Title
On the Fundamentals of Power Allocation Strategies for Non-Orthogonal Multiple Access Downlink Systems

Author
Oviedo, Jose Armando

Publication Date
2020

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ON THE FUNDAMENTALS OF POWER ALLOCATION STRATEGIES FOR NON-ORTHOGONAL MULTIPLE ACCESS DOWNLINK SYSTEMS

A dissertation submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL AND COMPUTER ENGINEERING

by

José Armando Oviedo

March 2020

The Dissertation of José Armando Oviedo is approved:

Hamid R. Sadjadpour, Chair

Leila Parsa

Ahmet Ali Yanik

Quentin Williams
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

On the Fundamentals of Power Allocation Strategies for Non-Orthogonal Multiple Access Downlink Systems

by

José Armando Oviedo

The fundamentals of non-orthogonal multiple access (NOMA) power allocation strategies for downlink wireless transmissions are investigated, based on improving the performance over downlink transmissions using orthogonal multiple access (OMA). First, the scenario where the base-station possesses the perfect channel state information (CSI) is considered. For this scenario, the power allocation region is defined for a two-user downlink NOMA system. The expressions for the ergodic capacities and outage probabilities are derived for the case when the two users are pair randomly, each user with independently identically distributed (i.i.d.) fading channel gains. These results are then extended to the case where a cell-center user and cell-edge user are paired together, and a closed-form expression for the gap in ergodic capacity between NOMA and OMA is derived when SNR is large. This scenario is then extended to the general multi-user case, and it is proved that there always exists a power allocation strategy which allows all users to achieve a higher capacity when compared to OMA. In the second scenario, the effects of users with cached files on power allocation are investigated. When a user with weak channel condition has cached a file which a user with strong channel condition is requesting, these two users downlink transmissions are paired. An approximation of the optimum power allocation is derived, and the union-outage probability of this system is shown to improve over conventional NOMA and OMA. In the final scenario, the complete description of power allocation strategies is derived for more realistic wireless systems, where it is not assumed that the base-station possesses perfect CSI. Based on the target rates of each user, the fundamental properties of the power allocation strategy are derived, and the approach for selecting a strategy which improves the outage probability of each user is outlined. The existence of the power allocation strategies and their optimal energy efficiency is proved to be functions of the target rates and the OMA system parameters which they are compared to.

Dedicated to my mother, step-father, and older brother, who sacrificed so much to create this reality for me at a time when it seemed so bleak, and to my father, who would be immensely proud of this accomplishment but left me too soon.

Acknowledgments

There are too many people to thank who helped me along this rough journey. From my UCSC colleagues, instructors, and friends, to the many coworkers and managers that were always supportive of me and my path, their impacts have all contributed to my growth, of which this dissertation is a direct result.

First and foremost, I want to thank my advisor Prof. Hamid Sadjadpour for suggesting this topic to me at a time when my ideas were stagnant, and for his patience with my progress.

I would also like to thank my committee members Prof. Yanik and Prof. Parsa for their helpful comments in improving this dissertation. I would also like to thank my former lab-mates Mohsen Karimzadeh and Bita Azimdoost for setting the bar high, and showing me the hard work that is required to finish a doctorate.

Last, but definitely not least, I must thank my partner Lilinoe Harbottle, who has stood by my side through all of the difficult times and sleepless nights, and helped me maintain the strength and focus required to get through all of the difficult stages I endured in order to achieve this goal.

Punk rock till I die!

Chapter 1

Introduction

Non-orthogonal multiple access (NOMA) is a multiple access technique that has been considered for future wireless networks for improving several system performance metrics when compared to orthogonal multiple access (OMA). These performance metrics include the overall number of users served simultaneously in a cell, the spectral efficiency, sum-rate capacity, the outage probability, and energy efficieny, among others [1–4]. By transmitting users' downlink signals using superposition coding (SC) over the same time-frequency resources, and successive interference cancellation (SIC) at the users' receivers, the capacity of the channel can be achieved [6]. Superposition coding is when different signals are transmitted simultaneously over the same time-frequency resources, with each signal being allocated a portion of the total transmit power, allowing the base-station to multiplex the users over the power-domain (PD). At each user's receiver, the signals are received and are completely overlapping over the shared time-frequency resources, thus SIC is used to detect and decode the signals in order of the largest transmit power to the lowest [6]. Thus, the relationship between the power allocated to the different signals transmitted using SC will affect the ability of a SIC enabled receiver to successfully decode the received signals and obtain its own information.

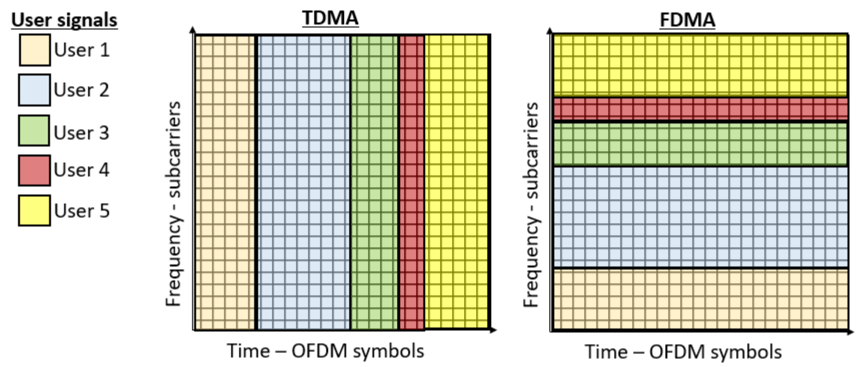

Figure 1.1: TDMA (left) and FDMA (right) orthogonal resource allocation for a BS serving five users over a shared time-frequency resource.

1.1 Introduction to Multiple Access Techniques

OMA is defined as a multiple access approach that schedules multiple users in non-overlapping time slots, i.e. time-division multiple access (TDMA), or frequency bands, i.e. frequency-division multiple access (FDMA), during a transmission time period. Let the signals scheduled for K users be transmitted by a base-station over a transmission time period T and bandwidth B. Therefore each user's signal transmission is assigned either an orthogonal fraction of the transmission time period T, or an orthogonal fraction of the bandwidth B. One difference between TDMA and FDMA is that with TDMA, each transmission can be allocated all of the transmit power for that time duration, while with FDMA the transmit power must be shared between all of the orthogonal frequency bands, and hence FDMA has the additional complexity of a power allocation problem [7]. Both TDMA and FDMA concepts are illustrated in figure 1.1

NOMA on the other hand schedules the transmission of the K users' signals simultaneously over the entire transmission period T and bandwidth B. Since the total transmit power must be shared between the K users, a fraction $a_k \in (0, 1)$ of the transmit power is allocated to user k, and $\sum_{k=1}^{K} a_k \leq 1$. Therefore, when

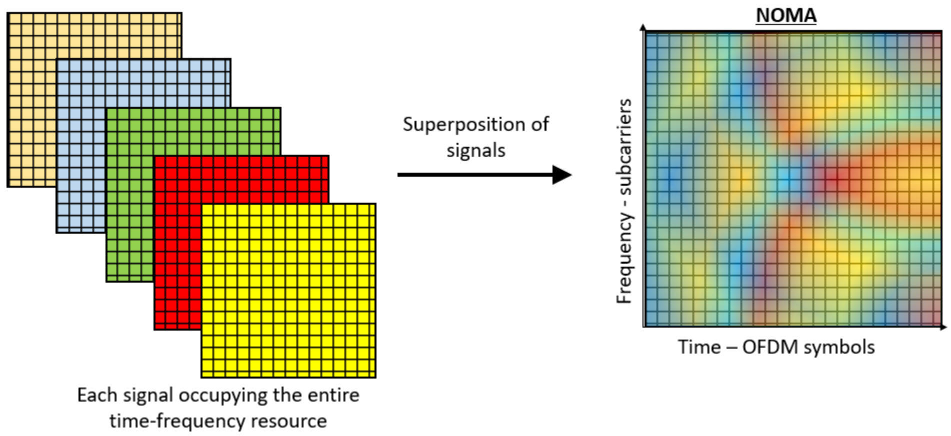

Figure 1.2: Five signals, each shown occupying the entire time-frequency resource (left), superposed over the time-frequency resource (right).

the base-station uses SC at the transmitter to transmit K users' signals over the shared time-frequency resource, and and users equipped with SIC enabled receivers are used to obtain their respective desired signals. This concept of superposing multiple signals over a shared time-frequency resource is illustrated in figure 1.2, where 5 users each have their information mapped over the entire time-frequency resource, causing every signal to be overlaid on the same frequencies at the same time.

1.2 NOMA System Model Overview

A BS which transmits K downlink signals to K users using NOMA, will allocate the transmit power a fraction of the transmit power a_k to each signal, and then will take the sum of these signals before transmitting the superposition of the signals to all K users. As an example given in figure 1.3 for $K = 2$ and 4-point quadrature amplitude modulated (4-QAM) complex symbols, the superposition of the two signals form a composite constellation on the complex plane. This is illustrated in figure 1.3 for the case. For each constellation point of user m's

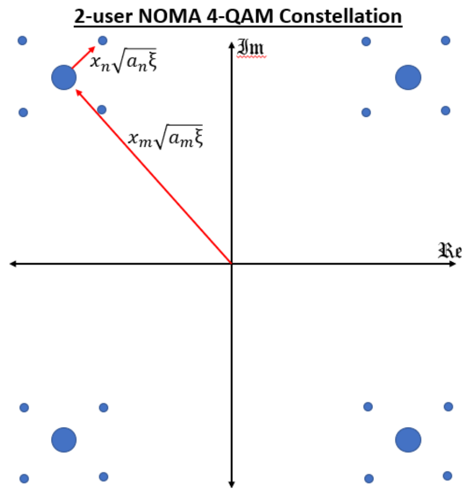

Figure 1.3: Example of a superposition of constellation points from two users' signals.

signal, there are 4 constellation points for user n's signal, forming the 16 superposition constellation points represented by the smaller circles. Note how the smaller constellation points seem like small interference points of the larger constellation points. This disparity in amplitudes of the SC signal is a key factor in being able to perform SIC at the receiver.

In figure 1.4, a block diagram representation of a DL NOMA system is shown. On the left, the BS's SC procedure is illustrated, where the K signals share the total transmit power and added together. The signal travels through each users wireless channel, with random complex valued gain G_k. Each receiver is enabled with SIC capability in order for each user to have the capability to extract its own signal.

Figure 1.5 illustrates the procedure used by an SIC enabled receiver, where the signals received with the highest transmit power are detected and decoded first. The SC signal is sent through a detector, which determines the symbols received, and the symbols are then used to obtain the information message of the respective signal. If the desired information message w_n is not obtained by

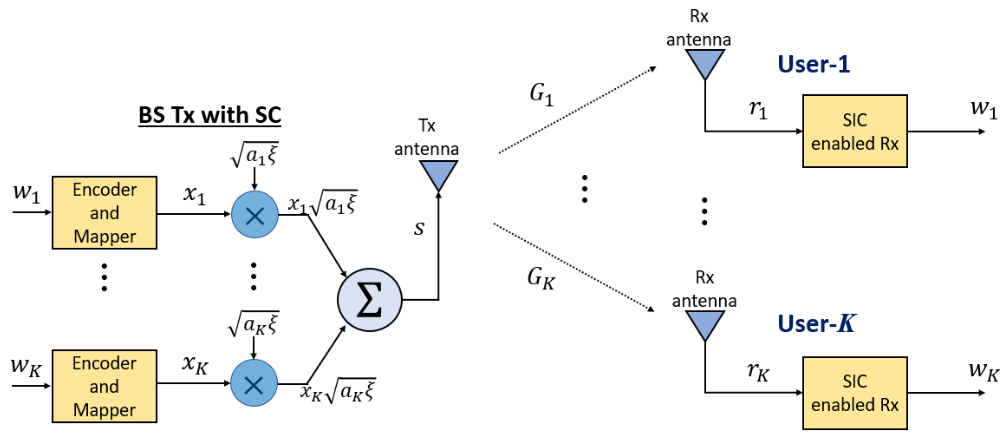

Figure 1.4: A downlink NOMA system where the BS transmits the signals using SC, and each user's receiver is enabled with SIC.

user n, the decoded message is sent the SIC operation block, which will subtract the signal associated with w_n from the SC signal. These steps are repeated until w_n is obtained at user n. The details of the SIC operation, which include re-encoding of the decoded information message, channel gain and power scaling, and subtraction from the superposed signal, are shown in figure 1.6. A signal which is not desired but is decoded is subtracted from the composite signal, and a new composite signal which contains the superposition of the remaining signals is the output of the SIC operation. An example of the symbol constellation point detection of user n is illustrated in figure 1.7. As can be seen from figure, user n must perform SIC in order to remove the interference cause by the signal of user m. First user n detects the constellation point in the upper-left quadrant as the symbol which represents the signal for user m. Then after performing SIC to remove this interference arrives at a new constellation, from which user n then detects its own symbol as the constellation point in the upper-right quadrant.

The details this section are shown in a generalized context. In the subsequent chapters, the specific channel models used in the studies are given in details, along with the assumptions related to the results. However, certain aspects mentioned

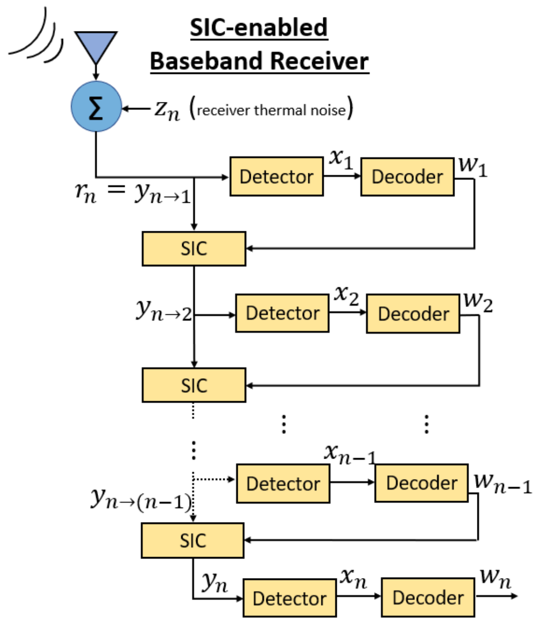

Figure 1.5: Block diagram of receiver with SIC capability.

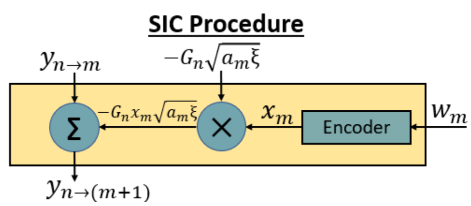

Figure 1.6: Codeword re-encoding, amplification, and subtraction steps of SIC.

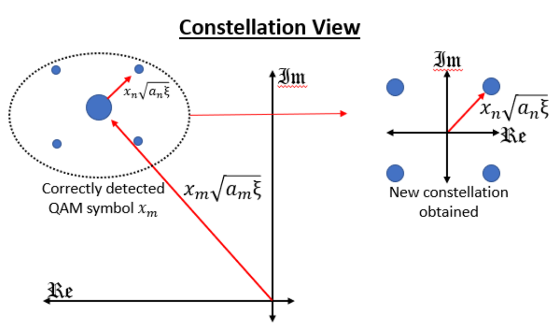

Figure 1.7

in this section are considered out of the scope of this study, which include signal encoding, mapping, detection, decoding, modulation, and thus these are considered to be performed ideally in this study.

1.3 Some initial motivating works

Non-orthogonal access approaches using SC for future wireless cellular networks were mentioned in [5] as a way to increase single user rates when compared to CDMA. Schaepperle and Ruegg [9] evaluated the performance of non-orthogonal signaling using SC and SIC in single antenna OFDMA systems using very little modifications to the existing standards, as well as how user pairing impacts the throughput of the system when the channel gains become increasingly disparate. This was then applied [10] to OFDMA wireless systems to evaluate the performance of cell edge user rates, proposing an algorithm that attempts to increase the average throughput and maintain fairness. These works do not assume to have the exact channel state information at the transmitter.

The concept of NOMA is evaluated through simulation for full channel state information at the transmitter (CSIT) in the uplink [11] and downlink [12], where the throughput of the system is shown to be on average always better than OMA when considering a fully defined cellular system evaluation, with both users occupying all of the bandwidth and time, and was compared to FDMA with each user being assigned an orthogonal channel. In [13], the downlink system performance throughput gains are evaluated by incorporating a complete simulation of an LTE cellular system (3GPP).

1.4 Overview of contained contributions

In this work, it is fundamentally established that the power allocated to the different users' signals depends on a few system parameters. In chapter 2, a NOMA approach called Fair-NOMA is introduced, where the fundamental relationship between the group of K NOMA users' channel gains and the power allocation strategy is described for a downlink NOMA system in which each user has greater channel capacity than the OMA case. The power allocation region, sum-rate capacity improvement of NOMA over OMA, and outage probability are described in closed-form for the case when the transmit signal-to-noise ratio is very large. Then the power allocation strategies are described in detail for the general K-user downlink NOMA system to improve each user's channel capacity over OMA.

In chapter 3, the focus shifts to an scheme called cache-aided NOMA (CA-NOMA) which applies NOMA to downlink systems where users at the edge of the network can possess cached files which are requested by other users in the same cell. In this case, the base-station can pair the two users together and transmit their signals using NOMA, such that one user who is requesting a file is paired with another user who possesses a cache of the same file. When both users are assigned a common quality-of-service (QoS) rate R_0, the base-station will transmit the two users' signals using SC, with power allocation designed so that the signal of the user who possesses the cached file is first in the SIC decoding order. It is shown that the cached file combined with NOMA enables both users to perform SIC, thereby enhancing the power allocation region and allowing for more opportunity to improve the union-outage of the system.

In chapter 4, the fundamental relationship between information target rates and power allocation strategies for K-user NOMA systems is investigated. When users have individual QoS target rates associated with their downlink transmissions, the base-station can design the entire power allocation strategy by using the set of target rates. This becomes especially practical since when compared to the majority of work on downlink NOMA power allocation strategies, which rely on the exact channel gain values of all users, this fundamental relationship between target rates and power allocation strategy improves the outage probability of all users compared to OMA, regardless of the channel gain values. Furthermore, the power allocation strategy is easily computed without the use of a complex algorithm, and is proved to always exist. The outage probability events are used to derive this relationship, with outage probabilities being the metric associated with target rates in wireless system deployments. This means that a parameter that is always known to the base-station and its associated performance metric form the fundamental relationship that enable designing power allocation strategies that outperform OMA.

In each of these chapters, a comparison between theoretical results and simulation results are provided in order to further demonstrate the validity of the theoretical results. These comparisons are for specific examples of system parameters, and they help highlight the performance improvement of NOMA over OMA for the different performance metrics considered. In the scope of this work, it is assumed that certain receiver functions, such as signal detection and channel estimation at the receiver are ideal.

Chapter 2

A Fair Power Allocation Approach to Multi-user Downlink NOMA Systems

A NOMA approach that always outperforms OMA called Fair-NOMA is introduced. In Fair-NOMA, each mobile user is allocated its share of the transmit power such that its capacity is always greater than or equal to the capacity that can be achieved using OMA. In other words, the "fairness" is defined by not compromising the performance of any single user with respect to the baseline OMA performance. For any channel gains of the two users, the set of possible power allocation coefficients are derived. For the infimum and supremum of this set, the individual capacity gains and the sum-rate capacity gain are derived. It is shown that the ergodic sum-rate capacity gain approaches 1 bps/Hz when the transmit power increases for the case when pairing two random users with i.i.d. channel gains. The outage probability of this approach is derived and shown to be better than OMA.

The Fair-NOMA approach is applied to the case of pairing a near base-station user and a cell-edge user and the ergodic capacity gap is derived as a function of total number of users in the cell at high SNR. This is then compared to the conventional case of fixed-power NOMA with user-pairing. Finally, Fair-NOMA is extended to K users and it is proven that the capacity can always be improved for each user, while using less than the total transmit power required to achieve OMA capacities per user.

The results in this chapter have been published in [22]

2.1 Previous Work for Motivating Fair-NOMA

The concept of NOMA was shown to achieve the capacity of the channel for a two-user downlink wireless system by Cover and Thomas [6]. The existence of a set of power allocation coefficients that allow all of the participating users to achieve capacity at least as good as OMA was suggested in [7] for the case when the BS possesses full CSI.

Fairness in NOMA systems is addressed in some works. The uplink case in OFDMA systems is addressed in [15] by using an algorithm that attempts to maximize the sum throughput, with respect to OFDMA and power constraints. The fairness is not directly addressed in the problem formulation, but is evaluated using Jain's fairness index. In [16], a proportional fair scheduler and user pair power allocation scheme is used to achieve fairness in time and rate. In [17], fairness is achieved in the max-min sense, where users are paired such that their channel conditions are not too disparate, while the power allocation maximizes the rates for the paired users.

Ding et. al. [18] provide an analysis for fixed-power NOMA (F-NOMA) and cognative radio NOMA (CR-NOMA). In F-NOMA, with a cell that has N total users, it is shown that the probability that NOMA outperforms OMA asymptotically approaches 1. In CR-NOMA, a primary user is allowed all of the time and bandwidth, unless an opportunistic secondary user exists with a stronger channel condition relative to the primary user, such that transmitting both of their signals will not reduce the primary user's SINR below some given threshold. It is shown that the diversity order of the n-th user is equal to the order of the weaker m-th user, leading to the conclusion that this approach benefits from pairing the two users with the strongest channels.

The main contribution of this chapter is to demonstrate that NOMA capacity can fundamentally outperform OMA capacity for each user, regardless of the channel gain values of the users, and to derive exactly the exact power allocation strategy to achieve this, based on their channel gains. Furthermore, the improvements provided by Fair-NOMA over OMA for various performance metrics and scenarios are derived, giving a firm analytical answer as to how much performance improves when Fair-NOMA is used.

In this chapter, it is assumed that the OMA system model used to compare to NOMA is the TDMA model which allocates equal transmit durations to each user during the transmission time period T. This equates to half of the transmission time period for each user in the case of two-user NOMA, while in the general K-user NOMA case each user is allocated $1/K$ of the total transmission time period. Note that these are equivalent to the FDMA cases where each user is assigned equal parts of the bandwidth and total transmit power.

2.2 System Model and Capacity

A BS serving two mobile users during a transmission time period possesses the full knowledge of the CSI, i.e. knows the exact value of the channel gain

$|G_n|$ between the BS transmit antennas and the users' receive antennas. In the discussion that follows, it is assumed that both the BS and users each possess one antenna, i.e. it is a single-input-single-output (SISO) system. This assumption does not affect the outcome of the power allocation strategy itself, but does affect the analyses' results.

The BS transmits to user n a signal x_n over the wireless channel with Rayleigh fading channel gain $G_n \in \mathbb{C}$ with SNR gain p.d.f. $f_{|G|^2}(w) = \frac{1}{\beta}e^{-\frac{w}{\beta}}$, and receiver noise is complex-normal distributed $z_n \sim \mathcal{CN}(0, 1)$.

In the two-user OMA case, each user's signals are allocated half of transmission time period T and transmit SNR ξ. The received signal for each user is $y_n = G_n\sqrt{\xi}x_n + z_n, n = 1, 2$. If $\mathbb{E}[|x_n|^2] = 1$, the information capacity of each user is $C_n^{\text{oma}} = \frac{1}{2}\log_2\left(1 + \xi|G_n|^2\right)$. The sum-rate capacity for OMA is therefore $S_{\text{O}} = C_1^{\text{O}} + C_2^{\text{O}}$.

For the case of two-user NOMA, it is assumed that user 2's channel SNR gain is greater than user 1's channel SNR gain, i.e. $|G_2|^2 > |G_1|^2$). Therefore, the BS will allocate the transmit power such that user-2 can perform SIC at the receiver by treating its own signal as noise and decoding user 1's signal first. Meanwhile, user 1 will only attempt to decode its own signal and treat user 2's signal as noise. Let the power allocation coefficient for user 2 be $a \in (0, 1)$, such that user 1 is allocated $1 - a$ of the total transmit power. The transmitted SC signal is $\sqrt{(1-a)}x_1 + \sqrt{a}x_2$, then the received SC signals are

$$r_n = G_n(\sqrt{(1-a)\xi}x_1 + \sqrt{a\xi}x_2) + z_n, n = 1, 2. \tag{2.1}$$

User 1 will only attempt to decode its own signal, so the signal used to detect and decode user 1's signal $y_1 = r_1$, and thus the channel information capacity of user

1 is given by

$$C_1^{\text{noma}}(a) = \log_2\left(1 + \frac{(1-a)\xi|G_1|^2}{a\xi|G_1|^2+1}\right). \tag{2.2}$$

Meanwhile, user 2 must first detect and decode user 1's signal, then use it to perform SIC in order to detect and decode its own signal. The signal user 2 uses to perform SIC is $y_{2\to1} = r_2$. The channel capacity for user 2 to decode user 1's signal is

$$C_{2\to1}^{\text{noma}}(a) = \log_2\left(1 + \frac{(1-a)\xi|G_2|^2}{a\xi|G_2|^2+1}\right). \tag{2.3}$$

Since $|G_2|^2 > |G_1|^2 \Rightarrow C_2^{\text{noma}}(a) > C_1^{\text{noma}}(a)$, the SIC procedure will always be successful in the event that user 1 also succeeds in decoding its own signal. The signal obtained after the SIC procedure is $y_2 = r_2 - G_2\sqrt{(1-a)\xi}x_1 = G_2\sqrt{a\xi}x_2 + z_n$, and thus the channel capacity to decode its own signal is given by

$$C_2^{\text{noma}}(a) = \log_2\left(1 + a\xi|G_2|^2\right). \tag{2.4}$$

The sum-rate capacity for NOMA is therefore $S_{\text{N}}(a) = C_1^{\text{noma}}(a) + C_2^{\text{noma}}(a)$. These capacity expressions are used in each case of OMA and NOMA to find the values of a that make NOMA "fair."

2.3 Analysis of Two-User Fair-NOMA

In order for user 1 NOMA capacity to be greater than or equal to OMA capacity, it must be true that $C_1^{\text{noma}}(a) \geq C_1^{\text{oma}}$. Solving this inequality for a gives $a \leq \frac{\sqrt{1+\xi|G_1|^2}-1}{\xi|G_1|^2}$. Similarly, for user 2 when $C_2^{\text{noma}}(a) \geq C_2^{\text{oma}}$ results in $a \geq \frac{\sqrt{1+\xi|G_2|^2}-1}{\xi|G_2|^2}$. Both the upper and lower bounds on the transmit power fraction

a to achieve better sum and individual capacities have the form $a(u) = (\sqrt{1+\xi u} - 1)/(\xi u)$.

Define

$$a_{\text{inf}} = \frac{\sqrt{1+\xi|G_2|^2} - 1}{\xi|G_2|^2} \text{ and } a_{\text{sup}} = \frac{\sqrt{1+\xi|G_1|^2} - 1}{\xi|G_1|^2}. \tag{2.5}$$

Then by proposition 1 in [19], it is clear that if $|G_2|^2 > |G_1|^2 \Rightarrow a_{\text{sup}} > a_{\text{inf}}$. The Fair-NOMA power allocation region is therefore defined as $\mathcal{A}_{\text{FN}} = [a_{\text{inf}}, a_{\text{sup}}]$, and selecting any $a \in \mathcal{A}_{\text{FN}}$ gives $C_1^{\text{noma}}(a) \geq C_1^{\text{oma}}$, $C_2^{\text{noma}}(a) \geq C_2^{\text{oma}}$, and $S_{\text{noma}}(a) > S_{\text{oma}}$. Since the sum-rate capacity $S_{\text{noma}}(a)$ is a monotonically increasing function of a, then $a_{\text{sup}} = \arg\max_{a \in \mathcal{A}_{\text{FN}}} (C_2^{\text{noma}}(a))$ also maximizes $S_{\text{noma}}(a)$. The sum-rate capacity of NOMA is strictly larger than the sum-rate capacity of OMA because at the least one of the users' capacities always increases.

Theorem 2.3.1. *For a two-user NOMA system that allocates power fraction* $1-a$ *to user 1 and* a *to user 2, such that* $a \in \mathcal{A}_{FN}$*, the sum-rate* $S_{noma}(a)$ *is a monotonically increasing function of both* $|G_1|^2$ *and* $|G_2|^2$*.*

Proof. See appendix 2.7.1. □

This result implies that as the channel gain for the weaker user increases, the total capacity increases while the power allocation to the stronger user decreases. This means that, as the channel gain $|G_1|^2$ increases towards the value of $|G_2|^2$, then the capacity gain by user 1 is greater than the capacity loss by user 2. In the extreme case where $|G_1|^2 \to |G_2|^2$, then $a_{\text{sup}} \to a_{\text{inf}}$, and both $C_1^{\text{noma}}(a)$ and $C_2^{\text{noma}}(a) \to C_2^{\text{oma}}$. In other words, the Fair-NOMA capacity is upper bounded by the capacity obtained by allocating all of the transmit power to the stronger user. This is somewhat related to the multiuser diversity concept result in [20] for OMA systems, which suggests allocating all the transmit power to the stronger users will increase the overall capacity of the network.

