\text{P}-\text{PS}: \quad \max_{\boldsymbol{\theta}} \sum_{f\in\mathcal{F}} \mathcal{R}_f \tag{6.46a}$$

$$\text{s.t.} \quad (6.34\text{d}), (6.34\text{j}). \tag{6.46b}$$

Algorithm 5: Proposed Iterative Hungarian Method

Input: $\{\mathcal{R}_{m,f,n}\,|\forall m \in \mathcal{M}, f \in \mathcal{F}, n \in \mathcal{N}\}$ and the maximum number of iterations J_2;
Initialize: Iteration index $t = 1$ and UC policy $\mathcal{X}^{0^\star}$ feasible to (6.42);
Compute: $\mathcal{X}^{1^\star}, \mathcal{X}^{2^\star}$, and $\mathcal{X}^{3^\star}$;
if $\mathcal{X}^{1^\star} = \mathcal{X}^{2^\star} = \mathcal{X}^{3^\star}$ **then**
  $\mathcal{X}^\star = \mathcal{X}^{3^\star}$;
end if
else
  $t = 1$ and Check $= 0$;
  while *(*$t \leq J_2$ and Check $= 0$*)* **do**
    Find Ω_{3t} based on $\mathcal{X}^{3t^\star}$ and solve P.1 − UC using the *Hungarian method* to obtain $\mathcal{X}^{3t+1^\star}$;
    if $\mathcal{X}^{3t-1^\star} = \mathcal{X}^{3t^\star} = \mathcal{X}^{3t+1^\star}$ **then**
      $\mathcal{X}^\star = \mathcal{X}^{3t+1^\star}$ and Check $= 1$;
    end if
    else
      Find Ω_{3t+1} based on $\mathcal{X}^{3t+1^\star}$ and solve P.2 − UC using the *Hungarian method* to obtain $\mathcal{X}^{3t+2^\star}$;
      if $\mathcal{X}^{3t^\star} = \mathcal{X}^{3t+1^\star} = \mathcal{X}^{3t+2^\star}$ **then**
        $\mathcal{X}^\star = \mathcal{X}^{3t+2^\star}$ and Check $= 1$;
      end if
      else
        Find Ω_{3t+2} based on $\mathcal{X}^{3t+2^\star}$ and solve P.3 − UC using the *Hungarian method* to obtain $\mathcal{X}^{3t+3^\star}$;
        if $\mathcal{X}^{3t+1^\star} = \mathcal{X}^{3t+2^\star} = \mathcal{X}^{3t+3^\star}$ **then**
          $\mathcal{X}^\star = \mathcal{X}^{3t+3^\star}$ and Check $= 1$;
        end if
      end if
    end if
    $t = t + 1$;
  end while
end if
Output: UC policy $\mathcal{X}^\star$;

Since the RIS is located at the edge of the cell to improve the CEUs' performance, the objective of (6.46) is to maximize only the sum-rate of the CEUs. It can be easily seen that P − PS has a non-concave objective function and non-convex constraints, i.e., (6.34d), (6.34j), which is also challenging to be solved. To overcome this issue, we first reformulate problem P − PS into a rank-one constrained optimization problem through change-of-variables and matrix lifting. Then, we design a DC representation for the rank-one constraint. Finally, an efficient SCA algorithm is developed to solve the resulting problem.

6.5.4.2 Rank-One Constraint Optimization problem

Let us define $\boldsymbol{v} \triangleq [\phi_1(\theta_1), \ldots, \phi_L(\theta_L)]^H$ and by applying the change of variables $\boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{1,R} = \boldsymbol{v}^H \boldsymbol{\Phi}$, where $\boldsymbol{\Phi} = \text{diag}(\boldsymbol{h}_{R,f}^H) \boldsymbol{h}_{1,R} \in \mathbb{C}^{L \times 1}$ and $\boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{2,R} = \boldsymbol{v}^H \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = \text{diag}(\boldsymbol{h}_{R,f}^H)\ \boldsymbol{h}_{2,R} \in \mathbb{C}^{L \times 1}$, we obtain

$$\begin{aligned} |h_{1,f} + \boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{1,R}|^2 &= |h_{1,f} + \boldsymbol{v}^H \boldsymbol{\Phi}|^2, \text{ and} \\ |h_{2,f} + \boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{2,R}|^2 &= |h_{2,f} + \boldsymbol{v}^H \boldsymbol{\Psi}|^2. \end{aligned} \tag{6.47}$$

By defining an auxiliary variable τ, then an equivalent representation of the achievable rate R_f of CEU_f, $\forall f \in \mathcal{F}$, can be obtained as follows,

$$\mathcal{R}_f = \log \Bigg[1 + \frac{\alpha_{1,m,f}^* P_1 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{1,f} \right) + |h_{1,f}|^2 \right) + \alpha_{2,n,f}^* P_2 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{2,f} \right) + |h_{2,f}|^2 \right)}{(1 - \alpha_{1,m,f}^*) P_1 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{1,f} \right) + |h_{1,f}|^2 \right) + (1 - \alpha_{2,n,f}^*) P_2 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{2,f} \right) + |h_{2,f}|^2 \right) + \sigma^2} \Bigg], \tag{6.48}$$

where

$$\boldsymbol{Q}_{1,f} = \begin{bmatrix} \boldsymbol{\Phi} \boldsymbol{\Phi}^H & \boldsymbol{\Phi} h_{1,f}^H \\ h_{1,f} \boldsymbol{\Phi}^H & 0 \end{bmatrix}, \quad \boldsymbol{Q}_{2,f} = \begin{bmatrix} \boldsymbol{\Psi} \boldsymbol{\Psi}^H & \boldsymbol{\Psi} h_{2,f}^H \\ h_{2,f} \boldsymbol{\Psi}^H & 0 \end{bmatrix}, \tag{6.49}$$

and $\bar{\boldsymbol{v}}^H = [\boldsymbol{v}, \quad \tau]$, in which we used the fact that $\bar{\boldsymbol{v}}^H \boldsymbol{Q}_z \bar{\boldsymbol{v}} = \text{tr}\left(\boldsymbol{Q}_z \bar{\boldsymbol{v}} \bar{\boldsymbol{v}}^H \right)$ for all $z \in \{\{1, f\}, \{2, f\}\}$. In addition, let $\boldsymbol{U} \triangleq \bar{\boldsymbol{v}} \bar{\boldsymbol{v}}^H$, which satisfies $\text{rank}(\boldsymbol{U}) = 1$ and $\boldsymbol{U} \succeq \boldsymbol{0}$. It is worth to mention that in (6.48), the PA coefficients $\alpha_{1,m,f}^*$ and $\alpha_{2,n,f}^*$ represent the optimal PA of the *coordinated NOMA cluster* (CCU_m^1, CEU_f, CCU_n^2) after obtaining the optimal UC policy $\mathcal{X}^\star$. Maximizing the CEUs sum-rate in their current form still implies a non-concave objective function with non-convex constraints. By following the same change of variables approach proposed in [186], we define the following exponential slack variables

$$\begin{aligned} e^{y_f} &= P_1 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{1,f} \right) + |h_{1,f}|^2 \right) + P_2 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{2,f} \right) + |h_{2,f}|^2 \right) + \sigma^2, \\ e^{q_f} &= (1 - \alpha_{1,m,f}^*) P_1 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{1,f} \right) + |h_{1,f}|^2 \right) + (1 - \alpha_{2,n,f}^*) P_2 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{2,f} \right) + |h_{2,f}|^2 \right) + \sigma^2. \end{aligned} \tag{6.50}$$

Afterwards, by applying simple exponential and logarithmic properties, the optimization problem of the PS matrix can be rewritten as follows.

$$\text{OPT}-\text{PS}: \quad \max_{\boldsymbol{U}, \boldsymbol{y}, \boldsymbol{q}} \sum_{f \in \mathcal{F}} (y_f - q_f) \tag{6.51a}$$

$$\text{s.t.} \quad y_f - q_f \geq R_f^{\text{th}}, \qquad \forall f \in \mathcal{F}, \tag{6.51b}$$

$$P_1 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{1,f} \right) + |h_{1,f}|^2 \right) + P_2 \left(\text{tr}\left(\boldsymbol{U} \boldsymbol{Q}_{2,f} \right) + |h_{2,f}|^2 \right) + \sigma^2 \geq e^{y_f}, \quad \forall f \in \mathcal{F}, \tag{6.51c}$$

$$(1-\alpha^*_{1,m,f})P_1\left(\text{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f})+|h_{1,f}|^2\right)+(1-\alpha^*_{2,n,f})P_2\left(\text{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f})+|h_{2,f}|^2\right)+\sigma^2$$
$$\leq e^{q_f}, \qquad \forall f\in\mathcal{F}, \tag{6.51d}$$
$$\boldsymbol{U}\succeq 0, \tag{6.51e}$$
$$[\boldsymbol{U}]_{l,l}=1, \qquad \forall\, l\in[\![1,L+1]\!], \tag{6.51f}$$
$$\text{rank}(\boldsymbol{U})=1, \tag{6.51g}$$

where $\boldsymbol{y}=[y_1,\ldots,y_F]^T$ and $\boldsymbol{q}=[q_1,\ldots,q_F]^T$. After these mathematical manipulations, OPT − PS is still a non-convex optimization problem due to the non-convexity of constraint (6.51d) and of the rank-one constraint in (6.51g). In order to tackle the non-convexity of constraint (6.51d), we use SCA to perform a first-order Taylor approximation as $e^{q_f}=e^{\bar{q}_f}(q_f-\bar{q}_f+1)$, where the linearization is made around a point $\bar{q}_f$ that satisfies the constraints of problem (6.51). Consequently, problem (6.51) can be rewritten as

$$\text{OPT}-\text{PS}: \max_{\boldsymbol{U},\boldsymbol{y},\boldsymbol{q}} \sum_{f\in\mathcal{F}}(y_f-q_f) \tag{6.52a}$$
$$\text{s.t. } (1-\alpha^*_{1,m,f})P_1\left(\text{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f})+|h_{1,f}|^2\right)+(1-\alpha^*_{2,n,f})P_2\left(\text{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f})+|h_{2,f}|^2\right)+\sigma^2$$
$$\leq e^{\bar{q}_f}(q_f-\bar{q}_f+1), \quad \forall\, f\in\mathcal{F}, \tag{6.52b}$$
$$(6.51\text{b}),(6.51\text{c}),(6.51\text{e})-(6.51\text{g}). \tag{6.52c}$$

Considering the rank-one constraint, one widely adopted technique in the literature to deal with such constraint is the SDR technique. By dropping the rank-one constraint, the resulting optimization problem ends up to a convex SDP, which can be efficiently solved by existing convex optimization solvers such as CVX [131]. However, when the dimension of the optimization variable is large, the SDR method has a low probability of returning a rank-one solution and it often fails to reach it [142, 187]. Hence, the Gaussian randomization should be applied after the SDR technique to obtain a rank-one sub-optimal solution [142, 187]. Unfortunately, the obtained sub-optimal solution may not be feasible to meet the QoS constraints [142,187], and hence, it causes early stopping in the AO approach. To circumvent this issue, we design instead a DC representation for this rank-one constraint, which is detailed in the following part.

6.5.4.3 DC Representation for Rank-One Constraints

The key idea behind the DC representation can be presented as follows. For a PSD matrix $\boldsymbol{U}\in\mathbb{C}^{L+1\times L+1}$, the rank-one constraint indicates that $\sigma_1(\boldsymbol{U})>0$ and $\sigma_l(\boldsymbol{U})=0$, for all $l\in[\![2,L+1]\!]$, where for $l\in[\![1,L+1]\!]$, $\sigma_l(\boldsymbol{U})$ denotes the lth largest eigenvalue value

of $\boldsymbol{U}$. Based on this, the rank-one constraint can be represented through the following proposition.

***Proposition* 4.** *For a PSD matrix $\boldsymbol{U} \in \mathbb{C}^{L+1\times L+1}$ satisfying $\mathrm{tr}(\boldsymbol{U}) > 0$, we have*

$$\mathrm{rank}(\boldsymbol{U}) = 1 \Leftrightarrow \|\boldsymbol{U}\|_* - \|\boldsymbol{U}\|_2 = 0, \tag{6.53}$$

where $\|\boldsymbol{U}\|_2 = \sigma_1(\boldsymbol{U})$ and $\|\boldsymbol{U}\|_* = \sum_{l=1}^{L+1} \sigma_l(\boldsymbol{U})$ represent the spectral and the nuclear norms of $\boldsymbol{U}$, respectively.

Proof. The complete proof can be found in [142, Proposition 3]. ■

According to **Proposition 4**, the non-convex rank-one constraint imposed on $\boldsymbol{U}$ in (6.51g) can be equivalently replaced by a DC function that can be added as a penalty term to the objective function in (6.52a) of problem OPT − PS [126]. Consequently, problem OPT − PS can be expressed as

$$\mathrm{PS}-1: \quad \max_{\boldsymbol{U},\boldsymbol{y},\boldsymbol{q}} \ \sum_{f\in\mathcal{F}}(y_f - q_f) - \varsigma\left(\|\boldsymbol{U}\|_* - \|\boldsymbol{U}\|_2\right) \tag{6.54a}$$

$$\text{s.t.} \quad (6.52\text{b}), (6.51\text{b}), (6.51\text{c}), (6.51\text{e}), (6.51\text{f}), \tag{6.54b}$$

where $\varsigma > 0$ is a penalty factor that forces the equality constraint for the rank-one constraint [188]. Therefore, when the penalty term reaches zero, i.e., $(\|\boldsymbol{U}\|_* - \|\boldsymbol{U}\|_2) = 0$, a rank-one solution can be obtained [188]. Similar to [188], we can initialize the penalty coefficient ς with a small value and gradually increase it until the penalty term reaches zero as follows: $\varsigma = \omega\varsigma$, where $\omega > 1$ [188]. Furthermore, since $\|\cdot\|_2$ is a convex function, the problem in (6.54) is still non-convex optimization problem. Here, we can apply SCA to solve it in an iterative manner. Particularly, by linearizing the convex term, all what remains is to solve the following optimization problem.

$$\mathrm{PS}-2: \quad \max_{\boldsymbol{U},\boldsymbol{y},\boldsymbol{q}} \ \sum_{f\in\mathcal{F}}(y_f - q_f) - \varsigma\left(\|\boldsymbol{U}\|_* - \left\langle \partial_{\boldsymbol{U}^{[r-1]}} \left\|\boldsymbol{U}^{[r-1]}\right\|_2, \boldsymbol{U}\right\rangle\right), \tag{6.55a}$$

$$\text{s.t.} \quad (6.52\text{b}), (6.51\text{b}), (6.51\text{c}), (6.51\text{e}), (6.51\text{f}), \tag{6.55b}$$

where $\boldsymbol{U}^{[r-1]}$ is the obtained solution at iteration $r-1$, $\partial_{\boldsymbol{U}^{[r-1]}} \left\|\boldsymbol{U}^{[r-1]}\right\|_2$ is the sub-gradient of the spectral norm at the point $\boldsymbol{U}^{[r-1]}$, and $\langle ., . \rangle$ denotes the inner product that is given by $\langle \boldsymbol{V}, \boldsymbol{Z} \rangle = \Re(\mathrm{tr}(\boldsymbol{V}^H \boldsymbol{Z}))$. Note that $\partial_{\boldsymbol{U}^{[r-1]}} \|\boldsymbol{U}\|_2$ can be efficiently evaluated as $\boldsymbol{u}_1 \boldsymbol{u}_1^H$, where $\boldsymbol{u}_1$ is the eigenvector corresponding to the largest singular value $\sigma_1(\boldsymbol{U})$ [142]. Given an initial value of $\boldsymbol{U}^{[0]}$ and by iteratively solving problem (6.55) until the penalty term reaches zero, we can obtain an exact rank-one solution for the PS matrix. One practical stopping criterion for problem (6.55) is given by $\|\boldsymbol{U}\|_* - \|\boldsymbol{U}\|_2 < \epsilon_{\mathrm{DC}}$, where $\epsilon_{\mathrm{DC}} > 0$ is a sufficiently small constant [142].

Algorithm 6: Proposed Algorithm for the PS Matrix Design

Input: $\boldsymbol{Q}_{1,f}, \boldsymbol{Q}_{2,f}, \mathcal{P}^*, \mathcal{X}^\star, P_1, P_2, \sigma^2, R_f^{\text{th}}$, maximum number of iteration J_3, iteration index $r = 1$, and maximum tolerance ϵ and ϵ_{DC};

while *($r \leq J_3$)* **do**

  if $\max_f \left[\bar{q}_f^{[r]} - \bar{q}_f^{[r-1]}\right] \leq \epsilon$ and penalty term $\left\|\boldsymbol{U}^{[r]}\right\|_* - \left\|\boldsymbol{U}^{[r]}\right\|_2 \leq \epsilon_{\text{DC}}$ **then**

    break;

  end if

  else

    Compute the subgradient $\partial_{\boldsymbol{U}^{[r-1]}} \|\boldsymbol{U}\|_2$ and obtain a solution $\boldsymbol{U}^{[r]}$ by solving (6.55);

    Update $\bar{\boldsymbol{q}}^{[r]} := \boldsymbol{q}^{[r]}$;

    Update $\varsigma^{[r+1]} = \omega\varsigma^{[r]}$;

    Increment $r := r + 1$;

  end if

end while

Obtain $\bar{\boldsymbol{u}}^\star$ using Cholesky decomposition, where $\boldsymbol{U}^{[r]} = \bar{\boldsymbol{v}}^\star \left(\bar{\boldsymbol{v}}^\star\right)^H$;

Obtain the PS matrix $\boldsymbol{\Theta}^* = \text{diag}\left(\left(\boldsymbol{v}^\star\right)^H\right)$, where $\boldsymbol{v}^\star = [\bar{\boldsymbol{v}}^\star]_{1:L} \,/\, [\bar{\boldsymbol{v}}^\star]_{L+1}$;

Based on the above analysis, the overall algorithm for solving the original PS matrix optimization problem $\text{P}-\text{PS}$ in (6.46) is summarized in **Algorithm** 6. In addition, the convergence behavior of the proposed iterative algorithm is discussed in the following proposition.

***Proposition* 5.** *The generated sequence* $\{\boldsymbol{U}^{[r]}\}$ *by iteratively solving problem* (6.55) *is strictly decreasing and the sequence* $\{\boldsymbol{U}^{[r]}\}$ *converges to a critical point of* $\text{PS}-2$ *from an arbitrary initial point* $\boldsymbol{U}^{[0]}$.

Proof. The complete proof can be found in [187, Proposition 1]. ■

At this point, the overall proposed AO algorithm for the network sum-rate maximization problem OPT in (6.34) for the considered RIS-aided CoMP NOMA network is summarized in Algorithm 7. It is noteworthy that according to the convergence analysis in section 6.5.3-B and ***Proposition 5***, one can deduce that the objective function of problem $\text{P}-\text{PA}$ in **Algorithm** 7 is monotonically non-decreasing over iterations. Therefore, since the objective value is upper bounded due to the limited transmit power at the BSs, **Algorithm** 7 is guaranteed to converge.

Algorithm 7: Proposed AO Algorithm for RIS-assisted CoMP NOMA networks

Initialize: phase-shift matrix for RIS $\boldsymbol{\Theta}^{(j)}$, the tolerance ξ, the maximum number of iteration J_1, network sum-rate $\mathcal{R}_{\text{Sum}}^{[j]} = 0$ and set the current iteration number as $j = 0$;

while *($j \leq J_1$ or $\mathcal{R}_{\text{Sum}}^{[j+1]} - \mathcal{R}_{\text{Sum}}^{[j]} > \xi$)* **do**

Step 1: Power allocation optimization

With a given value of $\boldsymbol{\Theta}^{[j]}$, check the feasibility conditions according to ***Theorem 6.2***;

if P − PA in (6.36) is feasible for this *coordinated NOMA cluster* **then**

Update the optimal PA for this potential *coordinated NOMA cluster* $\mathcal{P}_{m,f,n}^{[j+1]}$ based on ***Theorem 6.3***;

end if

Step 2: User clustering optimization

With given values of all possible $\mathcal{P}^{[j+1]}$ and $\boldsymbol{\Theta}^{[j]}$, update $\mathcal{X}^{[j+1]}$ via **Algorithm** 5;

Step 3: Phase-shift optimization

With given $\mathcal{P}^{[j+1]}$ and $\mathcal{X}^{[j+1]}$, update $\boldsymbol{\Theta}^{[j+1]}$ via **Algorithm** 6;

end while

Output: the solution $\mathcal{P}^{[j]}$, $\mathcal{X}^{[j]}$, and $\boldsymbol{\Theta}^{[j]}$;

6.5.5 RIS-Assisted CoMP NOMA Networks: Practical Implementation

Here, we provide the main steps for implementing RIS-assisted CoMP NOMA in a cellular network.

1. ***User association and classification:*** Each user first detects the reference signal from the serving BS and the other BS, which is in this case an interfering BS. Then, the estimated RSS measures from the two BSs are reported to the serving BS and then forwarded to the central processing unit. If the measured RSS values from the two BSs are less distinctive, this user is deemed as a CEU, and therefore, it associates with the two-BSs and the CoMP transmission is triggered. Otherwise, if the reported RSS values are highly distinctive, the user is served by only one BS. Finally, the central processing unit reports the user classification to the BSs.

2. ***Channel estimation for all the wireless links:*** In this step, the CSI for all the considered wireless links should be estimated. We rely on efficient channel estimation techniques that have been developed in the literature for RIS-enabled wireless networks. These techniques can be adopted in our proposed system model to obtain an accurate CSI [136, 137, 182].

3. ***Power allocation, user clustering, and phase-shift optimization:*** In this step, at the central processing unit, PA coefficients, UC policy, and PS matrix optimization are obtained and forwarded to the two BSs as well as the RIS controller so that the PS for each meta atom can be accordingly tuned.

4. ***User tracking and channel measurements:*** In this step, the two BSs keep tracking frequently the RF channel measurements and reports along with the estimates of the users' positions. Afterward, based on the received information, the central processing unit will check if there are any changes in either the RSS or the user positions (some users are mobile and can change from CCUs to CEUs or vice versa). If it is the case, the first three steps should be repeated.

