Applications of Power-Domain Non-Orthogonal Multiple Access in Wireless Communication Networks

by

Guoxin Li

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Communications

Department of Electrical and Computer Engineering
University of Alberta

Abstract

The expected explosive data traffic and scarcity of available spectrum in wireless networks place higher demands on spectrum utilization. As one of the advanced techniques for improving spectral efficiency, non-orthogonal multiple access (NOMA) has raised considerable attention recently. Although many works have been conducted to demonstrate the advantages of NOMA in various communication scenarios, there are still some open problems. Our research attempts to fill some research gaps by investigating protocol design, resource allocation, and performance analysis for the applications of power-domain NOMA in cooperative networks, cognitive radio networks, and wireless powered networks.

We first investigate how to enjoy the benefits of cooperative technique for NOMA-based systems while addressing the possible loss of spectral efficiency caused by the fixed relaying. Due to the constraint of half-duplex, the relay needs to occupy half of the transmission block to forward information. Therefore, cooperative NOMA networks may suffer a loss of spectral efficiency if the relay always participates in the cooperation. To solve this concern, we propose an incremental cooperative NOMA protocol in which the relay is invoked for cooperation in a more reasonable way.

We next study an overlay cognitive NOMA network which can be treated as a type of NOMA networks with user priority difference. In this type of networks, how to improve the performance of the lower-priority user while ensuring the quality-of-service requirements of the higher-priority user becomes an appropriate research target, which is quite different from the conventional NOMA networks where sum rate or user fairness is more concerned. This observation motivates us to study the joint optimal design of power allocation and decoding order suitable for overlay cognitive NOMA networks. Specifically, we propose two joint optimal strategies based on the availability of channel state information at the transmitter.

We then solve two optimization problems in a cooperative NOMA network with simultaneous wireless information and power transfer. Different from current works where the solely power-splitting (PS) or the solely time-splitting (TS) scheme is employed to harvest

energy, we assume that the energy-constrained relay employs the hybrid PS and TS scheme to harvest energy since the hybrid PS and TS scheme is expected to outperform the PS and the TS schemes by exploiting both the degree-of-freedom in the time and power domains. Depending on the availability of channel state information at the transmitter, we successfully address two optimization problems which realize maximization of energy efficiency and minimization of system outage probability, respectively.

Finally, we address a fair resource allocation problem involved in a power beacon-assisted downlink NOMA network. In this network, the energy-constrained source first collects energy from the power beacon and then uses it to communicate with a pair of users. By designing an alternating optimization algorithm based on the statistical channel state information, we successfully obtain the joint optimal time allocation and power allocation ratios that can maximize the worse throughput of the paired users.

Preface

This thesis is an original work by Guoxin Li. Parts of the thesis have been published or submitted to journals, which are indicated below.

The work of Chapter 2 is published as "G. Li, *et al.*, 'Cooperative NOMA with Incremental Relaying: Performance Analysis and Optimization,' in *IEEE Trans. Veh. Technol*, vol. 67, no. 11, pp. 11291–11295, Nov. 2018".

The work of Chapter 3 has been submitted to IEEE Transactions on Vehicular Technology in July 2019 as "G. Li, *et al.*, 'Channel-Aware Power Allocation and Decoding Order in Overlay Cognitive NOMA Networks'.

The work of Chapter 4 has been submitted to IEEE Transactions on Communications in August 2019 as "G. Li, *et al.*, 'Optimal Designs for Relay-Assisted NOMA Networks with Hybrid SWIPT Scheme'.

The work of Chapter 5 has been submitted to IEEE Transactions on Vehicular Technology in August 2019 as "G. Li, *et al.*, 'Resource Allocation in Power Beacon-Assisted Non-Orthogonal Multiple Access Networks'.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Hai Jiang. Thanks for providing me with this precious opportunity to study at such a friendly, lovely, and beautiful country. Thanks for your high standards and strict requirements for my work. Your scrupulous academic attitude deeply affects me, and I believe that will benefit me throughout my whole career and life. It's my great honour to be your student.

Many thanks to my thesis examining committee members Dr. Witold A. Krzymień, Dr. Masoud Ardakani, Dr. Huazhou Li, and Dr. Xianbin Wang for their insightful suggestions helping me to improve the quality of the thesis. Special thanks to all the teachers who delivered courses in the first two years during my PhD study, which helps me to lay a solid academic foundation.

I would like to thank all my collaborators, Dr. Saman Atapattu, Dr. Yulin Hu, Dr. Yuzhen Huang, and Dr. Deepak Mishra, for their helpful discussions and kind advice. I also would like to thank all my colleagues, including but not limited to Lijie Huang, Jie Gao, Fudong Li, Xiao Lu, Lu Lv, Ziling Wei, Keyu Wu, and Long Yang, for helping me in life and study.

I acknowledge the financial support from Alberta Innovates-Graduate Student Scholarship and also like to thank the Department of Electrical and Computer Engineering for providing me with the teaching assistant positions. These job opportunities not only give me financial support but also help me to accumulate valuable teaching experience.

Last but not least, I would like to thank my parents for their continuous support and selfless love. You keep encouraging me whenever I doubt myself. You always be the most trusted reliance when I am lonely and helpless. From the bottom of my heart, I thank them all profusely and will be eternally grateful.

Contents

List of Tables

List of Figures

Glossary of Terms

Acronyms	Definition
AF	amplify-and-forward
AWGN	additive white Gaussian noise
CCN	conventional cooperative non-orthogonal multiple access
CCDF	complementary cumulative distribution function
CDF	cumulative distribution function
CSI	channel state information
CSIT	channel state information at the transmitter
DF	decode-and-forward
DP-DD	dynamic power allocation and dynamic decoding order
DP-FD	dynamic power allocation and fixed decoding order
EH	energy harvesting
F-PA-DO	full channel state information power allocation and decoding order
FP-FD	fixed power allocation and fixed decoding order
HAP	hybrid access point
H-SWIPT	hybrid simultaneous wireless information and power transfer
HSCN	hybrid simultaneous wireless information and power transfer based cooperative non-orthogonal multiple access
i.i.d.	independent and identically distributed
i.n.i.d.	independent and not necessarily identically distributed
ICN	incremental cooperative non-orthogonal multiple access
IoT	Internet of Things
MRC	maximal ratio combining
NOMA	non-orthogonal multiple access
NBT	non-orthogonal multiple access based transmission

OMA	orthogonal multiple access
OBT	orthogonal multiple access based transmission
PB	power beacon
PCN	power-splitting based cooperative non-orthogonal multiple access
PDF	probability distribution function
P-PA-DO	partial channel state information power allocation and decoding order
PS	power-splitting
QoS	quality-of-service
SIC	successive interference cancellation
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
SWIPT	simultaneous wireless information and power transfer
TCN	time-splitting based cooperative non-orthogonal multiple access
TS	time-splitting

Glossary of Notations

$\|\|\cdot\|\|$	the Frobenius norm
$\|\cdot\|$	the absolute value
$f_X(x)$	the PDF of a random variable X
$F_X(x)$	the CDF of a random variable X
$\bar{F}_X(x)$	the CCDF of a random variable X
$\Pr\{X\}$	the probability of an event X
$K_\nu(\cdot)$	the ν-th order modified Bessel function of the second kind, $K_\nu(x) = \int_0^\infty e^{-x\cosh t}\cosh\nu t dt$
$\Gamma(\cdot)$	the gamma function, $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$
$\Gamma(\cdot,\cdot)$	the upper incomplete gamma function, $\Gamma(\alpha,x) = \int_x^\infty e^{-t}t^{\alpha-1}dt$
$\gamma(\cdot,\cdot)$	the lower incomplete gamma function, $\gamma(\alpha,x) = \int_0^x e^{-t}t^{\alpha-1}dt$
$\mathbb{E}[\cdot]$	the expectation operator
$[x]^+$	the larger value in x and 0, $[x]^+ = \max\{x,0\}$
$\binom{n}{k}$	the number of k-element subsets out of an n-element set, $\binom{n}{k} = \frac{n!}{(n-k)!k!}$
$X \sim \mathcal{CN}(0,\sigma^2)$	a circularly symmetric complex Gaussian random variable X with variance σ^2

Chapter 1

Introduction

1.1 Spectrally Efficient Communications

With the rapid development of wireless communication technology, future wireless networks are envisioned to support a diversity of usage scenarios, such as the Internet of Things (IoT), augmented virtual reality, and the tactile internet, which certainly poses an extremely high demand on data rates. As shown in Fig. 1.1, the recent report by Cisco Visual Networking Index [1] forecasts that mobile data and Internet traffic will be 77.5 exabytes/month in 2022, which means a sevenfold increase from 2017 to 2022. To accommodate such explosive data traffic, we can explore new spectrum (millimeter wave frequencies, 30-300 GHz) or apply some advanced techniques to further improve spectral efficiency.

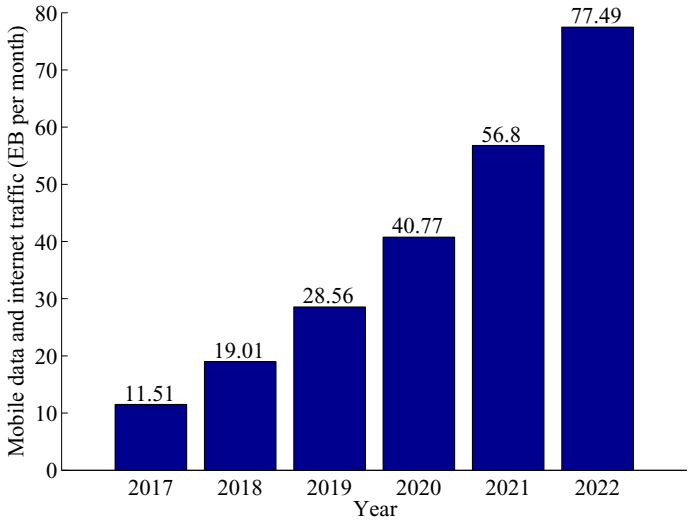


Figure 1.1: Mobile data and Internet traffic from 2017 to 2022 [1].

In the past few years, several emerging techniques, such as full-duplex [2], massive multiple-input multiple-output (MIMO) [3], and non-orthogonal multiple access (NOMA) [4], have been proposed to improve spectral efficiency from different perspectives. In particular, full-duplex improves spectral efficiency by allowing simultaneous transmission and reception at transceivers. Massive MIMO, which deploys a large-scale antenna at the base station, utilizes a huge degree-of-freedom in the spatial domain to generate receive and transmit diversity gain to enhance spectral efficiency. Compared to the orthogonal multiple access (OMA) techniques, such as frequency-division multiple access, time-division multiple access, and orthogonal frequency-division multiple access, NOMA promises to improve spectral efficiency by allowing more than one user to access the same orthogonal resource block (e.g., time, frequency and so on) via power or code domain multiplexing [5–8]. A schematic of the difference between NOMA and OMA is shown in Fig. 1.2.

Compared to power-domain NOMA, code-domain NOMA can introduce additional spreading gain and shaping gain at the expense of higher implementation complexity and wider signal bandwidth [4, 9]. Considering that the base station can afford more complex computation unit, more code-domain NOMA schemes are proposed for uplink transmissions. This thesis focuses on the implementation of the power-domain NOMA on the downlink transmissions, and unless otherwise stated, we hereafter refer to power-domain NOMA as NOMA. Next, we will introduce the NOMA technique in detail.

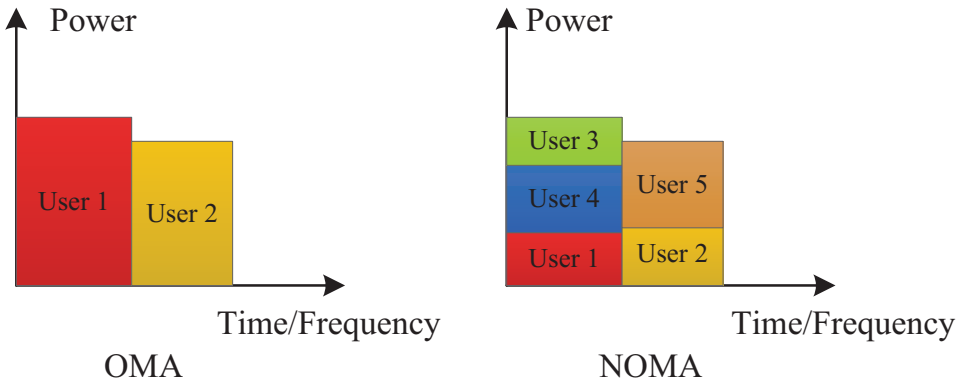


Figure 1.2: Illustration of OMA and NOMA.

1.2 NOMA Scheme

1.2.1 Key Technologies of NOMA

The two key techniques in NOMA transmission systems are the superposition coding at the transmitter side and the successive interference cancellation (SIC) at the receiver side.

- Superposition coding, first proposed in [10], is a technique that allows the transmitter to transmit multiple users' information by one block. It is well known that superposition coding can achieve the capacity on a Gaussian scalar broadcast channel [11].

- SIC and parallel interference cancellation are two major categories of interference cancellation for multi-user systems. Compared to parallel interference cancellation, SIC has lower complexity and performs in an iterative manner [12]. Specifically, receivers with SIC first decode the signal with the largest power and then subtract it from the received composite signal; the same procedure will be performed several times until the signal with the smallest power can be detected without interference [13]. It has been shown in [14] that SIC can achieve the boundaries of Shannon capacity region for both the uplink and downlink multi-user channels.

1.2.2 An Example of NOMA

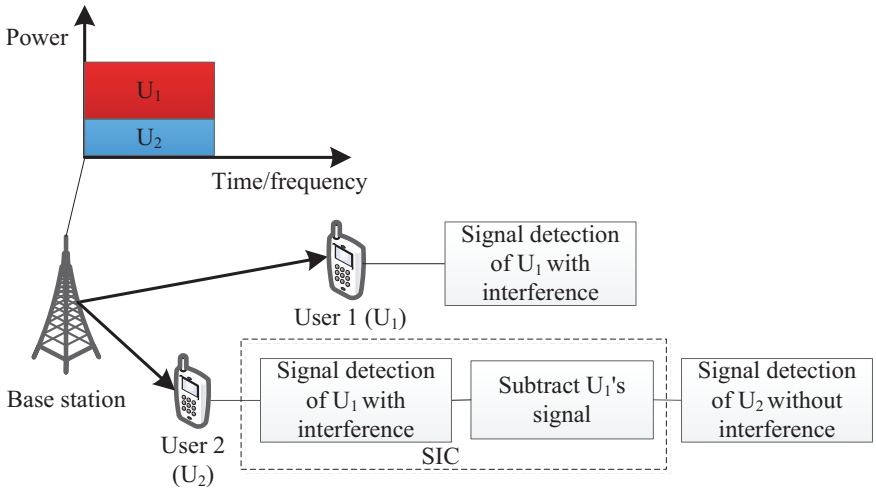


Figure 1.3: A two-user downlink NOMA network.

To facilitate the understanding of NOMA, we here take a two-user downlink NOMA network as an example, as shown in Fig. 1.3. In this network, the base station intends to communicate with two users, U_1 and U_2. The base station first generates a superimposed message $\sum_{n=1}^{2}\sqrt{\alpha_n P_b}x_n$, where x_n is the message for user n, P_b denotes the transmit power of the base station, and α_n is the power allocation coefficient. Assume that users are ordered by their channel conditions and user 1 has a weaker connection to the base station, i.e., $|h_2|^2 > |h_1|^2$. To ensure that user 1 can directly detect x_1 by treating x_2 as interference, the base station allocates more power to user 1, i.e., $\alpha_1 > 0.5$ [15]. After receiving the superimposed signal, user 1 directly decodes x_1 and treats x_2 as noise, then its achievable rate for x_1 is given by[1]

$$R_{1\to 1} = \log_2\left(1 + \frac{\alpha_1 P_b |h_1|^2}{\alpha_2 P_b |h_1|^2 + N_0}\right), \tag{1.1}$$

where $|h_n|^2$ is the channel gain of user n and N_0 denotes the power of the additive white Gaussian noise (AWGN) at user n, $n = 1, 2$.

Since x_1 is allocated with more power, user 2 also first decodes x_1 and treats x_2 as noise, then its achievable rate for x_1 can be written as

$$R_{2\to 1} = \log_2\left(1 + \frac{\alpha_1 P_b |h_2|^2}{\alpha_2 P_b |h_2|^2 + N_0}\right). \tag{1.2}$$

Let $\tilde{R}_1$ denote the targeted data rate of user 1. If user 2 correctly decodes x_1, i.e., $R_{2\to 1} \geq \tilde{R}_1$, the component of x_1 can be subtracted from the received signal by applying SIC. Then x_2 can be decoded without inter-user interference, and its achievable rate at user 2 is given by

$$R_{2\to 2} = \log_2\left(1 + \frac{\alpha_2 P_b |h_2|^2}{N_0}\right). \tag{1.3}$$

1.2.3 Two Key Problems in NOMA-Based Systems

In the previous subsection, we have introduced that user 2 needs to decode x_1 before decoding itself information. To ensure that user 2 can correctly decode x_1 and successfully perform SIC with a non-zero probability, i.e, $\Pr\left\{R_{2\to 1} \geq \tilde{R}_1\right\} > 0$, we have

[1]Note that (1.1)-(1.3) are derived based on the assumptions that channel codes with arbitrarily long codewords are applied, and both the input and noise are Gaussian distributed. All the analysis throughout this thesis is performed based on these two assumptions.

the following constraint

$$\lim_{P_b|h_2|^2\to\infty} R_{2\to1} = \log_2\left(1+\frac{\alpha_1}{\alpha_2}\right) > \tilde{R}_1. \tag{1.4}$$

From (1.4), a constraint regarding the power allocation factor α_1 can be derived as $\alpha_1 > \frac{2^{\tilde{R}_1}-1}{2^{\tilde{R}_1}}$. This observation demonstrates that the importance of the right choice for the power allocation factors under a given user ordering. Therefore, power allocation and user ordering are two key issues for NOMA-based systems. Since the decoding order at receiver sides relies on user ordering, we use the two terms of decoding order and user ordering interchangeably throughout this thesis.

1.2.4 Main Advantages of NOMA

To show the advantages of NOMA, we here briefly study the achievable rate of each user in a two-user downlink OMA network. If the base station communicates with the two users with the OMA technique, the achievable rate of user n ($n = 1, 2$) is given by

$$R_n = \beta_n \log_2\left(1+\frac{P_b|h_n|^2}{N_0}\right), \tag{1.5}$$

where β_n denotes the proportion of the resource block allocated to user n ($n = 1, 2$) with $0 < \beta_n < 1$ and $\beta_1 + \beta_2 = 1$.

Comparing the achievable rates of each user for NOMA and OMA, we can find that each NOMA user can access the whole resource block at the cost of controllable interference. In the high signal-to-noise ratio (SNR) regime, i.e., $\frac{P_b}{N_0} \to \infty$, the sum rate of the two users for NOMA and OMA can be respectively approximated as

$$\begin{aligned}\log_2\left(1+\frac{\alpha_1 P_b|h_1|^2}{\alpha_2 P_b|h_1|^2+N_0}\right) &+ \log_2\left(1+\frac{\alpha_2 P_b|h_2|^2}{N_0}\right)\\ &\approx \log_2\left(1+\frac{\alpha_1}{\alpha_2}\right) + \log_2\left(\frac{\alpha_2 P_b|h_2|^2}{N_0}\right) = \log_2\left(\frac{P_b|h_2|^2}{N_0}\right),\end{aligned} \tag{1.6}$$

and

$$\beta_1\log_2\left(1+\frac{P_b|h_1|^2}{N_0}\right) + \beta_2\log_2\left(1+\frac{P_b|h_2|^2}{N_0}\right) \approx \beta_1\log_2\left(\frac{P_b|h_1|^2}{N_0}\right) + \beta_2\log_2\left(\frac{P_b|h_2|^2}{N_0}\right). \tag{1.7}$$

Since $|h_2|^2 > |h_1|^2$, it is easy to check that the sum rate of NOMA is higher than that of OMA. In addition to this sum-rate comparison in the high SNR regime, D. Tse *et*

al. have shown that in a two-user uplink or downlink AWGN channel, NOMA is not only superior to OMA in terms of sum rate, but also superior to OMA in terms of fair rate[2] [16].

In addition to the superiority in terms of spectral efficiency (sum throughput of users) and user fairness, NOMA also shows its advantages in the following aspects:

- Latency and massive connectivity: Since NOMA can serve multiple users simultaneously, it is clear that massive connectivity is possible to realize and users can experience low latency.
- Good compatibility: NOMA is compatible with the current wireless networks since it takes advantage of the power domain to achieve multiplexing gain and does not need significant modifications to existing network architectures.

1.3 Applications of NOMA

So far, NOMA has received considerable attention from both academia and industry due to its advantages, and research on it has been extended to various wireless networks. Next, three important application scenarios of NOMA are introduced.

1.3.1 Cooperative NOMA Networks

Compared to OMA-based transmission, the coverage of NOMA-based transmission is relatively limited since each NOMA user is allocated only a fraction of total transmit power. To expand the coverage of NOMA-based transmission, one of the effective ways is to integrate the cooperative techniques into NOMA networks, generating the cooperative NOMA networks. Cooperative communication has been a hot research topic for the past two decades since it can increase the coverage probability, throughput and transmission reliability of the networks by leveraging the spatial diversity gain to combat the effect of wireless fading [17]. The key idea of cooperative communication is to introduce a certain number of intermediate nodes to assist the communication between the source and destination. By combining copies from different paths, the reception reliability of the destination gets improved owing to the

[2]Fair rate refers to the lowest data rate among multiple users and is the criterion for evaluating fair performance in multiuser systems.

spatial diversity gain. In this subsection, three basic relaying protocols and two types of cooperative NOMA networks are respectively introduced.

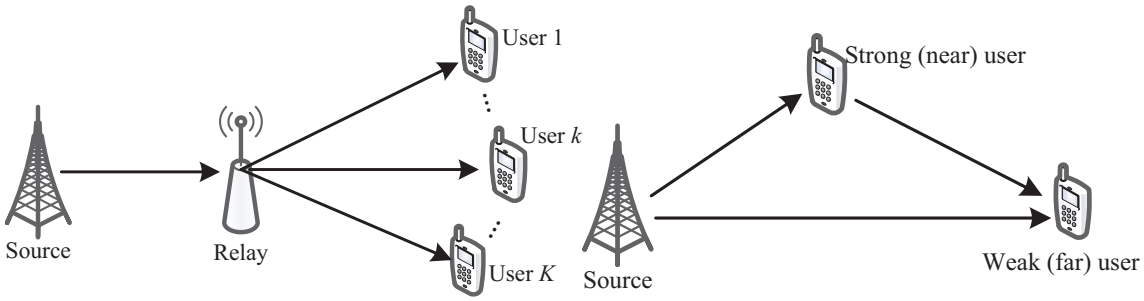
1.3.1.1 Relaying Protocols

Currently, there are three widely adopted relaying protocols: amplify-and-forward (AF), decode-and-forward (DF), and incremental relaying [17]. The relay in AF protocol simply amplifies the received signal and forwards the amplified signal to the destination. The relay in DF protocol first tries to decode the received signal. If the message is correctly decoded, the relay re-encodes the message and then forwards it; otherwise, the relay keeps silent in the forwarding phase. It is expected that the DF protocol outperforms the AF protocol since the AF protocol forwards the received message as well as the received noise. However, the implementation complexity of the DF protocol is also higher as the relay needs to decode message and re-encode message.

The AF and DF protocols can be viewed as fixed relaying protocols as the relay forwards information all the time regardless of the channel condition between the source and the destination. Note that fixed relaying underutilizes the degree-of-freedom of the channel as relaying is unnecessary when the channel condition between the source and the destination is good enough to support a successful transmission. Relaying generally occupies half of the transmission block due to the constraint of half-duplex, which may lower the spectral efficiency of the network. To alleviate this concern, one can apply the incremental relaying protocol. In this protocol, the relay is invoked for cooperation only when the channel condition between the source and the destination is below a predefined threshold. This can be achieved by introducing additional 1-bit feedback sent by the destination. Compared to the AF and DF protocols, incremental relaying protocol improves spectral efficiency by sufficiently use the degree-of-freedom of the channel.

1.3.1.2 Relay-Aided and User-Aided Cooperative NOMA Networks

Depending on the cooperation types, current works on cooperative NOMA networks can be classified into two categories. As shown in Fig. 1.4(a), the first category relies on the dedicated relay cooperation, where dedicated relays are deployed to assist the



(a) Relay-aided cooperative NOMA network. (b) User-aided cooperative NOMA networks.

Figure 1.4: Cooperative NOMA networks.

communication between the source and NOMA users [18]. As shown in Fig. 1.4(b), the second category relies on the user cooperation, where the NOMA users having strong connections to the source serve as relays to help the NOMA users having weak connections to the source [19]. The motivation behind this kind of cooperation is that according to the principle of NOMA, the strong users need to decode the weak users' information before decoding their desired information. Therefore, when the strong users successfully detect the weak users' information, they can be exploited to help the weak users.

1.3.2 Cognitive NOMA Networks

With the increasing diversity of wireless applications and services, spectrum resources are becoming more and more scarce since most of the frequency bands are exclusively allocated to specific services. On the other hand, studies show that most of the allocated spectrum is far from being fully utilized. Therefore, wireless communication faces the contradiction between "shortage" and "waste" of spectrum resources, which is attributed to the fixed exclusive spectrum allocation method. To break this dilemma, the idea of cognitive radio is proposed, in which the secondary users (i.e., unlicensed users) is allowed to opportunistically access the spectrum licensed to the primary users (i.e., licensed users) under certain conditions [20, 21]. Driven by this fact, cognitive radio networks have received considerable attention in the past twenty years. Since cognitive radio and NOMA can improve spectral efficiency from different perspectives, it is promising to further enhance spectrum utilization by combining these two techniques [22–24]. In this subsection, we first introduce three basic architectures of cognitive radio, and then describe a kind of cognitive NOMA networks.

1.3.2.1 Basic Architectures of Cognitive Radio

According to the access manners of the secondary users, cognitive radio has three basic operation modes: interweave, underlay, and overlay [25]. To be specific, interweave cognitive radio does not allow concurrent primary and secondary transmission, and the secondary users can access the licensed spectrum only when the spectrum is idle. To implement interweave cognitive radio, spectrum sensing technique needs to be employed to detect the status of spectrum occupancy. In underlay cognitive radio, the concurrent transmission is allowed as long as the interference of the secondary transmission to the primary network is below a regulated threshold. The secondary users in overlay cognitive radio access the spectrum by using part of its power to assist (relay) the primary transmission, and the remaining power is used for transmitting the secondary message.

1.3.2.2 Overlay Cognitive NOMA Networks

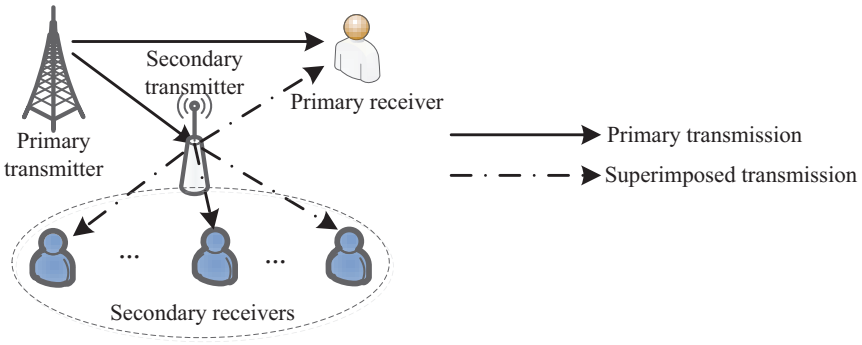


Figure 1.5: An overlay cognitive NOMA network.

Fig. 1.5 illustrates an overlay cognitive NOMA network where the secondary users access the licensed spectrum in a cooperation way. Generally, the cooperation takes place in two phases. In the first phase, the primary transmitter broadcasts a primary message. In the second phase, the secondary transmitter sends a superimposed signal consisting of the primary message received in the first phase and the secondary messages.

1.3.3 Wireless Powered NOMA Networks

The realization of IoT relies on the dense deployment of low-cost devices (sensors) which are generally energy-constrained due to their limited size. Therefore, IoT

networks can be viewed as energy-constrained networks in which how to prolong network lifetime becomes a major concern. To solve this concern, one can periodically replace the batteries of IoT devices. However, periodical battery replacement may be uneconomical and even impractical (e.g., sensors embedded in human bodies) and hazardous (e.g., sensors deployed in a toxic environment) [26, 27]. In this regard, energy harvesting (EH) becomes an appealing technique as it can harvest energy from ambient environments. Current research on EH-based networks mainly focus on two kinds of energy sources: natural renewable energy sources (e.g., tidal, solar, and wind), and man-made radio frequency signals. Compared to the first kind of energy sources which is heavily affected by weather, and thus unstable and random, the second kind of energy sources is controllable and can carry information and energy simultaneously, and thus more suitable for the networks with quality-of-service (QoS) requirements.

With the ability to simultaneously increase the lifetime and spectral efficiency for energy-constrained networks, the integration of wireless power transfer and NOMA has received considerable attention recently. In this subsection, we first introduce three basic operation modes of wireless power transfer, and then present two types of wireless powered NOMA networks.

1.3.3.1 Operation Modes of Wireless Power Transfer

Currently, three canonical operation modes of wireless power transfer are widely studied [28].

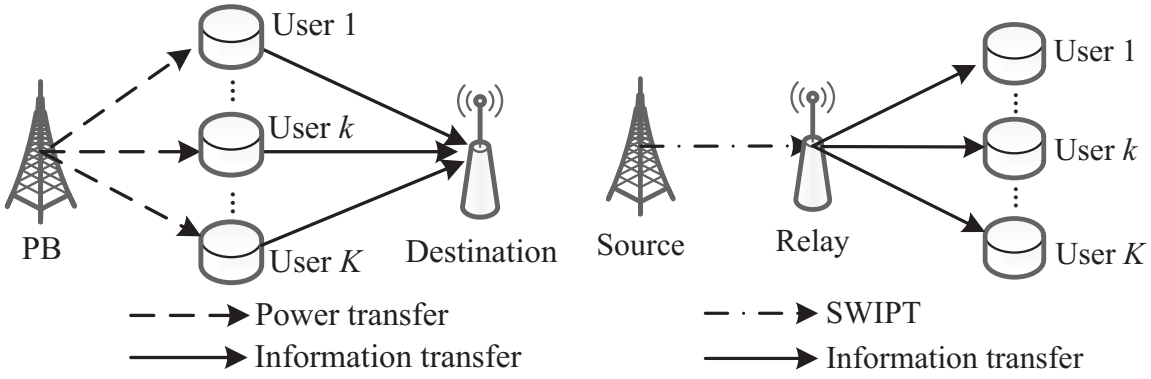
The first one is the hybrid access point (HAP)-assisted networks where users first extract energy from the energy-bearing signals sent by the HAP in the downlink, and then send information to the HAP with the harvested energy [29] in the uplink. A challenging issue in HAP-assisted networks is to design an efficient multiple access scheme alleviating the doubly near-far effect to optimize user fairness. The doubly near-far effect comes from the fact that users away from the HAP collect less energy than users close to the HAP on one hand, and need to consume more energy to achieve a better fairness on the other hand.

To avoid the doubly near-far effect in HAP-assisted networks, one of the ways is to separate the co-located energy transmitter and the information receiver at the HAP,

forming the power beacon (PB)-assisted networks, in which the PB is only responsible for charging the energy-constrained nodes [30]. Another advantage of PB-assisted networks is that PB has low cost and is applicable to large-scale deployment since PB does not need backhaul links or complex computation unit to process data.

In simultaneous wireless information and power transfer (SWIPT) networks, the energy-constrained receivers can simultaneously extract energy and information from the received information-bearing signals [26]. Zhou *et al.* first propose two practical SWIPT schemes, i.e., the power-splitting (PS) and time-splitting (TS) schemes [31]. In particular, these two schemes split the received signal into two parts in the power domain and time domain, respectively, where one part is used for EH and the other is for information processing. Based on these two basic SWIPT schemes, Atapattu *et al.* propose a hybrid SWIPT (H-SWIPT) scheme in which the EH-based nodes are allowed to split the received signal in both the power and time domains for EH [32]. Note that by dynamically adjusting the TS and PS ratios, the H-SWIPT scheme gets reduced to the PS or TS scheme in some special cases. More importantly, the H-SWIPT scheme is expected to outperform the TS and PS schemes since it can exploit both the degree-of-freedom in time and power domains.

1.3.3.2 PB-Assisted and SWIPT-Assisted NOMA Networks



(a) PB-assisted Uplink NOMA networks. (b) Cooperative SWIPT NOMA networks.

Figure 1.6: PB-assisted and SWIPT-assisted NOMA networks.

Figs. 1.6(a) and 1.6(b) respectively present a PB-assisted uplink NOMA network and a cooperative SWIPT NOMA network. Specifically, in the former, the PB first transfers energy to the energy-constrained users in the downlink, then the users use

the harvested energy to transmit information to the destination in a NOMA way. In the latter, the information source communicates with the users under the help of an energy-constrained relay. Therefore, the relay first extracts energy and information simultaneously from the received information-bearing signals by applying different SWIPT schemes, and then uses the harvested energy to forward the information to the users in a NOMA way.

1.4 Research Problems and Contributions

Although considerable efforts have been made to study the applications of NOMA in various wireless networks, there are still some open problems that need to be solved. In this thesis, we focus on the applications of NOMA in cooperative networks, cognitive networks, and wireless powered networks, and study four different problems.

1.4.1 Application of Incremental Relaying in Cooperative NOMA Networks

In conventional user-aided cooperative NOMA networks, the strong user serving as a relay generally employs the fixed relaying protocol to help the weak user. However, due to the half-duplex constraint, the strong user needs to occupy half of the transmission block to forward information, which may lead to a loss of spectral efficiency. This observation motivates us to study how to enjoy the benefits of the cooperative technique in NOMA-based systems while avoiding the possible loss of spectral efficiency caused by the fixed relaying. Recall that in conventional cooperative networks, incremental relaying can achieve higher spectral efficiency than the fixed relaying since it can avoid unnecessary cooperation by checking the channel condition between the source and the destination. Inspired by this fact, we introduce the idea of incremental relaying into cooperative NOMA networks and propose an incremental cooperative NOMA protocol in Chapter 2. Specifically, the proposed protocol is new and practical, which only introduces an additional 1-bit feedback compared to the conventional cooperative NOMA protocol. For the proposed protocol, we derive exact or tightly approximated closed-form expressions of the outage probability of each user and the overall system. We prove that the proposed protocol outperforms the conventional one in terms of each user's outage probability and the system outage probability. In

addition, asymptotic outage behavior of the proposed protocol is studied to derive the diversity order of each user and the optimal power allocation strategy that minimizes the system outage probability.

1.4.2 Channel-Aware Power Allocation and Decoding Order in Overlay Cognitive NOMA Networks

In overlay cognitive NOMA networks, the primary and secondary users are grouped together to perform NOMA transmission. Note that the primary users have a higher priority to be served since they are the licensed users of the spectrum, thus the use of the spectrum by the secondary users (unlicensed users) should improve or at least not compromise the performance of the primary users. In this context, how to optimize the performance of the secondary users while ensuring different QoS requirements of the primary users becomes a problem worth studying. This research target is quite different from the conventional NOMA networks without user priority difference where sum rate or fair rate is more concerned. As we mentioned in Section 1.2.3, user ordering and power allocation are two key problems in NOMA-based systems since they jointly decide the achievable rate of each user. Therefore, although the issue of joint optimal user ordering and power allocation has been sufficiently studied in conventional NOMA networks, it is necessary to reconsider this issue in overlay cognitive NOMA networks because these two types of NOMA networks have quite different research targets.

For NOMA networks with user priority difference, several works have been done and some heuristic designs of user ordering and power allocation have been proposed. However, all of them fail to optimize the performance of the secondary users while strictly satisfying the QoS requirements of the primary users. To fill this research gap, in Chapter 3, we study the joint optimal design of user ordering and power allocation for an overlay cognitive NOMA network. Specifically, depending on the availability of channel state information at the transmitter (CSIT), we study two different optimization problems and propose two different strategies. When full CSIT is available, we intend to minimize the outage probability of the secondary user while achieving the target data rate of the primary user. The joint optimal user ordering and power allocation in this case is obtained in closed form. When only partial CSIT

is available, we intend to minimize the outage probability of the secondary user while ensuring that the outage probability of the primary user is less than a predetermined threshold. In this case, the joint optimal solution of user ordering and power allocation is obtained in semi-closed form. Moreover, closed-form outage probability expressions of the primary user and secondary user for the proposed strategies are derived.

1.4.3 Optimal Designs for Relay-Assisted NOMA Networks with Hybrid SWIPT Scheme

In Section 1.3.1, we have introduced that the application of cooperative technique can efficiently extend the coverage of NOMA-based communication. However, it should be noticed that this advantage is realized at the cost of additional energy consumption at the relay, which may hinder intermediate nodes participating in cooperation, especially for those energy-constrained nodes. To alleviate this concern, SWIPT technique that enables the receivers to simultaneously extract energy and information from the received signal can be integrated at energy-constrained relays.

As we have introduced in Section 1.3.3.1, there are three SWIPT schemes, i.e., the PS, the TS, and the H-SWIPT schemes. So far, relay-assisted NOMA networks with the PS or the TS scheme have been widely studied. However, there has been no reported literature investigating the relay-assisted NOMA networks with the H-SWIPT scheme. The H-SWIPT scheme is superior to the PS and TS schemes since it can exploit both the degree-of-freedom in the time and power domains. In view of this fact, we introduce the H-SWIPT scheme into relay-assisted NOMA networks and study two different resource allocation problems in Chapter 4.

We aim to realize different targets under different availability of CSIT by jointly optimizing the transmit power of the source, the PS and TS ratios, the power allocation ratios at the source and the relay, and the user ordering. In particular, when full CSIT is available, we intend to realize a successful transmission with the minimum energy consumption at the source. When only partial CSIT is available, we aim to minimize the system outage probability. Although the first optimization problem is non-convex, we successfully find the joint optimal solution using the efficient bisection method. For the second optimization problem, the joint optimal solution can be obtained with a 1-D search.

1.4.4 Fair Resource Allocation in PB-Assisted NOMA Networks

In downlink NOMA networks, power allocation plays an important role in determining user fairness. Take a two-user downlink as an example, more power should be allocated to the weak user from a fairness standpoint. In addition to power allocation, in PB-assisted downlink NOMA networks, time allocation is another key factor affecting the performance of user fairness. This is because in PB-assisted NOMA networks, the energy-constrained source first collects energy from the PB, and then sends information to the information receivers. Given one resource block, if the source spends too much time on harvesting energy, then there left less time for information transmission. On the other hand, if the EH time is too short, the harvested energy may be too little to support long-term information transmission. Therefore, to optimize the fairness performance of the PB-assisted networks, there exists a balance in the time allocation of energy transmission and information transmission.

Based on the aforementioned observations, in Chapter 5, we study a fair resource allocation problem in a PB-assisted downlink NOMA network by jointly optimizing the time allocation and power allocation ratios. Although the formulated optimization problem is non-convex, we design an alternating optimization algorithm to approximate the joint optimal solution.

1.5 Thesis Outline

The thesis is organized as follows. Chapter 2 introduces the incremental relaying protocol into a user-aided cooperative NOMA network and studies the optimal power allocation strategy. In Chapter 3, the joint optimal design of user ordering and power allocation for an overlay cognitive NOMA network is discussed. Chapters 4 and 5 solve different resource allocation problems involved with the cooperative SWIPT NOMA networks and PB-assisted NOMA networks. Chapter 6 concludes our research results and proposes some possible future research directions.

Chapter 2

Cooperative NOMA with Incremental Relaying: Performance Analysis and Optimization

In conventional cooperative NOMA networks, loss of spectral efficiency may occur due to the half-duplex constraint. To address this issue, we propose an incremental cooperative NOMA (ICN) protocol for a two-user downlink network. In particular, this protocol allows the source to adaptively switch between a direct NOMA transmission mode and a cooperative NOMA transmission mode according to a 1-bit feedback from the far user. We analytically prove that the proposed ICN protocol outperforms the conventional cooperative NOMA protocol. In addition, an optimal power allocation strategy at the source is studied to minimize the asymptotic system outage probability. Finally, numerical results validate our theoretical analysis, present insights, and quantify the enhancement achieved over the benchmark scheme.

2.1 Introduction

In conventional user-aided cooperative NOMA networks, the strong user adopts the fixed relaying protocol to forward the weak user's information. Although the cooperation is expected to improve the weak user's performance by increasing the diversity gain, the spectral efficiency of the networks may suffer a loss. This is because the strong user in the conventional cooperative NOMA (CCN) protocol [19] needs half of its time to forward information due to the constraint of half-duplex. Note that

the major drawback of the CCN protocol is that the relay always participates in the cooperation regardless of the channel condition between the source and the weak user, which fails to make efficient use of the degree-of-freedom of the channel. To efficiently exploit the degree-of-freedom of the channel in a two-user downlink NOMA network, the work in [33] proposes a new cooperative protocol, termed as relaying with NOMA backhaul. In this protocol, the source can adaptively adjust the time durations of NOMA transmission and relay transmission based on global instantaneous channel state information (CSI). However, global instantaneous CSI at the source may be difficult or costly to obtain in practice. This observation motivates us to propose a new and practically viable cooperative protocol for a two-user downlink NOMA network to improve spectral efficiency of the CCN protocol.

Recall that in conventional cooperative networks, the incremental relaying protocol [17] is widely adopted since it can achieve higher spectral efficiency by introducing a negligible 1-bit feedback overhead to avoid unnecessary cooperation. Specifically, the incremental relaying protocol invokes a relay for cooperation only when the source-to-destination channel gain is below a predetermined threshold. Inspired by this feature, in this chapter we propose an ICN protocol for a two-user downlink NOMA network with only statistical CSI at the source. In this protocol, the strong user works as a half-duplex relay only when the weak user broadcasts a 1-bit negative feedback.

The main contributions of this chapter can be summarized as follows. 1) We propose a new and practical cooperative protocol for two-user downlink NOMA networks. To the best of our knowledge, the proposed ICN is the first time that the incremental relaying protocol is introduced into NOMA networks. 2) For the proposed ICN protocol, we derive exact or tightly approximated closed-form expressions of the outage probability of each user and the overall system. We prove that the ICN protocol outperforms the CCN protocol in terms of each user's outage probability and the system outage probability. 3) Asymptotic outage behavior of the ICN protocol is studied to derive the diversity order of each user and an optimal power allocation strategy that minimizes the system outage probability. 4) Valuable insights regarding the ICN protocol are provided through detailed theoretical analysis and numerical results.

Table 2.1: Description of symbols used for key parameters in Chapter 2.

Symbol	Meaning
$h_i, i = 1, 2$	Channel coefficient from the source to user i
h_3	Channel coefficient from user 1 (the near user) to user 2 (the far user)
$x_i, i = 1, 2$	Message for user i
P_s	Transmit power of the source
P_r	Transmit power of user 1
$\alpha_i, i = 1, 2$	Power allocation factor at the source
$\gamma_{1,2}$	Received SINR at user 1 to detect the message of user 2
$\gamma_{1,1}$	Received SNR at user 1 to detect the message of user 1
$\gamma_{2,2}$	Received SINR at user 2 to detect the message of user 2 in the direct NOMA transmission mode
$\gamma_{2,2}^{\text{MRC}}$	Received SINR at user 2 to detect the message of user 2 in the cooperative NOMA transmission mode
γ_{th}	Decoding threshold in the direct NOMA transmission mode
γ_{th}'	Decoding threshold in the cooperative NOMA transmission mode
$P_i^{\text{ICN}}, i = 1, 2$	Exact outage probability of user i in the proposed ICN protocol
$P_{1\&2}^{\text{ICN}}$	Exact overall system outage probability in the proposed ICN protocol
$P_{i,\text{asy}}^{\text{ICN}}, i = 1, 2$	Asymptotic outage probability of user i in the proposed ICN protocol

2.2 System Model

We consider a two-user downlink NOMA scenario with a source (S) and two users: user 1 (U_1) is the near user while user 2 (U_2) is the far user. Similar to [34, 35], the two users are ordered according to their distance to S. Thus, U_1 and U_2 are treated as the strong user and the weak user, respectively. All the channels suffer Rayleigh fading. Let h_1, h_2 and h_3 denote the channel coefficients from S to U_1, S to U_2, and U_1 to U_2, respectively, where $h_i \sim \mathcal{CN}(0, \Omega_i)$ $(i = 1, 2, 3)$. We assume that channel coefficients remain unchanged during one transmission block, but may vary from one transmission block to another. Next we introduce the proposed ICN protocol in details.

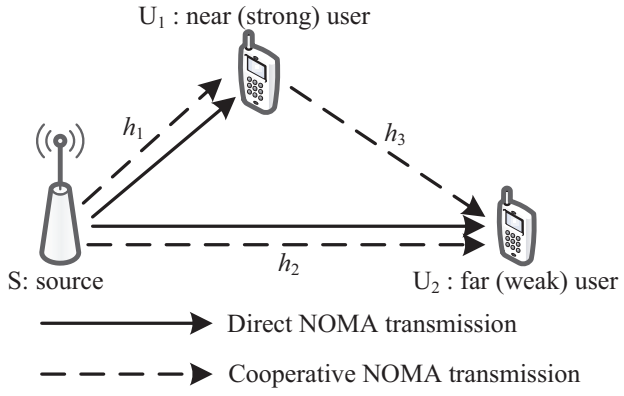


Figure 2.1: System model.

2.2.1 Incremental Cooperative NOMA Protocol

At the beginning of each transmission block, S broadcasts a pilot signal to U_1 and U_2. Based on the received pilot signal, U_2 performs channel estimation of h_2 and compares it with a predefined threshold. If U_2 judges that it can correctly decode its desired message through direct transmission, it feedbacks a 1-bit positive acknowledgement to S and U_1. After receiving the positive acknowledgement feedback, S adopts a *direct NOMA transmission mode*, i.e., it sends the superimposed signal to U_1 and U_2 within the whole transmission block. If U_2 finds that it is unable to decode its desired message without U_1's cooperation, it feedbacks a 1-bit negative acknowledgement to S and U_1. Upon hearing the negative acknowledgement feedback, S adopts a *cooperative NOMA transmission mode*, i.e., it broadcasts the superimposed signal in the first half of the transmission block, and then U_1 decodes U_2's message and forwards it in the second half of the transmission block.

To identify the difference between our proposed ICN and the CCN protocols, here we briefly review the CCN protocol [19]. In the CCN protocol, the transmission block is divided into two phases with equal duration. During the first phase, S sends the superimposed signal to U_1 and U_2, and U_1 decodes U_2's message and forwards it in the second phase. Compared to the CCN protocol, our proposed ICN protocol is essentially an adaptive protocol which can adaptively switch between the direct NOMA transmission mode and the cooperative NOMA transmission mode based on a 1-bit indicator.[1]

2.2.2 Signal Model

1) Direct NOMA transmission mode: S sends a superimposed signal to U_1 and U_2, which occupies the whole transmission block. The resulted signal at U_n is defined by

$$y_n = \sqrt{\alpha_1 P_s} h_n x_1 + \sqrt{\alpha_2 P_s} h_n x_2 + w_n,\ n = 1, 2, \tag{2.1}$$

where P_s is the transmit power of S, x_n denotes the message for U_n, α_n is the power allocation factor for x_n with $\alpha_1 + \alpha_2 = 1$, and w_n is the AWGN at U_n with zero mean

[1] In the CCN protocol, both S and U_1 need to send pilot signals, for channel estimation at the receiver side(s). In the ICN protocol, only S sends a pilot signal in the direct NOMA transmission mode, while both S and U_1 send pilot signals in the cooperative NOMA transmission mode. Thus, the signaling overhead of the two protocols are comparable to each other.

and variance σ^2.