In contrast, with the increase in $|G_2|^2$, $C_2^{\text{noma}}(a_{\text{sup}})$ increases and hence the capacity gains from Fair-NOMA increase. Therefore with Fair-NOMA, as is the same with the previously obtained result for fixed-power allocation NOMA, when $|G_2|^2 - |G_1|^2$ increases, $S_{\text{noma}}(a) - S_{\text{oma}}$ also increases [6, 18]. This will be further exemplified in Section 2.4, Theorem 2.4.1.

2.3.1 Expected Value of Fair-NOMA Capacity

The expected value of the Fair-NOMA capacities of the two users depend on the power allocation coefficient a. In order to determine the bounds of this region, the expected value of capacity of each user is derived for the cases of a_{inf} and a_{sup} and compared with that of OMA.

Since the channels of the two users are i.i.d. random variables, let the two users selected have channel SNR gains of $|G_i|^2$ and $|G_j|^2$, where $f_{|G|^2}(t) = \frac{1}{\beta} e^{-\frac{t}{\beta}}, t < 0$. Since we call the user with weaker (stronger) channel gain user 1 (user 2), then $|G_1| = \min\{|G_i|^2, |G_j|^2\}$ and $|G_2|^2 = \max\{|G_i|^2, |G_j|^2\}$. Therefore, the joint pdf of $|G_1|^2$ and $|G_2|^2$ is

$$f_{|G_1|^2,|G_2|^2}(t_1, t_2) = \frac{2}{\beta^2} e^{-\frac{t_1+t_2}{\beta}}, 0 < t_1 < t_2. \tag{2.6}$$

It is shown [19] that the ergodic capacities and the sum-rate of users in OMA are

$$\mathbb{E}[C_1^{\text{oma}}] = \frac{e^{\frac{2}{\beta\xi}}}{\ln(4)} E_1\left(\frac{2}{\beta\xi}\right), \tag{2.7}$$

$$\mathbb{E}[C_2^{\text{oma}}] = \frac{e^{\frac{1}{\beta\xi}}}{\ln(2)} E_1\left(\frac{1}{\beta\xi}\right) - \frac{e^{\frac{2}{\beta\xi}}}{\ln(4)} E_1\left(\frac{2}{\beta\xi}\right), \tag{2.8}$$

$$\mathbb{E}[S_{\text{oma}}] = \frac{e^{\frac{1}{\beta\xi}}}{\ln(2)} E_1\left(\frac{1}{\beta\xi}\right) \tag{2.9}$$

where $E_1(x) = \int_x^\infty u^{-1} e^{-u} du$ is the well-known exponential integral. Note that $\mathbb{E}[C_1^{\text{oma}}] = \mathbb{E}[C_1^{\text{noma}}(a_{\text{sup}})]$ and $\mathbb{E}[C_2^{\text{oma}}] = \mathbb{E}[C_2^{\text{noma}}(a_{\text{inf}})]$.

It is also shown [19] that

$$\begin{aligned}\mathbb{E}\left[C_1^{\text{noma}}(a_{\text{inf}})\right] =& \frac{3e^{\frac{2}{\beta\xi}}}{\ln(4)} E_1\left(\frac{2}{\beta\xi}\right) - \int_0^\infty \frac{2}{\beta\ln(2)} \cdot \exp\left(-\frac{x}{\beta}\left(\frac{\sqrt{1+\xi x}-2}{\sqrt{1+\xi x}-1}\right)\right) \\ &\times \left(E_1\left(\frac{x}{\beta(\sqrt{1+\xi x}-1)}\right) - E_1\left(\frac{x\sqrt{1+\xi x}}{\beta(\sqrt{1+\xi x}-1)}\right)\right) dx,\end{aligned} \tag{2.10}$$

and

$$\begin{aligned}\mathbb{E}[C_2^{\text{noma}}(a_{\text{sup}})] =& \frac{e^{\frac{2}{\beta\xi}}}{\ln(4)} E_1\left(\frac{2}{\beta\xi}\right) + \int_0^\infty \frac{2}{\beta\ln(2)} \\ &\times \exp\left(-\frac{x}{\beta}\left(\frac{\sqrt{1+\xi x}-2}{\sqrt{1+\xi x}-1}\right)\right) E_1\left(\frac{x\sqrt{1+\xi x}}{\beta(\sqrt{1+\xi x}-1)}\right) dx.\end{aligned} \tag{2.11}$$

At high SNR ($\xi \gg 1$), the approximate capacities are

$$C_n^{\text{oma}} \approx \frac{1}{2}\log_2(\xi|G_n|^2), n = 1,2 \tag{2.12}$$

$$C_1^{\text{noma}}(a_{\text{inf}}) \approx \frac{1}{2}\log_2(\xi|G_2|^2), \tag{2.13}$$

$$C_2^{\text{noma}}(a_{\text{sup}}) \approx \log_2\left(\sqrt{\frac{\xi}{|G_1|^2}}|G_2|^2\right). \tag{2.14}$$

This implies that when $\xi \gg 1$, $C_1^{\text{noma}}(a_{\text{sup}}) \approx C_2^{\text{oma}}$. The high SNR approximations lead to following result for the difference in the expected capacity gains, i.e., $\Delta S(a) = S_{\text{noma}}(a) - S_{\text{oma}}$.

Theorem 2.3.2. *For a two-user downlink SISO NOMA system with* $|G_n|^2 \sim \text{Exponential}(\frac{1}{\beta})$*,* $n = 1,2$*, and at high SNR regime, the sum-rate capacity gap has expected value* $\mathbb{E}[\Delta S(a)] \approx 1$ *bps/Hz,* $\forall a \in \mathcal{A}_{FN}$*.*

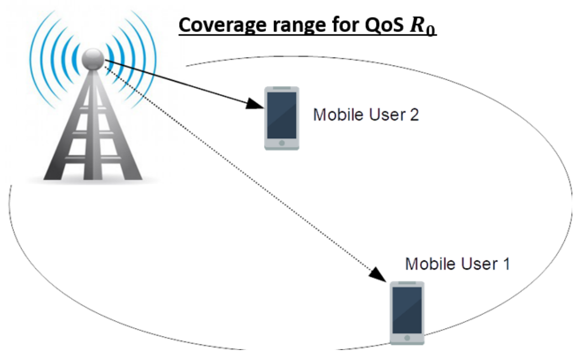

Figure 2.1: Users in are in coverage if they can support a minimum rate R_0, which related to the geographical are of a cell.

Proof. See appendix 2.7.2. □

This interesting result means that when the transmit power approaches infinity, the average increase in sum capacity is the same for both $a_{\inf}$ and $a_{\sup}$ and is equal to 1 bps/Hz. Equivalently, it means that $\forall a \in \mathcal{A}_{\text{FN}}$, both users experience an expected increase in capacity over OMA of c and $1-c$ where $c \in [0,1]$.

2.3.2 Outage Probability of Fair-NOMA

Suppose that the minimum rate that is allowed by the system to transmit a signal is R_0. The probability that a user cannot achieve this rate with any coding scheme is given by $\Pr\{C < R_0\}$. A user with channel SNR gain which cannot support the minimum rate R_0 is considered to be *out of coverage*, as illustrated by figure 2.1. As with the average capacity analysis, the outage performance of NOMA is analyzed by looking at $a_{\inf}$ and $a_{\sup}$, and then draw logical conclusions from that.

The outage probability of user 1 using OMA is given by

$$p_{\text{oma},1}^{\text{out}} = \Pr\left\{\log_2(1+\xi|G_1|^2)^{1/2} < R_0\right\} \tag{2.15}$$

$$= \int_0^{\frac{4^{R_0}-1}{\xi}} \int_{t_1}^{\infty} \frac{2}{\beta^2} e^{-\frac{t_1+t_2}{\beta}} dt_2 dt_1 \tag{2.16}$$

$$= 1 - \exp\left(-\tfrac{2(4^{R_0}-1)}{\beta\xi}\right).$$

For user 2 using OMA, the outage probability is given by

$$p_{\text{oma},2}^{\text{out}} = \Pr\left\{\log_2(1+\xi|G_2|^2)^{1/2} < R_0\right\} \tag{2.17}$$

$$= \int_0^{\frac{4^{R_0}-1}{\xi}} \int_{t_1}^{\frac{4^{R_0}-1}{\xi}} \frac{2}{\beta^2} e^{-\frac{t_1+t_2}{\beta}} dt_2 dt_1, \tag{2.18}$$

$$=1 + \exp\left(-\tfrac{2(4^{R_0}-1)}{\beta\xi}\right) - 2\exp\left(-\tfrac{4^{R_0}-1}{\beta\xi}\right).$$

Denote the NOMA outage probability for user n as $p_{\text{noma},n}^{\text{out}}(a)$ such that $p_{\text{noma},n}^{\text{out}}(a) = \Pr\{C_n^{\text{noma}}(a) < R_0\}$ for $n = 1, 2$. It should be obvious that $p_{\text{noma},1}^{\text{out}}(a_{\text{sup}}) = p_{\text{oma},1}^{\text{out}}$ and $p_{\text{noma},2}^{\text{out}}(a_{\text{inf}}) = p_{\text{oma},2}^{\text{out}}$. The outage probabilities $p_{\text{noma},1}^{\text{out}}(a_{\text{inf}})$ and $p_{\text{noma},2}^{\text{out}}(a_{\text{sup}})$ are provided in the following theorem.

Theorem 2.3.3. *Outage Probabilities $p_{noma,1}^{out}(a_{inf})$ and $p_{noma,2}^{out}(a_{sup})$:*

(a) The outage probability for user 1 when $a = a_{inf}$ is given by

$$p_{noma,1}^{out}(a_{inf}) = 1 + e^{-\frac{\alpha_2}{\beta}} - \frac{2}{\beta}\int_{\alpha_2}^{\infty} e^{-\frac{t(\alpha_1+1)}{\beta}} dt, \tag{2.19}$$

where α_1 and α_2 are defined as

$$\alpha_1 = \tfrac{2^{R_0}-1}{\xi t + 2^{R_0}(1-\sqrt{1+\xi t})},$$

$$\alpha_2 = \tfrac{4^{R_0}-2}{2\xi} + \sqrt{\tfrac{4^{R_0}-1}{\xi^2} + \tfrac{(4^{R_0}-2)^2}{4\xi^2}}.$$

(b) The outage probability for user 2 when $a = a_{sup}$ is given by

$$p_{noma,2}^{out}(a_{sup}) = 1 + e^{-\frac{2(4^{R_0}-1)}{\beta\xi}} - 2e^{-\frac{2(2^{R_0}-1)}{\beta\xi}} + (2^{R_0}-1) \times e^{\frac{(2^{R_0}-3)^2}{4\beta\xi}} \cdot \sqrt{\frac{\pi}{\beta\xi}} \left[\operatorname{erfc}\left(\frac{2^{R_0}+1}{2\sqrt{\beta\xi}}\right) - \operatorname{erfc}\left(\frac{3(2^{R_0})-1}{2\sqrt{\beta\xi}}\right)\right]. \tag{2.20}$$

Proof. See appendices 2.7.3 and 2.7.4. □

There is no closed form solution for the integral in $p_{\text{noma},1}^{\text{out}}(a_{\text{inf}})$, however it can be easily computed by a computer.

2.3.3 Comparison of Theoretical and Simulations Results

Fair-NOMA theoretical results are compared with simulation when $\beta = 1$. In figure 2.2, the capacity of NOMA is compared with that of OMA for both users, including the high SNR approximations. As can be seen, the theoretical derivations match the simulation results. The performances of $C_1^{\text{noma}}(a_{\text{inf}})$ and $C_2^{\text{noma}}(a_{\text{sup}})$ are plotted. The simulation of $\mathbb{E}[C_1^{\text{noma}}(a_{\text{inf}})]$ matches the theoretical result in equation (2.10), and the simulation of $\mathbb{E}[C_2^{\text{noma}}(a_{\text{sup}})]$ matches the theoretical result in equation (2.11). The high SNR approximations show to be very close for values of $\xi > 25$ dB. Since $C_2^{\text{oma}} = C_2^{\text{noma}}(a_{\text{inf}})$ and $C_1^{\text{oma}} = C_1^{\text{noma}}(a_{\text{sup}})$, it is apparent from the plots that the gain in performance is always approximately 1 bps/Hz for one of the users and also the sum capacity when using Fair-NOMA [19].

Figure 2.3 plots the outage probabilites of OMA and NOMA for different values of a and for $R_0 = 2$ bps/Hz. The probability of user 1 experiencing an outage is clearly greater than for user 2. However, the reduction of the outage probability for user 1 using $a = a_{\text{inf}}$ becomes significant as ξ increases, to the effect of nearly 1 order of magnitude drop-off when ξ is really large. The outage probability reduction for user 2 is not as significant as the improvement made by

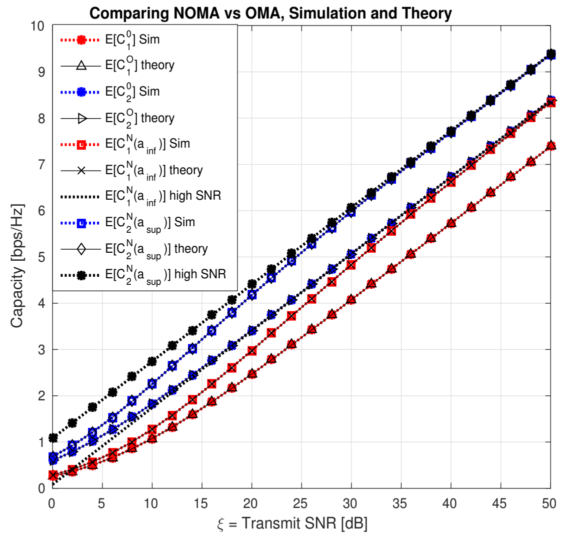

Figure 2.2: Comparing the capacity of NOMA and OMA

user 1. However, when $a = a_{\text{sup}}$, the same outage probability can be obtained using NOMA with ξ approximately 2 dB less than is required when using OMA. Thus, even when the power allocation coefficient a is restricted to being in $\mathcal{A}_{\text{FN}}$, the probability of users to be able to achieve their minimum service requirement rates R_0 is improved, and especially improved for the weaker channel gain. Even when the power allocation coefficient $a = (a_{\text{inf}} + a_{\text{sup}})/2$, the outage probabilities of both users improves significantly when using NOMA.

2.4 Fair-NOMA in Opportunistic User-Pairing

It has been suggested in [18] that the best NOMA performance is obtained when user channel conditions are most disparate, i.e. pairing the user with the weakest channel condition and the user with the strongest channel condition together. However, it is not known what the expected capacity gap is in this case, particularly for the case when both users are allocated power such that they both always outperform their OMA performance. Thus, the concept of Fair-NOMA is

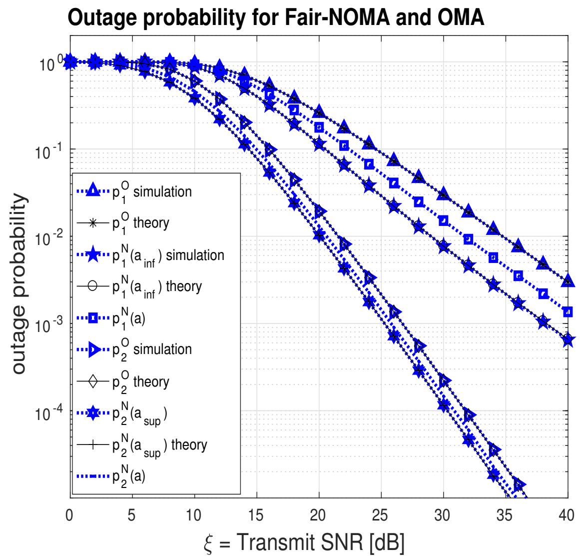

Figure 2.3: Outage probabilities of NOMA and OMA as functions of ξ.

applied to this user-pairing approach and the sum-rate capacity gap is analyzed.

Suppose there exists a set of K mobile users in a cell, and two of these users can be scheduled during the same transmission period. It is of particular interest to select the users that have the largest difference in channel SNR gain. If the channel SNR gains of the K users are i.i.d. Exponential$(\frac{1}{\beta})$, and the two selected users have the minimum and maximum channel SNR gains, how much of an improvement in the sum-rate capacity will be observed by using NOMA versus OMA? This scenario is depicted in figure 2.4.

2.4.1 Analysis of Fair-NOMA with Opportunistic User-Pairing

Let $|G_0|^2 = \min(|G_1|^2, \ldots, |G_K|^2)$ and $|G_M|^2 = \max(|G_1|^2, \ldots, |G_K|^2)$. In order to compute the expected sum-rate capacity, the joint CDF $F_{|G_0|^2,|G_M|^2}(t_0, t_M)$

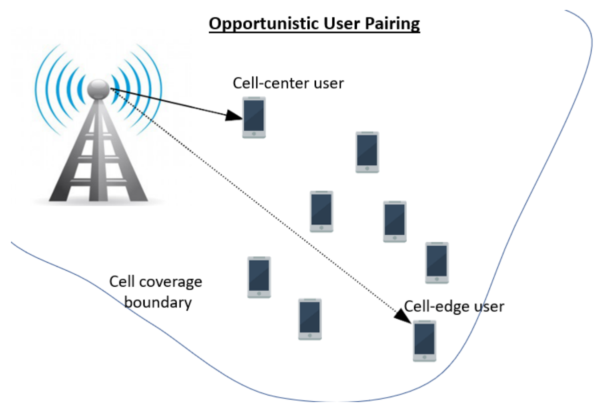

Figure 2.4: A cell-center user and a cell-edge user are paired for downlink NOMA transmissions.

and PDF of $f_{|G_0|^2,|G_M|^2}(t_0, t_M)$ are needed. It is easily shown that

$$\Pr\{|G_M|^2 < t_M\} = \Pr\{|G_0|^2 < t_0, |G_M|^2 < t_M\} + \Pr\{|G_0|^2 > t_0, |G_M|^2 < t_M\},$$

$$\Rightarrow F_{|G_0|^2,|G_M|^2}(t_0, t_M) = \Pr\{|G_0|^2 < t_0, |G_M|^2 < t_M\}$$
$$= \Pr\{|G_M|^2 < t_M\} - \Pr\{|G_0|^2 > t_0, |G_M|^2 < t_M\}. \quad (2.21)$$

For Rayleigh fading channels, the first term on the right in equation (2.21) is the CDF of the maximum of K i.i.d. exponential random variables, which is given by

$$\Pr\{|G_M|^2 < t_M\} = (1 - e^{-\frac{t_M}{\beta}})^K. \quad (2.22)$$

The second term can be easily computed.

$$\Pr\{|G_0|^2 > t_0, |G_M|^2 < t_M\} = \int_{t_0}^{t_M} \cdots \int_{t_0}^{t_M} \prod_{k=1}^{K} \frac{e^{-\frac{t_k}{\beta}}}{\beta} dt_k$$
$$= (e^{-\frac{t_0}{\beta}} - e^{-\frac{t_M}{\beta}})^K$$

Therefore, the joint CDF is given by

$$F_{|G_0|^2,|G_M|^2}(t_0, t_M) = (1 - e^{-\frac{t_M}{\beta}})^K - (e^{-\frac{t_0}{\beta}} - e^{-\frac{t_M}{\beta}})^K, \tag{2.23}$$

and the joint PDF is

$$f_{|G_0|^2,|G_M|^2}(t_0, t_M) = \frac{K(K-1)}{\beta^2} e^{-\frac{t_0+t_M}{\beta}} (e^{-\frac{t_0}{\beta}} - e^{-\frac{t_M}{\beta}})^{K-2}, \tag{2.24}$$

for $0 < t_0 < t_M$. The following theorem provides the sum-rate capacity increase of NOMA when $\xi |G_0|^2 \gg 1$.

Theorem 2.4.1. *Let $\{|G_1|^2, \ldots, |G_K|^2\}$ be the i.i.d. exponentially distributed SISO channel SNR gains of K users, such that the two users selected for transmission together have the minimum and maximum channel SNR gains. When $\xi\ |G_0|^2 \gg 1$, the sum-rate capacity increase from OMA to NOMA for $a = a_{sup}$ is*

$$\mathbb{E}[\Delta S(a_{sup})] \approx \frac{1}{2}\log_2(K) + \frac{1}{2}\sum_{m=2}^{K}\binom{K}{m}(-1)^m \log_2(m). \tag{2.25}$$

Proof. See appendix 2.7.5. □

Remark 1. *This result is similar to the result obtained for the 2-by-2 MIMO case in Lemma 2, equation 33 in [23], except a fixed-power allocation approach was used there, whereas the result above uses a Fair-NOMA power allocation approach. Although the expected capacity gap $\mathbb{E}[\Delta S(a)]$ increases when selecting $a = a_{sup}$, caution should be used when utilizing the fixed-power approach to not set a too close to the value of a_{inf}. An approximation of the capacity gap using $\mathbb{E}[\Delta S(a_{inf})]$*

for large ξ and K is given as

$$\mathbb{E}[\Delta S(a_{inf})] \approx \frac{e^{\frac{K}{\beta\xi}}}{\ln(4)} E_1\left(\frac{K}{\beta\xi}\right) - \log_2\left(1 + \frac{\sqrt{1+\xi(\psi(K+1)+\gamma)} - 1)}{K(\psi(K+1)+\gamma)}\right),$$

where $\psi(w) = \Gamma'(w)/\Gamma(w)$ is the digamma function, $\Gamma(w) = \int_0^\infty u^{w-1}e^{-u}du$ is the gamma function, and $\gamma = -\int_0^\infty e^{-u}\ln(u)du$ is the Euler-Mascheroni constant. It can be seen in figure 2.7 that as the number of users increases, the expected capacity gap actually decreases. Therefore, even for fixed-power allocation approaches to NOMA, a should be selected to be greater than a_{inf} for the case of pairing minimum and maximum channel gain users.

This result shows that the sum-rate capacity difference increases as a function of K. However, this increase is slow. Nonetheless, there is a fundamental limit to the amount the capacity can increase when using Fair-NOMA, while maintaining the capacity of the weaker user equal to the capacity using OMA.

It is important to note that as the number of mobile users becomes very large, while pairing the strongest and weakest users together will give us the greatest increase in sum-rate capacity, it does not maximize sum-rate capacity itself. This can be seen from theorem 2.3.1, which states that the sum-rate capacity actually increases as the channel gain of the weaker user monotonically increases. A practical way of viewing this issue is that, as the number of users K increases, the weakest user has channel gain that in probability is too weak to achieve the quality of service threshold rate R_0. Should no outage rate be specified, the weaker user achieves such a low capacity, that the stronger user contributes most of the capacity, while using only half the transmit power, according to proposition 1 from [19]. Hence, a little more than half of the transmit power is nearly wasted.

2.4.2 Comparing Simulation Results with Analysis

For the simulation results, the performance of Fair-NOMA combined with opportunistic user-pairing is compared to the performance of OMA and fixed-power NOMA. The simulations are run for different values of ξ and K. For fixed-power NOMA, the power allocation coefficient is a constant value of $a = \frac{1}{5}$, such that the weaker user is allocated $\frac{4}{5}$ of the transmit power.

Figure 2.5 shows the average capacities of both the weakest and strongest users versus K and for each case of $a = a_{\text{inf}}$ and a_{sup}. The capacity of the stronger user is shown to exhibit the effects of multiuser diversity, since not only does its channel gain grow as K increases, but also the power allocated also increases when $a = a_{\text{sup}}$, thus providing the increase in capacity predicted in equation (2.25). In the case of $a = a_{\text{inf}}$ the capacity is initially shown to increase as K increases, due to a_{inf} decreasing with $|G_M|^2$ according to proposition 1 from [19]. However, as K continues to increase, the weakest users capacity eventually begins to decrease due to its channel gain being the minimum of a large number of users, and thus this term begins to dominate the capacity behavior.

The sum-rate capacity for Fair-NOMA with $a = a_{\text{sup}}$, fixed-power NOMA with $a = \frac{1}{5}$, and OMA are shown in Fig. 2.6. As expected, the sum-rate capacity for each user at lower values of ξ performs best when applying Fair-NOMA when compared to fixed-power NOMA. This is because Fair-NOMA always guarantees a capacity increase, i.e. with probability 1, while fixed-power NOMA only achieves higher capacity with probability as given in [18]. However, as ξ increases, both capacities of Fair-NOMA and fixed-power NOMA approach the same value asymptotically. This agrees with the result obtained that at high SNR, the capacity gain should reach a limit when $\xi \to \infty$, no matter how much extra power is allocated to the stronger user.

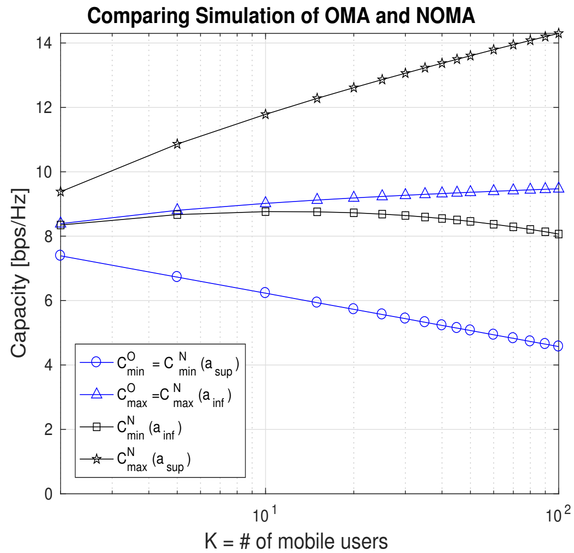

Figure 2.5: Ergodic capacity with opportunistic user-pairing, $\xi = 50$ dB

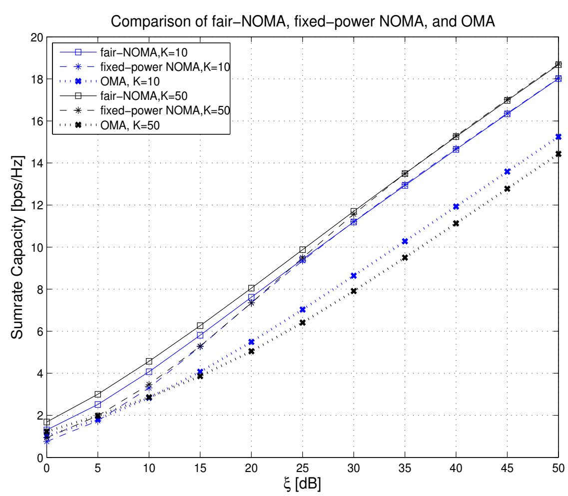

Figure 2.6: Comparison of Fair-NOMA, fixed-power NOMA, and OMA

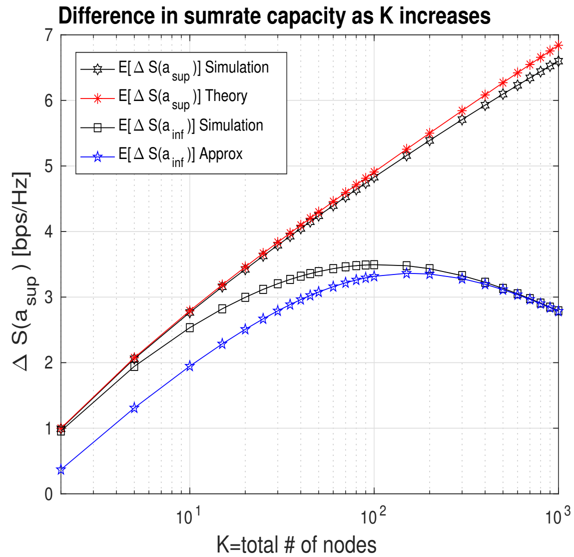

Figure 2.7: Difference in ergodic capacity with opportunistic user-pairing; $\xi = 50$ dB

Equation (2.25) shows that the capacity gain made by pairing the nearest and furthest cell-edge users is slow in K, and is due to the combined gain in capacity achieved by the strongest user and loss in capacity by the weakest user. This makes sense from multiple points of view. The expected value of power allocation coefficient $a_{\text{sup}} \to \frac{1}{2}$ when K is large, due to the selection of the user with the weakest channel gain. In other words, the weakest user needs less power in NOMA to achieve the same capacity as with OMA as K increases. Hence, more power goes to the stronger user. Figure 2.7 plots the simulation of $\mathbb{E}[\Delta S(a_{\text{sup}})]$ for $\xi = 50$ dB, and the approximation given by (2.25). Notice that the simulation and approximation seem to slightly diverge as the number of users increases. This is because the approximation in (2.25) needs a sufficiently large value of ξ as the number of users increases for the simulation and approximation to become tighter. However, $\xi = 50$ dB was used because it is a large but still realistic value of ξ. Since the number of users K cannot become arbitrarily large, the approximation remains tight for realistic values of ξ and K.

2.5 Multi-user NOMA Systems

So far, the treatment of Fair-NOMA has focused on the two-user case. However, downlink NOMA can be extended to more than two users paired on a common resource. If a BS serves K users in a cell, there can be more opportunity to experience the benefits of NOMA when compared to the two-user case. However, it must be determined whether or not there exists a fair power allocation approach for the general K-user case.

