Cooperative Communications for 5G Wireless Networks and Beyond.

Phuc Huu Dinh

A Thesis

In the Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Master of Applied Science (Electrical and Computer Engineering) at

Concordia University

Montreal, Quebec, Canada

March 2020

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Phuc Huu Dinh

Entitled: **Cooperative Communications for 5G Wireless Networks and Beyond.**

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

Master of Applied Science (Electrical and Computer Engineering)

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

Signed by the Final Examining Committee:

_____ Chair
Dr. Jun Cai

_____ External Examiner
Dr. Emad Shihab

_____ Examiner
Dr. Jun Cai

_____ Supervisor
Dr. Chadi Assi

Approved by _____
Yousef R. Shayan, Chair
Department of Electrical and Computer Engineering

March 02, 2020

_____
Amir Asif, Dean
Faculty of Engineering and Computer Science

Abstract

Cooperative Communications for 5G Wireless Networks and Beyond.

Phuc Huu Dinh

Cooperative communication is an appealing technique stemming from the information-theoretic notion of cooperative diversity and having gradually evolved into a mainstream design paradigm in 4G LTE-Advanced. By exploiting cooperation among multiple transmission/reception nodes the technique reveals tremendous benefits, both in theoretical research and practical deployment, to enhance network performance in terms of throughput, reliability, latency, and network coverage. Due to these desirable attributes, it is expected that the technique will continue to be utilized in the coming generations of wireless networks. However, a major challenge to enable cooperative communication in the context of 5G and beyond is the advent of new wireless technologies and network architectures. The emergence of these technologies, on the one hand, enables a plethora of new applications. On the other hand, they entail new network operational constraints, which hinders the cooperation among network nodes. To address this challenge, it is crucial to study cooperative communication in specific network scenarios where these new technologies are employed to gain new insights their impact on in-network cooperation. Motivated by this, the thesis studies the role of cooperative communication in two futuristic network scenarios, encompassing two novel wireless technologies, namely non-orthogonal multiple access (NOMA) and unmanned aerial vehicles (UAVs).

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Dr. Chadi Assi, who has always been an inspiration for his knowledge and work ethics. The two years working with him to complete this thesis has been educational and fulfilling.

I would also like to thank Dr. Sanaa Sharafeddine, Dr. Ali Ghrayeb, and Dr. Wessam Ajib, and Dr. Tri Minh Nguyen, the mentors that I had the chance to work with during the making of this thesis.

In addition, I have been incredibly lucky to have the most amazing colleagues in the Lab. Their encouragement and advice have been indispensable at times of hardship and frustration.

I also greatly appreciate the opportunity to work with my senior and best friend, Mr. Mohamed Amine Arfaoui. His enthusiastic support is crucial to the making of this thesis.

Most importantly, I would like to thank my parents, who have always believed in me.

Contents

List of Figures

List of Abbreviations

BS	Base Station
CoMP	Coordinated multipoint
CB	Coordinated beamforing
C-NOMA	Cooperative Nonorthogornal
CS	Coordinated scheduling
DC	Difference-of-convex
FD	Full-duplex
HD	Half-duplex
i.i.d.	Independent and identically-distributed
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
NOMA	Nonorthogornal Multiple Access
pmf	Probability mass function
RF	Radio Frequency
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
UAV	Unmanned aerial vehicles
w.r.t.	With respect to
ZFBF	Zero-Forcing beamforming

Notations

s	Scalar value s
$\mathbf{x}$	Vector $\mathbf{x}$
$\mathbf{A}$	Matrix $\mathbf{A}$
$(.)^\intercal$	Transpose
$(.)^\dagger$	Hermitian transpose
$\lvert\,.\,\rvert$	Absolute value
$\lVert . \rVert$	Euclidean norm
$\lVert . \rVert_F$	Frobenius norm
$\lVert . \rVert_\infty$	Infinity norm
$\circ$	Hadamard product
Bdiag($\mathbf{A}_1$,....,$\mathbf{A}_n$)	Block diagonal matrix that connects matrices $\mathbf{A}_1$,....,$\mathbf{A}_n$
σ^2	Variance of normal distribution
$\mathcal{N}\left(0,\sigma^2\right)$	Normal distribution with zero mean and variance σ^2
$\mathbb{E}\left(\cdot\right)$	Expected value
$\mathbb{1}\left(\cdot\right)$	Indicator function

Chapter 1

Introduction

In this chapter, we provide the background knowledge of two core technologies that are used in this thesis, namely non-orthogonal multiple access and coordinated multipoint. Then, a summary of our contributions in this thesis is presented.

1.1 Non-orthogonal Multiple Access (NOMA)

1.1.1 Evolution of multiple access techniques

Since the first generation (1G) of wireless networks, engineers and scientists have been exploring novel techniques for users to effectively access the shared wireless channel. These techniques are collectively referred to as multiple (or channel) access methods. Throughout the evolution of wireless networks, each generation is usually associated with a signature multiple access technique as depicted in Fig. 1.1, from frequency division multiple access (FDMA) in 1G, time division multiple access (TDMA) in 2G, code division multiple access (CDMA) in 3G to orthogonal frequency-division multiple access (OFDMA). These multiple access methods are chosen to be compatible with the other underlying techniques and the specific requirements for each generation. However, a common trait of these pre-5G channel access is that they are all orthogonality-based. In other words, the users access the channel via orthogonal times/frequencies/codes to avoid multiple access interference. For this reason, this type of multiple access methods is generally called orthogonal multiple access (OMA) However, as future wireless networks are expected to enable a variety

of new applications, which can be categorized into enhanced Mobile Broadband (eMBBB), massive Machine Type Communications (mMTC), and Ultra-Reliable Low Latency Communications (URLLC) [1], the number of orthogonal resource blocks gradually become inadequate to accommodate a large number of users and applications, which inspired the advent of non-orthogonal multiple access (NOMA).

Generally, NOMA can be classified into code-domain NOMA and power-domain NOMA [2]. However, in the scope of this thesis, we limit of investigation to power-domain NOMA, which generally provides more tractable analytical framework, as a starting point to facilitate the incorporation of cooperative communications. The joint study of code-domain NOMA and cooperative communications should be interesting and is left for future works.

The operation of power-domain NOMA is illustrated in Fig. 1.2 for two users. Additionally, we illustrate OMA (FDMA) as a comparison. To enable transmission for two users in the same frequency (time), the BS applies superposition to multiplex the two transmitted signals with a certain power coefficient. Then, at the user end, the user with higher channel gain apply successive interference cancellation (SIC) to cancel the interfering signal of the other users and can detect its signal interference-free; on the other hand, the user with lower channel gain will treat the other user's signal as noise and decode its signal. Power-domain NOMA has been proven to significantly enhance network performance in terms of user fairness, coverage and capacity due to its ability to multiplex multiple transmitted signals into the same resource block and decode them at the receivers [1]

1.1.2 SINR analysis of multi-user NOMA networks

We hereby introduce a basic downlink analysis for a general multi-user NOMA-based communication system. We focus on downlink communications. Our objective is to derive the expression of SINR, which is one of the most important metrics in wireless communication systems. Note that some other important metrics, including achievable rate, outage probability, and coverage can be derived directly as functions of SINR. We consider downlink communications including one BS and a set $\mathcal{N}$ of users, indexed by $\{1,2...N\}$. Without loss of generality, we assume these users have descending channel gains, e.g, $|h_1|^2 \geq |h_2|^2 \geq ... \geq |h_N|^2$, where $h_k, \forall k \in \mathcal{N}$ is the channel

[1] Note that in the remaining of this thesis, the term "NOMA" implies power-domain NOMA for convenience.

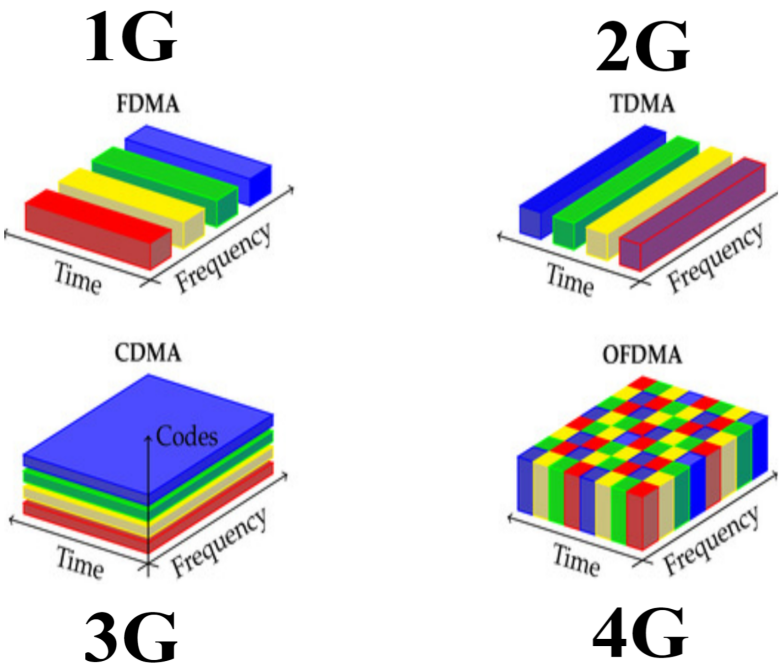


Figure 1.1: Illustration of OMA techniques from 1G to 4G.

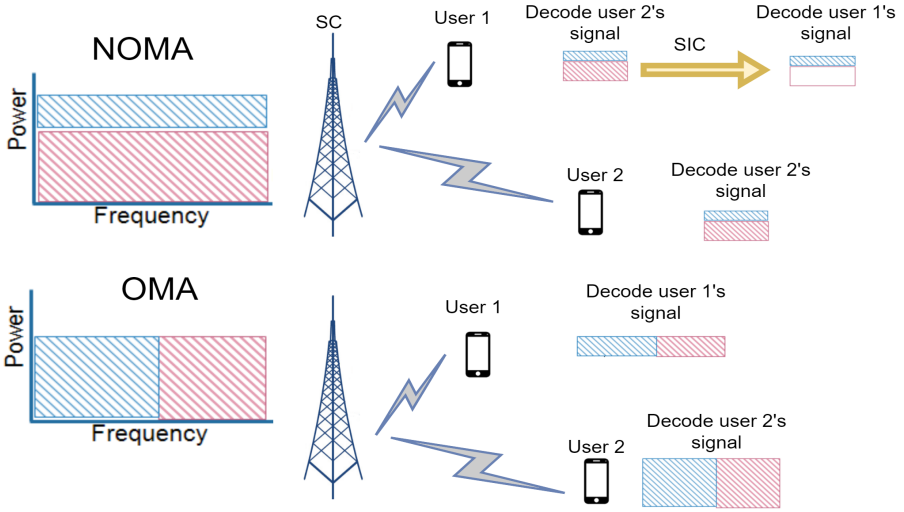


Figure 1.2: Illustration of 2-user power-domain NOMA as compared to OMA (FDMA).

coefficient of user k. As aforementioned at the BS, all the intended signals for the users are encoded into one super-positioned signal using SC. The encoded signal after SC can be written as follows

$$s = \sum_{i=1}^{N} \sqrt{P_i} s_i, \tag{1.1}$$

where P_i is the power allocated to user i. Note that in order to successfully decode the signals using SIC, the following condition on power allocation must hold.

$$P_1 \leq P_2 \leq ... \leq P_N, \tag{1.2}$$

In other word the allocated power has an inverse relationship to the channel gains. After broadcasting s, the received signal at users k can be written as follows

$$y_k = h_k \sum_{i=1}^{N} \sqrt{P_i} s_i + n_k, \tag{1.3}$$

Where n_k is AWGN (assuming to be zero-mean and unit-variance for simplicity). According to NOMA principle, users k decodes and cancels the signals of all the users that have lower channel gains while treating the signals of the other users as noise to decode its own message. After SIC is applied, the SINR for user k to decode its own message is as follows

$$SINR_k = \frac{P_k \mid h_k \mid^2}{\sum\limits_{i=1}^{k-1} P_i \mid h_k \mid^2 + 1} \tag{1.4}$$

and the SINR for user k to decode the intended message for user l ($l > k$) can be written as follows

$$SINR_{k \to l} = \frac{P_l \mid h_k \mid^2}{\sum\limits_{i=1}^{l-1} P_i \mid h_k \mid^2 + 1} \tag{1.5}$$

User clustering for NOMA: One important notice is that although the concept of NOMA can be applied to any number of users, it is a common practice to keep the number of non-orthogonal users small. The reason is that perfect SIC is not usually the case in practice. At each user, the impact of imperfect SIC to decode one user's message will propagate more errors for all the messages decoded for the other users that have higher channel gains, resulting in high bit error rates on all the users. This effect is called *error propagation* and is more severe as the number of users increases. Hence, it is more practical to assign a small number of users to a NOMA group and use OMA over different NOMA groups.

1.2 Coordinated multipoint (CoMP)

1.2.1 CoMP as a core technology in 4G

Coordinated multipoint (CoMP) is one of the widely used wireless technologies that leverage the concept of cooperative communication [3]. As opposed to SISO, where only one communication channel is established between a user and a base station or conventional MIMO, which is still susceptible to path loss, CoMP can offer resistance to both small-scale fading and path loss, as depicted in Fig. 1.3. The operating principle of CoMP is interference mitigation across multiple transmission/reception points. Furthermore, with proper cooperation among multiple transmission points, signals from different points may also be exploited as constructive signals to enhance reception diversity of cell-edge users. With these attributes, CoMP has been standardized by 3GPPP and adopted in 4G LTE Advanced. For 5G and future networks, the study of CoMP has become timely with the emergence of a variety of network architecture, including heterogeneous networks (HetNets) [4], cloud radio access networks (C-RAN) [5], or UAV-assisted wireless networks [6]. These types of networks have different topologies and operational constraints for which the application of conventional CoMP in LTE networks is not straightforward.

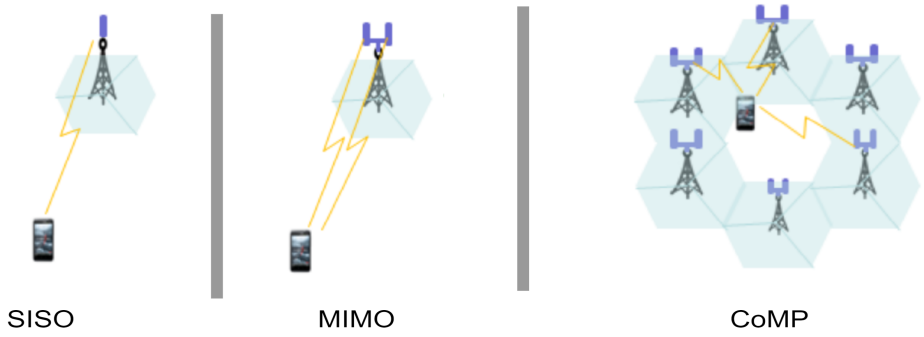


Figure 1.3: CoMP vs SISO vs MIMO

1.2.2 Different types of CoMP

Generally, there are three types of CoMP as illustrated in Fig. 1.4

Joint transmission (JT): the technique resorts to joint transmissions of multiple transmission points to replace interfering signals with meaningful signals. This technique is particularly useful

to improve received signals at cell-edge users. The technique requires the requested data to be available at all transmission points.

Coordinated scheduling/coordinated beamforming (CS/CB): the origin of CS/CB techniques can be traced back to the SINR level problem, whose objective is to maximize the minimum SINR. While CB determines the power or beamforming coefficients to achieve a higher common SINR, the CS approach clusters the transmission points and determine among that which transmission points serving which users.

Dynamic point transmission (DPS): the technique has the lowest complexity among all the other CoMP-based technique. The users in DPS may simply change its association to the access point based on its channel gain.

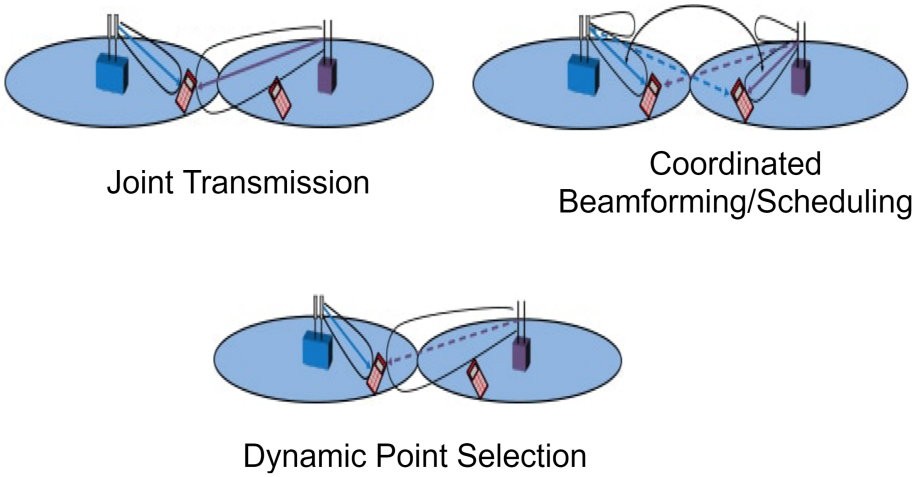


Figure 1.4: Different types of CoMP

1.3 Contributions

Our first major contribution is to study the performance of cooperative non-orthogonal multiple access (C-NOMA), a fusion of NOMA and cooperative communications in a cellular downlink system. The system model consists of a base station (BS) serving multiple users, where users with good channel quality can assist the transmissions between the BS and users with poor channel quality through either half-duplex (HD) or full-duplex (FD) device-to-device (D2D) communications.

We formulate and solve a novel optimization problem that jointly determines the optimal D2D user pairing and the optimal power control scheme, where the objective is maximizing the achievable sum rate of the whole system while guaranteeing a certain quality of service (QoS) for all users. The formulated problem is a mixed-integer non-linear program (MINLP) which is generally NP-hard. To overcome this issue, we reconstruct the original problem into a bi-level optimization problem that can be decomposed into two sub-problems to be solved independently. The outer problem is a linear assignment problem which can be efficiently handled by the well-known Hungarian method. The inner problem is still a non-convex optimization problem for which finding the optimal solution is challenging. However, we derive the optimal power control policies for both the HD and the FD schemes in closed-form expressions, which makes the computational complexity of the inner problems polynomial for every possible pairing configurations. These findings solve ultimately the original MILNP in a timely manner that makes it suitable for real-time and low latency applications. Our simulation results show that the proposed framework outperforms a variety of proposed schemes in the literature and that it can obtain the optimal pairing and power control policies for a network with 100 users in a negligible computational time.

Our second major contribution is that we investigate cooperative downlink transmissions in a wireless communication system enabled by a swarm of unmanned aerial vehicles (UAVs) which are spatially dispatched to cooperatively deliver requested contents to ground users. First, we propose a communication scheme that exploits the flexible deployment of UAVs as well as their cooperative transmissions to improve in-network user admission. Unlike previous literature, a practical operational constraint of limited storage capacity for UAVs is considered. Then, from the knowledge that cooperation among UAVs depends on the availability of the contents in their limited storage space, we propose a novel joint optimization problem to determine the content placement, location planning, user admission decision and transmit beamforming to maximize the number of users experiencing a minimum required rate, so-called admitted users. Since the formulated problem is a mixed-integer non-linear program which is generally non-deterministic polynomial-time hard, we proposed a framework that is developed on the basis of difference-of-convex (DC) programming to transform the original problem into a series of approximate convex problems which can be iteratively solved until convergence. Our extensive simulation results reveal that the proposed scheme

outperforms other schemes that have been introduced in previous work and reflect a notable trend that deploying more cooperative UAVs with fewer resources (power and storage capacity) is more efficient than deploying fewer UAVs with more resources. In particular, in one of our collected results, the total communication power can be reduced by roughly 40 dB when doubling the number of cooperative UAVs.

Chapter 2

A Low-Complexity Framework for Joint User Pairing and Power Control for Cooperative NOMA in 5G and Beyond Cellular Networks

2.1 Introduction

2.1.1 Motivation

The total data traffic is expected to become about 49 exabytes per month by 2021, while in 2016, it was approximately 7.24 exabytes per month [7]. With this dramatic increase, fifth-generation (5G) networks and beyond must urgently provide high data rates, seamless connectivity and ultra-low latency communications [8–10]. In addition, with the emergence of the Internet-of-things (IoT) networks, the number of connected devices to the internet is increasing exponentially [11,12]. This fact implies not only a significant increase in data traffic, but also the emergence of some IoT services with crucial requirements. Such requirements include higher data rates, higher connection density and ultra reliable low latency communication (URLLC). Hence, traditional radio-frequency (RF)

networks, which are already crowded, are unable to satisfy these high demands [13]. Network densification [14, 15] has been proposed as a solution to increase the capacity and coverage of 5G networks. However, with the continuous dramatic growth in data traffic, researchers from both industry and academia are trying to explore new network architectures, new transmission techniques and new communication technologies to meet these demands. Among the new communication technologies that have been proposed as auspicious solutions for 5G and beyond are non-orthogonal multiple access (NOMA), device-to-device (D2D) communication, half duplex (HD) and full-duplex (FD) communications.

NOMA is capable of supporting more users than the number of available orthogonal resources [2], thereby leading to higher spectral efficiency and user fairness when compared to standard orthogonal multiple access (OMA) techniques.[1] The principle of NOMA leverages the concept of superposition coding (SPC) at the transmitter, to multiplex users in power-domain, and successive interference cancellation (SIC) at the receiver [16]. However, as a standalone technique, NOMA still cannot adequately fulfil the demanding specifications of 5G networks and beyond. In fact, an inherent limitation of NOMA lies in the requirement that the allocated power to a user with poor channel conditions needs to be high for successful decoding of the superimposed signal. This requirement generally reduces the spectral (and power) efficiency since the poor channel will absorb a large portion of the available power budget. Thus, recent research trends aim to modularize and integrate NOMA with other advanced and 5G-enabled techniques, such as multiple-input-multiple-output (MIMO) [17], hybrid automatic repeat request (HARQ) [18], and device-to-device (D2D) communications [19].

A recent appealing extension for NOMA is cooperative NOMA (C-NOMA), taking advantage of desirable attributes of NOMA, FD/HD, and D2D communications. For a practical scenario, each user with a high channel gain, which is referenced as strong user, can act as relay to assist communications between the transmitter and a user with a poor channel gain, which is referenced as weak user. Each weak user can then combine both signals coming from the transmitter and from the associated strong user. The novelty of C-NOMA revolves around adding the D2D communication

[1] In this chapter, the term "NOMA" is restricted to power-domain NOMA as distinct from its code-domain NOMA counterpart.

factor as a new degree of diversity to the direct downlink transmission. Intuitively speaking, each weak user can receive its message either from the BS or from another user, depending on which one is more favorable.

2.1.2 Literature Review

The performance gain of NOMA comes at the expense of design complexity. Therefore, extensive performance analysis and efficient performance optimization scheme under a wide range of network scenarios have been studied in [20–29]. In [20], the authors presented the basis of NOMA and the relations of NOMA to a few network paradigms such as MIMO or cognitive radio (CV). As a major aspect of NOMA literature, resource allocation for NOMA-based networks are investigated in [21–25]. To mitigate the impact of error propagation due to imperfect SIC and to reduce the overall complexity, pairing (and clustering) is essential for networks that have large number of users. Motivated by this, optimal pairing policy for NOMA has been studied in [26–28]. User pairing does not only offer enhanced scalability and modularity for system optimization but also entails potential usage of D2D communications among users to further enhance system throughput, energy efficiency and fairness. Optimal pairing scheme for sum-rate maximization for multi-user NOMA networks has been proposed in [30]. Thus, the literature on NOMA is considered rich, yet NOMA still cannot adequately fulfil the demanding specifications of 5G networks and beyond in terms of high data rates and massive connectivity.

