# Throughput Enhancement and Power Optimization in NOMA-based Multiuser Multicast Systems

Sareh Majidi Ivari

A Thesis In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Department of Electrical and Computer Engineering

Concordia University

Montreal,Quebec

September 2023

**CONCORDIA UNIVERSITY**
**SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Sareh Majidi Ivari

Entitled: Throughput Enhancement and Power Optimization in NOMA-based Multiuser Multicast Systems

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

**Doctor Of Philosophy** (Electrical and Computer Engineering)

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

Signed by the final examining committee:

Chair
Dr. Amir G. Aghdam

External Examiner
Dr. Shahrokh Valaee

External to Program
Dr. Hovhannes Harutyunyan

Examiner
Dr. Walaa Hamouda

Examiner
Dr. Dongyu Qiu

Thesis Supervisor (s)
Dr. Mohammad Reza Soleymani

Dr. Yousef R. Shayan

Approved by

Dr. Jun Cai Chair of Department or Graduate Program Director

September 2023
Date of Defence

Dr. MOURAD DEBBABI Dean, Gina Cody School of Engineering and Computer Science

## ABSTRACT

**Throughput Enhancement and Power Optimization in NOMA-based Multiuser Multicast Systems**

**Sareh Majidi Ivari, Ph.D.**
**Concordia University, 2023**

In recent years, Non-Orthogonal Multiple Access (NOMA) has emerged as a promising technique for enhancing the capacity and throughput of wireless communication systems. This thesis investigates the potential of NOMA in improving the performance of multiuser multicast systems, focusing on multibeam satellite communication systems in the forward link, throughput enhancement, and power optimization. We propose a novel framework that combines a NOMA scheme with multibeam architecture and frequency reuse in multicast transmission. The proposed framework enhances system throughput by optimizing power allocation.

First, we present a comprehensive review of the principles and techniques related to NOMA and multibeam multicast systems, highlighting their unique challenges and potential benefits. Next, we introduce our proposed framework in 4-color frequency reuse satellite systems. In 4-color frequency reuse, each user receives signals from other co-channel beams. However, the level of isolation is such that the interbeam interference can be treated as background noise without significant performance degradation. This means that there is no collaboration between beams, and each beam can be isolated from the rest. Therefore, NOMA is considered in single-beam multicast satellite communication systems. The optimum power allocation to maximize the minimum fairness rate and sum-rate is derived for a given user clustering in a single beam. Moreover, an optimum user clustering is derived, which improves the system throughput.

Next, we investigate our proposed framework in full frequency reuse satellite systems under perfect channel state information at the transmitter (CSIT). The proposed framework integrates the NOMA scheme in multicast multibeam architecture. Linear precoding techniques, such as zero-forcing (ZF) and minimum mean square error (MMSE), are used to cancel interbeam interference while NOMA is applied on a beam basis. NOMA and linear precoding are adopted for the proposed framework in multicast transmission. A low-complexity user scheduling is proposed to deal with the trade-offs between optimum user scheduling for linear precoding and the NOMA scheme. Moreover, a low-complexity linear precoding in multicast transmission is proposed based on unicast linear precoding methods and a mapper which deals with the lack of spatial degrees of freedom. To improve the performance of linear precoding, we present three mappers, where the proposed singular-value-decomposition (SVD) mapper demonstrates the best performance.

To improve system throughput, power allocation should be optimized. In this thesis, we consider two objective functions: max-min fairness rate (MMF) and sum-rate. This thesis introduces a technique for addressing the non-convex MMF optimization issue in the proposed framework by employing auxiliary variables to convert it into a semi-definite programming problem, which can then be resolved using linear programming solvers. This thesis also suggests an approach to tackle the non-convex sum-rate maximization goal function in MB-MC-NOMA systems by constructing Lagrangian multipliers concerning the constraints. By employing quadratic transformations on the sum-of-ratios, the problem is restructured within an iterative sum-rate power optimization algorithm.

This thesis considers a realistic scenario with imperfect CSIT. To combat the effect of imperfect CSIT in multibeam multicast satellite communication systems, a rate-splitting approach is proposed. An averaging rate (AR) framework for MMF rate and sum-rate optimization considering ICST is proposed. To render the formulated MMF and sum-rate problems convex, we utilize the Weighted Minimum Mean Square Error (WMMSE) method. We first derive a rate-WMMSE relationship and then, using this relationship along with a low-complexity solution based on Alternating Optimization (AO), we transform the problems into equivalent convex ones.

To validate the effectiveness of our proposed frameworks, we conduct extensive simulations and comparisons with state-of-the-art schemes. The results demonstrate significant improvements in throughput and power efficiency, confirming the potential of NOMA-based multiuser multicast systems for future wireless communication networks.

Finally, we discuss potential future research directions, including the integration of the proposed frameworks in the cellular networks, calculating the transmitter and receiver complexity of the proposed techniques, considering higher layers of RS. This thesis contributes to the ongoing development of next-generation wireless communication systems, paving the way for more efficient and reliable data transmission in multiuser multicast environments.

## DEDICATION

This document is dedicated to my parents, who have always loved and supported me unconditionally. I am certain that my mom watches over me always.

To my wonderful husband and friend, Amir, whose endless love, support, and encouragement embrace me every day.

To our lovely daughters, Emma and Maya, whose presence is a continual source of my energy.

## ACKNOWLEDGMENTS

I would like to start by extending my deep appreciation to my distinguished advisors, Prof. Yousef R. Shayan and Prof. M. Reza Soleymani. Their consistent guidance, patience, and encouragement have been instrumental throughout my Ph.D. process. Their mentorship has profoundly broadened my perspectives and skills. I'm profoundly grateful for the trust they placed in me, allowing me to chase my visions while meticulously guiding my academic growth. My association with them has honed my skills in research, effective articulation, analytical thinking, and the craft of precise and lucid writing.

I am also grateful to my Ph.D. defense committee members: Prof. Walla Hamouda, Prof. Dongyu Qiu, and Prof. Hovhannes A. Harutyunyan. Their insights and the time they dedicated at various stages of my graduate journey are deeply valued. I extend special gratitude to the external examiner, Prof. Shahrokh Valaee from University of Toronto, for his invaluable technical contributions.

I extend a special thank you to my friends and colleagues at the Wireless Communication Lab of Concordia University, with a particular mention to my dear friend Nazli.

I would like to extend my gratitude to my dear parents. I would not be where I am today without their love, support, and encouragement. While my mom is no longer with us, her influence continues to guide me. My heartfelt thanks go to Amir, my confidant and husband, who has consistently been a beacon of love and motivation. I'm deeply grateful to my daughters, Emma and Maya, for accompanying me on this exploratory journey. Lastly, I'd like to express my gratitude to my family, especially my sister, Zahra, for her unwavering support which has been foundational.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AO | Alternating Optimization |
| AR | Averaging Rate |
| BC | Broadcast Channel |
| CSIT | Channel State Information at the Transmitter |
| DPC | Dirty Paper Coding |
| FDMA/TDMA/CDMA | Frequency/Time/Code Division Multiple Access |
| MB-MC-NOMA | Multibeam Multicast NOMA |
| MC-NOMA | Multicast NOMA |
| MC-RS-NOMA | Multicast Rate-splitting NOMA |
| MIMO | Multiple-Input Multiple-Output |
| MMF | Max-Min Fairness |
| MMSE | Minimum Mean Square Error |
| NOMA | Non-orthogonal Multiple Access |
| OMA | Orthogonal Multiple Access |
| PSD | Power Spectral Density |
| QoS | Quality of Service |
| RS | rate-splitting |
| SFPB | Single Feed per Beam |
| SIC | Successive Interference Cancellation |
| SNR | Signal-to-Noise Ratio |
| SUD | Single User Detection |
| SVD | Singular Value Decomposition |
| WMMSE | Weighted MMSE |
| ZF | Zero-Forcing |

## LIST OF SYMBOLS

| | |
|---|---|
| $K$, $N_t$ | Number of beams and antenna feeds |
| $I_k$, $M$ | Number of users in beam $k$ and per group |
| $G$ | Number of multicasting groups of users per beam |
| $\mathcal{U}_k$, $\mathcal{U}_k^t$ | Set of all users and in time slot $t$ in beam $k$ |
| $\mathcal{U}_{k,\mathcal{A}}^t$, $\mathcal{U}_{k,\mathcal{B}}^t$ | Indices of users in groups $\mathcal{A}$ and $\mathcal{B}$ of beam $k$ per time slot |
| $\mathbf{h}_{k,Y}^i$ | Channel vector of $i$-th user in group $\mathcal{A}$ or $\mathcal{B}$ of beam $k$ |
| $\mathbf{f}_{k,Y}^i$, $\overline{\mathbf{h}}_{k,Y}^i$ | Model the fading effect and the channel vector |
| $z_{k,Y}^i$, $w_{k,Y}^i$ | Lognormally distributed line-of-sight and Rayleigh distributed components |
| $\theta_{k,Y}^{i,\mathrm{LoS}}$, $\theta_{k,Y}^{i,\mathrm{MP}}$ | phase of LoS and MP |
| $a_l^i$, $G_R$ | gain from the $l$-th feed to the $i$-th user per beam and Reception antenna gain |
| $d_k^i$ | Distance between $i$-th user in beam $k$ and satellite |
| $\lambda$, $K_B$, | Carrier wavelength, Boltzmann constant, |
| $T$, and $B_W$ | receiver noise temperature, and carrier bandwidth |
| $\Phi_l^i$ | Time-varying phase between $l$-th feed and $i$-th user per beam |
| $\theta_{RF}^i$ | Phase rotation of $i$-th user |
| $\theta_{LNB}^i$ | Phase contribution of the receiver low noise block downconverters |
| $\theta_{PL,l}$ | Payload oscillator phase offsets |
| $\xi_{k,Y}^i$, $\theta_{k,Y}^i$ | Power gain and phase of rain attenuation |
| $\mathbf{x}$ | Transmitted Signal |
| $p_k$ | Allocated power to beam $k$ |
| $\mathbf{W}$, $\mathbf{w}_k$ | Precoding matrix and Precoding vector of beam $k$ |
| $\alpha_k$ | Variable splits power between groups in beam $k$ |
| $s_{k,\mathcal{A}}$,$s_{k,\mathcal{B}}$ | Transmitted symbols for users in groups $\mathcal{A}$ and $\mathcal{B}$ |
| $P_T$ | Maximum available power in the satellite payload |

| | |
|---|---|
| $\mathbf{h}_{k,\mathcal{A}}^{j}$ | $j$-th user's channel vector in groups $\mathcal{A}$ of beam $k$ |
| $\mathbf{h}_{k,\mathcal{B}}^{l}$ | $l$-th user's channel vector in groups $\mathcal{B}$ of beam $k$ |
| $n_{k,\mathcal{A}}^{j}$ ,$n_{k,\mathcal{B}}^{l}$ | Additive white Gaussian noise |
| $d_{k}^{ij}$ | Euclidean distance between users in beam $k$ |
| $\mathbf{g}_k$ | Representative Channel vector of users in beam $k$ |
| $\Sigma_k$ | Singular value matrix of beam $k$ |
| $\mathbf{U}_k\mathbf{U}_k$ | Left(Right)-singular vector |
| $\mathrm{SNR}_k^i$ | SNR of the $i$-th user in beam $k$ |
| $\Gamma_{k,\mathcal{A}}\Gamma_{k,\mathcal{B}}$ | Minimum signal-to-co channel interference plus noise for group $\mathcal{AB}$ of beam $k$ |
| $\hat{\mathrm{SINR}}_{k,\mathcal{A}}\hat{\mathrm{SINR}}_{k,\mathcal{B}}$ | Minimum SINR of users in group $\mathcal{AB}$ of beam $k$ |
| $R_{k,\mathcal{A}}R_{k,\mathcal{B}}$ | Achievable rate of users in group $\mathcal{AB}$ |
| $c_{k,\mathcal{A}}c_{k,\mathcal{B}}$ | Weight of users in group $\mathcal{AB}$ of beam $k$ |
| $p_{k,\mathcal{A}}p_{k,\mathcal{B}}$ | Allocated power to users in group $\mathcal{AB}$ of beam $k$ |
| $\mathcal{C}_k$ | Achievable rate region in beam $k$ |
| $\beta_k$ | Priority factor |
| $G_{k,\mathcal{A}}$, $G_{k,\mathcal{B}}$ | Channel transfer functions |
| $b_{k,A}$, $b_{k,B}$ | Scaling factors |
| $S_{k,\mathcal{A}}S_{k,\mathcal{B}}$ | Optimum PSDs of two groups of users in beam $k$ |
| $\hat{S}_{k,\mathcal{A}}\hat{S}_{k,\mathcal{B}}$ | Optimum PSDs of two groups of users $k$ in beam in the equivalent channel |
| $p_c$ | Allocated power to the common part |
| $\alpha_{k,g}$ | fraction of power allocated to groups g of beam $k$ |
| $\mathbf{w}_c$ | precoding vector of common message |
| $t$ | fraction of power allocated to the private part |
| $\gamma_{c,i}$ | SINR pf common part of $i$-th user |
| $R_c$ | Common rate |
| $c_{k,g}$ | Portion of common rate of group $g$ in beam $k$ |

| | |
|---|---|
| $\gamma_i$ | SINR of private part of $i$-th user |
| $r_{k,g}$ | Rate of private part of users in group $g$ of beam $k$ |
| $R_{k,g}$ | Achievable Rate of users in group $g$ of beam $k$ |
| $\hat{\mathbf{h}}_i, \tilde{\mathbf{h}}_i$ | Estimated channel state and channel estimation error |
| $\sigma^2_{e,i}$ | normalized CSIT error variance |
| $\eta$ | CSIT quality parameter |
| $g_{c,i}, g_i$ | User $i$ MSEs equalizers |
| $\varepsilon_{c,i}, \varepsilon_i$ | User $i$ MSEs estimations |
| $u_{c,i}, u_i$ | User $i$ MSEs weights |
| $\xi_{c,i}, \xi_i$ | User $i$ weighthed MSEs |

# CHAPTER 1
# Introduction

## 1.1 Literature Survey and Motivation

Satellite communication systems play a crucial role in developing the next generation of wireless communications, 5G [1, 2]. The potential for terrestrial wireless networks to become less congested by utilizing satellite technologies allows for anytime-anywhere connectivity.

To achieve broadband interactive data traffic and provide extensive data rates in satellite communication systems, utilization of two key techniques is inevitable; full frequency reuse and multibeam architecture [3]. Full frequency reuse allows for efficient spectrum utilization by reusing the same frequency bands across different beams, maximizing the available bandwidth. This technique enables higher capacity and improved system performance.

In addition, the implementation of a multibeam architecture further enhances the system's capabilities. With a multibeam architecture, the satellite is equipped with multiple beams, each covering a specific geographic area [4]. This allows for simultaneous transmission to multiple users in different regions, enabling better coverage and capacity allocation. By dynamically adjusting beam shapes and power allocation, multibeam architectures can efficiently manage resources and adapt to varying traffic demands and user distributions.

It is important to note that the combination of full frequency reuse and multibeam architecture, while bringing significant advantages to satellite communication systems such as high data rates, increased capacity, and improved quality of service, can also introduce interbeam interference [4]. This interference occurs when signals from adjacent beams overlap, leading to potential performance degradation if not properly managed. Therefore, effective interference mitigation techniques need to be employed to minimize the impact of interbeam interference and optimize system performance.

According to state-of-the-art technologies in DVB-S2X [5], the multibeam design includes a multicast framework due to the large codewords. The transmitter sends a single coded frame to multiple users within a beam simultaneously. It means that users within the same beam share the same precoding vector, thus, the coding gain increases. The multibeam multicast adopts the physical layer (PHY) multigroup multicast transmission, and different beams represent different groups of users [6]. Besides the interbeam interference, the multibeam satellite systems face other obstacles, such as the per-feed available power constraints, the uncertainty of channel status information at the transmitter (CSIT), and the overloaded regime [7]. According to the literature, there are three methods to cancel interbeam interference including linear precoding, non-orthogonal multiple access (NOMA), and rate-splitting (RS) which are discussed in the following sections.

### 1.1.1 Linear precoding

One technique to mitigate interbeam interference is linear precoding, which effectively mitigates interbeam interference in underloaded systems and under perfect CSIT conditions [8]. Linear precoding involves applying a linear transformation to the transmitted signals at the satellite's beams to suppress interference and enhance the received signals at the intended users. In linear precoding, the interference is detected and canceled at the transmitter and the receiver treats it as background noise.

Employing the linear precoding technique for multibeam satellite communication systems involves the initial design of zero-forcing (ZF) and minimum mean square error (MMSE) precoding [9]. These precoding techniques, commonly utilized in scenarios with one user per frame, are adapted from linear precoding approaches employed in multiuser multiple-input-multiple-output (MIMO) systems [10, 11, 12]. ZF is a straightforward method that employs a prefiltering operation using the channel pseudo-inverse, effectively eliminating inter-user interference [10]. MMSE, on the other hand, builds upon ZF by taking into account the noise variance to further enhance performance, particularly in low signal-to-noise ratio (SNR) conditions [11, 12].

The multicast fashion entails a modification of the overall precoding scheme since multiple users within the same frame share the same precoding vector, and each user requires a particular SINR to stay connected. Therefore, applying the linear precoding in the multicast framework needs a particular scheduling scheme to include multiple users data, with various SINR rates, within the same frame with the same precoding filter [4, 13, 14, 15, 16, 17]. Such scheduling optimizations in designing the multicast linear precoding result in a trades-off between the complexity and the performance in precoding techniques [18, 19, 20, 21, 22]. The first attempt of designing a multicast multigroup precoding was a modification of the regularized channel inversion [18], where the equivalent channel for each multicast group is the average of the users' channels. This technique has low complexity and spectral efficiency. To improve the performance, a method based on the singular value decomposition (SVD) is proposed in [19] to compute the precoding vectors. The precoding matrix is constructed row by row, where each row is the null space projection to reject the other beams' interference. This method achieves higher spectral efficiency, whereas the computational complexity is increased.

Another approach in designing the linear precoding matrix is based on sum-rate optimization, as discussed in [20]. This approach aims to maximize the overall system throughput while considering per-antenna power constraints. However, it is important to note that this approach often comes with increased computational complexity.

In [21], a frame-based technique is proposed, which considers joint sum-rate optimization with per-antenna power constraints and user scheduling. This technique aims to maximize the sum rate of the system while efficiently managing power allocation and user selection. Another approach, presented in [22], proposes a two-stage linear precoding design that achieves higher spectral efficiency and lower complexity compared to other techniques. These advancements in linear precoding techniques have contributed to enhancing the overall system performance.

In the context of employing linear precoding in a multicasting framework, the design of a low-complexity and high-performance linear precoding scheme presents significant challenges.

Additionally, an important issue that needs to be addressed is the user scheduling in multicast linear precoding systems. This thesis focuses on addressing these issues.

### 1.1.2 Non-orthogonal Multiple Access

In addition to linear precoding, NOMA scheme has emerged as another technique to mitigate interbeam interference. The NOMA scheme relies on the superposition coding at the transmitter and the successive interference cancellation (SIC) at the receiver [23, 24]. The NOMA scheme is suitable for the overloaded systems and serves a number of users that could be much higher than the number of feeds [25]. In this case, the SIC receiver is able to cope with the strongest interfering signal that generally cannot be completely removed by mentioned linear precoding techniques. In the NOMA scheme, users could be ordered after the precoding based on their effective scalar channel from the weakest to the strongest users. Users can decode the messages of the weaker users in a successive manner [26, 27]. Therefore, the strongest receiver can decode all messages. Consequently, the NOMA can transform the multi-antenna non-degraded channel into a single-antenna degraded channel. Moreover, it is well established that NOMA offers an impressive throughput boost compared to orthogonal multiple access (OMA) techniques.

The NOMA scheme in the multicast framework of the terrestrial networks is investigated in [28, 29, 30]. In [28], a cooperative NOMA scheme is designed for a unicast-multicast system. In the first phase, the base station transmits a superposed message containing all users information. Then, in the second phase, a multicast user is selected to forward the information intended by unsuccessfully decoded users (multicast or unicast). In [29], the NOMA scheme is studied in a two-layered multicast system. The multicast users are divided into the base layer and the enhanced layer. Also, an algorithm is proposed to maximize the sum-rate while the minimum rate of the base layer is guaranteed. However, in [30], the cooperative NOMA is designed for a system with only multicast users. In the second phase of the proposed system, successfully received users forward the received message to the remaining users through a device-to-device multicast way.

The application of NOMA in satellite communication systems has been extensively researched [24, 25, 31, 32, 33, 34, 35]. These studies discuss various non-orthogonal schemes and payload architectures suitable for satellite environments, highlighting the potential advantages of NOMA in satellite multibeam communications. However, further investigation is needed to fully understand and optimize the benefits of NOMA in satellite communication systems.

Some research focuses on addressing the challenges of implementing NOMA in multibeam satellite systems. For instance, [31] investigates joint precoding of signals in NOMA systems, while exploring the use of simultaneous non-unique detection (SND) to enhance spectral efficiency. Novel scheduling algorithms that consider both successive interference cancellation (SIC) and SND strategies are proposed to improve system performance and spectral efficiency.

In the context of multibeam satellite communication, [32] proposes a cooperative NOMA scheme where beams collaborate to serve users at the beam edges, utilizing the strongest co-channel interference (CCI) as additional information. [33] introduces NOMA to enhance frequency reuse and mitigate intra-beam interference by formulating a max-min resource allocation problem. The study proposes a suboptimal algorithm to optimize the Offered Capacity to requested Traffic Ratio (OCTR), demonstrating the potential benefits of incorporating NOMA in multibeam satellite systems.

Additionally, [34] presents a geographical NOMA-based multiuser beamforming (NOMA-BF) scheme for improving spectral efficiency in multibeam satellite-based Internet of Things (IoT) systems. The paper explores the advantages of NOMA in enhancing spectral efficiency. While [35] provides a comprehensive overview of NOMA's application in various satellite architectures, addressing the availability, coverage, and efficiency requirements of 5G networks.

Considerable research has been conducted on the application of the NOMA scheme in various scenarios. However, to the best of our knowledge, there is no research that specifically addresses the use of NOMA in beam basis of a multibeam multicast satellite communication system. This thesis

aims to fill this research gap by investigating the potential benefits and challenges of implementing the NOMA scheme in such a scenario.

### 1.1.3 Rate-splitting

RS is a versatile framework that offers a flexible approach to interference mitigation in multiuser communication systems, making it more generalized compared to linear precoding or NOMA [36]. In RS, the message for each user is divided into a common part and a private part. The simplest form of RS, known as 1-layer RS, combines the common parts into a single common stream, while encoding the private messages as individual private streams [37]. At the receiver, the common stream is decoded and removed from the received signal using SIC. Subsequently, the private stream intended for a specific user is decoded, treating the other interfering streams as noise [37].

RS has demonstrated its effectiveness in mitigating interference and achieving good performance, even in scenarios with imperfect CSIT and in overloaded regimes [8, 38]. The flexible nature of RS enables it to adapt to different levels of interference. It can automatically switch between linear precoding and NOMA by adjusting the powers and contents of the common and private streams, depending on the strength of the interference. Thus, RS acts as a bridge between linear precoding and NOMA, encompassing any possible hard switching between the two techniques.

RS has been extensively studied in the context of multiuser MIMO systems with perfect CSIT [38, 39]. It enables the exploitation of multiuser interference, leading to higher spectral efficiency and increased system capacity, and improve MMF rate performance in scenarios with perfect CSIT [8]. However, the assumption of perfect CSIT may not hold in practical scenarios, leading to the investigation of robust RS algorithms to mitigate the impact of imperfect CSIT [39, 40].

In the context of multicast transmission, RS has gained attention due to the challenges associated with multicast transmission [41]. Multicast transmission presents unique challenges, as it involves simultaneously transmitting the same information to multiple users. The problem

of MMF transmit beamforming in the multigroup multicasting framework under perfect CSIT is studied and showed that the RS outperforms the conventional linear precoding [41].

RS has also been investigated in the context of satellite communication systems, as highlighted in [42, 43, 44]. In [42], the focus is on superposition coding (SC) and RS in a two-beam satellite communication system, where orthogonal multiple access schemes like TDMA are considered within each beam. The study examines the performance of SC and RS in this specific satellite configuration.

In [43], RS is evaluated in the context of multibeam multicast satellite communication systems. The authors address a per-feed power-constrained max-min fair (MMF) problem, taking into account various qualities of CSIT, including both perfect and imperfect CSIT. They assess the performance of RS under different CSIT scenarios and validate its potential in improving system performance compared to existing techniques.

Building upon the previous work, [44] considers a more realistic scenario by incorporating imperfect CSIT in both underloaded and overloaded systems. The proposed RS framework is investigated in the context of multibeam multicast systems, and the results demonstrate significant performance improvements over existing techniques.

All the existing research mentioned focuses solely on the application of RS technique for mitigating interbeam interference in multicast transmission. However, none of them consider the combination of RS with other techniques to further improve system performance. Ongoing research continues to explore the potential of RS and its performance in practical satellite communication systems such as NOMA-based multibeam multicast satellite systems.

## 1.2 Objectives and Thesis Contributions

In the previous section, we discussed the most important interbeam interference cancellation techniques in satellite communication systems such as linear precoding, NOMA, and RS. While each technique has its benefits and limitations, it is important to carefully consider their characteristics and trade-offs in the context of multibeam multicast satellite communication systems.

Linear precoding offers advantages such as low complexity and the ability to perfectly mitigate interference under perfect CSIT and underloaded conditions. However, its performance may degrade under imperfect CSIT and in overloaded scenarios, which are limitations of this technique. Furthermore, designing multicast linear precoding is not straightforward, and there exists a trade-off between complexity and performance.

NOMA provides benefits in efficiently handling an overloaded regime and achieving high spectral efficiency. However, it is important to address some limitations associated with NOMA. As the number of users increases, the complexity of NOMA systems escalates, posing challenges in terms of implementation and processing. Additionally, the performance of NOMA is influenced by signal-to-noise ratio (SNR) imbalances among users, which can impact overall system performance.

The RS is a promising solution to mitigate interbeam interference, and it has better performance than the other techniques, even in systems under imperfect CSIT. RS offers a flexible approach to interference mitigation and can automatically adapt to different levels of interference by adjusting the common and private streams. However, it has higher computational complexity. The additional complexity arises from the need to design the RS and decoding strategies.

In conclusion, considering the benefits and limitations of each technique, using only one method to mitigate interbeam interference in a fully overloaded multicast multibeam satellite system is ineffective. Combining two techniques gives more flexibility to optimize the Max-Min fairness (MMF) rate and the sum-rate. Further research is also needed to address the limitations, optimize the performance, and explore the application of multicast NOMA and linear precoding and RS in satellite communication systems.

### 1.2.1 Objectives

The primary aim of this thesis is to enhance the performance of multibeam multicast satellite communication systems by employing a combination of techniques to mitigate interbeam interference. The research focuses on evaluating and optimizing three distinct frameworks.

The first framework of this thesis, known as MC-NOMA, focuses on the analysis of NOMA within a single beam in the context of multicast transmission. This framework extensively investigates the effectiveness of NOMA and its impact on various system performance metrics, such as the MMF rate and total achievable sum-rate. The goal is to gain insights into the advantages and limitations of implementing NOMA on a beam basis in multicast scenarios.

In the second framework, denoted as MB-MC-NOMA, the integration of linear precoding and NOMA techniques is explored in multibeam multicast satellite communication systems. Linear precoding is employed to mitigate interbeam interference and cancel out unwanted signals between beams, while NOMA is applied within each beam to further enhance system performance in terms of MMF rate and sum-rate. By combining these two techniques, the aim is to achieve improved interference cancellation capabilities through linear precoding and leverage the benefits of NOMA in a multibeam multicast scenario.

The third framework, MC-RS-NOMA, focuses on the integration of RS and NOMA techniques in multibeam multicast satellite communication systems under imperfect CSIT. The objective is to exploit the benefits of RS in handling imperfect CSIT and combine it with NOMA to improve system performance, particularly in terms of MMF rate and total achievable sum-rate. By combining these two techniques, the goal is to enhance the overall system performance and address the challenges of interbeam interference in multibeam multicast satellite communication systems.

The proposed frameworks aim to achieve a more effective and efficient system for multicast multibeam satellite communications. The research will address the limitations of existing approaches and provide valuable insights into the optimization of system performance in various CSIT scenarios. Based on the comprehensive analysis and research conducted, the subsequent subsections discuss the specific objectives and contributions of this thesis in a clear and concise manner.

### 1.2.2 Thesis Contributions

Based on the defined objectives, this thesis is dedicated to the development of novel frameworks in the field of multibeam multicast satellite communication systems. Through the research process, several significant contributions have been made to advance this field. The key contributions of this thesis can be summarized as follows:

- Deriving achievable rates of MC-NOMA framework and the optimum power allocation to maximize MMF rate and sum-rate with considering Quality-of-Service (QoS). Moreover, an optimum user clustering is proposed for improving the performance .
- Designing a user scheduling method to improve the performance of both multicast linear precoding and the MC-NOMA scheme in the second framework, MB-MC-NOMA. The proposed user scheduling method obtains the optimal compromise between SINR imbalance and the co-linearity of the channel vectors.
- Designing the multicast linear precoding technique for the MB-MC-NOMA framework. To deal with the lack of spatial degrees of freedom in the multicast linear precoding, we propose a new formation of the composite channel matrix as the equivalent virtual channel. We present three different mappers. The mappers are governed by SVD, SNR, and averaging.
- Optimizing the power allocation to maximize the MMF rate and the weighted sum-rate in the MB-MC-NOMA with and without considering the QoS.
- Developing the achievable rate region for the MB-MC-NOMA scheme, where the optimal PSD is efficiently computed based on the proposed equivalent channel and water-filling algorithm for the weighted rate sum.
- Applying rate splitting for the first time to the multibeam multicast NOMA satellite communication systems under imperfect CSIT assumption, called MC-RS-NOMA. The rate splitting is used to cancel interbeam interference and combat the effect of the imperfect CSIT. Moreover, NOMA is considered on a beam basis to improve spectral efficiency. The achievable data rates of the common and private parts are derived.

- Formulating the MMF rate and sum-rate optimization problems of the MC-RS-NOMA under imperfect CSIT assumption using AR framework. This thesis employs the weighted MMSE (WMMSE) approach to make the formulated MMF and sum-rate problems convex. First, a rate-WMMSE relationship is derived. Then, using the rate-WMMSE relationship, and a low-complexity solution based on alternating optimization (AO), the problems are transferred into equivalent convex problems.

## 1.3 Thesis Organization

This thesis is organized as follows. Chapter 2 provides the necessary background on the main topics referenced throughout the thesis. In Chapter 3, the MC-NOMA framework is presented, including the derivation of achievable data rates and proposed power allocation methods. In this chapter, an optimal user clustering approach to maximize system performance is also discussed. Chapter 4 introduces the proposed framework for multibeam satellite communication systems, MB-MC-NOMA. An optimal user scheduling scheme is developed to optimize the performance of linear precoding and the NOMA scheme in multicast transmission. Additionally, the chapter explores multicast linear precoding based on unicast linear precoding, introducing three mappers: averaging, maximum SNR, and SVD. In Chapter 5, an optimal power allocation scheme to maximize minimum rate and sum-rate in the MB-MC-NOMA framework is proposed. The chapter also includes the derivation of the achievable rate region. Chapter 6 focuses on the MC-RS-NOMA framework under imperfect CSIT. This chapter investigates the theoretical findings presented throughout the thesis and validates them through simulations across various scenarios. The thesis is concluded in Chapter 7, which provides a summary of the key findings and contributions. Additionally, potential avenues for future research are identified.

# CHAPTER 2
# Background

## 2.1 Introduction

This chapter is providing a comprehensive analysis of the fundamental topics that form the basis of the subsequent chapters. It delves into the exploration of the Non-Orthogonal Multiple Access (NOMA) scheme, which plays a pivotal role in the proposed frameworks. By examining this scheme, we gain a deeper understanding of the proposed farmework. The chapter focuses on investigating linear precoding techniques, which are suboptimal strategies that can achieve the capacity of the broadcast channel. These techniques serve as the primary building blocks in the proposed frameworks of this thesis. Furthermore, the chapter delves into the water filling technique, providing an in-depth analysis of its application and impact on the system. Additionally, the chapter addresses the crucial aspect of CSIT uncertainty model and its scaling with SNR. By discussing these topics, we shed light on the challenges and considerations associated with imperfect knowledge of the channel state.