Consider an OMA system, where K users have their information transmitted over K orthogonal time slots (or frequency bands) during a total time of T (and bandwidth B). For each user k, the capacity of user k is given by

$$C_k^{\text{oma}} = \frac{1}{K}\log_2(1+\xi|G_k|^2), \forall k = 1,\ldots,K. \tag{2.26}$$

When applying NOMA to this system, the information of each user occupies the entire time T (bandwidth B) simultaneously. Hence, a superposition coding strategy must be used, in which all K users must share the total transmit power ξ. User k must perform SIC of each message that is intended for the other users l that have weaker channel conditions than user k. The channel gains are ordered as $|G_1|^2 < \cdots < |G_K|^2$. Lets define the power allocation coefficients $\{b_1,\ldots,b_K\}$, where b_k is the power allocation coefficient for user k and

$$\sum_{k=1}^{K} b_k \leq 1. \tag{2.27}$$

Therefore, the capacity of user k for $1 \leq k \leq K$ is given by

$$C_k^{\text{noma}}(b_1,\ldots,b_K) = \log_2\left(1+\frac{b_k\xi|G_k|^2}{1+\xi|G_k|^2\sum_{l=k+1}^{K} b_l}\right). \tag{2.28}$$

In order for $C_k^{\text{noma}}(b_1, \ldots, b_K) > C_k^{\text{oma}}$, the inequality must be solved for b_k assuming that equation (2.27) is true. Since user K does not receive any interference power after decoding all of the other users' messages, solving for b_K is straight forward.

$$C_K^{\text{oma}} \leq C_K^{\text{noma}}(b_1, \ldots, b_K) \Rightarrow b_K \geq \frac{(1+\xi|G_K|^2)^{\frac{1}{K}} - 1}{\xi|G_K|^2}. \tag{2.29}$$

For users $k = 1, \ldots, K-1$, the power allocation for each user is conditioned on $C_k^{\text{oma}} \leq C_k^{\text{noma}}(b_1, \ldots, b_K)$ which results in

$$b_k \geq \frac{[(1+\xi|G_k|^2)^{\frac{1}{K}} - 1]\left(1+\xi|G_k|^2 \sum_{l=k+1}^{K} b_l\right)}{\xi|G_k|^2}. \tag{2.30}$$

As expected, the power allocation of the users with weaker channel gains depends on the power allocation of the users with stronger channel gains. The power allocation strategy $(b_1, \ldots, b_K)$ assigns to each signal the minimum power allocation required in order for each user $k = 1, \ldots, K$ to achieve exactly the same capacity using NOMA as with OMA.

Notice that in the above derivation, the total power allocation was not necessarily used. Consider the case where $\sum_{k=1}^{K} a_k = 1$ and the case where user 1 capacity in OMA and NOMA are equal. Therefore, $C_1^{\text{oma}} = C_1^{\text{noma}} \Rightarrow$

$$\begin{aligned}
&\Rightarrow \log_2(1+\xi|G_1|^2)^{\frac{1}{K}} = \log_2\left(1 + \frac{a_1\xi|G_1|^2}{1+\xi|G_1|^2\sum_{l=2}^{K} a_l}\right) \\
&\Rightarrow (1+\xi|G_1|^2)^{\frac{1}{K}} = \frac{1+\xi|G_1|^2}{1+\xi|G_1|^2(1-a_1)}.
\end{aligned} \tag{2.31}$$

Solving for a_1 gives

$$a_1 = \frac{1 + \xi|G_1|^2 - (1 + \xi|G_1|^2)^{\frac{K-1}{K}}}{\xi|G_1|^2}. \tag{2.32}$$

Note that both sides of equation (2.31) are greater than 1 which means $0 < a_1 < 1, \forall \xi > 0$. Define $A_1 = 1 - a_1$ as the sum of the interference coefficients to user 1. Therefore,

$$A_1 = \frac{(1 + \xi|G_1|^2)^{\frac{K-1}{K}} - 1}{\xi|G_1|^2}, \tag{2.33}$$

and $0 < A_1 < 1$. In general, the power allocation coefficient required for the NOMA capacity of user k to equal the OMA capacity of user k can be derived by solving the equation

$$\begin{aligned} C_k^{\text{oma}} &= C_k^{\text{noma}}(a_1, \ldots, a_K) \\ \Rightarrow (1 + \xi|G_k|^2)^{\frac{1}{K}} &= \frac{1 + A_{k-1}\xi|G_k|^2}{1 + (A_{k-1} - a_k)\xi|G_k|^2}, \end{aligned} \tag{2.34}$$

$\forall k \in \{2, \ldots, K\}, \xi > 0$, where $A_{k-1} = 1 - \sum_{l=1}^{k-1} a_l$. The following theorem for the set of power allocation coefficients $\{a_1, \ldots, a_K\}$ arises from solving equation (2.34).

Theorem 2.5.1. *If the set of power allocation coefficients $\{a_1, \ldots, a_K\}$ are derived from equations (2.31) and (2.34), then*

$$a_k \in (0, 1), \qquad \textit{and} \qquad \sum_{k=1}^{K} a_k \leq 1, \tag{2.35}$$

meaning that there always exists a power allocation strategy for NOMA such that the capacities of each user can be at least as great as the capacity of OMA, when the channel gains are known at the base-station.

Proof. See appendix 2.7.6. □

This is an important theorem, because it sets the precedent for the existence of a set of NOMA power allocation coefficients that (i) achieves at least OMA capacity for every user in the current transmission time period, and (ii) allows for at least one user to have a capacity greater than OMA capacity.

The power allocation coefficient a_k considers interference received from users with higher channel gains to be at a maximum. However, the coefficients b_k consider the minimum power allocation. Note that in power allocation for multiuser Fair-NOMA using a_k coefficients, the allocation process begins with user having weakest channel (first user) and allocate enough power to have a capacity of at least equal to OMA capacity for the first user. Then, the process continues with the next user until all the power is allocated amongst all users, i.e., the last user with strongest channel receives the remaining power allocation that results in higher capacity than OMA for that user. When b_k coefficients are used for power allocations, the power allocation process begins with the user with strongest channel, K^{th} user and assign enough power to achieve the same capacity as OMA for that user. The process then continues with the next user until the process reaches the first user. Therefore, it is clear that

$$b_k < a_k \tag{2.36}$$

$$\text{and } C_k^{\text{noma}}(b_1, \ldots, b_K) < C_k^{\text{noma}}(a_1, \ldots, a_K), \forall k. \tag{2.37}$$

Hence, theorem 2.5.1 highlights that there always exists a power allocation scheme in the general multiuser NOMA case that always achieves higher capacity than OMA, while keeping the total transmit power to ξ. This minimum power allocation requirement is demonstrated in figure 2.8. The most interesting aspect of

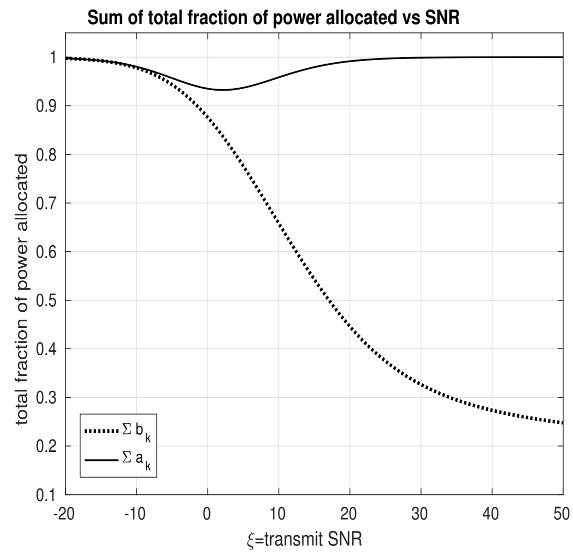

Figure 2.8: Minimum total power allocation in NOMA required to achieve capacity equal to OMA per user, $K = 5$

this result is that the same capacity of OMA can be achieved using Fair-NOMA with potentially much less total transmit power by using b_k coefficients. This can be useful if the purpose of NOMA is to minimize the total transmit power in the network.

2.6 Conclusions of Fair-NOMA

Fair-NOMA approach is introduced which allows two paired users to achieve capacity greater than or equal to the capacity with OMA. Given the power allocation set $\mathcal{A}_{\text{FN}}$ for this scheme, the ergodic capacity for the infimum and supremum of this set is derived for each user, and the expected asymptotic capacity gain is found to be 1 bps/Hz. The outage probability was also derived and it is shown that when $a = a_{\text{inf}}$, the outage performance of the weaker user significantly improves over OMA, where as the outage performance of the stronger user improves by at most roughly 2dB.

Fair-NOMA is applied to opportunistic user-pairing and the exact capacity gain is computed. The performance of Fair-NOMA is compared with a fixed-power NOMA approach to show that even when the power allocation coefficient $a = a_{\text{sup}}$ becomes less than the fixed-power allocation coefficient, the capacity gain is the same at high SNR, while Fair-NOMA clearly outperforms the fixed-power approach at low SNR. Finally, for a K-user downlink NOMA system, it is demonstrated that there always exists a power allocation strategy where each user can perform at least as well as in OMA, when the BS knows the exact channel gains.

The concept of Fair-NOMA can be extended to MIMO systems. In [23], a similar result is found for the approximate expected capacity gap of a 2-user 2-by-2 MIMO NOMA system. In order to eliminate the existing possibility that NOMA does not outperform OMA in capacity for any user, the Fair-NOMA approach can be applied to users that are utilizing the same degree of freedom from the BS. By ordering the composite channel gains, which include the transmit and receive beamforming applied to the channel, K users on the same transmit beam can have their signals superpositioned, and then SIC can be done at their receivers to obtain their own signal with minimum interference. Receive beamforming is used to eliminate the interference from the transmit beams' signals that are along the remaining spatial degress of freedom. The power allocation region can then be derived in the same manner as in Section 2.5, and NOMA can then be used to either increase the capacity gap as is done in [23], or to minimize the transmit power required to achieve the same capacity as in OMA, similar to what was done in Section 2.5.

Lastly, it is important to note that although this approach can be used for any downlink NOMA system (not only for SISO, but also MIMO), this approach has the limitation that it requires the BS to possess knowledge of the channel gain values, which is unrealistic in a real cellular system deployments. Therefore, the approach to designing power allocation stategies for downlink NOMA systems with a BS not possessing the channel gain value is provided in chapter 4.

2.7 Proofs of Fair-NOMA Results

2.7.1 Proof of Theorem 2.3.1

Proof. For the case when $a = a_{\text{inf}}$, the proof is trivial. Proving for the case when $a = a_{\text{sup}}$ then suffices to show it is true for all $a \in \mathcal{A}_{\text{FN}}$, because the $C_1^{\text{N}}(a)$ performance is lower-bounded by the case when $a = a_{\text{sup}}$, while for $C_2^{\text{N}}(a)$ the performance will only improve for $a > a_{\text{inf}}$. In order for S to be monotonically increasing function of $|G_i|^2$, it must be shown that $dS/d|G_i|^2 > 0$, $\forall |G_i|^2$. In the case of $|G_2|^2$, $C_1^{\text{N}}(a_{\text{sup}})$ does not factor in, so $\frac{dS}{d|G_2|^2} = \frac{dC_2^{\text{N}}(a_{\text{sup}})}{d|G_2|^2} = \frac{a_{\text{sup}}\xi}{a_{\text{sup}}\xi|G_2|^2} > 0$, $\forall |G_2|^2$. The case of $|G_1|^2$ goes as follows.

$$\frac{dS}{d|G_1|^2} = \frac{1}{\ln 2}\left[\frac{\xi}{2(1+\xi|G_1|^2)} + \frac{\frac{\xi|G_2|^2}{2|G_1|^2\sqrt{1+\xi|G_1|^2}} + \frac{|G_2|^2}{|G_1|^4}(1-\sqrt{1+\xi|G_1|^2})}{1+\frac{|G_2|^2}{|G_1|^2}(\sqrt{1+\xi|G_1|^2}-1)}\right]$$

$$= \frac{\left\{\begin{array}{l} \xi(|G_1|^4 + |G_1|^2|G_2|^2[(1+\xi|G_1|^2)^{\frac{1}{2}} - 1] + 2|G_2|^2(1+\xi|G_1|^2) \\ +\xi|G_1|^2|G_2|^2(1+\xi|G_1|^2)^{\frac{1}{2}} - 2|G_2|^2(1+\xi|G_1|^2)^{\frac{3}{2}} \end{array}\right\}}{2|G_1|^2\ln(2)(1+\xi|G_1|^2)(|G_1|^2+|G_2|^2(\sqrt{1+\xi|G_1|^2}-1))}$$

The numerator above can be simplified as

$$\xi|G_1|^4 + |G_2|^2\left[2\xi|G_1|^2\sqrt{1+\xi|G_1|^2} + \xi|G_1|^2 + 2 - 2(1+\xi|G_1|^2)^{\frac{3}{2}}\right].$$

The value inside the square brackets above can be simplified to

$$2 + \xi|G_1|^2 - 2\sqrt{1 + \xi|G_1|^2}.$$

Since $2 + \xi|G_1|^2 - 2\sqrt{1 + \xi|G_1|^2} \geq 0$ because $\xi^2|G_1|^4 \geq 0$, then

$$\begin{aligned}
&\Rightarrow \xi|G_1|^4 + |G_2|^2(2 + \xi|G_1|^2 - 2\sqrt{1 + \xi|G_1|^2}) > 0 \\
&\Rightarrow \frac{dS}{d|G_1|^2} > 0.
\end{aligned}$$

Since $|G_1|^2 < |G_2|^2$, $\frac{dS}{d|G_1|^2} > 0$, and $\frac{dS}{d|G_2|^2} > 0$, then S is a monotonically increasing function with respect to $|G_1|^2$ and $|G_2|^2$. □

2.7.2 Proof of Theorem 2.3.2

Proof. For $\xi \gg 1$,

$$\Delta C_1(a_{\text{inf}}) \approx \Delta C_2(a_{\text{sup}}) \approx \frac{1}{2}\log_2(|G_2|^2) - \frac{1}{2}\log_2(|G_1|^2).$$

The expected value of $\Delta C_1(a_{\text{inf}})$ and $\Delta C_2(a_{\text{sup}})$ is then

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{2}\log_2(\frac{|G_2|^2}{|G_1|^2})\right] \\
\approx &\int_0^\infty\int_0^{t_2} \frac{1}{\beta^2}e^{-\frac{t_1+t_2}{\beta}}\log_2(t_2)dt_1dt_2 - \int_0^\infty\int_{t_1}^\infty \frac{1}{\beta^2}e^{-\frac{t_1+t_2}{\beta}}\log_2(t_1)dt_2dt_1 \\
= &\int_0^\infty \frac{1}{\beta}e^{-\frac{t_2}{\beta}}\log_2(t_2)dt_2 - \int_0^\infty \frac{1}{\beta}e^{-\frac{2t_2}{\beta}}\log_2(t_2)dt_2 - \int_0^\infty \frac{1}{\beta}e^{-\frac{2t_1}{\beta}}\log_2(t_1)dt_1 \\
\overset{(a)}{=} &\int_0^\infty \frac{1}{\beta}e^{-\frac{t_2}{\beta}}\log_2(t_2)dt_2 - 2\int_0^\infty \frac{1}{\beta}e^{-\frac{2t_2}{\beta}}\log_2(t_2)dt_2 \\
\overset{(b)}{=} &\int_0^\infty \frac{1}{\beta}e^{-\frac{t_2}{\beta}}\log_2(t_2)dt_2 - \int_0^\infty \frac{1}{\beta}e^{-\frac{t}{\beta}}\log_2(t)dx + \int_0^\infty \frac{1}{\beta}e^{-\frac{t}{\beta}}\log_2(2)dx \\
\overset{(c)}{=} &1,
\end{aligned}$$

where (a) is true because the second two integrals are actually the same integral adding together, (b) is true by making the substitution $t = 2t_2$ for the second integral in the previous line, and (c) is true because the first two integrals are the same integral subtracting each other. Since $S_{\mathrm{N}}(a)$ is a monotonically increasing function of a, and $S_{\mathrm{N}}(a_{\inf}) - S_{\mathrm{O}} = \Delta C_1(a_{\inf})$ and $S_{\mathrm{N}}(a_{\sup}) - S_{\mathrm{O}} = \Delta C_2(a_{\sup})$, then when $\xi \gg 1$, $\Delta S = S_{\mathrm{N}}(a) - S_{\mathrm{O}} \approx 1, \forall a \in \mathcal{A}_{\mathrm{FN}}$.

□

2.7.3 Proof of proposition 2.3.3(a)

Proof. The outage probability $p_{\mathrm{N},1}^{\mathrm{out}}(a_{\inf})$ is given by

$$p_{\mathrm{N},1}^{\mathrm{out}}(a_{\inf}) = \Pr\left\{\log_2\left(\frac{1+\xi|G_1|^2}{1+a_{\inf}\xi|G_1|^2}\right) < R_0\right\} \tag{2.38}$$

$$= \Pr\left\{|G_1|^2 < \frac{|G_2|^2(2^{R_0}-1)}{\xi|G_2|^2 + 2^{R_0}(1-\sqrt{1+\xi|G_2|^2})}\right\} \tag{2.39}$$

Since $|G_1|^2 < |G_2|^2$, then there are two cases as

$$|G_2|^2 \lesseqgtr \frac{|G_2|^2(2^{R_0}-1)}{\xi|G_2|^2 + 2^{R_0}(1-\sqrt{1+\xi|G_2|^2})} = \alpha_1. \tag{2.40}$$

Solving above for $|G_2|^2$ gives

$$\Longrightarrow |G_2|^2 \lesseqgtr \frac{4^{R_0}-2}{2\xi} + \sqrt{\frac{4^{R_0}-1}{\xi^2} + \frac{(4^{R_0}-2)^2}{4\xi^2}} = \alpha_2. \tag{2.41}$$

This allows the event in equation (2.39) to be written as the two mutually exclusive events given in

$$\left\{|G_1|^2 < \alpha_1\right\} = \left\{|G_1|^2 < |G_2|^2, |G_2|^2 < \alpha_2\right\} \bigcup \left\{|G_1|^2 < \alpha_1, |G_2|^2 > \alpha_2\right\}. \tag{2.42}$$

The probability in equation (2.39) can then be written as $\Pr\{|G_1|^2 < \alpha_1\} = \Pr\{|G_1|^2 < |G_2|^2, |G_2|^2 < \alpha_2\} + \Pr\{|G_1|^2 < \alpha_1, |G_2|^2 > \alpha_2\}$. The first probability is equal to

$$\begin{aligned} \Pr\{|G_1|^2 < |G_2|^2, |G_2|^2 < \alpha_2\} &= \int_0^{\alpha_2}\int_0^{t_2} \frac{2}{\beta} e^{-\frac{t_1+t_2}{\beta}} dt_1 dt_2 \\ &= 1 + e^{-\frac{2\alpha_2}{\beta}} - 2e^{-\frac{\alpha_2}{\beta}}. \end{aligned} \tag{2.43}$$

The second probability is found to be

$$\begin{aligned} \Pr\{|G_1|^2 < \alpha_1, |G_2|^2 > \alpha_2\} &= \int_{\alpha_2}^{\infty}\int_0^{\alpha_1} \frac{2}{\beta} e^{-\frac{t_1+t_2}{\beta}} dt_1 dt_2 \\ &= 2e^{-\frac{\alpha_2}{\beta}} - \frac{2}{\beta}\int_{\alpha_2}^{\infty} e^{-\frac{t_2+\alpha_1}{\beta}} dt_2, \end{aligned} \tag{2.44}$$

where the integral in equation (2.44) has no known closed-form solution. Combining equations (2.43) and (2.44), gives

$$p_{\mathrm{N},1}^{\mathrm{out}}(a_{\mathrm{inf}}) = 1 + e^{-\frac{2\alpha_2}{\beta}} - \frac{2}{\beta}\int_{\alpha_2}^{\infty} e^{-\frac{t_2+\alpha_1}{\beta}} dt_2 \tag{2.45}$$

□

2.7.4 Proof of proposition 2.3.3(b)

Proof. The outage probability $p_{\mathrm{N},2}^{\mathrm{out}}(a_{\mathrm{sup}})$ is given by

$$p_{\mathrm{N},2}^{\mathrm{out}}(a_{\mathrm{sup}}) = \Pr\left\{\log_2\left(1 + \frac{|G_2|^2}{|G_1|^2}(\sqrt{1+\xi|G_1|^2} - 1)\right) < R_0\right\} \tag{2.46}$$

$$= \Pr\left\{|G_1|^2 > \frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1}\right\}. \tag{2.47}$$

Since $0 < |G_1|^2 < |G_2|^2$ is always true, then the domain of $|G_2|^2$ that makes the statement $\frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1} > 0$ true or false must be found, and thus gives us two intervals for $|G_2|^2$, which can be derived from the following relationship

$$\frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1} \lesseqgtr 0 \tag{2.48}$$

$$\Longrightarrow |G_2|^2 \lesseqgtr \frac{2(2^{R_0}-1)}{\xi}. \tag{2.49}$$

For the case of $|G_2|^2 < \frac{2(2^{R_0}-1)}{\xi}$, which gives $\frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1} < 0$, the event is explicitly written as

$$\mathcal{A}_1^{\text{out}} = \left\{0 < |G_1|^2 < |G_2|^2, 0 < |G_2|^2 < \frac{2(2^{R_0}-1)}{\xi}\right\}. \tag{2.50}$$

For the case of $|G_2|^2 > \frac{2(2^{R_0}-1)}{\xi}$, the interval for $|G_1|^2$ is $\frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1} < |G_1|^2 < |G_2|^2$, so it must also be true that $|G_2|^2 > \frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1}$. This gives $|G_2|^2 < \frac{(2^{R_0}-1)^2+2(2^{R_0}-1)}{\xi} = \frac{4^{R_0}-1}{\xi}$, and therefore the interval for this event is explicitly written as

$$\mathcal{A}_2^{\text{out}} = \left\{\frac{\xi|G_2|^4 - 2|G_2|^2(2^{R_0}-1)}{(2^{R_0}-1)^2} < |G_1|^2 < |G_2|^2, \frac{2(2^{R_0}-1)}{\xi} < |G_2|^2 < \frac{4^{R_0}-1}{\xi}\right\}$$

Now the probability above can be derived by computing the probabilities of the two disjoint regions as

$$\Pr\left\{|G_1|^2 > \frac{\xi|G_2|^4}{(2^{R_0}-1)^2} - \frac{2|G_2|^2}{2^{R_0}-1}\right\} = \Pr\{\mathcal{A}_1^{\text{out}}\} + \Pr\{\mathcal{A}_2^{\text{out}}\}. \tag{2.51}$$

The first probability is computed by

$$\Pr\{\mathcal{A}_1^{\text{out}}\} = \int_0^{\frac{2(2^{R_0}-1)}{\xi}} \int_0^{t_2} \frac{2}{\beta^2} e^{-\frac{t_1+t_2}{\beta}} dt_1 dt_2$$

$$= 1 + e^{-\frac{4(2^{R_0}-1)}{\beta\xi}} - 2e^{-\frac{2(2^{R_0}-1)}{\beta\xi}}. \tag{2.52}$$

Let $K_1 = \frac{2(2^{R_0}-1)}{\xi}$ and $K_2 = \frac{4^{R_0}-1}{\xi}$. Then the second probability is given by

$$\Pr\{\mathcal{A}_2^{\text{out}}\} = \int_{K_1}^{K_2} \int_{\frac{\xi t_2^2 - 2t_2(2^{R_0}-1)}{(2^{R_0}-1)^2}}^{t_2} \frac{2}{\beta^2} e^{-\frac{t_1+t_2}{\beta}} dt_1 dt_2 \tag{2.53}$$

$$= \int_{K_1}^{K_2} \frac{2}{\beta} \left(e^{-\left(\frac{t_2}{\beta} + \frac{\xi t_2^2 - 2t_2(2^{R_0}-1)}{\beta(2^{R_0}-1)^2}\right)} - e^{-\frac{2t_2}{\beta}} \right) dt_2 \tag{2.54}$$

The second term in the integral in equation (2.54) can be easily computed to be

$$\int_{K_1}^{K_2} \frac{2}{\beta} e^{-\frac{2t_2}{\beta}} dt_2 = -e^{-\frac{2(4^{R_0}-1)}{\beta\xi}} + e^{-\frac{4(2^{R_0}-1)}{\beta\xi}}. \tag{2.55}$$

The first integral in equation (2.54) is computed by completing the square in the exponent as

$$\int_{K_1}^{K_2} \frac{2}{\beta} e^{-\left(\frac{t_2}{\beta} + \frac{\xi t_2^2 - 2t_2(2^{R_0}-1)}{\beta(2^{R_0}-1)^2}\right)} dt_2 = \int_{K_1}^{K_2} \frac{2}{\beta} e^{-\frac{\xi}{\beta(2^{R_0}-1)^2}\left(t_2^2 + \frac{t_2(2^{R_0}-1)(2^{R_0}-3)}{\xi}\right)} dt_2. \tag{2.56}$$

Let $\phi = \frac{(2^{R_0}-1)(2^{R_0}-3)}{\xi}$ Since

$$t_2^2 + t_2\phi = \left(t_2 + \frac{\phi}{2}\right)^2 - \left(\frac{\phi}{2}\right)^2, \tag{2.57}$$

then equation (2.56) equals

$$= \frac{2}{\beta} e^{\frac{(2^{R_0}-3)^2}{4\beta\xi}} \int_{K_1}^{K_2} e^{-\frac{\xi}{\beta(2^{R_0}-1)^2}\left(t_2 + \frac{(2^{R_0}-1)(2^{R_0}-3)}{2\xi}\right)^2} dt_2. \tag{2.58}$$

By using the substitution

$$u(t_2) = \frac{1}{2^{R_0}-1}\sqrt{\frac{\xi}{\beta}}\left(t_2 + \frac{(2^{R_0}-1)(2^{R_0}-3)}{2\xi}\right), \tag{2.59}$$

the integral in equation (2.58) equals

$$= \frac{2}{\beta} e^{\frac{(2^{R_0}-3)^2}{4\beta\xi}} \int_{u(K_1)}^{u(K_2)} e^{-u^2} \cdot (2^{R_0}-1)\sqrt{\frac{\beta}{\xi}} du. \tag{2.60}$$

$$= (2^{R_0}-1) e^{\frac{(2^{R_0}-3)^2}{4\beta\xi}} \sqrt{\frac{\pi}{\beta\xi}} \cdot [\mathrm{erfc}\,(u(K_1)) - \mathrm{erfc}\,(u(K_2))], \tag{2.61}$$

where $u(t)$ is obtained by equation (2.59), and thus $u(K_1) = \frac{2^{R_0}+1}{2\sqrt{\beta\xi}}$, $u(K_2) = \frac{3(2^{R_0})-1}{2\sqrt{\beta\xi}}$, and $\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty e^{-u^2} du$ is the complementary error function. Thus, combining equations (2.52, 2.55, 2.61) results in equation (2.20). □

2.7.5 Proof of Theorem 2.4.1

Proof. When $\xi|G_0|^2 \gg 1$, $\Delta S(a_{\text{sup}}) \approx \frac{1}{2}(\log_2(|G_M|^2) - \log_2(|G_0|^2))$. Therefore by equation (2.62),

$$\begin{aligned}
&\mathbb{E}\left[\Delta S(a_{\text{sup}})\right] \\
&\approx \int_0^\infty \int_0^{t_M} \frac{1}{2}\log_2(t_M) \frac{K(K-1)}{\beta^2} e^{-\frac{t_0+t_M}{\beta}} (e^{-\frac{t_0}{\beta}} - e^{-\frac{t_M}{\beta}})^{K-2} dt_0 dt_M \\
&\quad - \int_0^\infty \int_{t_0}^\infty \frac{1}{2}\log_2(t_0) \frac{K(K-1)}{\beta^2} e^{-\frac{t_0+t_M}{\beta}} (e^{-\frac{t_0}{\beta}} - e^{-\frac{t_M}{\beta}})^{K-2} dt_M dt_0 \qquad (2.62) \\
&= \int_0^\infty \log_2(t_M) \frac{K}{2\beta} e^{-\frac{t_M}{\beta}} (1 - e^{-\frac{t_M}{\beta}})^{K-1} dt_M - \int_0^\infty \log_2(t_0) \frac{K}{2\beta} e^{-\frac{Kt_0}{\beta}} dt_0 \\
&= \int_0^\infty \log_2(t_M) \frac{K}{2\beta} e^{-\frac{t_M}{\beta}} \sum_{n=0}^{K-1} \binom{K-1}{n} (-1)^n e^{-\frac{nt_M}{\beta}} dt_M - \int_0^\infty \log_2(t_0) \frac{K}{2\beta} e^{-\frac{Kt_0}{\beta}} dt_0 \\
&= \sum_{n=0}^{K-1} \int_0^\infty \log_2\left(\frac{t}{n+1}\right) \frac{K}{2(n+1)\beta} \binom{K-1}{n} (-1)^n e^{-\frac{t}{\beta}} dt - \int_0^\infty \log_2\left(\frac{t}{K}\right) \frac{1}{2\beta} e^{-\frac{t}{\beta}} dt
\end{aligned}$$

$$=\int_0^{\infty}\log_2\left(\frac{K}{t}\cdot\prod_{n=0}^{K-1}\left(\frac{t}{n+1}\right)^{\binom{K}{n+1}(-1)^n}\right)\frac{1}{2\beta}e^{-\frac{t}{\beta}}dt$$
$$=\frac{1}{2}\log_2(K)+\frac{1}{2}\sum_{m=1}^{K}\binom{K}{m}(-1)^m\log_2(m).$$