By combining NOMA with HD/FD and D2D communications, C-NOMA takes the advantage of the SIC process combined with decode-and-forward procedure at users with high channel conditions to increase the reception diversity at the users that experience severe channel fading, and hence increasing the total throughput [31]. As opposed to NOMA, the research on C-NOMA is still far from being mature. One observation is that both the pairing policy and the power allocation scheme of C-NOMA systems are not a direct inheritance of those of conventional NOMA. One of the key challenges of investigating C-NOMA lies in the complicated achievable rate expressions which capture the characteristics of NOMA as well as the HD/FD decode-and-forward procedures. Specifically, the achievable rate of a given user within a C-NOMA system is a non-concave and non-differentiable function, which hinders the direct application of derivative-based numerical

optimization frameworks [32]. In [33–37], the performance of C-NOMA, measured in terms of outage probability, error performance and capacity, were investigated. However, these works focused mainly on two-user setting for simplicity and tractability purposes. In addition, the adopted performance metric is the max-min achievable rate, which is known to be bandwidth-inefficient since the majority of resource is allocated to the user with poor channel gain to maintain fairness. A limited amount of works in the literature have formally studied the performance of multi-user C-NOMA [38–41]. In [38, 39], the authors studied the performance of C-NOMA in terms of outage performance. However, these works have either employed random or heuristic distance-based pairing, which limits the performance gains of C-NOMA. Additionally, in [39], simultaneous wireless information and power transfer was integrated with a C-NOMA system to reduce the energy consumption at the users. However, the pairing policy and power allocation problem were not investigated. Recently, Obeed *et al*, [41] studied the joint problem of user pairing and power allocation with HD links and a fixed relaying power for hybrid radio-frequency (RF) and visible light communication (VLC) systems. However, assuming a fixed relaying power is not a practical assumption for realistic scenarios. In fact, assuming that the strong user uses the maximum allowed relaying power is technically unappealing, since power at user devices is in general small, the energy efficient of C-NOMA schemes should be considered carefully instead of assuming fixed powers at the relaying devices.

2.1.3 Contributions

We investigate in this chapter the performance of C-NOMA in a cellular downlink system that consists of a BS that wants to serve multiple users within a region of service. In this system, users that have the capability of either HD or FD communications can assist the transmissions between the BS and users with poor channel quality through D2D communications. We formulate and solve an optimization problem that jointly determines the optimal D2D user pairing policy, the optimal power control scheme, and the best D2D communication mode, i.e., HD or FD, where the objective is maximizing the achievable sum rate of the whole system while guaranteeing a certain quality of service (QoS) for all users.

In light of the above background, and to the best of our knowledge, this problem has not been

investigated in the literature. Our contributions can be summarized as follows

We propose a modular formulation for that can be applied for all possible relaying schemes, such as HD, FD or hybrid HD/FD for C-NOMA system. The formulated problem ends up to be a mixed-integer non-convex program, which involves not only binary decision variables but also non-smooth utility functions. Thus, derivative-based convexification methods can not be directly applicable [42].[2] Alternatively, we rewrite the problem into a bi-level optimization model, consisting of one outer problem and one inner problem.[3].

- The formulated problem ends up to be a mixed-integer non-convex program, which involves not only binary decision variables but also non-smooth utility functions. Thus, derivative-based convexification methods can not be directly applicable [42].[4] Alternatively, we rewrite the problem into a bi-level optimization model, consisting of one outer problem and one inner problem.

- To propose an efficient framework, we observe that the outer problem is a classic linear assignment problem. Thus, standard matching algorithms such as the Hungarian method can be applied. For the inner problem, we derive the feasibility conditions of the formulated problem as relations between the power resources, the QoS requirements, and the channel state information (CSIs). Then, we derive the closed-form solutions of the power control scheme for both the HD and FD cases. These closed-form solutions, with computational complexity $\mathcal{O}(1)$, facilitates the application of the Hungarian method, and thus, the whole problem can be solved in a polynomial time.

The rest of the chapter is organized as follows. Section 2.2 presents the system model. Section 2.3 presents the achievable rates analysis. Section 2.4 presents the optimal pairing policy and power control scheme. Sections 2.5 and 2.6 presents the simulation results and the conclusion, respectively.

[2]Note that most previous works considering this class of problems typically assume overprovisioning of resources, i.e, high power and low minimum-rate requirement, to guarantee the feasibility to apply proposed algorithms.

[3]It is important to clarify that the proposed bilevel optimization method is fundamentally different from BCD method proposed by [41] although both involve decomposition of the original problem into subproblems. Besides, [41] does not consider adaptive D2D relaying.

[4]Note that most previous works considering this class of problems typically assume overprovisioning of resources, i.e, high power and low minimum-rate requirement, to guarantee the feasibility to apply proposed algorithms.

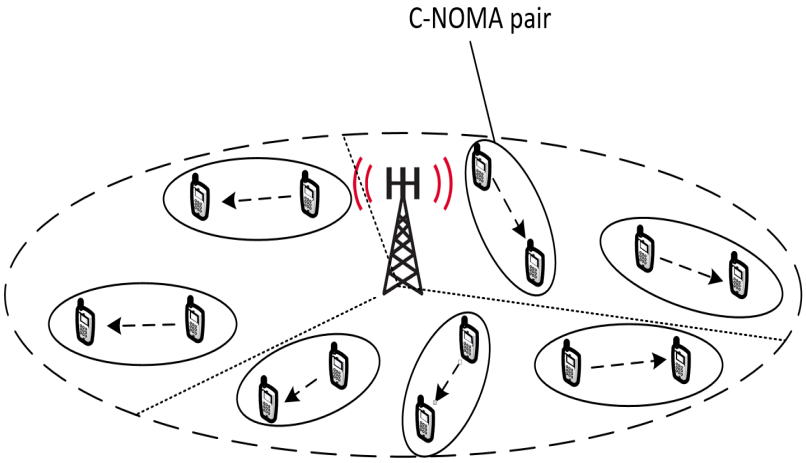


Figure 2.1: Proposed downlink C-NOMA for a cellular network. The network is divided into $M = 3$ sectors that are served via orthogonal channels.

2.2 System Model

2.2.1 Network Model

We consider a standard single-cell cellular system that consists of a BS equipped with M antennas and serving M disjoint sectors, where each sector is served with one antenna as shown in Fig. 2.1. In this chapter, we investigate the performance of one typical sector within which we assume that there exist $2K$ spatially dispersed active users (AUs). It is important to highlight here that considering an odd number of AUs does not affect the generality of the system model, since if the number of AUs is even, we still can adopt the same network model by adding an extra virtual AU which has zero channel gain. The AUs can be divided based on their channel gains into two disjoint sets of users, which we denote by $\mathscr{S}$ and $\mathscr{W}$. The set $\mathscr{S}$ contains the AUs that have high channel gain and such users are referred to as "strong AUs". On the other hand, the set $\mathscr{W}$ contains the users with low channel gains and these users are referred to as "weak AUs". According to C-NOMA principle [31], each weak AU from $\mathscr{W}$ is paired with exactly one strong AU from $\mathscr{S}$ and the resulting distinct pairs of (strong AU, weak AU) are served simultaneously over orthogonal and equally divided frequency subchannels in order to cancel the inter-pair interference. Moreover, the AUs within each pair are served using NOMA principle, where additionally, the strong AU can assist the communication between the BS and the weak AU. Obviously, the system performance depends

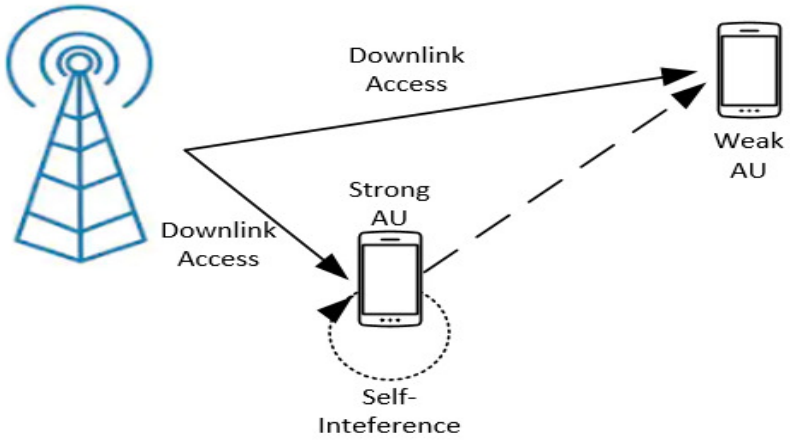


Figure 2.2: Illustration of the transmission phases within a FD C-NOMA pair.

on the pairing configurations and the communication links within each pair, which is the focus of this chapter. In the following subsection, we investigate the transmission model within each pair of AUs.

2.2.2 Transmission Model

According to C-NOMA principle, and as presented in Fig. 2.2, the transmission model within each pair of AUs consists of two phases that are detailed as follows.

(1) The NOMA downlink transmission phase: the BS applies superposition coding (SC) on the messages intended to the strong and weak AUs and transmit the superimposed message to both of them [43]. Then, following NOMA principle, the weak AU treats the interference from the strong AUs as noise and decode its own message [43].

(2) The D2D cooperative relaying phase: the strong AU first applies SIC to decode the intended message for weak AU. Second, it subtracts the decoded message of the weak AU from its own reception. Then, it decodes its own message from its resulting interference-free reception [16]. Finally, it relays this decoded message for the weak AU through a D2D channel, therefore enhancing the signal reception diversity [44].

For the case of HD communication, the two transmission phases occur on consecutive resource blocks. However, for the case of FD communication, they occur on the same resource block, which

comes with the cost of inducing SI at the strong AU (the dotted line in Fig. 2.2).

2.3 Achievable Rate Analysis

In this section, the analysis of the achievable rates for a downlink C-NOMA system for both HD and FD cases are presented. Consider a pair of AUs $(m,n) \in [\![1,K]\!]^2$, where m and n denote the indices of the strong and the weak AU, respectively, within the pair. For both cases of HD and FD relaying and at each channel use, the resulting signal at the BS after applying SC is expressed as

$$s = \sqrt{\alpha_{m,n} P_{\mathrm{BS}}}\, s_n + \sqrt{(1-\alpha_{m,n})\, P_{\mathrm{BS}}}\, s_m, \tag{2.1}$$

where s_m and s_n represent the intended messages of the strong AU and the weak AU, respectively, such that $\mathbb{E}(|s_n|^2) = \mathbb{E}(|s_m|^2) = 1$, P_{BS} represents the total average power available at the BS and $\alpha_{m,n} \in [0,1]$ is the power allocation factor, i.e., $(1-\alpha_{m,n})\, P_{\mathrm{BS}}$ and $\alpha_{m,n} P_{\mathrm{BS}}$ represent the transmit powers allocated to the strong AU and the weak AU, respectively.

2.3.1 HD C-NOMA

For the case of HD C-NOMA, the received signal at the strong AU at each channel use is given by

$$y_m = h_m \left(\sqrt{\alpha_{m,n} P_{\mathrm{BS}}}\, s_n + \sqrt{(1-\alpha_{m,n})\, P_{\mathrm{BS}}}\, s_m \right) + \omega_m, \tag{2.2}$$

where h_m is the channel gain from the BS to the strong AU and ω_m is a zero-mean unit-variance additive Gaussian noise (AWGN). At the strong AU, SIC is applied to decode the message s_n of the weak AU. Therefore, the achievable rate of the strong AU to decode the message intended for the weak AU can be expressed as

$$R^{\mathrm{H}}_{m,n} = \frac{1}{2} \log_2 \left(1 + \frac{\alpha_{m,n} P_{\mathrm{BS}} \gamma_m}{(1-\alpha_{m,n}) P_{\mathrm{BS}} \gamma_m + 1} \right), \tag{2.3}$$

where $\gamma_m = h_m^2$. After subtracting the message s_n from its reception, the strong AU can decode its own message with an achievable rate that is given by

$$R_m^{\mathrm{H}} = \frac{1}{2}\log_2\left(1 + (1-\alpha_{m,n})P_{\mathrm{BS}}\gamma_m\right). \tag{2.4}$$

On the other hand, the received signal at the weak AU from the BS in the direct transmission phase is expressed as

$$y_n = h_n\left(\sqrt{\alpha_{m,n}P_{\mathrm{BS}}}s_n + \sqrt{(1-\alpha_{m,n})P_{\mathrm{BS}}}s_m\right) + \omega_n, \tag{2.5}$$

where h_n is the channel gain from the BS to the weak AU and ω_n is a zero-mean unit-variance AWGN. In the cooperative relaying phase, the strong AU forwards the message s_n to the weak AU. Thus, the received signal at the weak AU in this phase is given by

$$y_n = h_{m,n}^{\mathrm{d}}\sqrt{P_{m,n}^{\mathrm{d}}}s_n + \omega_n, \tag{2.6}$$

where $h_{m,n}^{\mathrm{d}}$ is the channel gain from the strong AU to the weak AU and $P_{m,n}^{\mathrm{d}}$ is the D2D transmit relaying power. Since the weak AU receives duplicate messages from both the BS and the strong AU, repetition decoding (RD) can be applied to decode the replicated information [45]. In other words, the effective signal-to-interference-plus-noise-ratio (SINR) at the weak AU is the summation of the SINRs of the transmission links from the BS to the weak AU and from the strong AU to the weak AU. In this case, the achievable rate of the weak AU is expressed as

$$R_n^{\mathrm{RD}} = \frac{1}{2}\log\left(1 + P_{m,n}^{\mathrm{d}}\gamma_{m,n}^{\mathrm{d}} + \frac{\alpha_{m,n}P_{\mathrm{BS}}\gamma_n}{(1-\alpha_{m,n})P_{\mathrm{BS}}\gamma_n + 1}\right), \tag{2.7}$$

where $\gamma_{m,n}^{\mathrm{d}} = \left(h_{m,n}^{\mathrm{d}}\right)^2$ and $\gamma_n = h_n^2$. Finally, in line with the results of [46], the resulting achievable rate of the weak AU from the cooperative diversity of the BS and the strong AU is expressed as

$$R_n^{\mathrm{H}} = \min\left(R_n^{\mathrm{RD}}, R_{m,n}^{\mathrm{H}}\right). \tag{2.8}$$

2.3.2 FD C-NOMA

For the case of FD C-NOMA, the received signal at the strong AU is given by

$$y_m = h_m \left(\sqrt{\alpha_{m,n} P_{\mathrm{BS}}} s_n + \sqrt{(1-\alpha_{m,n}) P_{\mathrm{BS}}} s_m \right) + h_m^{\mathrm{SI}} \sqrt{P_{m,n}^{\mathrm{d}}} s_n + \omega_m, \tag{2.9}$$

where h_m^{SI} represents the SI channel gain at the strong AU (the dotted line in Fig. 2.2). Thus, the achievable rate of the strong AU to decode the message of the weak AU is expressed as

$$R_{m,n}^{\mathrm{F}} = \log_2 \left(1 + \frac{\alpha_{m,n} P_{\mathrm{BS}} \gamma_m}{(1-\alpha_{m,n}) P_{\mathrm{BS}} \gamma_m + P_{m,n}^{\mathrm{d}} \gamma_m^{\mathrm{SI}} + 1} \right), \tag{2.10}$$

where $\gamma_m^{\mathrm{SI}} = \left(h_m^{\mathrm{SI}}\right)^2$. Then, after successfully decoding and canceling the message s_n of the weak AU, the strong AU decodes his own message s_m. Therefore, the achievable rate of the strong user to decode its own message is expressed as

$$R_m^{\mathrm{F}} = \log_2 \left(1 + \frac{(1-\alpha_{m,n}) P_{\mathrm{BS}} \gamma_m}{P_{m,n}^{\mathrm{d}} \gamma_m^{\mathrm{SI}} + 1} \right). \tag{2.11}$$

Afterwards, the strong AU forwards the message intended for the weak AU. Assuming that the processing delay caused by the SIC process at the strong AU is small, the weak AU receives the message s_n from the BS and from the strong user at approximately the same channel use [47]. Therefore, the total received signal at the weak AU is given by

$$y_n = h_n \left(\sqrt{\alpha_{m,n} P_{\mathrm{BS}}} s_n + \sqrt{(1-\alpha_{m,n}) P_{\mathrm{BS}}} s_m \right) + h_{m,n}^{\mathrm{d}} \sqrt{P_{m,n}^{\mathrm{d}}} s_{\mathrm{w}} + \omega_n, \tag{2.12}$$

Following [48], we assume that the weak AU can successfully co-phase and combine the signals from the BS and the strong AU by a proper diversity combining technique such as the maximum ratio combining (MRC). In this case, the effective SINR of the weak AU is the summation of the SINRs resulting from decoding its received message from the BS and the strong AU. Consequently,

the achievable rate of the weak AU when applying MRC can be written as

$$R_n^{\mathrm{MRC}} = \log_2\left(1 + P_{m,n}^{\mathrm{d}}\gamma_{m,n}^{\mathrm{d}} + \frac{\alpha_{m,n}P_{\mathrm{BS}}\gamma_n}{(1-\alpha_{m,n})P_{\mathrm{BS}}\gamma_n + 1}\right), \tag{2.13}$$

Based on the above analysis and the results of [49], the resulting achievable rate of the weak AU is given by

$$R_n^{\mathrm{F}} = \min\left(R_{m,n}^{\mathrm{F}},\, R_n^{\mathrm{MRC}}\right). \tag{2.14}$$

2.4 Optimal Pairing Policy and Power Control Scheme

2.4.1 Problem Statement

The objective of this chapter is maximizing the sum rate of a downlink C-NOMA cellular sector consisting of a set of $2K$ AUs, while guaranteeing a certain QoS for each user. The maximization is performed with respect to the power allocation at the BS and the D2D links, the user pairing policy and the D2D transmission mode, i.e., HD or FD. This objective is expressed in a formal optimization problem in the following subsection.

2.4.2 Problem Formulation

Assuming that each sector consists of $2K$ AUs, the number of possible pairing configurations is $(2K-1)!! = (2K-1)\times(2K-3)\times(2K-5)\times\cdots\times 1$. Therefore, exhaustively trying all possible configurations is not practically appealing. To overcome this issue, and in order to provide a scalable solution, we instead reduce the pairing policy into a linear assignment problem, where each AU from the group of strong AUs $\mathscr{S}$ is paired with one AU from the group of weak AUs $\mathscr{W}$. Consequently, the joint pairing and power control for the sum-rate maximization problem of the overall downlink C-NOMA system can be given by the following optimization problem.

$$\mathscr{P}: \quad R^* = \max_{\mathbf{B},\alpha,\mathbf{P}^{\mathrm{d}}} \sum_{m=1}^{K}\sum_{n=1}^{K} b_{m,n}R_{m,n}\left(\alpha_{m,n}, P_{m,n}^{\mathrm{d}}\right), \tag{2.15a}$$

$$\text{s.t. } 0 \le \alpha_{m,n} \le b_{m,n},\ \forall\, m,n \in [\![1,K]\!], \tag{2.15b}$$

$$0 \le P_{m,n}^{\mathrm{d}} \le b_{m,n}\bar{P}^{\mathrm{d}},\ \forall\, m,n \in [\![1,K]\!], \tag{2.15c}$$

$$b_{m,n} \in \{0,1\},\ \forall\, m,n \in [\![1,K]\!], \tag{2.15d}$$

$$\sum_{m=1}^{K} b_{m,n} = 1,\ \forall\, m,n \in [\![1,K]\!], \tag{2.15e}$$

$$\sum_{n=1}^{K} b_{m,n} = 1,\ \forall\, m,n \in [\![1,K]\!], \tag{2.15f}$$

$$b_{m,n} R_m(\alpha_{m,n}, P^{\mathrm{d}}_{m,n}) \geq R_{\mathrm{th}},\ \forall\, m,n \in [\![1,K]\!], \tag{2.15g}$$

$$b_{m,n} R_n(\alpha_{m,n}, P^{\mathrm{d}}_{m,n}) \geq R_{\mathrm{th}},\ \forall\, m,n \in [\![1,K]\!], \tag{2.15h}$$

where $R_{m,n}$ represents the sum rate of the pair (m,n) that is expressed as

$$R_{m,n} = R_m + R_n, \tag{2.16}$$

such that the rate functions $(R_m, R_n) = \left(R^{\mathrm{H}}_m, R^{\mathrm{H}}_n\right)$ for the case of HD C-NOMA and $(R_m, R_n) = \left(R^{\mathrm{F}}_m, R^{\mathrm{F}}_n\right)$ for the case of FD C-NOMA, $b_{m,n}$ is the pairing decision variables, where $b_{m,n} = 1$ indicates that AU m in set $\mathscr{S}$ is paired with AU n in set $\mathscr{W}$ and $b_{m,n} = 0$, otherwise, $\mathbf{B} = \{b_{m,n} | m,n \in [\![1,K]\!]\}$, $\alpha_{m,n}$ is the "potential power allocation coefficient" of the pair (m,n), $\alpha = \{\alpha_{m,n} | m,n \in [\![1,K]\!]\}$, $P^{\mathrm{d}}_{m,n}$ denotes the "potential D2D transmit power" within the pair (m,n), $\mathbf{P}^{\mathrm{d}} = \left\{P^{\mathrm{d}}_{m,n}, | m,n \in [\![1,K]\!]\right\}$ and $\bar{P}^{\mathrm{d}}$ denote the D2D power budget at each AU's device. Constraints (2.15b) and (2.15c) ensure that $\alpha^{(\mathrm{m,n})} \in [0,1]$ and $P^{\mathrm{d}}_{m,n} \in \left[0, \bar{P}^{\mathrm{d}}\right]$, respectively, when AU m is paired with AU n. Constraints (2.15e) and (2.15f) ensure that each AU from each group can be paired with only one AU from the other group. Constraints (2.15g)-(2.15h) ensure that the paired AUs have each an achievable rate greater than a minimum achievable rate R_{th}, whichs guarantee the QoS constraint. Note that, in our particular problem, we assume that the minimum rate achievable rate is set to be high enough so that the optimal power control always lies in the SIC-stable region.

Problem $\mathscr{P}$ is a mixed-integer non-linear program (MINLP), which is generally NP-hard. Most of the previous literature resorts to iterative numerical methods such derivative-based methods or block coordinate descent (BCD) and/or off-the-shelf optimization solvers to solve this class of problems. However, BCD, which was previously adopted in [41] is known to have poor convergence properties as shown in [50] and the derivative-based methods, such as successive convex approximation (SCA), cannot be directly applied here due to the binary pairing decision variable and the

non-smooth rate functions. To overcome this issue, we solve problem $\mathscr{P}$ using the concept of bilevel optimization [51].