## 2.2 Broadcast Channel

The forward link of satellite communication systems can be modeled as a broadcast channel (BC). A BC is a wireless communication system that consists of one transmitter (or information source) and multiple uncoordinated receivers. In the context of information theory, a BC encompasses various scenarios where the transmitter communicates independent messages to the receivers, or a combination of independent and common messages, among other possibilities [45, 46]. The term ”broadcasting” refers to the transmission of information over a shared medium.

In the case where both the transmitter and receivers are equipped with a single antenna, the communication scenario is known as a multiuser single-input single-output (SISO) system. The capacity of the multiuser SISO channel, which can be modeled as a degraded-broadcast channel,

is well-known [45]. In the degraded-broadcast channel, users can be ordered in terms of their channel norms. Non-orthogonal transmission techniques can achieve the capacity of the multiuser SISO channel, and will be further discussed in detail in subsection 2.3.2.

On the other hand, if the transmitter is equipped with multiple antennas, $N_t$, it is referred to as a multiple-input multiple-output (MIMO) system. We use the term MIMO loosely to describe the scenario where the transmitter is equipped with multiple antennas while the receivers may have only a single antenna each. Unlike the degraded broadcast channel, the MIMO broadcast channel is generally not characterized by a specific ordering of the channels. As a result, the capacity region of the MIMO broadcast channel is not generally known or determined.

Consider a $J$-user MIMO BC, and the received complex baseband signal at the $j$-th receiver during the $t$-th channel use mathematically expressed as

$$y_j t = \mathbf{h}_j t \mathbf{x} t \; n_j t, \tag{2.1}$$

where $\mathbf{h}_j t \in \mathbb{C}^{1\times N_t}$ is the channel vector between the transmitter and the receiver. Moreover, $j \in 1, J$ and $t \in \; 0, 1.$ $\mathbf{x}t \in \mathbb{C}^{N_t \times 1}$ and $n_j t$ are the transmitted signal and the additive white Gaussian noise. Moreover, the input signal is constrained by an average power limit per channel use, which is defined as

$$\mathbb{E}\{\|\mathbf{x}\|^2\} \leq P_T \tag{2.2}$$

where $P_T$ is the maximum available power at the transmitter.

### 2.2.1 Capacity of MIMO Broadcast Channel

In this case, the capacity region of the MIMO BC, subject to a power constraint, is defined as the closure of the set of all achievable rate tuples $r_1, ..., r_J$ that satisfy the power constraint. The capacity region represents the tradeoff between the achievable rates for different receivers and is influenced by the type of transmission employed. In the non-ergodic case, the capacity region is affected by the instantaneous channel state, while in ergodic transmission, it is determined by the long-term properties of the channel.

The accuracy and availability of CSIT play a crucial role in defining the capacity region [47, 48, 49]. In the case of perfect CSIT, the capacity region of the MIMO BC can be achieved using a non-linear coding strategy called Dirty Paper Coding (DPC) [50]. However, when CSIT is imperfect, the capacity region is generally unknown.

The capacity region of a communication system provides valuable insights into its performance, and various scalar performance measures can be derived from it. One commonly used measure is the sum-capacity, which is defined as

$$C_{\text{Sum-rate}} = \max_{r_1,\ldots,r_J \in \mathcal{C}} \sum_{j=1}^{J} r_j \tag{2.3}$$

which represents the maximum achievable sum-rate. Although operating at the sum-capacity ensures optimal system throughput, it may not guarantee fairness among users. In some cases, users with poor channel conditions may experience significant resource deprivation while resources are allocated to users with better channel conditions to maximize overall throughput. To address this, the Max-Min Fair (MMF) capacity is introduced as a performance measure that considers fairness among users. The MMF capacity is defined as

$$C_{\text{MMF}} = \max_{r_1,\ldots,r_J \in \mathcal{C}} \min_{j} r_j. \tag{2.4}$$

The performance measure described as the MMF capacity is also referred to as the symmetric-capacity in literature [51]. This is because the symmetric rate tuple, where all users achieve the same rate denoted as $C_{\text{MMF}}$, is considered optimal in terms of achieving fairness among users. The objectives of this thesis revolve around optimizing the measures mentioned in equations (2.3) and (2.4), or their suboptimal versions that may not necessarily satisfy the maximization operators. These objectives are central to the design considerations explored in this research.

### 2.2.2 Precoding techniques

Precoding is a signal processing technique commonly employed in MIMO systems to improve the performance and achieve the capacity. The basic idea behind precoding is to pre-process (or

shape) the transmitted signals in a way that mitigates the adverse effects of the wireless channel, such as interference, fading, and noise. This improves the overall system performance and data rate. In the following sections, two types of precoding are presented, non-linear precoding and linear precoding.

**Non-linear precoding**

Non-linear precoding methods generally offer better performance in terms of data rate and error performance compared to linear methods, but they come at the cost of higher computational complexity.

- Tomlinson-Harashima Precoding (THP): THP is one of the earliest and best-known non-linear precoding techniques. It effectively mitigates inter-symbol interference by performing modulo operations and backward decision-feedback equalization. It is particularly effective in scenarios where the channel matrix is ill-conditioned or near singular. These are the situations where linear precoders might perform poorly.
- Vector Perturbation (VP): This method introduces a perturbation vector to the transmitted signal vector such that the received signal after the channel aligns closely with lattice points, reducing quantization noise and hence improving performance. It achieves near-optimal performance at the expense of increased complexity due to the search for the best perturbation vector.
- Dirty Paper Coding (DPC): This is an information-theoretic technique wherein the transmitter has knowledge of the interference and cancels it by coding the information in a way that the receiver does not see the interference. In practice, exact DPC is computationally complex, but there are approximations and practical schemes that aim to achieve the benefits of DPC without the high complexity. DPC enables the achievement of the MIMO BC by employing distinct and suitable codewords for each receiver [50]. The transmitter begins by selecting an appropriate codeword specifically intended for Receiver 1. Subsequently, the codeword for Receiver 2 is chosen based on the complete knowledge (without any causality

constraints) of the first codeword intended for Receiver 1. As a result, Receiver 2 does not see the codeword intended for Receiver 1 as an interference signal. This process continues iteratively, ensuring that the last receiver does not experience the codewords intended for other receivers as interference. However, the first receiver perceives all the other codewords as interference.

### 2.2.3 Linear precoding

Although DPC is a theoretically optimal concept, its practical implementation is considered highly complex and challenging [52, 53]. As an alternative approach, linear precoding, also known as Beamforming (BF), has emerged as a suboptimal but more practical strategy. In linear precoding, each message is independently encoded into a data stream and then mapped to transmit antennas using a precoding vector consisting of beamforming weights. This simplifies the design of codes for MIMO broadcast channels, making the problem more tractable and less complex. Throughout this discussion, we assume perfect CSIT. In the following subsection, we discuss two most important linear precoding techniques which are considered in this thesis.

#### Zero-forcing

The zero-Forcing (ZF) is a simple and practical alternative precoding used to cancel the interbeam interference and is useful for the high SNR regime. The ZF precoding technique achieves maximum capacity under perfect CSIT. The beamforming vectors in the ZF, $\mathbf{W_{ZF}}$, are computed from the pseudo-inverse of the composite channel matrix $\mathbf{H} = \mathbf{h}_1\mathbf{h}_2...\mathbf{h}_J$ [10]:

$$\mathbf{W_{ZF}} = 1\sqrt{\gamma_{\mathbf{ZF}}}\left(\mathbf{H}^{\dagger}\right) = 1\gamma_{\mathbf{ZF}}\left(\left(\mathbf{H}^{H}\mathbf{H}\mathbf{H}^{H}\right)^{-1}\right). \tag{2.5}$$

The precoding matrix should be divided by

$$\gamma_{\mathbf{ZF}} = \max_{n}\text{diag}\mathbf{W_{ZF}}\mathbf{W_{ZF}}^{H} \tag{2.6}$$

**MMSE**

The Minimum Mean Square Error (MMSE) approach shows low computational complexity and good sum-rate performance [54]. The precoding matrix is given by

$$\mathbf{W}_{\mathbf{MMSE}} = 1\sqrt{\gamma_{\mathbf{MMSE}}}\left(\left(\mathbf{H}^{H}\mathbf{H}\,\frac{K}{P_{\mathrm{T}}}\mathbf{I}_{K}\right)^{-1}\mathbf{H}^{H}\right), \tag{2.7}$$

where $\mathbf{I}_K$ is the $K$-dimensional identity matrix. To control the power and satisfy the power constraints, the precoding matrix should be divided by,

$$\gamma_{\mathbf{MMSE}} = \max_{n}\left(\mathrm{diag}\left(\mathbf{W}_{\mathbf{MMSE}}\left(\mathbf{W}_{\mathbf{MMSE}}\right)^{H}\right)\right). \tag{2.8}$$

## 2.3 Information-theoretic views of the NOMA scheme

From an information-theoretic standpoint, users in the NOMA scheme share the same resource elements, including time, frequency, space, and code [45, 46]. The NOMA technique offers a superior rate region compared to OMA techniques. Examples of OMA techniques include frequency division multiple access (FDMA), time division multiple access (TDMA), and code division multiple access (CDMA). It has been demonstrated that non-orthogonal transmission is the optimal choice for both the uplink and downlink scenarios.

### 2.3.1 Capacity region of two-user MAC (Uplink)

The capacity region of a two-user multiple access channel (MAC) can be achieved through non-orthogonal transmission. In the uplink channel scenario, multiple users simultaneously transmit their data to a common base station (BS) using the same time and frequency resources. The BS applies the SIC to decode the signal from the strongest user. Single user detection (SUD) is then used to decode the message from the other user. The capacity region of the two-user MAC is obtained using the non-orthogonal multiple access (NOMA) scheme and represents the set of

non-negative $R_1, R_2$ values that satisfy certain conditions such that

$$R_1 \leq \mathcal{C}\gamma_1$$

$$R_2 \leq \mathcal{C}\gamma_2$$

$$R_1\ R_2 \leq \mathcal{C}\gamma_1\ \gamma_2$$

where $\gamma_i$ is the received signal-to-noise ratio (SNR) for user $i$ and $\mathcal{C}x = 0.5\log_2 1\ x$.

### 2.3.2 Capacity region of two-user BC (Downlink)

Similar to the two-user MAC, the capacity region of the two-user BC is known and is achieved via non-orthogonal transmission in which both users signals are served at the same time and in the same frequency band. In particular, the BS sends the superposition coding of messages and the user with the stronger channel gain (usually the one closer to the BS) uses SIC to decode its signal free of interference, while the user with the weaker channel gain treats the signal of the stronger users as noise. The capacity region of the two-user BC is the set of non-negative $R_1, R_2$ such that

$$R_1 \leq \mathcal{C}\alpha\gamma_1$$
$$R_2 \leq \mathcal{C}\frac{1-\alpha\gamma_2}{\alpha\gamma_2\ 1}$$

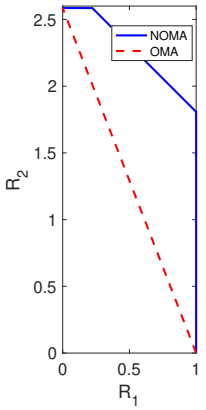
where $\alpha$ is the fraction of the BS power allocated to user 1s data (strongest user). Figure 2–2 and Figure2–1b show the capacity region of NOMA and OMA in uplink and downlink, respectively. The capacity that OMA can achieve in both uplink and downlink for $\beta \in 0, 1$ is only

$$R_1 \leq \beta\mathcal{C}\gamma_1$$

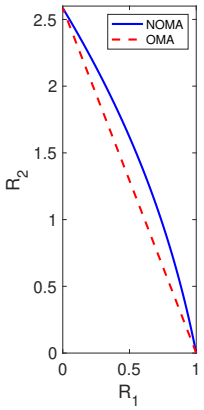
$$R_2 \leq 1 - \beta\mathcal{C}\gamma_2$$

## 2.4 Capacity of Multiple Access Channel (MAC) with ISI Using water-filling approach

In this thesis, the water-filling approach is employed to derive the achievable rate region of the proposed MB-MC-NOMA framework. To enhance our understanding of this approach, this section concentrates on examining the capacity of the MAC with inter-symbol interference




(a) Capacity region of two-user MAC (Uplink)




(b) Capacity region of two-user BC (Downlink)

Figure 2–1: Capacity region of two-user in Uplink and Downlink, $\gamma_1 = 1, \gamma_2 = 5$

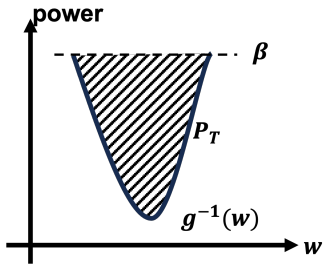



Figure 2–2: Single-user water-filling scheme

(ISI), which is accomplished by utilizing the water-filling technique. Extensive research on the capacity region of the Gaussian MAC with ISI has been conducted in the reference [55]. The study employs a multiuser water-filling scheme to achieve the capacity region in this context. Although the details of the study are provided in the referenced work, we can provide a brief overview of the key findings.

Before examining the analysis of the capacity of a two-user MAC with ISI, let's first examine the water-filling approach for a single user with a transfer function denoted as $Gw$. Figure 2–2 shows the water-filling scheme for a single user channel. The square of the magnitude of the channel transfer function over the noise power spectral density (PSD), denoted as $gw = |Gw|^2 Nw$, where $Nw$ is the noise PSD. The $g^{-1}w$ is the bottom of the water-filling container and the fixed amount of water (power), $P_T$, is poured into the container.

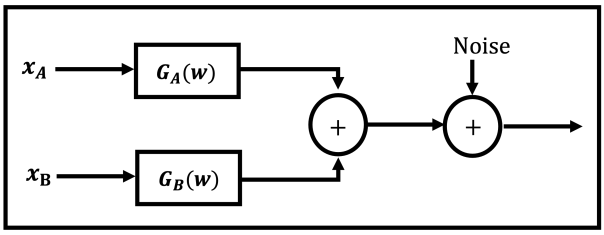



Figure 2–3: Two-user multiple access channel model

According to the water-filling scheme, the optimal PSD, denoted as $Sw$ and depicted by the shaded area in Figure 2–2, can be determined by solving the following equations:

$$Sw = \beta - g^{-1}w \tag{2.9}$$

$$P_T = \frac{1}{\pi}\int_0^{\pi} Swdw \tag{2.10}$$

where the  sign indicates that $Sw \geq 0$.

The capacity of memoryless Gaussian MAC is well-established, but evaluating the capacity region of a MAC with ISI requires a different approach. In this case, the channel needs to be decomposed into parallel memoryless channels, which is more challenging compared to single-user channels due to the interdependence between the channels. To overcome this challenge, a water-filling scheme can be employed to visualize the PSD distribution across the frequency domain. This scheme allows us to graphically represent the optimal total PSD. By incorporating the successive cancellation concept into the water-filling scheme, we can determine the optimal PSD for each user over the entire bandwidth.

Figure 2–3 depicts the channel model for the two-user Gaussian MAC with ISI, where $G_{\mathcal{A}}w$ and $G_{\mathcal{B}}w$ represent the transfer functions of the channels. In our analysis, we specifically investigate a two-user MAC with both similar and dissimilar transfer functions. We provide an explanation of the case with similar transfer functions to enhance understanding, but our primary focus is on the more relevant case with dissimilar transfer functions.

### 2.4.1 Two-user Channel with the same transfer function

In the case of similar transfer functions of two-user MAC, the same approach employed for the single user MAC can be applied to determine the optimal PSD. $g^{-1}w = g_{\mathcal{A}}^{-1}w = g_{\mathcal{B}}^{-1}w$ will be used as the bottom of the containers for $S_{\mathcal{A}}w, S_{\mathcal{B}}w$, and $S_{\mathcal{A}}w\ S_{\mathcal{B}}w$ water diagrams seperately while satisfying the following equations

$$S_{\mathcal{A}}w\ S_{\mathcal{B}}w = \beta - g^{-1}w \tag{2.11}$$

$$P_{\mathcal{A}}\ P_{\mathcal{B}} = \frac{1}{\pi}\int_0^{\pi} S_{\mathcal{A}}w\ S_{\mathcal{B}}w dw \tag{2.12}$$

$$S_i w = \beta_i - g^{-1}w \tag{2.13}$$

$$P_i = \frac{1}{\pi}\int_0^{\pi} S_i w dw \tag{2.14}$$

where $i = \mathcal{A}$ or $\mathcal{B}$ and $P_{\mathcal{A}}\ P_{\mathcal{B}} = P_T$.

### 2.4.2 Two-user Channel with different transfer function

In this scenario, the objective is to determine the optimal PSD for the two users that maximizes a weighted sum-rate, denoted as $\beta R_{\mathcal{A}}\ 1-\beta R_{\mathcal{B}}$, where $\beta \in 0,1$. However, it is not possible to find the optimal PSD for each user separately using two separate water-filling diagrams, as they would interfere with each other. To overcome this challenge, an equivalent channel is introduced, combining both $g_{\mathcal{A}}^{-1}w$ and $g_{\mathcal{B}}^{-1}w$. This equivalent channel allows us to determine the optimal sum of the PSDs as well as the individual PSDs for each user.

To merge the two water-filling diagrams, an equivalent channel can be utilized, as depicted in Figure 2–4. This equivalent channel scales the original channels in order to facilitate the combination of the two water-fillings into a single diagram. The optimal PSD for the equivalent channel, which maximizes the weighted sum rate, can be obtained by solving the following equations:

$$C = \{R_{\mathcal{A}}, R_{\mathcal{B}} \in \mathcal{R}^2 : \beta R_{\mathcal{A}}\ 1-\beta R_{\mathcal{B}} \leq C\beta\} \tag{2.15}$$

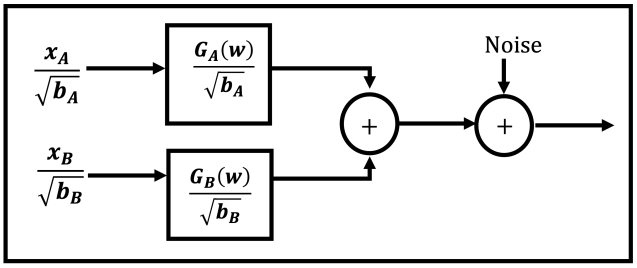



Figure 2–4: Equivalent two-user multiple access channel model

where $C\beta$ is defined as

$$C\beta = \begin{cases} \beta FS_{\mathcal{A}}, S_{\mathcal{B}}\ 1-2\beta F0, S_{\mathcal{B}} & \text{if}\beta \in 0, 0.5 \\ 1-\beta FS_{\mathcal{A}}, S_{\mathcal{B}}\ 2\beta - 1FS_{\mathcal{A}}, 0 & \text{if}\beta \in 0.5, 1 \end{cases} \tag{2.16}$$

where $FS_{\mathcal{A}}, S_{\mathcal{B}}$ can be written as

$$FZ_{\mathcal{A}}, Z_{\mathcal{B}} = \frac{1}{2\pi}\int_0^{\pi} \log\left[1\ Z_{\mathcal{A}} w g_{\mathcal{A}} w\ Z_{\mathcal{B}} w g_{\mathcal{B}}\right] dw \tag{2.17}$$

The rationale behind these equations is based on the SIC methodology, which outlines how to divide the combined water-filling diagram for the equivalent channel. The user with a weaker channel is assigned lower priority and is decoded first, treating the signal from the second user as interference. The first user's signal is then reconstructed and subtracted from the total signal, enabling the decoding of the second user's signal without any influence from the first user.

For example, when $\beta \in 0, 0.5$, where User 1 is considered low priority, the corresponding data is recovered while treating the signal from User 2 as interference, i.e., $\beta FS_1, S_2$. On the other hand, User 2, as the higher priority user, decodes its data as if there were no User 1, i.e., $1-\beta F0, S_2$. Similar principles apply to cases where $\beta \in 0.5, 1$. In the case of equal priority with $\beta = 0.5$, the sum rate is divided between users based on their respective channel conditions.

Figure 2–5 illustrates a representative water-filling diagram for $\beta \in 0, 0.5$. With a fixed water level of $\beta$, the bottom of the container is determined as the minimum of the two curves, $b_{\mathcal{A}} g_{\mathcal{A}}^{-1} w$ and $b_{\mathcal{B}} g_{\mathcal{B}}^{-1} w\ 2\beta - 1$. Moreover, the parameters $b_A$ and $b_B$ are adjusted so that the total amount

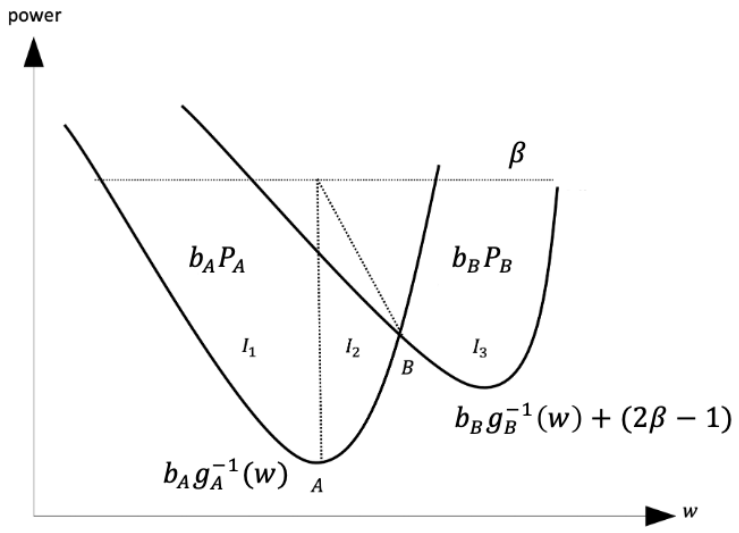



Figure 2–5: Water-filling diagram for two-user MAC with different transfer functions

of water (power) is equal to $b_{\mathcal{A}}P_{\mathcal{A}}$ $b_{\mathcal{B}}P_{\mathcal{B}}$. Then areas $I_1, I_2, I_3$ define the optimal PSD for the two users.

To determine the optimal scales $b_{\mathcal{A}}$ and $b_{\mathcal{B}}$ that maximize the sum rate, we have employed the approach described in [56]. This methodology utilizes a geometrical water-filling method to achieve the desired outcome. The main objective is to maximize the weighted sum-rate of the two users, as stated in equation (2.15), while ensuring that each user's power constraints are satisfied.

By iteratively adjusting the value of $\beta$, both the water-filling diagram shown in Figure 2–5 undergoes change. Consequently, considering the channel transfer function and the PSD of each user, various combinations of scaling parameters will be computed. The optimal scales $b_{\mathcal{A}}$ and $b_{\mathcal{B}}$ are then determined by selecting the values that maximize the weighted sum-rate.

For the two-user MAC, where the capacity region is rectangular, a rectangular region is defined in the two-dimensional coordinate system for each value of $\beta$. This region is utilized to calculate the relative pairs $b_{\mathcal{A}}, b_{\mathcal{B}}$. The outer boundary of this region is determined by considering single-user detection independently, without considering the presence of the other user. This process yields the maximum values of $b_{\mathcal{A}}$ and $b_{\mathcal{B}}$, denoted as $b_{\mathcal{A},max}$ and $b_{\mathcal{B},max}$, respectively,

which can be mathematically expressed as:

$$_{w} b_{\mathcal{A},max} g_{\mathcal{A}}^{-1} w - b_{\mathcal{A}} P_{\mathcal{A}} = 0, \tag{2.18}$$

$$_{w} b_{\mathcal{B},max} g_{\mathcal{B}}^{-1} w \; 2\beta - 1 - b_{\mathcal{B}} P_{\mathcal{B}} = 0. \tag{2.19}$$

The optimal pair $b_{\mathcal{A}}, b_{\mathcal{B}}$ can be determined through an iterative process. Initially, the pair $b_{\mathcal{A}}, b_{\mathcal{B}}$ is set to the maximum threshold pair $b_{\mathcal{A},max}, b_{\mathcal{B},max}$. Then, the pair is adjusted iteratively until both the individual power constraints and the total power constraints are satisfied:

$$_{w} S_{\mathcal{A}} - b_{\mathcal{A}} P_{\mathcal{A}} = 0, \tag{2.20}$$

$$_{w} S_{\mathcal{B}} - b_{\mathcal{B}} P_{\mathcal{B}} = 0, \tag{2.21}$$

$$_{w} S_{\mathcal{A}} \; S_{\mathcal{B}} - b_{\mathcal{A}} P_{\mathcal{A}} \; b_{\mathcal{B}} P_{\mathcal{B}} = 0. \tag{2.22}$$

By varying the priority factor $\beta$ from 0 to 1, we calculate different points in the sum-rate capacity. By associating all the resulting points, we can determine the final capacity region.

## 2.5 CSIT uncertainty Model

This research work assumes that the receiver has perfect channel state information available (CSIR), enabling accurate knowledge of the channel conditions. However, the accuracy of the CSIT can vary due to various factors, including estimation errors in time-division duplexing (TDD) systems [57, 58], quantization errors in frequency-division duplexing (FDD) systems [59], and information staleness caused by transmission delays [60]. These imperfections in CSIT are taken into account and considered during the analysis and evaluation of the proposed schemes in this thesis. The imperfect CSIT of user-$i$ is modelled by

$$\mathbf{h}_i = \hat{\mathbf{h}}_i \; \tilde{\mathbf{h}}_i, \quad i \in \mathcal{U} \tag{2.23}$$

where $\mathbf{h}_i$ is the channel coefficient of $i$-th user. $\hat{\mathbf{h}}_i$ and $\tilde{\mathbf{h}}_i$ denote estimated channel state and the corresponding channel estimation error at the transmitter, respectively. CSIT uncertainty (channel

estimation error) can be characterized by a conditional density $f\left(\mathbf{h}|\hat{\mathbf{h}}\right)$, which is known at the transmitter [61].

We define $\Upsilon_i = \mathbb{E}|\mathbf{h}_i|^2$, $\hat{\Upsilon}_i = \mathbb{E}|\hat{\mathbf{h}}_i|^2$, and $\tilde{\Upsilon}_i = \mathbb{E}|\tilde{\mathbf{h}}_i|^2$ for each user. According to the principles of orthogonality, $\hat{\mathbf{h}}_i$ and $\tilde{\mathbf{h}}i$ are uncorrelated, and $\tilde{\mathbf{h}}i$ has a zero mean. Thus, we can express $\Upsilon_i$ as the sum of $\hat{\Upsilon}_i$ and $\tilde{\Upsilon}_i$, yielding $\hat{\Upsilon}_i = 1 - \sigma_{e,i}^2\Upsilon_i$ and $\tilde{\Upsilon}_i = \sigma_{e,i}^2\Upsilon_i$ for some $\sigma_{e,i}^2 \in 0, 1$. Here, $\sigma_{e,i}^2$ represents the normalized error variance of the CSIT [61, 62]. A value of $\sigma_{e,i}^2 = 1$ corresponds to no instantaneous CSIT, while $\sigma_{e,i}^2 = 0$ represents perfect instantaneous CSIT.

### 2.5.1 CSIT scalling with SNR

The error variance of the CSIT scales with the SNR as $\sigma_{e,i}^2 = P_T^{-\eta_i}$, where $P_T$ is the total available transmit power and the noise power is assumed to be 1 [63]. The parameter $\eta_i$ is the CSIT quality scaling factor for user $i$, which quantifies the degradation of CSIT as the SNR increases. In this thesis, it is assumed that all users have identical normalized error variances, denoted as $\sigma_e^2$, which are given by $\sigma_e^2 = P_T^{-\eta}$.

The CSIT quality scaling factor, $\eta$, represents the relation between the number of feedback bits and the SNR. A value of $\eta = 0$ indicates a fixed number of feedback bits regardless of the SNR, while $\eta = \infty$ corresponds to an infinite number of feedback bits. The scaling factors are truncated to the range $0, 1$ for practical considerations. In the context of multiplexing gain, a value of $\eta = 1$ represents perfect CSIT, where the interference caused by multiple users can be reduced to the level of noise. It should be noted that the CSIT quality scaling factor, $\eta$, has various interpretations in addition to its relationship with limited feedback, such as its relation to the Doppler process in delayed or outdated CSIT [60, 64].

# CHAPTER 3
# NOMA in single beam multicast satellite systems

## 3.1 Introduction

In recent years, the demand for extensive data rates in broadband satellite communication systems has increased due to the growing data traffic. To cope with this demand and achieve high data throughput, the utilization of two important techniques is inevitable: non-orthogonal techniques and frequency reuse [25].

Power domain NOMA has been proven advantageous for improving user fairness and the attainable data rate when users are served with a significant signal-to-noise (SNR) imbalance [25]. Multicast transmission, which embeds more than one user's information into the same frame, can also be used in satellite communication to make the most efficient use of satellite resources [28]. However, multicast transmission and NOMA have been studied separately in previous research.

This chapter aims to investigate the first framework which is a combination of NOMA and multicast transmission in satellite communication, referred to as multicast NOMA (MC-NOMA). Specifically, we examine the performance of MC-NOMA in a single beam of the 4-color frequency reuse of satellite communication. Furthermore, we consider the optimization of power allocation and user clustering to maximize the system performance.

Two power allocation optimization problems are considered in this chapter: maximizing minimum fairness and maximizing sum-rate with quality of service (QoS). The proposed MC-NOMA scheme benefits from the theory developed in NOMA. The proposed optimum user clustering method aims to maximize the system performance, as user clustering impacts the system throughput.

The power allocation and user clustering problems are decoupled into two separate optimization problems, allowing us to solve them separately. Specifically, for the power allocation

problem, we consider user clustering to be fixed, while for a given power allocation, we optimize user clustering to improve system performance.

## 3.2 System model

Consider the forward link of a multibeam satellite system that tessellates the coverage area into $K$ beams. The frequency is reused across the coverage area according to a 4-color pattern. Due to the frequency reuse, each user receives the signals from the other co-channel beams. However, the level of isolation is such that the interference can be treated as a background noise without significant performance degradation. This means that there is no collaboration between beams and each beam can be isolated form the rest. From the information theory it is known that the power domain NOMA can be applied on a beam basis to increase the sum-rate with respect to orthogonal schemes, such as time and frequency division multiplexing. When MC-NOMA comes into play, each beam creates $G$ groups.

This chapter considers $G = 2$, which we called groups $\mathcal{A}$ and $\mathcal{B}$. Let $\mathcal{U}_k = \{1, ..., 2M\}$ gather the users' indices in beam $k$. In each beam, $2M$ single antenna users form two multicasting groups of $M$ users, as shown in Figure 3–1, groups $\mathcal{A}$ and $\mathcal{B}$. To group $2M$ users, $\mathcal{U}_k$ should be divided into two disjoint sets $\mathcal{U}_{k,\mathcal{A}}, \mathcal{U}_{k,\mathcal{B}}$, with cardinality of $M$.

### 3.2.1 Channel model

The Land mobile satellite (LMS) model is used in this thesis to model the propagation conditions [65]. The channel is considered constant during a frame transmission. Therefore, the channel coefficient of $i$-th user in beam $k$ is defined as follows

$$h_k^i = f_k^i \overline{\mathrm{h}}_k^i \tag{3.1}$$

where $f_k^i$ describes the fading effects. The channel obeys the Loo distribution [65]. The Loo model assumed that the line-of-sight (LoS) components is lognormally distributed, while the multipath component's attenuation is Rayleigh distributed. Therefore, the fading effect is defined as

$$f_k = z_k e^{j\theta_k^{\mathrm{LoS}}} \; w_k e^{j\theta_k^{\mathrm{multipath}}} \tag{3.2}$$

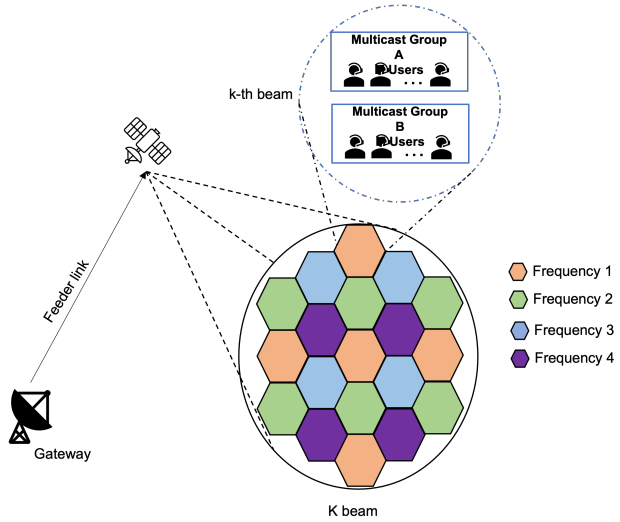



Figure 3–1: System model of the proposed MC-NOMA in a 4-colour frequency reuse of the satellite systems

where $z_k$ is lognormally distributed, $w_k$ is Rayleigh distributed, and $\theta_k^{\text{LoS}}$ and $\theta_k^{\text{multipath}}$ are uniformly distributed between 0 and $2\pi$. The mean, the standard deviation, and the average power parameters for the distribution functions are chosen from [66]. The rest of the effects are modeled by $\bar{\text{h}}_k^i$ that is defined as follows

$$\bar{\text{h}}_k^i = \frac{\sqrt{G_R} a_k^i e^{j\psi_k^i}}{4\pi \frac{d_k^i}{\lambda} \sqrt{K_B T_R B_W}} \tag{3.3}$$

where $G_R$ is the receiver antenna gain, $a_k^i$ is the gain from $k$-th feed to the $i$-th user at beam k. In addition, $e^{j\psi_k^i}$ represents the time varying phase due to the beam radiation pattern and radiowave propagation. $d_k^i$ is the distance between $i$-th user at beam $k$ and the satellite. Finally, $\lambda$, $K_B$, $T_B$, and $B_W$ are the carrier wavelength, the Boltzmann constant, the receiver noise temperature, and the carrier bandwidth, respectively. Note that the channel is normalized to the noise power. Hence, the noise terms in (3.4) and (6–1) have unit variance.

