PhD-FSTM-2025-094

The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 17 September 2025 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITE DU LUXEMBOURG
EN INFORMATIQUE

by

Kha Hung NGUYEN
Born on 6 May 1996 in Hà Nội, Vietnam

DYNAMIC SPECTRUM MANAGEMENT FOR EMERGING INTEGRATED SATCOM AND 5G NETWORKS

Dissertation defense committee

Dr. Björn Ottersten, Chairman
Professor, University of Luxembourg
Dr. Symeon Chatzinotas, Dissertation Supervisor
Professor, University of Luxembourg
Dr. Eva Lagunas
Assistant Professor, University of Luxembourg,
Dr. Joel Grotz
Senior Manager, SES, Luxembourg
Dr. Hien Quoc Ngo
Professor, Queen's University Belfast, United Kingdom

Dynamic Spectrum Management for Emerging Integrated SatCom and 5G Networks

Kha Hung NGUYEN

To my family, my love, and my best friends.

Abstract

The rapid growth of mobile traffic and emerging service requirements have accelerated the evolution of wireless networks, particularly through the integration of satellite communication (SatCom) systems and terrestrial networks (TNs). This convergence has led to Integrated Satellite-Terrestrial Networks (ISTNs), envisioned as a promising enabler of seamless, ubiquitous, and high-capacity connectivity in future 6G networks. By leveraging the strengths of both TNs and satellite networks (SatNets), ISTNs are able to enhance coverage, support direct-to-device (D2D) services, provide satellite-based backhaul, and improve spectrum utilization. However, several key challenges arise due to the dynamic behavior of non-geostationary satellites-especially low Earth orbit (LEO) satellites (LSats), heterogeneous terrestrial environments, uneven traffic demands, spectrum coexistence, and traditional fixed spectrum regulation.

Addressing these challenges necessitates a careful study of various technical issues. To efficiently deliver traffic demands under system constraints, critical technique issues must be addressed, including: (i) bandwidth (BW) allocation across services, systems, and satellite-based backhaul links, (ii) link association among users (UEs), base stations (BSs), LSats, (iii) ensuring seamless connectivity, (iv) comprehensive resource management, encompassing resource block (RB) assignment and power control; and (v) effective traffic management. These issues introduce various problems which must be addressed.

Driven by ISTN scenarios identified by standardization bodies and the aforementioned technique issues, within the scope of this thesis, three critical ISTN problems are considered: (1) backhaul-link BW allocation and two-tier user association, (2) throughput enhancement and seamless handover in C-band 5G-ISTNs, and (3) resource management in dynamic spectrum sharing ISTNs.

The first problem investigates the ISTN scenarios wherein a LEO constellation provides backhaul links for TNs. To efficiently deliver data from UEs to BSs and from BSs to LSats,

BW allocation backhaul links, a time-window-based optimization framework is proposed to minimize transmission time through jointly managing two-tier UE-BS-LSat association, backhaul BW allocation, subchannel assignment, and power control. To efficiently solve this challenging non-convex mixed-integer non-linear programming (MINLP) problem, a centralized algorithm is proposed by leveraging successive-convex-approximation (SCA) and compressed-sensing techniques. Furthermore, a decentralized algorithm is developed to support scalable and efficient deployment as well as to offload computations from central node.

The second problem explores ISTNs supporting the co-primary TN-SatNet spectrum use in urban environments, specifically for automotive UEs with both TN and SatNet direct-to-vehicle access. Accordingly, a joint UE association (UA) and power control mechanism is developed to maximize system throughput and minimize connection state changes. To deal with this problem efficiently, two algorithms are developed: an SCA-based algorithm addresses the full time-window problem while a prediction-based one enables the optimization in sequential sub-time-windows.

The third problem introduces a digital-twin (DT)-aided dynamic spectrum sharing (DSS) framework for D2D-enabled ISTNs, where the TN and SatNet share the same 5G-NR band. The goal is to minimize system congestion via joint optimization of BW slicing, traffic steering, and radio resource management (RRM). This framework comprises two optimization problems: (i) a joint-RA problem based on DT-predicted data for network and resource decisions, and (ii) a refinement problem that updates decisions using real-time feedback. Corresponding algorithms are designed to align with the proposed system architecture.

Extensive simulations are conducted under practical settings, especially ray-tracing (RayT)-based evaluations using realistic 3D maps of London for the second and third problems. Numerical results demonstrate the superior performance of the proposed algorithms for each of the three problems compared to conventional benchmarks in terms of transmission time reduction, throughput improvement, minimized connection transition events, and congestion alleviation, respectively. Overall, this thesis contributes efficient RRM designs for ISTNs across diverse scenarios and highlights promising directions for future ISTN research.

Preface

This Ph.D. thesis was conducted from May 2022 to September 2025 at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg. It was carried out under the supervision and mentoring of Prof. Symeon Chatzinotas, Prof. Eva Lagunas, Dr. Joel Grotz, and Dr. Vu Nguyen Ha. The periodic evaluation of the Ph.D. Thesis was duly performed by the CET members constituting the supervisors at SnT, University of Luxembourg, Luxembourg and Dr. Joel Grotz from SES, Luxembourg.

Collaboration

This work was made possible through a valuable collaboration with SES S.A. Additionally, I completed two internship periods at SES S.A. (June-September 2023 and June-September 2024), both under the supervision of Dr. Joel Grotz.

Support of the Thesis

This work was supported by the Luxembourg National Research Fund (FNR) through the Project INtegrated Satellite–TeRrestrial Systems for Ubiquitous Beyond 5G CommunicaTions (INSTRUCT) under Grant IPBG19/14016225/INSTRUCT. Additionally, the substantial support provided by SIGCOM is gratefully acknowledged, contributing significantly to the research conducted in this thesis.

Publications

The original publications and technique reports that have been carried out during the doctoral program are listed below. The publication category is referred in text as J-Journal, C-Conference, and B-Book Chapter while technical reports are denoted by TR.

Journal Papers

[J1] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Joint Two-tier User Association and Resource Management for Integrated Satellite-Terrestrial Network", in *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 16648-16665, Nov. 2024

[J2] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Enhanced Throughput and Seamless Handover Solutions for Urban 5G-Automotive Supported C-band Integrated Satellite-Terrestrial Networks", in *IEEE Transaction on Communications*, July 2025.

[J3] **H. Nguyen-Kha**, V. N. Ha, T. T. Nguyen, E. Lagunas, S. Chatzinotas, and J. Grotz, "Digital-Twin-Aided Dynamic Spectrum Sharing and Resource Management in Integrated Satellite-Terrestrial Networks", submitted to *IEEE Journal on Selected Areas in Communications*, 2025.

Conference Papers

[C1] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Two-Tier User Association and Resource Allocation Design for Integrated Satellite-Terrestrial Networks", in *Proc. 2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, Rome, Italy, 2023, pp. 1234-1239.

[C2] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "LEO-to-User Assignment and Resource Allocation for Uplink Transmit Power Minimization", in *Proc. IEEE WSA & SCC 2023*, Braunschweig, Germany, 2023, pp. 1-6.

[C3] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "An Experimental Study of C-Band Channel Modeling in Integrated Satellite and Terrestrial Systems", in *Proc. IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, Madrid, Spain, 2024.

[C4] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Seamless 5G Automotive Connectivity with Integrated Satellite Terrestrial Networks in C-Band", in *Proc. IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, Washington DC, USA, Oct 2024.

[C5] **H. Nguyen-Kha**, V. N. Ha, T. T. Nguyen, E. Lagunas, S. Chatzinotas, and J. Grotz, "DT-Aided Resource Management in Integrated Satellite-Terrestrial Networks", in *Proc. 2025 IEEE Global Communications Conference (GLOBECOM)*, Taiwan, Dec 2025.

Book Chapter

[B1] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Dynamic Spectrum Sharing for Integrated Satellite-Terrestrial Networks", book chapter in "Non-Terrestrial Networks: Paving the Way towards Global Connectivity" (ELSEVIER Nov 2024).

Technique reports - Internship with SES

[TR1] **H. Nguyen-Kha** and J. Grotz, "3GPP RAN100 summary and link budget assessment implementation for ISTN scenarios", SES-Internship technique report, Sept. 2023.

[TR2] **H. Nguyen-Kha** and J. Grotz, "Develop link evaluation tool for direct-to-device links", SES-Internship technique report, Sept. 2024.

Publications beyond Thesis

During my doctoral program, I have made contributions to other works on Edge AI implementation topic which are outlined in the following publications.

[J4] T.-D. Le and et al., "Onboard Satellite Image Classification for Earth Observation: A Comparative Study of ViT Models", *submitted to IEEE J-STARS*, 2025.

[C6] T. T. Nguyen and et al., "A Semantic-Loss Function Modeling Framework With Task-Oriented Machine Learning Perspectives", in *IEEE ICMLCN*, Spain, 2025.

[C7] T.-D. Le and et al., "GLUSE: Enhanced Channel-Wise Adaptive Gated Linear Units SE for Onboard Satellite Earth Observation Image Classification", in *arXiv:2504.12484*, 2025.

Acknowledgments

Firstly, I want to express my sincere gratitude to all who have supported me throughout my doctoral studies. In particular, I extend my deepest appreciation to my supervisor, Prof. Symeon Chatzinos, as well as Prof. Eva Lagunas, Dr. Vu Nguyen Ha, and Dr. Joel Grotz, for their invaluable guidance and the numerous opportunities they provided for my both professional and personal development. Special thanks to Prof. Eva Lagunas and Dr. Vu Nguyen Ha for their continuous efforts and dedication, which have played a vital role in my development. Their persistent support and mentoring have been the fundamental to my achievements during this Ph.D. journey. Additionally, I am grateful to Prof. Symeon Chatzinotas and Dr. Joel Grotz for their unwavering support through the highs and lows of my Ph.D. life, helping me navigate challenges and celebrate successes. Last but not least, I am deeply thankful to my parents, my brother, my love Linh, and my great friends for being next to me during my difficult time.

Hung Nguyen-Kha

Luxembourg, 1 August 2025

Contents

List of Figures

List of Tables

Abbreviations

2D	Two Dimensional
3D	Three Dimensional
5G	Fifth Generation of Mobile Communications
6G	Sixth Generation of Mobile Communications
BW	Bandwidth
BWA	Bandwidth Allocation
BWP	Bandwidth Part
CC	Connection Change
CDF	Cumulative Distribution Function
CINR	Carrier to Interference plus Noise Ratio
CN	Core Network
CNC	Center Network Controller
CSI	Channel State Information
CU	Central Unit
D2D	Direct-to-Device
DL	Downlink
DSS	Dynamic Spectrum Sharing
DT	Digital Twin
DU	Distributed Unit
ESA	European Space Agency
FCC	Federal Communications Commission
FDD	Frequency Division Duplexing
FR1	Frequency Range 1
FR2	Frequency Range 2
FSS	Fixed-Satellite Services

GEO Geostationary Earth Orbit

gNB gNodeB

GW Gateway

HAP High Altitude Platforms

HEO Highly Elliptical Orbit

HO Handover

IaSI intra-system interference

ISL Inter-Satellite Link

ISTN Integrated Satellite-Terrestrial Communication Network

ISyI inter-system interference

LEO Low Earth Orbit

LOS Line of Sight

LSat Low Earth Orbit Satellite

MAPE Mean Absolute Percentage Error

MEO Medium Earth Orbit

MINLP Mixed Integer Nonlinear Programming

NextG Next Generation

NGSO Non-Geostationary Orbit

NLOS Non-Line of Sight

NMS Network Management System

NOMA Non-Orthogonal Multiple Access

NR New Radio

NTN Non-Terrestrial Network

PL Path Loss

QoE Quality of Experience

QoS Quality of Service

RA Resource Allocation

RB Resource Block

RF Radio Frequency

RFB Radio Frequency Band

RRM Radio Resource Management

SatCom Satellite Communication

SatNet	Satellite Communication Network
SC	Sub Channel
SCA	Successive Convex Approximation
SF	Sub-Frame
SINR	Signal to Interference Plus Noise Ratio
SR	Sum-Rate
SRI	satellite radio interface
TAP	Terrestrial Network Access Point
TCU	TN Center Unit
TDD	Time Division Duplexing
TLE	Two-Line Element
TN	Terrestrial Network
TS	Time Slot
UE	User
UL	Uplink
URLLC	Ultra-Reliable Low-Latency Communication
VSAT	Very Small Aperture Terminal

Chapter 1

Introduction

RECENT years have witnessed the explosive growth in mobile traffic and emerging service demands, fundamentally reshaping wireless communication. According to Ericsson Mobility Report [1], total mobile network data traffic is expected to reach about 430 Exabytes (EB) per month by 2030, representing 2.6-fold increase compared to 2024. Furthermore, while the 5G mobile data traffic accounted for 35% of the total traffic with the average monthly data consumption of 19 Gigabytes (GB) per device in 2024, it is predicted to grown to 80% with 36 GB per device by 2030. Beyond these quantitative increases, the demand for seamless and ubiquitous connectivity has become paramount, enabling devices to maintain consistent connections and transition smoothly across networks regardless of location [2]. These escalating demands have been a primary driver for the evolution of wireless communication, particularly pushing the integration of satellite communication (SatCom) networks (SatNets) with terrestrial network (TN) systems. In particular, integrated satellite-terrestrial communication networks (ISTNs) have garnered significant attention in both industry and academia due to their numerous advantages, particularly for providing seamless and ubiquitous connectivity [2–4]. In particular, TNs offer high-speed data rates and reliable connections, they often fall short in delivering high-quality connections in sparsely populated or remote regions, such as rural, remote areas, and oceans, due to infrastructure limitations, difficulty in implementation, and economic constraints. Furthermore, due to the extremely high demand and massive connectivity in several areas such as dense urban regions and event hotspots, TNs can be overloaded and fail to ensure the connection for entire users (UEs). Conversely, non-terrestrial-networks (NTNs), especially SatNets, can support global and ubiquitous connectivity with a high data rate [5, 6]. As a result, ISTNs have emerged as

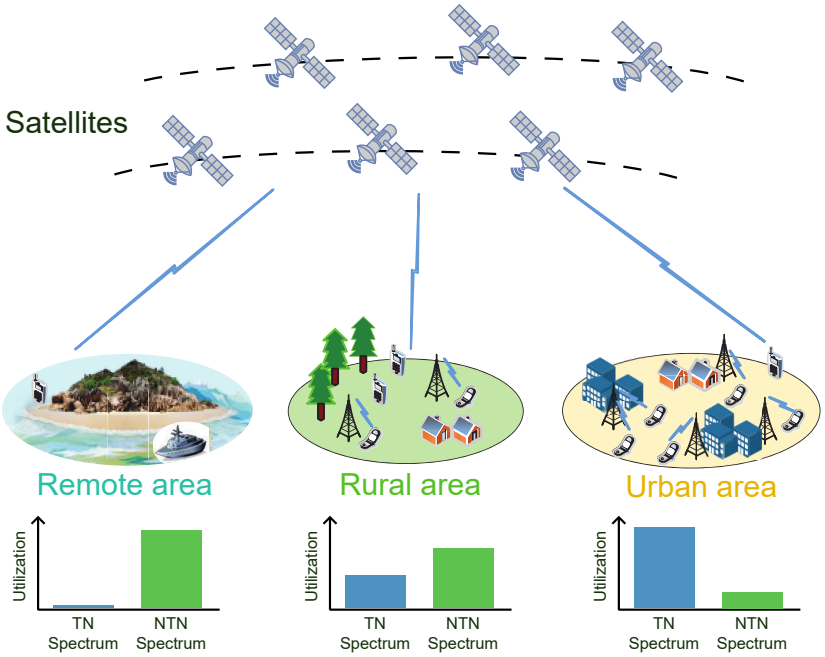


Figure 1.1: Spectrum utilization in integrated satellite and terrestrial networks.

a promising solution to leverage the strengths of both systems in the future sixth-generation (6G) wireless network, wherein satellites can provide the complementary connection, coverage enhancement, offloading for TNs, and satellite-based backhaul links. Furthermore, integration of TN and SatNet systems enables the coordination between the two systems to improve the overall system performance.

Traditionally, TNs and SatNets operate in predefined frequency bands. However, the uneven service and UE demands among ground areas may lead to inefficient spectrum utilization. Particularly, as per example depicted in Fig. 1.1, spectrum usage patterns vary across remote, rural, and urban areas. Specifically, SatCom is particularly well-suited for wide-area coverage, serving UEs in geographically dispersed and economically underserved regions where TN deployment may not be viable. Consequently, spectrum utilization for SatNets tends to be higher in remote and rural areas.

Conversely, TN spectrum utilization is typically high in urban environments, driven by high population densities and elevated data traffic demands, often leading to overutilization despite dense infrastructure. In contrast, TN spectrum is frequently underutilized in rural and remote areas due to sparse populations and limited demand. Furthermore, deploying TN backhaul infrastructure in such regions may pose challenges due to geographical constraints and economic inefficiency. Consequently, these areas may suffer from insufficient TN

connectivity meanwhile the TN spectrum is wasted.