2.4.3 Bi-level Optimization

In problem $\mathscr{P}$, it can be observed that the rate value at the optimal power control can be computed for all different pairing configurations. Precisely, let $\left\{ \left(b^*_{m,n}, \alpha^*_{m,n}, P^{\mathrm{d}^*}_{m,n}\right) \middle| m,n \in [\![1,K]\!] \right\}$ denotes the set of optimal user pairing policy and power control scheme, which are the solutions of problem $\mathscr{P}$. For all $m,n \in [\![1,K]\!]$ if $b^*_{m,n} = 0$, then $\left(\alpha^*_{m,n}, P^{\mathrm{d}^*}_{m,n}\right) = (0,0)$. However, if $b^*_{m,n} = 1$, then $\left(\alpha^*_{m,n}, P^{\mathrm{d}^*}_{m,n}\right)$ should be the optimal solutions of the power control scheme of the pair of users (m,n). In other words, For all $m,n \in [\![1,K]\!]$, if assume that the the users (m,n) are paired together and that we can obtain their optimal power control scheme $\left(\alpha^*_{m,n}, P^{\mathrm{d}^*}_{m,n}\right)$, problem $\mathscr{P}$ becomes a linear assignment problem and it remains to determine the optimal pairing policy $\left(b^*_{m,n}\right)_{1\leq m,n\leq K}$. Hence, since we aim to find the optimal power control that maximizes the total achievable sum rate of the C-NOMA system, we can reduce the feasible set of power control for problem $\mathscr{P}$ to the set of power control that maximizes the sum rate within a pair of users. Consequently, based on this observation, we rewrite the original problem in to a bi-level optimization [52] problem as follows.

$$\mathscr{P}_{\text{outer}} : R^* = \max_{\mathbf{B}} \sum_{n=1}^{K} \sum_{m=1}^{K} b_{m,n} R_{m,n}(\alpha^*_{m,n}, P^{\mathrm{d}^*}_{m,n}), \tag{2.17a}$$

$$\text{s.t.} \quad (2.15\text{d}) - (2.15\text{f}) \tag{2.17b}$$

where $\alpha^*_{m,n}$ and $P^{\mathrm{d}^*}_{m,n}$ are parameters obtained by solving the following problem

$$\mathscr{P}_{\text{inner}} : \quad R^*_{m,n} = \max_{\alpha_{m,n}, P^{\mathrm{d}}_{m,n}} R_{m,n}(\alpha_{m,n}, P^{\mathrm{d}}_{m,n}), \tag{2.18a}$$

$$\text{s.t. } 0 \leq \alpha_{m,n} \leq 1, \ \forall\, m,n \in [\![1,K]\!], \tag{2.18b}$$

$$0 \leq P^{\mathrm{d}}_{m,n} \leq \bar{P}^{\mathrm{d}}, \ \forall\, m,n \in [\![1,K]\!], \tag{2.18c}$$

$$R_m\left(\alpha_{m,n}, P^{\mathrm{d}}_{m,n}\right) \geq R_{\text{th}}, \ \forall\, m,n \in [\![1,K]\!], \tag{2.18d}$$

$$R_n\left(\alpha_{m,n}, P^{\mathrm{d}}_{m,n}\right) \geq R_{\text{th}}, \ \forall\, m,n \in [\![1,K]\!], \tag{2.18e}$$

for each pair of AUs $(m,n) \in [\![1,K]\!]^2$. The inner problem $\mathscr{P}_{\text{inner}}$ is a power control problem for a given pair of AUs and it defines a new set of feasible solutions for the outer problem $\mathscr{P}_{\text{outer}}$, which is a linear assignment problem. Nevertheless, since we need to solve problem $\mathscr{P}_{\text{inner}}$ for all possible combinations of strong and weak AUs, a computational efficient approach for solving $\mathscr{P}_{\text{inner}}$ needs to be investigated, which is the focus of the following section.

2.4.4 Power Control For a C-NOMA Pair

In this section, our objective is solving problem $\mathscr{P}_{\text{inner}}$ for both HD and FD C-NOMA cases. We consider first the HD C-NOMA case and then we investigate the case of FD C-NOMA.

HD C-NOMA

In this part, we consider the case of HD C-NOMA and hence the sum-rate $R_{m,n}\left(\alpha_{m,n}, P^{\text{d}}_{m,n}\right) = R^{\text{H}}_{m}\left(\alpha_{m,n}, P^{\text{d}}_{m,n}\right) + R^{\text{H}}_{n}\left(\alpha_{m,n}, P^{\text{d}}_{m,n}\right)$. Let us investigate first the feasibility conditions of the inner problem $\mathscr{P}_{\text{inner}}$. These conditions are detailed in the following theorem.

***Theorem* 1.** *For the case of HD C-NOMA, the inner problem $\mathscr{P}_{\text{inner}}$ is feasible if and only if the following conditions hold.*

$$P_{\text{BS}} \geq \frac{(\delta^{\text{H}}_{\text{th}})^2 + \delta^{\text{H}}_{\text{th}}}{\gamma_m}, \tag{2.19a}$$

$$\bar{P}^{\text{d}} \geq P^{\min}_{m,n}. \tag{2.19b}$$

where $\delta^{\text{H}}_{\text{th}} = 2^{2R_{\text{th}}} - 1$ *and*

$$P^{\min}_{m,n} = \frac{((\delta^{\text{H}}_{\text{th}})^2 + \delta^{\text{H}}_{\text{th}}(\gamma_m + P_{\text{BS}}\gamma_n) - P_{\text{BS}}\gamma_n\gamma_m}{\gamma^{\text{d}}_{m,n}(\delta^{\text{H}}_{\text{th}}\gamma_n + \gamma_m)}. \tag{2.20}$$

Proof. Please see Appendix A □

The conditions (2.19a) and (2.19b) are necessary to ensure that there exists a power control decision for the inner problem $\mathscr{P}_{\text{inner}}$ in case of HD C-NOMA. In addition, these conditions describe the relations between the available power budgets at the BS and at the strong user, the channel gains

of both users and the minimum QoS requirements that need to be satisfied in order to make the inner problem $\mathscr{P}_{\text{inner}}$ feasible. However, in the case where at least one of the conditions can not be fulfilled, the two possible options to guarantee the required QoS for both users is either reducing the minimum required rate R_{th} or increasing the power budgets P_{BS} and $\bar{P}^{\text{d}}$. Now, assuming that the inner problem $\mathscr{P}_{\text{inner}}$ is feasible, the optimal power control scheme is presented in the following theorem.

***Theorem* 2.** *For the case of HD C-NOMA, the optimal power control scheme, which is defined by the couple* $\left(\alpha^*_{m,n}, P^{\text{d}^*}_{m,n}\right)$*, for the inner problem* $\mathscr{P}_{\text{inner}}$*, is expressed in equation* (2.21) *on top of this page, where*

$$\left(\alpha^*_{m,n}, P^{\text{d}^*}_{m,n}\right) = \begin{cases} \left(\dfrac{(\delta^{\text{H}}_{\text{th}} - \bar{P}^{\text{d}}\gamma^{\text{d}}_{m,n})(P_{\text{BS}}\gamma_n + 1)}{P_{\text{BS}}\gamma_n(\delta^{\text{H}}_{\text{th}} - \gamma^{\text{d}}_{m,n}\bar{P}^{\text{d}} + 1)}, \bar{P}^{\text{d}}\right), & \text{if } \bar{P}^{\text{d}} \in \left[P^{\min}_{m,n}, P^{int}_{m,n}\right] \\ \left(\dfrac{\delta^{\text{H}}_{\text{th}}(P_{\text{BS}}\gamma_m + 1))}{P_{\text{BS}}\gamma_m(\delta^{\text{H}}_{\text{th}} + 1)}, P^{int}_{m,n}\right), & \text{if } \bar{P}^{\text{d}} \in \left[P^{int}_{m,n}, +\infty\right] \end{cases} \tag{2.21}$$

$$P^{int}_{m,n} = \frac{P_{\text{BS}}}{\gamma^{\text{d}}_{m,n}} \frac{\delta^{\text{H}}_{\text{th}}(\delta^{\text{H}}_{\text{th}} + 1)(\gamma_m - \gamma_n)}{P_{\text{BS}}\gamma_m(\delta^{\text{H}}_{\text{th}} + P_{\text{BS}}\gamma_n + 1) - \delta^{\text{H}}_{\text{th}}P_{\text{BS}}\gamma_n} \tag{2.22}$$

Proof. Please see Appendix A □

FD C-NOMA

In this part, we consider the case of FD C-NOMA and hence $R_{m,n}\left(\alpha_{m,n}, P^{\text{d}}_{m,n}\right) = R^{\text{F}}_m\left(\alpha_{m,n}, P^{\text{d}}_{m,n}\right) + R^{\text{F}}_n\left(\alpha_{m,n}, P^{\text{d}}_{m,n}\right)$. In this case, the feasibility conditions of the inner problem $\mathscr{P}_{\text{inner}}$ are detailed in the following theorem.

***Theorem* 3.** *FFor the case of FD C-NOMA, the inner problem* $\mathscr{P}_{\text{inner}}$ *is feasible if and only if the following conditions hold.*

$$\textit{Condition 1: } \Delta_1 \geq 0 \wedge b1 < 0 \wedge b_3 \leq b_2 \wedge \bar{P}^{\text{d}} \geq b_3, \tag{2.23a}$$

$$\textit{Condition 2: } \Delta_1 \geq 0 \wedge b1 \geq 0, \tag{2.23b}$$

$$\textit{Condition 3: } \Delta_1 < 0, \tag{2.23c}$$

where the parameters Δ_1, b_1, b_2, *and* b_3 *are defined in* (2.24)-(2.27) *on top of next page, in which* $\delta_{\text{th}}^{\text{F}} = 2^{R_{\text{th}}} - 1$ *and* Δ_2 *is expressed as shown in* (2.28) *on top of next page.*

Proof. Please see Appendix B □

$$\begin{aligned}\Delta_1 = &\left[P_{\text{BS}}\left(\gamma_m\gamma_{m,n}^{\text{d}} - \gamma_n\gamma_m^{\text{SI}}(\delta_{\text{th}}^{\text{F}})^2 - \gamma_n\gamma_m^{\text{SI}}\delta_{\text{th}}^{\text{F}} + \gamma_n\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}}\right)\right]^2 \\ &- 4P_{\text{BS}}^3\gamma_m\gamma_m^{\text{SI}}\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}}\left[P_{\text{BS}}\gamma_m\gamma_n - \gamma_m\delta_{\text{th}}^{\text{F}} - \gamma_n(\delta_{\text{th}}^{\text{F}})^2 - \gamma_n\delta_{\text{th}}^{\text{F}}\right],\end{aligned} \tag{2.24}$$

$$b_1 = \frac{\left[-\sqrt{\Delta_1} + P_{\text{BS}}^2\gamma_m\gamma_n\gamma_{m,n}^{\text{d}} + P_{\text{BS}}(\gamma_m\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}} + \gamma_m\delta_{\text{th}}^{\text{F}} + \gamma_m\gamma_{m,n}^{\text{d}} + \gamma_n\gamma_m^{\text{SI}}(\delta_{\text{th}}^{\text{F}})^2 + \gamma_n\gamma_m^{\text{SI}}\delta_{\text{th}}^{\text{F}} - \gamma_n\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})\right]}{(2P_{\text{BS}}^2\gamma_m\gamma_m^{\text{SI}}\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})} \tag{2.25}$$

$$b_2 = \frac{\left[\Delta_2 + P_{\text{BS}}^2\gamma_m\gamma_n\gamma_{m,n}^{\text{d}} + P_{\text{BS}}(\gamma_m\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}} + \gamma_m\delta_{\text{th}}^{\text{F}} + \gamma_m\gamma_{m,n}^{\text{d}} + \gamma_n\gamma_m^{\text{SI}}(\delta_{\text{th}}^{\text{F}})^2 + \gamma_n\gamma_m^{\text{SI}}\delta_{\text{th}}^{\text{F}} - \gamma_n\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})\right]}{(2P_{\text{BS}}^2\gamma_m\gamma_m^{\text{SI}}\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})}, \tag{2.26}$$

$$b_3 = \frac{\left[\sqrt{\Delta_1} + P_{\text{BS}}^2\gamma_m\gamma_n\gamma_{m,n}^{\text{d}} + P_{\text{BS}}(\gamma_m\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}} + \gamma_m\delta_{\text{th}}^{\text{F}} + \gamma_m\gamma_{m,n}^{\text{d}} + \gamma_n\gamma_m^{\text{SI}}(\delta_{\text{th}}^{\text{F}})^2 + \gamma_n\gamma_m^{\text{SI}}\delta_{\text{th}}^{\text{F}} - \gamma_n\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})\right]}{(2P_{\text{BS}}^2\gamma_m\gamma_m^{\text{SI}}\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})}, \tag{2.27}$$

$$\begin{aligned}\Delta_2 = &\left[P_{\text{BS}}^2\gamma_m\gamma_n\gamma_{m,n}^{\text{d}} + P_{\text{BS}}(\gamma_m\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}} + \gamma_m\delta_{\text{th}}^{\text{F}} + \gamma_m\gamma_{m,n}^{\text{d}} + \gamma_n\gamma_m^{\text{SI}}(\delta_{\text{th}}^{\text{F}})^2 + \gamma_n\gamma_m^{\text{SI}}\delta_{\text{th}}^{\text{F}} - \gamma_n\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}})\right]^2 \\ &- 4P_{\text{BS}}^3\gamma_m\gamma_m^{\text{SI}}\gamma_{m,n}^{\text{d}}\delta_{\text{th}}^{\text{F}}\left[\gamma_{\text{S}}(\delta_{\text{th}}^{\text{F}})^2 + \gamma_{\text{S}}\delta_{\text{th}}^{\text{F}} - \gamma_{\text{W}}(\delta_{\text{th}}^{\text{F}})^2 - \gamma_{\text{W}}\delta_{\text{th}}^{\text{F}}\right]\end{aligned} \tag{2.28}$$

Theorem 3 provides the feasibility conditions of the inner problem $\mathscr{P}_{\text{inner}}$, which ensure that there exists at least a value of $\alpha_{m,n}$ and $P_{m,n}^{\text{d}}$ that satisfy the constraints (2.18d)-(2.18e). Now, based on the results of theorem 3, the closed-form solution of the optimal power control scheme of problem for the case of FD C-NOMA is given in the following theorem.

***Theorem* 4.** *For FD C-NOMA, the optimal power control scheme of the inner problem* $\mathscr{P}_{\text{inner}}$ *is given by the couple* $\left(\alpha_{m,n}^*, P_{m,n}^{\text{d}^*}\right)$, *where* $P_{m,n}^{\text{d}^*}$ *is expressed as shown in* (2.29) *on top of next page, in which* $\mathbb{1}(\cdot)$ *denotes the indicator function, and* $\alpha_{m,n}^* = f\left(P_{m,n}^{\text{d}^*}\right)$, *such that the function* f *is given,*

for all $x \in \mathbb{R}$, *by*

$$P_{m,n}^{\mathrm{d}^*} = \begin{cases} \bar{P}^{\mathrm{d}} \times \mathbb{1}\left(0 \le \bar{P}^{\mathrm{d}} \le b_2\right) + b_2 \times \mathbb{1}\left(b_2 \le \bar{P}^{\mathrm{d}}\right) \\ \textit{if } \Delta < 0, \\ \bar{P}^{\mathrm{d}} \times \mathbb{1}\left(0 \le \bar{P}^{\mathrm{d}} \le \max(b_1, b_2)\right) + \max(b_1, b_2) \times \mathbb{1}\left(\max(b_1, b_2) \le \bar{P}^{\mathrm{d}}\right), \\ \textit{if } \Delta \ge 0 \wedge b_1 \ge 0, \\ \bar{P}^{\mathrm{d}} \times \mathbb{1}\left(\max(0, b_3) \le \bar{P}^{\mathrm{d}} \le b_2\right) + b_2 \times \mathbb{1}\left(b_2 \le \bar{P}^{\mathrm{d}}\right) \\ \textit{if } \Delta \ge 0 \wedge b_1 \le 0 \wedge b_3 \le b_2, \end{cases} \tag{2.29}$$

$$f(x) = \frac{\delta_{\mathrm{th}}^{\mathrm{F}}(\gamma_m^{\mathrm{SI}} x + \gamma_m P_{\mathrm{BS}} + 1)}{\gamma_m P_{\mathrm{BS}}(\gamma_{\mathrm{th}} + 1)}. \tag{2.30}$$

Proof. Please see Appendix B □

2.4.5 Pairing Policy and Proposed Algorithm

After deriving the optimal power control schemes for each pair and the corresponding achievable rate, we can proceed to apply the Hungarian method to determine the optimal pairing configurations. Let $\mathbf{g}$ be the $2K \times 1$ vector that contains the channel gains from the BS to the AUs in a way that are sorted in an ascending order, where entries $g_1, g_2, ..., g_K$ represent the channel gains of the weak AUs and $g_{K+1}, g_{K+2}, ..., g_{2K}$ represent the channel gains of the strong AUs. In addition, let $\mathbf{D}$ be the $K \times K$ matrix that contains the D2D channel gains and let $\mathbf{s}$ be the $K \times 1$ vector that contains the SI channel gains. The input of Hungarian algorithm is a $K \times K$ cost matrix $\mathbf{C}$ and its output is the pairing matrix $\mathbf{B}^*$ where $\mathbf{B}^*(m,n) = 1$ indicates that the AU m in the strong AUs set $\mathscr{S}$ is paired with the AU n in the weak AUs set $\mathscr{W}$, and $\mathbf{B}^*(m,n) = 0$ otherwise. In our algorithm, we define the cost of pairing two users as the opposite value of the sum rate obtained by solving the inner problem $\mathscr{P}_{\mathrm{inner}}$. After the computation of $\mathbf{C}$, the Hungarian algorithm is applied to solve the outer problem $\mathscr{P}_{\mathrm{outer}}$. The proposed algorithm is highlighted in Algorithm 1. After computing the matrix $\mathbf{B}^*$, it is straightforward to solve problem $\mathscr{P}$ since the optimal value $\alpha_{m,n}^*$ and $P_{m,n}^{\mathrm{d}^*}$ within each pairs have been found in step 2 of Algorithm 1. After applying algorithm 1, we obtain the optimal pairing policy and the optimal power control scheme within each pair, which is the final

Algorithm 1 Low-complexity algorithm .

1. **Estimate** channel gain vector $\mathbf{g}$, D2D channel gain matrix $\mathbf{D}$, and the SI channel gain vector $\mathbf{s}$.
2. **Compute** the optimal rate for each pairing configuration $R^*_{m,n}$ following theorem 2 for HD case and theorem 4 for FD case.
3. **Compute** the cost matrix $\mathbf{C}$ with $\mathbf{C}(m,n) = -R^*_{m,n}$.
4. **Solve** the optimal pairing matrix $\mathbf{B}^*$ using the Hungarian algorithm with cost matrix $\mathbf{C}$.

solution for problem $\mathscr{P}$.

2.4.6 Mode Selection

Mode selection has been proposed in previous works on relaying systems in general and C-NOMA based system in particular [35, 53]. Mode selection assumes that the user device have the capability of both FD and HD relaying and can switch from one mode to the other (hybrid devices). The idea of mode selection stems from the fact that FD relaying does not necessarily perform better than HD relaying especially with the presence of high SI. Since we can compute the resulting optimal sum-rate value for both FD C-NOMA and HD C-NOMA based on theorem 2 and theorem 4, determining the optimal mode is straightforward. In fact, we just need to select the mode that gives the highest resulting sum-rate from the power control schemes derived in these theorems.

2.4.7 Complexity Analysis

Complexity is worth bringing into discussion, since the complexity of the proposed method seems extreme and it requires the computation of the rate values for all possible pairing configurations. However, it is worth mentioning that, due to the closed-form solution obtained in Theorems 2 and 4, the computational complexity of obtaining the optimal sum rate for a given pair is approximately $\mathscr{O}(1)$. Thus, even for all the possible configurations, the computational complexity of obtaining all the sum rates is $\mathscr{O}(K^2)$. With the addition of the Hungarian method, the total computational complexity of our proposed algorithm is approximately $\mathscr{O}(K^2+K^3)$ or $\mathscr{O}(K^3)$ for large values of K. Clearly, the overall computational complexity depends more on the Hungarian method than the computation of the cost matrix $\mathbf{C}$.

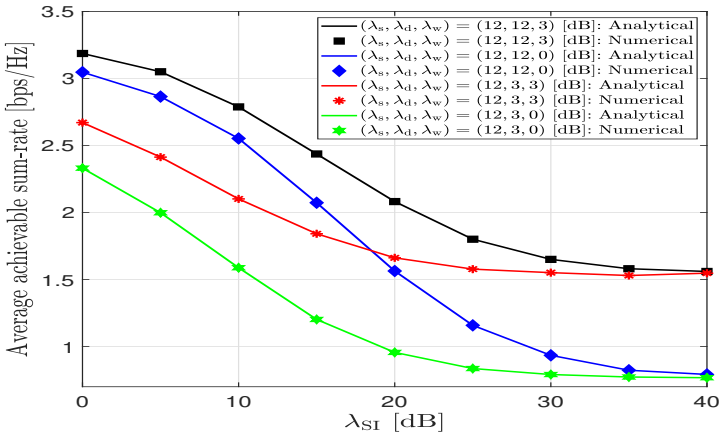


Figure 2.3: Analytical and numerical average achievable sum-rate versus the average self interference channel λ_{SI}.

2.5 Simulation Results

In this section, we will validate the proposed scheme. We assume that the channel gains of the strong AU h_m, the weak AU h_n, the D2D link $h^{\mathrm{d}}_{m,n}$, and the SI h^{SI}_m follows independent Rayleigh distributions with scale parameters, σ_{s}, σ_{w}, σ_{d} and σ_{SI}, respectively. Therefore, their associated squares γ_m, γ_n, $\gamma^{\mathrm{d}}_{m,n}$ and γ^{SI}_m follow independent exponential distributions with means $\lambda_{\mathrm{s}} = (2\sigma_{\mathrm{s}})^{-1}$, $\lambda_{\mathrm{w}} = (2\sigma_{\mathrm{w}})^{-1}$, $\lambda_{\mathrm{d}} = (2\sigma_{\mathrm{d}})^{-1}$, and $\lambda_{\mathrm{SI}} = (2\sigma_{\mathrm{SI}})^{-1}$, respectively. The required QoS is defined by the minimum achievable rate threshold that is given by $R_{\mathrm{th}} = 1$ [bps/Hz]. Simulation results are performed over 10^5 independent Monte-Carlo trials on the channel gain realizations.

Fig. 2.3 presents the analytical and numerical average sum rate for one pair of users with a FD C-NOMA transmission scheme versus the average self interference channel λ_{SI} for different means of the channel gain of the strongest AU λ_{s}, the weakest AU λ_{w} and the D2D link λ_{d}. The analytical results are obtained through the closed-form power control scheme derived in theorem 4 whereas the numerical results are obtained by solving problem $\mathscr{P}_{\mathrm{inner}}$ using an off-the-shelf optimization solver.[5] This figure shows that the analytical results match perfectly the numerical results, which validate the optimality of the power control scheme derived in theorem 4. In addition, Fig. 2.3 shows that when the mean of the SI channel increases, the average sum rate decreases. This observation is

[5]The adopted solver is fmincon, which is a predefined matlab solver [54]. In addition, 100 distinct initial points were generated in order to converge to the optimal solution.