□

2.7.6 Proof of proposition 2.5.1

Proof. It is already established that $a_1, A_1 \in (0,1)$. The power allocation coefficient for user 2 is found by the equation

$$(1+\xi|G_2|^2)^{\frac{1}{K}} = \frac{1+A_1\xi|G_2|^2}{1+(A_1-a_2)\xi|G_2|^2}$$
$$\implies a_2 = \frac{(1+A_1\xi|G_2|^2)[(1+\xi|G_2|^2)^{\frac{1}{K}}-1]}{\xi|G_2|^2(1+\xi|G_2|^2)^{\frac{1}{K}}}. \quad (2.63)$$

If the following is true

$$1 < \frac{1+A_1\xi|G_2|^2}{1+(A_1-a_2)\xi|G_2|^2} < 1+A_1\xi|G_2|^2, \quad (2.64)$$

then clearly $a_2 \in (0, A_1)$. However, for equation (2.63) and inequality (2.64) to be true, it must be true that

$$1 < (1+\xi|G_2|^2)^{\frac{1}{K}} < 1+A_1\xi|G_2|^2. \quad (2.65)$$

It is trivial to show that $1 < (1+\xi|G_2|^2)^{\frac{1}{K}}, \forall\xi, |G_2|^2 > 0$. To show that $(1+\xi|G_2|^2)^{\frac{1}{K}} < 1+A_1\xi|G_2|^2, \forall\xi > 0$, the inequality is rearranged so that

$$\gamma_2 < A_1, \quad (2.66)$$

where

$$\gamma_k = \frac{(1+\xi|G_k|^2)^{\frac{1}{K}} - 1}{\xi|G_k|^2}. \tag{2.67}$$

The inequality $\gamma_2 < A_1$ is clearly true because $\gamma_2 < \gamma_1$, and $\gamma_1 < A_1$ because $(1+\xi|G_1|^2)^{\frac{m}{K}} < (1+\xi|G_1|^2)^{\frac{K-1}{K}}, \forall m < K-1$. Therefore, equation (2.63) and inequality (2.64) are true. In a similar manner, in order for the power allocation coefficient a_k for user k to be less than total interference A_{k-1} received by user $k-1$, the following must be true:

$$a_k = \frac{(1+A_{k-1}\xi|G_k|^2)[(1+\xi|G_k|^2)^{\frac{1}{K}} - 1]}{\xi|G_k|^2(1+\xi|G_k|^2)^{\frac{1}{K}}}, \tag{2.68}$$

$$1 < \frac{1+A_{k-1}\xi|G_k|^2}{1+(A_{k-1}-a_k)\xi|G_k|^2} < 1+A_{k-1}\xi|G_k|^2, \tag{2.69}$$

$$1 < (1+\xi|G_k|^2)^{\frac{1}{K}} < 1+A_{k-1}\xi|G_k|^2. \tag{2.70}$$

Equation (2.68) is true by solving eq. (2.34), while (2.69) states that $a_k \in (0, A_{k-1})$ and (2.70) requires that user k's OMA capacity is feasible within $a_k \in (0, A_{k-1})$, given the channel condition of user k. Therefore, (2.70) leads to

$$\gamma_k < A_{k-1} = A_{k-2} - a_{k-1} = \frac{A_{k-2} - \gamma_{k-1}}{(1+\xi|G_{k-1}|^2)^{\frac{1}{K}}}$$
$$\Longrightarrow A_{k-2} > \frac{(1+\xi|G_k|^2)^{\frac{1}{K}} - 1}{\xi|G_k|^2}(1+\xi|G_{k-1}|^2)^{\frac{1}{K}} + \frac{(1+\xi|G_{k-1}|^2)^{\frac{1}{K}} - 1}{\xi|G_{k-1}|^2} \tag{2.71}$$

Since the function

$$f(t) = \frac{(1+t)^{\frac{m}{K}} - 1}{t}, \forall m < K, m \text{ and } K \in \mathbb{N} \tag{2.72}$$

is a monotonically decreasing function of t, then

$$\frac{(1+\xi|G_k|^2)^{\frac{1}{K}}-1}{\xi|G_k|^2}(1+\xi|G_{k-1}|^2)^{\frac{1}{K}}+\frac{(1+\xi|G_{k-1}|^2)^{\frac{1}{K}}-1}{\xi|G_{k-1}|^2} \tag{2.73}$$

$$<\frac{(1+\xi|G_{k-1}|^2)^{\frac{1}{K}}-1}{\xi|G_{k-1}|^2}(1+\xi|G_{k-1}|^2)^{\frac{1}{K}}+\frac{(1+\xi|G_{k-1}|^2)^{\frac{1}{K}}-1}{\xi|G_{k-1}|^2} \tag{2.74}$$

$$=\frac{(1+\xi|G_{k-1}|^2)^{\frac{2}{K}}-1}{\xi|G_{k-1}|^2} \tag{2.75}$$

$$< A_{k-2} = A_{k-3} - a_{k-2}$$

$$=\frac{A_{k-3}-\gamma_{k-2}}{(1+\xi|G_{k-2}|^2)^{\frac{1}{K}}} \tag{2.76}$$

$$\Longrightarrow A_{k-3} > \frac{(1+\xi|G_{k-1}|^2)^{\frac{2}{K}}-1}{\xi|G_{k-1}|^2}(1+\xi|G_{k-2}|^2)^{\frac{1}{K}}+\frac{(1+\xi|G_{k-2}|^2)^{\frac{1}{K}}-1}{\xi|G_{k-2}|^2}. \tag{2.77}$$

The inequality in (2.77) has the same form as the inequality in (2.71), so the same steps taken in inequalities (2.73) and (2.75) can be used repeatedly, until the following is obtained

$$\frac{(1+\xi|G_1|^2)^{\frac{k-1}{K}}-1}{\xi|G_1|^2} \leq A_1, \tag{2.78}$$

which is true $\forall k \leq K$. Hence, this series of inequalities shows that the transmit power allocation coefficient a_k required for user k to achieve OMA capacity is always less than the total interference coefficient received by user $k-1$, which equals the total fraction of power available for users $k, \ldots, K$.

$\square$

Chapter 3

Downlink Cache-Aided NOMA Systems

Caching and non-orthogonal multiple access (NOMA) are two candidate technologies to help the next generation of wireless cellular systems [2] to meet future high throughput demands. Power-domain NOMA relies on superposition coding (SC) at the transmitter, and successive interference cancellation (SIC) at the receiver [7] in order to achieve capacity in multi-user wireless systems. Given the advanced storage and processing capabilities of current and future user terminals, edge caching can be leveraged in the downlink by users with weaker channels in a NOMA system. This approach is called cache-aided NOMA [27].

3.1 Introduction to Cache-Aided NOMA

In a cache-aided NOMA system, a user with stronger channel can use SIC to remove the interference caused by the signal intended for the user with weaker channel, while the user with weaker channel can use cache-aided interference cancellation (CA-IC) if it caches the information being requested by the stronger user.

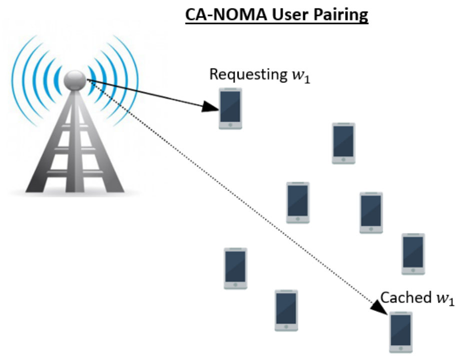

Figure 3.1: A cell-edge user possesses a cache of the information being requested by a cell-center user, thus the BS pairs them for a DL NOMA transmission.

The rate for the stronger user does not have to be within the capacity region of the weaker user's wireless channel in order to perform CA-IC, unlike the case of the stronger user for performing SIC on the weaker user's signal. Thus this scenario creates interference free channel at each user terminal for decoding its respective signal. An scenario is illustrated in figure 3.1.

The results in this chapter have been motivated by the results in [27], but have been generalized as a result of further studies.

3.1.1 Previous work on caching and NOMA

Caching at the edge of networks, including small Base Stations (BSs) and device-to-device communications for future 5G wireless networks is described in [24], where it was shown that caching is an energy efficient solution to help the network improve coverage probabilities while maximizing spectral efficiency. In [25], it is demonstrated that using decentralized coded caches at user terminals increases the throughput capacity of the network, while alleviating the burden of network traffic caching at BSs and helper nodes.

Recently, two techniques that combine caching and NOMA have been introduced. The work in [27] focused on introducing the possibility of leveraging caching to facilitate the function of a downlink NOMA system for users with i.i.d. channel gains. The difference between this work and [27] is that this work considers the users to possess different distributions based on their path losses (β), which benefits CA-NOMA but also creates new problems. The work in [26] focuses on utilizing NOMA to facilitate the caching of content on servers. These two works demonstrate that caching and NOMA are two technologies that should be utilized in conjunction. This work aims to demonstrate that a power allocation region always exists that allows CA-NOMA to outperform regular NOMA and OMA, and finds the union-outage probability minimizing power allocation coefficient.

3.2 CA-NOMA System Model and Capacity

Consider a two-user orthogonal multiple access (OMA) single-input-single-output (SISO) system. Let the total transmit time period be T, where users are allocated equal non-overlapping time slots of length $T/2$, and allocated the total transmit SNR ξ for that slot[1]. For each user $k = 1, 2$, if the BS is transmitting signal x_k ($\mathbb{E}[|x_k|^2] = 1$), the channel SNR gain between the BS and user k is G_k ($|G_k|^2 \sim \text{Exponential}(\frac{1}{\beta_k})$), where β_k is the long-term average of the channel SNR gain $|G_k|^2$ given by how far user k is from the BS. Each receiver has noise $z_k \sim \mathcal{CN}(0, 1)$, so the received signal at user k is given by $y_k = \sqrt{\xi} G_k x_k + z_k$. Since the time duration is $\frac{1}{2}$, then the capacity of user k in bps/Hz is given by $C_k^{\text{oma}} = \frac{1}{2} \log_2(1 + \xi |G_k|^2)$.

Without loss of generality, user 1 is labeled as the user farther from the BS, and user 2 the close one to the BS, leading to the long-term statistic of the channel SNR

[1] The same OMA formulation can be made in the case of frequency division instead of time division.

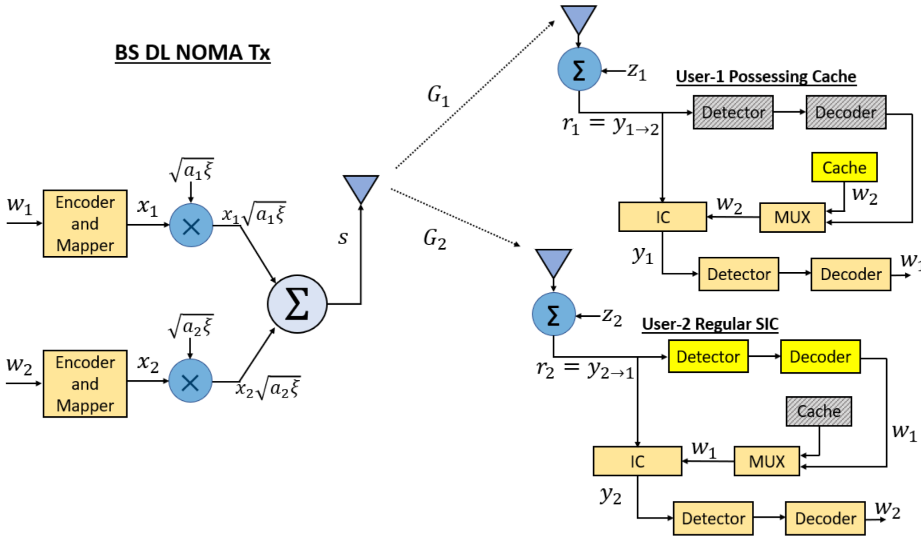

Figure 3.2: Block diagram of DL CA-NOMA system.

gains to have the relationship $\beta_1 < \beta_2$. In a two-user NOMA system, where the BS is transmitting the two signals using SC, the signal x_i is carrying the information for user i. The power allocation coefficients for users 1 and 2 are $1 - a$ and a, respectively. The SC signal transmitted by the BS is then $\sqrt{(1-a)\xi}x_1 + \sqrt{a\xi}x_2$, thus the raw received signal at user i is

$$r_i = G_i(\sqrt{(1-a)\xi}x_1 + \sqrt{a\xi}x_2) + z_i, i = 1, 2. \tag{3.1}$$

As shown in figure 3.2, user 2 receives r_2 and detects user 1's signal, decodes it, reconstructs x_1, and performs SIC in order to obtain y_2. However, in the event that user 1 possesses a cache of the data requested by user 2, it can perform CA-IC on r_1, thus obtaining y_2. The signals and associated capacities are thus given by

$$y_1 = \sqrt{(1-a)\xi}G_1x_1 + z_1, \;\; y_2 = \sqrt{a\xi}G_2x_2 + z_2, \tag{3.2}$$

and

$$C_1 = \log_2(1 + (1-a)\xi|G_1|^2), \;\; C_2 = \log_2(1 + a\xi|G_2|^2). \tag{3.3}$$

However, the capacity for user 2 to decode and subtract user 1's signal using SIC is

$$C_{2\to 1} = \log_2\left(1 + \frac{(1-a)\xi|G_2|^2}{a\xi|G_2|^2 + 1}\right) \tag{3.4}$$

When using regular NOMA, it is always true that if $|G_1|^2 < |G_2|^2 \to C_1(a) < C_{2\to 1}(a)$. However, since CA-NOMA allows user 1 to subtract user 2's interference, this relationship of capacities no longer holds for CA-NOMA. Thus the effects of CA-NOMA on outage performances is investigated.

3.3 CA-NOMA for Achieving a Minimum QoS Rate

Suppose that users 1 and 2 have target rates R_1 and R_2, respectively, where $R_1 < R_2$. The following definition gives the outage events for users 1 and 2.

Definition 1. *For the following sets*

$$\mathcal{B}_1 = \{C_1 < R_1\} = \left\{|G_1|^2 < \tfrac{2^{R_1}-1}{(1-a)\xi}\right\},$$
$$\mathcal{B}_{21} = \{C_{2\to 1} < R_1\} = \left\{|G_2|^2 < \tfrac{2^{R_1}-1}{\xi(1-a2^{R_1})}\right\},\; \mathcal{B}_{22} = \{C_2 < R_2\} = \left\{|G_2|^2 < \tfrac{2^{R_2}-1}{a\xi}\right\},$$

the outage event for user 1 is $\mathcal{B}_1$ *and the outage event for user 2 is* $\mathcal{B}_2 = \mathcal{B}_{21} \cup \mathcal{B}_{22}$.

For user 2, it is clear that if $\mathcal{B}_{21}$ occurs, then the set $\mathcal{B}_{22}$ becomes unrealizable. If any of these outage events occurs, then the system will fail to perform a successful NOMA transmission. The union outage probability, which is a function of a, is then given by

$$p_{\text{U-out}}(a) = 1 - \Pr\{\overline{\mathcal{B}_1} \cap \overline{\mathcal{B}_2}\}. \tag{3.5}$$

For comparison, the OMA outage events are given by

$$\mathcal{B}_k^{\text{oma}} = \{C_k^{\text{oma}} < R\} = \left\{|G_k|^2 < \tfrac{4^{R_k}-1}{\xi}\right\} \tag{3.6}$$

The simplest approach to power allocation in NOMA systems is fixed-power allocation. The optimum power allocation coefficient a_{opt} is then chosen to minimize $p_{\text{U-out}}(a)$. However, defining the constraints for finding a_{opt} is more complex compared to [27], because having $R_1 \neq R_2$ in general can give rise to the possibility for both users to simultaneously have better individual outage performances, while that was not possible in [27].

3.3.1 Fixed power allocation conditions for successful CA-NOMA

Suppose a fixed-power NOMA system is employed. The following proposition outlines the fundamental power allocation requirements needed for CA-NOMA, and provides the foundation for computing $p_{\text{U-out}}(a)$.

Proposition 1. *Let the events $\mathcal{B}_1$, $\mathcal{B}_{21}$, $\mathcal{B}_{22}$, and $\mathcal{B}_2$ be defined as definition 1. Also, let $R_1 < R_2$, and the set $\mathcal{A} = \{a : \mathcal{B}_{21} \subset \mathcal{B}_{22}\}$.*

1. *If $\mathcal{B}_{21} \subseteq \mathcal{B}_{22}$, then $\mathcal{B}_2 = \mathcal{B}_{22}$ and*

$$\mathcal{A} = \left(0, \frac{2^{R_2} - 1}{2^{R_1+R_2} - 1}\right), \tag{3.7}$$

where $\mathcal{B}_{21} \subseteq \mathcal{B}_{22}$ *occurs when* $a = \sup\{\mathcal{A}\}$.

2. *If* a *is selected such that* $\mathcal{B}_1 \subseteq \mathcal{B}_1^{oma}$ *and* $\mathcal{B}_2 = \mathcal{B}_{22}$*, then*

$$a \leq \min\left\{\frac{2^{R_2}-1}{2^{R_1+R_2}-1}, \frac{2^{R_1}}{2^{R_1}+1}\right\}, \tag{3.8}$$

and the minimum is always given by $\frac{2^{R_2}-1}{2^{R_1+R_2}-1}$ *if* $R_1 \geq \log_2 \varphi$*, where* $\varphi = \frac{1+\sqrt{5}}{2}$ *is the* golden ratio.

3. *If* $\mathcal{B}_2 \subseteq \mathcal{B}_2^{oma}$*, then* $a \geq \frac{1}{2^{R_2}+1}$*, and the set*

$$\mathcal{A}_{\text{CAN}} = \left(\frac{1}{2^{R_2}+1}, \frac{2^{R_2}-1}{2^{R_1+R_2}-1}\right) \tag{3.9}$$

is always a non-empty subset of $\mathcal{A}$.

4. *For* $\forall a \in \mathcal{A}$, $\frac{2^{R_1}-1}{(1-a)\xi} < \frac{2^{R_2}-1}{a\xi}$.

Proof. See appendix. □

Since SIC requires event $\mathcal{B}_{21}$ to be true in order to make event $\mathcal{B}_{22}$ possible, it is preferred that $\mathcal{B}_{21} \subseteq \mathcal{B}_{22}$, which is the condition stated item (1) in the above proposition. Item (2) shows the condition for user 1 to have better outage performance when compared to OMA, and that as long as $R_1 \geq \log_2 \varphi \approx 0.694$bps/Hz, user 1's outage performance is always better than OMA $\forall a \in \mathcal{A}$. Furthermore, it is easy to verify that for regular NOMA (without caching), a should still lie within $\mathcal{A}$, and user 1's outage performance is always superior for CA-NOMA over regular NOMA. Finally, item (3) is for the case when the engineer desires to have both users 1 and 2 simultaneously outperform OMA, and item (4) will help set up the calculation for $p_{\text{U-out}}(a)$ in (3.5).

3.3.2 Outage Probability of CA-NOMA

Suppose the cell-edge user possesses the cache of the information that the cell-center user is requesting. Also, assume that the cell-center user has a greater instantaneous channel SNR gain and its information is transmitted a higher target rate. The joint p.d.f. of $|G_1|^2$ and $|G_2|^2$, given that $|G_1|^2 < |G_2|^2$, can be easily shown to be

$$f_{|G_1|^2,|G_2|^2}(w_1, w_2) = \frac{\beta_1 + \beta_2}{\beta_1 \beta_2^2} e^{-\frac{w_1}{\beta_1}} e^{-\frac{w_2}{\beta_2}}, \tag{3.10}$$

Since $p_{\text{U-out}}(a) = 1 - \Pr\{\overline{\mathcal{B}_1} \cap \overline{\mathcal{B}_2}\}$, and

$$\Pr\{\overline{\mathcal{B}_1} \cap \overline{\mathcal{B}_2}\} = \Pr\{\overline{\mathcal{B}_1}\} - \Pr\{\overline{\mathcal{B}_1} \cap \mathcal{B}_2\}, \tag{3.11}$$

where

$$\Pr\{\overline{\mathcal{B}_1}\} = e^{-\frac{2^{R_1}-1}{(1-a)\xi} \cdot \frac{\beta_1+\beta_2}{\beta_1\beta_2}}, \tag{3.12}$$

and

$$\Pr\{\overline{\mathcal{B}_1} \cap \mathcal{B}_2\} = -\frac{\beta_1 + \beta_2}{\beta_2} e^{-\frac{2^{R_1}-1}{(1-a)\xi\beta_1}} e^{-\frac{2^{R_2}-1}{a\xi\beta_2}} + \frac{\beta_1}{\beta_2} e^{-\frac{2^{R_2}-1}{a\xi} \cdot \frac{\beta_1+\beta_2}{\beta_1\beta_2}} + e^{-\frac{2^{R_1}-1}{(1-a)\xi} \cdot \frac{\beta_1+\beta_2}{\beta_1\beta_2}}, \tag{3.13}$$

then

$$p_{\text{U-out}}(a) = 1 + \frac{\beta_1}{\beta_2} e^{-\frac{2^{R_2}-1}{a\xi} \cdot \frac{\beta_1+\beta_2}{\beta_1\beta_2}} - \frac{\beta_1 + \beta_2}{\beta_2} e^{-\frac{2^{R_1}-1}{(1-a)\xi\beta_1}} e^{-\frac{2^{R_2}-1}{a\xi\beta_2}}. \tag{3.14}$$

The following lemma provides a tight approximation for the optimum power allocation coefficient that minimizes equation (3.14), for reasonably high SNR values.

Lemma 1. *Let* $a_{opt} = \arg\min_{a \in \mathcal{A}} p_{U\text{-}out}(a)$. *For large* ξ, a_{opt} *is tightly approxi-*

mated by

$$a_{opt} \approx \min\{\tfrac{1}{2}\left(\sqrt{2\eta} - \sqrt{2\eta - 4(\tfrac{K_2}{2} + \tfrac{K_1}{2\sqrt{2\eta}+\eta})}\right) - \tfrac{L_3}{4}, \sup(\mathcal{A})\}, \tag{3.15}$$

where $K_1 = L_1 - \frac{1}{2}L_2L_3 + \frac{1}{8}L_3^3$, $K_2 = L_2 - \frac{3}{8}L_3^2$, *and*

$$\begin{aligned} L_1 &= -\frac{2^{R_2}-1}{\beta_2\xi}\left(\frac{3(2^{R_2}-1)}{2^{R_1}-1}+1\right), \\ L_2 &= \frac{3(2^{R_2}-1)}{\beta_2\xi}\left(\frac{2^{R_2}-1}{2^{R_1}-1}+1\right) \\ L_3 &= \frac{2^{R_1}-1}{\beta_1\xi} - \frac{2^{R_2}-1}{\beta_2\xi}\left(\frac{2^{R_2}-1}{2^{R_1}-1}+1\right) - 1, \end{aligned} \tag{3.16}$$

and η *is defined in the appendix.*

Proof. See appendix. □

3.4 Comparison of theoretical and simulation results

In figure 3.3, the propositions made in proposition 1 are demonstrated completely. The plot of $p_{\text{U-out}}(a)$ clearly shows that a should always be selected to be in $\mathcal{A}$. Once a is outside this region, $p_{\text{U-out}}(a)$ quickly approaches 1. Figure 3.4 demonstrates the significant performance gap in $p_{\text{U-out}}(a)$ for CA-NOMA over both regular NOMA and OMA for the case of users with very disparate path losses.

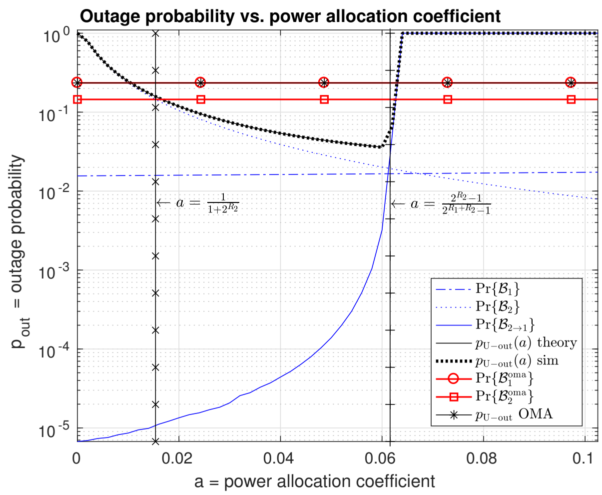

Figure 3.3: Outage probabilities are plotted vs. a; $\xi = 30$ dB, $\beta_1 = 1$, $\beta_2 = 20$, $R_1 = 4$ bps/Hz, $R_2 = 6$ bps/Hz

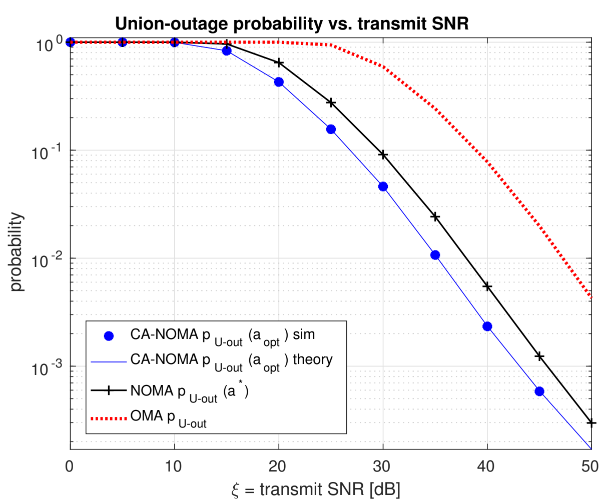

Figure 3.4: Outage probabilities are plotted vs. ξ, with a^* being optimum for regular NOMA; $\beta_1 = 1$, $\beta_2 = 100$, $R_1 = 4$ bps/Hz, $R_2 = 8$ bps/Hz

3.5 Conclusion and Future Work

The fundamental requirements on power allocation for CA-NOMA are outlined, and their effects on the set of power allocation coefficients that allow this system to function are derived. It is proven and verified through analysis and simulations that the set of power allocation coefficients and the optimum value always exist and always improve the system performance. This work focuses on the downlink case. A similar investigation of leveraging caching in uplink NOMA systems is needed.

3.6 Proofs

The proofs of the items in the proposition are provided below.

3.6.1 Proof of proposition 1

1. If $\mathcal{B}_{21} \subseteq \mathcal{B}_{22}$, then

$$\frac{2^{R_1} - 1}{\xi(1 - a2^{R_1})} \geq \frac{2^{R_2} - 1}{a\xi}. \tag{3.17}$$

   Solving the above inequality for a provides the result.

2. $\mathcal{B}_1 \subseteq \mathcal{B}_1^{\text{oma}}$ implies $\frac{2^{R_1}-1}{(1-a)\xi} \leq \frac{4^{R_1}-1}{\xi}$. Solving for a gives

$$a \leq \frac{2^{R_1}}{2^{R_1} + 1},$$

   which increases as R_1 increases. Since $\mathcal{B}_2 = \mathcal{B}_{22}$ implies that $a \leq \frac{2^{R_2}-1}{2^{R_1+R_2}-1}$, the minimum of these two constraints is determined by

$$\frac{2^{R_1}}{2^{R_1}+1} \lesseqgtr \frac{2^{R_2}-1}{2^{R_1+R_2}-1} \tag{3.18}$$

$$\Rightarrow 2^{2R_1} - 2^{R_1} + 2^{-R_2} - 1 \lesseqgtr 0 \tag{3.19}$$

$$\Rightarrow R_1 \lesseqgtr \log_2\left(\frac{1+\sqrt{5-4\times 2^{-R_2}}}{2}\right) \tag{3.20}$$

Letting $R_2 \to \infty$ gives the desired result.

3. $\mathcal{B}_2 \subseteq \mathcal{B}_2^{\text{oma}}$ implies

$$\frac{2^{R_2}-1}{a\xi} \leq \frac{4^{R_2}-1}{\xi} \;\Rightarrow\; a \geq \frac{1}{2^{R_2}+1}. \tag{3.21}$$

The set $\mathcal{A}_{\text{FN}}$ is always non-empty if

$$\frac{1}{2^{R_2}+1} < \frac{2^{R_2}-1}{2^{R_1+R_2}-1},$$

which is always true as long as $R_1 < R_2$.