### 3.2.2 Signal model

If we focus the attention on beam $k$, the received signal at each group in beam $k$ is expressed as follows

$$y^{j}_{k,\mathcal{A}} = h^{j}_{k,\mathcal{A}} \left( \sqrt{\alpha_k p_k} s_{k,\mathcal{A}} \; \sqrt{1-\alpha_k p_k} s_{k,\mathcal{B}} \right) I^{j}_{k,\mathcal{A}} \; n^{j}_{k,\mathcal{A}} \quad j \in \mathcal{U}^{t}_{k,\mathcal{A}} \tag{3.4}$$

$$y^{l}_{k,\mathcal{B}} = h^{l}_{k,\mathcal{B}} \left( \sqrt{\alpha_k p_k} s_{k,\mathcal{A}} \; \sqrt{1-\alpha_k p_k} s_{k,\mathcal{B}} \right) I^{l}_{k,\mathcal{B}} \; n^{l}_{k,\mathcal{B}} \quad l \in \mathcal{U}^{t}_{k,\mathcal{B}} \tag{3.5}$$

where $j$ and $l$ superscripts refer to the $j$-th and the $l$-th user in groups $\mathcal{A}$ and $\mathcal{B}$, respectively. $\mathcal{U}^{t}_{k,\mathcal{A}}$ ($\mathcal{U}^{t}_{k,\mathcal{B}}$) gathers the indices of those users that form group $\mathcal{A}$ ($\mathcal{B}$) in time slot $t$. For the remainder of this chapter, we will omit the variable $t$ for the sake of simplicity. The cardinality of each group is $M$. Hence, there are $2M$ users to be served. The coefficients $h^{j}_{k,\mathcal{A}}, h^{l}_{k,\mathcal{B}}$ denote the channel associated with the reference beam for users in groups $\mathcal{A}$ and $\mathcal{B}$, respectively. Hence, $I^{j}_{k,\mathcal{A}}$ and $I^{l}_{k,\mathcal{B}}$ represent the co-channel interference that comes from the adjacent beams. Note that $p_k$ is the transmit power of beam $k$ and $s_{k,\mathcal{A}}, s_{k,\mathcal{B}}$ are the transmitted symbols that are intended for users in groups $\mathcal{A}$ and $\mathcal{B}$, respectively. To be concise, symbol indices are omitted. According to the key concept of NOMA, the transmitted signal is formed by the superposition of two signals, i.e., $s_k = \sqrt{\alpha_k} s_{k,\mathcal{A}} \; \sqrt{1-\alpha_k} s_{k,\mathcal{B}}$. The term $\alpha_k \in 0, 1$ is a variable that controls the power split. Finally, $n^{j}_{k,\mathcal{A}}$ and $n^{l}_{k,\mathcal{B}}$ are the additive noise terms that contaminate the reception of users in each group. The interference plus noise terms, i.e. , $I_{k,\mathcal{A}} \; n_{k,\mathcal{A}}$ and $I_{k,\mathcal{B}} \; n_{k,\mathcal{B}}$ are distributed as $\mathcal{CN}0, N^{j}_{k,\mathcal{A}}$ and $\mathcal{CN}0, N^{l}_{k,\mathcal{B}}$, respectively.

Following the NOMA approach under the assumption that users in group $\mathcal{B}$ experience better the channel conditions than those in group $\mathcal{A}$, it follows that for fairness (unlike sum-rate maximization) more power is allocated to users of group $\mathcal{A}$. Therefore $\alpha_k \geq 0.5$ and users of groups $\mathcal{A}$ and $\mathcal{B}$ can perform SUD and SIC, respectively. Without loss of generality, maximum achievable rates under the Gaussian signaling in beam $k$ are

$$R_{k,\mathcal{A}} = \min_{j \in \mathcal{U}_{k,\mathcal{A}}} \log_2 \left( 1 \; \frac{\alpha_k \mathrm{SINR}^{j}_{k,\mathcal{A}}}{1 \; 1 - \alpha_k \mathrm{SINR}^{j}_{k,\mathcal{A}}} \right). \tag{3.6}$$

$$R_{k,\mathcal{B}} = \min_{l \in \mathcal{U}_{k,\mathcal{B}}} \log_2 \left(1\ 1 - \alpha_k \mathrm{SINR}^l_{k,\mathcal{B}}\right), \tag{3.7}$$

if

$$\min_{l \in \mathcal{U}_{k,\mathcal{B}}} \log_2 \left(1\ \frac{\alpha_k \mathrm{SINR}^l_{k,\mathcal{B}}}{1\ 1 - \alpha_k \mathrm{SINR}^l_{k,\mathcal{B}}}\right) \geq R_{k,\mathcal{A}}. \tag{3.8}$$

The rates have been compactly expressed as a function of signal-to-interference-plus-noise ratio (SINR) defined as follows:

$$\mathrm{SINR}^j_{k,\mathcal{A}} = \frac{p_k \left(h^j_{k,\mathcal{A}}\right)^2}{N^j_{k,\mathcal{A}}}, \mathrm{SINR}^l_{k,\mathcal{B}} = \frac{p_k \left(h^l_{k,\mathcal{B}}\right)^2}{N^l_{k,\mathcal{B}}}. \tag{3.9}$$

It is worth mentioning that if the function $fx$ is strictly increasing for $x \geq 0$, then $\min fx = f\min x$. It can be verified that functions (3.6) and (3.7) are strictly increasing for $\mathrm{SINR} \geq 0$ if $0 \leq \alpha_k \leq 1$. Consequently, we can introduce

$$\Gamma_{k,A} = \min_{j \in \mathbf{I}_A} \mathrm{SINR}^j_{k,A} \tag{3.10}$$

$$\Gamma_{k,B} = \min_{l \in \mathbf{I}_B} \mathrm{SINR}^l_{k,B}, \tag{3.11}$$

to compactly express the rates as follow

$$R_{k,A} = \log_2 \left(1\ \frac{\alpha_k \Gamma_{k,A}}{1\ 1 - \alpha_k \Gamma_{k,A}}\right), \tag{3.12}$$

$$R_{k,B} = \log_2 \left(1\ 1 - \alpha_k \Gamma_{k,B}\right). \tag{3.13}$$

which are formulated under the assumption

$$\log_2 \left(1\ \frac{\alpha_k \Gamma_{k,B}}{1\ 1 - \alpha_k \Gamma_{k,B}}\right) \geq R_{k,A}. \tag{3.14}$$

This inequality guarantees that $s_{k,A}$ and $s_{k,B}$ can be decoded by users of group B and it is equivalent to

$$\Gamma_{k,B} \geq \Gamma_{k,A}. \tag{3.15}$$

If the condition in (3.15) is not satisfied, then the users of group B can not apply SIC and not decode the signal of interest in the absence of interference. In such a case, the roles should be exchanged so that users belonging to group A and B perform SIC and single-user decoding, respectively. In this chapter, it is assumed that the condition in (3.15) is satisfied and more power is allocated to the users of group A. In the following sections, we study the power allocation to maximize the minimum rate and sum-rate.

## 3.3 Power allocation in MC-NOMA

In this thesis, we investigate the optimal power allocation to maximize the MMF rate and the sum-rate.

### 3.3.1 MMF rate analysis

For a given user clustering, we pose an optimization problem optimizing the power is equivalent to optimize the $\alpha$. In this case the goal is to maximize the fairness between users of a single beam. The optimum power allocation of $\alpha_k$ in beam $k$ between groups A and B to maximize the minimum fairness is formulated as follows

$$\begin{aligned} &\underset{\alpha_k}{\text{argmax}} \quad && \min\{R_{k,\mathcal{A}}, R_{k,\mathcal{B}}\} \\ &\text{subject to} \quad && \alpha_k \in 0, 1, \Gamma_{k,\mathcal{B}} \geq \Gamma_{k,\mathcal{A}} \end{aligned}$$

The problem of power allocation to maximize the minimum fairness for a unicast NOMA scheme is known [27, 67, 68, 69], however it is unknown for a MC-NOMA scheme. The NOMA scheme enables a flexible management of the users achievable rates and provides an efficient way to enhance the user fairness. In this section an optimal power allocation to achieve the MMF rate between users of a single beam in MC-NOMA scheme is studied. Using Equations (3.12) and (3.13), the optimization problem turns into the MMF rate optimization in unicast NOMA. Therefore, the problem is solved using the method proposed in [67]. The optimum $\alpha_k$ is equal to $\alpha_k = \alpha_k^*$, namely,

$$\alpha_k^* = \frac{2\Gamma_{k,\mathcal{A}}\Gamma_{k,\mathcal{B}}\ \Gamma_{k,\mathcal{A}}\ \Gamma_{k,\mathcal{B}} - \sqrt{\Gamma_{k,\mathcal{A}}\ \Gamma_{k,\mathcal{B}}{}^2\ 4\Gamma_{k,\mathcal{A}}^2\Gamma_{k,\mathcal{B}}}}{2\Gamma_{k,\mathcal{A}}\Gamma_{k,\mathcal{B}}} \tag{3.16}$$

The optimum $\alpha_k$ is given in a closed form. In addition, if $\alpha \neq \alpha^*$, then it can be verified that the fairness is degraded. Therefore, achievable rates for users of groups $\mathcal{A}$ and $\mathcal{B}$ in beam k are

$$R_k = R_{k,\mathcal{A}}\ R_{k,\mathcal{B}} = 2\log_2\left(\frac{\Gamma_{k,\mathcal{A}} - \Gamma_{k,\mathcal{B}}\ \sqrt{\Gamma_{k,\mathcal{A}}\ \Gamma_{k,\mathcal{B}}{}^2\ 4\Gamma_{k,\mathcal{A}}^2\Gamma_{k,\mathcal{B}}}}{2\Gamma_{k,\mathcal{A}}}\right) \tag{3.17}$$

The Equation (3.17) shows that the MC-NOMA provides absolute fairness for two group of users in beam k.

### 3.3.2 Sum-rate with QoS

The sum of Equations (3.12) and (3.13) is a strictly decreasing function for $0 \leq \alpha_k \leq 1$. Therefore, the minimum of $\alpha_k$ maximizes the sum-rate without any constraint and it means that all power should be allocated to the users of group B. In order not to shut down weaker users we place a constraint on minimum rate. The SR maximization with constraints in unicast NOMA is studied is in [67], [70]. In this case the power allocation problem is given by

$$\begin{aligned} \max_{\alpha_k} \quad & R_{k,A}\ R_{k,B} \\ \text{subject to} \quad & \alpha_k \in 0, 1, R_{k,A} \geq R_{k,A}^{\text{OMA}} \end{aligned}$$

It is considered that $Rk, \mathcal{A}^{\text{OMA}}$ is equal to the rate that users in group $\mathcal{A}$ would achieve if groups are served in an orthogonal multiple access fashion, i.e., $R_{k,\mathcal{A}}^{\text{OMA}} = 0.5\log_2 1\ \Gamma_{k,A}$. From the constraint, it can be infererred that

$$\frac{1\ \Gamma_{k,A} - \sqrt{1\ \Gamma_{k,A}}}{\Gamma_{k,A}} \leq \alpha_k^* \tag{3.18}$$

By design, $\alpha_k^*$ is always greater than 0.5 and lower than one because $\Gamma_{k,A}$ is positive. The optimization problem can be written as follows:

$$\begin{aligned} \max_{\alpha_k} \quad & R_{k,A}\ R_{k,B} \\ \text{subject to} \quad & \alpha \in \left[\frac{1\ \Gamma_{k,A} - \sqrt{1\ \Gamma_{k,A}}}{\Gamma_{k,A}}, 1\right] \end{aligned}$$

The maximum of sum-rate is achieved by minimum of $\alpha_k$ and is $\alpha_k^{*}* = \frac{1\Gamma_{k,A} - \sqrt{1\Gamma_{k,A}}}{\Gamma_{k,A}}$. Therefore the maximum sum-rate that is maximized subject to the constraint becomes

$$R_k = R_{k,A}\ R_{k,B} = \frac{1}{2}\log_2 1\ \Gamma_{k,A}\ \log_2\left(\frac{\Gamma_{k,A} - \Gamma_{k,B}\ \Gamma_{k,B}\sqrt{1\ \Gamma_{k,A}}}{\Gamma_{k,A}}\right) \tag{3.19}$$

### 3.4 User clustering in MC-NOMA

In this section, we study the clustering of users into two groups, $\mathcal{A}$ and $\mathcal{B}$, for a given power allocation in each beam. For convenience, we assume that all clusters have $M$ users. User clustering can generally be classified into two types: random clustering and ordered clustering.

Consider a specific partitioning partitioning $P$ over the indices, as follows:

$$P^t : \mathcal{U}_k \longrightarrow \mathcal{U}_{k,\mathcal{A}}^t, \mathcal{U}_{k,\mathcal{B}}^t \tag{3.20}$$

where,

$$\mathcal{U}_k = \{1, 2M\}, M \geq 2$$
$$\mathcal{U}_{k,\mathcal{A}}^t = \{i_j^t | i_j^t \in \mathcal{U}_k, \forall j \in 1, M\}$$
$$\mathcal{U}_{k,\mathcal{B}}^t = \mathcal{U}_k - \mathcal{U}_{k,\mathcal{A}}^t$$
$$t \in \left\{1\cdot, \frac{2M!}{M!^2}\right\}$$

The partitioning $t$ is selected such that

$$\mathcal{U}_{k,\mathcal{A}}^t \subset \mathcal{U}_k, \mathcal{U}_{k,\mathcal{B}}^t \subset \mathcal{U}_k, \mathcal{U}_{k,\mathcal{A}}^t\ \mathcal{U}_{k,\mathcal{B}}^t = \emptyset, \mathcal{U}_{k,\mathcal{A}}^t\ \mathcal{U}_{k,\mathcal{B}}^t = \mathcal{U}_k,$$

Using the properties of the clustering, in the following two kind of clustering are studied.

#### 3.4.1 Random clustering

In this category, the set $\mathcal{U}_k = \{1, ..., 2M\}$ is divided into two disjoint groups,$\mathcal{U}_{k,\mathcal{A}}, \mathcal{U}_{k,\mathcal{B}}$, randomly and without any criterion.

In random clustering, the group that has index associated with the lowest SINR is labeled as group $\mathcal{U}_{k,\mathcal{A}}$ and the other one as group $\mathcal{U}_{k,\mathcal{B}}$. Therefore, users of groups $\mathcal{A}$ and $\mathcal{B}$ perform SUD and SIC, respectively.

### 3.4.2 Ordered clustering

In this section a clustering method is derived to optimize the performance of the MC-NOMA. In this method, at each time the clustering is done based on judicious user selection. The problem can be formulated as follows:

$$\begin{aligned} &\underset{\mathcal{U}_A^t, \mathcal{U}_B^t}{\text{argmax}} \quad \mathrm{R}_{\mathrm{k},\mathcal{A}}\, \mathrm{R}_{\mathrm{k},\mathcal{B}} \\ &\text{s.t.} \qquad \mathcal{U}_{k,\mathcal{A}}^t\, \mathcal{U}_{k,\mathcal{B}}^t = \mathcal{U}_k, \mathcal{U}_{k,\mathcal{A}}^t\, \mathcal{U}_{k,\mathcal{B}}^t = \emptyset, \min_{j \in \mathcal{U}_{k,\mathcal{A}}^t} \mathrm{SINR}_{k,\mathcal{A}}^j \leq \min_{l \in \mathcal{U}_{k,\mathcal{B}}^t} \mathrm{SINR}_{k,\mathcal{B}}^l \end{aligned}$$

The rates $R_{k,\mathcal{A}}$ and $R_{k,\mathcal{B}}$ are formulated in the equations (3.12) and (3.13), respectively. The constraints indicate how the clustering should be made so that (3.15) is satisfied.

**Proposition 1.** *The optimal clustering which maximizes the sum-rate must satisfy this inequality*

$$\max_{j \in \mathcal{U}_{k,\mathcal{A}}^t} SINR_{k,\mathcal{A}}^j \leq \min_{l \in \mathcal{U}_{k,\mathcal{B}}^t} SINR_{k,\mathcal{B}}^l \tag{3.21}$$

*Proof.* Since the sum-rate is strictly increasing for $\Gamma_{k,\mathcal{A}} \geq 0$ and $\Gamma_{k,\mathcal{B}} \geq 0$, thus the maximum sum-rate is achieved if $\Gamma_{k,\mathcal{A}}$ and $\Gamma_{k,\mathcal{B}}$ are maximized without violating the condition that $\Gamma_{k,\mathcal{A}} \leq \Gamma_{k,\mathcal{B}}$.

Consider two partitioning $t_{opt}$ and $t_0$. We denote $t_{opt}$ the optimal partitioning, which satisfies (6.1). for convenience and to be consistent with the notation of the chapter, let us assume that $\mathcal{U}_{k,\mathcal{A}}^{t_{opt}}$ and $\mathcal{U}_{k,\mathcal{A}}^{t_0}$ gather the indices of the weak users, while $\mathcal{U}_{k,\mathcal{B}}^{t_{opt}}$ and $\mathcal{U}_{k,\mathcal{B}}^{t_0}$ identify the strong users. Now suppose that

$$\min_{l \in \mathcal{U}_{k,\mathcal{B}}^{t_{opt}}} \mathrm{SINR}_{k,\mathcal{B}}^l \leq \min_{l \in \mathcal{U}_{k,\mathcal{B}}^{t_0}} \mathrm{SINR}_{k,\mathcal{B}}^l$$

$$\min_{j \in \mathcal{U}_{k,\mathcal{A}}^{t_{opt}}} \mathrm{SINR}_A = \min_{j \in \mathcal{U}_{k,\mathcal{A}}^{t_0}} \mathrm{SINR}_A$$

Table 3–1: Simulation parameters.

| | |
|---|---|
| Carrier Frequency | 20GHz |
| Orbit | GEO |
| G/T | 17.68 dB/K |
| user location distribution | uniform |
| Beam radiation pattern | Provided by ESA |
| Beam Radius | 140Km |

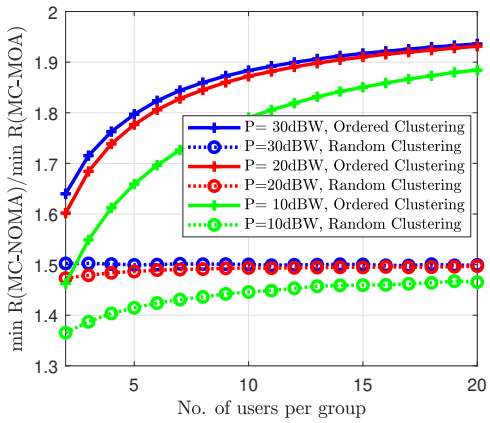
By grouping users differently, it becomes evident that the sum-rate, which is governed by the weakest users, would decrease. This result contradicts the initial hypothesis and thus, (3.21) must be satisfied. This concludes the proof. □

### 3.5 Simulation results

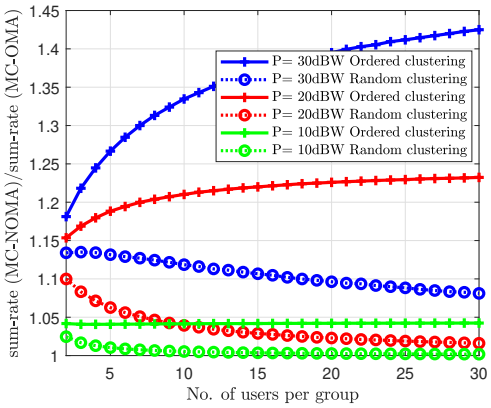
In this section, we present some numerical results of the proposed power allocation and clustering algorithms for a single beam MC-NOMA scheme, according to the different optimization criteria. We consider the forward link of a 4-color frequency reuse satellite communication systems. The simulation model consists of a single beam and the interference from the other beams are considered as the background noise. For the LMS channel model, we have used the statistical information provided in [66] for the ka band and the intermediate shadowing. The parameters of the simulation are given in the table 3–1.

Figure 3–2 compares the performance of MC-NOMA with MC-OMA (where groups of users are served in different time slots) in terms of MMF rate and sum-rate with QoS, for different numbers of users per group, denoted by $M$.

Figure 3–2a shows the achievable maximum fairness rate in MC-NOMA compared to MC-OMA under two different types of clustering for various numbers of users per group. The figure presents the results for three different transmit powers. As expected, MC-NOMA outperforms the OMA scheme, and ordered clustering outperforms random clustering. The simulation results show that as the number of users per group increases for a given transmit power, the MC-NOMA scheme achieves more gain compared to the MC-OMA scheme. The simulation results shows that




(a) Comparison MMF rate of MC-NOMA with MC-OMA




(b) Comparison sum-rate with QoS of MC-NOMA with MC-OMA

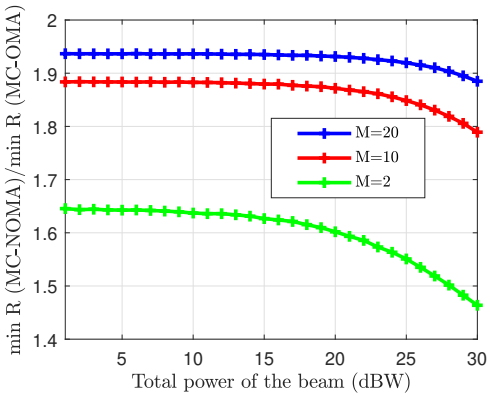
Figure 3–2: Comparison the performance of MC-NOMA with MC-OMA versus number of users per group.

the gain increases the other way around, i.e. 20% increases if the number of users increases from 2 to 20 per group.

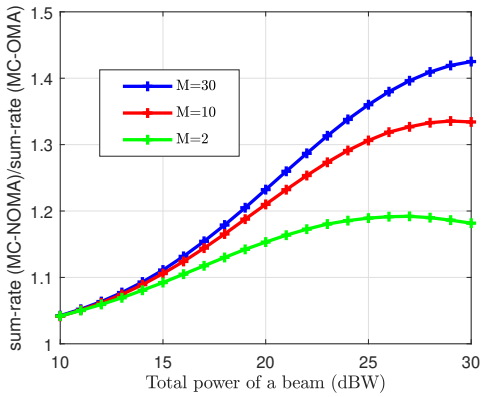
Figure 3–2b shows the achievable maximum sum-rate with QoS in MC-NOMA compared to MC-OMA under two different types of clustering for various numbers of users per group. The figure presents the results for three different transmit powers. MC-NOMA has better performance compared to OMA when ordered clustering is used, and the gain increases with an increasing number of users per group. However, the gain of MC-NOMA compared to OMA decreases as the number of users per group increases under random clustering, and MC-NOMA does not achieve any gain compared to MC-OMA. The simulation results show that the gain increases by 7% if the number of users increases from 2 to 20 per group under ordered clustering.

Figure 3–3 demonstrates the performance of MC-NOMA with MC-OMA (where groups of users are served in different time slots) in terms of MMF rate and sum-rate with QoS, for different power allocated to beam $k$ denoted by $p_k$."

Figure 3–3a illustrates the achievable MMF rate in MC-NOMA compared to the MC-OMA scheme for different total power of the beam ($p_k$) and number of users per group ($M$) when ordered




(a) Comparison MMF rate of MC-NOMA with MC-OMA




(b) Comparison sum-rate with QoS of MC-NOMA with MC-OMA

Figure 3–3: Comparison the performance of MC-NOMA with MC-OMA versus different transmit power per beam.

clustering is used. The MC-NOMA achieves more gain than OMA with an increasing number of users. However, the gain of MC-NOMA over OMA decreases with an increasing power of beams. The reduction in the gain of MC-NOMA over OMA with increasing total power is negligible for a higher total number of users per group. Specifically, the gain of MC-NOMA over MC-OMA decreases by 15% with an increasing power of the beam from 1 dBW to 30 dBW when the number of users per group is 2.

Figure 3–3b presents the performance of MC-NOMA over OMA under different numbers of users for different powers of the beam. In this set of simulations, only ordered clustering is considered. The results indicate that the gain of MC-NOMA over OMA increases as the power and the number of users per group increase. Specifically, the gain increases by 40% as the power increases from 10 dBW to 30 dBW when $M = 10$.

## 3.6 Conclusion

In this chapter, we investigated the performance of the MC-NOMA scheme in the forward link of satellite communication with a 4-color frequency reuse pattern. We analyzed the performance of multicast NOMA in a single beam and showed that the attainable data rates are based on the minimum SINR in each group. To optimize the power allocation for different performance metrics,

we derived the optimum power allocation using existing methods. Additionally, we proposed an ordered clustering method in which users are ordered based on their SINR in each beam and then clustered into two groups. The ordered clustering method increases the SINR imbalance between groups of users, maximizing performance.

Our simulation results demonstrated that MC-NOMA outperforms MC-OMA in terms of minimum rate and sum-rate, especially when ordered clustering is used. Specifically, the minimum-rate and sum-rate of MC-NOMA can be increased by a factor of 2 and 1.45, respectively, with respect to MC-OMA. Moreover, we showed that utilizing user clustering can improve the performance of MC-NOMA by up to 30% compared to random clustering.

In conclusion, the proposed MC-NOMA scheme with ordered clustering and optimum power allocation can significantly improve the performance of satellite communication systems in terms of data rates and user fairness. Our study also highlights the importance of considering user clustering and NOMA together to make the most efficient use of satellite resources. Next chapter explores the application of MC-NOMA in the multibeam satellite communication systems.

# CHAPTER 4
# NOMA in multibeam multicast satellite systems with perfect CSIT: Optimizing User-scheduling and Linear precoding

## 4.1 Introduction

This chapter explores the use of full frequency reuse with the multibeam architecture in satellite communication systems to improve spectral efficiency. While 4-color frequency reuse was explored in the previous chapter, full frequency reuse offers even greater potential for improving spectral efficiency. However, implementing full frequency reuse in the multibeam architecture presents challenges due to interbeam interference in adjacent beams.

Fortunately, recent advancements in DVB-S2X [5] have led to the development of a multicast framework that is well-suited to the multibeam architecture. By embedding multiple users in a single frame, this framework allows for significant coding gain improvements. However, all users whose data are transmitted in the same multicast frame must accept a rate no higher than the achievable rate for the weakest user scheduled in the frame.

To address interbeam interference, interference management strategies are critical. Linear precoding has been identified as a critical element for interference cancellation for the next-generation multibeam satellite systems. This technique can be used to counteract interbeam interference generated by co-located beams employing the same frequency. However, implementing the precoding system will require hardware upgrades to both the ground and user segments. The ground equipment will include operations involved in the precoding procedure, such as precoding matrix computation and multiplication, channel state information feedback processing, and more. User equipment must perform new synchronization and channel feedback activities not included in non-precoding systems.

The hardware upgrade required for precoding presents an excellent opportunity to enhance the satellite system capacity with other promising techniques. One such technique is non-orthogonal multiple access (NOMA), which has been extensively investigated in a 4-color frequency reuse in previous chapter. The NOMA scheme can improve spectral efficiency by sending more than one multicasting frame simultaneously under certain conditions of power imbalance. NOMA offers a throughput increase compared to OMA techniques. In this chapter, we introduce the second framework proposed in this thesis, which combines linear precoding and the NOMA scheme in the context of multicast transmission. This framework is referred to as the multibeam multicast NOMA-based (MB-MC-NOMA) scheme.

In this chapter, we investigate the proposed MB-MC-NOMA scheme in the forward link of satellite communication systems. The user scheduling, multicast linear precoding (MC-linear precoding), and multicast NOMA (MC-NOMA) are adopted for the MB-MC-NOMA scheme. The proposed user scheduling for MB-MC-NOMA improves the performance of both MC-linear precoding and MC-NOMA scheme. The MC-linear precoding is designed to improve performance while keeping the complexity low. The proposed linear precoding is obtained from a unicast design by computing the composite channel matrix, which is a virtual channel that does not necessarily have a physical meaning. To this end, the users' channel vectors to be served in a given beam are mapped into a single vector to deal with the lack of spatial degrees of freedom. This chapter presents three different mappers; singular-value-decomposition (SVD), signal-to-noise-ratio (SNR), and averaging mappers. In addition, using the results from chapter 3, the MC-NOMA is adopted for the MB-MC-NOMA by optimizing the user grouping and power allocation in each beam.

## 4.2 System model

Consider a Ka-band multibeam multicast satellite communication system that tessellates the coverage area of $K$ beams in the forward link. A single geostationary orbit (GEO) satellite provides service for multiple single-antenna users, as shown in Figure 4–1. It is assumed that

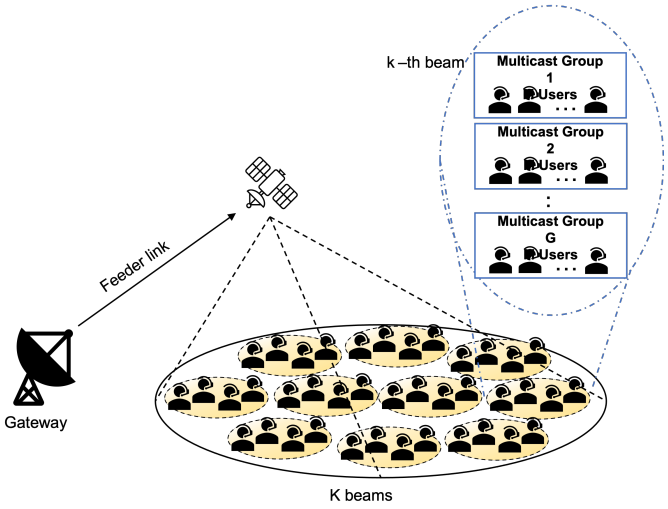



Figure 4–1: System model of the proposed MB-MC-NOMA scheme with $N_t$ antenna feeds over the coverage area of $K$ beams and each beam has $G$ multicast groups of $M$ users per time slot.

a single gateway with a noiseless feeder link between the gateway and satellite is employed. Moreover, $N_t$ denotes the number of antenna feeds. This thesis considers a satellite architecture such as Eutelsat Ka-Sat with a single feed per beam architecture (SFPB) [71]. Therefore, one feed is required to generate one beam (i.e., $N_t = K$). Since the multibeam satellite system is practically overloaded, each beam simultaneously serves more than one user. Let $\mathcal{U}_k$ and $\mathcal{U}_k^t$ denote the set of all users in beam $k$ and the set of users in beam $k$ in time slot $t = \{1, ..., T\}$, respectively. In the multicasting transmission, $|\mathcal{U}_k^t| > 1$ users are served in a single frame per beam per time slot. In this chapter, we consider that each beam contains more than one multicasting group of users per time slot to improve the system capacity and spectral efficiency. The groups of users are served with different multicasting frames and in the NOMA framework in which they share the same frequency and time resources. Let $G$ denote the number of multicasting groups of users. Therefore, $|\mathcal{U}_k^t|$ single antenna users are scheduled in $G > 1$ multicast groups of $M$ users per beam per time slot as shown in Figure 4–1. In this case, the transmitter with $N_t = K$ antennas serves $MG \times K$ single antenna users simultaneously, $N_t \ll MG \times K$. To group $MG$ users, $\mathcal{U}_k^t$ should be divided into $G$ groups of indices, $\mathcal{U}_{k,1}^t, ..., \mathcal{U}_{k,G}^t$. The user grouping has the following properties for

$\mathcal{K} \in \{1...K\}$:

$$|\mathcal{U}_{k,1}^t| = ... = |\mathcal{U}_{k,G}^t| = M,$$
$$\mathcal{U}_{k,1}^t \cap \mathcal{U}_{k,2}^t ... \cap \mathcal{U}_{k,G}^t = \emptyset,$$
$$\mathcal{U}_{k,1}^t \cup \mathcal{U}_{k,2}^t ... \cup \mathcal{U}_{k,G}^t = \mathcal{U}_k^t.$$

Furthermore, $\mathcal{U}_k^t$ is selected from a larger collection of users, $\mathcal{U}_k$, which specifies the set of all users indices that should be served. This user selection is referred to as user scheduling in this work, and it will be described in the following section. Without loss of generality, in this chapter we consider that each beam contains two groups of users, $G = 2$, group $\mathcal{A}$ and group $\mathcal{B}$. The extension to more than two groups is straightforward. It is shown in [26] that the performance gain in NOMA grows as the number of groups increases. For limited resources such as time and frequency (e.g. a multicasting frame size of $M$ users), the other techniques can only serve $M$ users. However, the NOMA scheme with the same resources can support $G$ times multicasting frames of size $M$ which means $G \times M$ users. Therefore, by increasing the number of groups of users, the NOMA scheme achieves higher spectral efficiency than other techniques.