According to the NOMA principle, U_n first decodes x_2 upon observing y_n. Denote $\gamma_{n,2}$ as the received signal-to-interference-pulse-noise ratio (SINR) at U_n to decode x_2, and then $\gamma_{n,2}$ is given by $\gamma_{n,2} = \frac{\alpha_2 \rho_s |h_n|^2}{\alpha_1 \rho_s |h_n|^2 + 1}$, where $\rho_s = P_s/\sigma^2$ denotes the transmit signal-to-noise ratio (SNR) of S. After U_1 successfully decodes x_2 and performs SIC, the received SNR to detect x_1 at U_1, denoted by $\gamma_{1,1}$, is $\gamma_{1,1} = \alpha_1 \rho_s |h_1|^2$.

2) Cooperative NOMA transmission mode: Here the entire transmission block consists of two phases with equal duration. In the first phase, the received signal at U_n is the same as defined in (2.1), and the received SINR at U_n for message x_2 is also given as $\gamma_{n,2}$ defined in the direct NOMA transmission mode. If U_1 successfully decodes x_2 and performs SIC in the first phase, its received SNR to detect x_1 is given as $\gamma_{1,1}$ defined in the direct NOMA transmission mode. Then, in the second phase, U_1 forwards the re-encoded x_2 to U_2. The corresponding received signal at U_2 in the second phase can be expressed as $y_2' = \sqrt{P_r} h_3 x_2 + w_2$, where P_r is the transmit power of U_1. Finally, U_2 combines the observed signals y_2 and y_2' using the maximal ratio combining (MRC), and thus, the received SINR at U_2 to decode x_2 after MRC is given by $\gamma_{2,2}^{\text{MRC}} = \frac{\alpha_2 \rho_s |h_2|^2}{\alpha_1 \rho_s |h_2|^2 + 1} + \rho_r |h_3|^2$, where $\rho_r = P_r/\sigma^2$ is U_1's transmit SNR.

2.3 Outage Performance Analysis and Optimization

For each user, an outage event happens when the received SINR (or SNR) is below a pre-determined decoding threshold. Note that the decoding thresholds of the direct NOMA transmission and the cooperative NOMA transmission modes are different. In the direct NOMA transmission mode, the decoding threshold is $\gamma_{\text{th}} = 2^R - 1$ with R being the target rate of x_1 and x_2. In the cooperative NOMA transmission mode, the threshold is $\gamma_{\text{th}}' = 2^{2R} - 1$.

2.3.1 Outage Probability Analysis

2.3.1.1 Near User

According to the ICN protocol, the outage probability of U_1 can be expressed as

$$\begin{aligned} P_1^{\text{ICN}} = 1 &- \Pr\left\{\gamma_{2,2} \geq \gamma_{\text{th}}, \gamma_{1,2} \geq \gamma_{\text{th}}, \gamma_{1,1} \geq \gamma_{\text{th}}\right\} \\ &- \Pr\left\{\gamma_{2,2} < \gamma_{\text{th}}, \gamma_{1,2} \geq \gamma_{\text{th}}', \gamma_{1,1} \geq \gamma_{\text{th}}'\right\}, \end{aligned} \tag{2.2}$$

where $\gamma_{2,2} \geq \gamma_{\text{th}}$ indicates that the system works in the direct NOMA transmission mode, and $\gamma_{2,2} < \gamma_{\text{th}}$ indicates that the system works in the cooperative NOMA transmission mode. As $\gamma_{2,2}$ is independent from $\gamma_{1,2}$ and $\gamma_{1,1}$, (2.2) can be rewritten as

$$\begin{aligned} P_1^{\text{ICN}} = 1 - &\underbrace{\Pr\{\gamma_{2,2} \geq \gamma_{\text{th}}\}}_{Q_1} \underbrace{\Pr\{\gamma_{1,2} \geq \gamma_{\text{th}}, \gamma_{1,1} \geq \gamma_{\text{th}}\}}_{Q_2} \\ &- \underbrace{\Pr\{\gamma_{2,2} < \gamma_{\text{th}}\}}_{\bar{Q}_1} \underbrace{\Pr\{\gamma_{1,2} \geq \gamma'_{\text{th}}, \gamma_{1,1} \geq \gamma'_{\text{th}}\}}_{Q_3}, \end{aligned} \tag{2.3}$$

where $\bar{Q}_1 = 1 - Q_1$. It is easy to verify that $Q_1 = Q_2 = 0$ for $\frac{1}{1+\gamma_{\text{th}}} \leq \alpha_1 < 1$, and $Q_3 = 0$ for $\frac{1}{1+\gamma'_{\text{th}}} \leq \alpha_1 < 1$. Thus, $P_1^{\text{ICN}} = 1$ for $\frac{1}{1+\gamma_{\text{th}}} \leq \alpha_1 < 1$. When $0 < \alpha_1 < \frac{1}{1+\gamma_{\text{th}}}$, Q_2 is given by

$$Q_2 = \Pr\left\{|h_1|^2 \geq \frac{\gamma_{\text{th}}}{\rho_s(\alpha_2 - \gamma_{\text{th}}\alpha_1)}, |h_1|^2 \geq \frac{\gamma_{\text{th}}}{\alpha_1 \rho_s}\right\} = \Pr\left\{|h_1|^2 \geq \frac{\gamma_{\text{th}}}{\rho_s \Theta}\right\} = e^{-\frac{\gamma_{\text{th}}}{\rho_s \Omega_1 \Theta}}, \tag{2.4}$$

where $\Theta \triangleq \min\{\theta, \alpha_1\}$ and $\theta \triangleq \alpha_2 - \gamma_{\text{th}}\alpha_1$. Q_2 is derived using the fact that $|h_i|^2$ $(i = 1, 2, 3)$ follows exponential distribution with mean Ω_i. Following similar steps, we have $Q_1 = e^{-\frac{\gamma_{\text{th}}}{\rho_s \Omega_2 \theta}}$ for $0 < \alpha_1 < \frac{1}{1+\gamma_{\text{th}}}$, and $Q_3 = e^{-\frac{\gamma'_{\text{th}}}{\rho_s \Omega_1 \Theta'}}$ for $0 < \alpha_1 < \frac{1}{1+\gamma'_{\text{th}}}$, where $\Theta' \triangleq \min\{\theta', \alpha_1\}$ and $\theta' \triangleq \alpha_2 - \gamma'_{\text{th}}\alpha_1$. Substituting the results of Q_1, Q_2 and Q_3 into (2.3), a closed-form expression of U_1's outage probability is given by

$$P_1^{\text{ICN}} = \begin{cases} 1 - e^{-\frac{\gamma_{\text{th}}}{\rho_s}\left(\frac{1}{\Omega_1 \Theta} + \frac{1}{\Omega_2 \theta}\right)} - e^{-\frac{\gamma'_{\text{th}}}{\rho_s \Omega_1 \Theta'}} + e^{-\frac{\gamma_{\text{th}}}{\rho_s \Omega_2 \theta}} e^{-\frac{\gamma'_{\text{th}}}{\rho_s \Omega_1 \Theta'}}, & 0 < \alpha_1 < \frac{1}{1+\gamma'_{\text{th}}}, \\ 1 - e^{-\frac{\gamma_{\text{th}}}{\rho_s}\left(\frac{1}{\Omega_1 \Theta} + \frac{1}{\Omega_2 \theta}\right)}, & \frac{1}{1+\gamma'_{\text{th}}} \leq \alpha_1 < \frac{1}{1+\gamma_{\text{th}}}, \\ 1, & \frac{1}{1+\gamma_{\text{th}}} \leq \alpha_1 < 1. \end{cases} \tag{2.5}$$

2.3.1.2 Far User

The outage probability of U_2 with the ICN protocol is given by

$$\begin{aligned} P_2^{\text{ICN}} &= \Pr\{\gamma_{2,2} < \gamma_{\text{th}}, \gamma_{1,2} < \gamma'_{\text{th}}\} + \Pr\left\{\gamma_{2,2} < \gamma_{\text{th}}, \gamma_{1,2} \geq \gamma'_{\text{th}}, \gamma_{2,2}^{\text{MRC}} < \gamma'_{\text{th}}\right\} \\ &= \underbrace{\Pr\{\gamma_{2,2} < \gamma_{\text{th}}\}}_{\bar{Q}_1} \underbrace{\Pr\{\gamma_{1,2} < \gamma'_{\text{th}}\}}_{Q_4} + \underbrace{\Pr\{\gamma_{1,2} \geq \gamma'_{\text{th}}\}}_{\bar{Q}_4} \underbrace{\Pr\left\{\gamma_{2,2} < \gamma_{\text{th}}, \gamma_{2,2}^{\text{MRC}} < \gamma'_{\text{th}}\right\}}_{Q_5}, \end{aligned} \tag{2.6}$$

where $\bar{Q}_4 = 1 - Q_4$. Similar to U_1's outage probability, the outage probability of U_2 is also segmented regarding α_1 as follows.

When $\frac{1}{1+\gamma_{\text{th}}} \leq \alpha_1 < 1$, we have $P_2^{\text{ICN}} = 1$ since $\bar{Q}_1 = Q_4 = 1$. When $\frac{1}{1+\gamma'_{\text{th}}} \leq \alpha_1 < \frac{1}{1+\gamma_{\text{th}}}$, we have $Q_4 = 1$ and thus, $P_2^{\text{ICN}} = \bar{Q}_1 = 1 - e^{-\frac{\gamma_{\text{th}}}{\rho_s \Omega_2 \theta}}$, which is an

increasing function of α_1. Now we derive P_2^{ICN} over the region $\alpha_1 \in \left(0, \frac{1}{1+\gamma'_{\text{th}}}\right)$, where $Q_4 = 1 - e^{-\frac{\gamma'_{\text{th}}}{\rho_s \Omega_1 \theta'}}$ and Q_5 can be derived as

$$\begin{aligned} Q_5 =& \Pr\left\{\frac{\alpha_2\rho_s|h_2|^2}{\alpha_1\rho_s|h_2|^2+1} < \gamma_{\text{th}}, \rho_r|h_3|^2 + \frac{\alpha_2\rho_s|h_2|^2}{\alpha_1\rho_s|h_2|^2+1} < \gamma'_{\text{th}}\right\} \\ =& \int_0^{\frac{\gamma_{\text{th}}}{\rho_s\theta}} F_{|h_3|^2}\left(\frac{\gamma'_{\text{th}}}{\rho_r} - \frac{\alpha_2\rho_s x}{\rho_r\left(\alpha_1\rho_s x+1\right)}\right) f_{|h_2|^2}(x)\, dx \\ =& 1 - e^{-\frac{\gamma_{\text{th}}}{\rho_s\Omega_2\theta}} - \underbrace{\int_0^{\frac{\gamma_{\text{th}}}{\rho_s\theta}} e^{-\frac{1}{\rho_r\Omega_3}\left(\gamma'_{\text{th}} - \frac{\alpha_2\rho_s x}{\alpha_1\rho_s x+1}\right)} \frac{1}{\Omega_2} e^{-\frac{x}{\Omega_2}} dx}_{Q_6}. \end{aligned} \tag{2.7}$$

Though it is difficult to derive a closed-form expression for Q_6, we can obtain an approximation for it. By replacing the variable $x = \frac{\gamma_{\text{th}}}{2\rho_s\theta}(t+1)$ in Q_6 and using Gaussian-Chebyshev quadrature [36, Eq. 25.4.38], we have

$$Q_6 = \frac{\gamma_{\text{th}}}{2\rho_s\Omega_2\theta}\int_{-1}^{1} e^{-\frac{g(t)}{\rho_r\Omega_3}} e^{-\frac{\gamma_{\text{th}}(t+1)}{2\rho_s\Omega_2\theta}} dt \approx \frac{\gamma_{\text{th}}}{2\rho_s\Omega_2\theta}\frac{\pi}{K}\sum_{k=1}^{K}\sqrt{1-\xi_k^2} e^{-\frac{g(\xi_k)}{\rho_r\Omega_3}} e^{-\frac{\gamma_{\text{th}}(\xi_k+1)}{2\rho_s\Omega_2\theta}}, \tag{2.8}$$

where K is a parameter to balance accuracy and complexity, $\xi_k = \cos\left(\frac{2k-1}{2K}\pi\right)$, and $g(x) = \gamma'_{\text{th}} - \frac{\gamma_{\text{th}}(x+1)\alpha_2}{\gamma_{\text{th}}(x+1)\alpha_1+2\theta}$. Substituting (2.8) into (2.7), we can obtain an approximation of Q_5.

Combining the results for Q_1, Q_4 and Q_5, and after some algebraic manipulations, a closed-form expression of approximated P_2^{ICN} over the region $\alpha_1 \in \left(0, \frac{1}{1+\gamma'_{\text{th}}}\right)$ is given by

$$P_2^{\text{ICN}} \approx 1 - e^{-\frac{\gamma_{\text{th}}}{\rho_s\Omega_2\theta}} - e^{-\frac{\gamma'_{\text{th}}}{\rho_s\Omega_1\theta'}} Q_6, \tag{2.9}$$

where Q_6 is given by (2.8).

From the above derivations, we know that P_1^{ICN} and P_2^{ICN} are both equal to 1 when $\frac{1}{1+\gamma_{\text{th}}} \leq \alpha_1 < 1$. Thus, in the sequel we only focus on the remaining region, i.e., $0 < \alpha_1 < \frac{1}{1+\gamma_{\text{th}}}$.

2.3.1.3 Overall System

Similar to [19], the system outage is defined as the event when one user or both users in the system are in outage. Thus, the system outage probability with the ICN

protocol can be expressed as

$$P_{1\&2}^{\mathrm{ICN}} = 1 - \Pr\left\{\gamma_{2,2} \geq \gamma_{\mathrm{th}}, \gamma_{1,2} \geq \gamma_{\mathrm{th}}, \gamma_{1,1} \geq \gamma_{\mathrm{th}}\right\} - \Pr\left\{\gamma_{2,2} < \gamma_{\mathrm{th}}, \gamma_{1,2} \geq \gamma_{\mathrm{th}}', \gamma_{1,1} \geq \gamma_{\mathrm{th}}', \gamma_{2,2}^{\mathrm{MRC}} \geq \gamma_{\mathrm{th}}'\right\}. \tag{2.10}$$

Following similar procedures to those in the derivations of P_1^{ICN} and P_2^{ICN}, a closed-form approximation of the system outage probability can be given as

$$P_{1\&2}^{\mathrm{ICN}} = \begin{cases} 1 - e^{-\frac{\gamma_{\mathrm{th}}}{\rho_s}\left(\frac{1}{\Omega_1\Theta} + \frac{1}{\Omega_2\theta}\right)} - Q_6 e^{-\frac{\gamma_{\mathrm{th}}'}{\rho_s\Omega_1\Theta'}}, 0 < \alpha_1 < \frac{1}{1+\gamma_{\mathrm{th}}'}, \\ 1 - e^{-\frac{\gamma_{\mathrm{th}}}{\rho_s}\left(\frac{1}{\Omega_1\Theta} + \frac{1}{\Omega_2\theta}\right)}, \quad \frac{1}{1+\gamma_{\mathrm{th}}'} \leq \alpha_1 < \frac{1}{1+\gamma_{\mathrm{th}}}, \end{cases} \tag{2.11}$$

where Q_6 is given by (2.8). Comparing the expressions of P_1^{ICN} and $P_{1\&2}^{\mathrm{ICN}}$ given in (2.5) and (2.11), respectively, we find that the outage probability of U_1 is identical to the system outage probability when $\alpha_1 \in \left[\frac{1}{1+\gamma_{\mathrm{th}}'}, \frac{1}{1+\gamma_{\mathrm{th}}}\right)$. In other words, when the overall system is in outage, it also means that U_1 is in outage. This is due to the following two facts: 1) The system works in the direct NOMA transmission mode only when U_2 can correctly decode its desired information (which means that U_2 has no outage). In this case, U_1 in outage also leads to an outage of the overall system. 2) When U_2 requests cooperation (which indicates that the target rate of U_2 cannot be achieved in the direct NOMA transmission mode), if $\alpha_1 \in \left[\frac{1}{1+\gamma_{\mathrm{th}}'}, \frac{1}{1+\gamma_{\mathrm{th}}}\right)$, we have $\gamma_{1,2} < \gamma_{\mathrm{th}}'$, i.e., U_1 fails to decode x_2, which results in an outage at both U_1 and U_2.

2.3.2 Outage Performance Comparison with the CCN protocol

We denote the outage probability of U_1, U_2, and the overall system in the CCN protocol by P_1^{CCN}, P_2^{CCN} and $P_{1\&2}^{\mathrm{CCN}}$, respectively. Following the CCN protocol details from [19] along with the expressions of P_1^{ICN}, P_2^{ICN} and $P_{1\&2}^{\mathrm{ICN}}$ given in (2.2), (2.6) and (2.10), respectively, we have

$$\begin{aligned} P_1^{\mathrm{ICN}} &< 1 - \Pr\left\{\gamma_{2,2} \geq \gamma_{\mathrm{th}}, \gamma_{1,2} \geq \gamma_{\mathrm{th}}', \gamma_{1,1} \geq \gamma_{\mathrm{th}}'\right\} - \Pr\left\{\gamma_{2,2} < \gamma_{\mathrm{th}}, \gamma_{1,2} \geq \gamma_{\mathrm{th}}', \gamma_{1,1} \geq \gamma_{\mathrm{th}}'\right\} \\ &= 1 - \Pr\left\{\gamma_{1,2} \geq \gamma_{\mathrm{th}}', \gamma_{1,1} \geq \gamma_{\mathrm{th}}'\right\} = P_1^{\mathrm{CCN}}, \end{aligned} \tag{2.12}$$

$$P_2^{\mathrm{ICN}} < \Pr\left\{\gamma_{2,2} < \gamma_{\mathrm{th}}', \gamma_{1,2} < \gamma_{\mathrm{th}}'\right\} + \Pr\left\{\gamma_{1,2} \geq \gamma_{\mathrm{th}}', \gamma_{2,2}^{\mathrm{MRC}} < \gamma_{\mathrm{th}}'\right\} = P_2^{\mathrm{CCN}}, \tag{2.13}$$

and

$$\begin{aligned} P_{1\&2}^{\mathrm{ICN}} &< 1 - \Pr\left\{\gamma_{2,2} \geq \gamma_{\mathrm{th}}, \gamma_{1,2} \geq \gamma'_{\mathrm{th}}, \gamma_{1,1} \geq \gamma'_{\mathrm{th}}, \gamma_{2,2}^{\mathrm{MRC}} \geq \gamma'_{\mathrm{th}}\right\} \\ &\quad - \Pr\left\{\gamma_{2,2} < \gamma_{\mathrm{th}}, \gamma_{1,2} \geq \gamma'_{\mathrm{th}}, \gamma_{1,1} \geq \gamma'_{\mathrm{th}}, \gamma_{2,2}^{\mathrm{MRC}} \geq \gamma'_{\mathrm{th}}\right\} \\ &= 1 - \Pr\left\{\gamma_{1,2} \geq \gamma'_{\mathrm{th}}, \gamma_{1,1} \geq \gamma'_{\mathrm{th}}, \gamma_{2,2}^{\mathrm{MRC}} \geq \gamma'_{\mathrm{th}}\right\} = P_{1\&2}^{\mathrm{CCN}}. \end{aligned} \tag{2.14}$$

Therefore, it can be concluded that the ICN protocol outperforms the CCN protocol in terms of each user's outage probability and the system outage probability.

2.3.3 System Outage Probability Minimization and Diversity Order Analysis

In this subsection, we first investigate the asymptotic outage performance of the ICN protocol when $\rho_s \to \infty$ and $\rho_r = \lambda\rho_s$ with $0 < \lambda \leq 1$. Based on the asymptotic analysis, an optimal power allocation strategy that minimizes the system outage probability is developed, and the diversity order of each user is derived as well.

2.3.3.1 System Outage Probability Minimization

As $\rho_s \to \infty$, we have $\gamma_{2,2}^{\mathrm{MRC}} \to \frac{\alpha_2}{\alpha_1} + \rho_r |h_3|^2 > \gamma'_{\mathrm{th}}$ for $0 < \alpha_1 < \frac{1}{1+\gamma'_{\mathrm{th}}}$, which indicates that $\Pr\left\{\gamma_{2,2}^{\mathrm{MRC}} > \gamma'_{\mathrm{th}}\right\} \to 1$, and thus, $P_{1\&2}^{\mathrm{ICN}}$ converges to P_1^{ICN} based on (2.2) and (2.10). Together with the fact that U_1's outage probability is identical to the system outage probability when $\alpha_1 \in \left[\frac{1}{1+\gamma'_{\mathrm{th}}}, \frac{1}{1+\gamma_{\mathrm{th}}}\right)$, it can be concluded that the system outage probability converges to U_1's outage probability as $\rho_s \to \infty$. Noting this key observation, in the following we focus on the minimization of U_1's outage probability.

When $\rho_s \to \infty$, applying $e^{-x} \overset{x\to 0}{\simeq} 1-x$ into (2.5), we can derive the asymptotic outage probability of U_1 as

$$P_{1,\mathrm{asy}}^{\mathrm{ICN}} \simeq \begin{cases} \frac{\gamma_{\mathrm{th}}}{\rho_s \Omega_1 \Theta} + \frac{\gamma_{\mathrm{th}}\gamma'_{\mathrm{th}}}{\rho_s^2 \Omega_1 \Omega_2 \theta \Theta'}, 0 < \alpha_1 < \frac{1}{1+\gamma'_{\mathrm{th}}}, \\ \frac{\gamma_{\mathrm{th}}}{\rho_s}\left(\frac{1}{\Omega_1\Theta} + \frac{1}{\Omega_2\theta}\right), \frac{1}{1+\gamma'_{\mathrm{th}}} \leq \alpha_1 < \frac{1}{1+\gamma_{\mathrm{th}}}. \end{cases} \tag{2.15}$$

Substituting the expressions of θ, Θ, and Θ' into (2.15), $P_{1,\mathrm{asy}}^{\mathrm{ICN}}$ can be further expressed

as

$$P_{1,\text{asy}}^{\text{ICN}} \simeq \begin{cases} \frac{\gamma_{\text{th}}}{\rho_s \Omega_1} f_1(\alpha_1), 0 < \alpha_1 < \frac{1}{2+\gamma'_{\text{th}}}, \\ \frac{\gamma_{\text{th}}}{\rho_s \Omega_1} f_2(\alpha_1), \frac{1}{2+\gamma'_{\text{th}}} \le \alpha_1 < \min\left\{\frac{1}{2+\gamma_{\text{th}}}, \frac{1}{1+\gamma'_{\text{th}}}\right\}, \\ \frac{\gamma_{\text{th}}}{\rho_s \Omega_1} f_3(\alpha_1), \min\left\{\frac{1}{2+\gamma_{\text{th}}}, \frac{1}{1+\gamma'_{\text{th}}}\right\} \le \alpha_1 < \frac{1}{1+\gamma'_{\text{th}}}, \\ \frac{\gamma_{\text{th}}}{\rho_s} f_4(\alpha_1), \frac{1}{1+\gamma'_{\text{th}}} \le \alpha_1 < \max\left\{\frac{1}{2+\gamma_{\text{th}}}, \frac{1}{1+\gamma'_{\text{th}}}\right\}, \\ \frac{\gamma_{\text{th}}}{\rho_s} f_5(\alpha_1), \max\left\{\frac{1}{2+\gamma_{\text{th}}}, \frac{1}{1+\gamma'_{\text{th}}}\right\} \le \alpha_1 < \frac{1}{1+\gamma_{\text{th}}}, \end{cases} \tag{2.16}$$

in which we have

$$f_1(\alpha_1) = \frac{1}{\alpha_1} + \frac{\gamma'_{\text{th}}}{\rho_s \Omega_2 (1 - \alpha_1 (1 + \gamma_{\text{th}}))\alpha_1}, \tag{2.17}$$

$$f_2(\alpha_1) = \frac{1}{\alpha_1} + \frac{\gamma'_{\text{th}}}{\rho_s \Omega_2 (1 - \alpha_1 (1 + \gamma_{\text{th}}))(1 - \alpha_1 (1 + \gamma'_{\text{th}}))}, \tag{2.18}$$

$$f_3(\alpha_1) = \frac{1}{1 - \alpha_1 (1 + \gamma_{\text{th}})} \left[1 + \frac{\gamma'_{\text{th}}}{\rho_s \Omega_2 (1 - \alpha_1 (1 + \gamma'_{\text{th}}))}\right], \tag{2.19}$$

$$f_4(\alpha_1) = \frac{1}{\Omega_1 \alpha_1} + \frac{1}{\Omega_2 (1 - \alpha_1 (1 + \gamma_{\text{th}}))}, \tag{2.20}$$

$$f_5(\alpha_1) = \frac{1}{1 - \alpha_1 (1 + \gamma_{\text{th}})} \left(\frac{1}{\Omega_1} + \frac{1}{\Omega_2}\right). \tag{2.21}$$

For $f_1(\alpha_1)$: It can be shown that $\frac{1}{\alpha_1(1-\alpha_1(1+\gamma_{\text{th}}))}$ monotonically decreases with $\alpha_1 \in \left(0, \frac{1}{2+2\gamma_{\text{th}}}\right)$. Thus, $f_1(\alpha_1)$ is a decreasing function over $\alpha_1 \in \left(0, \frac{1}{2+\gamma'_{\text{th}}}\right)$ since $\frac{1}{2+\gamma'_{\text{th}}} \le \frac{1}{2+2\gamma_{\text{th}}}$.

For $f_2(\alpha_1)$: $f_2(\alpha_1)$ is a convex function of α_1 due to the facts that $\frac{1}{\alpha_1}$, $\frac{1}{1-\alpha_1(1+\gamma_{\text{th}})}$ and $\frac{1}{1-\alpha_1\left(1+\gamma'_{\text{th}}\right)}$ are convex functions of α_1 and that the sum of convex functions is still a convex function. The first-order derivative of $f_2(\alpha_1)$ is given by

$$\frac{df_2(\alpha_1)}{d\alpha_1} = -\frac{1}{\alpha_1^2} + \frac{\mathrm{a}\left(\mathrm{b}(1 - \mathrm{c}\alpha_1) + \mathrm{c}(1 - \mathrm{b}\alpha_1)\right)}{\left((1 - \mathrm{b}\alpha_1)(1 - \mathrm{c}\alpha_1)\right)^2}, \tag{2.22}$$

where $\mathrm{a} = \frac{\gamma'_{\text{th}}}{\rho_s \Omega_2}$, $\mathrm{b} = 1 + \gamma_{\text{th}}$ and $\mathrm{c} = 1 + \gamma'_{\text{th}}$. From (2.22), we can easily verify that $\frac{df_2(\alpha_1)}{d\alpha_1}|_{\alpha_1 \to 0} < 0$ and $\frac{df_2(\alpha_1)}{d\alpha_1}|_{\alpha_1 \to \frac{1}{1+\gamma'_{\text{th}}}} > 0$. Since $f_2(\alpha_1)$ is a convex function, the critical point of $f_2(\alpha_1)$, denoted as δ, must lie in the interval $\left(0, \frac{1}{1+\gamma'_{\text{th}}}\right)$, and is the root of $\frac{df_2(\alpha_1)}{d\alpha_1} = 0$ that falls in $\left(0, \frac{1}{1+\gamma'_{\text{th}}}\right)$.[2] Thus, for $\frac{1}{2+\gamma'_{\text{th}}} \le \alpha_1 < \min\left\{\frac{1}{2+\gamma_{\text{th}}}, \frac{1}{1+\gamma'_{\text{th}}}\right\}$, the minimal point of $P_{1,\text{asy}}^{\text{ICN}}$ is at $\alpha_1 = \beta_1$ with $\beta_1 \triangleq \max\left\{\frac{1}{2+\gamma'_{\text{th}}}, \min\left\{\delta, \frac{1}{2+\gamma_{\text{th}}}\right\}\right\}$.

[2]Note that $\frac{df_2(\alpha_1)}{d\alpha_1}$ can be transformed to a quartic function of α_1, and the procedures in [37] can be used to find closed-form roots of $\frac{df_2(\alpha_1)}{d\alpha_1} = 0$.

For $f_3(\alpha_1)$: $f_3(\alpha_1)$ is an increasing function of α_1.

For $f_4(\alpha_1)$: Like $f_2(\alpha_1)$, $f_4(\alpha_1)$ is also a convex function of α_1, whose critical point can be obtained as $\alpha_1 = \frac{1}{1+\psi+\gamma_{\rm th}}$, where $\psi = \sqrt{\frac{\Omega_1(1+\gamma_{\rm th})}{\Omega_2}}$. Thus, for $\frac{1}{1+\gamma'_{\rm th}} \leq \alpha_1 < \max\left\{\frac{1}{2+\gamma_{\rm th}}, \frac{1}{1+\gamma'_{\rm th}}\right\}$, the minimal point of $P_{1,\rm asy}^{\rm ICN}$ is at $\alpha_1 = \beta_2$ with $\beta_2 \triangleq \max\left\{\frac{1}{1+\psi+\gamma_{\rm th}}, \frac{1}{1+\gamma'_{\rm th}}\right\}$.

For $f_5(\alpha_1)$: $f_5(\alpha_1)$ is an increasing function of α_1.

Combing all above observations, we conclude that $P_{1,\rm asy}^{\rm ICN}$ achieves its global minimum value at $\alpha_1 = \beta_1$ if $\frac{\gamma_{\rm th}}{\rho_s\Omega_1}f_2(\beta_1) < \frac{\gamma_{\rm th}}{\rho_s}f_4(\beta_2)$, or at $\alpha_1 = \beta_2$ otherwise.

2.3.3.2 Diversity Order of Each User

From (2.15), we can observe that the diversity order of $\rm U_1$ is 1, which is the full diversity order for $\rm U_1$.

As $\rho_s \to \infty$, the asymptotic outage probability of $\rm U_2$ over the region $\alpha_1 \in \left[\frac{1}{1+\gamma'_{\rm th}}, \frac{1}{1+\gamma_{\rm th}}\right)$ can be easily derived as $P_{2,\rm asy}^{\rm ICN} = \bar{Q}_1 \simeq \frac{\gamma_{\rm th}}{\rho_s\Omega_2\theta}$, which illustrates that the diversity order of $\rm U_2$ in this region is 1. The reason for the diversity loss is that in this region of α_1, $\rm U_1$ cannot work in the cooperative mode since $\gamma_{1,2} < \gamma'_{\rm th}$, and thus, it fails to provide assistance to $\rm U_2$.

Now we focus on the derivation of $P_{2,\rm asy}^{\rm ICN}$ when $0 < \alpha_1 < \frac{1}{1+\gamma'_{\rm th}}$. As $\rho_s \to \infty$, Q_6 in (2.7) can be approximated as

$$\begin{aligned} Q_6 &\overset{\rm (i)}{\simeq} \int_0^{\frac{\gamma_{\rm th}}{\rho_s\theta}} \left(1 - \frac{1}{\rho_r\Omega_3}\left(\gamma'_{\rm th} - \frac{\alpha_2\rho_s x}{(\alpha_1\rho_s x+1)}\right)\right)\frac{1}{\Omega_2}e^{-\frac{x}{\Omega_2}}dx \\ &\overset{\rm (ii)}{\simeq} \left(1 - e^{-\frac{\gamma_{\rm th}}{\rho_s\Omega_2\theta}}\right)\left(1 - \frac{\gamma'_{\rm th}}{\rho_r\Omega_3}\right) + \frac{\gamma_{\rm th}}{2\rho_s\rho_r\Omega_2\Omega_3\theta}\frac{\pi}{K} \\ &\quad \times \sum_{k=1}^{K}\sqrt{1-\xi_k^2}\frac{\alpha_2\gamma_{\rm th}(\xi_k+1)}{\alpha_1\gamma_{\rm th}(\xi_k+1)+2\theta}e^{-\frac{\gamma_{\rm th}(\xi_k+1)}{2\rho_s\Omega_2\theta}}, \end{aligned} \tag{2.23}$$

where step (i) is obtained by using $e^{-x} \overset{x\to 0}{\simeq} 1-x$, and step (ii) is achieved by applying the Gaussian-Chebyshev quadrature. Now substituting (2.23) into (2.7) and applying $e^{-x} \overset{x\to 0}{\simeq} 1-x$ again, we have $Q_5 \simeq \frac{\gamma_{\rm th}\Xi}{\lambda\rho_s^2\Omega_2\Omega_3\theta}$, where Ξ is given by

$$\Xi = \gamma'_{\rm th} - \frac{\pi}{2K}\sum_{k=1}^{K}\sqrt{1-\xi_k^2}\frac{\alpha_2\gamma_{\rm th}(\xi_k+1)}{\alpha_1\gamma_{\rm th}(\xi_k+1)+2\theta}. \tag{2.24}$$

In addition, an approximation of $\bar{Q}_1Q_4$ in (2.6) can be easily obtained as $\bar{Q}_1Q_4 \simeq \frac{\gamma_{\rm th}\gamma'_{\rm th}}{\rho_s^2\Omega_1\Omega_2\theta\theta'}$. To this end, by combining the approximate results for $\bar{Q}_1Q_4$ and Q_5, the

asymptotic outage probability of U_2 over the region $\alpha_1 \in \left(0, \frac{1}{1+\gamma'_{\text{th}}}\right)$ is given by

$$P_{2,\text{asy}}^{\text{ICN}} \simeq \frac{1}{\rho_s^2}\left(\frac{\gamma_{\text{th}}\gamma'_{\text{th}}}{\Omega_1\Omega_2\theta\theta'} + \frac{\gamma_{\text{th}}\Xi}{\lambda\Omega_2\Omega_3\theta}\right). \tag{2.25}$$

According to (2.25), it is clear that in region $\alpha_1 \in \left(0, \frac{1}{1+\gamma'_{\text{th}}}\right)$, U_2 achieves its full diversity order of two.

2.4 Numerical Results

Now numerical investigation is carried out to verify the analytical results and present some non-trivial design insights. Unless otherwise specified, the following parameters are used: $\Omega_1 = \Omega_3 = 0.1$, $\Omega_2 = 0.01$, $\rho_s = \rho_r$, and $K = 10$.

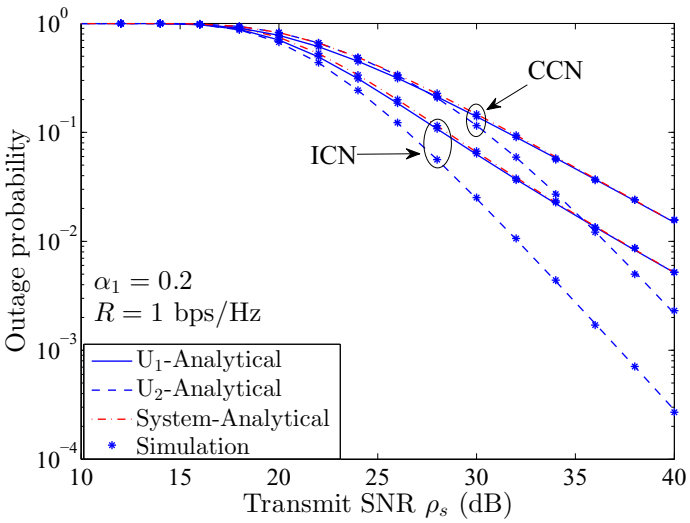


Figure 2.2: Outage performance of the ICN and CCN protocols.

Fig. 2.2 compares outage performance of the proposed ICN protocol against the CCN protocol.[3] A close match between the analytical and simulation results in Fig. 2.2 verifies the accuracy of our analysis. Fig. 2.2 also shows that both the ICN and CCN protocols achieve a full diversity order for each user. Further, we can observe that the proposed ICN protocol is superior to the CCN protocol in terms of each user's outage probability and the system outage probability, which is consistent with our analysis in Section 2.3.2.

[3]Here we compare our ICN protocol with the CCN protocol as only statistical CSI is needed at the source in both protocols, whereas for the optimal cooperative NOMA transmission protocol proposed in [33], the source requires global instantaneous CSI.

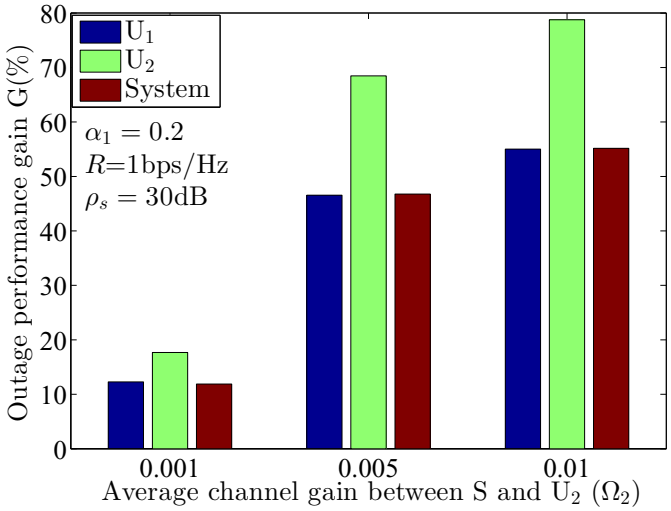


Figure 2.3: Outage performance improvement of ICN over CCN.

Fig. 2.3 shows the outage performance gain of the ICN protocol relative to the CCN protocol, and we define *performance gain* of the ICN protocol relative to the CCN protocol as $\mathrm{G}\,(\%) = 100 \times \left(1 - \frac{P_{\Delta}^{\mathrm{ICN}}}{P_{\Delta}^{\mathrm{CCN}}}\right)$, where $\Delta \in \{1, 2, 1\&2\}$. It is obvious that U_2 has the highest performance gain, while the performance gains of U_1 and the system are almost the same. Note that this observation is also verified by Fig. 2.2. All the performance gains shrink as Ω_2 decreases, because S in the ICN protocol tends to transmit information in the cooperative NOMA transmission mode as the channel from S to U_2 deteriorates.

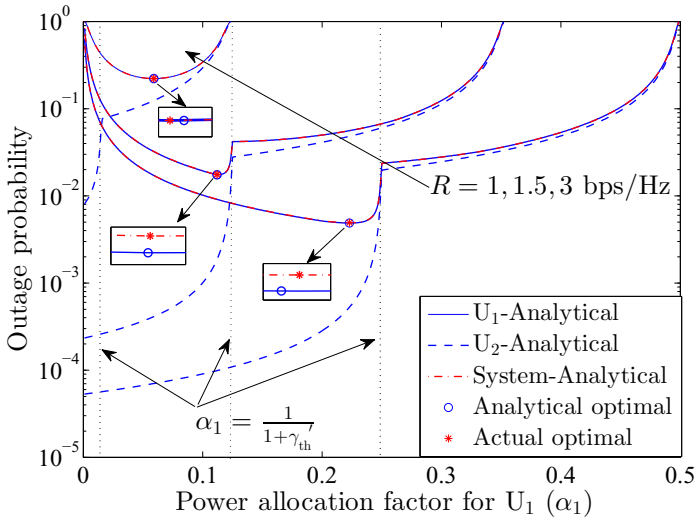


Figure 2.4: Outage performance of the ICN protocol for varying α_1 ($\rho_s = 40$dB).

Fig. 2.4 investigates the impact of power allocation factor α_1 on the outage perfor-

mance of the network. It can be observed that the outage probability of U_2 increases with α_1, while the outage probability of U_1 first decreases and then increases with α_1. The reasons are as follows. With a higher α_1, α_2 is lower, and thus, the chance that U_2 can successfully decode its information in the direct NOMA transmission mode is lower. Further, in the cooperative NOMA transmission mode, a lower α_2 means the chance that U_1 correctly decodes U_2's message is lower, and thus, the chance that U_1 can help U_2 to achieve U_2's target rate is lower. Therefore, the outage probability of U_2 increases with α_1. The outage probability of U_1 is affected by two factors as follows. Factor 1: A higher α_1 means more power for U_1's signal, which tends to decrease its outage probability. Factor 2: As aforementioned, a higher α_1 also means the chance that U_1 correctly decodes U_2's message is lower, or in other words, the chance that U_1 performs SIC is lower, which tends to increase U_1's outage probability. When α_1 is low, Factor 1 dominates, and thus, U_1's outage probability decreases with α_1. When α_1 increases beyond a point, Factor 2 dominates, and thus, U_1's outage probability increases with α_1. From Fig. 2.4, we can see that the analytical approximation of the optimal α_1 (which minimizes $P_{1,\text{asy}}^{\text{ICN}}$) is close to the actual optimal value (which is the point of α_1 that minimizes the system outage probability). It is worth noticing that when $R = 3$bps/Hz, the optimal α_1 lies in the region $\left[\frac{1}{1+\gamma'_{\text{th}}}, \frac{1}{1+\gamma_{\text{th}}}\right)$, which indicates that to minimize system outage probability, the system should stay in the direct NOMA transmission mode in this case. When $R = 1$bps/Hz and $R = 1.5$bps/Hz, the optimal α_1 is smaller than $\frac{1}{1+\gamma'_{\text{th}}}$, and thus, the best system outage performance is achieved by adaptively switching its transmission mode according to the quality of direct link to U_2.

2.5 Summary

In this chapter, we have proposed a novel and practical cooperative protocol for two-user downlink NOMA networks. We have analytically proved that the proposed ICN protocol outperforms the CCN protocol in terms each user's outage probability and the system outage probability. In addition, we have discussed the optimal power allocation strategy to optimize the system outage performance in the high SNR region. Numerical results have validated our analysis and demonstrated valuable insights.

Chapter 3

Channel-Aware Power Allocation and Decoding Order in Overlay Cognitive NOMA Networks

To further improve spectral efficiency, this chapter considers an overlay cognitive NOMA network consisting of a pair of primary users and a pair of secondary users. In this network, communication between the pair of primary users is achieved with the help of the secondary transmitter. In return, the secondary transmitter can communicate with the secondary receiver by sending a superimposed signal consisting of its own message and the received primary message. We intend to minimize the outage probability of the secondary system under different QoS constraints of the primary system by jointly optimizing the power allocation factor at the secondary transmitter and the decoding order at the receivers. When full CSI is available at the secondary transmitter, we derive the optimal power allocation factor and decoding order in closed form, which can minimize the outage probability of the secondary system while guaranteeing the target rate of the primary system. When only statistical CSI is available at the secondary transmitter, we derive the optimal power allocation factor and decoding order in semi-closed form, which minimizes the outage probability of the secondary system while guaranteeing that the outage probability of the primary system is not more than a predetermined threshold. For the proposed strategies, outage probability expressions of the primary and secondary systems are derived in closed forms. Numerical results are shown to verify the analytical results and demonstrate the superiority of the proposed strategies in different CSI-availability scenarios.

3.1 Introduction

3.1.1 Background and Motivation

It can be noticed that in overlay cognitive NOMA networks, the primary receiver should be given higher priority while the secondary receiver should be given lower priority to perform NOMA transmission. Therefore, overlay cognitive NOMA networks can be viewed as multiuser networks with user priority difference. For this kind of multiuser networks, an appropriate research target, referred to as **priority target**, is to get the best performance of the lower priority user while satisfying the QoS requirements of the higher priority user. The priority target is quite different from conventional target for conventional multiuser networks without user priority difference, in which sum rate or fair rate is more concerned.

It is worth mentioning that in NOMA transmissions to multiple users, the achievable rate of the desired information at each user (receiver) is highly affected by the power allocation factor at the transmitter and the decoding order at the receiver since NOMA is implemented by employing superposition coding at the transmitter and SIC at receivers. However, the power allocation and decoding order strategy for conventional NOMA networks does not work in multiuser networks with user priority difference due to the aforementioned priority target.

- The power allocation strategy may not work: In conventional NOMA networks, it is not possible for the source to allocate all the power to one user to guarantee its service, while this may be allowed in NOMA networks with user priority difference to satisfy the higher priority user's QoS requirement.

- The decoding order may not work: Conventional NOMA networks suggest that users should be ordered by their channel conditions and the decoding order at the receiver sides should be from the weakest user's message to the strongest user's message [15]. However, in multiuser networks with user priority difference, when the higher priority user has a better connection from the source (compared with the lower priority user) but its channel condition is not good enough to support the higher priority user in decoding both higher priority and lower priority users' messages, the higher priority user should decode its desired message first in order

to achieve its QoS target.

In NOMA networks with user priority difference, a typical setting is to use a fixed power allocation and fixed decoding order (FP-FD) strategy in which the power allocated to the higher priority user is fixed and always larger than that allocated to the lower priority user, and the decoding order is also fixed and always from the higher priority user's message to the lower priority user's message [38–41]. These FP-FD strategies do not guarantee the QoS of the higher priority user, and do not provide the optimal performance for the lower priority user. There are limited works in the literature considering dynamic power allocation and/or decoding order. Dynamic power allocation and fixed decoding order (DP-FD) strategies are employed in [42–44] where the decoding order is fixed as in [38–41], but the power allocation factor can be dynamically adjusted to best serve the lower priority user while satisfying the QoS requirement of the higher priority user. Although these strategies can strictly satisfy the QoS requirement of the higher priority user, they do not provide the optimal performance for the lower priority user since the degree-of-freedom in decoding order is lost. To further improve the performance of the lower priority user, the authors in [45, 46] propose dynamic power allocation and dynamic decoding order (DP-DD) strategies in which the decoding order can be dynamically adjusted as in conventional NOMA networks without user priority difference [15] (i.e., the decoding order is from the weaker user's message to the stronger user's message). The DP-DD strategies in [45, 46] improve the performance of the lower priority user by loosening the constraint of decoding order, but they do not strictly guarantee the QoS requirement of the higher priority user. As a summary, the above strategies do not satisfy our aforementioned priority target, i.e., to get the best performance of the lower priority user while strictly ensuring the QoS demands of the higher priority user. To fill this gap, we study the joint optimization of decoding order and power allocation for NOMA networks with user priority difference, to satisfy our priority target.

3.1.2 Key Contributions and Organization

We consider an overlay cognitive NOMA network consisting of a pair of primary users and a pair of the secondary users. We have a target information rate for the primary system as well as a target information rate for the secondary system. We study

the joint optimal design of power allocatin (at the secondary transmitter side) and decoding order (at the primary receiver and the secondary receiver sides), to meet our priority target.

In a *full CSI case*, i.e., when the secondary transmitter has the instantaneous CSI of links from the secondary transmitter to the primary receiver and the secondary transmitter to the secondary receiver), we guarantee the primary system's target rate while minimizing the outage probability of the secondary system. On the other hand, due to some reasons such as reducing the signalling overhead [47–49], the secondary transmitter may only have the statistical CSI of links from the secondary transmitter to the primary receiver and the secondary transmitter to the secondary receiver, referred to as the *partial CSI case*, and thus, it is impossible to guarantee the primary system's target rate. Accordingly, we guarantee a target outage performance of the primary system, while minimizing the outage probability of the secondary system. The key contributions of this work can be summarized as follows:

- Novel system setting: Different from existing works, we guarantee the QoS requirement of the primary system and get the best performance of the secondary system.

- Closed-form (or semi-closed-form) optimal solution: For the full CSI case, we derive the optimal power allocation factor and decoding order in closed form. For the partial CSI case, we derive the optimal power allocation factor and decoding order in semi-closed form.

- Closed-form performance analysis: We derive closed-form outage probability expressions of the primary and secondary systems for the proposed strategies. Numerical results are provided to verify the analytical results and show the superiority of the proposed strategies over the existing strategies.

The rest of this chapter is organized as follows. Section 3.2 illustrates the system model and the transmission protocol. Sections 3.3 and 3.4 introduce the proposed optimal power allocation and decoding order strategies for the full CSI and partial CSI cases, respectively, and analyze their outage performance. Numerical results are provided in Section 3.5 to verify the analytical results.

Table 3.1: Description of symbols used for key parameters in Chapter 3.