Therefore, thanks to the ubiquitous advantage of SatNets, satellites can provide backhaul links for TNs or direct-to-device (D2D) connectivity to enhance and expand network coverage. Furthermore, coordinated spectrum sharing and resource management between the TN and SatNet can improve spectrum efficiency and boost overall system throughput. Driven by efforts of standardization bodies such as the International Telecommunication Union (ITU), 3rd Generation Partnership Project (3GPP), U.S. Federal Communications Commission (FCC), and European Space Agency (ESA), spectrum sharing between networks is increasingly being implemented to enhance utilization efficiency. The spectrum sharing regulations are described in Section 1.1.3.

1.1 Background

Traditionally, Satellite Communication (SatCom) has primarily focused on providing dedicated satellite services, such as Fixed Satellite Services (FSS) and Mobile Satellite Services (MSS). This includes offering internet connectivity to maritime, aerial, and terrestrial terminals, as well as enabling media broadcasting services [7]. However, spurred by the aforementioned user demands and emerging service requirements, SatCom is now evolving beyond these conventional roles to support the evolving 5G network. In particular, Non-Terrestrial Networks (NTN), which encompass SatNets, are actively being studied and standardized to support 5G New Radio (NR) [8, 9]. This strategic shift is further underscored by the significant attention garnered by the integration of SatNets with Terrestrial Networks (TNs), a development that leverages the complementary strengths of both systems, as previously discussed. In the context of this thesis focusing on ISTN scenarios supporting 5G NR, this section first provides the basic background of SatCom systems, especially satellite integration for 5G. Subsequently, ISTN scenarios are discussed. Especially, there exist cases where SatNets and TNs share the same frequency band. Hence, regulations relevant to spectrum sharing between SatNet and TN systems are subsequently examined.

Within the scope of this thesis, which specifically investigates ISTN scenarios supporting 5G NR, this section establishes the relevant foundational background. Initially, the fundamental aspects of SatCom systems, with a particular focus on satellite integration for 5G, are presented. This is followed by discussions on various ISTN scenarios. Given the prevalence of

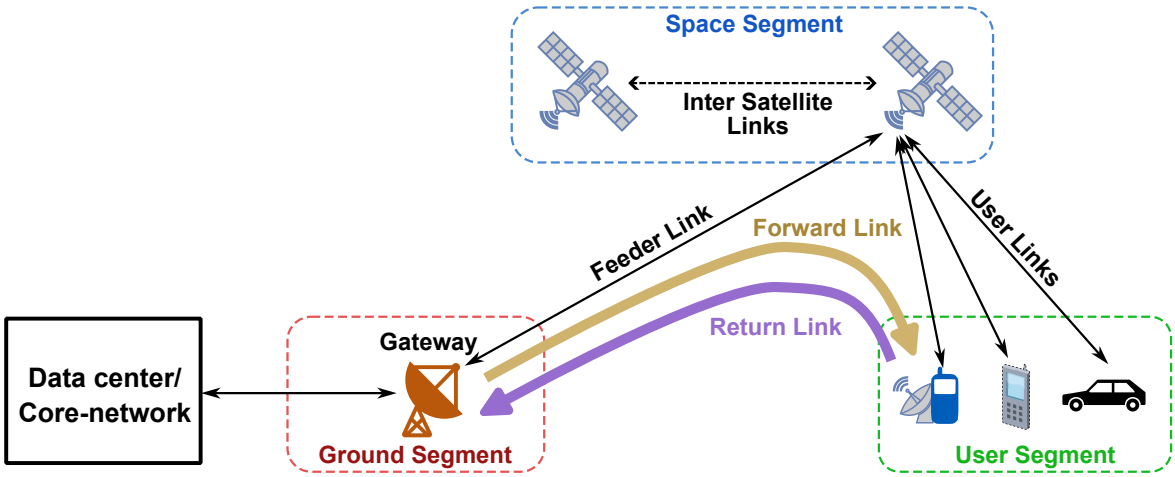


Figure 1.2: Basic SatCom system architecture.

ISTN instances where SatNets and TNs operate in shared spectrum, the relevant regulations pertaining to spectrum sharing between these systems are subsequently described.

1.1.1 Satellite Integration for 5G

SatCom System Architecture

A typical satellite communication system comprises three primary components: the space segment, which includes the satellite constellation; the ground segment, consisting of gateway (GW) stations; and the user segment, which involves user terminals deployed on both fixed and mobile platforms, as illustrated in Fig. 1.2. These components are interconnected through various types of communication links, which are defined as follows:

- **Feeder link:** The communication link between ground gateway stations and satellites.
- **User link (or service link):** The communication link between satellites and UEs.
- **Forward link:** The link direction from GW to UEs via satellites.
- **Return link:** The link direction from UEs to GW via satellites.
- **Inter-satellite link (ISL):** The link connecting satellites to each other.

In the context of SatCom systems, the term user equipment (UE) generally refers to either 3GPP-compliant devices or very small aperture terminals (VSATs). Depending on the deployment scenario, satellite UEs may serve as end-user terminals or as relay nodes, such

as terrestrial base stations. To further enhance the performance and flexibility of satellite constellations, ISLs can be employed, enabling multiple satellites to cooperate in executing complex tasks. ISLs also help reduce ground segment complexity by enabling communication between ground gateways (GWs) through any satellite in the constellation via ISLs. ISLs can be implemented using either traditional radio frequency (RF) or optical wireless communication. These links may be established between satellites within the same orbital plane or across different orbital planes, supporting more robust and adaptive network topologies.

This thesis specifically investigates providing services to end-users through both TNs and SatNets. Consequently, the primary contributions of this work are centered on the service link within the SatNet.

Satellite Payload Types: In SatCom systems, satellite platforms are typically equipped with either transparent or regenerative payloads. In a transparent (or bent-pipe) payload configuration, the satellite performs basic signal processing tasks such as RF filtering, frequency conversion, and amplification, without altering the baseband content. In contrast, regenerative payloads incorporate more advanced processing capabilities, including on-board execution of gNB (gNodeB) functions, allowing the satellite to demodulate, decode, and regenerate the signal before retransmission. A more detailed discussion of these satellite architectures is provided in subsequent sections.

Satellite Network Platform Types

SatNet platforms are typically categorized based on orbital altitude, each exhibiting distinct characteristics such as orbit type and beam footprint size. The primary SatNet classifications include Geostationary Earth Orbit (GEO), Medium Earth Orbit (MEO), Low Earth Orbit (LEO), and Highly Elliptical Orbit (HEO) satellites, as summarized in Table 1.1. Among these, GEO, MEO, and LEO represent the three basic satellite orbit types.

A GEO satellite is capable of covering approximately one-third of the Earth's surface, excluding polar regions. Therefore, a GEO constellation can cover more than 99% of the global population and economic activity. In contrast, LEO and MEO satellites, due to their lower altitudes and relative motion with respect to the Earth's surface, require larger constellations to achieve such global coverage and to ensure continuous service. Unlike these three orbits, which are typically circular, certain applications utilize elliptical orbits with high eccentricity, commonly known as Highly Elliptical Orbits (HEO). These elongated orbits offer the benefit

Table 1.1: Types of Satellite Orbits [8]

Platforms	**Altitude range**	**Orbit**	**Typical beam footprint size**
LEO satellite	300 - 1500 km	Circular around the earth	100 - 1000 km
MEO satellite	7000 - 25000 km	Circular around the earth	100 - 1000 km
GEO satellite	35786 km	Fixed projection on earth	200 - 3500 km
HEO satellite	400 - 50000 km	Elliptical around the earth	200 - 3500 km

of prolonged dwell times over specific regions during the satellite's ascent to, and descent from, apogee.

Among various satellite orbits, LEO has recently garnered significant attention. While LEO satellites inherently introduce high dynamics into the SatNet due to their rapid movement, they offer unique advantages such as low latency connectivity and high channel capacity due to their relatively low altitude compared to other orbits. Consequently, LEO satellites are predominantly considered in ISTN systems, as evidenced by recent studies [10–17].

Satellite Beam Configurations

A satellite cell is defined as the ground area illuminated by a satellite's beam footprint. These cells can be categorized into three main types based on their dynamic characteristics:

- **Earth fixed beam:** This type of beam footprint remains geographically constant over time. It is the typical configuration for GEO satellites, where the satellite's fixed position relative to the Earth's surface ensures continuous and static illumination of a specific ground area.
- **Quasi-Earth fixed beam:** Primarily employed by NGSO satellites, such as those in the LEO and MEO, equipped with steerable antennas. These antennas enable the beam to be directed towards and maintain coverage over a fixed geographical point on the ground for the duration of the satellite's visibility, typically lasting several minutes. While the illuminated area remains fixed, the beam's shape may dynamically adjust with the NGSO satellite's movement.
- **Moving-Earth fixed beam:** Also typically deployed on NGSO (MEO/LEO) satellites, but utilizing non-steerable antennas. As the satellite moves, its beam continuously sweeps across the geographical area on the ground. This results in dynamic, shifting geographical cells that track the satellite's trajectory.

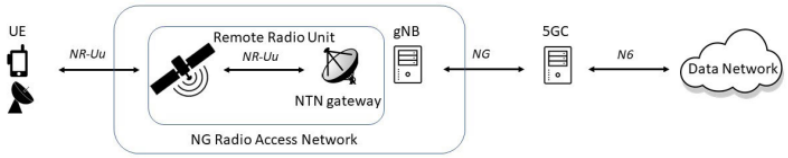


Figure 1.3: SatNet RAN architecture with transparent satellite [8, 18].

In the context of ISTNs involving LEO satellites, quasi-Earth fixed beam or moving-Earth fixed beam configuration can be utilized. Given the ability of steerable antenna, a quasi-Earth fixed beam configuration offers flexibility, enabling longer coverage time. However, this configuration also introduces complexity in the system, which requires a complicated beam management mechanism. In contrast, the moving-earth fixed beam configuration can operate with a lower complexity. Additionally, although the illuminated beam cell is non-steerable and moving along the satellite movement, the coverage can still be ensured for a sufficient constellation size. Hence, this configuration is adopted in many recent works such as [10–13, 15].

SatNet-based RAN Architecture

Based on 3GPP TS 38.821 [8], this section presents different SatNet-based RAN architecture types that involve the transparent, representative gNB, and representative gNB-DU (Distributed Unit) satellite payloads.

RAN architecture with transparent payload: In a transparent payload configuration as depicted in Fig. 1.3, the satellite performs frequency conversion and RF amplification for both uplink and downlink directions, effectively functioning as an analog RF repeater. Consequently, the satellite relays the NR-Uu radio interface between the feeder link and the service link, without terminating the NR-Uu interface. The satellite radio interface (SRI) used on the feeder link is NR-Uu, indicating that the satellite merely forwards the interface rather than processing it. All necessary functions for handling the NR-Uu interface are supported by the GW on the ground. Furthermore, multiple transparent satellites can be associated with a single ground-based gNB, enabling flexible and scalable network deployments.

RAN architecture with representative gNB payload: In this context, the satellite payload is equipped with signal regeneration capabilities for transmissions received from Earth as depicted in Fig. 1.4. Specifically, the following radio interfaces are involved:

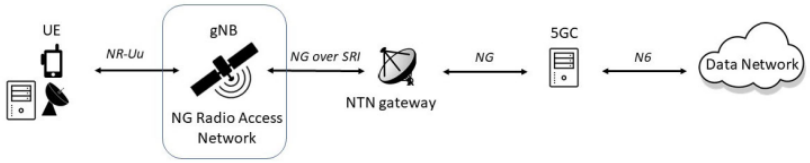


Figure 1.4: SatNet RAN architecture with representative satellite - gNB payload [8, 18].

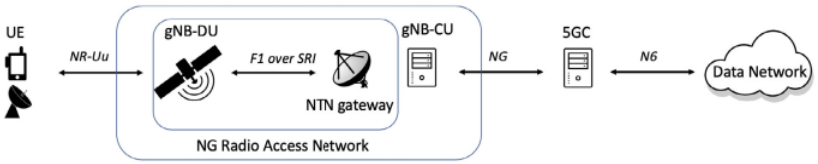


Figure 1.5: SatNet RAN architecture with representative satellite - gNB-DU payload [8, 18].

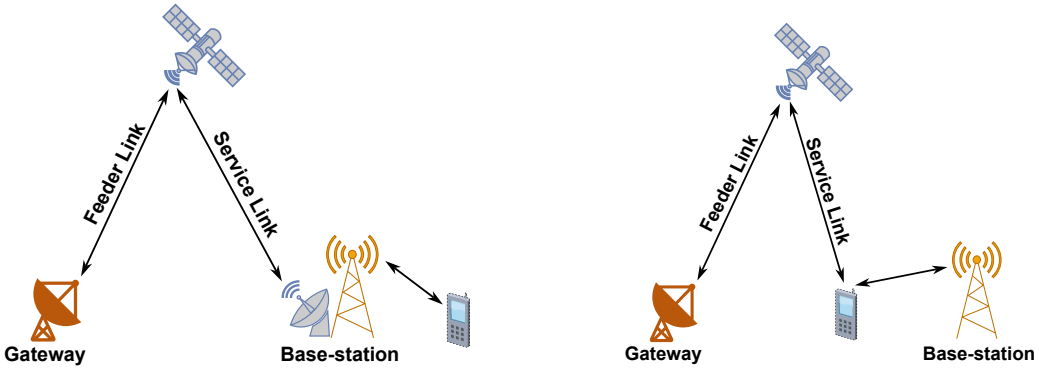
- NR-Uu: The standard radio interface used on the service link between the UE and the satellite.
- Satellite Radio Interface (SRI) over the feeder link between the SatNet GW and the satellite.

The SRI is the transport link for communication between the SatNet GW and the satellite. In addition to these interfaces, the satellite payload supports ISLs, which enable direct communication between satellites.

RAN architecture with representative gNB-DU payload: The NG-RAN logical architecture with CU and DU functional split, as defined in 3GPP TS 38.401, serves as the reference framework for NTN scenarios. In this architecture as depicted in Fig. 1.5, the satellite payload performs signal regeneration for transmissions received from Earth. The system incorporates the following interfaces:

- NR-Uu: The radio interface between the satellite and the UE, used on the service link.
- SRI on the feeder interface between the satellite and the SatNet GW, responsible for carrying the F1 protocol.

The SRI serves as a 3GPP-specified transport link for the logical F1 interface, enabling communication between the CU and DU. The SatNet GW is a transport network layer node and supports all necessary transport protocols for enabling this connectivity. Moreover, the satellite payload can support ISLs for communication between satellites.



(a) Satellite-based backhaul link case. (b) Direct-to-Device case.

Figure 1.6: Typical ISTN scenarios.

1.1.2 Integrated satellite terrestrial networks

This section provides typical use cases of SatNet in ISTNs which are proposed and discussed by 3GPP and ETSI in [8, 19–21]. Particularly, as depicted in Fig. 1.6, two primary ISTN scenarios are identified: (i) satellites serving as a backhaul link for TN systems and (ii) satellites enabling D2D connectivity for UEs.

Satellite-based backhaul scenario

In rural or remote areas, the deployment of TNs is often challenging or impractical, primarily due to the difficulties associated with establishing backhaul links. These challenges arise from geographic constraints, such as mountainous terrain or sparse infrastructure, as well as economic inefficiencies, that make such deployments cost-prohibitive. Given their inherent advantages of ubiquitous coverage, wide-area reach, and high capacity, satellites offer a viable solution by providing backhaul links for base-stations (BSs) in these underserved regions as depicted in Fig. 1.6a. To meet the high-capacity requirements of such backhaul links, satellite communications in these scenarios typically operate over high-frequency bands, such as the Ku-band or Ka-band.

Direct-to-Device scenario

Satellites can enable D2D connectivity for UEs, including both handheld and vehicle devices, thereby enhancing network coverage and alleviating traffic load on TNs, as illustrated in Fig. 1.6b. In such ISTN scenarios, satellites and terrestrial BSs can jointly serve UEs on the ground, offering improved service continuity and reliability. Furthermore, as explored in

ISTN use cases studied by 3GPP [9], MediaTek [22], and the US Federal Communications Commission (FCC) [23], SatNets may operate in the same or adjacent frequency bands as TNs. This co-frequency or adjacent-band operation forms the coexistence of SatNet and TN.

1.1.3 Spectrum Sharing Band

Regarding the industrial perspective, this section outlines the current and expected spectrum regulations across various frequency bands relevant to TN and SatNet coexistence in the D2D ISTN scenarios described above. Specifically, it examines the coexistence of satellite services and 5G New Radio (NR) within the mid-band, upper mid-band, and high-band frequency ranges.

Mid-Band

The mid-band spectrum is the frequency in the range of $1-6$ GHz. Mid-band spectrum is seen as the perfect frequency range for 5G since it can provide high data speed at a significantly long distance, offering a balance between coverage and capacity. Hence, the mid-band spectrum is widely designed in the world for 5G, and is described as the ideal spectrum for 5G by GSMA. This section describes the spectrum sharing in L, S, and C bands in the regulation aspects.

L-Band: In L-band ($1-2$ GHz), regarding the 5G NR standard, there are many defined NR bands in L-band. Regarding the spectrum sharing system aspect, Table 1.2 presents the NR bands for TNs and non-TNs (NTNs) which are overlap or adjacent. It is noted that bands n254 and n255 are defined to be used for 5G NR NTN. Regarding the spectrum sharing between satellite and terrestrial networks, the NTN band n255 and TN bands n24 and n99 fully overlap in both transmission and in uplink, respectively. Additionally, NTN band n254 is adjacent to TN bands n24 and n99 in the uplink. In summary, the spectrum sharing between 5G NR bands for TN and NTN systems in the L-band can be illustrated in Figure 1.7. The utilization of these bands can cause interference and adjacent channel leakage power, which need to be managed in the spectrum-sharing system.