### 4.2.1 Channel Model

This chapter focuses on the LMS channel in the Ka band [66, 72, 73]. The LMS model proposed in [66] is employed, which is based on the Loo probability density function [65]. The LMS model is traditionally used for mobile users in L-band. Recent advancements in mobile terminal antenna technology and the increasing demand for broadband services and higher data rates have made the utilization of the Ka-band applicable for mobile satellite services [73].

We operate under the assumption that the noise and interference within the feeder link between the satellite and the gateway are negligible. Furthermore, the channel is deemed to be constant throughout the frame transmission with respect to mobility. Therefore, the channel vector can be characterized as follows:

$$\mathbf{h}_{k,Y}^i = \mathbf{f}_{k,Y}^i \circ \overline{\mathbf{h}}_{k,Y}^i \tag{4.1}$$

where $\circ$ is the Hadamard product. $i \in \mathcal{U}^t_{k,Y}$, for $Y = \{\mathcal{A}, \mathcal{B}\}$ and $k = 1, ..., K$. Moreover, $f^i_{k,Y}$ and $\overline{\mathbf{h}}^i_{k,Y}$ model the fading effect and the channel vector, respectively. The fading effect obeys the Loo distribution and is defined as

$$f^i_{k,Y} = z^i_{k,Y} e^{j\theta^{i,\mathrm{LoS}}_{k,Y}} \; w^i_{k,Y} e^{j\theta^{i,\mathrm{MP}}_{k,Y}} \tag{4.2}$$

Let $z^i_{k,Y}$ denote the lognormally distributed line-of-sight (LoS) component, and $w^i_{k,Y}$ denote the Rayleigh distributed multipath (MP) component. Moreover, $\theta^{i,\mathrm{LoS}}_{k,Y}$ and $\theta^{i,\mathrm{MP}}_{k,Y}$ are phases of the LoS and MP which uniformly distributed between 0 and $2\pi$.

The channel vector $\overline{\mathbf{h}}^i_{k,Y}$ is given by

$$\overline{\mathbf{h}}^i_{k,Y} = \frac{\sqrt{G_R} r^i_{k,Y} a^i_1 e^{j\Phi^i_1}, ..., a^i_K e^{j\Phi^i_K}}{4\pi \frac{d^i_k}{\lambda} \sqrt{K_B T B_W}} \tag{4.3}$$

where $a^i_l$ is the gain from the $l$-th feed to the $i$-th user, $G_R$ is the reception antenna gain, and $d^i_k$ denotes the distance between the $i$-th user at beam $k$ and the satellite. $r^i_{k,Y}$ is the atmospheric fading parameter. Due to the working at high frequency band, such as Ka-band, the channel is significantly affected by atmospheric fading and mainly rain attenuation, which is defined as [74]

$$r^i_{k,Y} = {\xi^i_{k,Y}}^{-\frac{1}{2}} e^{\theta^i_{k,Y}} \tag{4.4}$$

where $\xi^i_{k,Y}$ denotes the power gain of rain attenuation which is lognormally distributed, i.e. $\ln\left(20\log_{10}\xi^i_{k,Y}\right) \sim \mathcal{N}\mu, \sigma^2$. Moreover, $\theta^i_{k,Y}$ is the uniform distributed phase.

Additionally, $\Phi^i_l$ denotes the time-varying phase caused by the beam radiation pattern and radio wave propagation. The phase value, $\Phi^i_l$, consists of

$$\Phi^i_l = \theta^i_{RF} \; \theta^i_{LNB} \; \theta_{PL,l} \tag{4.5}$$

where $\theta^i_{RF} = \frac{2\pi}{\lambda} d^i_k$ is the phase rotation due to the radiofrequency signal propagation and depends on the user distance to the satellite. $\theta^i_{LNB}$ is the phase contribution of the receiver low noise block downconverters assumed to be Gaussian with zero mean and standard deviation of 0.24 degrees and $\theta_{PL,l}$ which are the payload oscillator phase offsets which are assumed to be Gaussian with

zero mean and standard deviation that is usually around 2 degrees. Finally, the carrier wavelength, Boltzmann constant, receiver noise temperature, and carrier bandwidth are represented by $\lambda$, $K_B$, $T$, and $B_W$, respectively. The channel has been adjusted to the noise power. This chapter assumes perfect channel state information at the transmitter (CSIT) is available.

### 4.2.2 Signal Model

In the multibeam satellite communication system , each user receives the interfere signals from the other beams due to the full frequency reuse. Therefore, interference mitigation techniques shall be used to cancel the interbeam interference. The linear precoding techniques would be applied at the transmitter side to revert the interbeam interference [4]. To more improve the system capacity and the spectral efficiency, we consider that each beamforming vector conveys information for more than one group of users in each beam. Therefore, users receive inter-group interference per beam and the MC-NOMA scheme can be applied to revert the inter-group interference. According to the MC-NOMA scheme in the power domain, the transmitter sends the superposition of messages and strong users apply the SIC to cancel inter-group interference. The combination of the MC-NOMA with the linear precoding leads to MB-MC-NOMA.

According to the literature, there are two possibilities for merging linear precoding and NOMA: beamformer-based structure and cluster-based structure [75]. Each beamforming vector in the beamformer-based structure serves a single group of users, while each vector in the cluster-based structure serves multiple groups of users. Because of the lack of the spatial degree of freedom, we investigate a cluster-based structure where each beam has just one precoding vector instead of two. As a result, in the proposed MB-MC-NOMA method, the transmitted signal is

$$\mathbf{x} = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{w}_k\sqrt{\alpha_k}s_{k,\mathcal{A}}\ \sqrt{1-\alpha_k}s_{k,\mathcal{B}} = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{w}_k s_k, \tag{4.6}$$

where $p_k$ is the allocated power to beam $k$ and $\alpha_k$ controls the percentage of power allocated to each group of users in beam $k$. Symbols $s_{k,\mathcal{A}}$ and $s_{k,\mathcal{B}}$ convey transmitted symbols to the users in group $\mathcal{A}$ and group $\mathcal{B}$. The vector $\mathbf{w}_k$ precodes the users' symbols in beam $k$. The transmitted

signal is constrained to

$$\sum_{k=1}^{K} p_k ||\mathbf{w}_k||^2 \leq P_{\mathrm{T}} \tag{4.7}$$

where $P_{\mathrm{T}}$ is the maximum available power in the satellite payload. Since in this chapter we consider two groups of users, then the received signal by the $j$-th user in group $\mathcal{A}$ and the $l$-th user in group $\mathcal{B}$ are expressed as follows

$$y_{k,\mathcal{A}}^{j} = \mathbf{h}_{k,\mathcal{A}}^{j} \mathbf{w}_k \sqrt{p_k} \sqrt{\alpha_k} s_{k,\mathcal{A}} \ \sqrt{1-\alpha_k} s_{k,\mathcal{B}} \ \mathbf{h}_{k,\mathcal{A}}^{j} \sum_{n=1, n\neq k}^{K} \mathbf{w}_n \sqrt{p_n} s_n \ n_{k,\mathcal{A}}^{j}, \quad j \in \mathcal{U}_{k,\mathcal{A}}^{t} \tag{4.8}$$

$$y_{k,\mathcal{B}}^{l} = \mathbf{h}_{k,\mathcal{B}}^{l} \mathbf{w}_k \sqrt{p_k} \sqrt{\alpha_k} s_{k,\mathcal{A}} \ \sqrt{1-\alpha_k} s_{k,\mathcal{B}} \ \mathbf{h}_{k,\mathcal{B}}^{l} \sum_{n=1, n\neq k}^{K} \mathbf{w}_n \sqrt{p_n} s_n \ n_{k,\mathcal{B}}^{l}, \quad l \in \mathcal{U}_{k,\mathcal{B}}^{t} \tag{4.9}$$

In terms of notation, $\mathbf{h}_{k,\mathcal{A}}^{j} \in \mathbb{C}^{1\times K}$ and $\mathbf{h}_{k,\mathcal{B}}^{l} \in \mathbb{C}^{1\times K}$ denote the $j$-th and the $l$-th user's channel vector in groups $\mathcal{A}$ and group $\mathcal{B}$, respectively. Finally, $n_{k,\mathcal{A}}^{j}$ and $n_{k,\mathcal{B}}^{l}$ are the additive noise terms that contaminate the reception of users in each group. Without loss of generality, the noise terms $n_{k,\mathcal{B}}^{l}$ and $n_{k,\mathcal{A}}^{j}$ have the same distribution as $\mathcal{CN}0, 1$ and perfect CSIT is available at the transmitter.

Algorithm 0 explains the steps of designing the MB-MC-NOMA scheme. In this thesis, all variables $(g_k, \mathbf{w}_k, p_k, \alpha_k)$ are referred to $(g_k^t, \mathbf{w}_k^t, p_k^t, \alpha_k^t)$, where the superscript $t$ is dropped off for simplicity. In the following section, we thoroughly analyze each step.

**Algorithm 1** Steps of design

**Input**: $\mathcal{U}_k$
**Outputs**:$\mathcal{U}_k^t$,$\mathbf{w}_k$,$\mathcal{U}_{k,\mathcal{A}}^t$,$\mathcal{U}_{k,\mathcal{B}}^t$,$p_k$,$\alpha_k$
**for** $t = 1, 2, \ldots, T$ **do**
    **for** $k = 1, 2, \ldots, K$ **do**
        $\mathcal{U}_k^t \in \mathcal{U}_k \& |\mathcal{U}_k^t| = 2M$ (User scheduling)
        $\mathbf{g}_k^t \leftarrow \mathcal{U}_k^t$ (Mapping)
        $\mathbf{w}_k^t \leftarrow \mathbf{g}_k^t$ (MC-linear precoding)
        $\mathcal{U}_{k,\mathcal{A}}^t$,$\mathcal{U}_{k,\mathcal{B}}^t \in \mathcal{U}_k^t$ (User grouping)
        $p_k^t, \alpha_k^t \leftarrow \mathcal{U}_k^t, \mathbf{w}_k^t$ (Power allocation)
    **end for**
**end for**

## 4.3 MB-MC-NOMA

In this section, we explain the steps of algorithm 1, except the power allocation, which will be discussed in the next chapter.

### 4.3.1 User scheduling

This thesis considers that the transmitter has access to all users' channel coefficient estimations. To estimate the channel coefficients, one pilot (reference signal) per beam is needed because of the long codewords and multicast transmission in satellite systems. Moreover, it is considered there is no pilot contamination because of the high directivity of the beams. Each beam may receive interference only from its adjacent beams. Therefore, the pilot of the adjacent beams can be considered orthogonal (e.g., orthogonal reference signals or different measuring timings) [7]. Moreover, it is assumed users have equal priority. As a results, the user scheduling in the MB-MC-NOMA is selecting $\mathcal{U}_k^t$ out of $\mathcal{U}_k$ per beam per time slot, where $|\mathcal{U}_k^t| \ll |\mathcal{U}_k|$. The user scheduling in the MB-MC-NOMA scheme is not straightforward because the requirements for the optimum user scheduling in the MC-linear precoding and the MC-NOMA contradict each other.

The performance of the MC-linear precoding improves if the Euclidean distance between users' channel vectors in each beam decreases [54]. The Euclidean distance between the $i$-th and the $j$-th user in beam $k$ is calculated as

$$d_k^{ij} = \|\mathbf{h}_k^i - \mathbf{h}_k^j\|^2, \{i, j\} \in \mathcal{U}_k, \tag{4.10}$$

where the channel vector of the $i$-th user is defined as $\mathbf{h}_k^i = h_{k1}^i, h_{k2}^i, ..., h_{kK}^i$.

On the other hand, the MC-NOMA technique has been shown to improve the user data rate when groups of users to be serviced have a large SNR imbalance [24]. The SNR of the $i$-th user in beam $k$ is defined as

$$\mathrm{SNR}_k^i = \|h_{kk}^i\|^2. \tag{4.11}$$

As a result, the SNR imbalance in beam $k$ between the $i$-th and the $j$-th user is $\|h_{kk}^i\|^2 \|h_{kk}^j\|^2$. Therefore, these two criteria are incompatible.

We propose a low complexity scheduling method to deal with the trade-off between the Euclidean norm and the SNR imbalance. In the proposed method, the set of $\mathcal{U}_k$ is firstly divided into two subsets according to the users' SNR. The new subsets are labeled as the subsets of the strong and weak users. Next, $M$ users with the lowest Euclidean distance are selected from each subset. Searching for the $M$ users with the lowest Euclidean norm is not an exhaustive search. The first user in each subset is chosen randomly, and then $M-1$ users are added one by one. At the end, the indices of the selected users are gathered in the $\mathcal{U}_k^t$. Therefore, on each time slot, $\mathcal{U}_k^t$ contains two groups of users with the highest possible SNR imbalance, and each group has $M$ users with a high degree of co-linearity in their channel vectors. Algorithm 0 explains the proposed user scheduling procedure.

**Algorithm 2** User scheduling

**Input**: Users terminals ($\mathcal{U}_k$
**Outputs**; $\mathcal{U}_k^t$
**for** $k = 1, 2, \ldots, K$ **do**
    **for** $i = 1, 2, \ldots, 2M$ **do**
        $\mathrm{SNR}_k^i = \|h_{kk}^i\|^2$
    **end for**
    $\mathrm{SNR}_k = \mathrm{sortSNR}_k^i$
    Based on the resulting $\mathrm{SNR}_k$, divide $\mathcal{U}_k$ into two subsets: subset of strong users and subset of weak users
    Choose $M$ users with the lowest Euclidean distance from subset of strong users
    Choose $M$ users with the lowest Euclidean distance from subset of weak users
    Gather the indices of the $2M$ selected users in $\mathcal{U}_k^t$
**end for**

### 4.3.2 MC-linear precoding

The precoding matrix, $\mathbf{W} = \left[\mathbf{w}_1 \mathbf{w}_2 ... \mathbf{w}_K\right]$ in the unicast transmission is a function of the composite channel matrix, $\mathbf{H} = \mathbf{h}_1^T, ..., \mathbf{h}_K^T$, where $\mathbf{h}_k$ is the channel vector of a single user in beam $k$. Unlike the unicast transmission, designing the precoding matrix in the MC transmission is not straightforward.

One way to design a low-complexity linear precoding in the MC transmission is to mimic the linear precoding techniques in the unicast transmission [18]. However, constructing the channel

matrix is challenging because a matrix rather than a vector characterizes each beam. The matrix $\mathbf{C}_k = \left[\mathbf{h}_k^{1T}, ..., \mathbf{h}_k^{2MT}\right]$ gathers the channel vectors of all users in beam $k$. Therefore, function $f$ is needed to map the $\mathbf{C}_k$ into the vector $\mathbf{g}_k$ namely,

$$f : \mathbf{C}_k \longrightarrow \mathbf{g}_k. \tag{4.12}$$

Then the composite channel matrix becomes $\mathbf{G} = \left[\mathbf{g}_1^T \mathbf{g}_2^T ... \mathbf{g}_K^T\right]^T$ after performing the mapping on each beam. It's worth noting that this approach has low complexity, and the performance of the linear precoding depends on the mapper. Therefore, the most challenging part of this approach is searching for an optimum mapper.

Next, a linear precoding technique is used to produce the linear precoding matrix after building $\mathbf{G}$. In this thesis, we select MMSE and ZF techniques which are discussed in section (2.2.3) to generate the precoder matrix.

The complexity of the ZF and MMSE linear precoding in the multicast transmission for $N_t$ number of antenna elements and $GMK$ number of single-antenna users that are grouped into $GK$ multicast groups is equal to $\mathcal{O}MGN_tK^2\ N_t^3\ KN_t^2\ K^2N_t\ \frac{1}{2}MGKN_t^2$ [19]. In the following subsections, we present three mappers to construct the composite channel matrix and analyze the presented mappers' complexity.

**Mapping by averaging**

The averaging mapper is the most common approach for calculating the composite channel matrix in the MC satellite communication systems. In this method, the mapper finds the average of the user channel vectors per beam [4]. As a result, $\mathbf{g}_k$ is calculated as

$$\mathbf{g}_k = \frac{i \in \mathcal{U}_k^t \mathbf{h}_k^i}{GM}. \tag{4.13}$$

The number of operations to calculate the average of $GM$ users is $GM$ per beam. Therefore, the complexity of the mapping by averaging in $K$ beams is $\mathcal{O}GMK$.

**Mapping by SNR**

The proposed mapper in [70] influences us to present the mapping by SNR. The proposed mapper in [70] selects the strongest user in each cluster of the MIMO-NOMA system. The channel vectors of the selected users generate the beamforming vectors. The strongest user defines as the user with the largest channel gain in each cluster.

Using this result, we present the SNR mapper, which chooses the user with the highest SNR in each beam. The channel vector of the selected user constructs the $\mathbf{g}_k$. The SNR of $i$-th user in beam $k$ is given by (4.11).

To find the user with the highest SNR out of $GM$ users, $GM$ comparison operations are needed. Therefore, the complexity of the mapping by SNR in $K$ beams is $\mathcal{O}GMK$.

**Mapping by SVD**

We propose a mapper based on the singular value decomposition (SVD) approach. In the proposed mapper, first the SVD of the $\mathbf{C}_k^H\mathbf{C}_k$ is calculated

$$\mathbf{C}_k^H\mathbf{C}_k = \mathbf{U}_k\Sigma_k\mathbf{V}_k^H, \tag{4.14}$$

where $\Sigma_k$ is the singular value matrix and $\mathbf{U}_k$ ($\mathbf{V}_k$) gathers the left-singular vectors (right-singular vectors). Then the right or left singular vector associated with the highest singular value is selected. The selected singular vector in beam $k$ constructs the $\mathbf{g}_k$. We propose the SVD mapper aiming to maximize the energy spread over the users.

The complexity of the mapping by SVD for $GM$ users in $K$ beams is $\mathcal{O}G^2M^2K$. By comparing the complexity of the three presented mappers, we conclude that the mapping by averaging and the mapping by SNR have the same complexity of order $GMK$. However, the mapping by SVD has a higher complexity of order $G^2M^2K$. Moreover, the complexity of the linear precoding technique and the mapping function in the multicast transmission increases with the increasing number of beams and users and groups per beam. In the next section, we investigate the MC-NOMA scheme and derive the achievable data rates.

### 4.3.3 MC-NOMA scheme

Following the MC-NOMA approach, the weak and strong users perform SUD and SIC, where the weak and strong users are grouped as Group $\mathcal{A}$ and $\mathcal{B}$, respectively. The users' grouping is based on effective channel gains. The users effective channel gains are calculated and sorted in beam $k$, $|h_k^1 \mathbf{w}_k| \leq ... \leq |h_k^{2M} \mathbf{w}_k|$. Then the users' indices are divided into two groups and labeled as $\mathcal{U}_{k,\mathcal{A}}^t$ and $\mathcal{U}_{k,\mathcal{B}}^t$. Therefore, it is assumed that the effective channel gains of users have the following order:

$$\min_{j \in \mathcal{U}_{k,\mathcal{A}}^t} |\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2 < \min_{l \in \mathcal{U}_{k,\mathcal{B}}^t} |\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_k|^2. \tag{4.15}$$

Let's first define the minimum received signal-to-co channel interference plus noise in beam $k$ as

$$\Gamma_{k,\mathcal{A}} = \min_{j \in \mathcal{U}_{k,\mathcal{A}}^t} \frac{p_k |\mathbf{h}_{k,A}^j \mathbf{w}_k|^2}{1 \sum_{n=1, n \neq k}^{K} p_n |\mathbf{h}_{k,A}^j \mathbf{w}_n|^2} \tag{4.16}$$

$$\Gamma_{k,\mathcal{B}} = \min_{l \in \mathcal{U}_{k,\mathcal{B}}^t} \frac{p_k |\mathbf{h}_{k,B}^l \mathbf{w}_k|^2}{1 \sum_{n=1, n \neq k}^{K} p_n |\mathbf{h}_{k,B}^l \mathbf{w}_n|^2}, \tag{4.17}$$

and the minimum received signal-to-interference plus noise as

$$\hat{\text{SINR}}_{k,\mathcal{A}} = \min_{j \in \mathcal{U}_{k,\mathcal{A}}^t} \text{SINR}_{k,\mathcal{A}}^j = \min_{j \in \mathcal{U}_{k,\mathcal{A}}^t} \frac{\alpha_k p_k |\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2}{1 \; 1 - \alpha_k p_k |\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2 \sum_{n=1, n \neq k}^{K} p_n |\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_n|^2}, \tag{4.18}$$

$$\hat{\text{SINR}}_{k,\mathcal{B}} = \min_{l \in \mathcal{U}_{k,\mathcal{B}}^t} \text{SINR}_{k,\mathcal{B}}^l = \min_{l \in \mathcal{U}_{k,\mathcal{B}}^t} \frac{1 - \alpha_k p_k |\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_k|^2}{1 \sum_{n=1, n \neq k}^{K} p_n |\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_n|^2}. \tag{4.19}$$

The MB-MC-NOMA scheme in beam k is a degraded broadcast channel. Therefore, the capacity rates under the Gaussian signaling in beam $k$ are given in the equations (4.20) and (4.21),

where $\mathcal{C}x = \log_2 1\ x$:

$$R_{k,\mathcal{A}} \leq IX_{k,\mathcal{A}}; Y_{k,\mathcal{A}} = \mathcal{C}\left(\frac{\alpha_k \Gamma_{k,\mathcal{A}}}{1\ \left(1-\alpha_k\right)\Gamma_{k,\mathcal{A}}}\right) = \log_2\left(1\ \frac{\alpha_k \Gamma_{k,\mathcal{A}}}{1\ \left(1-\alpha_k\right)\Gamma_{k,\mathcal{A}}}\right) = \log_2\left(1\ \hat{\text{SINR}}_{k,\mathcal{A}}\right) \tag{4.20}$$

$$R_{k,\mathcal{B}} \leq IX_{k,\mathcal{B}}; Y_{k,\mathcal{B}}|X_{k,\mathcal{A}} = \mathcal{C}\left(\left(1-\alpha_k\right)\Gamma_{k,\mathcal{B}}\right) = \log_2\left(1\ \left(1-\alpha_k\right)\Gamma_{k,\mathcal{B}}\right) = \log_2\left(1\ \hat{\text{SINR}}_{k,\mathcal{B}}\right) \tag{4.21}$$

It is important to remark that the user grouping satisfies

$$\log_2\left(1\ \frac{\alpha_k \Gamma_{k,\mathcal{B}}}{1\ \left(1-\alpha_k\right)\Gamma_{k,\mathcal{B}}}\right) \geq R_{k,\mathcal{A}} \tag{4.22}$$

Expression (4.22) will be simplified as

$$\Gamma_{k,\mathcal{B}} \geq \Gamma_{k,\mathcal{A}} \tag{4.23}$$

Therefore, this chapter assumes that the condition (4.23) is satisfied. According to results of chapter 3, the user grouping is optimized if the imbalance between $\Gamma_{k,\mathcal{A}}$ and $\Gamma_{k,\mathcal{B}}$ is maximized. To maximize the imbalance between $\Gamma_{k,\mathcal{A}}$ and $\Gamma_{k,\mathcal{B}}$ the condition $\max\limits_{j \in \mathcal{U}_{k,\mathcal{A}}}^{t} \text{SINR}_{k,\mathcal{A}}^{j} \leq \min\limits_{l \in \mathcal{U}_{k,\mathcal{B}}}^{t} \text{SINR}_{k,\mathcal{B}}^{l}$ should be satisfied.

The next step in designing the MB-MC-NOMA is power allocation. In this chapter, we consider that $p_k$ is given and $\alpha_k$ is optimized to maximize the users' fairness. The optimum $\alpha_k$ in terms of fairness is given in equation (3.16). This value for $\alpha_k$ guarantees that $R_{k,\mathcal{A}} = R_{k,\mathcal{B}}, \forall k \in \mathcal{K}$.

It is worth mentioning the complexity of the receiver for the proposed MB-MC-NOMA is determined based on the complexity of the MMSE detector. The computational complexity of the MMSE detector is $\mathcal{O}G^2$ [76], where $G$ is the number of NOMA groups.

### 4.4 Simulation results

In this section, we evaluate the proposed MB-MC-NOMA scheme. The parameters of the simulation are given in Table 4–1. We have used the statistical information provided in [66] to model the LMS channel for the ka band and the intermediate shadowing. According to [5], the elevation angle in ka band is fixed at $30 - 35$ to calculate the statistical parameters. The

| Carrier Frequency | 20GHz |
|---|---|
| Orbit | GEO |
| $G_R$/T | 17.68 dB/K |
| user location distribution | uniform |
| Beam radiation pattern | Provided by ESA |
| EIRP/beam | 63dBW |
| $B_W$ | 500MHz |
| Number of beams | 7 |

Table 4–1: Simulation parameters.

radiation pattern is based on the ESA database, which can be modeled based on the parabolic antenna pattern using the Bessel function, where the gain depends on the angle of bore-sight [77]. Moreover, the database considers two basic traffic models: location-based and time-based traffic models. The location-based traffic model is based on geographical factors. The time-based traffic model represents the demand for communication at different times of the day. It assumes that users are distributed uniformly in the coverage area and have the same demand over the day. Therefore, both location-based and time-based traffic models have a uniform distribution over the coverage area and the time. In addition, the database considers the atmospheric fading and, more specifically, the rain attenuation for each user. The users' SINR suffers a scaling that depends on respective rain attenuation. As such, the loss in terms of minimum rate and sum rate, depending on the atmospheric statistics and rain fading statistics, is considered to be a constant value affecting all the scenarios similarly [78].

In this chapter, we aim to evaluate the performance of the proposed MB-MC-NOMA scheme and compare it with an orthogonal multiple access scheme in multibeam multicast (MB-MC-OMA) scheme. The MB-MC-NOMA scheme serves two groups of users with indices $\mathcal{U}^t_{k,\mathcal{A}}$ and $\mathcal{U}^t_{k,\mathcal{B}}$ in each time slot. As a benchmark, we consider a TDMA in MB-MC-OMA scheme that divides each time slot into two sub-time slots, serving $\mathcal{U}^t_{k,\mathcal{A}}$ and $\mathcal{U}^t_{k,\mathcal{B}}$ separately in an orthogonal fashion. This results in the computation of two different precoding matrices considering both user subsets. Remarkably, we use the same precoding technique for the OMA case.

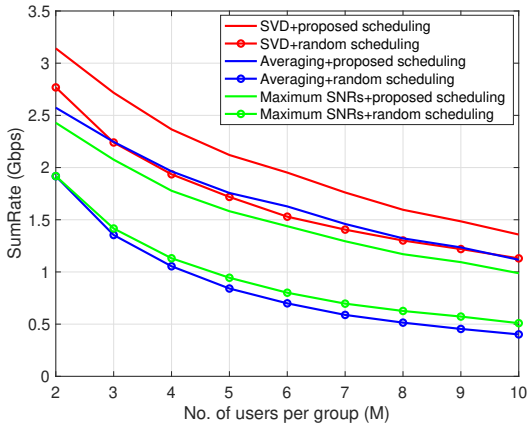



Figure 4–2: Performance of MB-MC-NOMA under different mappers and number of users per group

In this chapter, we consider only the MMSE precoding technique. MMSE precoding is a widely used technique for linear precoding that provides a good balance between complexity and performance. We use MMSE precoding to compare the performance of the MB-MC-NOMA scheme with the TDMA scheme. Our simulation results will provide insights into the advantages and limitations of both schemes.

Figure 4–2 displays the performance comparison of the MB-MC-NOMA scheme with various MC-linear precoding designs and user scheduling techniques, while considering a variable number of users per beam $M$. In the simulations, $\alpha_k$ is optimized to maximize the fairness per beam. The results demonstrate that for all precoding methods, the proposed scheduling technique outperforms the pure random user scheduling method. This difference is particularly significant for the maximum SNR mapping and average mapping.

Regarding precoding design, the proposed SVD-based approach yields the highest sum-rate values for all $M$ values considered in the simulation. Specifically, the SVD-based approach offers a sum-rate gain of at least 15% and up to 50% compared to the maximum SNR and average mapping methods for certain $M$ values. The results indicate that the proposed MB-MC-NOMA scheme outperforms the MB-MC-OMA schemes when SVD is used as the mapper with the proposed scheduling. The MB-MC-NOMA shows a 25% sum-rate gain over the MB-MC-OMA. However,

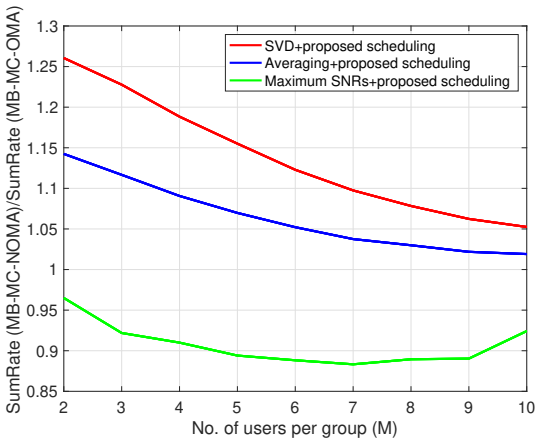



Figure 4–3: Comparison of performance of MB-MC-NOMA and MB-MC-OMA schemes under different mappers and number of users per group

this gain reduces as the maximum SNR is used as the mapper. Moreover, it does not gain over MB-MC-OMA when the averaging mapper is used.

Figure 4–3 compares the performance of two multiple access schemes, MB-MC-NOMA and MB-MC-OMA, is compared in terms of sum-rate ratio with maximum possible fairness per beam. The figure considers different precoding techniques and the proposed scheduling technique is used for both schemes. The optimized fairness factor $\alpha_k$ is used to maximize the fairness in each beam. The results show that the proposed MB-MC-NOMA scheme outperforms the MB-MC-OMA schemes when SVD is used as the mapper with the proposed scheduling. In fact, MB-MC-NOMA shows a 25% sum-rate gain over MB-MC-OMA. However, the gain reduces when the maximum SNR is used as the mapper, and MB-MC-NOMA does not gain over MB-MC-OMA when the averaging mapper is used. The proposed mapper and scheduling methods have been shown to improve the performance of the MB-MC-NOMA scheme significantly compared to the existing methods, such as the maximum SNR and channel averaging. Based on the proposed mapper and scheduling methods, the proposed MB-MC-NOMA scheme performs much better than the MB-MC-OMA scheme. However, this improvement comes at the cost of increased complexity at the receiver and the transmitter.

## 4.5 Conclusion

In conclusion, this chapter presented the MB-MC-NOMA framework for the forward link of multibeam satellite communication systems. The proposed scheme aims to optimize the performance of both MC-linear precoding and MC-NOMA by introducing a carefully designed user scheduling strategy. Additionally, three different mappers were proposed to address the lack of spatial degrees of freedom in MC transmission. User grouping is also optimized to improve the performance of the NOMA scheme.

The simulation results demonstrated that our proposed MB-MC-NOMA scheme with SVD mapper and the proposed user scheduling outperformes the MB-MC-OMA scheme in terms of both MMF rate and sum-rate with QoS. The proposed scheme provides better spectral efficiency and higher throughput for multibeam satellite systems.

Overall, the proposed MB-MC-NOMA scheme shows great potential for improving the performance of multibeam satellite systems, and we believe that it can significantly contribute to the development of next-generation satellite communication systems. Next chapter considers the power allocation optimizing in the proposed MB-MC-NOMA.

