UNIVERSITÉ DU LUXEMBOURG

PhD-FSTM-2024-050
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 11/07/2024 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN *INFORMATIQUE*

by

Lin CHEN
Born on 28 May 1990 in Hubei, (China)

DYNAMIC RESOURCE MANAGEMENT OPTIMIZATION FOR FLEXIBLE SATELLITE PAYLOADS

Dissertation defence committee

Dr Eva LAGUNAS, dissertation supervisor
Research Scientist, Université du Luxembourg

Dr Symeon CHATZIONTAS, Vice Chairman
Professor, Université du Luxembourg

Dr Bhavani Shankar MYSORE, Chairman
Assistant Professor, Université du Luxembourg

Dr Carlos MOSQUERA
Professor, University of Vigo, Spain

Dr Carlos-Faouzi BADER
Professor, Technology Innovation Institute, United Arab Emirates

"Wanderer above the Sea of Fog" – Caspar David Friedrich

"Life is a journey, there is always a way!"

Acknowledgements

I would like to thank myself who have taken me here.

I would also like to express my sincere gratitude to all who have helped me during my journey towards the Doctor of Philosophy. I remember and appreciate everything you have done for me.

This work was supported by the Luxembourg National Research Fund (FNR) under the project "FlexSAT: Resource Optimization for Next Generation of Flexible SATellite Payloads" (C19/IS/13696663).

Abstract

This dissertation addresses the topic of dynamic resource management optimization for flexible satellite payloads. The capability to flexibly allocate on-board resources over the service coverage is becoming a must for future broadband satellite systems. The trend for future satellite flexible payloads is to assign resources in an intelligent manner according to the heterogeneous traffic demands. In particular, this thesis focuses on Beam Hopping (BH)-enabled systems, where a subset of beams can be illuminated at a given time. Conventional BH illumination pattern design provides all available spectrum to a selected set of beams as long as they are not adjacent to each other (to avoid inter-beam interference). In this thesis, advanced BH illumination pattern designs are explored and assessed. First, we address the BH design for geostationary satellites and target the shortcomings of adjacent beam avoidance requirements of conventional BH illumination designs. In particular, we propose a dynamic beam illumination scheme combined with selective precoding, where only sub-sets of beams that are subject to strong inter-beam interference implement interference mitigation techniques like precoding. The formulated binary quadratic programming (BQP) problem is proved to converge to a local optimum solution. Next, we propose a two-stage framework with a probabilistic perspective that exploits flexibility in the time and power domains at the same time. The first stage addresses the coupling relationship between power and beam activation probability. Conditioned on the optimal solution, we reformulate the problem and prove its convexity, and lastly propose a method for solving it iteratively. The second stage designs the detailed beam illumination pattern by mapping the beam activation probability obtained in the first stage. Last but not least, we address the BH illumination design problem for Lower Earth Orbit (LEO) satellite constellation systems, where a virtual cell on the ground can be served by multiple beams belonging to multiple satellites. As a consequence, we extend the flexibility in the time and power domains with the flexibility of serving a cell with different satellites. This additional degree of flexibility is shown to minimize transmission energy consumption. More precisely, we address the demand satisfaction constraint with the load coupling model, where parameters are given by their expectations. Based on the optimum condition, the three-variable problem is reformulated to an inverse matrix optimization problem with two variables. We prove the convexity of the objective and propose an iterative algorithm to solve the problem.

Table of Contents

List of Figures

List of Tables

I Introduction

1 Background and Motivation

The roll-out of the next generation of wireless communication systems is expected to deliver faster internet access and increased capacity, providing customized services in a variety of use cases [2]. Despite the global growth of digital technologies, the United Nations (UN) has recently announced in the General Assembly that half of the world's population still has no internet access [3]. The reason behind this is the fact that deployment of fiber and general terrestrial infrastructure is not feasible (or is not worth the investment) in remote areas, or the ground equipment is with high probability subject to disruption by man-made events.

Satellite communication (SatCom) systems have been identified as a key solution to provide connectivity everywhere. It has the advantage of wide coverage and supporting terminals located in remote areas, making them a vital component of global communication infrastructure [4]. As a key component of SatCom, conventional High Throughput Satellite (HTS) systems have employed multi-beam technology, where satellites are equipped with antenna architectures able to form high-directional spot-beam serving terminals located in geographical virtual cells [5]. However, the provided capacity is equally distributed across the multiple beams, which generates a waste of limited onboard resources considering the heterogeneous traffic demands of beams [6]. Therefore, it is of great interest to develop fully reconfigurable resource allocation schemes that can smartly provide capacity matching the required traffic demands of beams so as to improve the system's cost efficiency.

The recent advances in space technology have opened the door to unprecedented flexibility and adaptability to on-board satellite resources. As highlighted by the major satellite industry experts in Europe [7], "The continuous development of new technologies and the huge increase in satellite interest and investment, witnessed in recent time, have indeed pushed the satellite communication potentialities towards higher limits that now need to be explored to support the efficient and sustainable development of new markets and smart services". As pointed out, making full flexibility of onboard resources has been identified as the major challenge that needs to be resolved to unleash the potential of next-generation SatCom systems [7].

The most advanced GEO HTS systems are claimed to be reconfigurable. One of the flagship flexible HTS satellites, the so-called Eutelsat Quantum, developed under an ESA Partnership Project with the satellite operator Eutelsat and the prime manufacturer Airbus, was launched in July 2021. Eutelsat Quantum is claimed to be the first commercial fully flexible software-defined satellite in the world [8]. Coverage, spectrum, and capacity all can be reconfigured in orbit via its innovative reconfigurable payloads, which can efficiently serve any application. According to the technical capabilities of the Eutelsat Quantum [9], spot beams can be hopped to spatially diverse regions rapidly and seamlessly. With a similar vision in mind, SES, the satellite operator, worked with Thales Alenia Space to manufacture SES-17. The satellite was launched in October 2021 and incorporates a digital transparent processor (DTP), enabling unique features, such as reconfigurable resource allocation, to meet real-time traffic demands [10]. Additionally, most of the industry-led projects are still in the testing phase, where the algorithms to unleash the potential of the system are still in the early stages. For instance, the European Commission has recently launched two 3-year projects related to optimal on-board resource management [11, 12].

Beam Hopping (BH) is one of the technologies that enable the systems to allocate onboard resources flexibly. Unlike traditional fixed-beam systems, the BH-enabled systems only activate a subset of spot beams within a designated period, known as "dwell time". These active beams can be changed dynamically, following the illumination pattern which determines the status of a beam at a time slot for all beams during the time window and is designed according to the required uneven traffic demands. BH offers numerous benefits, among which the most significant is that it offers flexibility in the

time and power domains to hop beams, and thus provides capacity to different beams as needed. Moreover, activating fewer beams at any given time requires fewer onboard Radio Frequency (RF) chains, which can reduce the spacecraft's weight and size. Ultimately, this reduction may result in decreased launch costs.

BH-enabled SatCom technology has reached a level of maturity, significantly bolstered by industry support. This support has led to adaptations, such as the standardization of satellite air interfaces, to accommodate BH synchronization needs [13, 14]. An example of this technology in action is the Eutelsat Quantum satellite, which employs forward link BH. While the technical capability for beam illumination and configuration exists, the effective management of these functions remains an active area of research and development. This is mainly because crafting an optimal beam time-activation plan and allocating power effectively is a complex task. Specifically, it is challenging to properly allocate the time and the limited power across the active beams to meet the dynamic demands while satisfying the satellite payload hardware limitations.

BH technology offers the operator natural flexibility of onboard resources, i.e. the allocation of time and power of beams, for GEO SatCom systems, where the virtual cells are usually one-to-one linked with the radiated beams when it comes to serving information. While a virtual cell could be served by multiple beams within a Low Earth Orbit (LEO) satellite constellation system, each being linked to one satellite, which provides extra flexibility in allocating required loads of beams serving the same cell [15]. LEO satellite constellation systems have developed rapidly recently, such as Starlink and Eutelsat OneWeb. Usually, the systems consist of thousands of satellites, and each satellite is designed to provide communication service to users located in its coverage. Moreover, to guarantee global coverage, the coverage area of adjacent satellites has significant overlap [16].

Furthermore, precoding is another technology that increases the systems' throughput. Similar to the situation in the terrestrial domain, the rapid development of data-hungry services has also resulted in a spectrum scarcity context in the satellite domain. As a consequence, the satellite digital broadcasting standard (DVB-S2X) introduced the possibility of using the channel state information feedback to implement precoding techniques and enable efficient spectrum management while increasing the spectrum reuse across satellite spot beams [17, 18]. The feasibility and potential of precoding applied

to HTS systems have been recently validated via live experiments on a GEO satellite scenario [19], confirming its relevance for future HTS deployments.

It is worth highlighting that SatCom systems should carefully manage energy consumption. Satellite payloads are equipped with solar panels that convert solar energy into electricity, which powers wireless transmission and other operations [20]. Given that these panels can only collect a limited amount of power, it is crucial for these systems to address the issue of energy efficiency. Additionally, excessive energy consumption can have adverse effects on the payload's mass and lifespan [21]. As a consequence, the minimization of transmit power has a significant impact on general non-terrestrial communication platforms, e.g. [22].

In recent years, flexible resource allocation strategies to match uneven traffic demands of user terminals for SatCom systems have attracted considerable attention. In [23], the authors propose allocating power based on traffic demands and channel conditions. An energy-aware power allocation problem for non-BH systems is formulated in the work [21], which aims to minimize both unmet system capacity and total radiated power. Both [23] and [21] consider only the flexibility of onboard resources in the power domain. The benefits of the application of BH technology to GEO SatCom systems have also been demonstrated in academic works, which explore flexibility in the time domain (in general). Conventional BH was conceived to exploit the full available spectrum (i.e. full frequency reuse) over a subset of selected beams, ensuring that the geographical distance between selected beams is far enough to work under a noise-limited scenario [24]. The design of illumination pattern with conventional BH usually involves binary variables. Therefore, the formulated problem falls within the general mixed integer non-linear programming problem [25], which is difficult to solve. To increase the systems' throughput, the activation of an adjacent set of beams (referred to as a cluster) was investigated within the European Space Agency (ESA) [26] and proposed in [27, 28], with the so-called Cluster Hopping scheme, where linear precoding [29] was considered to mitigate the intra-beam interference. However, there exists a trade-off between the diversity of available clusters and computational costs, limiting the performance of this method. Therefore, it is worth modeling a general framework that releases the full potential of joint design of precoding and flexibility of onboard resources including the time-domain illumination pattern and the transmit power.

Moreover, power allocation is another domain that provides flexibility in resource allocation, which has been evaluated for non-beam-hopping SatCom systems in the literature. In [23], the authors proposed allocating power based on traffic demands and channel conditions. An energy-aware power allocation problem for non-BH systems is formulated in the work [21], which aims to minimize both unmet system capacity and total radiated power. Although [30] proposed jointly allocating power and frequency carriers to minimize the weighted objective of energy consumption and frequency occupation, the proposed design does not fully utilize the precious spectrum resources. Consequently, it would be interesting to exploit the full flexibility in the time and power domains for the communication systems, which are not only able to reuse the full frequency but also provide potential in energy consumption. Additionally, The load coupling model, originally analyzed by [31], characterizes the coupling relationships among the temporal occupation of beams. The model has been employed in [32], where the power and temporal load of beams are studied for cellular networks to minimize energy consumption. However, [32] does not consider the constraint on the maximal number of simultaneously active beams, where all beams are available all the time. While there is a constraint of the maximal number of simultaneously active beams for SatCom systems, the scenario in [32] is a special case of the one that needs to be solved. Therefore, it is interesting to exploit the potential of joint design of flexibility in the time and power domain with the load coupling model subject to the practical constraint for the BH-enabled SatCom system.

Furthermore, the methods proposed by the above-mentioned works focus on GEO SatCom systems. In the literature, some works address the resource allocation problem for LEO SatCom, e.g. [33, 34, 35]. [33] jointly optimize the power, time, and bandwidth to maximize the throughput for multiple satellites. [34] summarizes heuristic algorithms for the resource allocation problem for LEO SatCom. Additionally, [35] proposes to minimize the latency of the systems. [36] propose to integrate satellite with base stations to provide communication service. However, these works either consider the energy consumption problem or fail to utilize the extra flexibility that an on-ground virtual cell being served by multiple beams, each per satellite. Although [15] proposes a method utilizing this flexibility, the proposed method is conditioned on balancing the loads of each satellite, which lacks theoretical verification. Moreover, as far as the au-

thors' knowledge, no paper has simultaneously utilized the full flexibility (time, power, required loads of beams) for the LEO SatCom systems.

2 Objectives

In order to take full advantage of the flexibility of available resources onboard in improving computational cost and energy efficiencies, this thesis proposes new techniques for beam-hopping-enabled satellite communication systems, including illumination pattern design algorithms and joint design of power and time optimization framework for GEO satellite system, and joint optimization of power, time and required loads of beams for LEO satellite constellation systems. Specifically, the objectives of our research include:

- **Objective 1:** Considering that large geographical areas covered by GEO satellites may aggregate high-traffic demands, thus involving more than one adjacent beam to be simultaneously activated, we consider the design of the temporal illumination pattern to meet heterogeneous traffic demands of beams, where precoding technique is employed to mitigate inter-beam interference only for those beams that subject to strong interference. By utilizing the precoding technique, interference from other beams is kept within an acceptable level, which helps to increase the spectral and energy efficiencies of the system.

- **Objective 2:** Making full advantage of flexibility in both time and power domains of onboard resources so as to minimize the total transmission energy consumption while satisfying both the uneven traffic demands and limited hardware constraints for GEO satellite communication systems.

- **Objective 3:** Taking into account that on-ground virtual cells can be served by multiple LEO satellites, besides the flexibility in the time and power domains, it is essential to optimize the flexibility in balancing the load across satellites and beams that are providing coverage to the same on-ground virtual cell. This optimization aims to minimize the total energy consumption while meeting the heterogeneous traffic demands of cells for LEO satellite constellation systems.

3 Major Contributions of the Thesis

The main contributions of the thesis are as follows:

- **To address Objective 1:** We propose a general framework and its mathematical model to support dynamic and flexible cluster hopping for beam-hopping-enabled GEO satellite systems, where geographically adjacent beams are allowed to be activated simultaneously whenever needed according to the required traffic demands of beams. The optimization aims to minimize the usage of precoding while meeting both uneven demands satisfaction of beams and limited hardware of payload constraints. First, we propose a novel method to approximate the demands in the number of time slots, then the formulated binary quadratic programming (BQP) problem is efficiently addressed by three methodologies, among which the proposed multiplier penalty and majorization minimization (MPMM) method is proved to converge to a local optimum. Compared with the conventional cluster hopping method, the proposed framework efficiently reduces the required computational cost for precoding while satisfying the constraints.

- **To address Objective 2:** We propose a two-stage framework to jointly design power allocation and beam scheduling for beam-hopping-enabled GEO satellite systems. Within the framework, by utilizing the mean-field theory, a probabilistic reformulation becomes available, paving the way for addressing the intrinsic coupling between beam power and beam illumination pattern. Besides, for the reformulated problem, we analyze the optimality conditions and develop an iterative method to yield an optimal solution. We also develop a systematic mapping scheme that converts the probabilistic solution obtained in the first stage into a deterministic one satisfying all constraints. To ensure the practicality of the solution, beam hopping latency is also integrated into the scheme. Numerical simulation results validate the theoretical findings: i) the system consumes minimal energy when all available beams are active during the time window; ii) increasing the maximal number of active beams could reduce energy consumption. Furthermore, given that our mathematical model is based on the Shannon formula, we acknowledge a potential performance loss due to coding modulation methods in real-world applications. To address this concern, our study includes a method

specifically designed to compensate for such performance loss, ensuring the completeness and applicability of our approach in practical scenarios.

- **To address Objective 3:** We propose a joint design of power allocation, beam scheduling, and required load assignments for beam hopping-enabled LEO satellite constellation systems. Employing the load coupling model, the energy minimization problem is formulated from a probabilistic perspective. We analyze the demand satisfaction constraint, which provides the condition to prove the convexity of the objective. Moreover, we analyze the impact of systematic parameters on the problem's feasibility and performance and validate with numerical simulations the theoretical findings: i) compared with fixed loads of beams, the flexibility in assigning loads reduces energy consumption; ii) the proposed method flexibly assigns loads of beams making the infeasible problem feasible with fixed assignment of loads.

4 Organization of the Thesis

This is a cumulative thesis that constraints 3 published or submitted journal papers. The following of the thesis is structured as follows. Chapter II summarizes the work appended in Chapter VIII, which introduces the methods of flexible resource allocation in the time domain for GEO satellite systems. In Chapter III, we introduce the methods of joint optimization of resources in time and power domains for GEO satellite systems, which are detailed in the work appended in Chapter IX. Following, Chapter IV introduces methods utilizing full flexibility in the time, power, and load domains of onboard resources for LEO satellite constellation systems, as detailed in the work appended in Chapter X. Finally, the thesis is concluded, and future research is recommended in Chapter V.

II Flexible Resource Allocation in Time Domain for Beam Hopping-Enabled GEO Satellite Communication Systems

In this chapter, we explore the flexibility in the time domain of resource allocation for Beam Hopping-enabled GEO satellite communication systems. In addition, to mitigate the strong interference from adjacent beams, we utilize the precoding technique. Specifically, we propose a general framework for the illumination pattern design, where the transmission of activated beams in separating clusters can be jointly precoded. The objective is to minimize the interference-based penalty with the aim of reducing the use of precoding while constraining the system to satisfy a certain beam demand in a given time window. Such technical design can be stated into a Binary Quadratic Programming. Whenever high demand expands over multiple adjacent beams, the solution from the proposed framework considers precoding to deal with the resulting inter-beam interference. With such a selective precoding mechanism, complexity at the ground segment is reduced where precoding operation can be considered flexible.

To linearize the BQP problem, we introduce a procedure to convert the beam demand constraint into a more tractable notation involving required illuminated time slots per beam. Next, we present different ways to convexify the BQP problem. Inspired by methods proposed in [37, 38], we reformulate the BQP into a Semi-Definite Programming (SDP) form which can be solved efficiently by employing some standard optimization solver tools. We also propose a novel solving framework MPMM which integrates

Multiplier Penalty (MP) [39] and Majorization-Minimization (MM) [40] methods. In particular, we convexify the problem by penalizing the binary constraint, where the penalty is given by its augmented Lagrangian function. Subsequently, the penalized-form problem is solved in a sequence, which drives the solution to binary values and guarantees to approach the local optimum. Since the previously proposed methods prioritize performance versus computational complexity, we complement this paper by proposing a heuristic greedy algorithm that provides a sub-optimal but efficient solution. Please refer to Chapter VIII for more detail.

III Flexible Resource Allocation in Time and Power Domains for Beam Hopping-Enabled GEO Satellite Communication Systems

In Chapter II, we investigate the methods that flexibly allocate the resources in the time domain for GEO satellite communication systems. In this chapter, besides the time, we also consider the flexibility in the power domain such that energy consumption is minimized.

To address the challenges, we propose a two-stage framework to solve the combinatorial non-convex problem. In the first stage, utilizing the mean-field theory [41, 42], we reinterpret the beam illumination patterns with beam activation probabilities. Consequently, the corresponding power will be the average for the whole time window. Accordingly, the original problem is reformulated into a new one, where the average powers and activation probabilities of the beams are the optimization variables. An iterative method is also proposed to solve this problem optimally. In the second stage, the activation probabilities of the solution obtained in the first stage are mapped into the discrete beam illumination patterns by solving a binary quadratic programming problem. Please refer to Chapter IX for more detail.

IV Flexible Resource Allocation in Time and Power and Load Domains for Beam Hopping-Enabled LEO Satellite Communication Systems

In the previous chapters, we have made full flexibility to allocate resources onboard for GEO satellite communication systems. In this chapter, we focus on LEO satellite constellation communication systems. In addition to examining time and power considerations, we investigate the methods that flexibly allocate loads of beams serving the same virtual on-ground cell, the feature of LEO systems, to minimize energy consumption.

To tackle the challenges, first, we formulate an ideal illumination pattern design problem, within which we employ the load coupling model to deal with the complicated demand-matching constraints, within which the three variables are coupled. To address the formulated problem, we find a necessary condition for the optimum, based on which we find that power is a function of activation probability and the required loads of beams. The reformulated problem is solved by the proposed iterative method. Second, the obtained activation probability is mapped into the illumination pattern by solving binary quadratic programming. Please refer to Chapter X for more detail.

V | Conclusion

1 Summary

In this dissertation, we have focused on the time flexibility provided by Beam-Hopping in high throughput satellite systems. In particular, this thesis addressed the resource management design, including beam illumination pattern, power allocation and load balancing between multiple satellites. Theoretical analysis and numerical simulations have demonstrated the effectiveness of the proposed discrete illumination pattern design method and the two-stage framework with a probabilistic perspective for GEO satellite systems, and the full flexibility optimization for the LEO satellite constellation systems. The main contributions of the dissertation are summarized as follows:

- We propose an analytical framework, a class of binary quadratic programming problems, to support dynamic beam illumination design considering selective precoding for the next generation of time-flexible GEO satellite broadband systems. Three algorithms are proposed to solve the problem: (i) SDP-based approach, (ii) MPMM methodology, and (iii) low-complexity greedy algorithm. All three methods target cross-beam interference minimization, such that the number of beams that need to be precoded is kept to a minimum in an attempt to reduce system complexity. An extensive evaluation has been carried out based on numerical simulations. The results have shown interesting gains provided by the proposed algorithm with respect to the relevant benchmark schemes. In particular, the proposed framework provides an efficient solution to deal with high-demand areas while keeping the precoding-related complexity low.

- We propose a novel two-stage framework for optimizing energy consumption through the joint design of power and time slot allocation. The framework is designed to achieve optimal performance by addressing various challenges. In the first stage, we utilize mean-field theory to extract the activation probability and reformulate the mathematical model into an inverse matrix optimization problem. This reformulation enables us to convert the problem into a convex form, which has been thoughtfully analyzed and solved efficiently using a proposed iterative method. In the second stage, we employ the MPMM method to map the activation probability into the illumination pattern. Additionally, we introduce a compensation method to mitigate the performance loss resulting from the discrepancy between practical adaptive coding modulation and the ideal Shannon formula. Overall, this step yields a deterministic and practical solution for the considered beam-hopping satellite system. To validate our theoretical findings, we conduct numerical simulations. The results demonstrate that our proposed method surpasses the benchmark in terms of energy consumption and demand-matching performance.

- We propose a method to jointly optimize power, required loads, and time slot allocation to minimize total energy consumption while satisfying the traffic demands of cells in LEO constellation satellite systems. There are two difficulties within the problem: i)satisfaction of uneven traffic demands of beams, ii)and binary illumination pattern design. We propose to address it by decomposing it into two problems. First, we employ the load coupling model to simplify the demand satisfaction constraint, where parameters are represented with their expectations. Subsequently, the problem is reformulated into an inverse matrix optimization form, which is approximated by a convex problem. We propose an iterative method to obtain a high-quality solution to the original problem. With the power, activation probability, and the required loads of beams obtained from the first problem, we employ the MPMM method to design the practical illumination pattern. Moreover, we propose demand compensation methods to eliminate systematic performance loss caused by the gap between the Shannon formula and the DVB-S2X standard. Finally, simulations validate our theoretical findings and prove that the proposed method outperforms the baseline.

2 Limitations and Outlook

The limitations of the proposed algorithms are twofold. First, the proposed algorithms are conditioned on the precise given traffic demands of beams (cells), which, however, usually have to be predicted in advance in practice. Thus, the demand-matching performance is also based on the precise traffic prediction algorithm. The other limitation of the proposed algorithms comes from the complexity of the solving methods when it comes to real-time practical application for large-scale systems. All three chapters employ the MPMM method to design the discrete illumination pattern, which is a double-loop iterative method. The computational complexity of the inner loop is $\mathcal{O}((MN)^3)$, and there is a trade-off between the systematic performance and convergence condition (the computational time). In this initial work, we have aimed at achieving the best performance at the cost of complexity, to help us understand the achievable gains.

Although there are limitations, both limitations have the potential to be addressed through mature learning-based methods. First, the learning-based methods have shown their advantages in extracting the hidden structure of historical data, with which one could make good predictions. Second, although the proposed methods may not suit real-time applications, they could be accelerated with the use of machine learning. In particular, a learning model can be trained to mimic the performance of the optimization-based methods presented in this thesis. Once the model is trained (which can be done offline), the trained model can be used to provide the resource allocation outcomes. Using machine learning for complex optimization acceleration is well-known in the literature and has shown great potential.

3 Acknowledgement of Previous and Related Work

I would like to acknowledge the previous work [27] which is titled "*Precoded Cluster Hopping in Multi-Beam High Throughput Satellite Systems*". Based on the innovative idea proposed by this paper, I have started exploring the potential of the idea and proposed methods presented in this dissertation.

VI | References

[1] **Chen, Lin** et al. "The next generation of beam hopping satellite systems: Dynamic beam illumination with selective precoding". In: *IEEE Transactions on Wireless Communications* 22.4 (2022), pp. 2666–2682.

[2] Marco Giordani et al. "Toward 6G networks: Use cases and technologies". In: *IEEE Communications Magazine* 58.3 (2020), pp. 55–61.

[3] United Nations. *Deputy secretary-general remarks at the general assembly*. https://www.un.org/press/en/2021/dsgsm1579.doc.htm. [Online; accessed 21-Dec-2021]. Apr. 2020.

[4] Oltjon Kodheli et al. "Satellite communications in the new space era: A survey and future challenges". In: *IEEE Communications Surveys & Tutorials* 23.1 (2020), pp. 70–109.

[5] Juan Andrés Vásquez-Peralvo et al. "Flexible beamforming for direct radiating arrays in satellite communications". In: *IEEE Access* (2023).

[6] Ana I Perez-Neira et al. "Signal processing for high-throughput satellites: Challenges in new interference-limited scenarios". In: *IEEE Signal Processing Magazine* 36.4 (2019), pp. 112–131.

[7] NetWorld 2020. *White Paper: SatCom Resources for Smart and Sustainable Networks and Services*. https://www.networld2020.eu/satcom-wg/. [Online; accessed 07-Feb-2021]. 2019.

[8] EUTELSAT. *High Throughput Satellite - Eutelsat KaSat*. http://www.eutelsat.com.

[9] Eutelsat Quantum. *Technical Capabilities of Eutelsat Quantum, more information available at:* https://issuu.com/eutelsat/docs/eutelsat_sat_quantum_brochure_1020. Accessed: 2022-06.

[10] *SES-17 satellite, more information available at:* https://www.ses.com/press-release/ses-17-successfully-launched-ariane-5. Accessed: 2022-06.

[11] H2020 DYNASAT project. *Dynamic Spectrum Sharing and Bandwidth-Efficient Techniques for High Throughput MIMO Satellite Systems*. https://www.dynasat.eu/. [Online; accessed Jun-2022]. Dec., 2020.

[12] H2020 ATRIA project. *AI Powered Ground Segment Control for Flexible Payloads*. https://www.atria-h2020.eu/. [Online; accessed Jun-2022]. 2020.

[13] A. Morello and N. Alagha. "DVB-S2X Air Interface Supporting Beam Hopping Systems". In: *25th Ka and Broadband Communications Conference, (Ka-2019), Sorrento, Italy*. Oct. 2019, pp. 1–5.

[14] Christian Rohde et al. "Beam-hopping system configuration and terminal synchronization schemes". In: *Proceedings of the 37th International Communications Satellite Systems Conference (ICSSC-2019)*. 2019, pp. 1–13. DOI: 10.1049/cp.2019.1229.

[15] Zhiyuan Lin et al. "Multi-satellite beam hopping based on load balancing and interference avoidance for NGSO satellite communication systems". In: *IEEE Transactions on Communications* 71.1 (2022), pp. 282–295.

[16] Oltjon Kodheli, Alessandro Guidotti, and Alessandro Vanelli-Coralli. "Integration of Satellites in 5G through LEO Constellations". In: *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE. 2017, pp. 1–6.

[17] Pantelis-Daniel Arapoglou et al. "DVB-S2X-enabled precoding for high throughput satellite systems". In: *International Journal of Satellite Communications and Networking* 34.3 (2016), pp. 439–455.

[18] Malek Khammassi, Abla Kammoun, and Mohamed-Slim Alouini. "Precoding for high throughput satellite communication systems: A survey". In: *IEEE Communications Surveys & Tutorials* (2023).

[19] Jevgenij Krivochiza et al. "End-to-end precoding validation over a live GEO satellite forward link". In: *IEEE Access* 11 (2021), pp. 41556–41564.

[20] Tobias Posielek. "An Energy Management Approach for Satellites". In: *69th International Astronautical Congress, Bremen, Germany*. Oct. 2018, pp. 1–5.

[21] Christos N Efrem and Athanasios D Panagopoulos. "Dynamic energy-efficient power allocation in multibeam satellite systems". In: *IEEE Wireless Communications Letters* 9.2 (2019), pp. 228–231.

[22] Yuan Yang et al. "Towards Energy-Efficient Routing in Satellite Networks". In: *IEEE Journal on Selected Areas in Communications* 34.12 (2016), pp. 3869–3886. DOI: 10.1109/JSAC.2016.2611860.

[23] Jihwan P Choi and Vincent WS Chan. "Optimum power and beam allocation based on traffic demands and channel conditions over satellite downlinks". In: *IEEE Transactions on Wireless Communications* 4.6 (2005), pp. 2983–2993.

[24] A. Freedman, D. Rainish, and Y. Gat. "Beam hopping how to make it possible". In: *Proc. of Ka and Broadband Communication Conference, Bologna, Italy*. Oct. 2015.

[25] Pietro Belotti et al. "Mixed-integer nonlinear optimization". In: *Acta Numerica* 22 (2013), pp. 1–131.

[26] ESA Project FlexPreDem. *Demonstrator of Precoding Techniques for Flexible Broadband Satellite Systems*. https://artes.esa.int/projects/flexpredem. [Online; accessed 21-Dec-2021]. 2020.

[27] Mirza Golam Kibria et al. "Precoded cluster hopping in multi-beam high throughput satellite systems". In: *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE. 2019, pp. 1–6.

[28] Eva Lagunas et al. "Precoded cluster hopping for multibeam GEO satellite communication systems". In: *Frontiers in Signal Processing* 1 (2021), p. 721682.

[29] Miguel Ángel Vázquez et al. "Precoding in multibeam satellite communications: Present and future challenges". In: *IEEE Wireless Communications* 23.6 (2016), pp. 88–95.

[30] Tedros Salih Abdu et al. "Flexible resource optimization for GEO multibeam satellite communication system". In: *IEEE Transactions on Wireless Communications* 20.12 (2021), pp. 7888–7902.

[31] Iana Siomina and Di Yuan. "Analysis of cell load coupling for LTE network planning and optimization". In: *IEEE Transactions on Wireless Communications* 11.6 (2012), pp. 2287–2297.

[32] Chin Keong Ho et al. "Power and load coupling in cellular networks for energy optimization". In: *IEEE Transactions on Wireless Communications* 14.1 (2014), pp. 509–519.

[33] Shengjun Guo et al. "An efficient multi-dimensional resource allocation mechanism for beam-hopping in LEO satellite network". In: *Sensors* 22.23 (2022), p. 9304.

[34] Yuanpeng Li et al. "Overview of beam hopping algorithms in large scale LEO satellite constellation". In: *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE. 2021, pp. 1345–1351.

[35] Leyi Lyu and Chenhao Qi. "Beam position and beam hopping design for LEO satellite communications". In: *China Communications* 20.7 (2023), pp. 29–42.

[36] Tong Li et al. "Pattern design and power management for cognitive LEO beaming hopping satellite-terrestrial networks". In: *IEEE Transactions on Cognitive Communications and Networking* (2023).

[37] Peng Wang et al. "Large-scale binary quadratic optimization using semidefinite relaxation and applications". In: *IEEE transactions on pattern analysis and machine intelligence* 39.3 (2016), pp. 470–485.

[38] Otto Nissfolk et al. "A mixed integer quadratic reformulation of the quadratic assignment problem with rank-1 matrix". In: *Computer Aided Chemical Engineering*. Vol. 31. Elsevier, 2012, pp. 360–364.

[39] Dimitri P Bertsekas. "On penalty and multiplier methods for constrained minimization". In: *SIAM Journal on Control and Optimization* 14.2 (1976), pp. 216–235.

[40] Linlong Wu, Prabhu Babu, and Daniel P Palomar. "Cognitive radar-based sequence design via SINR maximization". In: *IEEE Transactions on Signal Processing* 65.3 (2016), pp. 779–793.

[41] Manfred Opper and David Saad. *Advanced mean field methods: Theory and practice*. MIT press, 2001.

[42] Christopher M Bishop. “Pattern recognition and machine learning”. In: *Springer google schola* 2 (2006), pp. 1122–1128.

[43] Vaibhav Kumar Gupta et al. “Combining time-flexible geo satellite payload with precoding: The cluster hopping approach”. In: *IEEE Transactions on Vehicular Technology* 72.12 (2023), pp. 16508–16523.

[44] **Chen, Lin** et al. “Satellite broadband capacity-on-demand: Dynamic beam illumination with selective precoding”. In: *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE. 2021, pp. 900–904.

[45] **Chen, Lin** et al. “Adaptive Resource Allocation for Satellite Illumination Pattern Design”. In: *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*. IEEE. 2022, pp. 1–6.

VII | Appendix

1 Relevant publications

1.1. Articles in this dissertation

- Research Paper 1: The Next Generation of Beam Hopping Satellite Systems: Dynamic Beam Illumination with Selective Precoding.
- Research Paper 2: Joint Power Allocation and Beam Scheduling in Beam-Hopping Satellites: A Two-Stage Framework with a Probabilistic Perspective.
- Research Paper 3: Joint Time Slot, Power and Required Load Optimization for Energy-Efficient LEO Constellation Satellite Systems.

1.2. Other peer-reviewed publications not included in this dissertation

1.2.1. Journal publications

- Vaibhav Kumar Gupta, Vu Nguyen Ha, Eva Lagunas, Hayder Al-Hraishawi, **Chen, Lin**, and Symeon Chatzinotas. "Combining time-flexible geo satellite payload with precoding: The cluster hopping approach". In: *IEEE Transactions on Vehicular Technology* 72.12 (2023), pp. 16508–16523.

1.2.2. Conference publications

- **Chen, Lin**, Eva Lagunas, Symeon Chatzinotas, and Björn Ottersten. "Satellite broadband capacity-on-demand: Dynamic beam illumination with selective precoding". In: *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE. 2021, pp. 900–904.
- **Chen, Lin**, Eva Lagunas, Lei Lei, Symeon Chatzinotas, and Björn Ottersten. "Adaptive Resource Allocation for Satellite Illumination Pattern Design". In: *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*. IEEE. 2022, pp. 1–6.

2 Individual contribution to the included research papers

This thesis contains material from 2 papers published or under review in the following peer-reviewed journals, 1 journal paper is to be submitted soon, in which the student is listed as the main author.

Research Paper 1 is published as **Chen, Lin**, Vu Nguyen Ha, Eva Lagunas, Linlong Wu, Symeon Chatzinotas, and Björn Ottersten. "The next generation of beam hopping satellite systems: Dynamic beam illumination with selective precoding". In: *IEEE Transactions on Wireless Communications* 22.4 (2022), pp. 2666–2682. In this paper, my contributions are as follows:

- I formulated the problem to be addressed and discussed it with the co-authors.
- I proposed methods, conducted simulations, prepared the manuscript drafts, and revised the manuscript according to comments from the co-authors and reviewers.

Research Paper 2 is accepted as Lin Chen, Linlong Wu, Eva Lagunas, Anyue Wang, Lei Lei, Symeon Chatzinotas, Björn Ottersten. "Joint Power Allocation and Beam Scheduling in Beam-Hopping Satellites: A Two-Stage Framework with a Probabilistic Perspective", by *IEEE Transactions on Wireless Communications*. In this paper, my contributions are as follows:

- I formulated the problem to be addressed.
- I proposed methods, conducted simulations, prepared the manuscript drafts, and revised the manuscript according to comments from the co-authors and reviewers.

Research Paper 3 is a journal paper titled "Joint Time Slot, Power and Required Load Optimization for Energy-Efficient LEO Constellation Satellite Systems" which is to be submitted to *IEEE Transactions on Wireless Communications* in June. The authors are Lin Chen, Eva Lagunas, Symeon Chatzinotas. In this paper, my contributions are as follows:

- I formulated the problem to be addressed.
- I proposed methods, conducted simulations, prepared the manuscript drafts, and revised the manuscript according to comments from the co-authors.

VIII | Research Paper 1 – *Beam Hopping with Selective Precoding*

Full Title:
The next generation of beam hopping satellite systems: Dynamic beam illumination with selective precoding

Authors:
Lin Chen, Vu Nguyen Ha, Eva Lagunas, Linlong Wu, Symeon Chatzinotas, Björn Ottersten

Published in:
IEEE Transactions on Wireless Communications

Abstract:
Beam Hopping (BH) is a popular technique considered for next-generation multi-beam satellite communication system which allows a satellite focusing its resources on where they are needed by selectively illuminating beams. While beam illumination plan can be adjusted according to its needs, the main limitation of convectional BH is the adjacent beam avoidance requirement needed to maintain acceptable levels of interference. With the recent maturity of precoding technique, a natural way forward is to consider a dynamic beam illumination scheme with selective precoding, where large areas with high-demand can be covered by multiple active precoded beams. In this paper, we mathematically model such beam illumination design problem employing an interference-based penalty function whose goal is to avoid precoding whenever possible subject to beam demand satisfaction constraints. The problem can be written as a

binary quadratic programming (BQP). Next, two convexification frameworks are considered namely: (i) A Semi-Definition Programming (SDP) approach particularly targeting BQP type of problems, and (ii) Multiplier Penalty and Majorization-Minimization (MPMM) based method which guarantees to converge to a local optimum. Finally, a greedy algorithm is proposed to alleviate complexity with minimal impact on the final performance. Supporting results based on numerical simulations show that the proposed schemes outperform the relevant benchmarks in terms of demand matching performance while minimizing the use of precoding.

Keywords:
Dynamic Beam Illumination, Selective Precoding, User Scheduling, Binary Quadratic Programming

DOI:
10.1109/TWC.2022.3213418

The Next Generation of Beam Hopping Satellite Systems: Dynamic Beam Illumination with Selective Precoding

Lin Chen, Vu Nguyen Ha, Eva Lagunas, Linlong Wu,
Symeon Chatzinotas, Björn Ottersten

Abstract

Beam Hopping (BH) is a popular technique considered for the next-generation multi-beam satellite communication system which allows a satellite focusing its resources on where they are needed by selectively illuminating beams. While the beam illumination plan can be adjusted according to its needs, the main limitation of conventional BH is the adjacent beam avoidance requirement needed to maintain acceptable levels of interference. With the recent maturity of the precoding technique, a natural way forward is to consider a dynamic beam illumination scheme with selective precoding, where large areas with high demand can be covered by multiple active precoded beams. In this paper, we mathematically model such beam illumination design problem employing an interference-based penalty function whose goal is to avoid precoding whenever possible subject to beam demand satisfaction constraints. The problem can be written as a binary quadratic programming (BQP). Next, two convexification frameworks are considered namely: (i) A Semi-Definition Programming (SDP) approach particularly targeting BQP type of problems, and (ii) Multiplier Penalty and Majorization-Minimization (MPMM) based method which guarantees to converge to a local optimum. Finally, a greedy algorithm is proposed to alleviate complexity with minimal impact on the final performance. Supporting results based on numerical simulations show that the proposed schemes outperform the relevant benchmarks in terms of demand-matching performance while minimizing the use of precoding.

Index Terms

Dynamic Beam Illumination, Selective Precoding, User Scheduling, Binary Quadratic Programming

The preliminary results of this manuscript were presented in EUSIPCO 2021 [1]. This work has been supported by the Luxembourg National Research Fund (FNR) under the project "FlexSAT: Resource Optimization for Next Generation of Flexible SATellite Payloads" (C19/IS/13696663). The authors are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, 1855 Luxembourg Ville, Luxembourg. The corresponding author is Lin Chen(lin.chen@uni.lu).

I. Introduction

The roll-out of the next generation of wireless communication systems is expected to deliver faster internet access and increased capacity, providing customized services in a variety of use cases [2]. Despite the global growth of digital technologies, the United Nations (UN) has recently announced in the General Assembly that half of the world's population still has no internet access [3]. It is because of that there are still many remote locations where fiber and general terrestrial infrastructure cannot be deployed (or are not worth the investment), or where the ground equipment is with high probability subject to disruption by man-made events.

Exploiting satellite communications has been identified as a key solution to deliver ubiquitous and high-quality connectivity anywhere in the globe [4]. Conventional High-Throughout Satellite (HTS) systems have employed the spot beam technology with which satellite capacity is equally distributed across the multiple beams and contiguous coverage over a specific region can be provided [5]. While HTS with multi-beam architecture has dramatically improved the satellite system throughput, there have been increasing interests in developing fully reconfigurable satellite schemes that can smartly allocate the high capacity *"hot-spot"* areas [6].

The recent advances in space technology have opened the door to unprecedented flexibility and adaptability to satellite resources. As highlighted by the major satellite industry experts in Europe [7], "the continuous development of new technologies and the huge increase in satellite interest and investment, witnessed in recent time, have indeed pushed the satellite communication potentialities towards higher limits that need now to be explored to support the efficient and sustainable development of new markets and smart services". In the same document [7], spectrum usage and smart resource management are identified as major research challenges that need to be resolved to unleash the potential of the next-generation satellite communication system.