6.5.6 Computational Complexity Analysis

To quantify the computational complexity of the proposed AO Algorithm for the considered RIS-assisted CoMP NOMA network, which is presented in **Algorithm 7**, the computational complexity of the PA optimization, the UC optimization, and the PS matrix design need to be evaluated. First, for the PA solution, since we obtained a closed-form solution as shown in ***Theorem 6.3***, the computational complexity of having the optimal achievable sum-rate for an arbitrary *coordinated NOMA cluster* is approximately $\mathcal{O}(1)$. Thus, the computational complexity for obtaining the optimal achievable sum-rates of all the possible *coordinated NOMA cluster* configurations is $\mathcal{O}(MFN)$. Then, for the UC optimization, and according to [184], the computational complexity of *iterative Hungarian method*, which is presented in **Algorithm** 5, is $\mathcal{O}(\Gamma^3)$, where $\Gamma = \max(M, F, N)$. Finally, regarding the PS matrix design, it is an SDP problem that can be solved using interior point methods [149]. Hence, based on [149, Theorem 3.12], the order of the computational complexity of an SDP problem with k SDP constraints that contain an $a \times a$ PSD matrix is given by $\mathcal{O}\left(\sqrt{a}\log\left(1/\zeta\right)\left(ka^3 + k^2a^2 + k^3\right)\right)$, where $\zeta > 0$ is the solution accuracy. For the PS matrix problem in hands, we have $a = L + 1$ and $k = L + 3F + 1$. Hence, the computational complexity of the PS matrix design is approximate $\mathcal{O}\left(J_3\log\left(1/\zeta\right)\left(L^{4.5} + F^2L^{2.5} + L^{3.5}F\right)\right)$. As a result, the overall complexity of the proposed AO algorithm is $\mathcal{O}\left(J_1\left(MFN + J_2\Gamma + J_3\log\left(1/\zeta\right)\left(L^{4.5} + F^2L^{2.5} \quad + L^{3.5}F\right)\right)\right)$, which is a polynomial time complexity.

6.5.7 Simulation Results and Discussion

In this section, the efficacy of the proposed system model and the solution approach are evaluated through the following three points. First, we validate the optimal closed-form expressions of the PA coefficients. Second, we assess the proposed UC policy, i.e. the *iterative Hungarian method*, in comparison with the optimal UC scheme that is based on the exhaustive search method. Third, we evaluate the performance of the proposed RIS-assisted CoMP NOMA networks in comparison with four benchmark schemes under different system parameters by varying the number of RIS's elements, the rate threshold at the cellular users, and the BS transmit power. These benchmark schemes are presented as follows.

1. ***The CoMP NOMA scheme***: This scheme was proposed in [9,67] and it is similar to the proposed scheme but without including the assistance of the RIS. Hence, with the objective of maximizing the network sum-rate, the optimal PA coefficients at the BSs that derived in ***Theorem 6.3*** and the proposed UC policy, which is based on the *iterative Hungarian method*, are adopted for this scheme.

2. ***Random Phase-shift Matrix***: This scheme is similar to the proposed scheme, but a random RIS's PS matrix is considered instead of the designed one. In this case, the same PA coefficients and UC policy, which are presented in Section IV, are adopted for this scheme. However, the phase-shifts of the RIS's elements are independent and uniformly generated within $[0, 2\pi]$.

3. ***The RIS-assisted CoMP OMA scheme***: This scheme was proposed in [160], where the CCUs and CEUs are allocated orthogonal radio channels. With the objective of maximizing the network sum-rate, the PA and PS matrix sub-problems are solved using the AO approach. Moreover, after applying the AO approach, the PA sub-problem is convex, and hence, it can be efficiently solved by existing convex optimization solvers such as CVX. Meanwhile, the PS matrix at the RIS can be obtained based on our method presented in Section V. Note that, the UC phase is not required.

4. ***The RIS-assisted NOMA scheme***: This scheme is studied in [165, 173] in which there is no coordination between the BSs and each cell will invoke NOMA to

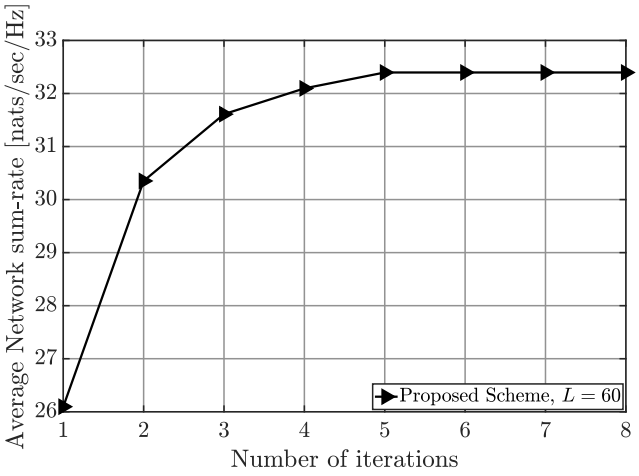


Figure 6.9: Convergence of the proposed AO algorithm.

serve the associated users by constructing a two-user cluster; meanwhile the RIS is deployed at the edge of the cell to enhance the performance of the CEUs. With the objective of maximizing the sum-rate at the first cell, and similar to our proposed scheme, we invoke the AO approach to obtain the PA, UC, and PS matrix for the first cell considering that the transmission of the other BS acts as an ICI at both the CCUs and CEUs. We derive a closed-form expression for the PA coefficients and we employ the *Hungarian method* to tackle the UC problem. On the other hand, the PS matrix at the RIS is obtained based on our method presented in Section V.

6.5.7.1 Simulation Settings

In the 3-dimensional Cartesian coordinates system, the locations of BS_1 and BS_2 are assumed to be $(0\text{m}, 0\text{m}, 10\text{m})$ and $(140\text{m}, 0\text{m}, 10\text{m})$, respectively, where the radius of each cell is assumed to be equal to 80 m. The RIS is located in the overlapped area of the two cells, precisely at $(70\text{m}, 20\text{m}, 0\text{m})$. The CCUs^1 and the CCUs^2 are randomly distributed within two discs centered at their nearby BSs, with inner and outer radii of 20 m and 40 m, respectively. Meanwhile, the CEUs are randomly distributed within the overlapped area of the two cells and at a distance of 20 m from the RIS. Unless otherwise stated, we assume that the numbers of CCUs^1, CCUs^2, and CEUs are $M = N = F = 3$ users.[3] Moreover, $\forall i \in \mathcal{I}$, $m \in \mathcal{M}$, $f \in \mathcal{F}$ $n \in \mathcal{N}$, the path-loss exponents of the wireless links are $\eta_{i,m} = \eta_{i,n} = 3.5$, $\eta_{i,R} = \eta_{R,f} = 2.2$, and $\eta_{i,f} = 4$, whereas the Racian factors

[3]In these simulation settings, we consider the number of CCUs and CEUs are equal. This assumption is due to the fact that the maximum number of potential *coordinated NOMA clusters* is the minimum between the cardinals of the three available sets of users, i.e. $\min(M, F, N)$.

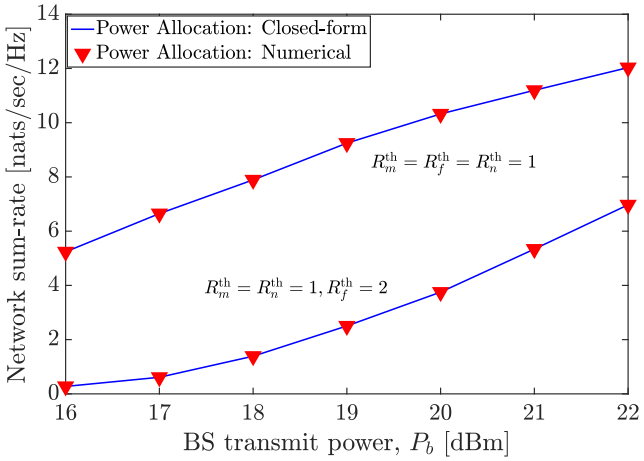


Figure 6.10: Analytical and numerical network sum-rate.

$\kappa_{i,R} = \kappa_{R,f} = 3$ dB. We assume that the transmit power by BS_1 and BS_2 are equal, where $P_1 = P_2 = P_b$ and that the noise power at the cellular users $\sigma^2 = -90$ dBm [13, 113]. All the results are obtained by averaging over 2000 channel realizations.

6.5.7.2 Convergence Analysis

Fig. 6.9 presents the convergence behavior of the proposed AO algorithm for the proposed RIS-assisted CoMP NOMA network versus the iteration index when the number of RIS elements $L = 60$, the rate thresholds of all users $R_m^{\text{th}} = R_f^{\text{th}} = R_n^{\text{th}} = 1$ nats/sec/Hz, and the transmit power at the BSs $P_b = 18$ dBm. It can be seen from this figure that the proposed AO Algorithm converges in about 5 iterations, which shows the low computational complexity of the proposed solution [13].

6.5.7.3 Validation of The Closed-form Expressions of The PA Coefficients

Fig. 6.10 presents the numerical and analytical average achievable sum-rate for one *coordinated NOMA cluster*. The analytical results are obtained by applying both ***Theorem 1*** and ***Theorem 6.3***, whereas the numerical results are acquired by solving problem $\text{P} - \text{PA}$ using one optimization solver.[4] It can be observed from Fig. 6.10 that the analytical results perfectly match the numerical results which proves the optimality of the PA expressions derived by ***Theorem 6.3***. It can be also seen from Fig. 6.10 that when the rate requirements for the CEU increases, the average achievable sum-rate per *coordinated NOMA cluster* decreases. This observation is due to the following two reasons. First,

[4]fmincon is used to solve the problem $\text{P} - \text{PA}$, which is a built-in function in MATLAB [107]. Moreover, 100 different initial points are generated in order to guarantee convergence to the optimal solution.

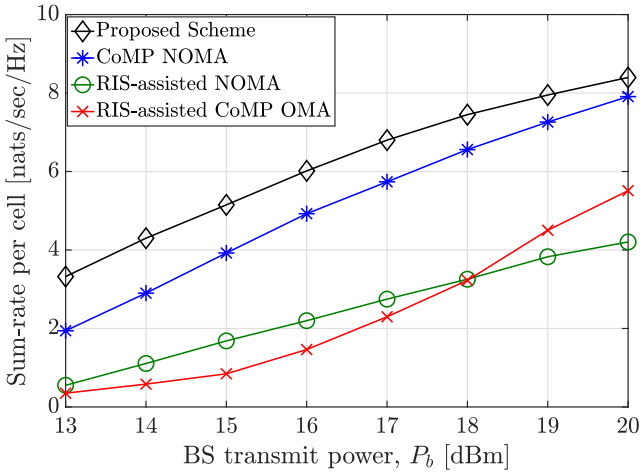


Figure 6.11: Sum-rate per cell versus the BS transmit power, when $L = 60$ elements, $R_m^{\text{th}} = R_f^{\text{th}} = R_n^{\text{th}} = 1$ nats/sec/Hz.

increasing the rate requirement at the CEU leads to increasing the failure probability of satisfying the feasibility conditions in ***Theorem 6.2***, which results in a high probability of having zero achievable sum-rate for the associated *coordinated NOMA cluster*. Second, a high rate requirement at the CEU means that high PA coefficients should be allocated by the BSs. In this case, small PA coefficients are allocated to the associated CCUs, which results in a lower achievable sum-rate per cluster.

6.5.7.4 Sum-Rate per Cell

Fig. 6.11 presents the sum-rate per cell achieved by the proposed scheme, CoMP NOMA scheme, RIS-assisted NOMA scheme, and RIS-assisted CoMP OMA scheme versus the BS transmit power considering a single *coordinated NOMA cluster*. For a fair comparison between the four schemes, the sum-rate per cell is adopted where only the achievable sum-rate of one CCU along with the CEU is considered. It can be seen from this figure that the proposed system outperforms that of the three baseline schemes. This is because of two main reasons. First, our proposed model combats the severe path-loss at the CEU by allowing the two BSs to exploit the RIS in serving this user by properly adjusting its PS matrix. Therefore, a stronger combined channel gain can be achieved. Second, our proposed model also allows cooperation between the two BSs through CoMP transmission, which can mitigate and benefit from the ICI by converting the interference signal into a useful signal. In addition, for a high transmit power at BSs, one can see that the RIS-assisted NOMA without CoMP has the worst performance. This is due to the fact that the

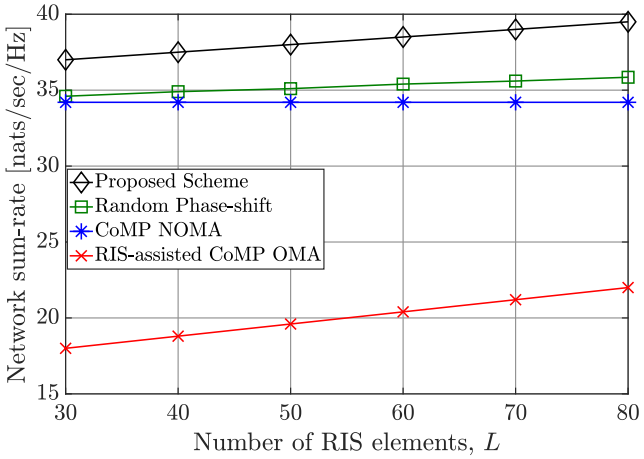


Figure 6.12: Network sum-rate versus number of RIS's elements L, where $R_m^{\text{th}} = R_f^{\text{th}} = R_n^{\text{th}} = 1$ nats/sec/Hz and $P_b = 20$ dBm.

CEU experiences severe ICI, which deteriorates its performance. In order to evaluate the performance of the proposed two-cell network, we will exclude the RIS-assisted NOMA baseline scheme in the following results.

6.5.7.5 Network Sum-Rate Performance

Fig. 6.12 presents the average network sum-rate versus the number of RIS reflecting elements. First, one can see that the network sum-rate achieved by the RIS-based schemes increases when the number of RIS's reflecting elements increases while the performance of CoMP NOMA scheme (without RIS) remains unchanged. This is due to the fact that a larger number of RIS's passive elements leads to a strong combined channel gains at the CEUs, and hence, a strong passive array gains. Second, the orthogonal allocation of the radio channel in CoMP OMA with RIS, which results in doubling the required signal-to-noise-ratio threshold, makes the CoMP OMA scheme exhibits the worst performance, which validates the necessary of invoking NOMA at each BS. For instance, when the number of reflecting elements is equal to 60, the proposed RIS-assisted CoMP NOMA scheme can achieve around 9.5%, 12.5%, and 88% gain over the random phase-shift matrix, the CoMP NOMA without RIS, and the RIS-assisted CoMP OMA schemes, respectively.

Fig. 6.13 illustrates the relationships between the average network sum-rate and the CEUs rate threshold R_f^{th}. It can be seen from this figure that the proposed scheme outperforms the other three benchmark schemes under various required rate settings. Moreover, it can be observed that when the data rate threshold increases, the system

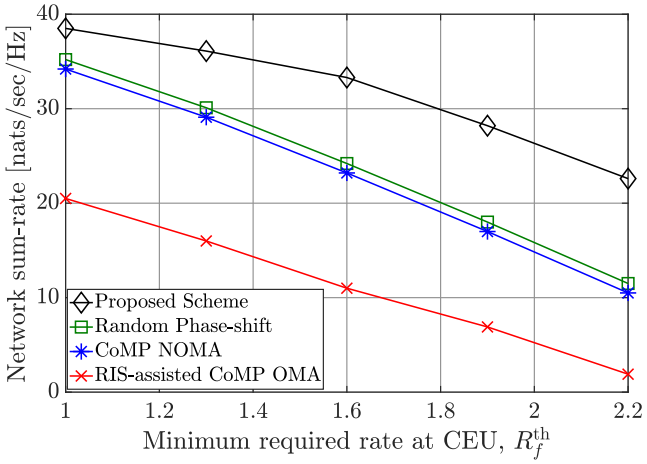


Figure 6.13: Network sum-rate versus rate threshold at CEUs, where $R_m^{\text{th}} = R_n^{\text{th}} = 1$ nats/sec/Hz, $P_b = 20$ dBm and $L = 60$.

performance in terms of the network sum-rate decreases. This can be explained as follows. In order to have a better network performance, the feasibility condition in ***Theorem 6.2*** should be maintained for each possible *coordinated NOMA cluster*. This can be attained by increasing the transmit power, by having low required QoS, and/or by enhancing the channel gain at the CEUs. As a result, for low rate requirements at the CEUs, more *coordinated NOMA clusters* are constructed and most of the power will be allocated to the CCUs, which translates into a higher achievable network sum-rate. In addition, for a low rate threshold at the CEUs, the minimum fraction of power allocated to the CEUs to having a successful SIC is almost enough to achieve their required QoS, and therefore, the improvement are almost coming from enhancing the achievable rate at the CEUs using RIS. This explains the small gain between the proposed RIS-assisted CoMP NOMA model and the CoMP NOMA model without RIS.

On the other hand, when the rate threshold at the CEUs increases, the probability of achieving the target QoS and SIC constraints decreases. In such conditions, most of the CEUs are not admitted to the network and most of the *coordinated NOMA clusters* are in an outage, which leads to a deterioration in the network performance. The gap between the proposed model and the model without RIS is, however, high. This is due to the enhancement the RIS brings to the channel coefficients of the CEUs, which makes the proposed model more robust against the increase in the QoS of the users, and, therefore, more cellular users can be admitted to the network. In other words, due to the integration of CoMP, NOMA, and RIS technology, the proposed model provides higher connectivity

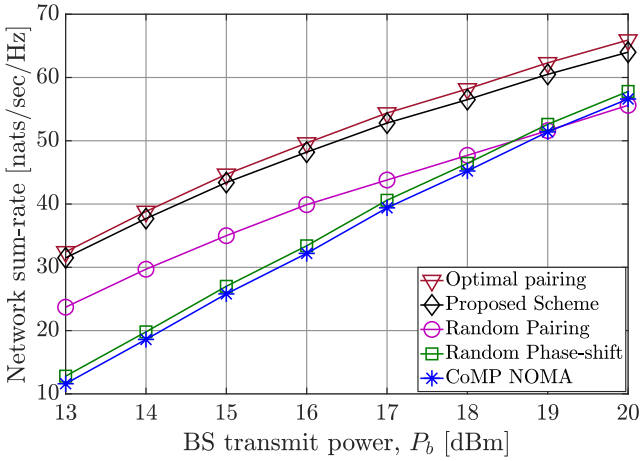


Figure 6.14: Network sum-rate versus transmit power, where $R_m^{\text{th}} = R_f^{\text{th}} = R_n^{\text{th}} = 1$ nats/sec/Hz, $M = F = N = 4$, and $L = 60$.

compared to the benchmark schemes.

Fig. 6.14 depicts the average network sum-rate versus the maximum transmit power at the BSs for different UC policies. In order to validate the performance of the proposed UC policy, i.e., the *iterative Hungarian method*, we compare its performance with two other UC schemes, which can be explained as follows.

- The *optimal clustering scheme*: In this scheme, the UC solution is obtained through an exhaustive search over all possible *coordinated NOMA clusters*. In fact, using the exhaustive search, all the possible combinations of (CCU_m^1, CEU_f, CCU_n^2), for all $m \in \mathcal{M}$, $f \in \mathcal{F}$, $n \in \mathcal{N}$ are tested. Then, the optimal UC policy is selected.

- The *random clustering scheme:* In this scheme, a random UC policy is randomly generated first to construct different *Coordinated NOMA clusters*. Then, the PA and the PS matrix sub-problems are solved based on the proposed AO approach.

Fig. 6.14 shows that the performance of proposed *iterative Hungarian method* is very close to the one of the optimal UC policy. For instance, when the BSs power $P_b = 16$ dBm, the proposed scheme gets around 97% of the network sum-rate achieved by the optimal UC policy, which validates the high performance of the proposed UC policy. It is noteworthy that we have tested the performance of the proposed UC policy for large numbers of users and we concluded that it achieved the same efficiency compared to the optimal UC scheme. It can be also observed from Fig. 6.14 that the random clustering scheme achieves a better performance than the random phase-shift scheme in most of the

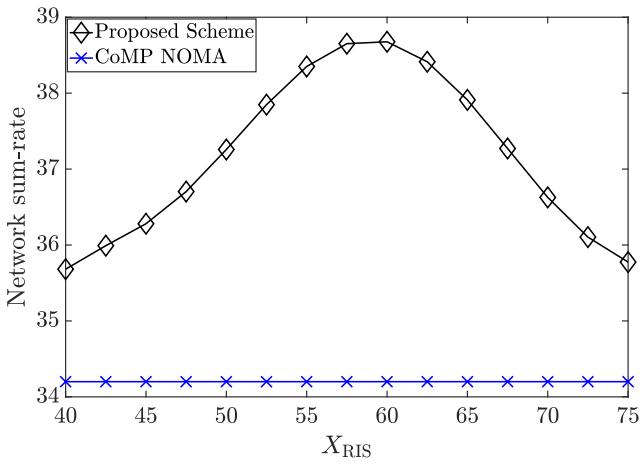


Figure 6.15: Illustration of the impact of the RIS location on the network sum-rate, when $L = 60$, $R_m^{\text{th}} = R_f^{\text{th}} = R_n^{\text{th}} = 1$ nats/sec/Hz.

BS transmit power values. This demonstrated that optimizing the PS matrix at the RIS is dominant over-optimizing the UC policy.

It can be also seen from Fig. 6.14 that the network sum-rate is improved when the BS transmit power increases. However, one can remark that the gain between the proposed RIS-assisted CoMP NOMA model and the CoMP NOMA model without RIS is reduced. The reason behind such behavior is that enhancing the network performance mainly comes from improving the performance of the CCUs. Thus, at low transmit power, enhancing the channel gain of the CEUs by properly adjusting the PS of the RIS leads to reduce in the allocated power at the BS to the CEUs, which reflects on improving the CCUs performance, and hence, an improved network performance. However, when the BS transmit power increases, the minimum allocated power to maintain the NOMA and SIC principles is almost enough to achieve the target QoS constraints at the CEUs. Therefore, the effect of enhancing the channel gains of the CEUs through the RIS on the CCUs performance will reduce. As a result, at high transmit power, most of the gain between the system with and without RIS comes from the enhancement of the CEUs performance.