Symbol	Meaning
h_1	Channel coefficient from the primary transmitter to the secondary transmitter
h_2	Channel coefficient from the secondary transmitter to the primary receiver
h_3	Channel coefficient from the secondary transmitter to the secondary receiver
P_{t}	Transmit power at the primary and secondary transmitters
x_{p}	Message for the primary receiver
x_{s}	Message for the secondary receiver
α	Power allocation factor
$\hat{\alpha}^{\mathbf{i}}, i = 1, 2$	Optimal power allocation factor under the i-th decoding order in the full CSIT case
$\check{\alpha}^{\mathbf{i}}, i = 1, 2$	Optimal power allocation factor under the i-th decoding order in the partial CSIT case
$R^{\mathbf{1}}_{2,x_{\mathrm{p}}}(\alpha)$	Achievable rate of the primary message at the primary receiver under the first decoding order
$R^{\mathbf{1}}_{3,x_{\mathrm{p}}}(\alpha)$	Achievable rate of the primary message at the secondary receiver under the first decoding order
$R^{\mathbf{1}}_{3,x_{\mathrm{s}}}(\alpha)$	Achievable rate of the secondary message at the secondary receiver under the first decoding order
$R^{\mathbf{2}}_{2,x_{\mathrm{s}}}(\alpha)$	Achievable rate of the secondary message at the primary receiver under the second decoding order
$R^{\mathbf{2}}_{3,x_{\mathrm{s}}}(\alpha)$	Achievable rate of the secondary message at the secondary receiver under the second decoding order
$R^{\mathbf{2}}_{2,x_{\mathrm{p}}}(\alpha)$	Achievable rate of the primary message at the secondary receiver under the second decoding order
$\tilde{R}_{\mathrm{p}}$	Target rate of the primary system
$\tilde{R}_{\mathrm{s}}$	Target rate of the secondary system
α^{F}	Power allocation factor in the proposed F-PA-DO strategy
δ^{F}	Decoding order in the proposed F-PA-DO strategy
α^{P}	Power allocation factor in the proposed P-PA-DO strategy
δ^{P}	Decoding order in the proposed P-PA-DO strategy
$R^{\mathrm{F}}_{2,x_{\mathrm{p}}}$	Achievable rate of the primary message at the primary receiver with the proposed F-PA-DO strategy
$R^{\mathrm{F}}_{3,x_{\mathrm{s}}}$	Achievable rate of the secondary message at the secondary receiver with the proposed F-PA-DO strategy
$P^{\mathrm{F}}_{\mathrm{p}}$	Outage probability of the primary system with the proposed F-PA-DO strategy
$P^{\mathrm{F}}_{\mathrm{s}}$	Outage probability of the secondary system with the proposed F-PA-DO strategy
$P^{\mathbf{i}}_{\mathrm{p}}(\alpha), i = 1, 2$	Outage probability of the primary system with the power allocation factor α and the i-th decoding order
$P^{\mathbf{i}}_{\mathrm{s}}(\alpha), i = 1, 2$	Outage probability of the secondary system with the power allocation factor α and the i-th decoding order
$P^{\mathrm{P}}_{\mathrm{p}}$	Outage probability of the primary system with the proposed P-PA-DO strategy
$P^{\mathrm{P}}_{\mathrm{s}}$	Outage probability of the secondary system with the proposed P-PA-DO strategy

3.2 System Model and Transmission Protocol

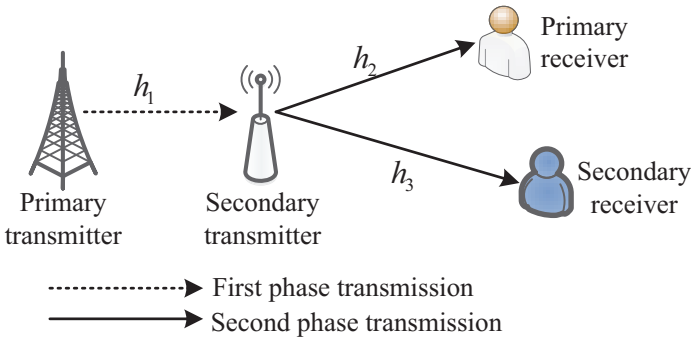


Figure 3.1: System model and NOMA-based transmission protocol.

3.2.1 Network Architecture and Channel Model

As depicted in Fig. 3.1, we consider an overlay cognitive NOMA network consisting of a pair of primary users (the primary transmitter and the primary receiver) and a pair of secondary users (the secondary transmitter and the secondary receiver). Due to a heavy shadowing effect, the direct link between the primary transmitter and the primary receiver is not available [40]. Thus, the primary transmitter has to seek the help from the secondary transmitter to communicate with the primary receiver. In return, the secondary transmitter is allowed to access the licensed spectrum to send its own message to the secondary receiver by using NOMA. All nodes have one antenna and work in the half-duplex mode. Specifically, the secondary transmitter employs the DF protocol to process the signal sent by the primary transmitter. Let h_1, h_2 and h_3 denote the channel coefficients of the primary transmitter to the secondary transmitter, the secondary transmitter to the primary receiver, and the secondary transmitter to the secondary receiver, respectively. We assume that all the channels experience Nakagami-m fading, i.e., $h_i \sim Nakagami\,(m_i, \Omega_i)$[1], $\forall i \in \{1, 2, 3\}$, and all the channel coefficients remain unchanged within one block time. However, they may change independently from one block to the next. Moreover, all the receivers can perfectly access local instantaneous CSI [48, 49] and are corrupted by the AWGN with an identical variance σ^2.

[1] $X \sim Nakagami\,(a, b)$ denotes that a random variable X follows Nakagami-m distribution with fading severity parameter a and average fading power b.

3.2.2 Overlay Cognitive NOMA Transmission Protocol

The whole transmission block is equally divided into two phases. In the first phase, the primary transmitter broadcasts a primary message x_{p} with a transmit power P_{t} and $\mathbb{E}\left[|x_{\mathrm{p}}|^2\right] = 1$. Then, the received signal at the secondary transmitter can be expressed as

$$y_1 = \sqrt{P_{\mathrm{t}}} h_1 x_{\mathrm{p}} + n_1, \tag{3.1}$$

where n_1 is the AWGN at the secondary transmitter. According to (3.1), the achievable rate of x_{p} at the secondary transmitter is given by $R_1 = \frac{1}{2}\log_2\left(1 + \rho_{\mathrm{t}}|h_1|^2\right)$, where the factor "$\frac{1}{2}$" accounts for the fact that the first phase transmission occupies half transmission block, and $\rho_{\mathrm{t}} \triangleq \frac{P_{\mathrm{t}}}{\sigma^2}$ is the transmit SNR.

Letting $\tilde{R}_{\mathrm{p}}$ denote the target rate of the primary system, then the condition for the secondary transmitter to correctly decode x_{p} can be expressed as $R_1 \geq \tilde{R}_{\mathrm{p}}$. If the secondary transmitter correctly decodes x_{p}, it uses transmit power P_{t} to send a superimposed signal consisting of the primary message x_{p} and the secondary message x_{s} in the second phase, and the observed signals at the primary receiver and the secondary receiver are respectively given by

$$y_2 = \sqrt{\alpha P_{\mathrm{t}}} h_2 x_{\mathrm{p}} + \sqrt{(1-\alpha) P_{\mathrm{t}}} h_2 x_{\mathrm{s}} + n_2, \tag{3.2}$$

and

$$y_3 = \sqrt{\alpha P_{\mathrm{t}}} h_3 x_{\mathrm{p}} + \sqrt{(1-\alpha) P_{\mathrm{t}}} h_3 x_{\mathrm{s}} + n_3, \tag{3.3}$$

where $\alpha \in (0, 1]$ denotes the power allocation factor of x_{p} at the secondary transmitter, and n_2 and n_3 are the AWGN at the primary receiver and the secondary receiver, respectively. Depending on the decoding order at the primary and secondary receivers, we have different achievable rates of x_{p} and x_{s}.

1) First decoding order, $x_{\mathrm{p}} \rightarrow x_{\mathrm{s}}$: In the first decoding order, the primary receiver and the secondary receiver decode x_{p} first while treating x_{s} as noise. The achievable rate of x_{p} at the primary receiver and the secondary receiver are respectively given by

$$R_{2,x_{\mathrm{p}}}^{\mathbf{1}}(\alpha) = \frac{1}{2}\log_2\left(1 + \frac{\alpha\rho_{\mathrm{t}}|h_2|^2}{(1-\alpha)\rho_{\mathrm{t}}|h_2|^2 + 1}\right), \tag{3.4}$$

and

$$R_{3,x_{\mathrm{p}}}^{\mathbf{1}}(\alpha)=\frac{1}{2}\log_2\left(1+\frac{\alpha\rho_{\mathrm{t}}|h_3|^2}{(1-\alpha)\,\rho_{\mathrm{t}}|h_3|^2+1}\right), \tag{3.5}$$

where the superscript "**1**" is the indicator for the first decoding order.

Provided that $R_{3,x_{\mathrm{p}}}^{\mathbf{1}}(\alpha)\geq\tilde{R}_{\mathrm{p}}$, i.e., the secondary receiver correctly decodes x_{p} and successfully performs SIC, the achievable rate of x_{s} at the secondary receiver is given by

$$R_{3,x_{\mathrm{s}}}^{\mathbf{1}}(\alpha)=\frac{1}{2}\log_2\left(1+(1-\alpha)\,\rho_{\mathrm{t}}|h_3|^2\right). \tag{3.6}$$

2) Second decoding order, $x_{\mathrm{s}}\rightarrow x_{\mathrm{p}}$: In the second decoding order, the primary receiver and the secondary receiver decode x_{s} first while treating x_{p} as noise. The achievable rate of x_{s} at the primary receiver and the secondary receiver for the second decoding order are respectively given by

$$R_{2,x_{\mathrm{s}}}^{\mathbf{2}}(\alpha)=\frac{1}{2}\log_2\left(1+\frac{(1-\alpha)\,\rho_{\mathrm{t}}|h_2|^2}{\alpha\rho_{\mathrm{t}}|h_2|^2+1}\right), \tag{3.7}$$

and

$$R_{3,x_{\mathrm{s}}}^{\mathbf{2}}(\alpha)=\frac{1}{2}\log_2\left(1+\frac{(1-\alpha)\,\rho_{\mathrm{t}}|h_3|^2}{\alpha\rho_{\mathrm{t}}|h_3|^2+1}\right), \tag{3.8}$$

where the superscript "**2**" is the indicator for the second decoding order.

Let $\tilde{R}_{\mathrm{s}}$ denote the target rate of the secondary system. Conditioned on $R_{2,x_{\mathrm{s}}}^{\mathbf{2}}(\alpha)\geq\tilde{R}_{\mathrm{s}}$, i.e., the primary receiver correctly decodes x_{s} and successfully performs SIC, the achievable rate of x_{p} at the primary receiver is given by

$$R_{2,x_{\mathrm{p}}}^{\mathbf{2}}(\alpha)=\frac{1}{2}\log_2\left(1+\alpha\rho_{\mathrm{t}}|h_2|^2\right). \tag{3.9}$$

3.3 Joint Optimal Design of Power Allocation and Decoding Order in Full CSI Case

For the full CSI case, the secondary transmitter knows the instantaneous CSI of the links from the secondary transmitter to the primary receiver and the secondary transmitter to the secondary receiver. In this case, we intend to guarantee the target rate of the primary system[2] and minimize the outage probability of the secondary

[2]Ensuring the target rate of the primary system can be also interpreted as the primary receiver correctly decodes x_{p}.

system, by jointly optimizing the power allocation factor and the decoding order. Note that an outage event is defined as an event when the achievable rate of the desired message is less than the target rate [39, 40, 44, 46]. Therefore, in the full CSI case, minimizing the outage probability of the secondary system is equivalent to maximizing the achievable rate of $x_{\rm s}$ in the secondary system. In this section, we first present a novel full-CSI power allocation and decoding order strategy, denoted as F-PA-DO strategy. In the strategy, we derive the joint optimal power allocation factor and decoding order, which maximizes the achievable rate of $x_{\rm s}$ in the secondary system while satisfying the target rate requirement of the primary system.[3] Then, the outage performance of the primary and secondary systems for the proposed strategy are analyzed.

3.3.1 Proposed F-PA-DO Strategy

To find the joint optimal power allocation factor and decoding factor, we first derive the optimal power allocation factors for two different decoding orders, and then pick up the optimal decoding order with larger achievable rate of $x_{\rm s}$ in the secondary system, as well as the corresponding power allocation factor.

The optimal power allocation factor for the first decoding order has been studied in [42–44]. According to [44, Eq. (20)], the optimal power allocation factor of $x_{\rm p}$ at the secondary transmitter for the first decoding order is given by[4]

$$\hat{\alpha}^{\mathbf{1}} = 1 - \left[\frac{\rho_{\rm t}\tilde{h} - \gamma_{\rm th}^{\rm p}}{\rho_{\rm t}\tilde{h}\left(1+\gamma_{\rm th}^{\rm p}\right)}\right]^{+}, \tag{3.10}$$

where $\tilde{h} \triangleq \min\left\{|h_2|^2, |h_3|^2\right\}$ and $\gamma_{\rm th}^{\rm p} \triangleq 2^{2\tilde{R}_{\rm p}} - 1$.

We can find from (3.10) that $\hat{\alpha}^{\mathbf{1}} = 1$ when $\tilde{h} \leq \frac{\gamma_{\rm th}^{\rm p}}{\rho_{\rm t}}$, which indicates that the secondary transmitter allocates all its available power to serve the primary system when the target rate of the primary system cannot be satisfied or the secondary receiver fails to perform SIC in the first decoding order.

[3] At a transmission block, if it is impossible to satisfy the target rate requirement of the primary system, then the the secondary transmitter assigns all its transmit power for the primary message $x_{\rm p}$ in the transmission block.

[4] It has been shown in [42–44] that for the first decoding order, the power allocation factor given in (3.10) can maximize the achievable rate of the secondary system under the QoS constraint of the primary system. Readers may refer to [43, 44] for detailed process.

Next, we try to find the optimal power allocation factor for the second decoding order, which has not been studied in the literature.

Lemma 1. *In the second decoding order, the target rate of the primary system can be satisfied only when* $|h_2|^2 \geq \frac{\gamma_{th}}{\rho_t}$, *where* $\gamma_{th} \triangleq \gamma_{th}^p + \gamma_{th}^s + \gamma_{th}^p \gamma_{th}^s$ *and* $\gamma_{th}^s \triangleq 2^{2\tilde{R}_s} - 1$. *When this constraint holds, the optimal power allocation factor of* x_p *at the secondary transmitter for the second decoding order is given by*

$$\hat{\alpha}^{\mathbf{2}} = \frac{\gamma_{th}^p}{\rho_t |h_2|^2}. \tag{3.11}$$

Proof. First, to achieve the target rate of the primary system in the second decoding order, the following two conditions should be satisfied

$$R_{2,x_\mathrm{s}}^{\mathbf{2}}(\alpha) \geq \tilde{R}_\mathrm{s}, \quad R_{2,x_\mathrm{p}}^{\mathbf{2}}(\alpha) \geq \tilde{R}_\mathrm{p}, \tag{3.12}$$

where the first constraint ensures that the primary receiver successfully performs SIC by decoding x_s, and the second constraint ensures that the primary receiver correctly decodes x_p after the interference of x_s is removed. Now substituting (3.7) and (3.9) into (3.12), we can derive a constraint regarding the power allocation factor α, which is given by

$$\frac{\gamma_\mathrm{th}^\mathrm{p}}{\rho_\mathrm{t}|h_2|^2} \leq \alpha \leq \frac{\rho_\mathrm{t}|h_2|^2 - \gamma_\mathrm{th}^\mathrm{s}}{\rho_\mathrm{t}|h_2|^2(1+\gamma_\mathrm{th}^\mathrm{s})}. \tag{3.13}$$

A valid α can be found from (3.13) only when $\frac{\rho_\mathrm{t}|h_2|^2 - \gamma_\mathrm{th}^\mathrm{s}}{\rho_\mathrm{t}|h_2|^2(1+\gamma_\mathrm{th}^\mathrm{s})} \geq \frac{\gamma_\mathrm{th}^\mathrm{p}}{\rho_\mathrm{t}|h_2|^2}$, which can be simplified as $|h_2|^2 \geq \frac{\gamma_\mathrm{th}}{\rho_\mathrm{t}}$. Here completes the proof of the first part in *Lemma* 1.

In addition, we can tell from (3.8) that the achievable rate of x_s at the secondary receiver in the second decoding order $R_{3,x_\mathrm{s}}^{\mathbf{2}}(\alpha)$ is a strict decreasing function of α. Combining this observation with the constraint given in (3.13), we know that $R_{3,x_\mathrm{s}}^{\mathbf{2}}(\alpha)$ achieves its maximum at $\alpha = \frac{\gamma_\mathrm{th}^\mathrm{p}}{\rho_\mathrm{t}|h_2|^2}$, which completes the proof of the second part in *Lemma* 1. □

After presenting *Lemma* 1, we are ready to introduce the proposed F-PA-DO strategy which outlines the optimal design of power allocation and decoding order design to maximize the achievable rate of x_s in the secondary system under the target rate constraint of the primary system.

Proposition 1. *When the secondary transmitter has full CSI, the joint optimal power allocation factor and decoding order, denoted as* $\left(\alpha^F, \delta^F\right)$*, is given by*

$$\left(\alpha^F, \delta^F\right) = \begin{cases} (\hat{\alpha}^{\mathbf{2}}, \mathbf{2}), \textit{if } |h_2|^2 \geq \max\left\{\frac{\gamma_{th}}{\rho_t}, |h_3|^2\right\}, \\ (\hat{\alpha}^{\mathbf{1}}, \mathbf{1}), \textit{else}, \end{cases} \tag{3.14}$$

where the superscript "F" refers to the proposed F-PA-DO strategy, and $\hat{\alpha}^{\mathbf{1}}$ *and* $\hat{\alpha}^{\mathbf{2}}$ *have been given in* (3.10) *and* (3.11)*, respectively.*

Proof. According to *Lemma* 1, we know that the target rate of the primary system can be satisfied under the second decoding order only when $|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_{\mathrm{t}}}$. If $|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_{\mathrm{t}}}$, the maximal achievable rate of x_{s} in the secondary system under the second decoding order is $R^{\mathbf{2}}_{3,x_{\mathrm{s}}}(\hat{\alpha}^{\mathbf{2}})$. By comparing $R^{\mathbf{2}}_{3,x_{\mathrm{s}}}(\hat{\alpha}^{\mathbf{2}})$ with $R^{\mathbf{1}}_{3,x_{\mathrm{s}}}(\hat{\alpha}^{\mathbf{1}})$ which is the maximal achievable rate of x_{s} in the secondary system under the first decoding order, we can find that $R^{\mathbf{1}}_{3,x_{\mathrm{s}}}(\hat{\alpha}^{\mathbf{1}}) \leq R^{\mathbf{2}}_{3,x_{\mathrm{s}}}(\hat{\alpha}^{\mathbf{2}})$ when $|h_2|^2 \geq |h_3|^2$. Therefore, the second decoding order is preferred only when $|h_2|^2 \geq \max\left\{\frac{\gamma_{\mathrm{th}}}{\rho_{\mathrm{t}}}, |h_3|^2\right\}$. □

Remark 1. *Proposition 1 illustrates that in NOMA networks with user priority difference, to get the best performance of the lower priority user while achieving the QoS requirement of the higher priority user, the decoding order should be decided by the relative channel conditions of the higher priority and lower priority users and the absolute channel condition of the higher priority user.*

With the proposed F-PA-DO strategy, the achievable rate of x_{p} and x_{s} at the primary and secondary receivers are respectively given by

$$R^{\mathrm{F}}_{2,x_{\mathrm{p}}} = \begin{cases} \tilde{R}_{\mathrm{p}}, \text{if } |h_2|^2 \geq \frac{\gamma^{\mathrm{p}}_{\mathrm{th}}}{\rho_{\mathrm{t}}}, \\ \frac{1}{2}\log_2\left(1 + \rho_{\mathrm{t}}|h_2|^2\right), \text{else}, \end{cases} \tag{3.15}$$

and

$$R^{\mathrm{F}}_{3,x_{\mathrm{s}}} = \begin{cases} \frac{1}{2}\log_2(1+\Lambda_1), \text{if } |h_2|^2 \geq \max\left\{\frac{\gamma_{\mathrm{th}}}{\rho_{\mathrm{t}}}, |h_3|^2\right\}, \\ \frac{1}{2}\log_2(1+\Lambda_2), \text{else}, \end{cases} \tag{3.16}$$

where $\Lambda_1 \triangleq \frac{|h_3|^2\left(\rho_{\mathrm{t}}|h_2|^2 - \gamma^{\mathrm{p}}_{\mathrm{th}}\right)}{|h_2|^2 + \gamma^{\mathrm{p}}_{\mathrm{th}}|h_3|^2}$ and $\Lambda_2 \triangleq \left[\frac{|h_3|^2\left(\rho_{\mathrm{t}}\tilde{h} - \gamma^{\mathrm{p}}_{\mathrm{th}}\right)}{\tilde{h}\left(1+\gamma^{\mathrm{p}}_{\mathrm{th}}\right)}\right]^+$.

To this end, we have finished the joint optimal design of power allocation and decoding order for the full CSI case. Next, we will analyze the system performance with the proposed F-PA-DO strategy.

3.3.2 Outage Performance Analysis

3.3.2.1 Primary system

According to the transmission protocol, an outage event for the primary system happens when either the secondary transmitter or the primary receiver fails to decode x_{p}, i.e., the achievable rate of x_{p} at the secondary transmitter or the primary receiver is less than $\tilde{R}_{\mathrm{p}}$.

Corollary 1. *The outage probability of the primary system for the proposed F-PA-DO strategy is given by*

$$P_p^F = 1 - \Psi\left(m_1, \frac{\gamma_{th}^p}{\rho_t \Omega_1}\right) \Psi\left(m_2, \frac{\gamma_{th}^p}{\rho_t \Omega_2}\right), \tag{3.17}$$

where $\Psi(x, y) \triangleq \frac{\Gamma(x,y)}{\Gamma(x)}$ *denotes the regularized gamma function.*

Proof. Please see section 3.7.1. □

3.3.2.2 Secondary system

An outage event of the secondary system happens when the secondary transmitter fails to decode x_{p} or the secondary receiver fails to decode x_{s}, i.e., the achievable rate of x_{p} at the secondary transmitter is less than $\tilde{R}_{\mathrm{p}}$ or the achievable rate of x_{s} at the secondary receiver is less than $\tilde{R}_{\mathrm{s}}$.

Corollary 2. *The outage probability of the secondary system for the proposed strategy can be approximated as*

$$\begin{aligned} P_s^F \approx 1 &- \Psi\left(m_1, \frac{\gamma_{th}^p}{\rho_t \Omega_1}\right) \\ &\times \left(\Upsilon\left(m_2, \Omega_2, m_3, \Omega_3, \gamma_{th}^s\right) + \Upsilon\left(m_3, \Omega_3, m_2, \Omega_2, \gamma_{th}^p\right) - \prod_{l_3=2}^{3} \Psi\left(m_{l_3}, \frac{\gamma_{th}}{\rho_t \Omega_{l_3}}\right)\right), \end{aligned} \tag{3.18}$$

where function $\Upsilon(\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e})$ *is defined as*

$$\begin{aligned} \Upsilon(\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e}) \triangleq\ & e^{-\frac{(\gamma_{th}-\mathrm{e})\mathrm{a}}{\rho_t \mathrm{b}} - \frac{\mathrm{ec}}{\rho_t \mathrm{d}}} \left(\frac{\mathrm{a}}{\rho_t \mathrm{b}}\right)^{\mathrm{b}} \frac{1}{\Gamma(\mathrm{a})} \sum_{l_1=0}^{\mathrm{c}-1} \sum_{l_2=0}^{\mathrm{a}-1+l_1} \frac{1}{l_1!} \left(\frac{\mathrm{ec}}{\rho_t \mathrm{d}}\right)^{l_1} \binom{\mathrm{a}-1+l_1}{l_2} \\ & \times (\gamma_{th} - \mathrm{e})^{\mathrm{a}-1+l_1-l_2} \left[\Phi\left(\frac{\mathrm{ce}(\gamma_{th}-\mathrm{e})}{\rho_t \mathrm{d}}, \frac{\mathrm{a}}{\rho_t \mathrm{b}}, l_2 - l_1\right)\right. \\ & \left. - \sum_{n=1}^{N} \frac{\mathrm{e}\pi}{2N} \sqrt{1-(\theta_n)^2} \left(\mathrm{e}\frac{\theta_n+1}{2}\right)^{l_2-l_1} e^{-\frac{(\theta_n+1)\mathrm{ae}}{2\rho_t \mathrm{b}} - \frac{2(\gamma_{th}-\mathrm{e})\mathrm{c}}{\rho_t \mathrm{d}(\theta_n+1)}}\right]. \end{aligned} \tag{3.19}$$

In (3.19), $\Phi\left(\mathrm{a},\mathrm{b},\mathrm{c}\right) \triangleq 2\left(\frac{\mathrm{a}}{\mathrm{b}}\right)^{\frac{\mathrm{c}+1}{2}} K_{\mathrm{c}+1}\left(\sqrt{4\mathrm{ab}}\right)$, $\theta_n \triangleq \cos\left(\frac{(2n-1)\pi}{2N}\right)$, *and* N *is a complexity-accuracy tradeoff parameter.*

Proof. See section 3.7.2. □

Remark 2. *By checking the first-order partial derivative of* P_s^F *with respect to (w.r.t.)* Ω_j $(j = 1, 2)$ *and* $\tilde{R}_p$*, we can find that* P_s^F *is a decreasing function and an increasing function of* Ω_j *and* $\tilde{R}_p$*, respectively. This can be explained by the fact that larger* Ω_j *or smaller* $\tilde{R}_p$ *can make the target rate of the primary system easier to be satisfied. As a result, the secondary transmitter can allocate more power to the secondary message, which consequently improves the outage performance of the secondary system.*

3.4 Joint Optimal Design of Power Allocation and Decoding Order with Partial CSI

For the partial CSI case, the secondary transmitter knows only the statistical CSI of links from the secondary transmitter to the primary receiver and the secondary transmitter to the secondary receiver. We intend to minimize the outage probability of the secondary system while guaranteeing that the outage probability of the primary system is not more than a predefined threshold [50] denoted as ς.[5] To achieve this target, we first analyze the outage performance of the primary and secondary systems with the given power allocation factor and decoding order. Then, based on the derived closed-form expressions of the outage probabilities, we present a novel partial-CSI power allocation and decoding order strategy, denoted as P-PA-DO strategy, which derives the optimal power allocation factor and decoding order.

3.4.1 Outage Performance with the Given Power Allocation Factor and Decoding Order

Lemma 2. *Given the power allocation factor* α *and the first decoding order, the outage probability of the primary and secondary systems are respectively given by*

$$P_p^{\mathbf{1}}(\alpha)=\begin{cases}1-\Psi\left(m_2,\frac{\xi}{\rho_t\Omega_2}\right)\Theta, \textit{ if } \frac{\gamma_{th}^p}{1+\gamma_{th}^p}<\alpha\leq 1,\\ 1, \textit{ else},\end{cases} \tag{3.20}$$

[5] If it is impossible to guarantee that the outage probability of the primary system is not more than ς, then the secondary transmitter assigns all its transmit power for the primary message.

and

$$P_s^{\mathbf{1}}(\alpha)=\begin{cases}1-\Psi\left(m_3,\frac{\xi'}{\rho_t\Omega_3}\right)\Theta, \textit{ if } \frac{\gamma_{th}^p}{1+\gamma_{th}^p}<\alpha<1,\\ 1, \textit{ else},\end{cases} \tag{3.21}$$

where $\xi \triangleq \frac{\gamma_{th}^p}{\alpha\left(1+\gamma_{th}^p\right)-\gamma_{th}^p}$, $\xi' \triangleq \max\left\{\xi, \frac{\gamma_{th}^s}{1-\alpha}\right\}$, *and* $\Theta \triangleq \Psi\left(m_1, \frac{\gamma_{th}^p}{\rho_t\Omega_1}\right)$.

Under the second decoding order with the power allocation factor α*, the outage probability of the primary and secondary systems are respectively given by*

$$P_p^{\mathbf{2}}(\alpha)=\begin{cases}1-\Psi\left(m_2,\frac{\zeta'}{\rho_t\Omega_2}\right)\Theta, \textit{ if } 0<\alpha<\frac{1}{1+\gamma_{th}^s},\\ 1, \textit{ else},\end{cases} \tag{3.22}$$

and

$$P_s^{\mathbf{2}}(\alpha)=\begin{cases}1-\Psi\left(m_3,\frac{\zeta}{\rho_t\Omega_3}\right)\Theta, \textit{ if } 0<\alpha<\frac{1}{1+\gamma_{th}^s},\\ 1, \textit{ else},\end{cases} \tag{3.23}$$

where $\zeta \triangleq \frac{\gamma_{th}^s}{1-\alpha\left(1+\gamma_{th}^s\right)}$ *and* $\zeta' \triangleq \max\left\{\zeta, \frac{\gamma_{th}^p}{\alpha}\right\}$.

Proof. See section 3.7.3. □

3.4.2 Proposed P-PA-DO Strategy

Based on the analytical results derived in the previous subsection, we are ready to introduce the novel P-PA-DO strategy. To obtain the joint optimal power allocation factor and decoding order, we first try to find the optimal power allocation factors for the two different decoding orders. Then, the joint optimal solution of power allocation factor and decoding order can be decided by comparing the outage probability of the secondary system under different optimal power allocation factors for the two different decoding orders.

Lemma 3. *In the first decoding order, the outage constraint of the primary system can be satisfied when* $P_p^{\mathbf{1}}(1) \leq \varsigma$*, and the optimal power allocation factor is* $\check{\alpha}^{\mathbf{1}} = \max\left\{\epsilon, 1-\frac{\gamma_{th}^s}{\gamma_{th}}\right\}$*, where* ς *denotes the outage probability constraint of the primary system, and* ϵ *denotes the unique solution of* $P_p^{\mathbf{1}}(\alpha) = \varsigma$ *that lies in the interval* $\left(\frac{\gamma_{th}^p}{\gamma_{th}^p+1}, 1\right]$*. In the second decoding order, the outage constraint of the primary system can be satisfied when* $P_p^{\mathbf{2}}\left(\frac{\gamma_{th}^p}{\gamma_{th}}\right) \leq \varsigma$*, and the optimal power allocation factor is* $\check{\alpha}^{\mathbf{2}} = \vartheta$*, where* ϑ *is the unique solution of* $P_p^{\mathbf{2}}(\alpha) = \varsigma$ *that lies in the interval of* $\left(0, \frac{\gamma_{th}^p}{\gamma_{th}}\right]$.

Proof. In the first decoding order, it is easy to verify that the outage probability of the primary system $P_{\mathrm{p}}^{\mathbf{1}}(\alpha)$ (given in (3.20)) is a decreasing function of α due to the following two facts: $\Gamma(x,y)$ is a decreasing function of y according to its definition, and ξ is a decreasing function of α. Therefore, we know that the outage constraint of the primary system can be satisfied as long as $P_{\mathrm{p}}^{\mathbf{1}}(1) \leq \varsigma$ and the optimal power allocation factor for the first decoding order lies in the interval $[\epsilon, 1]$. In addition, we can tell that the outage probability of the secondary system for the first decoding order $P_{\mathrm{s}}^{\mathbf{1}}(\alpha)$ (given in (3.21)) attains its minimum at $\alpha = 1 - \frac{\gamma_{\mathrm{th}}^{\mathrm{s}}}{\gamma_{\mathrm{th}}}$ since $P_{\mathrm{s}}^{\mathbf{1}}(\alpha)$ is an increasing function of ξ', and ξ' decreases and increases as α increases within regions $\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}^{\mathrm{p}}+1}, 1 - \frac{\gamma_{\mathrm{th}}^{\mathrm{s}}}{\gamma_{\mathrm{th}}}\right]$ and $\left(1 - \frac{\gamma_{\mathrm{th}}^{\mathrm{s}}}{\gamma_{\mathrm{th}}}, 1\right)$, respectively. Combining these two observations, the desired optimal power allocatin factor for the first decoding order can be derived.

In the second decoding order, the outage probability of the primary system $P_{\mathrm{p}}^{\mathbf{2}}(\alpha)$ (given in (3.22)) attains its minimum at $\alpha = \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}$ since $P_{\mathrm{p}}^{\mathbf{2}}(\alpha)$ is an increasing function of ζ', and ζ' decreases and increases as α increases within regions $\left(0, \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right]$ and $\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}, \frac{1}{1+\gamma_{\mathrm{th}}^{\mathrm{s}}}\right)$, respectively. Based on this result, we can tell that the outage constraint of the primary system can be satisfied when $P_{\mathrm{p}}^{\mathbf{2}}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right) \leq \varsigma$, and the optimal power allocation factor for the second decoding order lies in the interval $[\vartheta, \varrho]$, where ϑ and ϱ are the two solutions of $P_{\mathrm{p}}^{\mathbf{2}}(\alpha) = \varsigma$ located in intervals of $\left(0, \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right]$ and $\left[\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}, \frac{1}{1+\gamma_{\mathrm{th}}^{\mathrm{s}}}\right)$, respectively. Combing this observation with another fact that the outage probability of the secondary system for the second decoding order $P_{\mathrm{s}}^{\mathbf{2}}(\alpha)$ (given in (3.23)) is an increasing function of α since ζ enlarges as α increases, we can derive the optimal power allocation factor for the second decoding order. □

Based on Lemma 3, we can decide the joint optimal power allocation factor and decoding order, denoted as $\left(\alpha^{\mathrm{P}}, \delta^{\mathrm{P}}\right)$, where the superscript "P" refers to the P-PA-DO strategy.

Proposition 2. *When only partial CSI is available at the secondary transmitter, the joint optimal power allocation factor and decoding order that minimizes the outage probability of the secondary system under the outage constraint of the primary system*

is given by

$$\left(\breve{\alpha}^P, \delta^P\right)=\begin{cases}(\breve{\alpha}^{\mathbf{2}}, \mathbf{2}), \textit{if } P_p^{\mathbf{2}}\left(\frac{\gamma_{th}^p}{\gamma_{th}}\right) \leq \varsigma, \\ (\breve{\alpha}^{\mathbf{1}}, \mathbf{1}), \textit{if } P_p^{\mathbf{1}}(1) \leq \varsigma < P_p^{\mathbf{2}}\left(\frac{\gamma_{th}^p}{\gamma_{th}}\right), \\ (1, \mathbf{1}), \quad \textit{if } P_p^{\mathbf{1}}(1) > \varsigma,\end{cases} \tag{3.24}$$

where $\breve{\alpha}^{\mathbf{1}}$ and $\breve{\alpha}^{\mathbf{2}}$ have been given in Lemma 3.

Proof. When the second decoding order meets the QoS requirement of the primary system, i.e., $P_{\mathrm{p}}^{\mathbf{2}}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right) \leq \varsigma$, the first decoding order also does. In this case, the joint optimal power allocation factor and decoding order can be decided by comparing the minimum outage probability of the secondary system for different decoding orders. It is easy to check that

$$P_{\mathrm{s}}^{\mathbf{2}}\left(\alpha^{\mathbf{2}}\right) \leq P_{\mathrm{s}}^{\mathbf{2}}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right) = P_{\mathrm{s}}^{\mathbf{1}}\left(1-\frac{\gamma_{\mathrm{th}}^{\mathrm{s}}}{\gamma_{\mathrm{th}}}\right) \leq P_{\mathrm{s}}^{\mathbf{1}}\left(\alpha^{\mathbf{1}}\right). \tag{3.25}$$

Therefore, when $P_{\mathrm{p}}^{\mathbf{2}}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right) \leq \varsigma$, it is optimal to use the second decoding order with the power allocation factor as $\alpha^{\mathbf{2}}$.

When $P_{\mathrm{p}}^{\mathbf{1}}(1) \leq \varsigma < P_{\mathrm{p}}^{\mathbf{2}}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\gamma_{\mathrm{th}}}\right)$, the first decoding order can guarantee the outage constraint of the primary system while the second decoding order cannot. Thus, it is optimal to use the first decoding order with the power allocation factor as $\alpha^{\mathbf{1}}$.

When $P_{\mathrm{p}}^{\mathbf{1}}(1) > \varsigma$, i.e., neither the first decoding order nor the second decoding order can guarantee the outage constraint of the primary system, it is intuitive that the secondary transmitter should allocate all its available power to serve the primary system. □

Remark 3. *Proposition 2 indicates that when partial CSI is available at the secondary transmitter, to minimize the outage probability of the secondary system and guarantee the QoS requirement of the primary system, it is preferable to first decode the secondary message when $P_p^{\mathbf{2}}\left(\frac{\gamma_{th}^p}{\gamma_{th}}\right) \leq \varsigma$.*

Remark 4. *The optimal power allocation factor α^P given in Proposition 2 is a semi-closed form solution since it involves ϵ and ϑ which are the solutions of $P_p^{\mathbf{1}}(\alpha) = \varsigma$ and $P_p^{\mathbf{2}}(\alpha) = \varsigma$, respectively. Generally, closed-form expressions of ϵ and ϑ are very difficult to obtain, but they can be found in the following two special cases.*

- $m_2 = 1$: *When the link between the secondary transmitter and the primary receiver experiences Rayleigh fading, we have*

$$\epsilon = \frac{\gamma_{th}^p}{1+\gamma_{th}^p}\left(1+\frac{1}{\rho_t\Omega_2 \ln\frac{\Theta}{1-\varsigma}}\right), \vartheta = \frac{\gamma_{th}^p}{\rho_t\Omega_2 \ln\frac{\Theta}{1-\varsigma}}. \tag{3.26}$$

- $m_2 = 2$: *We have*

$$\epsilon = \frac{\gamma_{th}^p}{1+\gamma_{th}^p}\left(1+\frac{2}{\rho_t\Omega_2\psi}\right), \vartheta = \frac{2\gamma_{th}^p}{\rho_t\Omega_2\psi}, \tag{3.27}$$

where $\psi = -W\left(\frac{\varsigma-1}{e\Theta}\right) - 1$ *and* $W(\cdot)$ *is the Lambert W function.*

When the link between the secondary transmitter and the primary receiver experiences other general Nakagami-m fading, i.e., $m_2 \notin \{1, 2\}$*, obtaining closed-form expressions of* ϵ *and* ϑ *is very challenging. However, since* $P_p^{\mathbf{1}}(\alpha)$ *and* $P_p^{\mathbf{2}}(\alpha)$ *are increasing and decreasing functions of* α *for* $\alpha \in \left(\frac{\gamma_{th}^p}{\gamma_{th}^p+1}, 1\right]$ *and* $\alpha \in \left(0, \frac{\gamma_{th}^p}{\gamma_{th}}\right]$*, respectively, numerical results of* ϵ *and* ϑ *can be easily found using the bisection method.*

To this end, we have finished the joint optimal design of power allocation and decoding order when only partial CSI is available at the secondary transmitter. The outage probability of the primary and secondary systems for the proposed P-PA-DO strategy, denoted as $P_{\mathrm{p}}^{\mathrm{P}}$ and $P_{\mathrm{s}}^{\mathrm{P}}$, are respectively given by $P_{\mathrm{p}}^{\mathrm{P}} = P_{\mathrm{p}}^{\delta^{\mathrm{P}}}(\alpha^{\mathrm{P}})$ and $P_{\mathrm{s}}^{\mathrm{P}} = P_{\mathrm{s}}^{\delta^{\mathrm{P}}}(\alpha^{\mathrm{P}})$. Regarding $P_{\mathrm{s}}^{\mathrm{P}}$, we have the following observation.

Remark 5. *Similar to the full CSI case, the outage performance of the secondary system in the case of partial CSI is also affected by the parameters of the primary system. It can be easily shown that* P_s^P *is a non-increasing function of* Ω_1*,* Ω_2 *and* ς*.*

3.5 Numerical Results

This section provides numerical results to verify the analytical results presented in the previous sections, and highlights some nontrivial insights on the proposed strategies. Specifically, subsections 3.5.1 and 3.5.2 respectively evaluate the proposed F-PA-DO and P-PA-DO strategies. We assume that the average fading power of each channel is characterized by the path loss model, i.e., $\Omega_i = (d_i)^{-\varpi}$, $i = 1, 2, 3$, where ϖ is the path loss exponent, d_1, d_2 and d_3 are the distance from the primary transmitter to the secondary transmitter, the secondary transmitter to the primary receiver, and

the secondary transmitter to the secondary receiver, respectively. Unless otherwise stated, the system parameters are set as follows: Nakagami-m fading parameters $m_1 = m_2 = m_3 = 2$, path loss exponent $\varpi = 3$, target data rate of the primary and secondary systems $\tilde{R}_\mathrm{p} = \tilde{R}_\mathrm{s} = 1$bps/Hz, outage constraint of the primary system $\varsigma = 0.01$, and Gaussian-Chebyshev approximation term $N = 20$.

3.5.1 Performance Evaluation of F-PA-DO

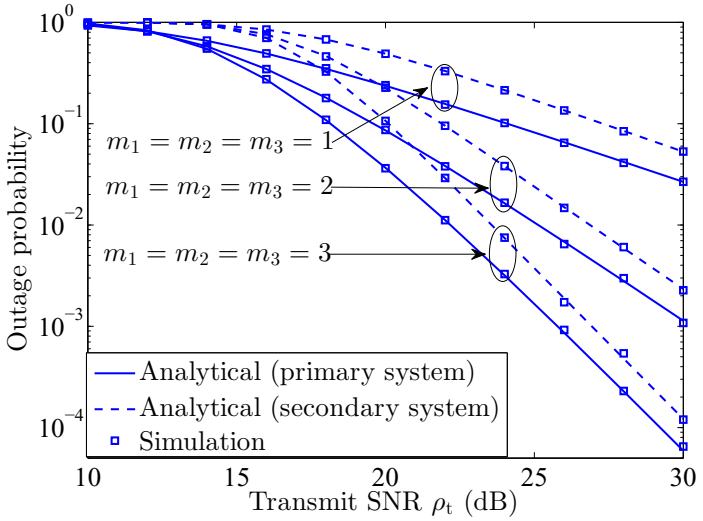


Figure 3.2: Outage performance of the proposed F-PA-DO for different values of fading parameters ($d_1 = d_2 = d_3 = 2$).

Fig. 3.2 verifies the analytically derived outage probability expressions (3.17) and (3.18) of the primary and secondary systems, respectively, for the proposed F-PA-DO strategy. It can be readily observed that our analytical results well coincide with the simulation results, which corroborates the accuracy of the theoretical analysis.

Since our analytical results have been verified, we will only present analytical results in the remaining part of this subsection.

Fig. 3.3 studies the impact of d_1 (distance between the primary transmitter and the secondary transmitter), d_2 (distance between the secondary transmitter and the primary receiver), and $\tilde{R}_\mathrm{p}$ (target rate of the primary system), on the outage performance of the secondary system. It is obvious that as d_1 or d_2 enlarges, i.e., Ω_1 or Ω_2 decreases, the outage performance of the secondary system deteriorates. Moreover, increasing the target rate of the primary system $\tilde{R}_\mathrm{p}$ significantly undermines the outage performance of the secondary system. These observations verify the conclusions

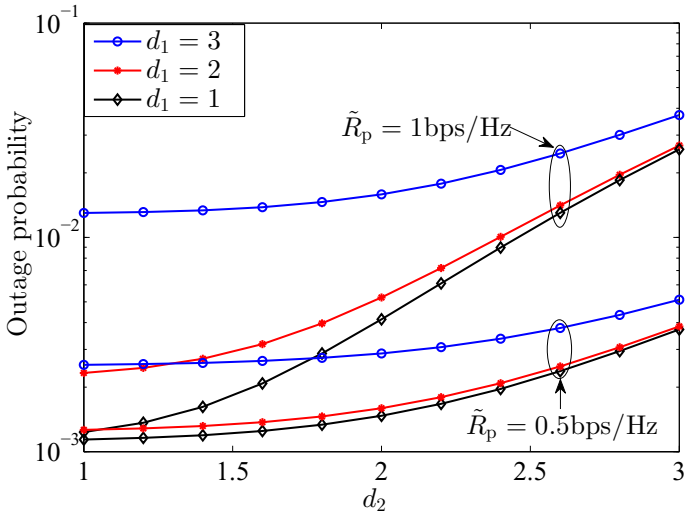
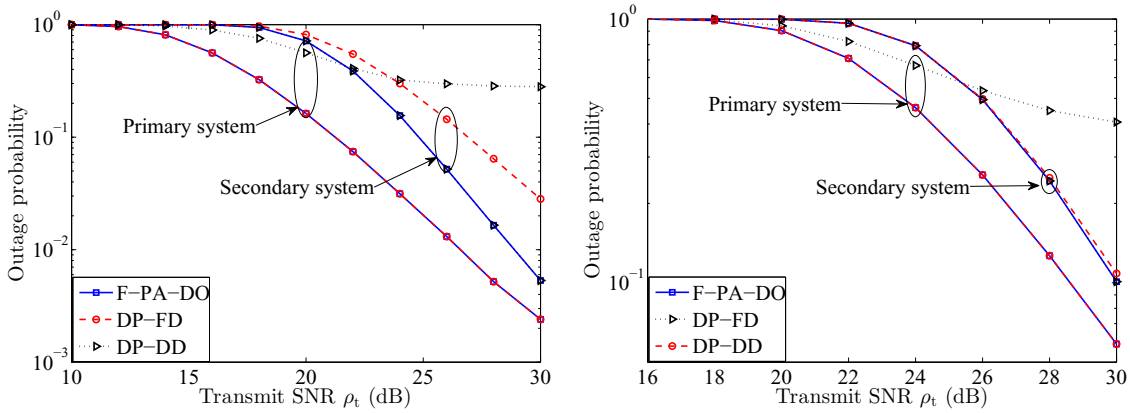


Figure 3.3: Impact of d_1, d_2, and $\tilde{R}_\mathrm{p}$ on the outage probability P_s^F of the secondary system ($\rho_\mathrm{t} = 30$dB, $d_3 = 2$).

drawn in *Remark* 2.



(a) $d_1 = d_2 = d_3 = 2$. (b) $d_1 = d_3 = 2$, $d_2 = 4$.

Figure 3.4: Outage probability of the primary and secondary systems in the F-PA-DO and benchmark strategies.

In Figs. 3.4(a) and 3.4(b), we make a comparison between our proposed F-PA-DO strategy and two benchmark strategies, i.e., DP-FD in [42–44] and DP-DD in [45, 46]. It can be seen from these two figures that the F-PA-DO strategy and the DP-FD strategy have the same outage performance for the primary system since they all promise to do their best effort to satisfy the QoS requirement of the primary system. Compared to the F-FA-DO and the DP-FD strategies, the DP-DD strategy has worse primary outage performance. For the secondary system, the F-PA-DO strategy and the DP-DD strategy have the same outage performance and are superior

to the DP-FD strategy. However, the performance gap between them is relatively small when $d_2 > d_3$. This is because when $d_2 > d_3$, the primary receiver is expected to have a weaker connection to the secondary transmitter. In this case, there is a high probability that these three strategies use the same power allocation factor and decoding order.

3.5.2 Performance Evaluation of P-PA-DO

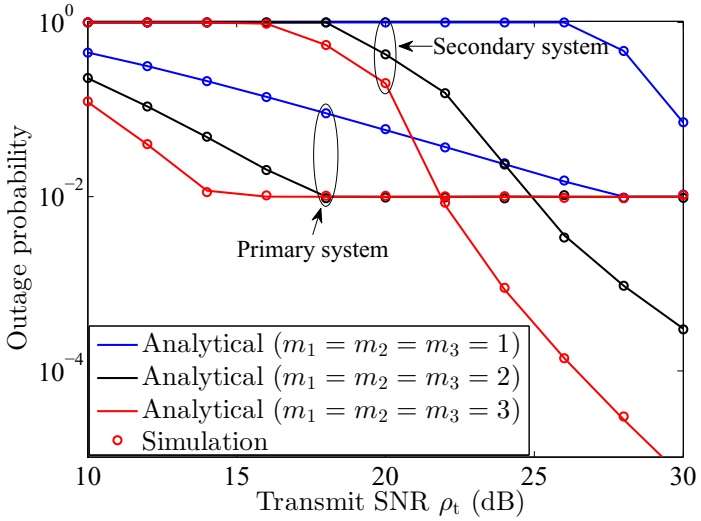


Figure 3.5: Outage performance of the proposed P-PA-DO strategy for different values of fading parameters ($d_1 = d_2 = 1$, $d_3 = 1.5$).

Fig. 3.5 verifies the analytically derived outage probability expressions of the primary and secondary systems in the proposed P-PA-DO strategy. It can be observed that our analytical results well match the simulation results. In addition, we can note that when the transmit SNR ρ_t increases, the outage probability of the secondary system remains at 1 until the QoS requirement of the primary system is satisfied (i.e., the outage probability of the primary system is not more than ς). This is because our proposed P-PA-DO strategy allows the secondary transmitter to allocate all the transmit power to the primary message when the outage probability of the primary system is larger than ς.