4. If

$$\frac{2^{R_1}-1}{(1-a)\xi} < \frac{2^{R_2}-1}{a\xi} \;\Rightarrow\; a < \frac{2^{R_2}-1}{2^{R_1}+2^{R_2}-2}. \tag{3.22}$$

However, since $a \in \mathcal{A} \Rightarrow a < \frac{2^{R_2}-1}{2^{R_1+R_2}-1}$, then

$$\frac{2^{R_2}-1}{2^{R_1+R_2}-1} < \frac{2^{R_2}-1}{2^{R_1}+2^{R_2}-2} \tag{3.23}$$

$$\Rightarrow 0 < 2^{R_1+R_2} - 2^{R_1} - 2^{R_2} + 1 = (2^{R_1}-1)(2^{R_2}-1), \tag{3.24}$$

which is true $\forall R_1, R_2 > 0$.

3.6.2 Proof of lemma 1

The minimum is found by taking the derivative with respect to a, which is

$$\frac{dp_{\text{U-out}}(a)}{da} = \frac{(\beta_1+\beta_2)(2^{R_2}-1)}{a^2\xi\beta_2^2} e^{-\frac{2^{R_1}-1}{a\xi}\cdot\frac{\beta_1+\beta_2}{\beta_1\beta_2}} + \frac{\beta_1+\beta_2}{\beta_2}\left(\frac{2^{R_1}-1}{(1-a)^2\xi\beta_1} - \frac{2^{R_2}-1}{a^2\xi\beta_2}\right) e^{-\frac{2^{R_1}-1}{(1-a)\xi\beta_1}-\frac{2^{R_2}-1}{a\xi\beta_2}}.$$

Using the power series of the exponential, $e^{-t} = \sum_{k=0}^{\infty}\frac{(-t)^k}{k!} = 1 - t + O(t^2)$, the above can be approximated for high SNR using the first two terms of the power series. Setting this approximation equal to zero and some algebraic manipulations gives the following

$$0 = a^4 + L_3a^3 + L_2a^2 + L_1a + L_0, \;\; L_0 = \frac{(2^{R_2}-1)^2}{\xi\beta_2(2^{R_1}-1)} \tag{3.25}$$

where L_1, L_2, and L_3 are defined in the lemma statement. This equation is the standard *quartic* equation, which is solved [29] by substituting $a = b - \frac{L_3}{4}$ and solving for b. This results in

$$0 = b^4 + K_2b^2 + K_1b + K_0, \tag{3.26}$$

where $K_0 = L_0 - \frac{1}{4}L_1L_3 + \frac{1}{16}L_2L_3^2 - \frac{3}{256}L_3^4$, and K_1 and K_2 are defined in the lemma statement. By adding $K_2^2/4$ to both sides of the equation and manipulating, the following is obtained

$$(b^2 + \frac{K_2}{2})^2 = \frac{K_2^2}{4} - K_1b - K_0. \tag{3.27}$$

Adding the quantity $\eta^2 + K_2\eta + 2b^2\eta$ to both sides and simplifying results in

$$(b^2 + \frac{K_2}{2} + \eta)^2 = 2b^2\eta - K_1b + \frac{K_2^2}{4} - K_0 + \eta^2 + K_2\eta. \tag{3.28}$$

The right side above is a perfect square if

$$0 = \eta^3 + K_2\eta^2 + (\frac{1}{4}K_2^2 - K_0)\eta - \frac{1}{8}K_1^2, \tag{3.29}$$

which is the general cubic equation, which is solved in [28] and demonstrated in [27]. Call $q_0 = \frac{1}{108}K_2^3 + \frac{1}{3}K_0K_2 - \frac{1}{8}K_1^2$ and $q_1 = -\frac{1}{12}K_2^2 - K_0$, then

$$\eta = -\frac{K_2}{3} + \sqrt[3]{q_0 + \left(q_0^2 + \frac{4}{27}q_1^3\right)^{\frac{1}{2}}} - \sqrt[3]{-q_0 + \left(q_0^2 + \frac{4}{27}q_1^3\right)^{\frac{1}{2}}}. \tag{3.30}$$

Using this value for η simplifies (3.28) into

$$b^2 - b\sqrt{2\eta} + \eta + \frac{K-2}{2} + \frac{K_1}{2\sqrt{2\eta}} = 0 \tag{3.31}$$

Solving this quadratic and substituting back into $a = b - \frac{L_3}{4}$, while also keeping in mind that a must be in $\mathcal{A}$ yields the result in equation (3.15).

Chapter 4

Fundamentals of Power Allocation Strategies for Downlink Multi-user NOMA with Target Rates

For downlink multi-user non-orthogonal multiple access (NOMA) systems with successive interference cancellation (SIC) receivers, and a base-station not possessing the instantaneous channel gains, the fundamental relationship between the target rates and power allocation is investigated. It is proven that the total interference from signals not removed by SIC has a fundamental upper limit which is a function of the target rates, and the outage probability equals one when exceeding this limit. The concept of well-behaved power allocation strategies is defined, and its properties are proven to be derived solely based on the target rates. The existence of power allocation strategies that enable NOMA to outperform OMA in per-user outage probability is proven, and are always well-behaved for the case when the outage probability performance of NOMA and OMA are equal for all users. The proposed SIC decoding order is then shown to be the most energy efficient. The derivation of well-behaved power allocation strategies that have improved outage probability performance over OMA for each user is outlined. Simulations validate the theoretical results, demonstrating that NOMA systems can always outperform OMA systems in outage probability performance, without relying on the exact channel gains.

The results in this chapter have been published in [43].

4.1 Introduction

Due to the rapidly increasing demand for higher data-rates, more connected users and devices, and diversity of deployments, power-domain non-orthogonal multiple access (NOMA) is being sought to help improve the capacity and user multiplexing of downlink (DL) and uplink cellular systems [1]. The 3rd Generation Partnership Project has already conducted the study items for both downlink NOMA for LTE-Advance [3], and for uplink NOMA in New Radio (NR) [4]. With the attention that NOMA receives from the academic, private, and standard sectors, it is only a matter of time before NOMA is implemented in future wireless system deployments.

Consider BS serving multiple users in a cell as depicted in figure 4.1. The BS's scheduler will determine whether the users can support the target rates required of their requested information, given a certain time-frequency resource allocation. Assuming that the BS determines that a set of users' channel can support their respective target rates, it will then schedule a NOMA downlink transmission to these users.

Although it is proven in [22] that there always exists a power allocation ap-

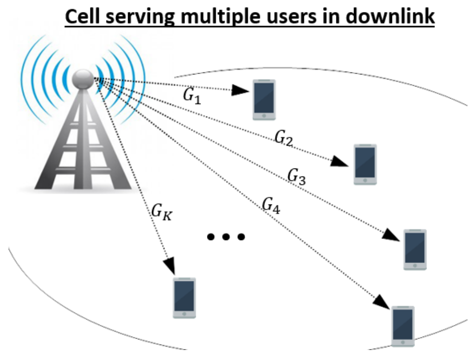

Figure 4.1: A BS serving multiple users in a cell in the downlink.

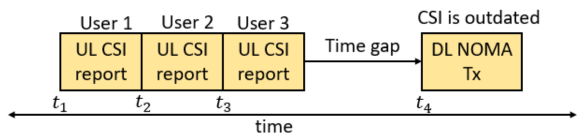

Figure 4.2: An example of three users sending their CSI reports to the BS, and the time-gap between reports and the downlink transmission.

proach for NOMA that can outperform OMA for the general multi-user NOMA case in terms of information capacity, this power allocation strategy relies on the BS having perfect instantaneous channel state information (CSI) at the transmitter, i.e. the exact channel gain value, which is not a realistic assumption in wireless system deployments. This is due to the limitations of the CSI formats that are fed back by the users to the BS, and time gaps between channel estimation by the BS and the associated downlink transmissions as shown in figure 4.2. In 4G and 5G system deployments [30], the BS determines whether a target rate can be supported based on realistic CSI formats (rank indicator, precoding matrix indicator, channel quality indicator, etc.), and selects the remaining transmission parameters which will accommodate the downlink transmission at the indicated target rate. Therefore, it is important that a DL NOMA system be able to determine the power allocation for all NOMA users based on the available information in real system deployments, and not the exact channel gain values.

In this chapter, it is assumed that the BS has determined that the target rates can be supported by the channels for all users, but that it does not possess the full CSI in the form of the channel gains, and must determine the NOMA power allocation strategy based on the available information. For the baseline OMA system which NOMA will be compared to, a general TDMA approach is used with each user being allocated a fraction of the total transmit time duration, and the BS uses full transmit power for each transmission. For the NOMA system, the transmit power allocation strategy and associated SIC decoding order are completely determined by the users' target data rates and the associated OMA transmit time durations. It is proven that a power allocation strategy must be such that the received interference coefficient for each signal must be below a fundamental threshold in order to not experience unavoidable outages. The concept of *well-behaved* power allocation strategies is defined, and shown that these strategies satisfy the interference requirement. It is then proven that there always exists a power allocation strategy such that all users will have NOMA outage probability performance equal to that of OMA, that the proposed SIC decoding order is the most energy efficient, and that such a strategy is always well-behaved. Finally, the approach to derive a well-behaved power allocation strategy is outlined, such that a user can achieve better outage probability performance with NOMA over OMA.

4.2 Previous Work and Current Contribution

The outage probability of NOMA was investigated in [8], where multiple users transmit simultaneously to multiple receivers using a uniform power allocation approach, and it was shown the outage probability is improved when NOMA is combined with H-ARQ vs OMA with H-ARQ. The authors in [14] showed that the power allocation and interference coefficients of each user are fundamentally dependent on the particular user's required rate, and thus the wrong selection of coefficients can lead to an outage with probability equal to 1. The usage of NOMA in a cognitive-radio inspired approach was mentioned in [18], where a user with weak channel condition is seen as the primary user and is provided as much power as needed in order to achieve its minimum rate, and the user with stronger channel is treated as the secondary user and receives any remaining power not allocated to the weaker user, and the outage probability of both is shown to clearly depend on pairing users with stronger channels.

A couple of works have focused on utilizing the rate achieved using OMA as the minimum rate required for NOMA, and the associated power allocation solution which achieves this condition. The region of power allocation coefficients that allow NOMA to outperform OMA in the downlink is first defined for the two-user case in [19]. The authors in [21] then use a power allocation approach in this region to analyze the outage performance and diversity orders of two paired users, according to their relative channel gains, and extend the work to the uplink case. In [22], the power allocation coefficients for a multi-user NOMA system which always outperforms OMA are proven to always have a sum less than or equal to 1, and hence a valid power allocation strategy for NOMA always exists that outperforms OMA in terms of capacity, while using less power than OMA. The work in [31] extends the concept of power allocation fairness with regards to NOMA compared to OMA, showing there always exists a power allocation for NOMA that allows the rate to outperform the rates of the generalized FDMA case with optimizing resource allocation. In [32], the authors prove that for any power and resource allocation in FDMA, there always exists a power allocation strategy that will provide a superior sum-rate and ergodic rate for NOMA over

OMA, while developing a user admission scheme to maintain a balance between the number of total admitted users and sum-rate performance.

More recent works have focused on optimizing the power allocation strategy. The authors in [33] propose a joint optimization of user pairing and power allocation by optimizing a cost function dependent on the instantaneous achievable rates and a metric based on proportional fairness. The scheduling and power allocation algorithm that solves the optimization problem is compared to the fractional transmit power control algorithm and shown to improve performance for the user with stronger channel, while performance is not always improved for the user with weaker channel. The work in [34] uses a new algorithm to solve the cognitive radio NOMA power allocation problem which can outperform the fractional transmit power algorithm for admitting secondary users into the network. In [35], the authors seek to optimize the sum-rate of a multi-user downlink NOMA system by using a constraint based on the total power allocated to the signals at each SIC stage, and its relation to the minimum required rate for each signal to be decoded. The authors in [36] studied several algorithms that solve the NOMA power allocation problem, and point out that not many existing works had considered the strict constraint for the power allocations to follow the order of SIC decoding in their algorithms. They proposed to incorporate the matching algorithm and optimum power allocation, and found that the constraint has a significant impact on the power allocation solution, which also yields superior performance over existing schemes. In [37], a new solution is proposed for a system that clusters users in order to solve the joint beamforming and power allocation problem by breaking the problem up into two separate sub-problems, where the goal is to maximize the sum-rate of each cluster. In [38], the authors propose a joint resource (bandwidth) and power allocation approach that optimizes a cost function that is an affine function of the power allocations and bandwidths. It is demonstrated that this algorithm can outperform the approach of simply optimizing the power allocations with fixed bandwidths, as well as the baseline OMA approach. The work in [39] proposes a joint beamforming and power allocation solution to a coordinated multi-point MIMO-NOMA system, where the intra-cell interference between clusters is cancelled through transmit beamforming, and the power allocation is designed to maximize one user's rate, while maintaining the rate of the second user. In [40], the authors derive a weighted sum-rate maximization algorithm to find the power allocation per subcarrier for a pair of users, and show that the performance of their approach improves as the diversity order of the system increases. In [41], the authors study the power allocation approach for multi-tiered cellular networks with cell-center users and a cell-edge user who is eligible for coordinated multi-point transmission. A joint power optimization algorithm is formulated, including target rates for each user, and due to the prohibitive complexity, a distributed power optimization problem is formulated and shown to exhibit near optimum performance. The constraint on the power allocation coefficients relies on a linear function derived from the SINR for each SIC stage.

In the previous works, the power allocation constraints either do not consider the necessary requirements for successful SIC performance, or use constraints that do not give the fundamental relationship between power allocation and outage, such that an outage event is certainly avoided. In other words, the target rates and power allocation required to ensure whether successful SIC performance is even feasible at each stage of SIC decoding for each user is not *directly* considered, and this can cause unnecessary unavoidable outages to occur for multiple users. In fact, this phenomenon was described and demonstrated in [27] for two-user cache-aided NOMA systems, and in [42] for multi-user downlink NOMA systems with a QoS constraint.

The main contribution of this chapter is to provide a comprehensive theoretical treatment of power allocation strategies, and how they are related to the SIC decoding order selected, the associated target rates of the users, and any associated OMA parameters that affect the design of the NOMA power allocation coefficients, while not relying on the channel gain value. In particular, this chapter provides the following:

- The fundamental maximum interference that a particular signal can tolerate from other NOMA users before its outage probability is equal to 1, regardless of how strong the channel SNR gains are;
- The definition of a *well-behaved* power allocation strategy, which causes the outage thresholds to be lesser for the signals of users that are earlier in the SIC decoding order. This condition is then shown to always lead to an acceptable value of NOMA interference;
- The baseline power allocation strategy which achieves outage performance equal to that of OMA for all users is derived in closed-form, and is used to prove that the proposed decoding order is energy efficiency optimal;
- The baseline power allocation strategy is used to derive the conditions for acquiring a power allocation strategy where all users have superior NOMA outage performance over OMA, the exact approach for increasing the power allocation beyond the baseline strategy is outlined in detail, and a quick example of a power allocation strategy that satisfies all of these conditions is presented along with its performance.

The necessity of such results in further studies of NOMA is clear, in the sense that when performing numerical studies of different algorithms applied to solve the power allocation problem for complex cellular deployments, the search space for the multi-user power allocation strategies can be greatly reduced to the subset of strategies that are well-behaved and improve the outage performance of NOMA over OMA.

4.3 System Model

Consider a wireless downlink system serving K users. The BS will transmit K multiplexed signals to the K users. Let the signal for user n be x_n, $n = 1, \ldots, K$, such that x_n is complex normally distributed with $\mathbb{E}[|x_n|] = 1$, and is transmitted with transmit SNR ξ through a wireless slow fading channel with SNR gain $|G_n|^2$. The channel gain G_n can be one of many fading channels, such as a Rayleigh fading channel $|G_n|^2 \sim \text{Exponential}(\beta_n)$, where the value β_n can depend on the distance from the BS, or a MIMO fading channel $\boldsymbol{H}_n$ with precoding vector $\boldsymbol{p}$ at the transmitter and detection vectors $\boldsymbol{v}_n$ at the receivers, such that the overall channel SNR gain is $|G_n|^2 = |\boldsymbol{v}_n \boldsymbol{H}_n \boldsymbol{p}|^2$. The channel gain is not assumed to be known at the BS.

In the case of OMA, the general TDMA model is used and this resource allocation is depicted in figure 4.3. For a normalized total transmit time duration, the fractional time duration allocated to user n is τ_n, such that $\sum_{n=1}^{K} \tau_n = 1$. The received signal at user n is given by $y_n = \sqrt{\xi} G_n x_n + z_n$, where $z_n \sim \mathcal{CN}(0, 1)$ is the receiver thermal noise. Since user n is allocated τ_n fraction of the total time resource, the capacity of user n using OMA is given by $C_n^{\text{oma}} = \tau_n \log_2(1 + \xi |G_n|^2)$.

For the NOMA system, user n has power allocation coefficient a_n, such that

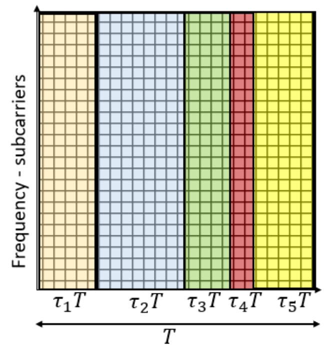

Figure 4.3: The time resource is allocated to each users' transmission by the BS.

$\sum_{n=1}^{K} a_n = 1$. The received signal at user n is

$$r_n = G_n \sum_{l=1}^{K} x_l \sqrt{a_l \xi} + z_n. \tag{4.1}$$

Using SIC, the receiver at user n, $n > 1$, will decode the messages of users $m < n$ in ascending order, starting with $m = 1$ (the SIC decoding order details are discussed in section 4.4). Therefore, user n will perform SIC on the signals of user $m = 1, \ldots, n-1$, which have the form

$$y_{n \to m} = G_n \left(\underbrace{x_m \sqrt{a_m \xi}}_{\text{user-}m\text{ signal for SIC}} + \sum_{l=m+1}^{K} x_l \sqrt{a_l \xi} \right) + z_n, \tag{4.2}$$

until it can obtain the intended signal for user n given by

$$y_n = G_n \left(\underbrace{x_n \sqrt{a_n \xi}}_{\text{user-}n\text{ signal}} + \sum_{l=n+1}^{K} x_l \sqrt{a_l \xi} \right) + z_n, \tag{4.3}$$

where $\sum_{l=n+1}^{K} x_l \sqrt{a_l \xi}$ are the signals that need not be decoded using SIC by user n in order to decode its own signal, and thus are treated as interference.

For the power allocation coefficients $a_1, \ldots, a_K$, the capacity of the channel for user $n < K$ is

$$C_n(a_1, \ldots, a_K) = \log_2 \left(1 + \frac{a_n \xi |G_n|^2}{1 + \xi |G_n|^2 \sum_{l=n+1}^{K} a_l} \right), \tag{4.4}$$

and user K has capacity

$$C_K(a_1, \ldots, a_K) = \log_2 \left(1 + a_K \xi |G_K|^2 \right). \tag{4.5}$$

Meanwhile, for each user n to achieve its capacity, it must have the capacity to decode the messages sent to all users $m < n$, and then subtract their signals from the composite signal received. The capacity of the channel which user n will use to decode user m's message is given by

$$C_{n \to m}(a_1, \ldots, a_K) = \log_2 \left(1 + \frac{a_m \xi |G_n|^2}{1 + \xi |G_n|^2 \sum_{l=m+1}^{K} a_l} \right). \tag{4.6}$$

4.4 Basics of NOMA power allocation for systems with target rates

Let each user n have its information transmitted at a target rate R_n. First, define the event when user n experiences an outage in an OMA system as

$$\mathcal{B}_n^{\text{oma}} = \{C_n^{\text{oma}} < R_n\} = \left\{ |G_n|^2 < \frac{2^{R_n/\tau_n} - 1}{\xi} \right\}. \tag{4.7}$$

Since the goal of NOMA is to outperform OMA with respect to certain metrics (outage probability in this study), the OMA parameter τ_n and associated outage events affect the selection of the NOMA power allocation strategy. Considering that the OMA outage event can be normalized by dividing by τ_n, yielding the

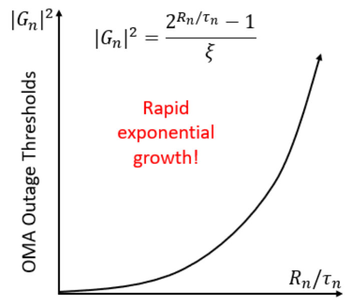

Figure 4.4: The OMA outage threshold grows exponentially with $\frac{R_n}{\tau_n}$, hence NOMA SIC decoding order should follow the ordering of this ratio.

normalized rate $\frac{R_n}{\tau_n}$, this normalized rate can be used as the quantity for determining the SIC decoding order. This quantity can also be seen from equation (4.7) as the determining factor for the value of the OMA outage threshold, due to the fact that the outage is an exponential function of this ratio as shown in figure 4.4. So selecting the decoding order based on increasing OMA outage thresholds seems intuitive, since the NOMA outage thresholds will be directly compared to the OMA outage thresholds when finding the power allocation strategy.

Let the ordering of the user indices follow the ordering of the relationship $\frac{R_n}{\tau_n}$, such that indices $(1, \ldots, K)$ correspond to $\frac{R_1}{\tau_1} < \cdots < \frac{R_K}{\tau_K}$. A user $n = 1, \ldots, K$, will experience an outage during the decoding process of its information if any of the following occurs:

$$C_n(a_1, \ldots, a_K) < R_n \quad \text{OR} \quad C_{n \to m}(a_1, \ldots, a_K) < R_m, \tag{4.8}$$

for any $m < n$. Define the following events based on the specific signal which user n needs to detect and decode, where $n = 2, \ldots, K$, and $m < n$,

$$\mathcal{B}_n = \{C_n(a_1, \ldots, a_K) < R_n\} \tag{4.9}$$
$$\mathcal{B}_{n \to m} = \{C_{n \to m}(a_1, \ldots, a_K) < R_m\}.$$

The NOMA outage event $\mathcal{B}_n^{\text{out}}$ at user n can then be described as

$$\mathcal{B}_n^{\text{out}} = \mathcal{B}_n \cup \left(\bigcup_{m=1}^{n-1} \mathcal{B}_{n \to m} \right). \tag{4.10}$$

Note that $\mathcal{B}_1^{\text{out}} = \mathcal{B}_1$ because user 1 does not perform SIC in order to decode its own signal.

4.4.1 Certain outage in NOMA transmissions

From the definition of the NOMA outages, the following theorem is obtained.

Theorem 4.4.1. *For a K-user DL NOMA system with user target rates $R_1, \ldots, R_K$ and power allocation coefficients $a_1, \ldots, a_K$, define $A_n = \sum_{l=n+1}^{K} a_l, \forall n = 1, \ldots, K-1$, which is the interference coefficient in the received signal that users $n, \ldots, K$ will use to detect and decode user n's information. If $\exists n$ such that $A_n > 2^{-\sum_{l=1}^{n} R_l}$, then for user n and $\forall l > n$,*

$$\Pr\{\mathcal{B}_n\} = \Pr\{\mathcal{B}_{l \to n}\} = 1, \tag{4.11}$$

and thus SIC will fail for all users $l = n, \ldots, K$.

Proof. See appendix 4.8.1. □

Theorem 4.4.1 demonstrates that there is a fundamental relationship between the set of target rates R_n and associated power allocation coefficients a_n, $n = 1, \ldots, K$. It also demonstrates that as these target rates increase, the values of a_n decrease rapidly, indicating that as the target rates increase for users earlier in the SIC decoding order, the amount of available power to the users later in the decoding order decreases.

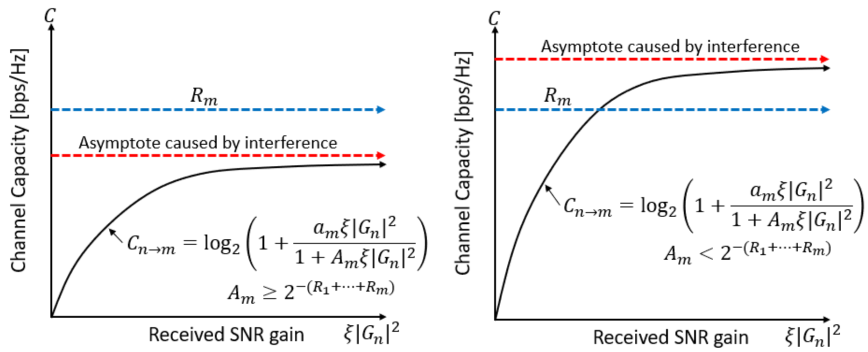

Figure 4.5: Left: The capacity to decode signal m will always fall below the target rate when A_m exceeds the threshold.
Right: The capacity to decode signal m can be met by channels with large enough received SNR gains when A_m is below the threshold.

Note that this does not indicate that the rate for user n is guaranteed if $A_n < 2^{-\sum_{l=1}^{n} R_l}$, since the total power allocation available to users $n, \ldots, K$ may be less than $2^{-\sum_{l=1}^{n} R_l}$ to begin with. In fact, a bound that is more case specific to the actually selected power allocation coefficients, as outlined in [14], is

$$A_{n-1} > A_n 2^{R_n}, n = 2, \ldots, K - 1. \tag{4.12}$$

However, although it is a more strict bound, it is dependent on the specific case of power allocation coefficients, whereas the bound provided in theorem 4.4.1 is a fundamental upper limit on the received interference coefficient that cannot be exceeded by any power allocation scheme. So a set of power allocation coefficients that satisfy equation (4.12) also satisfy theorem 4.4.1. An illustration of the channel capacity limit of an interference channel is provided in figure 4.5. The channel capacity of an interference channel will always have an upper limit, and hence the larger the interference, the lower this limit becomes.

With the assumption that the power allocation coefficients are selected such that $A_n = \sum_{l=n+1}^{K} a_l < 2^{-\sum_{l=1}^{n} R_l}$, it should also be noted that in order for the SIC process to reach a decoding stage n, it does so with a certain probability at each user $k \geq n$. In other words, if users $k = n, \dots, K$ are going to avoid an outage, they must sequentially decode messages $m = 1, \dots, n$ successfully in the process. Given the sequential nature of the decoding process, it is therefore desirable that the initial decoding stages have lower outage probabilities.

4.4.2 Well-behaved power allocation strategies

In light of theorem 4.4.1, the outage events $\mathcal{B}_n, n = 1, ..., K-1$, and $\mathcal{B}_{n\to m}, n = 2, ..., K$, can be rewritten as

$$\mathcal{B}_n = \left\{ |G_n|^2 < \frac{2^{R_n} - 1}{\xi(a_n - (2^{R_n} - 1)\sum_{l=n+1}^{K} a_l)} \right\} \tag{4.13}$$

$$\mathcal{B}_{n\to m} = \left\{ |G_n|^2 < \frac{2^{R_m} - 1}{\xi(a_m - (2^{R_m} - 1)\sum_{l=m+1}^{K} a_l)} \right\},$$

for $m = 1, \dots, n-1$, and $\mathcal{B}_K = \left\{ |G_K|^2 < \frac{2^{R_K}-1}{a_K \xi} \right\}$. This means that the overall outage event $\mathcal{B}_n^{\text{out}}$ can be expressed as

$$\mathcal{B}_n^{\text{out}} = \left\{ |G_n|^2 < \frac{2^{R_n} - 1}{\xi(a_n - (2^{R_n} - 1)\sum_{l=n+1}^{K} a_l)} \right\} \cup \left(\bigcup_{m=1}^{n-1} \left\{ |G_n|^2 < \frac{2^{R_m} - 1}{\xi(a_m - (2^{R_m} - 1)\sum_{l=m+1}^{K} a_l)} \right\} \right). \tag{4.14}$$

It is not desirable that a user n's outage probability be primarily dictated by the success or failure of the earlier decoding stages. Since each event in equation (4.14) is determined by a finite length interval in the form of $(0, \alpha) \subset \mathbb{R}^+$, it is clear that $\exists k, 1 \leq k \leq n$, such that $\forall m = 1, \dots, n$,

$$\frac{2^{R_k} - 1}{\xi(a_k - (2^{R_k} - 1)\sum_{l=k+1}^{K} a_l)} \geq \frac{2^{R_m} - 1}{\xi(a_m - (2^{R_m} - 1)\sum_{l=m+1}^{K} a_l)}, \tag{4.15}$$

$\Rightarrow \mathcal{B}_n^{\text{out}} = \mathcal{B}_k$. This leads to the following proposition.

Proposition 2. *For a K user DL NOMA system with target rates $R_1, \ldots, R_K$, if the associated power allocation coefficients $a_1, \ldots, a_K$ are selected such that $\mathcal{B}_n^{out} = \mathcal{B}_n, \forall n = 1, \ldots, K$, then*

$$a_n \geq a_{n+1} \frac{2^{R_{n+1}}(2^{R_n} - 1)}{2^{R_{n+1}} - 1}, n = 1, \ldots, K-1, \tag{4.16}$$

and $A_n = \sum_{l=n+1}^{K} a_l < 2^{-\sum_{l=1}^{n} R_l}$, satisfying the requirement from theorem 4.4.1.