Concerning the satellite industry's interest in the aforementioned challenges, next, we provide an overview of two of the most advanced GEO HTS systems developed recently. One of the flagship flexible HTS satellites, the so-called Eutelsat Quantum, developed under an ESA Partnership Project with the satellite operator Eutelsat and the prime manufacturer Airbus, was launched in July 2021. Eutelsat Quantum is claimed to be the first commercial fully flexible software-defined satellite in the world [8]. Coverage, spectrum, and capacity can all be reconfigured in-orbit via its innovative reconfigurable payload, to efficiently serve any applications and ensure optimal use of its resources. According to the technical capabilities of Eutelsat Quantum [9], beams can be

hopped to spatially diverse regions rapidly and seamlessly. With a similar vision in mind, SES, the satellite operator, worked with Thales Alenia Space to manufacture SES-17. The satellite was launched in October 2021 and incorporates a digital transparent processor (DTP), enabling unique features, such as re-configurable resource allocation, to meet real-time traffic demands [10]. In addition, most of the industry-led projects are still in the testing phase, where the algorithm to unleash the flexibility of satellite optimally is still in the early stages. For instance, the European Commission has recently launched two 3-year projects related to optimal on-board resource management [11, 12].

Furthermore, similar to the situation in the terrestrial domain, the rapid development of data-hungry services has also resulted in a spectrum scarcity context in the satellite domain. As a consequence, the satellite digital broadcasting standard (DVB-S2X) introduced the possibility of using precoding techniques to enable efficient spectrum management while increasing the spectrum reuse across satellite spot-beams [13, 14]. The feasibility and potential of precoding applied to HTS systems have been recently validated via live experiments on a GEO satellite scenario [15, 16], confirming its relevance for future HTS deployments. In this work, we therefore address a combination of the two aforementioned challenges, namely (i) optimization of payload flexibility; and (ii) spectrum reuse.

Within the flexible satellite payload architectures, this paper focuses on the so-called time-domain flexibility which is commonly implemented via beam-hopping (BH) over time. BH became promising in the early 2010s since this technique can provide a good compromise between complexity and cost. The most attractive feature of BH is the payload mass reduction which is reflected in a reduced cost. Essentially, a BH-enabled satellite scheme can activate a sub-set of beams at each time slot following a time-space transmission pattern and this mechanism can be repeated periodically. In such schemes, the bandwidth can be re-used fully across all activated beams and the inter-beam interference can be well-managed by avoiding the geographically-adjacent-beam activation. While BH provides a certain degree of flexibility, an extremely asymmetrical traffic demand scenario over the coverage may critically challenge the conventional BH methods. In particular, high-demand areas expanding over multiple adjacent beams necessitate clusters of beams to be simultaneously activated while making use of the full available spectrum. An example can be the surroundings of an international airport with multiple high-demand mobile platforms flying around or a highly dense populated area with multiple backhauling satellite terminals.

A. Related Works

The works presented in the literature related to multi-beam HTS systems involving BH can be classified into two main categories:

1) Conventional Beam Hopping: The benefits of BH applied to Geostationary (GEO) satellite systems have been well-demonstrated in multiple academic works. Additionally, BH has attracted much attention from some key industrial players, e.g. [17, 18], and has been taken into account in the updated DVB-S2X standard [19]. However, exploiting the conventional BH techniques has also raised some challenges. Conventional BH was conceived to exploit the full available spectrum (i.e. full frequency reuse) over a subset of selected beams, ensuring that the geographical distance between selected beams is far enough to work under a noise-limited scenario [20]. The main technical challenge addressed in the literature has been the design of effective illumination patterns, i.e. determining the different sets of beams that need to be activated at each time slot while trying to align the offered capacity with the beam traffic demands over time. The design of illumination patterns for conventional BH usually involves binary variables representing the beam activation simultaneously. Therefore, the problem typically falls within the general mixed integer non-linear programming problem (MINLP) [21], which is very difficult to solve. The authors in [22] considered genetic algorithm targeting a globally optimal solution at the expense of high computational time. In a similar vein, [23] proposed to employ a simulated annealing method which also requires a long time of implementation. As an alternative to optimization-based methods, the works in [24, 25] developed heuristic iterative procedures that operate in a much faster and more efficient fashion by sacrificing optimality. Clearly, the key challenge identified in early works is the fact that exploiting the beam-activation binary variables results in a large searching space which exponentially aggravates due to the increasing number of potential beams. Following the trends of Machine Learning (ML) applied to optimization problems, [26, 27] investigated the applicability of deep learning tools within the BH illumination pattern optimization procedure. In addition, the conventional BH methods normally focus on no-multiplexing transmission across activated beams, which limits its capability of coping with some irregular traffic-demand scenarios in the IoT era [28].

2) Cluster Hopping: The activation of an adjacent set of beams (referred to as a cluster) was investigated within the European Space Agency (ESA) [28] and proposed in [29, 30] with the so-called Cluster Hopping (CH) scheme, where linear precoding [14] was considered to

mitigate the resulting intra-cluster interference. The downside of the works in [29, 30] is the fact that the overall virtual multi-beam grid is split into a set of non-overlapping clusters of fixed size and shape. This is done to reduce the search space and exclude the cluster design from the optimization problem. While some practical guidelines about the clustering design have been discussed in [30], the problem remains largely unsolved, especially when considering the complexity added by the precoding within the clusters.

B. Our Contribution

In this paper, we propose a general framework for the illumination pattern design, where the transmission of activated beams in separating clusters can be jointly precoded. The objective is to minimize the interference-based penalty with the aim of reducing the use of precoding while constraining the system to satisfy a certain beam demand in a given time window. Such technical design can be stated into a Binary Quadratic Programming. Whenever high demand expands over multiple adjacent beams, the solution from the proposed framework considers precoding to deal with the resulting inter-beam interference. With such a selective precoding mechanism, complexity at the ground segment is reduced where precoding operation can be considered flexible.

To linearize the BQP problem, we first present a procedure to convert the beam demand constraint into a more tractable notation involving required illuminated time slots per beam. Next, we present different ways to convexify the BQP problem. First, inspired by the mathematical works in [31, 32], we reformulate the BQP into a Semi-Definite Programming (SDP) form which can be solved efficiently by employing some standard optimization solver tools. In another approach, we also propose a novel solving framework MPMM which integrates Multiplier Penalty (MP) [33] and Majorization-Minimization (MM) [34] methods. In particular, we relax the binary constraint and add its augmented Lagrangian function into the objective function by using the so-called penalty parameters. Then, the new penalty-form problem is solved iteratively by sequentially updating the penalty multipliers and driving the solution to binary values. In particular, in each iteration, we adopt the MM method to transfer the penalty-form problem into a sequence of simple problems, each of which can be solved optimally. The sequence generated by the optimal solutions of these simple problems is proved to converge to a stationary point. According to the convexity of the penalty-form problem, one also concludes that the stationary point is the optimal solution. Since the previously proposed methods prioritize performance

versus computational complexity, we complement this paper by proposing a heuristic greedy algorithm that provides a sub-optimal but efficient solution.

Our main contributions are summarized as follows.

- We propose a general framework and its mathematical formulation to support dynamic and flexible cluster hopping, where geographically adjacent beams are allowed to be simultaneously activated whenever needed according to its demand request. The resulting intra-cluster interference is mitigated with linear precoding, whose utilization is minimized by focusing the design on an interference-based penalty function.
- Based on probability theory, we propose an effective way to reformulate the beam demand constraint and convert it into the number of illuminated time slots required per beam in order to satisfy the demand. Such simplification convexifies the constraint and helps ease the tractability of the problem.
- Three different methodologies are proposed to address the BQP problem. We first make use of a novel SDP notation specifically designed for BQP problems. As a more accurate alternative, we propose an algorithm that integrates MP and MM methods. Finally, a novel heuristic algorithm is presented to rapidly provide a solution with acceptable performance.
- We provide a detailed complexity analysis for each of the proposed methods.
- Finally, an extensive numerical evaluation is carried out, where the proposed methods are compared with conventional BH and the previously proposed CH[1]. The results evidence the effectiveness of the proposed algorithms and demonstrate the flexibility of the proposed framework in adapting to any demand distribution.

Please note that, although this paper focuses its notation and simulations on GEO satellite systems, the methodology itself can be applicable to the beam placement problem encountered in NGEO constellations. However, the precoding application for distributed satellite swarms is still in the early stages of investigation and may need further discussion.

The remainder of this paper is organized as follows. In Section II, we present the system model. In Section III, we present the general formulation of the dynamic beam illumination design problem. To solve the problem, Section IV provides the method to simplify the non-convex demand constraints into the linear forms based on which the BH-design problem is reformulated as a BQP. In Section V, two efficient optimization-based algorithms and a greedy

[1]These benchmarks are detailed in the numerical results section.

mechanism are presented to deal with BQP. In Section VI, we present numerical simulations. Finally, Section VII concludes the paper. Notations used in this paper are summarized in Table I.

II. System Model

This paper studies the forward link of a bent-pipe multi-beam geostationary (GSO) satellite system, whose coverage area is divided into a virtual grid of N spot beams. In this system, the illumination pattern is designed over a specific BH window, which is periodically repeated over time. The BH window is divided into a set of M time slots (TSs), and within each TS no more than K beams ($K << N$) can be illuminated. The duration of one TS is denoted as T_s (seconds) which also represents the minimum dwelling time of the hopping mechanism. Let g_l be the traffic demand in bps of beam l, and $\mathbf{g} = [g_1, \ldots, g_N]^T$ represents the all-beam demand vector. For simplicity, a one-user-per-beam scenario is assumed, i.e. a single virtual user located inside the beam footprint (e.g., $4\ dB$ contour) is assumed whenever this beam is activated. Note that the single virtual user is assumed to aggregate the demand of the whole beam user density. The assumption of a single virtual user per beam is performed to abstract the user scheduling. User scheduling is out of the scope of the general BH design for different reasons. The multiplexing of multiple users is assumed to be performed in a time-division-multiple access (TDMA) fashion.

A. Channel Model

Let $\mathbf{H} \in \mathbb{C}^{N \times N}$ be the channel matrix containing all the channel coefficients of the forward link. In particular, the channel between antenna of the satellite payload corresponding to beam l and user k on the ground is modeled following the approach in [5], and can be written as,

$$H_{k,l} = \sqrt{G_R^{(k)} G_l\left(x_k, y_k\right) e^{j\phi_{k,l}}} \Big/ \left(4\pi \frac{d_k}{\lambda}\right) \tag{1}$$

where $G_R^{(k)}$ is the receiving antenna gain at user k; $G_l\left(x_k, y_k\right)$ stands for the beam pattern gain due to beam l at k which can be estimated according to the user's longitude x_k and latitude y_k; $\phi_{k,l}$ is the phase component associated with the antenna beam pattern; d_k represents the distance from the satellite to that user; λ denotes the wavelength of the carrier frequency band.

Doppler and absorption loss are intentionally not included in our model. The movement of the GEO satellite is maintained in a very tight box and has a negligible Doppler shift (note that daily maneuvering is performed to maintain the satellite in its position). Concerning the absorption loss, this would appear as a constant loss in our link budgeting thus not making an impact in our

TABLE I: Notations

Notation	Definition
M	Total number of TSs
N	Total number of beams
P_b	Transmit power per beam
B	Available full bandwidth
$\mathbf{g}$	Demand vector in [bps]
$\mathbf{d}$	Demand vector in [number of TSs]
$H_{k,l}$	Channel between beam l and user k
x_k	Longitude of user k
y_k	Latitude of user k
$G_R^{(k)}$	Gain of receiving antenna at user k
d_k	Distance from the satellite to user k
λ	Wavelength of the carrier frequency
$\lambda(\mathbf{A})$	Minimal eigenvalue of $\mathbf{A}$
$\mathbf{S}_j$	Coverage area of beam j
α	Predefined regularization factor for MMSE
$\hat{\mathbf{W}}_i$	MMSE precoding matrix for cluster i
$\mathbf{W}_i$	Normalized precoding matrix for cluster i
$\mathbf{W}$	Normalized precoding matrix for all beams
$\mathbf{w}_k^t$	Precoding vector for user k at TS t
$\mathbf{n}$	Zero-mean additive Gaussian noise
$\mathbf{h}_k$	Channel vector for user k
$\mathbf{s}[t]$	Transmitted symbol vector at TS t
$L[t]$	Number of clusters at TS t
τ	Blotzmann constant
δ_T^2	Power of thermal noise at temperature T
T_{Rx}	Clear sky noise temperature of the receiver
K	Maximal number of active beams
K_{avg}	Average number of active beams for an instance
$\Omega_{i,j}$	Penalty which is equal to $\omega_{i,j}$
ζ_n	Average achievable rate of beam n
κ	Clustering threshold
$\phi_{k,l}$	Phase component between beam l and user k
$G_l\left(x_k, y_k\right)$	Radiated gain due to beam l and user k
$x_{n,t}$	Illumination variable due to beam n in TS t
$\mathbf{x}_t$	Vector of all illumination variable in TS t
$\omega_{i,j}$	Influence factor from beam i to beam j

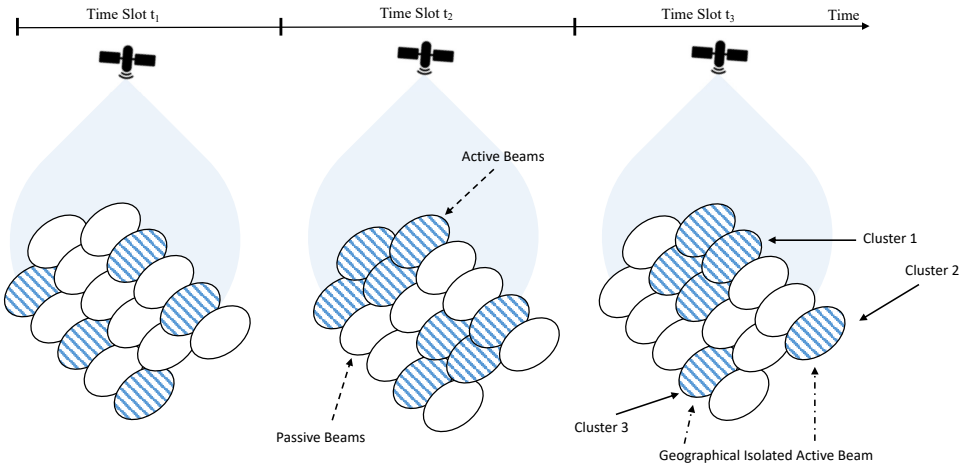


Fig. 1: Proposed flexible cluster hopping scheme

study. This is because the beam hopping window time is usually below a few hundred frames. For instance, considering a frame duration of 1.3 msec (the number of symbols in a super-frame is 612540, and its duration is about 1.3 msec for a 500 MHz bandwidth), and considering 256 frames, the total BH window-time is approximate to 330 msec. Clearly, the atmospheric loss has a longer coherence time.

B. Selective Precoding Strategy

As depicted in Fig. 1, we envision an illumination pattern design that dynamically activates clusters of beams. Whenever the cluster size is greater than one, precoding is needed to alleviate the inter-beam interference. The illumination pattern design is discussed in Section III. Herein, we detail how precoding is implemented for a given illumination pattern[2]. It is worth noting that the precoding operation entails significant complexity at the gateway side, which exponentially scales with the number of involved beams [35]. Hence, the precoding strategy should be designed smartly by grouping active beams into different precoded clusters, which are subsequently precoded independently. In what follows, we provide a detailed description of the two-step procedure of grouping beams in clusters and how the precoding matrix is designed for a specific cluster.

[2]Note that precoding refers to the exploitation of instantaneous CSI and it is implemented at the ground segment. It needs to be distinguished from the beam pattern formation, which is implemented onboard the satellite and is assumed to be fixed in this work.

1) Precoded-Clustering Strategy: Given the illumination pattern, at TS t, the definition of precoded clusters is given as follows. The general idea is that beams generating strong interference with each other should be grouped into one precoded cluster. First, we introduce the so-called influence factor $\omega_{i,j}$, which captures the impact of the inter-beam relative interference from beam i to beam j, and it is defined as

$$\omega_{i,j} = \frac{\iint_{\mathbf{S_j}} |H_{k,i}|^2 P_i dx_k\, dy_k}{\iint_{\mathbf{S_j}} |H_{k,j}|^2 P_j dx_k\, dy_k} \overset{(*)}{=} \frac{\iint_{\mathbf{S_j}} G_i\left(x_k, y_k\right) dx_k\, dy_k}{\iint_{\mathbf{S_j}} G_j\left(x_k, y_k\right) dx_k\, dy_k} \tag{2}$$

where P_i denotes the transmission power of beam i and $\mathbf{S}_j$ stands for coverage area of beam j which is defined by the beam contour at -4 dB from the maximum gain. As can be seen, $\omega_{i,j}$ represents the ratio of the interference power from beam i to beam j. Here, the power of signals is calculated as the average value over the coverage area of beam j, $(*)$ in (2) implies the simplified influence factor when the same transmission power for all activated beams and the same receiving gain at all users are assumed.

In a second step, these factors $\omega_{i,j}$ are compared to a predetermined threshold [3]. Two beams corresponding to an influence factor greater than the threshold will be located in one cluster. The proposed threshold-based clustering strategy is summarized **in Algorithm 1**. Particularly, in each TS, one starts by setting every activated beam as a separate cluster. Then, if there are any two beams in two separate clusters whose corresponding influence factor is greater than the threshold, these two corresponding clusters are merged into one. This process is iteratively repeated until there is no change in the clustering structure. Note that the outcome of Algorithm 1 classifies all active beams into clusters, some of which may contain one beam. Only those clusters with a size greater than one will be considered for the precoding process.

2) Precoding Design: Once the clusters are formed, the precoding matrix is designed separately for each of them. Let $L[t]$ be the number of clusters in TS t; here, all non-illuminated beams are grouped into a non-transmission cluster for convenience. Denote $\mathbf{W}_i \in \mathbb{C}^{c_i \times c_i}$ as the precoding matrix of cluster i where c_i stands for its cardinality. For the non-illuminated beams, the corresponding precoding matrix must be zeros since they are silent during that particular TS. For the clusters consisting of only one beam, the precoding matrix can be defined simply as $\sqrt{P_{\mathsf{b}}}$, where P_{b} denotes the per-beam transmit power. For the remaining clusters (the ones

[3]The value of the threshold can be easily set, as there is an evident abrupt drop in influence values for those beams that are not causing harmful interference.

that are formed with more than one active beam), their corresponding precoding matrices are obtained using the MMSE-based strategy as given in [14, 36]. In particular, the MMSE-based precoding matrix $\hat{\mathbf{W}}_i$ can be expressed as

$$\hat{\mathbf{W}}_i = \sqrt{P_{\mathtt{b}}}\mathbf{H}_i^H \left(\mathbf{H}_i\mathbf{H}_i^H + \alpha\mathbf{I}\right)^{-1} \tag{3}$$

where $\mathbf{H}_i$ represents the channel matrix of all the users in cluster i and α stands for a predefined regularization factor. Then, $\mathbf{W}_i$ is determined based on $\hat{\mathbf{W}}_i$ by normalizing every column vector of the matrix to meet the per-beam power constraints $P_{\mathtt{b}}$.

Let $z_k[t]$ be the received signal at user k in TS t, and $\mathbf{z}_i[t] = [z_k[t]|k \in \mathcal{C}_i]^T$ be the vector of received signals corresponding to cluster i. The vector $\mathbf{z}[t] = [\mathbf{z}_1^T[t], ..., \mathbf{z}_L^T[t]]^T$ including all users' received signal can be expressed as

$$\mathbf{z}[t] = \mathbf{H}\mathbf{W}[t]\mathbf{s}[t] + \mathbf{n} = \mathbf{H}\ \mathtt{BDiag}\left(\mathbf{W}_1, ..., \mathbf{W}_L\right)\mathbf{s}[t] + \mathbf{n} \tag{4}$$

where $\mathtt{BDiag}(*)$ stands for the block-diagonal matrix operation, $\mathbf{W}[t] \in \mathbb{C}^{N\times N}$ represents to the precoding matrix for all beams, $\mathbf{s}[t] \in \mathbb{C}^{N\times 1}$ denotes the transmitted symbol vector; and $\mathbf{n}$ stands for the noise vector. In this paper, the zero-mean additive Gaussian noise is assumed at all the users where $\mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right] = \sigma_T^2\mathbf{I}$ and $\sigma_T = \sqrt{\tau T_{Rx} B}$; τ denotes the Boltzmann constant and T_{Rx} is the clear sky noise temperature of the receiver [5].

III. BH Illumination Pattern Design Formulation

Let us denote $x_{n,t} \in \{0, 1\}$ as the binary assignment variable indicating the illumination of beam n in TS t. Then, the number of illuminated beams in TS t can be described as $\sum_{n=1}^{N} x_{n,t}$. Due to the typical payload mass limitations of a BH-enabled satellite, the number of active beams must remain not greater than the number of RF chains K, i.e., $\sum_{n=1}^{N} x_{n,t} \leq K, \forall t$.

The achievable rate of user n in TS t can be expressed as

$$R_n[t] = B\log_2\left(1 + \frac{x_{n,t}|\mathbf{h}_n\mathbf{w}_n^t|^2}{\sum_{k\neq n} x_{k,t}|\mathbf{h}_n\mathbf{w}_k^t|^2 + \sigma_T^2}\right) \tag{5}$$

where $\mathbf{w}_n^t$ denotes the precoding vector designed for user n in TS t, e.g. column of $\mathbf{W}[\mathbf{t}]$ corresponding to that user. Note that $\mathbf{W}[t]$ is also a function of illumination pattern $\mathbf{x}$ which determines the way to do clustering for precoding. Considering the average traffic demand of user n, g_n [bps], we can express the per-user demand constraint as,

$$\frac{1}{M}\sum_{t=1}^{M} R_n[t] \geq g_n \quad \text{[bps]}. \tag{6}$$

Algorithm 1 THRESHOLD-BASED CLUSTERING ALGORITHM

Initializaion $\kappa, \omega_{i,j}$
Let $\mathcal{B}^{\mathsf{ac}}[t]$ be the set of activated beams in TS t.
Define $\mathcal{C}_k = \{k\}$ initial cluster consisting user k ($k \in \mathcal{B}^{\mathsf{ac}}[t]$).
Define $\Phi[t] = \{\mathcal{C}_k | k \in \mathcal{B}^{\mathsf{ac}}[t]\}$ initial set of clusters, $\phi[t] = \mathcal{B}^{\mathsf{ac}}[t]$ initial set of cluster indices.
repeat
  for $k \in \phi[t]$ in ascending order **do**
    for $l \in \phi[t]$ that $l > k$ **do**
      if $\exists i \in \mathcal{C}_l$ that $\max_{j \in \mathcal{C}_k} \omega_{i,j} \geq \kappa$ **then**
        $\mathcal{C}_k = \mathcal{C}_k \cap \mathcal{C}_l$, $\Phi[t] = \Phi[t]/\{\mathcal{C}_l\}$, $\phi[t] = \phi[t]/\{l\}$.
      end if
    end for
  end for
until There is no change of $\Phi[t]$ and $\phi[t]$.

The main objective of this work is to develop a BH illumination pattern design such that the users' demands are satisfied while avoiding the use of precoding whenever possible (to ease the complexity burden). For this purpose, we shall avoid the strong cross-interference among the illuminated beam in every TS as much as possible. To formulate such a problem, we make use of the influence factors defined in (2) to generate a penalty matrix $\mathbf{\Omega} \in \mathbb{R}^{N \times N}$ as

$$[\mathbf{\Omega}]_{(i,j)} = \omega_{i,j}, \ 1 \leq i, j \leq N \tag{7}$$

where $[\mathbf{\Omega}]_{(i,j)}$ indicates the element on the i-th row and j-th column of $\mathbf{\Omega}$. Using this penalty

matrix, one can state the BH design problem as[4]

$$
\begin{aligned}
(\mathcal{P}_0): \underset{\mathbf{x}_1,\ldots,\mathbf{x}_M}{\texttt{minimize}} \quad & \sum_{t=1}^{M} \mathbf{x}_t^T \, \mathbf{\Omega} \, \mathbf{x}_t \\
\texttt{subject to} \quad & (C_1): \sum_{n=1}^{N} x_{n,t} \leq K, \ \forall t \\
& (C_2): \sum_{t=1}^{M} R_n[t] \geq M g_n, \ \forall n \\
& (C_3): x_{n,t} \in \{0,1\}, \ \forall t, n
\end{aligned}
\tag{8}
$$

where $\mathbf{x}_t = [x_{1,t}, x_{2,t}, ..., x_{N,t}]^T$. Problem $(\mathcal{P}_0)$ is an integer programming (IP) which is an NP problem in general. The challenge of solving this problem not only comes from the binary assignment variables but also from the non-convex function of $R_n[t]$ in constraint (C_2) which is a function of $\mathbf{x}$.

***Remark** 1: It is worth mentioning that problem $(\mathcal{P}_0)$ is stated in a general form in which $\mathbf{\Omega}$ can take any values. In particular, for different designing goals, $\mathbf{\Omega}$ can be determined carefully. Therefore, the BH strategy proposed in the following parts can stand in many schemes with appropriate penalty matrices.*

IV. Problem Reformulation

In this section, effective approaches for dealing with the challenging problem given by (8) are presented. Particularly, according to the idea of estimating the average supplied capacity, we first simplify the complicated non-convex traffic-demand constraints (C_2) to linear forms. Based on this, the problem is re-formulated as a binary quadratic programming (BQP).

A. Demand Constraint Simplification

The actual supplied capacity $R_n[t]$ is a non-convex function of the variable $\mathbf{x}_t$ and, therefore, it causes a big challenge for the BH protocol design. Some works in literature have suggested different frameworks to address this issue, such as simple interference-free relaxation given in [37], limiting the set of illuminated beams to avoid the strong interference in [29]. Considering a different approach to dealing with this issue, we aim to convert the demand g_n [bps] into a

[4]The satellite is assumed to be dimensioned according to the expected demands, i.e. to have enough resources to address the expected demand g_n, $\forall n$.

minimum number of TSs that each beam must be activated in order to meet its traffic demand. To do so, we first consider the following proposition.

Proposition 1: Let ζ_n be the average achievable rate of beam n, i.e., $\zeta_n \sum_{t=1}^{M} x_{n,t} = \sum_{t=1}^{M} R_n[t]$. Then, constraint (C_2) can be re-formulated as

$$(\hat{C}_4): \quad \sum_{t=1}^{M} \mathbf{e}_n^T \mathbf{x}_t \geq d_n, \ \forall n \tag{9}$$

where $\mathbf{e}_n$ denotes a vector in which the n-th component equals to one and all others are zeros,

$$d_n = \lceil M g_n / \zeta_n \rceil \tag{10}$$

and $\lceil * \rceil$ stands for the ceiling operator.

Proof: The proposition can be proved as follows. Since, $\zeta_n \sum_{t=1}^{M} x_{n,t} = \sum_{t=1}^{M} R_n[t]$ and $\mathbf{e}_n^T \mathbf{x}_t = x_{n,t}$, constraint (C_2) is equivalent to

$$\sum_{t=1}^{M} \mathbf{e}_n^T \mathbf{x}_t = \sum_{t=1}^{M} x_{n,t} \geq \frac{M g_n}{\zeta_n}, \ \forall n \tag{11}$$

Additionally, $\sum_{t=1}^{M} \mathbf{e}_n^T \mathbf{x}_t$ is an integer. Hence, the right hand side of (11) can be replaced by $d_n = \lceil M \frac{g_n}{\zeta_n} \rceil$. This has closed the proof. ■

In constraint $(\hat{C}_4)$, d_n can be considered as the minimum number of TSs in each of which beam n is illuminated. Next, the following theorem regards the relation between the optimal solution of problem $(\mathcal{P}_0)$ and d_n.

Theorem 1: Let $\mathbf{x}_t^\star$'s be the optimal solution of $(\mathcal{P}_0)$, then we have $\sum_{t=1}^{M} \mathbf{e}_n^T \mathbf{x}_t^\star = d_n, \ \forall n$, if ζ_n's are estimated accurately, which means constraints $(\hat{C}_4)$ hold for all beams.

Proof: This theorem can be proved easily by using the contradiction method. In particular, one assumes that there exists at least one beam that the corresponding constraint $(\hat{C}_4)$ does not hold. Denote this such beam as n^* which yields $\sum_{t=1}^{M} x_{n^*,t}^\star > d_{n^*}$. Selecting any TS t^* that $x_{n^*,t^*}^\star = 1$, we generate the new solution of $(\mathcal{P}_0)$, $\mathbf{x}_t'$s, that $x_{n,t}' = x_{n,t}^\star \ \forall (n,t) \neq (n^*, t^*)$ and $x_{n^*,t^*}' = 0$. It is easy to observe that $\mathbf{x}_t'$'s satisfies constraints $(\hat{C}_4)$. Moreover, this new solution meets the requirement of constraints (C_1) and (C_3) while resulting in the lower objective function. It follows by a contradiction since $\mathbf{x}_t^\star$'s is the optimal solution. Therefore, constraint $(\hat{C}_4)$ holds for all beams with any optimal solution of problem $(\mathcal{P}_0)$. ■

Thanks to Theorem 1, the following lemma can stand.

Lemma 1: If ζ_n's are estimated accurately, one can replace constraint (C_2) by the following one without changing the optimal solution of problem $(\mathcal{P}_0)$.

$$(C_4): \quad \sum_{t=1}^{M} \mathbf{e}_n^T \mathbf{x}_t = d_n,\ n = 1, \ldots, N \tag{12}$$

Due to this result, in what follows, we propose an iterative framework to estimate the average achievable rate ζ_n for each beam by appraising the expected interference.

1) Average Achievable Rate Estimation Framework: According to the mean field theory[38], we assume the uniform distribution of beams to be activated over the time window. Thanks to Lemma 1, the illuminating probability of beam n in a specific TS, e.g., TS t, can be given by

$$p_n = \text{Prob}\{x_{n,t} = 1 | t\} = d_n / M \tag{13}$$

Regarding the clustering and precoding processes, the expected interference to beam n can be expressed by taking into account the interference from the beams with low corresponding influence factors and their illuminating probability as $\psi_n = \sum_{j, j \in \mathcal{A}_n} p_j P_{\mathsf{b}} |H_{n,j}|^2$, where $\mathcal{A}_n = \{i \mid \Omega_{i,n} < \kappa, i \in \mathbb{N}, i \leq N, i \neq n\}$. Based on that, the expected average achievable rate of beam n can be described as

$$\zeta_n = B \log_2 \left(1 + \frac{P_{\mathsf{b}} |h_{n,n}|^2}{\psi_n + \sigma_T^2}\right) \tag{14}$$

Although one of $(d_n, p_n, \psi_n, \zeta_n)$'s can be defined if the others are given, determining the accurate values of these factors is very challenging. Exploiting the expectation maximization (EM) algorithm given in [38], we propose an iterative framework to estimate ζ_n's as summarized in **Algorithm 2**. Particularly, the algorithm initializes with zeros illuminating probability for all beams and repeatedly updating $(d_n, p_n, \psi_n, \zeta_n)$'s in each iteration. Each iteration processes two steps, namely, expectation (**E-Step**) and maximization (**M-Step**). The **E-Step** is called for updating the interference based on the illuminating probabilities of the previous iteration while the **M-Step** stands for calculating ζ_n's, d_n's, and adjusting the probabilities. The iterative process stops at the convergence according to the following proposition.

***Proposition** 2:* Algorithm 2 converges after a finite number of iterations.

Proof: As can be observed, p_n's increase while ζ_n's decrease in every iteration. Since p_n's are upper bounded by ones and ζ_n's are lower bounded by zeros. The iterative process must converge to a stable point after a finite number of iterations. ■

***Remark** 2: Note that if the required demand g_n is higher than ζ_n^{new} in a specific iteration, d_n^{new} will be re-set as M - the highest number of TSs. One also notices that problem $(\mathcal{P}_0)$ is*

Algorithm 2 AVERAGE CAPACITY ESTIMATION

1: **Initializaion** $P_{\mathsf{b}}, B, \sigma_T, \mathbf{H}, \mathbf{\Omega}, \kappa, \mathcal{A}_n, \mathbf{p}^{\mathsf{old}} = \mathbf{0}$

2: **while** $\| \boldsymbol{\zeta}^{\mathsf{new}} - \boldsymbol{\zeta}^{\mathsf{old}} \|^2 \leq \iota$ **do**

3: For all $n = 1, ..., N$, update:

$$\textbf{E Step:} \quad \psi_n^{\mathsf{new}} = \sum_{j, j \in \mathcal{A}_n} p_j^{\mathsf{old}} P_{\mathsf{b}} |h_{n,j}|^2 \tag{15}$$

$$\textbf{M Step:} \quad \zeta_n^{\mathsf{new}} = B \log_2 \left(1 + \frac{P_{\mathsf{b}} |h_{n,n}|^2}{\psi_n^{\mathsf{new}} + \sigma_T^2} \right) \tag{16}$$

$$d_n^{\mathsf{new}} = \lceil M \frac{g_n}{\zeta_n^{\mathsf{new}}} \rceil \tag{17}$$

$$p_n^{\mathsf{old}} = \frac{d_n^{\mathsf{new}}}{M} \tag{18}$$

4: **end while**

infeasible if K is smaller than the average number of active beams K_{avg} which is expressed as $K_{avg} = \lceil \sum_{\forall n} d_n / M \rceil$.

***Remark** 3: It is worth noting that (C_2) and (C_4) may be not equivalent if ζ_n is not estimated accurately. In addition, once ζ_n is well evaluated as in Algorithm 2, and d_n is calculated as in* (10)*, the unmet capacity of beam n must be smaller than, $\left[M g_n - d_n B \log_2 \left(1 + \frac{|h_{n,n}|^2 P_{\mathsf{b}}}{\sum_{k \neq n} |h_{k,n}|^2 P_{\mathsf{b}} + \sigma_T^2} \right) \right]^+$ since $\sum_{k \neq n} |h_{k,n}|^2 P_{\mathsf{b}}$ is the highest interference power suffering beam n in any TS.*

2) Problem Reformulation: For the sake of simplicity, we compact our notation by rearranging all TSs t into a single tall vector $\mathbf{x}^T = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T & \cdots & \mathbf{x}_M^T \end{bmatrix}$. Thanks to Lemma 1, problem $(\mathcal{P}_0)$ can be re-stated as

$$\begin{aligned} (\mathcal{P}_1): \quad & \underset{\mathbf{x}}{\texttt{minimize}} \quad \mathbf{x}^T \mathbf{A} \mathbf{x} \\ & \texttt{subject to} \, (C_5): \mathbf{B}\mathbf{x} \preceq K \cdot \mathbf{1}_M, \\ & (C_6): \mathbf{D}\mathbf{x} = \mathbf{d} \\ & (C_7): \mathbf{x} \in \{0,1\}^{NM} \end{aligned} \tag{19}$$

where $\mathbf{d} = [d_1, d_2, \ldots, d_N]^T$, $\mathbf{A} = \mathbf{I}_M \otimes \mathbf{\Omega}$, $\mathbf{B} = \mathbf{I}_M \otimes \mathbf{1}_N^T$, $\mathbf{D} = \mathbf{1}_M^T \otimes \mathbf{I}_N$. Herein, $\otimes$ denotes the Kronecker product, $\mathbf{1}_M$ stands for the vector with M one elements, and $\mathbf{I}_M$ is the identity matrix with dimension of M.

B. Objective Function Convexification

As can be observed, problem $(\mathcal{P}_1)$ is a BQP which is NP-hard in general. To ease the tractability of $(\mathcal{P}_1)$, we aim to characterize the objective function convexity by considering the following theorem.

***Theorem** 2:* For any value of a, problem $(\mathcal{P}_1)$ is equivalent to the following problem

$$(\mathcal{P}_a): \quad \underset{\mathbf{x}}{\texttt{minimize}} \quad \mathbf{x}^T(\mathbf{A}+a\mathbf{I})\mathbf{x} \quad \texttt{subject to } (C_5),\ (C_6),\ (C_7). \tag{20}$$

Proof: Due to the constraint (C_6), if $\mathbf{x}'$ is a solution of problem $(\mathcal{P}_1)$, then we have $a\mathbf{x}^T\mathbf{I}\mathbf{x} = a\sum_{n=1}^{N} d_n$ which is a constant. Then, $\mathbf{x}'$ must be a solution of problem $(\mathcal{P}_a)$. Inversely, it is easy to prove that any solution of $(\mathcal{P}_a)$ must be a solution of $(\mathcal{P}_1)$. Hence, $(\mathcal{P}_1)$ and $(\mathcal{P}_a)$ are equivalent for any value of a. ■

In addition, the convexity of the objective function $(\mathcal{P}_a)$ can be guaranteed if a is selected so that it is not less than $-\lambda(\mathbf{A})$ - the minimum eigenvalue of $\mathbf{A}$. Thanks to Theorem 2, we can state that problem $(\mathcal{P}_1)$ is equivalent to an integer QP with a convex objective function, i.e., $\mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x}$ where $\tilde{\mathbf{A}} = \mathbf{A} - \lambda(\mathbf{A})\mathbf{I}$. To this end, instead of solving $(\mathcal{P}_1)$, we will focus on the following

$$(\mathcal{P}_1): \quad \underset{\mathbf{x}}{\texttt{minimize}} \quad \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} \quad \texttt{subject to } (C_5),\ (C_6),\ (C_7). \tag{21}$$

V. Binary Quadratic Programming Optimization

In this section, three optimization approaches are introduced to deal with the BQP problem $(\tilde{\mathcal{P}}_1)$. Particularly, two efficient solving approaches using the SDP relaxation and MPMM methods, respectively. For completeness, a low-complexity greedy algorithm is also proposed. Finally, a complexity analysis of the proposed solution mechanisms is presented.

A. SDP-based Algorithm

Problem $(\tilde{\mathcal{P}}_1)$ corresponds to a BQP form, i.e. a problem involving a quadratic objective function with binary variables, which could be solved by relaxing the binary constraint [39]. Firstly, the binary constraint (C_7) is equivalent to two equations [40], i.e.

$$\mathbf{x} \in \{0,1\}^{MN} \iff \mathbf{X} = \mathbf{x}\mathbf{x}^T \text{ and } (C_8): \operatorname{diag}(\mathbf{X}) = \mathbf{x}. \tag{22}$$

Herein, the *"rank-one"* constraint $\mathbf{X}=\mathbf{x}\mathbf{x}^T$ can be further relaxed as [40]

$$\begin{cases}(C_9):\begin{bmatrix}\mathbf{X} & \mathbf{x}\\ \mathbf{x}^T & 1\end{bmatrix}\succeq 0 \text{ (equivalent to } \mathbf{X}\succeq\mathbf{x}\mathbf{x}^T)\\ (C_{10}):\mathbf{X}\in\mathbf{S}^n.\end{cases}\tag{23}$$

Then, problem $(\tilde{\mathcal{P}}_1)$ can be approximated to the following semidefinite problem,

$$\begin{aligned}(\mathcal{P}_{\mathsf{SDP}}):&\underset{\mathbf{x},\mathbf{X}}{\texttt{minimize}}\ \text{Tr}\left(\tilde{\mathbf{A}}\mathbf{X}\right)\\ &\texttt{subject to}\ (C_5),\ (C_6),\ (C_8),\ (C_9),(C_{10})\end{aligned}\tag{24}$$

Problem $(\mathcal{P}_{\mathsf{SDP}})$ in (24) can be solved efficiently by employing the advanced mixed-integer optimization toolboxes such as CVX [41]. If the matrix obtained by solving $(\mathcal{P}_{\mathsf{SDP}})$ is *"rank-one"*, then it provides the optimal solution to the problem. In the case of SDP providing a solution matrix whose rank is higher than one, the SDP-based branch and bound method [42, 43] can be applied to obtain the final solution.

B. MPMM Algorithm

In this section, we first introduce the general principles of MP and MM methods based on which a novel multiplier penalty and majorization-minimization (MPMM) algorithm is proposed to solve problem $(\tilde{\mathcal{P}}_1)$ efficiently. Then, the convergence of this approach is also discussed.

1) Multiplier Penalty Method: The MP method is an efficient approach for solving the constrained optimization problem. Considering, a general equality-constraint problem as follows,

$$\begin{aligned}&\underset{\mathbf{x}\in\mathcal{X}}{\texttt{minimize}}\ f(\mathbf{x})\\ &\texttt{subject to}\ \ h_i(\mathbf{x})=0, i=1,\ldots,m\end{aligned}\tag{25}$$

where $\mathcal{X}$ is a convex set. Following the MP method, this problem could be solved by minimizing the following sequential problems

$$\mathbf{x}^{[\ell]}=\arg\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})+\sum_{i=1}^{m}\eta_i^{[\ell]}h_i(\mathbf{x})+\frac{\rho^{[\ell]}}{2}\sum_{i=1}^{m}[h_i(\mathbf{x})]^2\tag{26}$$

where ℓ is the index of iteration, $\{\boldsymbol{\eta}^{[\ell]}\},\{\rho^{[\ell]}\}$ stand for sequences of penalty factors. Here, the penalty term is introduced by its augmented Lagrangian function. The feature of MP method is the way to update $\boldsymbol{\eta}^{[\ell]}$ step by step [44], which is given by

$$[\boldsymbol{\eta}^{[\ell+1]}]_i=\eta_i^{[\ell]}+\rho^{[\ell]}h_i(\mathbf{x}).\tag{27}$$

The result given in [33] also concludes that the sequence of $\{\boldsymbol{\eta}^{[\ell]}\}$ will converge to a fixed point at which $\{\mathbf{x}^{[\ell]}\}$ converges to a local optimum of problem (25).

2) Majorization-Minimization Method: The MM method is a well-known approach to dealing with a complicated problem by transferring it into a sequence of simple problems which can be solved effectively. The main idea of this scheme is to construct the surrogate function $u\left(\mathbf{x} \mid \mathbf{x}_{(k)}\right)$ which approximates the original objective function, then solve the constructed problems in sequence until convergence. For the general minimization problem, $\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$ where $\mathcal{X}$ is convex set. The constructed surrogate function $u\left(\mathbf{x} \mid \mathbf{x}_{(k)}\right)$ should satisfy

$$f(\mathbf{x}) \leq u\left(\mathbf{x} \mid \mathbf{x}_{(k)}\right), \forall \mathbf{x} \in \mathcal{X}, \text{ and } f\left(\mathbf{x}_{(k)}\right) \leq u\left(\mathbf{x}_{(k)} \mid \mathbf{x}_{(k)}\right) \tag{28}$$

Then the sequence of $\{\mathbf{x}_{(k)}\}$ is given by $\mathbf{x}_{(k+1)} = \arg\min_{\mathbf{x}\in\mathcal{X}} u\left(\mathbf{x} \mid \mathbf{x}_{(k)}\right)$, which will converge to a stationary point of the original problem [45]. If the problem is convex, then the stationary point is the global minimum.