Since our analytical results have been verified, we will only present analytical results in the following part of this subsection.

Fig. 3.6 studies the impact of d_1, d_2, and ς on the outage performance of the secondary system. It is easy to tell that the outage performance of the secondary

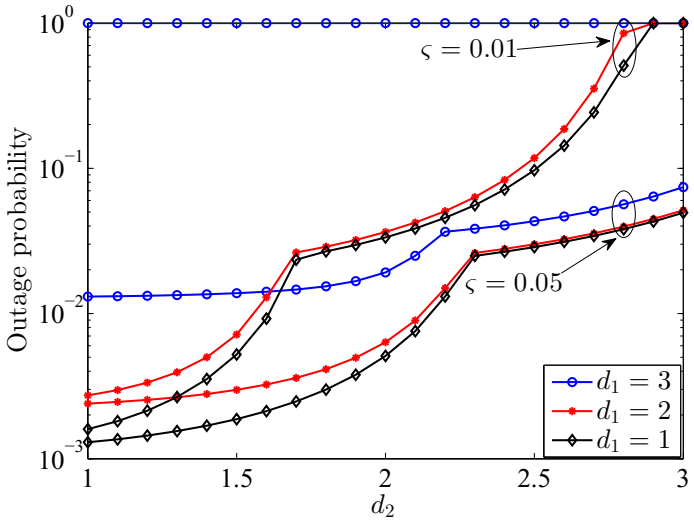


Figure 3.6: Impact of d_1, d_2, and ς on the outage performance of the secondary system ($\rho_{\mathrm{t}} = 30$dB, $d_3 = 2$).

system deteriorates with the increase of d_1 or d_2 or the decrease of ς. This can be explained by the fact that in order to satisfy the outage constraint of the primary system, a larger d_1 or d_2 or a smaller ς requires the secondary transmitter to allocate more power to the primary message and thereby, leave less power for its own signal transmission. In addition, in Fig. 3.6, we can also notice that there is a turning point for each curve except for the curve with $d_1 = 3$ and $\varsigma = 0.01$. The appearance of the turning point can be explained by the fact that as d_2 enlarges (i.e., the channel condition between the secondary transmitter and the primary receiver gets worse), the optimal decoding order changes from the second decoding order to the first decoding order.

Fig. 3.7 compares the outage performance of the proposed P-PA-DO strategy and two benchmark strategies. As existing DP-FD and DD-DP strategies in the literature need full CSI, here we compare with FP-FD strategies for the partial CSI case, as follows.

- FP-FD-1 Strategy: This is the FP-FD strategy adopted in [38–41], in which the power allocation factor for the primary message x_{p} is larger than that for the secondary message x_{s}, and the decoding order is from the primary message to the secondary message. In the simulation, the FP-FD-1 strategy takes the power allocation factor for the primary message as 0.8. Under this fixed decoding order

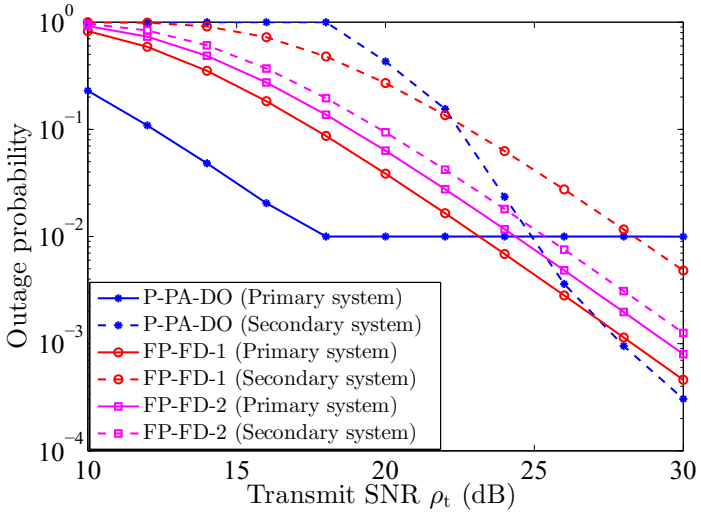


Figure 3.7: Outage performance of the P-PA-DO, FP-FD-1, and FP-FD-2 strategies ($d_1 = d_2 = 1$, $d_3 = 1.5$).

and power allocation factor, the outage probability of the primary and secondary systems can be easily derived as $P_{\rm p}^{\mathbf{1}}(0.8)$ and $P_{\rm s}^{\mathbf{1}}(0.8)$, where $P_{\rm p}^{\mathbf{1}}(\alpha)$ and $P_{\rm s}^{\mathbf{1}}(\alpha)$ have been given in (3.20) and (3.21), respectively.

- FP-FD-2: this is the FP-FD strategy adopted in [48, 51], in which the power allocation factor for the far user is larger than that for the near user, and the decoding order is from the far user's message to the near user's message. In the simulation, the secondary receiver is far user and the primary receiver is near user. So the FP-FD-2 strategy takes the second decoding order, and the power allocation factor for the primary message is set to be 0.15. Under this fixed decoding order and power allocation factor, the outage probability of the primary and secondary systems can be easily derived as $P_{\rm p}^{\mathbf{2}}(0.15)$ and $P_{\rm s}^{\mathbf{2}}(0.15)$, where $P_{\rm p}^{\mathbf{2}}(\alpha)$ and $P_{\rm s}^{\mathbf{2}}(\alpha)$ have been given in (3.22) and (3.23), respectively.

From Fig. 3.7, we have the following observations. When $\rho_{\rm t} < 18$ dB, none of the three strategies can guarantee the outage constraint of the primary system (i.e., guarantee that outage probability of the primary system is not more than $\varsigma = 0.01$). But our proposed P-PA-DO strategy achieves the best outage performance of the primary system. When $\rho_{\rm t}$ is between 18 dB and 24 dB, the proposed P-PA-DO strategy can guarantee the outage constraint of the primary system, while the two benchmark strategies cannot. When $\rho_{\rm t}$ is above 25 dB, all three strategies can guarantee the out-

age constraint of the primary system. But our proposed P-PA-DO strategy achieves the best outage performance for the secondary system.

3.6 Summary

This chapter has investigated the joint optimization of power allocation and decoding order in an overlay cognitive NOMA network to enhance the performance of the secondary system under different QoS demands of the primary system. When the secondary transmitter has full CSI, a novel F-PA-DO strategy is proposed, which can minimize the outage probability of the secondary system while guaranteeing the target rate of the primary system. When the secondary transmitter only has partial CSI, a novel P-PA-DO strategy is proposed, which is able to minimize the outage probability of the secondary system under the outage constraint of the primary system. Our analytical results have shown that in order to enhance the performance of the secondary system while satisfying the QoS requirement of the primary system, it is better to treat the secondary receiver as the "weak user" and let the primary receiver perform SIC when the channel condition between the secondary transmitter and the primary receiver exceeds a certain level. This observation is quite different from most of the current works on NOMA networks with user priority difference, in which the higher priority user is always treated as the "weak user" and the lower priority user performs SIC. Furthermore, simulation results have validated our analytical results and demonstrated the superiority of our proposed strategies, while shedding light on the interplay between key system parameters and the optimized outage performance.

3.7 Appendix

3.7.1 Proof of Corollary 1

We first give some useful properties which will be frequently invoked later. When the fading severity parameter, i.e., m_i, is an integer, the PDF and CCDF of $|h_i|^2$, $i = 1, 2, 3$, are respectively given by

$$f_{|h_i|^2}(x) = \left(\frac{m_i}{\Omega_i}\right)^{m_i} \frac{x^{m_i-1}}{\Gamma(m_i)} e^{-\frac{x m_i}{\Omega_i}}, x > 0, \tag{3.28}$$

and

$$\bar{F}_{|h_i|^2}(x) = \Psi\left(m_i, \frac{x}{\Omega_i}\right) = e^{-\frac{xm_i}{\Omega_i}} \sum_{t=0}^{m_i-1} \frac{1}{t!}\left(\frac{xm_i}{\Omega_i}\right)^t. \tag{3.29}$$

Now we proceed to the derivation of $P_{\mathrm{p}}^{\mathrm{F}}$. According to the outage definition for the primary system, the outage probability of the primary system can be evaluated as

$$\begin{aligned} P_{\mathrm{p}}^{\mathrm{F}} &= 1 - \Pr\left\{R_1 \geq \tilde{R}_{\mathrm{p}}\right\} \Pr\left\{R_{2,x_{\mathrm{p}}}^{\mathrm{F}} \geq \tilde{R}_{\mathrm{p}}\right\} \\ &= 1 - \Pr\left\{|h_1|^2 \geq \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}}\right\} \Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}}\right\} \\ &= 1 - \bar{F}_{|h_1|^2}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}}\right) \bar{F}_{|h_2|^2}\left(\frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}}\right) \\ &= 1 - \Psi\left(m_1, \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}\Omega_1}\right) \Psi\left(m_2, \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}\Omega_2}\right). \end{aligned} \tag{3.30}$$

3.7.2 Proof of Corollary 2

According to the outage definition for the secondary system, the outage probability of the secondary system for the proposed F-PA-DO strategy can be evaluated as

$$P_{\mathrm{s}}^{\mathrm{F}} = 1 - \Pr\left\{R_1 \geq \tilde{R}_{\mathrm{p}}\right\} \Pr\left\{R_{3,x_{\mathrm{s}}}^{\mathrm{F}} \geq \tilde{R}_{\mathrm{s}}\right\} = 1 - \Psi\left(m_1, \frac{\gamma_{\mathrm{th}}^{\mathrm{p}}}{\rho_{\mathrm{t}}\Omega_1}\right) Q_1, \tag{3.31}$$

where $Q_1 = \Pr\left\{R^{\mathrm{F}}_{3,x_\mathrm{s}} \geq \tilde{R}_\mathrm{s}\right\}$. According to the expression of $R^{\mathrm{F}}_{3,x_\mathrm{s}}$ given in (3.16), Q_1 can be calculated as follows

$$
\begin{aligned}
Q_1 &= \Pr\left\{|h_2|^2 \geq \max\left(\frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2\right), \Lambda_1 \geq \gamma^{\mathrm{s}}_{\mathrm{th}}\right\} \\
&\quad + \Pr\left\{|h_2|^2 < \max\left(\frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2\right), \Lambda_2 \geq \gamma^{\mathrm{s}}_{\mathrm{th}}\right\} \\
&= \Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_2|^2 \geq |h_3|^2 \geq \frac{\gamma^{\mathrm{s}}_{\mathrm{th}}|h_2|^2}{\rho_\mathrm{t}|h_2|^2 - \gamma^{\mathrm{ss}}_{\mathrm{th}}}\right\} \\
&\quad + \Pr\left\{|h_3|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2 \geq |h_2|^2 \geq \frac{\gamma^{\mathrm{p}}_{\mathrm{th}}|h_3|^2}{\rho_\mathrm{t}|h_3|^2 - \gamma^{\mathrm{pp}}_{\mathrm{th}}}\right\} \\
&= \Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_2|^2 \geq |h_3|^2 \geq \frac{\gamma^{\mathrm{s}}_{\mathrm{th}}|h_2|^2}{\rho_\mathrm{t}|h_2|^2 - \gamma^{\mathrm{ss}}_{\mathrm{th}}}\right\} && \text{(3.32a)} \\
&\quad + \Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2 \geq |h_2|^2 \geq \frac{\gamma^{\mathrm{s}}_{\mathrm{th}}|h_2|^2}{\rho_\mathrm{t}|h_2|^2 - \gamma^{\mathrm{ss}}_{\mathrm{th}}}\right\} && \text{(3.32b)} \\
&\quad + \Pr\left\{|h_3|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2 \geq |h_2|^2 \geq \frac{\gamma^{\mathrm{p}}_{\mathrm{th}}|h_3|^2}{\rho_\mathrm{t}|h_3|^2 - \gamma^{\mathrm{pp}}_{\mathrm{th}}}\right\} && \text{(3.32c)} \\
&\quad + \Pr\left\{|h_3|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_2|^2 \geq |h_3|^2 \geq \frac{\gamma^{\mathrm{p}}_{\mathrm{th}}|h_3|^2}{\rho_\mathrm{t}|h_3|^2 - \gamma^{\mathrm{pp}}_{\mathrm{th}}}\right\} && \text{(3.32d)} \\
&\quad - \Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2 \geq |h_2|^2 \geq \frac{\gamma^{\mathrm{s}}_{\mathrm{th}}|h_2|^2}{\rho_\mathrm{t}|h_2|^2 - \gamma^{\mathrm{ss}}_{\mathrm{th}}}\right\} && \text{(3.32e)} \\
&\quad - \Pr\left\{|h_3|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_2|^2 \geq |h_3|^2 \geq \frac{\gamma^{\mathrm{p}}_{\mathrm{th}}|h_3|^2}{\rho_\mathrm{t}|h_3|^2 - \gamma^{\mathrm{pp}}_{\mathrm{th}}}\right\} && \text{(3.32f)} \\
&= \underbrace{\Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2 \geq \frac{\gamma^{\mathrm{s}}_{\mathrm{th}}|h_2|^2}{\rho_\mathrm{t}|h_2|^2 - \gamma^{\mathrm{ss}}_{\mathrm{th}}}\right\}}_{I_1} && \text{(3.32g)} \\
&\quad + \underbrace{\Pr\left\{|h_3|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_2|^2 \geq \frac{\gamma^{\mathrm{p}}_{\mathrm{th}}|h_3|^2}{\rho_\mathrm{t}|h_3|^2 - \gamma^{\mathrm{pp}}_{\mathrm{th}}}\right\}}_{I_2} && \text{(3.32h)} \\
&\quad - \underbrace{\Pr\left\{|h_2|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}, |h_3|^2 \geq \frac{\gamma_{\mathrm{th}}}{\rho_\mathrm{t}}\right\}}_{I_3}, && \text{(3.32i)}
\end{aligned}
$$

where $\gamma^{\mathrm{ss}}_{\mathrm{th}} \triangleq \gamma_{\mathrm{th}} - \gamma^{\mathrm{s}}_{\mathrm{th}}$, $\gamma^{\mathrm{pp}}_{\mathrm{th}} \triangleq \gamma_{\mathrm{th}} - \gamma^{\mathrm{p}}_{\mathrm{th}}$, (3.32g) is obtained by combining (3.32a) and (3.32b), (3.32h) is obtained by combining (3.32c) and (3.32d), and (3.32i) is obtained by combining (3.32e) and (3.32f). By applying the CCDF of $|h_2|^2$ and $|h_3|^2$, we can

straightforwardly obtain

$$I_3 = \Psi\left(m_2, \frac{\gamma_{\rm th}}{\rho_{\rm t}\Omega_2}\right)\Psi\left(m_3, \frac{\gamma_{\rm th}}{\rho_{\rm t}\Omega_3}\right). \tag{3.33}$$

Now we proceed to the derivation of I_1, which can be expressed as

$$\begin{aligned} I_1 &= \int_{\frac{\gamma_{\rm th}}{\rho_{\rm t}}}^{\infty} \bar{F}_{|h_3|^2}\left(\frac{\gamma_{\rm th}^{\rm s} x}{\rho_{\rm t} x - \gamma_{\rm th}^{\rm ss}}\right) f_{|h_2|^2}(x)\, dx \\ &\overset{\text{(i)}}{=} \underbrace{\int_0^{\infty} \bar{F}_{|h_3|^2}\left(\frac{\gamma_{\rm th}^{\rm s}(u+\gamma_{\rm th}^{\rm ss})}{u\rho_{\rm t}}\right) f_{|h_2|^2}\left(\frac{u+\gamma_{\rm th}^{\rm ss}}{\rho_{\rm t}}\right)\frac{du}{\rho_{\rm t}}}_{I_{11}} \qquad (3.34a) \\ &\quad - \underbrace{\int_0^{\gamma_{\rm th}^{\rm s}} \bar{F}_{|h_3|^2}\left(\frac{\gamma_{\rm th}^{\rm s}(u+\gamma_{\rm th}^{\rm ss})}{u\rho_{\rm t}}\right) f_{|h_2|^2}\left(\frac{u+\gamma_{\rm th}^{\rm ss}}{\rho_{\rm t}}\right)\frac{du}{\rho_{\rm t}}}_{I_{12}}, \qquad (3.34b) \end{aligned}$$

where step (i) follows from a variable substitution $u = \rho_{\rm t} x - \gamma_{\rm th}^{\rm ss}$. By substituting the PDF of $|h_2|^2$ and CCDF of $|h_3|^2$ into (3.34a), we can solve the integral in I_{11} with the help of binomial expansion and [52, Eq. (3.471.9)]. The result of I_{11} is given by

$$I_{11} = \sum_{l_1=0}^{m_3-1}\sum_{l_2=0}^{m_2-1+l_1} \varepsilon_1 \Phi\left(\frac{\gamma_{\rm th}^{\rm s}\gamma_{\rm th}^{\rm ss} m_3}{\rho_{\rm t}\Omega_3}, \frac{m_2}{\rho_{\rm t}\Omega_2}, l_2 - l_1\right), \tag{3.35}$$

where the definition of $\Phi(\cdot,\cdot,\cdot)$ has been given in Corollary 2, and

$$\begin{aligned} \varepsilon_1 =& e^{-\frac{\gamma_{\rm th}^{\rm ss} m_2}{\rho_{\rm t}\Omega_2} - \frac{\gamma_{\rm th}^{\rm s} m_3}{\rho_{\rm t}\Omega_3}} \left(\frac{m_2}{\rho_{\rm t}\Omega_2}\right)^{m_2} \frac{1}{\Gamma(m_2)}\frac{1}{l_1!} \\ &\times \left(\frac{\gamma_{\rm th}^{\rm s} m_3}{\rho_{\rm t}\Omega_3}\right)^{l_1} \binom{m_2 - 1 + l_1}{l_2} (\gamma_{\rm th}^{\rm ss})^{m_2 - 1 + l_1 - l_2}. \end{aligned} \tag{3.36}$$

It is difficult to obtain an accurate closed-form expression of I_{12} given in (3.34b). However, we can resort to the Gaussian-Chebyshev quadrature method to find an approximation for I_{12} as follows

$$\begin{aligned} I_{12} &\overset{\text{(ii)}}{=} \frac{\gamma_{\rm th}^{\rm s}}{2\rho_{\rm t}} \int_{-1}^{1} \bar{F}_{|h_3|^2}\left(\frac{(v+1)\gamma_{\rm th}^{\rm s} + 2\gamma_{\rm th}^{\rm ss}}{(v+1)\rho_{\rm t}}\right) \\ &\quad \times f_{|h_2|^2}\left(\frac{(v+1)\gamma_{\rm th}^{\rm s} + 2\gamma_{\rm th}^{\rm ss}}{2\rho_{\rm t}}\right) dv \\ &\overset{\text{(iii)}}{\approx} \frac{\gamma_{\rm th}^{\rm s}\pi}{2N} \sum_{l_1=0}^{m_3-1}\sum_{l_2=0}^{m_2-1+l_1}\sum_{n=0}^{N} \sqrt{1-(\theta_n)^2}\,\varepsilon_1 \qquad (3.37) \\ &\quad \times \left(\gamma_{\rm th}^{\rm s}\frac{\theta_n+1}{2}\right)^{l_2-l_1} e^{-\frac{(\theta_n+1)m_2\gamma_{\rm th}^{\rm s}}{2\rho_{\rm t}\Omega_2} - \frac{2\gamma_{\rm th}^{\rm ss} m_3}{\rho_{\rm t}\Omega_3(\theta_n+1)}}, \end{aligned}$$

where step (ii) follows from a variable substitution $u = \frac{\gamma_{\rm th}^{\rm s}+1}{2}v$, and Gaussian-Chebyshev quadrature [36, Eq. (25.4.38)] is employed in step (iii).

Combining (3.35) with (3.37), we can obtain an approximation for I_1 in (3.32g), which is denoted as $\Upsilon\left(m_2, \Omega_2, m_3, \Omega_3, \gamma_{\rm th}^{\rm s}\right)$. Following similar steps, we can also find an approximation for I_2 in (3.32h), which is denoted as $\Upsilon\left(m_3, \Omega_3, m_2, \Omega_2, \gamma_{\rm th}^{\rm p}\right)$. Combing (3.33) with the two approximation results for I_1 and I_2, an approximation for Q_1 can be derived. Applying the result of Q_1 into (3.31), we can obtain $P_{\rm s}^{\rm F}$ given in (3.18).

3.7.3 Proof of Lemma 2

Here we only present the details regarding the derivation of $P_{\rm p}^{\mathbf{1}}(\alpha)$. The derivation details of $P_{\rm s}^{\mathbf{1}}(\alpha)$, $P_{\rm p}^{\mathbf{2}}(\alpha)$ and $P_{\rm s}^{\mathbf{2}}(\alpha)$ are omitted since they can be obtained by following the similar procedures in the derivation of $P_{\rm p}^{\mathbf{1}}(\alpha)$.

Under the first decoding order with the given power allocation factor α, the outage probability of the primary system can be evaluated as

$$\begin{aligned} P_{\rm p}^{\mathbf{1}}(\alpha) &= 1 - \Pr\Big\{R_1 \geq \tilde{R}_{\rm p}\Big\}\Pr\Big\{R_{2,x_{\rm p}}^{\mathbf{1}}(\alpha) \geq \tilde{R}_{\rm p}\Big\} \\ &= 1 - \Pr\left\{|h_1|^2 \geq \frac{\gamma_{\rm th}^{\rm p}}{\rho_{\rm t}}\right\}\Pr\left\{\frac{\rho_{\rm t}|h_2|^2}{\xi} \geq 1\right\} \\ &\overset{\rm (iv)}{=} 1 - \Psi\left(m_1, \frac{\gamma_{\rm th}^{\rm p}}{\rho_{\rm t}\Omega_1}\right)\Psi\left(m_2, \frac{\xi}{\rho_{\rm t}\Omega_2}\right), \end{aligned} \tag{3.38}$$

where step (iv) proceeds on the condition that $\xi = \frac{\gamma_{\rm th}^{\rm p}}{\alpha\left(1+\gamma_{\rm th}^{\rm p}\right)-\gamma_{\rm th}^{\rm p}} > 0$, i.e., $\frac{\gamma_{\rm th}^{\rm p}}{1+\gamma_{\rm th}^{\rm p}} < \alpha \leq 1$.

Chapter 4

Optimal Designs for Relay-Assisted NOMA Networks with Hybrid SWIPT Scheme

A relay-assisted NOMA network consisting of a source (S), an energy-constrained relay (R), and two users is studied in this chapter. Specifically, R adopts the more general and more powerful SWIPT scheme, i.e., the H-SWIPT scheme, to harvest energy. We provide optimal system designs for such network for both the scenario with full CSIT and the scenario with partial CSIT. In particular, when full CSIT is available, we intend to minimize the energy consumption at S while guaranteing the successful transmission (both users correctly receive the desired information) by jointly optimizing the transmit power of S, the TS and PS ratios, the power allocation ratios at S and R, and the user ordering for two users. When only partial CSIT is available, we aim at minimizing the system outage probability by jointly optimizing the TS and PS ratios, the power allocation ratios, and the user ordering. We solve the two problems with a bisection method and a one-dimension search, respectively. Analytical results demonstrate that users in the first scenario should be ordered by their instantaneous channel conditions, while in the second scenario, the optimal user ordering is jointly decided by the average channel gains and the decoding thresholds of two users. Finally, numerical results are provided to validate the analytical results and show the performance advantage of the H-SWIPT scheme with the proposed optimal designs.

4.1 Introduction

It is well known that relay-assisted networks have a wider coverage and thus making the long-distance NOMA transmission more reliable [53–57]. However, the advantage of relay-assisted networks brings at the cost of energy consumption at the relay nodes, which may hinder more devices from becoming relay candidates, especially for those energy-constrained devices. With the arrival of IoT era, a large number of sensor nodes that potentially can serve as relay nodes will be deployed everywhere. However, owing to the size limit, most of the sensor nodes carry small batteries, thus they may be reluctant to support other users' communication at their own energy expense. To alleviate this concern, SWIPT, one of the EH techniques enabling receivers to simultaneously extract energy and information from the received information-bearing signals, can be integrated into cooperative NOMA networks.

Up to now, relay-assisted NOMA networks with the PS or the TS scheme have been widely studied in [58–64]. However, to the best of our knowledge, there are no reported works involving the relay-assisted NOMA networks with the H-SWIPT scheme. Note that the H-SWIPT scheme is expected to outperform the TS and PS schemes since it can exploit both the degree-of-freedom in time and power domains. In view of this fact, we study a relay-assisted NOMA network with the H-SWIPT scheme, in which a source (S) communicates with two users (U_1 and U_2) under the help of an EH-based relay (R). Depending on the availability of the CSITs, we study two different resource allocation problems. Specifically, when S and R have full CSIT, i.e., instantaneous CSI of all the three links (we assume that there are no direct links between S and U_1/U_2), we aim at realizing a successful transmission with the minimum energy consumption at S by jointly optimizing the transmit power of S, TS and PS ratios, power allocation ratios at S and R, and user ordering for U_1 and U_2. Here, a successful transmission refers to the event that both U_1 and U_2 correctly decode the desired information. Since the acquisition of full CSIT involves considerable overhead, we also study a more practical scenario where transmitters are assumed to have only partial CSIT, i.e., the statistical CSI of all the three links. In this case, we aim at minimizing the system outage probability by jointly optimizing the TS and PS ratios, the power allocation ratios at S and R, and the user ordering.

Here, system outage refers to the event that either U_1 or U_2 fails to decode the desired information.

It is worth mentioning that although the optimal design for relay-assisted networks with the H-SWIPT scheme has been studied in [32], its results cannot be applied in our considered networks due to the following reasons. The work in [32] studies a relay-assisted single-user network with full CSIT and maximizes the throughput by jointly optimizing the TS and PS ratios. However, we consider a two-user downlink NOMA network with full CSIT and study an energy-efficiency maximization problem which involves more than just optimization of the TS and PS ratios. In addition, for the scenario with partial CSIT (which is not considered in [32]), a reliability optimal design is proposed in this work. Another thing needs to mention is that user ordering plays an important role in NOMA-based systems because it determines the SIC decoding order at receivers and thus affects the achievable rate of each user. However, in most of relay-assisted NOMA works, users are ordered by different predefined criteria which are not proved to be optimal [18, 38, 44, 48, 58, 59, 62, 65–67]. In this work, we drop the assumption of predefined user ordering criteria and analytically find the optimal user ordering by solving the optimization problems. The major contributions and results are summarized as follows:

- We introduce the more generalized and more powerful SWIPT scheme, i.e., the H-SWIPT scheme, into relay-assisted NOMA networks, and study two novel joint optimization problems.

- When full CSIT is available, we first obtain the closed-form expressions of the joint optimal solution when the transmit power of S and the TS ratio are fixed. Then, the optimal transmit power of S and the optimal TS ratio are found using a bisection method. Our analytical results demonstrate that when full CSIT is available, two users should be ordered according to the instantaneous channel conditions between them and R, which follows the user ordering criterion widely used in current NOMA works.

- When only partial CSIT is available, we successfully obtain the closed-form expressions of the joint optimal solution for a given TS ratio, and its corresponding system outage probability. Based on the closed-form expression of the system

outage probability, the optimal TS ratio is derived with a one-dimension (1-D) search. Analytical results reveal that to minimize the system outage probability, the optimal user ordering should be jointly decided by the average channel gains and the decoding thresholds of two users, which is a novel user ordering criterion.

- To provide benchmarks for the consider network with the H-SWIPT scheme, we also briefly discuss the optimal designs for the considered network with the PS and TS schemes under different CSIT-availability scenarios.

The rest of this chapter is organized as follows. Section 4.2 introduces the system model and the transmission protocol. Section 4.3 and 4.4 formulate the optimization problem and present the joint optimal solution for the full CSIT and the partial CSIT cases, respectively. In Section 4.5, numerical results are provided to verify the analytical results.

4.2 System Model

We consider a downlink network with a source (S), an EH-based relay (R), and two users (U_1, U_2), where each node has one antenna and works in a half-duplex mode. The direct link between S and U_j ($j = 1, 2$) does not exist due to physical obstacles or heavy shadowing, and thus the communication between S and two users is processed under the assistance of R [58, 59, 62–64]. In particular, R, which is energy-constrained, adopts the H-SWIPT scheme to simultaneously harvest energy and process information. Let h_0, h_1 and h_2 denote the channel coefficients of links S→R, R→U_1 and R→U_2, respectively. We consider a quasi-static channel model, and all the channel coefficients remain unchanged within each block of a time length of T and vary independently from one block to the next with $h_i \sim \mathcal{CN}(0, \Omega_i)$, $i = 0, 1, 2$, where Ω_i denotes the average channel gain. Without loss of generality, we hereafter assume that the block time T is normalized to 1. We further assume that all the receivers can perfectly access local CSI, i.e., R and U_j have instantaneous CSI of the links S→R and R→U_j, respectively. Now proceed to the description of the proposed hybrid SWIPT-based cooperative NOMA (HSCN) transmission protocol.

Table 4.1: Description of symbols used for key parameters in Chapter 4.

Symbol	Meaning
h_0	Channel coefficient from the source to the relay
$h_j, j = 1, 2$	Channel coefficient from the relay to user j
$x_j, j = 1, 2$	Message for user j
$\alpha_j, j = 1, 2$	Power allocation factor at the source
$\beta_j, j = 1, 2$	Power allocation factor at the relay
$\tilde{h}_i, i = 0, 1, 2$	Channel power gain
η	Energy conversion efficiency
τ	Time-splitting ratio for energy harvesting
θ	Power-splitting ratio for energy harvesting
P_{S}	Actual transmit power of the source
P_m	Maximum transmit power of the source
P_{R}	Transmit power of the relay
ς	User ordering indicator
U_{ς} and $\mathrm{U}_{\bar{\varsigma}}$	Weak user and strong user
$R_{\mathrm{R},x_{\varsigma}}$	Achievable rate of the weak user's message at the relay
$R_{\mathrm{R},x_{\bar{\varsigma}}}$	Achievable rate of the strong user's message at the relay
$R_{\mathrm{U}_j,x_{\varsigma}}, j = 1, 2$	Achievable rate of the weak user's message at user j
$R_{\mathrm{U}_{\bar{\varsigma}},x_{\bar{\varsigma}}}, j = 1, 2$	Achievable rate of the strong user's message at the strong user
$\widetilde{R}_j, j = 1, 2$	Target rate of user j
$(\alpha_{\varsigma^*}^*, \theta^*, \beta_{\varsigma^*}^*, \varsigma^*, \tau^*, P_{\mathrm{S}}^*)$	Joint optimal solution in the full CSIT case
$\alpha_{\varsigma}^{\circ}$	Optimal power allocation factor at the source in the full CSIT case with other parameters fixed
θ°	Optimal power-splitting ratio in the full CSIT case with other parameters fixed
ς°	Optimal user ordering in the full CSIT case with fixed P_{S} and τ
$\beta_{\varsigma^{\circ}}^{\circ}$	Optimal power allocation factor at the relay in the full CSIT case with fixed P_{S} and τ
P_1^{P}	System outage probability of the optimal PCN protocol in the full CSIT case
$(\alpha_{\varsigma^\star}^\star, \theta^\star, \varsigma^\star, \beta_{\varsigma^\star}^\star, \tau^\star)$	Joint optimal solution in the partial CSIT case
$\alpha_{\varsigma}^{\bullet}$	Optimal power allocation factor at the source in the partial CSIT case with other parameters fixed
$\theta^{\bullet}$	Optimal power-splitting ratio in the partial CSIT case with other parameters fixed
$\varsigma^{\bullet}$	Optimal user ordering in the partial CSIT case with fixed P_{S} and τ
$\beta_{\varsigma^{\bullet}}^{\bullet}$	Optimal power allocation factor at the relay in the partial CSIT case with fixed P_{S} and τ
$P_2^{\mathrm{H}}(\tau)$	System outage probability of the optimal HSCN protocol with a given τ in the partial CSIT case
$P_2^{\mathrm{T}}(\tau)$	System outage probability of the optimal TCN protocol with a given τ in the partial CSIT case
P_2^{P}	System outage probability of the optimal PCN protocol in the partial CSIT case

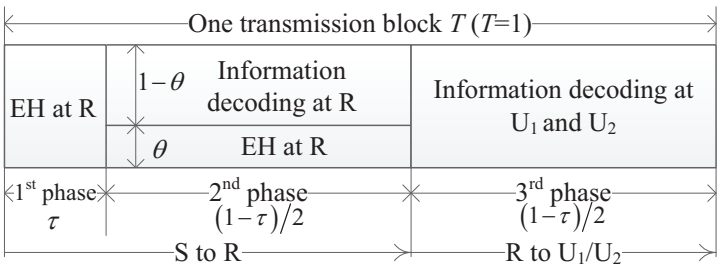


Figure 4.1: Hybrid SWIPT-based cooperative NOMA (HSCN) transmission protocol.

As depicted in Fig. 4.1, the HSCN transmission protocol consists of three phases. The first phase is dedicated for EH at R with a duration τ, where $0 \leq \tau < 1$ is the TS ratio for the first phase. Letting P_{S} denote the transmit power of S, then the harvested energy during the first phase is given by $E_1 = \eta\tau P_{\mathrm{S}}\tilde{h}_0$, where $0 < \eta < 1$ is the energy conversion efficiency and $\tilde{h}_i \triangleq |h_i|^2$, $i = 0, 1, 2$.

After the first phase, the remaining time is equally divided into two phases, each with a duration time given by $\frac{1-\tau}{2}$.[1] In the second phase, S sends a superimposed signal consisting of x_1 and x_2, where x_j is the message intended for U_j, $j = 1, 2$. The received signal at R can be expressed as

$$y_{\mathrm{R}} = \sqrt{\alpha_1 P_{\mathrm{S}}} h_0 x_1 + \sqrt{\alpha_2 P_{\mathrm{S}}} h_0 x_2 + n_{\mathrm{R}}, \tag{4.1}$$

where α_j is the power allocation factor for x_j with $\alpha_1 + \alpha_2 = 1$ and $n_{\mathrm{R}} \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN at R. After receiving y_{R}, R splits it into two parts, where θ $(0 \leq \theta < 1)$ portion of the signal is used for EH and the remaining part is used for information processing.

Given θy_{R}, the harvested energy at R during the second phase can be expressed as $E_2 = \eta P_{\mathrm{S}}\tilde{h}_0\theta\frac{1-\tau}{2}$. Therefore, the total harvested energy during the first two phases is given by

$$E_t = E_1 + E_2 = \eta P_{\mathrm{S}}\tilde{h}_0 \left(\tau + \frac{1-\tau}{2}\theta\right). \tag{4.2}$$

R employs the DF scheme to tackle the remaining part $(1-\theta)\, y_{\mathrm{R}}$. Assume that U_1 and U_2 is ordered as $(\mathrm{U}_\varsigma, \mathrm{U}_{\bar{\varsigma}})$, where ς denotes the user ordering indicator, U_ς and $\mathrm{U}_{\bar{\varsigma}}$ represent the weak and strong users, respectively, $(\varsigma, \bar{\varsigma}) \in \{(1,2), (2,1)\}$. According to the SIC decoding principle [15], R first decodes x_ς and treats $x_{\bar{\varsigma}}$ as noise, then the achievable rate of x_ς at R is given by

$$R_{\mathrm{R},x_\varsigma} = \frac{1-\tau}{2}\log_2\left(1 + \frac{\alpha_\varsigma \rho_{\mathrm{S}}\tilde{h}_0(1-\theta)}{\alpha_{\bar{\varsigma}}\rho_{\mathrm{S}}\tilde{h}_0(1-\theta)+1}\right), \tag{4.3}$$

where $\rho_{\mathrm{S}} = \frac{P_{\mathrm{S}}}{\sigma^2}$ denotes the transmit SNR of S. Let $\tilde{R}_j$ denote the targeted data rate of U_j, $j = 1, 2$. Conditioned on R successfully decodes x_ς (i.e., $R_{\mathrm{R},x_\varsigma} \geq \tilde{R}_\varsigma$) and

[1] To reduce implementation complexity, we here assume that the last two phases, i.e., information transfer from S to R and R to users, have an equal length so that S and R can use the same transmission rate. Note that this assumption is widely adopted in relay-assisted networks [38, 44, 48, 58, 62, 66–68].

performs SIC, the achievable rate of $x_{\bar{\varsigma}}$ at R is given by

$$R_{\mathrm{R},x_{\bar{\varsigma}}} = \frac{1-\tau}{2}\log_2\left(1+\alpha_{\bar{\varsigma}}\rho_{\mathrm{S}}\tilde{h}_0\left(1-\theta\right)\right). \tag{4.4}$$

If R successfully decodes x_1 and x_2, it will re-superimpose x_1 and x_2 and forward the new superimposed signal to U_1 and U_2 in the third phase, then the received signal at U_j $(j=1,2)$ can be written as

$$y_{\mathrm{U}_j} = \sqrt{\beta_1 P_{\mathrm{R}}}h_j x_1 + \sqrt{\beta_2 P_{\mathrm{R}}}h_j x_2 + n_{\mathrm{U}_j}, \tag{4.5}$$

where β_j is the power allocation factor for message x_j with $\beta_1+\beta_2=1$, $n_{\mathrm{U}_j} \sim \mathcal{CN}\left(0,\sigma^2\right)$ is the AWGN at U_j, and P_{R} is the transmit power of R. According to (4.2), P_{R} can be expressed as

$$P_{\mathrm{R}} = \frac{E_t}{\left(1-\tau\right)/2} = \eta P_{\mathrm{S}}\tilde{h}_0\left(\frac{2\tau}{1-\tau}+\theta\right). \tag{4.6}$$

Upon receiving y_{U_j}, U_j first decodes x_{ς} while treating $x_{\bar{\varsigma}}$ as noise, then the achievable rate of x_{ς} at U_j is given by

$$R_{\mathrm{U}_j,x_{\varsigma}} = \frac{1-\tau}{2}\log_2\left(1+\frac{\beta_{\varsigma}\rho_{\mathrm{R}}\tilde{h}_j}{\beta_{\bar{\varsigma}}\rho_{\mathrm{R}}\tilde{h}_j+1}\right), \tag{4.7}$$

where $\rho_{\mathrm{R}} = \frac{P_{\mathrm{R}}}{\sigma^2}$ denotes the transmit SNR of R. Provided that $\mathrm{U}_{\bar{\varsigma}}$ successfully decodes x_{ς} (i.e., $R_{\mathrm{U}_{\bar{\varsigma}},x_{\varsigma}} \geq \tilde{R}_{\varsigma}$) and performs SIC, the achievable rate of $x_{\bar{\varsigma}}$ at $\mathrm{U}_{\bar{\varsigma}}$ is given by

$$R_{\mathrm{U}_{\bar{\varsigma}},x_{\bar{\varsigma}}} = \frac{1-\tau}{2}\log_2\left(1+\beta_{\bar{\varsigma}}\rho_{\mathrm{R}}\tilde{h}_{\bar{\varsigma}}\right). \tag{4.8}$$

To this end, we have introduced the HSCN transmission protocol. Now let's focus on the optimal resource allocation policies under two different CSIT-availability scenarios.

4.3 Optimal Transmission Design with Full CSIT

In this section, we study the optimal resource allocation policy for the considered network with full CSIT. We first formulate the optimization problem, and then discuss the optimal design for the HSCN protocol. Finally, the optimal designs for the PS-based and TS-based transmission protocols are briefly discussed.

4.3.1 Problem Formulation

Note that when full CSIT is available at S, the transmit power of S can be adapted according to the instantaneous CSI under a peak transmit power constraint P_m. In such a scenario, we aim at minimizing the energy consumption at S while achieving successful transmission within each fading block. Here, successful transmission means that U_1 and U_2 correctly decode x_1 and x_2, respectively. According to our proposed HSCN transmission protocol, S works in the first two phases with a transmit power P_{S}, thus the total energy consumption at S within one fading block can be expressed as $P_{\mathrm{S}}\left(\tau+\frac{1-\tau}{2}\right)=P_{\mathrm{S}}\frac{1+\tau}{2}$. Accordingly, the transmit energy minimization problem can be formulated as

$$\mathbf{P1}: \min_{\alpha_\varsigma,\theta,\beta_\varsigma,\varsigma,\tau,P_{\mathrm{S}}} P_{\mathrm{S}}\frac{1+\tau}{2} \tag{4.9a}$$

$$\text{s.t. } R_{\mathrm{R},x_\varsigma}\geq\tilde{R}_\varsigma, R_{\mathrm{R},x_{\bar{\varsigma}}}\geq\tilde{R}_{\bar{\varsigma}}, \tag{4.9b}$$

$$R_{\mathrm{U}_\varsigma,x_\varsigma}\geq\tilde{R}_\varsigma, R_{\mathrm{U}_{\bar{\varsigma}},x_\varsigma}\geq\tilde{R}_\varsigma, R_{\mathrm{U}_{\bar{\varsigma}},x_{\bar{\varsigma}}}\geq\tilde{R}_{\bar{\varsigma}}, \tag{4.9c}$$

$$P_{\mathrm{S}}\leq P_m, \varsigma\in\{1,2\}, 0\leq\tau<1, 0\leq\theta<1, \tag{4.9d}$$

$$\frac{\gamma_\varsigma(\tau)}{1+\gamma_\varsigma(\tau)}<\alpha_\varsigma<1, \frac{\gamma_\varsigma(\tau)}{1+\gamma_\varsigma(\tau)}<\beta_\varsigma<1, \tag{4.9e}$$

where $\gamma_j(\tau)\triangleq 2^{\frac{2\tilde{R}_j}{1-\tau}}-1$, $j=1,2$. In **P1**, constraints in (4.9b)-(4.9c) ensure that a successful transmission is realized. In particular, constraints in (4.9b) guarantee that R successfully decodes x_1 and x_2, and constraints in (4.9c) guarantee that U_1 and U_2 successfully detect the desired information from the forwarded signal. The constraints regarding the power allocation factors α_ς and β_ς given in (4.9e) are derived from the fact that to realize a non-zero probability SIC at R and $\mathrm{U}_{\bar{\varsigma}}$, the upper bounds of $R_{\mathrm{R},x_\varsigma}$ and $R_{\mathrm{U}_{\bar{\varsigma}},x_\varsigma}$ should be larger than the targeted data rate $\tilde{R}_\varsigma$, i.e., $\lim_{\rho_{\mathrm{S}}\to\infty} R_{\mathrm{R},x_\varsigma}=\log_2\left(1+\frac{\alpha_\varsigma}{\alpha_{\bar{\varsigma}}}\right)>\tilde{R}_\varsigma$ and $\lim_{\rho_{\mathrm{R}}\to\infty} R_{\mathrm{U}_{\bar{\varsigma}},x_\varsigma}=\log_2\left(1+\frac{\beta_\varsigma}{\beta_{\bar{\varsigma}}}\right)>\tilde{R}_\varsigma$.

Next, we obtain the joint optimal solution to **P1**, which is denoted by $\left(\alpha_{\varsigma^*}^*,\theta^*,\beta_{\varsigma^*}^*,\varsigma^*,\tau^*,P_{\mathrm{S}}^*\right)$.

4.3.2 Joint Optimal Solution of the HSCN Protocol

Note that **P1** is non-convex and not easy to be solved by standard optimization solvers since the optimization variables are coupled in the objective function and the constraints (4.9b)-(4.9c). To efficiently address **P1**, we decompose it into several

subproblems and solve them one by one. We first derive the optimal power allocation ratio at S and the optimal PS ratio when other variables are fixed. Then, given the fixed TS ratio and the transmit power of S, we conduct the study of the optimal power allocation ratio at R and the optimal user ordering. Finally, the optimal transmit power of S and the optimal TS ratio are obtained based on the results of two previous steps.

4.3.2.1 Step 1 (Optimization of α_ς and θ)

Note that for DF-based relay networks, information transfer from R to users happens only when R correctly decodes the desired information and has available transmit power. In addition, it is easy to tell from (4.7) and (4.8) that the achievable rates of x_1 and x_2 at U_1 and U_2 monotonically increase as the transmit power of R grows. Based on these two observations, we first try to maximize the transmit power of R (P_{R}) while ensuring R correctly detects x_1 and x_2 by jointly optimizing the power allocation ratio at S α_ς and the PS ratio θ when other parameters are given. Accordingly, this problem can be formulated as

$$\mathbf{P1.1}: \max_{\theta,\alpha_\varsigma} P_{\mathrm{R}} \tag{4.10a}$$

$$\text{s.t.} \quad R_{\mathrm{R},x_\varsigma} \geq \tilde{R}_\varsigma, R_{\mathrm{R},x_{\bar{\varsigma}}} \geq \tilde{R}_{\bar{\varsigma}}, \tag{4.10b}$$

$$0 \leq \theta < 1, \frac{\gamma_\varsigma(\tau)}{1+\gamma_\varsigma(\tau)} < \alpha_\varsigma < 1. \tag{4.10c}$$

Proposition 3. ***P1.1*** *is feasible when* $\rho_S \geq \frac{\gamma(\tau)}{\tilde{h}_0}$*, and the optimal solutions to* ***P1.1****, denoted by* α_ς° *and* θ°*, are given by*

$$\alpha_\varsigma^\circ = 1 - \frac{\gamma_{\bar{\varsigma}}(\tau)}{\gamma(\tau)}, \theta^\circ = 1 - \frac{\gamma(\tau)}{\rho_S \tilde{h}_0}, \tag{4.11}$$

where $\gamma(\tau) = 2^{\frac{2(\tilde{R}_1+\tilde{R}_2)}{1-\tau}} - 1$.

Proof. Applying (4.3) and (4.4) into (4.10b), the two constraints in (4.10b) can be equivalently transformed as $\theta \leq 1 - \frac{g(\alpha_\varsigma)}{\rho_S \tilde{h}_0}$, where $g(\alpha_\varsigma) = \max\left\{\frac{\gamma_\varsigma(\tau)}{\alpha_\varsigma - \alpha_{\bar{\varsigma}}\gamma_\varsigma(\tau)}, \frac{\gamma_{\bar{\varsigma}}(\tau)}{\alpha_{\bar{\varsigma}}}\right\}$. From (4.6), we can tell that P_{R} is an increasing function of θ, so P_{R} can be maximized by minimizing $g(\alpha_\varsigma)$. Note that $\frac{\gamma_\varsigma(\tau)}{\alpha_\varsigma - \alpha_{\bar{\varsigma}}\gamma_\varsigma(\tau)}$ and $\frac{\gamma_{\bar{\varsigma}}(\tau)}{\alpha_{\bar{\varsigma}}}$ are respectively decreasing and increasing functions of α_ς. In addition, we have $\lim_{\alpha_\varsigma \to \frac{\gamma_\varsigma(\tau)}{\gamma_\varsigma(\tau)+1}} \frac{\gamma_\varsigma(\tau)}{\alpha_\varsigma - \alpha_{\bar{\varsigma}}\gamma_\varsigma(\tau)} = \infty > \frac{\gamma_{\bar{\varsigma}}(\tau)}{\alpha_{\bar{\varsigma}}}$ and

$\lim_{\alpha_\varsigma \to 1} \frac{\gamma_{\bar{\varsigma}}(\tau)}{\alpha_{\bar{\varsigma}}} = \infty > \frac{\gamma_\varsigma(\tau)}{\alpha_\varsigma - \alpha_{\bar{\varsigma}}\gamma_\varsigma(\tau)}$. Combining these two observations, we know that $g(\alpha_\varsigma)$ achieves its minimum when $\frac{\gamma_\varsigma(\tau)}{\alpha_\varsigma - \alpha_{\bar{\varsigma}}\gamma_\varsigma(\tau)} = \frac{\gamma_{\bar{\varsigma}}(\tau)}{\alpha_{\bar{\varsigma}}}$, from which we can derive the optimal power allocation factor at S α_ς°. After obtaining α_ς°, the minimum value of $g(\alpha_\varsigma)$ can be calculated as $\gamma(\tau)$. Therefore, the optimal PS ratio is $\theta^\circ = 1 - \frac{\gamma(\tau)}{\rho_{\mathrm{S}}\tilde{h}_0}$. Note that $0 \leq \theta < 1$, thus θ° exists only when $\rho_{\mathrm{S}} \geq \frac{\gamma(\tau)}{\tilde{h}_0}$. □

4.3.2.2 Step 2 (Optimization of β_ς and ς)

Provided that information transfer from S to R is successful, we now focus on the conditions required for the successful information transmission from R to users, which are given in (4.9c). First, by applying (4.7) and (4.8) to tackle the constraints in (4.9c), the three constraints can be equivalently converted into $\rho_{\mathrm{R}} - \varphi_\varsigma(\beta_\varsigma) \geq 0$, where $\varphi_\varsigma(\beta_\varsigma) \triangleq \max\left\{\frac{\vartheta_\varsigma}{\tilde{h}_\varsigma}, \frac{\psi_\varsigma}{\tilde{h}_{\bar{\varsigma}}}\right\}$, $\vartheta_\varsigma \triangleq \frac{\gamma_\varsigma(\tau)}{\beta_\varsigma(1+\gamma_\varsigma(\tau))-\gamma_\varsigma(\tau)}$, and $\psi_\varsigma \triangleq \max\left\{\vartheta_\varsigma, \frac{\gamma_{\bar{\varsigma}}(\tau)}{1-\beta_\varsigma}\right\}$. Then, substituting the optimal PS ratio given in (4.11) into (4.6), the transmit SNR of R can be expressed as $\rho_{\mathrm{R}} = \eta\rho_{\mathrm{S}}\tilde{h}_0\frac{1+\tau}{1-\tau} - \eta\gamma(\tau)$. With all these results, **P1** can be equivalently transformed as

$$\mathbf{P1.2}: \min_{\varsigma,\beta_\varsigma,\tau,P_{\mathrm{S}}} P_{\mathrm{S}}\frac{1+\tau}{2} \tag{4.12a}$$

$$\text{s.t.} \quad \eta\rho_{\mathrm{S}}\tilde{h}_0\frac{1+\tau}{1-\tau} - \eta\gamma(\tau) - \varphi_\varsigma(\beta_\varsigma) \geq 0, \tag{4.12b}$$

$$\frac{\gamma(\tau)}{\tilde{h}_0} \leq \rho_{\mathrm{S}} \leq \rho_m, 0 \leq \tau \leq \tau_1, \tag{4.12c}$$

$$\varsigma \in \{1,2\}, \frac{\gamma_\varsigma(\tau)}{1+\gamma_\varsigma(\tau)} < \beta_\varsigma < 1, \tag{4.12d}$$

where $\rho_m = \frac{P_m}{\sigma^2}$ and $\tau_1 = 1 - \frac{2\left(\tilde{R}_1+\tilde{R}_2\right)}{\log_2\left(\rho_m\tilde{h}_0+1\right)}$. Specifically, the constraint $\tau \leq \tau_1$ is given to ensure that $\frac{\gamma(\tau)}{\tilde{h}_0} \leq \rho_m$.