Fig. 6.15 presents the network sum-rate versus the location of the RIS X_{RIS} for CoMP NOMA network with and without the assist of the RIS. First, we set the coordinate of the RIS as (X_{RIS} m, 20 m, 0 m). One can see that the best placement for the RIS is within the overlapped area between the two cells, where the CEUs are distributed. This is basically due to the fact that the combined channel gains between the two BSs and the CEUs

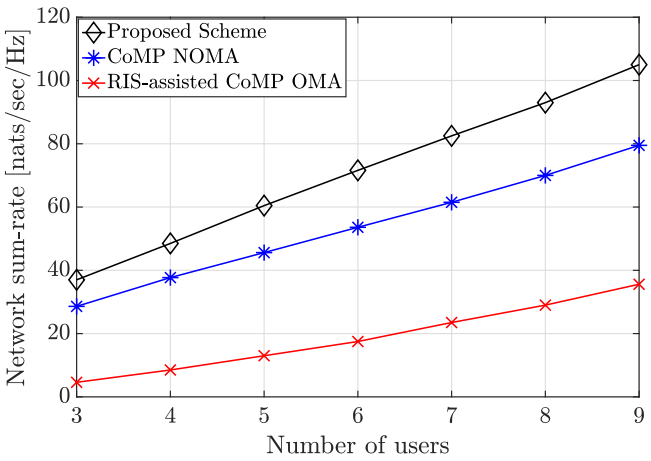


Figure 6.16: Network sum-rate vs the number of users, when $L = 60$, $R_m^{\text{th}} = R_f^{\text{th}} = R_n^{\text{th}} = 1.5$ nats/sec/Hz and $P_b = 20$ dBm.

will be boosted when the RIS is deployed near to the users-of-interest, i.e. the CEUs. Meanwhile, when the RIS is moved to the coverage areas of BS_1 or BS_2, the performance will be degraded. This is due to the combined channel gain with both BSs gets worse. Fig. 6.16 shows the effect of increasing the number of users (M, F, and N) on the network sum-rate performance achieved by the proposed scheme, the CoMP NOMA scheme, and the RIS-assisted CoMP OMA scheme. Note that, in this figure, we assume that number of CCUs^1, CCUs^2, and CEUs are equal. It is expected that when the number of CCUs^1, CCUs^2, and CEUs increases, the number of potential constructed *coordinated NOMA clusters* increases. This explains the improvement of the network sum-rate performance for the three schemes. Moreover, the proposed scheme provides the best performance due to the performance gains brought by integrating CoMP, NOMA, and RIS.

6.6 Summary

In this Chapter, we proved that the amalgamation between these three technologies, RIS, CoMP, and NOMA can form a promising paradigm for 6G networks due to their effectiveness of enhancing the spectral efficiency. First, we derived a closed-form expression for NSE, and then, the derived analytical results are assessed by Monte Carlo simulations and benchmarked with three baselines proposed in the literature. The obtained results demonstrated that the proposed RIS-empowered CoMP NOMA outperforms the state-of-the-art schemes in terms of both the CEU performance and the NSE.

Afterwards, with the goal of enhancing the performance of cellular systems in terms of

sum-rate, we investigated the joint PA, UC, and passive beamforming in an RIS-assisted CoMP NOMA network, which was formulated as an MINLP optimization problem. By invoking the AO technique, the non-convex network sum-rate optimization problem is decomposed into two sub-problems, a joint PA and UC optimization sub-problem and a PS optimization sub-problem, which are solved alternately. Finally, simulation results under various system parameter settings validated that through reconfiguring the wireless communication environment in favor of the CEUs, the RIS has the ability to improve the system performance in terms of the network sum-rate. Moreover, the proposed model with the assistance of the RIS provided significant gain compared to the same model without RIS when the transmit power at the BSs is low. Finally, it has been shown that the proposed RIS-assisted CoMP NOMA can achieve around 88% gain over the RIS-assisted CoMP OMA scheme.

It is important to mention that we consider in this work an accurate CSI, a perfect SIC process, and efficient capacity fronthaul links in order to present a proof of concept and to gain some insights from the proposed scheme. However, in practice, obtaining an accurate CSI requires a huge training signaling/overhead, the realization of a perfect SIC requires a complex hardware design at the users and the fronthaul capacity is limited. Therefore, considering an imperfect CSI and SIC along with limited capacity fronthaul links is indeed an interesting problem that can be considered as a potential future research direction in order to highlight the trade-off between the performance gain that can be achieved and the aforementioned challenges.

Chapter 7
Conclusions and Future Works

7.1 Conclusion

Unlike earlier generations of cellular networks, future wireless networks under different labels such as B5G or 6G are expected to seamlessly and ubiquitously connect everything, support very high data rates, and diverse requirements on reliability, latency, and capability of devices to support a myriad of applications such as augmented or virtual reality, ultra-high-definition movies, IoE and smart cities. This urges the development of unconventional solutions such as C-RAN, network densification, NOMA, FD communication, RIS, to name a few. NOMA has been deemed as a key enabler access technique for the 6G networks because of its ability to provide high spectral efficiency, low latency, and massive connectivity. In contrast to OMA, NOMA can simultaneously serve multiple users using the same transmission block (same frequency channel, time, and/or spatial domain) with the cost of intra-cell interference. Despite its numerous advantages, there are still several challenges that should be tackled. This thesis addressed the fundamentally challenging problems for integrating NOMA with other B5G/6G enabling technologies in future wireless networks in order to achieve high spectral efficiency, and energy-efficient communications, as well as increase the number of connected devices. A track record of publications resulting from the conducted research in this thesis indicates that this thesis has made significant contributions toward next-generation multiple access. We briefly present these achievements:

1. We have introduced analytical frameworks to investigate the application of CoMP transmission in C-NOMA/NOMA-based cellular networks. We invoked tools from stochastic geometry to derive analytical and closed-form expressions for CEU's outage probability and network spectral efficiency. Results obtained via both analyses and Monte-Carlo simulations demonstrated the efficacy of this integration compared to other baseline schemes.

2. We studied the joint user pairing and power control problem with the objective of maximizing the network sum-rate while guaranteeing the QoS requirement for the UEs, the SIC constraints, and the UEs power budget constraints. We showed that the proposed CoMP with FD C-NOMA outperformed the other two baseline schemes, especially when the SI channel gains is not sufficiently large regardless of the value of the transmit powers at the BS and the near UEs. Meanwhile, when the SI value is sufficiently large, it is recommended to use either CoMP with HD C-NOMA or CoMP NOMA as an access technique depending on the rate requirements, the transmit power at the BSs, and the transmit power at the near NOMA UEs. When the power transmit at the BS is sufficiently larger than the power transmit at the near UEs or when the rate requirement is sufficiently low, it is recommended to consider CoMP NOMA as a multiple access technique (no cooperation is required), otherwise, CoMP-assisted HD C-NOMA should be adopted.

3. We investigated the synergistic integration between RIS and HD/FD C-NOMA in order to enhance the network energy consumption toward "green" wireless networks. We demonstrated that the network energy consumption can be reduced by 50% compared to the current cellular system. Moreover, we showed that the potential gains brought by the proposed model can be significantly improved by increasing the number of the RIS elements and/or a well-optimized RIS location. More importantly, we showed that the optimal location of the RIS depended on the adopted multiple access technique at the BS.

4. We investigated the amalgamation between NOMA, CoMP, and RIS to improve and boost the performance of CEU in multi-cell wireless networks. We showed that this combination can provide a promising paradigm for the upcoming 6G networks. Simulation results under various system parameter settings validated that through controlling the wireless communication environment in favor of CEUs, RIS is able to improve system performance in terms of network sum-rate. Moreover, the proposed model with the assistance of the RIS provided significant gains compared to the same model without RIS, especially when the transmit power at the BSs is low.

5. We tackled the intensive computational complexity required to solve the proposed formulated optimization problems by deriving closed-form expressions for the power

allocation coefficients as well as the transmit power at the relaying user which has low computational complexity. In addition, regarding the other optimization variables such as passive beamforming at the RIS, we derived novel transformations and approximations to formulate convex optimization problems. Finally, we provided a low-computational complexity approach to tackle the user pairing problem in multi-cell NOMA networks in which the intertwined between the user pairings in the different cells has been considered for the first time.

7.2 Future Works

In this section, we shed light on some potential directions of research to further extend the derived results of this thesis:

1. **Channel Estimation for RIS-Empowered 6G Wireless Networks**
   To realize the full potential of the integration of RIS in NOMA-based cellular networks, there are several challenges that must be tackled. One fundamental challenge is channel estimation. In practice, to attain the promised performance gains by the RIS, accurate CSI acquisition is required [62]. Specifically, a proper configuration for the RIS's elements and the active beamforming at the BS critically depends on the CSI. Most of the prior works in the context of RIS focused primarily on designing efficient schemes for the RIS phase-shift matrix and implicitly assumed that the CSI of all wireless links is available at the BS, either totally or partially. There are two main reasons why the CSI acquisition is particularly challenging in RIS-aided wireless networks: 1) Unlike traditional transceiver architectures, an RIS is not equipped with transceiver chains. Therefore, conventional channel estimation techniques cannot be applied; 2) In order to achieve an adequate gain from a RIS, a sufficient number of elements is required. Consequently, a large number of channel coefficients needs to be estimated with reasonable accuracy. Therefore, a potential direction is to develop accurate, low-complexity, and low-training overhead channel estimation techniques.

2. **RIS Deployment Strategy**
   RIS deployment for multi-RIS is another fundamental and crucial problem in NOMA-based multi-cell wireless networks. However, the RIS position determination prob-

lem is not well-investigated in the literature. For a practical network setup, it is essential to study this critical problem due to the following reasons. 1) The RIS performance is very sensitive to the deployment location due to the severe path-loss experienced (the cascaded path-loss model). 2) The signal strength, at higher frequencies, is severely attenuated due to the blockage. Thus, the RIS should be deployed with a LoS with both the UE and the BS. 3) The more the number of deployed RISs, the more the signaling overhead due to the information exchange between the BS and the multiple RISs controllers is required. Therefore, an interesting direction is to investigate the fundamental problem of the RIS deployment strategies by providing the optimal number of RISs with their optimal locations, the active beamforming at the BS, and the passive beamforming (phase-shift matrix) at the RISs to enhance the network coverage performance.

3. **Rate Splitting Multiple Access for 6G Wireless Networks**

   Rate-Splitting Multiple Access (RSMA) has been envisioned as a contender non-orthogonal transmission mechanism for the next-generation wireless networks [189]. RSMA is a generalized joint framework of the space division multiple access (SDMA) and the NOMA and is capable of outperforming both SDMA and NOMA in terms of network energy efficiency, network spectral efficiency, and multiplexing gain [189]. Specifically, the main concept of RSMA is to split the user messages into common and private parts wherein the common parts are encoded into common streams, meanwhile, the private parts are encoded into private streams [189]. Note that the common streams are required to be decoded by multiple users whereas the private streams are needed to be decoded by the corresponding users [190–192]. Accordingly, RSMA provides a versatile interference management feature that allows for partial decoding and partial treatment of interference as noise [193]. Much research in the literature focused their attention to study the performance of RSMA considering the single-cell scenario. However, the performance of RSMA in multi-cell wireless networks has not been well investigated in the literature, which is considered a potential research direction.

Chapter 3

A1 Proof of Lemma 3.1

We start by evaluating the probability that a typical user associates with the nearest MRRH, which can be expressed as follows,

$$\begin{aligned}\mathcal{Q}_m &= \mathbb{P}[\mathcal{B}=y_m] = \mathbb{E}_{D_{1m}}\left[\mathbb{P}\left[\frac{P_m D_{1m}^{-\alpha_m}}{P_s D_{1s}^{-\alpha_s}} \geq \theta\right]\right], \\ &= \int_{\mathbb{R}_+} \mathbb{P}\left[D_{1s} \geq (\theta P_{sm})^{1/\alpha_s} r^{\alpha_{ms}}\right] f_{D_{1m}}(r_{1m}) dr_{1m}, \\ &\overset{(a)}{=} 2\pi\lambda_m \int_{\mathbb{R}_+} r_{1m} \exp\left[-\pi\left(\lambda_m r_{1m}^2 + \lambda_s (\theta P_{sm})^{\zeta_s} r_{1m}^{2\alpha_{ms}}\right)\right] dr_{1m}, \end{aligned} \tag{A.1}$$

where (a) obtains from the null probability of 2-D Poisson process with intensity λ and area πr^2, which states that $f_R(r) = 2\pi\lambda r \exp(-\pi\lambda r^2)$ [81].

Now, we evaluate the SRRH association probability which is calculated as follows

$$\mathcal{Q}_s = \mathbb{P}[\mathcal{B}=y_s] = \mathbb{E}_{D_{1s}}\left[\mathbb{P}\left[\frac{P_s D_{1s}^{-\alpha_s}}{P_{\text{In}} D_{\text{In}}^{-\alpha_{\text{In}}}} \geq \eta\right]\right]. \tag{A.2}$$

A typical user associates with the nearest SRRH, SRRH_1, only when the received signal from that SRRH is η times the maximum between the received signal from the nearest MRRH, MRRH_1 and the received signal from the second nearest SRRH, SRRH_2. Thus, the association probability in (A.2) can be written as

$$\begin{aligned}\mathcal{Q}_s &= \mathbb{P}\left[P_s D_{1s}^{-\alpha_s} \geq \eta \max(P_s D_{2s}^{-\alpha_s}, P_m D_{1m}^{-\alpha_m})\right], \\ &\overset{(a)}{=} \int_{\mathbb{R}_+} \mathbb{P}\left[D_{2s} \geq \eta^{\frac{1}{\alpha_s}} r_{1s}\right] \mathbb{P}\left[D_{1m} \geq (\eta P_{ms})^{\frac{1}{\alpha_m}} r_{1s}^{\alpha_{sm}}\right] f_{D_{1s}}(r_{1s}) dr_{1s}, \\ &\overset{(b)}{=} \int_{\mathbb{R}_+} \int_{\eta^{\frac{1}{\alpha_s}} r_{1s}}^{\infty} 2\pi\lambda_s r_{2s} \exp\left[-\pi\lambda_s\left(r_{2s}^2 - r_{1s}^2\right)\right] \exp\left[-\pi\lambda_m (\eta P_{ms})^{\zeta_m} r^{2\alpha_{sm}}\right] f_{D_{1s}}(r_{1s}) dr_{2s} dr_{1s}, \end{aligned} \tag{A.3}$$

where (a) follows from the independence of the two-tier distribution for a given value of D_{1s} and (b) comes from the joint PDF of the nearest n RRHs, which states that $f_{D_{1s},D_{2s}}(r_{1s}, r_{2s}) = (2\pi\lambda_s)^2 r_{1s} r_{2s} \exp(-\pi\lambda_s r_{2s}^2)$ [194] as well as using the null probability for the 2-D Poisson Process. After some mathematical manipulations, we can obtain the

SRRH association probability as in (3.8).

Now, we turn our attention towards the CoMP association probability especially when a typical user associates with the best two SRRHs which receives the strongest power from. This event occurs when the nearest SRRH is the serving RRH and the dominant interference resulting from SRRH_2. This association probability can be evaluated as follows,

$$\begin{aligned}
\mathcal{Q}_{ss} &= \mathbb{P}\left[\eta P_s D_{2s}^{-\alpha_s} > P_s D_{1s}^{-\alpha_s},\ P_s D_{2s}^{-\alpha_s} > P_m D_{1m}^{-\alpha_m}\right], \\
&= \mathbb{P}\left[D_{1s} > \left(\frac{1}{\eta}\right)^{\frac{1}{\alpha_s}} D_{2s}, D_{2s} < (P_{sm})^{\frac{1}{\alpha_s}} D_{1m}^{\alpha_{ms}}\right], \\
&= \int_{\mathbb{R}_+} \int_0^{P_{sm}^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}} \int_{\left(\frac{1}{\eta}\right)^{\frac{1}{\alpha_s}} r_{2s}}^{r_{2s}} f_{D_{1s},D_{2s}}(r_{1s}, r_{2s}) f_{D_{1m}}(r_{1m}) dr_{1s} dr_{2s} dr_{1m}. \qquad \text{(A.4)}
\end{aligned}$$

After some mathematical manipulations on (A.4), the association probability $\mathcal{Q}_{ss}$ can be given as in (3.9). *This completes the proof.*

A2 Proof of Lemma 3.2

Here, we need to find expressions for the PDF of the distance(s) from the serving RRH(s) to a typical user. For $f_U(u)$, since the event $U > u$ is the event of $D_{1m} > w$ given that UE_0 associates with the MRRH_1. The probability of $U > u$ can be given as

$$\begin{aligned}
\mathbb{P}\left[U > u\right] &= \mathbb{P}\left[D_{1s} > u | \mathcal{B} = \{y_m\}\right], \\
&= \frac{1}{\mathcal{Q}_m} \mathbb{P}\left[D_{1m} > u, \frac{P_m D_{1m}^{-\alpha_m}}{P_s D_{1s}^{-\alpha_s}} > \theta\right], \\
&= \frac{1}{\mathcal{Q}_m} \int\limits_{r_{1m}>u} \mathbb{P}\left[D_{1s} > (\theta P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}\right] f_{D_{1m}}(r_{1m}) dr_{1m}, \\
&\overset{(a)}{=} \frac{2\pi\lambda_m}{\mathcal{Q}_m} \int\limits_{r_{1m}>u} r_{1m} \exp\left[-\pi\left[\lambda_m r_{1m}^2 + \lambda_s (\theta P_{sm})^{\zeta_s} r_{1m}^{2\alpha_{ms}}\right]\right] dr_{1m}, \qquad \text{(A.5)}
\end{aligned}$$

where (a) comes from the null probability of 2-D Poisson process. Hence, to obtain the PDF $f_U(u)$, we use the resulting complementary cumulative distribution function (CCDF) in (A.5). By taking the derivative of the CCDF, i.e., $1-\mathbb{P}(U > u)$, which results in (3.11).

Secondly, considering a typical user associates with the nearest SRRH only, the CCDF, in this case, can be calculated as follows.

$$\begin{aligned}
&\mathbb{P}[V > v] = \mathbb{P}[D_{1s} > v | \mathcal{B} = \{y_s\}], \\
&= \frac{1}{\mathcal{Q}_s} \mathbb{P}\left[D_{1s} > v, \frac{P_s D_{1s}^{-\alpha_s}}{P_m D_{1m}^{-\alpha_m}} > \eta, D_{2s} > (\eta)^{1/\alpha_s} D_{1s}\right],
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\mathcal{Q}_s}\mathbb{P}\left[D_{1s} > v, D_{1m} > (\eta P_{ms})^{\frac{1}{\alpha_m}} {D_{1s}}^{\alpha_{sm}}, D_{2s} > (\eta)^{1/\alpha} D_{1s}\right], \\
&= \frac{1}{\mathcal{Q}_s}\int\limits_{r_{1s}>v} \mathbb{P}\left[D_{1m} > (\eta P_{ms})^{\frac{1}{\alpha_m}} r_{1s}^{\alpha_{sm}},\ D_{2s} > (\eta)^{1/\alpha_s} r_{1s}\right] f_{D_{1s}}(r_{1s})\, dr_{1s}, \\
&\overset{(a)}{=} \frac{1}{\mathcal{Q}_s}\int\limits_{r_{1s}>v} \mathbb{P}\left[D_{1m} > (\eta P_{ms})^{\frac{1}{\alpha_m}} r_{1s}^{\alpha_{sm}}\right]\mathbb{P}\left[D_{2s} > (\eta)^{1/\alpha_s} r_{1s}, D_{1s} = r_{1s}\right] dr_{1s}, \\
&\overset{(b)}{=} \frac{2\pi\lambda_s}{\mathcal{Q}_s}\int\limits_{r_{1s}>v} r_{1s}\exp\left[-\pi\left(\lambda_s\eta^{\zeta_s} r_{1s}^2 + \lambda_m\left(\eta P_{ms}\right)^{\zeta_m} r_{1s}^{2\alpha_{sm}}\right)\right] dr_{1s},
\end{aligned} \tag{A.6}
$$

where (a) follows from the independence of the distribution of the MRRHs and SRRHs. In addition, (b) obtains based on the null probability of the 2-D Poisson process and the joint PDF from a typical user to the SRRH_1 and SRRH_2. Since the CDF $F_V(v) = 1 - \mathbb{P}[V > v]$, the PDF $f_V(v)$ can be given using $f_V(v) = -\frac{d}{dv}\mathbb{P}[D_{1s} > v|\mathcal{B} = \{y_s\}]$. Thus, we can obtain the PDF of the distance to the serving SRRH as in (3.12).

Concerning the CoMP-UEs, we need to evaluate the joint distances distribution for the serving RRHs. Firstly, we start by finding the distribution of the joint distances to the nearest RRH from each tier, i.e., $f_{U,V}(u,v)$. In fact, we know that if the distance to the MRRH_1 is D_{1m}, the distance to the SRRH_1 D_{1s} is bounded as follows

$$
\left(\frac{P_{sm}}{\eta}\right)^{\frac{1}{\alpha_s}} D_{1m}^{\alpha_{ms}} \leq D_{1s} \leq \left(\theta P_{sm}\right)^{\frac{1}{\alpha_s}} D_{1m}^{\alpha_{ms}}, \tag{A.7}
$$

where, it can be obtained from (3.2) when $\mathcal{B} = \{y_m, y_s\}$. Thus, the conditional probability of $D_{1m} < u$ and $D_{1s} < v$ given that UE_0 associates with the nearest RRH from each tier can be calculated as shown in (A.8). Note that, to find the joint PDF expression, we find the derivative of the joint CDF, i.e., $\frac{\partial^2 F_{U,V}(u,v)}{\partial u \partial v}$. For computing the joint PDF in (3.13), by using the Leibniz rule, we directly calculate the joint PDF.

$$
\begin{aligned}
&\mathbb{P}[U \leq u, V \leq v] = \mathbb{P}[D_{1m} \leq w, D_{1s} \leq v|\mathcal{B} = \{y_m, y_s\}], \\
&= \frac{1}{\mathcal{Q}_{ms}}\mathbb{P}\Bigg[D_{1m} \leq u, D_{1s} \leq v, \frac{P_m D_{1m}^{-\alpha_m}}{P_s D_{1s}^{-\alpha_s}} < \theta, \frac{P_s D_{1s}^{-\alpha_s}}{P_m D_{1m}^{-\alpha_m}} < \eta, \\
&\qquad\qquad P_m D_{1m}^{-\alpha_m} > P_s D_{2s}^{-\alpha_s} D_{2s} \geq D_{1s}\Bigg], \\
&= \frac{1}{\mathcal{Q}_{ms}}\int\limits_{r_{1m}\leq u} \mathbb{P}\Bigg[D_{1s} \leq v,\ \left(\frac{P_{sm}}{\eta}\right)^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}} < D_{1s} < (\theta P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}, \\
&\qquad\qquad D_{2s} > (P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}, D_{2s} \geq D_{1s}\Bigg] f_{D_{1m}}(r_{1m}) dr_{1m}, \\
&= \frac{1}{\mathcal{Q}_{ms}}\Bigg[\int\limits_{\left(\frac{P_{ms}}{\theta}\right)^{1/\alpha_m} v^{\alpha_{sm}}}^{u} \mathbb{P}\left[\left(\frac{P_{sm}}{\eta}\right)^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}} \leq D_{1s} \leq v,\ D_{2s} > (P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}} D_{2s} \geq D_{1s}\right]
\end{aligned} \tag{A.8}
$$

$$\times f_{D_{1m}}(r_{1m})dr_{1m},$$

$$\overset{(a)}{=} \frac{1}{\mathcal{Q}_{ms}}\Bigg[\int_{\left(\frac{P_{ms}}{\theta}\right)^{1/\alpha_m} v^{\alpha_{sm}}}^{u}\Bigg[\int_{\left(\frac{P_{sm}}{\eta}\right)^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}}^{\min\left(v,(P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}\right)}\int_{(P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}}^{\infty} f_{D_{1s},D_{2s}}(r_{1s},r_{2s})dr_{1s}dr_{2s}+$$

$$\int_{(P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}}^{\min\left(v,(\theta P_{sm})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}\right)}\int_{r_{1s}}^{\infty} f_{D_{1s},D_{2s}}(r_{1s},r_{2s})dr_{1s}dr_{2s}\Bigg] f_{D_{1m}}(r_{1m})dr_{1m}\Bigg]. \tag{A.8}$$

In addition, throughout the calculation of the joint PDF, we consider that $D_{2s} \geq D_{1s}$. This is because the distance from a typical user to the SRRH_2 must be greater than the distance to SRRH_1. Moreover, in the first integral part of the (A.8) in (a), it can be observed that the received signal from the SRRH_1 is greater than the received signal from the MRRH_1 ($r_{1s} \leq (P_{ms})^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}$). This is because getting the value from the derivative of the first integral in (a), this condition $v \leq P_{sm}^{\frac{1}{\alpha_s}} r_{1m}^{\alpha_{ms}}$ must be achieved. This means that the serving RRH in this case is the SRRH_1 and the coordinated RRH is the MRRH_1. On the contrary, in the second case, the serving RRH is the MRRH_1 and the coordinated RRH is the SRRH_1 because of $v \geq (P_{ms})^{\frac{1}{\alpha_s}} u^{\alpha_{ms}}$.