3) Proposed MPMM Method: The challenge on solving $\tilde{\mathcal{P}}_1$ mainly comes from the binary constraint $(C7)$. To cope with this challenge, we aim to employ MP method to relax the binary constraint and deal with a sequence of penalty problems. In particular, the augmented Lagrangian function is added to the objective function with penalty parameters while the binary constraint $(C7)$ is relaxed to form the penalty problem as

$$(\mathcal{P}_{\mathsf{MPMM}}): \mathbf{x}^{[\ell]} = \arg\min_{\mathbf{x}\in\mathcal{Y}} f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right), \tag{29}$$

where $f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right) = \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} + \sum_i \eta_i^{[\ell]}(x_i - x_i^2) + \frac{\rho^{[\ell]}}{2}\sum_i (x_i - x_i^2)^2$ and $\mathcal{Y} = \{\mathbf{x}|\mathbf{B}\mathbf{x} \preceq K \cdot \mathbf{1}_M, \mathbf{D}\mathbf{x} = \mathbf{d}, 0 \leq x_i \leq 1, \forall i\}$. It is worth noting that $x_i \in \{0, 1\}$ is equivalent to $x_i - x_i^2 = 0, \forall i$. Moreover, the MP-based framework focus on solving $(\mathcal{P}_{\mathsf{MPMM}})$ iteratively and updating penalty parameters $\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}$ to drive the solution of $(\mathcal{P}_{\mathsf{MPMM}})$ to a point that $|x_i - x_i^2|$ is closed to zero. Here, the binary constraint is strengthened by adding $0 \leq x_i \leq 1$, i.e., $x_i^2 - x_i \leq 0$, $\forall i$ to the convex set $\mathcal{X}$. According to the MP-based framework given in [44], the penalty parameters can be updated as

$$[\eta^{[\ell+1]}]_i = [\eta^{[\ell]}]_i + \rho^{[\ell]}\left(x_i^{[\ell]} - \left(x_i^{[\ell]}\right)^2\right) \text{ and } \rho^{[\ell+1]} = \beta^{[\ell]}\rho^{[\ell]}, \tag{30}$$

where $\beta^{[\ell]}$ is the parameter to update ρ step by step. Usually, $\rho^{[\ell]}$ would be initialized with a small value and then increased with iterations. Note that $\rho^{[\ell]}$ can also keep fixed after certain iterations [33]. Due to the high order of variable $\mathbf{x}$ appearing in the objective function, it is very challenging to solve $(\mathcal{P}_{MPMM})$ directly. To overcome this issue, we employ the MM method by constructing a surrogate function and then finding the optimum solution in a sequence. The

Algorithm 3 MPMM ALGORITHM

Initialization: $\ell = 1, \boldsymbol{\eta}^{[1]} = \mathbf{0}, \rho^{[1]} = 1, \mathbf{x}_{(0)}^{[1]} = \mathbf{0}$
repeat
  Set $k = 0$.
  repeat
    Solve $\mathbf{x}_{(k+1)}^{[\ell]} = \arg\min_{\mathbf{x}\in\mathcal{X}} u\left(\mathbf{x} \mid \mathbf{x}_{(k)}^{[\ell]}, \boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right)$.
    Update $k = k + 1$.
  until Convergence
  Set $\mathbf{x}^{[\ell]} = \mathbf{x}_{(k)}^{[\ell]}$.
  Update $[\eta^{[\ell+1]}]_i = [\eta^{[\ell]}]_i + \rho^{[\ell]}\left(x_i^{[\ell]} - \left(x_i^{[\ell]}\right)^2\right)$, $\rho^{[\ell+1]} = \beta^{[\ell]}\rho^{[\ell]}$.
  Update $\ell = \ell + 1$.
until Convergence

proposed algorithm is summarized in Algorithm 3. The surrogate function is constructed by linearizing the binary constraint with its first-order Taylor series and is given by

$$\mathbf{x}_{(k+1)}^{[\ell]} = \arg\min_{\mathbf{x}\in\mathcal{Y}} u\left(\mathbf{x} \mid \mathbf{x}_{(k)}^{[\ell]}, \boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right) \tag{31}$$

where $u\left(\mathbf{x}|\mathbf{x}_{(k)}^{[\ell]},\boldsymbol{\eta}^{[\ell]},\rho^{[\ell]}\right) = \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x} + \sum_i \eta_i^{[\ell]}\left(1 - 2\left(x_i\right)_{(k)}^{[\ell]}\right)x_i + \frac{\rho^{[\ell]}}{2}\sum_i\left[\left(1 - 2\left(x_i\right)_{(k)}^{[\ell]}\right)x_i + \left(\left(x_i\right)_{(k)}^{[\ell]}\right)^2\right]^2$.

4) Convergence Analysis: Regarding the same approach analyzing the convergence of the MP method given in [33], one can prove the convergence of the proposed algorithm by addressing two facts: i) for given $\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}$, the MM procedure converges to the global optimum of $\mathbf{x}^{[\ell]}$; ii) the sequence $\{\boldsymbol{\eta}^{[\ell]}\}$ updated as in (30) converges to a fixed Lagrangian multiplier of $\tilde{\mathcal{P}}_1$.

Lemma 2: $\exists a \geq a_0$ such that the objective function $f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right)$ is convex.

The proof is presented in Appendix B. Then the convergence of sequence $\{\mathbf{x}_{(k)}^{[\ell]}\}$ will converge to the global optimum of $\mathbf{x}^{[\ell]}$.

C. Greedy Algorithm

In this section, we propose a heuristic approach to solving $(\tilde{\mathcal{P}}_1)$. The basic idea is to firstly solve the relaxed problem, i.e. $(\mathcal{P}_{\mathsf{rlx}})$, in which the binary variables are relaxed as continuous one, i.e.

$$(\mathcal{P}_{\mathsf{rlx}}) : \underset{\mathbf{x}\in\mathcal{Y}}{\mathtt{minimize}} \quad \mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x}. \tag{32}$$

Algorithm 4 GREEDY ALGORITHM

1: Round $\mathbf{x}^{\mathsf{con}}$ into binary solution, named $\mathbf{x}^{\mathsf{bin}}$, by setting d_n highest elements of the set $\left\{x_{n,t}^{\mathsf{con}}\right\}_{t=1}^{M}$ to 1 and setting the remaining elements to 0.
2: Calculate $\mathbf{B} \in \mathbb{R}^{N\times M}$ as $B_{n,t} = \sum\limits_{j\neq n}\left(\tilde{A}_{j,n} + \tilde{A}_{n,j}\right) x_{j,t}^{\mathsf{bin}}, \quad \forall (n,t).$
3: **while** $\mathbf{x}^{\mathsf{bin}}$ is not a feasible solution **do**
4: Denote $\mathcal{S}^{+} = \left\{ t \middle| 1 \leq t \leq M, \sum\limits_{n=1}^{N} x_{n,t}^{\mathsf{bin}} > K \right\}$, $\mathcal{S}^{-} = \left\{ t \middle| 1 \leq t \leq M, \sum\limits_{n=1}^{N} x_{n,t}^{\mathsf{bin}} < K \right\}$.
5: Let $\mathcal{T} = \left\{(n,t,u) | t \in \mathcal{S}^{+}, u \in \mathcal{S}^{-}, x_{n,t}^{\mathsf{bin}} = 1, x_{n,u}^{\mathsf{bin}} = 0\right\}$.
6: Solve $(n^*, t^*, u^*) = \arg\min_{(n,t,u)\in\mathcal{T}} B_{n,u} - B_{n,t}$.
7: Swap the values of $x_{n^*,t^*}^{\mathsf{bin}}$ and $x_{n^*,u^*}^{\mathsf{bin}}$ by setting $x_{n^*,t^*}^{\mathsf{bin}} = 0, x_{n^*,u^*}^{\mathsf{bin}} = 1$.
8: **end while**

The continuous optimal out-comes, denoted as $\mathbf{x}^{\mathsf{con}}$, are then rounded to binary solution, $\mathbf{x}^{\mathsf{bin}}$. The rounding mechanism is developed so that all the practical requirements of $(\tilde{\mathcal{P}}_1)$ are guaranteed. In addition, the approach also aims to minimize the total penalty. In particular, after solving the problem $(\mathcal{P}_{\mathsf{rlx}})$, for each user n, the first d_n highest elements among the set $\{x_{n,t}^{\mathsf{con}}\}_{t=1}^{M}$ are set to be ones while the others are down-rounded to zeros. Then, the illumination of beams over TSs can be further swapped to satisfy constraint (C_5) and keep the total penalty as small as possible. The greedy algorithm is presented in Algorithm 4 and described in Fig. 2.

D. Complexity Analysis

In this section, the complexity of our proposed algorithms are investigated based on the number of required operations.

1) Threshold-based Clustering Algorithm: Let $B_t = |\mathcal{B}^{\mathsf{ac}}[t]|$ be the number of activated beam in TS t where $|\mathcal{X}|$ stands for the cardinal number of set $\mathcal{X}$. As can be observed, Algorithm 1 consists of three loops, i.e., one "**repeat**" loop and two "**for**" loops, and it initializes with B_t clusters each of which contains one beam. In each iteration of the "**repeat**" loop, the number of clusters decreases if $\Phi[t]$ changes. Moreover, the "**repeat**" loop stops when there is no change of $\Phi[t]$. Therefore, the iteration number of the "**repeat**" loop must be less than B_t. Regarding "**for**" loops, we can observe that for each couple (k,l) in $\phi[t]$, one has to compare $\omega_{i,j}$ to κ for all $(i,j) \in \mathcal{C}_l \times \mathcal{C}_k$. Then, according to $|\phi[t]|, |\mathcal{C}_k| \leq B_t$ for all $k \in \mathcal{B}^{\mathsf{ac}}[t]$, the complexity of Algorithm 1 can be estimated as $\mathcal{O}\left(B_t^5\right)$.

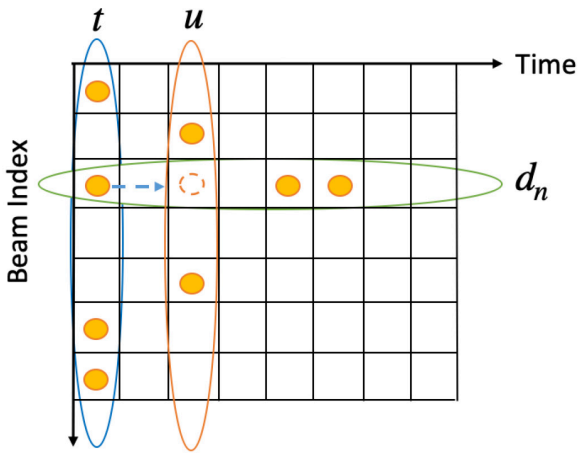


Fig. 2: The process of the Greedy Algorithm to design the illumination pattern. The demand of TSs of user n is d_n. At TS t, the number of active beams is greater than the limitation K, while the number is less than K at TS u. The exchange of active beam for user n will happen if this exchange would bring in a minimal increase of the objective function.

2) SDP-based Algorithm: As given in [46], the computational complexity involved in solving the SDP is $\mathcal{O}(\max(m,n)^4 n^{1/2})$ where n and m are the numbers of variables and constraints, respectively. As can be observed, these numbers corresponding to problem $(\mathcal{P}_{\mathsf{SDP}})$ are $(MN)^2 + MN$ and $M+N+1$, and $(MN)^2 + MN$ must be much greater than $M+N+1$. Therefore, the complexity of solving problem $(\mathcal{P}_{\mathsf{SDP}})$ by employing SDP method can be estimated as $\mathcal{O}\left([MN(MN+1)]^{4.5}\right)$.

3) MPMM Algorithm: MPMM algorithm consists of two loops where the inner loop attempts to solve problem (31) in each inner iteration while the outer loop aims to update the penalty parameters in each step as in (28). Generally, problem (31) is a convex QP with MN variables which can be solved in polynomial time with the complexity of $\mathcal{O}\left((MN)^3\right)$ [47]. Hence, the complexity of MPMM algorithm can be expressed based on the number of iterations as $I_{\mathsf{MPMM}}^{\mathsf{outer}}\left(I_{\mathsf{MPMM}}^{\mathsf{inner}}\mathcal{O}\left((MN)^3\right)+\mathcal{O}\left(MN\right)\right)$ where $I_{\mathsf{MPMM}}^{\mathsf{outer}}$ and $I_{\mathsf{MPMM}}^{\mathsf{inner}}$ are the average iteration numbers of outer and inner loops required in Algorithm 3 to solve problem $(\tilde{\mathcal{P}}_1)$, respectively.

4) Greedy Algorithm: The initial step of this algorithm attempts to solve the QP $(\mathcal{P}_1)$ with continuous variable $\mathbf{x}$ with the complexity of $\mathcal{O}\left((MN)^3\right)$ [47]. Then, it requires to calculate elements of matrix $\mathbf{B}$ before iteratively updating sets $\mathcal{S}^+$, $\mathcal{S}^-$, and $\mathcal{T}$. At the end of each iteration, the comparison procedure is processed to select the two specific beams in two different TSs for

swapping. It is worth noting that the cardinality of $\mathcal{T}$ cannot exceed MN and decrease after every iteration. Hence, the number of iterations in Algorithm 4 must be smaller than MN. Hence, the complexity of the greedy algorithm can be estimated based on that due to solving QP, calculating $\mathbf{B}$ and iterative process as $\mathcal{O}\left((MN)^3\right) + MN + (MN)^2 = \mathcal{O}\left((MN)^3\right)$.

VI. Numerical Results

Monte Carlo simulations are conducted to evaluate the performance of the proposed three algorithms in terms of demand matching, the total computation for precoding, and an average number of active beams per TS, compared with two benchmarks: conventional beam hopping (BH) and cluster hopping (CH). In particular, the conventional BH method is given in Appendix A while CH benchmark solution is proposed in [29, 30].

A. Simulation Setup

We consider a GEO satellite system with 67 spot beams, i.e., $N = 67$. The setting parameters are summarized in Table II. The simulation setup is the same as the one considered in [28, 30]. Unless mentioned otherwise, the number of TSs is set to $M = 20$. The traffic demand of all the users is generated uniformly at random between $400r$ and $1500r$ (Mbps), i.e., $400r \leq D_n \leq 1500r \ \forall n$. Herein, r represents the demand-density factor which is selected in $\{0.25, 0.3, 0.35, 0.4, 0.45\}$ where $r = 0.25$ implies the low demand setting while $r = 0.45$ refers to the high demand. For each selected value of r, 50 demand instances are generated for testing. A single representative user within each beam is assumed, which aggregates the overall beam demand.

B. Estimating Number of Time-Slots per-Beam Required to Satisfy Demand

Fig. 3 illustrates the convergence of Algorithm 2 where the evolution in terms of average numbers of activated beams per TS, i.e., K_{avg}, is shown with respect to iterations. In Fig. 3, we have illustrated the convergence for five demand instances based on the values of $r \in \{0.25, 0.3, 0.35, 0.4, 0.45\}$. It can be observed that the algorithm converges after $2-3$ iterations, where K_{avg} increases before saturating at constant values. Moreover, Fig. 3 also shows that the larger the value of r (i.e. the higher the demand), the higher the number of average numbers of activated beams per TS. According to Remark 2, the users' demand cannot be satisfied if K is less than K_{avg}; hence, unless mentioned otherwise, we set $K = K_{avg}$ in the subsequent simulations.

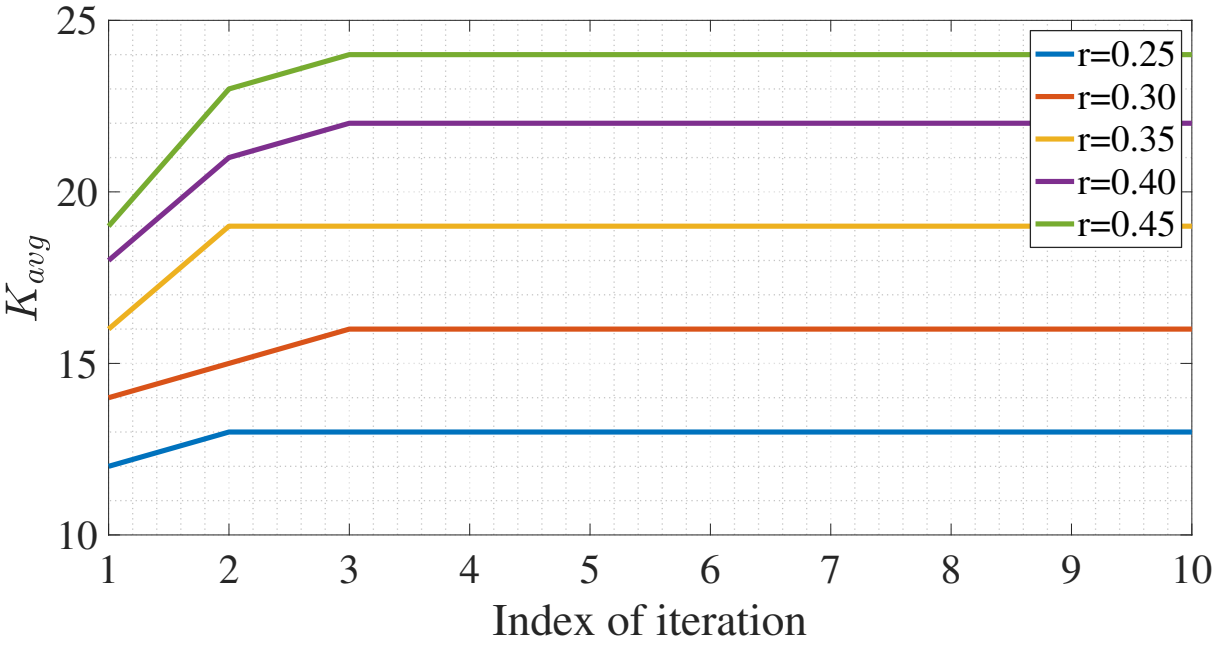


Fig. 3: Average number of active beams per TSs.

TABLE II: Simulation Parameters

Satellite Orbit	13°E (GEO)
Satellite Beam Power	80 W
OBO	3 dB
Addition Payload Loss	2 dB
Number of Virtual Beams (N)	67
Beam Radiation Pattern ($G_i(x,y)\, e^{j\phi_{x,y}}$)	Provided by ESA
Downlink Carrier Frequency	19.5 GHz
User Link Bandwidth, B	500 MHz
Roll-off Factor	20%
Temperature	50 K
Number of TSs (M)	20
Threshold do precoding(κ)	0.08

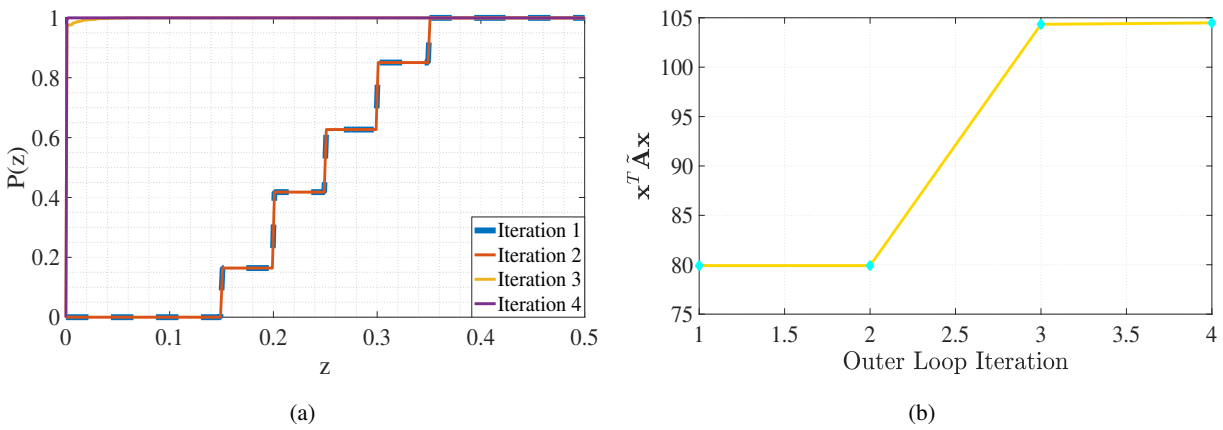
TABLE III: Clusters' Distribution ($r = 0.25$)

Size	SDP	MPMM	Greedy	BH	CH
1	4494	12549	12225	13061	0
2	1801	36	119	0	0
3	795	0	27	0	0
4	264	0	9	0	0
5	100	0	4	0	0
6	53	0	1	0	1814
7	25	0	1	0	186
8	8	0	1	0	0
9	3	0	0	0	0

[1] The total relative computation for precoding is 1.0000, 0.0031, 0.0418, 0.0000 and 4.9754 in the order of SDP, MPMM, Greedy, BH and CH.

C. Discussion on MPMM Convergence

1) The Convergence of MPMM Algorithm: In Fig. 4, we regard the convergence of Algorithm 3. In order to illustrate the convergence, three parameters are considered, i) $gap(x_i) = \min\{|x_i - 0|, |x_i - 1|\}$ which describes the minimal distance between the continuous element x_i and a binary variable; ii) $T(z) = \{x_i |\ gap(x_i) \leq z, i = 1, \ldots, MN\}$ is the set of elements



(a) (b)

Fig. 4: The Convergence of MPMM. (a) Geometric distribution of the elements of $\mathbf{x}$; (b) Evolution of the objective function $\mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x}$.

which belongs to the variable $\mathbf{x}$, whose $gap\,(x)$ is not greater than z; iii) $P\,(z) = \frac{|T(z)|}{MN}$ describes the percentage of the elements in variable $\mathbf{x}$ whose $gap\,(x)$ is less than z.

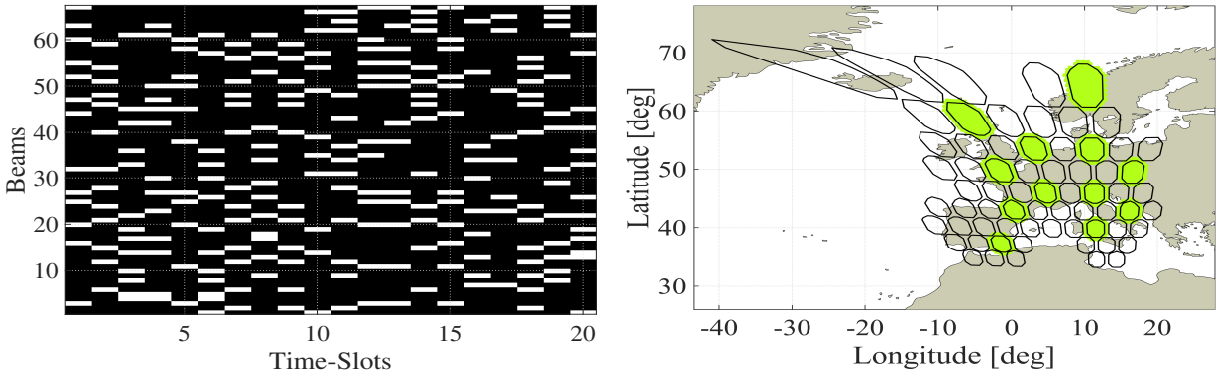
Fig. 4a shows the geometric distribution of elements in $\mathbf{x}$, i.e. $P\,(z)$ achieved in each outer-loop iteration. For the sake of clarity, in this simulation a single demand factor r is considered, being $r = 0.3$. It can be observed that the elements in $\mathbf{x}$ are closer to binary values after each iteration. As can be observed, the curves corresponding to iteration 3 and 4 illustrate that $P\,(z)$ is close to one with very small gaps. Certainly, this has confirmed that the final solution converges to binary variables.

Fig. 4b shows the variation of the objective function of problem $(\tilde{\mathcal{P}}_1)$, i.e., $\mathbf{x}^T\tilde{\mathbf{A}}\mathbf{x}$, achieved in every outer-loop iteration. The algorithm initiates with penalty parameters $\boldsymbol{\eta}^{[1]} = \mathbf{0}, \rho^{[1]} = 1$ which will increase after each iteration.

D. Performance Evaluation

In this section, we aim to evaluate the performance of the proposed algorithms in terms of two main aspects: (i) per-beam demand matching, and (ii) a number of beams that would require the implementation of precoding to deal with co-channel interference.

To begin with, Fig. 5a shows an example of an illumination pattern design obtained by implementing Algorithm 3 for a particular demand instance obtained with $r = 0.25$. In this figure, the white rectangles imply that the corresponding beams are illuminated while the blacks refer to the inactive ones in a specific TS. In addition, the illumination map of beams corresponding to



(a) (b)

Fig. 5: Example of Illumination Pattern obtained with MPMM. (a) Satellite illumination pattern; (b) Beam pattern at the 15-th TS.

TABLE IV: Clusters' Distribution $(r = 0.35)$

Size	SDP	MPMM	Greedy	BH	CH
1	3895	16562	14774	16865	0
2	1693	514	1207	0	0
3	1024	80	134	0	0
4	532	29	41	0	0
5	292	6	19	0	0
6	216	5	7	0	1814
7	119	1	8	0	186
8	91	0	0	0	0
9	51	0	3	0	0
10	35	0	2	0	0
11	16	0	1	0	0
12	10	0	1	0	0
13	5	0	0	0	0
14	1	0	0	0	0
15	1	0	0	0	0
16	1	0	0	0	0

[1] The total relative computation for precoding is 1.0000, 0.0273, 0.0788, 0.0000, and 1.2061 in the order of SDP, MPMM, Greedy, BH, and CH.

TS 15 of this simulation is illustrated in Fig. 5b, where the green areas represent the foot-prints of the illuminated beams.

Next, we consider the precoding utilization in the proposed algorithms in Table III, IV and V, which will have an impact on the system complexity. It is worth noting that the complexity

TABLE V: Clusters' Distribution ($r = 0.45$)

Size	SDP	MPMM	Greedy	BH	CH
1	3066	13076	11265	16963	0
2	1270	1585	2222	0	0
3	893	798	1028	0	0
4	532	432	425	0	0
5	373	171	207	0	0
6	267	148	114	0	1814
7	174	81	64	0	186
8	147	35	44	0	0
9	143	33	28	0	0
10	109	18	18	0	0
11	85	13	13	0	0
12	66	9	5	0	0
13	55	5	2	0	0
14	47	6	4	0	0
15	46	3	5	0	0
16	32	0	1	0	0
17	19	0	1	0	0
18	9	0	0	0	0
19	12	0	0	0	0
20	2	0	0	0	0
21	3	0	1	0	0
22	4	0	1	0	0
23	1	0	0	0	0

[1] The total relative computation for precoding is 1.0000, 0.1697, 0.1820, 0.0000, and 0.2829 in the order of SDP, MPMM, Greedy, BH, and CH.

TABLE VI: Average Number of Active Beams

	SDP	MPMM	Greedy	BH	CH
r=0.25	12.62	12.62	12.62	13.06	12.19
r=0.35	18.01	18.01	18.01	16.87	12.19
r=0.45	23.88	23.88	23.88	16.96	12.19

TABLE VII: Jain's Fairness Index

	SDP	MPMM	Greedy	BH	CH
r=0.25	0.9892	0.9852	0.9873	0.9842	0.8582
r=0.35	0.9924	0.9916	0.9911	0.9760	0.8582
r=0.45	0.9952	0.9955	0.9952	0.9651	0.8583

for MMSE-based precoding of a cluster of N beams is estimated as $O\left(N^3\right)$ in general [14, 36]. Therefore, we aim to demonstrate the precoding complexity corresponding to different BH mechanisms by illustrating the number of clusters with different sizes. In particular, Tables III, IV and V show the distribution of clusters in various sizes at different demand instances assuming $r = 0.25, 0.35$ and 0.45, respectively. In addition, based on the numbers given in these tables, the total computation for precoding due a specific BH mechanism can be estimated as $\mathcal{T} = \sum_{n_i=2}^{N} m_i \cdot (n_i)^3$, where n_i represents the size of clusters and m_i represents the total number of the corresponding clusters at a demand density instance. At the footnote of these tables, we show the total relative computation cost for precoding by comparing $\mathcal{T}$'s, where SDP is set as the baseline. Particularly, the precoding complexity due to a BH method is defined as the ratio of its $\mathcal{T}$ to that of SDP.

For the three proposed algorithms, these tables have demonstrated that: (i) the number of larger-size clusters is smaller, (ii) increasing the traffic demand results in a higher number of larger-size clusters. Interestingly, the MPMM method shows its superiority when it is on the smaller size of the cluster which would result in less computation for precoding. When $r = 0.25$, the total computation for precoding with the SDP method is more than 300 times that of MPMM. In addition, no adjacent beams will be illuminated simultaneously with the BH method, then there is an upper bound of the maximum number of activated beams. So the total number of active beams will not change much with the increase of the density of demands. The CH method predefines the clusters where the cluster size is 6 or 7.

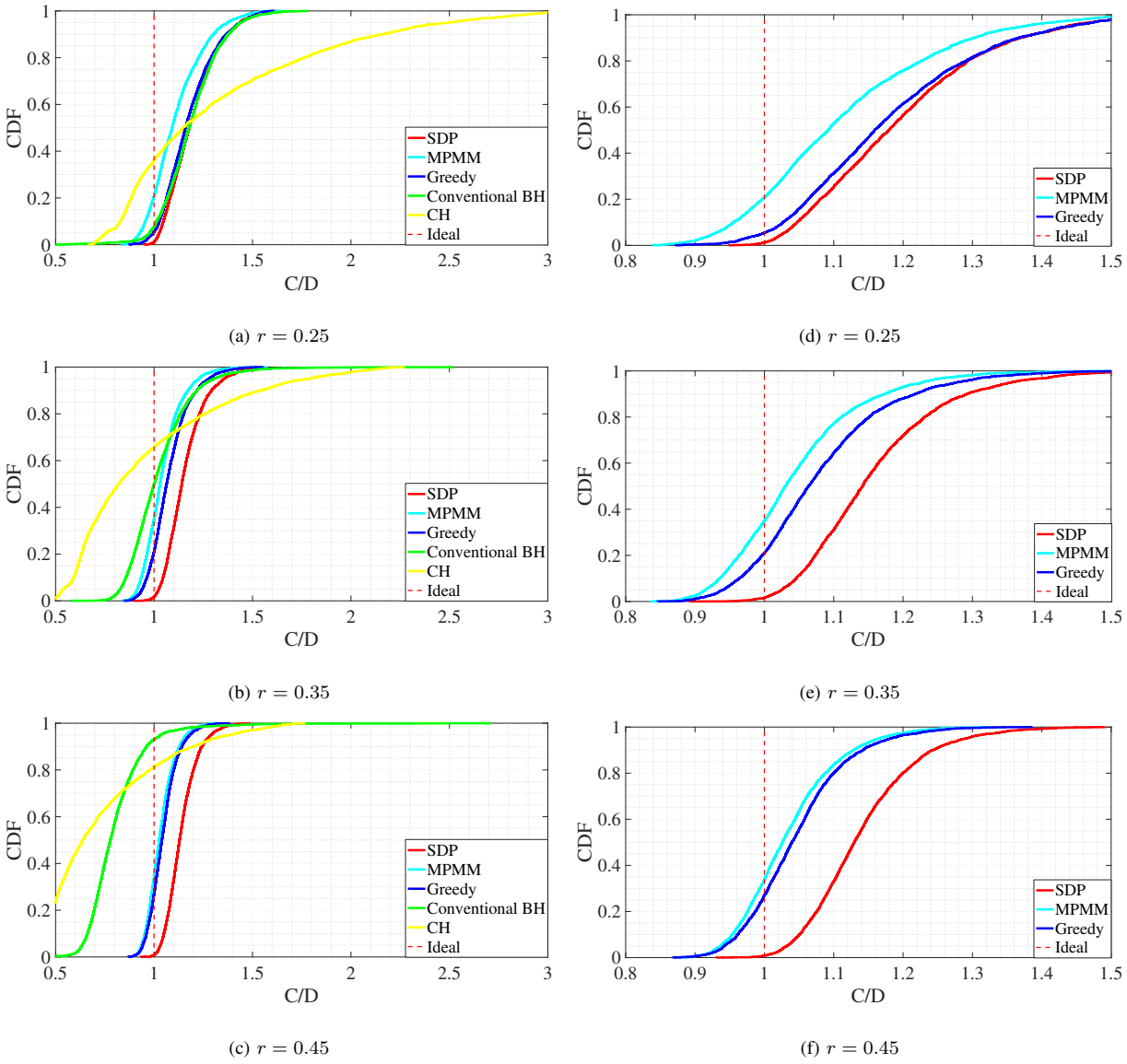
Next, we analyze the average number of active beams per TS, which determines the resulting interference as well as the operating power consumption of the satellite. In principle, one would like to minimize the number of active beams but make sure that the demand requirements are met. Table VI shows the average number of beams activated per TS for the different methodologies. For the proposed algorithms, the number of activated beams per TS is fixed and given by Algorithm 2, while the conventional methods provide different values. For the CH technique, the number of active beams is fixed and does not depend on the demands, which typically results in an inaccurate demand-matching performance. Regarding the BH technique, the number of active beams slightly increases as the demand increases, but the illumination design is limited to non-adjacent beams, and therefore the increase in the number of active beams is not so prominent. Unlike the benchmarks, the proposed techniques are more flexible in activating more beams and adapting to the demand increases. The results in Table VI match the distribution of cluster

numbers with different sizes depicted in Tables III, IV and V.

To evaluate the fairness of users' satisfaction corresponding to the proposed and benchmark methods, we consider Jain's Fairness Index proposed in [48]. The definition of the index is given as $\mathcal{J}(\mathbf{y}) = \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n \cdot \sum_{i=1}^{n} y_i^2}$ where y_i is the chosen metric and is given by $y_i = \frac{c_i}{g_i}$ in which c_i is defined as $\frac{1}{M}\sum_{t=1}^{M} R_n[t]$. This index aims to determine whether users are receiving a fair demand matching or not. "One" value of $\mathcal{J}(\mathbf{y})$ implies the highest fairness level among all users. The Jain's fairness indices achieved by implementing various BH mechanisms, the proposed and benchmark methods, for different demand factors are given in Table VII. From the results, it can be concluded that all three proposed methods can provide better fair indices than the benchmark, especially in the high-demand scenarios. In addition, it can be observed that the CH method stays at around 0.86 independent of the demand entry. Furthermore, Jain's index of the conventional BH method suffers from maintaining a good level of fairness as the demand increases.

In the subsequent results, we focus our evaluation on the capabilities of the proposed techniques to match the offered capacity with the actual demand. In particular, Fig. 6 illustrates the cumulative distribution function (CDF) of C/D – the ratio of the provided capacity of the beam to its required demand, for 3 different demand factors $r = 0.25, 0.35$ and 0.45. On the top of Fig. 6, we show the performance of the proposed methods with respect to the benchmarks. On the bottom of Fig. 6, the reader can find a zoom-in figure to better discern the performance of the proposed techniques. The vertical dashed line indicates the ideal scenario where $c_i/g_i = 1$, $\forall i$.

First, Fig. 6 confirms that the CH technique suffers from the limitation of pre-defined clustering shapes, which unavoidably illuminate low-demand beams with high-demand beams. On the other hand, the conventional BH is shown to experience significant degradation when the demand factor increases. This is because BH falls short of supplying enough capacity due to its inability to illuminate high-demand areas at once. Focusing on the proposed techniques, we can see from Fig. 6(b) and Fig. 6(c) that all three outperform the benchmarks, especially for moderate and high demand factors. It can also be observed that the SDP-based method provides $c_i/g_i > 1$ for almost all cases which implies that this approach can provide a capacity larger than the demand. It may not be expected in some specific circumstances that one avoids spending expensive network resources to serve the users much more than what satisfies them. The MPMM approach seems to provide a better trade-off as its curve is closer to the ideal case. The greedy algorithm provides performance in between SDP and MPMM methods and seems to be closer to the SDP solution



(a) $r = 0.25$ (d) $r = 0.25$

(b) $r = 0.35$ (e) $r = 0.35$

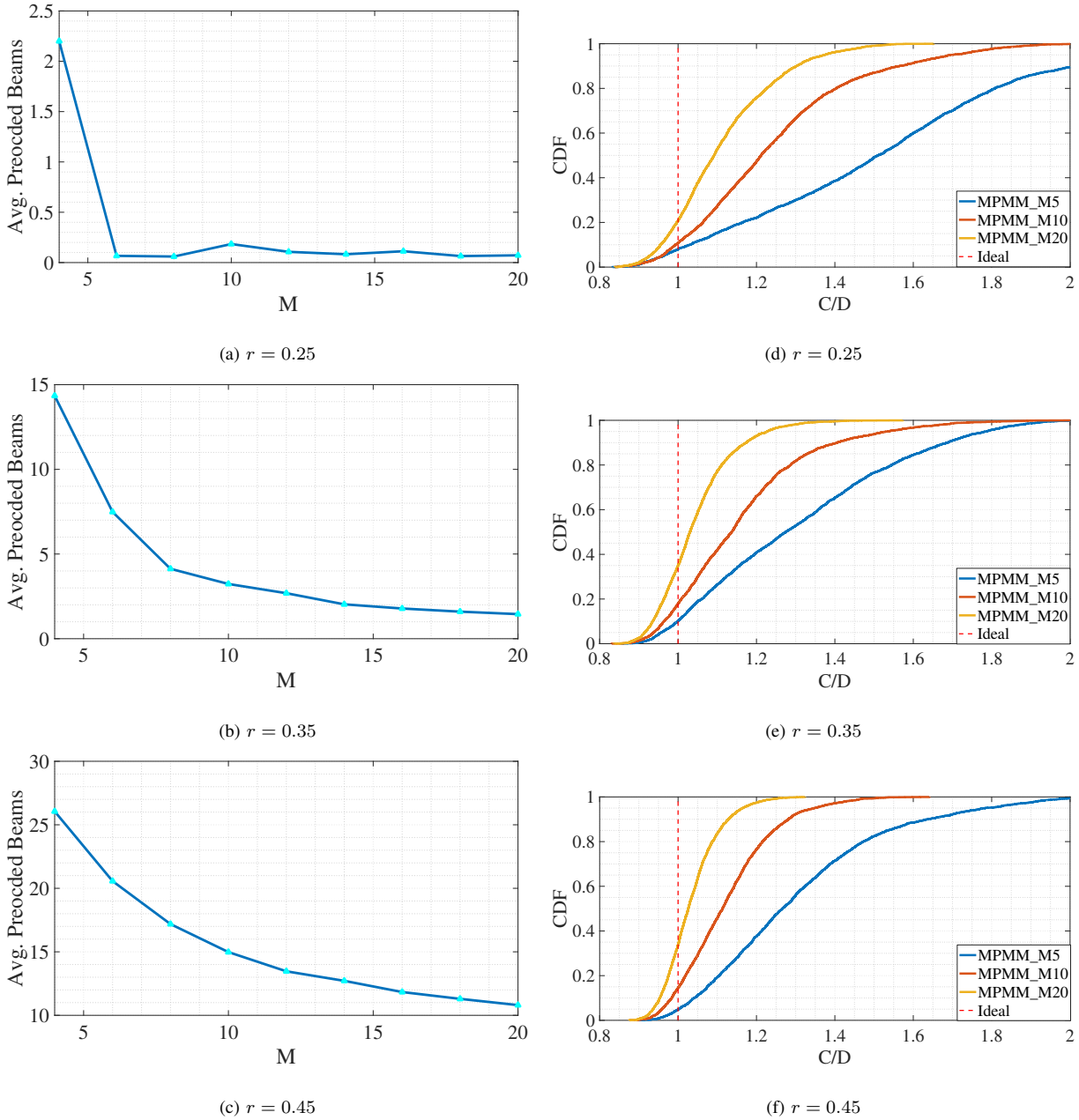
(c) $r = 0.45$ (f) $r = 0.45$

Fig. 6: Comparison of algorithms in demand matching at different density of demands.

for the low demand factor while it approaches the MPMM solution for the high demand factor.

E. Impact of number of TSs in BH window

Herein we evaluate the impact of the parameter M, which determines the number of TSs within a BH window. In particular, Fig. 7 evaluates the number of average precoded beams within a hopping window with respect to M shows the CDF of (c_i/g_i) for different values of M. In Fig. 7, we focus on the MPMM method's behavior, which was found to be the best in



(a) $r = 0.25$ (d) $r = 0.25$

(b) $r = 0.35$ (e) $r = 0.35$

(c) $r = 0.45$ (f) $r = 0.45$

Fig. 7: Comparison of M in terms of average number of active beams and demand matching at different density of demands.

demand matching fairness among users in the previous simulation result. As usual, we evaluate three different demand factors, i.e. $r = 0.25, 0.35$, and 0.45. From Fig. 7(a), Fig. 7(b), and Fig. 7(c), it can be observed that a higher value of M translates into a lower average number of precoded beams. This is an expected result where the more TSs are available, the less number of beams need to be activated simultaneously. Another interesting result is that the average number

of precoded beams increases with the users' demand. Focusing now on the CDF curves, depicted in Fig. 7(d), Fig. 7(e), and Fig. 7(f), we can observe that the longer window length can push the achievable capacity closer to the users' demand. This is because of that the higher value of M gains a higher degree of freedom in selecting the illuminating TSs for each beam to meet their demand.

F. Impact of imperfect CSI

The simulations above are based on perfect CSI which is unrealistic in practice. In this subsection, following [49], we model the CSI uncertainty with an additive complex Gaussian error with parameters (mean, standard deviation) as shown in Table VIII. Note that (I/N) in Table VIII denotes the Interference-to-Noise Ratio, which is a measurement of how strong are the signals (coming from different beams) to be measured.

Table IX compares the performance with perfect CSI and imperfect CSI in terms of estimated demand (d_n) and Jain's fairness index for different demand densities r, where the results are averaged for 50 Monte Carlo simulation. The only difference between each simulation is the channel information, one of which is with perfect channel and the other is with estimated channel. The first thing we observe is that the estimated average demand d_n is lower with imperfect CSI. The latter occurs due to the nullification of certain CSI components in the imperfect CSI (note that I/N lower than -10dB are not measured at all). The nullification of the channel matrix translates into a reduction of the assumed interference levels. This biased demand estimation could be compensated by modifying the way we calculate the average demand, maybe adding a margin, but this is out of the scope of this work. When comparing the values of Jain's fairness index in Table IX, we can observe an evident performance loss for the imperfect CSI case, which is justified essentially by the reduced estimated demand. Fig. **??** illustrates the demand matching for the same cases evaluated in Table IX for completeness. As expected, the figure shows that the perfect CSI scheme outperforms the imperfect CSI one where it can supply more beams as their demands than the other.