# CHAPTER 5
# NOMA in multibeam multicast satellite systems with perfect CSIT: Optimizing Power allocation and Rate region

## 5.1 Introduction

In the previous chapter, we introduced the second framework denoted as Multibeam Multicast Non-Orthogonal Multiple Access (MB-MC-NOMA) scheme in multibeam multicast satellite communication systems and explored its performance from user scheduling and MC-linear precoding perspectives. However, our analysis considered a fixed power allocation. In this chapter, we investigate further the power allocation problem for the proposed MB-MC-NOMA scheme with the goal of maximizing both the minimum fariness rate and the sum-rate of the system.

In the previous chapter, we compared the performance of the MB-MC-NOMA scheme with the Multibeam Multicast Orthogonal Multiple Access (MB-MC-OMA) scheme, which combines OMA schemes such as Time Division Multiple Access (TDMA) with linear precoding techniques. However, this comparison is not comprehensive, as various interference cancellation techniques have not yet been discussed. In this chapter, we will examine these additional interference cancellation methods to provide a more complete comparison between MB-MC-NOMA and other schemes.

Various techniques exist for managing interbeam interference in multibeam multicast systems, including linear precoding, NOMA, and rate-splitting (RS). This chapter compares the performance of the proposed MB-MC-NOMA framework with MC-linear precoding, MC-RS, and MB-MC-OMA schemes. Results show that despite the benefits of these techniques using only one technique to mitigate interbeam interference in multicast multibeam satellite systems is suboptimal. Combining two techniques provides more flexibility for optimizing the MMF rate and sum-rate.

This chapter formulates generic transmit power constraint optimization problems for achieving max-min fairness (MMF) and sum-rate maximization in the MB-MC-NOMA scheme. Power allocation to each antenna is optimized according to respective objective functions.

Additionally, this chapter proposes a method for solving the non-convex MMF optimization problem using auxiliary variables to transform the problem into a semi-definite programming problem, which is solved using linear program solvers. A method for solving the non-convex sum-rate maximization objective function in MB-MC-NOMA systems is also proposed based on forming the Lagrangian multipliers with respect to the constraints. The problem is reformulated by applying quadratic transforms on sum-of-ratios in an iterative sum-rate power optimization algorithm.

The achievable rate region of the MB-MC-NOMA scheme is derived using the proposed power optimization techniques. Due to the dependency of broadcasting power and respective capacity rates, per-beam and per-user optimal power allocation could not be investigated separately. Therefore, the equivalent channel and water-filling algorithm for the weighted sum-rate in MB-MC-NOMA are developed, and the optimal transmit power spectral density (PSD) for groups of users within multiple beams is efficiently computed.

In this chapter, we consider the system model, channel model, and signal model as explained in chapter 4.

## 5.2 Max-min fairness Analysis

In this section, we study the power allocation to maximize the minimum fairness (MMF) rate in MB-MC-NOMA scheme. The NOMA scheme enables a flexible management of the users achievable rates and provides an efficient way to enhance user fairness. In this section, we study the optimal power allocation to achieve MMF in the MB-MC-NOMA system. The problem is

defined as

$$\mathcal{P}_1 : \underset{p_k,\alpha_k}{\text{argmax}} \min_{k=1,\ldots,K} \{R_{k,\mathcal{A}}, R_{k,\mathcal{B}}\} \tag{5.1a}$$

s.t.

$$\alpha_k \in 0, 1, \quad \forall k \in \mathcal{K} \tag{5.1b}$$

$$\sum_{k=1}^{K} p_k \|\mathbf{w}_k\|^2 \leq P_{\mathrm{T}} \tag{5.1c}$$

$$p_k \geq 0, \quad \forall k \in \mathcal{K} \tag{5.1d}$$

where $R_{k,\mathcal{A}}$ ($R_{k,\mathcal{B}}$), the achievable rates by users in group $\mathcal{A}$ ($\mathcal{B}$), are given in equation 4.20 (4.21). The problem $\mathcal{P}_1$ is nonconvex. To solve this optimization problem, we need to restate (5.1), as the minimum operation in the objective function is not a convex function. First, we denote $\alpha_k p_k = p_{k,\mathcal{A}}$ and $1 - \alpha_k p_k = p_{k,\mathcal{B}}$ in the equations (4.20) and (4.21). Next, we add an auxiliary variable $R_{\mathrm{MMF}}$ and convert the problem $\mathcal{P}_1$ to a new constrained optimization problem as

$$\mathcal{P}_2 : \underset{p_{k,\mathcal{A}}, p_{k,\mathcal{B}}, R_{\mathrm{MMF}}}{\text{argmax}} \quad R_{\mathrm{MMF}} \tag{5.2a}$$

s.t.

$$R_{k,\mathcal{A}} \geq R_{\mathrm{MMF}}, R_{k,\mathcal{B}} \geq R_{\mathrm{MMF}}, \quad \forall k \in K \tag{5.2b}$$

$$p_{k,\mathcal{A}}\ p_{k,\mathcal{B}} = p_k, \quad \forall k \in K \tag{5.2c}$$

$$p_{k,\mathcal{A}} \geq 0, p_{k,\mathcal{B}} \geq 0, \quad \forall k \in K \tag{5.2d}$$

$$\sum_{k=1}^{K} p_k \|\mathbf{w}_k\|^2 \leq P_{\mathrm{T}} \tag{5.2e}$$

The problem $\mathcal{P}_2$ is still a non-convex optimization problem. In this section, we further modify the optimization problem $\mathcal{P}_2$ to obtain an equivalent semi-definite programming problem. We first

define an auxiliary variable $e = 2^{R_{\text{MMF}}} - 1$. Then, we find a convex approximation as

$$\mathcal{P}_3 : \max_{p_{k,\mathcal{A}}, p_{k,\mathcal{B}}, e} \quad \text{e} \tag{5.3a}$$

s.t.

$$\text{SINR}_{k,\mathcal{A}}^{j} \geq e, \forall k \in K, \forall j \in \mathcal{U}_{k,\mathcal{A}}^{t} \tag{5.3b}$$

$$\text{SINR}_{k,\mathcal{B}}^{l} \geq e, \forall k \in K, \forall l \in \mathcal{U}_{k,\mathcal{B}}^{t} \tag{5.3c}$$

$$5.2c, 5.2d, 5.2e \tag{5.3d}$$

where $\text{SINR}_{k,\mathcal{A}}^{j}$ and $\text{SINR}_{k,\mathcal{B}}^{l}$ are given in the equations (4.18) and (4.19). The problem $\mathcal{P}_3$ is convex. For fixed $e$, the problem $\mathcal{P}_3$ is a linear program (LP) and can be efficiently solved by several LP solvers such as the CVX toolbox. Therefore, for a fixed $e$ the optimum values for $p_{k,\mathcal{A}}$ and $p_{k,\mathcal{B}}$ which satisfy the constraints would be achieved. Then, one can exploit the bisection method to search the optimal e. The optimal power allocation for the maximum possible value of $e$ is the answer. It is well established that the MB-MC-NOMA scheme offers more flexibility other than the other techniques such as rate-splitting, linear precoding, and NOMA to increase the minimum rate. In the next section, we investigate the sum-rate maximization optimization.

### 5.3 Sum-rate maximization

In this section we study the power allocation to maximize the sum-rate in the MB-MC-NOMA system with or without the QoS constraint.

### 5.3.1 Weighted sum-rate maximization in the MB-MC-NOMA

In this section we study the weighted sum-rate maximization. The problem is defined as

$$\mathcal{S}_1 : \underset{p_{k,\mathcal{A}}, p_{k,\mathcal{B}}}{\operatorname{argmax}} \sum_{k=1}^{K} \{c_{k,\mathcal{A}} R_{k,\mathcal{A}} \; c_{k,\mathcal{B}} R_{k,\mathcal{B}}\} \tag{5.4a}$$

s.t.

$$p_{k,\mathcal{A}} \; p_{k,\mathcal{B}} = p_k, \forall k \in K \tag{5.4b}$$

$$p_{k,\mathcal{A}} \geq 0, p_{k,\mathcal{B}} \geq 0, \forall k \in K \tag{5.4c}$$

$$\sum_{k=1}^{K} p_k \|\mathbf{w}_k\|^2 \leq P_{\mathrm{T}} \tag{5.4d}$$

where $c_{k,\mathcal{A}}, p_{k,\mathcal{A}}$ and $c_{k,\mathcal{B}}, p_{k,\mathcal{B}}$ are the weight of users and the power allocated to users of groups $\mathcal{A}$ and $\mathcal{B}$ in beam $k$, respectively. $R_{k,\mathcal{A}}$ and $R_{k,\mathcal{B}}$ are given in the equations (4.20) and (4.21). The optimization problem $\mathcal{S}_1$ is a nonconvex problem.

In [79], a novel quadratic transform is proposed to solve the weighted sum-rate maximization for the MIMO systems. We use the proposed method as a base and adapt it to our method, MB-MC-NOMA. First, we introduce new variables $\Phi_{k,\mathcal{A}}$ and $\Phi_{k,\mathcal{B}}$ denoting the minimum SINR of groups $\mathcal{A}$ and $\mathcal{B}$, respectively. The problem can be rewritten as

$$\mathcal{S}_2 : \underset{p_{k,\mathcal{A}}, p_{k,\mathcal{B}}}{\operatorname{argmax}} \sum_{k=1}^{K} c_{k,\mathcal{A}} \log_2 1 \; \Phi_{k,\mathcal{A}} \; c_{k,\mathcal{B}} \log_2 1 \; \Phi_{k,\mathcal{B}} \tag{5.5a}$$

s.t.

$$5.4b, 5.4c, 5.4d \tag{5.5b}$$

$$\Phi_{k,\mathcal{A}} = \min_{j \in \mathcal{U}^t_{k,\mathcal{A}}} \frac{p_{k,\mathcal{A}} |\mathbf{h}^j_{k,\mathcal{A}} \mathbf{w}_k|^2}{1 \; p_{k,\mathcal{B}} |\mathbf{h}^j_{k,\mathcal{A}} \mathbf{w}_k|^2 \sum_{n=1, n\neq k}^{K} p_n |\mathbf{h}^j_{k,\mathcal{A}} \mathbf{w}_n|^2} \tag{5.5c}$$

$$\Phi_{k,\mathcal{B}} = \min_{l \in \mathcal{U}^t_{k,\mathcal{B}}} \frac{p_{k,\mathcal{B}} |\mathbf{h}^l_{k,\mathcal{B}} \mathbf{w}_k|^2}{1 \sum_{n=1, n\neq k}^{K} p_n |\mathbf{h}^l_{k,\mathcal{B}} \mathbf{w}_n|^2} \tag{5.5d}$$

Then, we form the Lagrangian function with respect to the constraints (5.5c) and (5.5d) as

$$\begin{aligned} Lp,\Phi,\lambda \quad = \quad & \sum_{k=1}^{K} c_{k,\mathcal{A}} \log_2 1 \ \Phi_{k,\mathcal{A}} - \lambda_{k,A}\Phi_{k,\mathcal{A}} - \min_{j\in\mathcal{U}_{k,\mathcal{A}}^t} \frac{p_{k,\mathcal{A}}|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2}{1 \ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2 \ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_n|^2} \\ & \sum_{k=1}^{K} c_{k,\mathcal{B}} \log_2 1 \ \Phi_{k,\mathcal{B}} - \lambda_{k,B}\Phi_{k,\mathcal{B}} - \min_{l\in\mathcal{U}_{k,\mathcal{B}}^t} \frac{p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_k|^2}{1 \ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_n|^2}, \end{aligned} \tag{5.6}$$

where $\lambda_{k,A}$ and $\lambda_{k,B}$ are the dual variables. Since $\partial L \partial \Phi_{k,\mathcal{A}} = 0$ and $\partial L \partial \Phi_{k,\mathcal{B}} = 0$ at the optimum, we have

$$\Phi_{k,\mathcal{A}} = \frac{c_{k,\mathcal{A}}}{\lambda_{k,A}} - 1, \Phi_{k,\mathcal{B}} = \frac{c_{k,\mathcal{B}}}{\lambda_{k,B}} - 1 \tag{5.7}$$

Combining the above equations with (4.18) and (4.19), we arrive at a relationship between the optimal dual variable $\lambda_{k,A}(\lambda_{k,B})$ and the vector $\mathbf{p} = \{p_{1,\mathcal{A}}, p_{1,\mathcal{B}}, ..., p_{K,\mathcal{A}}, p_{K,\mathcal{B}}\}$:

$$\lambda_{k,\mathcal{A}} = \max_{j\in\mathcal{U}_{k,\mathcal{A}}^t} \frac{c_{k,\mathcal{A}} \left(1 \ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2 \ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_n|^2\right)}{1 \ \sum_{n=1}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_n|^2} \tag{5.8}$$

and

$$\lambda_{k,\mathcal{B}} = \max_{l\in\mathcal{U}_{k,\mathcal{B}}^t} \frac{c_{k,\mathcal{B}} \left(1 \ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_n|^2\right)}{1 \ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_n|^2 \ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_k|^2} \tag{5.9}$$

Now, the original optimization problem $(\mathcal{S}_1)$ can be thought of as an optimization of the Lagrangian (5.6) with appropriate $\lambda$. But the optimal $\lambda_{k,\mathcal{A}}$ and $\lambda_{k,\mathcal{B}}$ are related to $\mathbf{p}$ through (5.8) and (5.9). We can substitute the above optimal $\lambda_{k,\mathcal{A}}$ and $\lambda_{k,\mathcal{B}}$ to arrive at a new form of the objective function, denoted as $f_q\mathbf{p},\Phi$, is given in equation (5.10).

$$\begin{aligned} f_q\mathbf{p},\Phi = & \sum_{k=1}^{K} c_{k,\mathcal{A}} \log_2 1 \ \Phi_{k,\mathcal{A}} - c_{k,\mathcal{A}}\Phi_{k,\mathcal{A}} \min_{j\in\mathcal{U}_{k,\mathcal{A}}^t} \frac{c_{k,\mathcal{A}} 1 \ \Phi_{k,\mathcal{A}} p_{k,\mathcal{A}}|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_k|^2}{1 \ \sum_{n=1}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j \mathbf{w}_n|^2} \\ & \sum_{k=1}^{K} c_{k,\mathcal{B}} \log_2 1 \ \Phi_{k,\mathcal{B}} - c_{k,\mathcal{B}}\Phi_{k,\mathcal{B}} \min_{l\in\mathcal{U}_{k,\mathcal{B}}^t} \frac{c_{k,\mathcal{B}} 1 \ \Phi_{k,\mathcal{B}} p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_k|^2}{1 \ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_n|^2 \ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^l \mathbf{w}_k|^2} \end{aligned} \tag{5.10}$$

In essence, the original problem ($\mathcal{S}_1$) is now equivalently reformulated as

$$\mathcal{S}_3 : \underset{\mathbf{p},\Phi}{\operatorname{argmax}} f_q\mathbf{p},\Phi \tag{5.11a}$$

$$\text{s.t.} \quad 5.4b, 5.4c, 5.4d \tag{5.11b}$$

Now, the variable $\mathbf{p}$ is outside of logarithm function. However, the optimization problem $\mathcal{S}_3$ is still a nonconvex problem. Two terms in $f_q\mathbf{p},\Phi$ take the form of a sum-of-ratios programming problem. The quadratic transform can be applied on the sum-of-ratios because both numerators and denominators are positive for any values of $\mathbf{p}$ [79]. The optimization problem for the ratio

$$\max_x \quad {}_k f_k \frac{Ax}{Bx} \tag{5.12a}$$

$$\text{s.t.} \quad x \in \mathcal{X} \tag{5.12b}$$

is equivalent to

$$\max_x \quad {}_k f_k 2y\sqrt{Ax} - y^2 Bx \tag{5.13a}$$

$$\text{s.t.} \quad x \in \mathcal{X} \tag{5.13b}$$

Using the quadratic transform,, the optimization problem $\mathcal{S}_3$ is transferred into problem $\mathcal{S}_4$ in (5.14a).

$$\mathcal{S}_4 : \underset{\mathbf{p},\Phi,\mathbf{y}}{\operatorname{argmax}} \qquad f_r\mathbf{p},\Phi,\mathbf{y} \tag{5.14a}$$

s.t.

$$5.4b, 5.4c, 5.4d \tag{5.14b}$$

where $f_r\mathbf{p},\Phi,\mathbf{y}$ is

$$f_r\mathbf{p},\Phi,\mathbf{y} = \sum_{k=1}^{K} c_{k,\mathcal{A}}\log_2 1\ \Phi_{k,\mathcal{A}} - c_{k,\mathcal{A}}\Phi_{k,\mathcal{A}}\ c_{k,\mathcal{B}}\log_2 1\ \Phi_{k,\mathcal{B}} - c_{k,\mathcal{B}}\Phi_{k,\mathcal{B}}$$
$$\sum_{k=1}^{K}\min_{j\in\mathcal{U}_{k,\mathcal{A}}^t}\left(2y_{k,A}^j\sqrt{c_{k,\mathcal{A}}1\ \Phi_{k,\mathcal{A}}p_{k,\mathcal{A}}|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_k|^2} - {y_{k,A}^j}^2 1\ \sum_{n=1}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_n|^2\right)$$
$$\sum_{k=1}^{K}\min_{l\in\mathcal{U}_{k,\mathcal{B}}^t}\left(2y_{k,\mathcal{B}}^l\sqrt{c_{k,\mathcal{B}}1\ \Phi_{k,\mathcal{B}}p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^j\mathbf{w}_k|^2} - {y_{k,B}^l}^2 1\ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^j\mathbf{w}_n|^2\ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^j\mathbf{w}_k|^2\right) \tag{5.15}$$

$f_r p,\Phi,y$ would be maximized over variables $y_{k,\mathcal{A}},\Phi_{k,\mathcal{A}},y_{k,\mathcal{B}},\Phi_{k,\mathcal{B}}$ and $p_{k,\mathcal{A}},p_{k,\mathcal{B}}$ in an iterative manner. When all the other variables are fixed, the optimal $y_{k,\mathcal{A}}(y_{k,\mathcal{B}})$ can be obtained by setting $\partial f_r\partial y_{k,\mathcal{A}}$ ($\partial f_r\partial y_{k,\mathcal{B}}$) to zero, i.e.,

$${y_{k,A}^j}^* = \frac{\sqrt{c_{k,\mathcal{A}}1\ \Phi_{k,\mathcal{A}}p_{k,\mathcal{A}}|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_k|^2}}{1\ \sum_{n=1}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_n|^2} \tag{5.16}$$

and

$${y_{k,B}^l}^* = \frac{\sqrt{c_{k,\mathcal{B}}1\ \Phi_{k,\mathcal{B}}p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^j\mathbf{w}_k|^2}}{1\ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^j\mathbf{w}_n|^2\ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^j\mathbf{w}_k|^2} \tag{5.17}$$

We substitute the above optimal y expression in $\partial f_r\partial\Phi$, then find the optimal $\Phi$ by setting $\partial f_r\partial\Phi_{k,A}$ and $\partial f_r\partial\Phi_{k,B}$ to zero to get

$$\Phi_{k,\mathcal{A}}^* = \min_{j\in\mathcal{U}_{k,\mathcal{A}}^t}\frac{p_{k,\mathcal{A}}|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_k|^2}{1\ p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_k|^2\ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^j\mathbf{w}_n|^2} \tag{5.18}$$

and

$$\Phi_{k,\mathcal{B}}^* = \min_{l\in\mathcal{U}_{k,\mathcal{B}}^t}\frac{p_{k,\mathcal{B}}|\mathbf{h}_{k,\mathcal{B}}^l\mathbf{w}_k|^2}{1\ \sum_{n=1,n\neq k}^{K} p_n|\mathbf{h}_{k,\mathcal{B}}^l\mathbf{w}_n|^2} \tag{5.19}$$

By fixing $y_{k,\mathcal{A}},\Phi_{k,\mathcal{A}},y_{k,\mathcal{B}},\Phi_{k,\mathcal{B}}$, the next step is optimizing the power allocation, $p_{k,\mathcal{A}}$ and $p_{k,\mathcal{B}}$ by setting $\partial f_r\partial p_{k,\mathcal{A}}$ and $\partial f_r\partial p_{k,\mathcal{B}}$ to zero. In the next iteration, $y_{k,\mathcal{A}},\Phi_{k,\mathcal{A}},y_{k,\mathcal{B}},\Phi_{k,\mathcal{B}}$ would be

updated using the optimized power from the previous iteration. The iteration continues until it converges. The iterative Algorithm 0 shows the steps of solving the optimum power allocation.

**Algorithm 3** Power allocation to maximize sum-rate in MB-MC-NOMA

1: **Input**: Users terminals ($\mathcal{U}_k$
2: **Outputs**; $p_{k,\mathcal{A}}, p_{k,\mathcal{B}}$
3: **Initialization**: Initialize $p$
4: **Repeat**
5: 1- Update $\Phi_{k,\mathcal{A}}$ and $\Phi_{k,\mathcal{B}}$ by equations (5.18) and (5.19)
6: 2- Update $y^j_{k,\mathcal{A}}$ and $y^l_{k,\mathcal{B}}$ by equations (5.16) and (5.17)
7: 3- Update $p$ by Solving the optimization problem $\mathcal{S}_4$
8: **Convergence**

Note: for $c_{k,\mathcal{A}} = c_{k,\mathcal{B}}$, the sum-rate would be maximized by allocating the zero power to the low-profile users in each beam. In addition, it is possible that the power allocated to a beam would be zero.

### 5.3.2 Weighted sum-rate maximization with QoS in the MB-MC-NOMA

This section contains the study of the sum-rate maximization with QoS constraint. The QoS constraint guarantees to provide the QoS even for weak users in all beams. The problem is defined as

$$\mathcal{S}_5 : \underset{p_{k,\mathcal{A}}, p_{k,\mathcal{B}}}{\operatorname{argmax}} \sum_{k=1}^{K} \{c_{k,\mathcal{A}} R_{k,\mathcal{A}}\ c_{k,\mathcal{B}} R_{k,\mathcal{B}}\} \tag{5.20a}$$

s.t.

$$\sum_{k=1}^{K} p_{k,\mathcal{A}}\ p_{k,\mathcal{B}} \|\mathbf{w}_k\|^2 \leq P_{\mathrm{T}}$$

$$R_{k,\mathcal{A}} \geq R^{\mathrm{OMA}}_{k,\mathcal{A}} \tag{5.20b}$$

The QoS constraint (5.20b) implies that

$$p_{k,\mathcal{A}} \geq \frac{2^{R_{k,\mathcal{A}}^{\mathrm{OMA}}}-1}{2^{R_{k,\mathcal{A}}^{\mathrm{OMA}}} \min\limits_{j\in\mathcal{U}_{k,\mathcal{A}}^{t}} \frac{|\mathbf{h}_{k,\mathcal{A}}^{j}\mathbf{w}_k|^2}{1+\sum\limits_{n=1}^{K} p_n|\mathbf{h}_{k,\mathcal{A}}^{j}\mathbf{w}_n|^2}}. \tag{5.21}$$

This problem is nonconvex and is a hard problem. One way to solve this problem is assuming the power budget for each beam $p_k$ is given. Therefore, $p_{k,\mathcal{A}}$ and $p_{k,\mathcal{B}}$ can be optimized for each beam separately. The power can be equally allocated to each beam. However, this power allocation may not maximize the sum-rate.

In this chapter, we proposed to use the power allocation among each beam using the weighted sum-rate maximization without constraint and the power allocation results of algorithm 3. Therefore, the power allocated to each beam first maximizes the sum-rate. Then, the optimized $p_k$ is used to satisfy the constraint (5.20b) or the equation (5.21).

## 5.4 Achievable rate region

In this section, we investigate the achievable rate region of the MB-MC-NOMA scheme in the forward link of the satellite communication systems. First, we determine the achievable rate region of the k-th beam in Theorem VI.1. The result of Theorem VI.1 can be extended to $K$ multibeam. The channel corresponding to the proposed MB-MC-NOMA in a beam, e.g., beam $k$, is a broadcasting channel that is shown as an example in Figure 5–1. According to the definition of broadcasting channels with one transmitter and multiple receivers, Figure 5–1 shows the channel with one transmitter and two groups of receiving users as an instance. The transmitter is considered to send signal $X$, that is received at the receivers after going through respective channels, e.g., with channel transfer functions $G_{k,\mathcal{A}}$ and $G_{k,\mathcal{B}}$. The noise is added separately, e.g., $n_{k,A}$ and $n_{k,B}$, where the added noise implies the noise experienced by the users with the highest noise levels in each group of users, i.e., worst channel conditions. Therefore, the achievable rate region in beam $k$ is defined as

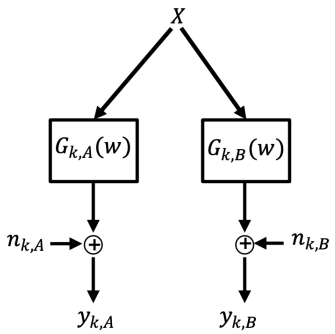



Figure 5–1: Single beam $k$ of the MB-MC-NOMA scheme

**Theorem 5.4.1.** *The achievable rate region of the Gaussian MB-MC-NOMA for two-group of users in beam k is*

$$\mathcal{C}_k = \left\{ \left(R_{k,\mathcal{A}}, R_{k,\mathcal{B}}\right) \in IR^2 : \beta_k R_{k,\mathcal{A}} \left(1-\beta_k\right) R_{k,\mathcal{B}} \leq \mathcal{C}_k \beta_k, \forall \beta_k \in 0,1 \right\} \tag{5.22}$$

$R_{k,\mathcal{A}}$ and $R_{k,\mathcal{B}}$ are achievable data rates of users in group $\mathcal{A}$ and $\mathcal{B}$ which are given in the equations (4.20) and (4.21), respectively. For a given $\beta_k$, $\mathcal{C}_k\beta_k$ is the maximum achievable rate. The boundary of the achievable rate region makes up the points that fulfill the equality in (5.22). The value of $\beta_k$ might be considered a priority factor for the two groups of users. If $\beta_k > 0.5$, group $\mathcal{A}$ has higher priority and vice versa.

*Proof.* The NOMA scheme in the forward link is, by nature, a broadcast channel. Moreover, there is a dependency between users' powers and one power constraint at the transmitter in the MB-MC-NOMA scheme. We use the duality between Gaussian multiple access channels (MACs) and Gaussian broadcast channels (BCs), which is proved in [80] with the condition that the dual channels should have the same channel gains and the same noise power at all receivers. We use the duality and follow the proof as in [55] that is based on the Kuhn-Tucker theorem to the achievable rate region defined in (4.20) and (4.21). □

To achieve the boundary of the achievable rate region, we need to find the optimal power spectral density (PSD) for two-group of users. The PSD of two groups of users cannot be found

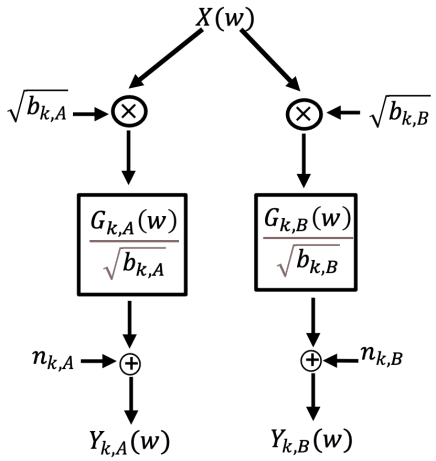



Figure 5–2: Equivalent channel idea of two-group of users in MB-MC-NOMA

separately in two water-filling diagrams because of dependency of the powers, $p_{k,\mathcal{A}}\ p_{k,\mathcal{B}} = p_k$. Thus, two water-filling diagrams cannot be combined because of the different water levels.

We use the idea of equivalent channel [55, 56] to combine two diagrams appropriately to maintain a single water level for all groups of users. The equivalent channel is shown in Figure5–2. The equivalent idea scales the channels to maintain the same water level. In Figure5–2, the scaling parameters are defined as $b_{k,\mathcal{A}}$ and $b_{k,\mathcal{B}}$. The capacities of the equivalent channel and the original channel are naturally identical. In contrast, the equivalent channel's optimum PSD is a scaled version of the original channel's optimal PSD. Figure 5–3 shows the water-filling scheme for two-group of users with different priorities. The water-filling diagrams overlap at some points.

The optimum PSDs of two groups of users in beam $k$ are defined as $S_{k,\mathcal{A}}w = \frac{\hat{S}_{k,\mathcal{A}}W}{b_{k,\mathcal{A}}}$ and $S_{k,\mathcal{A}}w = \frac{\hat{S}_{k,\mathcal{B}}W}{b_{k,\mathcal{B}}}$, where $\hat{S}_{k,\mathcal{A}}W$ and $\hat{S}_{k,\mathcal{B}}W$ are the PSDs of the users in the equivalent channels and

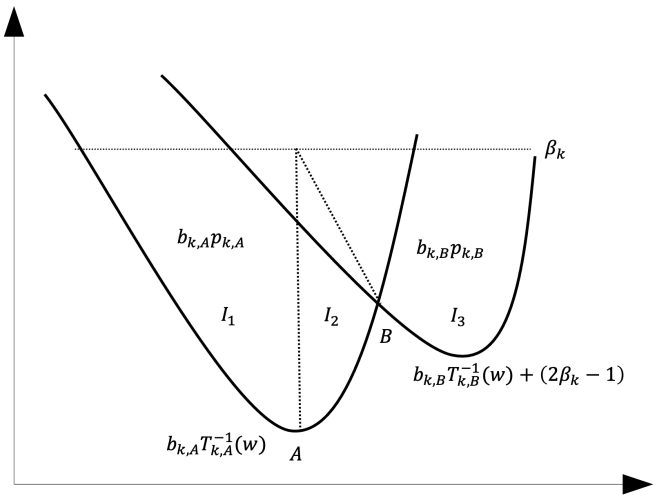



Figure 5–3: Water-filling diagram for two-group of users with different priorities

$b_{k,\mathcal{A}}$ and $b_{k,\mathcal{B}}$ are the scaling factors. The $\hat{S}_{k,\mathcal{A}}W$ and $\hat{S}_{k,\mathcal{B}}W$, if $\beta_k \in 0.5, 1$ are [55]

$$\hat{S}_{k,\mathcal{A}}w = \begin{cases} \beta_k - b_{k,\mathcal{A}}T_{k,\mathcal{A}}^{-1}w & \text{if} w \in I_1 \\ \beta_K - b_{k,\mathcal{A}}T_{k,\mathcal{A}}^{-1}w - \hat{S}_{k,\mathcal{B}}w & \text{if} w \in I_2 \\ 0 & \text{if} w \in I_3 \end{cases}$$

$$\hat{S}_{k,\mathcal{B}}w = \begin{cases} 0 & \text{if} w \in I_1 \\ 1 - \beta_k - b_{k,\mathcal{B}}T_{k,\mathcal{B}}^{-1}w.\gamma w & \text{if} w \in I_2 \\ 1 - \beta_k - b_{k,\mathcal{B}}T_{k,\mathcal{B}}^{-1}w & \text{if} w \in I_3 \end{cases} \tag{5.23}$$

where $\gamma w = \frac{2\beta_k - 1}{b_{k,\mathcal{A}}T_{k,\mathcal{A}}^{-1}w - b_{k,\mathcal{B}}T_{k,\mathcal{B}}^{-1}w}$, $T_{k,\mathcal{A}}w$ and $T_{k,\mathcal{B}}w$ are magnitude square of the channel transfer functions over the interference plus noise PSD. In addition, the individual power constraints of each group of users are:

$$\frac{1}{\pi}\int_0^{\pi} \hat{S}_{k,\mathcal{A}}w dw = b_{k,\mathcal{A}}p_{k,\mathcal{A}} \tag{5.24}$$

$$\frac{1}{\pi}\int_0^{\pi} \hat{S}_{k,\mathcal{B}}w dw = b_{k,\mathcal{B}}p_{k,\mathcal{B}} \tag{5.25}$$

$$b_{k,\mathcal{A}}p_{k,\mathcal{A}}\ b_{k,\mathcal{B}}p_{k,\mathcal{B}} = p_k. \tag{5.26}$$

The water-filling diagram and the corresponding equations can be extended to more than one beam. The achievable rate region for total $K$ beams can be achieved by maximizing $R_{k,\mathcal{A}}$ and $R_{k,\mathcal{B}}$. The corresponding achievable rate can be written as

$$\mathcal{C} = \{ \bigcup_{k=1}^{K} \left(\beta_k R_{k,\mathcal{A}} \ \left(1-\beta_k\right) R_{k,\mathcal{B}}\right) \leq \mathcal{C}\beta_1 ... \beta_k ... \beta_K, \forall \beta_k \in 0,1\}.$$

Variable $\beta_k$ is a priority factor for the two groups of users in beam $k$. We have studied the PSD optimization for two beams, which can be extended to more than two beams.