With a careful check for **P1.2**, we can find that the values of the power allocation ratio at R β_ς and the user ordering ς do not affect the objective function in **P1.2**, and they only affect the feasible region by deciding the value of $\varphi_\varsigma(\beta_\varsigma)$. It is straightforward to tell from (4.12b) that the feasible region of **P1.2** can be maximized by minimizing $\varphi_\varsigma(\beta_\varsigma)$. Therefore, given P_{S} and τ, the joint optimal ς and β_ς (denoted by ς° and $\beta_{\varsigma^\circ}^\circ$) can be found by minimizing $\varphi_\varsigma(\beta_\varsigma)$.

Proposition 4. *The optimal user ordering and the optimal power allocation ratio at*

R minimizing $\varphi_{\varsigma}(\beta_{\varsigma})$ *are respectively given by*

$$(\varsigma^{\circ}, \bar{\varsigma}^{\circ}) = \begin{cases} (1,2)\,, \textit{if } \tilde{h}_1 < \tilde{h}_2, \\ (2,1)\,, \textit{else}, \end{cases} \tag{4.13}$$

and

$$\beta_{\varsigma^{\circ}}^{\circ} = \frac{\tilde{h}_{\varsigma^{\circ}} \gamma_1(\tau)\, \gamma_2(\tau) + \tilde{h}_{\bar{\varsigma}^{\circ}} \gamma_{\varsigma^{\circ}}(\tau)}{\tilde{h}_{\varsigma^{\circ}} \gamma(\tau) + \left(\tilde{h}_{\bar{\varsigma}^{\circ}} - \tilde{h}_{\varsigma^{\circ}}\right) \gamma_{\varsigma^{\circ}}(\tau)}. \tag{4.14}$$

Proof. See Section 4.7.1. □

Remark 6. *Proposition 4 indicates that both* ς° *and* $\beta_{\varsigma^{\circ}}^{\circ}$ *are independent of the transmit power of R. In particular,* $\beta_{\varsigma^{\circ}}^{\circ}$ *is jointly decided by the decoding thresholds of two users and the instantaneous channel gains between R and two users, while* ς° *only depends on the instantaneous channel gains between R and two users, thus we have* $\varsigma^{*} = \varsigma^{\circ}$.

4.3.2.3 Step 3 (Optimization of τ and P_{S})

With Proposition 4, the value of $\varphi_{\varsigma}(\beta_{\varsigma})$ is given by $\frac{\gamma(\tau)}{\tilde{h}_{\bar{\varsigma}^{\circ}}} + \frac{\tilde{h}_{\bar{\varsigma}^{\circ}} - \tilde{h}_{\varsigma^{\circ}}}{\tilde{h}_1 \tilde{h}_2} \gamma_{\varsigma^{\circ}}(\tau)$. Substituting this result into (4.12b) and after some simplifications, **P1.2** can be reformulated as

$$\mathbf{P1.3}: \min_{\tau, P_{\mathrm{S}}} P_{\mathrm{S}} \frac{1+\tau}{2} \tag{4.15a}$$

$$\text{s.t.} \quad \widehat{V}_1(\tau) \le \rho_{\mathrm{S}} \le \rho_m, 0 \le \tau \le \tau_1, \tag{4.15b}$$

where $\widehat{V}_1(\tau) = \max\left\{V_1(\tau), \frac{\gamma(\tau)}{\tilde{h}_0}\right\}$ with $V_1(\tau) = \left(\gamma(\tau) \frac{\eta \tilde{h}_{\bar{\varsigma}^{\circ}} + 1}{\eta \tilde{h}_0 \tilde{h}_{\bar{\varsigma}^{\circ}}} + \gamma_{\varsigma^{\circ}}(\tau) \frac{\tilde{h}_{\bar{\varsigma}^{\circ}} - \tilde{h}_{\varsigma^{\circ}}}{\eta \tilde{h}_0 \tilde{h}_1 \tilde{h}_2}\right) \frac{1-\tau}{1+\tau}$.

It is obvious that the objective function in **P1.3** increases as P_{S} grows, thus the optimal transmit power minimizing $P_{\mathrm{S}} \frac{1+\tau}{2}$ is $P_{\mathrm{S}} = \widehat{V}_1(\tau)\, \sigma^2$. Based on this observation, **P1.3** reduces to a 1-D optimization problem, which can be expressed as

$$\mathbf{P1.4}: \min_{0 \le \tau \le \tau_1} \widehat{V}_1(\tau) \frac{1+\tau}{2} \tag{4.16a}$$

$$\text{s.t.} \quad \widehat{V}_1(\tau) = \rho_{\mathrm{S}} \le \rho_m. \tag{4.16b}$$

Before solving **P1.4**, the feasible conditions of **P1.4** are first given.

Proposition 5. *The feasible conditions of* ***P1.4*** *are* $\rho_m\tilde{h}_0 > \gamma(0)$ *and* $V_2(\tau_2) \geq 0$, *where*

$$\tau_2 = \begin{cases} 0, \textit{ if } V_3(0) \leq 0, \\ \min\{\tau_1, \tau_3\}, \textit{ else}, \end{cases} \tag{4.17}$$

$$V_2(\tau) = \eta\rho_m\tilde{h}_0\frac{1+\tau}{1-\tau} - \frac{\eta\tilde{h}_{\bar{\varsigma}^\circ}+1}{\tilde{h}_{\bar{\varsigma}^\circ}}\gamma(\tau) - \frac{\tilde{h}_{\bar{\varsigma}^\circ}-\tilde{h}_{\varsigma^\circ}}{\tilde{h}_1\tilde{h}_2}\gamma_{\varsigma^\circ}(\tau), \tag{4.18}$$

$$V_3(\tau) = \rho_m - 2^{\frac{2(\tilde{R}_1+\tilde{R}_2)}{1-\tau}}\frac{\eta\tilde{h}_{\bar{\varsigma}^\circ}+1}{\eta\tilde{h}_0\tilde{h}_{\bar{\varsigma}^\circ}}\ln 2^{\tilde{R}_1+\tilde{R}_2} - 2^{\frac{2\tilde{R}_{\varsigma^\circ}}{1-\tau}}\frac{\tilde{h}_{\bar{\varsigma}^\circ}-\tilde{h}_{\varsigma^\circ}}{\eta\tilde{h}_0\tilde{h}_1\tilde{h}_2}\ln 2^{\tilde{R}_{\varsigma^\circ}}, \tag{4.19}$$

and τ_3 *is the unique solution of* $V_3(\tau) = 0$ *in the interval* $(0,1)$ *that can be found using a bisection method [69].*

Proof. **P1.4** is feasible as long as a τ satisfying $\widehat{V}_1(\tau) = \max\left\{V_1(\tau), \frac{\gamma(\tau)}{\tilde{h}_0}\right\} \leq \rho_m$ can be found within the region $[0,\tau_1]$. According to the expression of τ_1 given in (4.12), $\tau_1 > 0$ requires $\rho_m\tilde{h}_0 > \gamma(0)$. In addition, we have $\frac{\gamma(\tau)}{\tilde{h}_0} \leq \rho_m$ when $\tau \leq \tau_1$. Therefore, a valid τ exists when $\min\limits_{0\leq\tau\leq\tau_1} V_1(\tau) \leq \rho_m$. After some simplifications, $\min\limits_{0\leq\tau\leq\tau_1} V_1(\tau) \leq \rho_m$ can be equivalently transformed into $\max\limits_{0\leq\tau\leq\tau_1} V_2(\tau) \geq 0$. Next, the properties of $V_2(\tau)$ will be studied.

The first-order derivative of $V_2(\tau)$ with respect to (w.r.t.) τ is given by $\frac{dV_2(\tau)}{d\tau} = \frac{2\eta\tilde{h}_0}{(1-\tau)^2}V_3(\tau)$. It is easy to tell that $V_3(\tau)$ is a decreasing function of τ within the region $\tau \in [0,1)$ and $\lim\limits_{\tau\to 1^-} V_3(\tau) < 0$. Therefore, $V_2(\tau)$ is a monotonically decreasing function of τ if $V_3(0) \leq 0$; otherwise, $V_2(\tau)$ first increases and then decreases as τ grows. Based on this observation, we know that the maximum value of $V_2(\tau)$ within the region $\tau \in [0,\tau_1]$ is achieved at $\tau = \tau_2$, and thus **P1.4** is feasible when $\rho_m\tilde{h}_0 > \gamma(0)$ and $V_2(\tau_2) \geq 0$. □

When the feasible conditions given in Proposition 5 are satisfied, the optimal solution to **P1.4** can be found via a 1-D search over the region $[0,\tau_1]$. Although the complexity for a 1-D search is acceptable, it is still time-consuming since we need to search the optimal solution in each transmission block. Next we present a more efficient way to find the optimal solution. To proceed, an important proposition is first introduced.

Proposition 6. *The optimal solution to* ***P1.4*** *must be achieved when* $V_1(\tau) \geq \frac{\gamma(\tau)}{\tilde{h}_0}$ *and lies in the region* $\tau \in [\tau_l, \min\{\tau_r, \tau_r'\}]$, *where* $\tau_l = 0$ *if* $V_2(0) \geq 0$, *otherwise* τ_l *is*

the root of $V_2(\tau)=0$ *in the interval* $(0,\tau_3)$, τ_r *is the root of* $V_2(\tau)=0$ *in the interval* $(\tau_l,1)$, *and* $\tau_r'=\tau_1$ *if* $V_1(\tau)\geq\frac{\gamma(\tau)}{\tilde{h}_0}\ \forall\tau\in[0,\tau_1]$, *otherwise* $\tau_r'=\tau_e$ *and* τ_e *is the smallest root of* $V_1(\tau)=\frac{\gamma(\tau)}{\tilde{h}_0}$ *in the interval* $[0,\tau_1]$.

Proof. We first try to find the feasible range of τ according to the constraint $\widehat{V}_1(\tau)\leq\rho_m$. First, $\frac{\gamma(\tau)}{\tilde{h}_0}\leq\rho_m$ requires $\tau\in[0,\tau_1]$. The constraint $V_1(\tau)\leq\rho_m$ can be equivalently transformed into $V_2(\tau)\geq 0$. In the proof process of Proposition 5, we have shown that $V_2(\tau)$ either monotonically decreases or first increases and then decreases. Combining this result with another fact that $\lim_{\tau\to1^-}V_2(\tau)=-\infty$ (this is because $\lim_{\tau\to1^-}\gamma(\tau)(1-\tau)=\infty$), we know that $V_2(\tau)\geq 0$ when $\tau\in[\tau_l,\tau_r]$. To this end, we know that $\widehat{V}_1(\tau)\leq\rho_m$ when $\tau\in[\tau_l,\min\{\tau_r,\tau_1\}]$. Next we will discuss the optimal solution to **P1.4**.

It is easy to tell that the optimal τ minimizing $\widehat{V}_1(\tau)\frac{1+\tau}{2}$ must be in the $[\tau_l,\min\{\tau_r,\tau_1\}]$ interval if $V_1(\tau)\geq\frac{\gamma(\tau)}{\tilde{h}_0}\ \forall\tau\in[0,\tau_1]$; otherwise, the optimal τ must be less than or equal to τ_e due to the following three facts: 1) $V_1(0)>\frac{\gamma(0)}{\tilde{h}_0}$, 2) $\widehat{V}_1(\tau)>\frac{\gamma(\tau_e)}{\tilde{h}_0}$ $\forall\tau\in(\tau_e,\tau_1)$ (this is because $\widehat{V}_1(\tau)=\max\left\{V_1(\tau),\frac{\gamma(\tau)}{\tilde{h}_0}\right\}$ and $\frac{\gamma(\tau)}{\tilde{h}_0}$ is an increasing function of τ), and 3) $\frac{1+\tau}{2}$ is an increasing function of τ. It is worth to notice that τ_e is larger than τ_l. This can be proved by contradiction as follows. If $\tau_e\leq\tau_l$, we have $\frac{\gamma(\tau_e)}{\tilde{h}_0}=V_1(\tau_e)>\rho_m$ and $\tau_e>\tau_1$, which is contradictory to the prerequisite that $\tau_e\leq\tau_1$. Combining all these observations, we know that the optimal solution must lies in $\tau\in[\tau_l,\min\{\tau_r,\tau_r'\}]$, in which we have $V_1(\tau)\geq\frac{\gamma(\tau)}{\tilde{h}_0}$. □

Based on Proposition 6, the optimal solution to **P1.4** can be presented as follows.

Proposition 7. *The optimal solution to* ***P1.4*** *is* $\tau^*=\tau_2$ *if* $V_2(\tau_2)=0$. *When* $V_2(\tau_2)>0$, *the optimal solution is given by*

$$\tau^*=\begin{cases}0, \textit{if } V_4(0)\geq 0 \textit{ and } V_1(0)\leq\rho_m,\\ >0\begin{cases}\tau_m, \textit{if } V_4(0)<0 \textit{ and } \frac{\gamma(\tau_m)}{\tilde{h}_0}\leq V_1(\tau_m)\leq\rho_m,\\ \tau_r', \textit{if } V_1(\tau_m)\leq\rho_m \textit{ and } V_1(\tau_m)<\frac{\gamma(\tau_m)}{\tilde{h}_0},\\ \tau_l, \textit{if } V_1(\tau_m)>\rho_m \textit{ and } \tau_m<\tau_l,\\ \min\{\tau_r,\tau_1\}, \textit{if } V_1(\tau_m)>\rho_m, \tau_m>\tau_r \textit{ and } V_1(\tau_m)\geq\frac{\gamma(\tau_m)}{\tilde{h}_0},\\ \min\{\tau_r,\tau_r'\}, \textit{if } V_1(\tau_m)>\rho_m, \tau_m>\tau_r \textit{ and } V_1(\tau_m)<\frac{\gamma(\tau_m)}{\tilde{h}_0},\end{cases}\end{cases}\tag{4.20}$$

where τ_m *is the unique root of* $V_4(\tau) = 0$ *lying in* $(0, 1)$ *and*

$$V_4(\tau) = \frac{\eta\tilde{h}_{\bar{\varsigma}^\circ} + 1}{2\eta\tilde{h}_0\tilde{h}_{\bar{\varsigma}^\circ}}\left(2^{\frac{2(R_1+R_2)}{1-\tau}}\frac{\ln 4^{R_1+R_2}}{1-\tau} - \gamma(\tau)\right) + \frac{\tilde{h}_{\bar{\varsigma}^\circ} - \tilde{h}_{\varsigma^\circ}}{2\eta\tilde{h}_0\tilde{h}_1\tilde{h}_2}\left(2^{\frac{2R_{\varsigma^\circ}}{1-\tau}}\frac{\ln 4^{R_{\varsigma^\circ}}}{1-\tau} - \gamma_{\varsigma^\circ}(\tau)\right). \tag{4.21}$$

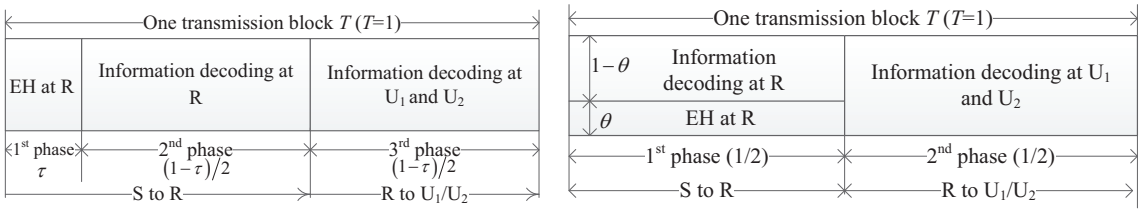
Proof. See Section 4.7.2. □

Remark 7. *It is worth mentioning that the optimal TS ratio* τ^* *given in* (4.20) *is a semi-analytical solution. Although closed-form expressions of* τ_l, τ_r, τ_r', *and* τ_m *are difficult to obtain in general, the numerical results can be efficiently found using the bisection method.*

Remark 8. *Note that* $\tau^* = 0$ *suggests that to realize a successful transmission with the minimum energy consumption at S, it is preferred to reduce the H-SWIPT scheme to the PS scheme. We have shown that* $\tau^* = 0$ *when* $V_4(0) \geq 0$ *and* $V_1(0) \leq \rho_m$. *This may happen when two users have higher targeted data rates and the maximum allowable transmit power is large.*

To this end, we have found the optimal TS ratio τ^*. After deriving τ^*, the optimal transmit power of S is given by $P_{\mathrm{S}}^* = V_1(\tau^*)\sigma^2$. By substituting τ^* and P_{S}^* into (4.11), (4.13), and (4.14), we can obtain $\alpha_{\varsigma^*}^*$, θ^*, ς^* and $\beta_{\varsigma^*}^*$ accordingly.

4.3.3 Optimal Designs for the TS-based and PS-based Transmission Protocols



(a) TS-based transmission. (b) TS-based transmission.

Figure 4.2: TS-based and PS-based cooperative NOMA transmission protocols.

The TS-based cooperative NOMA (TCN) and PS-based cooperative NOMA (PCN) transmission protocols are illustrated in Fig. 4.2(a) and Fig. 4.2(b), respectively. Specifically, the optimal design for the TCN protocol involves the setting of the power allocation ratios at S and R, the TS ratio, the user ordering, and the transmit power

of S. The optimal design for the PCN protocol involves the setting of the power allocation ratios at S and R, the PS ratio, the user ordering, and the transmit power of S.

It is worth mentioning that the optimal power allocation ratio at S for the HSCN protocol (given in (4.11)) can still be applied in the TCN and PCN protocols to ensure that R can successfully decode both messages. In addition, the optimal PS ratio for the HSCN protocol (given in (4.11)) can be applied in the PCN protocol to maximize the harvested energy at R. Although different SWIPT schemes result different transmit power of R, we have shown in Proposition 4 that the optimal power allocation ratio at R and the optimal user ordering are independent of the transmit power of R. Therefore, the results for the HSCN protocol given in Proposition 4 can also be applied in the TCN and PCN protocols. Based on these observations, for the TCN protocol, we only need to care about the design of the transmit power of S and the TS ratio, and for the PCN protocol, only the transmit power of S need to be optimized.

4.3.3.1 TCN Protocol

By following the analysis in the HSCN protocol, we can obtain the feasible conditions for the TCN protocol realizing a successful transmission as $\rho_m \tilde{h}_0 > \gamma(0)$ and $\bar{V}_2(\bar{\tau}_2) \geq 0$, where $\bar{V}_2(\tau) = \frac{2\tau\eta\rho_m\tilde{h}_0}{1-\tau} - \frac{\gamma(\tau)}{\tilde{h}_{\bar{\varsigma}^\circ}} - \frac{\tilde{h}_{\bar{\varsigma}^\circ}-\tilde{h}_{\varsigma^\circ}}{\tilde{h}_1\tilde{h}_2}\gamma_{\varsigma^\circ}(\tau)$,

$$\bar{\tau}_2 = \begin{cases} 0, \text{if } \bar{V}_3(0) \leq 0, \\ \min\{\tau_1, \bar{\tau}_3\}, \text{else}, \end{cases} \tag{4.22}$$

$$\bar{V}_3(\tau) = \rho_m - 2^{\frac{2(\tilde{R}_1+\tilde{R}_2)}{1-\tau}} \frac{\ln 2^{\tilde{R}_1+\tilde{R}_2}}{\eta\tilde{h}_0\tilde{h}_{\bar{\varsigma}^\circ}} - 2^{\frac{2\tilde{R}_{\varsigma^\circ}}{1-\tau}} \frac{\tilde{h}_{\bar{\varsigma}^\circ} - \tilde{h}_{\varsigma^\circ}}{\eta\tilde{h}_0\tilde{h}_1\tilde{h}_2} \ln 2^{\tilde{R}_{\varsigma^\circ}}, \tag{4.23}$$

and $\bar{\tau}_3$ is the unique solution of $\bar{V}_3(\tau) = 0$ in the interval $(0, 1)$ that can be found using a bisection method.

For a given τ, the optimal transmit SNR of S is $\max\left\{\bar{V}_1(\tau), \frac{\gamma(\tau)}{\tilde{h}_0}\right\}$ if $\max\left\{\bar{V}_1(\tau), \frac{\gamma(\tau)}{\tilde{h}_0}\right\} \leq \rho_m$, where $\bar{V}_1(\tau) = \left(\frac{\gamma(\tau)}{\eta\tilde{h}_0\tilde{h}_{\bar{\varsigma}^\circ}} + \gamma_{\varsigma^\circ}(\tau)\frac{\tilde{h}_{\bar{\varsigma}^\circ}-\tilde{h}_{\varsigma^\circ}}{\eta\tilde{h}_0\tilde{h}_1\tilde{h}_2}\right)\frac{1-\tau}{2\tau}$; otherwise, the given τ cannot satisfy the feasible conditions. After this, with a 1-D search over the region $\tau \in (0, \tau_1]$, we can obtain the optimal TS ratio of the TCN protocol minimizing $P_S\frac{1+\tau}{2}$.

4.3.3.2 PCN Protocol

Following the analysis for the HSCN protocol, the feasible condition for the PCN protocol can be easily derived as $V_1(0) \leq \rho_m$. When this condition is satisfied, the optimal transmit power of S is $V_1(0)\sigma^2$, where $V_1(0)$ has been given in (4.15).

4.3.4 System Outage Probability of the Optimal HSCN, TCN and PCN Protocols

For all these three protocols, outage events happen when the feasible condition(s) cannot be satisfied. Since the feasible conditions for the HSCN and TCN protocols are semi-analytical, it is difficult to get closed-form expressions of the system outage probability for those two protocols. However, the system outage probability of the PCN protocol can be derived because its feasible condition has been explicitly given.

Corollary 3. *When full CSIT is available, the system outage probability of the optimal PCN protocol is given by*

$$P_1^P \approx 1-e^{-\frac{\gamma(0)}{\rho_m\Omega_0}}\frac{\pi}{\Omega_1\Omega_2}\sum_{j=1}^{2}\sum_{n_j=0}^{N_j}\frac{\sqrt{1-\delta_{n_j}^2}}{N_j}\frac{\Phi_1\left(\gamma_j(0),\delta_{n_j}\right)}{\Phi_2\left(\Omega_j,\Omega_{\bar{j}},\delta_{n_j}\right)} \times K_2\left(2\sqrt{\Phi_1\left(\gamma_j(0),\delta_{n_j}\right)\Phi_2\left(\Omega_j,\Omega_{\bar{j}},\delta_{n_j}\right)}\right), \tag{4.24}$$

where $\Phi_1(x,y) \triangleq \frac{\gamma(0)-x}{\eta\rho_m\Omega_0}+\frac{2x}{\eta\rho_m\Omega_0(y+1)}$, $\Phi_2(x,y,z) \triangleq \frac{z+1}{2x}+\frac{1}{y}$, $\delta_{n_j} \triangleq \cos\left(\frac{(2n_j-1)\pi}{2N_j}\right)$, $(j,\bar{j}) \in \{(1,2),(2,1)\}$ *and* N_j *is a complexity-accuracy tradeoff parameter.*

Proof. See Section 4.7.3. □

4.4 Optimal Transmission Design with Partial CSIT

In this section, we consider a more practical scenario where only the statistical CSI is available at the transmitters. We first formulate the optimization problem for the HSCN protocol, and then derive the joint optimal solution when the TS ratio is given, based on which, a closed-form expression of the system outage probability is obtained. With the derived closed-form expression of the system outage probability, the optimal TS ratio can be found via a 1-D search. In addition, the optimal designs for the TCN and PCN protocols in the partial CSIT case are analyzed to provide benchmarks.

4.4.1 Problem Formulation

Since only statistical CSI is available at the transmitters, our target in this case is to minimize the system outage probability. It is intuitive that the larger the transmit power of S, the smaller the system outage probability. Hence, the transmit power of S is set to be P_m, and our remaining task is to jointly optimize the power allocation ratios at S and R, the TS and PS ratios, and the user ordering. In fact, the minimization of the system outage probability is equivalent to maximizing the probability of successful transmission. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ denote the events that successful transmission from S to R and R to $\mathrm{U}_1/\mathrm{U}_2$, respectively. Accordingly, the optimization problem in the partial CSIT case can be formulated as

$$\mathbf{P2}: \max_{\alpha_\varsigma,\theta,\varsigma,\beta_\varsigma,\tau} \Pr\left\{\mathcal{S}_1\mathcal{S}_2\right\} \tag{4.25a}$$

$$\text{s.t. } \varsigma \in \{1,2\}, 0 \le \tau < 1, 0 \le \theta < 1, \frac{\gamma_\varsigma(\tau)}{1+\gamma_\varsigma(\tau)} < \alpha_\varsigma < 1, \frac{\gamma_\varsigma(\tau)}{1+\gamma_\varsigma(\tau)} < \beta_\varsigma < 1, \tag{4.25b}$$

Particularly, the probability of $\mathcal{S}_1$ and $\mathcal{S}_2$ are respectively given by

$$\begin{aligned}\Pr\left\{\mathcal{S}_1\right\} &= \Pr\left\{R_{\mathrm{R},x_\varsigma} \ge \tilde{R}_\varsigma, R_{\mathrm{R},x_{\bar{\varsigma}}} \ge \tilde{R}_{\bar{\varsigma}}, P_{\mathrm{S}} = P_m\right\} \\ &\overset{\text{(i)}}{=} \Pr\left\{\theta \le 1 - \frac{1}{\rho_m \tilde{h}_0} \max\left\{\frac{\gamma_\varsigma(\tau)}{\alpha_\varsigma - \alpha_{\bar{\varsigma}}\gamma_\varsigma(\tau)}, \frac{\gamma_{\bar{\varsigma}}(\tau)}{\alpha_{\bar{\varsigma}}}\right\}\right\},\end{aligned} \tag{4.26}$$

and

$$\begin{aligned}\Pr\left\{\mathcal{S}_2\right\} &= \Pr\left\{R_{\mathrm{U}_\varsigma,x_\varsigma} \ge \tilde{R}_\varsigma, R_{\mathrm{U}_{\bar{\varsigma}},x_\varsigma} \ge \tilde{R}_\varsigma, R_{\mathrm{U}_{\bar{\varsigma}},x_{\bar{\varsigma}}} \ge \tilde{R}_{\bar{\varsigma}}, P_{\mathrm{S}} = P_m\right\} \\ &\overset{\text{(ii)}}{=} \Pr\left\{\eta\rho_m|h_0|^2\left(\frac{2\tau}{1-\tau}+\theta\right) - \varphi_\varsigma(\beta_\varsigma) \ge 0\right\},\end{aligned} \tag{4.27}$$

where step (i) involves the results given in (4.3)-(4.4), step (ii) involves the results given in (4.6)-(4.8), $\varphi_\varsigma(\beta_\varsigma) \triangleq \max\left\{\frac{\vartheta_\varsigma}{\tilde{h}_\varsigma}, \frac{\psi_\varsigma}{\tilde{h}_{\bar{\varsigma}}}\right\}$, $\vartheta_\varsigma \triangleq \frac{\gamma_\varsigma(\tau)}{\beta_\varsigma(1+\gamma_\varsigma(\tau))-\gamma_\varsigma(\tau)}$, and $\psi_\varsigma \triangleq \max\left\{\vartheta_\varsigma, \frac{\gamma_{\bar{\varsigma}}(\tau)}{1-\beta_\varsigma}\right\}$. We next seek to find the joint optimal solution to **P2**, which is denoted by $\left(\alpha_{\varsigma^\star}^\star, \theta^\star, \varsigma^\star, \beta_{\varsigma^\star}^\star, \tau^\star\right)$.

4.4.2 Joint Optimal Solution of the HSCN Protocol

Similar to the full CSIT case, we also first determine the optimal power allocation ratio at S and the optimal PS ratio when the remaining variables are given, which are respectively denoted by $\alpha_\varsigma^\bullet$ and $\theta^\bullet$. In **P1.1**, we have discussed how to maximize

the transmit power of R while guaranteeing the successful transmission from S to R, and presented the joint optimal solution $\left(\alpha_{\varsigma}^{\circ}, \theta^{\circ}\right)$ in (4.11). We can find that when P_{S}, τ and ς are fixed, $\alpha_{\varsigma}^{\circ}$ is also fixed and the setting of θ° only requires R to have the instantaneous CSI of the link S→R. This observation indicates that $\alpha_{\varsigma}^{\circ}$ and θ° are still applicable in the partial CSIT case, thus we have

$$\alpha_{\varsigma}^{\bullet} = \alpha_{\varsigma}^{\circ} = 1 - \frac{\gamma_{\bar{\varsigma}}(\tau)}{\gamma(\tau)}, \theta^{\bullet} = \theta^{\circ}|_{\rho_{\mathrm{S}}=\rho_m} = 1 - \frac{\gamma(\tau)}{\rho_m \tilde{h}_0}. \tag{4.28}$$

Substituting (4.28) into (4.26) and (4.27), the probability of $\mathcal{S}_1$ and $\mathcal{S}_2$ can be respectively re-expressed as $\Pr\{\mathcal{S}_1\} = \Pr\left\{\rho_m \tilde{h}_0 \geq \gamma(\tau)\right\}$ and $\Pr\{\mathcal{S}_2\} = \Pr\{\rho_{\mathrm{R}}^{\star} \geq \varphi_{\varsigma}(\beta_{\varsigma})\}$, where $\rho_{\mathrm{R}}^{\star} = \eta \rho_m \tilde{h}_0 \frac{1+\tau}{1-\tau} - \eta\gamma(\tau)$. Then, **P2** can be equivalently transformed as

$$\mathbf{P2.1}: \max_{\varsigma,\tau,\beta_{\varsigma}} \Pr\{\mathcal{S}_1\mathcal{S}_2\} = \Pr\left\{\rho_m \tilde{h}_0 \geq \gamma(\tau), \rho_{\mathrm{R}}^{\star} \geq \varphi_{\varsigma}(\beta_{\varsigma})\right\} \tag{4.29a}$$

$$\text{s.t. } \varsigma \in \{1,2\}, 0 \leq \tau < 1, \frac{\gamma_{\varsigma}(\tau)}{1+\gamma_{\varsigma}(\tau)} < \beta_{\varsigma} < 1, \tag{4.29b}$$

To solve **P2.1**, we first try to find the optimal user ordering and the optimal power allocation ratio at R for a given τ, which are denoted by $\varsigma^{\bullet}$ and $\beta_{\varsigma^{\bullet}}^{\bullet}$, respectively.

Proposition 8. *For a given τ, the optimal user ordering and the optimal power allocation ratio at R minimizing the system outage probability are respectively given by*

$$(\varsigma^{\bullet}, \bar{\varsigma}^{\bullet}) = \begin{cases} (1,2), \textit{if} \sqrt{\frac{\Omega_1}{\Omega_2}} \leq \Psi, \\ (2,1), \textit{else}, \end{cases} \tag{4.30}$$

and

$$\beta_{\varsigma^{\bullet}}^{\bullet} = \frac{\sqrt{\Omega_{\bar{\varsigma}^{\bullet}}} + \xi_{\varsigma^{\bullet}}\sqrt{\Omega_{\varsigma^{\bullet}}}}{\zeta_{\varsigma^{\bullet}}\sqrt{\Omega_{\varsigma^{\bullet}}} + \sqrt{\Omega_{\bar{\varsigma}^{\bullet}}}}, \tag{4.31}$$

where $\Psi = \frac{-\lambda_2 + \sqrt{\lambda_2^2 + 4\lambda_1^2}}{2\lambda_1}$, $\lambda_1 = \gamma_2(\tau)\xi_1(\zeta_2 - \xi_2)$, $\lambda_2 = 2(\xi_2 - \xi_1)$, $\xi_j \triangleq \sqrt{\frac{\gamma_1(\tau)\gamma_2(\tau)}{1+\gamma_j(\tau)}}$, $\zeta_j \triangleq \sqrt{\frac{\gamma_{\bar{j}}(\tau)(1+\gamma_j(\tau))}{\gamma_j(\tau)}}$, *and* $(j, \bar{j}) \in \{(1,2),(2,1)\}$.

Proof. See Section 4.7.4. □

Remark 9. *Proposition 8 suggests that in the partial CSIT case, the optimal user ordering minimizing the system outage probability is jointly decided by the average channel gains and the decoding thresholds of U_1 and U_2. Furthermore, we can find that*

the optimal user ordering and the optimal power allocation factor at R are independent of the transmit power of R. Therefore, our proposed user ordering and power allocation strategies can be directly applied in any two-user downlink NOMA networks that aim at minimizing the system outage probability in a Rayleigh fading environment with partial CSIT.

Remark 10. *Based on* (4.30)*, the optimal user ordering for the following two special cases are studied to provide more valuable insights.*

1) $\Omega_1 = \Omega_2$*, two users have an identical average channel gain: It is easy to check that* $\lambda_2 \leq 0, \Psi \geq 1$ *when* $\tilde{R}_1 \leq \tilde{R}_2$ *and* $\lambda_2 > 0, \Psi < 1$ *when* $\tilde{R}_1 > \tilde{R}_2$*. Therefore, when* $\Omega_1 = \Omega_2$*, the optimal user ordering is* $(1, 2)$ *if* $\tilde{R}_1 \leq \tilde{R}_2$*; otherwise, the optimal user ordering is* $(2, 1)$*. This user ordering criterion suggests that when* $\Omega_1 = \Omega_2$*, user with a smaller decoding threshold should be treated as a weaker user. This is understandable since this operation is helpful for performing SIC. Note that the distance-based user ordering criterion adopted in [40, 47, 48, 51, 70] is invalid when* $\Omega_1 = \Omega_2$*. Hence, our proposed user ordering criterion can be more generally applied in partial CSIT scenarios.*

2) $\Omega_1 < (>) \, \Omega_2$ *and* $\tilde{R}_1 \leq (\geq) \, \tilde{R}_2$*, the average channel gain and targeted data rate of* U_1 *are smaller (larger) than that of* U_2*: When* $\Omega_1 < \Omega_2$ *and* $\tilde{R}_1 \leq \tilde{R}_2$*, we have* $\sqrt{\frac{\Omega_1}{\Omega_2}} < 1$ *and* $\Psi \geq 1$*, thus the optimal user ordering is* $(1, 2)$*. Similarly, we can show that the optimal user ordering is* $(2, 1)$ *when* $\Omega_1 > \Omega_2$ *and* $\tilde{R}_1 \geq \tilde{R}_2$*. Therefore, in these two cases, our proposed user ordering criterion follows the distance-based user ordering criterion adopted in [40, 47, 48, 51, 70].*

With the derived $\alpha_\varsigma^\bullet$, $\theta^\bullet$, $\varsigma^\bullet$ and $\beta_{\varsigma^\bullet}^\bullet$, the system outage probability of the optimal HSCN protocol for a given TS ratio τ can be derived in a closed-form.

Corollary 4. *The system outage probability of the optimal HSCN protocol for a given* τ*, denoted by* $P_2^H(\tau)$*, is given by*

$$P_2^H(\tau) = 1 - \Upsilon_1 + \Upsilon_2, \tag{4.32}$$

where

$$\Upsilon_1 = e^{-\frac{(1-\tau)\eta\gamma(\tau)}{\eta\rho_m\Omega_0(1+\tau)}}\sqrt{\frac{4\upsilon_{\varsigma^\bullet}(1-\tau)}{\eta\rho_m\Omega_0(1+\tau)}}K_1\left(\sqrt{\frac{4\upsilon_{\varsigma^\bullet}(1-\tau)}{\eta\rho_m\Omega_0(1+\tau)}}\right), \tag{4.33}$$

$$\Upsilon_2 \approx \frac{1-\tau}{\eta\rho_m\Omega_0(1+\tau)}e^{-\frac{(1-\tau)\eta\gamma(\tau)}{\eta\rho_m\Omega_0(1+\tau)}}\frac{\pi\varpi}{2N_3}\sum_{n_3=1}^{N_3}\sqrt{1-\delta_{n_3}^2}e^{-\frac{2\upsilon_{\varsigma^\bullet}}{(\delta_{n_3}+1)\varpi}}e^{-\frac{(1-\tau)(\delta_{n_3}+1)\varpi}{2\eta\rho_m\Omega_0(1+\tau)}}, \tag{4.34}$$

$$\upsilon_j \triangleq \frac{\zeta_j\sqrt{\Omega_j}+\sqrt{\Omega_{\bar{j}}}}{\sqrt{\Omega_1\Omega_2}}\left(\frac{\gamma_j(\tau)}{\sqrt{\Omega_j}}+\frac{\gamma_{\bar{j}}(\tau)}{\sqrt{\Omega_{\bar{j}}}(\zeta_j-\xi_j)}\right), (j,\bar{j}) \in \{(1,2),(2,1)\}, \tag{4.35}$$

$\varpi = \frac{\eta\gamma(\tau)2\tau}{1-\tau}$, $\delta_{n_3} \triangleq \cos\left(\frac{(2n_3-1)\pi}{2N_3}\right)$ *and* N_3 *is a complexity-accuracy tradeoff parameter.*

Proof. See Section 4.7.5. □

After deriving $P_2^{\mathrm{H}}(\tau)$, the optimal TS ratio $\tau^\star$ can be obtained via a 1-D search over the region $\tau \in [0,1)$. Subsequently, $\alpha_{\varsigma^\star}^\star$, $\theta^\star$, $\varsigma^\star$ and $\beta_{\varsigma^\star}^\star$ can be obtained by substituting $\tau^\star$ into (4.28), (4.30) and (4.31).

Remark 11. *Compared to the full CSIT case, 1-D search in the partial CSIT case is more acceptable since the search only needs to be performed once as long as the statistical CSI of the network remains unchanged.*

4.4.3 Optimal Designs for the TCN and PCN Protocols

4.4.3.1 TCN Protocol

For a given TS ratio τ, the derived optimal user ordering and the optimal power allocation ratios for the HSCN protocol can be directly applied in the TCN protocol. Therefore, we here only need to find the optimal TS ratio for the TCN protocol. Following the analysis in the HSCN protocol, the system outage probability of the optimal TCN protocol for a given τ is given by

$$P_2^{\mathrm{T}}(\tau) = 1-\Upsilon_1'+\Upsilon_2', 0<\tau<1, \tag{4.36}$$

where $\Upsilon_1' = \sqrt{\frac{2\upsilon_{\varsigma^\bullet}(1-\tau)}{\eta\rho_m\Omega_0\tau}}K_1\left(\sqrt{\frac{2\upsilon_{\varsigma^\bullet}(1-\tau)}{\eta\rho_m\Omega_0\tau}}\right)$,

$$\Upsilon_2' \approx \frac{1-\tau}{2\eta\rho_m\Omega_0\tau}\frac{\pi\varpi}{2N_4}\sum_{n_4=1}^{N_4}\sqrt{1-\delta_{n_4}^2}e^{-\frac{2\upsilon_{\varsigma^\bullet}}{(\delta_{n_4}+1)\varpi}}e^{-\frac{(1-\tau)(\delta_{n_4}+1)\varpi}{4\eta\rho_m\Omega_0\tau}}, \tag{4.37}$$

$\delta_{n_4} \triangleq \cos\left(\frac{(2n_4-1)\pi}{2N_4}\right)$ and N_4 is a complexity-accuracy tradeoff parameter.

Based on (4.36), the optimal TS ratio for the TCN protocol can be found via a 1-D search over the region $\tau \in (0, 1)$. Subsequently, by substituting the derived TS ratio into (4.28), (4.30) and (4.31), we can derive all the optimal parameters for the TCN protocol.

4.4.3.2 PCN Protocol

The joint optimal solution of the PCN protocol can be easily derived by substituting $\tau = 0$ into (4.28), (4.30) and (4.31). Furthermore, the system outage probability of the optimal PCN protocol is given by $P_2^{\mathrm{P}} = P_2^{\mathrm{H}}(0)$.

4.5 Numerical Results

In this section, we provide simulation results to validate our analytical model. We assume that the average channel gain is characterized by a path-loss model $\Omega_i = \varrho d_i^{-\omega}$ [29], $i = 0, 1, 2$, where ϱ refers to the signal attenuation corresponding to a distance of 1 meter, d_0 is the distance from S to R, and d_j is the distance from R to U_j, $j = 1, 2$. Without otherwise specified, the system parameters are set as follows: the path loss exponent is $\omega = 2.7$, signal attenuation is $\varrho = 10^{-3}$, energy conversion efficiency is $\eta = 0.5$, noise power is $\sigma^2 = -100$dBm, distance from S to R is $d_0 = 5$m, distance from R to U_j is $d_j = 10$m, and two users have the same targeted data rate $\tilde{R}_1 = \tilde{R}_2 = \tilde{R}_t$.

4.5.1 Results with Full CSIT

We start with Fig. 4.3 to depict the system outage probability of different optimal transmission protocols with full CSIT. It can be easily seen that the HSCN protocol outperforms the PCN and TCN protocols. In addition, we can observe that the analytical results for the PCN protocol given in (4.24) well match with the simulation results. Comparing the TCN and PCN protocols, we can find that the TCN protocol is better than the PCN protocol when $\tilde{R}_t = 0.1$bps/Hz. However, when $\tilde{R}_t$ is higher, the PCN protocol shows its superiority over the TCN protocol, and it even has the same system outage performance as the HSCN protocol when $\tilde{R}_t = 1$bps/Hz. This observation confirms our conclusion in Remark 8 that the H-SWIPT scheme reduces to the PS scheme when the two users have higher targeted data rates.

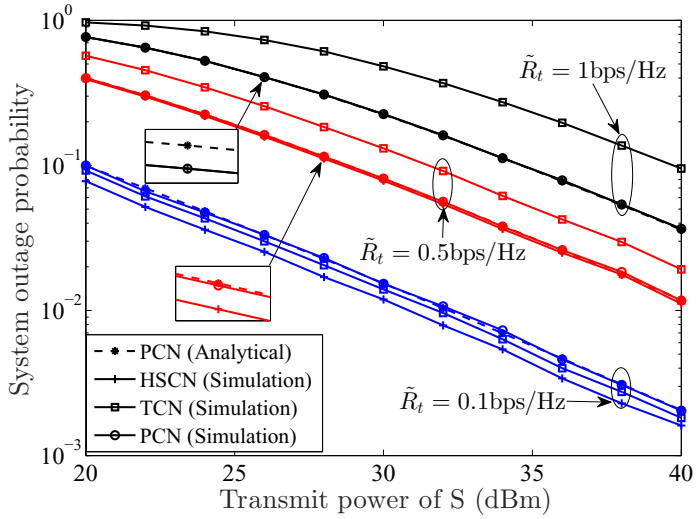


Figure 4.3: System outage probability of different optimal transmission protocols with full CSIT.

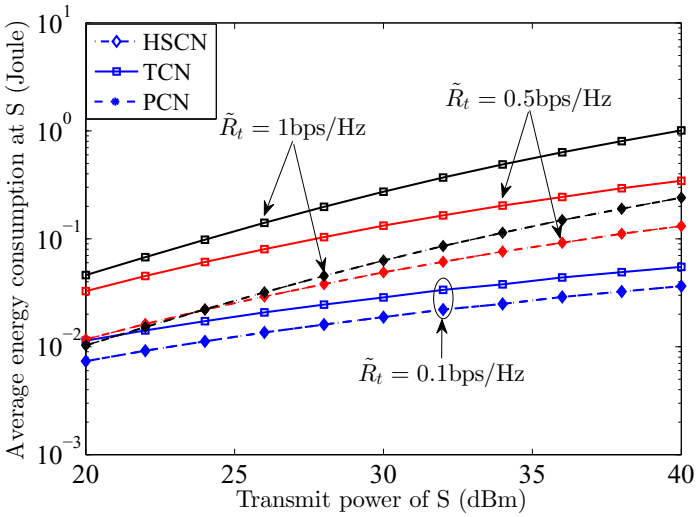


Figure 4.4: Average energy consumption of different optimal transmission protocols.

In Fig. 4.4, we compares the average energy consumption of different optimal transmission protocols. To make a fair comparison, we count the energy consumed of each protocol when all three protocols can achieve successful transmission. We can see that the energy consumption of the HSCN and PCN protocols are identical and less than that of the TCN protocol. This observation suggests that when all these protocols promise to realize a successful transmission, it is preferable to choose the HSCN or the PCN protocol as they are more energy efficient.

4.5.2 Results with Partial CSIT

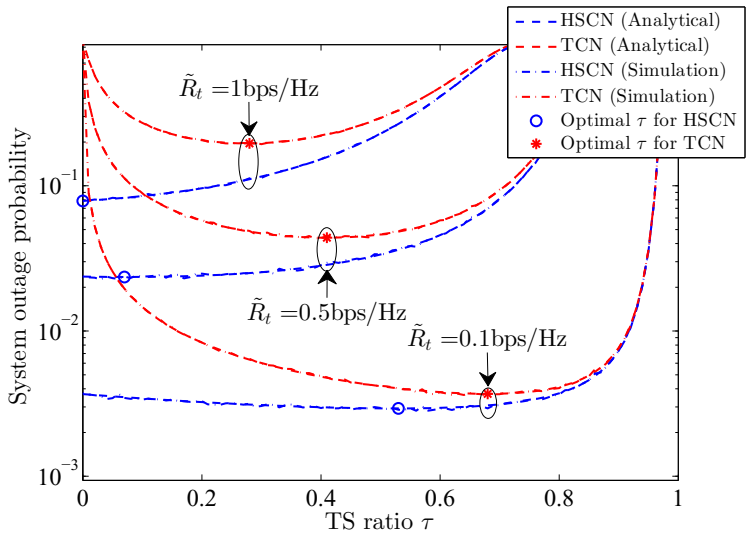


Figure 4.5: System outage probability of the HSCN and TCN protocols with different TS ratios ($P_{\mathrm{S}} = 40$dBm).