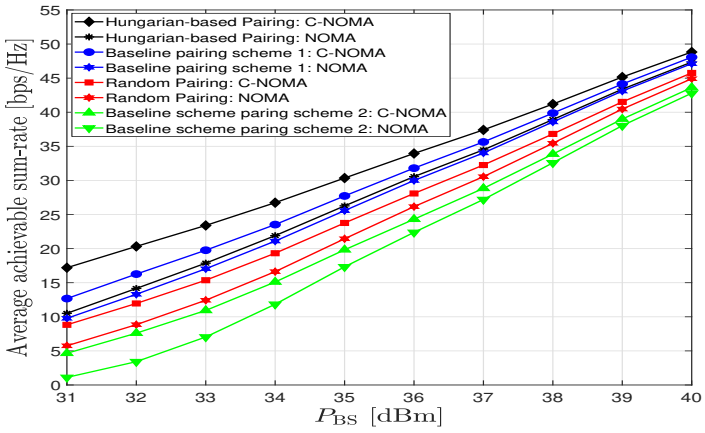


Figure 2.4: Average achievable sum-rate versus the power budget at the base station P_{BS} for C-NOMA and NOMA with different pairing schemes ($\lambda_{\mathrm{s}} = 10$ dB, $\lambda_{\mathrm{w}} = 0$ dB, $\lambda_{\mathrm{d}} = 6$ dB, $\lambda_{\mathrm{SI}} = 0$ dB, $\bar{P}^{\mathrm{d}} = 30$ dBm)

expected since as shown in equations (10) and (11), the achievable rates of the strongest AU and the weakest AU are decreasing with respect to λ_{SI}. On the other hand, by comparing the cases where $(\lambda_{\mathrm{s}}, \lambda_{\mathrm{d}}, \lambda_{\mathrm{w}}) = (12, 12, 3)$ [dB] and $(12, 3, 3)$ [dB], this figure shows that, when the mean of the SI channel increases, the potential of the D2D relaying decreases. The same observation holds when comparing the cases $(\lambda_{\mathrm{s}}, \lambda_{\mathrm{d}}, \lambda_{\mathrm{w}}) = (12, 12, 0)$ [dB] and $(12, 3, 0)$ [dB]. This observation makes sense as well, since when λ_{SI} increases, the SI at the strongest AU increases, and thus, the only way to alleviate its effect is decreasing the transmit D2D relaying power $P^{\mathrm{d}}_{m,n}$. This makes C-NOMA converges to the conventional NOMA.

Fig. 2.4 presents the average sum rate achieved by the FD C-NOMA and conventional NOMA with different pairing schemes versus the power budget at the base station P_{BS}. For conventional NOMA, different pairs of AUs are served via OMA, where the AUs within each pair are served via NOMA with no cooperation between users. Besides, we compare the proposed user pairing policy with three different pairing schemes. For baseline pairing scheme 1, we pair the kth AU $(1 \leq k \leq K)$ with the $(2K-k+1)$th AU, e.g, the weakest AU is paired with the strongest AU, the second weakest is paired with the second strongest, and so on [27]. For baseline pairing scheme 2, we pair the kth AU $(1 \leq k \leq K)$ with the $(K+k)$th AU, e.g, the kth weakest AU in the set of weak AUs $\mathscr{W}$ is paired with the kth weakest AU in the set of strong AUs $\mathscr{S}$. For random pairing scheme, two randomly selected AU in each set are paired with each other. Fig. 2.4 shows that the proposed user pairing

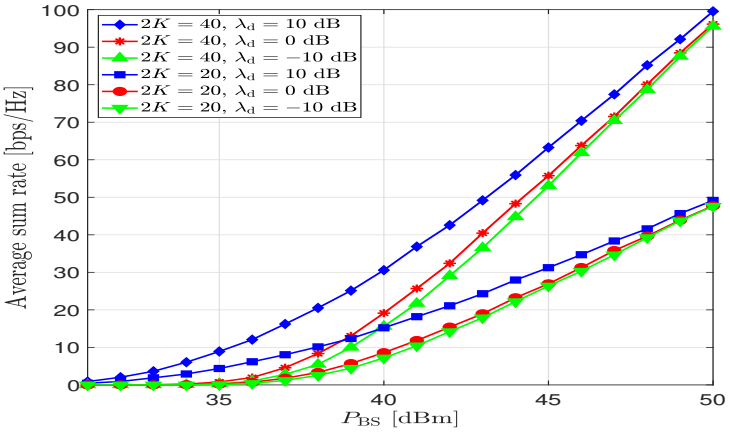


Figure 2.5: Average sum rate versus the power budget at the base station P_{BS} for C-NOMA and different means of the D2D channel gain λ_{d} ($\lambda_{\text{s}} = 10$ dB, $\lambda_{\text{w}} = 6$ dB $\lambda_{\text{SI}} = 6$ dB, $\bar{P}^{\text{d}} = 30$ dBm)

policy outperforms the baseline schemes. In addition, this figure shows that enabling cooperation between users within each pair improves the performance of the system.

Fig. 2.5 presents the average sum rate achieved by FD C-NOMA and the Hungarian method versus the power budget at the base station P_{BS} when the total number of users is $2K = 20$ and $2K = 40$ and for different values of the mean D2D channel λ_D. This figure shows that the average sum rate increases when the total number of users increases and when the mean D2D channel λ_D increases. This observation is expected since, as shown in equation (13) the achievable rate of the weak AU increases when λ_D increases. Note that, in Fig. 2.3, Fig. 2.4 and Fig. 2.5, the performance of HD C-NOMA follows similar pattern to the one of FD C-NOMA and this is why it is omitted here.

In Fig. 2.6, we evaluate the performance of the system when the AUs can operate under different relaying modes. In total, we compare between three relaying modes, which are HD relaying, FD relaying and hybrid relaying with mode selection as presented in subsection 2.4.6. In addition, we consider the cases where, within each pair, either the strong AU forwards the message of the weak AU with all its available D2D power budget $\bar{P}^{\text{d}}$, i.e., without power adaptation, or it employs the closed-form solution of the optimal transmit D2D power $P^{\text{d}^*}_{m,n}$ obtained in theorem 2 for the HD case and in theorem 4 for the FD case. Three observations can be remarked from Fig. 2.6. First, with

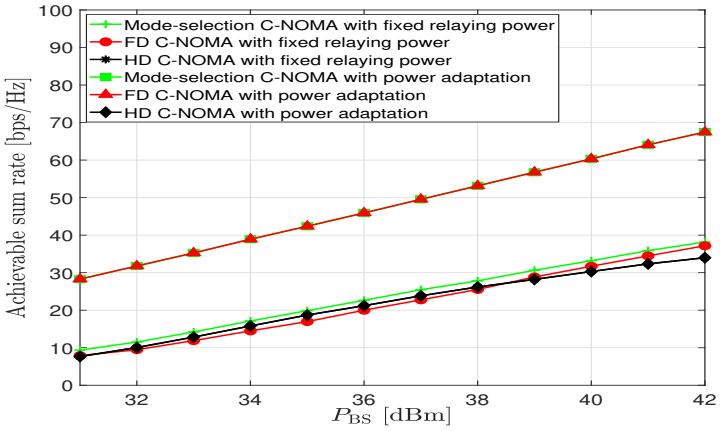


Figure 2.6: Comparisons between different relaying modes of cooperation in C-NOMA system for fixed and adaptive relaying power control with different power budget at the BS P_{BS} ($\lambda_{\mathrm{s}} = 10$ dB, $\lambda_{\mathrm{w}} = 6$ dB, $\lambda_{\mathrm{d}} = \lambda_{\mathrm{SI}} = 6$ dB).

power adaptation and assuming the given parameters, the best relaying mode is FD mode, which explains why C-NOMA with mode selection and FD C-NOMA have the same performance. Second, the HD C-NOMA is not affected by the power adaptation at each strong AU, since even with power adaptation, HD C-NOMA will always choose to transmit with maximum power. This observation is expected because HD C-NOMA operates without the induction of SI at the strong users. Thus, the higher power the strong AUs can transmit with, the higher the achievable sum rate is. The third observation is that, for the case when there is no power adaptation at the strong AUs, HD C-NOMA can outperform FD C-NOMA although FD C-NOMA can operate over the whole subchannel assigned. Obviously, without a proper power adaption that is based on the channel conditions, the transmit power at the strong AUs will adversely affect its achievable rate due to SI. This trend is highly observable when the BS power budget P_{BS} is lower than D2D power budget $\bar{P}^{\mathrm{d}}$, which means that the impact of the D2D relaying transmission becomes comparable to that of the downlink access transmission. In this situation, the relaying mode becomes an important feature since the SI channel gain h_{SI} and the adopted D2D transmit power $P^{\mathrm{d}}_{m,n}$ will significantly impact the overall achievable sum rate. When the power P_{BS} becomes higher, the downlink access transmission can be enough to fulfil the minimum QoS constraints and the D2D cooperative transmission becomes less influential on the sum rate of the overall system; therefore FD C-NOMA will prevail HD C-NOMA since the downlink transmission can occupy double the bandwidth resource.

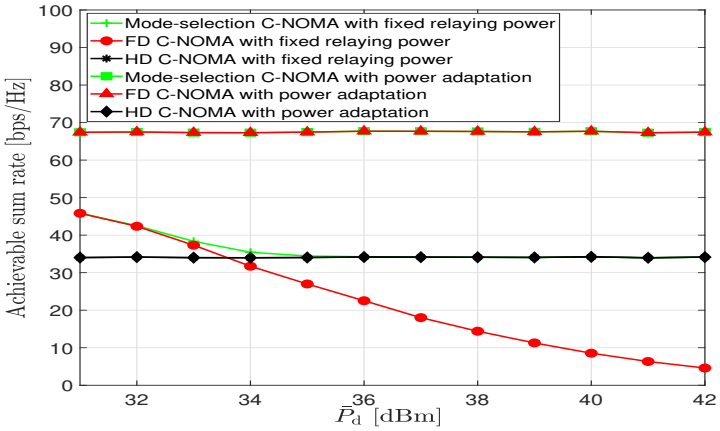


Figure 2.7: Comparisons between different relaying modes of cooperation in C-NOMA system for fixed and adaptive relaying power control with different power budget at the user device $\bar{P}^{\mathrm{d}}$ ($\lambda_{\mathrm{s}} = 10$ dB, $\lambda_{\mathrm{w}} = 6$ dB, $\lambda_{\mathrm{d}} = 6$ dB, $\lambda_{\mathrm{SI}} = 0$ dB, $P_{\mathrm{BS}} = 42$ dBm).

Fig. 2.7 presents the average achievable sum rate of the overall C-NOMA system versus the D2D power budget $\bar{P}^{\mathrm{d}}$ for the three aforementioned relaying modes, with and without power adaptation at the strong AUs. Similar to Fig. 2.6, for the case with power adaptation, FD C-NOMA and mode-selection C-NOMA have the same performance and stay constant even when the D2D power budget increases. The reason is that the relaying power in these schemes will adapt to the various channel conditions and operate at the optimal power. In other words, it is not necessary to exhaust the D2D transmit power as in the fixed relaying scheme for two reasons. First, it is energy-inefficient for the users' devices. Second, the overall sum rate decreases when there is no power adaptation. Finally, it is also worth noting that for the case of fixed relaying power scheme and low D2D power budget, mode-selection C-NOMA converges to FD C-NOMA, whereas it converges to HD C-NOMA at high D2D power budget.

In Fig. 2.8, we compare the computational time of the proposed algorithm in 1 with one of the SCA-based scheme. In the SCA-based scheme, the inner problem $\mathscr{P}_{\mathrm{inner}}$ is solved using successive linear relaxations of all the non-convex functions until convergence, where each relaxation is the input to an off-the-shelf optimization solver. This method has been repeatedly used for many non-convex optimization problems. Despite giving decent performance, it can be seen that the computational time for SCA-based method scale exponentially when the number of AUs increases while the computational time of our proposed scheme is negligible. For instance, when the number

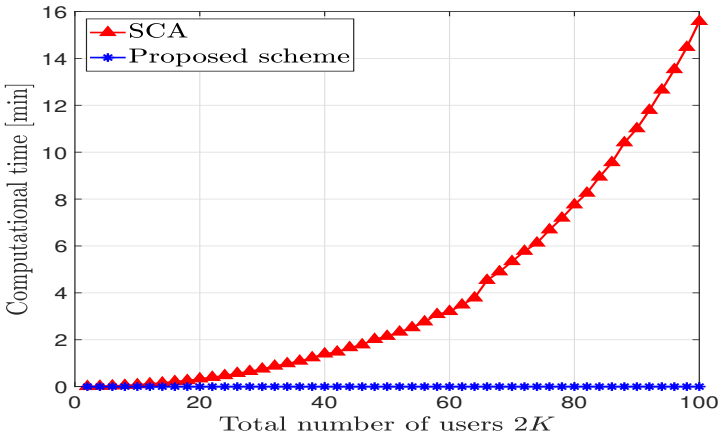


Figure 2.8: Computational time of the proposed scheme versus that of SCA-based approach .

of users is 100, it takes about 16 minutes to compute the optimal result using SCA-based scheme, whereas it takes about 10 ms to compute the optimal solution using the proposed algorithm.

2.6 Conclusion

In this chapter studied the performance of C-NOMA in a cellular downlink system, in which users that have the capability of both HD and FD communications can assist the transmissions between the BS and users with poor channel quality through D2D communications. A novel optimization problem that jointly determines the optimal D2D user pairing and the optimal power control scheme was formulated and solved, where the objective is maximizing the achievable sum rate of the whole system while guaranteeing a certain quality of service (QoS) for all users. A two-step policy was proposed to solve the problem in a polynomial time. First, the closed-form expression of the optimal power control scheme that maximizes the sum rate of a given pair of users with a required QoS was derived for both cases of FD and HD communications. Then, using the derived expressions, the Hungarian method was adopted as a user pairing policy, which provided the optimal pairing strategy. The simulation results showed that the proposed scheme always outperforms some existing schemes in the literature.

Moreover, this chapter has shown promising compatibility of cooperative communications with some of the enabling technologies for 5G and beyond. Precisely, cooperative further leverage the

inherent advantages of NOMA, D2D communications, and full-duplex communications to provides considerable performance gain. In the next chapter, we attempt to employ cooperative communications by extending coordinated multipoint, a core technology of 4G, with the assistance of unmanned aerial vehicles, a newly proposed communication platform for future wireless networks.

Chapter 3

Joint Location and Beamforming Design for Cooperative UAVs with Limited Storage Capacity

3.1 Introduction

Future wireless communication technologies are under development to adapt to different application scenarios of 5G and beyond (5G/B5G). Specifically, Enhanced Mobile Broadband (eMBB), Massive Machine-Type Communication (mMTC) and Ultra-Reliable Low-Latency Communication (URLLC), also called mission-critical communications, are widely considered the key services in 5G [1]. In other words, 5G networks are expected to accommodate a huge number of mobile devices with heterogeneous and stringent requirements in terms of data rate, latency, and reliability. To meet the heightened expectations for 5G, numerous wireless technologies have been developed including mmWave communications [55], massive MIMO [56], or ultra-densification [57]. However, capacity-enhancing solutions such as installing more cell sites (ultra-densification) or using additional spectrum (mmWave) are costly and may sometimes lead to underutilization of resources due to the time-varying network traffic. Thus, an alternative is to integrate a highly flexible communication platform that can be swiftly deployed to adapt to the dynamic traffic demand, yet ensure

adequate quality of service (QoS).

Recently, the advent of wireless communication systems enabled by unmanned aerial vehicles (UAVs) has attracted significant attention. Compared to the conventional cellular network infrastructure such as ground base stations (GBSs), UAVs are highly mobile, much cheaper and faster to deploy, and able to establish favorable line-of-sight (LoS) communication channels due to their high-altitude deployment for more reliable transmissions [6, 58]. Due to these desirable attributes, UAVs can offer a wide range of applications including coverage expansion [59], data collection [60], computation offloading [61], and mobile relaying [62]. Despite their enormous potential, UAVs reveal some technical limitations hindering their operations in wireless networks such as the SWaP (size, weight, and power) constraints [63]. In other worlds, UAVs can only operate with a low power budget and carry a limited payload of storage, computing, and communication devices. Another major challenge is to manage the severe interference induced by the LoS channels in systems of multiple UAVs [64]. To address these challenges, a proper deployment for UAVs needs to be simultaneously considered with effective communication techniques.

With the recent advances of distributed antenna systems (DAS), multicast beamforming [3] has leveraged the cooperative communications over multiple transmission/reception points and has been realized to be an effective interference mitigation technique and been incorporated in LTE Advanced Releases. While interference management techniques such as power allocation depend only on the handling of signal amplitude, beamforming exploits both amplitudes and phases of the transmitter-receiver channels for proper beamforming design to improve spatial multiplexing gain. In the context of UAV communications, the UAVs can operate as a virtual distributed antenna array whose elements (UAVs) can be swiftly relocated to offer an additional design degree to the original beamforming optimization problem and further enhance macrodiversity gain of DAS. For instance, UAVs can fly closer to target users to reduce the impact of path loss and inter-UAV interference. As a result, joint optimization of beamforming design (beamformers) and location planning for cooperative UAVs, subject to their operational constraints in terms of limited power and storage capacity is an appealing, yet challenging problem to address.

3.1.1 Related work

Recently, extensive efforts have been dedicated to the study of UAVs' utility in wireless communication systems. In [65] and [66] the authors studied the optimal trajectory of UAVs for computational offloading and data collection, respectively. Besides trajectory design, several works have studied the optimal placement of UAVs for different use cases. Typically, the authors in [67] proposed a low-complexity algorithm for 3-D UAV placement for maximum coverage. Similarly, the positioning problem of UAVs is studied in [68], where a neural network based technique is used to minimize the delay in heterogeneous wireless networks, while the authors in [69] considered the optimal placement of relaying UAVs for maximum communication reliability. The work [70] drew a comparison between the performance of UAVs acting as aerial base stations and traditional terrestrial base stations in terms of average sum rate and transmit power.

One of the major obstacles to exploit the full potential of UAVs lies in their limited flight time. Unlike fixed terrestrial communication infrastructure such as GBSs, the performance of UAV systems is fundamentally limited by the on-board energy [6]. Inspired by this, several works focus on the aspect of energy-efficiency and service time for UAVs. Typically, the authors in [71] studied the communication between a UAV and a ground terminal by optimizing the UAV's trajectory to maximize the energy efficiency as a function of throughput and energy consumption. In [72], a UAV's trajectory is designed to minimize its mission completion time while still ensuring all the ground nodes can recover the disseminated file. However, in these works, the communication systems include only one UAV. In terms of security, [73] performed an analysis for cache-enabled UAVs to assist secure transmission in hyperdense networks.

Besides UAV-related literature, it is worth noting that interference management with limited storage capacity at the transmitters has been an well-established research topic for networks of fixed infrastructure. Typically, the authors in [74] leveraged tools from stochastic geometry and difference-of-convex (DC) programming to maximize successful transmission probability. However, [74] utilized the multiple antennas at the receivers as a diversity technique without considering multiple cooperative transmissions by exploiting duplicate files at the transmitters. In addition, the

swift and cost-effective deployment of UAVs shows more promises for sudden variations for network demand traffic since installing additional infrastructures, such as cache helpers as in [74] or GBSs, is often slow or even economically infeasible.

3.1.2 Motivations and contributions

Although there is a fair amount of research on UAV-enabled wireless communication system, most considered single-UAV systems [65–67, 69, 71, 72, 75–77]. Multi-UAV scenarios are considered in [78] and [79]. However, these works simplify interference management by considering power allocation or using orthogonal resource blocks, which do not fully exploit in-network spatial multiplexing gain. Most recently, Liu *et al* have considered zero-forcing beamforming (ZFBF) for an uplink multi-UAV setup by utilizing multiple receptions at UAVs to improve the minimum throughput [80]. In this chapter, we instead focus on multicast beamforming for downlink transmissions, where joint (cooperative) transmissions are influenced not only by location planning but also by strategic content placement for the UAVs. The advantages of our proposed UAV-enabled cooperative communication can be listed as follows.

- **High scalability**: the number of UAVs (antenna elements) can be adjusted on-demand and is not limited by space constraint as in centralized antenna systems. Thus, UAVs can be flexibly added (removed) to (from) the system based on the time-variant network traffic to meet users' QoS requirements or to avoid resource underutilization.

- **Enhanced reliability**: by deploying multiple cooperative UAVs, we alleviate the cost when one node (UAV) is down by reconstructing network topology.

- **Higher energy efficiency**: by replacing a high-power centralized antenna by a group of low-power cooperative UAVs, we reduce the path loss and increase the presence of LoS channels. Hence, each UAV consumes less energy and achieves longer flight time.

In this work, we consider the communication system from a content-centric viewpoint, where the central entities are the "named contents". This viewpoint is especially relevant to the context of 5G and B5G networks with a huge number of subscribers, where the system is no longer a host-centric

network but mostly a distribution network [81]. Our contributions in this work can be summarized as follows. [1].

- We propose a network architecture that incorporates multiple cooperative UAVs to serve ground users. In our system, we employ multicast beamforming as the interference management scheme. The UAVs serve the ground users requesting different contents via spatial multiplexing and joint transmissions to achieve their required QoS in terms of transmission rates. The users that experience the (minimum) required QoS are referred to as "admitted users".

- Since the cooperation among UAVs depends on the availability of the contents on each UAV and the location planning, we formulate the joint problem of determining content placement, location planning, admission strategy, and beamforming design to maximize the number of admitted users.

- Our formulation is a mixed-integer non-convex optimization problem which is very difficult to solve in polynomial time. Thus, we propose a framework based on the concept of difference-of-convex (DC) programming to approximate it into a series of convex problems and successively solve them until the results converge.

- We perform extensive numerical results to show that our scheme outperforms previous communication schemes proposed in the literature and offer some insights by varying some of the system parameters. A notable trend in our results is that increasing the number UAVs remarkably reduce the power consumption per UAVs for the same user admission. This finding offers a possible solution to save power consumption and improve the UAVs' lifetime, which is a major concern in UAV communications.

The rest of the chapter is organized as follows. Section 3.2 introduces the system model. The problem formulation is presented in 3.3. Section 3.4 proposes the solution approach. Section 3.5 presents the development of our low-complexity algorithm. Section 3.6 presents and discusses our numerical results. Finally, conclusions are drawn in Section 3.7.

[1] Part of this work has been presented at the IEEE Wireless Communications and Networking Conference 2019 [82]

3.2 System Model

3.2.1 Spatial model

We consider a wireless communication system which employs multiple single-antenna UAVs, each integrated with a storage device of limited size and playing the role of aerial base stations to provide delay-tolerant wireless services to the ground users as in Fig. 3.1. Let us denote $\mathscr{U} = \{0,...,U\}$ as the set of UAVs , $\mathscr{N} = \{1,...,N\}$ as the set of users within a given geographical region, and $\mathscr{K} = \{1,...,K\}$ as the set of contents to be requested, on which we will elaborate further in section 3.2.2. We divide set $\mathscr{N}$ into multiple subsets $\mathscr{N}_k, \forall k \in \mathscr{K}$, where $\mathscr{N}_k$ indicates the group of users requesting the k-th content belonging to set $\mathscr{K}$, which we will call group k from now on. To serve ground users by the UAVs, we apply âĂIJ*fly-then-hover-and-transmit*âĂİ operation as in [83], where rotary-wing UAVs are a good option for deployment. In other words, based on users' locations and content request, the UAVs fly to designated locations and hover there to transmit the requested contents in a cooperative manner. Multiple UAVs can perform beamforming and joint transmission based on the availability of the contents in their storage devices to improve the quality of received signals at the user ends. We consider the relative positions of the users and UAVs to be quasi-static within the period of service. The locations of UAVs can be dynamically adjusted at the next period of service when users' locations change or a new swarm of UAVs can be deployed while the first swarm returns to some docking stations to recharge. For simplicity, each UAV is deployed at a fixed altitude H and denote the coordinate of the i-th UAV as $p_i = \{x_i, y_i, H\}, \forall i \in \mathscr{U}$. For notational convenience, let $\mathbf{p} = \{p_i, \forall i \in \mathscr{U}\}$ be the location vector of all UAVs. Then, let $\{x_{jk}, y_{jk}, 0\}, \forall k \in \mathscr{K}$ denote the position of the j-th user of group k. As a result, the distance between the i-th UAV and the j-th user of group k is $d_{ijk} = \sqrt{(x_{jk}-x_i)^2 + (y_{jk}-y_i)^2 + H^2}$.