G. Impact of Random User Location

The assumption of a single virtual user per beam is performed to abstract the user scheduling. However, the assumed location of such a virtual user may have some impact on the final performance. For instance, having the virtual user on the beam edge will not have a strong

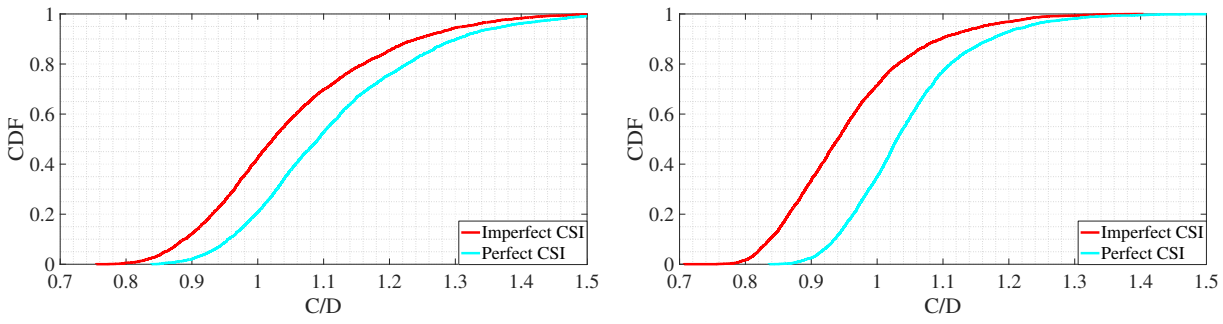
TABLE VIII: Mean and Standard Deviation of the channel estimate (reference signal SNR=10 dB) [49]

I/N	Amplitude		Phase	
[dB]	Mean	Std.	Mean(°)	Std.
10	0	0	0	0
5	0.02	0.005	0	0.2
0	0.04	0.010	0	0.5
-5	0.10	0.020	0	1.0
-10	0.16	0.030	3	2.0
-11	No lock	No lock	No lock	No lock

TABLE IX: Comparison between Perfect and Imperfect CSI

r	Avg. Demand			Jain's Index	
	Imperfect	Perfect	Percentage	Imperfect	Perfect
0.25	11.95	12.62	0.9470	0.9831	0.9852
0.35	16.39	18.01	0.9098	0.9883	0.9916
0.45	21.03	23.88	0.8808	0.9915	0.9955

[1] The percentage is the ratio of estimated average demand in the number of TSs with imperfect CSI to that with perfect CSI.



(a) (b)

Fig. 8: Comparison between perfect and imperfect channels in demand matching at different demand densities.

TABLE X: Comparison in Jain's Index

r	Jain's Index	
	Random	Centered
0.45	0.9878	0.9955
0.50	0.9839	0.9961
0.55	0.9775	0.9961
0.60	0.9691	0.9956

impact whenever the active beam is isolated. However, for high-demand instances, we expect the edge users to impact the selective precoding and generate a higher miss-match between the

estimated capacity and the actual supplied capacity. For the latter, estimating the capacity with a user on the edge will provide lower capacity than a user in the beam center, therefore requiring more TSs to satisfy the demand.

To evaluate this, we have run some results by randomly selecting the virtual user location within its -3 dB beamwidth. For the sake of comparison purposes, for each instance of random user location, the same beam traffic demand (in bps) as the user in the beam center is assumed. In addition, there are 50 instances, each of whose demands are randomly generated. Table X compares the performance in terms of demand satisfaction between random users and centered users at different densities of demand. It can be observed that, as the demand r increases, the error in the demand matching increases when the non-centered virtual user is considered.

VII. Conclusion

In this paper, we propose an analytical framework, a class of BQP problems, to support dynamic beam illumination design considering selective precoding for the next generation of time-flexible satellite broadband systems. Three algorithms are proposed to solve the problem: (i) SDP-based approach, (ii) MPMM methodology, and (iii) low-complexity greedy algorithm. All three methods target cross-beam interference minimization, such that the number of beams that need to be precoded is kept to a minimal in an attempt to reduce system complexity.

An extensive evaluation has been carried out based on numerical simulations. The results have shown interesting gains provided by the proposed algorithm with respect to the relevant benchmark schemes. In particular, the proposed framework provides an efficient solution to deal with high-demand areas while keeping the precoding-related complexity low.

Appendix A
Conventional Beam-Hopping Method

The conventional BH method is one of the methods to design the illumination pattern and is developed by solving the following problem

$$(\mathcal{P}_{\mathsf{Conv}}) : \underset{\mathbf{X},\, t}{\texttt{maximize}} \quad t \tag{33a}$$

$$\texttt{subject to} \quad (S_1) : \sum_{i=1}^{M} \mathbf{x}_i \odot \boldsymbol{\zeta} M \succeq t \cdot \boldsymbol{\Delta} \tag{33b}$$

$$(S_2) : X\left(i,:\right) + X\left(j,:\right) \preceq \mathbf{1}_M^T, \forall \left(i,j\right) \in \mathcal{B} \tag{33c}$$

$$(S_3): \mathbf{1}_N^T \mathbf{X} \preceq K \cdot \mathbf{1}_M^T \tag{33d}$$

$$(S_4): \mathbf{x}_i \in \{0,1\}^N \tag{33e}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M]$ and $\mathcal{B} = \{(i,j)\,|Q_{i,j} = 1, i = 1, \cdots, N; j = 1, \cdots, N\}$, $\mathbf{Q}$ is the adjacent matrix of the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\boldsymbol{\zeta} \in \mathcal{R}^N$ represents the estimated capacity for all the beam, which is given by (14) where no inter-beam interference is considered. Herein, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is defined as follows. Each beam center is considered as a vertex $v \in \mathcal{V}$ and any two vertices are connected with an edge $e \in \mathcal{E}$ if those vertices represent geographically adjacent beams. To solve the problem, one needs to estimate the capacity $\boldsymbol{\zeta}$ first and then calculate the set $\mathcal{B}$ according to the adjacent matrix of the graph. In the last, the problem ($\mathcal{P}_{\mathsf{Conv}}$) could be solved by advanced optimization toolboxes such as CVX[41]. After calculating the illumination pattern $\mathbf{X}$, one could exactly calculate the capacity for each beam.

APPENDIX B
CONVERGENCE OF MM PROCEDURE

The Hessian matrix of $f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right)$ is given by

$$\begin{aligned}\nabla^2 f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right) =& \mathbf{A} + a\mathbf{I} - \mathtt{Diag}\left(2\eta_1^{[\ell]}, ..., 2\eta_{MN}^{[\ell]}\right) \\ &+ 6\rho^{[\ell]}\boldsymbol{\Phi} - \frac{\rho^{[\ell]}}{2}\mathbf{I}\end{aligned} \tag{34}$$

where $\mathtt{Diag}\,(*)$ represents the diagonal matrix operator, and $\boldsymbol{\Phi} = \mathtt{Diag}\left(\left(x_1 - \frac{1}{2}\right)^2, ..., \left(x_{MN} - \frac{1}{2}\right)^2\right)$.

Since the sequences $\{\rho^{[\ell]}\}$ and $[\boldsymbol{\eta}^{[\ell]}]_i = \sum_{k=0}^{\ell-1} \rho^{[k]} \left(x_i^{[k]} - \left(x_i^{[k]}\right)^2\right)$ are bounded, then $\exists a \geq a_0$ such that $\nabla^2 f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right) \succeq 0$. So $f\left(\mathbf{x}|\boldsymbol{\eta}^{[\ell]}, \rho^{[\ell]}\right)$ is convex. Additionally, the set $\mathcal{Y}$ is convex. Therefore, the resulting stationary point of the problem is the global optimum point.

REFERENCES

[1] L. Chen, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Satellite broadband capacity-on-demand: Dynamic beam illumination with selective precoding," in *European Signal Processing Conference (EUSIPCO)*, 2021, pp. 900–904.

[2] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55–61, 2020.

[3] United Nations, "Deputy secretary-general remarks at the general assembly," https://www.un.org/press/en/2021/dsgsm1579.doc.htm, Apr. 2020, [Online; accessed 21-Dec-2021].

[4] O. Kodheli *et al.*, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys Tutorials*, vol. 23, no. 1, pp. 70–109, 2021.

[5] 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," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 112–131, 2019.

[6] T. S. Abdu, S. Kisseleff, E. Lagunas, and S. Chatzinotas, "Flexible resource optimization for GEO multibeam satellite communication system," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7888–7902, 2021.

[7] NetWorld, "NetWorld2020 - Satellite working group, white paper, more information available at:,," https://www.networld2020.eu/wp-content/uploads/2019/11/satcom-ss-whitepaper_v1_final.pdf?x85154, accessed: 2022-06.

[8] European Space Agency, "Eutelsat Quantum, more information available at:," https://www.esa.int/Applications/Telecommunications_Integrated_Applications/Eutelsat_Quantum, 2021.

[9] Eutelsat Quantum, "Technical Capabilities of Eutelsat Quantum, more information available at:," https://issuu.com/eutelsat/docs/eutelsat_sat_quantum_brochure_1020, accessed: 2022-06.

[10] "SES-17 satellite, more information available at:," https://www.ses.com/press-release/ses-17-successfully-launched-ariane-5, accessed: 2022-06.

[11] H2020 DYNASAT project, "Dynamic Spectrum Sharing and Bandwidth-Efficient Techniques for High Throughput MIMO Satellite Systems," https://www.dynasat.eu/, Dec., 2020, [Online; accessed Jun-2022].

[12] H2020 ATRIA project, "AI Powered Ground Segment Control for Flexible Payloads," https://www.atria-h2020.eu/, 2020, [Online; accessed Jun-2022].

[13] P.-D. Arapoglou, A. Ginesi, S. Cioni, S. Erl, F. Clazzer, S. Andrenacci, and A. Vanelli-Coralli, "Dvb-s2x-enabled precoding for high throughput satellite systems," *International Journal of Satellite Communications and Networking*, vol. 34, no. 3, pp. 439–455, 2016. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1122

[14] M. A. Vazquez, A. Perez-Neira, D. Christopoulos, S. Chatzinotas, B. Ottersten, P. Arapoglou, A. Ginesi, and G. Tarocco, "Precoding in multibeam satellite communications: Present and future challenges," *IEEE Wireless Communications*, vol. 23, no. 6, pp. 88–95, December 2016.

[15] L. M. Marrero, J. Duncan, J. Querol, N. Maturo, J. Krivochiza, S. Chatzinotas, and B. Ottersten, "Differential phase compensation in over-the-air precoding test-bed for a multi-beam satellite," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 1325–1330.

[16] J. Krivochiza, J. C. M. Duncan, J. Querol, N. Maturo, L. M. Marrero, S. Andrenacci, J. Krause, and S. Chatzinotas, "End-to-end precoding validation over a live geo satellite forward link," *IEEE Access*, pp. 1–1, 2021.

[17] SatixFy Ltd., "Beam hopping: Ready technology," https://www.satixfy.com/beam-hopping/, Apr. 2020, [Online; accessed 21-Dec-2021].

[18] H. Fenech, A. Sonya, A. Tomatis, V. Soumpholphakdy, and J. L. S. Merino, *Eutelsat quantum: A game changer*. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2015-4318

[19] C. Rohde, R. Wansch, G. Mocker, A. Trutschel-Stefan, L. Roux, E. Feltrin, H. Fenech, and N. Alagha, "Beam-hopping over-the-air tests using DVB-S2X super-framing," in *36th International Communications Satellite Systems Conference (ICSSC 2018)*, 2018, pp. 1–7.

[20] A. Freedman, D. Rainish, and Y. Gat, "Beam hopping how to make it possible," in *Proc. of Ka and Broadband Communication Conference, Bologna, Italy*, Oct. 2015.

[21] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," *Acta Numerica*, vol. 22, 2013.

[22] P. Angeletti, D. Fernandez, and R. Rinaldo, "Beam hopping in multi-beam broadband satellite systems: System performance and payload architecture analysis," in *AIAA International Communications Satellite Systems Conference*. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2006-5376

[23] G. Cocco, T. de Cola, M. Angelone, Z. Katona, and S. Erl, "Radio resource management optimization of flexible satellite payloads for DVB-S2 systems," *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 266–280, 2018.

[24] R. Alegre-Godoy, N. Alagha, and M. A. VÃązquez Castro, "Offered capacity optimization mechanisms for multi-beam satellite systems," in *2012 IEEE International Conference on Communications (ICC)*, 2012, pp. 3180–3184.

[25] J. Anzalchi, A. Couchman, P. Gabellini, G. Gallinaro, L. D'Agristina, N. Alagha, and P. Angeletti, "Beam hopping in multi-beam broadband satellite systems: System simulation and performance comparison with non-hopped systems," in *Advanced Satellite Multimedia Systems Conference(ASMS) and Signal Processing for Space Communications (SPSC) Workshop*, 2010, pp. 248–255.

[26] J. Lei and M. Ã. VÃązquez-Castro, "Multibeam satellite frequency/time duality study and capacity optimization," *Journal of Communications and Networks*, vol. 13, no. 5, pp. 472–480, 2011.

[27] L. Lei, E. Lagunas, Y. Yuan, M. G. Kibria, S. Chatzinotas, and B. Ottersten, "Beam illumination pattern design in satellite networks: Learning and optimization for efficient beam hopping," *IEEE Access*, vol. 8, pp. 136 655–136 667, 2020.

[28] ESA Project FlexPreDem, "Demonstrator of Precoding Techniques for Flexible Broadband Satellite Systems," https://artes.esa.int/projects/flexpredem, 2020, [Online; accessed 21-Dec-2021].

[29] M. G. Kibria, E. Lagunas, N. Maturo, D. Spano, and S. Chatzinotas, "Precoded cluster hopping in multi-beam high throughput satellite systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[30] E. Lagunas, M. G. Kibria, H. Al-Hraishawi, N. Maturo, and S. Chatzinotas, "Precoded Cluster Hopping for Multibeam GEO Satellite Communication Systems," *Frontiers in Signal Processing*, vol. 1, p. 6, 2021. [Online]. Available: https://www.frontiersin.org/article/10.3389/frsip.2021.721682

[31] P. Wang, C. Shen, A. v. d. Hengel, and P. H. S. Torr, "Large-scale binary quadratic optimization using semidefinite relaxation and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 3, pp. 470–485, 2017.

[32] "A mixed integer quadratic reformulation of the quadratic assignment problem with rank-1 matrix," in *International Symposium on Process Systems Engineering*, ser. Computer Aided Chemical Engineering, I. A. Karimi and R. Srinivasan, Eds. Elsevier, 2012, vol. 31, pp. 360–364.

[33] D. P. Bertsekas, "On penalty and multiplier methods for constrained minimization," *SIAM Journal on Control and Optimization*, vol. 14, no. 2, pp. 216–235, 1976.

[34] L. Wu, P. Babu, and D. P. Palomar, "Cognitive radar-based sequence design via sinr maximization," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 779–793, 2016.

[35] S. Kisseleff, E. Lagunas, J. Krivochiza, J. Querol, N. Maturo, L. M. Marrero, J. Merlano-Duncan, and S. Chatzinotas, "Centralized gateway concept for precoded multi-beam geo satellite networks," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, 2021, pp. 1–6.

[36] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part i: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, Jan 2005.

[37] A. Ginesi, E. Re, and P. Arapoglou, "Joint beam hopping and precoding in hts systems," in *9th Int. Conf. on Wireless and Satellite Systems (WiSATS)*, 2017.

[38] C. M. Bishop, "Pattern recognition and machine learning," *Machine learning*, vol. 128, no. 9, 2006.

[39] O. Nissfolk, "Binary quadratic optimization," Ph.D. dissertation, Abo Akademi University, Finland, 2016.

[40] A. Billionnet, S. Elloumi, and M.-C. Plateau, "Quadratic 0–1 programming: tightening linear or quadratic convex reformulation by use of relaxations," *RAIRO-Operations Research*, vol. 42, no. 2, pp. 103–121, 2008.

[41] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[42] T. Gally, M. E. Pfetsch, and S. Ulbrich, "A framework for solving mixed-integer semidefinite programs," *Optimization Methods and Software*, vol. 33, no. 3, pp. 594–632, 2018.

[43] Jupyter, "How to solve binary quadratic problems with Mosek, more information available at:," https://nbviewer.org/github/MOSEK/Tutorials/blob/master/binary-quadratic/binquad.ipynb.

[44] D. P. Bertsekas, "Multiplier methods: A survey," *Automatica*, vol. 12, no. 2, pp. 133–145, 1976.

[45] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for non-smooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[46] Z. Luo, W. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[47] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[48] R. K. Jain, D.-M. W. Chiu, W. R. Hawe *et al.*, "A quantitative measure of fairness and discrimination," *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*, 1984.

[49] ESA project OPTIMUS, TN2:, "Physical and mac layer performance estimation of the selected techniques," https://wwwfr.uni.lu/snt/research/sigcom/projects/optimus, July 2017.

IX | Research Paper 2 – *A Two-stage Framework for Beam Hopping*

Full Title:
Joint Power Allocation and Beam Scheduling in Beam-Hopping Satellites: A Two-Stage Framework with a Probabilistic Perspective

Authors:
Lin Chen, Linlong Wu, Eva Lagunas, Anyue Wang, Lei Lei, Symeon Chatzinotas, Björn Ottersten

Submitted to: (Major Revision)
IEEE Transactions on Wireless Communications

Abstract:
Beam-hopping (BH) technology, integral to multi-beam satellite systems, adapts beam activation to the variable communication demands of terrestrial users. The optimization of power allocation and beam illumination scheduling constitutes the core design challenge in BH systems, especially under the constraint of a limited number of simultaneously active beams due to restricted radio frequency chain availability. This paper proposes a two-stage BH design solution, which minimizes energy consumption in BH satellite communications while accommodating the heterogeneous demands of users. The first stage addresses the coupling variables of power and beam status by recasting the allocation and scheduling problem through a statistical lens, thus breaking down the intricate relationship between variables. To manage the resulting non-convex challenge, we propose an iterative method that capitalizes on the optimality conditions inherent

to this problem. This method is designed to procure a statistically-informed solution that aligns with our reformulated interpretation. Subsequently, the second stage maps this solution into a concrete beam illumination schedule, employing binary quadratic programming techniques. A penalty-based iterative method is applied, ensuring convergence to a locally optimal solution. Through numerical simulations, the proposed framework has been validated for its efficacy in improving energy efficiency and accurately matching demands.

Keywords:
Energy minimization, power control, beam hopping, binary quadratic programming

Joint Power Allocation and Beam Scheduling in Beam-Hopping Satellites: A Two-Stage Framework with a Probabilistic Perspective

Lin Chen, Linlong Wu, Eva Lagunas, Anyue Wang, Lei Lei,
Symeon Chatzinotas, Björn Ottersten

Abstract

Beam-hopping (BH) technology, integral to multi-beam satellite systems, adapts beam activation to the variable communication demands of terrestrial users. The optimization of power allocation and beam illumination scheduling constitutes the core design challenge in BH systems, especially under the constraint of a limited number of simultaneously active beams due to restricted radio frequency chain availability. This paper proposes a two-stage BH design solution, which minimizes energy consumption in BH satellite communications while accommodating the heterogeneous demands of users. The first stage addresses the coupling variables of power and beam status by recasting the allocation and scheduling problem through a statistical lens, thus breaking down the intricate relationship between variables. To manage the resulting non-convex challenge, we propose an iterative method that capitalizes on the optimality conditions inherent to this problem. This method is designed to procure a statistically-informed solution that aligns with our reformulated interpretation. Subsequently, the second stage maps this solution into a concrete beam illumination schedule, employing binary quadratic programming techniques. A penalty-based iterative method is applied, ensuring convergence to a locally optimal solution. Through numerical simulations, the proposed framework has been validated for its efficacy in improving energy efficiency and accurately matching demands.

Index Terms

Energy minimization, power control, beam hopping, binary quadratic programming

This work was supported by the Luxembourg National Research Fund (FNR) under the project "FlexSAT: Resource Optimization for Next Generation of Flexible SATellite Payloads" (C19/IS/13696663). Lin Chen, Linlong Wu, Eva Lagunas, Anyue Wang, Symeon Chatzinotas, and Björn Ottersten are with the Interdisciplinary Centre for Security, Reliability, and Trust (SnT), University of Luxembourg, Luxembourg. The work of L. Wu was supported by the Luxembourg National Research Fund (FNR) through the CORE INTER project under C20/IS/14799710/SENCOM. Lei Lei is with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China. *(Corresponding author: Linlong Wu, Lin Chen)*

I. Introduction

Satellite communications (SatCom) are widely acknowledged for their capability to offer pervasive connectivity and reliable data transfer, utilizing high-altitude satellites to encompass extensive geographic regions. Historically used for television broadcasting through single-beam Geostationary (GEO) satellites, the technology has advanced significantly since the 2000s. The evolution has led to multi-beam high-throughput satellites that operate in the higher frequency ranges (i.e., Ku and Ka bands) and cater to the broadband market [1]. Nowadays, SatCom's capabilities to connect anyone, anywhere, and anytime are considered a key feature to extend terrestrial 5G cellular connectivity to unreachable areas [2].

Current broadband satellite communication systems typically use a static beam configuration where the coverage is evenly divided into hundreds of spot-beam areas. To avoid interfering among beams, the operator schedules the active beams using an orthogonal frequency reuse strategy. However, this method is frequency-inefficient, which has motivated the development of more adaptable payload designs [3, 4].

To address this challenge, the concept of Beam-Hopping (BH)-enabled satellite communication systems has been introduced [5]. Unlike traditional fixed-beam systems, BH systems only activate a subset of spot beams within a designated period, known as "dwell time." These active beams can be changed dynamically, following a beam illumination plan that responds to varying demands. Beam-hopping offers numerous benefits, the most significant being the temporal flexibility it provides in distributing capacity to different beams as needed. Additionally, activating fewer beams at any given time requires fewer onboard Radio Frequency (RF) chains, which can reduce spacecraft weight and size. Ultimately, this reduction may result in decreased launch costs.

BH satellite communication technology has reached a level of maturity, significantly bolstered by industry support. This support has led to adaptations, such as the standardization of satellite air interfaces, to accommodate BH synchronization needs [6, 7]. An example of this technology in action is the Eutelsat Quantum satellite, which employs forward link BH. While the technical capability for beam illumination and configuration exists, the effective management of these functions remains an active area of research and development. This is mainly because crafting an optimal beam time-activation plan and allocating power effectively is complex. Specifically, it is challenging to properly allocate the limited satellite power across all active beams to meet the dynamic demands while satisfying the satellite payload hardware limitations.

SatCom systems should carefully manage energy consumption. Satellite payloads are equipped with solar panels that convert solar energy into electricity, which powers wireless transmission and other operations [8]. Given that these panels can only collect a limited amount of power, it is crucial for these systems to address the issue of energy efficiency. Additionally, excessive energy consumption can have adverse effects on the payload's mass and lifespan [9]. As a consequence, the minimization of transmit power has become a significant area of research for general non-terrestrial communication platforms, e.g. [10].

In recent years, flexible resource allocation strategies for satellite communication systems have attracted considerable attention. For instance, [11] offers a design and analysis of the beam illumination pattern for BH systems, aiming to optimally match the traffic demands. In [12], [13], and [14] BH is combined with interference mitigation techniques to enable the simultaneous activation of adjacent beams, using the same spectral resources. Notably, [14] proposed to design the illumination pattern by penalizing the activation of adjacent beams, with the aim of reducing the usage of precoding techniques whenever possible. Other works, like [15], and [16], proposed to design the illumination pattern via deep learning techniques. These works focus only on exploring the flexibility in the time domain, which does not fully utilize the capabilities of onboard resources. Power allocation for non-BH satellite systems has also been evaluated. In [17], the authors propose allocating power based on traffic demands and channel conditions. An energy-aware power allocation problem for non-BH systems is formulated in the work [9], which aims to minimize both unmet system capacity and total radiated power. Both [17] and [9] consider only the flexibility of onboard resources in the power domain. Although [18] proposed jointly allocating power and frequency carriers to minimize the weighted objective of energy consumption and frequency occupation, the proposed design does not fully utilize the precious spectrum resources. The load coupling model, originally analyzed by [19], characterizes the coupling relationships among the temporal occupation of beams. The model has been employed, for instance, in [20], the work most closely related to ours, where the power and load of beams are studied for cellular networks to minimize energy consumption. However, [20] does not consider the constraint on the maximal number of simultaneously active beams, an issue that is addressed in our work.

Based on the existing works, we consider the joint design of beam illumination patterns and power allocation to optimize the energy efficiency of the BH satellite system, which has a limit on the maximum number of simultaneously active beams. Intrinsically, the beam pattern and

the corresponding power assignment are coupled, both of which contribute to the achievable capacity and therefore influence the overall energy consumption. To the best of our knowledge, this challenge has not been effectively addressed in the literature to date.

To address the challenge stated above, we propose a two-stage framework to solve the combinatorial nonconvex problem. In the first stage, utilizing the mean-field theory [21, 22], we reinterpret the beam illumination patterns with beam activation probabilities, consequent to which the corresponding power will be the average one for the whole time window. Accordingly, the original problem is reformulated into a new one, where the average powers and activation probabilities of the beams are the optimization variables. An iterative method is also proposed to solve this problem optimally. In the second stage, the activation probabilities of the solution obtained in the first stage are mapped into the discrete beam illumination patterns by solving a binary quadratic programming problem. Our major contributions are summarized as follows:

- We propose a two-stage framework to jointly design power allocation and beam scheduling for BH satellite systems. Within the framework, by utilizing the mean-field theory, a probabilistic reformulation becomes available which paves the way for addressing the intrinsic coupling between beam power and beam illumination pattern. Besides, for the reformulated problem, we analyze the optimality conditions and develop an iterative method to yield a globally optimal solution.
- We also develop a systematic mapping scheme that converts the probabilistic solution obtained in the first stage into a deterministic one satisfying all constraints. To ensure the practicality of the solution, beam hopping latency is also integrated into the scheme.
- Numerical simulation results validate the theoretical findings: i) the system consumes minimal energy when all available beams are active during the time window; ii) increasing the maximal number of active beams could reduce energy consumption. Furthermore, given that our mathematical model is based on the Shannon formula, we acknowledge a potential performance loss due to coding modulation methods in real-world applications. To address this concern, our study includes a method specifically designed to compensate for such performance loss, ensuring the completeness and applicability of our approach in practical scenarios.

The remainder of the paper is organized as follows. In section II, we present the system model and formulate the energy minimization problem from a deterministic perspective. In section III,

TABLE I: Glossary of notations

Notation	Definition
B	Available full bandwidth
K	Maximal number of simultaneously active beams
M	Total number of TSs within the time window
N	Total number of beams
$\mathbf{d}$	Vector of demands in bps
$\hat{\mathbf{d}}$	Vector of demands in [number of TSs]
$\mathbf{H}$	The channel matrix
$p_{n,t}$	Transmit power of beam n at TS t
p_n	Average transmit power of beam n
$\mathbf{p}$	Vector of the average transmit power
$\mathbf{P}$	Penalty matrix
$x_{n,t}$	Beam activation indicator n at TS t
$\mathbf{y}$	Vector of consumed energy of beams
$\boldsymbol{\rho}$	Vector of activation probability of beams
ΔT	Duration of TS
$\sigma(\cdot)$	Spectral radius of the matrix
σ_T^2	Power of thermal noise
$g(\rho_n)$	$\rho_n(2^{\frac{d_n}{B\rho_n}}-1)$
$\mathbf{G}_i$	$\mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}_i)), \quad i=0,1,2$
τ	Latency trade-off factor
ϵ	Demand compensation coefficient
$\mathtt{dim}(\cdot)$	the dimension of the vector
$\lfloor\cdot\rfloor$	The minimal integer of the value

we reformulate the problem from a statistical perspective. Section IV, we focus on solving the statistical problem, and we simply introduce the method to convert the statistical solution to design the deterministic illumination pattern in Section V. In Section VI, we validate the proposed framework with numerical simulations. The conclusion is made in Section VII.

II. System Model and Problem Formulation

In this section, we first introduce the system model for the considered beam-hopping satellite system and then formulate the problem of energy efficiency, which minimizes the total consumed energy while satisfying both the limited onboard resources and on-ground user demands.

A. System Model for Beam-Hopping Satellite

A bent-pipe multi-beam geostationary orbit (GEO) satellite system is considered for the forward-link transmission. The system covers the area of service by N spot beams, each of which reuses the same spectrum of the bandwidth B. The BH time window for the system is composed of M consecutive time slots (TSs) with time-slot duration $\triangle T$. The considered system configurations are compliant with the satellite DVB-S2(X) standard [23] air interface, where the duration of the superframe [24] corresponds to the minimum switching time $\triangle T$ defined above. For such a system, we aim to design the beam illumination pattern, which assigns these N beams over the M time slots within the time window.

Let $\mathbf{H} \in \mathbb{C}^{N\times N}$ be the channel matrix containing all the channel coefficients of the forward link. In particular, the channel coefficient corresponding to the n-th satellite beam to the user k is modeled following the approach in [25], and can be written as,

$$H_{n,k} = \sqrt{G_k^{Rx} G_n^{Tx}(x_k, y_k)} e^{j\phi_{n,k}} / (4\pi \frac{D_k}{\lambda}) \tag{1}$$

where G_k^{Rx} denotes the receive antenna gain of the user k; $G_n^{Tx}(x_k, y_k)$ denotes the transmit antenna gain from the n-th satellite beam to the user k, which is located at the longitude x_k and latitude y_k; $\phi_{n,k}$ is the phase component associated with the n-th satellite beam and the user k; λ denotes the wavelength of the carrier frequency band; D_k denotes the distance between the satellite and the user k.

Note that our main focus is on beam-demand satisfaction, concentrating on beam scheduling rather than ground user scheduling, which is beyond the scope of this work. Accordingly, a super-user terminal is modeled, which essentially aggregates all user's demands and is served by one of the N beams. This abstraction is compliant with the unicast transmission mode typically employed in the DVB-S2(X) air interface, where the multiplexing of users is done via Time-Division Multiple Access (TDMA). Thus given this one-to-one mapping between the beam and super-user, the index n interchangeably refers to both the satellite beam and super-user terminal.

We denote the traffic demand for beam n as d_n with unit bps, and the vector of demands of all beams as $\mathbf{d} = [d_1, \ldots, d_N]^T$. We denote the transmit power for beam n at TS t by $p_{n,t}, \forall n = 1, \ldots, N, t = 1 \ldots, M$. The activation of beam n at TS t is indicated by $x_{n,t} \in \{0, 1\}$ with $x_{n,t} = 1$ if the beam is activated and $x_{n,t} = 0$ otherwise.

The achievable rate of the terminal in beam n at TS t is

$$R_n[t] = B \log_2(1 + \frac{p_{n,t}|H_{n,n}|^2}{\sum_{i\neq n} p_{i,t}|H_{i,n}|^2 + \sigma_T^2}), \tag{2}$$

where $\sigma_T^2 = \tau T_{Rx} B$ represents the thermal noise power with τ being the Boltzmann constant, and T_{Rx} being the clear sky noise temperature of the receiver [25].

Therefore, the demand constraint is

$$\frac{1}{M}\sum_{t=1}^{M} R_n[t] \geq d_n,\ \forall n, \tag{3}$$

which ensures that the average achievable rate within the M TSs is no less than the requested traffic demand of the n-th beam.

B. Problem Formulation of Energy Minimization

We formulate a resource allocation problem to jointly optimize power allocation and beam illumination to minimize the total consumed energy while satisfying uneven beam demands. Therefore, the problem, denoted as $\mathcal{P}_0$, is formulated as

$$(\mathcal{P}_0): \underset{p_{n,t}}{\texttt{minimize}} \quad \sum_{t=1}^{M}\sum_{n=1}^{N} p_{n,t} \tag{4a}$$

$$\texttt{subject to} \quad (C_1): \frac{1}{M}\sum_{t=1}^{M} R_n[t] \geq d_n,\ \forall n, \tag{4b}$$

$$(C_2): \sum_{n=1}^{N} x_{n,t} \leq K,\ \forall t, \tag{4c}$$

$$(C_3): p_{n,t} \geq 0,\ \forall n, t, \tag{4d}$$

$$(C_4): x_{n,t} = \texttt{sign}(p_{n,t}),\ \forall n, t, \tag{4e}$$

where $R_n[t]$ is defined by Eq. (2), and $\texttt{sign}(\cdot)$ is the sign function. The problem is subject to four constraints: C_1, the demand constraint, C_2, the maximum number of simultaneously active beams, C_3, the non-negativity of $p_{n,t}$, and C_4, the binarization of $x_{n,t}$. Among these constraints, C_2 is incroporated. Due to the mass limit of the satellite payload, the maximum number of simultaneously active beams should not exceed the number of digital RF chains, denoted as $K (K \leq N)$. The constraint is a key feature of BH-based SatCom systems, as it enables a significant reduction in satellite mass. $\mathcal{P}_0$ is a non-convex mixed-integer programming problem, which poses significant challenges in finding a solution.

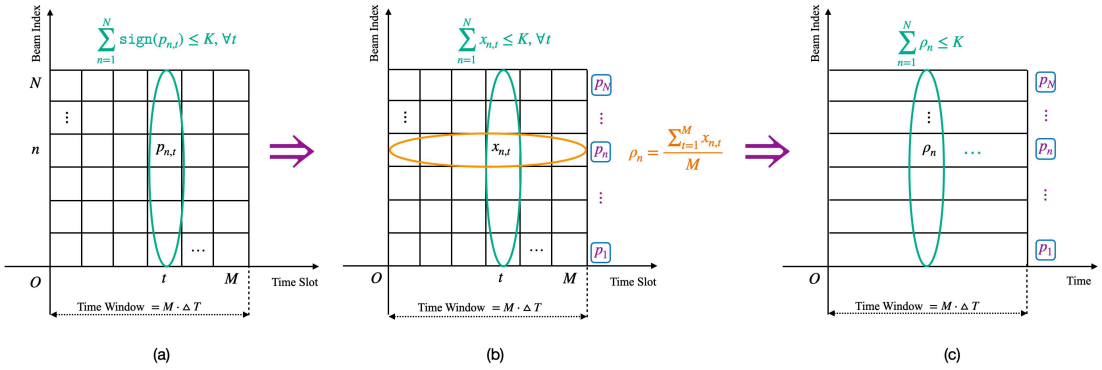


Fig. 1: The formulation and reformulation of the problem. (a) the original problem $\mathcal{P}_0$; (b) the intermediate status of the problem conversion; (c) the reformulated problem $\mathcal{P}_1$.

III. Problem Reformulation via Mean Field Theory

For a conventional approach to mixed integer programming problems, it first relaxes the binary variable $x_{n,t}$ to be continuous and linearizes the demand constraint with respect to $p_{n,t}$. Subsequently, the variables are alternatively updated with iterations.

However, the variables of $\mathcal{P}_0$ are coupled, making this typical approach less suitable. Furthermore, popular learning methods such as deep learning often require a large dataset for model training. Given the aforementioned challenges, properly addressing the strong coupling relationship between $p_{n,t}$ and $x_{n,t}$ is crucial for efficiently solving the problem $\mathcal{P}_0$. To achieve this, we employ the Mean Field Theory (MFT) [21, 22] to decouple the power and beam activation indicator.

In the mean-field method, the mutual influence between random variables is replaced by an effective field, which acts independently on each random variable [21]. Considering that the power variable $p_{n,t}$ serves dual functions — represents both the power value and potentially indicating the status of the beam at a given instance of TS— these variables are mutually influenced in the time domain via the demand constraint C_1. To simplify the problem, we apply the Mean Field Theory method, which replaces the power vector variable with its average power across all beams. This is followed by replacing the indicator vector variable with its activation probability. We summarize the reformulation progress graphically in Fig. 1.

Specifically, suppose that we have obtained the optimal solution to $\mathcal{P}_0$, denoted by $(p_{n,t}, x_{n,t}), \forall n, t$. We define p_n and ρ_n as the average power and activation probability, respectively, over the

entire time window for beam n across all beams. The variable ρ_n represents the probability that beam n is activated during the time window of the optimal solution, and is defined as $\rho_n = \frac{\sum_{t=1}^{M} x_{n,t}}{M} \in (0, 1]$.

Accordingly, the total energy consumption can be reformulated as the total expected consumed energy, $\boldsymbol{\rho}^T \mathbf{p} \cdot \triangle T$, where $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_N]^T$ and $\mathbf{p} = [p_1, \ldots, p_N]^T$. The number of simultaneously active beams can be expressed as the expected number of active beams, $\boldsymbol{\rho}^T \mathbf{1}$. Regarding the demand constraint C_1, it can be reformulated to ensure the expected capacity meets its demand as follows.

$$\rho_n B \log_2(1 + \frac{p_n |H_{n,n}|^2}{\sum_{i \neq n} \rho_i p_i |H_{i,n}|^2 + \sigma_T^2}) \geq d_n, \ \forall n. \tag{5}$$

which can be further rewritten as

$$\rho_n \geq f_n(\boldsymbol{\rho}, \mathbf{p}) \triangleq \frac{d_n}{B \log_2(1 + \frac{p_n |H_{n,n}|^2}{\sum_{i \neq n} \rho_i p_i |H_{i,n}|^2 + \sigma_T^2})}, \ \forall n. \tag{6}$$

Therefore, $\mathcal{P}_0$ can be reformulated as

$$(\mathcal{P}_1) : \underset{\boldsymbol{\rho}, \mathbf{p}}{\texttt{minimize}} \quad \boldsymbol{\rho}^T \mathbf{p} \tag{7a}$$

$$\texttt{subject to} \quad (\hat{C}_1) : \boldsymbol{\rho} \succeq \mathbf{f}(\boldsymbol{\rho}, \mathbf{p}), \tag{7b}$$

$$(\hat{C}_2) : \boldsymbol{\rho}^T \mathbf{1} \leq K, \tag{7c}$$

$$(\hat{C}_3) : \mathbf{p} \succ \mathbf{0}, \tag{7d}$$

$$(\hat{C}_4) : \mathbf{0} \prec \boldsymbol{\rho} \preceq \mathbf{1}, \tag{7e}$$

where $\mathbf{f}(\boldsymbol{\rho}, \mathbf{p}) = [f_1(\boldsymbol{\rho}, \mathbf{p}), \ldots, f_N(\boldsymbol{\rho}, \mathbf{p})]^T$. The curled inequality symbol $\succeq$ (and its strict form $\succ$) is used to denote a generalized component-wise inequality between vectors. The vectors $\mathbf{0}$ and $\mathbf{1}$ are the ones with all elements being 0 and 1, respectively.

Note that the solution of $\mathcal{P}_1$ is in terms of activation probabilities of the N beams, which needs to be converted into discrete illumination pattern. This conversion method will be developed in Section V. In the next section, we will focus on solving $\mathcal{P}_1$.

IV. First-stage: Probabilistic Beam Hopping Solution

In this section, our focus is on solving the non-convex problem $\mathcal{P}_1$. The conventional approach might involve linearizing constraint $\hat{C}_1$ first and then updating the two variables through iterations [18]. However, this approach often leads to performance loss due to the approximation error

between the original and linearized versions and is typically time-consuming. In this work, we propose a novel method that optimally solves the problem. Briefly, we first prove that the fixed point satisfying $\boldsymbol{\rho} = \mathbf{f}(\boldsymbol{\rho}, \mathbf{p})$ is a necessary condition for the optimum of $\mathcal{P}_1$. Based on this necessary condition, the optimal power can be expressed as a function of the optimal activation probability. Consequently, the objective of the problem boils down to an inverse matrix optimization problem, which we have further demonstrated is equivalent to a convex problem.

A. Necessary Condition for the Feasible Problem $\mathcal{P}_1$

According to [26], a fixed point of a function is a point that the function maps to itself, i.e. $x = 1$ is a fixed point of the function $f(x) = x^2$ because $f(1) = 1$. In nonlinear programming, the characteristics of the fixed point can be utilized to devise iterative methods for finding solutions [27]. The following theorem indicates that the solution to $\mathcal{P}_1$ is a fixed point.

***Theorem** 1:* Given $\mathbf{d}, \mathbf{H}$, and K, the solution to the problem $\mathcal{P}_1$, if it exists, is the fixed point of the equations: $\boldsymbol{\rho} = \mathbf{f}(\boldsymbol{\rho}, \mathbf{p})$.

Proof: The proof is in Appendix A. ■

Theorem 1 illustrates the coupling relationship between power and activation probability. Intuitively, a change in demand for one beam can strongly affect the solutions of adjacent beams and, consequently, the entire system. This observation validates the theorem and leads to the following corollary.

Corollary 1: Denoting the relationship between the optimal activation probability $\boldsymbol{\rho}$ and the optimal power $\mathbf{p}$ as $\mathbf{p} = \hat{\mathbf{f}}(\boldsymbol{\rho})$, the function $\hat{\mathbf{f}}(\cdot)$ is a one-to-one mapping.