Fig. 4.5 checks the analytical results ((4.32) and (4.36)) for the optimal HSCN and TCN protocols in the partial CSIT case. It can be seen that the analytical results have a good match with the simulation results. We can also notice that the optimal TS ratios for both protocols decrease as the targeted data rate $\tilde{R}_t$ increases. This phenomenon can be explained from the energy-saving perspective. It is intuitive that the decrease of $\tilde{R}_t$ makes successful transmission easier to realize and therefore consumes less energy. Since the total energy consumption in each fading block can be expressed as $P_{\mathrm{S}}\frac{1+\tau}{2}$ and the transmit power of S in the partial CSIT case is fixed as P_m, less energy consumption means smaller TS ratio. Furthermore, it can be observed that the optimal TS ratio for the HSCN protocol is equal to zero when $\tilde{R}_t = 1$bps/Hz, which indicates that the H-SWIPT scheme reduces to the PS scheme in this case.

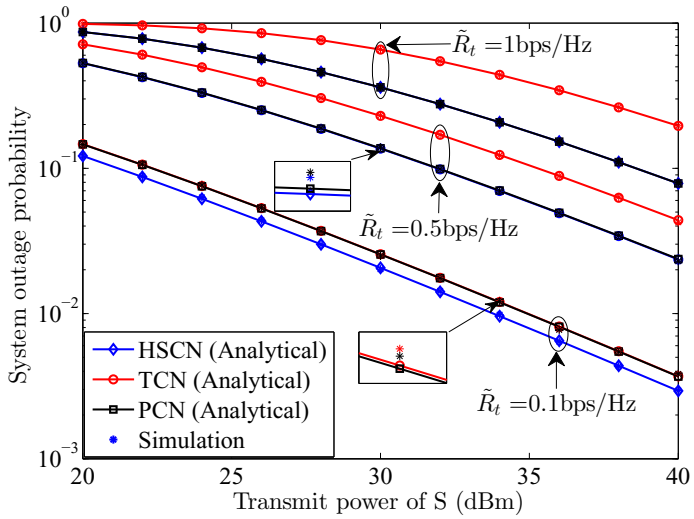


Figure 4.6: System outage probability of different optimal transmission protocols with partial CSIT.

Fig. 4.6 plots the system outage probability of different optimal transmission protocols with partial CSIT. It is obvious that the optimal HSCN protocol always has a smaller system outage probability than the other two protocols. We can also find that the PCN protocol is superior to the TCN protocol, and it even has the same system outage probability as the HSCN protocol when $\tilde{R}_t = 1$bps/Hz. This is expected since in Fig. 4.5 we have shown that the optimal HSCN protocol reduces to the PCN protocol when $\tilde{R}_t = 1$bps/Hz. In addition, we can see that the performance gap between the PCN and TCN protocols gradually reduces as the targeted data rates decrease, which also can be observed in the full CSIT case.

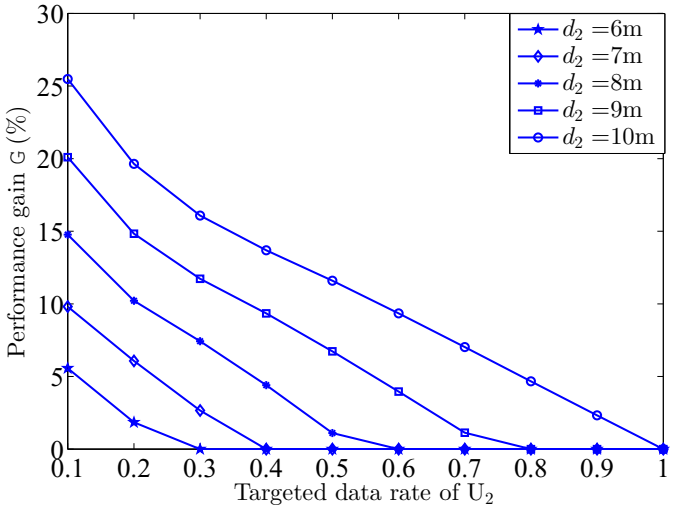


Figure 4.7: Performance comparison of the HSCN protocol under two different user ordering criteria.

To show the importance of user ordering, Fig. 4.7 compares the performance of the HSCN protocol under our derived optimal user ordering and the conventional user ordering. Specifically, the conventional user ordering is widely adopted in current works in which users are ordered based on their distances to the relay or the source [40, 47,48,51,70]. Given the conventional user ordering, other optimal system parameters can be easily obtained by following our analysis given in Section 4.4.2. Let P^{con} denote the system outage probability of the HSCN protocol with the conventional user ordering. Then, the performance gain can be defined as $\mathsf{G}\,(\%) = 100 \times \left(\frac{P^{\text{con}} - P_2^{\text{H}}(\tau^\star)}{P^{\text{con}}}\right)$. We have shown in Section 4.4.2 that when $\Omega_1 < (>)\,\Omega_2$ and $\tilde{R}_1 \leq (\geq)\,\tilde{R}_2$, our proposed user ordering criterion is identical to the conventional one, and thus there would be no performance gap. In view of this, here we only study the case $d_2 \leq d_1 = 10\text{m}$ (i.e., $\Omega_1 \leq \Omega_2$) and $\tilde{R}_2 < \tilde{R}_1 = 1\text{bps/Hz}$. Based on this setting, the conventional user ordering selects U_1 as the weak user and lets U_2 perform SIC. From the figure, we can see that for a given d_2, the performance gain gradually reduces to 0 as $\tilde{R}_2$ increases, and for a given $\tilde{R}_2$, the performance gain non-decreases as d_2 increases. This observation indicates that for a given d_2, our proposed user ordering is different from the conventional one when $\tilde{R}_1 - \tilde{R}_2$ exceeds a certain threshold, and this threshold gradually decreases as d_2 increases. We can also notice that the largest performance gain appears when $d_1 = d_2 = 10\text{m}$, $\tilde{R}_1 = 1\text{bps/Hz}$, and $\tilde{R}_2 = 0.1\text{bps/Hz}$. Those observations remind us that in the partial CSIT case, a prudent choice for the user ordering is important when the two users have similar distances to the relay and a larger difference in targeted data rates.

4.6 Summary

In this chapter, we have introduced the more powerful and more general SWIPT scheme, i.e., the H-SWIPT scheme, into relay-assisted NOMA networks and provided two novel resource allocation designs. In particular, for the full CSIT case, we aim at minimizing the energy consumption while realizing the required transmissions by jointly optimizing the power allocation ratios at S and R, the TS and PS ratios, the user ordering, and the transmit power of S. Although the resulting optimization problem is non-convex, we successfully find the joint optimal solution via the low-

complexity bisection method. On the other hand, for the case with only partial CSIT, we intend to minimize the system outage probability by jointly optimizing the power allocation ratios at S and R, the TS and PS ratios, and the user ordering. We successfully solve this problem with a 1-D search. Our analytical results show that the derived optimal user ordering in the full CSIT case follows the user ordering strategy widely adopted in current works, while the derived optimal user ordering in the partial CSIT case is different from the one applied in current works and can be more generally applied.

4.7 Appendix

4.7.1 Proof of Proposition 4

To find the optimal user ordering and the optimal power allocation factor at R, we first find the optimal power allocation factors minimizing $\varphi_\varsigma(\beta_\varsigma) = \max\left\{\frac{\vartheta_\varsigma}{\tilde{h}_\varsigma}, \frac{\psi_\varsigma}{\tilde{h}_{\bar{\varsigma}}}\right\}$ under different user ordering ($\varsigma = 1, 2$), then the joint optimal solution can be decided by comparing the minimum values of $\varphi_1(\beta_1)$ and $\varphi_2(\beta_2)$. It is easy to check that when $\varsigma = 1$, $\vartheta_1 = \frac{\gamma_1(\tau)}{\beta_1(1+\gamma_1(\tau))-\gamma_1(\tau)}$ is a decreasing function of β_1 for $\beta_1 \in \left(\frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}, 1\right)$, and $\psi_1 = \max\left\{\vartheta_1, \frac{\gamma_2(\tau)}{1-\beta_1}\right\}$ is equal to ϑ_1 when $\beta_1 \in \left(\frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}, 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}\right)$. When $\beta_1 \in \left(1-\frac{\gamma_2(\tau)}{\gamma(\tau)}, 1\right)$, ψ_1 increases with respect to (w.r.t.) β_1. After elaborating the properties of ϑ_1 and ψ_1, the minimum value of $\varphi_1(\beta_1)$ can be decided according to the relationship between $\tilde{h}_1$ and $\tilde{h}_2$.

1) $\tilde{h}_1 > \tilde{h}_2$: In this case, it is easy to show that $\frac{\psi_1}{\tilde{h}_2}$ is always larger than $\frac{\vartheta_1}{\tilde{h}_1}$. Therefore, the minimum value of $\varphi_1(\beta_1)$ is equal to the minimum value of $\frac{\psi_1}{\tilde{h}_2}$, which is achieved at $\beta_1 = 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}$ and the minimum value is $\frac{\gamma(\tau)}{\tilde{h}_2}$.

2) $\tilde{h}_1 \leq \tilde{h}_2$: In this case, $\frac{\psi_1}{\tilde{h}_2} = \frac{\vartheta_1}{\tilde{h}_2}$ is smaller than $\frac{\vartheta_1}{\tilde{h}_1}$ when $\beta_1 \in \left(\frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}, 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}\right)$. Since ψ_1 and ϑ_1 respectively increases and decreases w.r.t. β_1 when $\beta_1 \in \left[1-\frac{\gamma_2(\tau)}{\gamma(\tau)}, 1\right)$, and $\lim_{\beta_1 \to 1^-} \psi_1 = \infty > \lim_{\beta_1 \to 1^-} \vartheta_1 = \gamma_1(\tau)$, we know that $\varphi_1(\beta_1)$ attains its minimum when $\frac{\vartheta_1}{\tilde{h}_1} = \frac{\gamma_2(\tau)}{\tilde{h}_2(1-\beta_1)}$. After some simplifications, the minimum value $\varphi_1(\beta_1)$ is given by $\frac{\gamma(\tau)}{\tilde{h}_2} + \frac{\tilde{h}_2-\tilde{h}_1}{\tilde{h}_1\tilde{h}_2}\gamma_1(\tau)$, which can be achieved at $\beta_1 = \frac{\tilde{h}_1\gamma_1(\tau)\gamma_2(\tau)+\tilde{h}_2\gamma_1(\tau)}{\tilde{h}_1\gamma(\tau)+\left(\tilde{h}_2-\tilde{h}_1\right)\gamma_1(\tau)}$.

To this end, the minimum value of $\varphi_1(\beta_1)$ can be concluded as

$$\min\{\varphi_1(\beta_1)\} = \begin{cases} \frac{\gamma(\tau)}{\tilde{h}_2}, \text{if } \tilde{h}_1 > \tilde{h}_2, \\ \frac{\gamma(\tau)}{\tilde{h}_2} + \frac{\tilde{h}_2-\tilde{h}_1}{\tilde{h}_1\tilde{h}_2}\gamma_1(\tau), \text{else.} \end{cases} \tag{4.38}$$

Following the similar procedures in the derivation of $\min\{\varphi_1(\beta_1)\}$, we can also find the minimum value of $\varphi_2(\beta_2)$, which is given by

$$\min\{\varphi_2(\beta_2)\} = \begin{cases} \frac{\gamma(\tau)}{\tilde{h}_1}, \text{if } \tilde{h}_1 \leq \tilde{h}_2, \\ \frac{\gamma(\tau)}{\tilde{h}_1} + \frac{\tilde{h}_1 - \tilde{h}_2}{\tilde{h}_1 \tilde{h}_2}\gamma_2(\tau), \text{else.} \end{cases} \tag{4.39}$$

Based on (4.38) and (4.39), we can easily check that $\min\{\varphi_1(\beta_1)\} \leq \min\{\varphi_2(\beta_2)\}$ if $\tilde{h}_1 \leq \tilde{h}_2$; otherwise, $\min\{\varphi_1(\beta_1)\} > \min\{\varphi_2(\beta_2)\}$. Based on this observation, the joint optimal user ordering and power allocation factor at R minimizing $\varphi_\varsigma(\beta_\varsigma)$ can be obtained.

4.7.2 Proof of Proposition 7

As we have stated in Proposition 5, **P1.4** is feasible when $V_2(\tau_2) \geq 0$. In view of this, we discuss the optimal solution to **P1.4** based on the value of $V_2(\tau_2)$. First, $V_2(\tau_2) = 0$ indicates that τ_2 is the only one feasible point. Therefore, the optimal solution is $\tau^* = \tau_2$ in this case. Next, we study the optimal solution when $V_2(\tau_2) > 0$. Based on Proposition 6, the optimization problem in **P1.4** can be re-formulated as $\min\limits_{\tau_l \leq \tau \leq \min\{\tau_r, \tau_r'\}} V_1(\tau)\frac{1+\tau}{2}$. Let τ_m denote the optimal τ minimizing $V_1(\tau)\frac{1+\tau}{2}$ within the region $[0, 1)$. The first-order derivative of $V_1(\tau)\frac{1+\tau}{2}$ w.r.t. τ is $V_4(\tau)$. By checking the first-order derivative of $V_4(\tau)$ w.r.t. τ, we can find that $V_4(\tau)$ is an increasing function of τ. Combining this result with another fact that $\lim\limits_{\tau \to 1^-} V_4(\tau) = \infty$, we can tell that $V_1(\tau)\frac{1+\tau}{2}$ is an increasing function of τ and $\tau_m = 0$ if $V_4(0) \geq 0$; otherwise, $V_1(\tau)\frac{1+\tau}{2}$ first decreases and then increases, and its minimum is achieved at $\tau_m > 0$. Depending on the value of τ_m, the optimal τ minimizing $V_1(\tau)\frac{1+\tau}{2}$ within the region $[\tau_l, \min\{\tau_r, \tau_r'\}]$ can be decided in the following six cases.

1) $V_4(0) \geq 0$ and $V_1(0) \leq \rho_m$: $V_4(0) \geq 0$ indicates that $\tau_m = 0$, and $V_1(0) \leq \rho_m$ indicates that $\tau_l = 0$. Therefore, we know that $\tau_m \in [\tau_l, \min\{\tau_r, \tau_r'\}]$ and the optimal τ is $\tau^* = \tau_m = 0$.

2) $V_4(0) < 0$ and $\frac{\gamma(\tau_m)}{\tilde{h}_0} \leq V_1(\tau_m) \leq \rho_m$: $V_4(0) < 0$ suggests that $\tau_m > 0$, and $\frac{\gamma(\tau_m)}{\tilde{h}_0} \leq V_1(\tau_m) \leq \rho_m$ indicates that $\tau_m \in [\tau_l, \min\{\tau_r, \tau_r'\}]$. Therefore, the optimal τ is $\tau^* = \tau_m > 0$.

3) $V_1(\tau_m) \leq \rho_m$ and $V_1(\tau_m) < \frac{\gamma(\tau_m)}{\tilde{h}_0}$: $V_1(\tau_m) \leq \rho_m$ indicates that $\tau_m \in [\tau_l, \tau_r]$, and $V_1(\tau_m) < \frac{\gamma(\tau_m)}{\tilde{h}_0}$ suggests that $\tau_m > \tau_r'$. Therefore, we have $\tau_r' < \tau_m \leq \tau_r$. Since $V_1(\tau)\frac{1+\tau}{2}$ is a decreasing function w.r.t. $\tau \in [0, \tau_m]$, we can tell that the optimal τ

minimizing $V_1(\tau)\frac{1+\tau}{2}$ within the region $[\tau_l, \min\{\tau_r, \tau_r'\}]$ is τ_r' and $V_1(\tau) - \frac{\gamma(\tau)}{\tilde{h}_0}$ is a decreasing function w.r.t. $\tau \in [0, \tau_m]$. Specifically, a bisection method can be applied to find the root of $V_1(\tau) = \frac{\gamma(\tau)}{\tilde{h}_0}$ within the region $[0, \tau_m]$, i.e., τ_e. If $\tau_e \geq \tau_1$, then $\tau_r' = \tau_1$; otherwise, $\tau_r' = \tau_e$.

4) $V_1(\tau_m) > \rho_m$ and $\tau_m < \tau_l$: Since $V_1(\tau)\frac{1+\tau}{2}$ is an increasing function w.r.t. $\tau \in [\tau_m, 1)$ and $\tau_m < \tau_l$, we know that the optimal τ minimizing $V_1(\tau)\frac{1+\tau}{2}$ within the region $[\tau_l, \min\{\tau_r, \tau_r'\}]$ is τ_l.

5) $V_1(\tau_m) > \rho_m$, $\tau_m > \tau_r$ and $V_1(\tau_m) \geq \frac{\gamma(\tau_m)}{\tilde{h}_0}$: Since $V_1(\tau) - \frac{\gamma(\tau)}{\tilde{h}_0}$ is a decreasing function w.r.t. $\tau \in [0, \tau_m]$ and $V_1(\tau_m) \geq \frac{\gamma(\tau_m)}{\tilde{h}_0}$, we have $V_1(\tau) \geq \frac{\gamma(\tau)}{\tilde{h}_0}$ w.r.t. $\tau \in [0, \tau_m]$. Combining this result with another condition that $\tau_m > \tau_r$, we know that $\tau_r' > \tau_r$ if $\tau_1 > \tau_r$; otherwise, $\tau_r > \tau_r' = \tau_1$. Therefore, the optimal τ minimizing $V_1(\tau)\frac{1+\tau}{2}$ within the region $[\tau_l, \min\{\tau_r, \tau_r'\}]$ is $\min\{\tau_r, \tau_1\}$ since $V_1(\tau)\frac{1+\tau}{2}$ is a decreasing function w.r.t. $\tau \in [0, \tau_m]$.

6) $V_1(\tau_m) > \rho_m$, $\tau_m > \tau_r$ and $V_1(\tau_m) < \frac{\gamma(\tau_m)}{\tilde{h}_0}$: $V_1(\tau_m) < \frac{\gamma(\tau_m)}{\tilde{h}_0}$ indicates that $\tau_m > \tau_r'$. Therefore, we have $\tau_m > \max\{\tau_r, \tau_r'\}$. In this case, the optimal τ minimizing $V_1(\tau)\frac{1+\tau}{2}$ within the region $[\tau_l, \min\{\tau_r, \tau_r'\}]$ is $\min\{\tau_r, \tau_r'\}$ since $V_1(\tau)\frac{1+\tau}{2}$ is a decreasing function w.r.t. $\tau \in [0, \tau_m]$. Specifically, $\tau_r' = \tau_1$ if $\tau_e \geq \tau_1$; otherwise, $\tau_r' = \tau_e$, where τ_e is the root of $V_1(\tau) = \frac{\gamma(\tau)}{\tilde{h}_0}$ lying in the region $[0, \tau_m]$ that can be found with a bisection method.

4.7.3 Proof of Corollary 3

According to the feasible condition given in Section 4.3.3.2, the system outage probability of the optimal PCN protocol can be evaluated as

$$P_1^{\mathrm{P}} = 1 - \underbrace{\Pr\{V_1(0) \leq \rho_m \,|\varsigma^\circ = 1\}}_{I_1} - \underbrace{\Pr\{V_1(0) \leq \rho_m \,|\varsigma^\circ = 2\}}_{I_2}. \tag{4.40}$$

According to the expression of $V_1(\tau)$ given in (4.15), I_1 can be re-expressed as

$$\begin{aligned} I_1 &= \Pr\left\{\tilde{h}_0 \geq \frac{\gamma(0)}{\rho_m} + \frac{\gamma(0)-\gamma_1(0)}{\eta\rho_m\tilde{h}_2} + \frac{\gamma_1(0)}{\eta\rho_m\tilde{h}_1}, \tilde{h}_1 < \tilde{h}_2\right\} \\ &= e^{-\frac{\gamma(0)}{\rho_m\Omega_0}} \int_0^\infty e^{-\frac{\gamma(0)-\gamma_1(0)}{\eta\rho_m\Omega_0 y}} f_{\tilde{h}_2}(y) \underbrace{\int_0^y e^{-\frac{\gamma_1(0)}{\eta\rho_m\Omega_0 x}} f_{\tilde{h}_1}(x)dx}_{I_{11}} dy, \end{aligned} \tag{4.41}$$

where the second step is processed by applying the CCDF of $\tilde{h}_0$, $\bar{F}_{\tilde{h}_i}(x) = e^{-\frac{x}{\Omega_i}}$, $i = 0, 1, 2$. To proceed, we need to calculate I_{11}. It is very difficult to get an exact closed-

form expression of I_{11}. However, by resorting to Gaussian-Chebyshev quadrature [36, Eq. 25.4.38], an approximation of I_{11} can be obtained as follows

$$\begin{aligned} I_{11} &= \int_0^y \frac{1}{\Omega_1} e^{-\frac{\gamma_1(0)}{\eta\rho_m\Omega_0 x}} e^{-\frac{x}{\Omega_1}} dx = \frac{y}{2\Omega_1}\int_{-1}^{1} e^{-\frac{2\gamma_1(0)}{\eta\rho_m\Omega_0(t+1)y}} e^{-\frac{(t+1)y}{2\Omega_1}} dt \\ &\approx \frac{\pi y}{2N_1\Omega_1}\sum_{n_1=0}^{N_1}\sqrt{1-\delta_{n_1}^2}\, e^{-\frac{2\gamma_1(0)}{\eta\rho_m\Omega_0(\delta_{n_1}+1)y}} e^{-\frac{(\delta_{n_1}+1)y}{2\Omega_1}}, \end{aligned} \tag{4.42}$$

where the second step is processed with variable substitution $x = \frac{t+1}{2}y$. Now substituting (4.42) into (4.41) and then applying [52, Eq. 3.471.9] to solve the integral regarding y, I_1 can be approximated as

$$\begin{aligned} I_1 \approx & e^{-\frac{\gamma(0)}{\rho_m\Omega_0}} \frac{\pi}{\Omega_1\Omega_2} \sum_{n_1=0}^{N_1} \frac{\sqrt{1-\delta_{n_1}^2}}{N_1} \frac{\Phi_1\left(\gamma_1(0), \delta_{n_1}\right)}{\Phi_2\left(\Omega_1, \Omega_2, \delta_{n_1}\right)} \\ & \times K_2\left(2\sqrt{\Phi_1\left(\gamma_1(0), \delta_{n_1}\right)\Phi_2\left(\Omega_1, \Omega_2, \delta_{n_1}\right)}\right). \end{aligned} \tag{4.43}$$

An approximation of I_2 also can be obtained by following the similar steps in the derivation of I_1. Substituting the approximations of I_1 and I_2 into (4.40), we can derive the approximation of P_1^{P} as shown in (4.24).

4.7.4 Proof of Proposition 8

In **P2.1**, the successful transmission probability $\Pr\{\mathcal{S}_1\mathcal{S}_2\}$ can be calculated as

$$\begin{aligned} \Pr\{\mathcal{S}_1\mathcal{S}_2\} &= \Pr\left\{\rho_m\tilde{h}_0 \geq \gamma(\tau), \rho_{\mathrm{R}}^\star \geq \max\left\{\frac{\vartheta_\varsigma}{\tilde{h}_\varsigma}, \frac{\psi_\varsigma}{\tilde{h}_{\bar{\varsigma}}}\right\}\right\} \\ &= \Pr\left\{\tilde{h}_0 \geq \frac{\gamma(\tau)}{\rho_m}, \rho_{\mathrm{R}}^\star > 0, \tilde{h}_\varsigma \geq \frac{\vartheta_\varsigma}{\rho_{\mathrm{R}}^\star}, \tilde{h}_{\bar{\varsigma}} \geq \frac{\psi_\varsigma}{\rho_{\mathrm{R}}^\star}\right\} \\ &= \Pr\left\{\rho_{\mathrm{R}}^\star > \varpi, \tilde{h}_\varsigma \geq \frac{\vartheta_\varsigma}{\rho_{\mathrm{R}}^\star}, \tilde{h}_{\bar{\varsigma}} \geq \frac{\psi_\varsigma}{\rho_{\mathrm{R}}^\star}\right\} \\ &= \int_{\varpi}^{\infty} \bar{F}_{\tilde{h}_\varsigma}\left(\frac{\vartheta_\varsigma}{x}\right) \bar{F}_{\tilde{h}_{\bar{\varsigma}}}\left(\frac{\psi_\varsigma}{x}\right) f_{\rho_{\mathrm{R}}^\star}(x)\, dx \\ &= \int_{\varpi}^{\infty} e^{-\frac{\phi_\varsigma(\beta_\varsigma)}{x}} f_{\rho_{\mathrm{R}}^\star}(x)\, dx, \end{aligned} \tag{4.44}$$

where $\varpi = \frac{\eta\gamma(\tau)2\tau}{1-\tau}$, $\phi_\varsigma(\beta_\varsigma) = \frac{\vartheta_\varsigma}{\Omega_\varsigma} + \frac{\psi_\varsigma}{\Omega_{\bar{\varsigma}}}$, and (4.44) is obtained by applying the CCDF of $\tilde{h}_\varsigma$ and $\tilde{h}_{\bar{\varsigma}}$.

From (4.44), it can be observed that the values of ς and β_ς only affect $\phi_\varsigma(\beta_\varsigma)$. Combining this observation with another fact that $e^{-\frac{\phi_\varsigma(\beta_\varsigma)}{x}}$ is a decreasing function of

$\phi_\varsigma(\beta_\varsigma)$, we know that the optimal ς and β_ς maximizing the successful transmission probability can be found by minimizing $\phi_\varsigma(\beta_\varsigma)$. To proceed, we first seek to find the optimal power allocation ratio at R for different user ordering, then the joint optimal power allocation ratio at R and user ordering can be decided by making a comparison. When $\varsigma = 1$, $\phi_\varsigma(\beta_\varsigma) = \phi_1(\beta_1)$ can be expressed as

$$\begin{aligned}\phi_1(\beta_1) &= \frac{1}{\Omega_1}\frac{\gamma_1(\tau)}{\beta_1(1+\gamma_1(\tau))-\gamma_1(\tau)} + \frac{1}{\Omega_2}\max\left\{\frac{\gamma_1(\tau)}{\beta_1(1+\gamma_1(\tau))-\gamma_1(\tau)}, \frac{\gamma_2(\tau)}{1-\beta_1}\right\} \\ &= \begin{cases}\phi_{11}(\beta_1), \frac{\gamma_1(\tau)}{1+\gamma_1(\tau)} < \beta_1 \leq 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}, \\ \phi_{12}(\beta_1), 1-\frac{\gamma_2(\tau)}{\gamma(\tau)} < \beta_1 < 1,\end{cases}\end{aligned} \tag{4.45}$$

where

$$\phi_{11}(\beta_1) = \frac{\gamma_1(\tau)}{\beta_1(1+\gamma_1(\tau))-\gamma_1(\tau)}\left(\frac{1}{\Omega_1}+\frac{1}{\Omega_2}\right), \tag{4.46}$$

$$\phi_{12}(\beta_1) = \frac{1}{\Omega_1}\frac{\gamma_1(\tau)}{\beta_1(1+\gamma_1(\tau))-\gamma_1(\tau)} + \frac{1}{\Omega_2}\frac{\gamma_2(\tau)}{1-\beta_1}. \tag{4.47}$$

It is easy to tell that $\phi_{11}(\beta_1)$ and $\phi_{12}(\beta_1)$ are decreasing and convex functions of β_1, respectively. In addition, since $\phi_{12}(\beta_1) \to \infty$ when $\beta_1 \to \frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}$ or $\beta_1 \to 1$, we know that the critical point of $\phi_{12}(\beta_1)$, i.e., the root of $\frac{d\phi_{12}(\beta_1)}{d\beta_1} = 0$, must lies in the region $\left(\frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}, 1\right)$. After some algebraic manipulations, we can obtain the critical point of $\phi_{12}(\beta_1)$ as $\beta_1^\diamond = \frac{\sqrt{\Omega_2}+\xi_1\sqrt{\Omega_1}}{\zeta_1\sqrt{\Omega_1}+\sqrt{\Omega_2}}$. Combining all these observations, we know that if $\beta_1^\diamond < 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}$, i.e., $\chi = \sqrt{\frac{\Omega_1}{\Omega_2}} > \zeta_1$, $\phi_1(\beta_1)$ decreases and increases w.r.t. $\beta_1 \in \left(\frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}, 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}\right)$ and $\beta_1 \in \left(1-\frac{\gamma_2(\tau)}{\gamma(\tau)}, 1\right)$, respectively, and its minimum is achieved at $\beta_1 = 1-\frac{\gamma_2(\tau)}{\gamma(\tau)}$; otherwise, $\phi_1(\beta_1)$ decreases and increases w.r.t. $\beta_1 \in \left(\frac{\gamma_1(\tau)}{1+\gamma_1(\tau)}, \beta_1^\diamond\right)$ and $\beta_1 \in (\beta_1^\diamond, 1)$, respectively, and achieves its minimum at $\beta_1 = \beta_1^\diamond$. After obtaining the optimal β_1, the minimum value of $\phi_1(\beta_1)$, denoted as $\phi_1^{\min}$, can be calculated as

$$\phi_1^{\min} = \begin{cases}\upsilon_1, \text{if } \chi \leq \zeta_1, \\ \gamma(\tau)\left(\frac{1}{\Omega_1}+\frac{1}{\Omega_2}\right), \text{else},\end{cases} \tag{4.48}$$

where υ_1 has been given in (4.35). Following the similar procedures, we can also find the minimum value of $\phi_2(\beta_2)$, denoted as $\phi_2^{\min}$, which is given by

$$\phi_2^{\min} = \begin{cases}\upsilon_2, \text{if } \chi > (\zeta_2)^{-1}, \\ \gamma(\tau)\left(\frac{1}{\Omega_1}+\frac{1}{\Omega_2}\right), \text{else},\end{cases} \tag{4.49}$$

where υ_2 has been given in (4.35).

Now the optimal user ordering can be decided by comparing $\phi_1^{\min}$ and $\phi_2^{\min}$. Specifically, the relationship between $\phi_1^{\min}$ and $\phi_2^{\min}$ can be decided in the following three cases.

1) $\chi \leq (\zeta_2)^{-1} < \zeta_1$*:* In this case, it is easy to check that $\phi_1^{\min} = \upsilon_1 < \gamma(\tau)\left(\frac{1}{\Omega_1}+\frac{1}{\Omega_2}\right) = \phi_2^{\min}$.

2) $\chi \geq \zeta_1 > (\zeta_2)^{-1}$*:* In this case, we have $\phi_2^{\min} = \upsilon_2 < \gamma(\tau)\left(\frac{1}{\Omega_1}+\frac{1}{\Omega_2}\right) = \phi_1^{\min}$.

3) $(\zeta_2)^{-1} < \chi < \zeta_1$*:* In this case, $\phi_1^{\min} = \upsilon_1$ and $\phi_2^{\min} = \upsilon_2$. After a series simplifications, $\upsilon_1 - \upsilon_2$ can be re-expressed as $\upsilon_1 - \upsilon_2 = \frac{\lambda_1\chi^2+\lambda_2\chi-\lambda_1}{\Omega_1\Omega_2(\zeta_2-\xi_2)(\zeta_1-\xi_1)}$, where λ_1 and λ_2 are given in Proposition 8. Since $\zeta_j > \xi_j$ $(j = 1, 2)$, the sign of $\upsilon_1 - \upsilon_2$ can be analyzed by discussing $\lambda_1\chi^2 + \lambda_2\chi - \lambda_1$, which can be viewed as a function of χ. As $\lambda_1 > 0$, there is only one positive root, denoted as Ψ, that makes $\lambda_1\chi^2 + \lambda_2\chi - \lambda_1 = 0$. By resorting to the quadratic formula, Ψ is derived as $\frac{-\lambda_2+\sqrt{\lambda_2^2+4\lambda_1^2}}{2\lambda_1}$. To this end, we know that $\upsilon_1 - \upsilon_2$ is negative and positive in the cases of $(\zeta_2)^{-1} < \chi < \Psi$ and $\Psi \leq \chi < \zeta_1$, respectively.

Combining the results for all these three cases, we can conclude that $\phi_1^{\min}$ is smaller than and larger than $\phi_2^{\min}$ in the cases of $\chi \leq \Psi$ and $\chi > \Psi$, respectively. Based on this result, the optimal power allocation ratio at R and the optimal user ordering can be decided.

4.7.5 Proof of Corollary 4

With the derived $\varsigma^\bullet$ and $\beta_{\varsigma^\bullet}^\bullet$, the minimum value of $\phi_\varsigma(\beta_\varsigma)$ is $\upsilon_{\varsigma^\bullet}$ as given in (4.35). Substituting this result into (4.44), the system outage probability of the optimal HSCN protocol can be calculated as

$$\begin{aligned} P_2^{\mathrm{H}}(\tau) &= 1 - \Pr\{\mathcal{S}_1\mathcal{S}_2\} = 1 - \int_{\varpi}^{\infty} e^{-\frac{\upsilon_{\varsigma^\bullet}}{x}} f_{\rho_{\mathrm{R}}^\star}(x)\, dx \\ &= 1 - \underbrace{\int_0^{\infty} e^{-\frac{\upsilon_{\varsigma^\bullet}}{x}} f_{\rho_{\mathrm{R}}^\star}(x)\, dx}_{\Upsilon_1} - \underbrace{\int_0^{\varpi} e^{-\frac{\upsilon_{\varsigma^\bullet}}{x}} f_{\rho_{\mathrm{R}}^\star}(x)\, dx}_{\Upsilon_2}. \end{aligned} \tag{4.50}$$

To proceed, we need the PDF of $f_{\rho_{\mathrm{R}}^\star}(x)$. As $\rho_{\mathrm{R}}^\star = \eta\rho_m\tilde{h}_0\frac{1+\tau}{1-\tau} - \eta\gamma(\tau)$, the PDF of

$f_{\rho^{\star}_{\mathrm{R}}}(x)$ can be calculated in the following way

$$\begin{aligned} f_{\rho^{\star}_{\mathrm{R}}}(x) &= \frac{dF_{\rho^{\star}_{\mathrm{R}}}(x)}{dx} = \frac{d}{dx} F_{\tilde{h}_0}\left(\frac{(x+\eta\gamma(\tau))(1-\tau)}{\eta\rho_m(1+\tau)}\right) \\ &= \frac{1-\tau}{\eta\rho_m(1+\tau)} f_{\tilde{h}_0}\left(\frac{(x+\eta\gamma(\tau))(1-\tau)}{\eta\rho_m(1+\tau)}\right) \\ &= \frac{1-\tau}{\eta\rho_m\Omega_0(1+\tau)} e^{-\frac{x(1-\tau)}{\eta\rho_m\Omega_0(1+\tau)}} e^{-\frac{(1-\tau)\eta\gamma(\tau)}{\eta\rho_m\Omega_0(1+\tau)}}. \end{aligned} \tag{4.51}$$

By applying (4.51) into (4.50) and resorting [52, Eq. 3.471.9] to solve the integral, Υ_1 can be derived as (4.33). It is difficult to obtain an exact closed-form expression of Υ_2, however, we can resort to the Gaussian-Chebyshev quadrature [36, Eq. 25.4.38] to approximate Υ_2. Following the similar steps in the derivation of I_{11} given in (4.42), we can obtain an approximation of Υ_2 as given in (4.34). Combining the derived results for Υ_1 and Υ_2, the closed-form expression of $P_2^{\mathrm{H}}(\tau)$ can be derived as shown in (4.32).

Chapter 5

Fair Resource Allocation in PB-Assisted NOMA Networks

In this chapter, we study a wireless powered network consisting of one PB, one EH-based source, and multiple users. To improve spectral efficiency of the network as well as for practical implementation consideration, two users are paired to perform NOMA transmission. Specifically, the NOMA-based transmission protocol consists of two phases, where the first phase is dedicated to power transfer from the PB to the source, and the second phase is dedicated to information transfer from the source to a pair of users. We derive exact and asymptotic closed-form expressions of the average throughput for each paired user. Then, the joint optimization problem for the time and power allocation is investigated to find the optimal fair performance of the paired users. To provide a benchmark, the optimal resource allocation with an OMA-based transmission protocol is also derived. Finally, simulation results confirm the validity of analytical derivations and show that the considered network with the NOMA transmission is superior to that with the OMA transmission, especially when the transmit power of the PB is low and the paired users have significant differences in channel gain.

5.1 Introduction

In the previous chapter, we have introduced a SWIPT-assisted NOMA network. It should be noticed that SWIPT systems are more suitable for short-distance wireless power transfer due to the fact that the information decoder and the energy harvester have a large gap of the operational sensitivity [71]. To solve this concern, we can adopt

the dedicated energy transmitter, termed as PB, to extend the coverage of wireless power transfer. As we mentioned earlier, PB-assisted network is one of the three network architectures in practical wireless powered communication networks. PB has low cost and is applicable to large-scale deployment since it does not need backhaul links or complex computation to process data. Therefore, PB-assisted network has raised wide attention [72–79]. However, all of these works focus on the single-user scenarios or the OMA-based multi-user scenarios.

To our best knowledge, there are very limited efforts on PB-assisted NOMA networks except [80, 81]. The work in [80] studies a PB-assisted two-user downlink NOMA network and aims to minimize the PB's transmit power by jointly optimizing the beamformer at the PB and the EH-based source. By taking the circuit consumption at the EH-based users into account, the work in [81] studies a PB-assisted multiple-user uplink network and compares the spectral efficiency of NOMA and OMA.

In this paper, we consider a PB-assisted multiple-user downlink network consisting of one PB, one EH-based source, and multiple information receivers (users). Specifically, a NOMA-based transmission (NBT) protocol is proposed, in which the source first harvests energy from the PB, and then transmits superimposed information to two selected users, called paired users.[1] We intend to maximize the worse throughput of two paired users by allocating time and power resources based on statistical CSI. To achieve this, we jointly optimize the time allocation ratio for EH and power allocation ratio for NOMA transmissions. Since the joint optimization problem is nonconvex, we decompose it into two subproblems, i.e., to get optimal power allocation for a given time allocation ratio and to get optimal time allocation for a given power allocation ratio, respectively. We derive exact closed-form expressions of the throughput of the two paired users, and based on the derivations, we show that the two subproblems are two individual generalized-convex problems, and the optimal solutions can be obtained using a bisection method and a golden-section search method, respectively. Furthermore, based on the asymptotic closed-form expressions of the throughput, we present some further insights regarding how the system parameters affect the two

[1] The reason for selecting only two users is two-fold: First, it is not realistic to let all the users be served simultaneously since NOMA network is still an interference-limited system [19, 82]. Second, since users in NOMA networks have to decode other users' information to perform SIC, the implementation complexity at the transmitter and receiver sides increases as the number of users increases.

optimal solutions. After solving the two subproblems, an alternating optimization algorithm is designed to approximate the joint optimal solution. To provide a rigorous benchmark, we also design an orthogonal multiple access (OMA)-based transmission (OBT) protocol and try to find its optimal performance by jointly optimizing two time allocation ratios. Exact closed-form expression of the throughput of each paired user with the OBT protocol is provided, based on which, we show that the optimization problem for the OBT protocol is also a generalized-convex problem that can be solved with a golden-section search method. We numerically study the impact of user pairing on the fair performance gain of NBT over OBT. Results show that the NBT protocol outperforms the OBT protocol in terms of the worse throughput of the two paired users, especially when the transmit power of PB is low and the paired users have large differences in channel gain.

The main contributions of this work can be summarized as follows:

- Low complexity: In our work, the joint optimal time allocation and power allocation ratios are derived based on only statistical CSI. Thus, for a group of multiple users, the formulated problem needs to be solved only once. After the problem is solved, we can apply the found time allocation and power allocation ratios in every fading block. Thus, in each fading block, no computation is needed for solving the optimization problem.

- Closed-form expressions of the outage probability of each selected user: Although many works [39, 65, 83, 84] have studied multi-user downlink NOMA networks without EH and presented the closed-form expressions of the outage probability for the ordered user under different fading environments, all of them assume that transmit power of the source/base station is a constant value and channels between the source and users experience independent and identically distributed (i.i.d.) fading. However, in our paper, transmit power of the source in our paper is a random variable that depends on the channel condition of power transfer and other system parameters. Further, we consider a more general scenario in which channels between the source and users experience independent but not necessarily identically distributed (i.n.i.d.) fading. These differences make the derivation of outage probability of the selected users much more chal-

lenging than those in the literature, as can be observed from our derivations in Section 5.3.1.

- Insightful asymptotic analysis: asymptotic expressions of the throughput of the selected users are derived, which reveal that in the NBT protocol, the achievable diversity order for each selected user is identical to its channel gain order, i.e., the user with the n-th worst connection from the source achieves a diversity order of n.

- The OBT protocol and analysis: The OBT protocol, based on OMA, provides a rigorous benchmark for the NBT protocol. Further, when it is not feasible to implement NOMA (for example, due to users' limitation in signal processing capability), our analysis provides a guideline for the optimal system setting.

The rest of this chapter is organized as follows. Section 5.2 introduces the system model and the NBT protocol. In Section 5.3, the average throughput of the NBT protocol for each selected user is analyzed. Section 5.4 and 5.5 investigate the optimization of user fairness for the NBT and OBT protocols, respectively. In Section 5.6, numerical results are provided to verify the analytical results.

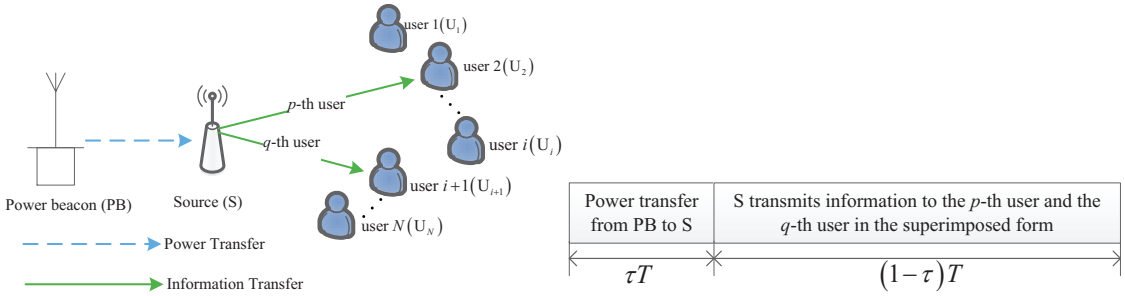
5.2 System Model

As shown in Fig. 5.1(a), an energy-constrained source S, powered by a PB, communicates with N users ($\mathrm{U}_1, \mathrm{U}_2, \cdots, \mathrm{U}_N, N \geq 2$). To expand the coverage of wireless power transfer, the PB is equipped with M antennas to perform energy beamforming, while all the other nodes are equipped with a single antenna. Generally, Rician fading is used to model the wireless power transfer channel to capture the effect of the line-of-sight path [85, 86]. In this work, however, we assume that the channels between the PB and S exhibit Nakagami-m fading which has a tight approximation to the Rician distribution [87]. More importantly, this assumption facilitates mathematical analysis, and hence can provide us more insights [73, 74]. The channels between S and U_i ($1 \leq i \leq N$) experience i.n.i.d. Rayleigh fading with $g_i \sim \mathcal{CN}(0, \Omega_i)$, where g_i denotes the channel coefficient of the link $\mathrm{S} \rightarrow \mathrm{U}_i$. In addition, all the channel coefficients are assumed to be unchanged within one block time T, but vary independently

Table 5.1: Description of symbols used for key parameters in Chapter 5.

Symbol	Meaning
M	Number of antennas at the power beacon
N	Number of users
$\mathbf{h}$	Channel vector between the power beacon and the source
$g_i, 1 \leq i \leq N$	Channel coefficient from the source to user i
$\lvert\tilde{g}_i\rvert^2, 1 \leq i \leq N$	Channel power gain of the i-th weakest user
P_b	Transmit power of the power beacon
P_c	Circuit power consumption at the source
P_s^{N}	Transmit power of the source in the NBT protocol
$x_n, n = p, q, 1 \leq p < q \leq N$	Message for the n-th weakest user
τ	Time allocation ratio for energy harvesting in the NBT protocol
α	Power allocation ratio for the q-th user
$R_{n,x_p}^{\mathrm{N}}, n = p, q$	Achievable rate of the p-th user's message at the n-th user in the NBT protocol
R_{q,x_q}^{N}	Achievable rate of the q-th user's message at the q-th user in the NBT protocol
R_t	Target rate of two paired users in the NBT protocol
$\bar{R}_n^{\mathrm{N}}(\tau, \alpha), n = p, q$	Average throughput of the n-th user in the NBT protocol
$\mathrm{P}_n^{\mathrm{N}}(\tau, \alpha), n = p, q$	Outage probability of the n-th user in the NBT protocol
$\bar{R}_{n,\infty}^{\mathrm{N}}(\tau, \alpha), n = p, q$	Asymptotic throughput of the n-th user in the NBT protocol
$\mathrm{P}_{n,\infty}^{\mathrm{N}}(\tau, \alpha), n = p, q$	Asymptotic outage probability of the n-th user in the NBT protocol
$\bar{R}_{\mathrm{F}}^{\mathrm{N}}(\tau, \alpha)$	Worse throughput of two paired users in the NBT protocol
(τ^*, α^*)	Joint optimal time allocation and power allocation ratios
$\alpha^\star$	Optimal power allocation ratio for a given time allocation ratio
$\tau^\star$	Optimal time allocation ratio for a given power allocation ratio
β_1	Time allocation ratio for energy harvesting in the OBT protocol
β_2	Time allocation ratio during information transfer in the OBT protocol
$\bar{R}_n^{\mathrm{O}}(\beta_1, \beta_2), n = p, q$	Average throughput of the n-th user in the OBT protocol
$\mathrm{P}_n^{\mathrm{O}}(\beta_1), n = p, q$	Outage probability of the n-th user in the OBT protocol
$\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\tau, \alpha)$	Worse throughput of two paired users in the OBT protocol
(β_1^*, β_2^*)	Joint optimal time allocation ratios in the OBT protocol
$\beta_2^\star$	Optimal time allocation ratio of β_2 for a given β_1

from block to block.



(a) System model. (b) NOMA-based transmission protocol.

Figure 5.1: System model and transmission protocol.

Let $|\tilde{g}_i|^2$ $(1 \leq i \leq N)$ denote the i-th weakest user's channel gain, i.e., $|\tilde{g}_1|^2 \leq |\tilde{g}_2|^2 \leq \cdots \leq |\tilde{g}_N|^2$. Hereafter, we use the phrase "the n-th user" to refer to the n-th weakest user. For practical implementation, we assume that two users are paired (selected) from the N sorted users for NOMA transmission [19, 82]. Furthermore, to study the impact of user pairing on the fair performance gain of NOMA over OMA, we here consider a general case in which the p-th and q-th users are paired with $1 \leq p < q \leq N$. Next we introduce the NBT protocol in detail, which contains two phases as shown in Fig. 5.1(b).

5.2.1 Wireless Power Transfer Phase

During the first phase of time period $\tau T\,(0 < \tau < 1)$, the PB transmits an energy signal to S using the maximal ratio transmission technique and S extracts energy from the received signal. Assuming that perfect CSI of the link PB→S is available at the PB, the harvested energy at S in the first phase can be expressed as [73]

$$E_h = \eta P_b \|\mathbf{h}\|^2 \tau T, \tag{5.1}$$

where η $(0 < \eta < 1)$ is the energy conversion efficiency, P_b is the transmit power of the PB, and $\mathbf{h} = [h_1, h_2, \cdots, h_M]$ is the channel vector between the PB and S. In particular, the entries of $\mathbf{h}$ follow i.n.i.d. Nakagami-m fading with fading parameter m and $\mathbb{E}\left[|h_\varsigma|^2\right] = \Omega_0$, $1 \leq \varsigma \leq M$. Letting P_c denote the circuit power consumption at S, the transmit power of S in the second phase can be written as

$$P_s^{\mathrm{N}} = \left[\frac{E_h}{(1-\tau)\,T} - P_c\right]^+ = \left[\frac{\eta P_b \|\mathbf{h}\|^2 \tau}{1-\tau} - P_c\right]^+, \tag{5.2}$$

where the superscript "N" refers to the NBT protocol.