Proof. See appendix 4.8.2. □

The above proposition provides the relationship between the power allocation strategies and the desired outage events. In other words, the outage probability to decode user n's information should not be determined by an outage event during a SIC stage, but by the outage event of its own signal. Also, note that this condition also favors the decoding probability of all users whose signals are earlier in the SIC decoding order, as it places a lesser upper bound on the amount of NOMA interference received. From here on, any power allocation strategy which satisfies proposition 2, and by extension theorem 4.4.1, will be defined as being a *well-behaved* strategy.

The concept of *well-behaved* is not simply a preference, but an essential component for selection of an efficient NOMA power allocation strategy which aims to improve the outage performance over OMA for any user n without having their performance sabotaged by an earlier SIC decoding stage m. For example, suppose $\exists m$ and n, $m < n$, for a non-well-behaved power allocation strategy such that $\mathcal{B}_n \subset \mathcal{B}_{m\to n}$, then $\mathcal{B}_n^{\text{out}} = \mathcal{B}_{m\to n}$. Furthermore, let the power allocation strategy be such that $\mathcal{B}_m = \mathcal{B}_m^{\text{oma}}$. This means that the outage probability of user n is no longer a function a_n because $\mathcal{B}_n^{\text{out}} = \left\{ |G_n|^2 < \frac{2^{R_m}-1}{\xi(a_m - (2^{R_m}-1)\sum_{l=m+1}^{K} a_l)} \right\}$ remains constant. Hence, power allocation to user n can essentially be increased without any benefit to performance, which is something that should be avoided.

4.4.3 NOMA power allocation strategies that achieve outage performance equal to OMA

Another requirement for a NOMA power allocation strategy is that the outage probability performance is equal to or better than the outage probability performance of OMA. First, the following power allocation strategies are formally defined in the following.

Definition 2. *For a user n:*

(i) The power allocation coefficient a_n^{oma} is defined as the exact power allocation required such that user n achieves the same outage probability performance as it would achieve using OMA. In other words, $\mathcal{B}_n = \mathcal{B}_n^{oma}$;

(ii) The power allocation coefficient $\tilde{a}_n^{oma}$ is defined as the minimum power allocation such that $\mathcal{B}_n = \mathcal{B}_n^{oma}$, which can only be applied when all users $l = n+1, \ldots, K$ also have power allocation $\tilde{a}_l^{oma}$;

(iii) The interference coefficient $A_n^{oma} = \sum_{l=n+1}^{K} \tilde{a}_l^{oma}$.

Any power allocation strategy that improves the outage probability performance over OMA can be written as $(a_1^{\mathrm{oma}} + \epsilon_1, \ldots, a_K^{\mathrm{oma}} + \epsilon_K)$. If $\epsilon_n = 0, \forall n = 1, \ldots, K$, then all users will achieve the same outage probability performance as OMA, and $a_n^{\mathrm{oma}} = \tilde{a}_n^{\mathrm{oma}}, \forall n = 1, \ldots, K$. The following theorem shows that there always exists a power allocation strategy such that the NOMA outage probabilities for all users are equal to or less than the respective OMA outage probabilities.

Theorem 4.4.2. *For a K-user DL NOMA system with target rates $R_1, \ldots, R_K$, there always exists a power allocation strategy $(a_1, \ldots, a_K)$ with associated SIC decoding order $(1, \ldots, K)$ such that $\mathcal{B}_n \subseteq \mathcal{B}_n^{oma}, \forall n = 1, \ldots, K$. Furthermore, $\exists$ at least one user n such that $\mathcal{B}_n \subset \mathcal{B}_n^{oma}$, meaning that the NOMA outage probability performance can always be at least as good or better than the OMA outage probability performance for every user.*

Proof. See appendix 4.8.3. □

According to equation (4.41), $(\tilde{a}_1^{\text{oma}}, \ldots, \tilde{a}_K^{\text{oma}})$ is given by $\tilde{a}_K^{\text{oma}} = \frac{2^{R_K}-1}{2^{R_K/\tau_K}-1}$ and

$$\tilde{a}_n^{\text{oma}} = \frac{2^{R_n}-1}{2^{R_n/\tau_n}-1} + \frac{2^{R_n}-1}{2^{R_n}} \sum_{l=n+1}^{K} \frac{2^{R_l}-1}{2^{R_l/\tau_l}-1} \prod_{k=n}^{l-1} 2^{R_k},$$

$n = 1, \ldots, K-1$. Based on the previous theorem, it is clear that $\sum_{n=1}^{K} \tilde{a}_n^{\text{oma}} < 1$, and that the improvement of the outage probability performance of NOMA over OMA is based on the design of the additional power allocation ϵ_n to each coefficient a_n^{oma}, and the strategy $(\tilde{a}_1^{\text{oma}}, \ldots, \tilde{a}_K^{\text{oma}})$ is the starting point. A consequence of theorem 4.4.2 is that it can be used to highlight the fact that the decoding order based on increasing values of $\frac{R_n}{\tau_n}$ is an essential component of the power allocation strategy.

Corollary 1. *Let the user indices $1, \ldots, K$ be assigned such that they follow the relationship $\frac{R_1}{\tau_1} < \cdots < \frac{R_K}{\tau_K}$. Also, define a SIC decoding order $(\sigma(1), \ldots, \sigma(K))$, such that $(\sigma(1), \ldots, \sigma(K))$ is a permutation of the sequence $(1, \ldots, K)$. For all SIC decoding orders $(\sigma(1), \ldots, \sigma(K))$ which have associated power allocation strategies $(\tilde{a}_{\sigma(1)}^{oma}, \ldots, \tilde{a}_{\sigma(K)}^{oma})$ such that NOMA achieves equal outage performance to OMA,* *the SIC decoding order*

$$(1, \ldots, K) = \arg \min_{(\sigma(1), \ldots, \sigma(K))} \sum_{n=1}^{K} \tilde{a}^{oma}_{\sigma(n)}. \tag{4.17}$$

In other words, the most energy efficient power allocation strategy is the one with SIC decoding order

$$(\sigma(1), \ldots, \sigma(K)) = (1, \ldots, K). \tag{4.18}$$

.

Proof. See appendix 4.8.4. □

This corollary states that the most energy efficient power allocation strategy which enables NOMA outage performance equal to that of OMA is based on the SIC decoding order which follows the increasing order of $\frac{R_n}{\tau_n}$. The most important aspect of this result is that this SIC decoding order provides the most power allocation headroom in order to improve the outage performance of NOMA over OMA. In the case that user m has power allocation greater than $\tilde{a}^{\mathrm{oma}}_m$, then clearly all users $n = 1, \ldots, m-1$ will have to allocate additional power in order for $\mathcal{B}_n = \mathcal{B}^{\mathrm{oma}}_n$. Furthermore, any power allocation strategy should be demonstrated to be well-behaved. The fundamental properties of well-behaved NOMA power allocation strategies which demonstrate better outage probability performance over OMA are discussed in the next section.

4.5 Well-behaved power allocation strategies that improve NOMA outage probability performance over OMA

In order to determine how to construct a well-behaved power allocation strategy which improves NOMA outage probability performance over OMA, the power allocation strategy that satisfies theorem 4.4.2 must be generalized. Since a power allocation coefficient for user n's signal can be described by $a_n = a_n^{\text{oma}} + \epsilon_n, \forall n$, $a_n^{\text{oma}} = \frac{2^{R_n}-1}{2^{KR_n}-1} + (2^{R_n}-1)A_n$ (where $A_n \geq A_n^{\text{oma}}$), and $a_K^{\text{oma}} = \tilde{a}_K^{\text{oma}}$, then

$$a_K = \tilde{a}_K^{\text{oma}} + \epsilon_K \tag{4.19}$$

$$a_{K-1} = a_{K-1}^{\text{oma}} + \epsilon_{K-1} = \tilde{a}_{K-1}^{\text{oma}} + \epsilon_{K-1} + (2^{R_{K-1}}-1)\epsilon_K$$

$$a_n = \tilde{a}_n^{\text{oma}} + \epsilon_n + (2^{R_n}-1)\left(\epsilon_{n+1} + \sum_{l=n+2}^{K} \epsilon_l \prod_{k=n+1}^{l-1} 2^{R_k}\right),$$

for $n = 1, \ldots, K-2$. Note that by definition 2, $a_n = a_n^{\text{oma}}$ iff $\epsilon_n = 0$, and $a_n^{\text{oma}} = \tilde{a}_n^{\text{oma}}$ iff $\epsilon_l = 0, \forall l = n+1, \ldots, K$. Furthermore, the portion of the interference coefficient caused by the terms $\epsilon_l, l = n+1, \ldots, K$ (the expression in the parenthesis above) can be expressed as

$$c_n = \epsilon_{n+1} + \sum_{l=n+2}^{K} \epsilon_l \prod_{k=n+1}^{l-1} 2^{R_k}. \tag{4.20}$$

So the general interference coefficient for user n can be written as $A_n = A_n^{\text{oma}} + c_n$.

The total available power allocation coefficient for user n is a function of $\epsilon_m, m = 1, \ldots, n-1$. This is because in a DL NOMA system, the goal is to improve the overall outage performance, and the outage performance of the users later in the SIC decoding order is more difficult to improve, as shown by the coefficient c_n. Thus, improving the performance of users with signals earlier in the SIC decoding order does not come at an additional cost for users later in the decoding order. The total available power allocation coefficient for user n is then found by noting that

$$A_{\text{tot}}^{\text{oma}} + \epsilon_1 + \sum_{m=1}^{n-1} \epsilon_m \prod_{k=1}^{m-1} 2^{R_k} \leq 1. \tag{4.21}$$

The sum of the additional power allocation for users $m = 1, \ldots, n-1$, is given by

$$d_n = \epsilon_1 + \sum_{l=2}^{n-1} \epsilon_l \prod_{k=1}^{l-1} 2^{R_k}. \tag{4.22}$$

So the additional power allocation coefficient ϵ_n for user n is a function of d_n.

Using the generalized expression of the power allocation strategy that satisfies theorem 4.4.2, the properties of ϵ_n can be found such that the power allocation strategy is well-behaved.

Theorem 4.5.1. *For users $1, \ldots, K$ with target rates $R_1, \ldots, R_K$, which are scheduled to receive signals with power allocation strategy $(a_1^{oma}+\epsilon_1, \ldots, a_K^{oma}+\epsilon_K)$, the power allocation strategy is well-behaved if each user n has one or the other of the following conditions:*

(a) $a_{n-1} = a_{n-1}^{oma}$ and $a_n = a_n^{oma}$, meaning ϵ_{n-1} and $\epsilon_n = 0$, for any $n = 2, \ldots, K$;

(b) or

$$0 < \epsilon_n < \min \left\{ \begin{array}{l} \epsilon_{n-1} \dfrac{2^{R_n}-1}{2^{R_{n-1}}-1} + \dfrac{2^{R_n}-1}{2^{R_{n-1}/\tau_{n-1}}-1} - \dfrac{2^{R_n}-1}{2^{R_n/\tau_n}-1}, \\ (1-A_{tot}^{oma}) \displaystyle\prod_{l=1}^{n-1} 2^{-R_l} - \sum_{m=1}^{n-1} \epsilon_m \prod_{l=m}^{n-1} 2^{-R_l} \end{array} \right\}. \tag{4.23}$$

Proof. See appendix 4.8.5. □

Now that the fundamental properties of well-behaved additional power allocation strategies beyond $\tilde{a}_n^{\text{oma}}$ have been described in detail, the method for selecting/designing a power allocation strategy can be discussed. Specifically, the selection/design of the power allocation strategy is completely focused on the selection of $\epsilon_n, n = 1, \ldots, K$. In other words, if an algorithm is designed to minimize the overall outage probability performance with respect to the power allocation strategy, and subject to the constraints that the performance of each user outage is better than the OMA performance, then the variables to be solved for are $(\epsilon_1, \ldots, \epsilon_K)$, and the constraints are given by theorem 4.5.1. These constraints are linear with coefficients based on the target rates $R_1, \ldots, R_K$ and OMA time durations $(\tau_1, \ldots, \tau_K)$.

However, a simpler but not optimal approach can be used to determine a power allocation strategy such that it satisfies theorem 4.5.1 by using the definition of being *well-behaved*, hence enhancing the outage probability performance of each user with respect to OMA. This is accomplished by noting that if

$$\epsilon_{n-1} > \epsilon_n 2^{R_n} \frac{2^{R_{n-1}} - 1}{2^{R_n} - 1}, n = 2, \ldots, K \tag{4.24}$$

then the power allocation strategy is well-behaved. This can be accomplished using the total addition power allocated to all users $m = 1, \ldots, n-1$ caused by adding ϵ_n to user n's power allocation. When user n has ϵ_n added to its power allocation coefficient, the BS must also add to the power allocation coefficient of users $m = 1, \ldots, n-1$ in order to maintain their outage performance, according to equations (4.53) and (4.54). This amount can be easily seen to be

$$\epsilon_n^{\text{tot}} = \epsilon_n \prod_{l=1}^{n-1} 2^{R_l}. \tag{4.25}$$

Therefore, by setting

$$\epsilon_{n-1}^{\text{tot}} = \epsilon_n^{\text{tot}} 2^{R_n} \frac{2^{R_{n-1}} - 1}{2^{R_n} - 1}, n = 2, \ldots, K \tag{4.26}$$

it can easily be shown that equation (4.24) is satisfied. Let $S = \sum_{n=1}^{K} R_n$. Solving the $K-1$ equations in equation (4.26) for ϵ_K, and using the fact that the sum of the additional power allocation coefficients is bounded by $1 - A_{\text{tot}}^{\text{oma}}$, yields

$$\begin{aligned} &\sum_{n=1}^{K} \epsilon_n^{\text{tot}} = 1 - A_{\text{tot}}^{\text{oma}} \\ \Rightarrow &\epsilon_K = (1 - A_{\text{tot}}^{\text{oma}}) \frac{2^{R_K} - 1}{2^S - 1} \prod_{l=1}^{K-1} 2^{-R_l} \\ &\epsilon_1 = (1 - A_{\text{tot}}^{\text{oma}}) \frac{2^{R_1} - 1}{2^S - 1} \prod_{l=2}^{K} 2^{R_l}, \\ &\epsilon_n = (1 - A_{\text{tot}}^{\text{oma}}) \frac{2^{R_n} - 1}{2^S - 1} \prod_{l=1}^{n-1} 2^{-R_l} \prod_{l=n+1}^{K} 2^{R_l}, \end{aligned} \tag{4.27}$$

for $n = 2, \ldots, K-1$.

This simple strategy will use all of the power allocation available, while improving the outage probability performance of all K users when employing NOMA over OMA. Note that this strategy also heavily distributes the remaining available power allocation coefficient $1 - A_{\text{tot}}^{\text{oma}}$ in favor of the users whose signals are earlier in the SIC decoding order. This is in line with what is expected with DL NOMA systems with SIC enabled receivers, where users whose signals are decoded first will have their interference removed, and thus the additional power allocation coefficient ϵ_n will also improve the SIC performance of users $l = n+1, \ldots, K$. While users whose signals are later in the SIC decoding order cause interference which in turn causes all users $m = 1, \ldots, n$ to have their power allocation coefficient bumped up in order to maintain the same performance, and thus creating the case where less additional power allocation is actually available and gains are marginal.

4.6 Comparison of theoretical and simulation results

For the simulation results, two different fading channel scenarios are used to demonstrate the validity of the theoretical results. The first fading channel model is the K SISO Rayleigh fading channel, with channel gains $h_{n_1}, \ldots, h_{n_K}$, such that $|h_{n_i}| \sim \text{Exponential}(1), i = 1, \ldots, K$. The channel SNR gains $|G_1|^2, \ldots, |G_K|^2 = \text{sort}(|h_{n_1}|^2, \ldots, |h_{n_K}|^2)$, where the sort function sorts the channel SNR gains in ascending order. Therefore, $|G_1|^2 < \ldots < |G_K|^2$. This is conceptually the same model used in [14, 18, 22], where the ordering of i.i.d. Rayleigh fading channel gains are used to represent the position of a user within a cell, and thus outage probabilities and diversity orders are derived from the distribution of this ordering. For the simulations using this channel model, the ordering of the channel gains and that of the SIC decoding order follow the same trend, so the user with weakest channel has its signal decoded first by all users, then the second weakest user, and so on. This channel model from here on is referred to as channel model 1.

The second channel model used is the MIMO Rayleigh fading channel model with i.i.d. fading channel gains between the different transmit-receive antenna pairs. A common precoding vector $\boldsymbol{p}$, $\|\boldsymbol{p}\| = 1$, is used to transmit to K users using M antennas, where $\boldsymbol{p}$ is not a function of the channel gains[1]. The signal passes through user n's $N \times M$ channel matrix $\boldsymbol{H}_n$ where the channel from transmit antenna i to receive antenna j is $h_{j,i} \sim \mathcal{CN}(0, \beta_n)$, and each user n with N receive

[1]In cellular deployments, the precoder is typically selected from a set of predetermined vectors, based on CSI feedback

antennas uses the optimum detection vector $\boldsymbol{v}_n = \boldsymbol{p}^{\mathsf{H}}\boldsymbol{H}_n^{\mathsf{H}}/\|\boldsymbol{H}_n\boldsymbol{p}\|$. This gives a channel SNR gain of $|G_n|^2 = \|\boldsymbol{H}_n\boldsymbol{p}\|^2$. The channel SNR gain has distribution Erlang(β_n, N), with expected value $\mathbb{E}\left[|G_n|^2\right] = N\beta_n$. For the simulations that use this channel model, $M = 2$, $N = 3$, $\boldsymbol{p}$ is selected randomly isotropically, and the β_n are not selected with any relationship to the target rates in order to demonstrate that the channel gain ordering has no bearing on the validity of the results. Therefore, $(\beta_1 = 0.5, \beta_2 = 1.4, \beta_3 = 0.8, \beta_4 = 1.7, \beta_5 = 1.1)$. This channel model from here on is referred to as channel model 2.

For all simulation plots, there are $K = 5$ users, the target rates are $(R_1 = 0.5, R_2 = 1.2, R_3 = 0.9, R_4 = 1.3, R_5 = 1.1)$ bps/Hz, and the OMA time durations are $(\tau_1 = 0.15, \tau_2 = 0.30, \tau_3 = 0.20, \tau_4 = 0.20, \tau_5 = 0.15)$. As mentioned previously, the decoding order must be such that $r_n = \frac{R_n}{\tau_n}$ is increasing, so since $(r_1 = \frac{10}{3}, r_2 = 4, r_3 = 4.5, r_4 = 6.5, r_5 = \frac{22}{3})$, the indices for the rates and time durations above are as such.

Figure 4.6 and figure 4.7 demonstrate the phenomenon described in theorem 4.4.1. For a power allocation strategy such that the interference coefficient A_n received when attempting to decode signal x_n exceeds the value given in theorem 4.4.1, then the outage probability is equal to 1, regardless of the channel strength and SNR. As can be seen in figure 4.6, for each signal to be decoded, the outage probabilities are lesser for users with stronger channels. In figure 4.7, the same phenomenon is observed even though the users have more receive antennas to increase their received SNR. In this case user 4 has the strongest channel statistically, so user 4 always has the least outage probabilities when the interference is below the certain outage threshold.

Figures 4.8 and 4.9 demonstrate the outage probability performance for NOMA compared to OMA, when NOMA uses both the power allocation strategy

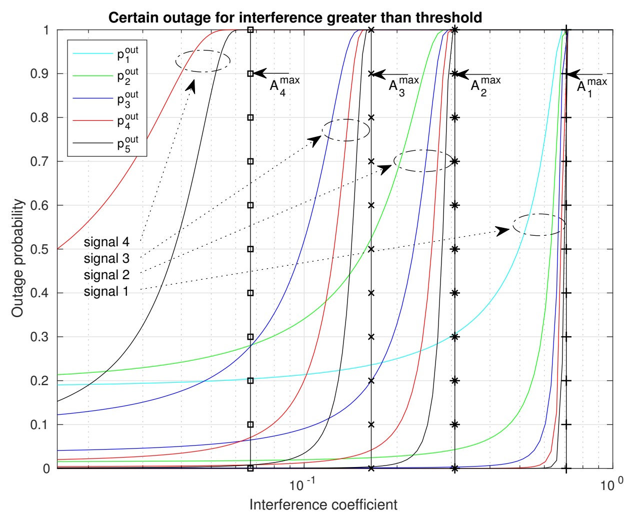

Figure 4.6: Channel model 1: Certain outage when interference exceeds limits in theorem 4.4.1; $\xi = 10\text{dB}$

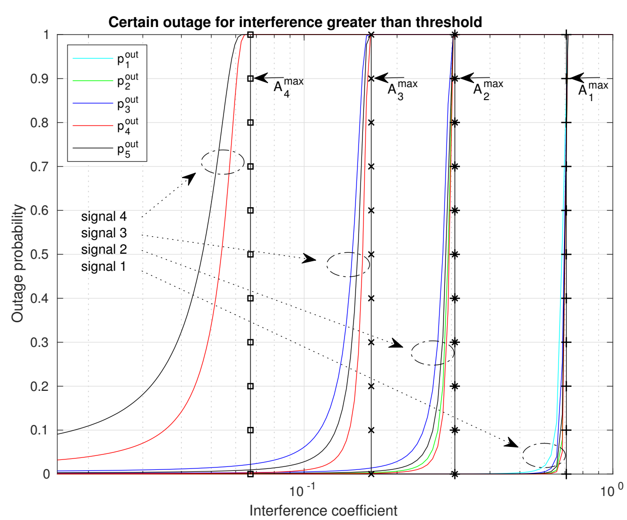

Figure 4.7: Channel model 2: Certain outage when interference exceeds limits in theorem 4.4.1; $\xi = 10\text{dB}$

$(\tilde{a}_1^{\text{oma}}, \ldots, \tilde{a}_5^{\text{oma}})$ in order to demonstrate the validity of the power allocation result of theorem 4.4.2 (blue curves), and $(a_1^{\text{oma}} + \epsilon_1, \ldots, a_5^{\text{oma}} + \epsilon_5)$ with ϵ_n selected according to the power allocation strategy described in equation (4.27) to ensure the power allocation coefficients are well-behaved (red curves). Clearly when the power allocation strategy for NOMA is $(\tilde{a}_1^{\text{oma}}, \ldots, \tilde{a}_5^{\text{oma}})$, the outage probability is exactly equal to that of OMA. However, as proven in theorem 4.4.2 the sum of the power allocation is less than 1. In fact, for this particular case it is ≈ 0.5036, which means that roughly only half of the maximum transmit SNR is needed to have the outage performance of NOMA equal that of OMA. In figure 4.8, there is a large difference in performance between NOMA (red curves) and OMA outage probabilities for users $K = 1, 2, 3$, while for users $K = 4, 5$ the gap is not so big. The same phenomenon is observed in figure 4.9, even though the ordering of the users' channel gains is not considered in the SIC decoding order. It makes sense that the gap in outage probability performance decreases for users whose signals are decoded towards the end of the SIC procedure. For example, if the BS tries to improve user 5's outage probability performance using NOMA over OMA by allocating ϵ_5 additional power allocation coefficient to its signal, while keeping the outage performance of the other users the same as OMA, the BS also has to increase the power allocation coefficient of user 4 by $c_4 = \epsilon_5(2^{R_4} - 1)$, and for user n by $c_n = \epsilon_5(2^{R_n} - 1)\prod_{k=n+1}^{4} 2^{R_k}, n = 1, 2, 3$, just so that they can have the same performance as OMA. So the amount of additional power allocation that the BS has available for a signal that is decoded later in the SIC procedure becomes less.

In figures 4.10 and 4.11, the well-behaved property of the strategy derived is demonstrated by plotting the outage probabilities for each signal to be decoded by each user in the SIC procedure. For example, user 5 must decode signals 1, 2, 3, and 4 before it can decode its own signal, and the outage probability

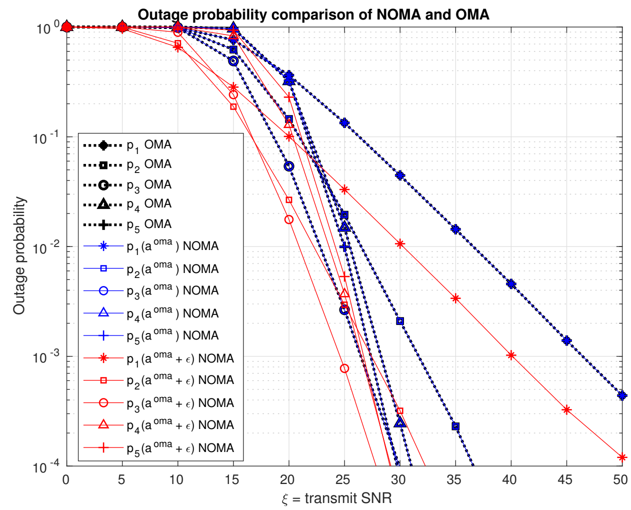

Figure 4.8: Channel model 1: Comparing NOMA performance according to power allocation strategies that perform equal to OMA (theorem 4.4.2) and better than OMA

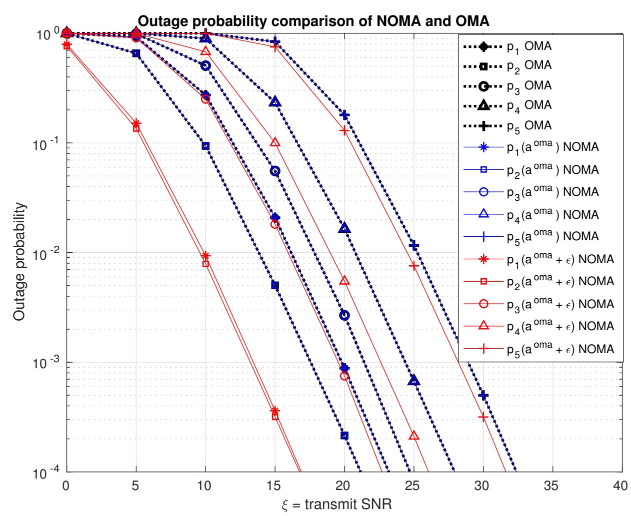

Figure 4.9: Channel model 2: Comparing NOMA performance according to power allocation strategies that perform equal to OMA (theorem 4.4.2) and better than OMA

performances are better for the signals earlier in the SIC procedure. A similar phenomenon is observed for user 4 with signals 1, 2, and 3, and so for the other users. This is consistent with what was stated regarding the overall outage event for a user, as it should not be bounded by the outage event of an earlier signal in the SIC procedure. In other words, the probability of outage should always be better for the decoding of the signals that are earlier in the SIC procedure. For figure 4.10, the outage probability for decoding a specific signal, say signal 1, is better for the users with stronger channels, as can be seen by the blue diamond curve belonging to user 5 being the best for decoding signal 1, and the black diamond curve belonging to user 1 being the worst, which is still better than user 1's outage probability curve for OMA as shown in 4.8. The same phenomenon is observed in figure 4.11, except that here the user with statistically the strongest channel gain is user 4, and accordingly the red diamond curve belonging to user 4 outperforms all of the other diamond curves. In this plot, even though user 5 is only the third strongest channel out of all, it has the signal that is decoded last among all other signals, and thus decodes all four other signals first, yet its outage probabilities for the first four decodings still demonstrate a well-behaved power allocation, while the outage curve with the blue star is still better than its OMA outage curve given (both seen in figure 4.9).

4.7 Conclusion and Future Work

In this chapter, it was demonstrated that for downlink NOMA systems with a BS which does not have knowledge of the exact channel gains, the power allocation strategy must be carefully designed in order to avoid certain outages for multiple users. Furthermore, it was demonstrated that a well-behaved power allocation strategy which has the same exact outage performance as OMA always exists, such

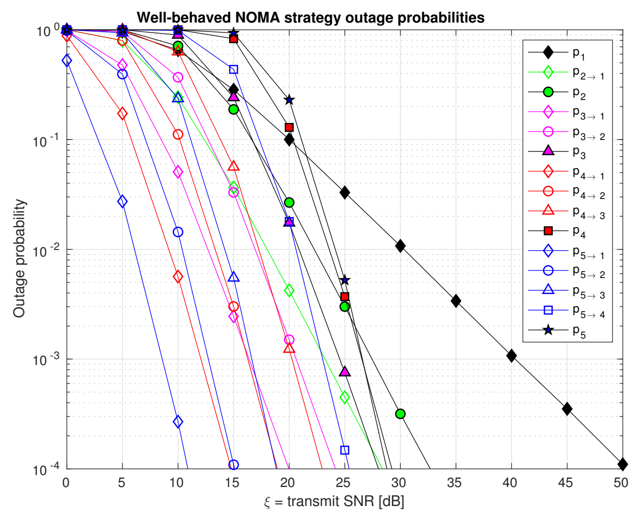

Figure 4.10: Channel model 1: Demonstration of well-behaved strategy behavior for NOMA

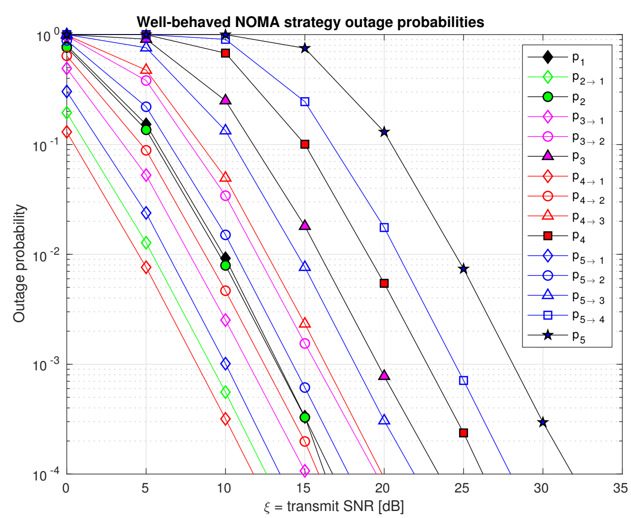

Figure 4.11: Channel model 2: Demonstration of well-behaved strategy behavior for NOMA

that it depends only on the target rates and their relative OMA time durations, and it is derived in closed form. The proposed SIC decoding order based on increasing values of $\frac{R_n}{\tau_n}$ was shown to be the most energy efficient. Lastly, the approach for designing a power allocation strategy which can always outperform OMA in terms of outage probability was outlined, and the associated properties of such a strategy were derived. The validity of these theoretical results are then substantiated with the simulation results, which show the outage performances for various power allocation strategies to exhibit these fundamental characteristics outlined in the chapter.