Proof: Theorem 1 states that the fixed point is the necessary condition of the solution to the $\mathcal{P}_1$. Given the fixed point equation $\rho_n = f_n(\boldsymbol{\rho}, \mathbf{p}), \forall n$, we can derive the followings.

$$d_n = \rho_n B \log_2(1 + \frac{p_n |H_{n,n}|^2}{\sum_{i \neq n} \rho_i p_i |H_{i,n}|^2 + \sigma_T^2}) \tag{8a}$$

$$\Rightarrow \quad p_n = (2^{\frac{d_n}{B\rho_n}} - 1)(\sum_{i \neq n} \rho_i p_i \frac{|H_{i,n}|^2}{|H_{n,n}|^2} + \frac{\sigma_T^2}{|H_{n,n}|^2}) \tag{8b}$$

$$\Rightarrow \rho_n p_n = \rho_n (2^{\frac{d_n}{B\rho_n}} - 1)(\sum_{i \neq n} \rho_i p_i \frac{|H_{i,n}|^2}{|H_{n,n}|^2} + \frac{\sigma_T^2}{|H_{n,n}|^2}) \tag{8c}$$

$$\Rightarrow \quad y_n = g(\rho_n)(\sum_{i \neq n} y_i A_{i,n} + b_n) \tag{8d}$$

where we denote matrix $\mathbf{A}$ with elements $A_{n,n} = 0$, and $\forall i \neq n, A_{i,n} = \frac{|H_{i,n}|^2}{|H_{n,n}|^2}$; $b_n = \frac{\sigma_T^2}{|H_{n,n}|^2}$; $g(\rho_n) = \rho_n(2^{\frac{d_n}{B\rho_n}} - 1)$; and $y_n = \rho_n p_n$, representing the consumed energy of the n-th beam.

According to (8d), we denote the vector of all beams' consumed energies by $\mathbf{y} = [y_1, \ldots, y_N]^T$ and establish its relationship with the vector $\boldsymbol{\rho}$ compactly as follows:

$$\mathbf{y} = (\mathbf{I} - \mathbf{GA})^{-1}\mathbf{Gb} \tag{9a}$$

$$\Rightarrow \mathbf{p} = \mathtt{Diag}(\boldsymbol{\rho}^{-1})(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{Gb} \tag{9b}$$

where $\mathbf{G} = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}))$, and $\mathbf{g}(\boldsymbol{\rho}) = [g(\rho_1), \ldots, g(\rho_N)]^T$. $\mathtt{Diag}(\mathbf{x})$ denotes the diagonal matrix with diagonal elements given by vector $\mathbf{x}$, and $\mathbf{I}$ is the identity matrix. Finally, $\mathbf{b} = [b_1, \ldots, b_N]^T$.

Eq. (9b) indicates that the relationship between the optimal power and the activation probability is a one-to-one mapping and thereby completes the proof. ■

Based on Eq. (9a), $\mathcal{P}_1$ is equivalent to

$$(\mathcal{P}_2): \underset{\boldsymbol{\rho}}{\mathtt{minimize}} \quad \mathbf{1}^T(\mathbf{I} - \mathbf{G}(\boldsymbol{\rho})\mathbf{A})^{-1}\mathbf{G}(\boldsymbol{\rho})\mathbf{b} \tag{10a}$$

$$\mathtt{subject\ to} \quad (\hat{C}_2): \boldsymbol{\rho}^T\mathbf{1} \leq K \tag{10b}$$

$$(\tilde{C}_3): \sigma(\mathbf{G}(\boldsymbol{\rho})\mathbf{A}) < 1 \tag{10c}$$

$$(\hat{C}_4): \mathbf{0} \prec \boldsymbol{\rho} \preceq \mathbf{1} \tag{10d}$$

where $\mathbf{G}(\boldsymbol{\rho}) = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}))$ is a function of $\boldsymbol{\rho}$ and will be denoted as $\mathbf{G}$ for simplicity. $\sigma(\mathbf{X})$ denotes the spectral radius of the matrix $\mathbf{X}$ – the largest absolute eigenvalue of $\mathbf{X}$.

The constraint $\tilde{C}_3$ in $\mathcal{P}_2$ corresponds to the constraint $\hat{C}_3$ in $\mathcal{P}_1$. Note that $\forall \boldsymbol{\rho} \succ \mathbf{0}$, there holds $\mathbf{G} \succ \mathbf{0}$. Additionally, $\mathbf{b} \succ \mathbf{0}, \mathbf{A} \succ \mathbf{0}$ for all times. According to Eq. (9b), given $\boldsymbol{\rho} \succ \mathbf{0}$, the positivity of $\mathbf{p}$ is equivalent to the positivity of the matrix $(\mathbf{I} - \mathbf{GA})^{-1}$ which is determined by the spectral radius of the matrix $\mathbf{GA}$. The following Theorem 2 provides a detailed explanation.

***Theorem** 2:* Given $\mathbf{d}, \mathbf{H}$ and K, a feasible positive solution of the Eq. (9b) exists if and only if there exists a $\boldsymbol{\rho}$ such that the spectral radius of the matrix $\mathbf{GA}$ is less than 1, i.e. $\sigma(\mathbf{GA}) < 1$.

Proof: The proof is given by the [28, Theorem 1]. ■

Theorem 2 establishes the necessary condition for the feasibility of the solution to $\mathcal{P}_2$.

B. Convexity Analysis of Problem $\mathcal{P}_2$

$\mathcal{P}_2$ involves matrix inverse and spectral radius, making it challenging to solve in general. However, by analyzing and exploiting the special structures of the constraints and the objective

function, we demonstrate that solving $\mathcal{P}_2$ is equivalent to solving a convex problem. This insight paves the way for obtaining the global optimal solution to $\mathcal{P}_1$.

Lemma 1: Given B, d, the function $g(z) = z(2^{\frac{d}{Bz}} - 1)$ is monotonically decreasing and convex in the domain $z \in (0, +\infty)$

Proof: The proof is in Appendix B. ■

Lemma 2: Given $\mathbf{H}, \mathbf{d}$, the set $\mathcal{Y} = \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) \leq 1, \mathbf{0} \preceq \boldsymbol{\rho} \preceq \mathbf{1}\}$ is convex.

Proof: The proof is in Appendix C. ■

In the proof presented in Appendix C, we define the condition to determine whether the given parameters $\mathbf{H}, \mathbf{d}$ are reasonable. When given reasonable parameters, the set $\mathcal{Y}$ is non-empty.

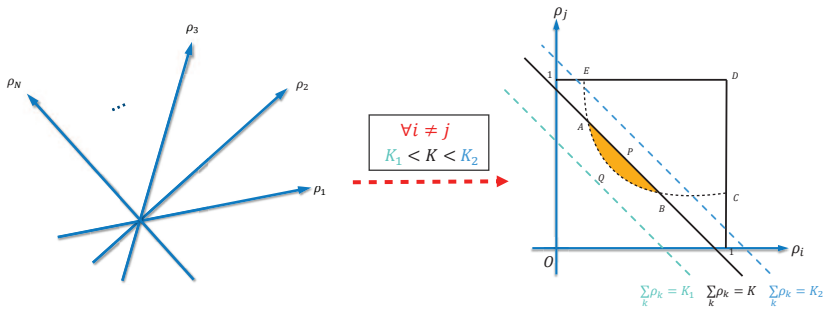


Fig. 2: The feasible zone for $\mathcal{P}_3$ is indicated by the yellow shading. By increasing the parameter K from K_1 to K_2, the previously infeasible problem becomes feasible.

***Theorem** 3:* The solution to $\mathcal{P}_2$ is equivalent to the solution to a convex problem, which is given by

$$\mathcal{P}_3 : \underset{\boldsymbol{\rho} \in \hat{\mathcal{Y}}}{\texttt{minimize}} \quad L(\boldsymbol{\rho}) \tag{11}$$

where $\hat{\mathcal{Y}} = \mathcal{Y} \cap \{\boldsymbol{\rho} | \boldsymbol{\rho}^T \mathbf{1} \leq K\}$, and

$$L(\boldsymbol{\rho}) = \begin{cases} \mathbf{1}^T (\mathbf{I} - \mathbf{GA})^{-1} \mathbf{Gb} & \boldsymbol{\rho} \in \hat{\mathcal{Y}} \setminus \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) = 1\} \\ +\infty & \boldsymbol{\rho} \in \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) = 1\} \end{cases} \tag{12}$$

Proof: The proof is in Appendix D. ■

Previously, we explored the problem theoretically. Next, we will graphically illustrate the potential characteristics of the optimum, providing the basis for mapping the statistical solution to the deterministic one in Section V.

Lemma 3: Given reasonable $\mathbf{d}, \mathbf{H}$, i.e. $\mathcal{Y} \neq \emptyset$, the feasibility of $\mathcal{P}_3$ depends on the parameter K, which represents the maximum number of active beams. If K is small, the problem may be infeasible. However, increasing K may render an initially infeasible problem feasible.

Proof: The feasible set of $\mathcal{P}_3$ is the intersection of two convex sets, i.e., $\hat{\mathcal{Y}} = \mathcal{Y} \cap \{\boldsymbol{\rho} | \boldsymbol{\rho}^T \mathbf{1} \leq K\}$, thus it is convex.

According to *Lemma* 2, given reasonable $\mathbf{H}, \mathbf{d}$, the set $\mathcal{Y}$ is convex and non-empty. Therefore, whether the set $\hat{\mathcal{Y}}$ is empty is determined by the parameter K. Additionally, for $\forall \boldsymbol{\rho}_1 \succeq \boldsymbol{\rho}_2$, it holds that $\mathbf{G}_1\mathbf{A} \leq \mathbf{G}_2\mathbf{A}$ (The symbol $\geq$ is defined in Appendix C). Consequently, based on [29, Corollary 1.5 p-27], we have $\sigma(\mathbf{G}_1\mathbf{A}) \leq \sigma(\mathbf{G}_2\mathbf{A})$.

Based on these two characteristics, we illustrate how K affects the set $\hat{\mathcal{Y}}$ using Fig. 2. As seen in the figure, the left side displays the N dimensions of the variable $\boldsymbol{\rho}$ and the right side illustrates the slice of set $\hat{\mathcal{Y}}$ cut by the plane determined by any two dimensions of $\boldsymbol{\rho}$, i.e. $\rho_i O \rho_j, \forall i \neq j$. Given reasonable $\mathbf{d}, \mathbf{H}$, the zone $EAQBCD$ is non-empty, which denotes the slice of the set $\mathcal{Y}$. The dashed curve $EAQBC$ denotes the boundary on which $\sigma(\mathbf{GA}) = 1$. The line AB denotes the boundary on which $\boldsymbol{\rho}^T\mathbf{1} = K$. The yellow shadow with its boundaries, i.e. $AQBP$, is the slice of the set $\hat{\mathcal{Y}}$, which denotes the slice of the feasible zone cut.

When K is small, i.e. $K = K_1$, the zone $AQBP$ is $\emptyset$, denoting the infeasibility of $\mathcal{P}_3$. However, by increasing K to K_2, the line AB will move parallel to the direction of vector $\vec{OD}$, which would result in the feasibility of the problem. Thereby, we complete the proof. ■

Lemma 4: The total consumed energy $L(\boldsymbol{\rho})$ is monotonically decreasing with respect to the variable $\boldsymbol{\rho} \in \hat{\mathcal{Y}}$.

Proof: The proof is in Appendix E. ■

***Theorem** 4:* The system achieves minimal energy consumption when all available beams are active within the time window; furthermore, increasing the maximum number of simultaneously active beams further reduces the energy consumed.

Proof: The proof is in Appendix F. ■

Theorem 4 demonstrates that the solution to $\mathcal{P}_3$, if it exists, will be located on the boundary where $\boldsymbol{\rho}^T\mathbf{1} = K$. This theorem also suggests that increasing the maximum number of active beams reduces energy consumption. It is important to note that there is an inherent constraint on K, which must not exceed the total number of beams due to the satellite hardware constraints. Choosing an appropriate value for K is a complex task that involves considerations such as size,

weight, power (SWaP), and the distribution of traffic demands. However, these considerations are beyond the scope of this paper.

C. Iterative Optimization of Problem $\mathcal{P}_3$

For solving $\mathcal{P}_3$, we will start by rewriting the spectral radius constraint using its equivalent expression. Then, we will develop a method based on Successive Convex Approximation (SCA) [30] to tackle the resulting problem.

In general, the SCA method addresses a complicated problem by iteratively approximating the function value. When the original problem is convex, SCA can find the global optimum [30]. The objective function can be expressed as

$$L(\boldsymbol{\rho}) = \mathrm{Tr}\{(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{GB}\} \tag{13}$$

where $\mathbf{B} = \mathbf{b} \cdot \mathbf{1}^T$, and $\mathrm{Tr}\cdot$ represents the trace of the matrix. Given $\boldsymbol{\rho}_0$ and subsequently defining $\mathbf{G}_0 = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}_0))$, we approximate inverse matrix $(\mathbf{I} - \mathbf{GA})^{-1}$ using its first-order Taylor expansion with respect to the variable $\mathbf{G}$ [31]. This allows us to approximate the total consumed energy as follows:

$$L(\boldsymbol{\rho}|\boldsymbol{\rho}_0) \approx \mathrm{Tr}\{\underbrace{(\mathbf{I} - \mathbf{G}_0\mathbf{A})^{-1}}_{\mathbf{T}}\mathbf{GB} + (\mathbf{I} - \mathbf{G}_0\mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \mathbf{G}_0\mathbf{A})^{-1}\mathbf{GB}(\mathbf{G} - \mathbf{G}_0)\} \tag{14a}$$

$$= \mathbf{g}^T(\boldsymbol{\rho})\underbrace{\mathtt{Diag}(\mathbf{1}^T\mathbf{TATB})}_{\boldsymbol{\Omega}}\mathbf{g}(\boldsymbol{\rho}) + \underbrace{[\mathbf{1}^T\mathbf{TB} - \mathbf{1}^T\mathbf{TATG}_0\mathbf{B}]}_{\hat{\mathbf{b}}^T}\mathbf{g}(\boldsymbol{\rho}) = \mathbf{g}^T(\boldsymbol{\rho})\boldsymbol{\Omega}\mathbf{g}(\boldsymbol{\rho}) + \mathbf{g}^T(\boldsymbol{\rho})\hat{\mathbf{b}} \tag{14b}$$

Since $\boldsymbol{\Omega}$ is a diagonal matrix of positive elements, the function $y(\mathbf{z}) = \mathbf{z}^T\boldsymbol{\Omega}\mathbf{z}$ is convex and non-decreasing when $\mathbf{z} \succ \mathbf{0}$. As the function $g(\rho)$ is convex, and following the composition rules outlined in [32], the composite function $L(\boldsymbol{\rho}|\boldsymbol{\rho}_0)$ with respect to the variable $\boldsymbol{\rho}$ is also convex. Furthermore, according to [33], $\sigma(\mathbf{X}) \leq t$ is equivalent to $\sigma(\mathbf{X}) \leq t,\ \sigma(-\mathbf{X}) \leq t$, which is further equivalent to

$$\begin{pmatrix} t\mathbf{I} & \mathbf{X} \\ \mathbf{X}^T & t\mathbf{I} \end{pmatrix} \succeq 0. \tag{15}$$

By substituting the equivalent expression into the constraint $\sigma(\mathbf{GA}) \leq 1$ in $\mathcal{P}_3$, we obtain the subproblem to be solved at each SCA iteration as follows:

$$(\mathcal{P}_4): \underset{\boldsymbol{\rho}}{\texttt{minimize}} \quad L(\boldsymbol{\rho}|\boldsymbol{\rho}_0) \tag{16a}$$

$$\texttt{subject to} \quad (\hat{C}_2): \boldsymbol{\rho}^T\mathbf{1} \leq K, \tag{16b}$$

$$(\tilde{C}_3): \begin{pmatrix} \mathbf{I} & \texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}))\mathbf{A} \\ \mathbf{A}^T\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho})) & \mathbf{I} \end{pmatrix} \succeq 0, \tag{16c}$$

$$(C_4): \mathbf{0} \preceq \boldsymbol{\rho} \preceq \mathbf{1}, \tag{16d}$$

which is convex and can be efficiently solved using CVX [34]. An overview of the proposed approach for solving $\mathcal{P}_3$ is provided in **Algorithm** 1.

D. Computational Complexity

The overall computational complexity of Algorithm 1 scales linearly with the number of outer loop iterations. Given $\boldsymbol{\rho}_0$, the non-symmetric exponential cone optimization of $\mathcal{P}_4$ is solved using CVX [35] with an advanced solver, which employs a primal-dual interior point method with a worst-case computational complexity of $\mathcal{O}((MN)^3)$ [36, 37].

Algorithm 1 INVERSE MATRIX OPTIMIZATION

Initialization $\boldsymbol{\rho}_0 = \mathbf{1}, K, \mathbf{A}, \mathbf{d}, \mathbf{H}, k = 0$
repeat
  if $\sigma(\mathbf{G}_k\mathbf{A}) < 1$ **then**
    Set $k = k + 1$, Update $\boldsymbol{\Omega}, \hat{\mathbf{b}}$
    Solve $\boldsymbol{\rho}_{k+1} = \arg \underset{\boldsymbol{\rho}\in\hat{\mathcal{Y}}}{\texttt{minimize}}\ L(\boldsymbol{\rho}|\boldsymbol{\rho}_k)$
  else
    Return: Infeasible
  end if
until $\frac{\|\boldsymbol{\rho}_{k+1}-\boldsymbol{\rho}_k\|_2}{\texttt{dim}(\boldsymbol{\rho}_k)} \leq 10^{-3}$

V. SECOND-STAGE: STATISTIC-TO-DETERMINISTIC SOLUTION MAPPING

In this section, we first introduce the method to convert the acquired activation probability into discrete demands. Subsequently, we design the illumination pattern, where beam hopping latency is also considered.

A. Conversion to Discrete Demands

Recall that in the previous section, we obtained the optimal beam activation probability and power. To implement this resource allocation solution in a satellite system, we still need to convert them into deterministic illumination operations. This involves selecting time slots for beam activation and assigning the corresponding power.

We present a rounding scheme that converts the continuous probability into discrete demand in the number of TS. According to Theorem 4, the optimum is located on the boundary where $\boldsymbol{\rho}^T\mathbf{1} = K$. Therefore, the rounding algorithm aims to find an integer point, i.e. the number of selected active TS for all beams, that is not only close to $M\boldsymbol{\rho}$ but also sums up to MK. The rounding scheme is detailed in the **Algorithm** 2.

Algorithm 2 ROUNDING

1: **Input**: $\boldsymbol{\rho}^\star$, M, K
2: The lower bound of the demands $\hat{\mathbf{d}} = \lfloor \boldsymbol{\rho}^\star \cdot M \rfloor$
3: The residual $\mathbf{o} = \boldsymbol{\rho}^\star \cdot M - \hat{\mathbf{d}}$
4: Sort the residual $\mathbf{o}$ in descending order and find the index of its first $KM - \mathbf{1}^T\hat{\mathbf{d}}$ elements.
5: Modify the demands $\hat{\mathbf{d}}$ by adding 1 to the selected indexes of beams.

Recall that the relationship between the optimal power and the corresponding activation probability is a one-to-one mapping. Consequently, by applying the rounding algorithm, the power also needs to be modified given that the beam activation probability changes. Specifically, the modification is expressed as

$$\hat{\boldsymbol{\rho}} = \frac{\hat{\mathbf{d}}}{M}, \quad \hat{\mathbf{p}} = \mathtt{Diag}(\hat{\boldsymbol{\rho}}^{-1})(\mathbf{I} - \hat{\mathbf{G}}\mathbf{A})^{-1}\hat{\mathbf{G}}\mathbf{b} \tag{17}$$

where $\hat{\mathbf{d}}$ is the vector of converted discrete demands and $\hat{\mathbf{G}} = \mathtt{Diag}(\mathbf{g}(\hat{\boldsymbol{\rho}}))$.

B. Mapping to Deterministic Illumination Pattern

With the required power and discrete demand for all beams, the next step is to design the illumination pattern. This involves choosing specific TS from the entire time window to meet the required number of TS for each beam and assigning the necessary power to them. In a prior study [38], a heuristic method was proposed to randomly assign the demands while satisfying the constraints.

However, as discussed in Section III, the problem reformulation is based on the assumption that we have obtained the optimal solution. In other words, the activation probability denotes the probability of a beam being activated during the BH window of the optimal solution, which minimizes consumed energy while meeting the constraints. We follow the optimal assumption to design the deterministic solution $x_{n,t}, \forall n, t$. Given $(\hat{\mathbf{p}}, \hat{\boldsymbol{\rho}})$, the consumed energy is determinate. In this context, we propose focusing on demand-matching performance while satisfying the constraint of a maximum number of active beams. Specifically, our proposed method penalizes the activation of any two beams simultaneously, aiming to minimize the total penalty while adhering to the constraints.

The penalty $\hat{\mathbf{P}} \in \mathbb{R}^{N \times N}$ defines the relative accumulated interference from one beam to the other. Specifically, the interference from the i-th beam to the j-th beam is defined as

$$\hat{P}_{i,j} = \frac{\iint_{\mathbf{S_j}} |H_{i,k}|^2 \, p_i dx_k \, dy_k}{\iint_{\mathbf{S_j}} |H_{j,k}|^2 \, p_j dx_k \, dy_k} \tag{18}$$

where (x_k, y_k) represents the longitude and latitude of user k which is in the beam j; and $\mathbf{S}_j$ stands for the coverage of beam j which is defined by the beam contour at $-3dB$ from the maximum gain.

Moreover, the operator must also take into account the minimization of the beam-hopping latency when designing the illumination pattern in practice. In the context of the considered system, beam-hopping latency of the beam refers to the number of times the beam switches from on to off and from off to on in a given time window. Given the required number of active time slots for the beam within a specific time window, latency increases more when the frequency of the switch turning on and off is higher. This increase in latency is due to the multiple operations required across various layers, including but not limited to the physical and MAC layers, which result in additional time consumption.

Based on the above illustration, the goal to avoid the latency increase is equivalent to reducing the total number of switching times of the designed beam pattern, which can be expressed as

$$\sum_{n=1}^{N} \sum_{t=1}^{M-1} \parallel x_{n,t} - x_{n,t+1} \parallel^2 \tag{19a}$$

$$= \sum_{n=1}^{N} \{2\hat{d}_n - x_{n,1}^2 - x_{n,M}^2 - \sum_{t=1}^{M-1} 2x_{n,t}x_{n,t+1}\} \tag{19b}$$

$$= - \mathbf{x}^T \tilde{\mathbf{P}} \mathbf{x} + \text{Const} \tag{19c}$$

where $\mathbf{x}_t = [x_{1,t}, \ldots, x_{N,t}]^T$ is the binary vector indicating the status of all beams at TS t, and $\mathbf{x} = [\mathbf{x}_1^T, \ldots, \mathbf{x}_M^T]^T$. The matrix $\tilde{\mathbf{P}}$ is characterized as a block tridiagonal matrix. When $M = 2$ and $M \geq 3$, its expressions can be given respectively by

$$\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{I}_N & \mathbf{I}_N \\ \mathbf{I}_N & \mathbf{I}_N \end{bmatrix}_{2N \times 2N} \tag{20a}$$

$$\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{I}_N & \mathbf{I}_N & \mathbf{0}_N & \ldots & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{I}_N & \mathbf{0}_N & \mathbf{I}_N & \ldots & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{I}_N & \mathbf{0}_N & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{I}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \ldots & \mathbf{I}_N & \mathbf{0}_N & \mathbf{I}_N \\ \mathbf{0}_N & \mathbf{0}_N & \ldots & \mathbf{0}_N & \mathbf{I}_N & \mathbf{I}_N \end{bmatrix}_{MN \times MN} \tag{20b}$$

where $\mathbf{I}_N$ signifies the identity matrix and $\mathbf{0}_N$ represents the zero matrix, both of which are of dimensions $N \times N$."

Therefore, the mapping problem is defined as follows:

$$(\mathcal{P}_m) : \underset{\mathbf{x}}{\texttt{minimize}} \quad \mathbf{x}^T(\mathbf{P} - \iota\tilde{\mathbf{P}})\mathbf{x} \tag{21a}$$

$$\texttt{subject to} \quad (C_5) : \mathbf{E}_t\mathbf{x} \preceq K \cdot \mathbf{1}_M \tag{21b}$$

$$(C_6) : \mathbf{E}_b\mathbf{x} = \hat{\mathbf{d}} \tag{21c}$$

$$(C_7) : \mathbf{x} \in \{0,\ 1\}^{MN} \tag{21d}$$

where $\mathbf{P} = \texttt{Diag}(\hat{\mathbf{P}}, \ldots, \hat{\mathbf{P}}) \in \mathbb{R}^{MN \times MN}$ is the block diagonal penalty matrix and ι is the trade-off factor between the interference penalty and the operation cost.

The matrix $\mathbf{E}_t = \mathbf{I}_M \otimes \mathbf{1}_N^T$ serves as the selective matrix, with each row selecting all the active beams at TS t, where $\otimes$ denotes the Kronecker product; and $\mathbf{1}_N$ denotes a column vector of dimension N with all elements equal to 1. Additionally, $\mathbf{E}_b = \mathbf{1}_M^T \otimes \mathbf{I}_N$ is another selective matrix, with each row selecting all the active TS for the beam. Lastly, $\hat{\mathbf{d}}$ is the converted discrete demands in terms of the number of TS.

$\mathcal{P}_m$ is a binary quadratic programming (BQP) task, efficiently addressed using the multiplier penalty and majorization-minimization (MPMM) method proposed in our previous work [14].

In summary, Fig. 3 illustrates the procedures of the proposed framework. Given the request demands $\mathbf{d}$, the framework provides the designed illumination pattern $\mathbf{x}$ and the corresponding

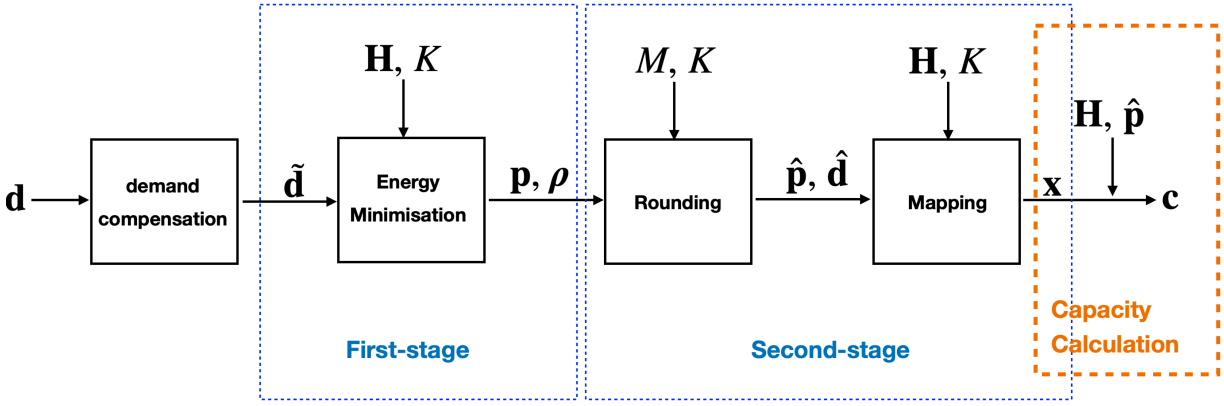


Fig. 3: The proposed framework.

power of beams $\hat{\mathbf{p}}$ such that the provided capacity $\mathbf{c}$ matches the demands as closely as possible. The preprocessing demand compensation part will be explained in Section VI-C.

VI. Numerical Results

In this section, we conduct various simulations to evaluate the proposed framework. First, we demonstrate the influence of system parameters on the final performance. Specifically, we show the convergence of the proposed solving method and validate the theoretical findings. Second, the ideal capacity given by the Shannon formula [39] is undermined in practical implementations, we propose a preprocessing scheme to compensate in advance for the difference between the ideal and practical capacities. Third, we analyze the impact of different latency trade-off factors on system-level performance. Lastly, to assess the performance of the proposed framework, we compare it with state-of-the-art alternatives in terms of energy consumption and demand-matching performance, also providing computational time complexity.

A. Simulation Setup

The parameter settings of the GEO satellite are summarized in TABLE II. The traffic demands of all beams are randomly generated and follow a uniform distribution between $400r$ and $1500r$ ($Mbps$), that is, $400r \leq d_n \leq 1500r,\ \forall n$. Herein r stands for the demand density factor, selected from the set $\{0.1, 0.2, 0.3, 0.4, 0.5\}$. For each selected r, 50 demand instances are generated for testing. A single user in the beam center is assumed. To the best of our knowledge, the method described in [14] represents the state-of-the-art and is thus chosen as the baseline. For the

TABLE II: Summary of System Parameters

Satellite orbit	13°E (GEO)
Additional payload loss	2 dB
Number of virtual beams (N)	67
Beam radiation pattern ($G_n^{Tx}(x_k, y_k)\, e^{j\phi_{n,k}}$)	Provided by ESA
Downlink carrier frequency	19.5 GHz
User link bandwidth, (B)	500 MHz
Roll-off factor	20%
Temperature	50 K
Number of TSs (M)	20

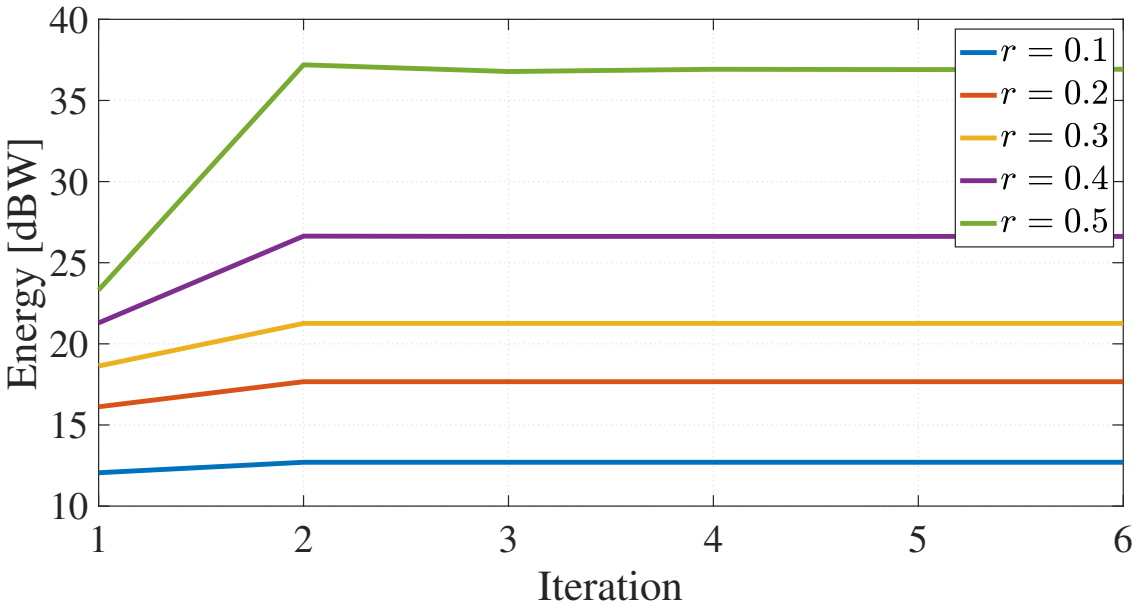


Fig. 4: The convergence of the proposed method. The maximal number of simultaneously active beams $K = 26$, and consumed energy is calculated by $10 \log 10(\boldsymbol{\rho}^T \mathbf{p})$.

baseline, we set the power of each beam as 60 W with 3 dB output back-off (OBO). [1] Unless stated otherwise, all the parameter settings are applied to simulations.

[1]The first step of the baseline is to convert the demands in *bps* into demands in the number of time slots by approximation, within which the expected interference from adjacent beams will be omitted because of the use of precoding. Since there is no precoding in this paper, we would not take this part into account when conducting approximation in the baseline.

B. Performance of the Proposed Method

1) Convergence of Algorithm 1: Fig. 4 shows the convergence of the proposed Inverse Matrix Optimization Algorithm 1 at different demand densities. As observed, the algorithm converges to the optimum within approximately two to three iterations across all demand densities. We set the initial point to $\boldsymbol{\rho}_0 = \mathbf{1}$, representing the largest value of $\boldsymbol{\rho}$, which results in the minimal objective value and the spectral radius of the matrix. However, this initial point falls outside the feasible zone unless $K = N$. Consequently, the consumed energy initially increases due to the objective function's monotonically decreasing nature.

2) Impact of the maximum number of simultaneously active beams: TABLE III and IV illustrate the influence of the maximum number of simultaneously active beams, denoted as K, on the total consumed energy and the corresponding optimal spectral radius of the matrix $\mathbf{GA}$, respectively.

The results yield the following findings and corresponding analysis: As shown in TABLE III, the total consumed energy decreases with an increase in the maximum number of simultaneously active beams across all demand densities. This observation validates Theorem 4. Furthermore, we confirm *Lemma* 3. Specifically, $\mathcal{P}_3$ becomes feasible when the value of K surpasses a certain threshold. For instance, at a demand density of $r = 0.4$, $\mathcal{P}_3$ is infeasible with the initially set K, but becomes feasible when K is increased to 20. Lastly, comparing TABLE III with TABLE IV, we find that energy consumption has a positive correlation with the spectral radius of the matrix.

3) Correlation among Parameters Density Distribution: We define the parameter density distribution as the normalized distribution of the parameter of the beam across all beams. Fig. 5 intuitively demonstrates the correlation among the density distribution of parameters such as demands, power, activation probability, and consumed energy.

Initially, the given demands are decomposed into power and activation probability with the proposed framework. Subsequently, the consumed energy can be calculated by multiplying the power by the activation probability. As expected, the resulting energy density distribution has great similarity with the given demands density distribution, which validates our proposed framework. Surprisingly, we also find that the probability density distribution has a positive correlation with that of the demands, which suggests an efficient heuristic approach to allocate the limited maximal number of simultaneously active beams.

TABLE III: Influence of K on energy

	$r = 0.1$	$r = 0.2$	$r = 0.3$	$r = 0.4$	$r = 0.5$
$K = 10$	14.89	24.43	NaN	NaN	NaN
$K = 15$	13.69	20.34	27.46	NaN	NaN
$K = 20$	13.12	18.77	23.43	39.74	NaN
$K = 25$	12.80	17.91	21.71	27.87	NaN
$K = 30$	12.58	17.37	20.72	25.34	30.80
$K = 35$	12.43	17.00	20.07	23.96	27.55
$K = 40$	12.32	16.73	19.60	23.07	25.93
$K = 45$	12.23	16.52	19.25	22.44	24.90
$K = 50$	12.16	16.36	19.00	21.97	24.21

[1] NaN represents that there is no optimal solution in the case.

[2] The energy is given by $10 \log 10(\boldsymbol{\rho}^T \mathbf{p})$

TABLE IV: Influence of K on spectral radius

	$r = 0.1$	$r = 0.2$	$r = 0.3$	$r = 0.4$	$r = 0.5$
$K = 10$	0.097	0.477	NaN	NaN	NaN
$K = 15$	0.075	0.265	0.644	NaN	NaN
$K = 20$	0.066	0.201	0.417	0.970	NaN
$K = 25$	0.062	0.171	0.326	0.669	NaN
$K = 30$	0.059	0.155	0.278	0.530	0.801
$K = 35$	0.057	0.144	0.249	0.452	0.655
$K = 40$	0.056	0.136	0.230	0.402	0.567
$K = 45$	0.055	0.131	0.216	0.369	0.511
$K = 50$	0.054	0.128	0.207	0.346	0.474

[1] NaN represents that there is no optimal solution in the case.

[2] The spectral radius is $\sigma(\mathbf{GA})$ of the optimal solution.

C. Demand Precompensation

As illustrated in Fig. 3, the proposed two-stage framework aims to design power $\hat{\mathbf{p}}$ and illumination pattern $\mathbf{x}$ for the satellite communication system such that the satellite's capacities (outputs, $\mathbf{c}$) for the ground users match correspondingly the required traffic demands (inputs, $\mathbf{d}$) of these users. In the first stage of the proposed framework, the formulated problem is conditioned on the Shannon Formula, which indicates that the capacity y and the SINR (x) are related by $y = \log_2(1 + x)$ as shown in the blue curve in Fig. 6(a).

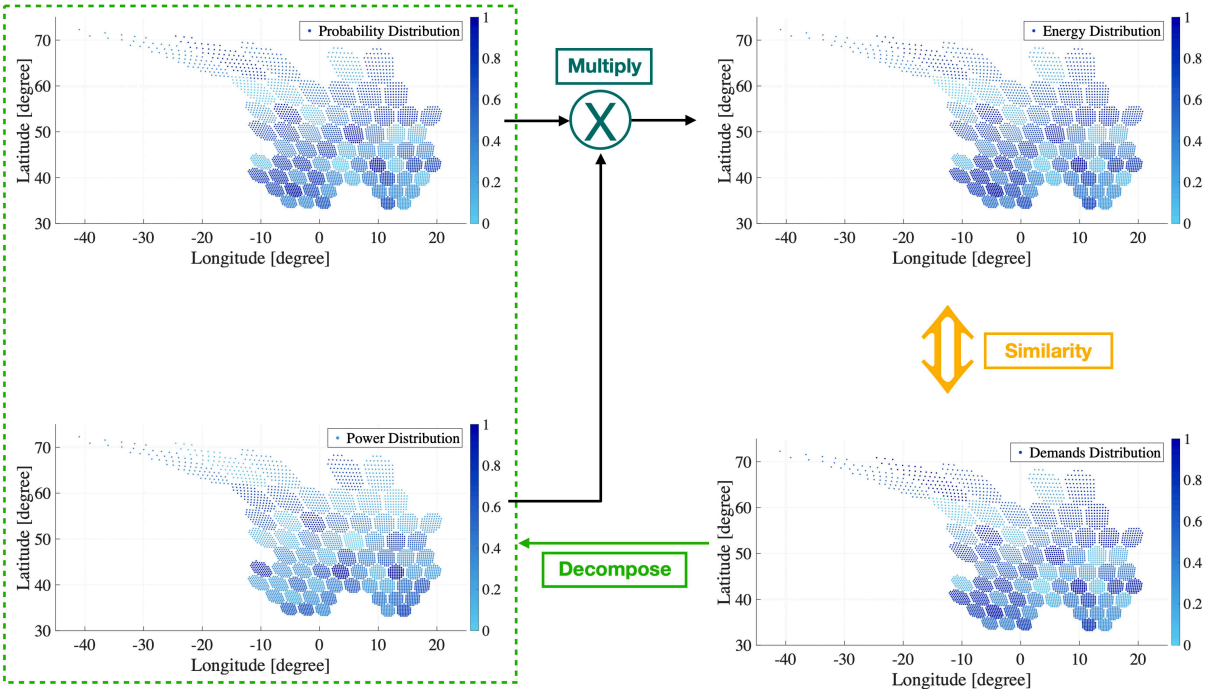


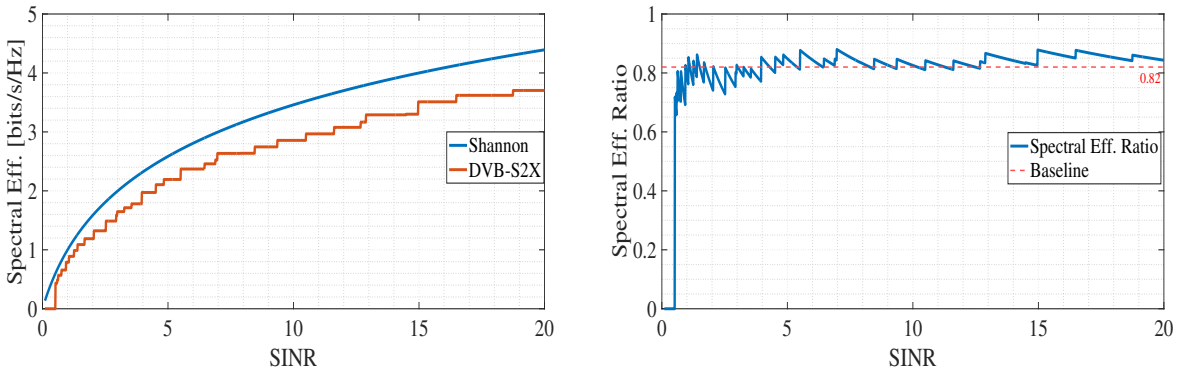
Fig. 5: Density distribution of parameters of an instance. The demand density is at $r = 0.3$.

However, in practice, the operator modulates the signal based on the range of the SINR following the DVB-S2X standard [23] when transmitting the information. Due to these modulations, there is a gap between the Shannon capacity and the DVB-S2X actual output, as seen clearly in Fig. 6(a). Taking a step further, we plot the ratio of DVB-S2X output to Shannon capacity along with SINR, which is the blue curve in Fig. 6(b). It is observed that the information transmission can be conducted well only when the SINR exceeds a particular level. Specifically, when the SINR exceeds 5 dB, the observed ratio surpasses 0.8. However, as long as the ratio is not 1, there is always a transmission loss systematically.

To compensate for the gap, a straightforward approach is to increase the input demands. This adjustment can be expressed as $\tilde{\mathbf{d}} = \mathbf{d}/\epsilon$, where ϵ is the chosen compensation coefficient. With a smaller coefficient, the required demands become higher, resulting in increased energy consumption.

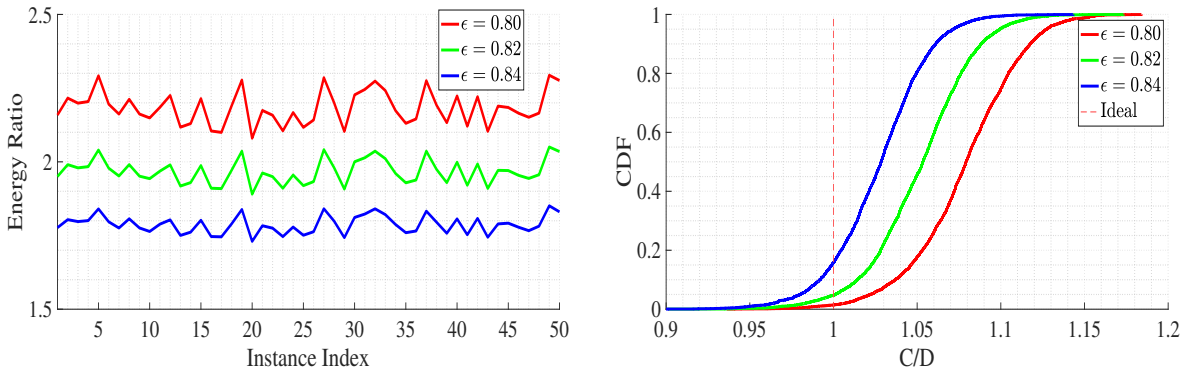
For performance comparison of energy consumption and demand-matching, we choose three different coefficients. The performance metric of energy ratio is defined as

$$\mathrm{ER}_1 = \frac{\boldsymbol{\rho}_\epsilon^{\star T}\mathbf{p}_\epsilon^\star}{\boldsymbol{\rho}_1^{\star T}\mathbf{p}_1^\star} \tag{22}$$



(a) (b)

Fig. 6: The comparison between Shannon and DVB-S2X standard in terms of spectral efficiency. (a) the spectral efficiency of DVB-S2X standard and Shannon capacity; (b) the ratio of DVB-S2X standard output to Shannon capacity.