S-Band: The S-band is defined as the frequency range from $2-4$ GHz. Currently, the S-band is widely used to provide terrestrial mobile services. In the context of 5G NR, there are many bands defined for TNs in 3GPP TS 38.104 [24]. Recently, under the development of NTN 5G NR, bands n256 (both DL and UL) and n254 DL have been defined in 3GPP TS 38.101-5 [25]. In particular, Table 1.3 shows the defined 5G NR bands for NTNs and TNs in

Table 1.2: Overlap and adjacent TN and NTN NR-bands in L-band.

Band	**Downlink frequency range (MHz)**	**Uplink frequency range (MHz)**	**Duplex mode**	**Note**
n24	1525 - 1559	1626.5 - 1660.5	FDD	TN
n99		1626.5 - 1660.5	SUL	TN
n254	2483.5 - 2500 (S-band)	1610 - 1626.5	FDD	NTN
n255	1525 - 1559	1626.5 - 1660.5	FDD	NTN

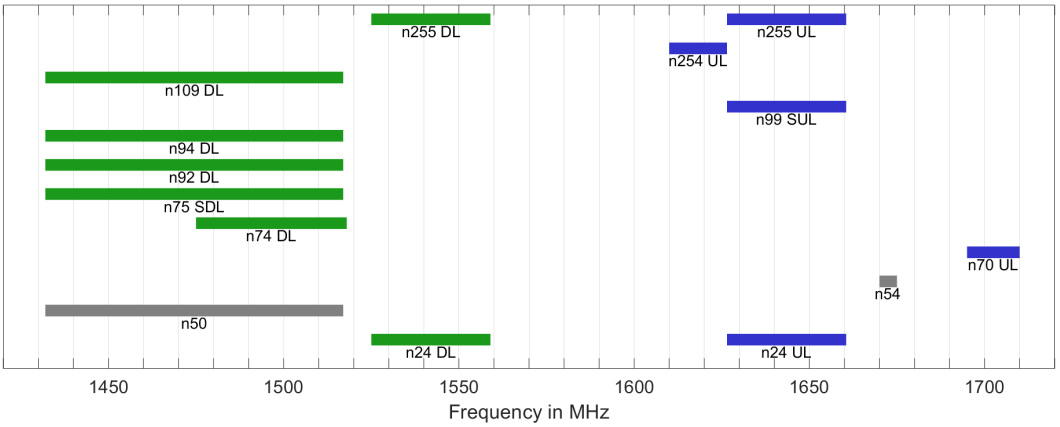


Figure 1.7: Spectrum sharing between 5G NR bands for TNs and NTNs in L-band.

the S-band frequency range.

Regarding the spectrum-sharing aspect between TN and NTN systems, we focus on the overlap and adjacent bands that are defined for TNs and NTNs. The overlap and adjacency of the TN and NTN band in the S-band can be depicted in Figure 1.8. It can be seen that the NTN bands overlap and are adjacent to many TN bands, which can potentially cause harmful interference in the spectrum-sharing ISTNs. For instance, band n256 UL is fully overlapped with bands n65 UL and partially overlapped with bands n70 DL, n25 DL, and n2 DL, while it is adjacent to bands n95 SDL, n34, and n1 UL. Additionally, band n256 DL fully overlaps with bands n66 DL and n65 DL, while it is adjacent to band n1 DL. Depending on the scenario, the interference in spectrum-sharing ISTNs in the S-band may be critical due to the overlapped bandwidth portion, and even the adjacency between bands due to the aggregated interference in large-scale system. The interference due to adjacent channel leakage power in the coexistence of TN and NTN is also studied in 3GPP TR 38.863 [9].

C-band: The IEEE defines the C-band as ranging $4-8$ GHz, while the US FCC defines it as 3.7 to 4.2 GHz. Previously, the C-band was allocated exclusively for fixed-satellite services (FSS) in many countries, such as Europe and the United States. However, due to spectrum

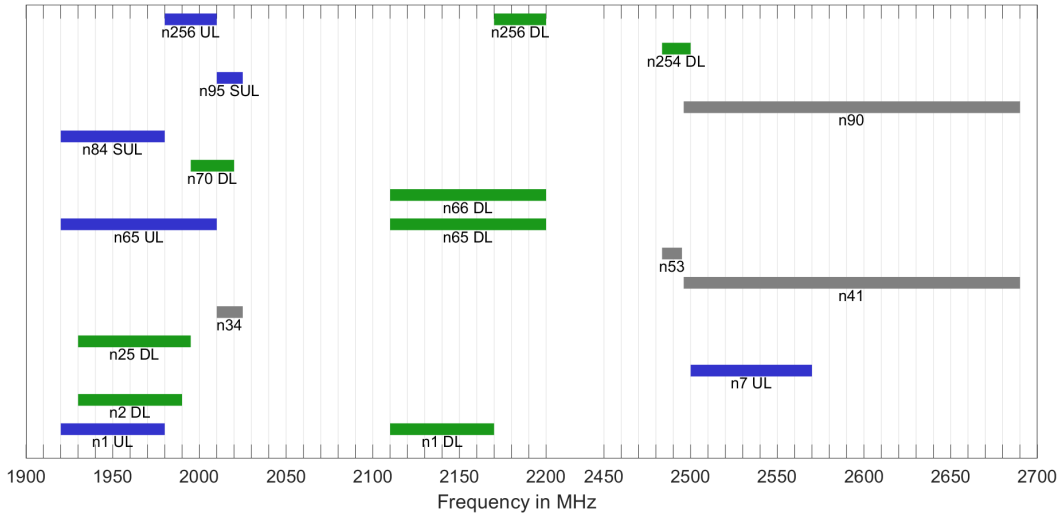


Figure 1.8: Spectrum sharing between 5G NR bands for TNs and NTNs in S-band.

Table 1.3: NR-bands in S-band.

Band	**Downlink frequency range (MHz)**	**Uplink frequency range (MHz)**	**Duplex mode**	**Note**
n1	2110 - 2170	1920 - 1980	FDD	TN
n7	2620 - 2690	2500 - 2570	FDD	TN
n30	2350 - 2360	2305 - 2315	FDD	TN
n34	2010 - 2025	2010 - 2025	TDD	TN
n38	2570 - 2620	2570 - 2620	TDD	TN
n40	2300 - 2400	2300 - 2400	TDD	TN
n41	2496 - 2690	2496 - 2690	TDD	TN
n48	3550 - 3700	3550 - 3700	TDD	TN
n53	2483.5 - 2495	2483.5 - 2495	TDD	TN
n65	2110 - 2200	1920 - 2010	FDD	TN
n66	2110 - 2200	1710 - 1780	FDD	TN
n70	1995 - 2020	1695 - 1710	FDD	TN
n90	2496 - 2690	2496 - 2690	TDD	TN
n95		2010 - 2025	SUL	TN
n97		2300 - 2400	SUL	TN
n254	2483.5 - 2500	1610 - 1626.5 (L-band)	FDD	NTN
n256	2170 - 2200	1980 - 2010	FDD	NTN

scarcity, TNs are now utilizing the C-band for terrestrial mobile services.

In addition to spectrum scarcity, the evolution of ISTNs has led to significant interest in using the C-band for 5G mobile services in both TNs and NTNs. In Europe, the C-band, spanning 400 MHz (3.4 – 3.8 GHz), is the primary band used for 5G systems for balancing the coverage and data speed [26]. In the US, the FCC has decided to reallocate 280 MHz of the C-band (3.7 – 3.98 GHz) for 5G systems and 20 MHz for a guard band [27]. However, the

coexistence of 5G and satellite systems in the C-band raises potential cross-system interference issues, as studied in recent works such as [28,29]. In addition, the interference channel between satellites and TNs has also been studied recently in [30]. Subsequently, the co-primary use between TNs and NTNs in the spectrum-sharing ISTNs has been studied in [31] wherein the resource allocation problem is considered. Regarding the 5G NR standard, 3GPP has defined sub-bands n77 and n78 within the C-band [25] accessible in many systems.

Regarding the industrial vision, the SES satellite company has recently proposed the spectrum pairing options of DL and UL transmissions for NTN systems [32]. Particularly, while the band $3.4 - 3.7$ GHz in the C-band can be used for the NTN DL transmission, the NTN UL one can operate in: 1) existing NTN bands in L or S bands, 2) band $3.7 - 4.2$ GHz, 3) band $3.4 - 3.7$ GHz, and 4) band $5.925 - 7.025$ GHz. As can be foreseen, opening C-band access will play a significant role in future ISTNs.

Upper Mid-Band

The upper mid-band spectrum, ranging approximately from 7 to 24 GHz, is designated as Frequency Range 3 (FR3) by the 3GPP, also called the centimeter wave (cmWave) band. This band is considered promising for balancing coverage and bandwidth, addressing the spectral limitations of mid-bands while offering better propagation and penetration than mmWave frequencies [33]. Consequently, it has recently garnered significant attention for commercial cellular communications and is regarded by the industry as a prime candidate for next-generation wireless networks. However, it also necessitates considering the effect on the incumbent system on this band [34,35]. Regarding these potential benefits, 3GPP has initiated studies on these bands [36]. Furthermore, the Federal Communications Commission has also identified the upper mid-band as crucial for meeting the rapid growth of data and service demands [37].

In the context of the United States, regarding the use of the 12 GHz band for TNs, the FCC released proposals and notices for the $12.2 - 12.7$ GHz band in 2021 and for the $12.7 - 13.25$ GHz band in 2022 [38,39]. These notices sought comments on adding 500 MHz for TNs in the $12.2 - 12.7$ GHz band and on using TNs to improve spectrum utilization in the $12.7 - 13.25$ GHz band while protecting incumbent users. Subsequently, in May 2023, the FCC unanimously voted to adopt new rules for the $12.2 - 13.25$ GHz band. These rules propose using certain services in the $12.7 - 13.25$ GHz band and seek comments on sharing options for the $12.2 - 12.7$ GHz band.

Table 1.4: NR operating bands in FR2-TN [24].

Band	**Uplink/Downlink operating band (MHz)**	**Duplex mode**
n257	26500 - 29500	TDD
n258	24250 - 27500	TDD
n259	39500 - 43500	TDD
n260	37000 - 40000	TDD
n261	27500 - 28350	TDD
n262	47200 - 48200	TDD
n263	57000 - 71000	TDD

Regarding existing systems, incumbents operate in the 12.2 – 13.25 GHz band. Specifically, the 12.2 – 12.7 GHz portion hosts broadcasting satellite service (BSS) including direct broadcast satellite (DBS) with the highest priority, non-geostationary FSS (NGSO FSS), and fixed service (FS). Additionally, the 12.7 – 13.25 GHz band includes FS, FSS, and mobile services (MSs), although usage in this portion is more sparse.

Subsequently, after reviewing comments from various stakeholders, the use of high-power terrestrial mobile broadband in the 12.2 – 12.7 GHz band was declined due to the potential for critical harmful interference with incumbent users. Conversely, a notice was released to explore the possibility of repurposing some or all of the 12.7 – 13.25 GHz band for high-power terrestrial mobile services, indicating that this band could be used for future TNs. However, despite the sparse utilization of the 12.7 – 13.25 GHz band, high-density deployment of high-power TNs could cause critical harmful interference with incumbent satellite services in this band or even in adjacent bands. This highlights the necessity of managing interference between ISTN systems in this band.

High-band

The high band in range 24 – 40 GHz includes the portions of Ka-band (26 – 40) GHz, FR2-TN (24 – 71) and the FR2-NTN (17 – 30) GHz, respectively. This band consists of the co-existence of satellite, TN 5G NR, and NTN 5G NR services. Currently, there are many FSSs in this band, which are provided through 17.7 – 21.2 GHz in DL and 27.5 – 31 GHz in UL. On the other hand, regarding the 5G NR FR2, many bands are also shared with FSS around the frequencies of 18 GHz and 28 GHz.

First, the defined bands in 5G NR FR2 for TNs [24] are described as in Table 1.4. Compared the operation band between FSS in Ka-band and FR2 for TNs, it can be seen that the

Table 1.5: NR operating bands in FR2-NTN [25].

Band	**Uplink (UL) operating band (MHz)**	**Downlink (DL) operating band (MHz)**	**Duplex mode**
n512	27500 - 30000	17300 - 20200	FDD
n511	28350 - 30000	17300 - 20200	FDD
n510	27500 - 28350	17300 - 20200	FDD

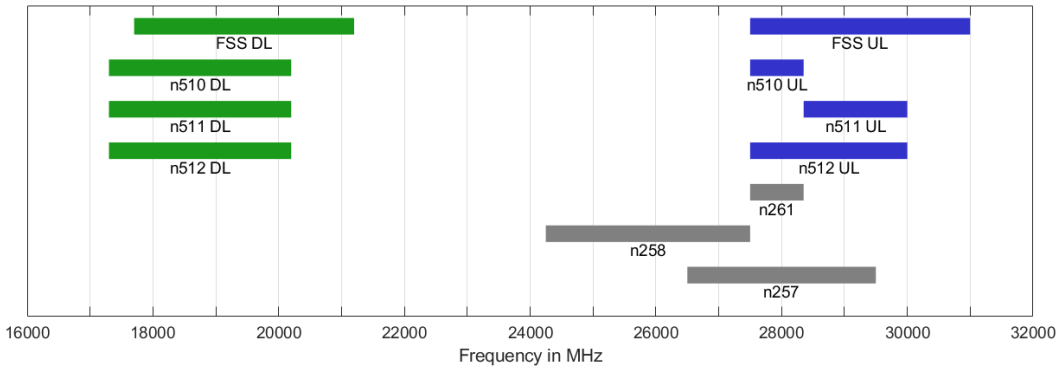


Figure 1.9: Spectrum sharing between FSS, 5G NR TNs and NTNs in high-band.

spectrum around the frequency of 28 GHz is shared between the uplink transmission of satellite services and the bands in 5G NR FR2-TN. In particular, the FSS uplink band $27.5 - 31$ GHz overlaps with bands n257 and n261, and is adjacent to band n258.

Besides, the operation bands in 5G NR FR2-NTN [25] are listed as Table 1.5. Difference from the bands in FR2-TN whose duplex mode is time-division-duplexing (TDD), that of bands in FR2-NTN is frequency-division-duplexing (FDD). In addition, these bands share the spectrum with FSS in both UL and DL. In particular, the FSS UL band $27.5 - 31$ GHz fully overlaps with bands n510 and n512, and partially overlaps with band n511. On the other hand, the FSS DL band $17.5 - 21.2$ GHz partially overlaps with all bands in FR2-NTN.

In addition, compared to the 5G FR2 bands for NTNs and TNs, the spectrum around 28 GHz is shared between UL NTN bands and TN bands. For instance, the UL band of NTN band n512 fully overlaps with TN bands n257 and n261, and is adjacent to NTN band n258, while the UL band of NTN band n511 partially overlaps with TN bands n257 and n258. In summary, the spectrum sharing between FSS, 5G NR TNs, and NTNs in the high-band can be illustrated in Figure 1.9. Due to the overlap and adjacent allocation among frequency bands, both co-channel and adjacent band interference need to be analyzed and addressed in the spectrum-sharing system.

1.2 Motivation and problem statement

According to the benefits of integrating TN and SatNet systems as mentioned above, ISTNs have emerged as a promising solution to leverage the strengths of both systems and enable effective coordination between them, thereby enhancing overall system performance. Furthermore, the LEO satellites provide unique advantages beyond ubiquitous and global connectivity thanks to their low altitude. These include reduced communication latency, improved channel capacity, and lower launch costs compared to MEO and GEO satellites. As a result, the integration of LEO satellites into TNs has garnered significant attention from both industry and academia.

However, the wide-area coverage provided by satellites in ISTNs introduces significant heterogeneity, stemming from diverse terrestrial environments and geographically uneven traffic demand, which complicates resource management. Furthermore, the inherent movement of LEO satellites poses dynamic system behavior, introducing further challenges to network management. Additionally, in D2D ISTN scenarios, where TN and SatNet may operate in the same frequency band, the coexistence further poses complex cross-system interference management challenges.

Addressing these challenges necessitates a careful study of various technical issues. To efficiently deliver traffic demands under system constraints, the following technique issues must be addressed:

- BW allocation: To adapt to uneven and dynamic traffic demands, appropriate BW allocation is crucial across services and between TN and SatNet systems. In satellite-based backhaul scenarios, BW allocation among backhaul links is particularly important to match link capacity with traffic demand across TN cells.

- Link association: In a sufficiently dense deployment, multiple options exist for establishing connections between nodes. Therefore, optimal UE-BS, UE-LSat, and BS-LSat associations are essential.

- Ensuring seamless connectivity: While the ISTN is envisioned as an enabler of seamless connectivity, the inherent LSat movement introduces dynamic issues, especially in complex environments like urban areas. Ensuring seamless connectivity with a minimum connection state transition in such scenarios is crucial.

- RB assignment: Effective RB assignment helps mitigate interference and aligns resource allocation with real-time traffic needs.
- Power control: Managing transmission power is crucial not only for limiting interference but also for optimizing energy efficiency and link reliability.
- Traffic management: Splitting and steering user traffic across optimal transmission paths enhances network efficiency and performance.