One thing that is not addressed in this chapter is the fact that the channel SNR gains can be used in the design of the power allocation coefficients. Comparing these results to the multi-user approaches similar to that in [22] can provide a very quick and simple assessment as to whether the channel SNR gains are strong enough to support the target rates. Further studies about how this type of phenomenon is exhibited and described theoretically in more complex cellular deployments is also critical, such as in multi-point and heterogeneous cellular networks. Lastly, a full treatment of the uplink scenario with regards to the power allocation strategy design is needed, as uplink NOMA is sought to be a vital deployment scenario for 5G cellular systems.

4.8 Proofs

The proofs for this chapter are as follows.

4.8.1 Proof of Theorem 4.4.1

Proof. For any specific user n, suppose that $A_{n-1} < 2^{-\sum_{l=1}^{n-1} R_l}$ and $A_n > 2^{-\sum_{l=1}^{n} R_l}$. Since $A_{n-1} = a_n + A_n$, it follows that

$$\begin{aligned} & a_n + A_n < 2^{-\sum_{l=1}^{n-1} R_l} \\ & \Longrightarrow a_n < 2^{-\sum_{l=1}^{n} R_l} - A_n < 2^{-\sum_{l=1}^{n-1} R_l} - 2^{-\sum_{l=1}^{n} R_l}. \end{aligned} \tag{4.28}$$

The events $\mathcal{B}_n$ and $\mathcal{B}_{k\to n}$ can be written in the form

$$\begin{aligned} & \log_2\left(1 + \frac{a_n \xi |G_k|^2}{A_n \xi |G_k|^2 + 1}\right) < R_n, k = n, \ldots, K \\ & \Longrightarrow \xi |G_k|^2 (a_n - (2^{R_n} - 1)A_n) < 2^{R_n} - 1. \end{aligned} \tag{4.29}$$

Since $a_n < 2^{-\sum_{l=1}^{n-1} R_l} - 2^{-\sum_{l=1}^{n} R_l}$ and $A_n > 2^{-\sum_{l=1}^{n} R_l}$,

$$\begin{aligned} a_n - (2^{R_n} - 1)A_n & < 2^{-\sum_{l=1}^{n;1} R_l} - 2^{-\sum_{l=1}^{n} R_l} - (2^{R_n} - 1)A_n \\ & < 2^{-\sum_{l=1}^{n-1} R_l} - 2^{-\sum_{l=1}^{n} R_l} - (2^{R_n} - 1)2^{-\sum_{l=1}^{n} R_l} \\ & = 0. \end{aligned} \tag{4.30}$$

Therefore solving equation (4.29) for $|G_k|^2$ leads to

$$\begin{aligned} & \xi |G_k|^2 (a_n - (2^{R_n} - 1)A_n) < 2^{R_n} - 1 \\ & \Longrightarrow |G_k|^2 > \frac{2^{R_n} - 1}{\xi (a_n - (2^{R_n} - 1)A_n)}. \end{aligned} \tag{4.31}$$

Therefore, since $\frac{2^{R_n}-1}{\xi(a_n-(2^{R_n}-1)A_n)} < 0 < |G_k|^2$, this condition makes $\Pr\{\mathcal{B}_n\} = \Pr\{\mathcal{B}_{k\to n}\} = 1$.

Now suppose that $A_n > 2^{-\sum_{l=n+1}^{K} R_l}, \forall n = 1, \ldots, K-1$, then it must be true that $A_1 > 2^{-R_1}$. This will avoid the previous impossible event. However, if this is true, then events $\mathcal{B}_1$ and $\mathcal{B}_{k\to 1}, k = 2, \ldots, K$, gives rise to the inequality

$$\xi |G_k|^2 (a_1 - (2^{R_1} - 1)A_1) < 2^{R_1} - 1, k = 1, \ldots, K, \tag{4.32}$$

where the value inside the parentheses must be greater than zero in order to avoid the certain outage situation from equation (4.32). Therefore,

$$\begin{aligned} &0 < a_1 - (2^{R_1} - 1)A_1 < a_1 - (2^{R_1} - 1)2^{-R_1} \\ \Rightarrow &a_1 > 1 - 2^{-R_1}. \end{aligned} \tag{4.33}$$

It must be true that $a_1 + A_1 \leq 1$ by definition of power allocation coefficients, however

$$a_1 + A_1 > (1 - 2^{-R_1}) + 2^{-R_1} = 1. \tag{4.34}$$

Therefore if $A_n > 2^{-\sum_{l=1}^{n} R_l}, \forall n = 1, \ldots, K-1$, then having $\Pr\{\mathcal{B}_n\} < 1$ and $\Pr\{\mathcal{B}_{k\to n}\} < 1$ requires $\sum_{n=1}^{K} a_n = a_1 + A_1 > 1$, which is not possible.

Hence, for any user n with $A_n > 2^{-\sum_{l=1}^{n} R_l}$, $\Pr\{\mathcal{B}_n\} = \Pr\{\mathcal{B}_{k\to n}\} = 1, k = n+1, \ldots, K$. □

4.8.2 Proof of Proposition 2

Proof. If $\mathcal{B}_n^{\text{out}} = \mathcal{B}_n, \forall n = 2, \ldots, K$, then it is true that

$$\frac{2^{R_m} - 1}{a_m - (2^{R_m} - 1)\sum_{l=m+1}^{K} a_l} \leq \frac{2^{R_n} - 1}{a_n - (2^{R_n} - 1)\sum_{l=n+1}^{K} a_l}, \tag{4.35}$$

$\forall m = 1, \ldots, n-1$. If this is true, then it is true that

$$\frac{2^{R_1}-1}{a_1 - (2^{R_1}-1)\sum_{l=2}^{K} a_l} \leq \cdots \leq \frac{2^{R_{K-1}}-1}{a_{K-1} - (2^{R_{K-1}}-1)\sum_{l=K}^{K} a_l} \leq \frac{2^{R_K}-1}{a_K}.$$

From the above, it is easy to show that

$$a_{n-1} \geq a_n \frac{2^{R_n}(2^{R_{n-1}}-1)}{2^{R_n}-1}, n = 2, \ldots, K. \tag{4.36}$$

To show that the condition above implies that $a_1, \ldots, a_K$ satisfy theorem 4.4.1, it is sufficient to show that any power allocation coefficients satisfying equation (4.36) satisfy the inequality

$$a_n - (2^{R_n}-1)A_n > 0, \tag{4.37}$$

based on equation (4.30), according to the theorem. So if equation (4.36) holds $\forall n = 2, \ldots, K$, then for any $n < K$ and $l > n$, it is easily shown that

$$\begin{aligned} a_l &< a_n \frac{2^{R_l}-1}{(2^{R_n}-1)\prod_{m=n+1}^{l} 2^{R_m}} \\ \Longrightarrow \sum_{l=n+1}^{K} a_l &< \frac{a_n}{2^{R_n}-1} \sum_{l=n+1}^{K} \frac{2^{R_l}-1}{\prod_{m=n+1}^{l} 2^{R_m}} \\ &= \frac{a_n}{2^{R_n}-1}\left(1 + \sum_{l=n+2}^{K} \prod_{m=n+1}^{l-1} \frac{1}{2^{R_m}} - \sum_{l=n+1}^{K} \prod_{m=n+1}^{l} \frac{1}{2^{R_m}}\right) \\ &= \frac{a_n}{2^{R_n}-1}\left(1 - \prod_{m=n+1}^{K} 2^{-R_m}\right). \end{aligned} \tag{4.38}$$

So

$$a_n - (2^{R_n}-1)\sum_{l=n+1}^{K} a_l > a_n - (2^{R_n}-1)\frac{a_n}{2^{R_n}-1}\left(1 - \prod_{m=n+1}^{K} 2^{-R_m}\right)$$

$$= a_n \prod_{m=n+1}^{K} 2^{-R_m} > 0, \forall n = 1, \ldots, K-1. \tag{4.39}$$

Hence, these power allocation coefficients satisfy the requirement in theorem 4.4.1.

□

4.8.3 Proof of Theorem 4.4.2

Proof. If $\mathcal{B}_n \subseteq \mathcal{B}_n^{\text{oma}}, \forall n = 1, \ldots, K$, it must at least be true that $\exists(a_1, \ldots, a_K)$ s.t. $\mathcal{B}_n = \mathcal{B}_n^{\text{oma}}, \forall n = 1, \ldots, K$, and then demonstrate that $\sum_{n=1}^{K} a_n < 1$. To show that $\exists(a_1, \ldots, a_K)$ s.t. $\mathcal{B}_n = \mathcal{B}_n^{\text{oma}}, \forall n = 1, \ldots, K$, begin with $n = K$ and equate

$$\frac{2^{R_K} - 1}{a_K \xi} = \frac{2^{R_K/\tau_K} - 1}{\xi} \quad \Longrightarrow \quad a_K = \frac{2^{R_K} - 1}{2^{R_K/\tau_K} - 1}. \tag{4.40}$$

Then for $n = 1, \ldots, K-1$, equate

$$\begin{aligned} \frac{2^{R_n} - 1}{\xi(a_n - (2^{R_n} - 1)\sum_{l=n+1}^{K} a_l)} &= \frac{2^{R_n/\tau_n} - 1}{\xi} \\ \Longrightarrow a_n - (2^{R_n} - 1) \sum_{l=n+1}^{K} a_l &= \frac{2^{R_n} - 1}{2^{R_n/\tau_n} - 1} \end{aligned} \tag{4.41}$$

This creates a recursive relationship which can be solved to find

$$a_n = \frac{2^{R_n} - 1}{2^{R_n/\tau_n} - 1} + \frac{2^{R_n} - 1}{2^{R_n}} \sum_{l=n+1}^{K} \frac{2^{R_l} - 1}{2^{R_l/\tau_l} - 1} \prod_{k=n}^{l-1} 2^{R_k}, \tag{4.42}$$

$n = 1, \ldots, K-1$. From here on, the power allocation strategy that satisfies equations (4.40, 4.42) will be called $(\tilde{a}_1^{\text{oma}}, \ldots, \tilde{a}_K^{\text{oma}})$. In order for this to be a valid power allocation strategy, the sum of the coefficients must be proven to always be less than or equal to 1. Let the interference coefficient for user n using this power allocation strategy be called A_n^{oma}, which can be found easily by noting from equation (4.41) that

$$\tilde{a}_n^{\text{oma}} - (2^{R_n} - 1)A_n^{\text{oma}} = \frac{2^{R_n} - 1}{2^{R_n/\tau_n} - 1} \tag{4.43}$$

$$\implies A_n^{\text{oma}} = \frac{1}{2^{R_n} - 1}\left(\tilde{a}_n^{\text{oma}} - \frac{2^{R_n} - 1}{2^{R_n/\tau_n} - 1}\right)$$

$$= \frac{2^{R_{n+1}} - 1}{2^{R_{n+1}/\tau_{n+1}} - 1} + \sum_{l=n+2}^{K} \frac{2^{R_l} - 1}{2^{R_l/\tau_l} - 1} \prod_{k=n+1}^{l-1} 2^{R_k}. \tag{4.44}$$

Define $\tau_n = \frac{b_n}{K}$ and $r_n = \frac{R_n}{b_n}$, so that $r_1 < \cdots < r_K$. Since the function $h(t) = (2^{bt} - 1)/(2^{Kt} - 1)$ is a monotonically decreasing function in t so long as $b < K$, then

$$\begin{aligned}
A_n^{\text{oma}} =& \frac{2^{R_{n+1}} - 1}{2^{R_{n+1}/\tau_{n+1}} - 1} + \sum_{l=n+2}^{K} \frac{2^{R_l} - 1}{2^{R_l/\tau_l} - 1} \prod_{k=n+1}^{l-1} 2^{R_k} \\
=& \frac{2^{b_{n+1}r_{n+1}} - 1}{2^{Kr_{n+1}} - 1} + \sum_{l=n+2}^{K} \tfrac{2^{b_l r_l} - 1}{2^{K r_l} - 1} \prod_{k=n+1}^{l-1} 2^{b_k r_k} \\
<& \frac{2^{b_{n+1}r_{n+1}} - 1}{2^{Kr_{n+1}} - 1} + \sum_{l=n+2}^{K-2} \frac{2^{b_l r_l} - 1}{2^{K r_l} - 1} \prod_{k=n+1}^{l-1} 2^{b_k r_k} + \frac{2^{b_{K-1}r_{K-1}} - 1}{2^{Kr_{K-1}} - 1} \prod_{k=n+1}^{K-2} 2^{b_k r_k} \\
&+ \frac{(2^{b_K r_{K-1}} - 1)2^{b_{K-1}r_{K-1}}}{2^{Kr_{K-1}} - 1} \prod_{k=n+1}^{K-2} 2^{b_k r_k} \\
=& \frac{2^{b_{n+1}r_{n+1}} - 1}{2^{Kr_{n+1}} - 1} + \sum_{l=n+2}^{K-2} \frac{2^{b_l r_l} - 1}{2^{K r_l} - 1} \prod_{k=n+1}^{l-1} 2^{b_k r_k} + \frac{2^{(b_{K-1}+b_K)r_{K-1}} - 1}{2^{Kr_{K-1}} - 1} \prod_{k=n+1}^{K-2} 2^{b_k r_k} \\
<& \\
&\vdots \\
<& \frac{2^{(b_{n+1}+\cdots+b_K)r_{n+1}} - 1}{2^{Kr_{n+1}} - 1}.
\end{aligned}$$

So given that $\tilde{a}_n^{\text{oma}} = \frac{2^{b_n r_n}-1}{2^{K r_n}-1} + (2^{b_n r_n}-1)A_n^{\text{oma}}$, then

$$\begin{aligned}
A_{\text{tot}}^{\text{oma}} &= \sum_{n=1}^{K} \tilde{a}_l^{\text{oma}} = \tilde{a}_1^{\text{oma}} + A_1^{\text{oma}} \\
&= \frac{2^{b_1 r_1}-1}{2^{K r_1}-1} + (2^{b_1 r_1}-1)A_1^{\text{oma}} + A_1^{\text{oma}} \\
&= \frac{2^{b_1 r_1}-1}{2^{K r_1}-1} + 2^{b_1 r_1} A_1^{\text{oma}} \\
&< \frac{2^{b_1 r_1}-1}{2^{K r_1}-1} + 2^{b_1 r_1} \frac{2^{(b_2+\cdots+b_K) r_2}-1}{2^{K r_2}-1} \\
&< \frac{2^{b_1 r_1}-1}{2^{K r_1}-1} + 2^{b_1 r_1} \frac{2^{(b_2+\cdots+b_K) r_1}-1}{2^{K r_1}-1} \\
&= \frac{2^{(b_1+\cdots+b_K) r_1}-1}{2^{K r_1}-1} \\
&= 1.
\end{aligned}$$

So clearly the sum is less than 1. To complete the proof, only one strategy that satisfies the conditions stated in the theorem is needed, so let user 1 have the power allocation coefficient $a_1 = a_1^{\text{oma}} + \epsilon_1$, such that

$$\epsilon_1 = 1 - A_{\text{tot}}^{\text{oma}} > 0, \tag{4.45}$$

which leads to $\mathcal{B}_1 \subset \mathcal{B}_1^{\text{oma}}$. □

4.8.4 Proof of Corollary 1

Proof. For any SIC decoding order $(\sigma(1), \ldots, \sigma(K))$, which is a permutation of $(1, \ldots, K)$, the power allocation strategy $(\tilde{a}_{\sigma(1)}^{\text{oma}}, \ldots, \tilde{a}_{\sigma(K)}^{\text{oma}})$ such that $\mathcal{B}_{\sigma(n)}^{\text{out}} = \mathcal{B}_{\sigma(n)}^{\text{oma}}, \forall n = 1, \ldots, K$, is given by $\tilde{a}_{\sigma(K)}^{\text{oma}} = \frac{2^{R_{\sigma(K)}}-1}{2^{R_{\sigma(K)}/\tau_{\sigma(K)}}-1}$ and

$$\tilde{a}_{\sigma(n)}^{\text{oma}} = \frac{2^{R_{\sigma(n)}}-1}{2^{R_{\sigma(n)}/\tau_{\sigma(n)}}-1} + \frac{2^{R_{\sigma(n)}}-1}{2^{R_{\sigma(n)}}} \sum_{l=n+1}^{K} \frac{2^{R_{\sigma(l)}}-1}{2^{R_{\sigma(l)}/\tau_{\sigma(l)}}-1} \prod_{k=n}^{l-1} 2^{R_{\sigma(k)}}, \tag{4.46}$$

$n = 1, \ldots, K-1$, and the sum[2] of this power allocation strategy is

$$\sum_{n=1}^{K} \tilde{a}_{\sigma(n)}^{\text{oma}} = \frac{2^{R_{\sigma(1)}} - 1}{2^{R_{\sigma(1)}/\tau_{\sigma(1)}} - 1} + \sum_{l=2}^{K} \frac{2^{R_{\sigma(l)}} - 1}{2^{R_{\sigma(l)}/\tau_{\sigma(l)}} - 1} \prod_{k=1}^{l-1} 2^{R_{\sigma(k)}}. \tag{4.47}$$

Suppose a SIC decoding order has consecutive SIC decoding stages m and $m+1$ such that $\sigma(m) > \sigma(m+1)$, i.e. the index of the user whose signal is decoded at SIC stage m is greater than the index of the user whose signal is decoded at SIC stage $m+1$, which means that $\frac{R_{\sigma(m)}}{\tau_{\sigma(m)}} > \frac{R_{\sigma(m+1)}}{\tau_{\sigma(m+1)}}$. If the SIC stages m and $m+1$ are reversed, such that the signal of user $m+1$ is now decoded before the signal of user m, while all other stages remain in the same order, then call $(\tilde{a}_{\sigma(1)}^{\prime\text{oma}}, \ldots, \tilde{a}_{\sigma(K)}^{\prime\text{oma}})$ the new power allocation strategy which has NOMA outage performance equal to OMA. It can easily be shown that $\tilde{a}_{\sigma(n)}^{\prime\text{oma}} = \tilde{a}_{\sigma(n)}^{\text{oma}}, \forall n \neq m, m+1$, and that $\tilde{a}_{\sigma(m)}^{\prime\text{oma}}$ and $\tilde{a}_{\sigma(m+1)}^{\prime\text{oma}}$ are given by

$$\begin{aligned}
\tilde{a}_{\sigma(m+1)}^{\prime\text{oma}} =& \frac{2^{R_{\sigma(m+1)}} - 1}{2^{R_{\sigma(m+1)}/\tau_{\sigma(m+1)}} - 1} \\
&+ (2^{R_{\sigma(m+1)}} - 1)\left[\frac{2^{R_{\sigma(m)}} - 1}{2^{R_{\sigma(m)}/\tau_{\sigma(m)}} - 1} + \sum_{l=m+2}^{K} \frac{2^{R_{\sigma(l)}} - 1}{2^{R_{\sigma(l)}/\tau_{\sigma(l)}} - 1} \prod_{\substack{k=m \\ k \neq m+1}}^{l-1} 2^{R_{\sigma(k)}}\right] \\
\tilde{a}_{\sigma(m)}^{\prime\text{oma}} =& \frac{2^{R_{\sigma(m)}} - 1}{2^{R_{\sigma(m)}/\tau_{\sigma(m)}} - 1} \\
&+ (2^{R_{\sigma(m)}} - 1)\left[\frac{2^{R_{\sigma(m+2)}} - 1}{2^{R_{\sigma(m+2)}/\tau_{\sigma(m+2)}} - 1} + \sum_{l=m+3}^{K} \frac{2^{R_{\sigma(l)}} - 1}{2^{R_{\sigma(l)}/\tau_{\sigma(l)}} - 1} \prod_{k=m+2}^{l-1} 2^{R_{\sigma(k)}}\right].
\end{aligned}$$

Taking the difference of the sums of the two power allocation strategies gives

$$\begin{aligned}
&\sum_{n=1}^{K} (\tilde{a}_{\sigma(n)}^{\text{oma}} - \tilde{a}_{\sigma(n)}^{\prime\text{oma}}) \qquad (4.48) \\
&= \left[\frac{(2^{R_{\sigma(m)}} - 1)(2^{R_{\sigma(m+1)}} - 1)}{2^{R_{\sigma(m+1)}/\tau_{\sigma(m+1)}} - 1} - \frac{(2^{R_{\sigma(m)}} - 1)(2^{R_{\sigma(m+1)}} - 1)}{2^{R_{\sigma(m)}/\tau_{\sigma(m)}} - 1}\right] \prod_{k=1}^{m-1} 2^{R_{\sigma(k)}} > 0,
\end{aligned}$$

[2] Note that this sum is only guaranteed to be less than 1 for the decoding order $(1, \ldots, K)$.

which is true because $\frac{R_{\sigma(m)}}{\tau_{\sigma(m)}} > \frac{R_{\sigma(m+1)}}{\tau_{\sigma(m+1)}}$, and thus power allocation strategy $\{\tilde{a}'^{\text{oma}}_{\sigma(n)}\}_{n=1}^{K}$ is more energy efficient than $\{\tilde{a}^{\text{oma}}_{\sigma(n)}\}_{n=1}^{K}$. Successively repeat this process of reversing the positions of all consecutive SIC decoding stages m and $m+1$ which have $\sigma(m) > \sigma(m+1)$, while keeping the SIC decoding order of every other stage constant, in order to successively obtain a more energy efficient power allocation strategy. This process is repeated until the SIC decoding order obtained is given by $(1, \ldots, K)$. □

4.8.5 Proof of Theorem 4.5.1

Proof. (a) Since $a_n^{\text{oma}} = \frac{2^{R_n}-1}{2^{R_n/\tau_n}-1} + (2^{R_n}-1)(A_n^{\text{oma}} + c_n)$, and ϵ_{n-1} and $\epsilon_n = 0$, then proposition 2 is used to show that the following is always true for $n = 2, \ldots, K$,

$$a_{n-1}^{\text{oma}} > a_n^{\text{oma}} 2^{R_n} \frac{2^{R_{n-1}}-1}{2^{R_n}-1} \tag{4.49}$$

$$\Longrightarrow \frac{2^{R_{n-1}}-1}{2^{R_{n-1}/\tau_{n-1}}-1} + (2^{R_{n-1}}-1)(A_{n-1}^{\text{oma}} + 2^{R_n} c_n) \tag{4.50}$$

$$> \left(\frac{2^{R_n}-1}{2^{R_n/\tau_n}-1} + (2^{R_n}-1)(A_n^{\text{oma}} + c_n) \right) 2^{R_n} \frac{2^{R_{n-1}}-1}{2^{R_n}-1}$$

$$\Longrightarrow \frac{1}{2^{R_{n-1}/\tau_{n-1}}-1} + (\tilde{a}_n^{\text{oma}} + A_n^{\text{oma}}) > \frac{2^{R_n}}{2^{R_n/\tau_n}-1} + 2^{R_n} A_n^{\text{oma}} \tag{4.51}$$

$$\Longrightarrow \frac{1}{2^{R_{n-1}/\tau_{n-1}}-1} > \frac{1}{2^{R_n/\tau_n}-1}, \tag{4.52}$$

which is true because $\frac{R_{n-1}}{\tau_{n-1}} < \frac{R_n}{\tau_n}$.

(b) For this case, the minimum allowable power allocation coefficient for users $m = 1, \ldots, n-1$ is a_m^{oma}. If $\epsilon_n > 0$, user n power allocation coefficient $a_n = a_n^{\text{oma}} + \epsilon_n$ leads to $\mathcal{B}_n \subset \mathcal{B}_n^{\text{oma}}$. The power allocation coefficient for user $n-1$ is then given by

$$a_{n-1} = \frac{2^{R_{n-1}}-1}{2^{R_{n-1}/\tau_{n-1}}-1} + (2^{R_{n-1}}-1)(A_{n-1}^{\text{oma}} + \epsilon_n) + \epsilon_{n-1}$$

$$= \tilde{a}_{n-1}^{\text{oma}} + (2^{R_{n-1}} - 1)\epsilon_n + \epsilon_{n-1}. \tag{4.53}$$

Then, the power allocation coefficient for users $m = 1, \ldots, n-2$ is found recursively starting from $n-2$ to be

$$a_m = \tilde{a}_m^{\text{oma}} + \epsilon_m + (2^{R_m} - 1)\left(\epsilon_{m+1} + \sum_{l=m+2}^{n} \epsilon_l \prod_{k=m+1}^{l-1} 2^{R_k}\right). \tag{4.54}$$

By noting that the sum of the power allocation coefficients[3] is less than or equal to 1,

$$\begin{aligned} 1 &\geq \sum_{m=1}^{K} a_m \\ &= \sum_{m=1}^{K} \tilde{a}_m^{\text{oma}} + d_n + \epsilon_n \prod_{l=1}^{n-1} 2^{R_l} \qquad (4.55) \\ &= A_{\text{tot}}^{\text{oma}} + d_n + \epsilon_n \prod_{l=1}^{n-1} 2^{R_l} \\ \Longrightarrow \epsilon_n &\leq (1 - A_{\text{tot}}^{\text{oma}}) \prod_{l=1}^{n-1} 2^{-R_l} - \sum_{m=1}^{n-1} \epsilon_m \prod_{l=m}^{n-1} 2^{-R_l} \qquad (4.56) \end{aligned}$$

It must also be true that

$$\begin{aligned} & a_{n-1} \geq a_n 2^{R_n} \frac{2^{R_{n-1}} - 1}{2^{R_n} - 1} \\ \Longrightarrow & \frac{2^{R_{n-1}} - 1}{2^{R_{n-1}/\tau_{n-1}} - 1} + \epsilon_{n-1} + (2^{R_{n-1}} - 1)(A_{n-1}^{\text{oma}} + c_n 2^{R_n} + \epsilon_n) \\ & \geq \left[\frac{2^{R_n} - 1}{2^{R_n/\tau_n} - 1} + \epsilon_n + (2^{R_n} - 1)(A_n^{\text{oma}} + c_n)\right] 2^{R_n} \frac{2^{R_{n-1}} - 1}{2^{R_n} - 1} \\ \Longrightarrow & \epsilon_n \leq \epsilon_{n-1} \frac{2^{R_n} - 1}{2^{R_{n-1}} - 1} + \frac{2^{R_n} - 1}{2^{R_{n-1}/\tau_{n-1}} - 1} - \frac{2^{R_n} - 1}{2^{R_n/\tau_n} - 1}. \qquad (4.57) \end{aligned}$$

The inequalities (4.56) and (4.57) hence yield the result in inequality (4.23). □

[3]Note that $c_n \geq 0$, with potential equality if no additional power allocation is available

4.9 Three-user Example of Theoretical Results

The following section gives a three-user example of the results provided in this chapter. This includes simple proofs for this special case that are easier to follow.

Suppose that three users have been determined to be suitable by the base-station to receive downlink transmissions of signals with specific target rates, which will occur within a transmission time period of T seconds. The suitability determined by the base-station is based on received CSI (e.g. RI, PMI, CQI, etc.) feedback from the users, as well as its own channel estimates based on prior uplink transmissions from these same users in the same bandwidth. Let the users be named user-1, user-2, and user-3. The target rate for user-1 is R_1, user-2 is R_2, and user-3 is R_3, all in terms of bits per second per Hz (bps/Hz).

Suppose that the base-station determines that the downlink transmissions to these users can be scheduled according to an orthogonal multiple access approach, TDMA, where the base-station allocates fractions of the transmission time τ_1 to user-1, τ_2 to user-2, and τ_3 to user-3 respectively, and the signals are transmitted with full transmit SNR ξ for their respective time duration. The selection of τ_1, τ_2, τ_3 can be any set positive values such that $\tau_1 + \tau_2 + \tau_3 = 1$, even the optimum values for this system. Assume the relationship $\frac{R_1}{\tau_1} < \frac{R_2}{\tau_2} < \frac{R_3}{\tau_3}$ determined the user indices.