(a) (b)

Fig. 7: Performance comparison with different compensation coefficients for the experiment on 50 instance with $r = 0.3$ and the value of K decided via the baseline [14]. (a) energy consumption performance; (b) demand matching performance.

where $\boldsymbol{\rho}_1^\star$ and $\mathbf{p}_1^\star$ are the optimal solutions to $\mathcal{P}_3$; $\boldsymbol{\rho}_\epsilon^\star$ and $\mathbf{p}_\epsilon^\star$ are the compensated ones with the parameter ϵ. The performance metric of demand matching is illustrated by the cumulative distribution function (CDF) of the ratio of the provided capacity to the required demand (C/D) of the beam.

In Fig. 7(a), the performance of the energy ratio is provided. It is observed that the proposed system would necessitate approximately 2.2 times the energy of the ideal system when $\epsilon = 0.80$,

whereas it is around 1.8 times when $\epsilon = 0.84$.

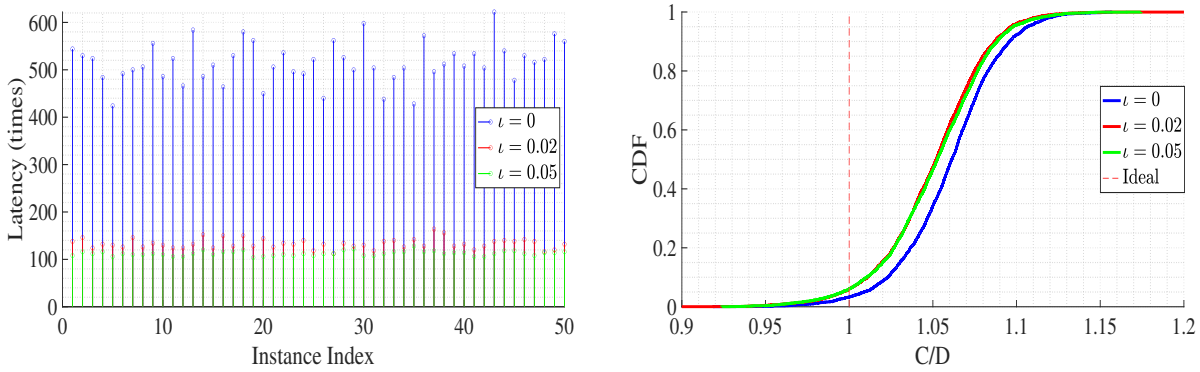
Fig. 7(b) demonstrates demand-matching performance with different compensation coefficients. The red dashed line indicates the ideal situation where the provided capacity perfectly matches the demand for all beams. It is observed that when $\epsilon = 0.80$, the demands of almost all beams are satisfied. However, with $\epsilon = 0.84$, more than 10% of the beams remain unsatisfied. Additionally, at $\epsilon = 0.82$, less than 5% of the beams are unsatisfied, with the satisfaction ratio ranging between 0.95 and 1. Based on the performance metrics shown in Fig. 7, we conclude that the selection of ϵ represents a trade-off between energy consumption and demand-matching performance. For all subsequent simulations, we will use $\epsilon = 0.82$.

D. Influence on Latency Trade-off Factor

Fig. 8 demonstrates the influence of the trade-off factor ι on the BH pattern design. Fig. 8(a) examines the impact of three factors on total latency at a demand density $r = 0.3$ across 50 instances. It is shown a significant reduction in latency when the trade-off factor exceeds 0. Moreover, as expected, the larger the factor, the lower the latency. Fig. 8(b) highlights the corresponding demand-matching performance. It is evident that, despite a latency penalty, all three curves of the demand-matching ratio fall within the range of 0.95 and 1.15, demonstrating the superior demand-matching performance of the proposed method. Additionally, as expected, the curve without latency penalties outperforms the latency-penalized curves. Specifically, fewer than 3% of the beams represented by the non-latency-penalized curve fail to meet the criteria, compared to over 6% of the beams depicted by the latency-penalized curve.

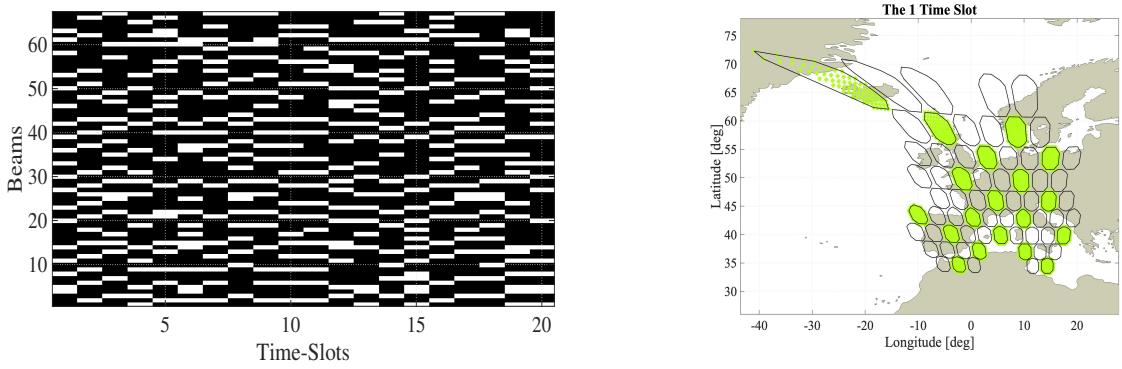
E. Designed Illumination Pattern

Fig. 9 shows an example of an illumination pattern designed by the proposed method at a particular demand instance obtained by $r = 0.3$. In Fig. 9(a), the white block refers to the corresponding active beams while the dark ones are passive. Moreover, the beam pattern corresponding to the first TS of the illumination pattern is illustrated in Fig. 9(b), where the green circle refers to the illuminated spot-beam.



(a) (b)

Fig. 8: The influence of trade-off factor on the performance. The demand density is at $r = 0.3$. (a) the total latency per instance; (b) the demand matching ratio per beam.



(a) (b)

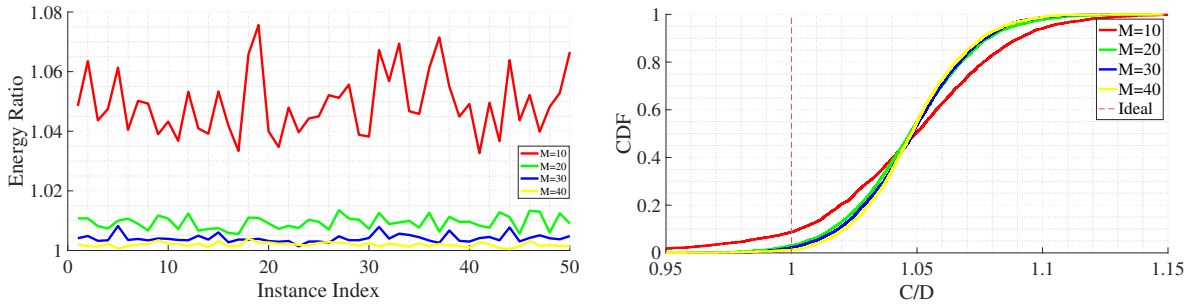
Fig. 9: The designed illumination pattern. (a) illumination pattern; (b) the specific beam pattern at one TS

F. Influence of Length of Time Window

The influence of the length of time window on the performance is also provided. We define the energy consumption performance metric as

$$\mathrm{ER}_2 = \frac{\hat{\boldsymbol{\rho}}^T \hat{\mathbf{p}}}{\boldsymbol{\rho}^{\star T} \mathbf{p}^{\star}} \tag{23}$$

where $\boldsymbol{\rho}^{\star}$ and $\mathbf{p}^{\star}$ denotes the optimal solution to $\mathcal{P}_3$, and $\hat{\boldsymbol{\rho}}$ and $\hat{\mathbf{p}}$ are the modified solutions obtained after rounding, defined by Eq. (17).



(a) (b)

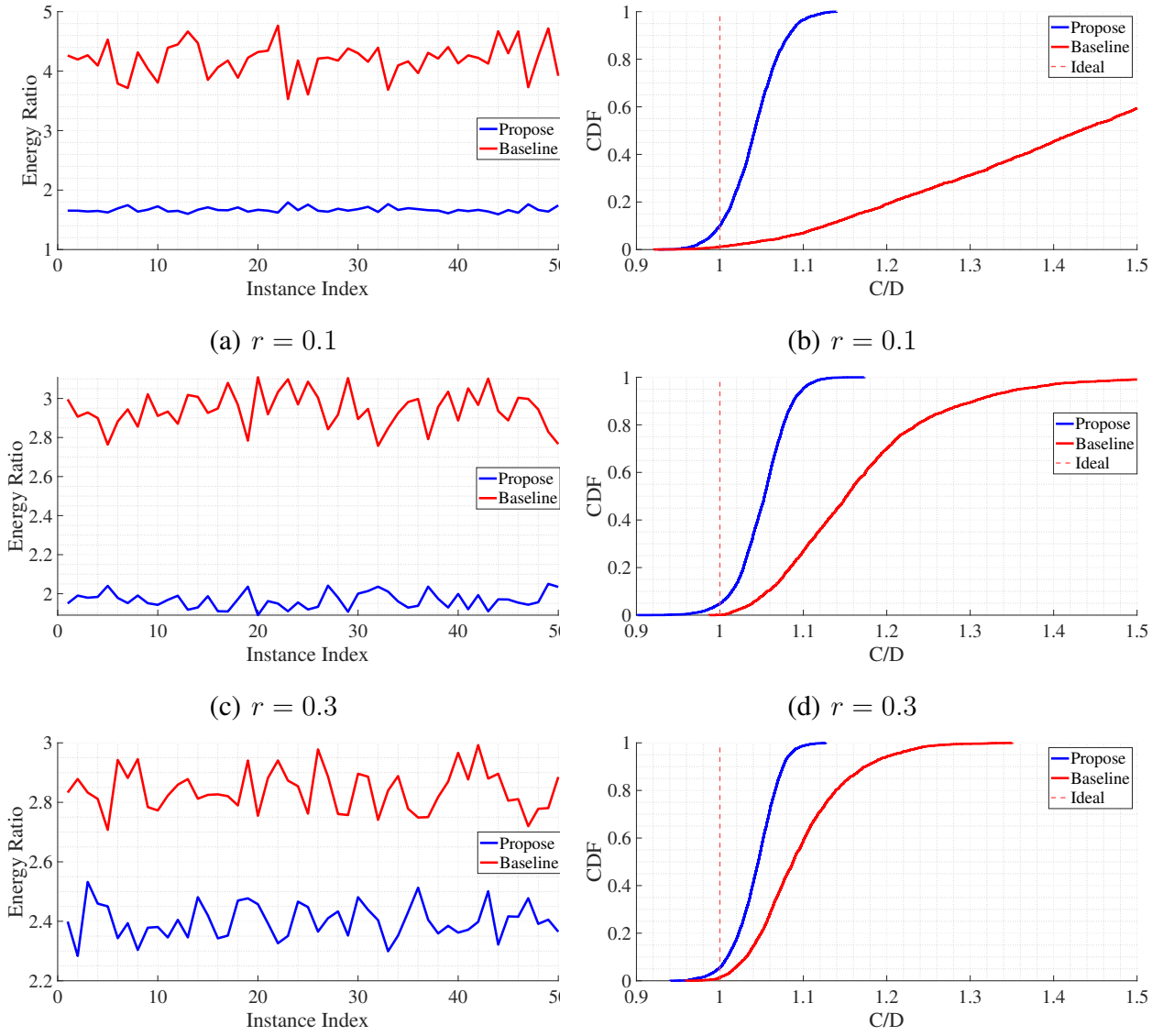
Fig. 10: The influence of M on the performance for the experiments on 50 instances with $r = 0.3$, $K = 21$ and $\epsilon = 0.82$. (a) the energy consumption ratio; (b) the demand matching

In Fig. 10(a), we can find that the higher the number of time slots (M) in the time window, the closer the distance between the approximated solution and the optimum. For instance, when M is small, i.e. $M = 10$, the total consumed energy increases by a maximum of 8% compared to the optimum. However, when M exceeds 20, the extra required energy stabilizes at less than 1.5% of the optimum. Notably, this amount drops to less than 0.5% when M exceeds 40. Fig. 10(b) illustrates the demand-matching performance. As observed in the figure, almost 10% of beams are not satisfied when $M = 10$, but more than 97% of beams are satisfied when M exceeds 20. Furthermore, the higher the M is, the better the demand-matching performance.

G. Comparison with the Baseline on Energy Efficiency and Demand Matching

Fig. 11 demonstrates the performance comparison with the baseline on energy consumption and demand matching across three different demand densities. The performance metric of energy consumption is ER_2 defined by Eq. (23) and the demand matching is by illustrating the CDF of the ratio of the provided capacity to the required demand of the beam. To ensure a fair comparison, both methods are constrained by the same value of K, determined by the method in baseline.

The left-hand side of the figures depicts the energy consumption performance. It is evident that the baseline method incurs significantly higher energy costs in all instances, amounting to approximately 2.5 times, 1.5 times, and 1.2 times the energy consumed by the proposed framework at demand densities $r = 0.1, 0.3, 0.5$, respectively. The right-hand side figures illustrate the demand-matching performance. Although the baseline successfully fulfills the requirements



(a) $r = 0.1$ (b) $r = 0.1$

(c) $r = 0.3$ (d) $r = 0.3$

(e) $r = 0.5$ (f) $r = 0.5$

Fig. 11: The performance comparison between the proposed method with the baseline.

for nearly all beams, it tends to exceed the necessary capacity. In contrast, the proposed method demonstrates greater stability in its performance. The demand-matching ratio generally falls within the range of 0.95 to 1.1 across all demand densities.

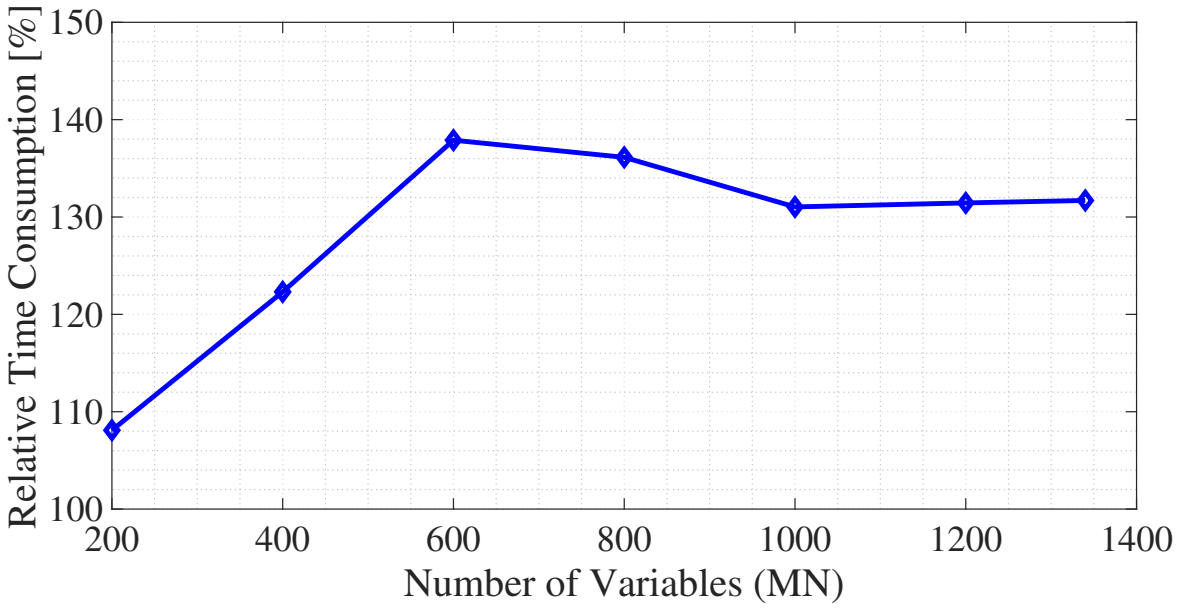


Fig. 12: The relative time-consumption comparison.

H. Comparison with Baseline on Computational Complexity

Fig. 12 presents the time consumption of the proposed framework relative to the baseline at the different number of variables. The y-axis value is the ratio of the time consumption of the proposed framework to that of the baseline, which is averaged over 50 instances at $r = 0.3,\ M = 20$. The number of variables is MN, where $N \in \{10, 20, 30, 40, 50, 60, 67\}$ is the number of beams and M is the number of TS in the time window. The convergence conditions of the iterative algorithms are given by $\frac{\|\mathbf{x}_{k+1}-\mathbf{x}_k\|_2}{\mathtt{dim}(\mathbf{x}_k)} \leq 10^{-3}$, where $\mathtt{dim}(\mathbf{x})$ denotes the dimension of the vector $\mathbf{x}$.

We also provide Table V, which compares the system performances of the proposed framework and the baseline. The Jain's Fairness Index is used to measure the satisfaction coverage of the users' demands and is defined as $f(\mathbf{y}) = \frac{(\sum_i y_i)^2}{\mathtt{dim}(\mathbf{y}) \sum_i y_i^2)}$, where the merit function $y_i = \frac{c_i}{d_i}$ is the ratio of the provided capacity to the request traffic demand of the beam [40]. The higher the index is, the better the scheme of the resource allocation would be. In addition, statistical parameters for the energy consumption ratio of the proposed framework to that of the baseline across 50 instances are also provided. Based on the results, we can see that, although the running time of the proposed framework is around 1.3 times that of the baseline, the system gets a notable performance improvement. Specifically, the proposed solution consumes only 65% of the energy

by the baseline while still providing slightly better demand-matching performance.

TABLE V: Performance comparison on resource allocation fairness and total consumed energy

	Number of Variables	200	400	600	800	1000	1200	1340
Jain's Fairness Index	Proposed	0.9994	0.9992	0.9992	0.9992	0.9991	0.9991	0.9990
	Baseline	0.9898	0.9904	0.9909	0.9920	0.9925	0.9930	0.9922
Consumed Energy	Mean	0.5119	0.6333	0.6159	0.6444	0.6647	0.6739	0.6766
	Min	0.3744	0.4839	0.5295	0.5608	0.5730	0.6125	0.6073
	Max	0.7931	0.8026	0.7801	0.7535	0.7564	0.7602	0.7583

[a] "Consumed Energy" denotes the ratio of energy consumption of the proposed framework to that of the baseline.

VII. Conclusion

In this paper, we propose a novel two-stage framework for optimizing energy consumption through the joint design of power and time slot allocation. The framework is designed to achieve optimal performance by addressing various challenges. In the first stage, we utilize mean-field theory to extract the activation probability and reformulate the mathematical model into an inverse matrix optimization problem. This reformulation enables us to convert the problem into a convex form, which has been thoughtfully analyzed and solved efficiently using a proposed iterative method. In the second stage, we employ the MPMM method to map the activation probability into the illumination pattern. Additionally, we introduce a compensation method to mitigate the performance loss resulting from the discrepancy between practical adaptive coding modulation and the ideal Shannon formula. Overall, this step yields a deterministic and practical solution for the considered beam-hopping satellite system. To validate our theoretical findings, we conduct numerical simulations. The results demonstrate that our proposed method surpasses the benchmark in terms of energy consumption and demand-matching performance.

Appendix A

Proof: Given $\mathbf{d}, \mathbf{H}$, and K, assume that the optimal solution to $\mathcal{P}_1$ is $(\boldsymbol{\rho}^\star, \mathbf{p}^\star)$. Consequently, this solution satisfies the demand constraint $\hat{C}_2$.

Suppose $(\boldsymbol{\rho}^\star, \mathbf{p}^\star)$ is not the fixed point, taking into account that it is the solution to the problem into account, then there $\exists n$ such that $\rho_n^\star > f_n(\boldsymbol{\rho}^\star, \mathbf{p}^\star)$ and $\rho_i^\star \geq f_i(\boldsymbol{\rho}^\star, \mathbf{p}^\star),\ \forall i \neq n$.

Consequently, there exists another point $(\hat{\boldsymbol{\rho}}, \mathbf{p}^{\star})$ such that $\hat{\rho}_n = f_n(\boldsymbol{\rho}^{\star}, \mathbf{p}^{\star})$ and $\hat{\rho}_i = \rho_i^{\star},\ \forall i \neq n$. Subsequently, according to the definition at Eq. (6), there would have $\hat{\rho}_i > f_i(\hat{\boldsymbol{\rho}}, \mathbf{p}),\ \forall i \neq n$, which proves that it is the feasible solution to the problem. Moreover, it consumes less energy than the supposed optimal solution, which contradicts the optimal assumption. Thus the optimal solution to $\mathcal{P}_1$ should be the fixed point of the equations. ■

APPENDIX B

Proof: The first-order and second-order derivatives of the function $g(z)$ are given by $g^{'}(z) = 2^{\frac{d}{Bz}}(1-\frac{d\ln 2}{Bz})-1$ and $g^{''}(z) = \frac{1}{z^3}\frac{(d\ln 2)^2}{B^2}2^{\frac{d}{Bz}}$ respectively. There always holds that $g^{''}(z) > 0, \forall z > 0$, so the first-order derivative of the function is monotonically increasing on the zone $(0, +\infty)$. Taking $g^{'}(+\infty) = 0$ into account, we would have $g^{'}(z) < 0, \forall z > 0$ and thereby completes the proof.

■

APPENDIX C

Proof: We define the given parameters are reasonable if the minimal spectral radius of the matrix is no greater than 1. If the condition is not met, the parameters are irreasonable. When given irreasonable parameters $\mathbf{H}, \mathbf{d}$, the set is empty, i.e. $\mathcal{Y} = \emptyset$, thus it is convex. In the following, we will first give the condition to check if the given parameters are reasonable, and then prove the convexity of the set when given reasonable parameters.

Firstly, we define $\mathbf{A} \geq \mathbf{B}$ if $A_{i,j} \geq B_{i,j},\ \forall i, j$, and $\mathbf{A} > \mathbf{B}$ if $\mathbf{A} \geq \mathbf{B},\ \mathbf{A} \neq \mathbf{B}$. According to [29, Corollary 1.5 on p.27], given matrix $\mathbf{A}, \mathbf{B}$, if $\mathbf{A} > \mathbf{B} > \mathbf{0}$, there would have $\sigma(\mathbf{A}) > \sigma(\mathbf{B})$, where $\mathbf{0}$ is the matrix all of whose elements are 0. Moreover, based on the *Lemma* 4, the function $g(z)$ is monotonically decreasing, thus the minimal spectral radius of the matrix $\mathbf{GA} = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}))\mathbf{A}$ is at the point $\boldsymbol{\rho} = \mathbf{1}$ as $\mathbf{1} \succeq \boldsymbol{\rho} \in \mathcal{Y}$. According to the previous definition, the condition is to check if the spectral radius of the matrix $\mathtt{Diag}(\mathbf{g}(\mathbf{1}))\mathbf{A}$ is less than 1.

Secondly, when given reasonable parameters, the set $\mathcal{Y}$ is non-empty. Suppose $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2 \in \mathcal{Y}$, then there have $\sigma(\mathbf{G}_1\mathbf{A}) \leq 1, \sigma(\mathbf{G}_2\mathbf{A}) \leq 1$, where $\mathbf{G}_i = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}_i)),\ i = 1, 2$.

Assume that $\boldsymbol{\rho} = \theta\boldsymbol{\rho}_1 + (1-\theta)\boldsymbol{\rho}_2, \forall\ \theta \in [0\ 1]$. According to [41, Remark 1.3], if the function $g(z)$ is log-convex, there would have $\sigma(\mathbf{GA}) \leq \sigma(\mathbf{G}_1\mathbf{A})^{\theta} \cdot \sigma(\mathbf{G}_2\mathbf{A})^{(1-\theta)} \leq 1$, which proves that the convexity of the set is conditioned on the log-convex of the function $g(z)$. In the following, we will prove the log-convex of the function $g(z)$.

According to [35], that a function is log-convex is equivalent to the positivity of the function $h(z) = g(z) \cdot g(z)^{''} - (g(z)^{'})^2$. The function $h(z)$ and its first-order derivative can be given by $h(z) = (t^2 - t)\frac{m^2}{z^2} - [t(1-m)-1]^2$ and $h^{'}(z) = \frac{m^2}{z^3}\{2t^2m(z^2-1) + t(m-4t) + z^2(2t-2t^2)\}$ respectively, where $t = 2^{\frac{d}{Bz}}, m = \frac{d\ln 2}{Bz}$. Considering that there would have $(z^2-1) \leq 0, (m-4t) < 0, (2t-2t^2) < 0,\ \forall\ 0 < z \leq 1$, subsequently $h^{'}(z) < 0$, which proves that the function $h(z)$ is strictly decreasing. Taking into account that $h(\infty) = 0$, thus $\forall z > 0, h(z) > 0$. According to [35], the function $g(z)$ is log convex, which thereby completes the proof.

■

APPENDIX D

Proof: According to *Lemma* 2, the set $\hat{\mathcal{Y}}$ is the intersection of two convex sets, thus it is convex [35]. In the following, we will first prove the convexity of $\mathcal{P}_3$ and then demonstrate that the solutions to $\mathcal{P}_2$ and $\mathcal{P}_3$ are equivalent.

The total derivatives of the objective function $L(\boldsymbol{\rho})$ is given by

$$\mathrm{d}L(\boldsymbol{\rho}) = \mathrm{Tr}\{\mathrm{d}[(\mathbf{I}-\mathbf{GA})^{-1}]\mathbf{GB} + (\mathbf{I}-\mathbf{GA})^{-1}\mathrm{d}[\mathbf{G}]\mathbf{B}\} \tag{24a}$$

$$= \mathrm{Tr}\{\underbrace{(\mathbf{A}(\mathbf{I}-\mathbf{GA})^{-1}\mathbf{GB} + \mathbf{B})((\mathbf{I}-\mathbf{GA})^{-1}}_{\mathbf{Z}}\mathrm{d}[\mathbf{G}]\} \tag{24b}$$

where $\mathbf{B} = \mathbf{b} \cdot \mathbf{1}^T$.

Note that $\mathbf{G} = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}))$ is a diagonal matrix, thus

$$\mathrm{d}[\mathbf{G}] = \mathtt{Diag}([g^{'}(\rho_1)d\rho_1, \ldots, g^{'}(\rho_N)d\rho_N]^T) \tag{25}$$

According to [42] and Eq. (24b, 25), the first-order derivative of the objective $L(\boldsymbol{\rho})$ is

$$\frac{\partial L}{\partial \rho_i} = Z_{i,i}g^{'}(\rho_i),\ \forall i \tag{26}$$

Subsequently the Hessian matrix of the objective is given by $\frac{\partial L^2}{\partial \rho_i \partial \rho_j} = \begin{cases} 0 & i \neq j \\ Z_{i,i}g^{''}(\rho_i) & i = j \end{cases}$, which is a diagonal matrix.

In the following, we will prove the positive of the elements in the diagonal of the matrix. First, according to *Lemma* 1, we have $g^{''}(\rho_i) > 0,\ \forall \rho_i > 0$. Second, when $\boldsymbol{\rho} \in \hat{\mathcal{Y}} \setminus \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) = 1\}$, there have $(\mathbf{I}-\mathbf{GA})^{-1} \succ \mathbf{0}$ because of $\sigma(\mathbf{GA}) < 1$. Then, according to the definition of $\mathbf{Z}$ in (24b), there have $Z_{i,i} > 0\ \forall i$. Taking these two parts into account, the Hessian matrix of

the objective is positive definite at the domain $\boldsymbol{\rho} \in \hat{\mathcal{Y}} \setminus \{\boldsymbol{\rho}|\sigma(\mathbf{GA}) = 1\}$. In addition, when $\boldsymbol{\rho} \in \{\boldsymbol{\rho}|\sigma(\mathbf{GA}) = 1\}$, there would have $\lim_{\sigma(\mathbf{GA})\to 1} L(\boldsymbol{\rho}) = +\infty$. To summarize, $\mathcal{P}_3$ is convex.

The difference between $\mathcal{P}_2$ and $\mathcal{P}_3$ is that the feasible set of $\mathcal{P}_3$ contains the boundary points, i.e. $\boldsymbol{\rho} \in \{\boldsymbol{\rho}|\sigma(\mathbf{GA}) = 1\} \cup \{\mathbf{0}\}$, while $\mathcal{P}_2$ does not. However, the optimum would not be located at these boundaries as $\lim_{\sigma(\mathbf{GA})\to 1} L(\boldsymbol{\rho}) = +\infty$. Thereby we complete the proof. ■

APPENDIX E

Proof: According to the Appendix D , the first order derivative of $L(\boldsymbol{\rho})$ is $\frac{\partial L}{\partial \rho_i} = Z_{i,i} g^{'}(\rho_i)$, where $\mathbf{Z} = (\mathbf{A}(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{GB} + \mathbf{B})((\mathbf{I} - \mathbf{GA})^{-1}$. For $\forall \boldsymbol{\rho} \in \hat{\mathcal{Y}}$, there would have $\sigma(\mathbf{GA}) \leq 1$, and thus $Z_{i,i} > 0,\ \forall i$. In addition, according to *Lemma* 1, $g^{'}(z) < 0, \forall z > 0$. To sum up, the first-order derivative of the objective is negative, so the total consumed energy is monotonically decreasing within the feasible zone. ■

APPENDIX F

Proof: Suppose $\boldsymbol{\rho}^{\star}$, the solution to $\mathcal{P}_3$, is not on the boundary, i.e. $(\boldsymbol{\rho}^{\star})^T\mathbf{1} < K$. Consequently, without losing generality, we can always construct a point $\boldsymbol{\rho}^{+}$ in the following manner: given any index n, the set $\rho_j^{+} = \rho_j^{\star}$, for all $j \neq n$, and $\rho_n^{+} > \rho_n^{\star}$. This configuration also satisfies the condition $(\boldsymbol{\rho}^{+})^T\mathbf{1} = K$.

Because of the monotonically decreasing of the function $g(z)$ and the corollary in [29, Corollary 1.5 p-27], there would have $\sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}^{\star}))) > \sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}^{+})))$. In addition, taking the constraint $\sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}^{\star}))) \leq 1$ into account, there has $\sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}^{+}))) < 1$. Consequently, the constructed point is in the feasible set, i.e. $\boldsymbol{\rho}^{+} \in \hat{\mathcal{Y}}$.

However, according to the *Lemma* 4, the objective $L(\boldsymbol{\rho})$ is monotonically decreasing. Thus there would have $L(\boldsymbol{\rho}^{\star}) > L(\boldsymbol{\rho}^{+})$ as $\boldsymbol{\rho}^{\star} \preceq \boldsymbol{\rho}^{+}$. This contradicts the optimal assumption. Therefore, if a solution to $\mathcal{P}_3$ exists, it would be located on the boundary. This implies that the system achieves minimal energy consumption when all available beams are activated simultaneously. With this, we complete the first part of the theorem.

As seen in Fig. 2, the line AB would move parallel to the direction of vector $\vec{OD}$ when K is increased, resulting in the solution of the former preceding the latter, i.e. $\boldsymbol{\rho}^{\star}_{former} \prec \boldsymbol{\rho}^{\star}_{latter}$. Considering that the objective function is monotonically decreasing, we deduce that $L(\boldsymbol{\rho}^{\star}_{former}) > L(\boldsymbol{\rho}^{\star}_{latter})$. This indicates that increasing the maximum number of active beams reduces the consumed energy, thereby completing the proof. ■

REFERENCES

[1] 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 *et al.*, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 70–109, 2020.

[2] G. Araniti, A. Iera, S. Pizzi, and F. Rinaldi, "Toward 6G non-terrestrial networks," *IEEE Network*, vol. 36, no. 1, pp. 113–120, 2022.

[3] G. Cocco, T. De Cola, M. Angelone, Z. Katona, and S. Erl, "Radio resource management optimization of flexible satellite payloads for DVB-S2 systems," *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 266–280, 2017.

[4] G. Taricco and A. Ginesi, "Precoding for flexible high throughput satellites: Hot-spot scenario," *IEEE Transactions on Broadcasting*, vol. 65, no. 1, pp. 65–72, 2019.

[5] A. Freedman, D. Rainish, and Y. Gat, "Beam hopping how to make it possible," in *Proc. of Ka and Broadband Communication Conference, Bologna, Italy*, Oct. 2015.

[6] A. Morello and N. Alagha, "DVB-S2X air interface supporting beam hopping systems," in *25th Ka and Broadband Communications Conference, (Ka-2019), Sorrento, Italy*, Oct. 2019, pp. 1–5.

[7] C. Rohde, D. Rainish, A. Freedman, G. Lesthievent, N. Alagha, D. Delaruelle, G. Mocker, and X. Giraud, "Beam-hopping system configuration and terminal synchronization schemes," in *Proceedings of the 37th International Communications Satellite Systems Conference (ICSSC-2019)*, 2019, pp. 1–13.

[8] T. Posielek, "An energy management approach for satellites," in *69th International Astronautical Congress, Bremen, Germany*, Oct. 2018, pp. 1–5.

[9] C. N. Efrem and A. D. Panagopoulos, "Dynamic energy-efficient power allocation in multibeam satellite systems," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 228–231, 2019.

[10] Y. Yang, M. Xu, D. Wang, and Y. Wang, "Towards energy-efficient routing in satellite networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3869–3886, 2016.

[11] P. Angeletti, D. F. Prim, and R. Rinaldo, "Beam hopping in multi-beam broadband satellite systems: System performance and payload architecture analysis," in *24th AIAA International Communications Satellite Systems Conference*. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2006-5376

[12] A. Ginesi, E. Re, and P. Arapoglou, "Joint beam hopping and precoding in hts systems," in *9th Int. Conf. on Wireless and Satellite Systems (WiSATS)*, 2017.

[13] M. G. Kibria, E. Lagunas, N. Maturo, D. Spano, and S. Chatzinotas, "Precoded cluster hopping in multi-beam high throughput satellite systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[14] L. Chen, V. N. Ha, E. Lagunas, L. Wu, S. Chatzinotas, and B. Ottersten, "The next generation of beam hopping satellite systems: Dynamic beam illumination with selective precoding," *IEEE Transactions on Wireless Communications*, 2022.

[15] L. Lei, E. Lagunas, Y. Yuan, M. G. Kibria, S. Chatzinotas, and B. Ottersten, "
" *IEEE Access*, vol. 8, pp. 136 655–136 667, 2020.

[16] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Satellite-terrestrial coordinated multi-satellite beam hopping scheduling based on multi-agent deep reinforcement learning," *IEEE Transactions on Wireless Communications*, 2024.

[17] J. P. Choi and V. W. Chan, "Optimum power and beam allocation based on traffic demands and channel conditions over satellite downlinks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2983–2993, 2005.

[18] T. S. Abdu, S. Kisseleff, E. Lagunas, and S. Chatzinotas, "Flexible resource optimization for geo multibeam satellite communication system," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7888–7902, 2021.

[19] I. Siomina and D. Yuan, "Analysis of cell load coupling for lte network planning and optimization," *IEEE Transactions*

on Wireless Communications, vol. 11, no. 6, pp. 2287–2297, 2012.

[20] C. K. Ho, D. Yuan, L. Lei, and S. Sun, "Power and load coupling in cellular networks for energy optimization," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 509–519, 2014.

[21] M. Opper and D. Saad, *Advanced mean field methods: Theory and practice*. MIT press, 2001.

[22] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.

[23] D. V. Broadcasting(DVB)Project, "Digital video broadcasting(DVB); second generation DVB interactive satellite system (DVB-S2X)," *ETSI EN 302 307 V1.4.1*, 2019.

[24] C. Rohde, R. Wansch, G. Mocker, A. Trutschel-Stefan, L. Roux, E. Feltrin, H. Fenech, and N. Alagha, "Beam-hopping over-the-air tests using DVB-S2X super-framing," in *36th International Communications Satellite Systems Conference (ICSSC 2018)*, 2018.

[25] 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," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 112–131, 2019.

[26] A. Granas and J. Dugundji, *Fixed point theory*. Springer, 2003, vol. 14.

[27] W. I. Zangwill, *Nonlinear programming: a unified approach*. Prentice-Hall Englewood Cliffs, NJ, 1969, vol. 52.

[28] C. K. Ho, D. Yuan, and S. Sun, "Data offloading in load coupled networks: A utility maximization framework," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 1921–1931, 2014.

[29] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

[30] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[31] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, no. 15, p. 510, 2008.

[32] M. Grant, S. Boyd, and Y. Ye, "Disciplined convex programming," *Global optimization: From theory to implementation*, pp. 155–210, 2006.

[33] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[34] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014.

[35] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[36] F. Curtis and J. Nocedal, "Steplength selection in interior-point methods for quadratic programming," *Applied mathematics letters*, vol. 20, no. 5, pp. 516–523, 2007.

[37] J. Dahl and E. D. Andersen, "A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization," *Mathematical Programming*, vol. 194, no. 1-2, pp. 341–370, 2022.

[38] L. Chen, E. Lagunas, L. Lei, S. Chatzinotas, and B. Ottersten, "Adaptive resource allocation for satellite illumination pattern design," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*. IEEE, 2022, pp. 1–6.

[39] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.

[40] R. K. Jain, D.-M. W. Chiu, W. R. Hawe *et al.*, "A quantitative measure of fairness and discrimination," *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*, vol. 21, p. 1, 1984.

[41] R. D. Nussbaum, "Convexity and log convexity for the spectral radius," *Linear Algebra and its Applications*, vol. 73, pp. 59–122, 1986.

[42] X.-D. Zhang, *Matrix analysis and applications*. Cambridge University Press, 2017.

X Research Paper 3 – *Beam Hopping for LEO Constellation Systems*

Full Title:
Joint Time Slot, Power and Required Load Optimization for Energy-Efficient LEO Constellation Satellite Systems

Authors:
Lin Chen, Linlong Wu, Eva Lagunas, Symeon Chatzinotas, Björn Ottersten

Submitted to:

Abstract:
The commercial development in telecommunication involves an intelligent system that integrates terrestrial base stations with an LEO constellation system, capable of providing communication services to users worldwide. Beam Hopping is one of the promising techniques that makes this intelligent system possible, wherein spot beams are selectively illuminated based on users' traffic demands. The conventional intelligent system aims to save transmitted energy by minimizing the gap between the provided capacity and the required uneven traffic, achieved through flexible resource allocation in both power and time domains. However, the serving area of adjacent satellites overlaps; conventional fixed connections between satellites and cells in the overlapping area waste flexibility. In this paper, we propose a novel optimization problem that maximizes the utilization of flexibility in the power and time of beams, as well as in the responsibilities of satellites to cells in overlapping areas, aiming to minimize the total consumed en-

ergy. The formulated inverse matrix minimization problem is proven to be convex and then optimally addressed by the proposed iterative method. Taking into account the gap between the output of DVB-S2 and the ideal Shannon Formula based on which the problem is formulated, we propose traffic compensation and SINR protection measures to guarantee the demand-matching performance for practical systems. Numerical simulations demonstrate the benefits brought by the proposed method in terms of energy saving and demand-matching performance for the LEO constellation system.

Keywords:
Energy Minimization, Power Control, Beam Hopping, LEO Constellation

Joint Time and Power and Required Load Optimization for Energy-Efficient LEO Satellite Constellation Systems

Lin Chen, Eva Lagunas, Symeon Chatzinotas

Abstract

The commercial development of telecommunication involves an intelligent system that integrates terrestrial base stations with the Low Earth Orbit (LEO) satellite constellation systems, capable of providing communication services to users worldwide. Beam Hopping is one of the promising techniques that makes this intelligent system possible, wherein spot beams are selectively illuminated based on users' traffic demands. LEO satellite constellation systems feature a virtual cell on the ground being served by multiple beams, each per satellite, however, this provided flexibility in optimizing the required loads of beams that have not been exploited in the conventional smart systems. In this paper, we propose a novel optimization problem that, besides the natural flexibility of resources onboard in the time and power domains, optimizes the required loads of beams to minimize energy consumption. We employ the load coupling model to address the demand satisfaction constraint, where parameters are represented by their expectations. Based on the optimum condition, the problem is reformulated as an inverse matrix minimization style, whose objective is then proved to be convex. An iterative method is proposed to solve the problem with a high-quality solution. Moreover, taking into account the gap between the output of the DVB-S2 standard and the ideal Shannon Formula based on which the problem is formulated, we propose traffic compensation and SINR protection measures to guarantee the demand-matching performance for practical systems. Numerical simulations demonstrate the benefits brought by the proposed method in terms of energy saving and demand-matching performance for the LEO satellite constellation system.

Index Terms

Energy Minimization, Power Control, Beam Hopping, LEO Constellation

This work was supported by the Luxembourg National Research Fund (FNR) under the project "FlexSAT: Resource Optimization for Next Generation of Flexible SATellite Payloads" (C19/IS/13696663). The authors are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, 1855 Luxembourg Ville, Luxembourg. The corresponding author is Lin Chen(lin.chen@uni.lu).

I. Introduction

Satellite communication (SatCom) systems play an indispensable role in modern life, providing global connectivity, supporting scientific research, and promoting economic growth [1]. It has the advantage of wide coverage and supporting terminals located in remote areas, making them a vital component of the global communication infrastructure [2]. LEO satellite constellation systems have developed rapidly recently, such as Starlink, Eutelsat OneWeb, and so on. Each satellite is designed to provide communication service to users located in its coverage. Moreover, to guarantee global coverage, the coverage area of adjacent satellites has common serving areas. How to utilize the limited onboard resources while satisfying the heterogeneous traffic demands of on-ground user terminals is a major problem in LEO constellation systems [3].

The next generation of high-throughput satellites is equipped with multiple antennas radiating spot beams that are featured with high gain and narrow beam width [4]. These spot beams make it possible to address the above problem, providing flexibility in allocating resources in time, power, and frequency domains to meet the uneven geographically distributed traffic demands of terminals. However, how to make full use of the provided flexibility motivates the development of the onboard intelligent resource allocation algorithm.