Finally, for the joint PDF $f_{V,W}(v,w)$ of a typical UE's distance to the cooperating SRRH_1 and SRRH_2, we know that if the distance to the SRRH_2 is D_{2s}, the distance to the SRRH_1 is bounded as follows:

$$\left(\frac{1}{\eta}\right)^{\frac{1}{\alpha_s}} D_{2s} \leq D_{1s} \leq D_{2s}. \tag{A.9}$$

This obtains from the fact that the SRRH_2 is always farther than the SRRH_1. In addition, the received signal from the nearest SRRH is lower than η times the received signal from the second nearest SRRH. Therefore, the conditional probability of $D_{1s} \leq v$ and $D_{2s} \leq w$ given that the user associates with the two strongest SRRHs can be written as shown in (A.10).

$$\begin{aligned}
&\mathbb{P}[V \leq v, W \leq w] = \mathbb{P}[D_{1s} \leq v, D_{2s} \leq w | \mathcal{B} = \{y_s, \bar{y}_s\}], \\
&= \frac{1}{\mathcal{Q}_{ss}} \mathbb{P}\left[D_{1s} \leq v, D_{2s} \leq w, \frac{P_s D_{1s}^{-\alpha_s}}{P_s D_{2s}^{-\alpha_s}} < \eta, P_s D_{2s}^{-\alpha_s} \geq P_m D_{1m}^{-\alpha_m}\right], \\
&= \frac{1}{\mathcal{Q}_{ss}} \int_{r_{1s} \leq v} \mathbb{P}\left[D_{2s} \leq w, D_{2s} < \eta^{\frac{1}{\alpha_s}} r_{1s}, D_{2s} < (P_{sm})^{\frac{1}{\alpha_s}} D_{1m}^{\alpha_{ms}}, D_{2s} > r_{1s}\right] f_{D_{1s}}(r_{1s}) dr_{1s},
\end{aligned}$$

$$= \frac{1}{\mathcal{Q}_{ss}} \left[\int_{(1/\eta)^{\frac{1}{\alpha_s}} w}^{v} \mathbb{P}\left(D_{2s} \leq w, D_{2s} < (P_{sm})^{\frac{1}{\alpha_s}} r_{1s} | D_{1s} = r_{1s} \right) f_{D_{1s}}(r_{1s}) dr_{1s} \right],$$

$$= \frac{1}{\mathcal{Q}_{ss}} \left[\int_{(1/\eta)^{\frac{1}{\alpha_s}} w}^{v} \int_{0}^{w} \int_{(P_{ms})^{\frac{1}{\alpha_m}} r_{2s}^{\alpha_{sm}}}^{\infty} f_{D_{2s}|D_{1s}}(r_{2s}) f_{D_{1m}}(r_{1m}) f_{D_{1s}}(r_{1s}) dr_{1m} dr_{2s} dr_{1s} \right], \quad \text{(A.10)}$$

By taking the derivative with respect to v and w and using the Leibniz rule for (A.10), we get the joint PDF as in (3.14).

A3 Proof of Lemma 3.3

Here, we evaluate the interference produces from the small-tier to a typical user that associates with the nearest MRRH. $\mathcal{L}_{I_{sm}}$ is the LT of a random variable I_{sm} conditioned on the distance to the closest RRH to a typical user. The LT of I_{sm} can be evaluated as follows,

$$\begin{aligned}
\mathcal{L}_{I_{sm}}(s) &= \mathbb{E}_{I_{sm}}\left[\exp[-sI_{sm}]\right], \\
&= \mathbb{E}_{\Phi_s,\{g_{i,0}\}}\left[\exp\left[-sP_s \sum_{i\in\Phi_s} g_{i,0} R^{-\alpha_s}\right]\right] \\
&= \mathbb{E}_{\Phi_s,\{g_{i,0}\}}\left[\prod_{i\in\Phi_s} \exp\left[-sP_s g_{i,0} R_i^{-\alpha_s}\right]\right], \\
&= \mathbb{E}_{\Phi_s}\left[\prod_{i\in\Phi_s} \mathbb{E}_{\{g_{i,0}\}}\left[\exp\left[-sP_s g_{i,0} R_i^{-\alpha_s}\right]\right]\right], \\
&\overset{(a)}{=} \exp\left[-2\pi\lambda_s \int_{x>\varrho} \left(1 - \mathbb{E}_{\{g_{i,0}\}}\left[\exp\left[-sP_s g_{i,0} x^{-\alpha_s}\right]\right]\right) x dx\right], \\
&\overset{(b)}{=} \exp\left[-2\pi\lambda_s \int_{x>\varrho} \left(1 - \left[\frac{1}{1+sP_s x^{-\alpha_s}}\right]\right) x dx\right], \\
&\overset{(c)}{=} \exp\left[-2\pi\lambda_s \int_{x>(\theta P_{sm})^{\frac{1}{\alpha_s}} u^{\alpha_{ms}}} \left(\left[\frac{1}{1+(sP_s x^{-\alpha_s})^{-1}}\right]\right) x dx\right]. \quad \text{(A.11)}
\end{aligned}$$

In the above (a) follows from the probability generating functional (PGFL) of PPP [90], which states that $\mathbb{E}\left[\prod_{y\in\Phi} f(y)\right] = \exp\left(-\lambda \int_{\mathbb{R}^2} (1 - f(y)) dy\right.$, (b) follows because the interference fading received power $g_{i,0} \sim \exp(1)$, and (c) comes from the closest interference resulting from the second tier is at least at distance $\varrho = (\theta P_{sm})^{\frac{1}{\alpha_s}} u^{\alpha_{ms}}$. After some

mathematical manipulations, the Laplace transform for I_{sm} can be obtained as in (3.18). Similarly, the Laplace transform of I_{mm}, I_{ms}, and I_{ss} can be obtained.

A4 Proof of Lemma 3.4

Based on the previous steps to derive the LT for the interference, we first calculate the interference from the macro-tier to a typical user attaching to the two nearest SRRHs. This interference produces from all the MRRHs nodes in the 2-D plane except the region defined by $\mathfrak{B}(0,(P_{ms})^{\frac{1}{\alpha_m}}w^{\alpha_{sm}})$, where $\mathfrak{B}(0,\rho)$ denotes a ball of radius ρ centered at the origin. This is obtained as follows,

$$\begin{aligned}\mathcal{L}_{\bar{I}_m} &= \exp\left[-2\pi\lambda_m \int\limits_{x>\varrho}\left(1-\left[\frac{1}{1+sP_m x^{-\alpha_m}}\right]\right)xdx\right]\\ &\overset{(a)}{=} \exp\left[-2\pi\lambda_m \int\limits_{x>(P_{ms})^{\frac{1}{\alpha_m}}w^{\alpha_{sm}}}\left(\left[\frac{1}{1+(sP_m x^{-\alpha_m})^{-1}}\right]\right)xdx\right]\\ &\overset{(b)}{=} \exp\left[-2\pi\lambda_m \int\limits_{\chi>(\frac{1}{sP_s})^{\zeta_m}w^{2\alpha_{sm}}}\left(\left[\frac{1}{1+\chi^{\frac{\alpha_m}{2}}}\right]\right)d\chi\right]\end{aligned} \tag{A.12}$$

where (a) obtains from the nearest interfering MRRH to a typical user should have a distance $x > (\frac{1}{sP_s})^{\zeta_m}w^{2\alpha_{sm}}$, and (b) is acquired by executing a change of variable $\chi = (sP_m)^{-\zeta_m}x^2$. By following the same steps to obtain $\bar{I}_m$, we can obtain $\bar{I}_s$, which is provided in (3.21).

A5 Proof of Lemma 3.5

Here, the interference from small-tier to a typical user depends on which RRH is the serving RRH and which one is the coordinated RRH. When the serving RRH is the MRRH, the $\tilde{I}_m$ and $\tilde{I}_s$ come from the other MRRHs on the plane except the region defined by $\mathfrak{B}(0,u)$ and the other SRRHs except the region defined by $\mathfrak{B}(0,v)$, respectively. However, when the SRRH is the serving RRH, the $\tilde{I}_m$ comes from the other MRRHs on the plane except the region defined by $\mathfrak{B}(0,u)$; while $\tilde{I}_s$ produces from the other SRRHs except the region defined by $\mathfrak{B}(0,P_{sm}^{\frac{1}{\alpha_s}}u^{\alpha_{ms}})$. Then, following the same procedures adopted in *Lemma 3.3* and *Lemma 3.4*, the LT of the interference can be obtained.

Chapter 4

B1 Power Allocation for CoMP-assisted HD C-NOMA

In this section, we provide the preliminary steps to derive a computationally efficient power allocation solution in a CoMP-assisted HD C-NOMA system. First, it is worth mentioning that the DT phase and the CT phase occur in two consecutive time-slots. In other words, the DT phase occurs in the first time-slot and the CT phase is executed in the second time-slot. Accordingly, the received `SINR` at the CCU associated with BS_1, i.e., CCU_m^1, to decode the message of the CEU_f and the received `SINR` at the CCU_m^1 to decode its own signal are given, respectively, in the DT phase as follows.

$$\delta_{1,m\longrightarrow f}^{\mathrm{HD}} = \frac{\alpha_{1,m,f}^{\mathrm{HD}}\gamma_{1,m}}{(1-\alpha_{1,m,f}^{\mathrm{HD}})\gamma_{1,m}+1}, \tag{B.1}$$

$$\delta_{1,m\longrightarrow m}^{\mathrm{HD}} = (1-\alpha_{1,m,f}^{\mathrm{HD}})\gamma_{1,m}, \tag{B.2}$$

where $\gamma_{1,m} = \frac{P_1|h_{1,m}|^2}{\mathcal{T}}$ and $\mathcal{T} \triangleq P_2|h_{2,m}|^2 + \sigma^2$. Thus, the achievable rate at CCU_m^1 to decode the message of CEU_f and to decode its own message can be written, respectively, as

$$\mathcal{R}_{1,m\longrightarrow f}^{\mathrm{HD}} = \frac{1}{2}\log\left(1+\delta_{1,m\longrightarrow f}^{\mathrm{HD}}\right), \tag{B.3}$$

$$\mathcal{R}_{1,m\longrightarrow m}^{\mathrm{HD}} = \frac{1}{2}\log\left(1+\delta_{1,m\longrightarrow m}^{\mathrm{HD}}\right). \tag{B.4}$$

Similarly, the achievable rate at CCU associated with the BS_2, i.e. CCU_n^2, to decode the message of CEU_f ($\mathcal{R}_{2,n\longrightarrow f}^{\mathrm{HD}}$) and to decode its own message ($\mathcal{R}_{2,n\longrightarrow n}^{\mathrm{HD}}$) can be obtained. Meanwhile, the received `SINR` at the CEU_f in the DT phase can be written as follows.

$$\delta_{1,f\longrightarrow f}^{\mathrm{HD}} = \frac{\alpha_{1,m,f}^{\mathrm{HD}}\gamma_{1,f} + \alpha_{2,n,f}^{\mathrm{HD}}\gamma_{2,f}}{(1-\alpha_{1,m,f}^{\mathrm{HD}})\gamma_{1,f} + (1-\alpha_{2,n,f}^{\mathrm{HD}})\gamma_{2,f}+1}, \tag{B.5}$$

where $\forall i \in \{1,2\}, \gamma_{i,f} = \frac{P_i|h_{i,f}|^2}{\sigma^2}$. On the other hand, the received `SINR` at the CEU_f in the CT phase can be written as follows.

$$\delta_{2,f\longrightarrow f}^{\mathrm{HD}} = \beta_{1,m,f}^{\mathrm{HD}}\gamma_{m,f} + \beta_{2,n,f}^{\mathrm{HD}}\gamma_{n,f}, \tag{B.6}$$

where $\gamma_{k,f} = \frac{P_{\max}|h_{k,f}|^2}{\sigma^2}$. Note that, in order to maximize the network sum-rate, we will allow the CCUs to transmit with the maximum power, i.e., $\beta_{1,m,f}^{\mathrm{HD}} = \beta_{2,n,f}^{\mathrm{HD}} = 1$. Thus, the

received SINR in the CT phase at the CEU$_f$ can be rewritten as follows.

$$\delta^{\mathrm{HD}}_{2,f\longrightarrow f} = \gamma_{m,f} + \gamma_{n,f}. \tag{B.7}$$

At the end of the CT phase, CEU$_f$ employs the MRC mechanism to combine these receptions in both the DT phase and the CT phase to decode its own signal, which can be as follows.

$$\begin{aligned}\delta_{\mathrm{MRC}} &= \delta^{\mathrm{HD}}_{1,f\longrightarrow f} + \delta^{\mathrm{HD}}_{2,f\longrightarrow f} \\ &= \frac{\alpha^{\mathrm{HD}}_{1,m,f}\gamma_{1,f} + \alpha^{\mathrm{HD}}_{2,n,f}\gamma_{2,f}}{(1-\alpha^{\mathrm{HD}}_{1,m,f})\gamma_{1,f} + (1-\alpha^{\mathrm{HD}}_{2,n,f})\gamma_{2,f} + 1} + \gamma_{m,f} + \gamma_{n,f},\end{aligned} \tag{B.8}$$

Since the two CCUs, CCU^1_m and CCU^2_n, operate in decode and forward relaying mode, the achievable rate of CEU$_f$ is limited by the weakest link and it is expressed as

$$\mathcal{R}^{\mathrm{HD}}_f = \min\{\mathcal{R}^{\mathrm{HD}}_{1,m\to f}, \mathcal{R}_{f,\mathrm{MRC}}, \mathcal{R}^{\mathrm{HD}}_{2,n\to f}\}, \tag{B.9}$$

where $\mathcal{R}_{f,\mathrm{MRC}} = \frac{1}{2}\log(1+\delta_{\mathrm{MRC}})$. After obtaining the achievable data rate expression for each user, we will follow the same steps given in Section 4.5.3 for deriving the feasibility conditions as well as a computationally efficient power allocation solution.

B2 Proof of the SINR Expression at CEU

Before we proceed, let us denote $h_{1,f} \triangleq h_0, h_{2,f} \triangleq h_1, h_{m,f} \triangleq h_2$, and $h_{n,f} \triangleq h_3$. We generally set the indices of the transmission nodes BS$_1$, BS$_2$, $CU1_m$, and $CU2_n$ as $0, 1, 2, 3$, respectively. In addition, let $\bar{P}_0 = \alpha_{1,m,f}P_1, \bar{P}_1 = \alpha_{2,n,f}P_2, \bar{P}_2 = \beta_{1,m,f}P_{\max}$, and $\bar{P}_3 = \beta_{2,n,f}P_{\max}$. The NCJT exploits the concept of cyclic prefix (CP), which ensures that the different timely-dispersed multi-path versions of the information-bearing signal do not cause inter-symbol interference. Based on that and according to our proposed system model that allows four nodes to jointly serve an CEU$_f$, the received OFDM signal, which is a discrete time-domain signal can be represented as

$$y_f[b] = \sum_{t=0}^{3}\sqrt{\bar{P}_t}(h_t \circledast z_f)[b] + i[b] + w_f[b], \qquad b = 1,\ldots,B, \tag{B.10}$$

where $z_f[b]$ is the message of EU_f, $i[b] = \sqrt{(1-\alpha_{1,m,f})P_1}h_{1,f}z_{1,m} + \sqrt{(1-\alpha_{2,n,f})P_2}h_{2,f}z_{2,n}$ is the INUI signal, $w_f[b]$ is an AWGN at EU_f, which is $\sim \mathcal{CN}(0,\sigma^2)$ distributed, B is the OFDM symbol length, and $\circledast$ denotes the circular convolution. The effective channel coefficient between the transmit node t and EU_f, which includes the distance-based path-loss $d_{tf}^{-\eta}$, the small scale fading g_{tf}, and the time offset due to the lack of synchronization between the four transmit nodes denoted as τ_t, can be represented, in time-domain, as, $h_t[b] = g_{tf}[b-\tau_t]d_{tf}^{-\frac{\eta}{2}}$. Note that, due to there is no prior phase alignment and a lack

of tight coordination, both the small-scale fading and the time offset are independent across the different transmission nodes. Based on that, the received *useful signal* in the frequency domain after applying the discrete Fourier transform (DFT) on the subcarrier c can be written as follows.

$$\begin{aligned} Y_f^{\text{use}}[c] &= \text{DFT}\left\{ \sum_{t=0}^{3} \sqrt{\bar{P}_t}(h_t \circledast z_f)[b] \right\}, \\ &= Z_f[c] \sum_{t=0}^{3} \sqrt{\bar{P}_t} |g_{tf}[c]| d_{tf}^{-\frac{\eta}{2}} e^{j\phi_{tf}[c] + j2\pi c \frac{\tau_t}{B}}, \end{aligned} \tag{B.11}$$

where ϕ_{tf} is the phase rotation of the complex-valued small-scale fading. Thus, and with CSI at the receiver, the useful received signal power at CEU$_f$ can be written as [41]

$$\begin{aligned} |Y_f^{\text{use}}[c]|^2 &= |Z_f[c]|^2 \sum_{t=0}^{3} \bar{P}_t |g_{tf}[c]|^2 d_{tf}^{-\eta} + \\ &|Z_f[c]|^2 \sum_{\substack{t,j=0 \\ t \neq j}}^{3} \frac{\sqrt{\bar{P}_t \bar{P}_j} g_{tf}[c] g_{jf}[c]}{d_{tf}^{-\eta/2} d_{jf}^{-\eta/2}} \mathcal{R}\left[e^{j\tilde{\phi}_{tj}[c]} e^{j2\pi c \frac{\tilde{\tau}_{tj}}{B}} \right], \end{aligned} \tag{B.12}$$

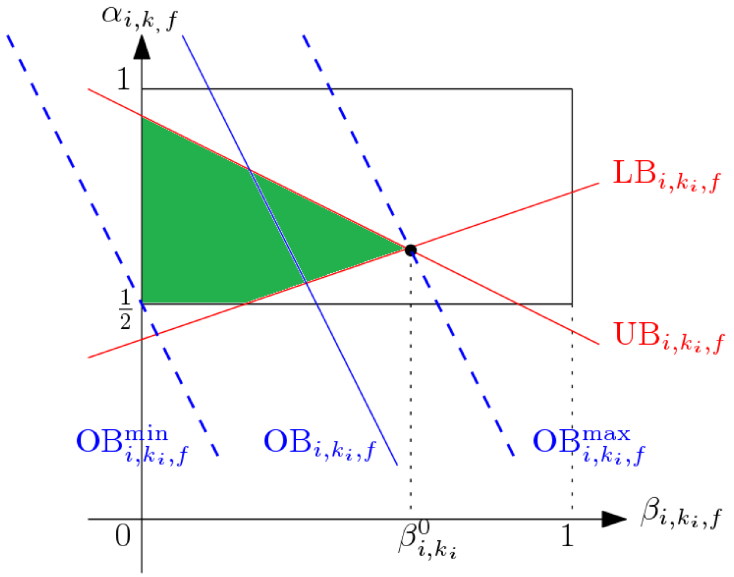
where $\tilde{\phi}_{tj}[c] \triangleq \phi_{tf}[c] - \phi_{jf}[c], \tilde{\tau}_{tj} \triangleq \tau_t - \tau_j$, and $\mathcal{R}(.)$ refers to the real part. Note that, when both the coherence bandwidth and $\tilde{\tau}_{tj}$ are relatively large, the modulation term (the two complex exponentials in (B.12)) causes the received signal power to change noticeably over a large number of subcarriers that have the same $g_{tf}[c] = g_{tf}$ and $\phi_{tf}[c] = \phi_{tf}$ [41]. In order to take the overall effect of the timing offset, the average of the power variations during the coherent bandwidth, i.e. spanning N_s subcarriers around subcarrier c is considered as follows

$$\begin{aligned} \overline{|Y_f^{\text{use}}[c]|^2} &= \sum_{t=0}^{3} \bar{P}_t |g_{tf}|^2 d_{tf}^{-\eta} \frac{1}{N_c} \sum_{a=0}^{N_c-1} |Z_f[a]|^2 + \\ &\sum_{\substack{t,j=0 \\ t \neq j}}^{3} \frac{\sqrt{\bar{P}_t \bar{P}_j} g_{tf}[c] g_{jf}[c]}{d_{tf}^{-\eta/2} d_{jf}^{-\eta/2}} \frac{1}{N_c} \sum_{a=0}^{N_c-1} |Z_f[a]|^2 \mathcal{R}\left[e^{j\tilde{\phi}_{tj}} e^{j2\pi a \frac{\tilde{\tau}_{tj}}{B}} \right], \\ &\approx \frac{1}{N} \sum_{t=0}^{3} \bar{P}_t |g_{tf}|^2 d_{tf}^{-\eta}, \end{aligned} \tag{B.13}$$

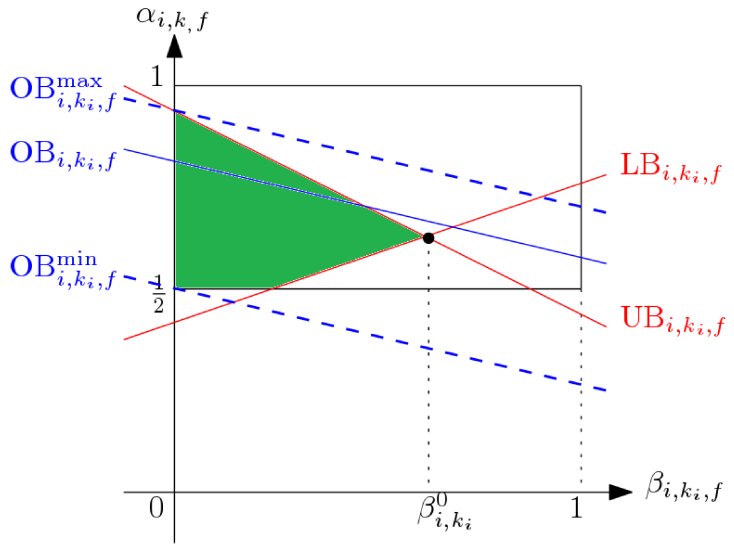
where the approximation in (B.13) is based on the fact that $1/N_c \sum_{a=0}^{N_c-1} |Z_f[a]|^2 \approx \mathbb{E}\left[z[b]\right]/N = 1/N$ and that $1/N_c \sum_{a=0}^{N_c-1} |Z_f[a]|^2 \mathcal{R}\left[e^{j\tilde{\phi}_{tj}} e^{j2\pi a \frac{\tilde{\tau}_{tj}}{B}} \right] \approx 0$ for large N_c [41]. Similarly and with the assumption that the interference signals are non-coherently superimposed at the CEU$_f$, the interference-plus-noise power can be readily obtained as

$$\overline{|I+W|^2} = \frac{1}{N} \sum_{t=0}^{1} (P_t - \bar{P}_t) |g_{tf}|^2 d_{tf}^{-\eta} + \frac{\sigma^2}{N}, \tag{B.14}$$

where I and W are the DFT for the $i[b]$ and $w_f[b]$, respectively. By combining (B.13) and (B.14), the `SINR` in (4.31) can be obtained.