To find the optimal PSD of the MB-MC-NOMA over $K$ beams, we consider $K = 2$, beam $k$ and beam $j$, and $\beta_k \in 0.5, 1$ and $\beta_j \in 0.5, 1$. Beams $k$ and $j$ have three disjoint frequency bands $\{I_1, I_2, I_3\}$ and $\{I_4, I_5, I_6\}$, respectively. However, there is overlap between these two set of frequency bands.

Next, we assume that $b_{j,\mathcal{A}} T_{j,\mathcal{A}}^{-1} w < b_{k,\mathcal{A}} T_{k,\mathcal{A}}^{-1} w$ and $b_{j,\mathcal{B}} T_{j,\mathcal{B}}^{-1} w < b_{k,\mathcal{B}} T_{k,\mathcal{B}}^{-1} w$. Then, the PSD of two beams are given in equations (5.27), (5.28), (5.29), (5.30). Where $b_{k,\mathcal{A}}$ ($b_{j,\mathcal{A}}$) and $b_{k,\mathcal{B}}$ ($b_{j,\mathcal{B}}$) are the scaling factor of users in group $\mathcal{A}$ and $\mathcal{B}$ in beam $k$ (beam $j$), respectively. $T_{k,\mathcal{A}} w$ ($T_{j,\mathcal{A}} w$) and $T_{k,\mathcal{B}} w$ ($T_{j,\mathcal{B}} w$) are magnitude square of the channel transfer functions over the interference plus noise PSD of users in groups $\mathcal{A}$ and $\mathcal{B}$ in beam $k$ (beam $j$), respectively.

$$\hat{S}_{k,\mathcal{A}} w = \begin{cases} \beta_K - b_{k,\mathcal{A}} T_{k,\mathcal{A}}^{-1} w - \hat{S}_{j,\mathcal{A}} w & \text{if } \ w \in I_{k,1}, I_{j,1} \| w \in I_{k,1}, I_{j,2} \\ \beta_K - b_{k,\mathcal{A}} T_{k,\mathcal{A}}^{-1} w - \hat{S}_{k,\mathcal{B}} w - \hat{S}_{j,\mathcal{A}} w & \text{if } \ w \in I_{k,2}, I_{j,1} \| w \in I_{k,2}, I_{j,2} \\ \beta_K - b_{k,\mathcal{A}} T_{k,\mathcal{A}}^{-1} w - \hat{S}_{j,\mathcal{B}} w & \text{if } \ w \in I_{k,1}, I_{j,3} \\ \beta_K - b_{k,\mathcal{A}} T_{k,\mathcal{A}}^{-1} w - \hat{S}_{k,\mathcal{B}} w - \hat{S}_{j,\mathcal{B}} w & \text{if } \ w \in I_{k,2}, I_{j,3} \\ 0 & \text{if } \ w \in I_{k,3}, I_{j,1} \| w \in I_{k,3}, I_{j,2} \| w \in I_{k,3}, I_{j,3} \end{cases} \tag{5.27}$$

$$\hat{S}_{k,\mathcal{B}}w = \begin{cases} 0 & \text{if } w \in I_{k,1}, I_{j,1} \| w \in I_{k,1}, I_{j,2} \| w \in I_{k,1}, I_{j,3} \\ 1-\beta_K - b_{k,\mathcal{B}} T_{k,\mathcal{B}}^{-1} w . \frac{2\beta_k - 1}{b_{k,\mathcal{A}} T_{k,\mathcal{A}}^{-1} w - b_{k,\mathcal{B}} T_{k,\mathcal{B}}^{-1} w} & \text{if } w \in I_{k,2}, I_{j,1} \| w \in I_{k,2}, I_{j,2} \| w \in I_{k,2}, I_{j,3} \\ 1-\beta_K - b_{k,\mathcal{B}} T_{k,\mathcal{B}}^{-1} w - \hat{S}_{j,\mathcal{A}} w & \text{if } w \in I_{k,3}, I_{j,1} \| w \in I_{k,3}, I_{j,2} \\ 1-\beta_K - b_{k,\mathcal{B}} T_{k,\mathcal{B}}^{-1} w - \hat{S}_{j,\mathcal{B}} w & \text{if } w \in I_{k,3}, I_{j,3} \end{cases} \tag{5.28}$$

$$\hat{S}_{j,\mathcal{A}}w = \begin{cases} \beta_j - b_{j,\mathcal{A}} T_{j,\mathcal{A}}^{-1} w & \text{if } w \in I_{k,1}, I_{j,1} \| w \in I_{k,2}, I_{j,1} \| w \in I_{k,3}, I_{j,1} \\ \beta_j - b_{j,\mathcal{A}} T_{j,\mathcal{A}}^{-1} w - \hat{S}_{j,\mathcal{B}} w & \text{if } w \in I_{k,1}, I_{j,2} \| w \in I_{k,2}, I_{j,2} \| w \in I_{k,3}, I_{j,2} \\ 0 & \text{if } w \in I_{k,1}, I_{j,3} \| w \in I_{k,2}, I_{j,3} \| w \in I_{k,3}, I_{j,3} \end{cases} \tag{5.29}$$

$$\hat{S}_{j,\mathcal{B}}w = \begin{cases} 0 & \text{if } w \in I_{k,1}, I_{j,1} \| w \in I_{k,2}, I_{j,1} \| w \in I_{k,3}, I_{j,1} \\ 1-\beta_j - b_{j,\mathcal{B}} T_{j,\mathcal{B}}^{-1} w . \frac{2\beta_j - 1}{b_{j,\mathcal{A}} T_{j,\mathcal{A}}^{-1} w - b_{j,\mathcal{B}} T_{j,\mathcal{B}}^{-1} w} & \text{if } w \in I_{k,1}, I_{j,2} \| w \in I_{k,2}, I_{j,2} \| w \in I_{k,3}, I_{j,2} \\ 1-\beta_j - b_{j,\mathcal{B}} T_{j,\mathcal{B}}^{-1} w & \text{if } w \in I_{k,1}, I_{j,3} \| w \in I_{k,2}, I_{j,3} \| w \in I_{k,3}, I_{j,3} \end{cases} \tag{5.30}$$

### 5.5 Numerical results

This section presents the numerical results of the proposed MB-MC-NOMA scheme considering the optimal power allocation to maximize the sum-rate and the minimum fairness. It is assumed that perfect CSIT is available at the transmitter. We have used the test bench provided in 4.4

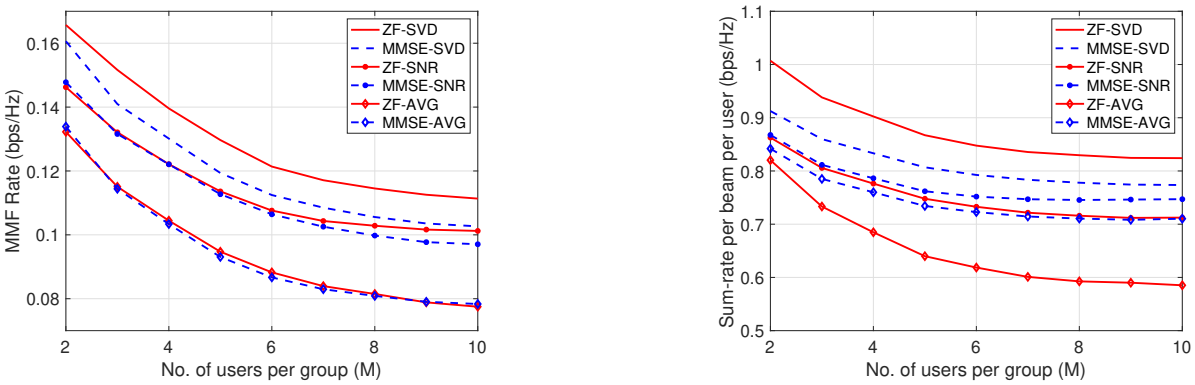
This chapter compares the performance of the MB-MC-NOMA with the MC-RS, the MC-linear precoding, and the MB-MC-OMA in terms of the minimum rate and sum-rate. We assume our proposed MB-MC-NOMA contains two groups of users per beam per time slot and $M$ users per group. In the MB-MC-OMA scheme, the orthogonal multiple access schemes such as TDMA

are used, where MB-MC-OMA contains $M$ users per beam per time slot. Moreover, the MC-linear precoding and the MC-RS contain $2M$ users per beam per time slot. To compare the performance of the techniques, we describe the framework used for each scheme and then present the simulation results.

In the MB-MC-NOMA scheme, the proposed user scheduling algorithm selects $\mathcal{U}_k^t$ out of $\mathcal{U}_k$ users to be served per beam per time slot. Then the mapper is applied to the channel coefficients of the selected users and maps the channel coefficients matrix to a vector. Next, the generated vectors are used to calculate the linear precoding matrix. In this chapter, we consider two linear precoding techniques: ZF and MMSE. Then, the effective channel gains of users are calculated, and the indices of users with lower effective channels are grouped and labeled as $I_{k,\mathcal{A}}$ and the remaining indices are labeled as $I_{k,\mathcal{B}}$. The last step is optimizing the power allocation according to different objective functions, max-min fairness, or maximum sum-rate.

In the MB-MC-OMA, MC-RS, and MC-linear precoding schemes, users with the lowest possible Euclidean distance are selected in each beam. Then, the mapper is applied to the channel coefficients of the selected users. Next, the generated vectors are used to calculate the linear precoding matrix. In the end, the power allocation is optimized to maximize the max-min fairness or sum-rate.

Figure 5–4a and Figure 5–4b demonstrate the MMF rate and sum-rate performance of the MB-MC-NOMA scheme for different numbers of users per group. These figures show the performance of the MB-MC-NOMA scheme when SVD, SNR, and average mapping approaches are applied for ZF and MMSE precoding techniques. We assume the SNR to be 8 dB. The simulation results show that the MMF rate and the sum-rate decrease as the number of users per group increases. In the multicasting transmission, the achievable rate is reduced as number of users per multicasting frame per beam increases. All users per beam share one precoding vector, even though they all have different channels. Therefore, the achievable rate is dictated by the user with the lowest SINR in the respective subframe.



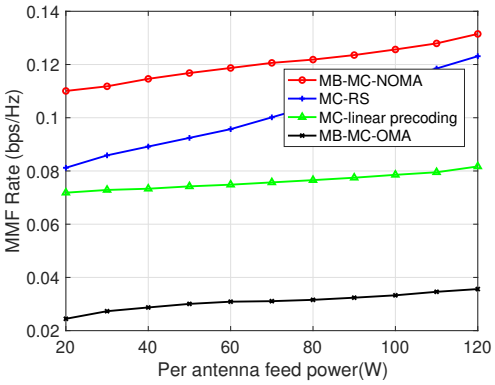

(a) MMF rate performance

(b) sum-rate performance

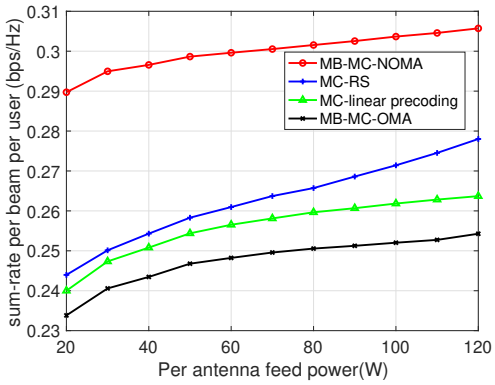
Figure 5–4: The MB-MC-NOMA performance versus number of users per group.

On the other hand, increasing the number of users per group increases the latency and spectral efficiency. Moreover, the results show that combining the ZF technique and the SVD mapper surpasses the other techniques and achieves the highest MMF rate and sum-rate in the MB-MC-NOMA scheme. The ZF outperforms the MMSE when the power is optimized to maximize fairness. The achievable rate in ZF is optimum for such cases, whereas it is suboptimal in the MMSE [81]. Moreover, the sum-rate is maximized by allocating the power to the strongest group per beam. Therefore, the SINR is high, and the ZF achieves a higher sum-rate than the MMSE for the high SINR.

Figure 5–5a and Figure 5–5b compare the MB-MC-NOMA scheme with the other techniques in terms of the MMF rate and the sum-rate versus per-feed transmit power. We assume eight users per group with the ZF precoding technique and the SVD mapper. It can be observed that the MB-MC-NOMA offers the highest MMF rate and the sum-rate in a fully overloaded system in the whole range of the transmit power. However, the gap between the MB-MC-NOMA and the MC-RS decreases with increasing the per-feed power. Due to the common part in the MC-RS, with the increased power, users can at least receive the common part without interference with higher SNR.




(a) MMF rate performance

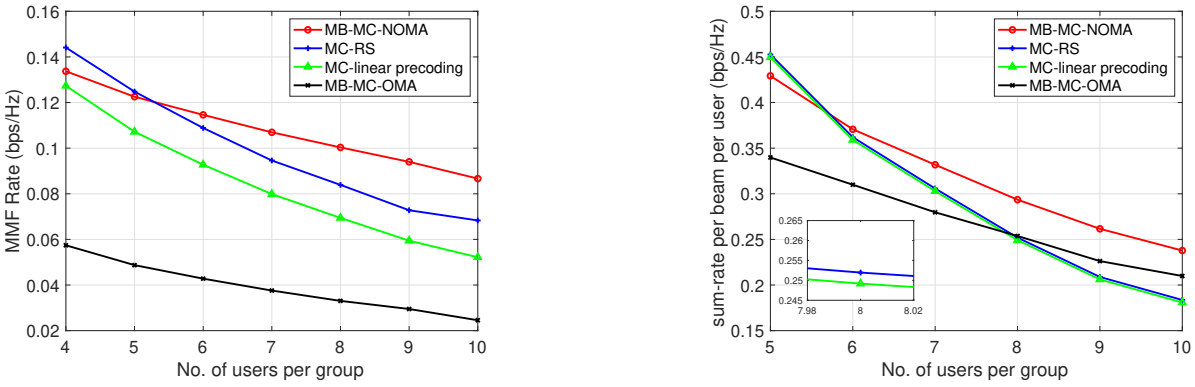



(b) sum-rate performance

Figure 5–5: MMF rate and sum-rate performance versus per-feed available power.

Figure 5–6a and Figure 5–6b compare the performance of the MB-MC-NOMA scheme with the other techniques in terms of MMF rate and sum-rate versus the different number of users per group. The ZF precoding technique is considered with the SVD mapping, and the SNR is assumed to be 15 dB. The results in Figure 5–6a and Figure 5–6b show that the MB-MC-NOMA have a higher MMF rate and sum-rate than the other techniques if the number of users per group is higher than 5 and 7, respectively. This chapter considers a maximum of 10 users per group per beam (for $G = 2$, 20 users per beam). However, there is no limit on the number of users per group per beam. The number of users can be as high as possible. In general in multicasting transmission, increasing the number of users in any given system reduces the achievable rate. However, the gain of our proposed scheme over the other solutions improves by increasing the number of users

The innovative part of our proposed solution is based on using NOMA and scheduling $G$-times more users at any time compared to the other methods, where $G$ is the number of scheduled groups. In other words, our proposed solution can be used with any state-of-the-art solutions ($M$ users per beam) and provide service to $G$-times more users ($G \times M$ users per beam). For the same number of users per beam, $M = 10$ ($G = 2$, $G \times M = 20$ number of users per beam), the MMF rate and the sum-rate of the MB-MC-NOMA gain up to 1.4 and 1.2 times over the MC-RS scheme, respectively.




(a) MMF rate performance

(b) sum-rate performance

Figure 5–6: MMF rate and sum-rate performance versus number of users per group.

In Figure 5–6b, MC-RS shows a slightly better performance than the MC-linear precoding. This confirms that while both methods achieve acceptable results in scenarios with a large number of antennas, they both get degraded as the ratio between the number of transmitter antennas and users decreases. The simulation results show that MB-MC-NOMA outperforms both methods even when the number of users per group is increased. It is worth mentioning that the sum-rate of the MB-MC-NOMA in Figure 5–5b and Figure 5–6b is calculated by considering the max-min fairness in each beam.

Up to here, the chapter proposes a method that maximizes the sum-rate with QoS in the MB-MC-NOMA scheme. In the proposed method, the power allocated to each beam should be optimized without considering the QoS. Then the power allocated to each group of users should be optimized to satisfy the QoS constraint. The power allocated to each beam without the QoS constraint can be optimized using the method proposed in section 5.3.1. Figure 5–7 compares the performance of the proposed method with the other methods, including the MB-MC-OMA, and considers equal power allocated to each beam. Two precoding techniques, ZF and MMSE with the SVD mapper, are considered for all techniques. As expected, the results for our proposed system show better performance for both ZF and MMSE compared to other technologies (OMA or equally allocated power). The MMSE and ZP almost follow each other, where ZF shows better

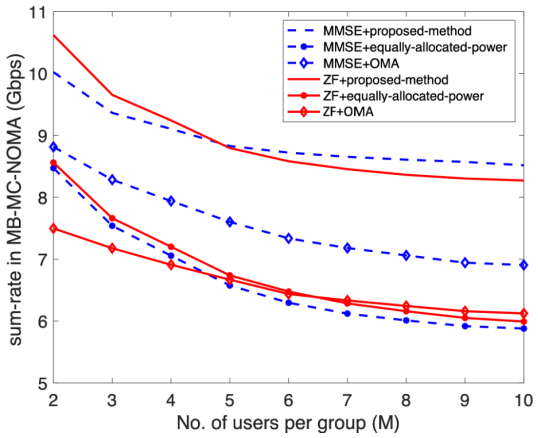



Figure 5–7: Sum-rate with QoS of MB-MC-NOMA

performance for fewer users, and MMSE achieves a higher sum-rate for a higher number of users. This can be intuitionally explained as by decreasing the number of users per group, the SINR increases. Therefore, the sum-rate of the ZF for high SINR outperforms the MMSE. However, with increasing the number of users, the SINR decreases, and the MMSE has a better performance than the ZF in terms of the sum-rate.

Finally, Figure 5–8 compares the achievable rate region of the MB-MC-NOMA and the MB-MC-OMA. The TDMA scheme is considered for the OMA techniques. To show the achievable rate region of the MB-MC-NOMA in two dimensions, we consider that all beams have the same priority. It means that $\beta_1 = \beta_2 = ... = \beta_K = \beta$. The SVD and ZF techniques are considered the mapper and precoding techniques, respectively. The results demonstrate that the MB-MC-NOMA has a higher achievable rate region than the MB-MC-OMA scheme for any number of users per group.

## 5.6 Conclusion

In conclusion, this chapter has proposed effective methods for addressing the non-convex optimization problems in the context of the MB-MC-NOMA scheme. By introducing auxiliary variables and utilizing semi-definite programming techniques, the non-convex MMF optimization problem has been transformed into a convex problem. Similarly, for the non-convex sum-rate

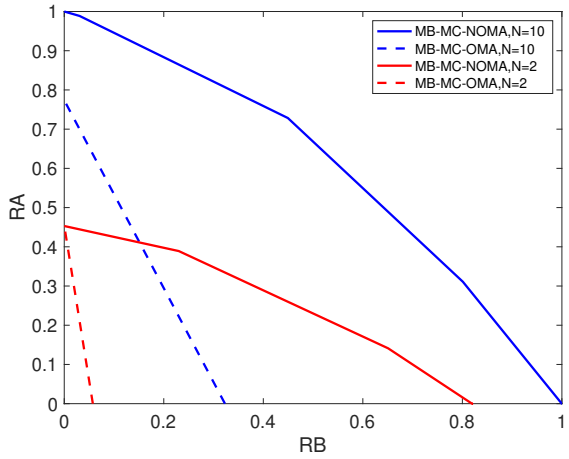



Figure 5–8: Capacity region of the MB-MC-NOMA and MB-MC-OMA

maximization objective, the Lagrangian multipliers and quadratic transforms have been employed to reformulate the problem and develop an iterative sum-rate power optimization algorithm.

Moreover, the achievable rate region of the MB-MC-NOMA scheme has been derived using the proposed power optimization techniques. Recognizing the interdependency between broadcasting power and capacity rates, the investigation of per-beam and per-user optimal power allocation has been conducted jointly. To this end, the equivalent channel and water-filling algorithm have been devised to efficiently compute the optimal transmit power spectral density (PSD) for groups of users across multiple beams.

Through comprehensive analysis and simulations, the effectiveness and efficiency of the proposed methods have been demonstrated. The derived achievable rate region and the optimized power allocation strategies contribute to enhancing the performance of the MB-MC-NOMA scheme in terms of both MMF rate and sum-rate. These findings pave the way for achieving improved fairness and transmission efficiency in multibeam multicast satellite communication systems.

# CHAPTER 6
# NOMA in multibeam multicast satellite systems with imperfect CSIT: A Rate-Splitting approach

## 6.1 Introduction

Chapter 4 and 5 investigate the performance of NOMA in a beam basis in multibeam multicast satellite communication systems while considering linear precoding to mitigate interbeam interference. This combination of NOMA and linear precoding in multibeam multicast is referred to as MB-MC-NOMA. Simulation results show that MB-MC-NOMA outperforms other interference mitigation techniques in an overloaded regime under perfect Channel State Information at the Transmitter (CSIT). However, in this chapter, we consider a realistic scenario where CSIT is imperfect, making the performance of linear precoding unreliable.

A more powerful and promising solution is Rate-Splitting (RS), which relies on superposition coding at the transmitter and Successive Interference Cancellation (SIC) at the receiver [37]. The RS is a promising solution to mitigate interbeam interference, with better performance than other techniques, even in systems under imperfect CSIT [43]-[44]. However, RS is more complex than linear precoding, which is one of its limitations.

In this chapter, we propose the use of RS to mitigate interbeam interference in a fully overloaded network under imperfect CSIT assumption, considering the NOMA scheme in each beam to improve capacity and spectral efficiency. The combination of RS and NOMA provides more flexibility to adjust power allocation, enhances performance, and improves the reliability of the system in realistic scenarios where CSIT is imperfect.

We investigate the challenges and trade-offs associated with the proposed framework, which we refer to as MC-RS-NOMA. In this chapter, we analyze the achievable data rates of the common and private parts of groups of users in MC-RS-NOMA. Moreover, we use precoding vectors for

the common and private parts to improve performance. The precoding vector of the common part is optimized to maximize the rate of the common message, while the precoding vectors of the private part are designed to cancel interbeam interference in the multicast framework.

Furthermore, we formulate the Max-Min fairness (MMF) rate and sum-rate optimization problems of MC-RS-NOMA under the imperfect CSIT assumption using the Averaging Rate (AR) framework. We employ the Weighted Minimum Mean Square Error (WMMSE) approach to make the formulated MMF and sum-rate problems convex. First, we derive a rate-WMMSE relationship, and then using the rate-WMMSE relationship and a low-complexity solution based on Alternating Optimization (AO), we transfer the problems into equivalent convex problems.

Overall, MC-RS-NOMA is a promising solution to mitigate interbeam interference in a fully overloaded multicast multibeam satellite system while improving capacity and spectral efficiency. This chapter provides a thorough analysis of the proposed scheme's performance under various scenarios, which can guide future research and development in the field of satellite communication systems.

## 6.2 System model

This chapter considers RS to mitigate interbeam interference in multibeam multicast satellite systems. We assume that the NOMA scheme is applied on a beam basis to improve the system capacity and spectral efficiency. Combining the NOMA with the RS in the multicast framework leads to MC-RS-NOMA.

Consider a Ka-band multibeam multicast satellite communication system that covers $K$ beams in the forward link as shown in Figure 4–1 depicts a single geostationary orbit (GEO) satellite that delivers service to multiple single-antenna users. The link between the single gateway and the satellite is considered to be noiseless. Additionally, $N_t$ represents the number of antenna feeds. This thesis considers one feed per beam, $N_t = K$.

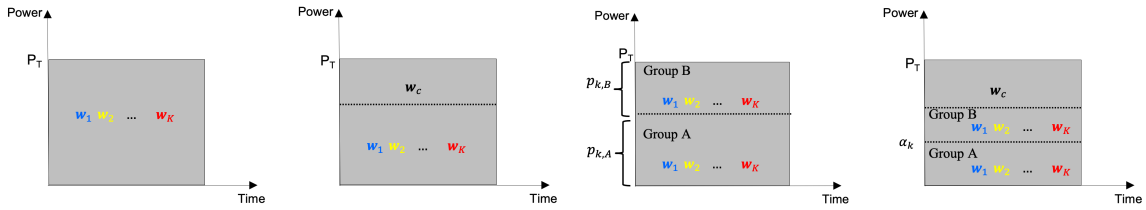
Let $\mathcal{K} = \{1, ..., K\}$ and $\mathcal{U} = \{1, ..., I\}$ gather the beams and users indices, where $K \ll I$. Let $\mathcal{U}_k$ denote the set of users belonging to beam $k$, for all $k \in \mathcal{K}$ and $G$ denote the number of groups

of users per beam. The groups indices in each beam are gathered in $\mathcal{G} = \{1, ..., G\}$. In each beam, $G \times M$ single antenna users form $G > 1$ multicast groups of $M$ users as shown in Figure 3.4. To group $G \times M$ users, $\mathcal{U}_k$ should be divided into $G$ disjoint groups of indices with cardinality of $M$, $\mathcal{U}_{k,1}, ..., \mathcal{U}_{k,G}$ where $\mathcal{U}_{k,1} \cup ... \cup \mathcal{U}_{k,G} = \mathcal{U}_k$. Moreover, this chapter considers imperfect CSIT available at the transmitter. The channel model is given in section 4.2.1.

### 6.2.1 Signal Model

The proposed MC-RS-NOMA is applied in the multibeam multicast satellite communication systems to mitigate interbeam interference and improve spectral efficiency and capacity. The scheme combines the RS and the NOMA in the multicast framework. According to the RS, each message is split into common and private parts. The common parts of all users are packed together and encoded into a common stream, $s_c$, and shared by all beams. In contrast, private parts are independently encoded into private streams for each beam, $s_k$. Based on NOMA, each beam contains $G$ groups of users. Therefore, $s_k$ is the superposition of $G$ private streams, $s_{k,1}, s_{k,2}, ..., s_{k,G}$. Figure 6–1 shows the time-power domain of the proposed MC-RS-NOMA in which $G$ is considered equal to two, group $\mathcal{A}$ and group $\mathcal{B}$.

Figure 6–1a shows the MC-Linear precoding in which users of beam $k$ are precoded using precoding vector $\mathbf{w}_k$. Figure 6–1b represents the MC-RS where $\mathbf{w}_c$ and $\mathbf{w}_k$ are the common and private precoding vectors, respectively, and the power $p_T$ is divided between these two parts. The combination of the linear precoding and the NOMA in the multicast framework, which is called MB-MC-NOMA, is shown in Figure 6–1c. In this figure, each beam contains two groups of users, groups $\mathcal{A}$ and $\mathcal{B}$. Users in groups $\mathcal{A}$ and $\mathcal{B}$ of beam $k$ are precoded using precoding vector $\mathbf{w}_k$. Figure 6–1d presents our proposed scheme, MC-RS-NOMA. In the proposed scheme, the power is divided into powers of the common and private parts. Moreover, the power of in each beam is divided between two groups of users, groups $\mathcal{A}$ and $\mathcal{B}$. All private messages of users in beam $k$ are precoded using corresponding $\mathbf{w}_k$, and the common messages are precoded by $\mathbf{w}_c$.




(a) MC-linear precoding (b) MC-RS (c) MB-MC-NOMA (d) MC-RS-NOMA

Figure 6–1: Time-power domain of different schemes in multibeam multicast satellite systems

As a result, in the proposed method, the transmitted signal in time unit is $\mathbf{x}t$, where the time units are omitted for simplicity of expression. Therefore, the transmitted signal is

$$\begin{aligned}\mathbf{x} &= \sqrt{p_c}\mathbf{w}_c s_c \sum_{j=1}^{K}\sum_{h=1}^{G}\mathbf{w}_j\sqrt{\alpha_{j,h}p_j}s_{j,h} \\ &= \sqrt{p_c}\mathbf{w}_c s_c\ \mathbf{w}_k\sqrt{\alpha_{k,g}p_k}s_{k,g}\sum_{h=1,h\neq g}^{G}\mathbf{w}_k\sqrt{\alpha_{k,h}p_k}s_{k,h}\sum_{j=1,j\neq k}^{K}\mathbf{w}_j\sqrt{p_j}s_j\end{aligned} \tag{6.1}$$

where $\mathbf{w}_c$ is the unit-norm precoding vector of the common message and $\mathbf{w}_k$ precodes the users' symbols in beam $k$. $p_c$ and $p_k$ are the allocated powers to users' common and private messages in beam $k$, respectively. Moreover, $\alpha_{k,g}$ ($\sum_{g=1}^{G}\alpha_{k,g}=1$) denote fraction of the power allocated to users of group $g$ in beam $k$. The transmitted signal is constrained to

$$p_c\sum_{k=1}^{K}p_k||\mathbf{w}_k||^2 \leq P_{\mathrm{T}} \tag{6.2}$$

where $P_{\mathrm{T}}$ is the maximum available power in the satellite payload. The received signal at user $i$ is $y_i = \mathbf{h}_i\mathbf{x}\ n_i, \forall i \in \mathcal{U}$. In terms of notation, $\mathbf{h}_i \in \mathbb{C}^{1\times N_t}$ is the channel vector between the transmitter and $i$-th user. This chapter defines $\nu i$ as mapping a user index to its corresponding beam and group indices, $\nu : i \longrightarrow k, g$. Therefore, the received signal by $i$-th user which is mapped to $k$-th beam and $g$-th group is expressed as

$$\begin{aligned}y_i &= \sqrt{p_c}\mathbf{h}_i\mathbf{w}_c s_c\ \sqrt{\alpha_{k,g}p_k}\mathbf{h}_i\mathbf{w}_k s_{k,g}\sum_{h=1,h\neq g}^{G}\sqrt{\alpha_{k,h}p_k}\mathbf{h}_i\mathbf{w}_k s_{k,h}\sum_{j=1,j\neq k}^{K}\sqrt{p_j}\mathbf{h}_i\mathbf{w}_j s_j\ n_i,\end{aligned} \tag{6.3}$$

where $n_i \sim \mathcal{CN}0, \sigma_i^2$ is the additive noise terms that contaminate the reception of $i$-th user. Without loss of generality, we assume noise variances are equal to one., i.e., $\sigma_i^2 = 1, i \in \mathcal{U}$.

According to the RS technique, each user firstly decodes the common stream $s_c$ and treats the private streams as noise. The SINR of the common part of user-$i$ is:

$$\gamma_{c,i} = \frac{p_c |\mathbf{h}_i \mathbf{w}_c|^2}{1 \sum_{j=1}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2}, \tag{6.4}$$

and its corresponding rate is $R_{c,i} = \log_2 1\ \gamma_{c,i}$. In the RS scheme, the common message, $s_c$, is shared among all beams and groups, and each user should be able to decode $s_c$. Therefore, the common rate is defined as

$$R_c = \min_{i \in \mathcal{I}} R_{c,i} \triangleq \sum_{k=1}^{K} \sum_{g=1}^{G} C_{k,g}, \tag{6.5}$$

where $C_{k,g}$ denotes the portion of common rate of group $g$ in beam $k$.

Following the MC-NOMA scheme in the power domain, different groups of users in a beam are allocated different power levels according to their channel conditions to obtain the maximum gain in system performance. The transmitter sends all users information by sending the superposition of messages. Such power allocation is also beneficial to separate different groups of users. Therefore, users can apply SIC to cancel interference from the weaker groups of users in a beam. However, the weak users perform single user detection (SUD) with considering the interference from the stronger users as the background noise.