Recently, Beam Hopping (BH) has drawn the researcher's attention, which selectively activates the beams according to the beams' traffic demands [5]. In other words, unlike the traditional fixed-beam system, the BH-enabled SatCom can dynamically activate the beams based on their traffic needs. Moreover, a list of superframe specifications is designed in the DVB-S2X standard [6] to support the application of BH technology in SatCom systems. SatCom systems benefit significantly from the utilization of beam hopping (BH) technology. This technology allows the system to fully exploit flexibility across four domains to serve user terminals: the time domain, power domain, frequency domain, and space domain. Additionally, selectively activating beams would reduce the interference from other beams, and consequently increase the spectral efficiency. Lastly, BH requires only a limited number of beams that can be activated simultaneously, which reduces the need for the radio frequency chains and, consequently the launch cost of the satellite.

Satellite communication systems should efficiently manage energy consumption, which is primarily provided by a combination of solar power and battery storage. The equipped solar panels transfer solar energy into electricity, which powers wireless communication transmission and other applications [7]. Taking into account that a solar panel can only collect limited energy,

increasing system energy efficiency helps to reduce the number of solar panels, and consequently the mass of the satellite and the cost of launch. Thus, it is important to minimize the energy consumption for satellite communication [8].

Flexibly allocating the onboard resources to match the uneven traffic demands of user terminals has drawn researchers' attention in recent years. A heuristic algorithm is proposed to allocate bandwidth per beam such that the provided capacity matches the traffic demands per beam in [9]. In [10], the researchers propose a power allocation method based on traffic distribution and channel conditions. Moreover, [11], [12], and [13] is proposed to combine the BH technique with the interference mitigation method such that the provided capacity satisfies the required traffic demands. Especially, [12] proposed a novel general framework to design the illumination pattern by penalizing any pair of simultaneously active beams with the aim of minimizing the total penalty. The above works consider neither the power nor the time domain, wasting not exploring the full flexibility of resources. Although, [14] proposed to jointly allocate the power and bandwidth per beam to minimize energy consumption while meeting the traffic demands, which does not fully utilize the rare spectral efficiency. The works as mentioned above only consider a single satellite, which does not suit LEO constellation systems. [15] deals with the multiple satellites problem, where the authors propose jointly optimizing resources in various domains. However, [15] failed to consider the required load allocation among beams serving the same cell. [16] proposed to address this issue by balancing the load per satellite first, which lacks theoretical validation. The load coupling model, originally analyzed by [17], characterizes the coupling relationships among the temporal occupation of beams. The model has been employed, for instance, in [18], the work most closely related to ours, where the power and load of beams are studied for cellular networks to minimize energy consumption. However, [18] does not consider the constraint on the maximal number of simultaneously active beams, an issue that is addressed in our work.

However, these works do not fully utilize the flexibility of onboard resources, wasting energy consumption. To address the problem, we propose to make full flexibility in time, power, and required loads of beams to minimize energy consumption while meeting the uneven traffic demands of cells. First, we formulate an ideal illumination pattern design problem, within which we employ the load coupling model to deal with the complicated demand-matching constraints, within which the three variables are coupled. To address the formulated problem, we find a necessary condition for the optimum, based on which we find that power is a function of

activation probability and the required loads of beams. Based on the equation, we reformulate the problem solved by the proposed iterative method. Second, the obtained activation probability is mapped into the illumination pattern by solving a binary quadratic programming problem.

Our major contributions are summarized as follows:

- We propose a joint design of power, time slot, and required loads of beams so as to minimize total transmission energy consumption while satisfying the uneven traffic demands for the LEO constellation systems for the first time. We prove the convexity of the objective, based on which we propose an iterative method to solve the problem.
- We propose methods to compensate for the systematic performance loss. There is an output gap between the practical coding modulation method and the Shannon formula, based on which the problem is formulated.
- Numerical simulations validate the theoretical finding that the infeasible problem with fixed required loads of beams becomes feasible when there is flexibility in the required loads of beams.

The remainder of the paper is organized as follows. In Section II, we present the system model. In Section III, we formulate the ideal pattern design problem. We analyze the problem and introduce methods to solve the problem in Section IV. In Section V, we analyze the impact of systematic parameters on the performance and introduce methods to implement practical illumination pattern design. Numerical simulations are conducted in Section VI, and lastly, we conclude in Section VII. Notations used in the paper are summarized in Table **??**.

II. System Model

In this section, we first introduce the considered BH-enabled regional multiple satellite systems. Next, we consider the load coupling model for the system.

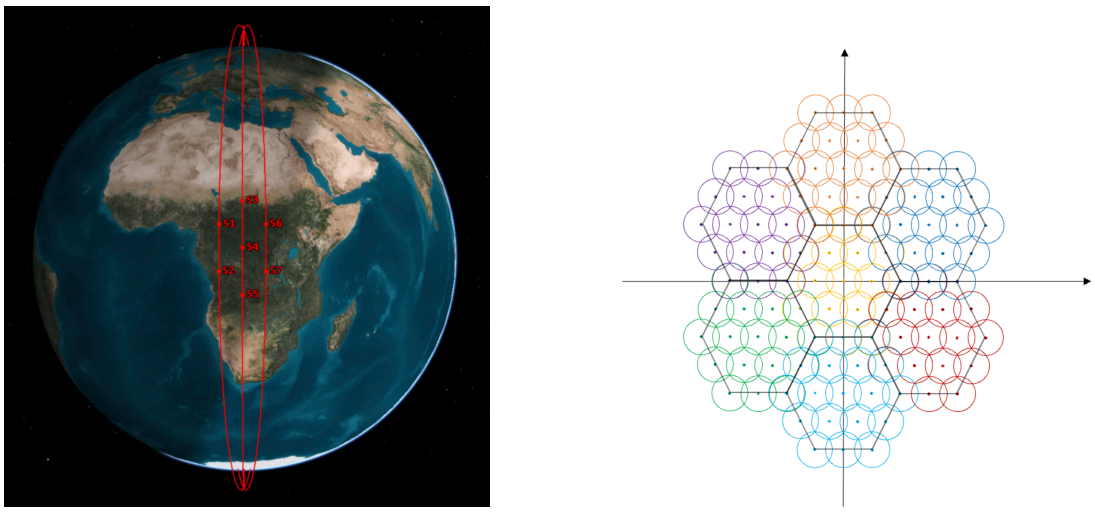
A. Preliminaries

In this paper, the forward-link transmission of regional multiple satellites is considered. The layout of satellites follows the standard Walker Delta constellation [19], where orbit planes at the same inclination and spread evenly around the equator, and satellites spread evenly around each orbit. Specifically, we denote $j,\ j \in \mathcal{N}_s = \{1, \ldots, N_s\}$ as the satellite j, and there are N_s satellites in total to serve a specific area covered by N_c virtual cells. The set of all virtual cells is denoted as $\mathcal{I}$, subsequently, it holds that $N_c = |\mathcal{I}|$, where $|\cdot|$ denotes the cardinality of the

TABLE I: Glossary of notations

Notation	Definition
$\mathcal{N}_s$	Set of index of satellites
N_s	Total number of satellites
$\mathcal{I}$	Set of index of virtual cells
N_c	Total number of cells
$\lvert\cdot\rvert$	cardinality of the set
$\mathcal{A}_i$	Set of index of satellites covering cell i
$\mathcal{B}_j$	Set of index of cells covered by satellite j
$J_1(x)$	Bessel function of the first kind and first order with argument x
a	Antenna's circular aperature
f	Frequency of operation
n_i^j	Index of the antenna on the satellite j, which is designed to serve the virtual cell i
ϑ_{i,n_i^j}	Angle between a ray from antenna n_i^j to the user terminal located in cell i and the bore sight of the antenna's main beam.
G_0	Maximal antenna gain
H_{i,n_i^j}	Channel between the terminal located in cell i and satellite j
q_n	Traffic demands of cell n in bps
B	Full spectrum bandwidth
M	Total number of time slots within the time window
ΔT	Duration of time slot
K_j	Maximal number of active beams of satellite j
$p_{i,j}$	Transmit power to cell i from satellite j
$\rho_{i,j}$	Activation probability to cell i from satellite j
$R_{i,j}$	Expected achievable rate of the user located in cell i from satellite j
δ_T^2	Thermal noise power
τ	Boltzmann constant
T_{Rx}	Clear sky noise temperature of receiver
$d_{i,j}$	Required load of the beam serving cell i, radiated by the antenna on satellite j
$\mathbf{p}_j$	Power vector of all beams radiated by satellite j
$\mathbf{p}$	Power vector of all beams of the systems
$\boldsymbol{\rho}_j$	Activation probability vector of all beams of satellite j
$\boldsymbol{\rho}$	Activation probability vector of all beams
$P_{\max}^j$	Maximal power of satellite j
$\mathrm{Tr}\{\cdot\}$	Trace of the matrix

set. The set of cells covered by the satellite j is denoted as $\mathcal{B}_j$. Thus, $\mathcal{I} = \bigcup_{j=1}^{N_s} \mathcal{B}_j$. We denote $\mathcal{A}_i, i \in \mathcal{I}$ as the set of index of satellites covering the cell i. Taking into account that some of the cells are covered by multiple satellites, subsequently, it holds that $|\mathcal{A}_i| \geq 1, \forall i$. An example of



(a) (b)

Fig. 1: The coverage of beams for regional multiple satellites. (a) regional multiple satellites; (b) the layout of beam coverage.

the layout of virtual cells for regional multiple satellites is illustrated in Fig. 1, where there are 7 satellites, each of which has 19 spot-beams serving corresponding cells and are in a different color. Cells located in the marginal areas of adjacent satellites can be served by all of these satellites.

The antenna gain pattern, corresponding to a typical reflector antenna with a circular aperture, following the work [20], is expressed as

$$G^{Tx}(\vartheta_{i,n_i^j}) = \begin{cases} 4G_0|\dfrac{J_1(ka\sin\vartheta_{i,n_i^j})}{ka\sin\vartheta_{i,n_i^j}}|^2, & 0 < |\vartheta_{i,n_i^j}| \leq \pi/2 \\ 1, & \vartheta_{i,n_i^j} = 0 \end{cases} \tag{1}$$

where $J_1(x)$ is the Bessel function of the first kind and first order with argument x; a is the antenna's circular aperture; $k = 2\pi f/c$ is the wave number; f is the frequency of operation; c is the speed of the light in vacuum; n_i^j denotes the index of the antenna on satellite j, which is designed to serve the virtual cell i; and ϑ_{i,n_i^j} is the angle between the ray from the antenna n_i^j to the user terminal located in cell i and the bore sight of the antenna's main beam. Lastly, G_0 is the maximal antenna gain.

The free space path loss is expressed as

$$L_{i,j} = (\frac{c}{4\pi f l_{i,j}})^2 \tag{2}$$

where $l_{i,j}$ describes the distance between the satellite j and the user within the cell i.

Assuming that the Doppler frequency offset caused by the satellite motion is compensated by the user and the sky is clear, thus rain attenuation is not considered. The channel coefficient between the user within the cell i and the corresponding antenna n_i^j of the satellite j, similar to the expression in [21], can be given by

$$H_{i,n_i^j} = \sqrt{G^{Tx}(\vartheta_{i,n_i^j})G^{Rx}L_{i,j}} \tag{3}$$

where G^{Rx} is the gain of receiver.

Note that our focus is on ground cell-demand satisfaction, concentrating on beam scheduling rather than user scheduling (which is beyond the scope of this work.). Accordingly, we consider a super-user terminal that essentially represents all the users covered within a cell, meaning that all users' demands are aggregated. We denote the traffic demand of cell n as q_n in bps, and demands of all cells as $\mathbf{q} = [q_1, \ldots, q_{N_c}]^T$. In order to fully utilize the spectrum resource, all beams reuse the full spectrum bandwidth B. In other words, the interference among simultaneously active beams can not be ignored. We consider a typical beam-hopping satellite system within which a BH window is composed of M consecutive time slots (TSs) with time duration ΔT. The goal of designing an illumination pattern refers to determining which beams are activated during a specific TS within the time window. To comply with the satellite DVB-S2(X) standard, the M TSs corresponds to the M superframes [6]. Due to the mass limitation of satellite payload, the maximal number of active beams should not exceed the number of radio frequency chains, dented as $K_j, \forall j$.

B. Load Coupling

We consider the load coupling model [18] to address the demand constraint. We define $p_{i,j}$ and $\rho_{i,j}$ as the average transmit power and activation probability, respectively, of the beam radiated by the corresponding antenna of satellite j to serve the user located in virtual cell i. Whenever the beam is activated, the transmit power of the beam is $p_{i,j}$, while the activation probability describes the probability of the beam being activated during the time window.

The expected achievable rate of the user located in cell i, which is served by satellite $j, j \in \mathcal{A}_i$,

can be expressed as

$$R_{i,j} = \rho_{i,j} B \log_2(1 + \frac{p_{i,j}|H_{i,n_i^j}|^2}{\underbrace{\sum_{s\in\mathcal{N}_s\setminus\mathcal{A}_i}\sum_{b\in\mathcal{B}_s}\rho_{b,s}p_{b,s}|H_{i,n_b^s}|^2}_{\textit{Term 1}} + \underbrace{\sum_{s\in\mathcal{A}_i}\sum_{b\in\mathcal{B}_s\setminus\{i\}}\rho_{b,s}p_{b,s}|H_{i,n_b^s}|^2}_{\textit{Term 2}} + \sigma_T^2}) \tag{4}$$

where *Term 1* denotes the expected interference from beams radiated by unrelated satellites, i.e. $s \in \mathcal{N}_s \setminus \mathcal{A}_i$; and *Term 2* denotes the expected interference from beams radiated by related satellites, i.e. $s \in \mathcal{A}_i$; σ_T^2 represents as the thermal noise power, dented as $\sigma_T^2 = \tau T_{Rx} B$, where τ denotes the Boltzmann constant, and T_{Rx} denotes clear sky noise temperature of receiver [21].

Taking into account that some cells are served by multiple satellites, we define $d_{i,j}$ as the required load of the beam serving cell i, radiated by the corresponding antenna of satellite j. In other words, when cell i is served by a solo satellite, we define that the required load of the beam is no less than the required traffic demand of the cell, denoted as $d_{i,j} \geq q_i, \mathcal{A}_i = \{j\}$; While when cell i is served by multiple satellites, the sum of the required load of beams serving cell i is no less than the required traffic, denoted as $\sum_{j\in\mathcal{A}_i} d_{i,j} \geq q_i$.

Subsequently, the demand constraint can be expressed as

$$\begin{cases} R_{i,j} \quad \geq d_{i,j}, \ \forall i, j \\ \sum_{j\in\mathcal{A}_i} d_{i,j} \geq q_i, \quad \forall i \end{cases} \tag{5}$$

which means that the expected rate of the beam served by the satellite should be greater than that of the required load, and can be equivalently rewritten as

$$\rho_{i,j} \geq f_{i,j}(\boldsymbol{\rho}, \mathbf{p}, \mathbf{d}) \triangleq \frac{d_{i,j}}{B\log_2(1 + \frac{p_{i,j}|H_{i,n_i^j}|^2}{\sum_{s\in\mathcal{N}_s\setminus\mathcal{A}_i}\sum_{b\in\mathcal{B}_s}\rho_{b,s}p_{b,s}|H_{i,n_b^s}|^2 + \sum_{s\in\mathcal{A}_i}\sum_{b\in\mathcal{B}_s\setminus\{i\}}\rho_{b,s}p_{b,s}|H_{i,n_b^s}|^2 + \sigma_T^2})} \tag{6}$$

where we denote $\mathbf{p}_j$ as the power vector for all beams radiated by satellite j, and $\mathbf{p} = [\mathbf{p}_1^T, \ldots, \mathbf{p}_{N_s}^T]^T$ as the power vector for all beams of the systems. Similarly, we define $\boldsymbol{\rho}_j$ and $\boldsymbol{\rho}$ as the activation probability vectors, and $\mathbf{d}_j$ and $\mathbf{d}$ as the required load vectors, for the beams of satellite j and the systems, respectively. Finally, for the sake of simplicity, in the same manner, we define $\mathbf{f}_j(\boldsymbol{\rho}, \mathbf{p}, \mathbf{d})$ and $\mathbf{f}(\boldsymbol{\rho}, \mathbf{p}, \mathbf{d})$.

III. Ideal Pattern Design Problem Formulation

Our objective is to minimize the total consumed energy while satisfying both the limited onboard resource constraints and heterogeneous traffic demands, where the consumed energy is given by its total expectation. The formulated problem is expressed as

$$
\begin{aligned}
(\mathcal{P}_1): \quad & \underset{\boldsymbol{\rho},\mathbf{p},\mathbf{d}}{\texttt{minimize}} \quad \boldsymbol{\rho}^T\mathbf{p} && \text{(7a)}\\
& \texttt{subject to} \quad (C_1): \boldsymbol{\rho} \succeq \mathbf{f}(\boldsymbol{\rho},\mathbf{p},\mathbf{d}), && \text{(7b)}\\
& \qquad (C_2): \sum_{j\in\mathcal{A}_i} d_{i,j} \geq q_i, \quad \forall i && \text{(7c)}\\
& \qquad (C_3): \boldsymbol{\rho}_j^T\mathbf{1} \leq K_j, \quad \forall j && \text{(7d)}\\
& \qquad (C_4): \boldsymbol{\rho}_j^T\mathbf{p}_j \leq P_{\max}^j, \quad \forall j && \text{(7e)}\\
& \qquad (C_5): \sum_{j\in\mathcal{A}_i} \rho_{i,j} \leq 1, \quad \forall i && \text{(7f)}\\
& \qquad (C_6): 0 \leq \rho_{i,j} \leq 1, \quad \forall i,j && \text{(7g)}\\
& \qquad (C_7): p_{i,j} \geq 0, \quad \forall i,j && \text{(7h)}\\
& \qquad (C_8): d_{i,j} \geq 0, \quad \forall i,j && \text{(7i)}
\end{aligned}
$$

where C_1 companions with C_2 form the demand constraint; C_3 is the maximal number of active beams, which means that the expected number of active beams should not exceed the maximum number of simultaneously active beams; C_4 is to limit the maximum power of each satellite, which is given by the expected power of the satellite for all satellites; C_5 prevents the cell from being simultaneously served by different satellites to avoid the interference; $C_6 \sim C_8$ are the natural constraints of the three variables, respectively. Note that we omit the constant ΔT in the objective, as it does not affect the solution.

Once we have solved $\mathcal{P}_1$, we map the continuous activation probabilities into binary illumination pattern, determining whether to activate a beam at a particular TS, which we will explain in detail in Section V-D. In the next section, we will focus on solving $\mathcal{P}_1$.

IV. Ideal Pattern Design Solution Method

Problem $\mathcal{P}_1$ is generally hard to solve, as the coupling relationships among the three vector variables within the demand satisfaction constraint C_1 and the maximum power per satellite constraint C_4. Typical optimization techniques would work, such as the Difference of Convex

Functions method [22] and Successive Convex Approximation method [23] are applied to deal with the demand constraint in work [14]. However, applying these methods to our problem directly would bring performance loss because of the coupling relationships among the variables. In this section, we propose a novel method to solve the problem optimally. Firstly, based on the necessary condition to the optimum of problem $\mathcal{P}_1$, we obtain the one-to-one mapping relationship from the activation probability and required load to the power of beams. Subsequently, we reformulate the problem, which is proved to be convex. Lastly, we propose an iterative method to solve the convex problem optimally.

A. Equivalent Problem $\mathcal{P}_2$

One of the main difficulties in solving problem $\mathcal{P}_1$ comes from the demand satisfaction constraint C_1. The following theorem helps to address this issue and subsequently simplify the problem.

***Theorem** 1:* Given $\mathbf{q}, \mathbf{H}$, the solution to problem $\mathcal{P}_1$, if it exists, is the fixed point of the equations: $\boldsymbol{\rho} = \mathbf{f}(\boldsymbol{\rho}, \mathbf{p}, \mathbf{d})$.

Proof: The proof is given by [Chapter IX, Theorem 1]. Do not forget to change the reference. ■

Theorem 1 indicates the coupling relationships among the three variables, i.e. $\boldsymbol{\rho}, \mathbf{p}$, and $\mathbf{d}$, which is the necessary condition of the solution to problem $\mathcal{P}_1$. Based on the fixed point equations, we derive the following.

$$\rho_{i,j} = \frac{d_{i,j}}{B\log_2(1+\frac{p_{i,j}|H_{i,n_i^j}|^2}{\sum_{s\in\mathcal{N}_s\setminus\mathcal{A}_i}\sum_{b\in\mathcal{B}_s}\rho_{b,s}p_{b,s}|H_{i,n_b^s}|^2+\sum_{s\in\mathcal{A}_i}\sum_{b\in\mathcal{B}_s\setminus\{i\}}\rho_{b,s}p_{b,s}|H_{i,n_b^s}|^2+\sigma_T^2})} \tag{8a}$$

$$p_{i,j}\rho_{i,j} = g(\rho_{i,j}, d_{i,j})\Big(\sum_{s\in\mathcal{N}_s\setminus\mathcal{A}_i}\sum_{b\in\mathcal{B}_s}\rho_{b,s}p_{b,s}\frac{|H_{i,n_b^s}|^2}{|H_{i,n_i^j}|^2}+\sum_{s\in\mathcal{A}_i}\sum_{b\in\mathcal{B}_s\setminus\{i\}}\rho_{b,s}p_{b,s}\frac{|H_{i,n_b^s}|^2}{|H_{i,n_i^j}|^2}+\frac{\sigma_T^2}{|H_{i,n_i^j}|^2}\Big) \tag{8b}$$

where $g(\rho_{i,j}, d_{i,j}) = \rho_{i,j}(2^{\frac{d_{i,j}}{B\rho_{i,j}}}-1)$.

According to Eq. (8b), we denote $y_{i,j} = p_{i,j}\rho_{i,j}$ as the expected consumed energy of the i-th beam of satellite j, $\mathbf{y}_j$ for all beams of satellite j, and $\mathbf{y} = [\mathbf{y}_1^T, \ldots, \mathbf{y}_{N_s}^T]^T$ for all beams. Subsequently, $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \boldsymbol{\rho}^T\mathbf{p} = (\mathbf{I}-\mathbf{G}\mathbf{A})^{-1}\mathbf{G}\mathbf{b} \tag{9a}$$

$$\Rightarrow \mathbf{p} = \texttt{Diag}(\boldsymbol{\rho}^{-1})(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{Gb} \tag{9b}$$

where $\mathbf{I}$ is the identity matrix; $\mathbf{G} = \texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}))$ is the diagonal matrix with the diagonal elements given by the vector $\mathbf{g}(\boldsymbol{\rho}, \mathbf{d})$; Vectors $\mathbf{g}(\boldsymbol{\rho}, \mathbf{d})$ and $\mathbf{b}$ are vectorized the same way as $\mathbf{y}$, whose elements are given by $g(\rho_{i,j}, d_{i,j}), \forall i, j$ and $b_{i,j} = \frac{\sigma_T^2}{|H_{i,n_i^j}|^2}, \forall i, j$, respectively; We define the element of matrix $\mathbf{A}$ as

$$A_{(b,s),(i,j)} = \begin{cases} 0 & b = i, s \in \mathcal{A}_i \\ \dfrac{|H_{i,n_b^s}|^2}{|H_{i,n_i^j}|^2} & \text{otherwise} \end{cases} \tag{10}$$

where (b, s) and (i, j) are the index of the row and column of the matrix, respectively, vectorized the same way as $\mathbf{y}$.

Eq. (9b) maps the power as a function of activation probability and the required load of all beams. Based on this equation, the equivalent problem to $\mathcal{P}_1$ can be expressed as

$$(\mathcal{P}_2): \quad \underset{\boldsymbol{\rho}, \mathbf{d}}{\texttt{minimize}} \quad \mathrm{Tr}\{(\mathbf{I} - \mathbf{G}(\boldsymbol{\rho}, \mathbf{d})\mathbf{A})^{-1}\mathbf{G}(\boldsymbol{\rho}, \mathbf{d})\mathbf{B}\} \tag{11a}$$

$$\texttt{subject to} \quad (C_2): \sum_{j \in \mathcal{A}_i} d_{i,j} \geq q_i, \qquad \forall i \tag{11b}$$

$$(C_3): \boldsymbol{\rho}_j^T \mathbf{1} \leq K_j, \qquad \forall j \tag{11c}$$

$$(\tilde{C}_4): \mathrm{Tr}\{(\mathbf{I} - \mathbf{G}(\boldsymbol{\rho}, \mathbf{d})\mathbf{A})^{-1}\mathbf{G}(\boldsymbol{\rho}, \mathbf{d})\mathbf{E}_j\mathbf{B}\} \leq P_{\max}^j \qquad \forall j \tag{11d}$$

$$(C_5): \sum_{j \in \mathcal{A}_i} \rho_{i,j} \leq 1, \qquad \forall i \tag{11e}$$

$$(C_6): 0 \leq \rho_{i,j} \leq 1, \qquad \forall i, j \tag{11f}$$

$$(\tilde{C}_7): \sigma(\mathbf{G}(\boldsymbol{\rho}, \mathbf{d})\mathbf{A}) \leq 1 \tag{11g}$$

$$(C_8): d_{i,j} \geq 0, \qquad \forall i, j \tag{11h}$$

where $\mathbf{B} = \mathbf{b}\mathbf{1}^T$; $\mathbf{1}$ denotes the vector with all of its elements are 1; $\mathrm{Tr}\{\cdot\}$ denotes the trace of the matrix; For simplicity, we denote $\mathbf{G} = \mathbf{G}(\boldsymbol{\rho}, \mathbf{d})$ in the rest of paper.

Based on C_1, the equivalent problem $\mathcal{P}_2$ simplifies $\mathcal{P}_1$ by substituting the power with activation probability and required loads of beams. Consequently, related constraints should be modified: i) we omit the constraint C_1; ii) we modify C_4 into $\tilde{C}_4$, where $\mathbf{E}_j$ is the diagonal matrix, selecting the corresponding index of all beams of satellite j; iii) we re-express the natural constraint on power as $\tilde{C}_7$, where $\sigma(\cdot)$ denotes the spectral radius of the matrix. Initially, the natural constraint on power is $\mathbf{p} \succeq \mathbf{0}, \mathbf{p} \neq \mathbf{0}$, taking into account that it holds the mapping relationship among

variables, i.e. Eq.(9b), and $\mathbf{b} \succeq \mathbf{0}, \mathbf{A} \geq 0, \mathbf{G} \geq 0$, where $\geq$ denotes the pointwise greater than or equal to, the constraint of non-negative of power is equivalent to $\tilde{C}_7$ [24, Theorem A.51]. Note that we keep $\sigma(\mathbf{GA}) = 1$ within $\tilde{C}_7$, which does not affect the solution, as the objective would approach infinity at this point.

B. Convex Analysis of Problem $\mathcal{P}_2$

Problem $\mathcal{P}_2$ is generally difficult to solve, as the matrix multiplication involves an inverse matrix within the objective and the spectral radius of the matrix within the constraint. In this section, we first prove the convex of the objective. To deal with the constraint of the spectral radius of the matrix, we approximate it with a convex one.

Lemma 1: The objective of problem $\mathcal{P}_2$ is convex.

Proof: The proof is in Appendix A. ■

As the objective is convex, taking into account that constraint $\tilde{C}_4$ is the multiplication of the objective and the corresponding selective matrix, limiting the maximum power of each satellite, thus it is convex [25]. The rest of the constraints, except for constraint $\tilde{C}_7$, are linear, and consequently convex [25].

According to [26, Theorem 3.12], the spectral radius is bounded by all norms, i.e. $\sigma(\mathbf{X}) \leq \|\mathbf{X}\|$, where $\|\cdot\|$ is any norm of the matrix. We approximate the spectral radius of the matrix with its 2-norm, denoted as

$$\sigma(\mathbf{X}) \leq \|\mathbf{X}\|_2 = \sqrt{\sigma(\mathbf{X})\sigma((\mathbf{X})^T)} \leq t \tag{12}$$

which is equivalent to

$$t^2\mathbf{I} - \mathbf{X}\mathbf{X}^T \succeq 0 \tag{13}$$

which, according to Schur complement [27], is further equivalent to

$$\begin{pmatrix} t\mathbf{I} & \mathbf{X} \\ \mathbf{X}^T & t\mathbf{I} \end{pmatrix} \succeq 0 \tag{14}$$

which is convex [25].

Note that the condition on the first equality holds in Eq. (12) is that, in the considered context, the matrix $\mathbf{X}$ is symmetric, i.e. $\mathbf{X} = \mathbf{X}^T$. Taking into account that $\mathbf{G} = \texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}))$ is diagonal and symmetric, and $\mathbf{A}$ is close to symmetric, thus the 2-norm is a good estimation of the spectral radius constraint.

C. Iterative Solving Method to Problem $\mathcal{P}_2$

Although the objective has been proven to be convex, as far as the author's knowledge, there is no way to solve it directly. For solving problem $\mathcal{P}_2$, based on the Successive Convex Approximation (SCA) method [28], we propose an iterative method to address the resulting problem. Generally, the SCA estimates the value of a complicated function with its approximation in a sequence. Moreover, if the problem is convex, SCA would obtain the optimum [28].

Given $(\boldsymbol{\rho}_0, \mathbf{d}_0)$, the first order Taylor's series of the objective can be expressed as

$$L(\boldsymbol{\rho}, \mathbf{d}|\boldsymbol{\rho}_0, \mathbf{d}_0) \approx \mathrm{Tr}\{\underbrace{(\mathbf{I}-\mathbf{G}_0\mathbf{A})^{-1}}_{\mathbf{T}}\mathbf{G}\mathbf{B} + (\mathbf{I}-\mathbf{G}_0\mathbf{A})^{-1}\mathbf{A}(\mathbf{I}-\mathbf{G}_0\mathbf{A})^{-1}\mathbf{G}\mathbf{B}(\mathbf{G}-\mathbf{G}_0)\} \tag{15a}$$

$$= \mathbf{g}^T(\boldsymbol{\rho}, \mathbf{d})\underbrace{\mathtt{Diag}(\mathbf{1}^T\mathbf{T}\mathbf{A}\mathbf{T}\mathbf{B})}_{\boldsymbol{\Omega}}\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}) + \underbrace{[\mathbf{1}^T\mathbf{T}\mathbf{B} - \mathbf{1}^T\mathbf{T}\mathbf{A}\mathbf{T}\mathbf{G}_0\mathbf{B}]}_{\hat{\mathbf{b}}^T}\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}) \tag{15b}$$

$$= \mathbf{g}^T(\boldsymbol{\rho}, \mathbf{d})\boldsymbol{\Omega}\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}) + \mathbf{g}^T(\boldsymbol{\rho}, \mathbf{d})\hat{\mathbf{b}} \tag{15c}$$

where $\mathbf{G}_0 = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}_0, \mathbf{d}_0))$.

Subsequently, the constraint $\tilde{C}_4$, maximum power for each satellite, can be expressed as

$$\mathbf{g}^T(\boldsymbol{\rho}, \mathbf{d})\mathbf{E}_j\boldsymbol{\Omega}\mathbf{E}_j\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}) \le P_{\max}^j, \quad \forall j \tag{16}$$

Consequently, the sub-problem to be solved in each iteration can be expressed as

$$(\mathcal{P}_3): \quad \underset{\boldsymbol{\rho}, \mathbf{d}}{\mathtt{minimize}} \quad L(\boldsymbol{\rho}, \mathbf{d}|\boldsymbol{\rho}_0, \mathbf{d}_0) \tag{17a}$$

$$\mathtt{subject\ to} \quad (C_2): \sum_{j\in\mathcal{A}_i} d_{i,j} \ge q_i, \qquad \forall i \tag{17b}$$

$$(C_3): \boldsymbol{\rho}_j^T\mathbf{1} \le K_j, \qquad \forall j \tag{17c}$$

$$(\hat{C}_4): \mathbf{g}^T(\boldsymbol{\rho}, \mathbf{d})\mathbf{E}_j\boldsymbol{\Omega}\mathbf{E}_j\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}) \le P_{\max}^j, \qquad \forall j \tag{17d}$$

$$(C_5): \sum_{j\in\mathcal{A}_i} \rho_{i,j} \le 1, \qquad \forall i \tag{17e}$$

$$(C_6): 0 \le \rho_{i,j} \le 1, \qquad \forall i,j \tag{17f}$$

$$(\hat{C}_7): \begin{pmatrix} \mathbf{I} & \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d})) \\ \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d})) & \mathbf{I} \end{pmatrix} \succeq 0 \tag{17g}$$

$$(C_8): d_{i,j} \ge 0, \qquad \forall i,j \tag{17h}$$

which is convex and can be efficiently solved by the CVX [29] with advanced solvers, i.e. Mosek [30].

The iterative procedures to obtain the activation probability, power, and required loads of beams to match the required traffic demands while minimizing the consumed energy as detailed in **Algorithm** 1.

Algorithm 1 OPTIMIZATION OF IDEAL PATTERN DESIGN

Initialization $\mathbf{H}, \mathbf{A}, \mathbf{q}, \boldsymbol{\rho}_0, \mathbf{d}_0, K_j, P_{\max}^j, \forall j, k = 0$
repeat
  if $\sigma(\mathbf{G}_k\mathbf{A}) < 1$ **then**
    Set $k = k + 1$, Update $\boldsymbol{\Omega}, \hat{\mathbf{b}}$
    Solve $(\boldsymbol{\rho}_{k+1}, \mathbf{d}_{k+1}) = \arg \underset{\boldsymbol{\rho},\mathbf{d}}{\texttt{minimize}}\ L(\boldsymbol{\rho}, \mathbf{d}|\boldsymbol{\rho}_k, \mathbf{d}_k)$
  else
    Return: Infeasible
  end if
until Convergence

The Algorithm 1 is proposed to solve problem $\mathcal{P}_1$. During the re-expressions, the only approximation comes from the estimation of the spectral radius of the matrix, which is replaced by its 2-norm. Consequently, the algorithm obtains a high-quality solution.

V. PRACTICAL PATTERN DESIGN IMPLEMENTATION

In Section IV, we formulate and solve the ideal beam hopping pattern design problem, which is conditioned on the infinite time window duration. However, in practice, the time window is limited. In other words, the operator needs to convert the continuous activation probability into the discrete required load in the number of TSs. In this section, we address the problems to be met in practice. First, we analyze the impact of systematical parameters on the performance, which paves the way for the implementation of the pattern design problem. Taking into account that the problem formulation in Section IV is based on the Shannon Formula, it would bring systematical performance loss. Then we propose a method to compensate for this loss in advance. Lastly, a rounding algorithm is proposed to convert activation probability into discrete required loads of beams, followed by the multiplier penalty and majorization-minimization (MPMM) algorithm to design the pattern.

A. Systematical Parameters Analysis

As power is mapped as a function of activation probability and the required load of beams, the energy consumption problem is jointly optimized by the two variables. In this section, firstly given the required loads of all beams, we analyze the impact of K_j, the maximum number of simultaneously active beams of each satellite, on the feasibility of the problem. Subsequently, we analyze the joint impact of the two variables on the performance of the problem.

1) Given required loads of beams:

Lemma 2: Given the channel $\mathbf{H}$, whether the set $\mathcal{X} = \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) \leq 1, \mathbf{0} \preceq \boldsymbol{\rho} \preceq \mathbf{1}\}$ is empty is determined by loads of beams, i.e. $\mathbf{d}$.

Proof: The proof is given by [Chapter IX, Lemma 2]. ■

Lemma 2 indicates that once allocating too much loads to beams, the problem is infeasible. Here, if the set $\mathcal{X}$ is non-empty, we define the allocated loads to beams are reasonable, and otherwise irreasonable. The following *Lemma* and Theorem indicate the impact of parameters K on the feasibility and energy minimization performance of the problem, respectively.

***Theorem** 2:* Given channel $\mathbf{H}$ and allocating reasonable loads to beams, whether the set $\mathcal{Y} = \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) \leq 1, \mathbf{0} \preceq \boldsymbol{\rho} \preceq \mathbf{1}\} \cap \{\boldsymbol{\rho} | \boldsymbol{\rho}_j^T \mathbf{1} \leq K_j, \forall j\}$ is empty is determined by the parameters $K_j, \forall j$. If the parameters exceed the threshold, the set is non-empty.

Proof: The proof is given by [Chapter IX, Lemma 3] ■

Lemma 3: Given channel $\mathbf{H}$, allocating reasonable loads to beams and providing sufficient power for each satellite, once the parameters $K_j, \forall j$ exceed the thresholds, problem $\mathcal{P}_3$ is feasible and the optimum locates on the boundary, i.e. $\boldsymbol{\rho}_j^T \mathbf{1} = K_j, \forall j$. Moreover, the higher the parameters are, the less the total consumed energy will be.

Proof: The proof is given by [Chapter IX, Theorem 4] ■

2) Jointly optimize loads and activation probability of beams:

Lemma 4: The optimum, if it exists, is located on the boundary, i.e. $\boldsymbol{\rho}_j^T \mathbf{1} = K_j, \forall j$.

Proof: The prove is given in Appendix B. ■

***Theorem** 3:* Given the infeasible problem with fixed required loads of beams, the flexibility in allocating required loads of beams helps to convert the infeasible problem into being feasible.

Proof:

In the following, we illuminate the impact of required loads of beams on the feasibility of problem $\mathcal{P}_3$. Suppose that satellites s, j serve the same cell i, i.e. $s, j \in \mathcal{A}_i$. Taking into account

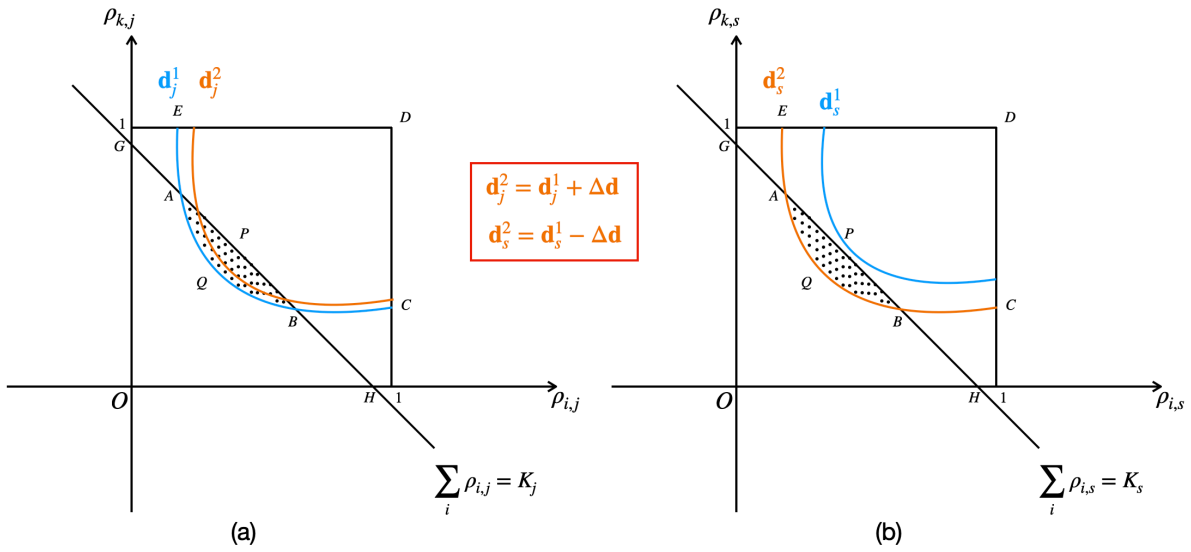


Fig. 2: The impact of flexibility in required load allocation on the feasibility of the problem. (a) increase the required loads for satellite j; (b) decrease the required loads for satellite s.

that, given $\mathbf{d}$, the set $\mathcal{X} = \{\boldsymbol{\rho} | \sigma(\mathbf{GA}) \leq 1, \mathbf{0} \preceq \boldsymbol{\rho} \preceq \mathbf{1}\}$ is convex [Chapter IX]. The feasible zones of the two satellites are shown in Fig. 2(a) and Fig. 2(b) respectively.

Specifically, in Fig. 2(a), the zone $EAQBCD$ indicates the slice of the set $\mathcal{X}$ cut by plane $\rho_{k,j}O\rho_{i,j}$, the zone $GOHBA$ indicates the slice of set $\mathcal{Y}_j$, and the dotted area $AQBP$ indicates the slice of the feasible zone. Initially, the required loads of beams are given by $\mathbf{d}_j^1$, consequently as shown in (a), the boundary of set $\mathcal{X}$ is given by the blue curve $EAQBC$. Once we increase the required loads of beams, i.e. $\mathbf{d}_j^2 = \mathbf{d}_j^1 + \Delta\mathbf{d}$, where $\Delta\mathbf{d} \succeq \mathbf{0}$, consequently it holds that $\sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}^2))) > \sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}^1)))$, which indicates that the set $\mathcal{X}_j$ will narrow down. The orange line describes the modified feasible zone.