Addressing these technical issues requires an in-depth investigation of complex and interdependent design problems. Despite progress in the literature [10–17, 40–48], several critical research gaps remain in addressing the above challenges in a holistic and practical manner. This thesis focuses on three challenging problems, each representing a subset of the above issues and arising from typical ISTN scenarios discussed in the previous sections, as follows:

1. Backhaul link BW allocation and two-tier UE association.
2. Throughput enhancement and seamless handover in 5G-ISTN systems
3. Resource management in dynamic spectrum sharing ISTNs. These problems are briefly outlined in the following subsections.

Problem 1: Backhaul-Link BW Allocation and Two-Tier UE Association

This problem focuses on satellite-based backhaul links within ISTN scenarios. Faced with the challenges of deploying terrestrial backhaul in remote or rural areas, LEO satellites are deployed to provide backhaul connectivity for TNs. In these scenarios, the LEO satellite movement poses significant challenges to system management. These challenges arise from dynamic network topology and time-varying backhaul link capacity. In the uplink transmission context, UEs first transmit their data demand to BSs, then BSs forward this data to LEO satellites via satellite-based backhaul links. Moreover, the limited coverage duration of each LEO satellite necessitates that UE data should be transmitted within the shortest possible time. Therefore, to efficiently transmit UE data demand, it is essential to minimize the overall transmission time by jointly optimizing two-tier UE-BS and BS-LEO satellite associations as well as BW allocation for the backhaul link, in addition to power allocation and TN sub-channel (SC) assignment.

Problem 2: Throughput Enhancement and Seamless Handover in C-Band 5G-ISTN Systems

This problem investigated the D2D ISTNs supporting vehicle UEs. In such D2D ISTN scenarios, UEs can establish direct connections with LEO satellites. As outlined in Section 1.1.3 regarding the spectrum regulation and envision, TN and SatNet can operate in the same frequency band. This coexistence, particularly in the C-band as supported by the US FCC [23, 27], enables seamless connectivity for UEs capable of accessing both TN and SatNet systems. However, such coexistence also introduces the critical issue of cross-system interference between the two network types.

In addition, due to the wide coverage area of LEO satellites, the underlying terrestrial environment varies significantly within the satellite footprint, especially in complex environments like urban areas. Managing such geographical heterogeneity is crucial for ensuring efficient network operation. In such dynamic ISTNs, especially in urban cases with mobile UEs (e.g., vehicle UEs), connection changes are required to maintain high throughput. To this end, the second problem focuses on maximizing sum-rate and minimizing the number of connection changes to reduce signaling overhead, while accounting for cross-system interference and environmental impacts on UE channel conditions.

Problem 3: Resource Management in Dynamic Spectrum Sharing ISTNs

In spectrum sharing ISTNs, alongside frequency bands exclusively allocated to TNs and SatNets, both systems can also co-primarily operate in a shared spectrum band to serve UEs capable of supporting both connectivity types. However, traditional fixed spectrum allocation strategies may result in inefficient spectrum utilization, as discussed earlier in this chapter. When coupled with spatially and temporally uneven UE traffic demand, such inefficiencies can lead to network congestion. Assuming UEs support multi-connectivity, their traffic can be split and dynamically routed to suitable access nodes. To ensure efficient traffic delivery, the third problem extends beyond UE association, resource block (RB) assignment, and power control by also jointly optimizing BW slicing across services and systems, along with traffic steering decisions, with the objective of minimizing overall system congestion.

1.3 Thesis Outline and Contributions

The remainder of this thesis is organized into four chapters. Specifically, Chapters 2, 3, and 4 present the core contributions of this thesis, each addressing one of the three problems in ISTNs as outlined in Section 1.2. Chapter 5 concludes the thesis and outlines potential directions for future research. These chapters are summarized as follows.

Chapter 2: Joint Two-tier User Association and Resource Management for ISTNs.

This chapter aims to address **Problem 1: Backhaul-Link BW Allocation and Two-Tier UE Association**, as outlined previously. In particular, this work studies a time-window-based two-tier association and RA framework for the ISTNs. Specifically, the core objective is to minimize the total time required to transmit all UE's data to the LEO constellation. The optimization problem jointly considers several critical technical aspects including UE-BS and BS-LEO associations, subchannel allocation, transmission power control, and bandwidth management. Subsequently, two tailored algorithms, one centralized and one decentralized, are designed to carefully address the problem ensuring optimal operation. Numerical results validate the effectiveness of the proposed methods, demonstrating superior performance in minimizing transmission time compared to conventional benchmark schemes. The material presented in this chapter has been partially published in [49–51].

[49] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Joint Two-tier User Association and Resource Management for Integrated Satellite-Terrestrial Network", in *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 16648-16665, Nov. 2024

[50] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Two-Tier User Association and Resource Allocation Design for Integrated Satellite-Terrestrial Networks", in *Proc. 2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, Rome, Italy, 2023, pp. 1234-1239.

[51] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "LEO-to-User Assignment and Resource Allocation for Uplink Transmit Power Minimization", in *Proc. IEEE WSA & SCC 2023*, Braunschweig, Germany, 2023, pp. 1-6.

Chapter 3: Enhanced Throughput and Seamless Handover Solutions for 5G-ISTNs.

This chapter aims to address **Problem 2: Throughput Enhancement and Seamless Handover in C-Band 5G-ISTN Systems**, as outlined above. In particular, this work studies a time-window (TW)-based user association and power control mechanism for ISTNs in urban environments wherein TNs and SatNets use the same RFB in the C-band. Specifically, a multi-objective of max-sum-rate (SR) and min-connection-state-change (CC) is considered. Especially, to consider the impact of environmental characteristics on channel condition, the utilized channel and antenna models are compatible with the ray-tracing (RayT) mechanism and the actual 3D maps. To solve the optimization problem efficiently, an iterative algorithm is proposed based on the SCA technique. For practical implementation aspects, a prediction-based algorithm is further proposed. Numerical results demonstrate the superiority of the proposed algorithms over benchmark schemes in addressing the joint SR-CC optimization problem. The contents of this chapter have been extracted and disseminated through [30,31, 52].

[52] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Enhanced Throughput and Seamless Handover Solutions for Urban 5G-Automotive Supported C-band Integrated Satellite-Terrestrial Networks", in *IEEE Transaction on Communications*, July 2025.

[30] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "An Experimental Study of C-Band Channel Modeling in Integrated Satellite and Terrestrial Systems", in *Proc. IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, Madrid, Spain, 2024.

[31] **H. Nguyen-Kha**, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Seamless 5G Automotive Connectivity with Integrated Satellite Terrestrial Networks in C-Band", in *Proc. IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, Washington DC, USA, Oct. 2024.

Chapter 4: Digital-Twin-based Resource Management for ISTNs.

This chapter aims to address **Problem 3: Resource Management in Dynamic Spectrum Sharing ISTNs**, as described previously. In particular, motivated by the challenges

discussed above and the D2D connectivity scenarios promoted by the ESA and U.S. FCC, this work studies novel digital-twin (DT)-aided dynamic spectrum sharing (DSS) mechanisms for ISTNs supporting D2D connectivity, where TNs and SatNets share the same 5G-NR RFBs. Given the uneven traffic demand across services and systems, the objective is to minimize system congestion by enhancing spectrum utilization through DSS and efficient radio resource management (RRM). Based on this model, two resource allocation (RA) optimization problems are formulated: (i) joint-RA, which leverages DT-predicted information, and (ii) refinement, which adjusts decisions using real-time system feedback. Correspondingly, two efficient algorithms aligned with the proposed system architecture are developed to solve these problems. Numerical results demonstrate that the proposed algorithms significantly outperform conventional greedy and heuristic benchmarks in reducing system congestion. The contents of this chapter have been extracted and disseminated through [53,54].

[53] **H. Nguyen-Kha**, V. N. Ha, T. T. Nguyen, E. Lagunas, S. Chatzinotas, and J. Grotz, "Digital-Twin-Aided Dynamic Spectrum Sharing and Resource Management in Integrated Satellite-Terrestrial Networks", submitted to *IEEE Journal on Selected Areas in Communications*, 2025.

[54] **H. Nguyen-Kha**, V. N. Ha, T. T. Nguyen, E. Lagunas, S. Chatzinotas, and J. Grotz, "DT-Aided Resource Management in Integrated Satellite-Terrestrial Networks", in *Proc. 2025 IEEE Global Communications Conference (GLOBECOM)*, Taiwan, Dec. 2025.

Chapter 5: Conclusions and Future Work.

This chapter presents a concise summary of the thesis's three main contributions and their corresponding conclusions. Furthermore, it discusses potential extensions of the current work and identifies open directions for future research.

Chapter 2

Joint Two-tier User Association and Resource Management for ISTNs

THIS chapter investigates the uplink transmission of an integrated satellite-terrestrial network, wherein the LEOSats provide backhaul services to isolated cellular BSs for forwarding mobile UE data to the core network. In this integrated system, the high mobility of LSats introduces significant challenges in managing RA, as well as the associations between UEs, BSs, and LSats for supporting users' demands efficiently, while also dynamically balancing the capacity of UE-BS access and BS-LEO backhaul links. Regarding these critical issues, the chapter aims to jointly optimize the two-tier UE-BS and BS-LSat association, sub-channel assignment, bandwidth allocation, and power control to meet users' demands in the shortest transmission time. This optimization problem, however, falls into the category of mixed-integer non-convex programming, making it very challenging and requiring advanced solution techniques to find optimal solutions. To tackle this complex problem efficiently, we first develop an iterative centralized algorithm by utilizing convex approximation and compressed-sensing-based methods to deal with binary variables. Furthermore, for practical implementation and to offload computation from the central processing node, we propose a *Dec-Alg* that can be implemented in parallel at local controllers and achieve efficient solutions. Numerical results are also illustrated to strengthen the effectiveness of our proposed algorithms compared to traditional greedy and benchmark algorithms.

2.1 Introduction

Recent years have witnessed a rapid development of terrestrial networks (TNs), marked by an increase in the number of connected devices, and the expansion of services and applications. This evolution necessitates the next-generation wireless communication networks be capable to meet the growth in cellular traffic demand and to satisfy various service requirements such as massive connectivity, seamless connectivity, and high reliability across global coverage [2–4,55,56]. In response to these pressing demands, TN infrastructure providers are compelled to deploy an increased number of BSs to enlarge network coverage and capacity. However, this solution poses considerable challenges, mainly due to the prohibitive costs associated with implementing backhaul links connecting these BSs to the core network [57]. It becomes more critical for the rural regions, suburban areas, mountainous terrains, and oceanic expanses due to the low economic efficiency [58]. To overcome these challenges, integrating satellites with cellular systems to establish integrated satellite-terrestrial networks (ISTNs) has been considered as a promising solution for future global connectivity [3, 7, 51, 59].

Previously, the satellite-based connection was considered a last resort for cellular backhaul due to bandwidth limitations and high costs; however, advancements in the satellite industry, particularly non-geostationary satellite constellations and ground segment innovations, have transformed satellite backhaul into a key solution, not only extending the mobile network coverage but also supporting high-quality broadband connectivity. Compared to geostationary (GEO) and medium-earth-orbit (MEO) systems, LSat constellations (including the very-low-earth-orbit) demonstrate additional benefits for certain applications when connecting the terrestrial BSs to the core network [60]. In particular, their lower altitude ensures lower latency in data transmission which is suitable for applications requiring near-real-time communication [61,62]. In addition, the lower altitudes which are equivalent to the higher channel gains, also make LSats to be capable of supporting high data rate transmissions [61].

Despite the substantial advantages of global coverage, higher channel gain, and low transmission latency, the integration of LSats and TNs also poses significant challenges in dynamic network management. Specifically, the inherent high mobility of LSats requires novel dynamic association strategies to establish stable two-hop links between UEs, BSs, and LSats. In addition, the RA mechanism should be designed carefully to maintain user traffic demands while dynamically balancing the capacity of UE-BS access and BS-LEO backhaul links. The novel

design should be well developed to ensure seamless handovers, manage power control, and mitigate cross-interference among all terrestrial access points [7, 59]. Hence, jointly optimizing the transmission and association strategies for LEO-based ISTNs has raised an important and challenging research topic.

2.1.1 Related Works

Recently, developing novel RA mechanisms for the ISTNs and user selection problems have received significant attention in research works [10–15, 40–44]. Particularly, the authors in [40] studied the service pricing and user selection problem, wherein evolutionary game-based framework is proposed to solve the user-computation provider association problem under the computation requirement of users. In [10] a LSat is deployed to assist the TNs for cache offloading, wherein power allocation, cache placement, and offloading decisions are optimized for energy-efficient maximization. However, only one LSat is considered in this work. The authors in [41] have proposed a frequency-reuse-based bandwidth-sharing solution for the ISTNs to extend the coverage area. In this scheme, each satellite beam uses a specific frequency while the terrestrial cells in its coverage reuse the frequency bands from other beams, however, only power and bandwidth allocation are optimized. In [11], the authors considered an ISTN supporting IoT uplink transmission. The transmission process consists of two hops, e.g., IoT devices to the central earth stations (CETs) and from CETs to the LSat. This work aimed to optimize the RA to minimize the completion time of each hop separately. However, the work in [11] solely addressed the one-tier LEO-BS associations, yet it is essential to consider the end-to-end UE-BS-LEO performance comprehensively. In particular, the user data is transmitted through two hops, i.e., UE-BS-LSat, and the aggregated data throughputs at each BS and LSat are limited by the backhaul link capacity and bandwidth limitation. Therefore, one must optimize the RA for both UE-BS access and backhaul links and manage the two-tier UE-BS-LEO associations regarding the dynamics of satellite movements to balance the traffic at each BS and LSat and improve the end-to-end performance of an ISTN.

As a result, different two-tier association and RA solutions for the ISTNs have been studied in recent works [12–15, 42, 43]. In [42], a novel joint offloading and user association solution was designed to maximize the total transmission rate for ISTNs supporting a disaster area that consists of a number of high-altitude-platforms (HAPs) and one LSat. In [43], a random access service for UEs in an ISTN is studied wherein every UE can choose either the BS or the

LSat to associate with for data transmission. In this work, UE-BS association, bandwidth, and power allocation are optimized separately for maximizing the transmission rate of the UEs. Note that only one LSat is considered in the systems of [42, 43]. Next, the ISTNs with multiple LSats were considered in [12, 13]. Specifically, [12] aimed to maximize the sum rate and also the number of served UEs in an ISTN wherein the LSats provide the backhaul access service to TNs including a macro cell and several small cell BSs. To obtain the solution to this multiple-object optimization problem, two matching algorithms were proposed. In a different approaching way, the authors in [13] discussed how two network operators, i.e, satellite operator (SO) and terrestrial operator (TO), cooperate to provide the radio access services to a set of users effectively. To do so, a Stackelberg game was formulated based on which the user association, bandwidth allocation, and service price are optimized so that the revenue equilibrium can be achieved. However, it is worth noting that the transmit power has been kept unchanged and not optimized in both [12, 13]. Furthermore, all [12, 13, 42, 43] only focused on RA frameworks for the snapshot model at a specific time where the system dynamic has not been investigated. Particularly, the high mobility of the LSats leads to the dynamic of the system, especially the variance of backhaul link capacity between BSs and LSats. Therefore, the window-based model should be considered, where the system is considered within subsequent windows of time, capturing the dynamics of the systems.

Subsequently, the window-based model regarding LSat's movement within a number of time slots has been studied to develop dynamic RA mechanisms for the ISTNs in [14, 15, 44]. The authors in [14, 15] discussed how the LSats use Ka-band to provide the backhaul link to TNs proficiently. A three-dimensional LEO constellation optimization algorithm was proposed to minimize the number of satellites under the seamless coverage and backhaul capacity requirements over time windows. However, these works were analyzed without considering the RA at the UE-BS tier. [44] proposed two dynamic window-based greedy algorithms which handily address the user association and enlarge the coverage areas while maximizing the total network capacity. In this work, multiple frequency access and users' demand have been not considered. To the best of our knowledge, the joint design of two-tier UE-BS-LEO association, sub-channel assignment, bandwidth allocation, and power control for ISTNs over time regarding the user demand and limited-backhaul constraints have not been thoroughly studied in the literature. Specifically, different design aspects of the existing frameworks studying two-tier UE-BS-LEO are summarized in Table 2.1. That motivates us to fill the gap

Table 2.1: Summary of works

Works	Model	Sat. No.	Objective function	QoS	2T UA	SA	TN PC	NTN PC	BWA
Ref. [42]	Snapshot	1	Max sum-rate	Rate-threshold	✓	✗	✗	✗	✗
Ref. [43]	Snapshot	1	Max sum-rate	Rate-threshold	✓	✗	✓	✗	✓
Ref. [12]	Snapshot	Multiple	Max-rate & Num. served UEs	None	✓	✓	✗	✓	✓
Ref. [13]	Snapshot	Multiple	The revenue equilibrium	None	✓	✗	✗	✗	✓
Ref. [15]	Snapshot	Multiple	Min Num. used satellites	Global coverage	✓	✗	✗	✗	✗
Ref. [44]	Window-based	Multiple	Max sum-rate	None	✓	✗	✗	✓	✗
Our work	Window-based	Multiple	Min completion time	User demand	✓	✓	✓	✓	✓
SA: Sub-channel assignment TN PC: Power control in TNs 2T UA: Two-tier user association NTN PC: Power control in NTNs BWA: Bandwidth allocation									

in this chapter.