3.2.2 Content request distribution model

The $\mathscr{K}$ represents the library of K contents requested by ground users. K is a finite positive integer and the content library follow a popularity order where content k is the k-th most requested one. The random variable k that follows a Zipf distribution for which the probability mass function is given by

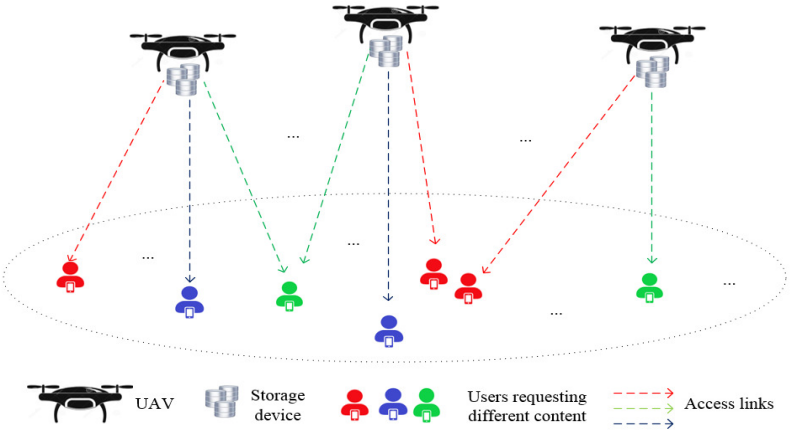


Figure 3.1: Cooperative UAVs serving users.

$$f(k) = \frac{k^{-\alpha}}{\sum_{i=1}^{K} i^{-\alpha}}, \tag{3.1}$$

where $0 < \alpha < 1$ is the shape parameter that characterizes the Zipf distribution and is considered a priori. The larger α is, the more likely the most popular contents are requested. In other words, α represents the level of concentration of the request distribution around the most popular content. The function $f(k)$ is interpreted as the likelihood that content k is requested. For simplicity, the size of each content is assumed to be 1 and the storage capacity at each UAV is characterized as the maximum number of contents the UAVs can store. However, the concept of our problem can be easily extended to inhomogeneous sizes of contents. We further assume that the content request distribution does not change while the UAVs are serving users. This is a valid assumption regarded by most previous literature.

3.2.3 Channel model

In this work, we consider free-space path-loss (FSPL) model with Rician small-scale fading following the authors in [84] and [77].[2]

[2]When the UAV's altitude is comparable to the cell radius as in [77], the probability of non-line-of-sight occurrence is negligible. This can be easily verified using the expression of LoS probability included in some relevant channel modeling literature such as [58].

Let us denote h_{ijk} as the AtG channel from the ith UAV to the jth user in the group k. Specifically, h_{ijk} is a composition of the small-scale fading coefficient $\tilde{h}_{ijk}$, $\mathbf{E}(|\tilde{h}_{ijk}|^2 = 1)$ and the path-loss coefficient θ_{ijk} which is given by

$$h_{ijk}(\theta_{ijk}, \tilde{h}_{ijk}) = \theta_{ijk}\tilde{h}_{ijk}, \tag{3.2}$$

where the path-loss θ_{ijk} is a function of the relative distance between the i-th UAV and the j-th user of group k and can be written as

$$\theta_{ijk} = d_{ijk}^{-n} = \theta_0 \sqrt{(x_{jk} - x_i)^2 + (y_{jk} - y_i)^2 + H^2}^{-n}, \tag{3.3}$$

where $n > 2$ is the path loss exponent, and θ_0 is the path loss at a reference distance 1 m. We assume $\tilde{h}_{ijk}$ follows a K-factor Rician distribution ($K > 0$). Hence, $\tilde{h}_{ijk}$ is composed of a deterministic component $\bar{h}_{ijk}$ ($\left|\bar{h}_{ijk}\right| = 1$), and a random scattering component $\hat{h}_{ijk}$ following a complex Gaussian distribution $\mathscr{CN}(0,1)$. The small-scale fading coefficient $\tilde{h}_{ijk}$ is given by

$$\tilde{h}_{ijk} = \sqrt{\frac{K}{1+K}}\bar{h}_{ijk} + \sqrt{\frac{1}{1+K}}\hat{h}_{ijk}, \tag{3.4}$$

3.2.4 Achievable rate analysis

We denote $s_k, \forall k \in \mathscr{K}$ as the signal with unit power, i.e., $\mathbb{E}\{s_k s_k^*\} = 1$ intended for the user group requesting c_k, $\mathbf{w}_k = [w_{1k}, w_{2k}, ..., w_{Uk}]^\intercal, \forall k \in \mathscr{K}$ as the beamforming vector from the U UAVs to the jth users of group k and $\mathbf{h}_{jk}(\theta_{jk}, \tilde{\mathbf{h}}_{jk}) = \left[h_{1jk}(\theta_{1jk}, \tilde{h}_{ijk}), \ldots, h_{Ujk}(\theta_{Ujk}, \tilde{h}_{ijk})\right]^\intercal$ as the vector that concatenates all the channels from U UAVs. To decide whether content k is placed in the i-th UAV, we introduce the binary variable $b_{ik}, \forall k \in \mathscr{K}, \forall j \in \mathscr{N}_k$. In particular, $b_{ik} = 1$ enforces that content k is stored in the i-th UAV and $b_{ik} = 0$ otherwise. Obviously, $b_{ik} = 0$ means the i-th UAV does not hold the content k and is therefore not able to cooperatively serve the users requesting content k. Thus, this UAV does not allocate any resource to these users. In other words, the beamforming applied on these users from i-th UAV is 0. Similarly, we denote $\mathbf{b}_k = [b_{1k}, b_{2k}, ..., b_{Uk}]^\intercal$ and $\mathbf{b} = \{\mathbf{b}_k, \forall k \in \mathscr{K}\}$

for notational convenience. Thus, the received signal at the j-th user of group k is writen as

$$y_{jk} = \mathbf{h}_{jk}(\theta_{jk}, \tilde{\mathbf{h}}_{jk})^{\dagger}(\mathbf{w}_k \circ \mathbf{b}_k)s_k + \sum_{\ell \neq k} \mathbf{h}_{jk}(\theta_{jk}, \tilde{\mathbf{h}}_{jk})^{\dagger}(\mathbf{w}_\ell \circ \mathbf{b}_\ell)s_l + n_{jk}, \tag{3.5}$$

In (3.5), the first term is the signal of interest at the j-th user of group k, while the second term is the interfering signals (intended for the user requesting different contents) and $n_{jk} \sim \mathscr{CN}(0, \sigma_0^2)$ is the additive white Gaussian Noise (AWGN). For notational convenience, we denote $\mathbf{w} = \{\mathbf{w}_k, \forall k \in \mathscr{K}\}$ and $\theta = \left\{\theta_{jk}, \forall k \in \mathscr{K}, \forall j \in \mathscr{N}_k\right\}$. Given the above assumption, the signal-to-interference-plus-noise ratio at the j-th user of group k is written as

$$\Gamma_{jk}(\mathbf{w}, \theta, \tilde{\mathbf{h}}_{jk}, \mathbf{b}) = \frac{\left|(\mathbf{h}_{jk}(\theta_{jk}, \tilde{\mathbf{h}}_{jk}))^{\dagger}(\mathbf{w}_k \circ \mathbf{b}_k)\right|^2}{\sum\limits_{\ell \neq k} \left|(\mathbf{h}_{jk}(\theta_{jk}, \tilde{\mathbf{h}}_{jk}))^{\dagger}(\mathbf{w}_\ell \circ \mathbf{b}_\ell)\right|^2 + \sigma_0^2}, \tag{3.6}$$

Then, by treating interference as noise, the achievable rate in bps/Hz at the user is

$$R_{jk}(\mathbf{w}, \theta, \tilde{\mathbf{h}}_{jk}, \mathbf{b}) = \log(1 + \Gamma_{jk}(\mathbf{w}, \theta, \tilde{\mathbf{h}}_{jk}\mathbf{b})), \tag{3.7}$$

To highlight the benefits of cooperative communication amongst limited-storage UAVs, let us perform some analysis on the following example:

*An illustration for cooperative transmission design for limited-storage UAV*s: In Fig. 3.2, we consider a system of 5 UAVs and a group of ground users requesting 20 different contents that are randomly distributed in the service region. We assume there are 20 contents to be requested by the users and the storage space at each UAV is 5 contents. The contents are loaded on each UAV and then the UAVs are dispatched to serve ground users. Note that in this example, although we assume a random content placement scheme for ease of illustration, a more strategic content placement is determined via the optimization problem we introduce later. The UAVs are able to serve the users via beamforming. Since the total storage space is 25, some UAVs can accommodate copies of the same contents and therefore can cooperate with each other to serve ground users. In Fig. 3.2, UAVs 1 and 2 store the same contents c_1 and c_2 and can support each other to serve the users requesting these file via a proper design of the weights w_{11}, w_{12}, w_{21}, w_{22}. Similarly, UAVs 2, 3 and 4 all store

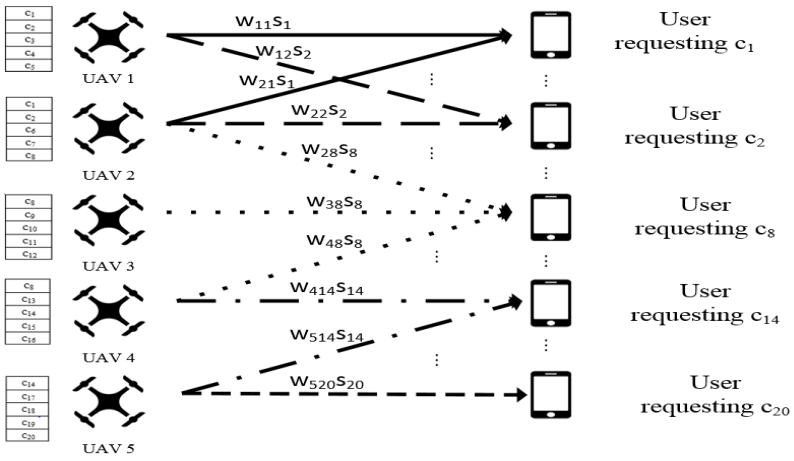


Figure 3.2: An illustration for cooperative transmission of UAVs with limited storage capacity

the content c_8 and thus can cooperate to serve users requesting it. The cooperation rule remains the same for UAVs 4 and 5. Since users requesting the duplicate contents stored in multiple UAVs benefit from the cooperation, the overall network performance is improved.

3.3 Problem Formulation

3.3.1 Formulation approach

Before formulating the problem, it is worth noting that the rate function is constituted by the channel coefficient $h_{ijk}(\theta_{ijk}, \tilde{h}_{ijk})$. However, the small-scale fading $\tilde{h}_{ijk}$ is a random complex number and is impossible to obtain prior to the deployment of UAVs. Without a deterministic value of small-scale fading, formulation of joint beamforming and location optimization cannot be done. Thus, our proposed approach is first considering the AtG channels are composed of the path loss only. In other words, we initially neglect the impact of small-scale fading and rewrite (3.2), (3.6), and (3.7) as functions of merely path loss as follows.

$$h_{ijk}(\theta_{ijk}) = \theta_{ijk}, \tag{3.8}$$

$$\Gamma_{jk}(\mathbf{w}, \theta, \mathbf{b}) = \frac{\left|(\mathbf{h}_{jk}(\theta_{jk}))^{\dagger}(\mathbf{w}_k \circ \mathbf{b}_k)\right|^2}{\sum\limits_{\ell \neq k} \left|(\mathbf{h}_{jk}(\theta_{jk}))^{\dagger}(\mathbf{w}_\ell \circ \mathbf{b}_\ell)\right|^2 + \sigma_0^2}, \tag{3.9}$$

$$R_{jk}(\mathbf{w}, \theta, \mathbf{b}) = \log(1 + \Gamma_{jk}(\mathbf{w}, \theta, \mathbf{b})), \tag{3.10}$$

With the deterministic form of the rate function (3.10), we can proceed to formulate the user admission maximization problem. Then, after solving the first problem, we can obtain the solution for user admission decision, content placement, beamforming and UAVs' locations. However, when deploying UAVs following the obtained solution, the rate quality may degrade with the existence of small-scale fading. To mitigate this potential rate loss, we introduce another supplement problem to adjust the beamformers in each channel realization. The two formulated problems are introduced subsequently. Note that although this approach is heuristic, the rate loss is insignificant when the UAVs' altitude is large since the path loss component would be the dominant factor ($K >> 1$ and $|\tilde{h}_{ijk}| \approx 1$).

3.3.2 User admission maximization

We first start with the user admission maximization problem. Let us introduce the admission decision variable $\mathbf{a} = \{a_{jk}, \forall j \in \mathcal{N}_k, \forall k \in \mathcal{K}\}$, where $a_{jk} = \{0,1\}$. $a_{jk} = 1$ enforces that the j-th user belonging to group k is admitted, and $a_{jk} = 0$ otherwise. The rate quality at each user depends on the beamforming vectors and path loss (or the UAVs' positions). Besides, the admission decision and the content placement, which influence the beamforming design, also affect the rate quality. As a result, our user admission maximization problem can be written as follows

$$\max_{\mathbf{a},\mathbf{b},\mathbf{w},\theta,\mathbf{p}} \quad \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{N}_k} a_{jk} \tag{3.11a}$$

$$\text{s.t.} \quad R_{jk}(\mathbf{w},\theta,\mathbf{b}) \geq a_{jk} R^k_{\min}, \forall j \in \mathcal{N}_k, \forall k \in \mathcal{K} \tag{3.11b}$$

$$\sum_{k\in\mathcal{K}} |b_{ik} w_{ik}|^2 \leq P_{\max}, \forall i \in \mathcal{U} \tag{3.11c}$$

$$a_{jk} = \{0,1\}, \forall j \in \mathcal{N}_k, \forall k \in \mathcal{K} \tag{3.11d}$$

$$b_{ik} = \{0,1\}, \forall i \in \mathcal{U}, \forall k \in \mathcal{K} \tag{3.11e}$$

$$\sum_{k\in\mathcal{K}} b_{ik} \leq S, \forall i \in \mathcal{U} \tag{3.11f}$$

$$\theta_{ijk}^{-1/n} \geq \tilde{\theta}_0 \sqrt{(x_{jk}-x_i)^2 + (y_{jk}-y_i)^2 + H^2}$$
$$, \forall i \in \mathcal{U}, \forall j \in \mathcal{N}_k, \forall k \in \mathcal{K} \tag{3.11g}$$

where constraint (3.11b) ensures a minimum rate for the admitted users, constraint (3.11c) presents the maximum power budget for each UAV, constraint (3.11f) implies that the total number of contents on each UAV should not exceed a maximum storage capacity S, constraint (3.11g) presents the relationship between the path loss and the UAV's position. It is noted that constraint (3.11g) is active and equivalent to (3.3) at the optimality for problem (3.11). Proof of this claim is similar to Appendix I in [85].

3.3.3 QoS maximization

After solving problem (3.11), the UAVs are deployed following the obtained optimal value $\mathbf{a}^* = \left\{a^*_{jk}, \forall j \in \mathcal{N}_k, \forall k \in \mathcal{K}\right\}$, content placement $\mathbf{b}^* = \left\{b^*_{ik}, \forall i \in \mathcal{U}, \forall k \in \mathcal{K}\right\}$, and UAVs' positions $\mathbf{p}^* = \{\mathbf{p}^*_i, \forall i \in \mathcal{U}\}$. However, with the presence of small-scale fading after deployment, the minimum required rate $R^k_{\min}$ may sometimes not be satisfied. Hence, the beamforming vectors can be adjusted at each channel realization according to the actual small-scale fading information, which can be obtained via pilot sequences. Let us denote $\mathbf{w}_n$ as the beamforming vectors at channel realization n, we introduce the following problem

$$\max_{\mathbf{w}_n} \quad \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{N}_k} R_{jk}(\mathbf{w}_n) \tag{3.12a}$$

$$\text{s.t.} \quad R_{jk}(\mathbf{w}_n) \geq a^*_{jk} \bar{R}^k_{\min}, \forall j \in \mathcal{N}_k, \forall k \in \mathcal{K} \tag{3.12b}$$

$$\sum_{k\in\mathcal{K}} \mid b^*_{ik} w_{ik,n} \mid^2 \leq P_{\max}, \forall i \in \mathcal{U} \tag{3.12c}$$

The purpose of problem (3.12) is to enhance the data rate at the admitted user by adjusting the beamforming $\mathbf{w}_n$ based on the instantaneous channel state information (CSI) $h_{ijk,n}(\theta_{ijk}, \tilde{h}_{ijk,n}) = \theta_{ijk}\tilde{h}_{ijk,n}$. However, problem (3.12) may sometimes be infeasible if the small-scale fading is severe. A countermeasure is to offset $R^k_{\min}$ in the initial problem by a small value χ^k, i.e., $R^k_{\min} = \bar{R}^k_{\min} + \chi^k$. For instance, if the actual target QoS requirement is 0.5 bps/Hz, we can set $R^k_{\min} = 0.55$ and $\chi^k = 0.05$. The value χ^k will depend on the Rician factor K, which presents the dominance of LoS components in the channel.

3.4 Solution Approach

First, we deal with problem (3.11). We propose the following equivalent transformation for problem (3.11) as follows

$$\max_{\substack{\mathbf{a},\mathbf{b},\mathbf{w},\\ \lambda,\theta,\mathbf{p}}} \quad \sum_{k\in\mathscr{K}}\sum_{j\in\mathscr{N}_k} a_{jk} \tag{3.13a}$$

$$\text{s.t.} \quad R_{jk}(\mathbf{w},\theta) \geq a_{jk}R^k_{\min}, \forall j \in \mathscr{N}_k, \forall k \in \mathscr{K} \tag{3.13b}$$

$$|\, w_{ik} \,|^2 \leq b_{ik}\lambda_{ik}, \forall i \in \mathscr{U}, \forall k \in \mathscr{K} \tag{3.13c}$$

$$\sum_{k\in\mathscr{K}} \lambda_{ik} \leq P_{\max}, \forall i \in \mathscr{U} \tag{3.13d}$$

$$(3.11d) - (3.11g)$$

Here, we introduce two new constraints (3.13c) and (3.13d) as a replacement for constraint (3.11c) to free the variables $\mathbf{b}$ from the rate function. Thus, the new rate function in (3.13b) can be written as $R_{jk}(\mathbf{w},\theta) = \log(1+\Gamma_{jk}(\mathbf{w},\theta))$ where $\Gamma_{jk}(\mathbf{w},\theta) = \frac{\left|(\mathbf{h}_{jk}(\theta_{jk}))^\dagger \mathbf{w}_k\right|^2}{\sum_{\ell\neq k}\left|(\mathbf{h}_{jk}(\theta_{jk}))^\dagger \mathbf{w}_\ell\right|^2+\sigma_0^2}$. It is worth noting that we also introduce a new variable $\lambda = \{\lambda_{ik}, \forall i \in \mathscr{U}, \forall k \in \mathscr{K}\}$ that indicates the soft-power level that the UAVs applied on the user group k (the optimized power when the binary variable b_{ik} is relaxed to continuous value). The constraints (3.13c) and (3.13d) convey the same meaning as constraint (3.11c). The reason behind our transformation is to smoothen the achievable rate function by removing the variable $\mathbf{b}$ as well as the dot-product operation. Since the new rate function, i.e., $R_{jk}(\mathbf{w},\theta)$, is now a smooth function, our later introduced DC-based approximation, which requires the derivation of gradient vectors, can be applied.

3.5 Low-complexity Algorithm

First and foremost, we acknowledge that problem (3.13) is non-convex because of the non-convexity of constraints (3.13b), (3.13c), (3.11g). In this section, we propose an algorithm developed on the basis of DC-programming and successive convex approximation (SCA) [86]. Since the rate function $R_{jk}(\mathbf{w},\theta)$ is not inherently presented as a difference of two convex functions, so-called

a DC function. We first apply DC decomposition to turn it into a DC form.

3.5.1 DC decomposition

$R_{jk}(\mathbf{w},\theta)$ can be written as

$$\underbrace{R_{jk}(\mathbf{w},\theta)+\xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2)}_{f_{jk}(\mathbf{w},\theta)}-\xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2),$$

An important observation is that if the parameter $\xi_{jk}>0$ is sufficiently large, function $f_{jk}(\mathbf{w},\theta)$ becomes convex with respect to $\mathbf{w}$ and θ due to the dominance of the strongly convex component $\xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2)$. Consequently, function $R_{jk}(\mathbf{w},\theta)$ turns into a DC function. Thus, we introduce the following theorem.

***Theorem* 5.** *For $\xi_{jk}\geq\xi_0$, where ξ_0 is derived in Appendix C, $f_{jk}(\mathbf{w},\theta)$ is strongly convex and $R_{jk}(\mathbf{w},\theta)$ becomes a DC function.*

Proof. Please refer to Appendix C. □

After applying the above decomposition, we can rewrite problem (3.13) as

$$\max_{\substack{\mathbf{a},\mathbf{b},\mathbf{w},\\ \lambda,\theta,\mathbf{p}}} \quad \sum_{j\in\mathcal{N}_k}\sum_{k\in\mathcal{K}} a_{jk} \tag{3.14a}$$

$$\text{s.t.} \quad f_{jk}(\mathbf{w},\theta)-\xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2)\geq a_{jk}R^k_{\min},\forall j\in\mathcal{N}_k,\forall k\in\mathcal{K} \tag{3.14b}$$

$$(3.11c)-(3.11g)$$

The left-hand side (LHS) of constraint (3.14b) now has DC form. Next, we introduce the DC-based approximation framework.