Users within a beam are initially grouped into $G$ sets. The grouping in this chapter is based on effective channel gains. The effective channel gains of users are computed and sorted within beam $k$ at the transmitter, as shown by the inequality $|\hat{\mathbf{h}}_{\mathcal{U}_k 1} \mathbf{w}_k| \leq \ldots \leq |\hat{\mathbf{h}}_{\mathcal{U}_k G \times M} \mathbf{w}_k|$. Then the users' indices are divided into G groups. It's important to note that the design of $w_k$ is independent of the user grouping. This will be discussed further in the following section. After user grouping, it

is assumed that the effective channel gains of users have the following order:

$$\min_{i_{k,1}\in\mathcal{U}_{k,1}} |\mathbf{h}_{i_{k,1}}\mathbf{w}_k|^2 < ... < \min_{i_{k,G}\in\mathcal{U}_{k,G}} |\mathbf{h}_{i_{k,G}}\mathbf{w}_k|^2. \tag{6.6}$$

At the receiver, users first decode and remove $s_c$ through SIC. Then users in group $g$ in beam $k$, $\forall k \in \mathcal{K}, \forall g \in \mathcal{G}$, perform SIC to decode $s_{k,h}, \forall h < g$ and remove it from the received signal. Finally, users apply SUD to decode $s_{k,g}$ by considering all the other interference streams as noise. Therefore, the SINR of $i$-th user is determined by

$$\gamma_i = \frac{\alpha_{k,g} p_k |\mathbf{h}_i \mathbf{w}_k|^2}{1 \sum_{h>g}^{G} \alpha_{k,h} p_k |\mathbf{h}_i \mathbf{w}_k|^2 \sum_{j=1, j\neq k}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2}, \tag{6.7}$$

In the multicast transmission, to guarantee all users can decode their messages, the user with the lowest SINR within a group dictates the rate of the corresponding group. Therefore, the achievable rate of group $g$ in beam $k$, $r_{k,g}$ , is defined by

$$r_{k,g} \triangleq \min_{i\in\mathcal{U}_{k,g}} R_i \tag{6.8}$$

Therefore, the rate of users in group $g$ are composed of $C_{k,g}$ and $r_{k,g}$ and written as

$$R_{k,g} = C_{k,g}\ r_{k,g}. \tag{6.9}$$

The sum-rate is $R_{\text{sum-rate}} = R_c \sum_{k=1}^{K} \sum_{g=1}^{G} r_{k,g}$.

It is important to remark that the user grouping in beam $k$ should satisfies inequality (6.10)

$$\min_{i\in\mathcal{U}_{k,g}} \log_2 \left( 1 \frac{\alpha_{k,g} p_k |\mathbf{h}_i \mathbf{w}_k|^2}{1 \sum_{h>g} \alpha_{k,h} p_k |\mathbf{h}_i \mathbf{w}_k|^2 \sum_{j=1, j\neq k}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2} \right) \leq .... \leq \min_{i\in\mathcal{U}_{k,G}} \log_2 \left( 1 \frac{\alpha_{k,g} p_k |\mathbf{h}_i \mathbf{w}_k|^2}{1 \sum_{h>g} \alpha_{k,h} p_k |\mathbf{h}_i \mathbf{w}_k|^2 \sum_{j=1, j\neq k}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2} \right), \tag{6.10}$$

which would be simplified as

$$\min_{i\in\mathcal{U}_{k,g}} \frac{p_k |\mathbf{h}_i \mathbf{w}_k|^2}{1 \sum_{j=1, j\neq k}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2} \leq .... \leq \min_{i\in\mathcal{U}_{k,G}} \frac{p_k |\mathbf{h}_i \mathbf{w}_k|^2}{1 \sum_{j=1, j\neq k}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2}, \forall k \in \mathcal{K}, \forall g \in \mathcal{G} \tag{6.11}$$

This chapter assumes that criterion (6.11) is met.

### 6.2.2 Precoder Design

In the proposed MC-RS-NOMA scheme, the linear precoding vectors should be designed for the private and common parts, $\mathbf{w}_k$ and $\mathbf{w}_c$, to mitigate interbeam interference and maximize the achievable rate of the common message, respectively. Designing the linear precoding $\mathbf{w}_k$ in the multicast transmission is not straightforward because a matrix rather than a vector characterizes each beam. A low complex and suboptimal precoder is investigated in section 4.3.2 under perfect CSIT assumption. The proposed precoder consists of two steps. First, the precoder maps the channel matrix of each beam to a vector, and the composite channel matrix consists of the generated vectors. Next, the linear precoding techniques generate the precoding vectors upon the composite channel matrix. In this chapter, the precoder vectors are generated in the presence of imperfect CSIT. The optimal precoders of the private messages under imperfect CSIT are still unknown. It is shown in [82, 83, 84] that the regularized zero-forcing (RZF) would be a suitable strategy for the precoders of private messages under imperfect CSIT assumption. The precoding matrix is given by

$$\mathbf{W_{RZF}} = 1\sqrt{\gamma_{\mathbf{RZF}}}\left(\left(\hat{\mathbf{G}}^H\hat{\mathbf{G}}\ \frac{K}{P_{\mathrm{T}}}\mathbf{I}_K\right)^{-1}\hat{\mathbf{G}}^H\right). \tag{6.12}$$

where $\mathbf{I}_K$ is the $K$-dimensional identity matrix. To control the power and satisfy the power constraints, the precoding matrix should be divided by,

$$\gamma_{\mathbf{RZF}} = \max_n\left(\mathrm{diag}\left(\mathbf{W_{RZF}}\left(\mathbf{W_{RZF}}\right)^H\right)\right). \tag{6.13}$$

The estimate composite channel matrix $\hat{\mathbf{G}} = \left[\hat{\mathbf{g}}_1^H\hat{\mathbf{g}}_2^H...\hat{\mathbf{g}}_K^H\right]^T$ is generated after performing the SVD mapping per beam. The SVD mapper maps the estimated channel matrix of users in beam $k$, $\hat{\mathbf{C}}_k = \left[\hat{\mathbf{h}}_{\mathcal{U}_k1}^H,...,\hat{\mathbf{h}}_{\mathcal{U}_k2M}^H\right]$ into the vector $\hat{\mathbf{g}}_k$.

The precoding vector of the common message, $\mathbf{w}_c$, is designed to maximize the achievable rate of the common message. Therefore, the optimization problem is defined as

$$\overline{\mathcal{D}}_1 : \max_{\mathbf{w}_c \in \mathcal{N}} \min_{i \in \mathcal{U}} \pi_i . |\mathbf{h}_i \mathbf{w}_c|^2 \tag{6.14a}$$

$$\text{s.t.} \quad \|\mathbf{w}_c\|^2 = 1 \tag{6.14b}$$

where $\pi_i$ is $p_c 1 \sum_{j=1}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2$. Given that there is no interference in receiving the common message, for a realization of $n \in \mathcal{N}$, it is appropriate to employ the matched filter form to maximize the signal power [85]. This is a standard approach when handling a single in the presence of noise. Consequently, the precoder of the common message is designed as follows:

$$\mathbf{w}_c = \sum_{i \in \mathcal{U}} a_i \hat{\mathbf{h}}_i^H . \tag{6.15}$$

The optimization problem, denoted as $\overline{\mathcal{D}}_1$, can be then transformed into the task of optimizing the coefficients $a_i$. By assuming $\sigma_{e,i}^2 = \sigma_e^2, \|\mathbf{h}_i\|^2 = 1, \|\mathbf{h}_i \hat{\mathbf{h}}_j^H\|^2 = 1 - \sigma_e^2 \varepsilon^2, \forall i \in \mathcal{U}, j \neq i$, and substituting (6.15) into (6.14), the problem $\mathcal{D}_1$ is equivalently transformed to $\mathcal{D}_2$

$$\overline{\mathcal{D}}_2 : \max_{a_i} \min_{i \in \mathcal{U}} \pi_i 1 - \sigma_e^2 \left( a_i^2 \ \varepsilon^2 \sum_{n=1, n \neq i}^{I} a_n^2 \right) \tag{6.16a}$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{U}} a_i^2 = \frac{1}{N_t} \tag{6.16b}$$

The optimal solution of problem $\mathcal{D}_2$ is obtained when all terms are equal [86], i.e., $\pi_i a_i^2$ $\pi_i \varepsilon^2 \sum_{n=1, n \neq i}^{I} a_n^2 = \pi_j a_j^2 \ \pi_j \varepsilon^2 \sum_{n=1, n \neq j}^{I} a_n^2, \forall i \neq j$. Therefore, the optimal precoding vector is achieved when all users experience the same common part SINR (6.4). In this thesis for simplicity and in order to obtain a more insightful and tractable asymptotic performance, we consider that $\pi_i = \pi_j, \forall i \neq j$, then the optimal $a_i$ is equal to $a_i^* = 1\sqrt{IN_t} = 1\sqrt{GMKN_t}$.

### 6.3 Power allocation optimization

In this section we study the optimal power allocation to maximize the MMF rate and sum-rate in the proposed MC-RS-NOMA scheme under imperfect CSIT. A stochastic Average Rate (AR)

framework [39] is used to formulate the optimization problems. In this section, we first define the AR framework.

For compactness, we define $\mathbf{H} = \mathbf{h}_1^H, \mathbf{h}_2^H, ..., \mathbf{h}_I^H$, $\hat{\mathbf{H}} = \hat{\mathbf{h}}_1^H, \hat{\mathbf{h}}_2^H, ..., \hat{\mathbf{h}}_I^H$, and $\tilde{\mathbf{H}} = \tilde{\mathbf{h}}_1^H, \tilde{\mathbf{h}}_2^H, ..., \tilde{\mathbf{h}}_I^H$ which implies $\mathbf{H} = \hat{\mathbf{H}}\ \tilde{\mathbf{H}}$. For a given $\hat{\mathbf{H}}$, and sample index set $\mathcal{N} = \{1, 2, ..., N\}$, a realization sample $\mathbb{H}^N \triangleq \left\{\mathbf{H}^n = \hat{\mathbf{H}}^n\ \tilde{\mathbf{H}}^n | \hat{\mathbf{H}}, n \in \mathcal{N}\right\}$ would be a sample of $N$ i.i.d realization drawn from a conditional distribution $f\left(\mathbf{H}|\hat{\mathbf{H}}\right)$. Taking each user separately, the marginal density of the $i$-th channel conditioned on its estimate writes as $f\left(\mathbf{h}_i|\hat{\mathbf{h}}_i\right)$. These realizations are available at the transmitter and used to estimate the ARs encountered by each user employing Sample Average Functions (SAFs). According to the strong law of large numbers, $N \longrightarrow \infty$, the ARs of the $i$-th user are calculated as follows:

$$\overline{R}_{c,i} = \lim_{N\to\infty} \overline{R}_{c,i}^N = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} R_{c,i}\left(\mathbf{H}^n\right), \tag{6.17}$$

$$\overline{R}_i = \lim_{N\to\infty} \overline{R}_{j,\mathcal{A}}^N = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} R_i\left(\mathbf{H}^n\right) \tag{6.18}$$

where $R_{c,i}\left(\mathbf{H}^n\right), R_i\left(\mathbf{H}^n\right), n \in \mathcal{N}$ are the rates based on the realization sample $\mathbf{H}^n$. In the following section, the optimization problems are formulated using the AR framework.

### 6.3.1 Problem Statement

We define the optimization problems in this section. The AR framework is used to formulate the MMF and sum-rate optimization problems under imperfect CSIT.

**Max-Min fairness Analysis**

The MMF optimization problem using the AR framework can be formulated as

$$\overline{\mathcal{P}}_1 : \underset{p,\alpha,\overline{\mathbf{c}}}{\text{argmax}} \quad \min_{k\in\mathcal{K}} \min_{g\in\mathcal{G}} \left\{ \overline{C}_{k,g} \min_{i\in\mathcal{U}_{k,g}} \overline{R}_i^N \right\} \tag{6.19a}$$

s.t.

$$\overline{R}_{c,i}^N \geq \sum_{k=1}^{K} \sum_{g=1}^{G} \overline{C}_{k,g}, \forall i \in \mathcal{I} \tag{6.19b}$$

$$\overline{C}_{k,g} \geq 0, \forall g \in \mathcal{G}, \forall k \in \mathcal{K} \tag{6.19c}$$

$$\alpha_{k,g} \in 0,1, \sum_{g=1}^{G} \alpha_{k,g} = 1, \forall k \in K \tag{6.19d}$$

$$p_c \sum_{k=1}^{K} p_k ||\mathbf{w}_k||^2 \leq P_{\mathrm{T}}, \forall k \in \mathcal{K} \tag{6.19e}$$

here $\overline{\mathbf{c}} = \overline{C}_{1,1}, ..., \overline{C}_{1,G}, ..., \overline{C}_{K,1}, ..., \overline{C}_{K,G}$ is the vector of Average common-rate portions. The constraint (6.19b) guarantees $s_c$ to be decoded by each user since the definition of the Average common rate is $\overline{R}_c = \sum_{k=1}^{K} \sum_{g=1}^{G} \overline{C}_{k,g} = \min_{i\in\mathcal{U}} \overline{R}_{c,i}$. Constraint (6.19c) implies that each portion of the Average common rate is non-negative. Constraints (6.19d) and (6.19e) are the power constraint. By solving Problem $\overline{\mathcal{P}}_1$, variables $\overline{\mathbf{c}}$, $p$, $\alpha$ are jointly optimized. Note that by fixing $p_c = 0$ and $\overline{\mathbf{c}} = 0$, the M-RS-NOMA scheme turns into MB-MC-NOMA technique.

**Sum-rate Analysis**

The sum-rate optimization is another problem which is addressed in this thesis. The sum-rate maximization under imperfect CSIT is also formulated using the AR framework as

$$\overline{\mathcal{S}}_1 : \underset{\overline{R}_c,t,\alpha_k}{\text{argmax}} \quad \overline{R}_c \sum_{k=1}^{K} \sum_{g=1}^{G} \min_{i\in I_{k,g}} \overline{R}_i^N \tag{6.20a}$$

s.t.

$$\overline{R}_{c,i}^N \geq \overline{R}_c, \forall i \in \mathcal{U} \tag{6.20b}$$

$$6.19d, 6.19e \tag{6.20c}$$

where $\overline{R}_c$ is an auxiliary variable. The constraint (6.20b) guarantees that all users can decode $s_c$.

Problems $\overline{\mathcal{P}}_1$ and $\overline{\mathcal{S}}_1$ are non-convex problems that are very challenging to solve because they contain superimposed rate expressions. The WMMSE approach is proposed in [39, 87] to solve non-convex problems containing superimposed rate expressions, i.e., RS. An alternating optimization algorithm based on the modified WMMSE is proposed in [43] to solve the RS's MMF optimization problem in a multibeam multicast satellite system.

This chapter uses the WMMSE approach to define the rates in terms of WMMSEs variables. Next, using this definition, the problems are transferred into a block-wise convex problem which can be solved iteratively through interior-point methods. In the following section, we derive the rate-WMMSE expressions.

### 6.3.2 Rate-WMMSE Relationship

In this section we establish the Rate-WMMSE relationship. Consider the estimate of $s_c$ in user $i$ as denoted by $\hat{s}_{c,i} = g_{c,i} y_i$, where $g_{c,i}$ is a scalar equalizer. Since the transmitter sends the superposition of $s_c$ and $s_{k,g}$, $\forall k \in \mathcal{K}, g \in \mathcal{G}$, user $i$ first decodes and removes $s_c$ from the received signal. Next, user $i$ which belongs to beam $k$ and group $g$ decodes and removes the signals of the weaker groups in beam $k$, $h < g$, through the SIC. However, the signals of the other beams and signals of the stronger groups in beam $k$ are considered as the background noise. Therefore, the estimate of $s_{k,g}$ is $\hat{s}_{k,g} = g_i \left( y_i - \sqrt{p_c}\mathbf{h}_i\mathbf{w}_c s_c - \sum_{h=1}^{h<g} \sqrt{\alpha_{k,h} p_k}\mathbf{h}_i\mathbf{w}_k s_{k,h} \right)$. The mean square errors (MSEs) of the common and private parts are defined as

$$\varepsilon_{c,i} = \mathbb{E}\left\{ |\hat{s}_{c,i} - s_{c,i}|^2 \right\} = |g_{c,i}|^2 T_{c,i} \; 1 - 2\Re\left\{ \sqrt{p_c} g_{c,i} \mathbf{h}_i \mathbf{w}_c \right\} \tag{6.21a}$$

$$\varepsilon_i = \mathbb{E}\left\{ \left| \hat{s}_{k,g} - s_{k,g} \right|^2 \right\} = |g_i|^2 T_i \; 1 - 2\Re\left\{ \sqrt{\alpha_{k,g} p_k} g_i \mathbf{h}_i \mathbf{w}_k \right\} \tag{6.21b}$$

where

$$T_{c,i} = p_c |\mathbf{h}_i \mathbf{w}_c|^2 \sum_{k=1}^{K} p_k |\mathbf{h}_i \mathbf{w}_k|^2 \; 1 \tag{6.22a}$$

$$T_i = p_k \sum_{h \geq g}^{G} \alpha_{k,h} |\mathbf{h}_i \mathbf{w}_k|^2 \sum_{j=1, j \neq k}^{K} p_j |\mathbf{h}_i \mathbf{w}_j|^2 \; 1 \tag{6.22b}$$

Moreover, we define the interference as

$$I_{c,i} = T_{c,i} - p_c|\mathbf{h}_i\mathbf{w}_c|^2, \tag{6.23}$$

$$I_i = T_i - \alpha_{k,g}p_k|\mathbf{h}_i\mathbf{w}_k|^2 \tag{6.24}$$

The optimum equalizers achieve by minimizing the MSEs over equalizers,

$$\frac{\partial \varepsilon_{c,i}}{g_{c,i}} = 0 \rightarrow g_{c,i}^{\mathrm{MMSE}} = \sqrt{p_c}\mathbf{h}_i\mathbf{w}_c T_i^{-1} \tag{6.25a}$$

$$\frac{\partial \varepsilon_i}{g_i} = 0 \rightarrow g_i^{\mathrm{MMSE}} = \sqrt{\alpha_{k,g}p_k}\mathbf{h}_i\mathbf{w}_k T_i^{-1} \tag{6.25b}$$

The MMSEs with optimum equalizers are $\varepsilon_{c,i}^{\mathrm{MMSE}} = \min\limits_{g_{c,i}} \varepsilon_{c,i} = T_{c,i}^{-1} I_{c,i}, \varepsilon_i^{\mathrm{MMSE}} = \min\limits_{g_i} \varepsilon_i = T_i^{-1} I_i$. Apparently, the SINRs can be expressed in the form of MMSEs, i.e. $\gamma = 1\varepsilon^{\mathrm{MMSE}} - 1$. So, the corresponding rates write as $R = -\log_2 1\varepsilon^{\mathrm{MMSE}}$. Now, the common and private augmented WMSEs are given by

$$\begin{aligned} \xi_{c,i} &= u_{c,i}\varepsilon_{c,i} - \log_2 u_{c,i}, \\ \xi_i &= u_i\varepsilon_i - \log_2 u_i, \end{aligned} \tag{6.26}$$

where $u_{c,i}, u_i > 0$ are weights associated with MSEs. In the following, we consider $\xi$s as WMSEs and, for simplicity, drop the "augmented". Then the optimal equalizers are substituted into the WMSEs, and we obtain

$$\xi_{c,i}\left(g_{c,i}^{\mathrm{MMSE}}\right) = \min_{g_{c,i}} \xi_{c,i} = u_{c,i}\varepsilon_{c,i}^{\mathrm{MMSE}} - \log_2 u_{c,i} \tag{6.27a}$$

$$\xi_i\left(g_i^{\mathrm{MMSE}}\right) = \min_{g_i} \xi_i = u_i\varepsilon_i^{\mathrm{MMSE}} - \log_2 u_i \tag{6.27b}$$

To minimize WMSEs over both equalizers and weights, $\frac{\partial \xi_{c,i}\left(g_{c,i}^{\mathrm{MMSE}}\right)}{\partial u_{c,i}} = 0$, $\frac{\partial \xi_i\left(g_i^{\mathrm{MMSE}}\right)}{\partial u_i} = 0$. Therefore, the optimum weights are

$$u_{c,i} = \left(\varepsilon_{c,i}^{\mathrm{MMSE}}\right)^{-1}, u_i = \left(\varepsilon_i^{\mathrm{MMSE}}\right)^{-1}. \tag{6.28}$$

We substitute them into (6.27a), (6.27b), leading to the Rate-WMMSE relationship

$$\xi_{c,i}^{\mathrm{MMSE}} = \min_{g_{c,i},u_{c,i}} \xi_{c,i} = 1 \log_2 \varepsilon_{c,i}^{\mathrm{MMSE}} = 1 - R_{c,i} \tag{6.29a}$$

$$\xi_{i}^{\mathrm{MMSE}} = \min_{g_{i},u_{i}} \xi_{i} = 1 \log_2 \varepsilon_{i}^{\mathrm{MMSE}} = 1 - R_{i} \tag{6.29b}$$

Considering imperfect CSIT, a deterministic SAF version of the Rate-WMMSE connection is developed so the average WMMSEs are

$$\overline{\xi}_{c,i}^{\mathrm{MMSE}N} = \frac{1}{N} \lim_{N\to\infty} \sum_{n=1}^{N} \xi_{c,i}^{\mathrm{MMSE}n} = 1 - \overline{R}_{c,i}^{N} \tag{6.30a}$$

$$\overline{\xi}_{i}^{\mathrm{MMSE}N} = \frac{1}{N} \lim_{N\to\infty} \sum_{n=1}^{N} \xi_{i}^{\mathrm{MMSE}n} = 1 - \overline{R}_{i}^{N} \tag{6.30b}$$

where $\xi_{c,i}^{\mathrm{MMSE}n}$ and $\xi_{i}^{\mathrm{MMSE}n}$ are associated with the $n$-th realization in $\mathbb{H}^N$. The sets of optimum MMSE equalizers associated with (6.30) are defined as $\mathbf{g}_{c,i}^{\mathrm{MMSE}} = \{g_{c,i}^{\mathrm{MMSE}n} | n \in \mathcal{N}\}$, $\mathbf{g}_{i}^{\mathrm{MMSE}} = \{g_{i}^{\mathrm{MMSE}n} | n \in \mathcal{N}\}$. Moreover, the sets of optimum weights are $\mathbf{u}_{c,i}^{\mathrm{MMSE}} = \{u_{c,i}^{\mathrm{MMSE}n} | n \in \mathcal{N}\}$, $\mathbf{u}_{i}^{\mathrm{MMSE}} = \{u_{i}^{\mathrm{MMSE}n} | n \in \mathcal{N}\}$. Therefore, in each realization in $\mathbb{H}^N$, the optimum equalizer and weights are calculated. The composite set of optimum equalizer and weights are defined as

$$\mathbf{G}^{\mathrm{MMSE}} = \left\{ \mathbf{g}_{c,i}^{\mathrm{MMSE}}, \mathbf{g}_{i}^{\mathrm{MMSE}} | i \in \mathcal{U}, \right\} \tag{6.31}$$

$$\mathbf{U}^{\mathrm{MMSE}} = \left\{ \mathbf{u}_{c,i}^{\mathrm{MMSE}}, \mathbf{u}_{i}^{\mathrm{MMSE}} | i \in \mathcal{U} \right\} \tag{6.32}$$

Using the Rate-WMMSE relationship, the optimization problems are rewritten using the WMMSE variables in the following section. The relationship between the achievable rates in the MC-RS-NOMA scheme and the WMMSE variables are given in the (6.30).

### 6.3.3 WMMSE Reformulation

In this section, we reformulate the optimization problems using the WMMSE expressions.

**Max-Min fairness Analysis**

Using the Rate-WMMSE relationship, and auxiliary variables, $\bar{z}$, $\mathbf{G}$, $\mathbf{U}$, $\bar{r}_g = \bar{r}_{1,g}, ..., \bar{r}_{K,g}$, the problem $\overline{\mathcal{P}}_1$ can be transferred into an equivalent convex WMMSE problem, $\overline{\mathcal{P}}_2$:

$$\overline{\mathcal{P}}_2 : \underset{p,\alpha,\bar{\mathbf{c}},\bar{z},\bar{r}_g}{\text{argmax}} \bar{z} \tag{6.33a}$$

s.t.

$$\overline{C}_{k,g}\ \bar{r}_{k,g} \geq \bar{z}, \forall k \in \mathcal{K}, \forall g = \{1, ..., G\} \tag{6.33b}$$

$$1 - \overline{\xi}_i^N \geq \bar{r}_{k,g}, \forall i \in \mathcal{U}_{k,g}, \forall k \in \mathcal{K}, \forall g = \{1, ..., G\} \tag{6.33c}$$

$$1 - \overline{\xi}_{c,i}^N \geq \sum_{k=1}^{K} \sum_{g=1}^{G} \overline{C}_{k,g}, \forall i \in \mathcal{U} \tag{6.33d}$$

$$6.19d, 6.19e \tag{6.33e}$$

where $\overline{\xi}_{c,i}$ and $\overline{\xi}_i$ are given in (6.26). It is worth to mention if $p^*, \alpha^*, \bar{\mathbf{c}}^*, \bar{z}^*, \mathbf{G}^*, \bar{r}_g^*, \mathbf{U}^*$ satisfies the KKT optimality conditions of $\overline{\mathcal{P}}_2$, $p^*, \alpha^*, \bar{\mathbf{c}}^*$ will satisfy the KKT optimality conditions of $\overline{\mathcal{P}}_1$.

**Sum-rate Analysis**

Motivated by the Rate-WMMSE relationships given in (6.30), and the auxiliary variables, $\overline{\xi_c}, \mathbf{U}, \mathbf{G}$, the problem $\overline{\mathcal{S}}_1$ is equivalently transferred into the convex problem $\overline{\mathcal{S}}_2$. The problem is reformulated as

$$\overline{\mathcal{S}}_2 : \underset{\overline{\xi}_c, p, \alpha_k}{\text{argmin}} \quad \overline{\xi}_c \sum_{k=1}^{K} \sum_{g=1}^{G} \max_{i \in \mathcal{U}_{k,g}} \overline{\xi}_i^N \tag{6.34a}$$

s.t.

$$\overline{\xi}_{c,i}^N \leq \overline{\xi}_c, \forall i \in \mathcal{U} \tag{6.34b}$$

$$6.19d, 6.19e \tag{6.34c}$$

where $\overline{\xi}_c$ refers to the common AWMSE. Noted problem $\overline{\mathcal{S}}_2$ and problem $\overline{\mathcal{S}}_1$ are equivalence. It means that for any point $p^*, \alpha^*, \overline{\xi}_c^*, \mathbf{G}^*, \mathbf{U}^*$ satisfying the KKT optimality conditions of problem $\overline{\mathcal{S}}_2$, $p^*, \alpha^*, \overline{\xi}_c^*$, satisfies the KKT optimality conditions of problem $\overline{\mathcal{S}}_1$.

Problems $\overline{\mathcal{P}}_2$ and $\overline{\mathcal{S}}_2$ are still non-convex. However, they are convex if two variables out of three variables, equalizer, weight, and power, are fixed. Considering this block-wise convexity property, we propose an Alternating Optimization algorithm to solve the problems $\overline{\mathcal{P}}_2$ and $\overline{\mathcal{S}}_2$.

### 6.3.4 Alternating Optimization Algorithm

The problems $\overline{\mathcal{P}}_2$ and $\overline{\mathcal{S}}_2$ remain non-convex for the entire set of optimization variables, which include $\alpha$, $p$, $\bar{\mathbf{c}}$, $\mathbf{U}$, and $\mathbf{G}$. However, they exhibit block-wise convexity, which can be leveraged to propose an alternating optimization algorithm. Each iteration of the algorithm consists of two steps: (1) updating $\mathbf{U}$ and $\mathbf{G}$ based on the value of $\mathbf{p}$ and $\alpha$ from the previous iteration, and (2) updating $\mathbf{p}$, $\alpha$, and $\bar{\mathbf{c}}$ using $\mathbf{U}$ and $\mathbf{G}$ obtained in step 1. We now provide a detailed explanation of these two steps.

**Step 1: Updating G, U**

In $l$-th iteration, all the equalizers and weights are updated according to the $\mathbf{p}, \alpha$ form the previous round, $l-1$, $\mathbf{G}p^{l-1}, \alpha^{l-1}$,$\mathbf{U}p^{l-1}, \alpha^{l-1}$. The corresponding SAFs $\overline{u}_{c,i}$, $\overline{u}_i$, $\overline{g}_{c,i}$, $\overline{g}_i$ are calculated by taking average over $N$ realization. To facilitate the next step, we introduce a set of variables are

$$t_{c,i} = u_{c,i}^n |q_{c,i}^n|^2, \qquad t_i = u_i^n |q_i^n|^2, \tag{6.35}$$

$$\Psi_{c,i}^n = t_{c,i} \mathbf{h}_i^{nH} \mathbf{h}_i^n, \qquad \Psi_i^n = t_i \mathbf{h}_i^{nH} \mathbf{h}_i^n, \tag{6.36}$$

$$f_{c,i}^n = u_{c,i}^n q_{c,i}^n \mathbf{h}_i^n \mathbf{w}_c^n, \quad f_i^n = u_i^n q_i^n \mathbf{h}_i^n \mathbf{w}_k^n \tag{6.37}$$

$$v_{c,i}^n = \log_2 u_{c,i}, \qquad v_i^n = \log_2 u_i \tag{6.38}$$

and the corresponding SAFs are calculated in the same way,

$$t_{c,i}^N, \Psi_{c,i}^N, f_{c,i}^N, v_{c,i}^N, t_i, \Psi_i^N, f_i^N, v_i^N \tag{6.39}$$

**Step 2: Updating $\mathbf{p}, \alpha$**

In the $l$-th iteration up to this step, we fix $\mathbf{G}$, $\mathbf{U}$, and the other introduced variables, which are obtained using the updated valusers of $\mathbf{U}, \mathbf{G}$. With these updated variables, in this step, the problems $\overline{\mathcal{P}}_2$ and $\overline{\mathcal{S}}_2$ transform into problems $\overline{\mathcal{P}}_3^l$ and $\overline{\mathcal{S}}_3^l$, which are convex problems. These problems can be solved using interior-point methods, allowing for the optimization of $\mathbf{p}$, $\alpha_{\mathbf{k}}$, and the other auxiliary variables.

$$\overline{\mathcal{P}}_3^l : \underset{\mathbf{p},\alpha,\bar{\mathbf{c}},\bar{z},\bar{r}_g,}{\operatorname{argmax}} \bar{z} \tag{6.40a}$$

s.t.

$$\overline{C}_{k_g}\ \bar{r}_{k_g} \geq \bar{z}, \quad \forall k \in \mathcal{K}, \forall g_k = \{1, ..., G_k\} \tag{6.40b}$$

$$1 - \bar{r}_{g_k} \geq \sum_{j=1, j\neq k}^{K} p_j \overline{\mathbf{w}}_j^H \overline{\Psi}_i^N \overline{\mathbf{w}}_j^N - 2\mathcal{R}\left\{\sqrt{\alpha_{g_k} p_k} \overline{f}_i^N\right\} \bar{t}_i^N\ \bar{u}_i^N - \bar{v}_i^N$$

$$\sum_{h \geq g} p_k \alpha_{h_k} \overline{\mathbf{w}}_k^H \overline{\Psi}_i^N \overline{\mathbf{w}}_k^N, \quad \forall i \in \mathcal{U}_{g_k}, \forall k \in \mathcal{K}, \forall g_k \in \mathcal{G}_k \tag{6.40c}$$

$$1 - \sum_{k=1}^{K} \sum_{g_k=1}^{G_k} \overline{C}_{g_k} \geq p_c \overline{\mathbf{w}}_c^H \overline{\Psi}_{c,i}^N \overline{\mathbf{w}}_c \sum_{k=1}^{K} p_k \overline{\mathbf{w}}_k^H \overline{\Psi}_{c,i}^N \overline{\mathbf{w}}_k\ \bar{t}_{c,i}^N - 2\mathcal{R}\left\{\sqrt{p_c} \overline{f}_{c,i}^N\right\}$$

$$\overline{u}_{c,i}^N - \bar{v}_{c,i}^N, \quad \forall i \in \mathcal{U} \tag{6.40d}$$

$$6.19d, 6.19e \tag{6.40e}$$

and

$$\overline{\mathcal{S}}_3^l : \underset{\overline{\xi}_c, \mathbf{p}, \alpha_{\mathbf{k}}}{\operatorname{argmin}} \quad \overline{\xi}_c \sum_{k=1}^{K} \sum_{g=1}^{G} \{\max_{i \in \mathcal{U}_{g_k}} \overline{\xi}_i\} \tag{6.41a}$$

s.t.

$$p_c \overline{\mathbf{w}}_c^H \overline{\Psi}_{c,i}^N \overline{\mathbf{w}}_c \sum_{k=1}^{K} p_k \overline{\mathbf{w}}_k^H \overline{\Psi}_{c,i}^N \overline{\mathbf{w}}_k\ \bar{t}_{c,i}^N - 2\mathcal{R}\left\{\sqrt{p_c} \overline{f}_{c,i}^N\right\} \overline{u}_{c,i}^N - \bar{v}_{c,i}^N \leq \overline{\xi}_c, \forall i \in \mathcal{U} \tag{6.41b}$$

$$6.19d, 6.19e \tag{6.41c}$$

where $\overline{\xi}_i$ is

$$\overline{\xi}_i = \sum_{j=1, j\neq k}^{K} p_j \overline{\mathbf{w}}_j^H \overline{\Psi}_i^N \overline{\mathbf{w}}_j^N \bar{t}_i^N - 2\mathcal{R}\left\{\sqrt{\alpha_{g_k} p_k} \bar{f}_i^N\right\} \sum_{h_k \geq g_k} p_k \alpha_{h_k} \overline{\mathbf{w}}_k^H \overline{\Psi}_i^N \overline{\mathbf{w}}_k^N$$
$$\overline{u}_i^N - \bar{v}_i^N, \forall i \in \mathcal{U}_{g_k}, \forall k \in \mathcal{K}, \forall g_k \in \mathcal{G}_k$$

As the iteration procedure continusers, the objective function in $\mathcal{P}_3$ or $\mathcal{S}_3$ grows until convergence. The proposed alternating optimization approach alternately optimizes the variables of the corresponding WMMSE problem $\overline{\mathcal{P}}_3$ and $\overline{\mathcal{S}}_3$. The proposed algorithm is guaranteed to converge as the objective function is bounded above for the specified power limitations.