2.1.2 Research Contributions

In this work, we study a time-window-based two-tier association and RA framework for the ISTNs. The data from UEs are forwarded to the LEO constellation systems through a two-hop transmission including terrestrial access links UE-BS and backhaul links BS-LSat. Regarding the quality of experience at UEs in terms of the data delivery time, this work aims to minimize the duration required to transfer all UEs' data to the LEO constellation. In this system, the technical issues encompassing UE-BS and BS-LEO associations, subchannel allocation, transmission power control, and bandwidth management are carefully addressed and designed to ensure optimal operation. The main contributions of this work can be summarized as follows:

- We study the window-based two-hop transmission model for LEO-based ISTNs under a practical system architecture, which features the dynamic topology and time-varying LEO-based backhaul link capacity. Regarding this, we develop a comprehensive problem optimizing the two-tier user association and RA jointly under the backhaul link capacity constraint. The objective is to offload all user data to the LEO constellation during the shortest duration. The optimization problem is formed as a challenging non-convex mixed-integer non-linear programming (MINLP) due to the combination of the discrete objective function, non-convex constraints, and coupling between binary and continuous variables.

- To deal with this difficult problem, we propose two efficient solutions by utilizing compressed-sensing and successive-convex-approximation methods, namely centralized and decentralized algorithms, which align two proposed network management approaches, respectively. In particular, for the performance purpose, we develop the centralized algo-

rithm which can be executed at central network controller, however, it may pose critical issues of practical implementation.

- For the implementation purpose, we further develop the decentralized algorithm that enables resource management to be optimized distributedly at local controllers with limited signaling between them. Consequently, we discuss how the decentralized algorithm can be efficiently implemented in the practical systems.

- For comparison purposes, we introduce a greedy-based algorithm and adapt an existing framework to suit our design requirements. The numerical results with practical parameters show the superiority of the proposed algorithms compared to the two reference baselines. Furthermore, the small outcome gap between the centralized and decentralized solutions also confirms the effectiveness of distributed designs.

The preliminary result of this manuscript was presented in [50] where only the centralized approach and the simple simulation were discussed. The rest of this chapter is organized as follows. Section 2.2 introduces the system model and problem formulation. In Section 2.3, one proposes two centralized and decentralized solution approaches. Next, benchmark algorithms and complexity analysis are described in Section 2.4. Subsequently, the numerical results are presented in Section 2.5 which is followed by the conclusion given in Section 2.6.

2.2 System Model and Problem Formulation

This work aims to study RA designs for the uplink transmission of a two-tier ISTN, where satellites of an LEO constellation connect to some isolated terrestrial BSs to forward the data from ground UEs located in a specified geographical area to the core network. One assumes that the capacity of the feeder links between the gateways (GWs) and LEO is large enough, i.e., Q/V-band or optical [63,64]; hence, this work only focuses on two-hop transmission from UEs to BSs, and from BSs to LSats. Let $\mathcal{N} \triangleq \{1, \ldots, N\}$ and $\mathcal{K} \triangleq \{1, \ldots, K\}$ be the sets of BSs and UEs, respectively. For convenient, we further denote $m-$th LSat, $n-$th BS and $k-$th UE as LEO_m, BS_n and UE_k, respectively. A multiple-time-slot transmission model is considered, in which every UE, i.e., UE_k, intends to transmit a data amount of D_k bits to the core network through terrestrial access and LEO-provided backhaul links. Here, we also assume that the duration of one time-slot (TS) is T_S (s). Regarding the coverage pattern of the NTN, the beam coverage can be moving with the satellite or fixed on the earth [65]. To sake simplicity, the

beam coverage is assumed to be moving with the LSat. Furthermore, to reduce the complexity, only the satellites in the constellation that can provide effective backhaul links to the BSs are considered. Herein, the backhaul link from an LSat to a BS is marked as effective if the BS's location is inside the 3-dB coverage of that satellite's beam pattern. In addition, one assumes that the beam pattern is fixed, i.e., without adaptive beamforming. Since LSats move, the set LSats having effective links to the BSs varies over time. Let $\mathcal{M}_t \triangleq \{1, \ldots, M_t\}$ be the set of such satellites in TS t. In this system, the connections between UEs to BSs are processed following the 5G-NR standard over the mid-bands, i.e., 1-6 GHz, while the BSs connect to LSats by using very-small-aperture-terminal (VSAT) antennas over the high-bands, e.g., Ka-band. In addition, the system is assumed to support delay-tolerant mechanisms. In particular, due to the dynamic of the system, especially the high mobility of LEO satellites [66], and the limitation of the system resource, link disruption can occur, i.e., temporary disconnection for certain UEs. Therefore, we assume that the system supports delay- and disruption-tolerant networking architecture, wherein each node can store the information and forward it later when the connectivity is restored [8,66,67]. On the other hand, the impact of link disruption and delay-requirement are discussed later in the numerical results.

2.2.1 Network Management Operation

A vision of ISTN architecture in this work is illustrated by Fig. 2.1. One assumes that the cellular BSs are deployed in clusters due to specific geographic conditions. In each cluster, a local controller (LC) is deployed which connects to all corresponding BSs through limited capacity links. Via these links, the control information can be exchanged based on which the transmission of BSs and UEs inside the cluster can be managed. On another side, the LEO constellation is operated by an NTN controller connecting to the GWs. In our vision, two network management structures can be considered for operating the cooperation between satellites and terrestrial networks, so-called centralized and decentralized.

- In the centralized structure, all LCs (including NTN one) are connected to a center network controller (CNC). This CNC can optimize all resource management functions, then, forward the controlling results to LCs for operation.

- In the decentralized structure, the distributed network operation can be processed by LCs with minimized information signaling among them via the central unit.

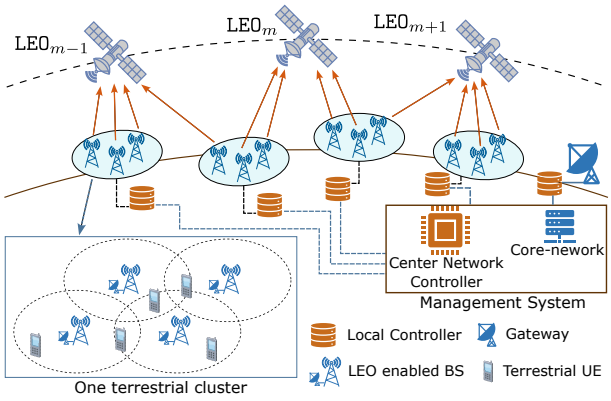


Figure 2.1: System architecture of the ISTNs.

2.2.2 Channel Model

In this section, we discuss the channel models of the UE access links and the Ka-band backhaul links. The channel coefficient from BS_n to LEO_m is modeled as

$$\bar{g}_{m,n}[t] = \frac{\sqrt{G^{\mathsf{LEO}} G^{\mathsf{BS}} \psi(\theta_{m,n}) \mathtt{PL}^{\mathsf{NTN}}_{m,n}[t]}}{\varrho^{\mathsf{c}}_{m,n}[t] \varrho^{\mathsf{r}}_{m,n}[t]} e^{j2\pi(tT_S \nu^t_m - f_c d_{m,n}[t]/c)}, \tag{2.1}$$

where $\varrho^{\mathsf{c}}_{m,n}[t]$ and $\varrho^{\mathsf{r}}_{m,n}[t]$ stand for attenuation due to cloud and rain, $\mathtt{PL}^{\mathsf{NTN}}_{m,n}[t] = 1/(4\pi f_c d_{m,n}[t]/c)^2$ and $d_{m,n}[t]$ denote the free space path loss and the distance from BS_n to LEO_m at TS t; f_c and c are the operation frequency and light speed. G^{LEO} and G^{BS} are the antenna gains of the LSats and BSs. ν^t_m denotes the Doppler shift due to the mobility of LEO_m at TS t [68]. One assumes that the Doppler shift can be compensated at the LSat payload [8]. Moreover, $\psi(\theta)$ is the normalized beam pattern expressed in [65] as $\psi(\theta) = 1$ if $\theta = 0$, and $\psi(\theta) = 4\left|\frac{J_1(ba\sin\theta)}{ba\sin\theta}\right|$ if $\theta \neq 0$ where a is the antenna aperture, $J_1(\cdot)$ is the first-order Bessel function, $\theta_{m,n}$ is the boresight angle at the viewpoint of LEO_m to BS_n, and $b = 2\pi f_c/c$. Besides, the channel coefficient of link $\mathsf{UE}_k - \mathsf{BS}_n$ over SC s at TS t is modeled as

$$\bar{h}_{n,k,s}[t] = \sqrt{\mathtt{PL}^{\mathsf{TN}}_{n,k}} \left(\sqrt{\frac{\kappa^{\mathsf{TN}}}{\kappa^{\mathsf{TN}}+1}} \bar{h}^{\mathsf{LOS}}_{n,k,s}[t] + \sqrt{\frac{1}{\kappa^{\mathsf{TN}}+1}} \bar{h}^{\mathsf{NLOS}}_{n,k,s}[t] \right), \tag{2.2}$$

where $\mathtt{PL}^{\mathsf{TN}}_{n,k}$ is the path-loss between UE_k and BS_n, $\bar{h}^{\mathsf{LOS}}_{n,k,s}[t]$ and $\bar{h}^{\mathsf{NLOS}}_{n,k,s}[t]$ are the LOS and NLOS small scale falding of link $\mathsf{UE}_k - \mathsf{BS}_n$ over SC s at TS t, respectively. κ^{TN} is the Rician factor for the $\mathsf{UE} - \mathsf{BS}$ link. Additionally, the random-walk process is exploited to generate

the *"time-correlation"* channels [69–71]. In particular, $\bar{h}^{\mathsf{NLOS}}_{n,k,s}[t]$ is generated as

$$\bar{h}^{\mathsf{NLOS}}_{n,k,s}[t] = (1-\rho)\bar{h}^{\mathsf{NLOS}}_{n,k,s}[t-1] + \rho\tilde{h}_{n,k,s}[t], \tag{2.3}$$

where $\bar{h}^{\mathsf{NLOS}}_{n,k,s}[0]$ and $\tilde{h}_{n,k,s}[t]$ are zero-mean unit-variance random variables, and ρ is the random-walk factor.

2.2.3 Transmission over the Terrestrial UE-BS Access Links

Based on the 5G-NR standard, we assume that each UE can connect to at most one BS while each BS can serve multiple UEs at each TS. In addition, the transmission bandwidth (BW) for the terrestrial network is divided into S orthogonal sub-channels (SCs). Here, the BW of each SC is $W_{\mathsf{SC}} = 180 \times 2^{\mu}$ (kHz) where μ stands for the pre-selected numerology [72, 73]. The set of SCs is denoted by $\mathcal{S} \triangleq \{1, \ldots, S\}$. Regarding the UE-BS association and the SC allocation for UEs, we introduce a binary variable $\boldsymbol{\alpha}[t] \triangleq [\alpha_{n,k,s}[t]]_{\forall(n,k,s)}$ as

$$\alpha_{n,k,s}[t] = \begin{cases} 1, & \text{if } \mathsf{BS}_n \text{ serves } \mathsf{UE}_k \text{ over SC } s \text{ at TS } t, \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

In this scheme, an intra-cell orthogonal SC allocation strategy is assumed. It means that at every BS, each SC is assigned to at most one UE. This assumption is cast by the following,

$$(C1): \quad \sum_{\forall k \in \mathcal{K}} \alpha_{n,k,s}[t] \leq 1, \forall (n,s) \in (\mathcal{N} \times \mathcal{S}), \forall t. \tag{2.5}$$

Due to the limited computation resource, one assumes that UE_k can be assigned to at most $\bar{S}_k$ SCs at each TS, which yields

$$(C2): \quad \sum_{\forall s \in \mathcal{S}} \alpha_{n,k,s}[t] \leq \bar{S}_k, \forall (n,k) \in (\mathcal{N} \times \mathcal{K}), \forall t. \tag{2.6}$$

Moreover, the one-BS association constraint can be stated as

$$(C3): \sum_{\forall n \in \mathcal{N}} \| \sum_{\forall s \in \mathcal{S}} \alpha_{n,k,s}[t] \|_0 \leq 1, \forall k \in \mathcal{K}, \forall t. \tag{2.7}$$

Denote $p_{n,k,s}[t]$ as the transmit power from UE_k to BS_n over SC s, it can describe the received signal at BS_n over SC s as

$$y_{n,s}^{\mathsf{BS},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]) = \sum_{\forall(k,n')} \sqrt{\alpha_{n',k,s}[t]p_{n',k,s}[t]}\bar{h}_{n,k,s}[t]s_k + \eta_n, \tag{2.8}$$

where $\mathbf{p}[t] \triangleq [p_{n,k,s}[t]]_{\forall(n,k,s)}$, $s_k[t]$ is UE_k's data symbol, $\mathbb{E}\{|s_k[t]|^2\} = 1$, and $\eta_n \sim \mathcal{CN}(0,\sigma_n^2)$ is the additive Gaussian noise at BS_n. Subsequently, the SINR of the UE_k's data received at BS_n over SC s and TS t can be expressed as

$$\gamma_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]) = \frac{\alpha_{n,k,s}[t]p_{n,k,s}[t]h_{n,k,s}[t]}{\sum_{\forall j\neq k} h_{n,j,s}[t]\left(\sum_{\forall i}\alpha_{i,j,s}[t]p_{i,j,s}[t]\right)+\sigma_n^2}, \tag{2.9}$$

where $h_{n,k,s}[t] = |\bar{h}_{n,k,s}[t]|^2$ is the corresponding channel gain. It is noting that, in (2.9), UEs that are not assigned SC s at BS_n do not cause the interference thanks to zero values of corresponding $\alpha_{n,k,s}[t]$. Let W_{SC} be the BW of one SC, we can describe UE_k's rate (nats/s) at BS_n over SC s as

$$R_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]) = W_{\mathsf{SC}}\ln(1+\gamma_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t])). \tag{2.10}$$

Then, the total achievable transmission data rate of UE_k in TS t can be written as

$$R_k^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]) = \sum_{\forall n\in\mathcal{N}}\sum_{\forall s\in\mathcal{S}} R_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]). \tag{2.11}$$

2.2.4 Transmission over the BS-LSat Backhaul Links

This subsection describes the transmission between the BS and the LSat. For convenience, the binary variables representing the BS-LEO association are introduced as

$$\mu_{m,n}[t] = \begin{cases} 1, & \mathsf{BS}_n \text{ is served by } \mathsf{LEO}_m \text{ at TS } t, \\ 0, & \text{otherwise.} \end{cases} \tag{2.12}$$

As described above, each BS can select at most one LSat at each TS to transmit the data, which yields

$$(C4): \quad \sum_{\forall m\in\mathcal{M}} \mu_{m,n}[t] \leq 1, \forall n\in\mathcal{N}, \forall t. \tag{2.13}$$

Denote $W_{m,n}^{\mathsf{BS}}[t]$ as a variable representing the BW allocation assigned to BS_n for transferring data to LEO_m. Letting $\bar{W}_m^{\mathtt{LEO}}[t]$ be the maximum BW at LEO_m in TS t, we have

$$(C5): \sum_{\forall n\in\mathcal{N}} \mu_{m,n}[t]W_{m,n}^{\mathsf{BS}}[t] \leq W_m^{\mathtt{LEO}}[t], \forall(m,t). \tag{2.14}$$

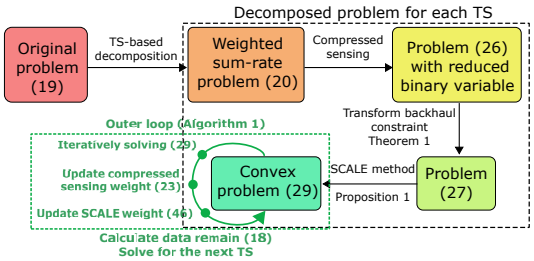


Figure 2.2: Flowchart for the centralized solution.