3.5.2 The DC-based approximation method

To describe the method, let us revisit the following DC-form constraint

$$f(\mathbf{x}) - g(\mathbf{x}) \geq 0 \tag{3.15}$$

where $f(\mathbf{x})$ and $g(\mathbf{x})$ are both convex with respect to variable $\mathbf{x}$. Note that in order for constraint (3.15) to characterize a feasible convex set, its left-hand side must be a concave function. Hence, the convex function $f(\mathbf{x})$ is the reason why constraint (3.15) is not convex. Assuming that $f(\mathbf{x})$ is differentiable, the first step of DC-based method is the linear approximation of $f(\mathbf{x})$ around the point $\mathbf{x}^{[\ell]}$ to convexify (3.15) into the following constraint

$$f_{\mathscr{T}}(\mathbf{x}^{[\ell]}) - g(\mathbf{x}) \geq 0 \tag{3.16}$$

(3.16) replaces $f(\mathbf{x})$ with its lower-bound first-order Taylor approximation $f_{\mathscr{T}}(\mathbf{x}^{[\ell]})$ around point $\mathbf{x}^{[\ell]}$. Hence, optimization problems with constraints following the form of (3.15) can be solved by iterative relaxation of non-convex constraints (3.15) into a convex approximation form as in (3.16). To apply this principle in our problem, we first acknowledge that problem (3.20) is nonconvex because of the convex function $f_{jk}(\mathbf{w},\theta)$ (with a sufficiently large value of ξ_{jk}) on the greater side of constraint (3.14b). Thus, we replace it by its first-order Taylor approximation $F_{jk}(\mathbf{w},\theta)$ around point $\mathbf{w}^{[\ell]}$ and $\theta^{[\ell]}$ as follows

$$F_{jk}(\mathbf{w},\theta;\mathbf{w}^{[\ell]},\theta^{[\ell]}) = f_{jk}(\mathbf{w}^{[\ell]},\theta^{[\ell]}) + \underbrace{\hat{f}_{jk}(\mathbf{w};\mathbf{w}^{[\ell]},\theta^{[\ell]}) + \check{f}_{jk}(\theta;\mathbf{w}^{[\ell]},\theta^{[\ell]})}_{\tilde{F}(\mathbf{w},\theta;\mathbf{w}^{[l]},\theta^{[\ell]})} - 2\xi_{jk}\underbrace{\mathrm{Re}\left((\mathbf{w}^{[\ell]})^{\dagger}\mathbf{w} - \left\|\mathbf{w}^{[\ell]}\right\|^2 + (\theta^{[\ell]})^{\dagger}\theta - \left\|\theta^{[\ell]}\right\|^2\right)}_{\tilde{F}(\mathbf{w},\theta;\mathbf{w}^{[\ell]},\theta^{[\ell]})}, \tag{3.17}$$

where

$$\hat{f}_{jk}(\mathbf{w};\mathbf{w}^{[\ell]},\theta^{[\ell]}) = \frac{2\mathrm{Re}\left((\mathbf{w}^{[\ell]})^{\dagger}\mathbf{H}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w} - (\mathbf{w}^{[\ell]})^{\dagger}\mathbf{H}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}^{[\ell]}\right)}{(\mathbf{w}^{[\ell]})^{\dagger}\mathbf{H}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}^{[\ell]} + \sigma_0^2} -$$

$$\frac{2\mathrm{Re}\left((\mathbf{w}^{[\ell]})^{\dagger}\tilde{\mathbf{H}}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}-(\mathbf{w}^{[\ell]})^{\dagger}\tilde{\mathbf{H}}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}^{[\ell]}\right)}{(\mathbf{w}^{[\ell]})^{\dagger}\tilde{\mathbf{H}}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}^{[\ell]}+\sigma_0^2}, \quad (3.18)$$

and

$$\breve{f}_{jk}(\theta;\mathbf{w}^{[\ell]},\theta^{[\ell]}) = \frac{2\mathrm{Re}\left((\theta^{[l]})^{\dagger}\Omega_{jk}\left(\mathbf{w}^{[\ell]}\right)\theta-(\theta^{[\ell]})^{\dagger}\Omega_{jk}\left(\mathbf{w}^{[\ell]}\right)\theta^{[l]}\right)}{(\mathbf{w}^{[\ell]})^{\dagger}\Omega_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}^{[\ell]}+\sigma_0^2} - \frac{2\mathrm{Re}\left((\theta^{[\ell]})^{\dagger}\tilde{\Omega}_{jk}\left(\mathbf{w}^{[\ell]}\right)\theta-(\theta^{[\ell]})^{\dagger}\tilde{\Omega}_{jk}\left(\mathbf{w}^{[\ell]}\right)\theta^{[\ell]}\right)}{(\mathbf{w}^{[\ell]})^{\dagger}\tilde{\Omega}_{jk}\left(\theta^{[\ell]}\right)\mathbf{w}^{[\ell]}+\sigma_0^2}, \quad (3.19)$$

The derivation of $\mathbf{H}_{jk}$, $\tilde{\mathbf{H}}_{jk}$, Ω_{jk}, and $\tilde{\Omega}_{jk}$ are as presented in Appendix C. After applying DC-based approximation, we obtain the following optimization problem:

$$\max_{\substack{\mathbf{a},\mathbf{b},\mathbf{w},\\ \lambda,\theta,\mathbf{p}}} \quad \sum_{j\in\mathcal{N}_k}\sum_{k\in\mathcal{K}} a_{jk} \quad (3.20a)$$

$$\text{s.t.} \quad F_{jk}(\mathbf{w},\theta;\mathbf{w}^{[\ell]},\theta^{[\ell]}) - \xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2)$$
$$\geq a_{jk}R^k_{\min}, \forall j\in\mathcal{N}_k, \forall k\in\mathcal{K} \quad (3.20b)$$
$$(3.13c), (3.13d), (3.11d)-(3.11g)$$

The above problem is still non-convex due to non-convex constraints (3.13c) and (3.11g) and the binary variables **a**, **b**. The tackle this, transformations and relaxations of these constraints and variables need to be applied.

3.5.3 Constraint transformation and relaxation.

First, we relax the binary variables **a** and **b** into continuous and replace (3.11d) and (3.11e) with the following constraints while still maintaining the equivalence

$$a_{jk}-a_{jk}^2 \leq 0, \quad (3.21)$$

$$0 \leq a_{jk} \leq 1, \quad (3.22)$$

$$b_{ik}-b_{ik}^2 \leq 0, \quad (3.23)$$

$$0 \leq b_{ik} \leq 1, \tag{3.24}$$

Since (3.21) and (3.23) are non-convex constraints, we replace the left-hand sides of them by their upper-bound first-order Taylor approximation around points $a_{jk}^{[\ell]}$ and $b_{jk}^{[\ell]}$ as follows

$$a_{jk} - (a_{jk}^{[\ell]})^2 - 2(a_{jk}^{[\ell]} a_{jk} - (a_{jk}^{[\ell]})^2) \leq 0, \tag{3.25}$$

$$b_{ik} - (b_{ik}^{[\ell]})^2 - 2(b_{ik}^{[\ell]} a_{ik,t} - (b_{ik}^{[\ell]})^2) \leq 0, \tag{3.26}$$

Thanks to the work [87], we acknowledge that the above constraints can render our DC-based iterative algorithm unable to compute a feasible solution when coupled with constraints (3.21) and (3.23). We make sure $a_{jk}^{[\ell]}$ and $b_{ik}^{[\ell]}$ are always feasible by introducing two slack variables μ_{jk} and ν_{ik} as follows

$$a_{jk} - (a_{jk}^{[\ell]})^2 - 2(a_{jk}^{[\ell]} a_{jk} - (a_{jk}^{[\ell]})^2) \leq \mu_{jk}, \tag{3.27}$$

$$b_{ik} - (b_{ik}^{[\ell]})^2 - 2(b_{ik}^{[\ell]} b_{ik,t} - (b_{ik}^{[\ell]})^2) \leq \nu_{ik}, \tag{3.28}$$

To maintain the equivalence of our problem, the variables μ_{jk} and ν_{ik}, associated with some penalty parameters, will be deducted from the objective function and the optimality of the problem occurs when μ_{jk} and ν_{ik} approach 0.

Next, with simple algebraic transformations, we can transform the constraint (3.13c) into the following

$$\sqrt{|\, w_{ik} \,|^2 + \frac{(b_{ik} - \lambda_{ik})^2}{4}} \leq \frac{b_{ik} + \lambda_{ik}}{2}, \tag{3.29}$$

After the transformation, it is obvious that (3.29) is the second-order cone representation of constraint (3.13c). Finally, to deal with constraint (3.11g), we replace the left-hand side by its lower-bound first-order Taylor approximation around point $\theta_{ijk}^{[\ell]}$ as follows

$$\theta_{ijk}^{[\ell]} - \frac{1}{\alpha} \theta^{\frac{-1-n}{n}} (\theta - \theta^{[\ell]}) \geq \tilde{\theta}_0 \sqrt{(x_{jk} - x_i)^2 + (y_{jk} - y_i)^2 + H^2}, \tag{3.30}$$

Let us denote $\tilde{\mathbf{M}} = \{\mu_{jk}, \forall j \in \mathscr{N}_k, \forall k \in \mathscr{K}\}$, $\tilde{\mathbf{N}} = \{\nu_{ik}, \forall i \in \mathscr{U}, \forall k \in \mathscr{K}\}$. By applying the above relaxation and approximation, we obtain the following convex optimization problem at ℓ-th iteration:

$$\max_{\substack{\mathbf{a},\mathbf{b},\mathbf{w},\\ \lambda,\theta,\mathbf{p}}} \quad \sum_{k\in\mathscr{K}}\sum_{j\in\mathscr{N}_k} a_{jk} - M^{[\ell]} \sum_{k\in\mathscr{K}}\sum_{j\in\mathscr{N}_k} \mu_{jk} - N^{[\ell]} \sum_{k\in\mathscr{K}}\sum_{i\in\mathscr{U}} \nu_{ik} \tag{3.31a}$$

$$\text{s.t.} \quad F_{jk}(\mathbf{w},\theta;\mathbf{w}^{[\ell]},\theta^{[\ell]}) - \xi_{jk}(\|\mathbf{w}\|^2 + \|\theta\|^2) \geq a_{jk} R^k_{\min}, \forall j \in \mathscr{N}_k \tag{3.31b}$$

$$\sum_{k\in\mathscr{K}} b_{ik} \leq S, \forall i \in \mathscr{U} \tag{3.31c}$$

$$\sqrt{|\, w_{ik}\,|^2 + \frac{(b_{ik}-\lambda_{ik})^2}{4}} \leq \frac{b_{ik}+\lambda_{ik}}{2}, \forall k \in \mathscr{K}, \forall i \in \mathscr{U} \tag{3.31d}$$

$$\sum_{k\in\mathscr{K}} \lambda_{ik} \leq P_{\max}, \forall i \in \mathscr{U} \tag{3.31e}$$

$$0 \leq a_{jk} \leq 1, \forall j \in \mathscr{N}_k, \forall k \in \mathscr{K} \tag{3.31f}$$

$$a_{jk} - (a^{[\ell]}_{jk})^2 - 2(a^{[\ell]}_{jk} a_{jk} - (a^{[\ell]}_{jk})^2) \leq \mu_{jk}, \forall j \in \mathscr{N}_k, \forall k \in \mathscr{K} \tag{3.31g}$$

$$0 \leq b_{jk,t} \leq 1, \forall i \in \mathscr{U}, \forall k \in \mathscr{K} \tag{3.31h}$$

$$b_{ik} - (b^{[\ell]}_{ik})^2 - 2(b^{[\ell]}_{ik} b_{ik,t} - (b^{[\ell]}_{ik})^2) \leq \nu_{ik}, \forall i \in \mathscr{U}, \forall k \in \mathscr{K} \tag{3.31i}$$

$$\tag{3.31j}$$

$$\theta^{[\ell]}_{ijk} - \frac{1}{\alpha}\theta^{(-1-n)/n}(\theta - \theta^{[\ell]}) \geq \tilde{\theta}_0 \sqrt{(x_{jk}-x_i)^2 + (y_{jk}-y_i)^2 + H^2}, \forall j \in \mathscr{N}_k, \forall k \in \mathscr{K} \tag{3.31k}$$

Note that in problem (3.31) the slack variables μ_{jk} and ν_{ik}, coupled with the penalty parameters $M^{[\ell]}$ and $N^{[\ell]}$, have been deducted from the original objective function. It is obvious that when μ_{jk} and ν_{ik} approach 0, the variables a_{jk} and b_{ik} converge to either 0 or 1. The values $M^{[\ell]}$ and $N^{[\ell]}$ are initiated at some small positive values and are updated by multiplying with values $\varepsilon_M, \varepsilon_N > 1$

to ensure that $\sum_{k\in\mathcal{K}}\sum_{j\in\mathcal{N}_k}\mu_{jk}$ and $\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{U}}\nu_{ik}$ approach 0 when $M^{[\ell]}$ and $N^{[\ell]}$ approach some large values $M_{\max}$ and $N_{\max}$.

At the ℓ-th iteration, the initial mixed-integer non-convex problem (3.11) has been approximated into a continuous second-order cone program (SOCP), which can be handled by modern optimization solvers [88]. Thus, the SCA method, which is widely known in the wireless communication and signal processing community, can be applied. Algorithm 2 outlines our method to solve problem (3.11). After solving problem (3.11), values $\mathbf{a}^*$, $\mathbf{b}^*$,$\mathbf{p}^*$ can be used to proceed to solve problem

Algorithm 2 DC-based SCA method

1: Initialize starting values of $\mathbf{a}^{[\ell]}$, $\mathbf{b}^{[\ell]}$, $\mathbf{w}^{[\ell]}$, $\theta^{[\ell]}$,$\mathbf{p}^{[\ell]}$, $M^{[\ell]}$, and $N^{[\ell]}$
2: Set $\ell := 0$;
3: **repeat**
4: Solve the convex problem (3.31) to obtain $\mathbf{a}^*$, $\mathbf{b}^*$, $\mathbf{w}^*$, θ^*,p^*;
5: Set $\ell := \ell + 1$;
6: Update $\mathbf{a}^{[\ell]} = \mathbf{a}^*$, $\mathbf{b}^{[\ell]} = \mathbf{b}^*$, $\mathbf{w}^{[\ell]} = \mathbf{w}^*$, $\theta^{[\ell]} = \theta^\star$,$\mathbf{p}^{[\ell]}$=$\mathbf{p}^*$;
7: Update $M^{[\ell]} = \min\{\varepsilon_M M^{[\ell-1]}, M_{\max}\}$ and $N^{[\ell]} = \min\{\varepsilon_N N^{[\ell-1]}, N_{\max}\}$;
8: **until** Convergence of the objective function (3.31a); $\mathbf{a}^* = \mathbf{a}^{[\ell]}$, $\mathbf{b}^* = \mathbf{b}^{[\ell]}$, $\mathbf{w}^* = \mathbf{w}^{[\ell]}$, $\theta^* = \theta^{[\ell]}$,$\mathbf{p}^* = \mathbf{p}^{[\ell]}$

(3.12). Since problem (3.12) is continuous and non-convex due to the function $R_{jk}(\mathbf{w}_n)$, we can apply DC decomposition and SCA the same as with function $R_{jk}(\mathbf{w},\theta)$ in problem (3.11) (θ is a now constant since we know p^*).

3.6 Numerical Results and Discussion

In this section, we present numerical results to evaluate the performance of our proposed scheme and draw some engineering insights from adjusting the system parameters. We also compare our scheme with non-cooperative power allocation scheme [78] and joint beamforming [82]. We first solve problem (3.11) with a value $R^k_{\min}$ greater than the actual minimum required rate $\bar{R}^k_{\min}$. Then, problem (3.13) is solved with value $\bar{R}^k_{\min}$ to determine the optimal beamforming at each channel realization. However, it is noted that $\bar{R}^k_{\min}$ and $R^k_{\min}$ can be roughly the same, especially when the Rician factor K is large.

Settings: we consider a wireless communication system with a circular coverage of radius $r =$ 100 m. There are 20 ground users that are randomly distributed within the circle. These users

request 6 different contents whose request probability follow the Zipf distribution. Unless otherwise stated, $\theta_0 = -50$ dB, n=2.3 , K= 12 dB, S=3, $U = 6$, $H = 100$ m and $\alpha = 0.6$. We assume the flexible LTE bandwidth of $W = 20$ MHz [89]. Thus, the AWGN noise power is $\sigma_0^2 = -174 + 10\log(W) \approx -100$ dBm since the UAVs are operating on the whole available bandwidth. The number of admitted users is averaged out over 100 randomly generated small-scale fading and user distributions and is rounded to the nearest integer. Note that although the users' QoS requirements are usually different from each other in practice, we consider the same minimum QoS requirement for all users for the sake of fair evaluation. For instance, if we set the requirement of a user group very low and the others very high, the system may only admit the group with the lower requirements and rejects the others despite varying the parameters.

In Fig. 3.3, we investigate the system performance under five scenarios with different (minimum) QoS requirements. Specifically, the QoS requirements, i.e., $\bar{R}^k_{\min}$, in scenarios 1, 2, 3, 4 and 5 is 0.1, 0.2, 0.4, 0.6 and 0.8 bps/Hz, respectively. Note that the QoS requirements have been normalized by W and the bitrate (in bps) is $W\bar{R}^k_{\min}$. Thus, the users admitted in scenarios 1, 2, 3, 4 and 5 can experience the bitrates 2, 4, 8, 12, and 16 Mbps, respectively. These bitrates can provide high-quality videos. For instance, the bitrates in scenarios 2, 3 and 4 are higher than the required rates for 480p, 720p and 1080p video quality, respectively [90].[3] As illustrated, the more we increase the minimum rate requirements, the fewer users the system can admit considering the same power budget. We also notice that the number of admitted users can be 0 when the power budget is low. The reason is the effect of path loss caused by the high-altitude deployment of UAVs makes the received signal very weak. In addition, increasing the power budget on UAVs leads to minor improvements on user admission as the QoS requirements become higher. For example, in scenario 5, no user is admitted at 18 dBm and only 3 are admitted at 28 dBm, which is 10-time higher power consumption. This is due to the following reasons. First, the received signals at the users experience significant path loss due to the UAVs' altitude and their own distance from the users. We will later show that increasing the number of UAVs can mitigate this trend since the UAVs can move closer to the users, especially those at the edge. Second, when the UAVs cannot support more users given their power budget, the power is reserved to increase the rate of the admitted users due to problem

[3]The redundant bandwidth can be reserved for cyclic prefix, signaling overhead, and pilot overhead.

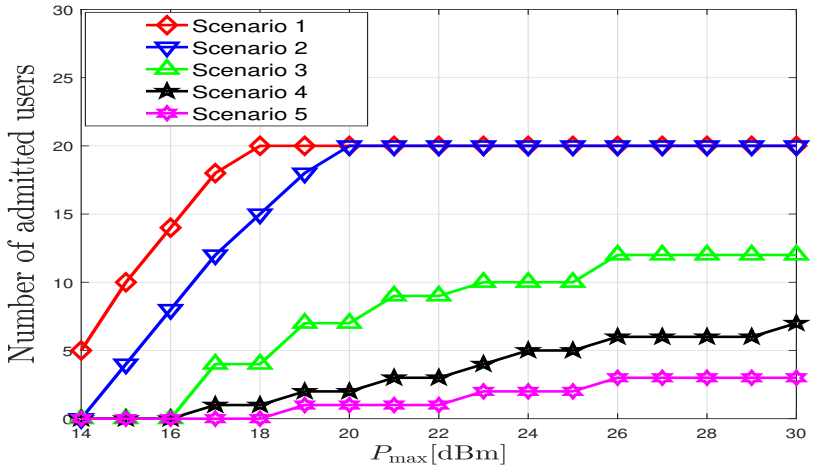


Figure 3.3: Number of admitted users versus the maximum power budget per UAV with different QoS requirements

(3.13).

As demonstrated in Fig. 3.3, increasing power is inefficient to enhance user admission. Thus, in Fig. 3.4, we increase the number of UAVs and see what is the minimum overall power budget (which is divided equally to each UAV) to admit all 20 users in each case. The QoS requirement is the same as Scenario 3 in Fig. 3.3. It is worth mentioning that the power budget in Fig. 3.4 is unrealistically increased up approximately 68 dBm per UAVs (70 dBm in total) for the sole purpose of highlighting the relationship between the number of UAVs and the power needed to admit all users. Fig. 3.4 shows that by adding 5 UAVs, the required total power can be reduced by 40 dBm (10^4 times), which is very significant. This is due to the following reasons. First, increasing the number of UAVs allows more UAVs to cooperate and boost received signal strength. Second, with more UAVs, each one has more freedom to fly closer to the users to significantly reduce the path loss. We also notice that this trend slows down gradually as we decrease the total power budget since after a certain point, the cooperation cannot overcome a fixed path-loss induced by the altitude $H = 100$ m when the overall power budget becomes too small. This observation offers a notable insight that more UAVs with lower power budget working in a cooperative manner consumes much less power than fewer UAVs high power to achieve the same user admission. Thus, employing more cooperative UAVs may to improve UAVs' lifetime.

In Fig. 3.5, we consider a fixed total storage capacity of 18 contents while deploying different

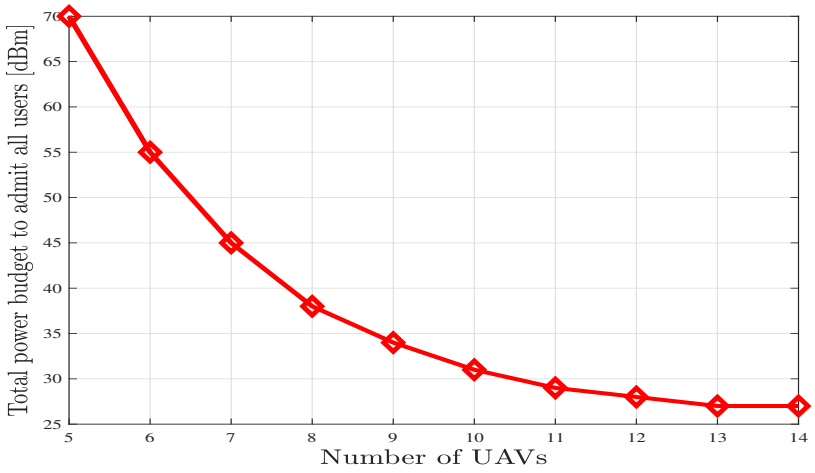


Figure 3.4: Minimum total power budget to admit all users versus different numbers of cooperative UAVs.

numbers of UAVs. The QoS requirement is the same as Scenario 3 in Fig. 3.3. As illustrated, deploying more UAVs with lower storage capacity on each UAV results in higher user admission. Thus, when considered as a resource, storage capacity plays a similar role to power budget per UAVs. Thus, from a more general perspective, Overall, Fig. 3.4 and Fig. 3.5 reveal that UAVs with limited capability (in terms of payload and power) can cooperatively work together and still outperform UAVs with much higher capability.

In Fig. 3.6, we observe the effect of storage capacity on the system performance while keeping the number of UAVs unchanged. The QoS requirement is as in scenario 3 of Fig. 3.3. Similar to the impact of the number of UAVs, more storage capacity results in higher user admission. The reason is that when we have more storage capacity, more UAVs can store the duplicate versions of the same contents and thus increase the cooperation among UAVs.

In Fig. 3.7, we draw the comparison among different interference mitigation schemes. The QoS' requirements are as in scenario 3 of Fig. 3.3. In the power allocation scheme, each UAV is assigned to serve each user group. We optimized the power that each UAV allocates to its corresponding user group and the UAVs' positions so as to mitigate inter-group interference. Some drawbacks of the power allocation scheme are considering no cooperation among UAVs based on content storage and only taking into consideration the amplitude of the transmitted signal while neglecting the phase components. Thus, the power allocation scheme performs poorly when the users are scattered. For

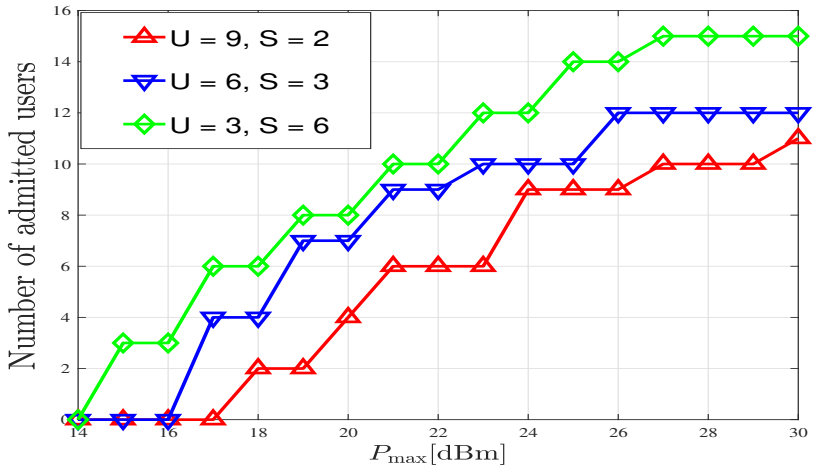


Figure 3.5: Number of admitted users versus power budget per UAV with the same total storage space and different number of UAVs.

the joint beamforming scheme [82], we also employ beamforming and location optimization as in our scheme. However, in this case, all the UAVs serve all the users via joint transmissions. As demonstrated in Fig 3.4, the system performance is supposed to be enhanced when more UAVs can serve one user group. However, this is only true when we have enough storage to store most of the contents to be requested (ideally all contents). Since we are considering a practical scenario with limited storage capacity, the UAVs can only store the most popular contents and the number of admitted users is upper-bounded by the total number of users requesting these contents. As illustrated, our proposed scheme outperforms the others since we are considering signal-level coordination (beamforming), UAVs' locations, and strategic content placement with limited storage capacity per UAV simultaneously.