## 6.4 Numerical results

In this section, the proposed algorithms in the multibeam multicast satellite systems are evaluated through simulations. The performance of the proposed scheme is investigated in terms of the MMF rate and the sum-rate. This chapter consider the test bench provided in section 4.4.

This chapter assumes the noise variance to be equal to one for all users, $\sigma_i^2 = 1, \forall i \in \mathcal{U}$ from which SNR is $p_T$. To model the CSIT uncertainty, entries of $\tilde{\mathcal{H}}$ are i.i.d complex Gaussian drawn from $\mathcal{CN}0, \sigma_e^2$ where $\sigma_e^2 = N_t^{-1}\sigma_{e,i}^2 = P^{-\eta}, \forall i \in \mathcal{U}$. The sample size $N$ is set to 1000. For each realization,$\mathbf{H}n$, the channel estimation is $\hat{\mathbf{H}}^n = \mathbf{H}\ \tilde{\mathbf{H}}^n$, where $\tilde{\mathbf{H}}^n$ follows the above CSIT error distribution.

Since in multibeam satellite communication systems, each antenna has its amplifier, this chapter assumes equal per-feed power. Therefore, power allocated to the common and private parts are given by $p_c = 1 - tP_T$, $p_k = \frac{tP_T}{K}$. Moreover, we consider two groups per beam, $G = 2$. The convex optimization problems are solved using the CVX toolbox.

### 6.4.1 MMF rate performance

This section compares the performance of the proposed method, MC-RS-NOMA with the other cutting edge methods, MC-RS and MB-MC-NOMA techniques, in terms of the MMF rate. The target is to maximize the minimum rate by optimizing the power allocation. The MMF rate

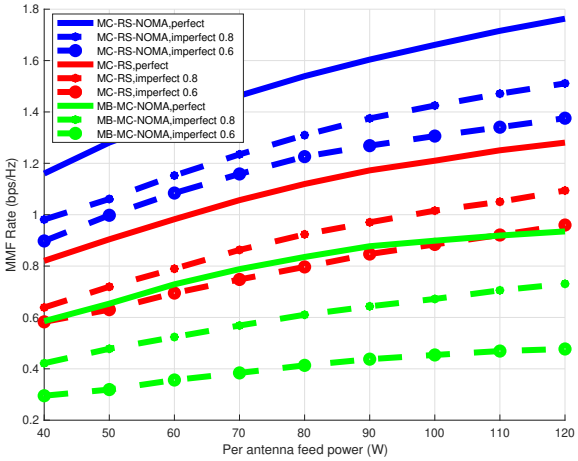



Figure 6–2: MMF rate performance versus per feed-power constraint, $I = 28, M = 2$ users

is calculated according to different power per feed, number of users per group, and CSIT quality parameter.

**Power per feed constraint**

Figure 6–2 shows the MMF rate versus different available transmit power per feed. We assume $I = 28$ users total, $M = 2$ users per group per beam, are served. Our proposed method, MC-RS-NOMA, has a more flexible architecture for non-orthogonal transmission and robust interference management. Thus it outperforms the other techniques in the whole range of per-feed available power. The gap between MC-RS-NOMA and the other techniques increases with higher power per antenna feed. If the CSIT is perfect, the MC-RS-NOMA scheme achieves up to 1.38 and 1.8 gains over MC-RS and MB-MC-NOMA schemes, respectively. For imperfect CSIT, it is observed that the gain of the MC-RS-NOMA over the MC-RS and MB-MC-NOMA is around 1.38 and 2 when $\eta = 0.8$, respectively. Consequently, the benefit of employing RS in multibeam satellite systems under imperfect CSIT is observed. It showed that MC-RS-NOMA and MC-RS are more resistant to CSIT uncertainty than MB-MC-NOMA. Moreover, MC-RS-NOMA has better performance than the MC-RS.

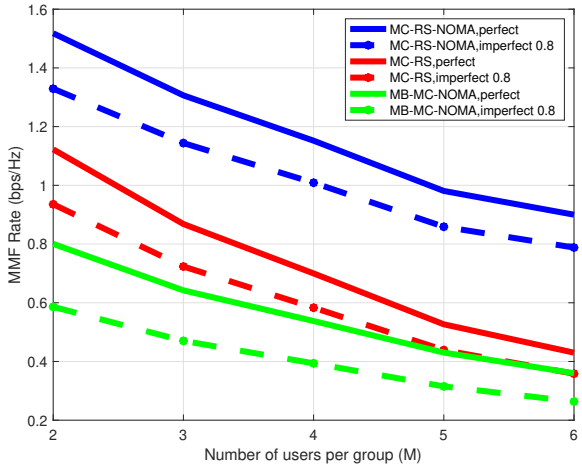



Figure 6–3: MMF rate performance versus number of users per group ($M$), $p_k = 120W$ and $\eta = 0.8$

**Number of users per group**

This section shows the effect of the number of users on the MMF rate performance. Figure 6–3 indicates the MMF rates versus the different number of users per group per beam. The power per feed is assumed to be $p_k = 120W$ and the CSIT quality parameter $\eta = 0.8$. Figure 6–3 shows The MMF rate decreases with the increasing number of users for all cases. Because all users within a beam share one precoding vector even though they all have different channels. Consequently, the user with the lowest SINR controls its beam rate. Despite this performance degradation, MC-RS-NOMA can still provide gains compared to MC-RS and MB-MC-NOMA schemes. The results also show that MC-RS-NOMA achieves more gain over the MC-RS scheme by increasing the number of users per group per beam, and the gap between them increases. For $M = 6$ (G = 2, G M = 12 number of users per beam), the MMF rate of the proposed MC-RS-NOMA gains around 2.25 times over the MC-RS scheme for the perfect CSIT.

**CSIT quality**

The influence of the CSIT uncertainty is shown in this section. Figure 6–4 depicts the MMF rate versus a wide range of CSIT quality. Here, we set the per-feed available transmit power to be $p_k = 80W$ and $p_k = 120W$. The number of users per group is equal to two, $M = 2$. The

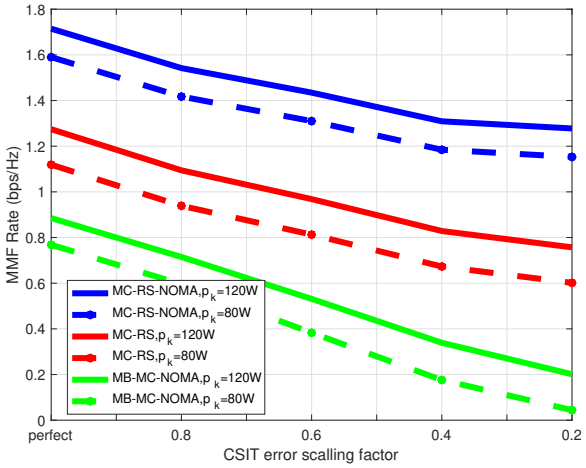



Figure 6–4: MMF rate performance versus CSIT uncertainty ($\eta$), $M = 2, I = 28$ users

MMF rate decreases as the CSIT quality parameter drops for all cases. The results shows that our proposed scheme has higher MMF rate However, the MMF rate gap between MC-RS-NOMA and MB-MC-NOMA gradually rises, indicating that the benefits of our proposed MC-RS-NOMA scheme become increasingly evident when CSIT quality degrades. It is shown that the MMF rate of the MC-RS-NOMA and MC-RS drops 1.28 and 1.625 times with increasing the CSIT error.

### 6.4.2 Sum-rate performance

In this section, we investigate the performance of the proposed MC-RS-NOMA scheme with the other cutting edge methods, MC-RS and MB-MC-NOMA techniques, in terms of the sum-rate. The sum-rate is calculated according to different power per feed, number of users per group, and CSIT quality parameter.

**Power per feed constraint**

Figure 6–6 shows the sum-rate rate versus different available transmit power per feed. This section considers $I = 28$ users total, $M = 2$ users per group per beam, are served. The results show that the MC-RS-NOMA performs better, and the gap with the other techniques increases by increasing the power. For imperfect CSIT, it is observed that the gain of the MC-RS-NOMA over the MC-RS and MB-MC-NOMA is around 1.67 and 1.175 when $\eta = 0.8$, respectively. If the

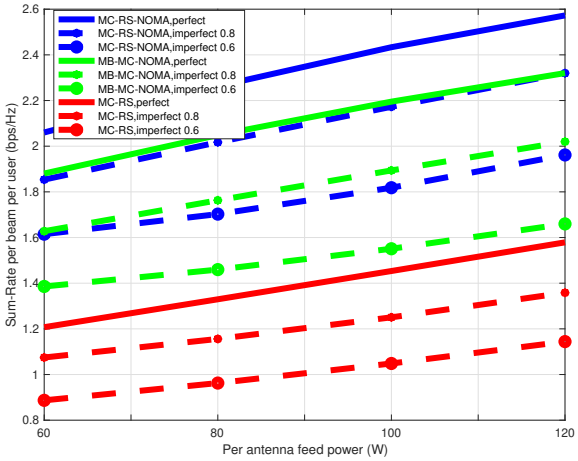



Figure 6–5: Sum-rate performance versus per feed-power constraint, $I = 28, M = 2$ users

CSIT is perfect, the MC-RS-NOMA scheme achieves up to 1.625 and 1.1 gains over MC-RS and MB-MC-NOMA schemes, respectively.

**Number of users per group**

Figure 6–3 illustrates the sum-rate versus the number of users per group, $M$, when the transmit power per antenna feed is $120W$ and the CSIT scaling factor is $\eta = 0.8$. As expected, the sum-rate drops with increasing $M$ for all cases. However, MC-RS-NOMA shows an explicit sum-rate gain compared to all other schemes for all $M$. For $M = 5$, the MC-RS-NOMA gain over the MC-RS goes up to 2 times for both perfect and imperfect CSIT.

**CSIT uncertainty**

In Figure 6–7, we further evaluate the sum-rate performance of the proposed MC-RS-NOMA with imperfect CSIT. The transmit power per antenna feed is considered to be $p_k = 120W$ and $p_k = 80W$ and two users per group, $M = 2$. The results show that the proposed MC-RS-NOMA has a much better performance for all CSIT error scaling parameters ($\eta$). The gain of MC-RS-NOMA over MB-MC-NOMA and MC-RS increases from 1.1 and 1.625 to 1.75 and 2 when $\eta$ drops from 1 to 0.2.

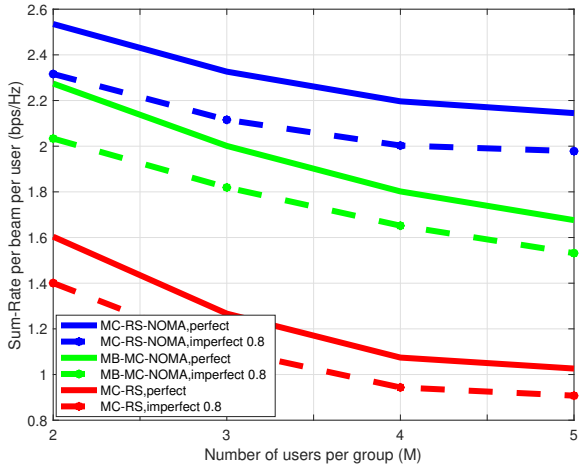



Figure 6–6: Sum-rate performance MMF rate performance versus number of users per group ($M$), $p_k = 120W$ and $\eta = 0.8$

All the simulation results show that the proposed MC-RC-NOMA has much better performance in terms of the MMF and the sum-rate in any conditions of a realistic scenario. However, the proposed scheme has more complexity than the MC-RS. Since the complexity of the MC-RS-NOMA at the receiver is $\mathcal{O}G^2$ time more than the MC-RS.

## 6.5 Conclusion

In summary, this chapter introduces the MC-RS-NOMA scheme as a solution to address the challenges posed by CSIT uncertainty in multibeam multicast satellite communication systems. The performance of the proposed scheme is evaluated in terms of max-min fairness and sum-rate metrics. To optimize the scheme, a modified WMMSE method and an AO algorithm are developed to maximize the MMF rate and sum-rate, respectively.

The proposed MC-RS-NOMA scheme demonstrates promising results compared to state-of-the-art techniques such as MB-MC-NOMA and MC-RS. It effectively tackles practical obstacles, including CSIT uncertainty, practical per-feed limits, and the overloaded regime, which are crucial considerations in real-world multibeam satellite communication scenarios.

The findings of this chapter highlight the potential of the MC-RS-NOMA scheme in enhancing the performance and efficiency of multibeam satellite communications. By addressing the impact

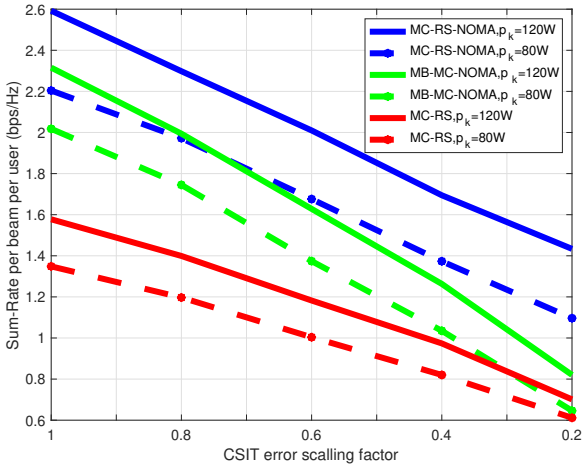



Figure 6–7: Sum-rate performance versus CSIT uncertainty ($\eta$), $M = 2, I = 28$ users

of CSIT uncertainty and considering practical constraints, the proposed approach offers significant improvements in terms of fairness and overall system capacity.

# CHAPTER 7
# Conclusion and Future Work

## 7.1 Conclusion

In conclusion, this thesis has focused on enhancing the performance of multibeam multicast satellite communication systems through the use of innovative frameworks and techniques. Three distinct frameworks, namely MC-NOMA, MB-MC-NOMA, and MC-RS-NOMA, were proposed and investigated to address the challenges posed by interbeam interference and imperfect channel state information.

In the MC-NOMA framework, the effectiveness of NOMA within a single beam in multicast transmission was extensively analyzed. The achievable data rates and power allocation methods were derived, and an optimal user clustering approach was proposed to maximize system performance. The results demonstrated the advantages of implementing NOMA in terms of improving the MMF rate and total achievable sum-rate.

The MB-MC-NOMA framework, which integrated linear precoding and NOMA techniques, aimed to further enhance system performance in a multibeam multicast scenario. The proposed user scheduling method improved the performance of both multicast linear precoding and the NOMA scheme. Multicast linear precoding techniques were designed, and different mappers were introduced to optimize the performance. Power allocation schemes were also proposed to maximize the MMF rate and weighted sum-rate. The achievable rate region of the MB-MC-NOMA scheme was derived, enabling efficient computation of the optimal transmit power spectral density for groups of users within multiple beams.

Finally, the MC-RS-NOMA framework was introduced to address the challenges of imperfect channel state information. Rate splitting was applied to cancel interbeam interference, and NOMA was considered on a beam basis to improve spectral efficiency. The achievable data rates of the

common and private parts were derived, and the optimization problems for MMF rate and sum-rate were formulated. The weighted minimum mean square error approach and alternating optimization were utilized to solve non-convex optimization problems.

Overall, the proposed frameworks and techniques in this thesis have shown promising results in improving the performance of multibeam multicast satellite communication systems. Theoretical analyses and extensive simulations have validated the effectiveness and benefits of the proposed approaches. The findings from this research contribute to the advancement of the field and provide valuable insights into the optimization of system performance in various CSIT scenarios.

## 7.2 Future Work

In this thesis, three frameworks have been proposed to enhance the throughput and performance of multibeam multicast satellite communication systems. While these frameworks have shown promising results, there are still several avenues for future research and development in this area.

Firstly, an interesting direction for future work is to explore L-layer Rate-splitting (RS). The focus of this thesis has been on the 1-layer RS technique, future work could delve into more than one layer L-RS frameworks, where $L > 1$. For example in 2-layer RS, the message of each user is split into three sub messages. In order to achieve greater flexibility in managing interference, the common streams are encoded into distinct layers. This extension would provide a deeper understanding of the benefits and trade-offs associated with L-layer RS schemes, enabling more flexible and efficient resource allocation strategies.

Also, it would be important to evaluate the complexity of the transmitter and receiver of the proposed frameworks. Lastly, it is essential to assess the Degree-of-Freedom (DoF) of the proposed frameworks and compare them with existing methods. The DoF, which is also referred to as the spatial multiplexing gain, provides a quantitative measure of how effectively the spatial dimension is utilized in a communication strategy. It represents the number or fraction of independent data streams that can be transmitted to a specific user. Calculating the DoF would provide insights

into the system's spectral efficiency and its ability to support multiple users simultaneously. This analysis would further validate the performance of the proposed frameworks and enable a comprehensive comparison with other techniques reported in the literature.

Overall, there is ample room for further research and development in the field of NOMA-based multiuser multicast systems, and the proposed frameworks can serve as a foundation for exploring these future directions.

### 7.3 Publications

The research conducted in this thesis has resulted in several publications that contribute to the field of multibeam multicast satellite communication systems. The following is a list of publications based on the results investigated in this thesis:

- S. M. Ivari, M. R. Soleymani, and Y. Shayan, "RS-Based MIMO-NOMA Systems in Multicast Framework" in "MIMO Communications - Fundamental Theory, Propagation Channels, and Antenna System", Dr. Ahmed Kishk and Dr. Xiaoming Chen,, 2023.
- S. M. Ivari, et al., On Optimal Power Allocation in Multibeam Multicast NOMA for Satellite Communication Systems, Published on IEEE Transactions on Aerospace and Electronic Systems.
- S. M. Ivari, et al., Precoding and Scheduling in Multibeam Multicast NOMA based Satellite Communication Systems, 2021 IEEE International Conference on Communications Workshops (ICC Workshops), 2021, pp. 1-6.
- S. M. Ivari, et al., Power Allocation and User Clustering in Multicast NOMA based Satellite Communication Systems," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), 2020, pp. 1-6.

References


[1] B. Evans, O. Onireti, T. Spathopoulos, and M. Imran, "The role of satellites in 5g," 09 2015.

[2] T. de Cola, A. Ginesi, G. Giambene, G. C. Polyzos, V. A. Siris, N. Fotiou, and Y. Thomas, "Network and protocol architectures for future satellite systems," vol. 12, pp. 1–161, 2017.

[3] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, J. Querol, L. Lei, T. X. Vu, and G. Goussetis, "Satellite communications in the new space era: A survey and future challenges," vol. 23, pp. 70–109, 2021.

[4] M. . Vzquez, A. Prez-Neira, D. Christopoulos, S. Chatzinotas, B. Ottersten, P.-D. Arapoglou, A. Ginesi, and G. Taricco, "Precoding in multibeam satellite communications: Present and future challenges," vol. 23, pp. 88–95, 2016.

[5] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup precoding and user scheduling for frame-based satellite communications," vol. 14, pp. 4695–4707, 2015.

[6] J. Wang, L. Zhou, K. Yang, X. Wang, and Y. Liu, "Multicast precoding for multigateway multibeam satellite systems with feeder link interference," vol. 18, pp. 1637–1650, 2019.

[7] A. I. Perez-Neira, M. A. Vazquez, M. B. Shankar, S. Maleki, and S. Chatzinotas, "Signal processing for high-throughput satellites: Challenges in new interference-limited scenarios," vol. 36, pp. 112–131, 2019.

[8] B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, and G. Caire, "Is NOMA efficient in multi-antenna networks? A critical look at next generation multiple access techniques," vol. abs/2101.04802, 2021.

[9] L. Cottatellucci, M. Debbah, G. Gallinaro, R. Mueller, M. Neri, and R. Rinaldo, "Interference mitigation techniques for broadband satellite systems," in *24th AIAA International Communications Satellite Systems Conference*, p. 5348, 2006.

[10] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," vol. 52, pp. 461–471, 2004.

[11] C.-B. Chae, D. Mazzarese, N. Jindal, and R. W. Heath, "Coordinated beamforming with limited feedback in the mimo broadcast channel," vol. 26, pp. 1505–1515, 2008.

[12] K. S. Gomadam, H. C. Papadopoulos, and C.-E. Sundberg, "Techniques for multi-user mimo with two-way training," in *2008 IEEE International Conference on Communications*, pp. 3360–3366, 2008.

[13] W. Wang, L. Gao, R. Ding, J. Lei, L. You, C. A. Chan, and X. Gao, "Resource efficiency optimization for robust beamforming in multi-beam satellite communications," vol. 70, pp. 6958–6968, 2021.

[14] C. Qi, H. Chen, Y. Deng, and A. Nallanathan, "Energy efficient multicast precoding for multiuser multibeam satellite communications," vol. 9, pp. 567–570, 2020.

[15] Y. Yan, W. Yang, B. Zhang, D. Guo, and G. Ding, "Outage constrained robust beamforming for sum rate maximization in multi-beam satellite systems," vol. 24, pp. 164–168, 2020.

[16] X. Zhang, J. Wang, C. Jiang, C. Yan, Y. Ren, and L. Hanzo, "Robust beamforming for multibeam satellite communication in the face of phase perturbations," vol. 68, pp. 3043–3047, 2019.

[17] E. Lagunas, V. N. Ha, T. V. Chien, S. Andrenacci, N. Mazzali, and S. Chatzinotas, "Multicast mmse-based precoded satellite systems: User scheduling and equivalent channel impact," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, pp. 1–6, 2022.

[18] G. Taricco, "Linear precoding methods for multi-beam broadband satellite systems," in *European Wireless 2014; 20th European Wireless Conference*, pp. 1–6, 2014.

[19] Y. C. B. Silva and A. Klein, "Linear transmit beamforming techniques for the multigroup multicast scenario," vol. 58, pp. 4353–4367, 2009.

[20] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Weighted fair multicast multigroup beamforming under per-antenna power constraints," vol. 62, pp. 5132–5142, 2014.

[21] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup precoding and user scheduling for frame-based satellite communications," vol. 14, pp. 4695–4707, 2015.

[22] V. Joroughi, M. . Vzquez, and A. I. Prez-Neira, "Generalized multicast multibeam precoding for satellite communications," vol. 16, pp. 952–966, 2017.

[23] I. Budhiraja, N. Kumar, S. Tyagi, S. Tanwar, Z. Han, M. J. Piran, and D. Y. Suh, "A systematic review on noma variants for 5g and beyond," vol. 9, pp. 85573–85644, 2021.

[24] A. I. Perez-Neira, M. Caus, M. A. Vazquez, and N. Alagha, "Noma schemes for multibeam satellite communications," 2018.

[25] A. I. Perez-Neira, M. Caus, and M. A. Vazquez, "Non-orthogonal transmission techniques for multibeam satellite systems," vol. 57, pp. 58–63, 2019.

[26] Z. Ding, F. Adachi, and H. V. Poor, "The application of mimo to non-orthogonal multiple access," vol. 15, pp. 537–552, 2016.

[27] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5g systems," vol. 22, pp. 1647–1651, 2015.

[28] L. Yang, J. Chen, Q. Ni, J. Shi, and X. Xue, "Noma-enabled cooperative unicastmulticast: Design and outage analysis," vol. 16, pp. 7870–7889, 2017.

[29] M. Chen and S. Li, "Power allocation for noma based layered multicast transmission," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, pp. 678–682, 2018.

[30] Y. Zhang, X. Wang, D. Wang, Y. Zhang, Q. Zhao, and Q. Deng, "Noma-based cooperative opportunistic multicast transmission scheme for two multicast groups: Relay selection and performance analysis," vol. 6, pp. 62793–62805, 2018.

[31] M. Caus, M. . Vzquez, and A. Prez-Neira, "Noma and interference limited satellite scenarios," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, pp. 497–501, 2016.

[32] N. A. K. Beigi and M. R. Soleymani, "Interference management using cooperative noma in multi-beam satellite systems," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2018.

[33] A. Wang, L. Lei, E. Lagunas, A. I. Prez Neira, S. Chatzinotas, and B. Ottersten, "On fairness optimization for noma-enabled multi-beam satellite systems," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, 2019.

[34] Y. Zhu, T. Delamotte, and A. Knopp, "Geographical noma-beamforming in multi-beam satellite-based internet of things," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2019.

[35] X. Yan, K. An, T. Liang, G. Zheng, Z. Ding, S. Chatzinotas, and Y. Liu, "The application of power-domain non-orthogonal multiple access in satellite communication networks," vol. 7, pp. 63531–63539, 2019.

[36] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-splitting multiple access: Fundamentals, survey, and future research trends," IEEE, 2022.

[37] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming sdma and noma," vol. 2018, pp. 1–54, Springer, 2018.

[38] B. Clerckx, Y. Mao, E. A. Jorswieck, J. Yuan, D. J. Love, E. Erkip, and D. Niyato, "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," vol. 41, pp. 1265–1308, 2023.

[39] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser miso systems with partial csit: A rate-splitting approach," vol. 64, pp. 4847–4861, 2016.

[40] B. Lee and W. Shin, "Max-min fairness precoder design for rate-splitting multiple access: Impact of imperfect channel knowledge," vol. 72, pp. 1355–1359, 2023.

[41] H. Joudeh and B. Clerckx, "Rate-splitting for max-min fair multigroup multicast beamforming in overloaded systems," vol. 16, pp. 7276–7289, 2017.

[42] M. Caus, A. Pastore, M. Navarro, T. Ramirez, C. Mosquera, N. Noels, N. Alagha, and A. I. Perez-Neira, "Exploratory analysis of superposition coding and rate splitting for multibeam satellite systems," in *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 1–5, 2018.

[43] L. Yin and B. Clerckx, "Rate-splitting multiple access for multibeam satellite communications," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2020.

[44] L. Yin and B. Clerckx, "Rate-splitting multiple access for multigroup multicast and multibeam satellite systems," vol. 69, pp. 976–990, 2021.

[45] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006.

[46] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[47] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by lapidoth, shamai and wigger on the collapse of degrees of freedom under finite precision CSIT," *CoRR*, vol. abs/1403.1541, 2014.

[48] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[49] A. Lapidoth, S. Shamai, and M. A. Wigger, "On the capacity of fading mimo broadcast channels with imperfect transmitter side-information," *ArXiv*, vol. abs/cs/0605079, 2006.

[50] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.

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

[52] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3417–3432, 2005.

[53] Y. Sun, Y. Yang, A. D. Liveris, V. Stankovic, and Z. Xiong, "Near-capacity dirty-paper code design: A source-channel coding approach," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3013–3031, 2009.

[54] M. . Vzquez and A. I. Prez-Neira, "Spectral clustering for beam-free satellite communications," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1030–1034, 2018.

[55] R. Cheng and S. Verdu, "Gaussian multiaccess channels with isi: capacity region and multiuser water-filling," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 773–785, 1993.

[56] C. Zeng, L. Hoo, and J. Cioffi, "Optimal water-filling algorithms for a gaussian multiaccess channel with intersymbol interference," in *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No.01CH37240)*, vol. 8, pp. 2421–2427 vol.8, 2001.

[57] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[58] T. L. Marzetta, "How much training is required for multiuser mimo?," in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, pp. 359–363, 2006.

[59] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, 2008.

[60] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser mimo achievable rates with downlink training and channel state feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2845–2866, 2010.

[61] N.-P. Nguyen, M. Zeng, O. A. Dobre, and H. V. Poor, "Securing massive mimo-noma networks with zf beamforming and artificial noise," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2019.

[62] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser miso systems: A rate-splitting approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6227–6242, 2016.

[63] N. Jindal, "Mimo broadcast channels with finite rate feedback," in *GLOBECOM '05. IEEE Global Telecommunications Conference, 2005.*, vol. 3, pp. 5 pp.–, 2005.

[64] C. Hao and B. Clerckx, "Degrees-of-freedom region of time correlated MISO broadcast channel with perfect delayed CSIT and asymmetric partial current CSIT," *CoRR*, vol. abs/1211.4381, 2012.

[65] C. Loo, "A statistical model for a land mobile satellite link," *IEEE Transactions on Vehicular Technology*, vol. 34, no. 3, pp. 122–127, 1985.

[66] F. Fontan, M. Vazquez-Castro, C. Cabado, J. Garcia, and E. Kubista, "Statistical modeling of the lms channel," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 6, pp. 1549–1567, 2001.

[67] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2744–2757, 2017.

[68] J. Cui, Z. Ding, and P. Fan, "A novel power allocation scheme under outage constraints in noma systems," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1226–1230, 2016.

[69] J. Choi, "Power allocation for max-sum rate and max-min rate proportional fairness in noma," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2055–2058, 2016.

[70] B. Kimy, S. Lim, H. Kim, S. Suh, J. Kwun, S. Choi, C. Lee, S. Lee, and D. Hong, "Non-orthogonal multiple access in a downlink multiuser beamforming system," in *MILCOM 2013 - 2013 IEEE Military Communications Conference*, pp. 1278–1283, 2013.

[71] T. de Cola, A. Ginesi, G. Giambene, G. C. Polyzos, V. A. Siris, N. Fotiou, and Y. Thomas, "Network and protocol architectures for future satellite systems," *Foundations and Trends in Networking*, vol. 12, no. 1-2, pp. 1–161, 2017.

[72] S. Rougerie, F. Lacoste, and B. Montenegro-Villacieros, "Mobile satellite propagation channels for ku and ka band," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, 2016.

[73] A. Al-Hourani and I. Guvenc, "On modeling satellite-to-ground path-loss in urban environments," *IEEE Communications Letters*, vol. 25, no. 3, pp. 696–700, 2021.

[74] G. Zheng, P.-D. Arapoglou, and B. Ottersten, "Physical layer security in multibeam satellite systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 852–863, 2012.

[75] Y. Liu, H. Xing, C. Pan, A. Nallanathan, M. Elkashlan, and L. Hanzo, "Multiple-antenna-assisted non-orthogonal multiple access," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 17–23, 2018.

[76] X. Wang and H. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded cdma," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, 1999.

[77] M. J. Hagh and M. R. Soleymani, "Raptor coding for non-orthogonal multiple access channels," in *2011 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2011.

[78] N. Mokari, P. Hajipour, L. Mohammadi, Z. Ghattan, and P. Sojoodi, "Resource allocation for non-delay-sensitive satellite services using adaptive coding and modulationmultiple-input and multiple-outputorthogonal frequency division multiplexing," *IET Communications*, vol. 10, 01 2016.

[79] K. Shen and W. Yu, "A coordinated uplink scheduling and power control algorithm for multicell networks," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, pp. 1305–1309, 2015.

[80] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of gaussian multiple-access and broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 768–783, 2004.

[81] N. Hassan and X. Fernando, "Massive mimo wireless networks: An overview," *Electronics*, vol. 6, no. 3, 2017.

[82] C. Hao, Y. Wu, and B. Clerckx, "Rate analysis of two-receiver miso broadcast channel with finite rate feedback: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 63, 06 2015.

[83] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexingthe large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[84] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, 2012.

[85] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A rate splitting strategy for massive mimo with imperfect csit," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4611–4624, 2016.

[86] Z. Xiang, M. Tao, and X. Wang, "Massive mimo multicasting in noncooperative cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1180–1193, 2014.

[87] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted mmse for mimo-bc beamforming design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, 2008.