During the second phase (Information Transmission Phase) with duration $(1-\tau)T$, S transmits a superimposed symbol to the paired users with the harvested energy. The received signal at the n-th ($n = p, q$) user is given by

$$y_n^{\mathrm{N}} = \tilde{g}_n \left(\sqrt{\alpha P_s^{\mathrm{N}}} x_q + \sqrt{(1-\alpha) P_s^{\mathrm{N}}} x_p \right) + w_n, \tag{5.3}$$

where x_n is the message for the n-th user, w_n is the AWGN at the n-th user with zero mean and variance σ^2, and $\alpha\,(0 < \alpha < 1)$ and $1-\alpha$ are the PA factors for x_q and x_p, respectively.

According to the SIC principle [15], after receiving the superimposed signal, the n-th user first decodes x_p and treats x_q as interference. As such, the achievable rate of x_p at the n-th user can be expressed as

$$R_{n,x_p}^{\mathrm{N}} = \log_2 \left(1 + \frac{(1-\alpha) P_s^{\mathrm{N}} |\tilde{g}_n|^2}{\alpha P_s^{\mathrm{N}} |\tilde{g}_n|^2 + \sigma^2} \right). \tag{5.4}$$

On the condition that the q-th user successfully decodes x_p, the achievable rate of x_q at the q-th user is given by

$$R_{q,x_q}^{\mathrm{N}} = \log_2 \left(1 + \frac{\alpha P_s^{\mathrm{N}} |\tilde{g}_q|^2}{\sigma^2} \right). \tag{5.5}$$

Based on (5.4) and (5.5), we will analyze the average throughput for each paired user in next section.

5.3 Throughput Analysis of the NBT Protocol

This work focuses on the application scenarios where users have a predetermined targeted data rate and S transmits information at a fixed rate. As such, the average throughput defined based on the outage probability [73, 88] is used as the performance metric. In this section, we first derive exact closed-form expressions of the outage probability and the average throughput for each selected user. Then, simple asymptotic expressions in the high SNR regime are obtained to analyze the diversity order of each selected user and provide some valuable insights.

5.3.1 Exact Analytical Results

For fairness consideration, we assume that the paired users have the same targeted data rate R_t [47, 84, 89]. As such, the average throughput of the n-th ($n = p, q$) user can be expressed as [73, 88]

$$\bar{R}_n^{\mathrm{N}}(\tau, \alpha) = R_t(1-\tau)\left(1 - \mathrm{P}_n^{\mathrm{N}}(\tau, \alpha)\right), \tag{5.6}$$

where $\mathrm{P}_n^{\mathrm{N}}(\tau, \alpha)$ is the outage probability of the n-th user. The outage is defined as the event that the achievable rate of the desired message is less than the targeted data rate. Note that when the circuit power consumption is taken into consideration, the outage event can happen in the two following cases. The first case is that the harvested energy at S cannot support the circuit operation, i.e., $P_s = 0$, which can be referred to as an energy outage. The second case is that the harvested energy is larger than the circuit power consumption, but the achievable rate of the desired message is less than the targeted data rate, which is referred to as a data transmission outage. Now, we are ready for calculating $\mathrm{P}_n^{\mathrm{N}}(\tau, \alpha)$.

Proposition 9. *An exact closed-form expression of the outage probability of the n-th ($n = p, q$) user is given by*

$$P_n^N(\tau, \alpha) = 1 + 2 \sum_{\mu=0}^{Mm-1} \underbrace{\sum_{v=n}^{N} \sum_{S_v} \sum_{\kappa=0}^{v}}_{N-v+\kappa\neq 0} \widetilde{\sum} \frac{(\rho_c)^{Mm-1-\mu}}{\Gamma(Mm)} \binom{Mm-1}{\mu} e^{-\frac{\psi\rho_c(1-\tau)}{\tau}} \times \left(\frac{\psi(1-\tau)}{\tau}\right)^{Mm-\frac{\mu+1}{2}} (\theta_n \Phi)^{\frac{\mu+1}{2}} K_{\mu+1}\left(2\sqrt{\frac{\theta_n \psi \Phi(1-\tau)}{\tau}}\right), \tag{5.7}$$

where $\rho_b \triangleq \frac{P_b}{\sigma^2}$, $\rho_c \triangleq \frac{P_c}{\sigma^2}$, $\psi \triangleq \frac{m}{\eta\rho_b\Omega_0}$, $\theta_p \triangleq \frac{\gamma_{th}}{1-\alpha(1+\gamma_{th})}$, $\theta_q \triangleq \max\{\theta_p, \frac{\gamma_{th}}{\alpha}\}$, $\gamma_{th} \triangleq 2^{R_t} - 1$, *and* $\binom{Mm-1}{\mu} \triangleq \frac{(Mm-1)!}{\mu!(Mm-\mu-1)!}$. *The summation* S_v *in* (5.7) *extends over all permutations* $(j_1, \cdots, j_N)$ *of* $(1, \cdots, N)$ *for which* $j_1 < \cdots < j_v$ *and* $j_{v+1} < \cdots < j_N$, $\widetilde{\sum} \triangleq \frac{(-1)^\kappa}{\kappa!} \underbrace{\sum_{t_1=1}^{v} \cdots \sum_{t_\kappa=1}^{v}}_{t_1\neq t_2\neq\cdots\neq t_\kappa}$, *and* $\Phi \triangleq \sum_{\nu=1}^{\kappa} \frac{1}{\Omega_{j_{t_\nu}}} + \sum_{l=v+1}^{N} \frac{1}{\Omega_{j_l}}$.

Proof. See section 5.8.1. □

It is easy to check that $\mathrm{P}_n^{\mathrm{N}}(\tau, \alpha) = 1$ when $\alpha \in \left[\frac{1}{1+\gamma_{\mathrm{th}}}, 1\right)$. Therefore, the following discussion proceeds under the constraint $0 < \alpha < \frac{1}{1+\gamma_{\mathrm{th}}}$. Substituting (5.7) into (5.6),

an exact closed-form expression of the average throughput for the n-th user can be obtained.

5.3.2 Asymptotic Analytical Results

In order to provide some intuitive insights, we now find an approximation of the average throughput for each selected user in the high SNR regime, i.e., $\rho_b \to \infty$. We first calculate the asymptotic outage probability of the n-th user.

Corollary 5. *As $\rho_b \to \infty$ and $P_b/P_c \to \infty$, the asymptotic outage probability of the n-th user is given by*

$$P_{n,\infty}^{N}(\tau,\alpha) \simeq \left(\frac{\theta_n \phi\,(1-\tau)}{\tau}\right)^n \varphi_n, \tag{5.8}$$

where $\phi \triangleq \frac{m}{\eta\rho_b\Omega_0(Mm-1)}$ and $\varphi_n \triangleq \sum_{S_n}\prod_{l=1}^{n}\frac{1}{\Omega_{j_l}}$.

Proof. See section 5.8.2. □

Remark 12. *According to the definition of diversity order [90], one can tell from* (5.8) *that when the circuit power at S can be ignored compared to the transmit power at the PB, the NBT protocol can help the n-th user to achieve a diversity order of n in the high SNR regime.*

Replacing $\mathrm{P}_n^{\mathrm{N}}(\tau,\alpha)$ in (5.6) with $\mathrm{P}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)$, we can derive an approximated average throughput in the high SNR range, denoted as $\bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)$.

5.4 User Fairness Maximization for the NBT Protocol

This section intends to optimize the user fairness of the NBT protocol, and the fair throughput is employed as the metric for evaluating the fair performance. Specifically, the fair throughput refers to the minimum throughput of the two selected users. In the NBT protocol, the optimal fair throughput can be found by jointly optimizing the power allocation and time allocation ratios, i.e., α and τ, and the optimization problem can be formulated as

$$\mathbf{P3}: \max_{\tau,\alpha} \bar{R}_{\mathrm{F}}^{\mathrm{N}}(\tau,\alpha) \tag{5.9}$$

$$\text{s.t. } 0<\alpha<\frac{1}{1+\gamma_{\mathrm{th}}}, 0<\tau<1, \tag{5.10}$$

where $\bar{R}_{\mathrm{F}}^{\mathrm{N}}(\tau,\alpha) \triangleq \min\limits_{n\in\{p,q\}} \bar{R}_{n}^{\mathrm{N}}(\tau,\alpha)$.

It is difficult to solve problem **P3** with standard optimization solvers since it is nonconvex in general. Thus, we first decouple the joint optimization problem into two subproblems, then individually optimize α and τ, and finally use an alternating optimization algorithm to approximate the joint optimal α and τ, denoted by α^* and τ^*, respectively.

5.4.1 Optimal Power Allocation for a Given Time Allocation Ratio

We first try to find the optimal power allocation factor $\alpha^\star$ for a given time allocation ratio τ, and the optimization problem can be formulated as

$$\mathbf{P4}: \max_{\alpha} \bar{R}_{\mathrm{F}}^{\mathrm{N}}(\tau,\alpha) \tag{5.11}$$

$$\text{s.t. } 0<\alpha<\frac{1}{1+\gamma_{\mathrm{th}}}. \tag{5.12}$$

To tackle problem **P4**, the following lemma is first introduced.

Lemma 4. *$\bar{R}_{p}^{N}(\tau,\alpha)$ is a decreasing function of α for $\alpha \in \left(0, \frac{1}{1+\gamma_{th}}\right)$, and $\bar{R}_{q}^{N}(\tau,\alpha)$ is an increasing function for $\alpha \in \left(0, \frac{1}{\gamma_{th}+2}\right)$ and a decreasing function for $\alpha \in \left(\frac{1}{\gamma_{th}+2}, \frac{1}{1+\gamma_{th}}\right)$, respectively.*

Proof. Since $F_X(x)$ is an increasing function of x, we can tell from (5.31) that the outage probability of the n-th user $\mathrm{P}_{n}^{\mathrm{N}}(\tau,\alpha)$ is an increasing function of θ_n. Along with another two facts that θ_p increases as α grows for $\alpha \in \left(0, \frac{1}{1+\gamma_{\mathrm{th}}}\right)$, and θ_q first decreases as α grows for $\alpha \in \left(0, \frac{1}{\gamma_{\mathrm{th}}+2}\right)$ and then increases for $\alpha \in \left(\frac{1}{\gamma_{\mathrm{th}}+2}, \frac{1}{1+\gamma_{\mathrm{th}}}\right)$, we know that $\mathrm{P}_{p}^{\mathrm{N}}(\tau,\alpha)$ increases as α increases, and $\mathrm{P}_{q}^{\mathrm{N}}(\tau,\alpha)$ first decreases and then increases as α increases. Combining this observation with another fact that $\bar{R}_{n}^{\mathrm{N}}(\tau,\alpha)$ increases as $\mathrm{P}_{n}^{\mathrm{N}}(\tau,\alpha)$ decreases, we can obtain Lemma 4. □

With the help of Lemma 4, we can solve problem **P4** and obtain the optimal power allocation factor $\alpha^\star$.

Proposition 10. *The unique optimal power allocation factor $\alpha^\star$ lies in $\left(0, \frac{1}{\gamma_{th}+2}\right)$ and satisfies $\bar{R}_{p}^{N}(\tau,\alpha^\star) = \bar{R}_{q}^{N}(\tau,\alpha^\star)$.*

Proof. When $\alpha \in \left[\frac{1}{\gamma_{\rm th}+2}, \frac{1}{1+\gamma_{\rm th}}\right)$, $\theta_p = \theta_q$ holds, thus, we know that $\mathrm{P}_p^{\rm N}(\tau,\alpha) > \mathrm{P}_q^{\rm N}(\tau,\alpha)$ and $\bar{R}_p^{\rm N}(\tau,\alpha) < \bar{R}_q^{\rm N}(\tau,\alpha)$ over $\alpha \in \left[\frac{1}{\gamma_{\rm th}+2}, \frac{1}{1+\gamma_{\rm th}}\right)$ due to the fact that $|\tilde{g}_p|^2 < |\tilde{g}_q|^2$. In addition, it is easy to check that $\bar{R}_p^{\rm N}(\tau,0) > \bar{R}_q^{\rm N}(\tau,0) = 0$. Combing these two observations with *Lemma* 4, we know that $\bar{R}_{\rm F}^{\rm N}(\tau,\alpha)$ achieves its maximum when $\bar{R}_p^{\rm N}(\tau,\alpha) = \bar{R}_q^{\rm N}(\tau,\alpha)$, which gives the optimal power allocation factor $\tau^\star$. □

Due to the presence of the multiple summation terms and the modified Bessel function in (5.7), it is impossible to derive $\alpha^\star$ in an explicit form. However, since $\bar{R}_p^{\rm N}(\tau,\alpha)$ and $\bar{R}_q^{\rm N}(\tau,\alpha)$ are monotonically decreasing and increasing functions of α for $\alpha \in \left(0, \frac{1}{\gamma_{\rm th}+2}\right)$, respectively, the numerical value of $\alpha^\star$ can be efficiently obtained by using a bisection method [69].

Next, we seek to find further insights on $\alpha^\star$ based on the asymptotic average throughput $\bar{R}_{n,\infty}^{\rm N}(\tau,\alpha)$. When $\alpha \in \left(0, \frac{1}{\gamma_{\rm th}+2}\right)$, let $\bar{R}_{p,\infty}^{\rm N}(\tau,\alpha) = \bar{R}_{q,\infty}^{\rm N}(\tau,\alpha)$, then the approximated $\alpha^\star$ can be found by solving the following equation

$$\frac{\alpha^q}{(1-\alpha(1+\gamma_{\rm th}))^p} = \vartheta^{q-p}\frac{\varphi_q}{\varphi_p}, \tag{5.13}$$

where $\vartheta \triangleq \gamma_{\rm th}\phi^{\frac{1-\tau}{\tau}}$. Based on (5.13), more valuable insights regarding $\alpha^\star$ are given.

Remark 13. *For $q = 2$, the equation in* (5.13) *can be easily solved, and $\alpha^\star$ is given by*

$$\alpha^\star \simeq \frac{-\vartheta\varphi_2(1+\gamma_{th}) + \sqrt{(\vartheta\varphi_2(1+\gamma_{th}))^2 + 4\vartheta\varphi_1\varphi_2}}{2\varphi_1}. \tag{5.14}$$

When $q = 3$ and $q = 4$, (5.13) *can be re-expressed as a cubic equation and a quartic equation of α, respectively, and thus the closed-form root can be obtained by following the procedures in [37]. It is impossible to obtain the closed-form solution of* (5.13) *when $q \geq 5$, but we can use the standard root-finding algorithms to find a numerical solution. Another non-trivial observation is that $\alpha^\star$ decreases as M, m, η, Ω_0, P_b or τ increases. This can be inferred from the following two facts: 1) The left-hand side of* (5.13) *increases as α grows, and the right-hand side of* (5.13) *increases as ϕ grows. 2) ϕ, given in* (5.8)*, decreases as M, m, η, Ω_0, P_b or τ grows.*

Note that the increase of M, m, η, Ω_0, P_b or τ leads to a larger transmit power of S. Thus, the aforementioned observation indicates that to achieve a better fair

throughput, power allocation should be more biased towards the weak user, i.e., the p-th user, as the transmit power of S increases. This conclusion also can be applied in conventional downlink NOMA networks without EH when we aim to find the optimal fairness between two paired users.

5.4.2 Optimal Time Allocation for a Given Power Allocation Ratio

In this subsection, we investigate how to find the optimal time allocation factor $\tau^\star$ for a given α to maximize the fair throughput, and the optimization problem can be written as

$$\mathbf{P5}: \max_{0<\tau<1} \bar{R}_{\mathrm{F}}^{\mathrm{N}}(\tau, \alpha). \tag{5.15}$$

Before solving problem **P5**, we first introduce two lemmas.

Lemma 5. *$\bar{R}_n^N(\tau, \alpha)$ is pseudoconcave[2] with respect to (w.r.t.) τ, and there exists an unique optimal τ_n^* that maximizes $\bar{R}_n^N(\tau, \alpha)$.*

Proof. See section 5.8.3. □

Lemma 6. *Consider a constrained maximization problem and let $\bar{x}$ be a feasible solution. If the objective function is pseudoconcave at $\bar{x}$, the constraint functions are differentiable and quasiconcave[3] at $\bar{x}$, and Karush-Kuhn-Tucker (KKT) conditions hold true at $\bar{x}$, then $\bar{x}$ is the global optimal solution to the maximization problem.*

Proof. This lemma can be easily proved according to [92, Theorem 4.3.8]. □

By invoking Lemma 5 and Lemma 6, we are able to solve problem **P5** and obtain the optimal time allocation factor $\tau^\star$.

Proposition 11. *The optimal time allocation factor $\tau^\star$ for problem **P5** is given by*

$$\tau^\star = \begin{cases} \tau_p^*, \text{if } \bar{R}_q^N\left(\tau_p^*, \alpha\right) \geq \bar{R}_p^N\left(\tau_p^*, \alpha\right) \\ \tau_q^*, \text{if } \bar{R}_p^N\left(\tau_q^*, \alpha\right) \geq \bar{R}_q^N\left(\tau_q^*, \alpha\right) \\ \tau^o, \text{else} \end{cases}, \tag{5.16}$$

[2]Function f is said to be pseudoconcave when the following conditions are satisfied: 1) f is differentiable; 2) f is defined on an open convex set; 3) for $\forall x_1, x_2$ in the open convex set, we have $\nabla f(x_2)(x_1 - x_2) > 0$ if $f(x_1) > f(x_2)$, with ∇f being the usual gradient of f [91, Definition 3.13]. If f is pseudoconcave, then $g = -f$ is pseudoconvex.

[3]Function f is called quasiconcave if the following conditions are satisfied: 1) f is defined on a convex set $\mathbb{C}$; 2) for any real value δ, function f's upper-level sets $\mathbb{U}(f, \delta) = \{x : x \in \mathbb{C}, f(x) \geq \delta\}$ are convex sets [91, Definition 3.1].

where τ^o is the solution $\bar{R}_p^N(\tau,\alpha) = \bar{R}_q^N(\tau,\alpha)$.

Proof. By introducing an auxiliary variable ϱ, problem **P5** can be equivalently transformed into a new problem

$$\begin{aligned} &\mathbf{P6}: \max_{\tau,\varrho} \ \varrho \\ &\text{s.t.} \ \ \bar{R}_p^{\mathrm{N}}(\tau,\alpha) \geq \varrho, \bar{R}_q^{\mathrm{N}}(\tau,\alpha) \geq \varrho, 0<\tau<1. \end{aligned} \tag{5.17}$$

According to *Lemma* 2, the constraint function $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)-\varrho$ in problem **P6** is pseudoconcave in τ and ϱ, and thus quasiconcave in τ and ϱ [91]. In addition, the objective function in problem **P6** is a linear function of ϱ. Following *Lemma* 6, we know that the global optimal solution $(\tau^\star, \varrho^\star)$ to problem **P6** can be obtained using the KKT conditions. The Lagrangian function associated with problem **P6** is first given by

$$\mathcal{L}(\tau,\varrho) = \varrho + \lambda_1\big(\bar{R}_p^{\mathrm{N}}(\tau,\alpha)-\varrho\big)+\lambda_2\big(\bar{R}_q^{\mathrm{N}}(\tau,\alpha)-\varrho\big), \tag{5.18}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the Lagrange multipliers. As such, the optimal solution $(\tau^\star, \varrho^\star)$ should satisfy the following KKT conditions

$$\begin{cases} \frac{\partial \mathcal{L}(\tau,\varrho)}{\partial \tau} = \lambda_1 \frac{\partial \bar{R}_p^{\mathrm{N}}(\tau,\alpha)}{\partial \tau} + \lambda_2 \frac{\partial \bar{R}_q^{\mathrm{N}}(\tau,\alpha)}{\partial \tau} = 0 \\ \frac{\partial \mathcal{L}(\tau,\varrho)}{\partial \varrho} = 1 - \lambda_1 - \lambda_2 = 0 \\ \lambda_1 \left(\bar{R}_p^{\mathrm{N}}(\tau,\alpha) - \varrho\right) = 0 \\ \lambda_2 \left(\bar{R}_q^{\mathrm{N}}(\tau,\alpha) - \varrho\right) = 0 \\ \bar{R}_p^{\mathrm{N}}(\tau,\alpha) \geq \varrho \\ \bar{R}_q^{\mathrm{N}}(\tau,\alpha) \geq \varrho \\ 0 < \tau < 1 \end{cases}. \tag{5.19}$$

According to (5.19), there are three candidate KKT points. *Candidate KKT point 1: $\lambda_1 = 0, \lambda_2 = 1$:* In this case, we have $\tau^\star = \tau_q^*$ and $\bar{R}_p^{\mathrm{N}}\left(\tau_q^*,\alpha\right) \geq \bar{R}_q^{\mathrm{N}}\left(\tau_q^*,\alpha\right) = \varrho^\star$. *Candidate KKT point 2: $\lambda_1 = 1, \lambda_2 = 0$:* In this case, we have $\tau^\star = \tau_p^*$ and $\bar{R}_q^{\mathrm{N}}\left(\tau_p^*,\alpha\right) \geq \bar{R}_p^{\mathrm{N}}\left(\tau_p^*,\alpha\right) = \varrho^\star$. *Candidate KKT point 3: $\lambda_1 > 0, \lambda_2 = 1-\lambda_1 > 0$:* In this case, the optimal ϱ and τ (denoted as τ^o) can be derived by solving the following equation system

$$\begin{cases} \lambda_1 \frac{\partial \bar{R}_p^{\mathrm{N}}(\tau,\alpha)}{\partial \tau} + (1-\lambda_1) \frac{\partial \bar{R}_q^{\mathrm{N}}(\tau,\alpha)}{\partial \tau} = 0 \\ \bar{R}_p^{\mathrm{N}}(\tau,\alpha) = \bar{R}_q^{\mathrm{N}}(\tau,\alpha) = \varrho \end{cases}. \tag{5.20}$$

Since the objective function in **P5** is pseudoconcave[4] and **P6** is equivalently transformed from **P5**, we know that **P6** has a unique global solution. Combing this fact

[4]Minimum of pseudoconcave functions is still pseudoconcave [91].

with Lemma 6, we know that only one of the three candidate KKT points exists. Therefore, the optimal time allocation factor $\tau^\star$ for problem **P5** can be derived as (5.16). □

Generally, it is impossible to derive τ_n^* and τ^o in an explicit form. However, since $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)$ is a pseudoconcave function of τ, we can efficiently find the numerical value of τ_n^* using a golden-section search method [93]. After obtaining τ_n^*, τ^o can be obtained using the bisection method over the region $\left(\min\left(\tau_p^*,\tau_q^*\right),\max\left(\tau_p^*,\tau_q^*\right)\right)$.

Now we seek to find further insights on $\tau^\star$ based on the asymptotic average throughput $\bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)$.

Corollary 6. *In the high SNR regime, $\tau^\star$ is a decreasing function of M, m, η, Ω_0 and P_b.*

Proof. According to Proposition 11, to complete the proof of Corollary 6, we need to show that τ_n^* $(n=p,q)$ and τ^o decrease with the increase of M, m, η, Ω_0 and P_b. We first find the approximation of τ_n^* in the high SNR regime.

The first-order and second-order partial derivative of $\bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)$ w.r.t. τ are respectively expressed as

$$\frac{\partial \bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)}{\partial \tau}=-R_t\left[1-\varphi_n\left(\frac{\phi\theta_n\left(1-\tau\right)}{\tau}\right)^n\left(1+\frac{n}{\tau}\right)\right], \tag{5.21}$$

and

$$\frac{\partial^2 \bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)}{\partial \tau^2}=-R_t\varphi_n(\phi\theta_n)^n n(n+1)\frac{(1-\tau)^{n-1}}{\tau^{n+2}}<0. \tag{5.22}$$

According to (5.22), we can tell that $\bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)$ is a concave function of τ. Along with another fact that $\left.\frac{\partial \bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)}{\partial \tau}\right|_{\tau\to 0^+}>0$ and $\left.\frac{\partial \bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)}{\partial \tau}\right|_{\tau=1}<0$, we know that the optimal τ_n^* maximizing $\bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)$ is the root of $\frac{\partial \bar{R}_{n,\infty}^{\mathrm{N}}(\tau,\alpha)}{\partial \tau}=0$, which can be found by solving the following equation

$$n\frac{(1-\tau)^n}{\tau^{n+1}}+\left(\frac{1-\tau}{\tau}\right)^n=\left(\frac{1}{\phi\theta_n}\right)^n\frac{1}{\varphi_n}. \tag{5.23}$$

Note that the left-hand and the right-hand sides of (5.23) decrease with τ and ϕ, respectively, which indicates that τ_n^* is an increasing function of ϕ. Next, we find the approximation of τ^o in the high SNR regime.

The approximation of τ^o can be found by solving $\bar{R}^{\mathrm{N}}_{p,\infty}(\tau,\alpha)=\bar{R}^{\mathrm{N}}_{q,\infty}(\tau,\alpha)$, which can be further simplified as (5.13). According to (5.13), we can easily tell that τ^o is an increasing function of ϕ. To this end, we know that τ_n^* and τ^o increase as ϕ grows. Since ϕ, given in (5.8), decreases as M, m, η, Ω_0 or P_b increases, we can obtain the conclusion given in Corollary 6. □

5.4.3 Joint Optimization of Time Allocation and Power Allocation Ratios

Based on the solutions to the two sub-optimization problems, an alternating optimization algorithm, as described in Algorithm 5.1, is designed to approximate the joint optimal solution (τ^*,α^*). Specifically, the proposed algorithm starts by initializing α as $\alpha[0]$, then we iteratively optimize α and τ by the following sequence: $\alpha[1]\rightarrow\tau[2]\rightarrow\alpha[2]\rightarrow\cdots\rightarrow\tau[k]\rightarrow\alpha[k]\rightarrow\cdots\rightarrow\tau^*\rightarrow\alpha^*$, $k\geq 1$. In each step of the iteration, we find the optimal $\tau[k](\alpha[k])$ based on $\alpha[k-1](\tau[k])$ which is given in the last iteration. This algorithm terminates when the improvement of the fair throughput is under an acceptable tolerance ϵ. In theory, the proposed alternating optimization algorithm fast converges to a stationary point of problem **P3** [94, Corollary 1].

Algorithm 5.1 Alternating optimization to approximate the joint optimal α^* and τ^*.

1: Initialization: Set $k=1$, $R^{\mathrm{N}}_{\mathrm{F}}(\tau[0],\alpha[0])=0$, $\epsilon>0$, and randomly select a pair of values τ and α as $(\tau[1],\alpha[1])$;
2: Calculate $\bar{R}^{\mathrm{N}}_{\mathrm{F}}(\tau[1],\alpha[1])$;
3: **while** $\bar{R}^{\mathrm{N}}_{\mathrm{F}}(\tau[k],\alpha[k])-\bar{R}^{\mathrm{N}}_{\mathrm{F}}(\tau[k-1],\alpha[k-1])>\epsilon$ **do**
4: $k=k+1$;
5: Find $\tau[k]$ for a given $\alpha[k-1]$ based on (5.16);
6: Find $\alpha[k]$ for a given $\tau[k]$ using a bisection method;
7: Calculate $\bar{R}^{\mathrm{N}}_{\mathrm{F}}(\tau[k],\alpha[k])$;
8: **end while**
9: Return with $\alpha^*=\alpha[k],\tau^*=\tau[k]$;

5.5 User Fairness Maximization for the OBT Protocol

In this section, we intend to find the optimal fair throughput of the OBT protocol. We first analyze the average throughput of the OBT protocol for each selected user. Then, we discuss how to jointly optimize two time allocation ratios to maximize the fair throughput.

5.5.1 Throughput Analysis of the OBT Protocol

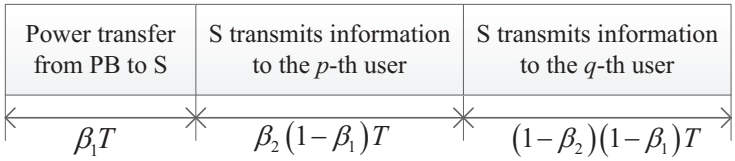


Figure 5.2: OMA-based transmission protocol.

As shown in Fig. 5.2, the OBT protocol consists of three phases. The first phase is still dedicated to wireless power transfer with a duration $\beta_1 T$, $0 < \beta_1 < 1$. The information transfer involves two disjoint phases where the transmission for the p-th user and the q-th user occupy $\beta_2 (1-\beta_1) T$ and $(1-\beta_2)(1-\beta_1) T$, respectively, $0 < \beta_2 < 1$. Recall that in the NBT protocol, the information transmission rate for each selected user is R_t, i.e., the total transmission rate at S is $2R_t$. For fair comparison, the transmission rate at S is still assumed to be $2R_t$ in the OBT protocol [83]. As such, the average throughput of the p-th user and the q-th user are respectively given by

$$\bar{R}_p^{\mathrm{O}}(\beta_1, \beta_2) = 2R_t \beta_2 (1-\beta_1)\big(1-\mathrm{P}_p^{\mathrm{O}}(\beta_1)\big), \tag{5.24}$$

$$\bar{R}_q^{\mathrm{O}}(\beta_1, \beta_2) = 2R_t (1-\beta_2)(1-\beta_1)\big(1-\mathrm{P}_q^{\mathrm{O}}(\beta_1)\big), \tag{5.25}$$

where the superscript "O" refers to the OBT protocol and $\mathrm{P}_n^{\mathrm{O}}(\beta_1)$ denotes the outage probability of the n-th user in the OBT protocol.

Proposition 12. *The closed-form expression of $P_n^O(\beta_1)$ can be obtained by replacing τ and θ_n in (5.7) with β_1 and $\gamma'_{th} \triangleq 2^{2R_t} - 1$, respectively.*

Proof. See Section 5.8.4. □

Next we seek to find the optimal fair throughput of the OBT protocol by jointly optimizing time allocation ratios β_1 and β_2.

5.5.2 Joint Optimization of Time Allocation Factors

For the OBT protocol, the optimal fair throughput can be found by jointly optimizing two time allocation ratios, i.e., β_1 and β_2. As such, the joint optimization problem

can be stated as

$$\mathbf{P7}: \max_{\beta_1,\beta_2} \bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1,\beta_2) \tag{5.26}$$

$$\text{s.t. } 0<\beta_1<1, 0<\beta_2<1, \tag{5.27}$$

where $\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1,\beta_2) \triangleq \min_{n\in\{p,q\}} \bar{R}_n^{\mathrm{O}}(\beta_1,\beta_2)$. To obtain the joint optimal solution of problem **P7**, denoted as (β_1^*,β_2^*), we first find the optimal β_2 for a given β_1.

Proposition 13. *For a given β_1, the optimal β_2, denoted as $\beta_2^\star$, is given by*

$$\beta_2^\star = \frac{1-P_q^O(\beta_1)}{2-P_p^O(\beta_1)-P_q^O(\beta_1)}. \tag{5.28}$$

Proof. According to (5.24) and (5.25), it is clear that $\bar{R}_p^{\mathrm{O}}(\beta_1,\beta_2)$ and $\bar{R}_q^{\mathrm{O}}(\beta_1,\beta_2)$ are increasing and decreasing functions, respectively, of β_2. In addition, it is trivial to check that $\bar{R}_q^{\mathrm{O}}(\beta_1,0) > \bar{R}_p^{\mathrm{O}}(\beta_1,0) = 0$ and $\bar{R}_p^{\mathrm{O}}(\beta_1,1) > \bar{R}_q^{\mathrm{O}}(\beta_1,1) = 0$. Thus, the optimal fair throughput is attained when $\bar{R}_p^{\mathrm{O}}(\beta_1,\beta_2) = \bar{R}_q^{\mathrm{O}}(\beta_1,\beta_2)$, from which $\beta_2^\star$ is derived as in (5.28). □

Remark 14. *Due to the fact that $|\tilde{g}_p|^2 < |\tilde{g}_q|^2$, it is trivial to verify that $P_q^O(\beta_1) < P_p^O(\beta_1)$, thus we know that $\beta_2^\star > 0.5$ according to* (5.28)*. This observation is quite intuitive since the weak (p-th) user should occupy more information transmission time than the strong (q-th) user to achieve better fairness.*

Now we seek to find more insights regarding $\beta_2^\star$ in the high SNR regime.

Corollary 7. *In the high SNR range, $\beta_2^\star$ is a decreasing function of M, m, η, Ω_0, P_b, $1/\gamma'_{th}$ and β_1.*

Proof. By replacing τ and θ_n in (5.8) with β_1 and γ'_{th}, respectively, the asymptotic outage probability of the OBT protocol for the n-th user is given by $\mathrm{P}_{n,\infty}^{\mathrm{O}} \simeq \chi^n\varphi_n$, where $\chi \triangleq \frac{\gamma'_{\mathrm{th}}m(1-\beta_1)}{\eta\rho_b\Omega_0(Mm-1)\beta_1}$. Substituting the expression of $P_{n,\infty}^{\mathrm{O}}$ into (5.28), we have $\beta_2^\star \simeq \frac{1-\chi^q\varphi_q}{2-\chi^p\varphi_p-\chi^q\varphi_q}$ in the high SNR regime. Taking the first-order partial derivative of $\beta_2^\star$ w.r.t. χ, we have

$$\frac{\partial\beta_2^\star}{\partial\chi} = \frac{p\chi^{p-1}\varphi_p - q\chi^{q-1}\varphi_q + \chi^{p+q-1}\varphi_p\varphi_q(q-p)}{\left(2-\chi^p\varphi_p-\chi^q\varphi_q\right)^2}. \tag{5.29}$$

As $\chi \to 0$ in the high SNR regime and $p < q$, we have $\frac{\partial \beta_2^\star}{\partial \chi} > 0$. Thus, we know that $\beta_2^\star$ is an increasing function of χ. Along with another fact that χ decreases as M, m, η, Ω_0, P_b, $1/\gamma'_{\text{th}}$ or β_1 increases, we can derive the conclusion given in Corollary 7. □

Remark 15. *The growth of M, m, η, Ω_0, P_b or β_1 leads to an increase of the transmit power of S. Therefore, Corollary 7 implies that in the high SNR regime, the optimal time allocation ratio in the information transmission phase, $\beta_2^\star$, gradually approaches to 0.5 as the transmit power of S increases. Note that this is quite different from the resource allocation strategy of the NBT protocol in the information transmission phase. According to Remark 3, we know that in the NBT protocol, as the transmit power of S increases, power allocation in the information transfer phase should be more biased towards the weak user in order to achieve better fairness.*

With the derived $\beta_2^\star$, the fair throughput of the OBT protocol can be expressed as

$$\bar{R}_{\text{F}}^{\text{O}}(\beta_1, \beta_2^\star) = 2R_t(1-\beta_1)\,\frac{\left(1-\text{P}_p^{\text{O}}(\beta_1)\right)\left(1-\text{P}_q^{\text{O}}(\beta_1)\right)}{2-\text{P}_p^{\text{O}}(\beta_1)-\text{P}_q^{\text{O}}(\beta_1)}. \tag{5.30}$$

Now the original joint optimization problem in **P7** reduces to the optimization of β_1. To obtain the joint optimal β_1^*, the property of $\bar{R}_{\text{F}}^{\text{O}}(\beta_1, \beta_2^\star)$ is first given.

Proposition 14. *$\bar{R}_F^O(\beta_1, \beta_2^\star)$ is a log-concave function of β_1.*

Proof. See section 5.8.5. □

According to Proposition 14, we know that the unique optimal β_1^* maximizing $\bar{R}_{\text{F}}^{\text{O}}(\beta_1, \beta_2^\star)$ can be found by using a golden-section search method. After finding β_1^*, the joint optimal β_2^* can be obtained by substituting β_1^* into (5.28).

5.6 Numerical Results

In this section, we first provide numerical results to verify the analytical results, and then highlight some nontrivial insights on the optimization of user fairness for both the NBT and OBT protocols. We assume that the variance of each channel is characterized by a path loss model [95], i.e., $\Omega_0 = \zeta\,(d_0)^{-\varpi}$, $\Omega_i = \zeta(d_i)^{-\varpi}$, $1 \le i \le N$, where ζ denotes the signal attenuation corresponding to a distance of 1 meter, d_0 is

the distance from the PB to S, d_i is the distance from S to U_i, and ϖ is the path-loss exponent. Unless otherwise stated, the system parameters are set as follows. The Nakagami-m fading parameter is $m = 4$, which corresponds to a Rician factor $K = 3 + \sqrt{12}$ [73]. The signal attenuation corresponding to a distance of 1 meter is $\zeta = -20$dB, and the path-loss exponent is $\varpi = 3$ [96]. The number of users is $N = 5$,[5] and the number of antennas at the PB is $M = 10$. Due to the relatively short power transfer distance, the distance between the PB and S is $d_0 = 4$m, and the energy conversion efficiency is $\eta = 0.5$. We assume that users are close to each other but the distances to the source are slightly different which are set to be $d_1 = 22$m, $d_2 = 24$m, $d_3 = 26$m, $d_4 = 28$m and $d_5 = 30$m.[6] Targeted data rate is $R_t = 2$bps/Hz. Circuit power consumption at S is $P_c = 0.5$mW, and noise power is $\sigma^2 = -100$dBm. In addition, the acceptable tolerance for all the algorithms, i.e., bisection, golden-section and alternating optimization, are set as 10^{-3}.

5.6.1 Verification of Outage Performance Analysis

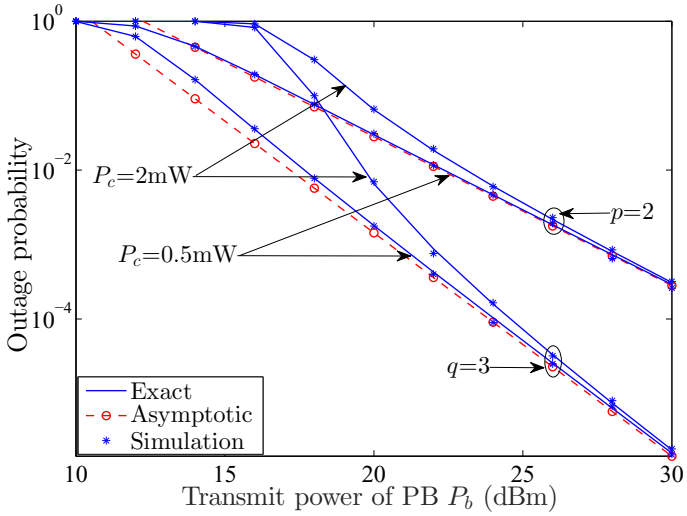


Figure 5.3: Outage probability of the NBT protocol for different circuit power consumption P_c ($\tau = 0.4$, $\alpha = 0.2$).

[5] A variety number of users from 2 to 50 have been considered in other works [39, 59, 65, 81, 83, 84, 97, 98], here we consider 5 users since it already provides us enough user pairing options to study the effect of user pairing on the fair performance gain of NOMA over OMA.

[6] In conventional downlink networks without EH, users can be away from the source as far as hundred of meters [95]. In our considered network, however, since the harvested energy at the source is very limited due to signal attenuation, we here limit the distance from the source to the user to no more than 30m, which also belongs to an indoor environment.

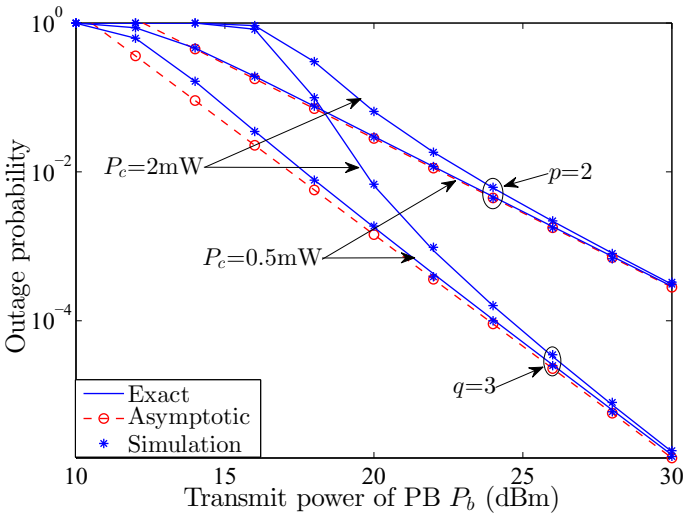


Figure 5.4: Outage probability of the OBT protocol for different circuit power consumption P_c ($\beta_1 = 0.4$).

In this subsection, we intend to examine the accuracy of the exact outage analysis and the asymptotic outage analysis. Figs. 5.3 and 5.4 illustrate the outage performance of the NBT and OBT protocols, respectively. It can be observed that our exact analytical results match well with the simulation results, and the asymptotic analytical results accurately coincide with the exact analytical results in the high SNR range. In addition, we can see that the circuit power consumption at S has a major impact on the outage performance in the medium SNR range. However, this effect gradually diminishes as P_b enlarges, i.e., $\frac{P_c}{P_b} \to 0$. As the simulation results corroborate our theoretical analysis, we will only use the proposed analytical results in the following subsections for the numerical investigation.

5.6.2 Key Design Insights for the NBT Protocol

Figs. 5.5 and 5.6 present insights on the optimal power allocation ratio $\alpha^\star$ for a given time allocation ratio τ. Fig. 5.5 plots the average throughput of each selected user versus power allocation ratio α. As reported in Lemma 4, $\bar{R}_p^{\mathrm{N}}(\tau, \alpha)$ decreases as α increases, and $\bar{R}_q^{\mathrm{N}}(\tau, \alpha)$ achieves its maximum value at point $\alpha = \frac{1}{2+\gamma_{\mathrm{th}}} = 0.2$. It can be also observed in Fig. 5.5 that the optimal fair throughput is reached when $\bar{R}_p^{\mathrm{N}}(\tau, \alpha^\star) = \bar{R}_q^{\mathrm{N}}(\tau, \alpha^\star)$, which confirms Proposition 10. Fig. 5.6 shows the impact of P_b and τ on $\alpha^\star$. It is easy to see that $\alpha^\star$ decreases as τ or P_b increases, which verifies Remark 13.

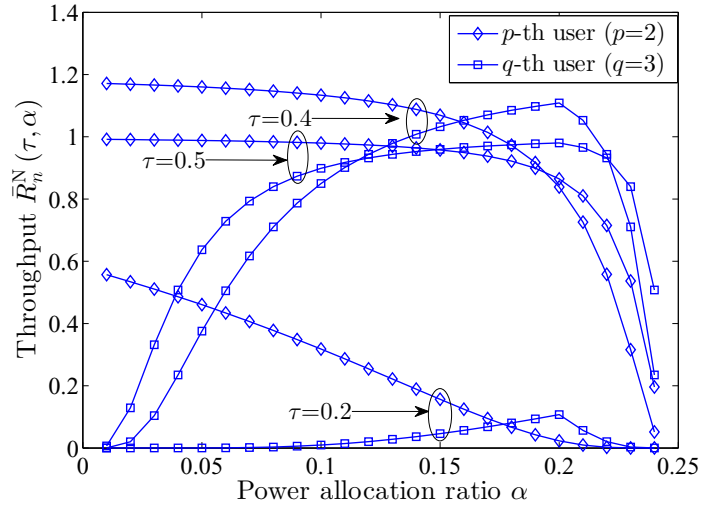


Figure 5.5: Throughput of each selected user for different time allocation ratios τ ($P_b = 15$dBm).

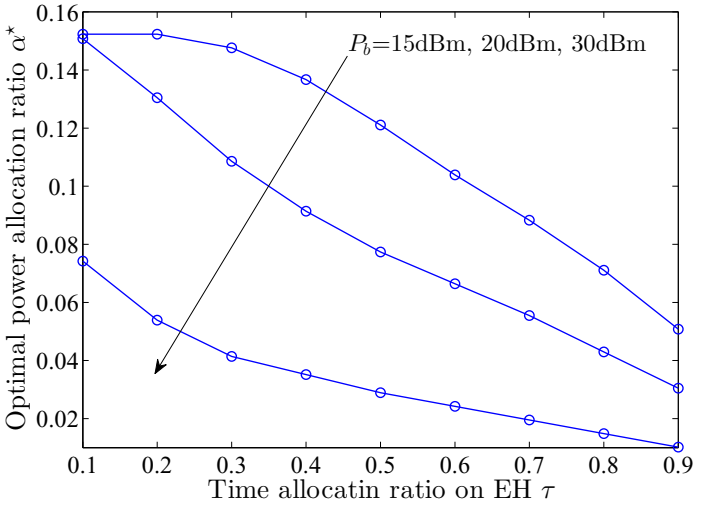


Figure 5.6: Optimal power allocation ratio $\alpha^\star$ for different transmit power of PB P_b ($p = 2, q = 3$).

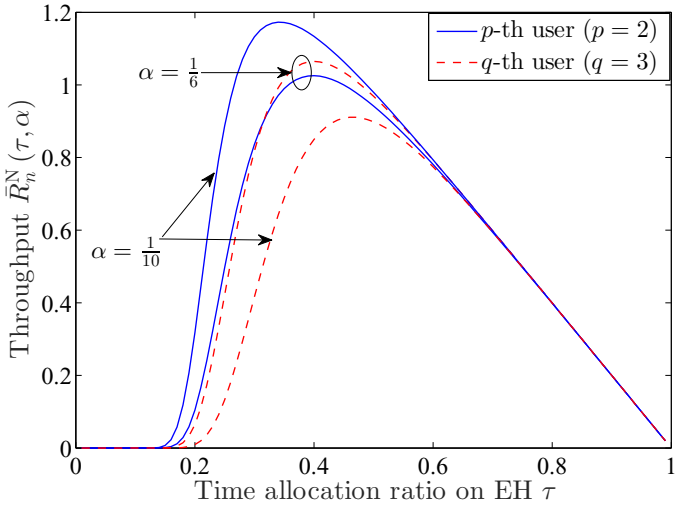


Figure 5.7: Throughput of each selected user for different power allocation ratios α ($P_b = 15$dBm).

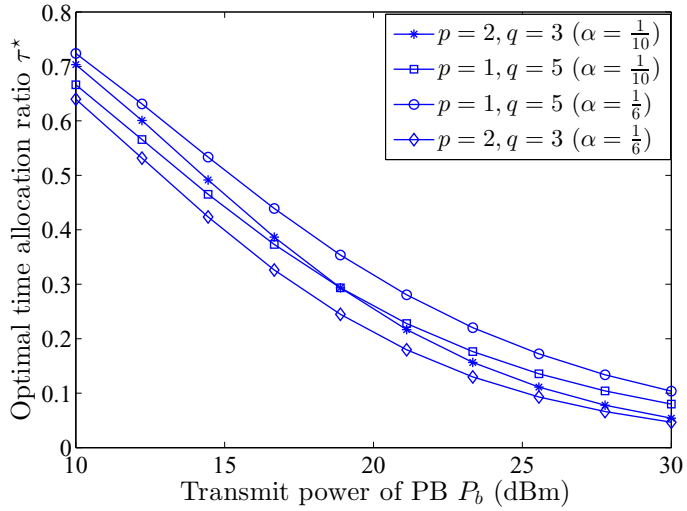


Figure 5.8: Optimal time allocation ratio $\tau^\star$ for different paired users and power allocation ratios α.

Figs. 5.7 and 5.8 present insights on the optimal time allocation ratio $\tau^\star$ for a given power allocation ratio α. Fig. 5.7 depicts the average throughput of each selected user versus time allocation ratio τ. From Fig. 5.7, we can tell that $\bar{R}_n^{\mathrm{N}}(\tau, \alpha)$ is a pseudoconcave funtion of τ, and there exists an unique τ_n^* that maximizes $\bar{R}_n^{\mathrm{N}}(\tau, \alpha)$, which confirms Lemma 5. In addition, we can notice from Fig. 5.7 that $\bar{R}_2^{\mathrm{N}}\left(\tau_2^*, \frac{1}{6}\right) < \bar{R}_3^{\mathrm{N}}\left(\tau_2^*, \frac{1}{6}\right)$ and $\bar{R}_2^{\mathrm{N}}\left(\tau_2^*, \frac{1}{10}\right) > \bar{R}_3^{\mathrm{N}}\left(\tau_2^*, \frac{1}{10}\right)$. Thus, according to (5.16), we have $\tau^\star = \tau_2^*$ when $\alpha = \frac{1}{6}$ and $\tau^\star = \tau_3^*$ when $\alpha = \frac{1}{10}$. Fig. 5.8 illustrates $\tau^\star$ for different paired users and different power allocation ratios. It can be seen that for the same two paired users, different values of α result in different values of $\tau^\star$. In addition, it can be observed that $\tau^\star$ decreases as P_b increases, which verifies Corollary 6.