(a) 1st case: $\frac{\gamma_{k_i,\mathrm{SI}}t}{\gamma_{i,k_i}} \leq \frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}$



(b) 2nd case: $\frac{\gamma_{k_i,\mathrm{SI}}t}{\gamma_{i,k_i}} \geq \frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}$

Figure B.1: Feasibility region.

B3 Proof of Theorem 4.1

The feasibility conditions of the power control problem $\mathcal{P}_{m,f,n}$ are the conditions that ensure the satisfaction of its constraints, which are constraints (4.40b)−(4.40g). For $i \in \mathcal{I}$, let us consider first the triple $(\alpha_{i,k_i,f}, \beta_{i,k_i,f}, Y_{i,k_i,f})$. Since $\mathrm{LB}_{i,k_i,f} \leq \alpha_{i,k_i,f} \leq \mathrm{UB}_{i,k_i,f}$, then $\beta_{i,k_i,f}$ should satisfy

$$0 \leq \beta_{i,k_i,f} \leq \beta^0_{i,k_i}, \tag{B.15}$$

where

$$\beta^0_{i,k_i} = \frac{\gamma_{i,k_i} - 2t - t^2}{\gamma_{k_i,\mathrm{SI}} t(2+t)}, \tag{B.16}$$

which corresponds to the abscissa of the intersection point between the bounds $\mathrm{UB}_{i,k_i,f}$ and $\mathrm{LB}_{i,k_i,f}$. Therefore, in order to have a feasible solution for $\beta_{i,k_i,f}$, the parameter β^0_{i,k_i} should satisfy $0 \leq \beta^0_{i,k_i}$. This can be achieved by ensuring that $0 \leq \gamma_{i,k_i} - 2t - t^2$, which is in turn satisfied iff

$$0 \leq t \leq -1 + \sqrt{1 + \gamma_{i,k_i}}. \tag{B.17}$$

Moreover, assuming that the condition in (B.17) holds, we have

$$\mathrm{UB}_{i,k_i,f}\left(\beta^0_{i,k_i}\right) = \mathrm{LB}_{i,k_i,f}\left(\beta^0_{i,k_i}\right) = \frac{1+t}{2+t} \in \left[\frac{1}{2}, 1\right]. \tag{B.18}$$

Therefore, the corresponding region of $\alpha_{i,k_i,f}$ is within the interval $\left[\frac{1}{2}, 1\right]$. Hence, the feasibility region of $(\alpha_{i,k_i,f}, \beta_{i,k_i,f})$ is non-empty if and only if the condition in (B.17) holds. Consequently, the feasibility region of the decision parameters $\{(\alpha_{i,k_i,f}, \beta_{i,k_i,f} | i \in \mathcal{I})\}$ is non empty if and only if

$$0 \leq t \leq -1 + \sqrt{1 + \min_{i \in \mathcal{I}} \gamma_{i,k_i}}, \tag{B.19}$$

which represents the first feasibility condition.

Now, let us assume that the condition in (B.17) holds and let us define the quantity $Y^{\max}_{i,k_i,f}$ for all $i \in \mathcal{I}$ as

$$Y^{\max}_{i,k_i,f} \triangleq \max_{\alpha_{i,k_i,f}, \beta_{i,k_i,f}} Y_{i,k_i,f}. \tag{B.20}$$

Hence, the second feasibility condition should satisfy

$$Y_f \leq \sum_{i=1}^{|\mathcal{I}|} Y^{\max}_{i,k_i,f}. \tag{B.21}$$

Based on Fig. B.1, $Y^{\max}_{i,k_i,f}$ is expressed as follows. For the case when $\frac{\gamma_{k_i,\mathrm{SI}} t}{\gamma_{i,k_i}} \leq \frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}$, and as shown in Fig. B.1(a), the maximum value of $Y_{i,k_i,f}$ is reached when the intersection between the bounds $\mathrm{UB}_{i,k_i,f}$ and $\mathrm{OB}_{i,k_i,f}$ occurs at the point with abscissa $\beta_{i,k_i,f} = \beta^{\max}_{i,k_i,f} = \min\left(1, \beta^0_{i,k_i}\right)$, i.e., when

$$\mathrm{UB}_{i,k_i,f}\left(\beta^{\max}_{i,k_i,f}\right) = \mathrm{OB}_{i,k_i,f}\left(\beta^{\max}_{i,k_i,f}, Y^{\max}_{i,k_i,f}\right), \tag{B.22}$$

which implies

$$Y^{\max}_{i,k_i,f} = \gamma_{i,f}(1+t)\left(\left(\frac{\gamma_{i,f}}{\gamma_{i,f}(1+t)} - \frac{\gamma_{k_i,\mathrm{SI}}}{\gamma_{i,k_i} t}\right)\beta^{\max}_{i,k_i,f} + \frac{\gamma_{i,k_i} - t}{\gamma_{i,k_i}}\right). \tag{B.23}$$

On the other hand, for the case when $\frac{\gamma_{k_i,\mathrm{SI}} t}{\gamma_{i,k_i}} \geq \frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}$, and as shown in Fig. B.1(b), the maximum value of $Y_{i,k_i,f}$ is reached when the intersection between the bounds $\mathrm{UB}_{i,k_i,f}$ and $\mathrm{OB}_{i,k_i,f}$ occurs at the point with abscissa $\beta_{i,k_i,f} = 0$, i.e., when

$$\mathrm{UB}_{i,k_i,f}(0) = \mathrm{OB}_{i,k_i,f}\left(0, Y^{\max}_{i,k_i,f}\right), \tag{B.24}$$

which implies

$$Y^{\max}_{i,k_i,f} = \gamma_{i,f}(1+t) \times \frac{\gamma_{i,k_i} - t}{\gamma_{i,k_i}}. \tag{B.25}$$

Therefore, by combining the expressions in (B.23) and (B.25), we obtain the expression of $Y^{\max}_{i,k_i,f}$ shown in equation (4.43), which completes the proof.

B4 Proof of Corollary 1

The value of $Y^{\max}_{i,k_i,f}$ was derived in Appendix B1. In this section, we derive the expression of $Y^{\min}_{i,k_i,f}$. Based on Fig. B.1, the lowest feasible value of $\alpha_{i,k_i,f}$ when $\beta_{i,k_i,f} = 0$, which we denote by $\alpha^0_{i,k_i,f}$, is given by

$$\alpha^0_{i,k_i,f} = \max\left(\frac{1}{2}, \frac{t(1+\gamma_{i,k_i})}{\gamma_{i,k_i}(1+t)}\right). \tag{B.26}$$

Consequently, $Y^{\min}_{i,k_i,f}$ is reached when the point $(\alpha_{i,k_i,f}, \beta_{i,k_i,f}) = \left(\alpha^0_{i,k_i,f}, 0\right)$ is located in the bound $\mathrm{OB}_{i,k_i,f}$, i.e., when

$$\alpha^0_{i,k_i,f} = \mathrm{OB}_{i,k_i,f}\left(0, Y^{\min}_{i,k_i,f}\right). \tag{B.27}$$

Therefore, from equation (B.27), we obtain

$$Y^{\min}_{i,k_i,f} = \max\left(\frac{\gamma_{i,f}\,(1+t)}{2}, \frac{t\gamma_{i,f}(1+\gamma_{i,k_i})}{\gamma_{i,k_i}}\right), \tag{B.28}$$

which completes the proof.

B5 Proof of Theorem 4.2

In this section, we derive the intersection points $(\alpha_{i,k_i,f}, \beta_{i,k_i,f})$ between the bound $\mathrm{OB}_{i,k_i,f}$ and the boundaries of the feasibility region except the upper bound $\mathrm{UB}_{i,k_i,f}$. Let $\beta^{\mathrm{int}}_{i,k_i,f}$ denotes the abscissa of the intersection point between the bounds $\mathrm{OB}_{i,k_i,f}$ and $\mathrm{LB}_{i,k_i,f}$. By solving the equality between the expressions of these bounds in (4.41) with respect to $\beta_{i,k_i,f}$, the expression of $\beta^{\mathrm{int}}_{i,k_i,f}$ can be obtained as

$$\beta^{\mathrm{int}}_{i,k_i,f} = \frac{\gamma_{i,k_i} Y_{i,k_i,f} - t\gamma_{i,f}(\gamma_{i,k_i}+1)}{\gamma_{k_i,\mathrm{SI}}\gamma_{i,f} + \gamma_{i,k_i}\gamma_{i,k_i,f}}. \tag{B.29}$$

Now, let $\beta^{c}_{i,k_i,f} = \min\left(\beta^{\max}_{i,k_i,f}, \beta^{\mathrm{int}}_{i,k_i,f}\right)$. We distinguish between the two cases $\frac{\gamma_{i,k_i}-t}{\gamma_{i,k_i}} \leq \frac{Y_{i,k_i,f}}{\gamma_{i,f}(1+t)}$ and $\frac{\gamma_{i,k_i}-t}{\gamma_{i,k_i}} \geq \frac{Y_{i,k_i,f}}{\gamma_{i,f}(1+t)}$. The intersection points within each case are given in the following subsections.

B5.1 1st case: $\frac{\gamma_{i,k_i}-t}{\gamma_{i,k_i}} \leq \frac{Y_{i,k_i,f}}{\gamma_{i,f}(1+t)}$

This case corresponds also to the case when $\frac{\gamma_{k_i,\mathrm{SI}}t}{\gamma_{i,k_i}} \leq \frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}$, which admits the scenario when the ordinate at the origin of the bound $\mathrm{OB}_{i,k_i,f}$ can be higher than the one of the upper bound $\mathrm{LB}_{i,k_i,f}$. In this case, and as it can be seen in Fig. B.1(a), there exists only one critical point that is expressed as

$$(\alpha_{i,k_i,f}, \beta_{i,k_i,f})_1 = \left(f_{i,k_i,f}\left(\beta^c_{i,k_i,f}, Y_{i,k_i,f}\right), \beta^c_{i,k_i,f}\right), \tag{B.30}$$

where the function $f_{i,k_i,f}(\cdot,\cdot)$ is expressed as

$$f_{i,k_i,f}(\beta, x) = \max\left(\frac{1}{2}, -\frac{\gamma_{i,k_i,f}}{\gamma_{i,f}(1+t)}\beta + \frac{x}{\gamma_{i,f}(1+t)}\right). \tag{B.31}$$

B5.2 2nd case: $\frac{\gamma_{i,k_i}-t}{\gamma_{i,k_i}} \geq \frac{Y_{i,k_i,f}}{\gamma_{i,f}(1+t)}$

In this case, and as it can be seen in Fig. B.1(b), the point $(\alpha_{i,k_i,f}, \beta_{i,k_i,f})_1$ in (B.30) is also a critical point. Moreover, an additional critical point will be added, which corresponds to the intersection between the bound $\mathrm{OB}_{i,k_i,f}$ and the line $\beta_{i,k_i,f} = 0$. This critical point is expressed as

$$(\alpha_{i,k_i,f}, \beta_{i,k_i,f})_2 = \left(f_{i,k_i,f}\left(0, Y_{i,k_i,f}\right), 0\right), \tag{B.32}$$

which completes the proof.

B6 Proof of Lemma 4.1

We assume that Φ^* is the output of the proposed matching algorithm. We prove that the output matching Φ^* is not blocked by any triple that does not exist in Φ^*. It can be shown from step 6 in **Algorithm 3** that every unmatched CEU_f is willing to propose itself to its most favorable $(\mathrm{CCU}^1_{m^*}, \mathrm{CCU}^2_{n^*})$ pair that have not rejected it in previous rounds. Furthermore, every unmatched CEU_f will continue to propose itself until it is accepted by an $(\mathrm{CCU}^1, \mathrm{CCU}^2)$ pair or traverses all the $(\mathrm{CCU}^1_m, \mathrm{CCU}^2_n)$ pair in its preference profile. Suppose that, in a certain round r, there is a triple $\Delta_r = (\mathrm{CCU}^1_m, \mathrm{CEU}_f, \mathrm{CCU}^2_n)$ that does not belong to Φ^*. In this case, there are two possibilities which can be described as follows: 1) CEU_f has never proposed itself to $(\mathrm{CCU}^1_m, \mathrm{CCU}^2_n)$ and 2) CEU_f proposes itself in an arbitrary round r to $(\mathrm{CCU}^1_m, \mathrm{CCU}^2_n)$ and is subsequently kicked out. The first possibility will prove this case by contradiction. According to the above description, CEU_f must prefer $\Phi^*(\mathrm{CEU}_f)$, which is its currently matching pair, to the $(\mathrm{CCU}^1_m, \mathrm{CCU}^2_n)$. This is because $\Phi^*(\mathrm{CEU}_f) \succ_{\mathrm{CEU}_f} (\mathrm{CCU}^1_m, \mathrm{CCU}^2_n)$ and $U_{\mathrm{CEU}_f,\Phi^*(\mathrm{CEU}_f)} > U_{\Delta_r}$, which contradicts

the blocking triple definition. On the other hand, the second possibility means that Δ_r is not a blocking triple. This is because if it is the case, then $(\text{CCU}_m^1, \text{CCU}_n^2)$ would have kept CEU_f instead of its former partners at round r to avoid the blocking triple. In other words, if $(\text{CCU}_m^1, \text{CCU}_n^2)$ rejects CEU_f, this indicates that one player, either CCU_m^1 or CCU_n^2, or both of them can construct another triple Δ_t with other CEU instead of CEU_f, which contradicts the definition of the blocking triple. As a result, the proposed UEs clustering algorithm converges to a final matching Φ^* and based on *Lemma 4.1*, if CECM algorithm converges to a matching Φ^*, then Φ^* is a stable matching.

B7 Proof of Theorem 4.3

With a detailed explanation of the CECM algorithm, we prove in this appendix that this matching algorithm will terminate after a limited number of rounds. In each round, each unmatched $EU_f, \forall f \in \mathcal{V}$ proposes itself to its most favorable $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ pair that is not rejected it in the previous rounds (step 6 in **Algorithm 3**). Note that, once the CEU_f sends a request to its most preferred pair, it removes this $(\text{CCU}_{m^*}^1, \text{CCU}_{n^*}^2)$ pair from its preference profile (step 17 in **Algorithm 3**). As a result, as the number of rounds increases, the number of pairs in the preference profile for the CEU_f decreases. It has been shown that the proposed algorithm ends when each CEU_f is matched with $(\text{CCU}_m^1, \text{CCU}_n^2)$ pair or is kicked out by all the pairs and so it has an empty preference profile. Since the preference profile for each CEU_f contains $M \times N$ pairs, the number of requests that each CEU_f makes is no larger than $M \times N$ and hence the total number of rounds is no more than $M \times N$. In the worst case, after each CEU_f passes over its preference profile, it has no request to make. In this situation, the proposed algorithm terminates with $M \times N$ rounds. Thus, the CECM algorithm can converge to a final matching Φ^* in a limited number of rounds and the final matching Φ^* is stable in accordance with *Lemma 4.1*.

Chapter 5

C1 Proof of Theorem 5.1

In this section, we present the proof of ***Theorem 5.1***. At the beginning, constraints (5.19b)-(5.19f) can be expressed as

$$0 \leq \alpha_{\mathrm{n}}^{\mathrm{HD}} \leq \alpha_{\mathrm{f}}^{\mathrm{HD}} \leq 1, \tag{C.1a}$$

$$\alpha_{\mathrm{f}}^{\mathrm{HD}} \leq 1 - \alpha_{\mathrm{n}}^{\mathrm{HD}}, \tag{C.1b}$$

$$0 \leq \beta^{\mathrm{HD}} \leq 1, \tag{C.1c}$$

$$\alpha_{\mathrm{n}}^{\mathrm{HD}} \geq \frac{t_{\mathrm{n}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}}, \tag{C.1d}$$

$$\alpha_{\mathrm{f}}^{\mathrm{HD}} \geq \alpha_{\mathrm{n}}^{\mathrm{HD}} t_{\mathrm{f}}^{\mathrm{HD}} + \frac{t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}}, \tag{C.1e}$$

$$\beta^{\mathrm{HD}} \geq \frac{t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{d}}} - \frac{\alpha_{\mathrm{f}}^{\mathrm{HD}} \gamma_{\mathrm{bf}}}{\gamma_{\mathrm{d}} \left(\alpha_{\mathrm{n}}^{\mathrm{HD}} \gamma_{\mathrm{bf}} + 1\right)}. \tag{C.1f}$$

Constraints (C.1a) and (C.1b) imply $0 \leq \alpha_{\mathrm{n}}^{\mathrm{HD}} \leq \frac{1}{2}$ In addition, constraints (C.1b) and (C.1e) imply $\alpha_{\mathrm{n}}^{\mathrm{HD}} \leq \frac{\gamma_{\mathrm{bn}} - t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}\left(t_{\mathrm{f}}^{\mathrm{HD}}+1\right)}$. Therefore, it can be concluded that $\alpha_{\mathrm{n}}^{\mathrm{HD}}$ should satisfy $\alpha_{\mathrm{min}}^{\mathrm{HD}} \leq \alpha_{\mathrm{n}}^{\mathrm{HD}} \leq \alpha_{\mathrm{max}}^{\mathrm{HD}}$, where $\alpha_{\mathrm{min}}^{\mathrm{HD}}$ and $\alpha_{\mathrm{max}}^{\mathrm{HD}}$ are expressed as shown in (5.24) in ***Theorem 5.1***. Hence, for problem $\mathrm{PC-HD}$ to be feasible, the condition $\alpha_{\mathrm{min}}^{\mathrm{HD}} \leq \alpha_{\mathrm{max}}^{\mathrm{HD}}$ should be satisfied. On the other hand, one can remark that the lowest value of β^{HD} that satisfies constraint (C.1f) is reached at the lowest feasible value of $\alpha_{\mathrm{n}}^{\mathrm{HD}}$ and the highest feasible value of $\alpha_{\mathrm{f}}^{\mathrm{HD}}$. In this context, based on constraint (C.1d), the lowest feasible value of $\alpha_{\mathrm{n}}^{\mathrm{HD}}$ is $\alpha_{\mathrm{n}}^{\mathrm{HD}} = \alpha_{\mathrm{min}}^{\mathrm{HD}} = \frac{t^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}}$ and, based on constraint (C.1b), the highest feasible value of $\alpha_{\mathrm{f}}^{\mathrm{HD}}$ is $\alpha_{\mathrm{n}}^{\mathrm{HD}} = 1 - \alpha_{\mathrm{min}}^{\mathrm{HD}} = 1 - \frac{t^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}}$. Therefore, the lowest feasible value of β^{HD} is given by $\beta_{\mathrm{min}}^{\mathrm{HD}}$ in (5.24). Consequently, for problem $\mathrm{PC-HD}$ to be feasible, the condition $\beta_{\mathrm{min}}^{\mathrm{HD}} \leq \beta_{\mathrm{max}}^{\mathrm{HD}}$ should be also satisfied, where $\beta_{\mathrm{max}}^{\mathrm{HD}} = 1$, which completes the proof.

C2 Proof of Theorem 5.2

In this section, we present the proof of ***Theorem 5.2***. With the goal of minimizing the total transmit power of the C-NOMA cellular system, two directions can be considered.

The first direction consists of minimizing the BS transmit power first and then minimizing the one of the near UE, while the second direction consists of minimizing the transmit power of the near UE first and then minimizing the one of the BS. Let us start with the first direction. As discussed in the previous appendix, the lowest feasible value of $\alpha_{\mathrm{n}}^{\mathrm{HD}}$ is $\alpha_{\mathrm{n},1}^{\mathrm{HD}} = \alpha_{\min}^{\mathrm{HD}}$. Based on this, we can see from constraints (5.19b) and (C.1e) that the lowest feasible value of $\alpha_{\mathrm{f}}^{\mathrm{HD}}$ is

$$\alpha_{\mathrm{f},1}^{\mathrm{HD}} = \max\left(\alpha_{\min}^{\mathrm{HD}}, \alpha_{\min}^{\mathrm{HD}} t_{\mathrm{f}}^{\mathrm{HD}} + \frac{t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{n}}}\right). \tag{C.2}$$

Therefore, by injecting these two values in constraint (C.1f), we conclude that the lowest value of β^{HD} is

$$\beta_1^{\mathrm{HD}} = \frac{1}{\gamma_{\mathrm{d}}}\left(t_{\mathrm{f}}^{\mathrm{HD}} - \frac{\alpha_{\mathrm{f},1}^{\mathrm{HD}}\gamma_{\mathrm{f}}}{\alpha_{\mathrm{n},1}^{\mathrm{HD}}\gamma_{\mathrm{f}} + 1}\right). \tag{C.3}$$

Now, considering the second direction, the lowest feasible value of β^{HD} is $\beta_1^{\mathrm{HD}} = 0$. on the other hand, recall that the lowest feasible value of $\alpha_{\mathrm{n}}^{\mathrm{HD}}$ is $\alpha_{\mathrm{n},2}^{\mathrm{HD}} = \alpha_{\min}^{\mathrm{HD}}$. Therefore, by injecting these two values in constraint (C.1f), we find that the lowest feasible value of $\alpha_{\mathrm{f}}^{\mathrm{HD}}$ is

$$\alpha_{\mathrm{f},2}^{\mathrm{HD}} = \frac{\left(\alpha_{\min}^{\mathrm{HD}}\gamma_{\mathrm{bn}} + 1\right) t_{\mathrm{f}}^{\mathrm{HD}}}{\gamma_{\mathrm{bn}}}. \tag{C.4}$$

Finally, the optimal power control scheme is the one among the above two schemes that minimize the total transmit power of the HD C-NOMA system, which completes the proof.