Correspondingly, the feasible zone of satellite s is illuminated by Fig. 2(b). As we can see, initially given $\mathbf{d}_s^1$, the set $\mathcal{X}_s$ is empty, i.e. $\mathcal{X}_s \cap \mathcal{Y}_s = \emptyset$. However, once we reduce the required loads of beams, i.e. $\mathbf{d}_s^2 = \mathbf{d}_s^1 - \Delta\mathbf{d}$, the space of the feasible zone would increase, resulting in the non-empty set of the feasible zone, i.e. $\mathcal{X}_s \cap \mathcal{Y}_s \neq \emptyset$. ■

Our analysis omits the constraint of the maximum power of each satellite as the power of a beam is determined by activation probability and the required loads of beams. In other words, the allocation of power is dominated by the allocation of activation probability and the required

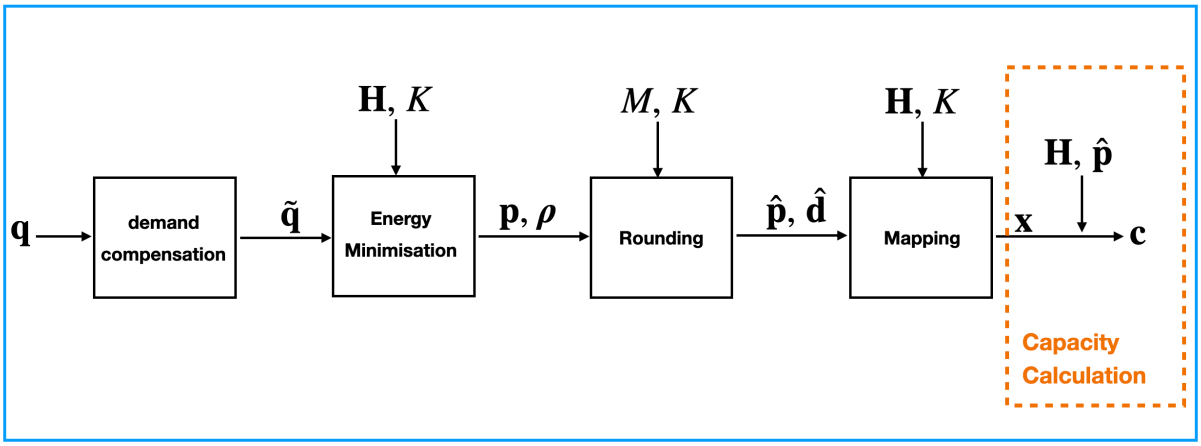


Fig. 3: The data flow of the proposed method.

loads of beams. Once the consumption of power of each satellite exceeds the preset threshold, it would require the satellite to adjust the required loads of beams serving the cell being able to be served by multiple satellites.

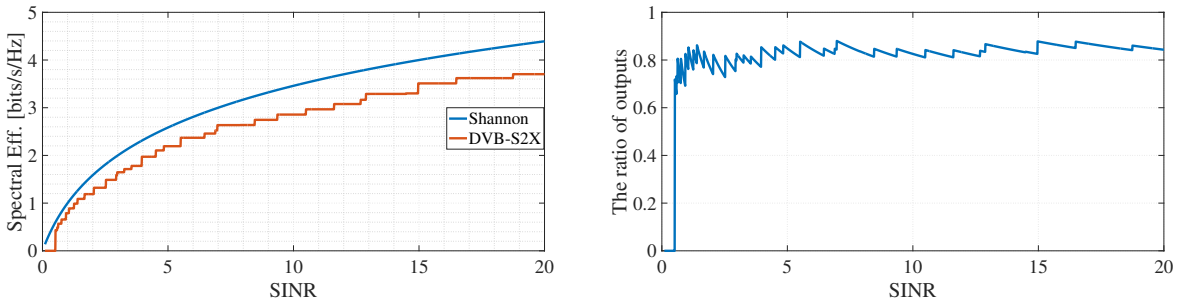
B. Demands Compensations

Fig. 3 illuminates the data flow of the proposed method. The goal of this work is to design the illumination pattern such that the provided capacity of cells $\mathbf{c}$ satisfies the required traffic $\mathbf{q}$ while minimizing the total consumed energy.

The formulation of $\mathcal{P}_1$ is based on the Shannon formula [31], where the relationship between output (capacity, y) and input (SINR, x) is expressed as $y = \log_2(1+x)$. However, in the practical satellite communication system, the signal should be modulated before transmission, following DVB-S2X standard [32]. Consequently, there exists a systematic performance loss between the provided capacity and the required traffic demands.

Fig. 4 indicates the gap in spectral efficiency between Shannon output and DVB-S2X. Specifically, Fig. 4(a) indicates the spectral output of the Shannon formula and DVB-S2X. To take a step further, Fig. 4(b) indicates the ratio of out of the DVB-S2X to that of the Shannon formula. As we can see, the ratio is less than 1 in whole the range of SINR, which indicates the systematic performance loss of the proposed method.

An intuitive way to compensate for this loss is demand compensation preprocessing, which requires more demands before putting into the energy minimization problem. The preprocessing



(a) (b)

Fig. 4: The comparison between Shannon output and DVB-S2X in terms of Spectral Efficiency. (a) the spectral output of Shannon formula and the DVB-S2X; (b) the ratio of the output of Shannon output to that of the DVB-S2X.

can be expressed as $\tilde{\mathbf{q}} = \mathbf{q}/\epsilon$, where ϵ is the demand compensation coefficient. It is shown in Fig. 4(b) that the ratio changes with the increase of SINR. We propose two methods to find an appropriate coefficient.

1) Adaptive demand compensation coefficient determination: With demand compensation preprocessing, the modified provided capacity constraint can be expressed as

$$\rho_{i,j}\epsilon B \log_2(1 + \texttt{SINR}_{i,j}) = d_{i,j}, \forall i, j \tag{18}$$

where $\texttt{SINR}_{i,j}$ is the expected SINR of beam i radiated by satellite j; ϵ denotes the demand compensation coefficient, which is related to the SINR of the beam. The equality is based on that, to minimize energy consumption, the algorithm is expected to provide capacity that is rightly equal to the required demand of the beam for all beams.

Suppose the expected SINR for all beams are equal, i.e. $\texttt{SINR} = \texttt{SINR}_{i,j}, \forall j \in \mathcal{N}_s, i \in \mathcal{B}_j$, consequently it holds that

$$\sum_{j\in\mathcal{N}_s}\sum_{i\in\mathcal{B}_j} \rho_{i,j}\epsilon B \log_2(1 + \texttt{SINR}) = \sum_{j\in\mathcal{N}_s}\sum_{i\in\mathcal{B}_j} d_{i,j} = \sum_{i\in\mathcal{N}_c} q_i \tag{19a}$$

$$\epsilon B \log_2(1 + \texttt{SINR}) = \frac{\sum_{i\in\mathcal{N}_c} q_i}{\sum_{j\in\mathcal{N}_s}\sum_{i\in\mathcal{B}_j} \rho_{i,j}} = \frac{\sum_{i\in\mathcal{N}_c} q_i}{\sum_{j\in\mathcal{N}_s} K_j} \tag{19b}$$

where the second equality of Eq. (19b) is based on the *Lemma* 4 that the optimum, if it exists, is located on the boundary.

Comparing the coefficient $\frac{\sum_{i\in\mathcal{N}_b} q_i}{\sum_{j\in\mathcal{N}_s} K_j}$ with the output of spectral efficiency in DVB-S2X standard [33, TABLE 1], one can determine the corresponding SINR, and subsequently the coefficient ϵ.

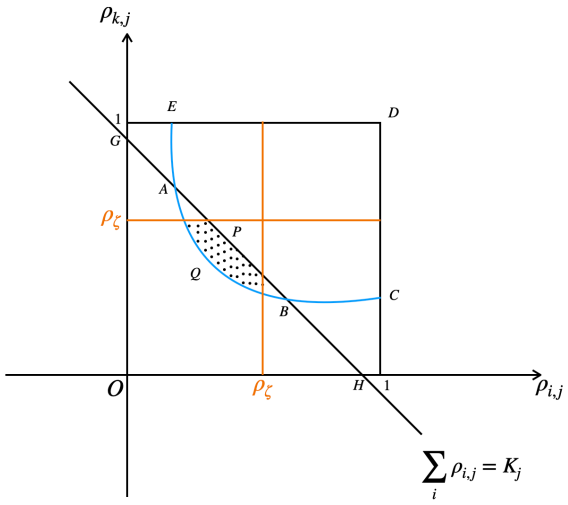


Fig. 5: The impact of the SINR requirement on the feasibility of problem $\mathcal{P}_3$.

2) SINR *requirement determination:* The operator may have requirements on modulation and coding schemes to increase the spectral efficiency, which consequently puts the requirement on the SINR of signals, i.e. $\texttt{SINR} \geq \zeta$. The corresponding demand coefficient is determined by checking Fig. 4(b) under the condition $\texttt{SINR} = \zeta$, which is denoted as ϵ_ζ. Consequently, the requirement on SINR is further equivalent to

$$\rho_{i,j} \leq \rho_\zeta \triangleq \frac{\sum_{i\in\mathcal{N}_c} q_i}{\epsilon_\zeta \log_2(1+\zeta) \sum_{j\in\mathcal{N}_s} K_j}, \forall i, j \tag{20}$$

Fig. 5 illuminates the impact of SINR requirement on the feasibility of problem $\mathcal{P}_3$, where the orange line indicates the constraint on SINR. As we can see, the orange line narrows down the feasible zone of the problem, which may result in infeasibility if the requirement is too high.

The constraint Eq. (20), denoted as C_9, is linear and thus convex. One can modify the problem $\mathcal{P}_3$ by adding the constraint C_9.

C. Approximation from Activation Probability to Demand

In this section, we propose two rounding schemes to convert the continuous activation probability into discrete demands in the number of TS, based on different demands compensation methods.

The proposed adaptive demand compensation scheme in Section V-B1 is based on the condition that the optimum is located on the boundary, i.e. $\boldsymbol{\rho}_j^T \mathbf{1} = K_j, \forall j$. Consequently, once employing

the adaptive demand compensation scheme, the goal of the approximation scheme is to find a discrete point, denoted as $\hat{\mathbf{d}}$, that is not only close to $M\boldsymbol{\rho}$ but also satisfies the equations $\hat{\mathbf{d}}_j^T \mathbf{1} = MK_j, \forall j$. The straightforward method is

- find the gap (residue) between the continuous optimum and its low bound, denoted as $M\boldsymbol{\rho} - \lfloor M\boldsymbol{\rho} \rfloor$.
- distribute the remaining numbers, i.e. $M \sum_{j \in \mathcal{N}_s} K_j - (\lfloor M\boldsymbol{\rho} \rfloor)^T \mathbf{1}$, to demands with the same number of beams sorted by the residue, plus one for each.

While the SINR requirement demand compensation scheme limits the signals' minimal expected SINR, the optimum solution may not be located on the boundary. Thus we compare the residue with a threshold to determine whether plus one to demand of the beam. Note that some beams serve the same cell, to guarantee the demand-matching performance for the cell, we reallocate the accumulated residue of these beams. First, we sort the residues of these beams and then distribute the accumulated residues to the beams in the order of the sorted list. Once the remaining residue is greater than 1, the demand of the beam plus one, then the residue is reduced by 1 and considering the next beam.

Recall that power is a function of activation probability and the required load of beams, i.e. Eq. (9b). Consequently, once the activation probability of beams changes, the power should be modified.

D. Discrete Illumination Pattern Design

Given discrete demands, power, and the required loads of beams, one has to design the illumination pattern, determining the status and assigning the corresponding power of a beam. A heuristic method is to assign the active beams randomly; however, this approach may not guarantee the demand-matching performance. In this context, we propose to penalize any two beams simultaneously being activated and design the illumination pattern by minimizing the total penalty while satisfying all the constraints.

Specifically, we denote $\hat{P}_{m,n}$ as the relative interference from beam m to beam n, which is defined by

$$\hat{P}_{m,n} = \frac{\iint_{\mathbf{S}_n} |H_{k,m}|^2 p_m dx_k \, dy_k}{\iint_{\mathbf{S}_n} |H_{k,n}|^2 p_n dx_k \, dy_k} \tag{21}$$

where (x_k, y_k) are the longitude and latitude of user k located in cell n, respectively; and $\mathbf{S}_n$ denotes the coverage of beam n which is defined by the -3 dB contour from the maximum gain.

TABLE II: LEO Constellation Systems Parameters

Height of satellite	h= 600 km
Downlink carrier frequency	19.5 GHz
User link bandwidth,	B=500 MHz
Longitude range of satellites	$[-0.11^\circ E,\ 0.11^\circ E]$
Latitude range of satellites	$[-0.13^\circ N,\ 0.13^\circ N]$
Antenna circular aperture	a = 1
Antenna Maximum Gain	$G_0 = 30$ dBi
Number of beams per satellite	19
Number of virtual cells	$N_c = 103$
Number of beams	$N_b = 113$
Number of cells served by two satellites	18
Number of cells served by three satellites	6
Temperature	50 K
Number of TSs	M=20

The pattern design problem can be expressed as

$$(\mathcal{P}_4): \underset{\mathbf{x}}{\texttt{minimize}} \quad \mathbf{x}^T\mathbf{P}\mathbf{x} \tag{22a}$$

$$\texttt{subject to} \quad (S_1): \mathbf{x}^T\mathbf{e}_j^t \le K_j, \quad \forall j,t \tag{22b}$$

$$(S_2): \mathbf{x}^T\mathbf{e}_b = \hat{d}_b, \quad \forall b \tag{22c}$$

$$(S_3): \mathbf{x}^T\mathbf{p}_j^t \le P_{\max}^j, \quad \forall j,t \tag{22d}$$

$$(S_4): \mathbf{x} \in \{0,\ 1\}^{MN} \tag{22e}$$

where $\mathbf{P} = \texttt{Diag}([\underbrace{\hat{\mathbf{P}}, \ldots, \hat{\mathbf{P}}}_{M}])$ is the diagonal block matrix; $\mathbf{e}_j^t, \forall j,t$ denotes the binary vector to select all beams that belong to satellite j at t-th TS; $\mathbf{e}_b, \forall b$ denotes the binary vector to select the b-th beam for all the whole time window; $\hat{d}_b$ denotes the discrete demands in number of TS of the b-th beam; $\mathbf{p}_j^t = \mathbf{e}_j^t \cdot \mathbf{p}_M$, where $\cdot$ denotes the pointwise multiplication and $\mathbf{p}_M = [\underbrace{\mathbf{p}^T, \cdots, \mathbf{p}^T}_{M}]^T$.

Problem $\mathcal{P}_4$ is a binary quadratic programming (BQP) problem, which can be efficiently addressed by the multiplier penalty and majorization-minimization (MPMM) method proposed in [12].

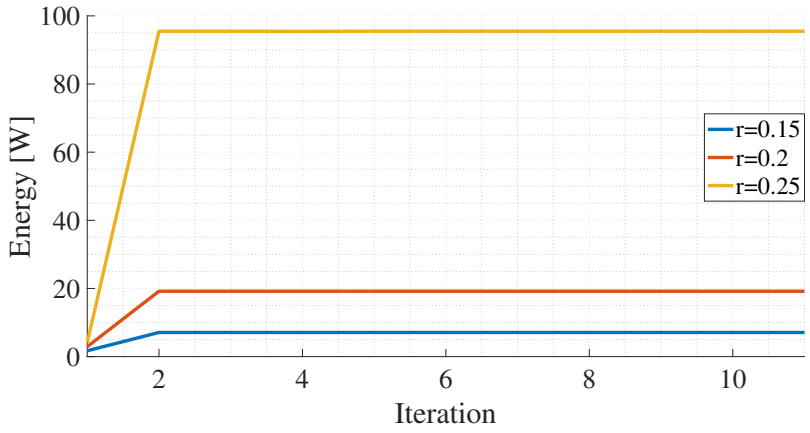


Fig. 6: The convergence of the proposed ideal pattern design problem. The energy is given by $\boldsymbol{\rho}^T \mathbf{p}$.

VI. Numerical Simulations

A. Simulation Setup

The parameter settings of the LEO constellation systems are based on [20, 34] and summarized in TABLE II. The traffic demands of all beams are randomly generated within the range of [400 Mbps, 1600Mbps]. Specifically, the demand of beams is given by $400 \cdot r \leq q_i \leq 1600 \cdot r, \forall i$, where r is the demand density parameter and selected from $\{0.15, 0.2, 0.25\}$. For each selected demand density parameter, 50 instances are generated for simulations. The maximum number of simultaneously active beams per satellite is the same and is $K_j = 5, \forall j$. The maximum power of satellites is set as $[20,\ 20,\ 24,\ 24,\ 20,\ 18,\ 16]\ W$. We do beam scheduling in this work, and we suppose one super user is located in the beam center. The baseline method jointly optimizes power and activation probability to minimize energy consumption. In this method, the traffic demands of a cell served by multiple satellites are equally allocated to the beams that can serve the cell. In other words, compared with the proposed method, the baseline method has no flexibility in traffic allocation among cells served by multiple satellites. As far as the authors know, the baseline method is state-of-the-art. Unless stated otherwise, all parameter settings are applied to simulations.

B. Performance of the Ideal Pattern Design Problem

TABLE III: Influence of ρ_ζ on energy consumption

	$\rho_\zeta = 0$	$\rho_\zeta = 3$	$\rho_\zeta = 4$	$\rho_\zeta = 5$
$r = 0.15$	3.12	4.53	5.54	7.09
$r = 0.20$	12.49	8.93	12.20	19.19
$r = 0.25$	32.12	19.39	26.86	95.50

[1] $\rho_\zeta = 0$ represents the adaptive demand compensation.

TABLE IV: Influence of ρ_ζ on spectral radius of the matrix

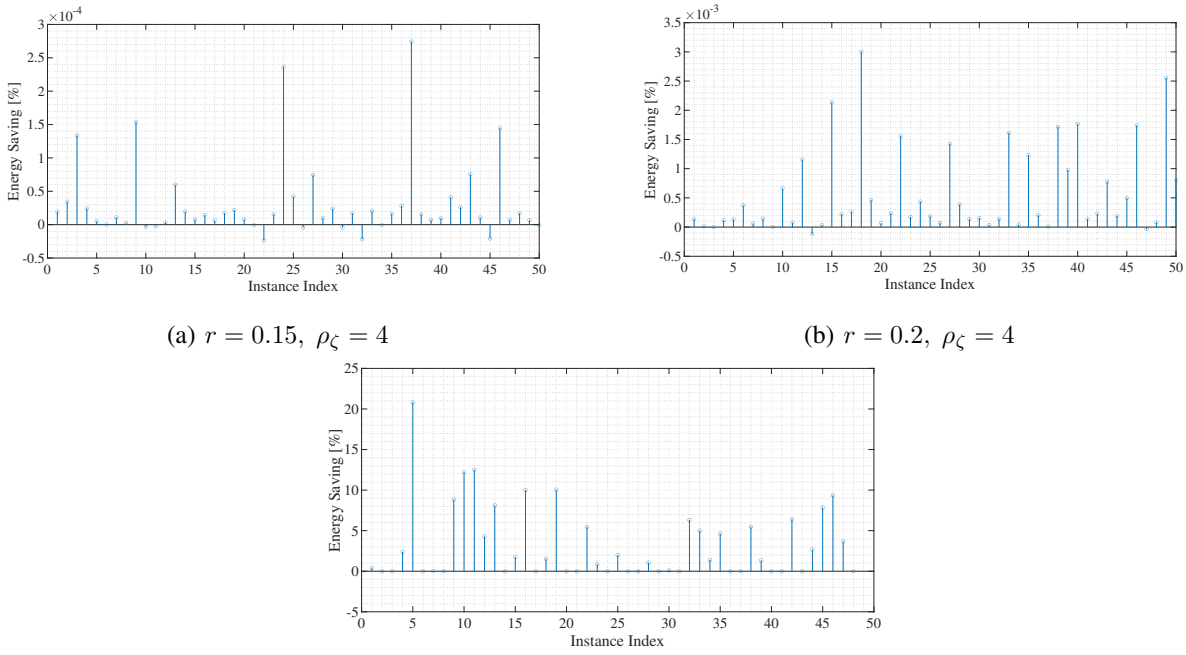
	$\rho_\zeta = 0$	$\rho_\zeta = 3$	$\rho_\zeta = 4$	$\rho_\zeta = 5$
$r = 0.15$	0.36	0.47	0.52	0.58
$r = 0.20$	0.71	0.64	0.71	0.80
$r = 0.25$	0.86	0.79	0.85	0.96

[1] $\rho_\zeta = 0$ represents the adaptive demand compensation.

1) Convergence: Fig. 6 illuminates the convergence of the Algorithm 1 at an instance of three demand densities respectively. The initial point is given by $\boldsymbol{\rho}_0 = \mathbf{1}$, which is the largest value of $\boldsymbol{\rho}$, consequently, it results in the minimal value of energy consumption. Moreover, it takes two to three times iterations for the algorithm to converge. Additionally, it is evident that higher demand density results in greater energy consumption. Especially, the total consumed energy would be below $10W$ when the demand density is $r = 0.1$. However, the system consumes around $100W$ energy when demand density increases to $r = 0.25$.

2) Demand Compensation Coefficient ρ_ζ: Moreover, we compare the influence of the choice of ρ_ζ on energy consumption and the corresponding spectral radius of the matrix for three different demand densities, each per instance, using Table III and Table IV, respectively. Specifically, we select 4 kinds of demand compensation methods, which include adaptive demand compensation, $\rho_\zeta = 0$, and three `SINR` requirement determination methods, i.e. $\rho_\zeta = 3, 4, 5$.

As we can observe, for any chosen ρ_ζ, the higher the demand density, the greater the consumed energy. Interestingly, the spectral radius of the matrix demonstrates a positive coordination with energy consumption. Moreover, the performance of the `SINR` requirement determination method is as expected. Higher `SINR` requirements result in a higher spectral radius of the matrix and consume greater energy. Although the adaptive demand compensation method has no limit on `SINR`, it consumes greater energy than methods with a low threshold of `SINR` at the cases of higher demand densities, i.e. $\rho_\zeta = 3,\ 4$ and $r = 0.2,\ 0.25$. This is because the method determines the compensation coefficient depends on the expected `SINR` of signals, higher demand density results in lower expected `SINR`, and subsequently requires more compensation demands.



(a) $r = 0.15,\ \rho_\zeta = 4$

(b) $r = 0.2,\ \rho_\zeta = 4$

(c) $r = 0.25,\ \rho_\zeta = 4$

Fig. 7: The energy consumption comparison with the baseline.

3) Energy consumption comparison: To compare the energy consumption performance, the merit energy saving is defined as

$$\text{Energy Saving} = \frac{\boldsymbol{\rho}_0^T \mathbf{p}_0 - \boldsymbol{\rho}_1^T \mathbf{p}_1}{\boldsymbol{\rho}_0^T \mathbf{p}_0} \cdot 100\% \tag{23}$$

where $(\boldsymbol{\rho}_0, \mathbf{p}_0)$ represents the solution to the problem with the baseline method and $(\boldsymbol{\rho}_1, \mathbf{p}_1)$ the proposed method.

Fig. 7 compares the energy consumption of the proposed method with the baseline at three different demand densities for 50 instances each, where the demand is decompensated with $\rho_\zeta = 4$. As observed, when demand density is not high, i.e. $r = 0.15, 0.2$, the advantage of the proposed method is not obvious. Both methods consume almost the same energy. This is because the low traffic demand of cells renders the extra flexibility provided by the proposed method in the domain of required loads of beams useless. However, when the demand density is high, i.e. $r = 0.25$, the proposed method shows great priority over the baseline. Generally, the proposed method would save about 10% of energy consumption, compared with the baseline. In some cases, this would be up to 20%. Especially, the figure shows 48 out of 50 instances for comparison. The rest of the 2 instances are infeasible with the baseline method, while they can

be solved with the proposed method. These findings validate the Theorem 3 that the flexibility in allocating required loads of beams helps to convert the infeasible problem into being feasible.

C. Pattern Design Problem

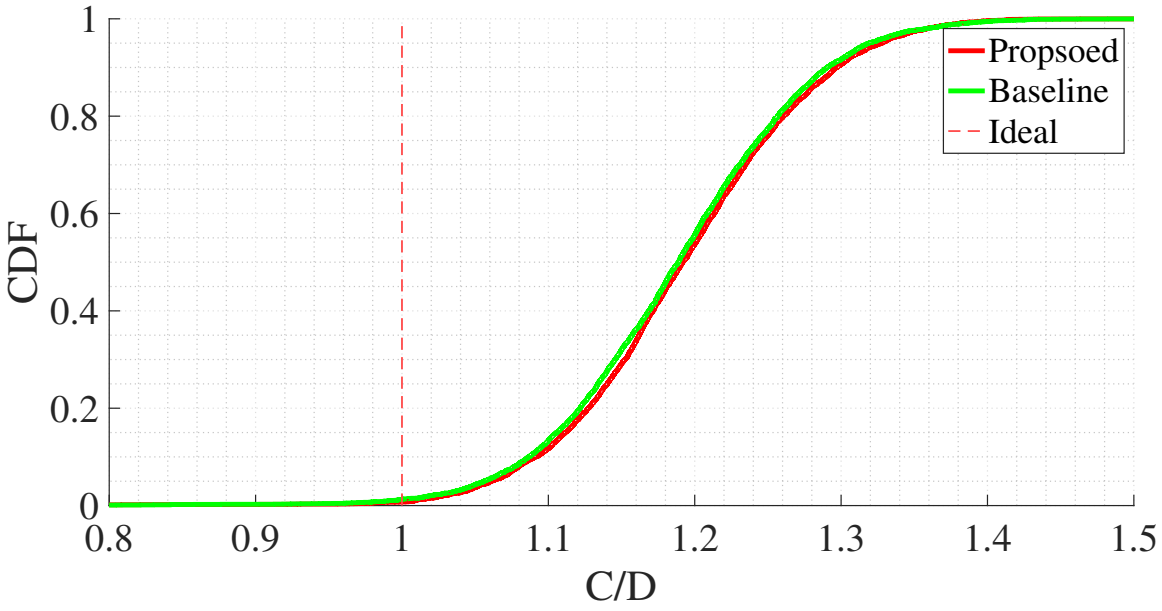


Fig. 8: Performance comparison with the baseline in terms of demand-matching.

We define the ratio of the provided capacity of the cell to that of the required traffic demand as the performance metric of demand matching, denoted as C/D. The performance is illustrated with the cumulative distribution function (CDF) of the ratio of all beams of all instances. Fig. 8 compares the demand-matching performance of the proposed method with the baseline. As we can see, almost all beams are satisfied. Note that both the baseline and the proposed method design the illumination pattern with the MPMM method, which results in a similar performance. However, as the proposed method consumes less energy for all instances, the proposed method outperforms the baseline.

VII. Conclusion

In this paper, we propose a method to jointly optimize power, required loads, and time slot allocation to minimize total energy consumption while satisfying the traffic demands of cells in LEO satellite constellation systems. There are two difficulties within the problem: i)satisfaction

of uneven traffic demands of beams, ii)and binary illumination pattern design. We propose to address it by decomposing it into two problems. First, we employ the load coupling model to simplify the demand satisfaction constraint, where parameters are represented with their expectations. Subsequently, the problem is reformulated into an inverse matrix optimization form, which is approximated by a convex problem. We propose an iterative method to obtain a high-quality solution to the original problem. With the power, activation probability, and the required loads of beams obtained from the first problem, we employ the MPMM method to design the practical illumination pattern. Moreover, we propose demand compensation methods to eliminate systematic performance loss caused by the gap between the Shannon formula and the DVB-S2X standard. Finally, simulations validate our theoretical findings and prove that the proposed method outperforms the baseline.

APPENDIX A

Lemma 1: The objective of problem $\mathcal{P}_2$ is convex.

Proof: Given the objective, i.e. $L((\mathbf{d}, \boldsymbol{\rho}))$, we have the following derivations, where the differentiation of the objective with respect to $\mathbf{G}$ is expressed as $\mathrm{d}L(\mathbf{d}, \boldsymbol{\rho})$.

$$L(\boldsymbol{\rho}, \mathbf{d}) = \mathrm{Tr}\{(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{GB}\} \tag{24a}$$

$$\mathrm{d}L(\boldsymbol{\rho}, \mathbf{d}) = \mathrm{Tr}\{\underbrace{(\mathbf{A}(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{GB} + \mathbf{B})(\mathbf{I} - \mathbf{GA})^{-1}}_{\mathbf{Z}}\mathrm{d}\mathbf{G}\} \tag{24b}$$

$$\mathrm{d}\mathbf{Z} = \underbrace{(\mathbf{ZA}(\mathbf{I} - \mathbf{GA})^{-1} + \mathbf{A}(\mathbf{I} - \mathbf{GA})^{-1}\mathbf{Z})}_{\mathbf{Y}}\mathrm{d}\mathbf{G} \tag{24c}$$

Note that $\mathbf{G} = \mathtt{Diag}(\mathbf{g}(\boldsymbol{\rho}, \mathbf{d}))$ is a diagonal matrix, accompanied with the above derivations, the second-order derivatives of the objective with respect to $\boldsymbol{\rho}$ and $\mathbf{d}$ can be expressed as

$$\frac{\partial^2 L}{\partial \rho_k^2} = Y_{k,k}\frac{\partial \mathbf{g}}{\partial \rho_k}\frac{\partial \mathbf{g}}{\partial \rho_k} + Z_{k,k}\frac{\partial^2 \mathbf{g}}{\partial \rho_k^2} = Y_{k,k}g^1_{\rho_k}g^1_{\rho_k} + Z_{k,k}g^2_{\rho_k} \tag{25a}$$

$$\frac{\partial^2 L}{\partial d_k^2} = Y_{k,k}\frac{\partial \mathbf{g}}{\partial d_k}\frac{\partial \mathbf{g}}{\partial d_k} + Z_{k,k}\frac{\partial^2 \mathbf{g}}{\partial d_k^2} = Y_{k,k}g^1_{d_k}g^1_{d_k} + Z_{k,k}g^2_{d_k} \tag{25b}$$

$$\frac{\partial^2 L}{\partial \rho_k \partial d_k} = Y_{k,k}\frac{\partial \mathbf{g}}{\partial \rho_k}\frac{\partial \mathbf{g}}{\partial d_k} + Z_{k,k}\frac{\partial^2 \mathbf{g}}{\partial \rho_k \partial d_k} = Y_{k,k}g^1_{\rho_k}g^1_{d_k} + Z_{k,k}g^2_{\rho_k d_k} \tag{25c}$$

$$\frac{\partial^2 L}{\partial d_k \partial \rho_k} = Y_{i,i}\frac{\partial \mathbf{g}}{\partial d_k}\frac{\partial \mathbf{g}}{\partial \rho_k} + Z_{k,k}\frac{\partial^2 \mathbf{g}}{\partial d_k \partial \rho_k} = Y_{k,k}g^1_{d_k}g^1_{\rho_k} + Z_{k,k}g^2_{d_k \rho_k} \tag{25d}$$

$$\Rightarrow \frac{\partial^2 L}{\partial d_k \partial \rho_k} = \frac{\partial^2 L}{\partial \rho_k \partial d_k} \tag{25e}$$

where ρ_k is the k-th element of $\boldsymbol{\rho}$ and d_k is the k-th element of $\mathbf{d}$; $\mathbf{g} = \mathbf{g}(\boldsymbol{\rho}, \mathbf{d})$, and $\frac{\partial \mathbf{g}}{\partial \rho_k} = \frac{\partial g(\rho_k, d_k)}{\partial \rho_k}$ denoted as $g^1_{\rho_k}$. In the same way, we denote $g^2_{\rho_k d_k}$ as $\frac{\partial^2 \mathbf{g}}{\partial \rho_k \partial d_k}$.

The Hessian matrix of the objective, which is a diagonal block, can be given by

$$\nabla L = \mathtt{Diag}(\nabla L_1, \ldots, \nabla L_k, \ldots \nabla L_{N_b}) \tag{26a}$$

$$\nabla L_k = \begin{bmatrix} \frac{\partial^2 L}{\partial \rho_k^2} & \frac{\partial^2 L}{\partial \rho_k \partial d_k} \\ \frac{\partial^2 L}{\partial d_k \partial \rho_k} & \frac{\partial^2 L}{\partial d_k^2} \end{bmatrix} \tag{26b}$$

where $N_b = \sum_{j \in \mathcal{N}_s} |\mathcal{B}_j|$ is the total number of beams.

Taking into account that the eigenvalue of the diagonal block matrix is the collection of the eigenvalue of each sub-matrix and the eigenvalues of a matrix are the solutions to the characteristic equation, i.e. $\mathtt{det}(\nabla L_k - \lambda \mathbf{I}) = 0$, [27]. In the following, we will prove the positive of the eigenvalues of any of the sub-matrix.

Note that $g^2_{\rho_k} = g^2_{d_k} (\frac{d_k}{\rho_k})^2$, $g^2_{\rho_k d_k} = -\frac{d_k}{\rho_k} g^2_{d_k}$, the characteristic equation of a sub-matrix can be expressed as

$$\mathtt{det}(\nabla L_k - \lambda \mathbf{I}) = \lambda^2 - \lambda(Y_{k,k} g^1_{\rho_k} g^1_{\rho_k} + Z_{k,k} g^2_{\rho_k} + Y_{k,k} g^1_{d_k} g^1_{d_k} + Z_{k,k} g^2_{d_k}) + Y_{k,k} Z_{k,k} g^2_{d_k} (g^1_{\rho_k} + \frac{d_k}{\rho_k} g^1_{d_k})^2 \tag{27}$$

Consequently, the discriminant of the quadratic equation in one unknown, i.e. $\mathtt{det}(\nabla L_k - \lambda \mathbf{I}) = 0$, can be expressed as

$$\Delta = (Y_{k,k} g^1_{\rho_k} g^1_{\rho_k} + Z_{k,k} f^2_{\rho_k} + Y_{k,k} g^1_{d_k} g^1_{d_k} + Z_{k,k} g^2_{d_k})^2 - 4 Y_{k,k} Z_{k,k} g^2_{d_k} (g^1_{\rho_k} + \frac{d_k}{\rho_k} g^1_{d_k})^2 \tag{28a}$$

$$= (Y_{k,k} (g^1_{\rho_k} g^1_{\rho_k} + g^1_{d_k} g^1_{d_k}) + Z_{k,k} (g^2_{\rho_k} + g^2_{d_k}))^2 - 4 Y_{k,k} Z_{k,k} g^2_{d_k} (g^1_{\rho_k} + \frac{d_k}{\rho_k} g^1_{d_k})^2 \tag{28b}$$

$$> 2 Y_{k,k} Z_{k,k} (g^1_{\rho_k} g^1_{\rho_k} + g^1_{d_k} g^1_{d_k})(g^2_{\rho_k} + g^2_{d_k}) - 4 Y_{k,k} Z_{k,k} g^2_{d_k} (g^1_{\rho_k} + \frac{d_k}{\rho_k} g^1_{d_k})^2 \tag{28c}$$

$$= 2 Y_{k,k} Z_{k,k} g^2_{d_k} \underbrace{((\frac{d_k}{\rho_k} g^1_{\rho_k} - g^1_{d_k})^2 - (g^1_{\rho_k} + \frac{d_k}{\rho_k} g^1_{d_k})^2)}_{h(t)} \tag{28d}$$

We define function $h(t)$ as shown in Eq.(28d), where $t = \frac{d_k}{\rho_k}$. In the following, we will prove the positive of the function.

According to Eq. (25), the function $h(t)$ can be expressed as

$$h(t) = (te^t - t^2 e^t - e^t - t)^2 - (e^t - 1)^2 \tag{29a}$$

$$= \underbrace{(t^2e^t - (t-2)e^t + t + 1)}_{h_1(t)} \underbrace{(t^2e^t - te^t + t + 1)}_{h_2(t)} \tag{29b}$$

Taking into account that $h_1(t)' = t^2e^t + (t+1)e^t + 1 > 0$, consequently, it holds that $h_1(t) \geq h_1(0) = 1$. Moreover, as $h_2''(t) = (t^2 + 3t)e^t \geq 0$, it holds that $h_2'(t) \geq h_2'(0) = 0$, and consequently, we have $h_2(t) > 0$. Finally, we have $h(t) > 0$, resulting in $\Delta > 0$.

Taking into account that it holds that $\lambda_1 + \lambda_2 > 0$, $\lambda_1\lambda_2 > 0$, and $\Delta > 0$, consequently, the characteristic equation has two positive solutions, which therefore completes the proof. ■

APPENDIX B

Lemma 2: The optimum, if it exists, is located on the boundary, i.e. $\boldsymbol{\rho}_j^T\mathbf{1} = K_j, \forall j$.

Proof: According to [26, Theorem 2.10], the spectral radius of a matrix can be defined as

$$\sigma(\mathbf{GA}) = \underset{\mathbf{x}}{\texttt{maximize}}\{\underset{x_k>0}{\texttt{minimize}}\frac{g(\rho_k, d_k)\mathbf{a}_k^T\mathbf{x}}{x_k}\} \tag{30}$$

where $\mathbf{A}^T = [\mathbf{a}_1, \cdots, \mathbf{a}_{N_b}]$.

According to the definition, taking into account that function $g(\rho, d)$ is monotonically decreasing with respect to ρ, consequently it holds that $\sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}^1, \mathbf{d})) \geq \sigma(\texttt{Diag}(\mathbf{g}(\boldsymbol{\rho}^2, \mathbf{d}))$ if there is $\boldsymbol{\rho}^1 \preceq \boldsymbol{\rho}^2$. Moreover, considering that the objective function is monotonically decreasing with respect to $\boldsymbol{\rho}$, each satellite would fully use the maximum number of simultaneously active beams, i.e. $K_j, \forall j$ to decrease the energy consumption. Consequently, the solution to problem $\mathcal{P}_3$, if it exists, is located at the boundary, where $\mathcal{Y}_j = \{\boldsymbol{\rho}_j | \boldsymbol{\rho}_j^T\mathbf{1} = K_j\}$. ■

REFERENCES

[1] M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad, "Satellite communications supporting internet of remote things," *IEEE Internet of Things Journal*, vol. 3, no. 1, pp. 113–123, 2015.

[2] 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 *et al.*, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 70–109, 2020.

[3] O. Kodheli, A. Guidotti, and A. Vanelli-Coralli, "Integration of satellites in 5g through leo constellations," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[4] S. Kisseleff, E. Lagunas, T. S. Abdu, S. Chatzinotas, and B. Ottersten, "Radio resource management techniques for multibeam satellite systems," *IEEE Communications Letters*, vol. 25, no. 8, pp. 2448–2452, 2020.

[5] A. Freedman, D. Rainish, and Y. Gat, "Beam hopping how to make it possible," in *Proc. of Ka and Broadband Communication Conference, Bologna, Italy*, Oct. 2015.

[6] C. Rohde, R. Wansch, G. Mocker, A. Trutschel-Stefan, L. Roux, E. Feltrin, H. Fenech, and N. Alagha, "Beam-hopping over-the-air tests using dvb-s2x super-framing," 2018.

[7] T. Posielek, "An energy management approach for satellites," in *69th International Astronautical Congress, Bremen, Germany*, Oct. 2018, pp. 1–5.

[8] Y. Yang, M. Xu, D. Wang, and Y. Wang, "Towards energy-efficient routing in satellite networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3869–3886, 2016.

[9] U. Park, H. W. Kim, D. S. Oh, and B. J. Ku, "Flexible bandwidth allocation scheme based on traffic demands and channel conditions for multi-beam satellite systems," in *2012 IEEE Vehicular Technology Conference (VTC Fall)*. IEEE, 2012, pp. 1–5.

[10] J. P. Choi and V. W. Chan, "Optimum power and beam allocation based on traffic demands and channel conditions over satellite downlinks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2983–2993, 2005.

[11] A. Ginesi, E. Re, and P.-D. Arapoglou, "Joint beam hopping and precoding in hts systems," in *Wireless and Satellite Systems: 9th International Conference, WiSATS 2017, Oxford, UK, September 14-15, 2017, Proceedings 9*. Springer, 2018, pp. 43–51.

[12] L. Chen, V. N. Ha, E. Lagunas, L. Wu, S. Chatzinotas, and B. Ottersten, "The next generation of beam hopping satellite systems: Dynamic beam illumination with selective precoding," *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2666–2682, 2022.

[13] V. K. Gupta, V. N. Ha, E. Lagunas, H. Al-Hraishawi, L. Chen, and S. Chatzinotas, "Combining time-flexible geo satellite payload with precoding: The cluster hopping approach," *IEEE Transactions on Vehicular Technology*, 2023.

[14] T. S. Abdu, S. Kisseleff, E. Lagunas, and S. Chatzinotas, "Flexible resource optimization for geo multibeam satellite communication system," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7888–7902, 2021.

[15] S. Guo, K. Han, W. Gong, L. Li, F. Tian, and X. Jiang, "An efficient multi-dimensional resource allocation mechanism for beam-hopping in leo satellite network," *Sensors*, vol. 22, no. 23, p. 9304, 2022.

[16] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Multi-satellite beam hopping based on load balancing and interference avoidance for ngso satellite communication systems," *IEEE Transactions on Communications*, vol. 71, no. 1, pp. 282–295, 2022.

[17] I. Siomina and D. Yuan, "Analysis of cell load coupling for lte network planning and optimization," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2287–2297, 2012.

[18] C. K. Ho, D. Yuan, L. Lei, and S. Sun, "Power and load coupling in cellular networks for energy optimization," *IEEE Transactions on wireless Communications*, vol. 14, no. 1, pp. 509–519, 2014.

[19] J. G. Walker, "Satellite constellations," *Journal of the British Interplanetary Society*, vol. 37, p. 559, 1984.

[20] 3rd Generation Partnership Project (3GPP), "Solutions for nr to support non-terrestrial networks (ntn)," 3GPP, Tech. Rep. Reference: 38.821, 2018.

[21] 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," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 112–131, 2019.

[22] X. Shen, S. Diamond, Y. Gu, and S. Boyd, "Disciplined convex-concave programming," in *2016 IEEE 55th conference on decision and control (CDC)*. IEEE, 2016, pp. 1009–1014.

[23] B. K. Sriperumbudur and G. R. Lanckriet, "On the convergence of the concave-convex procedure." in *Nips*, vol. 9, 2009, pp. 1759–1767.

[24] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of resource allocation in wireless networks: theory and algorithms*. Springer Science & Business Media, 2009, vol. 3.

[25] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[26] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

[27] J. N. Franklin, *Matrix theory*. Courier Corporation, 2012.

[28] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[29] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming, version 2.1," 2014.

[30] M. ApS, "Mosek optimization toolbox for matlab," *Users Guide and Reference Manual, Version*, vol. 4, no. 1, 2019.

[31] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[32] D. V. Broadcasting(DVB)Project, "Digital video broadcasting(DVB); second generation DVB interactive satellite system (DVB-S2X)," *ETSI EN 302 307 V1.4.1*, 2019.

[33] Broadcasting Satellite Services Organization, "Implementation guidelines for the second generation system for broadcasting, interactive services, news gathering, and other broadband satellite applications," Broadcasting Satellite Services Organization, Tech. Rep., 2020.

[34] 3rd Generation Partnership Project (3GPP), "Study on new radio (nr) to support non-terrestrial networks," 3GPP, Tech. Rep. Reference: 38.811, 2017.