Furthermore, high directivity VSAT antennas having a low side-lobe leakage are assumed at BSs based on which the inter-satellite interference can be ignored in this work [45]. Then, the achievable rate of BS_n at LEO_m in TS t is expressed as

$$R_{m,n}^{\mathsf{BS},t}(\mathbf{P}[t],\mathbf{W}[t],\boldsymbol{\mu}[t]) = \mu_{m,n}[t]W_{m,n}^{\mathsf{BS}}[t]\ln\left(1+\frac{P_n[t]g_{m,n}[t]}{W_{m,n}^{\mathsf{BS}}[t]\delta_m}\right), \tag{2.15}$$

where $\mathbf{W}[t] \triangleq [W_{m,n}^{\mathsf{BS}}[t]]_{\forall(m,n)}$, $\mathbf{P}[t] \triangleq [P_n[t]]_{\forall n}$, $P_n[t]$ denotes the transmit power of BS_n, $g_{m,n}[t] = |\bar{g}_{m,n}[t]|^2$ is the channel gain of link $\mathsf{BS}_n - \mathsf{LEO}_m$ in TS t, and δ_m is the noise power density at LEO_m. Therefore, the aggregated achievable rate of BS_n at TS t is given as

$$R_n^{\mathsf{BS},t}(\mathbf{P}[t],\mathbf{W}[t],\boldsymbol{\mu}[t]) = \sum_{\forall m\in\mathcal{M}} R_{m,n}^{\mathsf{BS},t}(\mathbf{P}[t],\mathbf{W}[t],\boldsymbol{\mu}[t]). \tag{2.16}$$

Moreover, to ensure the successful forwarding at each BS, the total data rate of associated UEs needs to be less than the backhaul link capacity, which yields the following constraint

$$(C6): \sum_{\forall(k,s)} R_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]) \le R_n^{\mathsf{BS},t}(\mathbf{P}[t],\mathbf{W}[t],\boldsymbol{\mu}[t]), \forall(n,t). \tag{2.17}$$

2.2.5 Problem Formulation

In this subsection, we formulate the optimization problem wherein the min-time objective is considered. It can be seen that, different from the terrestrial networks, the integrated LSat-terrestrial networks are highly dynamic due to the high movement of the LSats. The service time of BSs corresponding to one LSat is limited by its covering duration. As a result, the goal of this study is to develop a system that minimizes the time required to transfer the entire data of all UEs to the LEO constellation. To fulfill this technical objective, we first express the remaining data demand of UE_k right after TS t as

$$d_k[t] = \max\Big(0, D_k - \frac{T_S}{\ln 2}\sum_{u=1}^{t} R_k^{\mathsf{UE},u}(\mathbf{p}[u], \boldsymbol{\alpha}[u])\Big). \tag{2.18}$$

Thanks to (2.18), the joint two-tier association and RA design minimizing offloading duration for the ISTNs can be stated by the following optimization problem,

$$\begin{aligned}
&\min_{\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\alpha},\boldsymbol{\mu},v} \quad v \quad \text{s.t. constraints } (C1)-(C6), \\
&(C7): \sum_{\forall n\in\mathcal{N}}\sum_{\forall s\in\mathcal{S}} p_{n,k,s}[t] \le p_k^{\max}, \forall k \in \mathcal{K}, \forall t \le v, \\
&(C8): P_n[t] \le P_n^{\max}, \forall n \in \mathcal{N}, \forall t \le v, \\
&(C9): \alpha_{n,k,s}[t], \mu_{m,n}[t] \in \{0,1\}, \forall (m,n,k,s), \forall t \le v, \\
&(C10): d_k[v] = 0, \forall k \in \mathcal{K},
\end{aligned} \tag{2.19}$$

where $\mathbf{p} \triangleq \{\mathbf{p}[t]\}_{\forall t}, \mathbf{P} \triangleq \{\mathbf{P}[t]\}_{\forall t}, \mathbf{W} \triangleq \{\mathbf{W}[t]\}_{\forall t}, \boldsymbol{\alpha} \triangleq \{\boldsymbol{\alpha}[t]\}_{\forall t}$ and $\boldsymbol{\mu} \triangleq \{\boldsymbol{\mu}[t]\}_{\forall t}$. Constraints $(C7)$ and $(C8)$ represent the corresponding limited transmission power of UEs and BSs, respectively. Variable v stands for the TS where all UE's data demands are completely transmitted as described by constraint $(C10)$. Regarding the goal of offloading duration minimization, v should be minimized.

In practical ISTNs, problem (2.19) aims to determine a resource management strategy to offload the entire UEs data to the core network through the satellite links during the shortest time. Particularly, the movement of LSats leads to the dynamic of the LEO-based backhaul link capacity and the network topology, which poses difficulties in resource management. The outcome by solving (2.19) can provide the solution of UE-BS and BS-LSat associations, SC assignment, BW allocation and power control at each TS to address these dynamic issues and minimize the transmission time. However, this problem consists of both continuous and binary variables, along with non-convex constraints associated with achievable rate formulas and limited link capacities,which forms it into a challenging MINLP problem. Hence, we propose the solutions to solve this problem in the next sections.

2.3 Proposed Solutions

Regarding two network management structures discussed in Section 2.2.1, this section studies two frameworks to tackle the challenging problem (2.19) efficiently. The first solution operates solely at the CNC, whereas the latter is proposed to decompose the optimization process into

sub-tasks that can be executed in parallel at the LCs. Both approaches utilize compressed sensing, relaxation, and approximation methods as follows.

2.3.1 Centralized Solution

The overall centralized solution approach is summarized in Fig. 2.2, in which the details will be given subsequently.

Time-Slot-based Decomposition

It can be seen that the objective function brings the critical sparsity challenge to problem (2.19). Naturally, to upload all user demands in the shortest time, one should maximize the amount of transmitted data for each UE in every TS. Corresponding to the channel state in each TS, there is an achievable rate region for all UEs, i.e., $\boldsymbol{C}[t]$, which is the set of all possible rate vectors, $\left(R_1^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]),...,R_K^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t])\right)$. Regarding region $\boldsymbol{C}[t]$ and exploiting the approach presented in [74, 75], the window-based problem (2.19) can be transformed to a sequence of TS-based *"weighted sum-rate (SR) maximization"* problems each of which corresponds to a TS, e.g., TS t, as given follows,

$$\max_{\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\alpha},\boldsymbol{\mu}} \sum_{\forall(k,n,s)} \omega_k[t] R_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}[t],\boldsymbol{\alpha}[t]) \quad \text{s.t. } (C1)-(C9). \tag{2.20}$$

Herein, $\omega_k[t]$ is a design weight corresponding to UE_k's remaining data in TS t.

Remark 1. *To define the weight $\omega_k[t]$, one can consider the remaining data at every UE as a queuing model with no arrivals. Then, to upload all UE's data in the shortest time, the queue lengths of all UEs should reach zero at the same time after addressing the sequence of problems* (2.20)*. Then, all queues need to be served in fairness, in other words, the queue length corresponding to all UEs, i.e., the data remaining of all UEs, needs to be kept equal over TSs. Based on* ***Proposition 1*** *given in [74], one can achieve this goal by setting the weights to the queue lengths at each TS, i.e.,*

$$\omega_k[t] = d_k[t], \quad \forall(k,t). \tag{2.21}$$

Therefore, to tackle problem (2.19), one can sequentially solve problem (2.20) and update $\omega_k[t]$ after every TS until all UEs' data is uploaded to the LEO constellation successfully. Regarding TS t, it can be seen that problem (2.20) is still a non-convex MINLP owing to

the coupling between binary and continuous variables, as well as the sparsity terms and the non-convex functions relating to SNR and SINR.

Compressed-Sensing-based Relaxation

Regarding the relation between BS-UE association, SC assignment decision, and the continuous variables in TS t, one can see that:

- If UE_k is served by BS_n over SC s, the transmit power $p_{n,k,s}[t]$ is naturally positive, and vice versa, if UE_k is not served by BS_n over SC s, it should be $p_{n,k,s}[t] = 0$.
- Similarly, for the LEO-BS association, if BS_n is served by LEO_m, $W^{\mathsf{BS}}_{m,n}[t]$ is naturally positive, and vice versa, $W^{\mathsf{BS}}_{m,n}[t] = 0$ if BS_n is not served by LEO_m.

These observations yield the following results, $\alpha_{n,k,s}[t] = \|p_{n,k,s}[t]\|_0$, $\mu_{m,n}[t] = \|W^{\mathsf{BS}}_{m,n}[t]\|_0$, and

$$\|\sum_{\forall s} \alpha_{n,k,s}[t]\|_0 = \|\sum_{\forall s} p_{n,k,s}[t]\|_0 = \|p_{n,k}[t]\|_0, \tag{2.22}$$

for all (m,n,k,s,t), where $p_{n,k}[t] = \sum_{\forall s} p_{n,k,s}[t]$. Hence, $\boldsymbol{\alpha}$ and $\boldsymbol{\mu}$ can be represented by sparsity form of $\mathbf{p}$ and $\mathbf{W}$ efficiently. Subsequently, to address these sparsity terms, we tend to utilize an iterative compressed-sensing-based algorithm, so-called reweighted ℓ_1 minimization method [76, 77], which consists of the following process in each i^{th} iteration: (i) replacing the sparsity term $\|x\|_0$ by $1/\left(|x^{(i)}|^2+\epsilon\right)$ where ϵ is a small positive number; (ii) solving the approximate problem; (iii) updating the fixed point for the next iteration based on the recent achieved optimal solution. Employing this approach, ℓ_0-norm components in (2.22) can be approximated as

$$\|p_{n,k,s}[t]\|_0 = \zeta^{(i)}_{n,k,s}[t] p_{n,k,s}[t], \tag{2.22a}$$

$$\|p_{n,k}[t]\|_0 = \xi^{(i)}_{n,k}[t] p_{n,k}[t], \tag{2.22b}$$

$$\|W^{\mathsf{BS}}_{m,n}[t]\|_0 = \chi^{(i)}_{m,n}[t] W^{\mathsf{BS}}_{m,n}[t], \tag{2.22c}$$

for all (m,n,k,s,t), where $\zeta^{(i)}_{n,k,s}[t]$, $\xi^{(i)}_{n,k}[t]$ and $\chi^{(i)}_{m,n}[t]$ are the weights, which are updated based on the solution of $\mathbf{p}[t]$ and $\mathbf{W}[t]$ after the $(i-1)^{th}$ iteration, respectively as

$$\zeta^{(i)}_{n,k,s}[t] = \frac{1}{p^{(i-1)}_{n,k,s}[t] + \epsilon}, \quad \xi^{(i)}_{n,k}[t] = \frac{1}{p^{(i-1)}_{n,k}[t] + \epsilon}, \quad \chi^{(i)}_{m,n}[t] = \frac{1}{W^{\mathsf{BS},(i-1)}_{m,n}[t] + \epsilon}, \tag{2.23}$$

for all (m, n, k, s). Then, $(C1) - (C6)$ can be rewritten as

$$(\tilde{C}1): \sum_{\forall k \in \mathcal{K}} \zeta^{(i)}_{n,k,s}[t] p_{n,k,s}[t] \leq 1, \forall (n,s), \tag{2.23a}$$

$$(\tilde{C}2): \sum_{\forall s \in \mathcal{S}} \zeta^{(i)}_{n,k,s}[t] p_{n,k,s}[t] \leq \bar{S}_k, \forall (n,k), \tag{2.23b}$$

$$(\tilde{C}3): \sum_{\forall n \in \mathcal{N}} \xi^{(i)}_{n,k}[t] \sum_{\forall s \in \mathcal{S}} p_{n,k,s}[t] \leq 1, \forall k, \tag{2.23c}$$

$$(\tilde{C}4): \sum_{\forall m \in \mathcal{M}_t} \chi^{(i)}_{m,n}[t] W^{\mathsf{BS}}_{m,n}[t] \leq 1, \forall n, \tag{2.23d}$$

$$(\tilde{C}5): \sum_{\forall n \in \mathcal{N}} W^{\mathsf{BS}}_{m,n}[t] \leq W^{\mathtt{LEO}}_{m}[t], \quad \forall m \in \mathcal{M}_t, \tag{2.23e}$$

$$(\tilde{C}6): \sum_{\forall (k,s)} R^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}[t]) \leq R^{\mathsf{BS},t}_{n}(\mathbf{P}[t], \mathbf{W}[t]), \forall n, \tag{2.23f}$$

where $\boldsymbol{\alpha}[t]$ and $\boldsymbol{\mu}[t]$ in $(C1) - (C6)$ are omitted.

Thanks to the compressed-sensing approach [76], problem (2.20) can be addressed by iteratively updating $\zeta^{(i)}_{n,k,s}[t]$ and $\chi^{(i)}_{m,n}[t]$, and solving the following problem

$$\begin{aligned} \max_{\mathbf{p},\mathbf{P},\mathbf{W}} & \sum_{\forall (k,n,s)} \omega_k[t] R^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}[t]) \\ & \text{s.t. constraints } (\tilde{C}1) - (\tilde{C}6), (C7), (C8). \end{aligned} \tag{2.24}$$

Note that $R^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}[t])$ and $R^{\mathsf{BS},t}_{n}(\mathbf{P}[t], \mathbf{W}[t])$ are the rate functions with omitting the arguments of $\boldsymbol{\alpha}[t]$ and $\boldsymbol{\mu}[t]$.

Convexifying Constraint $(\tilde{C}6)$

It can be seen that problem (2.24) is still non-convex due to the non-convexity of $(\tilde{C}6)$. To address $(\tilde{C}6)$, we consider the following theorem.

Theorem 1. *The solution for* (2.24) *can be obtained by solving the following problem,*

$$\begin{aligned} \max_{\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\lambda}^{\mathsf{UE}},\boldsymbol{\lambda}^{\mathsf{BS}}} & \sum_{\forall k \in \mathcal{K}} \sum_{\forall n \in \mathcal{N}} \omega_k[t] \lambda^{\mathsf{UE}}_{n,k}[t] \\ & \text{s.t. } \textit{constraints } (\tilde{C}1) - (\tilde{C}5), (C7), (C8), \\ & (C11): \sum_{\forall k \in \mathcal{K}} \lambda^{\mathsf{UE}}_{n,k}[t] \leq \lambda^{\mathsf{BS}}_{n}[t], \forall n, \\ & (C12): \lambda^{\mathsf{UE}}_{n,k}[t] \leq \sum_{\forall s \in \mathcal{S}} R^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}[t]), \forall (n,k), \\ & (C13): \lambda^{\mathsf{BS}}_{n}[t] \leq \sum_{\forall m \in \mathcal{M}_t} R^{\mathsf{BS},t}_{m,n}(\mathbf{P}[t], \mathbf{W}[t]), \forall n, \end{aligned} \tag{2.25}$$

and calibrating $\mathbf{p}$ *so that* $(C12)$ *holds with equality. Herein,* $\boldsymbol{\lambda}^{\mathsf{UE}} \triangleq \{\lambda_{n,k}^{\mathsf{UE}}[t]\}$ *and* $\boldsymbol{\lambda}^{\mathsf{BS}} \triangleq \{\lambda_{n}^{\mathsf{BS}}[t]\}$ *are new variables, which are introduced as a lower bound of the UE and BS transmission rate functions, respectively. Additionally, the* $\mathbf{p}$*-calibrating process can be done by solving the following problem,*

$$\min_{\mathbf{p}} \sum_{\forall(n,k,s)} p_{n,k,s}[t] \text{ s.t. } (\tilde{C}1) - (\tilde{C}3), (C7), (C12), \tag{2.26}$$

which is well addressed in the literature [78, 79].

Proof: It can be easy to see that once $(C12)$ holds with equality, problems (2.24) and (2.25) are equivalent since the optimal solution of this problem is a feasible solution of the other and vice versa. For instance, let $(\mathbf{p}^*, \mathbf{P}^*, \mathbf{W}^*)$ be the optimal solution of problem (2.24). Then, denoting

$$\lambda_{n,k}^{\mathsf{UE}\prime}[t] = \sum_{\forall s \in \mathcal{S}} R_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}^*[t]), \quad \forall (n,k) \in (\mathcal{N} \times \mathcal{K}), \tag{2.26a}$$

$$\lambda_{n}^{\mathsf{BS}\prime}[t] = \sum_{\forall m \in \mathcal{M}_t} R_{m,n}^{\mathsf{BS},t}(\mathbf{P}^*[t], \mathbf{W}^*[t]), \quad \forall n \in \mathcal{N}, \tag{2.26b}$$

yields $(\mathbf{p}^*, \mathbf{P}^*, \mathbf{W}^*, \boldsymbol{\lambda}^{\mathsf{UE}\prime}, \boldsymbol{\lambda}^{\mathsf{BS}\prime})$ as a feasible solution of (2.25). On another hand, one assume $(\mathbf{p}', \mathbf{P}', \mathbf{W}', \boldsymbol{\lambda}^{\mathsf{UE}\prime}, \boldsymbol{\lambda}^{\mathsf{BS}\prime})$ be the optimal solution of problem (2.25). Solving (2.26) to obtain $\mathbf{p}''$, we can easily prove that $\lambda_{n,k}^{\mathsf{UE}\prime}[t] = \sum_{\forall s \in \mathcal{S}} R_{n,k,s}^{\mathsf{UE},t}(\mathbf{p}''[t]), \forall (n,k)$. Interestingly, $(\mathbf{p}'', \mathbf{P}', \mathbf{W}', \boldsymbol{\lambda}^{\mathsf{UE}\prime}, \boldsymbol{\lambda}^{\mathsf{BS}\prime})$ is also an optimal solution of problem (2.25) since this set satisfies all constraints and returns the same objective as $(\mathbf{p}', \mathbf{P}', \mathbf{W}', \boldsymbol{\lambda}^{\mathsf{UE}\prime}, \boldsymbol{\lambda}^{\mathsf{BS}\prime})$. Furthermore, $(\mathbf{p}'', \mathbf{P}', \mathbf{W}')$ also satisfies all constraints of (2.24), and it should be a feasible solution of this problem. ■

Based on Theorem 1, we can tackle problem (2.24) by solving problem (2.25) and calibrating $\mathbf{p}$ by addressing (2.26) if $(C12)$ does not hold with equality. In the following, we will discuss how to deal with problem (2.25) efficiently.

Successive Convex Approximation for Low-complExity (SCALE) Method

It can be seen, (2.25) is non-convex due to constraint $(C12)$. To address this issue, one exploits the SCALE approach - a successive convex approximation method introduced in [78,80] - to transform the UE rate function into a concave form as follows.