C3 Proof of Proposition 3

First, we denote $\mathcal{Q}(\mathbf{p}^{\mathrm{HD}}, \boldsymbol{\theta}^{\mathrm{HD}})$ as the objective value of $\mathrm{OPT}-\mathrm{HD}$ for a feasible solution $(\mathbf{p}^{\mathrm{HD}}, \boldsymbol{\theta}^{\mathrm{HD}})$. We also denote $(\mathbf{p}^{\mathrm{HD},[t]}, \boldsymbol{\theta}^{\mathrm{HD},[t]})$ as a feasible solution of $\mathrm{OPT}-\mathrm{HD}$ at the tth iteration. For a given $\boldsymbol{\theta}^{\mathrm{HD},[t]}$, in the $t+1$ iteration, we can obtain the optimal power allocation based on ***Theorem 5.2***, $\mathbf{p}^{\mathrm{HD},[t+1]}$, and we have

$$\mathcal{Q}(\mathbf{p}^{\mathrm{HD},[t+1]}, \boldsymbol{\theta}^{\mathrm{HD},[t]}) \overset{(a)}{<} \mathcal{Q}(\mathbf{p}^{\mathrm{HD},[t]}, \boldsymbol{\theta}^{\mathrm{HD},[t]}) \tag{C.5}$$

where (a) holds since for given $\boldsymbol{\theta}_{(1)}^{[t]}$ and $\boldsymbol{\theta}_{(2)}^{[t]}$, $\alpha_{\mathrm{n}}^{\mathrm{HD},[t+1]}, \alpha_{\mathrm{f}}^{\mathrm{HD},[t+1]}$, and $\beta^{\mathrm{HD},[t+1]}$ are the optimal solutions for the power allocation. For a given $\mathbf{p}^{\mathrm{HD},[t+1]}$, based on the proposed AO-based SCA algorithm if there exists a feasible solution $\boldsymbol{V}^{[t+1]}$ to problem $\mathcal{P}1$, it is also feasible $\mathrm{OPT}-\mathrm{HD}$ and we have $\mathcal{Q}(\mathbf{p}^{\mathrm{HD},[t+1]}, \boldsymbol{\theta}^{\mathrm{HD},[t]}) = \mathcal{Q}(\mathbf{p}^{\mathrm{HD},[t+1]}, \boldsymbol{\theta}^{\mathrm{HD},[t+1]})$ in which the equality holds due to the value of $\mathcal{Q}$ depends only on $\mathbf{p}^{\mathrm{HD}}$. According to the above discussion, we have $\mathcal{Q}(\mathbf{p}^{\mathrm{HD},[t+1]}, \boldsymbol{\theta}^{\mathrm{HD},[t+1]}) < \mathcal{Q}(\mathbf{p}^{\mathrm{HD},[t]}, \boldsymbol{\theta}^{\mathrm{HD},[t]})$. This demonstrates that $\mathcal{Q}$

is always decreasing over the iterations. In addition, since the total transmit power is lower bound because of the minimum required QoS for each user, the proposed algorithm is guaranteed to converge.

C4 Proof of Theorem 5.3

At the beginning, constraints (5.40b)-(5.40f) can be expressed as

$$\alpha_{\mathrm{n}}^{\mathrm{FD}} \geq \frac{\gamma_{\mathrm{SI}}}{\gamma_{\mathrm{bn}}}\beta^{\mathrm{FD}} + \frac{t_{\mathrm{n}}^{\mathrm{FD}}}{\gamma_{\mathrm{bn}}}, \tag{C.6a}$$

$$\alpha_{\mathrm{f}}^{\mathrm{FD}} \geq \alpha_{\mathrm{n}}^{\mathrm{FD}} t_{\mathrm{f}}^{\mathrm{FD}} + \frac{\gamma_{\mathrm{SI}} t_{\mathrm{f}}^{\mathrm{FD}}}{\gamma_{\mathrm{bn}}}\beta^{\mathrm{FD}} + \frac{t_{\mathrm{f}}^{\mathrm{FD}}}{\gamma_{\mathrm{bn}}}, \tag{C.6b}$$

$$\alpha_{\mathrm{f}}^{\mathrm{FD}} \geq \alpha_{\mathrm{n}}^{\mathrm{FD}} t_{\mathrm{f}}^{\mathrm{FD}} - \frac{\gamma_{\mathrm{d}} t_{\mathrm{f}}^{\mathrm{FD}}}{\gamma_{\mathrm{bf}}}\beta^{\mathrm{FD}} + \frac{t_{\mathrm{f}}^{\mathrm{FD}}}{\gamma_{\mathrm{bf}}}. \tag{C.6c}$$

Constraints (5.40b) and (5.40c) imply $0 \leq \alpha_{\mathrm{n}}^{\mathrm{FD}} \leq \frac{1}{2}$. Let B_1 be the bound whose expression is given in constraint (C.6a). In addition, constraints (5.40c) and (C.6b) imply

$$B_2: \ \alpha_{\mathrm{n}}^{\mathrm{HD}} \leq \frac{\gamma_{\mathrm{bn}} - t_{\mathrm{f}}^{\mathrm{HD}} - P_{\mathrm{n}}\gamma_{\mathrm{SI}} t_{\mathrm{f}}^{\mathrm{HD}}\beta^{\mathrm{FD}}}{\gamma_{\mathrm{bn}}\left(t_{\mathrm{f}}^{\mathrm{HD}} + 1\right)}. \tag{C.7}$$

Moreover, based on constraints (5.40c) and (C.6c), we obtain

$$B_3: \ \alpha_{\mathrm{n}}^{\mathrm{HD}} \leq \frac{\gamma_{\mathrm{bn}} - t_{\mathrm{f}}^{\mathrm{HD}} + P_{\mathrm{n}}\gamma_{\mathrm{d}}\beta^{\mathrm{FD}}}{\gamma_{\mathrm{bf}}\left(t_{\mathrm{f}}^{\mathrm{HD}} + 1\right)}. \tag{C.8}$$

The bounds B_1, B_2 and B_3 are, each, a function of the variables $\left(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}\right)$. Therefore, problem $\mathrm{PC}-\mathrm{FD}$ is feasible if and only if the bound B_1 can be lower than the bounds B_2 and B_3 simultaneously within the region $\left(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}\right) \in \left[0, \frac{1}{2}\right] \times [0, 1]$. The condition $B_1 \leq B_2$ implies that β^{FD} should satisfy

$$\beta^{\mathrm{FD}} \leq \frac{\gamma_{\mathrm{n}} - t_{\mathrm{f}}^{\mathrm{FD}} - t_{\mathrm{n}}^{\mathrm{FD}}\left(1 + t_{\mathrm{f}}^{\mathrm{FD}}\right)}{\gamma_{\mathrm{SI}} t_{\mathrm{n}}^{\mathrm{FD}}\left(1 + t_{\mathrm{f}}^{\mathrm{FD}}\right) + t_{\mathrm{f}}^{\mathrm{FD}}}. \tag{C.9}$$

On the other hand, the condition $B_1 \leq B_3$ implies that β^{FD} should satisfy

$$\beta^{\mathrm{FD}} \geq \frac{\gamma_{\mathrm{bf}} - t_{\mathrm{f}}^{\mathrm{FD}} - \frac{\gamma_{\mathrm{bf}}\left(1+t_{\mathrm{f}}^{\mathrm{FD}}\right)t_{\mathrm{n}}^{\mathrm{FD}}}{\gamma_{\mathrm{n}}}}{c_1 - c_2} \tag{C.10}$$

if $c_1 \leq c_2$ and

$$\beta^{\mathrm{FD}} \leq \frac{\gamma_{\mathrm{bf}} - t_{\mathrm{f}}^{\mathrm{FD}} - \frac{\gamma_{\mathrm{bf}}\left(1+t_{\mathrm{f}}^{\mathrm{FD}}\right)t_{\mathrm{n}}^{\mathrm{FD}}}{\gamma_{\mathrm{n}}}}{c_1 - c_2}, \tag{C.11}$$

if $c_1 \geq c_2$. Therefore, along with the fact that $\beta^{\mathrm{FD}} \in [0, 1]$, we obtain the first condition in ***Theorem 5.3***. Furthermore, since B_1 should be lower than the bounds B_2 and B_3 simultaneously within the region $\left(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}\right) \in \left[0, \frac{1}{2}\right] \times [0, 1]$, we conclude that $\frac{\gamma_{\mathrm{SI}} t_{\mathrm{n}}^{\mathrm{FD}}}{\gamma_{\mathrm{bn}}}\beta_{\min}^{\mathrm{FD}} + \frac{t_{\mathrm{n}}^{\mathrm{FD}}}{\gamma_{\mathrm{bn}}}$ should be lower than $\frac{1}{2}$, which is the second condition of ***Theorem 5.3*** and this completes the proof.

C5 Proof of Theorem 5.4

In this section, we present the proof of ***Theorem 5.4***. Since the objective function of problem PC − FD, the optimal solution is one of the intersection points of the boundaries of its feasibility region. The point with coordinates $\left(\alpha_{\mathrm{c}}^{\mathrm{FD}}, \beta_{\mathrm{c}}^{\mathrm{FD}}\right)$, defined in (5.47) represents the intersection point of the bounds B_2 and B_3. Therefore, depending on whether $\beta_{\mathrm{c}}^{\mathrm{FD}}$ is within the feasibility interval $\left[\beta_{\min}^{\mathrm{FD}}, \beta_{\max}^{\mathrm{FD}}\right]$, this intersection point is transformed into the point $\left(\alpha_0^{\mathrm{FD}}, \beta_0^{\mathrm{FD}}\right)$ in (5.48) in order to verify if this point should be considered as a candidate solutions or no. Based on this, the intersection points of the boundaries of the feasibility region of problem PC − FD are given in (5.49). Afterwards, for each intersection point $\left(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}\right)$, the corresponding $\alpha_{\mathrm{f}}^{\mathrm{FD}}$ can be obtained as shown in (5.50), which in turn constructs a candidate solution $\left(\alpha_{\mathrm{n}}^{\mathrm{FD}}, \alpha_{\mathrm{f}}^{\mathrm{FD}}, \beta^{\mathrm{FD}}\right)$. Consequently, the optimal solution of problem PC − FD is the candidate solution that produces the lowest objective function as presented in ***Theorem 5.4***, which completes the proof.

C6 Semi-Definite Relaxation

First, we apply change-of-variables with the SDR technique and Gaussian randomization method to find the phase-shift matrix. Define $\boldsymbol{v} = [v_1, \ldots, v_L]^H$, where $\forall l \in [\![1, L]\!]$, $v_l = e^{j\theta_l}$. Then, the constraints in (5.19g) are equivalent to the unit-modulus constraints, i.e., $|v_l|^2 = 1, \forall l \in [\![1, L]\!]$. By applying the change-of-variables $\boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{1,R} = \boldsymbol{v}^H \boldsymbol{\Phi}$, where $\boldsymbol{\Phi} = \mathrm{diag}(\boldsymbol{h}_{R,f}^H)\boldsymbol{h}_{1,R} \in \mathbb{C}^{L\times 1}$ and $\boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{2,R} = \boldsymbol{v}^H \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = \mathrm{diag}(\boldsymbol{h}_{R,f}^H)\boldsymbol{h}_{2,R} \in \mathbb{C}^{L\times 1}$, we have

$$\begin{aligned} |h_{1,f} + \boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{1,R}|^2 &= |h_{1,f} + \boldsymbol{v}^H \boldsymbol{\Phi}|^2, \\ |h_{2,f} + \boldsymbol{h}_{R,f}^H \boldsymbol{\Theta} \boldsymbol{h}_{2,R}|^2 &= |h_{2,f} + \boldsymbol{v}^H \boldsymbol{\Psi}|^2. \end{aligned} \tag{C.12}$$

By defining an auxiliary variable κ, then an equivalent representation of the achievable rate of the CEUs can be obtained as

$$\mathcal{R}_f = \log\left[1 + \frac{\alpha_{1,m,f}^* P_1 \mathrm{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f}) + \alpha_{2,n,f}^* P_2 \mathrm{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f})}{(1-\alpha_{1,m,f}^*) P_1 \mathrm{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f}) + (1-\alpha_{2,n,f}^*) P_2 \mathrm{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f}) + \sigma^2}\right], \tag{C.13}$$

where

$$\boldsymbol{Q}_{1,f} = \begin{bmatrix} \boldsymbol{\Phi}\boldsymbol{\Phi}^H & \boldsymbol{\Phi} h_{1,f}^H \\ h_{1,f}\boldsymbol{\Phi}^H & 0 \end{bmatrix}, \quad \boldsymbol{Q}_{2,f} = \begin{bmatrix} \boldsymbol{\Psi}\boldsymbol{\Psi}^H & \boldsymbol{\Psi} h_{2,f}^H \\ h_{2,f}\boldsymbol{\Psi}^H & 0 \end{bmatrix}, \qquad \bar{\boldsymbol{v}} = \begin{bmatrix} \boldsymbol{v} \\ \kappa \end{bmatrix}. \tag{C.14}$$

Note that $\bar{\boldsymbol{v}}^H \boldsymbol{Q}_z \bar{\boldsymbol{v}} = \mathrm{tr}(\boldsymbol{Q}_z \bar{\boldsymbol{v}} \bar{\boldsymbol{v}}^H)$ for all $z \in \{\{1, f\}, \{2, f\}\}$. In addition, define $\boldsymbol{U} = \bar{\boldsymbol{v}}\bar{\boldsymbol{v}}^H$, which needs to satisfy $\mathrm{rank}(\boldsymbol{U}) = 1$ and $\boldsymbol{U} \succeq \boldsymbol{0}$. It is worth to mention that in (C.13), the

PA coefficients $\alpha^*_{1,m,f}$ and $\alpha^*_{2,n,f}$ represent the optimal power coefficients of the *coordinated NOMA cluster* (CCU^1_m, CEU_f, CCU^2_n) after obtaining the optimal clustering policy $\boldsymbol{X}^\star$. Maximizing CEUs sum-rate in their current forms still results in a non-smooth objective function with non-convex constraints. After following the change-of-variables approach proposed in [186], we can now define the following exponential slack variables

$$\begin{aligned} e^{y_f} &= P_1 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f}) + P_2 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f}) + \sigma^2, \\ e^{q_f} &= (1-\alpha^*_{1,m,f}) P_1 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f}) + (1-\alpha^*_{2,n,f}) P_2 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f}) + \sigma^2. \end{aligned} \tag{C.15}$$

Next, after following simple exponential and logarithmic properties and dropping the rank-1 constraint on $\boldsymbol{U}$ through SDR, the maximization optimization problem for the phase-shift matrix design can be written as follow

$$\text{OPT} - \text{PS} : \max_{\boldsymbol{U},\boldsymbol{y},\boldsymbol{q}} \sum_{f\in\mathcal{F}} (y_f - q_f), \tag{C.16a}$$

$$\text{s.t.} \quad y_f - q_f \geq R^{\text{th}}_f, \qquad \forall f \in \mathcal{F}, \tag{C.16b}$$

$$P_1 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f}) + P_2 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f}) + \sigma^2 \geq e^{y_f}, \quad f \in \mathcal{F}, \tag{C.16c}$$

$$(1-\alpha^*_{1,m,f}) P_1 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{1,f}) + (1-\alpha^*_{2,n,f}) P_2 \text{tr}(\boldsymbol{U}\boldsymbol{Q}_{2,f}) + \sigma^2 \leq e^{q_f}, f \in \mathcal{F}, \tag{C.16d}$$

$$\boldsymbol{U} \succeq 0, \tag{C.16e}$$

$$[\boldsymbol{U}]_{l,l} = 1, \qquad \forall\, l \in [\![1, L+1]\!], \tag{C.16f}$$

where $\boldsymbol{y} = [y_1, \ldots, y_F]^T$, and $\boldsymbol{q} = [q_1, \ldots, q_F]^T$. The only non-convex constraint in (C.16) is (C.16d). In order to tackle this challenge, we resort to the SCA technique. Accordingly, we apply a first-order Taylor approximation such that $e^{q_f} = e^{\bar{q}_f}(q_f - \bar{q}_f + 1)$, where the linearization is made around $\bar{q}_f$. Consequently, (C.16) can be rewritten as

$$\text{OPT} - \text{PS} : \max_{\boldsymbol{U},\boldsymbol{y},\boldsymbol{q}} \sum_{f\in\mathcal{F}} (y_f - q_f) \tag{C.17a}$$

$$\text{s.t.} \quad (1-\alpha^*_{1,m,f}) P_1 \text{tr}\left(\boldsymbol{U}\boldsymbol{Q}_{1,f}\right) + (1-\alpha^*_{2,n,f}) P_2 \text{tr}\left(\boldsymbol{U}\boldsymbol{Q}_{2,f}\right) + \sigma^2 \leq e^{\bar{q}_f}(q_f - \bar{q}_f + 1), f \in \mathcal{F}, \tag{C.17b}$$

$$\text{(C.16b)}, \text{(C.16c)}, \text{(C.16e)}, \text{(C.16f)}. \tag{C.17c}$$

At this point, it can be observed that all the constraints in (C.17) are convex, which can be iteratively solved using software packages such as CVX. Finally, with the optimal solution of (C.17), i.e., $\boldsymbol{U}^* = \bar{\boldsymbol{v}}^* \bar{\boldsymbol{v}}^{*H}$, the optimal phase-shift matrix $\boldsymbol{\Theta}^*$ can be obtained. However, if $\text{rank}(\boldsymbol{U}^*) \neq 1$, the Gaussian randomization method will be applied to construct a rank-one solution from the higher-rank solution obtained by solving the relaxed problem. This can be achieved as follows.

Algorithm 8: Proposed Algorithm for PS Matrix Design

Input: $\boldsymbol{Q}_{1,f}, \boldsymbol{Q}_{2,f}, \mathcal{P}^*, \boldsymbol{X}^\star, P_1, P_2, \sigma^2, R_f^{\text{th}}$, maximum number of iteration J_3, iteration index $t = 1$, and maximum tolerance ϵ;

Define: $\text{err}^{[t]} = \max_f \left[\bar{q}_f^{[t+1]} - \bar{q}_f^{[t]}\right]$;

while *($t \leq J_3$)* **do**

  Solve (C.17) to obtain $\boldsymbol{U}^*, q_f, y_f, \forall f \in \mathcal{F}$ and evaluate $\text{err}^{[t]}$;

  if $\text{err}^{[t]} < \epsilon$ **then**

    break;

  end if

  else

    Update $\bar{\boldsymbol{q}}^{[t]} := \boldsymbol{q}^{[t]}$;

    Increment $t := t + 1$;

  end if

end while

Call **Algorithm 9** to calculate $\boldsymbol{\Theta}^*$;

Algorithm 9: Obtaining Phase-Shift Matrix $\boldsymbol{\Theta}^*$

Input: Optimal solution of $\boldsymbol{U}^*$;

Initialize the maximum number of candidate random vectors as $\mathcal{A}$;

if *(*$\text{rank}(\boldsymbol{U}^*) = 1$*)* **then**

  Calculate the eigenvalue λ and eigenvector $\boldsymbol{\vartheta}$ of $\boldsymbol{U}^*$ according to $\boldsymbol{U^*\vartheta} = \lambda\boldsymbol{\vartheta}$;

  Update $\boldsymbol{\Theta}^* := \text{diag}\{\sqrt{\lambda\boldsymbol{\vartheta}}\}$;

end if

else

  Obtain the eigenvalue decomposition using (C.18);

  for *$a = 1 : \mathcal{A}$* **do**

    Obtain a phase-shift matrix solution using (C.19) and (C.20);

  end for

  Find the optimal value of the phase-shift matrix, i.e., $\boldsymbol{\Theta}^*$ using (C.21);

end if

Output: The optimal phase-shift matrix $\boldsymbol{\Theta}^*$;

First, if $\text{rank}(\boldsymbol{U}^*) = 1$, the optimal phase-matrix can be obtained by finding the eigenvalue and eigenvector of $\boldsymbol{U^*}$. On the other hand, when $\text{rank}(\boldsymbol{U}^*) \neq 1$, the Gaussian randomization method is invoked, and the eigenvalue decomposition of $\boldsymbol{U}^*$ is defined as

$$\boldsymbol{U}^* = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^H, \tag{C.18}$$

where $\boldsymbol{V} = [\vartheta_1, \vartheta_2, \ldots, \vartheta_{L+1}]$ is a unitary matrix and $\boldsymbol{\Sigma} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{L+1})$ is a diagonal matrix, respectively. Then, we generate a random vector denoted as r_a, where $\boldsymbol{r}_a, \forall a \in [\![1, \mathcal{A}]\!]$ is a random vector that follows a circularly symmetric complex Gaussian (CSCG) distribution with a zero mean and a co-variance matrix $\boldsymbol{\Lambda} = \boldsymbol{I}_{L+1}$, i.e., $\mathbf{r} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Lambda})$ and $\mathcal{A}$ is the maximum generation of random vectors. Based on that, we can

obtain a sub-optimal solution to (C.16), denoting as

$$\bar{\boldsymbol{v}}_a = \boldsymbol{V}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{r}_a, \quad \forall a \in [\![1, \mathcal{A}]\!]. \tag{C.19}$$

Next, the candidate phase-shift matrix can be expressed as

$$\boldsymbol{\Theta}_a = \text{diag}\left\{\exp\left[j \arg\left(\frac{[\bar{\boldsymbol{v}}_a]_{1:L}}{[\bar{\boldsymbol{v}}_a]_{L+1}}\right)\right]\right\}, \quad \forall a \in [\![1, \mathcal{A}]\!], \tag{C.20}$$

where $[\boldsymbol{x}]_{1:L}$ denotes a vector having the first L elements in $\boldsymbol{x}$. With the obtained candidate set of phase-shift matrix $\boldsymbol{\Theta}_a | a \in [\![1, \mathcal{A}]\!]$, we can obtain the optimal one that maximizes the sum-rate of the cell-edge users as shown in (C.21).

$$\boldsymbol{\Theta}^* = \arg\max_a \sum_{f=1}^{F} \log\left[1+ \right.$$
$$\left. \frac{\alpha^*_{1,m,f} P_1 |h_{1,f} + \boldsymbol{h}_{R,f}\boldsymbol{\Theta}_a \boldsymbol{h}_{1,R}|^2 + \alpha^*_{2,n,f} P_2 |h_{2,f} + \boldsymbol{h}_{R,f}\boldsymbol{\Theta}_a \boldsymbol{h}_{2,R}|^2}{(1-\alpha^*_{1,m,f}) P_1 |h_{1,f} + \boldsymbol{h}_{R,f}\boldsymbol{\Theta}_a \boldsymbol{h}_{1,R}|^2 + (1-\alpha^*_{2,n,f}) P_2 |h_{2,f} + \boldsymbol{h}_{R,f}\boldsymbol{\Theta}_a \boldsymbol{h}_{2,R}|^2 + \sigma^2}\right], \tag{C.21}$$

The design approach for the phase-matrix is presented by **Algorithm 8**. At this point, the overall proposed alternating optimization algorithm for the network sum-rate maximization in RIS-aided CoMP NOMA cellular network is summarized in **Algorithm 4**. However, the SDR is invoked instead of the DC approach to obtain the phase-shift matrix.