Fig. 3.8 shows the impact of different values of α in our system. The QoS requirements are as in scenario 3 of Fig. 3.3. The value α, which has been presented in section 3.2.2, represents the degree that the content request distribution is concentrated on the most popular contents. In other words, the larger α, the larger the number of users requests the most common contents. As shown in Fig. 3.8, our system performance increases with α. Since there are more users requesting the most popular contents, it is more likely the contents are loaded in the UAVs as our objective is to maximize the number of served users. Thus, there are more copies of the same contents in the UAVs (constrained by the storage capacity), which enhances UAVs' cooperation. This trend has

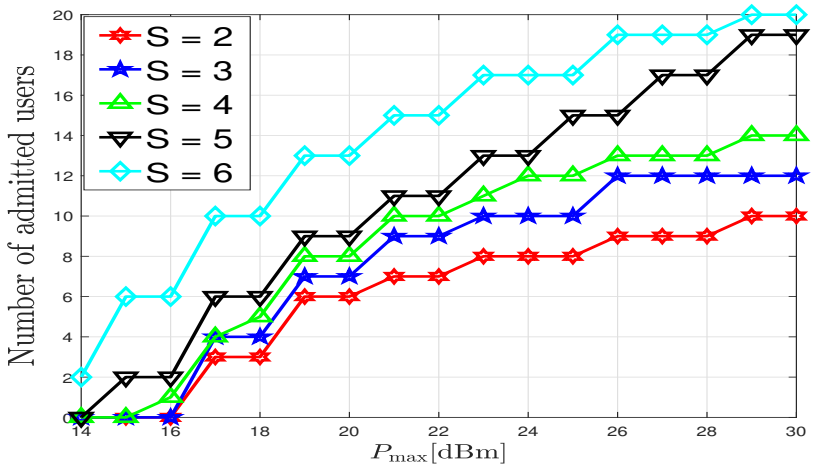


Figure 3.6: Number of admitted users versus power budget per UAV with the same total storage space and different number of UAVs.

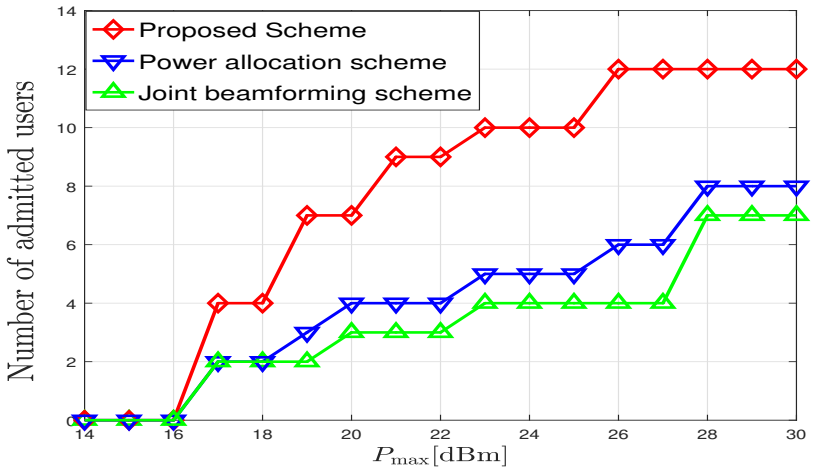


Figure 3.7: Number of admitted users versus the maximum power budget per UAV with different interference management scheme.

been referred to as higher *content reuse* in previous literature [5].

3.7 Conclusion

In this chapter, we have proposed and investigated a novel network architecture enabled by multiple content-aware UAVs. By leveraging the new promising technology of UAVs and the existing and established techniques such as distributed antenna system and multicast beamforming, we utilize a system of multiple content-aware UAVs with limited storage capacity that can cooperate with

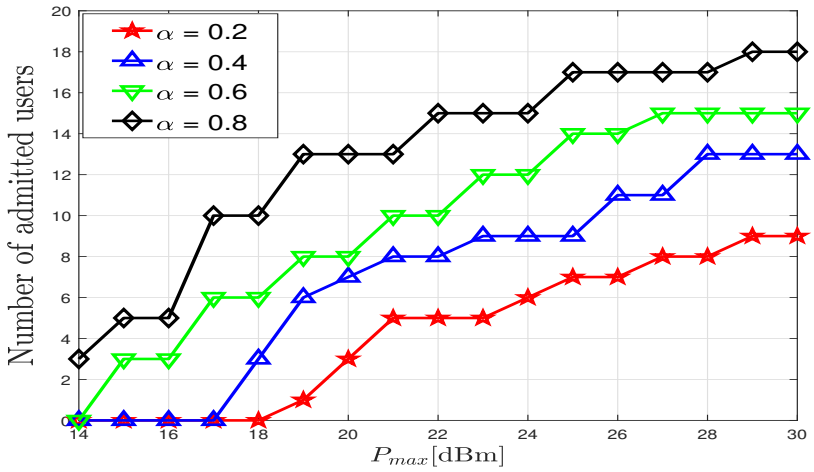


Figure 3.8: Number of admitted users versus the maximum power budget per UAV with different values of α.

each other to enhance the system performance. Further, by acknowledging the relationship among the UAVs' storage capacity, their locations, and their capability to cooperate, we formulate a novel optimization problem to maximize the user admission, taking into consideration some practical constraints. Since our formulated problem is a mixed-integer non-convex optimization problem whose search of the optimal solution cannot be guaranteed a polynomial time, we proposed a framework based on difference-of-convex programming to approximate it into a series of mixed-integer second-order cone programs and iteratively solve them until convergence. Our numerical results reveal the superiority of our scheme compared to previous work from the literature and offer notable insights into the system operation as a function of a wide range of design parameters.

Chapter 4

Final Conclusions and Future Work

In this thesis, we have investigated the performance of cooperative communications Incorporated by two states-of-the-art technologies, namely UAVs and NOMA. It is clear that when we endow conventional NOMA-enabled and UAV-enabled wireless networks with cooperation capability among multiple transmission and reception nodes, the network achieve higher performance gain than without cooperation. We have also analyzed why cooperative communication benefits in the specific networks scenarios in the Conclusion sections of Chapter 2 and Chapter 3. As a final conclusion of the thesis, we will discuss some of the limitation of our contributions as well as propose some promising extension for future work.

Some of the limitations and our proposed future work are as follows

- **Overhead:** coordination among multiple nodes requires tremendous information exchange on control channels, which may be costly resource-wise. An example of this is the resource allocated to channel estimation and CSI acquisitions and exchanging the information among BSs to determine the optimal beamforming vectors in Chapter 3. As the size of the network increases, the information required to coordinate all the nodes increase and more resource will be allocated to overhead than the actual payload data. This has not been discussed enough in this thesis. The key to overhead problem is to have a framework to reduce the degree of centralization since full coordination is not practical and is often not necessary. Hence, our future work will focus on an effective framework for our proposed problems but with partial

coordination due to overhead constraints.

- **Energy:** we have concluded in the previous chapter that cooperation can reduce the communication energy when we employ more nodes since this enhances the macro-diversity of the systems. However, as we employ more coordinated nodes, the cost to install and operate these nodes also increases energy usage. Thus, a more comprehensive energy model that captures different types of energy usage needs to be studied to practically evaluate the energy efficiency of cooperative communication. Furthermore, in C-NOMA, we utilize mobile users relaying nodes, which can drain their energy quickly. To this end, one solution is to complement/replace the users with intelligent reflecting surfaces (IRSs), which do not require to need energy emission to relay the transmitted signals [91].

Bibliography

[1] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, Sept 2018.

[2] S. R. Islam *et al.*, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, Oct. 2017.

[3] R. Irmer *et al.*, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[4] T. M. Nguyen, A. Yadav, W. Ajib, and C. Assi, "Resource allocation in two-tier wireless backhaul heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6690–6704, Oct. 2016.

[5] M. Tao, E. Chen, H. Zhou, and W. Yu, "Content-centric sparse multicast beamforming for cache-enabled cloud RAN," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6118–6131, Sep 2016.

[6] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, Nov. 2015.

[7] C. V. Networking, "Cisco global cloud index: Forecast and methodology, 2016–2021," White Paper, Cisco, San Jose, CA, USA Nov. 2019.

[8] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. on selected areas in commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[9] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. on selected areas in commun.*, vol. 35, no. 6, pp. 1201–1221, Apr. 2017.

[10] P. Pirinen, "A brief overview of 5G research activities," in *Proc. IEEE 5GU*, Akaslompolo, Finland, Nov. 2014.

[11] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, Jun. 2015.

[12] M. R. Palattella *et al.*, "Internet of things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, Feb. 2016.

[13] D. Tsonev, S. Videv, and H. Haas, "Towards a 100 Gb/s visible light wireless access network," *Optics express*, vol. 23, no. 2, pp. 1627–1637, Jan. 2015.

[14] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Commun. Magazine*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[15] X. Ge, S. Tu, G. Mao, C.-X. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 72–79, Mar. 2016.

[16] X. Zhang and M. Haenggi, "The performance of successive interference cancellation in random wireless networks," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 6368–6388, Oct 2014.

[17] L. Liu, C. Yuen, Y. L. Guan, and Y. Li, "Capacity-achieving iterative LMMSE detection for MIMO-NOMA systems," in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May. 2016.

[18] A. Li *et al.*, "Investigation on hybrid automatic repeat request (HARQ) design for NOMA with SU-MIMO," in *Proc. IEEE PIMRC*, Hong Kong, China, Sept. 2015.

[19] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, M. Elkashlan, and J. Alonso-Zarate, "NOMA-based D2D communications: Towards 5G," in *Proc. IEEE Globecom*, Washington, DC, USA, Dec. 2016.

[20] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, February 2017.

[21] X. Song, L. Dong, J. Wang, L. Qin, and X. Han, "Energy efficient power allocation for downlink NOMA heterogeneous networks with imperfect CSI," *IEEE Access*, vol. 7, pp. 39 329–39 340, 2019.

[22] J. Chen, J. Jia, Y. Liu, X. Wang, and A. H. Aghvami, "Optimal resource block assignment and power allocation for D2D-enabled NOMA communication," *IEEE Access*, vol. 7, pp. 90 023–90 035, 2019.

[23] K. Wang, Y. Liu, Z. Ding, A. Nallanathan, and M. Peng, "User association and power allocation for multi-cell non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2019.

[24] M. Zeng, A. Yadav, O. A. Dobre, and H. V. Poor, "Energy-efficient joint user-rb association and power allocation for uplink hybrid noma-oma," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5119–5131, June 2019.

[25] S. M. R. Islam, M. Zeng, O. A. Dobre, and K. Kwak, "Resource allocation for downlink NOMA systems: Key techniques and open issues," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 40–47, April 2018.

[26] W. Liang *et al.*, "User Pairing for Downlink Non-Orthogonal Multiple Access Networks Using Matching Algorithm," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5319–5332, Aug. 2017.

[27] L. Zhu, J. Zhang, Z. Xiao, X. Cao, and D. O. Wu, "Optimal User Pairing for Downlink Non-Orthogonal Multiple Access (NOMA)," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 328–331, Apr. 2019.

[28] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G non orthogonal multiple-access downlink transmissions," *IEEE Trans. on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, Sep. 2016.

[29] 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.

[30] L. Zhu, J. Zhang, Z. Xiao, X. Cao, and D. O. Wu, "Optimal user pairing for downlink non-orthogonal multiple access (noma)," *IEEE Wireless Communications Letters*, vol. 8, no. 2, April 2019.

[31] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Letters*, vol. 19, no. 8, pp. 1462–1465, Jun. 2015.

[32] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Lin. Alg. and its Applications*, vol. 284, pp. 193–228, Jan. 1998.

[33] F. Kara and H. Kaya, "On the error performance of cooperative-NOMA with statistical CSIT," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 128–131, Jan 2019.

[34] 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, June 2019.

[35] 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. on Wireless Commun.*, vol. 17, no. 1, pp. 506–519, Jan. 2018.

[36] B. Ning, W. Hao, A. Zhang, J. Zhang, and G. Gui, "Energy efficiencyâĂŞdelay tradeoff for a cooperative NOMA system," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 732–735, April 2019.

[37] T. E. A. Alharbi and D. K. C. So, "Full-duplex decode-and-forward cooperative non-orthogonal multiple access," in *Proc. IEEE VTC*, Porto, Portugal, Jun. 2018.

[38] N. Guo, J. Ge, Q. Bu, and C. Zhang, "Multi-user cooperative non-orthogonal multiple access scheme with hybrid full/half-duplex user-assisted relaying," *IEEE Access*, vol. 7, pp. 39 207–39 226, 2019.

[39] Y. Zhou, V. W. S. Wong, and R. Schober, "Dynamic decode-and-forward based cooperative noma with spatially random users," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3340–3356, May 2018.

[40] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, April 2016.

[41] M. Obeed *et al.*, "User pairing, link selection and power allocation for cooperative NOMA hybrid VLC/RF systems," *CoRR*, vol. abs/1908.10803, 2019. [Online]. Available: http://arxiv.org/abs/1908.10803

[42] J. Kaleva, A. TÃűlli, and M. J. Juntti, "Decentralized sum rate maximization with QoS constraints for interfering broadcast channel via successive convex approximation," *IEEE Trans. Signal Processing*, to appear.

[43] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, July 2012.

[44] K. R. Kumar and G. Caire, "Coding and decoding for the dynamic decode and forward relay protocol," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3186–3205, July 2009.

[45] M. Bossert, *Channel Coding for Telecommunications*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 1999.

[46] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec 2004.

[47] 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. on Wireless Commun.*, vol. 17, no. 1, pp. 506–519, Jan. 2018.

[48] S. Roy *et al.*, "Maximal-ratio combining architectures and performance with channel estimation based on a training sequence," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1154–1164, July 2004.

[49] J. Tang and X. Zhang, "Cross-layer resource allocation over wireless relay networks for quality of service provisioning," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 4, pp. 645–656, May 2007.

[50] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, Jun. 2001. [Online]. Available: http://dx.doi.org/10.1023/A:1017501703105

[51] J. F. Bard, *Practical bilevel optimization: algorithms and applications*. Springer Science & Business Media, 2013, vol. 30.

[52] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Annals of Operations Research*, vol. 153, pp. 235–256, 2007.

[53] Y. Li, T. Wang, Z. Zhao, M. Peng, and W. Wang, "Relay mode selection and power allocation for hybrid one-way/two-way half-duplex/full-duplex relaying," *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1217–1220, July 2015.

[54] S. Ebbesen, P. Kiwitz, and L. Guzzella, "A generic particle swarm optimization matlab function," in *2012 American Control Conference (ACC)*. IEEE, 2012, pp. 1519–1524.

[55] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter wave communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1616–1653, June 2018.

[56] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, February 2014.

[57] D. LÃşpez-PÃl'rez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2078–2101, June 2015.

[58] A. A. Khuwaja, Y. Chen, N. Zhao, M. Alouini, and P. Dobbins, "A survey of channel modeling for UAV communications," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2804–2821, Jul 2018.

[59] M. F. Sohail, C. Y. Leow, and S. Won, "Non-orthogonal multiple access for unmanned aerial vehicle assisted communication," *IEEE Access*, vol. 6, pp. 22 716–22 727, 2018.

[60] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 328–331, June 2018.

[61] T. Bai, J. Wang, Y. Ren, and L. Hanzo, "Energy-efficient computation offloading for secure UAV-edge-computing systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6074–6087, June 2019.

[62] S. Zeng, H. Zhang, K. Bian, and L. Song, "UAV relaying: Power allocation and trajectory optimization using decode-and-forward protocol," in *Proc. IEEE Int. Conf. Communications (ICC'18) Workshops*, 2018, pp. 1–6.

[63] X. Xu, Y. Zeng, Y. L. Guan, , and R. Zhang, "Overcoming endurance issue: UAV-enabled communications with proactive caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1231–1244, Jun 2018.

[64] U. Challita, W. Saad, and C. Bettstetter, "Interference management for cellular-connected UAVs: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2125–2140, April 2019.

[65] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a UAV-mounted cloudlet: Optimization of bit allocation and path planning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2049 – 2063, May 2017.

[66] J. Liu, X. Wang, B. Bai, and H. Dai, "Age-optimal trajectory planning for UAV-assisted data collection," in *Proc. IEEE Conf. on Computer Commun. (INFOCOM 2018) Workshops*, Honolulu,HI, Dec. 2016.

[67] M. Alzenad, A. El-Keyi, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station for maximum coverage of users with different QoS requirements," *IEEE Commun. Lett.*, vol. 7, no. 1, pp. 38 – 41, Feb 2018.

[68] V. Sharma, R. Sabatini, and S. Ramasamy, "UAVs assisted delay optimization in heterogeneous wireless networks," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2526 – 2529, Dec. 2016.

[69] Y. Chen *et al.*, "Optimum placement of UAV as relays," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 248–251, Feb 2018.

[70] M. M. Azari, F. Rosas, K.-C. Chen, and S. Pollin, "Joint sum-rate and power gain analysis of an aerial base station," in *Proc. IEEE Global Communications Conference (GLOBECOM 2016) Workshops*, Washington, DC, Dec. 2016.

[71] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3747–3760, Jun 2017.

[72] Y. Zeng, X. Xu, and R. Zhang, "Trajectory optimization for completion time minimization in UAV-enabled multicasting," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2233 – 2246, Jan. 2018.

[73] N. Zhao *et al.*, "Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2281 – 2294, May 2018.

[74] D. Jiang and Y. Cui, "Enhancing performance of random caching in large-scale wireless networks with multiple receive antennas," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2051–2065, April 2019.

[75] N. Rupasinghe, Y. YapÄścÄś, Ä. GÃijvenÃğ, and Y. Kakishima, "Non-orthogonal multiple access for mmWave drone networks with limited feedback," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 762–777, Jan 2019.

[76] H. Pan, S. C. Liew, J. Liang, Y. Shao, and L. Lu, "Network-coded multiple access on unmanned aerial vehicle," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2071–2086, Sep. 2018.

[77] A. A. Nasir *et al.*, "UAV-enabled communication using NOMA," *IEEE Trans. Commun.*, 2019.

[78] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2109–2121, Mar. 2018.

[79] L. Wang and S. Zhou, "Energy-efficient UAV deployment with flexible functional split selection," in *Proc. IEEE 19th International Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Anaheim, CA, June 2018.

[80] L. Liu, S. Zhang, and R. Zhang, "CoMP in the sky: UAV placement and movement optimization for multi-user communications," *IEEE Trans. Commun.*, Mar. 2019.

[81] M. Amade *et al.*, "Information-centric networking for the internet of things: challenges and opportunities," *IEEE Netw.*, vol. 30, no. 2, pp. 92–100, Mar 2016.

[82] P. Dinh, T. M. Nguyen, C. Assi, and W. Ajib, "Joint beamforming and location optimization for cooperative content-aware UAVs," in *Proc. IEEE Wireless Comm. Net. Conf. (WCNC'19)*, Marrakech, Morocco, Apr. 2019, pp. 1–7.

[83] H. He, S. Zhang, Y. Zeng, and R. Zhang, "Joint altitude and beamwidth optimization for UAV-enabled multiuser communications," *IEEE Wireless Commun. Lett.*, vol. 22, no. 2, pp. 344–347, Feb. 2018.

[84] Q. Yuan, Y. Hu, C. Wang, and Y. Li, "Joint 3D beamforming and trajectory design for UAV-enabled mobile relaying system," *IEEE Access*, vol. 7, pp. 26 488–26 496, 2019.

[85] T. M. Nguyen, A. Yadav, W. Ajib, and C. Assi, "Resource allocation in two-tier wireless backhaul heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6690–6704, Oct 2016.

[86] B. R. Marks and G. P. Wright, "Technical note–A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, pp. 681–683, Aug 1978.

[87] Q.-D. Vu, K.-G. Nguyen, and M. Juntti, "Max-min fairness for multicast multigroup multicell transmission under backhaul constraints," in *GLOBECOM 2016*, Washington, DC, Dec. 2016, pp. 1–6.

[88] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.*, 2019. [Online]. Available: http://docs.mosek.com/9.0/toolbox/index.html

[89] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-advanced: next-generation wireless broadband technology [invited paper]," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 10–22, June 2010.

[90] H. Riiser, "Adaptive bitrate video streaming over HTTP in mobile wireless networks," Ph.D. dissertation, University of Oslo, Norway, 2013.

[91] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov 2019.

Appendix A

Power control for HD C-NOMA

In this section, we present the proof of theorems 1 and 2. At first, note that for any value $\alpha \in [0,1]$, constraints (2.18d) and (2.18e) can be equivalently transformed into the following inequalities.

$$\max\left(A_{m,n}^{\mathrm{H}}, B_{m,n}^{\mathrm{H}}\left(P_{m,n}^{\mathrm{d}}\right)\right) \leq \alpha \leq C_{m,n}^{\mathrm{H}}, \tag{A.1}$$

where

$$\begin{aligned} A_{m,n}^{\mathrm{H}} &= \frac{\delta_{\mathrm{th}}^{\mathrm{H}}(P_{\mathrm{BS}}\gamma_m + 1)}{P_{\mathrm{BS}}\gamma_m(\delta_{\mathrm{th}}^{\mathrm{H}} + 1)}, \\ B_{m,n}^{\mathrm{H}}\left(P_{m,n}^{\mathrm{d}}\right) &= \frac{(P_{\mathrm{BS}}\gamma_n + 1)(\delta_{\mathrm{th}}^{\mathrm{H}} - P_{m,n}^{\mathrm{d}}\gamma_{m,n}^{\mathrm{d}})}{P_{\mathrm{BS}}\gamma_n(\delta_{\mathrm{th}}^{\mathrm{H}} + 1 - P_{m,n}^{\mathrm{d}}\gamma_{m,n}^{\mathrm{d}})}, \\ C_{m,n}^{\mathrm{H}} &= 1 - \frac{\delta_{\mathrm{th}}^{\mathrm{H}}(\gamma_m^{\mathrm{SI}} + 1)}{P_{\mathrm{BS}}\gamma_m}. \end{aligned} \tag{A.2}$$

Obviously, $A_{m,n}^{\mathrm{H}}$, $B_{m,n}^{\mathrm{H}}$, and $C_{m,n}^{\mathrm{H}}$ define together the boundaries of the region of the feasible solutions of α as presented in Fig. A.1. In addition, it can also be observed from equation (2.18a) that the sum rate is an decreasing function with respect to $\alpha_{m,n}$ and an increasing function with respect to $P_{m,n}^{\mathrm{d}}$, which can be confirmed through its partial derivatives. Consequently, to maximize this sum-rate function, the smallest and feasible value of $\alpha_{m,n}$ should be selected. By observing Fig. A.1, it can be realized that the pre-selected values of $\alpha_{m,n}$ are along the red and the blue lines, depending on the value of the relaying power budget $\bar{P}^{\mathrm{d}}$. Let $P_{m,n}^{\min}$ denotes the power value of the intersection point between the black and the red lines, which can be easily found by solving the equation $B_{m,n}^{\mathrm{H}} = C_{m,n}^{\mathrm{H}}$.