Proposition 1. *The equivalent problem of* (2.25) *at iteration* $(i+1)$*, TS* t *can be formulated*

Algorithm 1 PROPOSED CENTRALIZED ALGORITHM

1: Set $t = 1$.
2: **repeat**
3: **Initialization:** Set $i = 0$, and generate an initial starting point $\left(\bar{\mathbf{p}}^{(0)}[t], (\mathbf{W}[t])^{(0)}\right)$.
4: **repeat**
5: Calculate SCALE weights $a^{(i)}_{n,k,s}, b^{(i)}_{n,k,s}$, compressed-sensing weights $\zeta^{(i)}_{n,k,s}, \xi^{(i)}_{n,k}, \chi^{(i)}_{m,n}$, and $\mathbf{p}^{(i)}[t] = \exp(\bar{\mathbf{p}}^{(i)}[t])$.
6: Solve (2.27) to obtain $\left(\bar{\mathbf{p}}^{\star}[t], \mathbf{P}^{\star}[t], \mathbf{W}^{\star}[t]\right)$.
7: Update $\left(\bar{\mathbf{p}}^{(i+1)[t]}, \mathbf{W}^{(i+1)}[t]\right) := \left(\bar{\mathbf{p}}^{\star}[t], \mathbf{W}^{\star}[t]\right)$.
8: Set $i = i + 1$.
9: **until** Convergence
10: Calculate $d_k[t], \forall k$ based on (2.18).
11: Set $t = t + 1$.
12: **until** $d_k[t] = 0, \forall k$.
13: Recovery association variables $\boldsymbol{\alpha}$ and $\boldsymbol{\mu}$ by (2.32).
14: **Output:** The solution $\left(\mathbf{p}^{\star}, \mathbf{P}^{\star}, \mathbf{W}^{\star}, \boldsymbol{\alpha}^{\star}, \boldsymbol{\mu}^{\star}\right)$.

as

$$\max_{\bar{\mathbf{p}},\mathbf{P},\mathbf{W},\boldsymbol{\lambda}^{\mathsf{UE}},\boldsymbol{\lambda}^{\mathsf{BS}}} \quad \sum_{\forall k\in\mathcal{K}}\sum_{\forall n\in\mathcal{N}} \omega_k[t]\lambda^{\mathsf{UE}}_{n,k}[t] \tag{2.27}$$

$$\text{s.t. } \textit{constraints } (\tilde{C}4), (\tilde{C}5), (C8), (C11), (C13),$$

$$(\bar{C}1): \sum_{\forall k\in\mathcal{K}} \zeta^{(i)}_{n,k,s}[t]\exp(\bar{p}_{n,k,s}[t]) \le 1, \forall(n,s),$$

$$(\bar{C}2): \sum_{\forall s\in\mathcal{S}} \zeta^{(i)}_{n,k,s}[t]\exp(\bar{p}_{n,k,s}[t]) \le \bar{S}_k, \forall(n,k),$$

$$(\bar{C}3): \sum_{\forall n}\xi^{(i)}_{n,k}[t]\sum_{\forall s}\exp(\bar{p}_{n,k,s}[t]) \le 1, \forall k,$$

$$(\bar{C}7): \sum_{\forall n\in\mathcal{N}}\sum_{\forall s\in\mathcal{S}}\exp(\bar{p}_{n,k,s}[t]) \le p^{\max}_k, \forall k,$$

$$(\tilde{C}12): \lambda^{\mathsf{UE}}_{n,k}[t] \le \sum_{\forall s\in\mathcal{S}} \tilde{R}^{\mathsf{UE},t}_{n,k,s}(\bar{\mathbf{p}}[t]), \forall(n,k),$$

where $\bar{\mathbf{p}} \triangleq [\bar{\mathbf{p}}[t]]_{\forall t}$, $\bar{\mathbf{p}}[t] \triangleq [\bar{p}_{n,k,s}[t]]_{\forall(n,k,s)}$, $\bar{p}_{n,k,s}[t]$ *is a new variable which is used instead of* $p_{n,k,s}[t]$, *and*

$$\tilde{R}^{\mathsf{UE},t}_{n,k,s}(\bar{\mathbf{p}}[t]) = W_{\mathsf{SC}}\Big(a^{(i)}_{n,k,s}[t]\Psi^{t}_{n,k,s}(\bar{\mathbf{p}}[t]) + b^{(i)}_{n,k,s}[t]\Big), \tag{2.28}$$

in which $a^{(i)}_{n,k,s}[t]$ *and* $b^{(i)}_{n,k,s}[t]$ *are the updated coefficients at each iteration, and*

$$\Psi^{t}_{n,k,s}(\bar{\mathbf{p}}[t]) = \ln(h_{n,k,s}[t]) + \bar{p}_{n,k,s}[t] - \ln\Big(\sum_{\forall n'}\sum_{\forall k'\in\mathcal{K}\setminus k} h_{n,k',s}[t]e^{\bar{p}_{n',k',s}[t]} + \sigma^2_n\Big). \tag{2.29}$$

Proof: This part provides the proof of Proposition 1, wherein SCALE method [78] is used to convexify $(C12)$. Specifically, at iteration i, the UE rate function can be lower

bounded as

$$R^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}[t]) \geq W_{\mathsf{SC}}(a^{(i)}_{n,k,s}[t]\ln(\mathtt{SINR}^{t}_{n,k,s}) + b^{(i)}_{n,k,s}[t]), \tag{2.30}$$

where the coefficients $a^{(i)}_{n,k,s}[t]$ and $b^{(i)}_{n,k,s}[t]$ are defined as

$$a^{(i)}_{n,k,s}[t] = \frac{\mathtt{SINR}^{t,(i-1)}_{n,k,s}}{1+\mathtt{SINR}^{t,(i-1)}_{n,k,s}}, \quad b^{(i)}_{n,k,s}[t] = \log\left(1+\mathtt{SINR}^{t,(i-1)}_{n,k,s}\right) - a^{(i)}_{n,k,s}[t],$$

and $\mathtt{SINR}^{t,(i)}_{n,k,s} = \gamma^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}^{(i)}[t])$ is the SINR value at iteration i.

Subsequently, to convexify the function $\ln(\mathtt{SINR}^{t}_{n,k,s})$ in (2.30), one introduces new variables $\{\bar{p}_{n,k,s}[t]\}$ and replace $p_{n,k,s}[t]$ with $p_{n,k,s}[t] = \exp(\bar{p}_{n,k,s}[t])$, $\forall(n,k,s)$. Based on that, (2.30) can be rewritten as

$$R^{\mathsf{UE},t}_{n,k,s}(\mathbf{p}[t]) \geq \tilde{R}^{\mathsf{UE},t}_{n,k,s}(\bar{\mathbf{p}}[t]) := W_{\mathsf{SC}}(a^{(i)}_{n,k,s}[t]\Psi^{t}_{n,k,s}(\bar{\mathbf{p}}[t]) + b^{(i)}_{n,k,s}[t]). \tag{2.31}$$

It is worth noting that the log-sum-exp term in $\Psi^{t}_{n,k,s}(\bar{\mathbf{p}}[t])$ is convex, thus $\tilde{R}^{\mathsf{UE},t}_{n,k,s}(\bar{\mathbf{p}}[t])$ is naturally concave. Exploiting this approach, constraint $(C12)$ can be approximated as $(\tilde{C}12)$ in (2.31). In addition, regarding $\bar{p}_{n,k,s}[t]$, constraints $(\tilde{C}1)-(\tilde{C}3)$ and $(\tilde{C}7)$ are rewritten as $(\bar{C}1)-(\bar{C}3)$ and $(\bar{C}7)$ in (2.31) due to the transformation $p_{n,k,s}[t] = \exp(\bar{p}_{n,k,s}[t])$, respectively. It is worth noting that the replacing constraints $(\bar{C}1)-(\bar{C}3)$ and $(\bar{C}7)$ are convex thanks to the sum-exp forms. ■

Then, by exploiting the compressed-sensing and SCALE methods, as well as the valuable results addressed in the TS-based decomposition, Theorem 1, and Proposition 1, the solution for problem (2.19) can be obtained by iteratively solving problem (2.27) and updating the corresponding weights in every TS sequentially until all UEs' data have been delivered to LEO constellation. The proposed centralized algorithm is summarized in Algorithm 1. Regarding the compressed-sensing approach, the binary solution $\boldsymbol{\alpha}$ and $\boldsymbol{\mu}$ can be returned as

$$\alpha_{n,k,s}[t] = \begin{cases} 1, & \text{if } p_{n,k,s}[t] \geq \epsilon, \\ 0, & \text{otherwise,} \end{cases}, \quad \mu_{m,n}[t] = \begin{cases} 1, & \text{if } W^{\mathsf{BS}}_{m,n}[t] \geq \epsilon, \\ 0, & \text{otherwise.} \end{cases} \tag{2.32}$$

2.3.2 Decentralized Solution

Implementing the centralized solution may pose a significant challenge, as it requires all network information to be available at the CNC. This requirement may lead to execution delays

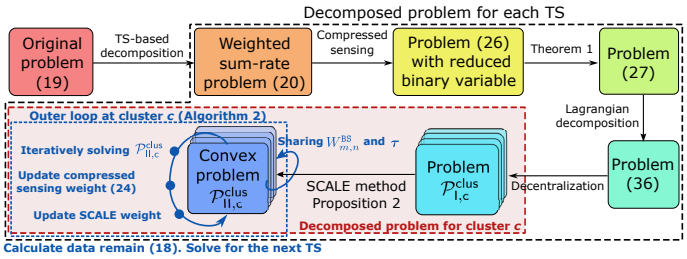


Figure 2.3: Flowchart for the decentralized solution.

and an overload of controller links, arising from the aggregated traffic needed for exchanging network information. In addition, centralized execution also leads to computation overload and difficulties in scaling up the algorithm for large networks. Therefore, to address these issues – including overload at the central network controller, the large amount of exchanged network information, execution delays, and scalability concerns – a decentralized algorithm is proposed in this section. Subsequently, we decompose the centralized optimization process into sub-problems each of which can be solved separately at the corresponding LC. The overall decentralized solution approach is summarized in Fig. 2.3.

Lagrangian-based Decomposition

Analyzing problem (2.25), one can see that the variable sets corresponding to clusters and NTN are coupled through constraint $(\tilde{C}5)$. To address this coupling, we propose to employ the Lagrangian relaxation method with the multiplexer $\boldsymbol{\tau} \triangleq [\tau_m[t]]_{\forall m,t>0}$ as follows. The Lagrangian function of (2.25) is given as

$$L(\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\lambda}^{\mathsf{UE}},\boldsymbol{\lambda}^{\mathsf{BS}},\boldsymbol{\tau}) = \sum_{\forall(k,n)} \omega_k[t]\lambda^{\mathsf{UE}}_{n,k}[t] + \sum_{\forall m\in\mathcal{M}_t} \tau_m[t](\bar{W}^{\mathsf{LEO}}_m[t] - \sum_{\forall n\in\mathcal{N}} W^{\mathsf{BS}}_{m,n}[t]). \tag{2.33}$$

Let Ω represent the tunnel of all variables, i.e., $\Omega \triangleq (\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\lambda}^{\mathsf{UE}},\boldsymbol{\lambda}^{\mathsf{BS}})$, $\mathcal{F}$ denote the feasible set of (2.25), i.e., $\mathcal{F} \triangleq \{\Omega | \Omega \text{ satisfies } (\tilde{C}1)-(\tilde{C}4),(C7),(C8),(C11)-(C13)\}$. The dual problem can be written as

$$\min_{\boldsymbol{\tau}} \; g_{\mathsf{Lag}}(\boldsymbol{\tau}) \text{ s. t. } \tau_m[t] \geq 0 \; \forall(m,t), \tag{2.34}$$

where $g_{\mathsf{Lag}}(\boldsymbol{\tau})$ is denoted as the dual function corresponding to (27) which is expressed as

$$g_{\mathsf{Lag}}(\boldsymbol{\tau}) = \sup_{\Omega\in\mathcal{F}} L(\Omega,\boldsymbol{\tau}) = \sup_{\Omega\in\mathcal{F}} \Big(\sum_{\forall(k,n)} \omega_k[t]\lambda^{\mathsf{UE}}_{n,k}[t] + \sum_{\forall m\in\mathcal{M}_t} \tau_m[t](W^{\mathsf{LEO}}_m[t] - \sum_{\forall n\in\mathcal{N}} W^{\mathsf{BS}}_{m,n}[t]) \Big).$$

For given $\boldsymbol{\tau}$ and $\{W_m^{\mathsf{LEO}}\}_{\forall m\in\mathcal{M}_t}$, the right-hand-side problem of (2.35) can be rewritten in a more trackable form as

$$\begin{aligned}\max_{\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\lambda}^{\mathsf{UE}},\boldsymbol{\lambda}^{\mathsf{BS}}} & \sum_{\forall(k,N)}\omega_k[t]\lambda_{n,k}^{\mathsf{UE}}[t]-\sum_{\forall m\in\mathcal{M}_t}\tau_m[t]\sum_{\forall n\in\mathcal{N}}W_{m,n}^{\mathsf{BS}}[t]\\ \text{s.t.}\quad & (\tilde{C}1)-(\tilde{C}4),(C7),(C8),(C11)-(C13).\end{aligned}\tag{2.35}$$

Herein, for a decomposition purpose, we assume that the inter-cluster interference is negligible due to the sufficiently long distance between clusters. Based on this assumption, the optimization for the TN layer can be executed at each cluster. As a result, problem (2.35) can be further decomposed into sub-problems which are solved separately at the local control units corresponding to the clusters. In particular, the optimization sub-problem for cluster c is formulated as

$$\begin{aligned}\left(\mathcal{P}_{\mathsf{l,c}}^{\mathsf{clus}}\right)\ \max_{\mathbf{p}_c,\mathbf{P}_c,\mathbf{W}_c,\boldsymbol{\lambda}_c^{\mathsf{UE}},\boldsymbol{\lambda}_c^{\mathsf{BS}}} & \sum_{\forall(k,n)}\omega_k[t]\lambda_{n,k}^{\mathsf{UE}}[t]-\sum_{\forall m\in\mathcal{M}_t}\tau_m[t]\sum_{\forall n\in\mathcal{N}_c}W_{m,n}^{\mathsf{BS}}[t]\\ \text{s.t.}\quad & (\tilde{C}1)_c-(\tilde{C}4)_c,(C7)_c,(C8)_c,(C11)_c,(C13)_c,\\ & (C14)_c:\lambda_{n,k}^{\mathsf{UE},c}[t]\le\sum_{\forall s\in\mathcal{S}}R_{n,k,s}^{\mathsf{UE},c,t}(\mathbf{p}_c[t]),\forall(n,k),\end{aligned}\tag{2.36}$$

where the low suffix $(*)_c$ is added to notations (CX) and variables to indicate that constraint (CX) and corresponding variables are applied for cluster c. In constraint $(C14)_c$, $R_{n,k,s}^{\mathsf{UE},c,t}(\mathbf{p}_c[t])=W_{\mathsf{SC}}\log_2(1+\gamma_{n,k,s}^{\mathsf{UE},c,t}(\mathbf{p}_c[t]))$, and

$$\gamma_{n,k,s}^{\mathsf{UE},c,t}(\mathbf{p}_c[t])=\frac{p_{n,k,s}[t]h_{n,k,s}[t]}{\sum_{\forall n'}\sum_{\forall k'\in\mathcal{K}_c\setminus\{k\}}h_{n,k',s}[t]p_{n',k',s}[t]+I_{n,s}^{c,(i)}+\sigma_n^2},\tag{2.37}$$

where $I_{n,s}^{c,(i)}=\sum_{\forall n'}\sum_{\forall j'\in\mathcal{K}\setminus\mathcal{K}_c}h_{n,j',s}[t]p_{n',j',s}^{(i)}[t]$ is the inter-cluster interference at BS n in cluster c over SC s.

Decentralized Algorithm

It can be seen that $(C14)_c$ is non-convex. The following proposition is to address this issue.