Chapter 6

D1 Proof of Theorem 6.2

The constraints of problem P − PA can be rewritten as

$$\max\left(\frac{1}{2}, \frac{(\gamma_{1,m}+1)\,t_m}{(t_m+1)\,\gamma_{1,m}}\right) \leq \alpha_{1,m,f} \leq \frac{\gamma_{1,m}-t_m}{\gamma_{1,m}}, \tag{D.1a}$$

$$\max\left(\frac{1}{2}, \frac{(\gamma_{2,n}+1)\,t_n}{(t_n+1)\,\gamma_{2,n}}\right) \leq \alpha_{2,n,f} \leq \frac{\gamma_{2,n}-t_n}{\gamma_{2,n}}, \tag{D.1b}$$

$$\alpha_{1,m,f}\,(1+t_f)\,\gamma_{1,f} + \alpha_{2,n,f}\,(1+t_f)\,\gamma_{2,f} \geq t_f\,(\gamma_{1,f}+\gamma_{2,f}+1)\,. \tag{D.1c}$$

Based on this, in order for constraint (D.1a) to be feasible, t_m should satisfy the condition $t_m \leq \min\left(\frac{\gamma_{1,m}}{2}, -1+\sqrt{1+\gamma_{1,m}}\right)$, which represents the first condition. Similarly, in order for constraints (D.1b) to be feasible, t_n should satisfy the condition $t_n \leq \min\left(\frac{\gamma_{2,n}}{2}, -1+\sqrt{1+\gamma_{2,n}}\right)$, which represents the second condition. Finally, in order to have feasible factors $\alpha_{1,m,f}$ and $\alpha_{2,n,f}$ that satisfy constraint (D.1c), then then maximum feasible values of $\alpha_{1,m,f}$ and $\alpha_{2,n,f}$, which are $\frac{\gamma_{1,m}-t_m}{\gamma_{1,m}}$ and $\frac{\gamma_{2,n}-t_n}{\gamma_{2,n}}$ respectively, should satisfy constraint (D.1c), i.e., $\frac{(\gamma_{1,m}-t_m)}{\gamma_{1,m}}\,(1+t_f)\,\gamma_{1,f} + \frac{(\gamma_{2,n}-t_n)}{\gamma_{2,n}}\,(1+t_f)\,\gamma_{2,f} \geq t_f\,(\gamma_{1,f}+\gamma_{2,f}+1)$, which represents the third condition and completes the proof.

D2 Proof of Theorem 6.3

Based on the proof in Appendix D1, the power control factors $\alpha_{1,m,f}$ and $\alpha_{2,n,f}$ satisfy

$$\max\left(-\frac{\gamma_{2,f}}{\gamma_{1,f}}\alpha_{2,n,f} + \frac{t_f\,(\gamma_{1,f}+\gamma_{2,f}+1)}{(1+t_f)\,\gamma_{1,f}}, \alpha_{1,m,f}^{\min}\right) \leq \alpha_{1,m,f} \leq \alpha_{1,m,f}^{\max}, \tag{D.2a}$$

$$\max\left(-\frac{\gamma_{1,f}}{\gamma_{2,f}}\alpha_{1,m,f} + \frac{t_f\,(\gamma_{1,f}+\gamma_{2,f}+1)}{(1+t_f)\,\gamma_{2,f}}, \alpha_{2,n,f}^{\min}\right) \leq \alpha_{2,n,f} \leq \alpha_{2,n,f}^{\max}, \tag{D.2b}$$

where $\alpha_{1,m,f}^{\min}, \alpha_{1,m,f}^{\max}, \alpha_{2,n,f}^{\min}$ and $\alpha_{2,n,f}^{\max}$ are expressed as shown in (6.41) in ***Theorem 6.3***. Consequently, one can conclude that $\alpha_{m,f,n}^{\min} \leq \alpha_{1,m,f} \leq \alpha_{m,f,n}^{\max}$, where

$$\alpha_{m,f,n}^{\min} = \max\left(\alpha_{1,m,f}^{\min}, -\frac{\gamma_{2,f}}{\gamma_{1,f}}\alpha_{2,n,f}^{\max} + \frac{t_f\,(\gamma_{1,f}+\gamma_{2,f}+1)}{(1+t_f)\,\gamma_{1,f}}\right), \tag{D.3a}$$

$$\alpha_{m,f,n}^{\max} = \min\left(\alpha_{1,m,f}^{\max}, -\frac{\gamma_{2,f}}{\gamma_{1,f}}\alpha_{2,n,f}^{\min} + \frac{t_f\,(\gamma_{1,f}+\gamma_{2,f}+1)}{(1+t_f)\,\gamma_{1,f}}\right). \tag{D.3b}$$

Based on this, since $R_f = R_f^{\text{th}}$, then solving problem $\mathrm{P}-\mathrm{PA}$ is reduced to solve the following optimization problem

$$\mathcal{P}'_{m,f,n}: \quad \max_{\alpha_{1,m,f}} \; f\left(\alpha_{1,m,f}\right), \tag{D.4a}$$

$$\text{s.t.} \quad \alpha_{m,f,n}^{\min} \leq \alpha_{1,m,f} \leq \alpha_{m,f,n}^{\min}, \tag{D.4b}$$

where the objective function $f(\cdot)$ is expressed

$$\begin{aligned} f\left(\alpha_{1,m,f}\right) &= \log\left[\left(1+\gamma_{1,m}\right) - \gamma_{1,m}\alpha_{1,m,f}\right] \\ &\quad + \log\left[\left(1+\gamma_{2,n}\right) - \gamma_{2,n}\left(-\frac{\gamma_{1,f}}{\gamma_{2,f}}\alpha_{1,m,f} + \frac{t_f\left(\gamma_{1,f}+\gamma_{2,f}+1\right)}{\left(1+t_f\right)\gamma_{2,f}}\right)\right], \qquad \text{(D.5)} \\ &= \log\left[\left(1+\gamma_{1,m}\right) - \gamma_{1,m}\alpha_{1,m,f}\right] \\ &\quad + \log\left[\left(1+\gamma_{2,n} - \frac{\gamma_{2,n}t_f\left(\gamma_{1,f}+\gamma_{2,f}+1\right)}{\left(1+t_f\right)\gamma_{2,f}}\right) - \frac{\gamma_{2,n}\gamma_{1,f}}{\gamma_{2,f}}\alpha_{1,m,f}\right]. \qquad \text{(D.6)} \end{aligned}$$

The gradient of the objective function with respect to $\alpha_{1,m,f}$ is expressed as

$$\frac{\partial f}{\partial \alpha_{1,m,f}}\left(\alpha_{1,m,f}\right) = \frac{2(\alpha_{1,m,f}-x_0)}{\left(\alpha_{1,m,f}-x_1\right)\left(\alpha_{1,m,f}-x_2\right)}, \tag{D.7}$$

where

$$\begin{aligned} x_0 &= \frac{1}{2}\left(\frac{1+\gamma_{1,m}}{\gamma_{1,m}} - \frac{\gamma_{2,f}\left(1+\gamma_{2,n}\right)}{\gamma_{1,f}\gamma_{2,n}} + \frac{t_f\left(\gamma_{1,f}+\gamma_{2,f}+1\right)}{\gamma_{1,f}\left(1+t_f\right)}\right), \\ x_1 &= \frac{1+\gamma_{1,m}}{\gamma_{1,m}}, \\ x_2 &= \frac{\gamma_{2,f}}{\gamma_{1,f}}\left(\frac{t_f\left(\gamma_{1,f}+\gamma_{2,f}+1\right)}{\gamma_{2,f}\left(1+t_f\right)} - \frac{1+\gamma_{2,n}}{\gamma_{2,n}}\right). \end{aligned} \tag{D.8}$$

Since $\alpha_{m,f,n}^{\max} \leq \alpha_{1,m,f}^{\max} \leq 1 \leq \frac{1+\gamma_{1,m}}{\gamma_{1,m}}$, we have $\alpha_{m,f,n}^{\max} \leq x_1$ and since $-\frac{\gamma_{2,f}}{\gamma_{1,f}}\alpha_{2,n,f}^{\max} + \frac{t_f\left(\gamma_{1,f}+\gamma_{2,f}+1\right)}{\left(1+t_f\right)\gamma_{1,f}} \leq \alpha_{m,f,n}^{\min}$ and $\alpha_{2,n,f}^{\max} \leq 1 \leq \frac{1+\gamma_{2,n}}{\gamma_{2,n}}$, we have $x_2 \leq \alpha_{m,f,n}^{\min}$. Therefore, the function $\alpha_{1,m,f} \mapsto \frac{2}{\left(\alpha_{1,m,f}-x_1\right)\left(\alpha_{1,m,f}-x_2\right)}$ is negative within the interval $[\alpha_{\min}, \alpha_{\max}]$. Hence, the sign of the gradient of f depends on the relation between x_0, $\alpha_{m,f,n}^{\min}$ and $\alpha_{m,f,n}^{\max}$. Based on this, the optimal solution $\alpha_{1,m,f}^*$ is expressed as $\alpha_{1,m,f}^* = \alpha_{m,f,n}^{\min}$ if $x_0 \leq \alpha_{m,f,n}^{\min}$, and $\alpha_{1,m,f}^* = x_0$ if $\alpha_{\min} \leq x_0 \leq \alpha_{\max}$ and $\alpha_{1,m,f}^* = \alpha_{m,f,n}^{\max}$ if $\alpha_{m,f,n}^{\max} \leq x_0$, which completes the proof.

Bibliography

[1] U. Cisco, "*Cisco Annual Internet Report (2018–2023) White Paper*," [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html, Mar. 2020.

[2] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G Networks: Use Cases and Technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.

[3] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.

[4] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G Wireless Communications: Vision and Potential Techniques," *IEEE Netw.*, vol. 33, no. 4, pp. 70–75, Jul./Aug. 2019.

[5] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The Road Towards 6G: A Comprehensive Survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, Feb. 2021.

[6] Y. Liu, S. Zhang, X. Mu, Z. Ding, R. Schober, N. Al-Dhahir, E. Hossain, and X. Shen, "Evolution of NOMA Toward Next Generation Multiple Access (NGMA) for 6G," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1037–1071, Apr. 2022.

[7] Y. Liu, W. Yi, Z. Ding, X. Liu, O. A. Dobre, and N. Al-Dhahir, "Developing NOMA to Next Generation Multiple Access (NGMA): Future Vision and Research Opportunities," *IEEE Wireless Commun.*, pp. 1–8, Jun. 2022.

[8] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart Radio Environments Empowered by Reconfigurable Intelligent Surfaces: How It Works, State of Research, and The Road Ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[9] M. Elhattab, M.-A. Arfaoui, and C. Assi, "CoMP Transmission in Downlink NOMA-Based Heterogeneous Cloud Radio Access Networks," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7779–7794, 2020.

[10] M. Elhattab, M. A. Arfaoui, and C. Assi, "A Joint CoMP C-NOMA for Enhanced Cellular System Performance," *IEEE Commun. Lett.*, vol. 24, no. 9, pp. 1919–1923, Sept. 2020.

[11] M. Elhattab, M. A. Arfaoui, and C. Assi, "Power allocation in CoMP-empowered C-NOMA networks," *IEEE Networking Letters*, vol. 3, no. 1, pp. 10–14, Mar. 2021.

[12] M. Elhattab, M. A. Arfaoui, and C. Assi, "Joint Clustering and Power Allocation in Coordinated Multipoint Assisted C-NOMA Cellular Networks," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3483–3498, May 2022.

[13] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "Reconfigurable Intelligent Surface Enabled Full-Duplex/Half-Duplex Cooperative Non-Orthogonal Multiple Access," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3349–3364, May 2022.

[14] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "Reconfigurable Intelligent Surface Assisted Coordinated Multipoint in Downlink NOMA Networks," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 632–636, Mar. 2021.

[15] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "RIS-Assisted Joint Transmission in a Two-Cell Downlink NOMA Cellular System," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1270–1286, Apr. 2022.

[16] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal Multiple Access for 5G and Beyond," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2347–2381, 2017.

[17] R. Steele and L. Hanzo, *Mobile radio communications: Second and third generation cellular and WATM systems: 2nd.* IEEE Press-John Wiley, 1999.

[18] K. Gilhousen, I. Jacobs, R. Padovani, A. Viterbi, L. Weaver, and C. Wheatley, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 303–312, 1991.

[19] J. Li, X. Wu, and R. Laroia, *OFDMA mobile broadband communications: A systems approach.* Cambridge University Press, 2013.

[20] O. Maraqa, A. S. Rajasekaran, S. Al-Ahmadi, H. Yanikomeroglu, and S. M. Sait, "A survey of rate-optimal power domain noma with enabling technologies of future wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2192–2235, 4th quarter 2020.

[21] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

[22] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A Survey of Non-Orthogonal Multiple Access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294–2323, 3rd quarter 2018.

[23] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-s. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2nd quarter 2017.

[24] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond.* Springer, 2019, vol. 159.

[25] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, 2014.

[26] Y. Liu, Z. Qin, M. Elkashlan, A. Nallanathan, and J. A. McCann, "Non-Orthogonal Multiple Access in Large-Scale Heterogeneous Networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2667–2680, Dec. 2017.

[27] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Nonorthogonal Multiple Access in Large-Scale Underlay Cognitive Radio Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 152–10 157, Dec. 2016.

[28] Y. Liu, M. Elkashlan, Z. Ding, and G. K. Karagiannidis, "Fairness of User Clustering in MIMO Non-Orthogonal Multiple Access Systems," *IEEE Commun. Lett.*, vol. 20, no. 7, pp. 1465–1468, Jul. 2016.

[29] Z. Ding, M. Peng, and H. V. Poor, "Cooperative Non-Orthogonal Multiple Access in 5G Systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.

[30] Z. Zhang, Z. Ma, M. Xiao, Z. Ding, and P. Fan, "Full-Duplex Device-to-Device-Aided Cooperative Nonorthogonal Multiple Access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4467–4471, May 2017.

[31] C. Zhong and Z. Zhang, "Non-Orthogonal Multiple Access With Cooperative Full-Duplex Relaying," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2478–2481, Dec. 2016.

[32] G. Liu, X. Chen, Z. Ding, Z. Ma, and F. R. Yu, "Hybrid Half-Duplex/Full-Duplex Cooperative Non-Orthogonal Multiple Access With Transmit Power Adaptation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 506–519, Jan. 2018.

[33] X. Yue, Y. Liu, S. Kang, A. Nallanathan, and Z. Ding, "Exploiting Full/Half-Duplex User Relaying in NOMA Systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 560–575, Feb. 2018.

[34] X. Yue, Y. Liu, S. Kang, A. Nallanathan, and Z. Ding, "Spatially Random Relay Selection for Full/Half-Duplex Cooperative NOMA Networks," *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3294–3308, Aug. 2018.

[35] M. Zeng, W. Hao, O. A. Dobre, and Z. Ding, "Cooperative NOMA: State of the Art, Key Techniques, and Open Challenges," *IEEE Network*, vol. 34, no. 5, pp. 205–211, Sep./Oct. 2020.

[36] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of Loopback Self-Interference in Full-Duplex MIMO Relays," *IEEE Trans. Signal Processing*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[37] S. Bassoy, H. Farooq, M. A. Imran, and A. Imran, "Coordinated Multi-Point Clustering Schemes: A Survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 743–764, 2nd quarter 2017.

[38] G. T. 36.819, "Coordinated multi-point operation for lte physical layer aspects," 2013.

[39] J. Lee, Y. Kim, H. Lee, B. L. Ng, D. Mazzarese, J. Liu, W. Xiao, and Y. Zhou, "Coordinated multipoint transmission and reception in LTE-advanced systems," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 44–50, Nov. 2012.

[40] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[41] R. Tanbourgi, S. Singh, J. G. Andrews, and F. K. Jondral, "A Tractable Model for Noncoherent Joint-Transmission Base Station Cooperation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4959–4973, Sept. 2014.

[42] H.-M. Wang, K.-W. Huang, and T. A. Tsiftsis, "Base Station Cooperation in Millimeter Wave Cellular Networks: Performance Enhancement of Cell-Edge Users," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 5124–5139, Nov. 2018.

[43] A. K. Gupta, R. Sahu, K. K. Chaurasia, and S. Tripathi, "On the Analysis of a Cellular Network With Macro-Diversity Using Stochastic Geometry," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3244–3248, Oct. 2021.

[44] Y. Sun, M. Peng, and S. Mao, "A Game-Theoretic Approach to Cache and Radio Resource Management in Fog Radio Access Networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10 145–10 159, Oct. 2019.

[45] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, and L. Hanzo, "The Non-Coherent Ultra-Dense C-RAN Is Capable of Outperforming Its Coherent Counterpart at a Limited Fronthaul Capacity," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2549–2560, 2018.

[46] A. Guo and M. Haenggi, "Spatial Stochastic Models and Metrics for the Structure of Base Stations in Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5800–5812, Nov. 2013.

[47] A. H. Sakr and E. Hossain, "Location-Aware Cross-Tier Coordinated Multipoint Transmission in Two-Tier Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6311–6325, Nov. 2014.

[48] J. Yao and N. Ansari, "QoS-Aware Joint BBU-RRH Mapping and User Association in Cloud-RANs," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 4, pp. 881–889, Dec. 2018.

[49] J. Tang, W. P. Tay, T. Q. S. Quek, and B. Liang, "System Cost Minimization in Cloud RAN With Limited Fronthaul Capacity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3371–3384, May 2017.

[50] L. Ferdouse, A. Anpalagan, and S. Erkucuk, "Joint Communication and Computing Resource Allocation in 5G Cloud Radio Access Networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 9122–9135, Sept. 2019.

[51] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Network*, vol. 29, no. 2, pp. 6–14, Mar./Apr. 2015.

[52] B. Dai and W. Yu, "Energy Efficiency of Downlink Transmission Strategies for Cloud Radio Access Networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 1037–1050, 2016.

[53] D. Zeng, J. Zhang, L. Gu, S. Guo, and J. Luo, "Energy-Efficient Coordinated Multipoint Scheduling in Green Cloud Radio Access Network," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9922–9930, Oct. 2018.

[54] M. Elhattab and W. Hamouda, "Performance Analysis for H-CRANs Under Constrained Capacity Fronthaul," *IEEE Networking Lett.*, vol. 2, no. 2, pp. 62–66, Jun. 2020.

[55] M. K. Elhattab, M. M. Elmesalawy, T. Ismail, H. H. Esmat, M. M. Abdelhakam, and H. Selmy, "A Matching Game for Device Association and Resource Allocation in Heterogeneous Cloud Radio Access Networks," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1664–1667, Aug. 2018.

[56] E. Bjornson, O. Ozdogan, and E. G. Larsson, "Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 90–96, Dec. 2020.

[57] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward Smart Wireless Communications via Intelligent Reflecting Surfaces: A Contemporary Survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2283–2314, 3rd quarter 2020.

[58] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan 2020.

[59] E. Bjornson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable Intelligent Surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, Mar. 2022.

[60] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd quarter 2021.

[61] J. Wang, W. Tang, Y. Han, S. Jin, X. Li, C.-K. Wen, Q. Cheng, and T. J. Cui, "Interplay Between RIS and AI in Wireless Communications: Fundamentals, Architectures, Applications, and Open Research Problems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2271–2288, Aug. 2021.

[62] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, and M. Gidlund, "Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6g wireless networks," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 184–191, Dec. 2021.

[63] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov, and S. Shamai, "Reconfigurable Intelligent Surfaces vs. Relaying: Differences, Similarities, and Performance Comparison," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 798–807, 2020.

[64] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research Directions," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, Jun. 2021.

[65] M. Di Renzo, "From walls at the office to sides of buildings, what if these surfaces can manage the way in which wireless signals propagate?" [Online]. Available: https://www.comsoc.org/publications/ctn/walls-office-sides-buildings-what-if-these-surfaces-can-manage-way-which-wireless, Sept. 2021.

[66] X. Gu, X. Ji, Z. Ding, W. Wu, and M. Peng, "Outage Probability Analysis of Non-Orthogonal Multiple Access in Cloud Radio Access Networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 149–152, Jan. 2018.

[67] M. S. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink Power Allocation for CoMP-NOMA in Multi-Cell Networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3982–3998, Sept. 2018.

[68] M. S. Ali, E. Hossain, and D. I. Kim, "Coordinated Multipoint Transmission in Downlink Multi-Cell NOMA Systems: Models and Spectral Efficiency Performance," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 24–31, Apr. 2018.

[69] Z. Zhang, H. Sun, and R. Q. Hu, "Downlink and Uplink Non-Orthogonal Multiple Access in a Dense Wireless Network," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2771–2784, Dec. 2017.

[70] K. S. Ali, M. Haenggi, H. ElSawy, A. Chaaban, and M.-S. Alouini, "Downlink Non-Orthogonal Multiple Access (NOMA) in Poisson Networks," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1613–1628, Feb. 2019.

[71] L. Qi, M. Peng, Y. Liu, and S. Yan, "Advanced User Association in Non-Orthogonal Multiple Access-Based Fog Radio Access Networks," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8408–8421, Dec. 2019.

[72] Q. Y. Liau and C. Y. Leow, "Successive User Relaying in Cooperative NOMA System," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 921–924, Jun. 2019.

[73] A. Tregancini, E. E. B. Olivo, D. P. M. Osorio, C. H. M. de Lima, and H. Alves, "Performance Analysis of Full-Duplex Relay-Aided NOMA Systems Using Partial Relay Selection," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 622–635, Jan. 2020.

[74] Y. Li, Y. Li, X. Chu, Y. Ye, and H. Zhang, "Performance Analysis of Relay Selection in Cooperative NOMA Networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 760–763, Apr. 2019.