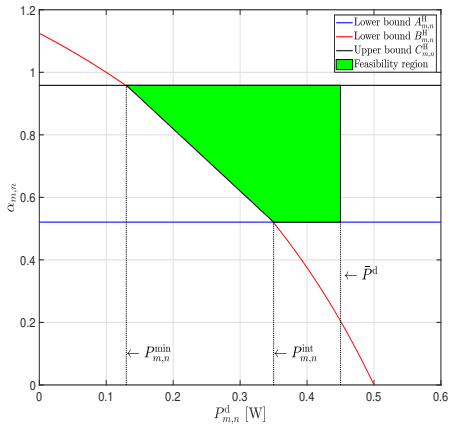


Figure A.1: Feasibility region of power control problem for HD C-NOMA.

Now, in order for the inner problem $\mathscr{P}_{\text{inner}}$ to be feasible, it is necessary that the feasibility region (green region in Fig. A.1) is non-empty. In other words, the value $P^{\bar{d}}_{m,n}$ needs to be greater then $P^{\min}_{m,n}$ and the upper bound C^{H} needs to be greater than 0. Solving these two inequalities implies the feasibility conditions in 2.19a-2.19b, which completes the proof of theorem 1.

From the above observation, it is obvious that the value $P^{d}_{m,n}$ should be as large as possible so that value α takes the minimum value. However, after the intersection between the red and the blue lines, which holds at the power value $P^{int}_{m,n}$, i.e., when $B^{H}_{m,n} = A^{H}_{m,n}$, the value of α will be constant even when $P^{int}_{m,n} \leq P^{d}_{m,n}$. Thus, when $\bar{P}^{d} \leq P^{int}_{m,n}$, the optimal value of $P^{d}_{m,n} = \bar{P}^{d}$ and the optimal value of $\alpha = B^{H}_{m,n}\left(\bar{P}^{d}\right)$. However, from a power efficiency perspective, when $P^{int}_{m,n} \leq \bar{P}^{d}$, the optimal value of $P^{d}_{m,n} = P^{int}_{m,n}$ and the optimal value of $\alpha = A^{H}_{m,n} = B^{H}_{m,n}\left(P^{int}_{m,n}\right)$. This scheme implies the closed-form solutions in (2.21), which completes the proof for theorem 2.

Appendix B

Power control for FD C-NOMA

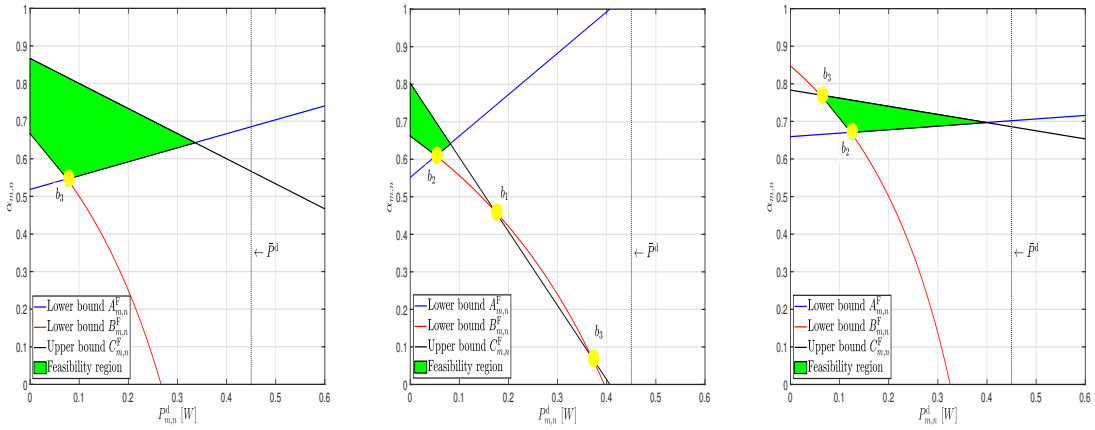
In this section, we present the proof of theorems 3 and 4.At first, note that for any value $\alpha \in [0,1]$, constraints (2.18d) and (2.18e) can be equivalently transformed into the following inequalities.

$$\max\left(A^{\mathrm{F}}_{m,n}, B^{\mathrm{F}}_{m,n}\right) \leq \alpha \leq C^{\mathrm{F}}_{m,n}, \tag{B.1}$$

where

$$\begin{aligned} A^{\mathrm{F}}_{m,n}\left(P^{\mathrm{d}}_{m,n}\right) &= \frac{\delta^{\mathrm{F}}_{\mathrm{th}}(P_{\mathrm{BS}}\gamma_m + \gamma^{\mathrm{SI}}_m P^{\mathrm{d}}_{m,n} + 1)}{P_{\mathrm{BS}}\gamma_m(\delta^{\mathrm{F}}_{\mathrm{th}} + 1)}, \\ B^{\mathrm{F}}_{m,n}\left(P^{\mathrm{d}}_{m,n}\right) &= \frac{(P_{\mathrm{BS}}\gamma_n + 1)(\delta^{\mathrm{F}}_{\mathrm{th}} - P^{\mathrm{d}}_{m,n}\gamma^{\mathrm{d}}_{m,n})}{P_{\mathrm{BS}}\gamma_n(\delta^{\mathrm{F}}_{\mathrm{th}} + 1 - P^{\mathrm{d}}_{m,n}\gamma^{\mathrm{d}}_{m,n})}, \\ C^{\mathrm{F}}_{m,n} &= 1 - \frac{\delta^{\mathrm{F}}_{\mathrm{th}}(\gamma^{\mathrm{SI}}_m + 1)}{P_{\mathrm{BS}}\gamma_m}. \end{aligned} \tag{B.2}$$

Obviously, A^{F} and B^{F} are functions of the relaying power $P^{\mathrm{d}}_{m,n}$, and along with $C^{\mathrm{F}}_{m,n}$, they define together the boundaries of the region of the feasible solutions of α. Precisely, the feasible region of α is characterized by the relation presented in (B.1). Obviously, there exit feasible solutions if and only if this region is non-empty. In order to characterize this feasible region, we use the geometric representations of (B.1) as shown in Fig. B.1(a), Fig. B.1(b) and Fig. B.1(a)B.1(c), where b_1 and b_3 are the lower and upper intersections between $B^{\mathrm{F}}_{m,n}$ and $C^{\mathrm{F}}_{m,n}$, respectively, and b_2 is the intersection between $A^{\mathrm{F}}_{m,n}$ and $B^{\mathrm{F}}_{m,n}$. It can be easily proven that $A^{\mathrm{F}}_{m,n}$ intersects with B^{F} at a positive value of $P^{\mathrm{d}}_{m,n}$ by solving the equation $A^{\mathrm{F}}_{m,n}\left(P^{\mathrm{d}}_{m,n}\right) = B^{\mathrm{F}}_{m,n}\left(P^{\mathrm{d}}_{m,n}\right)$. On the other hand, B^{F} and C^{F} may or may



(a) No intersection. (b) Positive lower intersection. (c) Negative lower intersection.

Figure B.1: Feasibility region power control problem for FD C-NOMA.

not intersect depending on the value Δ_1 in 2.24 that is arisen from solving the equation $B^{\mathrm{F}}_{m,n} = C^{\mathrm{F}}_{m,n}$. There are three cases where feasible solutions exist as follows.

(1) When there is no intersection between $B^{\mathrm{F}}_{m,n}$ and $C^{\mathrm{F}}_{m,n}$ (Fig. B.1(a)).

(2) when the two intersections of $B^{\mathrm{F}}_{m,n}$ and $C^{\mathrm{F}}_{m,n}$ occur at positive values of $P^{\mathrm{d}}_{m,n}$ (Fig. B.1(b)).

(3) when the lower intersection of $B^{\mathrm{F}}_{m,n}$ and $C^{\mathrm{F}}_{m,n}$ occurs at a negative value of $P^{\mathrm{d}}_{m,n}$ and their higher intersection (either at a positive or negative value of $P^{\mathrm{d}}_{m,n}$) is lower than the intersection between A^{F} and C^{F} (Fig. B.1(c)).

Based on this, we obtain the feasibility conditions as presented in (2.23) and it can be easily verified that there is no feasible solutions for α otherwise, which completes the proof for theorem 3.

After determining the feasibility conditions of α, we can proceed to find the optimal power control scheme. First, we note that the inner problem $\mathscr{P}_{\text{inner}}$ is non-convex due to constraints (2.18d) and (2.18e). However, it is widely known that the optimal power allocation scheme that maximizes the sum rate follows the water-filling power allocation policy. Thus, we choose the minimum feasible value of α for maximizing the objective function of $\mathscr{P}_{\text{inner}}$, and therefore, we the optimal value of $P^{\mathrm{d}}_{m,n}$ is the one that achieves the lowest feasible α. Based on this and on Fig. B.1(a), Fig. B.1(b) and Fig. B.1(a)B.1(c), when $\bar{P}^{\mathrm{d}} \leq b_2$, the optimal value of $P^{\mathrm{d}}_{m,n} = \bar{P}^{\mathrm{d}}$ and the

optimal value of $\alpha = B^{\mathrm{F}}_{m,n}\left(\bar{P}^{\mathrm{d}}\right)$. However, when $b_2 \leq \bar{P}^{\mathrm{d}}$, the optimal value of $P^{\mathrm{d}}_{m,n} = b_2$ and the optimal value of $\alpha = A^{\mathrm{F}}_{m,n}(b_2) = B^{\mathrm{F}}_{m,n}(b_2)$, which completes the proof for theorem 4.

Appendix C

Lipschitz Continuity and Convexity after DC Decomposition

First, we can rewrite $R_{jk}(\mathbf{w},\theta)$ as

$$R_{jk}(\mathbf{w},\theta) = \log(\sum_{k\in\mathscr{K}} \left|(\mathbf{h}_{jk}(\theta_{jk}))^{\dagger}\mathbf{w}_k\right|^2 + \sigma_0^2) - \log(\sum_{\ell\neq k} \left|(\mathbf{h}_{jk}(\theta_{jk}))^{\dagger}\mathbf{w}_\ell\right|^2 + \sigma_0^2), \tag{C.1}$$

The gradient vector of $R_{jk}(\mathbf{w},\theta)$ is defined as $\nabla R_{jk}(\mathbf{w},\theta) = \left[\nabla_{\mathbf{w}} R_{jk}(\mathbf{w},\theta); \nabla_{\theta} R_{jk}(\mathbf{w},\theta)\right]$, with

$$\nabla_{\mathbf{w}} R_{jk}(\mathbf{w},\theta) = \underbrace{\frac{2\mathbf{w}^{\dagger}\mathbf{H}_{jk}(\theta)}{\mathbf{w}^{\dagger}\mathbf{H}_{jk}(\theta)\mathbf{w}+\sigma_0^2}}_{\bar{h}_1(\mathbf{w},\theta)\bar{h}_2(\mathbf{w},\theta)} - \underbrace{\frac{2\mathbf{w}^{\dagger}\tilde{\mathbf{H}}_{jk}(\theta)}{\mathbf{w}^{\dagger}\tilde{\mathbf{H}}_{jk}\mathbf{w}+\sigma_0^2}}_{\underline{\mathrm{h}}_1(\mathbf{w},\theta)\underline{\mathrm{h}}_2(\mathbf{w},\theta)} \tag{C.2}$$

$$\nabla_{\theta} R_{jk}(\mathbf{w},\theta) = \underbrace{\frac{2\theta^{\dagger}\Omega_{jk}(\mathbf{w})}{\mathbf{w}^{\dagger}\Omega_{jk}(\mathbf{w})\mathbf{w}+\sigma_0^2}}_{\bar{h}_3(\mathbf{w},\theta)\bar{h}_2(\mathbf{w},\theta)} - \underbrace{\frac{2\theta^{\dagger}\tilde{\Omega}_{jk}(\mathbf{w})}{\mathbf{w}^{\dagger}\tilde{\Omega}_{jk}(\mathbf{w})\mathbf{w}+\sigma_0^2}}_{\underline{\mathrm{h}}_3(\mathbf{w},\theta)\underline{\mathrm{h}}_2(\mathbf{w},\theta)} \tag{C.3}$$

In (C.2), we define $\hat{\mathbf{H}}_{jk}(\theta) = \mathbf{h}_{jk}(\theta_{jk})(\mathbf{h}_{jk}(\theta_{jk}))^{\dagger}$ to support the notational purpose of $\mathbf{H}_{jk}(\theta) = \mathrm{Bdiag}\underbrace{\left(\hat{\mathbf{H}}_{jk}(\theta),\ldots,\hat{\mathbf{H}}_{jk}(\theta)\right)}_{U\text{elements}}$ and $\tilde{\mathbf{H}}_{jk}(\theta) = \mathrm{Bdiag}\underbrace{\left(\hat{\mathbf{H}}_{jk}(\theta),\ldots,0,..,\hat{\mathbf{H}}_{jk}(\theta)\right)}_{U\text{elements}}$, where 0 appears at the k-th element. In (C.3), we define $\omega_{jk}(\mathbf{w}_k) = \mathbf{w}_k \circ \mathbf{h}_k$, $\omega'_{jk} = \omega_{jk}(\mathbf{w})(\omega_{jk}(\mathbf{w}))^{\dagger}$ to support the notational purpose of $\Omega_{jk}(\mathbf{w}) = \sum_{k\in\mathscr{K}} \omega'_{jk}$ and $\tilde{\Omega}_{jk}(\mathbf{w}) = \sum_{l\in\mathscr{K}\setminus k} \omega'_{jl}$. Then, we attempt

to find Lipschitz constants for $\nabla_{\mathbf{w}} R_{jk}(\mathbf{w},\theta)$ and $\nabla_{\theta} R_{jk}(\mathbf{w},\theta)$. First, we have

$$\begin{aligned}
&\left\|\bar{h}_1(\mathbf{w},\theta)-\bar{h}_1(\mathbf{w}^{(0)},\theta^{(0)})\right\| = \|2\mathbf{w}^{\dagger}\mathbf{H}_{jk}(\theta)-2\mathbf{w}^{(0)\dagger}\mathbf{H}_{jk}\left(\theta^{(0)}\right)+ \\
&2\mathbf{w}^{\dagger}\mathbf{H}_{jk}\left(\theta^{(0)}\right)-2\mathbf{w}^{\dagger}\mathbf{H}_{jk}\left(\theta^{(0)}\right)\| \leq \left\|2\mathbf{w}^{\dagger}(\mathbf{H}_{jk}(\theta)-\mathbf{H}_{jk}\left(\theta^{(0)}\right))\right\| \\
&+\left\|2(\mathbf{w}-\mathbf{w}^{(0)})^{\dagger}\mathbf{H}_{jk}\left(\theta^{(0)}\right)\right\| \leq 2\|\mathbf{w}\|\left\|\text{Bdiag}\left(\left(\mathbf{h}_{jk}(\theta_{jk})-\mathbf{h}_{jk}(\theta_{jk})\right)\left(\mathbf{h}_{jk}(\theta_{jk})\right)^{\dagger}\right)\right\|_F \\
&+2\|\mathbf{w}\|\left\|\text{Bdiag}\left(\mathbf{h}_{jk}(\theta_{jk}^{(0)})\left(\mathbf{h}_{jk}(\theta_{jk})-\mathbf{h}_{jk}(\theta_{jk}^{(0)})\right)^{\dagger}\right)\right\|_F+2\left\|\mathbf{w}-\mathbf{w}^{(0)}\right\|\left\|\mathbf{H}_{jk}(1/H)\right\|_F \\
&\leq 2\|\mathbf{w}\|2U\left\|\mathbf{h}_{jk}(1/H)\right\|\left\|\tilde{\mathbf{h}}_{jk}\right\|_{\infty}\left\|\theta_{jk}-\theta_{jk}^{(0)}\right\|+2\left\|\mathbf{w}-\mathbf{w}^{(0)}\right\|\left\|\mathbf{H}_{jk}(1/H)\right\|_F \leq \\
&A_1\left(\left\|\theta-\theta^{(0)}\right\|+\left\|\mathbf{w}-\mathbf{w}^{(0)}\right\|\right) \qquad \text{(C.4)}
\end{aligned}$$

with

$$A_1 = \max\left\{4U^2P_{\max}\left\|\mathbf{h}_{jk}(1/H)\right\|\left\|\tilde{\mathbf{h}}_{jk}\right\|_{\infty}, 2\left\|\mathbf{H}_j(1/H)\right\|_F\right\}$$

In other words, the Lipschitz constant of $\bar{h}_1(\mathbf{v})$ is A_1. Similarly, we can derive

$$\left\|\bar{h}_2(\mathbf{w},\theta)-\bar{h}_2(\mathbf{w}^{(0)},\theta^{(0)})\right\| \leq B_1\left(\left\|\theta-\theta^{(0)}\right\|+\left\|\mathbf{w}-\mathbf{w}^{(0)}\right\|\right) \qquad \text{(C.5)}$$

where $B_1 = 1/N_0^2(A_1/2+UP_{\max}\left\|\mathbf{H}_j(1/H)\right\|_F$. Using (C.4) and (C.5), we can find the Lipschitz constant $\mathscr{A}_1$ of $\bar{h}_1(\mathbf{w},\theta)\bar{h}_2(\mathbf{w},\theta)$ as

$$\begin{aligned}
&\left\|\bar{h}_1(\mathbf{w},\theta)\bar{h}_2(\mathbf{w},\theta)-\bar{h}_1(\mathbf{w}^{(0)},\theta^{(0)})\bar{h}_2(\mathbf{w}^{(0)},\theta^{(0)})\right\| \quad \leq \underbrace{\left(2UP_{\max}\left\|\mathbf{H}_{jk}(1/H)\right\|_F B_1+1/N_0A_1\right)}_{\mathscr{A}_1}\times \\
&\left(\left\|\theta-\theta^{(0)}\right\|+\left\|\left(\mathbf{w}-\mathbf{w}^{(0)}\right)\right\|\right) \qquad \text{(C.6)}
\end{aligned}$$

Similarly, we can derive the Lipschitz constant of $\bar{h}_3(\mathbf{w},\theta)$ as

$$\left\|\bar{h}_3(\mathbf{w},\theta)-\bar{h}_3(\mathbf{w}^{(0)},\theta^{(0)})\right\| \leq C_1\left(\|\theta-\theta^{(0)}\|+\left\|\left(\mathbf{w}-\mathbf{w}^{(0)}\right)\right\|\right) \qquad \text{(C.7)}$$

where $C_1 = \max\left\{4U^2/H\left\|\tilde{\mathbf{h}}_{jk}\right\|_{\infty}^2P_{\max}, 2UP_{\max}\left\|\tilde{\mathbf{h}}_{jk}\right\|_{\infty}\right\}$. Then, we can also find the Lipschitz constant $\mathscr{B}_1$ of $\bar{h}_3(\mathbf{w},\theta)\bar{h}_3(\mathbf{w},\theta)$ as

$$\left\|\bar{h}_3(\mathbf{w},\theta)\bar{h}_2(\mathbf{w},\theta)-\bar{h}_3(\mathbf{w}^{(0)},\theta^{(0)})\bar{h}_2(\mathbf{w}^{(0)},\theta^{(0)})\right\|$$
$$\leq \underbrace{\left(2U/H2UP^{\max}\left\|\tilde{\mathbf{h}}_{jk}\right\|_{\infty}B_1+1/N_0C_1\right)}_{\mathscr{B}_1}\times\left(\left\|\theta-\theta^{(0)}\right\|+\left\|\left(\mathbf{w}-\mathbf{w}^{(0)}\right)\right\|\right) \quad \text{(C.8)}$$

By following the same steps, we can also conclude that the Lipschitz constants for the terms $\underline{\mathrm{h}}_k(\mathbf{w},\theta), k=1\ldots,3$ are A_1,B_1,C_1, respectively. Finally, we have

$$\left\|\nabla_{\mathbf{v}}R_{jk}(\mathbf{w},\theta)-\nabla_{\mathbf{v}}R_{jk}(\mathbf{w}^{(0)},\theta^{(0)})\right\|\leq\xi_0\left(\left\|\theta-\theta^{(0)}\right\|+\left\|\left(\mathbf{w}-\mathbf{w}^{(0)}\right)\right\|\right) \quad \text{(C.9)}$$

where $\xi_0=2\times\max\{\mathscr{A}_1,\mathscr{B}_1\}$. We now prove Theorem 5. According to (C.9), $R_{jk}(\mathbf{v},\theta)$ is ξ_0âĂŞsmooth. Thus, we have

$$\|R_{jk}(\mathbf{w},\theta)-R_{jk}(\mathbf{w}^{(0)},\theta^{(0)})-\nabla_{\mathbf{v}}R_{jk}(\mathbf{w}^{(0)},\theta^{(0)})^{\intercal}([\mathbf{w},\theta]-[\mathbf{w}^{(0)},\theta^{(0)}])\|$$
$$\leq\frac{\xi_0}{2}(||\theta-\theta^{(0)}||^2+||\mathbf{w}-\mathbf{w}^{(0)}||^2) \quad \text{(C.10)}$$

which means

$$R_{jk}(\mathbf{w},\theta)\geq-\frac{\xi_0}{2}(\left\|\theta-\theta^{(0)}\right\|^2+\left\|\left(\mathbf{w}-\mathbf{w}^{(0)}\right)\right\|^2)+$$
$$R_{jk}(\mathbf{w}^{(0)},\theta^{(0)})+\nabla R_{jk}(\mathbf{w}^{(0)},\theta^{(0)})^{\intercal}([\mathbf{w},\theta]-[\mathbf{w}^{(0)},\theta^{(0)}]) \quad \text{(C.11)}$$

Because of the strong convexity of $\xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2)$, we have

$$\xi_{jk}(\|\mathbf{w}\|^2+\|\theta\|^2)\geq\xi_{jk}\left\|\mathbf{w}^{(0)}\right\|^2+\left\|\theta^{(0)}\right\|^2)+2\xi_{jk}[\mathbf{w}^{(0)},\theta^{(0)}]^{\intercal}([\mathbf{w},\theta]-[\mathbf{w}^{(0)},\theta^{(0)}])$$
$$+\frac{\xi_{jk}}{2}(\left\|\theta-\theta^{(0)}\right\|^2+\left\|\left(\mathbf{w}-\mathbf{w}^{(0)}\right)\right\|^2) \quad \text{(C.12)}$$

From the two inequalities (C.11) and (C.12), we obtain

$$f_{jk}(\mathbf{w},\theta)\geq\frac{\xi_{jk}-\xi_0}{2}(||\theta-\theta^{(0)}||^2+||\mathbf{w}-\mathbf{w}^{(0)}||^2)+f_{jk}(\mathbf{w}^{(0)},\theta^{(0)})+$$

$$\nabla f_{jk,t}(\mathbf{w},\theta)(\mathbf{w}^{(0)},\theta^{(0)})^{\intercal}([\mathbf{w},\theta]-[\mathbf{w}^{(0)},\theta^{(0)}])$$

which means that $R_{jk}(\mathbf{w},\theta)$ is strongly convex.