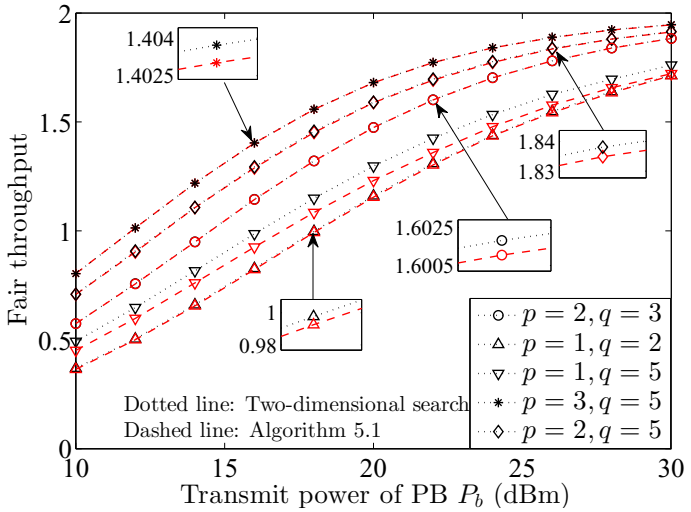


Figure 5.9: Fair throughput of the NBT protocol with different algorithms.

Fig. 5.9 shows the fair throughput obtained by the proposed alternating optimization algorithm, i.e., Algorithm 5.1. To examine the effectiveness of Algorithm 5.1, we also provide the fair throughput obtained by a two-dimension exhaustive search algorithm which ensures to find the global optimal solution. As shown in Fig. 5.9, in most cases, the fair throughput achieved by Algorithm 5.1 is very close to the optimal fair throughput achieved by the exhaustive search.

5.6.3 Key Design Insights for the OBT Protocol

Fig. 5.10 plots optimal time allocation ratio $\beta_2^\star$ versus time allocation ratio β_1 for different P_b. It can be seen that $\beta_2^\star$ monotonically decreases with the increase of β_1

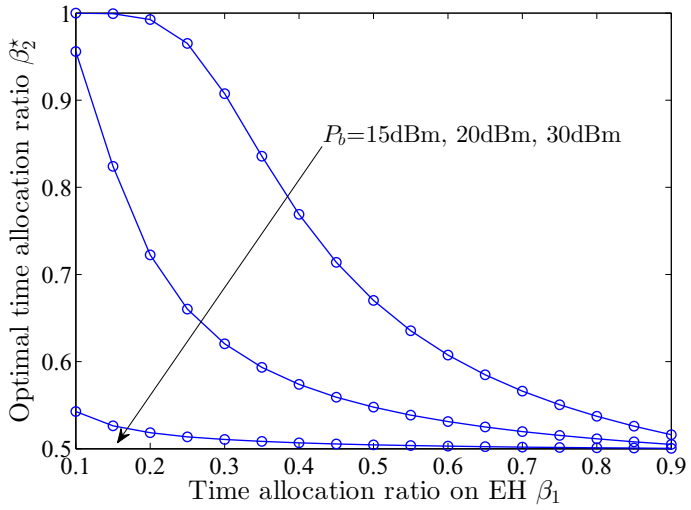


Figure 5.10: Optimal time allocation ratio $\beta_2^\star$ for different transmit power of PB P_b.

for the same P_b. In addition, for a fixed β_1, the larger P_b, the smaller $\beta_2^\star$. These two observations confirm the insights given in Corollary 7.

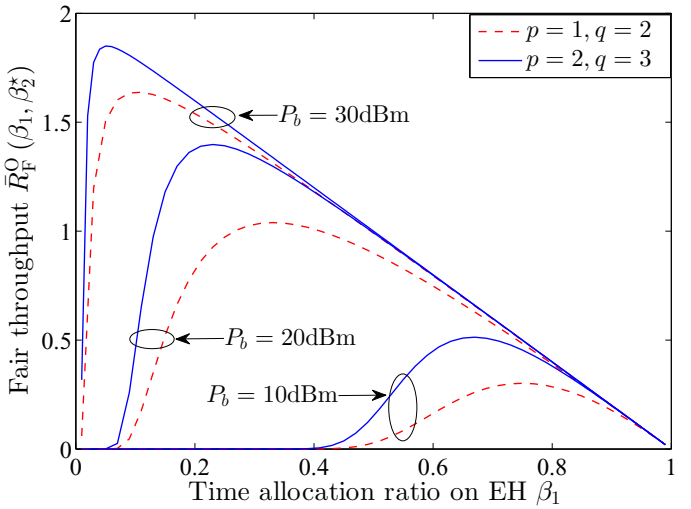


Figure 5.11: Fair throughput $\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1, \beta_2^\star)$ for different transmit power of PB P_b and paired users.

Fig. 5.11 shows the fair throughput $\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1, \beta_2^\star)$ (given in (5.30)) versus time allocation ratio β_1 for different transmit power of PB P_b. It is obvious that there exists an unique optimal β_1 that maximizes $\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1, \beta_2^\star)$, and the optimal β_1 decreases as P_b increases.

The fair throughput of the OBT protocol achieved by the golden-section search method and the two-dimensional search method are plotted in Fig. 5.12. From the figure, we can tell that the fair throughput with the golden-section search is identical

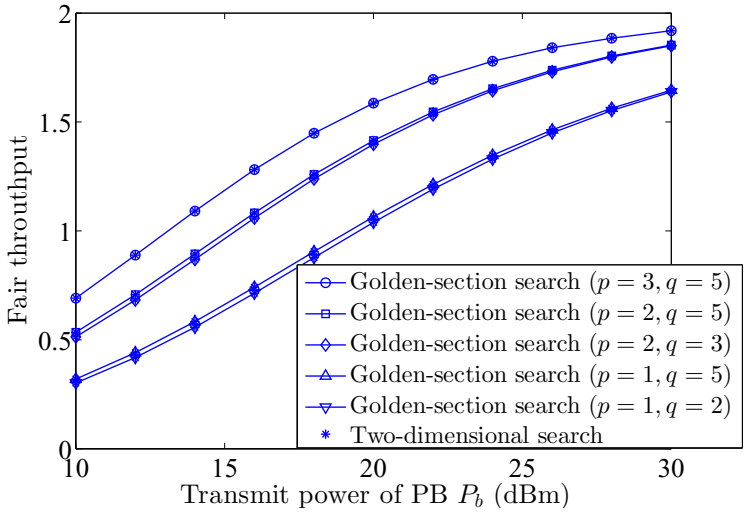


Figure 5.12: Fair throughput of the OBT protocol with different algorithms.

to that with the exhaustive search, which illustrates the effectiveness of the golden-section search method.

5.6.4 Fair Performance Comparison of the NBT and OBT Protocols

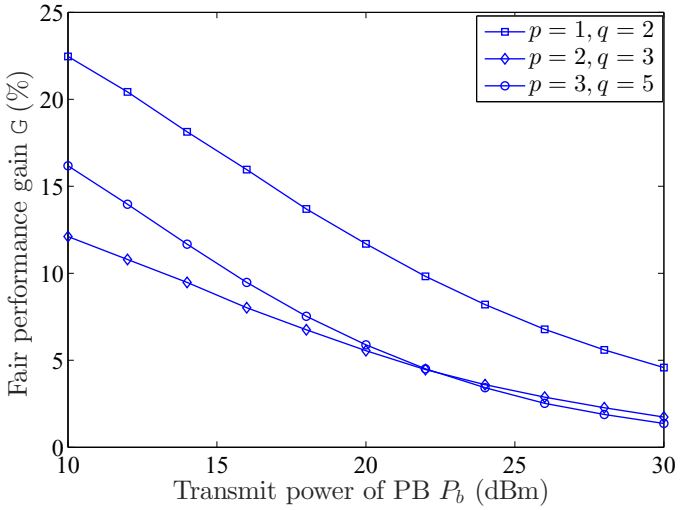


Figure 5.13: Fair performance gain G versus transmit power of PB.

In this subsection, we investigate the fair performance gain of the NBT protocol relative to the OBT protocol. Specifically, the fair performance gain is defined as $\mathsf{G}(\%) = 100 \times \left(\frac{\bar{R}_{\mathrm{F}}^{\mathrm{N}}(\tau^*, \alpha^*)}{\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1^*, \beta_2^*)} - 1 \right)$. From Fig. 5.13, we can easily tell that the NBT protocol outperforms the OBT protocol in terms of user fairness. However, the fair performance gain shrinks as P_b increases. Table 5.2 lists the fair performance gain for

Table 5.2: Fair performance gain of the NBT over the OBT protocols.

	$P_b = 10$dBm				$P_b = 30$dBm			
$\mathsf{G}(\%)$ p \ q	2	3	4	5	2	3	4	5
1	22.46	36.53	46.52	53.96	4.58	5.98	6.45	6.81
2	N/A	12.11	23.04	32.61	N/A	1.74	2.7	3.32
3	N/A	N/A	7.75	16.18	N/A	N/A	0.73	1.37
4	N/A	N/A	N/A	5.64	N/A	N/A	N/A	0.43

different paired users with $P_b = 10$dBm and $P_b = 30$dBm. It can be seen that the fair performance gain enlarges as the channel gain difference between the two selected users, i.e., $q - p$, increases. The fair performance gain reaches its maximum value 53.96% when $p = 1, q = 5$ and $P_b = 10$dBm. When $p = 4, q = 5$ and $P_b = 30$dBm, the fair performance gain attains its minimum value 0.43%.

5.7 Summary

This chapter has investigated the optimal user fairness of a PB-assisted downlink multiuser network with the NBT and OBT protocols. Specifically, the optimal user fairness of the NBT protocol can be found by jointly optimizing the time allocation and power allocation ratios. Since the joint optimization problem is nonconvex in general, the initial problem has been decomposed into two subproblems, i.e., optimal power allocation ratio for a given time allocation ratio and optimal time allocation ratio for a given power allocation ratio. After solving these two subproblems with linear search methods, we have applied an alternating optimization algorithm to approximate the joint optimal solution. The optimal user fairness of the OBT protocol has been successfully found by using the golden-section search method to find the joint optimal time allocation ratios. Numerical results have verified our analytical results and revealed that the NBT protocol outperforms the OBT protocol in terms of user fairness. In addition, the fair performance gain enlarges as the transmit power of PB decreases or the channel gain difference between the two selected users increases.

5.8 Appendix

5.8.1 Proof of Proposition 9

According to the transmission protocol, the outage probability of the n-th ($n = p, q$) user can be evaluated as

$$\begin{aligned} \mathrm{P}_n^{\mathrm{N}}(\tau, \alpha) &= \Pr\left\{P_s^{\mathrm{N}} = 0\right\} + \Pr\left\{R_{x_n} < R_t, P_s^{\mathrm{N}} > 0\right\} \\ &= \Pr\left\{X \le 0\right\} + \Pr\left\{X|\tilde{g}_n|^2 < \theta_n, X > 0\right\} \\ &= F_X(0) + \int_0^{\infty} F_{|\tilde{g}_n|^2}\left(\frac{\theta_n}{x}\right) f_X(x)\, dx, \end{aligned} \tag{5.31}$$

where $R_{x_p}^{\mathrm{N}} \triangleq R_{p,x_p}$, $R_{x_q}^{\mathrm{N}} \triangleq \min\left\{R_{q,x_p}^{\mathrm{N}}, R_{q,x_q}^{\mathrm{N}}\right\}$, and $X \triangleq \frac{\eta\tau\rho_b\|\mathbf{h}\|^2}{1-\tau} - \rho_c$. Note that the first part of (5.31) denotes the energy outage probability, and the second part of (5.31) denotes the data transmission outage probability. To proceed, we need to first obtain the CDF of ordered channel gain $|\tilde{g}_n|^2$ and the PDF of X. According to the order statistics theory [99, Eq. 5.2.1], the CDF of $|\tilde{g}_n|^2$ is given by

$$F_{|\tilde{g}_n|^2}(y) = \sum_{v=n}^{N} \sum_{S_v} \prod_{l=1}^{v} F_{|g_{j_l}|^2}(y) \prod_{l=v+1}^{N} \left[1 - F_{|g_{j_l}|^2}(y)\right], \tag{5.32}$$

where $F_{|g_{j_l}|^2}(y)$ is the CDF of unordered channel gain $|g_{j_l}|^2$. By applying multinomial expansion theorem, $\prod_{l=1}^{v} F_{|g_{j_l}|^2}(y)$ can be expanded as

$$\prod_{l=1}^{v} F_{|g_{j_l}|^2}(y) = \prod_{l=1}^{v}\left(1 - e^{-\frac{y}{\Omega_{j_l}}}\right) = \sum_{\kappa=0}^{v} \widetilde{\sum} e^{-\sum_{\nu=1}^{\kappa} \frac{y}{\Omega_{j_{t_\nu}}}}. \tag{5.33}$$

In addition, we have

$$\prod_{l=v+1}^{N}\left[1 - F_{|g_{j_l}|^2}(y)\right] = \prod_{l=v+1}^{N} e^{-\frac{y}{\Omega_{j_l}}} = e^{-\sum_{l=v+1}^{N} \frac{y}{\Omega_{j_l}}}. \tag{5.34}$$

Combing (5.33) and (5.34), $F_{|\tilde{g}_n|^2}(y)$ can be finally expressed as

$$F_{|\tilde{g}_n|^2}(y) = 1 + \underbrace{\sum_{v=n}^{N} \sum_{S_v} \sum_{\kappa=0}^{v} \widetilde{\sum}}_{N-v+\kappa\neq 0} e^{-\Phi y}. \tag{5.35}$$

To derive the PDF of X, the PDF of $\|\mathbf{h}\|^2$ is first given by

$$f_{\|\mathbf{h}\|^2}(z) = \frac{(m/\Omega_0)^{Mm}}{\Gamma(Mm)} z^{Mm-1} e^{-\frac{zm}{\Omega_0}}. \tag{5.36}$$

Using the method of transformations, the PDF of X can be derived as

$$f_X(x)=\frac{(x+\rho_c)^{Mm-1}}{\Gamma(Mm)}\left(\frac{\psi(1-\tau)}{\tau}\right)^{Mm} e^{-\frac{\psi(1-\tau)}{\tau}(x+\rho_c)}. \tag{5.37}$$

Substituting (5.35) and (5.37) into (5.31), we solve the integral in (5.31) with the help of [52, Eq. 3.471.9] and obtain the closed-form expression of $\mathrm{P}_n^{\mathrm{N}}(\tau,\alpha)$ shown in (5.7).

5.8.2 Proof of Corollary 5

As $\rho_b \to \infty$ and $P_b/P_c \to \infty$, the asymptotic outage probability of the n-th user is derived as

$$\begin{aligned}
\mathrm{P}_{n,\infty}^{\mathrm{N}}(\tau,\alpha) \simeq 1 &+ \underbrace{\sum_{v=n}^{N}\sum_{S_v}\sum_{\kappa=0}^{v}\widetilde{\sum}}_{N-v+\kappa\neq 0}\frac{e^{-\psi\rho_c\frac{1-\tau}{\tau}}}{\Gamma(Mm)}\left(\psi\rho_c\frac{1-\tau}{\tau}\right)^{Mm-1} \\
&+ \sum_{\mu=1}^{Mm-1}\underbrace{\sum_{v=n}^{N}\sum_{S_v}\sum_{\kappa=0}^{v}\widetilde{\sum}}_{N-v+\kappa\neq 0}\frac{\Gamma(\mu+1)}{\Gamma(Mm)}\binom{Mm-1}{\mu} \\
&\times e^{-\psi\rho_c\frac{1-\tau}{\tau}}\left(\psi\rho_c\frac{1-\tau}{\tau}\right)^{Mm-\mu-1}\left(1-\frac{\theta_n\psi\Phi(1-\tau)}{\mu\tau}\right) && \text{(5.38a)} \\
\simeq 1 &+ \underbrace{\sum_{v=n}^{N}\sum_{S_v}\sum_{\kappa=0}^{v}\widetilde{\sum}}_{N-v+l\neq 0}e^{-\psi\rho_c\frac{1-\tau}{\tau}}\left[1-\left(\frac{\theta_n\psi\Phi}{Mm-1}-\psi\rho_c\right)\frac{1-\tau}{\tau}\right] && \text{(5.38b)} \\
\simeq 1 &+ \underbrace{\sum_{v=n}^{N}\sum_{S_v}\sum_{\kappa=0}^{v}\widetilde{\sum}}_{N-v+l\neq 0}e^{-\frac{\theta_n\psi\Phi(1-\tau)}{(Mm-1)\tau}} && \text{(5.38c)} \\
= &\sum_{v=n}^{N}\sum_{S_v}\prod_{l=1}^{v}F_{|g_{j_l}|^2}\left(\frac{\theta_n\psi(1-\tau)}{(Mm-1)\tau}\right)\prod_{l=v+1}^{N}\left[1-F_{|g_{j_l}|^2}\left(\frac{\theta_n\psi(1-\tau)}{(Mm-1)\tau}\right)\right] && \text{(5.38d)} \\
\simeq &\sum_{S_n}\prod_{l=1}^{n}\frac{\theta_n\psi(1-\tau)}{(Mm-1)\Omega_{j_l}\tau} = \text{result in (5.8)}, && \text{(5.38e)}
\end{aligned}$$

where (5.38a) is obtained by applying the asymptotic expansion of $K_\omega(\cdot)$ [36, Eq. 9.6.11]

$$K_\omega(x)\overset{x\to 0}{\simeq}\begin{cases}x^{-1},\omega=1,\\ \frac{1}{2}\left(\frac{1}{2}x\right)^{-\omega}\left[\Gamma(\omega)\text{–}\Gamma(\omega-1)\frac{x^2}{4}\right],\omega\geq 2,\end{cases} \tag{5.39}$$

into (5.7), (5.38b) is obtained by expanding the summation $\sum_{\mu=1}^{Mm-1}$ and ignoring the higher order terms, (5.38c) is derived by using the Taylor series $1-x \overset{x\to 0}{\simeq} e^{-x}$, and (5.38e) is realized by using the Taylor series $e^{-x} \overset{x\to 0}{\simeq} 1-x$ and ignoring the higher order terms.

5.8.3 Proof of Lemma 5

Before proving *Lemma* 5, we first propose Lemma 7 and re-calculates $\mathrm{P}_n^{\mathrm{N}}(\tau,\alpha)$.

Lemma 7. *Let $f(x,y)$ be defined on the open convex set $C \subset \mathbb{R}^2$. If $f(x,y)$ is a positive differentiable and pseudoconcave (pseudoconvex) function of x, then the integration on y preserves the pseudoconcavity (pseudoconvexity) of $f(x,y)$ w.r.t. x.*

Proof. To prove that $\int f(x,y)\,dy$ is a pseudoconcave function of x, we need to show that for $x_1, x_2 \in C$, $\int f(x_1,y)\,dy > \int f(x_2,y)\,dy$ implies that $(x_1-x_2)\int \frac{\partial f(x_2,y)}{\partial x}dy > 0$.

Suppose $\forall x_1, x_2 \in C$ such that $\int f(x_1,y)\,dy > \int f(x_2,y)\,dy$, which indicates that $f(x_1,y) > f(x_2,y)$ since $f(x,y)$ is a positive function. According to the definition of pseudoconcave function, $f(x_1,y) > f(x_2,y)$ implies that $(x_1-x_2)\frac{\partial f(x_2,y)}{\partial x} > 0$, i.e., $(x_1-x_2)\int \frac{\partial f(x_2,y)}{\partial x}dy > 0$, which completes the proof. Similarly, we can also show that the integration preserves the pseudoconvexity of $f(x,y)$ w.r.t. x. □

From (5.31), the outage probability of the n-th user can be re-expressed as

$$\begin{aligned}
\mathrm{P}_n^{\mathrm{N}}(\tau,\alpha) &= F_X(0) + \int_0^{\infty}\left(F_X\left(\frac{\theta_n}{y}\right) - F_X(0)\right) f_{|\tilde{g}_n|^2}(y)\,dy \\
&= \int_0^{\infty} F_X\left(\frac{\theta_n}{y}\right) f_{|\tilde{g}_n|^2}(y)\,dy \\
&= \int_0^{\infty} F_{\|\mathbf{h}\|^2}\left(\frac{\xi_n\Omega_0(1-\tau)}{m\tau}\right) f_{|\tilde{g}_n|^2}(y)\,dy \\
&= \int_0^{\infty} \gamma\left(Mm, \xi_n\frac{1-\tau}{\tau}\right)\frac{f_{|\tilde{g}_n|^2}(y)}{\Gamma(Mm)}dy, \qquad (5.40)
\end{aligned}$$

where $\xi_n = \left(\frac{\theta_n}{y}+\rho_c\right)\frac{m}{\eta\rho_b\Omega_0}$. Note that (5.40) is obtained by invoking [52, Eq. 3.351.1] to derive the CDF of $F_{\|\mathbf{h}\|^2}\left(\frac{\xi_n\Omega_0(1-\tau)}{m\tau}\right)$.

Now we are ready for showing the pseudoconcavity of $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)$ w.r.t. τ. Based

on (5.40), the average throughput of the n-th user can be expressed as

$$\begin{aligned}\bar{R}_n^{\mathrm{N}}(\tau,\alpha) &= R_t(1-\tau)\left[1-\int_0^{\infty}\gamma\left(Mm,\xi_n\frac{1-\tau}{\tau}\right)\frac{f_{|\tilde{g}_n|^2}(y)}{\Gamma(Mm)}dy\right]\\ &= R_t\int_0^{\infty}(1-\tau)\left[\Gamma(Mm)-\gamma\left(Mm,\xi_n\frac{1-\tau}{\tau}\right)\right]\frac{f_{|\tilde{g}_n|^2}(y)}{\Gamma(Mm)}dy\\ &= R_t\int_0^{\infty}\underbrace{(1-\tau)\,\Gamma\left(Mm,\xi_n\frac{1-\tau}{\tau}\right)}_{f_n(\tau)}\frac{f_{|\tilde{g}_n|^2}(y)}{\Gamma(Mm)}dy. \end{aligned} \tag{5.41}$$

Taking the second-order derivative of $f_n(\tau)$ w.r.t. τ and after some basic algebraic manipulations, we have

$$f_n''(\tau)=\frac{\xi_n^{Mm}(1-\tau)^{Mm-1}e^{-\xi_n\frac{1-\tau}{\tau}}}{\tau^{Mm+3}\left[\xi_n(1-\tau)-(Mm+1)\tau\right]^{-1}}. \tag{5.42}$$

(5.42) shows that $f_n''(\tau)$ is positive and negative in intervals $\tau\in\left(0,\frac{\xi_n}{\xi_n+Mm+1}\right]$ and $\tau\in\left[\frac{\xi_n}{\xi_n+Mm+1},1\right)$, respectively. Combing this observation with the fact that differentiable convex is pseudoconvex and differentiable concave is pseudoconcave [91], we can infer that $f_n(\tau)$ is a pseudoconvex function and a pseudoconcave function of τ for $\tau\in\left(0,\frac{\xi_n}{\xi_n+Mm+1}\right]$ and $\tau\in\left[\frac{\xi_n}{\xi_n+Mm+1},1\right)$, respectively. Combining this result with Lemma 7, we know that there exists a value $\hat{\tau}_n$ such that $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)$ is respectively a pseudoconvex function and a pseudoconcave function of τ for $\tau\in(0,\hat{\tau}_n]$ and $\tau\in[\hat{\tau}_n,1)$. Note that $\bar{R}_n^{\mathrm{N}}(0,\alpha)=0$, which indicates that $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)$ is both pseudoconvex and pseudoconcave, i.e., pseudolinear, w.r.t. $\tau\in(0,\hat{\tau}_n]$. Based on the above discussions, we prove that $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)$ is a pseudoconcave function of τ. Furthermore, as $\bar{R}_n^{\mathrm{N}}(1,\alpha)=0$, it is clear that there exists a critical point denoted by τ_n^* which is the solution of $\frac{\partial\bar{R}_n^{\mathrm{N}}(\tau,\alpha)}{\partial\tau}=0$. According to [91, Theorem 3.39], we know that $\bar{R}_n^{\mathrm{N}}(\tau,\alpha)$ can be maximized when $\tau_n^*=\bar{\tau}_n$.

5.8.4 Proof of Proposition 12

The achievable rate of x_n for the OBT protocol is given by $R_{x_n}^{\mathrm{O}}=\log_2\left(1+\frac{P_s^{\mathrm{O}}|\tilde{g}_n|^2}{\sigma^2}\right)$, where P_s^{O} denotes the transmit power of S in the OBT protocol. By replacing τ in (5.2) with β_1, P_s^{O} can be expressed as $P_s^{\mathrm{O}}=\left[\frac{\eta P_b\|\mathbf{h}\|^2\beta_1}{1-\beta_1}-P_c\right]^+$. Based on the expressions of $R_{x_n}^{\mathrm{O}}$ and P_s^{O}, the outage probability of the n-th user for the OBT

protocol can be calculated as

$$\begin{aligned}\mathrm{P}_n^{\mathrm{O}}(\beta_1) &= \Pr\left\{P_s^{\mathrm{O}}=0\right\}+\Pr\left\{R_{x_n}^{\mathrm{O}}<2R_t, P_s^{\mathrm{O}}>0\right\}\\ &= \Pr\left\{X'\leq 0\right\}+\Pr\left\{X'|\tilde{g}_n|^2<\gamma'_{\mathrm{th}}, X'>0\right\}\\ &= F_{X'}(0)+\int_0^{\infty} F_{|\tilde{g}_n|^2}\left(\frac{\gamma'_{\mathrm{th}}}{x}\right) f_{X'}(x)\,dx, \end{aligned} \tag{5.43}$$

where $X' \triangleq \frac{\eta\beta_1\rho_b\|\mathbf{h}\|^2}{1-\beta_1}-\rho_c$. Note that (5.43) is quite similar to (5.31), and it can be derived by replacing τ and θ_n in (5.31) with β_1 and γ'_{th}. Therefore, the calculation of $\mathrm{P}_n^{\mathrm{O}}(\beta_1)$ can follow the calculation of $\mathrm{P}_n^{\mathrm{N}}(\tau,\alpha)$ shown in Appendix A. Finally, the closed-form expression of $\mathrm{P}_n^{\mathrm{O}}(\beta_1)$ can be derived by replacing τ and θ_n in $\mathrm{P}_n^{\mathrm{N}}(\tau,\alpha)$ (given in (5.7)) with β_1 and γ'_{th}.

5.8.5 Proof of Proposition 14

According to (5.30), $\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1,\beta_2^{\star})$ can be re-expressed as

$$\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1,\beta_2^{\star}) = 2R_t\frac{1-\beta_1}{\frac{1}{1-\mathrm{P}_p^{\mathrm{O}}(\beta_1)}+\frac{1}{1-\mathrm{P}_q^{\mathrm{O}}(\beta_1)}}. \tag{5.44}$$

To show the log-concavity of $\bar{R}_{\mathrm{F}}^{\mathrm{O}}(\beta_1,\beta_2^{\star})$, we first study the property of $1-\mathrm{P}_n^{\mathrm{O}}(\beta_1)$ $(n=p,q)$ in β_1. By replacing τ and θ_n in (5.40) with β_1 and γ'_{th}, $1-\mathrm{P}_n^{\mathrm{O}}(\beta_1)$ can be expressed as

$$1-\mathrm{P}_n^{\mathrm{O}}(\beta_1) = \int_0^{\infty}\Gamma\left(Mm,\xi'\frac{1-\beta_1}{\beta_1}\right)\frac{f_{|\tilde{g}_n|^2}(y)}{\Gamma(Mm)}dy. \tag{5.45}$$

where $\xi'=\left(\frac{\gamma'_{\mathrm{th}}}{y}+\rho_c\right)\frac{m}{\eta\rho_b\Omega_0}$.

Two important properties are first given: 1) $\Gamma(a,b)$ is log-concave w.r.t. b when $a\geq 1$ [100, Theorem 1]. 2) A positive non-increasing log-concave transformation of a positive convex function is log-concave [101, Lemma 4]. Invoking these two properties with the following facts that $Mm\geq 1$, $\Gamma(a,b)$ is a decreasing function of b, and $\xi'\frac{1-\beta_1}{\beta_1}$ is a positive and convex function of β_1, we can tell that $\Gamma\left(Mm,\xi'\frac{1-\beta_1}{\beta_1}\right)$ is a log-concave function of β_1. Since integration operation preserves the log-concavity [69], we further know that $1-\mathrm{P}_n^{\mathrm{O}}(\beta_1)$ is a log-concave function of β_1.

Recall that the inverse of a log-concave (log-convex) function is a log-convex (log-concave) function [69], and the summation operation preserves the log-convexity, we can infer that $\left(\frac{1}{1-\mathrm{P}_p^{\mathrm{O}}(\beta_1)}+\frac{1}{1-\mathrm{P}_q^{\mathrm{O}}(\beta_1)}\right)^{-1}$ is a log-concave function of β_1. In addition, it

is easy to prove that $1-\beta_1$ is a log-concave function of β_1. Since the product of two log-concave functions is a log-concave function [69], we finally arrive the conclusion that $\bar{R}_{\mathrm{F}}^{\mathrm{O}}\left(\beta_1, \beta_2^{\star}\right)$ is a log-concave function of β_1.

Chapter 6

Conclusions and Future Research

6.1 Conclusions

This thesis intends to solve some open problems resulted from the application of NOMA in cooperative networks, cognitive radio networks, and wireless powered networks.

In Chapter 2, to enjoy the benefits brought by the cooperation for NOMA-based systems while avoiding the possible loss of spectral efficiency caused by the fixed relaying, we propose an incremental cooperative NOMA protocol. In this protocol, the weak user sends a 1-bit feedback to indicate the strong user whether the cooperation is necessary or not. We theoretically prove that the proposed protocol is superior to the conventional one in terms of each user's outage probability and the system outage probability.

Chapter 3 studies the joint optimization problems of power allocation and user ordering for overlay cognitive NOMA networks. When full CSIT is available, we derive the joint optimal solution in closed-form, by which the outage probability of the secondary system is minimized while achieving the target data rate of the primary system. When partial CSIT is available, the joint optimal solution is obtained in semi-closed form, by which the outage probability of the secondary system is minimized while ensuring that the outage probability of the primary system is below or equal to a predetermined threshold.

To further improve the performance of the cooperative SWIPT NOMA networks, in Chapter 4, we introduce the more general and more powerful H-SWIPT scheme into the cooperative NOMA networks and discuss two optimal resource allocation policies.

When full CSIT is available, we intend to realize a successful transmission with the minimum energy consumption at the source. Although the formulated problem is non-convex, we successfully obtain the joint optimal solution with an efficient bisection method. When only partial CSIT is available, we aim to minimize the system outage probability. In this case, we show that the joint optimal solution can be found with a 1-D search.

Chapter 5 studies a fair resource allocation problem in a PB-assisted downlink NOMA network. Our goal is to use the statistical CSI to jointly optimize the time allocation and power allocation ratios, finally maximizing the worse throughput of the two paired users. Since the formulated problem is non-convex, we decompose the original problem into two subproblems. After solving these two subproblems with linear search methods, we design an alternating optimization algorithm to approximate the joint optimal solution.

6.2 Future Research

6.2.1 NOMA Networks with Joint User and Relay Cooperation

Various works including our two works presented in Chapters 2 and 4 have been done to study the cooperative NOMA networks. However, none of them has investigated the cooperative NOMA networks where the user cooperation and the dedicated relay cooperation can be jointly applied to maximize the efficacy of cooperative communications. It is worth mentioning that the combination of these two types of cooperation is practical and reasonable, since dedicated relays and direct links between the source and strong users can exist simultaneously [102–104]. More importantly, the diversity gain is expected to be increased by combining these two different cooperation mechanisms, which consequently enhances the overall system performance. Based on these two key observations, it is necessary to study the performance of cooperative NOMA networks integrating with the user cooperation and the dedicated relay cooperation.

6.2.2 Wireless Powered NOMA Networks with Energy Accumulation

In Chapters 4 and 5, we respectively investigate the performance of SWIPT-assisted and PB-assisted NOMA networks where the energy-constrained nodes are assumed

to exhaust the collected energy during the subsequent forwarding phase. Obviously, this strategy underutilizes the harvested energy. For example, when the channel conditions of the forwarding links are good, the energy-constrained nodes can use only part of the harvested energy to achieve successful transmission and save the remaining harvested energy for future use. Therefore, if the energy-constrained nodes have the energy storage capability and can adaptively adjust its energy use strategy according to the channel conditions, system performance is expected to be further improved. To model the state evolution of the EH-based nodes' battery, we can employ a finite-state Markov Chain. Note that there have been some preliminary results on wireless powered OMA networks with energy accumulation [105–113], but to our best of knowledge, there have been no reported works on wireless powered NOMA networks with energy accumulation.

6.2.3 NOMA-Based Systems with Imperfect CSI

All of the works presented in this thesis are processed based on the assumption that the receiver can perfectly access the CSI. However, in practice, perfect CSI can only be achieved with a very high system overhead, thus it is very likely that channel estimation error occurs. In view of this fact, it is necessary to investigate the impact of imperfect CSI on the NOMA-based systems.

Bibliography

[1] "Cisco visual networking index: forecast and trends, 2017–2022," Cisco, white paper, Nov. 2018.

[2] S. Hong, J. Brand, J. I. Choi *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, Feb. 2014.

[3] E. G. Larsson, O. Edfors, F. Tufvesson *et al.*, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] L. Dai, B. Wang, Y. Yuan *et al.*, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sept. 2015.

[5] M. AL-Imari, M. A. Imran, R. Tafazolli *et al.*, "Performance evaluation of low density spreading multiple access," in *Proc. IEEE Int. Wireless Commun. and Mobile Computing Conf. (IWCMC)*, Aug. 2012, pp. 383–388.

[6] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sept. 2013, pp. 332–336.

[7] D. Fang, Y. Huang, Z. Ding *et al.*, "Lattice partition multiple access: A new method of downlink non-orthogonal multiuser transmissions," in *Proc. IEEE Global Commun. Conf*, Washington D. C., US, Dec. 2016.

[8] S. Chen, B. Ren, Q. Gao *et al.*, "Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3185–3196, Apr. 2017.

[9] L. Dai, B. Wang, Z. Ding *et al.*, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tutorials.*, vol. 20, no. 3, pp. 2294–2323, Third quarter 2018.

[10] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[11] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[12] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wireless Commun.*, vol. 12, no. 2, pp. 19–29, 2005.

[13] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 5, pp. 796–807, 1994.

[14] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 364–375, 1996.

[15] Y. Saito, A. Benjebbour, Y. Kishiyama *et al.*, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sept. 2013, pp. 611–615.

[16] D. Tse and P. Viswanath, *Fundamentals of wireless communication.* Cambridge Univ. Press, 2005.

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

[18] J. Men and J. Ge, "Non-orthogonal multiple access for multiple-antenna relaying networks," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1686–1689, Oct. 2015.

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

[20] J. Mitola, G. Q. Maguire *et al.*, "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, 1999.

[21] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, 2005.

[22] Y. Liu, Z. Qin, M. Elkashlan *et al.*, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[23] F. Zhou, Y. Wu, Y. Liang *et al.*, "State of the art, taxonomy, and open issues on cognitive radio networks with NOMA," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 100–108, Apr. 2018.

[24] L. Lv, J. Chen, Q. Ni *et al.*, "Cognitive non-orthogonal multiple access with cooperative relaying: A new wireless frontier for 5G spectrum sharing," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 188–195, Apr. 2018.

[25] A. Goldsmith, S. A. Jafar, I. Maric *et al.*, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[26] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[27] X. Lu, P. Wang, D. Niyato *et al.*, "Wireless networks with RF energy harvesting: a contemporary survey," *IEEE Commun. Surveys Tutorials.*, vol. 17, no. 2, pp. 757–789, Secondquarter 2015.

[28] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.

[29] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[30] K. Huang and V. K. Lau, "Enabling wireless power transfer in cellular networks: architecture, modeling and deployment," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 902–912, Jan. 2014.

[31] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2013.

[32] S. Atapattu and J. Evans, "Optimal energy harvesting protocols for wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5789–5803, Aug. 2016.

[33] Y. Hu, M. C. Gursoy, and A. Schmeink, "Efficient transmission schemes for low-latency networks: NOMA vs. relaying," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*. IEEE, 2017, pp. 1–6.

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

[35] N. T. Do, D. B. D. Costa, T. Q. Duong *et al.*, "A BNBF user selection scheme for NOMA-based cooperative relaying systems with SWIPT," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 664–667, Mar. 2017.

[36] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York, NY, USA: Dover, 1972.

[37] S. Neumark, *Solution of cubic and quartic equations*. Oxford: Pergamon Press, 1965.

[38] Z. Ding, H. Dai, and H. V. Poor, "Relay selection for cooperative NOMA," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 416–419, Aug. 2016.

[39] L. Lv, Q. Ni, Z. Ding *et al.*, "Application of non-orthogonal multiple access in cooperative spectrum-sharing networks over Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5506–5511, Jun. 2017.

[40] L. Lv, L. Yang, H. Jiang *et al.*, "When NOMA meets multiuser cognitive radio: Opportunistic cooperation and user scheduling," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6679–6684, Jul. 2018.

[41] X. Yue, Y. Liu, S. Kang *et al.*, "Spatially random relay selection for full/half-duplex cooperative NOMA networks," *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3294–3308, Aug. 2018.

[42] Z. Ding, L. Dai, and H. V. Poor, "MIMO-NOMA design for small packet transmission in the Internet of Things," *IEEE Access*, vol. 4, pp. 1393–1405, 2016.

[43] Y. Yu, H. Chen, Y. Li *et al.*, "Antenna selection in MIMO cognitive radio-inspired NOMA systems," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2658–2661, Dec. 2017.

[44] Z. Yang, Z. Ding, Y. Wu *et al.*, "Novel relay selection strategies for cooperative NOMA," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10 114–10 123, Nov. 2017.

[45] Y. Yu, H. Chen, Y. Li *et al.*, "Antenna selection for MIMO non-orthogonal multiple access systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3158–3171, Apr. 2018.

[46] Y. Li, Y. Lia, X. Chu *et al.*, "Performance analysis of relay selection in cooperative NOMA networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 760–763, Apr. 2019.

[47] P. Xu and K. Cumanan, "Optimal power allocation scheme for non-orthogonal multiple access with α-fairness," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2357–2369, Oct. 2017.

[48] D. Wan, M. Wen, F. Ji *et al.*, "Cooperative NOMA systems with partial channel state information over Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 947–958, Mar. 2018.

[49] H. Xing, Y. Liu, A. Nallanathan *et al.*, "Optimal throughput fairness tradeoffs for downlink non-orthogonal multiple access over fading channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3556–3571, Jun. 2018.

[50] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, Jan. 2002.

[51] Z. Yang, Z. Ding, P. Fan *et al.*, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 654–667, Feb. 2016.

[52] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Elsevier/Academic Press, 2007.

[53] J. Chen, L. Yang, and M.-S. Alouini, "Physical layer security for cooperative NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4645–4649, May 2018.

[54] X. Liang, Y. Wu, D. W. K. Ng *et al.*, "Outage performance for cooperative NOMA transmission with an AF relay," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2428–2431, Nov. 2017.

[55] Y. Feng, S. Yan, C. Liu *et al.*, "Two-stage relay selection for enhancing physical layer security in non-orthogonal multiple access," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 6, pp. 1670–1683, Jun. 2018.

[56] X. Yue, Y. Liu, S. Kang *et al.*, "Exploiting full/half-duplex user relaying in NOMA systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 560–575, Feb. 2017.

[57] G. Liu, X. Chen, Z. Ding *et al.*, "Hybrid half-duplex/full-duplex cooperative non-orthogonal multiple access with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 506–519, Jan. 2018.

[58] Z. Yang, Z. Ding, P. Fan *et al.*, "The impact of power allocation on cooperative non-orthogonal multiple access networks with SWIPT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4332–4343, Jul. 2017.

[59] Y. Zhang, J. Ge, and E. Serpedin, "Performance analysis of a 5G energy-constrained downlink relaying network with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8333–8346, Dec. 2017.

[60] Y. Xu *et al.*, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, Sept. 2017.

[61] T. N. Do, D. B. da Costa, T. Q. Duong *et al.*, "Improving the performance of cell-edge users in MISO-NOMA systems using TAS and SWIPT-based cooperative transmissions," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 49–62, Mar. 2018.

[62] D.-T. Do, M. Vaezi, and T.-L. Nguyen, "Wireless powered cooperative relaying using NOMA with imperfect CSI," in *Proc. IEEE Global Commun. Conf.*, Abu Dhabi, UAE, Dec. 2018.

[63] L. Bariah, S. Muhaidat, and A. Al-Dweik, "Error probability analysis of NOMA-based relay networks with SWIPT," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1223–1226, Jul. 2019.

[64] X. Li, J. Li, and L. Li, "Performance analysis of impaired SWIPT NOMA relaying networks over imperfect Weibull channels," *IEEE Syst. J.*, to be published.

[65] J. Men, J. Ge, and C. Zhang, "Performance analysis of nonorthogonal multiple access for relaying networks over Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1200–1208, Feb. 2017.

[66] Y. Xiao, L. Hao, Z. Ma *et al.*, "Forwarding strategy selection in dual-hop NOMA relaying systems," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1644–1647, Aug. 2018.

[67] H. Liu, Z. Ding, K. J. Kim *et al.*, "Decode-and-forward relaying for cooperative NOMA systems with direct links," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8077–8093, Dec. 2018.

[68] P. Xu, Z. Yang, Z. Ding *et al.*, "Optimal relay selection schemes for cooperative NOMA," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7851–7855, Aug. 2018.

[69] S. Boyd and L. Vandenberghe, *Convex optimization*. New York, NY, USA: Cambridge University Press, 2004.

[70] L. Yang, J. Chen, Q. Ni *et al.*, "NOMA-enabled cooperative unicast-multicast: design and outage analysis," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7870–7889, Dec. 2017.

[71] K. Huang and X. Zhou, "Cutting the last wires for mobile communications by microwave power transfer," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 86–93, Jun. 2015.

[72] Y. Ma, H. Chen, Z. Lin *et al.*, "Distributed and optimal resource allocation for power beacon-assisted wireless-powered communications," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3569–3583, Oct. 2015.

[73] C. Zhong, X. Chen, Z. Zhang *et al.*, "Wireless-powered communications: performance analysis and optimization," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5178–5190, Dec. 2015.

[74] X. Jiang, C. Zhong, X. Chen *et al.*, "Secrecy performance of wirelessly powered wiretap channels," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3858–3871, Sept. 2016.

[75] X. Jiang, C. Zhong, Z. Zhang *et al.*, "Power beacon assisted wiretap channels with jamming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8353–8367, Dec. 2016.

[76] Z. Chen, L. Hadley, Z. Ding *et al.*, "Improving secrecy performance of a wirelessly powered network," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4996–5008, Nov. 2017.

[77] F. Zhao, H. Lin, C. Zhong *et al.*, "On the capacity of wireless powered communication systems over Rician fading channels," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 404–417, Jan. 2018.

[78] X. Zhou, J. Guo, S. Durrani *et al.*, "Power beacon-assisted millimeter wave ad hoc networks," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 830–844, Feb. 2018.

[79] N. P. Le, "Throughput analysis of power-beacon-assisted energy harvesting wireless systems over non-identical Nakagami-m fading channels," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 840–843, Apr. 2018.

[80] P. Deng, B. Wang, W. Wu *et al.*, "Transmitter design in MISO-NOMA system with wireless-power supply," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 844–847, Apr. 2018.

[81] Q. Wu, W. Chen, D. W. K. Ng *et al.*, "Spectral and energy-efficient wireless powered IoT networks: NOMA or TDMA?" *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6663–6667, Jul. 2018.

[82] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[83] Z. Ding, Z. Yang, P. Fan *et al.*, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[84] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[85] Y. Zeng and R. Zhang, "Optimized training design for wireless energy transfer," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 536–550, Feb. 2015.

[86] D. Mishra, S. De, and C. F. Chiasserini, "Joint optimization schemes for cooperative wireless information and power transfer over Rician channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 554–571, Feb. 2016.

[87] M. Nakagami, "The m-distribution, a general formula of intensity distribution of rapid fading," *Statistical Method of Radio Propagation*, pp. 3–34, 1960.

[88] A. A. Nasir, X. Zhou, S. Durrani *et al.*, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.

[89] P. Xu, Y. Yuan, Z. Ding *et al.*, "On the outage performance of non-orthogonal multiple access with 1-bit feedback," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6716–6730, Oct. 2016.

[90] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[91] M. Avriel, W. E. Diewert, S. Schaible *et al.*, *Generalized concavity*. Philadelphia, PA, USA: SIAM, 2010.

[92] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: theory and algorithms*. Hoboken, NJ, USA: Wiley, 2006.

[93] A. D. Belegundu and T. R. Chandrupatla, *Optimization concepts and applications in engineering*. Cambridge University Press, 2011.

[94] H. M. Wang, C. Wang, and D. W. K. Ng, "Artificial noise assisted secure transmission under training and feedback," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6285–6298, Dec. 2015.

[95] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[96] S. Lee and R. Zhang, "Distributed wireless power transfer with energy feedback," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1685–1699, Apr. 2017.

[97] P. D. Diamantoulakis, K. N. Pappi, Z. Ding *et al.*, "Wireless-powered communications with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8422–8436, Dec. 2016.

[98] H. Chingoska, Z. Hadzi-Velkov, I. Nikoloska *et al.*, "Resource allocation in wireless powered communication networks with non-orthogonal multiple access," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 684–687, Dec. 2016.

[99] H. A. David and N. Nagaraja, *Order Statistics, 3rd ed.* John Wiley & Sons, Inc., 2003.

[100] Á. Baricz, "Geometrically concave univariate distributions," *Journal of Mathematical Analysis and Applications*, vol. 363, no. 1, pp. 182–196, 2010.

[101] D. Mishra and S. De, "i^{2}RES: integrated information relay and energy supply assisted RF harvesting communication," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1274–1288, Mar. 2017.

[102] J.-B. Kim and I.-H. Lee, "Non-orthogonal multiple access in coordinated direct and relay transmission," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 2037–2040, Nov. 2015.

[103] C. Zhong and Z. Zhang, "Non-orthogonal multiple access with cooperative full-duplex relaying," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2478–2481, Dec. 2016.

[104] Z. Yu, C. Zhai, J. Liu *et al.*, "Cooperative relaying based non-orthogonal multiple access (NOMA) with relay selection," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 606–11 618, Dec. 2018.

[105] I. Krikidis, S. Timotheou, and S. Sasaki, "RF energy transfer for cooperative networks: data relaying or energy harvesting?" *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1772–1775, Nov. 2012.

[106] K. H. Liu, "Selection cooperation using RF energy harvesting relays with finite energy buffer," in *Proc. IEEE WCNC*, Apr. 2014, pp. 2156–2161.

[107] ——, "Performance analysis of relay selection for cooperative relays based on wireless power transfer with finite energy storage," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5110–5121, Jul. 2016.

[108] Y. Gu, H. Chen, Y. Li *et al.*, "Distributed multi-relay selection in accumulate-then-forward energy harvesting relay networks," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 74–86, Mar. 2018.

[109] I. Krikidis, "Relay selection in wireless powered cooperative networks with energy storage," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2596–2610, Dec. 2015.

[110] Z. Zhou, M. Peng, Z. Zhao *et al.*, "Wireless-powered cooperative communications: power-splitting relaying with energy accumulation," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 969–982, Apr. 2016.

[111] Z. Li, H. Chen, Y. Li *et al.*, "Incremental accumulate-then-forward relaying in wireless energy harvesting cooperative networks," in *Proc. IEEE GLOBECOM*, Dec. 2016, pp. 1–6.

[112] Y. Bi and H. Chen, "Accumulate and jam: Towards secure communication via a wireless-powered full-duplex jammer," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 8, pp. 1538–1550, Dec. 2016.

[113] G. Li and H. Jiang, "Performance analysis of wireless powered incremental relaying networks with an adaptive harvest-store-use strategy," *IEEE Access*, vol. 6, pp. 48 531–48 542, 2018.