[75] Y. Tian, A. R. Nix, and M. Beach, "On the Performance of Opportunistic NOMA in Downlink CoMP Networks," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 998–1001, May 2016.

[76] J. Choi, "Non-Orthogonal Multiple Access in Downlink Coordinated Two-Point Systems," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 313–316, Feb. 2014.

[77] Z. Liu, G. Kang, L. Lei, N. Zhang, and S. Zhang, "Power Allocation for Energy Efficiency Maximization in Downlink CoMP Systems with NOMA," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2017, pp. 1–6.

[78] C.-H. Liu and D.-C. Liang, "Heterogeneous Networks With Power-Domain NOMA: Coverage, Throughput, and Power Allocation Analysis," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3524–3539, May 2018.

[79] H. ElSawy, A. Sultan-Salem, M.-S. Alouini, and M. Z. Win, "Modeling and Analysis of Cellular Networks Using Stochastic Geometry: A Tutorial," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 167–203, 1st quarter 2017.

[80] P. Parida, H. S. Dhillon, and A. F. Molisch, "Downlink Performance Analysis of Cell-Free Massive MIMO with Finite Fronthaul Capacity," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Apr. 2018, pp. 1–6.

[81] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous Cellular Networks with Flexible Cell Association: A Comprehensive Downlink SINR Analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[82] M. Di Renzo, A. Zappone, T. T. Lam, and M. Debbah, "System-Level Modeling and Optimization of the Energy Efficiency in Cellular Networks-A Stochastic Geometry

Framework," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2539–2556, Apr. 2018.

[83] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and Analysis of K-Tier Downlink Heterogeneous Cellular Networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[84] R. Tao, W. Liu, X. Chu, and J. Zhang, "An Energy Saving Small Cell Sleeping Mechanism With Cell Range Expansion in Heterogeneous Networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2451–2463, May 2019.

[85] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522–2545, 2016.

[86] D. Zhai, R. Zhang, Y. Wang, H. Sun, L. Cai, and Z. Ding, "Joint User Pairing, Mode Selection, and Power Control for D2D-Capable Cellular Networks Enhanced by Nonorthogonal Multiple Access," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8919–8932, Oct. 2019.

[87] W. Wu, F. Zhou, R. Q. Hu, and B. Wang, "Energy-Efficient Resource Allocation for Secure NOMA-Enabled Mobile Edge Computing Networks," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 493–505, Jan. 2020.

[88] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[89] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[90] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A Tractable Approach to Coverage and Rate in Cellular Networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[91] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic Geometry for Modeling, Analysis, and Design of Multi-Tier and Cognitive Cellular Wireless Networks: A Survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 996–1019, 3rd quarter 2013.

[92] X. Lin, J. G. Andrews, and A. Ghosh, "Modeling, Analysis and Design for Carrier Aggregation in Heterogeneous Cellular Networks," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 4002–4015, Sept. 2013.

[93] X. Zhang and J. G. Andrews, "Downlink Cellular Network Analysis With Multi-Slope Path Loss Models," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1881–1894, May 2015.

[94] G. Zhao, S. Chen, L. Zhao, and L. Hanzo, "Joint Energy-Spectral-Efficiency Optimization of CoMP and BS Deployment in Dense Large-Scale Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4832–4847, Jul. 2017.

[95] J. Li, M. Peng, A. Cheng, Y. Yu, and C. Wang, "Resource Allocation Optimization for Delay-Sensitive Traffic in Fronthaul Constrained Cloud Radio Access Networks," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2267–2278, Dec. 2017.

[96] D. Fooladivanda and C. Rosenberg, "Joint Resource Allocation and User Association for Heterogeneous Wireless Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 248–257, Jan. 2013.

[97] F. Zhou, Y. Wu, R. Q. Hu, Y. Wang, and K. K. Wong, "Energy-efficient NOMA enabled heterogeneous cloud radio access networks," *IEEE Net.*, vol. 32, no. 2, pp. 152–160, Mar./Apr. 2018.

[98] Z. Ding, P. Fan, and H. V. Poor, "Impact of User Pairing on 5G Nonorthogonal Multiple-Access Downlink Transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[99] T. Han, J. Gong, X. Liu, S. M. R. Islam, Q. Li, Z. Bai, and K. S. Kwak, "On Downlink NOMA in Heterogeneous Networks With Non-Uniform Small Cell Deployment," *IEEE Access*, vol. 6, pp. 31 099–31 109, Jun. 2018.

[100] V. Garcia, Y. Zhou, and J. Shi, "Coordinated Multipoint Transmission in Dense Cellular Networks With User-Centric Adaptive Clustering," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4297–4308, Aug 2014.

[101] Y. Liu, "Wireless Information and Power Transfer for Multirelay-Assisted Cooperative Communication," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 784–787, Apr. 2016.

[102] V. Garcia, N. Lebedev, and J.-M. Gorce, "Capacity Outage Probability for Multi-Cell Processing Under Rayleigh Fading," *IEEE Commun. Lett.*, vol. 15, no. 8, pp. 801–803, Aug. 2011.

[103] R. Lei and D. Xu, "On the outage performance of JT-CoMP-CNOMA networks with SWIPT," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 432–436, Feb. 2021.

[104] J. Bae and Y. Han, "Joint power and time allocation for two-way cooperative NOMA," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 12 443–12 447, Dec. 2019.

[105] B. Ning, W. Hao, A. Zhang, J. Zhang, and G. Gui, "Energy Efficiency-Delay Tradeoff for a Cooperative NOMA System," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 732–735, Apr. 2019.

[106] L. Wang, Y. Ai, N. Liu, and A. Fei, "User Association and Resource Allocation in Full-Duplex Relay Aided NOMA Systems," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10 580–10 596, Dec. 2019.

[107] P. Hũu, M. A. Arfaoui, S. Sharafeddine, C. M. Assi, and A. Ghrayeb, "A Low-Complexity Framework for Joint User Pairing and Power Control for Cooperative NOMA in 5G and Beyond Cellular Networks," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6737–6749, Nov. 2020.

[108] W. Mei and R. Zhang, "Cooperative NOMA for downlink asymmetric interference cancellation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 884–888, June 2020.

[109] M. Hedayati and I. Kim, "CoMP-NOMA in the SWIPT Networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4549–4562, July 2020.

[110] M. R. Zamani, M. Eslami, M. Khorramizadeh, and Z. Ding, "Energy-Efficient Power Allocation for NOMA With Imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 1009–1013, Jan. 2019.

[111] X. Zhang and F. Wang, "Resource Allocation for Wireless Power Transmission Over Full-Duplex OFDMA/NOMA Mobile Wireless Networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 327–344, Feb. 2019.

[112] H. Wang, C. Liu, Z. Shi, Y. Fu, and R. Song, "Power minimization for two-cell IRS-aided NOMA systems with joint detection," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1635–1639, May 2021.

[113] Y. Sun, Z. Ding, X. Dai, and G. K. Karagiannidis, "A Feasibility Study on Network NOMA," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 4303–4317, Sept. 2018.

[114] M. Wu, Q. Song, L. Guo, and A. Jamalipour, "Joint User Pairing and Resource Allocation in a SWIPT-Enabled Cooperative NOMA System," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 6826–6840, July 2021.

[115] M. Obeed, H. Dahrouj, A. M. Salhab, S. A. Zummo, and M. S. Alouini, "User Pairing, Link Selection and Power Allocation for Cooperative NOMA Hybrid VLC/RF Systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1785–1800, Mar. 2021.

[116] Y. Xu, J. Cheng, G. Wang, and V. C. M. Leung, "Adaptive Coordinated Direct and Relay Transmission for NOMA Networks: A Joint Downlink-Uplink Scheme," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4328–4346, July 2021.

[117] W. Cheng, X. Zhang, and H. Zhang, "Full/half duplex based resource allocations for statistical quality of service provisioning in wireless relay networks," in *2012 Proceedings IEEE INFOCOM*, May 2012, pp. 864–872.

[118] J. F. Bard, *Practical bilevel optimization: algorithms and applications.* vol. 30, New York, NY, USA: Springer, 2013.

[119] B. Di, S. Bayat, L. Song, Y. Li, and Z. Han, "Joint User Pairing, Subchannel, and Power Allocation in Full-Duplex Multi-User OFDMA Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8260–8272, Dec. 2016.

[120] M. K. Elhattab, M. M. Elmesalawy, F. M. Salem, and I. I. Ibrahim, "Device-Aware Cell Association in Heterogeneous Cellular Networks: A Matching Game Approach," *IEEE Trans. Green Commun. Net.*, vol. 3, no. 1, pp. 57–66, Mar. 2019.

[121] M. Vaezi, G. A. Aruma Baduge, Y. Liu, A. Arafa, F. Fang, and Z. Ding, "Interplay Between NOMA and Other Emerging Technologies: A Survey," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 4, pp. 900–919, Dec. 2019.

[122] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless Communications Through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.

[123] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah, "Holographic MIMO Surfaces for 6G Wireless Networks: Opportunities, Challenges, and Trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.

[124] B. Zheng, Q. Wu, and R. Zhang, "Intelligent Reflecting Surface-Assisted Multiple Access With User Pairing: NOMA or OMA?" *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 753–757, Apr. 2020.

[125] H. Wang, C. Liu, Z. Shi, Y. Fu, and R. Song, "On Power Minimization for IRS-Aided Downlink NOMA Systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1808–1811, Nov. 2020.

[126] M. Fu, Y. Zhou, Y. Shi, and K. B. Letaief, "Reconfigurable Intelligent Surface Empowered Downlink Non-Orthogonal Multiple Access," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3802–3817, Jun. 2021.

[127] Y. Guo, Z. Qin, Y. Liu, and N. Al-Dhahir, "Intelligent Reflecting Surface Aided Multiple Access Over Fading Channels," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 2015–2027, Mar. 2021.

[128] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Exploiting Intelligent Reflecting Surfaces in NOMA Networks: Joint Beamforming Optimization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6884–6898, Oct. 2020.

[129] W. Ni, X. Liu, Y. Liu, H. Tian, and Y. Chen, "Resource Allocation for Multi-Cell IRS-Aided NOMA Networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4253–4268, July 2021.

[130] F. Fang, Y. Xu, Q.-V. Pham, and Z. Ding, "Energy-Efficient Design of IRS-NOMA Networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14 088–14 092, Nov. 2020.

[131] G. Yang, X. Xu, Y.-C. Liang, and M. D. Renzo, "Reconfigurable Intelligent Surface-Assisted Non-Orthogonal Multiple Access," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3137–3151, May 2021.

[132] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable Intelligent Surface Aided NOMA Networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, Nov. 2020.

[133] C. Zhang, W. Yi, Y. Liu, K. Yang, and Z. Ding, "Reconfigurable Intelligent Surfaces Aided Multi-Cell NOMA Networks: A Stochastic Geometry Model," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 951–966, Feb. 2022.

[134] Z. Xie *et al.*, "Modeling and Coverage Analysis for RIS-aided NOMA Transmissions in Heterogeneous Networks," [Online]. Available: https://arxiv.org/pdf/2104.13182, 2021.

[135] J. Zuo, Y. Liu, and N. Al-Dhahir, "Reconfigurable intelligent surface assisted cooperative non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6750–6764, Oct. 2021.

[136] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel Estimation for RIS-Empowered Multi-User MISO Wireless Communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Jun. 2021.

[137] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling Large Intelligent Surfaces With Compressive Sensing and Deep Learning," *IEEE Access*, vol. 9, pp. 44 304–44 321, Mar. 2021.

[138] Z. Ding and H. Poor, "A Simple Design of IRS-NOMA Transmission," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1119–1123, May 2020.

[139] E. Bjornson, O. Ozdogan, and E. G. Larsson, "Intelligent Reflecting Surface Versus Decode-and-Forward: How Large Surfaces are Needed to Beat Relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

[140] L. Chen, X. Qin, and G. Wei, "A Uniform-Forcing Transceiver Design for Over-the-Air Function Computation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 942–945, July 2018.

[141] E. Chen and M. Tao, "ADMM-Based Fast Algorithm for Multi-Group Multicast Beamforming in Large-Scale Wireless Systems," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2685–2698, Jun. 2017.

[142] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated Learning via Over-the-Air Computation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2022–2035, Mar. 2020.

[143] X. Yu, D. Xu, D. W. K. Ng, and R. Schober, "IRS-Assisted Green Communication Systems: Provable Convergence and Robust Optimization," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 6313–6329, Sept. 2021.

[144] H. Xie, J. Xu, and Y.-F. Liu, "Max-Min Fairness in IRS-Aided Multi-Cell MISO Systems With Joint Transmit and Reflective Beamforming," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1379–1393, Feb. 2021.

[145] A. Khalili, S. Zargari, Q. Wu, D. W. K. Ng, and R. Zhang, "Multi-Objective Resource Allocation for IRS-Aided SWIPT," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1324–1328, Jun. 2021.

[146] R. T. Rockafellar, *Convex analysis*. Princeton university press, 2015.

[147] P. D. Tao and L. T. H. An, "Convex analysis approach to DC programming: theory, algorithms and applications," *Acta mathematica vietnamica*, vol. 22, no. 1, pp. 289–355, 1997.

[148] K. Khamaru and M. Wainwright, "Convergence guarantees for a class of non-convex and non-smooth optimization problems," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80. PMLR, Jul. 2018, pp. 2601 – 2610.

[149] I. Pólik and T. Terlaky, "Interior point methods for nonlinear optimization," in *Nonlinear optimization*. Springer, 2010, pp. 215–276.

[150] S. Ebbesen, P. Kiwitz, and L. Guzzella, "A generic particle swarm optimization matlab function," in *2012 American Control Conference (ACC)*. IEEE, Oct. 2012, pp. 1519–1524.

[151] Z. Li, M. Hua, Q. Wang, and Q. Song, "Weighted Sum-Rate Maximization for Multi-IRS Aided Cooperative Transmission," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1620–1624, Oct. 2020.

[152] E. E. Haber, M. Elhattab, C. Assi, S. Sharafeddine, and K. K. Nguyen, "Latency and Reliability Aware Edge Computation Offloading via an Intelligent Reflecting Surface," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3947–3951, Dec. 2021.

[153] E. E. Haber, M. Elhattab, C. Assi, S. Sharafeddine, and K. K. Nguyen, "Latency and Reliability Aware Edge Computation Offloading in IRS-aided Networks," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 5035–5040.

[154] A. Al-Hilo, M. Samir, M. Elhattab, C. Assi, and S. Sharafeddine, "RIS-assisted UAV for timely data collection in IoT networks," *arXiv preprint arXiv:2103.17162*, 2021.

[155] A. Al-Hilo, M. Samir, M. Elhattab, C. Assi, and S. Sharafeddine, "Reconfigurable Intelligent Surface Enabled Vehicular Communication: Joint User Scheduling and Passive Beamforming," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2333–2345, Mar. 2022.

[156] M. Samir, M. Elhattab, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Optimizing Age of Information Through Aerial Reconfigurable Intelligent Surfaces: A Deep Reinforcement Learning Approach," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3978–3983, Apr. 2021.

[157] A. Muhammad, M. Elhattab, M. Shokry, and C. Assi, "Leveraging Reconfigurable Intelligent Surface to Minimize Age of Information in Wireless Networks," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 2525–2530.

[158] Y. Dai, J. Liu, M. Sheng, N. Cheng, and X. Shen, "Joint Optimization of BS Clustering and Power Control for NOMA-Enabled CoMP Transmission in Dense

Cellular Networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1924–1937, Feb. 2021.

[159] A. Rezaei, P. Azmi, N. M. Yamchi, M. R. Javan, and H. Yanikomeroglu, "Robust Resource Allocation for Cooperative MISO-NOMA-Based Heterogeneous Networks," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3864–3878, June 2021.

[160] M. Hua, Q. Wu, D. W. K. Ng, J. Zhao, and L. Yang, "Intelligent Reflecting Surface-Aided Joint Processing Coordinated Multipoint Transmission," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1650–1665, Mar. 2021.

[161] W. Wang, L. Yang, A. Meng, Y. Zhan, and D. W. K. Ng, "Resource Allocation for IRS-Aided JP-CoMP Downlink Cellular Networks With Underlaying D2D Communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4295–4309, Jun. 2022.

[162] J. Chen, Y. Xie, X. Mu, J. Jia, Y. Liu, and X. Wang, "Energy Efficient Resource Allocation for IRS Assisted CoMP Systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5688–5702, Jul. 2022.

[163] Y. Xiu, J. Zhao, W. Sun, M. D. 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. Wireless Commun.*, vol. 20, no. 12, pp. 8393–8409, Dec. 2021.

[164] Y. Li, M. Jiang, Q. Zhang, and J. Qin, "Joint Beamforming Design in Multi-Cluster MISO NOMA Reconfigurable Intelligent Surface-Aided Downlink Communication Networks," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 664–674, Jan. 2021.

[165] C. Zhang *et. al*, "Downlink analysis for reconfigurable intelligent surfaces aided NOMA networks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Dec. 2020, pp. 1–6.

[166] Y. Cheng, K. H. Li, Y. Liu, K. C. Teh, and H. Vincent Poor, "Downlink and Uplink Intelligent Reflecting Surface Aided Networks: NOMA and OMA," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3988–4000, June 2021.

[167] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Joint Deployment and Multiple Access Design for Intelligent Reflecting Surface Assisted Networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6648–6664, Oct. 2021.

[168] X. Xie, F. Fang, and Z. Ding, "Joint Optimization of Beamforming, Phase-Shifting and Power Allocation in a Multi-Cluster IRS-NOMA Network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7705–7717, Aug. 2021.

[169] J. Zuo, Y. Liu, E. Basar, and O. A. Dobre, "Intelligent Reflecting Surface Enhanced Millimeter-Wave NOMA Systems," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2632–2636, Nov. 2020.

[170] D. Zhang, Q. Wu, M. Cui, G. Zhang, and D. Niyato, "Throughput Maximization for IRS-Assisted Wireless Powered Hybrid NOMA and TDMA," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1944–1948, Oct. 2021.

[171] J. Zuo, Y. Liu, Z. Qin, and N. Al-Dhahir, "Resource Allocation in Intelligent Reflecting Surface Assisted NOMA Systems," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 7170–7183, Nov. 2020.

[172] Y. Dai and L. Lyu, "NOMA-Enabled CoMP Clustering and Power Control for Green Internet of Things Networks," *IEEE Access*, vol. 8, pp. 90 109–90 117, May 2020.

[173] Z. Ding, R. Schober, and H. V. Poor, "On the Impact of Phase Shifting Designs on IRS-NOMA," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1596–1600, Oct. 2020.

[174] S. Zhang and R. Zhang, "Intelligent Reflecting Surface Aided Multi-User Communication: Capacity Region and Deployment Strategy," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5790–5806, Sept. 2021.

[175] Z. Abdullah, G. Chen, S. Lambotharan, and J. A. Chambers, "A Hybrid Relay and Intelligent Reflecting Surface Network and Its Ergodic Performance Analysis," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1653–1657, Jun. 2020.

[176] F. J. Massey Jr, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[177] S. Atapattu, R. Fan, P. Dharmawansa, G. Wang, J. Evans, and T. A. Tsiftsis, "Reconfigurable intelligent surface assisted two–way communications: Performance analysis and optimization," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6552–6567, Oct. 2020.

[178] R. W. Heath Jr, T. Wu, Y. H. Kwon, and A. C. K. Soong, "Multiuser mimo in distributed antenna systems with out-of-cell interference," *IEEE Trans. Signal Processing*, vol. 59, no. 10, pp. 4885–4899, Oct. 2011.

[179] X. Yue and Y. Liu, "Performance Analysis of Intelligent Reflecting Surface Assisted NOMA Networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2623–2636, Apr. 2022.

[180] M. Elhattab, M. Kamel, and W. Hamouda, "Edge-aware remote radio heads cooperation for interference mitigation in heterogeneous C-RAN," *IEEE Trans. Veh. Technol.*, vol. 70, no. 11, pp. 12 142–12 157, Nov. 2021.

[181] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G Backhaul Challenges and Emerging Research Directions: A Survey," *IEEE Access*, vol. 4, pp. 1743–1766, Apr. 2016.

[182] Z.-Q. He and X. Yuan, "Cascaded Channel Estimation for Large Intelligent Metasurface Assisted Massive MIMO," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 210–214, Feb. 2020.

[183] A. Salem and L. Musavian, "NOMA in Cooperative Communication Systems With Energy-Harvesting Nodes and Wireless Secure Transmission," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1023–1037, Feb. 2021.

[184] T. Kim and M. Dong, "An Iterative Hungarian Method to Joint Relay Selection and Resource Allocation for D2D Communications," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 625–628, June 2014.

[185] M. K. Elhattab, M. M. Elmesalawy, and I. I. Ibrahim, "A Game Theoretic Framework for Device Association in Heterogeneous Cellular Networks With H2H/IoT Co-Existence," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 362–365, Feb. 2017.

[186] Z. Abdullah, G. Chen, S. Lambotharan, and J. A. Chambers, "Optimization of Intelligent Reflecting Surface Assisted Full-Duplex Relay Networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 363–367, Feb. 2021.

[187] S. Hua, Y. Zhou, K. Yang, Y. Shi, and K. Wang, "Reconfigurable Intelligent Surface for Green Edge Inference," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 2, pp. 964–979, Feb. 2021.

[188] X. Mu, Y. Liu, L. Guo, J. Lin, and H. V. Poor, "Intelligent reflecting surface enhanced multi-UAV NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3051–3066, Oct. 2021.

[189] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-Splitting Multiple Access: Fundamentals, Survey, and Future Research Trends," *IEEE Commun. Surveys Tuts.*, pp. 1–1, May 2022.

[190] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming SDMA and NOMA," *EURASIP journal on wireless communications and networking*, vol. 2018, no. 1, pp. 1–54, 2018.

[191] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-Splitting for Multi-Antenna Non-Orthogonal Unicast and Multicast Transmission: Spectral and Energy Efficiency Analysis," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8754–8770, Dec. 2019.

[192] S. Khisa, M. Almekhlafi, M. Elhattab, and C. Assi, "Full Duplex Cooperative Rate Splitting Multiple Access for a MISO Broadcast Channel With Two Users," *IEEE Commun. Lett.*, vol. 26, no. 8, pp. 1913–1917, Aug. 2022.

[193] Y. Mao, B. Clerckx, J. Zhang, V. O. K. Li, and M. A. Arafah, "Max-Min Fairness of K-User Cooperative Rate-Splitting in MISO Broadcast Channel With User Relaying," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6362–6376, Oct. 2020.

[194] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.