The base-station can instead schedule a NOMA downlink transmission for the signals of these same users instead, where the power allocation strategy is (a_1, a_2, a_3). The base-station selects an appropriate power allocation strategy (a_1, a_2, a_3) for the SIC decoding order $(1, 2, 3)$ such that:

- the total inferference coefficients of when decoding signals 1 and 2 must not exceed the thresholds $A_1 = 2^{-R_1}$ and $A_2 = 2^{-(R_1+R_2)}$, respectively;

- the power allocation strategy is well-behaved, which means that $\mathcal{B}_n^{\text{out}} = \mathcal{B}_n, n = 1, 2, 3$, and a clear example is given that a power allocation strategy that is not well-behaved is wasting transmit power without the additional gain in outage probability performance;
- the outage probability of this NOMA system will outperform the OMA system given above, and do so in the most energy efficient manner.

The following will demonstrate the above points regarding the NOMA power allocation strategy by first deriving the power allocation coefficients.

4.9.1 Total Interference of signals

If the power allocation strategy is such that when decoding the first signal, the total interference coefficient received $A_1 = a_2 + a_3 > 2^{-R_1}$, then the outage event of decoding this signal at the receiver of user-$n = 1, 2, 3$, is such that $\log_2(1 + \frac{a_1\xi|G_n|^2}{1+A_1\xi|G_n|^2}) < R_1$. So solving for the channel SNR gain $|G_n|^2$ gives

$$\begin{aligned} a_1|G_n|^2 &< (2^{R_1} - 1)(1 + A_1\xi|G_n|^2) \\ \Longrightarrow &\xi|G_n|^2[a_1 - A_1(2^{R_1} - 1)] < 2^{R_1} - 1 \\ \Longrightarrow &\xi|G_n|^2[1 - A_1 2^{R_1}] < 2^{R_1} - 1 \end{aligned}$$

Since $A_1 > 2^{-R_1} \Rightarrow A_1 2^{R_1} > 1$, so $1 - A_1 2^{R_1} < 1$. Therefore, the following is true

$$\Longrightarrow \frac{2^{R_1} - 1}{\xi[1 - A_1 2^{R_1}]} < 0 < |G_n|^2.$$

which means that there is an outage event $\forall |G_n|^2 > 0$ (indeed, all channel SNR gains are positive real numbers). Therefore, the interference coefficient $A_1 = a_2 + a_3 < R_1$ must hold true in order to not create a *certain outage*. In the same manner, if interference coefficient $A_2 = a_3 > 2^{-(R_1+R_2)}$, since the outage event at the receiver of user-$n = 2, 3$ is given by $\log_2(1 + \frac{a_2\xi|G_n|^2}{1+A_2\xi|G_n|^2}) < R_2$, then

$$\Longrightarrow \xi|G_n|^2[a_2 - A_2(2^{R_2} - 1)] < 2^{R_2} - 1$$
$$\Longrightarrow \xi|G_n|^2[A_1 - A_2 2^{R_2}] < 2^{R_2} - 1.$$

If $A_1 = a_2 + a_3 < 2^{-R_1}$ then $A_1 - A_2 2^{R_2} < 0$

$$\Longrightarrow \frac{2^{R_2} - 1}{\xi[A_1 - A_2 2^{R_2}]} < 0 < |G_n|^2,$$

Otherwise if $A_1 > 2^{-R_1}$ then stage 1 of the SIC decoding procedure will fail at all users.

4.9.2 Well-behaved Property

The power allocation strategy should be such that the coefficients have the property $a_1 \geq a_2 \frac{2^{R_2}(2^{R_1}-1)}{2^{R_2}-1}$ and $a_2 \geq a_3 \frac{2^{R_3}(2^{R_2}-1)}{2^{R_3}-1}$ so that the $\mathcal{B}_2^{\text{out}} = \mathcal{B}_2 \cup \mathcal{B}_{2\to 1} = \mathcal{B}_2$ and $\mathcal{B}_3^{\text{out}} = \mathcal{B}_3 \cup \mathcal{B}_{3\to 1} \cup \mathcal{B}_{3\to 2} = \mathcal{B}_3$. This can be demonstrated in the proof of proposition 1 to ensure that the interference coefficients do not exceed the upper-bounds described in theorem 1.

Also, this is important because of the following example. Let $a_3 = \tilde{a}_3^{\text{oma}}$, $a_2 = \tilde{a}_2^{\text{oma}} + \epsilon_2$ and $a_1 = \tilde{a}_1^{\text{oma}} + (2^{R_1} - 1)\epsilon_2$, such that $a_1 + a_2 + a_3 = 1$ and $\epsilon_2 > \frac{2^{R_2}-1}{2^{R_1/\tau_1}-1} - \frac{2^{R_2}-1}{2^{R_2/\tau_2}-1}$. This leads to $a_1 < a_2 \frac{2^{R_2}(2^{R_1}-1)}{2^{R_2}-1}$ and $a_2 \geq a_3 \frac{2^{R_3}(2^{R_2}-1)}{2^{R_3}-1}$, so then $\mathcal{B}_1^{\text{out}} = \mathcal{B}_1^{\text{oma}}$ and $\mathcal{B}_3^{\text{out}} = \mathcal{B}_3^{\text{oma}}$. However, $\mathcal{B}_2^{\text{out}} = \mathcal{B}_{2\to 1}$, because as is shown in the proof of theorem 3, and since $\mathcal{B}_2 = \left\{|G_2|^2 < \frac{2^{R_2}-1}{\xi[a_2 - a_3(2^{R_2}-1)]}\right\}, \mathcal{B}_{2\to 1} =$

$\left\{|G_2|^2 < \frac{2^{R_1}-1}{\xi[a_1-(a_2+a_3)(2^{R_1}-1)]}\right\}$ and

$$\begin{aligned}
&\frac{2^{R_2}-1}{\xi[a_2-a_3(2^{R_2}-1)]} < \frac{2^{R_1}-1}{\xi[a_1-(a_2+a_3)(2^{R_1}-1)]} \\
&\implies \mathcal{B}_2 \subset \mathcal{B}_{2\to 1}
\end{aligned}$$

$\implies \mathcal{B}_2^{\text{out}} = \mathcal{B}_2 \cup \mathcal{B}_{2\to 1} = \mathcal{B}_{2\to 1}$.

In plain terms, this means that the power allocation for user-2 can be reduced to the level such that $\epsilon_2 = \frac{2^{R_2}-1}{2^{R_1/\tau_1}-1} - \frac{2^{R_2}-1}{2^{R_2/\tau_2}-1}$, which leads to $\mathcal{B}_2 = \mathcal{B}_{2\to 1}$, meaning that the power allocation for user-2 can be reduced, while maintaining the same power allocation for users-1 and user-3, and still achieve the exact outage probability performance for all 3 users. In other words, a power allocation strategy that is not well-behaved is wasting power on one of the signals when compared to another existing power allocation strategy that achieves the same outage probability performance for all users.

4.9.3 Power Allocation so that NOMA has Equal Outage Performance to OMA for all Users

For the three users, the power allocation strategy that will outage probability performance equal to OMA (and to have sum less than 1 in the proof of theorem 2, and to be well-behaved and not violate the interference property in theorem 3) is given by

$$\begin{aligned}
\tilde{a}_1 &= \frac{2^{R_1}-1}{2^{R_1/\tau_1}-1} + (2^{R_1}-1)\left(\frac{2^{R_2}-1}{2^{R_2/\tau_2}-1} + 2^{R_2}\frac{2^{R_3}-1}{2^{R_3/\tau_3}-1}\right) \\
\tilde{a}_2 &= \frac{2^{R_2}-1}{2^{R_2/\tau_2}-1} + (2^{R_2}-1)\frac{2^{R_3}-1}{2^{R_3/\tau_3}-1} \\
\tilde{a}_3 &= \frac{2^{R_3}-1}{2^{R_3/\tau_3}-1}.
\end{aligned}$$

The summation of this power allocation strategy can be shown to be less than 1 as follows. The indices are assigned based on the relationship of $\frac{R_1}{\tau_1} < \frac{R_2}{\tau_2} < \frac{R_3}{\tau_3}$. Define $\tau_1 = \frac{b_1}{K}, \tau_2 = \frac{b_2}{K}, \tau_3 = \frac{b_3}{K}$, then $\frac{R_1}{\tau_1} = \frac{KR_1}{b_1}, \frac{R_2}{\tau_2} = \frac{KR_2}{b_2}, \frac{R_3}{\tau_3} = \frac{KR_3}{b_3}$. Defining $r_1 = \frac{R_1}{b_1}, r_2 = \frac{R_2}{b_2}, r_3 = \frac{R_3}{b_3}$, clearly $r_1 < r_2 < r_3$. Therefore, since $f(t) = \frac{2^{bt}-1}{2^{Kt}-1}, b < K$, is monotonically decreasing in t, then

$$
\begin{aligned}
a_1 + a_2 + a_3 &= \frac{2^{R_1} - 1}{2^{R_1/\tau_1} - 1} + 2^{R_1} \frac{2^{R_2} - 1}{2^{R_2/\tau_2} - 1} + 2^{R_1+R_2} \frac{2^{R_3} - 1}{2^{R_3/\tau_3} - 1} \\
&= \frac{2^{b_1 r_1} - 1}{2^{K r_1} - 1} + 2^{b_1 r_1} \frac{2^{b_2 r_2} - 1}{2^{K r_2} - 1} + 2^{b_1 r_1 + b_2 r_2} \frac{2^{b_3 r_3} - 1}{2^{K r_3} - 1} \\
&< \frac{2^{b_1 r_1} - 1}{2^{K r_1} - 1} + 2^{b_1 r_1} \frac{2^{b_2 r_2} - 1}{2^{K r_2} - 1} + 2^{b_1 r_1 + b_2 r_2} \frac{2^{b_3 r_2} - 1}{2^{K r_2} - 1} \\
&= \frac{2^{b_1 r_1} - 1}{2^{K r_1} - 1} + 2^{b_1 r_1} \frac{2^{(b_2+b_3) r_2} - 1}{2^{K r_2} - 1} \\
&< \frac{2^{b_1 r_1} - 1}{2^{K r_1} - 1} + 2^{b_1 r_1} \frac{2^{(b_2+b_3) r_1} - 1}{2^{K r_1} - 1} \\
&= \frac{2^{(b_1+b_2+b_3) r_1} - 1}{2^{K r_1} - 1} \\
&= 1.
\end{aligned}
$$

4.9.4 Energy Efficiency of Proposed SIC Decoding Order

Corollary 1 states that if a different SIC decoding order $(\sigma(1), \sigma(2), \sigma(3))$ is used, where $(\sigma(1), \sigma(2), \sigma(3))$ is any permutation of the sequence $(1, 2, 3)$, then a more energy efficient SIC decoding order can be found if $\sigma(1) > \sigma(2)$ or $\sigma(2) < \sigma(3)$, by switching the SIC decoding order. Then, the power allocation strategy that enables NOMA to achieve equal outage performance to OMA for each user is given by

$$
\begin{aligned}
\tilde{a}_{\sigma(1)} &= \frac{2^{R_{\sigma(1)}} - 1}{2^{R_{\sigma(1)}/\tau_{\sigma(1)}} - 1} + (2^{R_{\sigma(1)}} - 1) \left(\frac{2^{R_{\sigma(2)}} - 1}{2^{R_{\sigma(2)}/\tau_{\sigma(2)}} - 1} + 2^{R_{\sigma(2)}} \frac{2^{R_{\sigma(3)}} - 1}{2^{R_{\sigma(3)}/\tau_{\sigma(3)}} - 1} \right) \\
\tilde{a}_{\sigma(2)} &= \frac{2^{R_{\sigma(2)}} - 1}{2^{R_{\sigma(2)}/\tau_{\sigma(2)}} - 1} + (2^{R_{\sigma(2)}} - 1) \frac{2^{R_{\sigma(3)}} - 1}{2^{R_{\sigma(3)}/\tau_{\sigma(3)}} - 1}
\end{aligned}
$$

$$\tilde{a}_{\sigma(3)} = \frac{2^{R_{\sigma(3)}} - 1}{2^{R_{\sigma(3)}/\tau_{\sigma(3)}} - 1}.$$

The following permutations $(\sigma(1), \sigma(2), \sigma(3))$ of $(1, 2, 3)$ exist and can be the SIC decoding order for the system: $(2, 1, 3), (2, 3, 1), (1, 3, 2), (3, 1, 2), (3, 2, 1)$, and the identity permutation $(1, 2, 3)$. By the proof of corollary 1, the SIC decoding orders can be listed in order of least energy efficient to most energy efficient: $\{(3, 2, 1), (3, 1, 2), (1, 3, 2), (1, 2, 3)\}$ or $\{(3, 2, 1), (2, 3, 1), (2, 1, 3), (1, 2, 3)\}$ For example, the proof of corollary 1 can be used to show that the SIC decoding order $(3, 1, 2)$ is more energy efficient (i.e. uses less transmit power) than $(3, 2, 1)$ in order to enable NOMA to achieve equal outage probability performance to OMA. Call $S_{(3,2,1)}$ and $S_{(3,1,2)}$ the sums of the power allocation strategies given above for the SIC decoding orders $(3, 2, 1)$ and $(3, 1, 2)$, respectively. Then according to the proof of corollary 1 it is shown that $S_{(3,2,1)} - S_{(3,1,2)} > 0$ as follows

$$S_{(3,2,1)} = \frac{2^{R_3} - 1}{2^{R_3/\tau_3} - 1} + 2^{R_3} \frac{2^{R_2} - 1}{2^{R_2/\tau_2} - 1} + 2^{R_2+R_3} \frac{2^{R_1} - 1}{2^{R_1/\tau_1} - 1},$$
$$S_{(3,1,2)} = \frac{2^{R_3} - 1}{2^{R_3/\tau_3} - 1} + 2^{R_3} \frac{2^{R_1} - 1}{2^{R_1/\tau_1} - 1} + 2^{R_1+R_3} \frac{2^{R_2} - 1}{2^{R_2/\tau_2} - 1}$$

So $\implies S_{(3,2,1)} - S_{(3,1,2)} = 2^{R_3}(2^{R_1} - 1)(2^{R_2} - 1) \left(\frac{1}{2^{R_1/\tau_1} - 1} - \frac{1}{2^{R_2/\tau_2} - 1} \right) > 0$, which is true because $\frac{R_1}{\tau_1} < \frac{R_2}{\tau_2}$. Using the same steps, it is easy to show that $S_{(3,1,2)} - S_{(1,3,2)} > 0$ and $S_{(1,3,2)} - S_{(1,2,3)} > 0$. Therefore, it is clear that the SIC decoding order $(1, 2, 3)$ is more energy efficient than any of the possible SIC decoding orders in terms of the power allocation required for NOMA to achieve the same outage probility performance as OMA. Again, it should be noted that the sum of the power allocation strategies $(\tilde{a}_{\sigma(1)}, \tilde{a}_{\sigma(2)}, \tilde{a}_{\sigma(3)})$ is not guaranteed to be less than 1 for all other SIC decoding orders other than $(1, 2, 3)$, as was proven in theorem 2. Nonetheless, even if there are other SIC decoding orders with power allocation strategies $(\tilde{a}_{\sigma(1)}, \tilde{a}_{\sigma(2)}, \tilde{a}_{\sigma(3)})$ such that the sum is less than 1, the power headroom available in order to increase the power allocation beyond $(\tilde{a}_{\sigma(1)}, \tilde{a}_{\sigma(2)}, \tilde{a}_{\sigma(3)})$ is still less for all SIC decoding orders than that available for $(1, 2, 3)$, and thus the improvement of outage probability performance of NOMA over OMA is less than for the case of all other SIC decoding orders.

4.9.5 Power Allocation Strategy for NOMA to outperform Outage Probability of OMA

The last matter in this example is to demonstrate how to construct well-behaved power allocation strategy which allows NOMA to outperform the outage probability performance of OMA for all users, based on all the above results. After theorem 3, a simple approach is used based on the property of well-behaved power allocation (although not necessarily optimal). For the three-user case, this power allocation approach is given by calling $A_{\text{tot}}^{\text{oma}} = \tilde{a}_1^{\text{oma}} + \tilde{a}_2^{\text{oma}} + \tilde{a}_3^{\text{oma}}$, and then according to equation (4.27)

$$\begin{aligned}
\epsilon_1 &= (1 - A_{\text{tot}}^{\text{oma}})\frac{2^{R_1} - 1}{2^{R_1+R_2+R_3} - 1}2^{R_2+R_3}, \\
\epsilon_2 &= (1 - A_{\text{tot}}^{\text{oma}})\frac{2^{R_2} - 1}{2^{R_1+R_2+R_3} - 1}2^{-R_1+R_3} \\
\epsilon_3 &= (1 - A_{\text{tot}}^{\text{oma}})\frac{2^{R_3} - 1}{2^{R_1+R_2+R_3} - 1}2^{-R_1-R_2}.
\end{aligned}$$

As shown in this chapter and verified in the simulation results, this additional power allocation combined with the general description of the power allocation coefficients in equation (4.19) provide a simple framework for vastly improving the outage performance of downlink NOMA over OMA, without complex suboptimal searches or relying on the base-station having exact knowledge of all of the channel gains.

Chapter 5

Conclusion

In this dissertation, the fundamental approaches and properties of power allocation strategies are investigated for multi-user downlink wireless NOMA systems. For different cases of base-station channel state information and mobile user cached information, the impact on the approach for determining the feasible power allocation strategies were described, and the impact on performance measures were also analyzed.

In chapter 2, the concept of Fair-NOMA was introduced in order to ensure that all NOMA users' downlink transmissions can achieve at least the same capacity as they would in a downlink OMA transmission. The two-user case was used extensively in order to highlight the potential gains that can be achieved using NOMA. Fair-NOMA was used to demonstrate that performance gains can be achieved without focusing on user-pairing, and then was applied to the user-pairing case of the cell-center user and the cell-edge user. After focusing on the two-user case, the focus shifted to the K-user downlink scenario, and the fundamental existence of a Fair-NOMA power allocation strategy was proved to always exist. For the multi-user case, it was also demonstrated that the total power allocation required to achieve the same capacity as OMA for all users decreases as a function of the transmit SNR.

In chapter 3, the application of two candidate technologies for future wireless systems was considered by combining downlink NOMA transmissions with mobile user caching. The combined scheme called CA-NOMA was outlined, describing the basic form in which caching can assist the receive procedure in the downlink. When applied to a two-user downlink NOMA transmission, the impact on the power allocation set was investigated for the case where the two users have different QoS target rates. This led to the derivation of the desired operating set for the power allocation strategy. For the case of Rayleigh fading channels, the union-outage probability was derived and found in closed-form. The approximate optimum power allocation strategy to minimize the union-outage probability was then derived, and shown to be very tight to the true optimum power allocation strategy for reasonably high transmit SNR values.

In chapter 4, NOMA was investigated for the case where the BS does not possess the channel gain information perfectly. In line with more realistic cellular system deployments, it was assumed that the CSI is an approximation of the channel gain, and that the CSI ages between the CSI acquisition at the BS and the downlink transmission occassion. Therefore, this motivated a NOMA approach that does not rely on the channel gains, and thus it was determined that the main power allocation design parameter is the set of target rates of the users' downlink transmissions. Using the target rates, the fundamental relationship between the target rates and the maximum level of interference tolerable for a signal was described, with respect to the SIC decoding order. The concept of well-behaved power allocation strategies was then introduced and motivated, and it was proved that any well-behaved power allocation strategy always exists, and that they always align with the tolerable interference of the SIC stages. The strategy that allows NOMA to achieve the same outage probability performance as OMA was then outlined and proved to always exists for any set of target rates. This strategy was then used to prove that the SIC decoding order proposed, which was derived from the target rates and OMA time resource allocation, is the most energy efficient SIC decoding order out of all possible decoding orders. The procedure and restrictions for ensuring that a power allocation strategy which allows NOMA to achieve a better outage probability performance than OMA for all users was then described, and a simple approach to finding a power allocation strategy was provided that aligns with these findings. Finally, a three-user downlink NOMA example was provided in order to demonstrate the points in this chapter.

Bibliography

[1] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, Vol 105, No 12, Dec 2017

[2] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Vehicular Technology Magazine,* vol. 9, no. 1, pp. 71-78, Mar. 2014

[3] 3rd Generation Partnership Project (3GPP), "Study on downlink multi-user superposition transmission for LTE," TSG RAN Meeting 67, Tech. Rep. RP-150496, March 2015

[4] 3rd Generation Partnership Project (3GPP), "Study on non-orthogonal multiple access (NOMA) in NR," TSG RAN, technical report 38.802, v0.4.0, Nov 2018

[5] P. Wang, J. Xiao, L. Ping, "Comparison of Orthogonal and Non-Orthogonal Approaches to Future Wireless Cellular Systems," *IEEE Vehicular Technology Magazine*, Volume 1, Issue 3, pp. 4-11, Sept. 2006

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, U.S.A., 1991

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005

[8] J. Choi, "H-ARQ based Non-Orthogonal Multiple Access with Successive Interference Cancellation," *2008 IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1-5

[9] J. Schaepperle and A. Ruegg, "Enhancement of Throughput Fairness in 4G Wireless Access Systems by Non-Orthogonal Signaling," *Bell Labs Tecnical Journal*, Vol 13, Issue 4, pp 59-77, 2009.

[10] J. Schaepperle, "Throughput of a Wireless Cell Using Superposition Based Multiple-Acess with Optimized Scheduling," *Personal Indoor and Mobile Radio Communications, 21st International Symposium on*, pp 212-217, 2010.

[11] T. Takeda, K. Higuchi, "Enhanced User Fairness Using Non-orthogonal Access and SIC in Cellular Uplink," *IEEE Vehicular Technology Conference*, Fall 2011.

[12] S. Tomida, K. Higuchi, "Non-orthogonal Access with SIC in Cellular Downlink for User Fairness Enhancement," *ISPACS*, Dec 2011.

[13] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-Level Performance Evaluation of Downlink Non-orthogonal Multiple Access (NOMA)," *PIMRC 2013*, pp. 611-615.

[14] Z. Ding, Z. Yang, P. Fan, H.V. Poor, "On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users," *IEEE Signal Processing Letters*, Vol 21, No 12, pp 1501-1505, 2014

[15] M. Al-Imari, P. Xiao, M.A. Imran, R. Tafazolli, "Uplink Non-Orthogonal Multiple Access for 5G Wireless Networks," *Wireless Communications Systems, 11th International Symposium on*, pp 781-785, August 2014.

[16] F. Liu, P. Mähönen, M. Petrova, "Proportional Fairness-Based User Pairing and Power Allocation for Non-Orthogonal Multiple Access," *PIMRC, 26th Annual*, pp. 1306-1310, 2015.

[17] S. Timotheou and I. Krikidis, "Fairness for Non-Orthogonal Multiple Access in 5G Systems," *IEEE Signal Processing Letters*, Vol 22, No 10, October 2015.

[18] Z. Ding, P. Fan, H.V. Poor, "Impact of User Pairing on 5G Nonorthogonal Multiple-Access Downlink Transmissions," *IEEE Trans. on Veh. Tech.*, Vol 65, No 8, Aug. 2015

[19] J. A. Oviedo and H. R. Sadjadpour, A New NOMA Approach for Fair Power Allocation, *IEEE Conf. on Computer Commun. Workshops (INFOCOM WKSHPS)*, pp. 843847, Apr. 2016

[20] R. Knopp and P. Humblet, "Information Capacity and Power Control in Single-Cell Multiuser Communications," *Proc. of IEEE ICC*, Seattle, Washington, USA, June 18-22 1995.

[21] Z. Yang, Z. Ding, P. Fan, N. Al-Dhahir, "A General Power Allocation Scheme to Guarantee Quality of Service in Downlink and Uplink NOMA Systems," *IEEE Trans. Wireless Commun.*, Vol 15, No 11, Nov 2016

[22] J. A. Oviedo, H. R. Sadjadpour, "A Fair Power Allocation Approach to NOMA in Multi-user SISO Systems," *IEEE Trans. Veh. Tech.*, Vol 66, Issue 9, September 2017

[23] Z. Ding, F. Adachi, H.V. Poor, "The Application of MIMO to Non-Orthogonal Multiple Access," *IEEE Transactions on Wireless Communication*, Vol 15, No 1, Jan 2016

[24] B. Perabathini, E. Bastug, M. Kountouris, M. Debbah, A. Conte, "Caching at the Edge: a Green Perspective for 5G Networks," Workshop on Next Generation Green ICT, ICC 2015, pp 2830-5.

[25] M.K. Kiskani, H.R. Sadjadpour, "Throuput Analysis Decentralized Coded Content Caching," *Wireless Communications, IEEE Transactions on*, Vol 16, No 1, January 2017.

[26] Z. Ding, P. Fan, G. Karagiannidis, R. Schober, H. V. Poor, "NOMA Assisted Wireless Caching: Strategies and Performance Analysis," *IEEE International Conference on Communications (ICC), 2018*, invited paper

[27] J. A. Oviedo and H. R. Sadjadpour, "Leveraging edge caching in NOMA systems with QoS requirements ," *IEEE 15th Annu. Consum. Commun. and Netw. Conf. (CCNC)*, Jan 2018

[28] G. Cardano, "Ars Magna (The Rules of Algebra)," Dover (1992), 1545

[29] E. W. Weisstein, "Quartic Equation" *MathWorld–A Wolfram Web Resource*, `http://mathworld.wolfram.com/QuarticEquation.html`

[30] E. Dahlman, S. Parkvall, J. Skold, *5G NR: The Next Generation Wireless Access Technology*, Academic Press, 2018

[31] M. Zeng, A. Yadav, O. A. Dobre, H. V. Poor, "A Fair Individual Rate Comparison between MIMO-NOMA and MIMO-OMA," *2017 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 04-08, 2017

[32] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, H. V. Poor, "Capacity Comparison Between MIMO-NOMA and MIMO-OMA With Multiple Users in a Cluster," *IEEE J. on Sel. Areas Commun.*, Vol 35, No 10, pp 2413-2424, Oct 2017

[33] J. Mei, L. Yao, H. Long, K. Zheng, "Joint User Pairing and Power Allocation for Downlink Non-Orthogonal Multiple Access Systems," *2016 IEEE Int. Conf. on Commun. (ICC)*, Mobility and Wireless Networking Symposium, May 23-27, 2016

[34] M. Zeng, G. I. Tsiropoulosy, A. Dobre, M. H. Ahmedz, "Power Allocation for Cognitive Radio Networks Employing Non-orthogonal Multiple Access," *2016 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 04-08, 2016

[35] Z. Yang, W. Xu, C. Pan, Y. Pan, M. Chen, "On the Optimality of Power Allocation for NOMA Downlinks With Individual QoS Constraints," *IEEE Commun. Lett.*, Vol 21, No 7, pp 1649-1652, Jul 2017

[36] J. Zhu, J. Wang, Y. Huang, S. He, X. You, L. Yang, "On Optimal Power Allocation for Downlink Non-Orthogonal Multiple Access Systems," *IEEE J. Sel. Areas Commun.*, Vol 35, No 12, pp 2744-2757, Dec 2017

[37] Z. Xiao, L. Zhu, J. Choi, P. Xia, X. Xia, "Joint Power Allocation and Beamforming for Non-Orthogonal Multiple Access (NOMA) in 5G Millimeter Wave Communications," *IEEE Trans. on Wireless Commun.*, Vol 17, No 05, pp 2961-2974, May 2018

[38] Y. Wu, H. Mao, K. Ni, X. Feng, L. Qian, L. Huang, Z. Shi, "Resource optimisation for downlink nonorthogonal multiple access systems: a joint channel bandwidth and power allocations approach," *Inst. Eng. and Tech. J. Communications*, Vol 12, Issue 19, pp 2429-2437, Jul 2018

[39] X. Sun, N. Yang, S. Yan, Z. Ding, D. W. K. Ng, C. Shen, Z. Zhong, "Joint Beamforming and Power Allocation in Downlink NOMA Multiuser MIMO Networks," *IEEE Trans. on Wireless Commun.*, Vol 17, No 8, pp 5367-5381, Aug 2018

[40] F. Liu and M. Petrova, "Dynamic Power Allocation for Downlink Multi-Carrier NOMA Systems," *IEEE Commun. Lett.*, Vol 22, No 9, pp 1930-1933, Sept 2018

[41] Md S. Ali, E. Hossain, A. Al-Dweik, D. I. Kim, "Downlink Power Allocation for CoMP-NOMA in Multi-Cell Networks," *IEEE Trans. Commun.*, Vol 66, No 9, pp 3982-3998, Sept 2018

[42] J. A. Oviedo, H. R. Sadjadpour, "On the Power Allocation Limits for Downlink Multi-User NOMA with QoS," *2018 IEEE Int. Conf. on Commun. (ICC)*," Wireless Communications Symposium, May 21-25, 2018

[43] J. A. Oviedo, Hamid R. Sadjadpour, "Fundamentals of Power Allocation Strategies for Downlink Multi-User NOMA With Target Rates," *IEEE Trans. on Wireless Commun.*, Vol 19, No 3, pp 1906-1917, March 2020