Proposition 2. *Problem* $\left(\mathcal{P}_{\mathsf{l,c}}^{\mathsf{clus}}\right)$ *can be transformed into the following convex problem,*

$$\left(\mathcal{P}_{\mathsf{ll,c}}^{\mathsf{clus}}\right)\max_{\mathbf{p}_c,\mathbf{P}_c,\mathbf{W}_c,\boldsymbol{\lambda}_c^{\mathsf{UE}},\boldsymbol{\lambda}_c^{\mathsf{BS}}}\sum_{\forall(k,n)}\omega_k[t]\lambda_{n,k}^{\mathsf{UE}}[t]-\sum_{\forall m\in\mathcal{M}_t}\tau_m[t]\sum_{\forall n\in\mathcal{N}_c}W_{m,n}^{\mathsf{BS}}[t]$$

Algorithm 2 PROPOSED DECENTRALIZED ALGORITHM

Set $t = 1$.
repeat
  Initialization: Set $\ell = 1$. Generate an initial Lagrangian multiplier $\boldsymbol{\tau}$.
  repeat
    for Each cluster c **do**
      Set $i = 0$ and select $\left(\bar{\mathbf{p}}_c^{(0)}[t], \mathbf{W}_c^{(0)}[t]\right)$.
      repeat
        Calculate weights $a_{n,k,s}^{(i)}, b_{n,k,s}^{(i)}, \zeta_{n,k,s}^{(i)}, \xi_{n,k}^{(i)}, \chi_{m,n}^{(i)}$ for cluster c and $\mathbf{p}_c^{(i)}[t] = \exp(\bar{\mathbf{p}}_c^{(i)}[t])$.
        Solve $\mathcal{P}_{\mathsf{II,c}}^{\mathsf{clus}}$ to obtain $\left(\bar{\mathbf{p}}_c^{\star}[t], \mathbf{P}_c^{\star}[t], \mathbf{W}_c^{\star}[t]\right)$.
        Update $\left(\bar{\mathbf{p}}_c^{(i+1)}[t], \mathbf{W}_c^{(i+1)}[t]\right) := \left(\bar{\mathbf{p}}_c^{\star}[t], \mathbf{W}_c^{\star}[t]\right)$.
        Set $i = i + 1$.
      until Convergence
    end for
    Update $\boldsymbol{\tau}$ by (2.40).
    Set $\ell = \ell + 1$.
  until Convergence
  Calculate $d_k[t], \forall k$ based on (2.18).
  Set $t = t + 1$.
until $d_k[t] = 0, \forall k$.
Recovery association variables $\boldsymbol{\alpha}$ and $\boldsymbol{\mu}$ by (2.32).
Output: The solution $\left(\mathbf{p}^{\star}, \mathbf{P}^{\star}, \mathbf{W}^{\star}, \boldsymbol{\alpha}^{\star}, \boldsymbol{\mu}^{\star}\right)$.

$$\begin{aligned} &\text{s.t. } (\bar{C}1)_c - (\bar{C}3)_c, (\tilde{C}4)_c, (\bar{C}7)_c, (C8)_c, (C11)_c, (C13)_c, \\ &(\bar{C}14)_c : \lambda_{n,k}^{\mathsf{UE}}[t] \leq W_{\mathsf{SC}} \sum_{\forall s \in \mathcal{S}} a_{n,k,s}^{(i)} \Psi_{n,k,s}^{c,t}(\bar{\mathbf{p}}_c[t]) + b_{n,k,s}^{(i)}, \end{aligned} \tag{2.38}$$

where $\Psi_{n,k,s}^{c,t}(\bar{\mathbf{p}}_c[t])$ *is defined as*

$$\begin{aligned} \Psi_{n,k,s}^{c,t}(\bar{\mathbf{p}}_c[t]) &\triangleq \ln(h_{n,k,s}[t]) + \bar{p}_{n,k,s}[t] \\ &- \ln\Big(\sum_{\forall n'} \sum_{\forall k' \in \mathcal{K}_c \setminus k} h_{n,k',s}[t] \exp(\bar{p}_{n',k',s}[t]) + I_{n,s}^{c,(i)} + \sigma_n^2\Big). \end{aligned} \tag{2.39}$$

Proof: Using the similar approach described in Sub-Section 2.3.1, $(C14)$ can be approximately convexified by utilizing function $a\ln(x) + b$. Then, $(C14)_c$ can be rewritten as $(\bar{C}14)_c$ with specific values of $a_{n,k,s}^{(i)}$, $\Psi_{n,k,s}^{c,t}(\bar{\mathbf{p}}_c[t])$, and $b_{n,k,s}^{(i)}$. Additionally, exploiting the same approach introducing the new variables $\bar{\mathbf{p}}_c[t] \triangleq [\bar{p}_{n,k,s}[t]]_{\forall (n,k,s) \in (\mathcal{N}_c \times \mathcal{K}_c \times \mathcal{S})}$, one can replace $(C1)_c - (C3)_c$ and $(C7)_c$ by $(\bar{C}1)_c - (\bar{C}3)_c$ and $(\bar{C}7)_c$, respectively. ■

Remark 2. *Due to the relatively long distance separating the clusters, the inter-cluster interference is assumed to be very minimal. Therefore, our decentralized design treats these inter-cluster interference components as fixed colored noise in each iteration. However, these terms are still re-estimated after each iteration to evaluate the SINR at the users precisely.*

In sub-problem $\left(\mathcal{P}_{\mathsf{II,c}}^{\mathsf{clus}}\right)$, the second term of the objective function, given in (2.38), can be considered as the balancing term corresponding to the LEO BW allocation. Herein, the parameter $\boldsymbol{\tau}$ needs to be adjusted to drive the solutions of all problems $\left(\mathcal{P}_{\mathsf{II,c}}^{\mathsf{clus}}\right)$'s to meet constraint $(\tilde{C}5)$. One efficient conventional approach for this purpose is the use of the subgradient descent method [81]. In this method, the Lagrangian multiplier $\boldsymbol{\tau}$ can be updated as

$$\tau_m^{(\ell+1)}[t] = \tau_m^{(\ell)}[t] - \vartheta_\ell (W_m^{\mathsf{LEO}} - \sum_{\forall n \in \mathcal{N}} W_{m,n}^{\mathsf{BS}}), \tag{2.40}$$

where ϑ_ℓ is the step size. Moreover, this information will be disseminated to all LCs via a mechanism requiring limited signal transmission. The proposed decentralized algorithm is summarized in Algorithm 2. The convergence of this algorithm is discussed in the following.

Proposition 3. *Regarding the dual problem, one have:*

(i) The dual function g_{Lag} is a convex function.

(ii) Its sub-gradient at $\tau_m[t]$ is $W_m^{\mathsf{LEO}} - \sum_{\forall n \in \mathcal{N}} W_{m,n}^{\mathsf{BS}}$.

(iii) The sub-gradient updating method described in (2.40) is guaranteed to converge to the optimal solution of (2.34) if problem (2.35) is solved optimally and ϑ_ℓ's are set according to "the square summable but not summable step size rule" [81].

(iv) Once, the conditions given in (iii) are satisfied, the convergence rate of the sub-gradient method is $O(1/\sqrt{\ell})$.

Proof: The proposition can be proved briefly based on the results described in [81,82] as follows:

(i) It can be seen that $g_{\mathsf{Lag}}(\boldsymbol{\tau})$ is the pointwise supremum of a family of affine functions, $g_{\mathsf{Lag}}(\boldsymbol{\tau})$ is convex [82].

(ii) The choice of $W_m^{\mathsf{LEO}} - \sum_{\forall n \in \mathcal{N}} W_{m,n}^{\mathsf{BS}}$ as the sub-gradient for $\tau_m[t]$ is justified by the fact that $\tau_m[t]$ is the Lagrangian multiplier associated with the constraint $(\tilde{C}5)$.

(iii) Due to the results given in Section 2 of [81], the above sub-gradient method is guaranteed the convergence if the step sizes are set according to *"the square summable but not summable step size rule"* such as $\vartheta_\ell = 1/\ell$.

(iv) Using the sub-gradient descent method to solve the dual problem; hence, its convergence rate is $O(1/\sqrt{\ell})$ [81]. ■

Moreover, the duality gap is discussed in the following remark and proposition.

Remark 3. *The strong duality does not hold for the dual problem* (2.34) *and the corresponding primal problem* (2.25) *since* (2.25) *is non-convex. It is important to highlight that the use of Lagrangian decomposition enables the parallel optimization execution across local clusters, helping to overcome the challenges in the centralized solution implementation related to the overload of controller links, execution delays, and scalability.*

Proposition 4. *Let* $(\mathcal{P}_{\mathsf{II}})$ *and* $(\mathcal{D}_{\mathsf{II}})$ *be the aggregation of all* $(\mathcal{P}^{\mathsf{clus}}_{\mathsf{II,c}})$*'s and its dual problem, respectively. If the inter-cluster interference is neglectable, i.e.,* $I^c_{n,s} \approx 0$*, the strong duality holds for* $(\mathcal{P}_{\mathsf{II}})$ *and problem* (2.27) *in the centralized algorithm.*

Proof: Problem $(\mathcal{P}_{\mathsf{II}})$ formed by aggregating all cluster sub-problems $(\mathcal{P}^{\mathsf{clus}}_{\mathsf{II,c}})$, $\forall c$, can be formulated as

$$\begin{aligned}(\mathcal{P}_{\mathsf{II}}) \max_{\substack{\{\mathbf{p}_c,\mathbf{P}_c,\mathbf{W}_c,\\ \boldsymbol{\lambda}^{\mathsf{UE}}_c,\boldsymbol{\lambda}^{\mathsf{BS}}_c\}_{\forall c}}} & \sum_{\forall c}\Big(\sum_{\substack{\forall n\in\mathcal{N}_c\\ \forall k\in\mathcal{K}_c}} \omega_k[t]\lambda^{\mathsf{UE}}_{n,k}[t] - \sum_{\forall m\in\mathcal{M}_t}\tau_m[t]\sum_{\forall n\in\mathcal{N}_c} W^{\mathsf{BS}}_{m,n}[t]\Big)\\ \text{s.t.} & (\bar{C}1)_c - (\bar{C}3)_c, (\tilde{C}4)_c, (\bar{C}7)_c, (C8)_c, (C11)_c, (C13)_c, (\bar{C}14)_c, \forall c.\end{aligned}$$

By integrating the variables and index sets, the aggregation set of constraints $(\bar{C}1)_c - (\bar{C}3)_c$, $(\tilde{C}4)_c, (\bar{C}7)_c, (C8)_c, (C11)_c, (C13)_c, \forall c$ are equivalent to $(\bar{C}1) - (\bar{C}3), (\tilde{C}4)$, $(\bar{C}7)$, $(C8), (C11)$, $(C13)$, respectively. In addition, if we suppose that the inter-cluster interference is very minimal $I^c_{n,s} \approx 0$, the approximated rate function $\Psi^{c,t}_{n,k,s}(\mathbf{p}_c[t]) = \Psi^t_{n,k,s}(\mathbf{p}[t])$. Hence, the aggregation set of constraints $(\bar{C}14)_c, \forall c$ is equivalent to constraint $(\bar{C}12)$. Subsequently, problem $(\mathcal{P}_{\mathsf{II}})$ can be rewritten as

$$\begin{aligned}(\mathcal{P}_{\mathsf{II}}) \max_{\mathbf{p},\mathbf{P},\mathbf{W},\boldsymbol{\lambda}^{\mathsf{UE}},\boldsymbol{\lambda}^{\mathsf{BS}}} & \sum_{\forall(n,k)} \omega_k[t]\lambda^{\mathsf{UE}}_{n,k}[t] + \sum_{\forall m\in\mathcal{M}_t}\tau_m[t](W^{\mathtt{LEO}}_m - \sum_{\forall n\in\mathcal{N}} W^{\mathsf{BS}}_{m,n}[t])\\ \text{s.t. } & (\bar{C}1) - (\bar{C}3), (\tilde{C}4), (\bar{C}7), (C8), (C11), (\bar{C}12), (C13),\end{aligned}$$

wherein the term $\sum_{\forall m}\tau_m[t]W^{\mathtt{LEO}}_m$ is added to the objective to restore the relaxed form of BW constraint associated with $\boldsymbol{\tau}$, that does not change the optimality of problem $(\mathcal{P}_{\mathsf{II}})$.

Let denote $\mathcal{F}_{\mathsf{II}} \triangleq \{\Omega | \Omega \text{ satisfies } (\bar{C}1) - (\bar{C}3), (\tilde{C}4), (\bar{C}7), (C8), (C11), (\bar{C}12), (C13)\}$ is the feasible set of the relaxed problem $(\mathcal{P}_{\mathsf{II}})$, the dual problem of $(\mathcal{P}_{\mathsf{II}})$ can be expressed as

$$(\mathcal{D}_{\mathsf{II}}) \min_{\boldsymbol{\tau}} \sup_{\Omega\in\mathcal{F}_{\mathsf{II}}} L(\Omega,\boldsymbol{\tau}) \text{ s.t. } \tau_m[t] \geq 0, \forall(m,t), \tag{2.41}$$

where $L(\Omega, \boldsymbol{\tau}) = \sum_{\forall(n,k)} \omega_k[t]\lambda_{n,k}^{\mathsf{UE}}[t] + \sum_{\forall m \in \mathcal{M}_t} \tau_m[t](W_m^{\mathsf{LEO}} - \sum_{\forall n \in \mathcal{N}} W_{m,n}^{\mathsf{BS}}[t])$. Notably, problem $(\mathcal{D}_{\mathsf{II}})$ is also the dual problem of (2.27). Since the problem (2.27) is convex, the strong duality holds for $(\mathcal{D}_{\mathsf{II}})$ and (2.27). ■

Remark 4. *Thanks to Proposition 3, if problems* $\left(\mathcal{P}_{\mathsf{I,c}}^{\mathsf{clus}}\right)$ *are solved optimally and the Lagrangian variables* $\boldsymbol{\tau}$ *are updated as in* (2.40) *based on the bandwidth allocation solution* $\mathbf{W}$*, the convergence of the decentralized algorithm can be guaranteed. The strategy for exchanging and updating the Lagrangian variables* $\boldsymbol{\tau}$ *will be discussed in the subsequent section. Although Algorithm 2 may not obtain the optimal solution of* $\mathbf{W}$ *in each iteration due to the approximation attempt, the implementation of compressed sensing and SCALE methods in this algorithm–well-known as highly effective tools for dealing with sparsity and non-convex sum logarithm-rate maximization problems–can facilitate an efficient solution. Consequently, the convergence of Algorithm 2 is feasible, as demonstrated in Section 2.5. Furthermore, the distributed algorithm can be also set to stop after a certain number of iterations in a practical implementation.*

Decentralized Algorithm Implementation with Limited Signaling

To provide clarity on the practical implementation, Fig. 2.4 is presented to outline the execution steps and data exchange between the cluster LCs and CNC. Specifically, the process can be summarized as follows:

(a) After obtaining the initial Lagrangian multiplier $\boldsymbol{\tau}^{(0)}$ each cluster LC solves its corresponding sub-problem as described in Algorithm 2.

(b) Subsequently, the cluster LCs and the central control unit exchange the data of BW allocation that is used for Lagrangian multiplier calculation.

(c) The CNC calculates and broadcasts the new $\boldsymbol{\tau}$ to all LCs.

This process is reiterated until the convergence of the solution is attained. Moreover, as depicted in Fig. 2.1, one assumes the existence of a limited-capacity link between every LC and the CNC. Given the low-bit exchanging data ($W_{m,n}^{\mathsf{BS}}$ and $\boldsymbol{\tau}^{(\ell)}$), the signaling process can be carried out with low latency.

Remark 5. *Two proposed algorithms are designed based on the min-time optimization problem* (2.19)*, however, they can easily adapt to the max-sum-rate objective with a simple modification.*

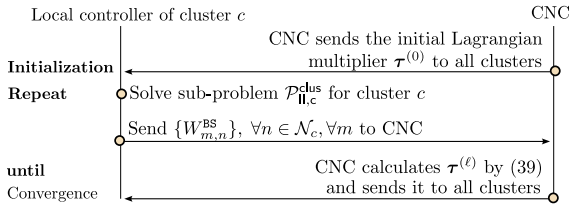


Figure 2.4: Exchange data between local cluster controllers and CNC.

Indeed, by removing the weights ω_k at the objective functions of (2.27) *and sub-problems $\mathcal{P}_{\mathsf{II},\mathsf{c}}^{\mathsf{clus}}$, the proposed centralized algorithm and decentralized algorithm can directly apply to solve the max-sum-rate optimization problem.*

2.4 Other Solutions and Complexity Analysis

2.4.1 Greedy Algorithm

Regarding BS-LEO links in each TS, each BS selects an LSat having the strongest channel gain for communication. Consequently, every LSat assigns equal BW to its associated BSs. Every BS spends all its power budget to forward UEs data, based on which the backhaul capacity R_n^{BS} of the BSs is defined. Then, in every TS, one updates the set of UEs asking for data transmission, i.e., $\mathcal{K}_{\mathtt{D}}$. Then, each UE in $\mathcal{K}_{\mathtt{D}}$ selects the BS offering the best average channel gain to associate with. At the BSs, SCs are assigned to the UEs in descending order of channel gain. Afterward, the water-filling algorithm is employed to optimize the transmission power at UEs without considering extra-cell interference. Herein, the maximum auxiliary power level at UEs needs to be calibrated if the capacity constraints of the BS-LSat backhaul link is violated. The Greedy Algorithm is detailed in Algorithm 3, wherein $R_n^{\mathtt{total}}$ is the total UE rate at BS_n.

2.4.2 Other Benchmark Solution

As shown in Table 2.1, Di *et al.* [12] studied some technical aspects similar to our work without optimizing the transmission power at UEs. However, [12] allocates only one sub-channel to each UE, letting each BS connect to multiple LSats at the same time, and accounting for inter-satellite interference which may be impractical in VSAT systems. Then, we aim to modify the algorithm proposed in [12], namely Di Algorithm, for a fair evaluation. The main modifications are: (i) allowing UEs to be assigned $\bar{S}$ SCs; (ii) equally allocating the transmission power at each UE over its assigned SCs; (iii) pushing each BS to connect to only

Algorithm 3 GREEDY-BASED ALGORITHM TO SOLVE (2.19)

1: Set $t = 1$ and $\mathcal{K}_{\mathtt{D}} = \mathcal{K}$.
2: **repeat**
3: **for** $n = 1 \rightarrow N$ **do**
4: BS_n associates with LEO_m with the best channel gain. Set $\mu_{m,n} = 1$ and $P_n[t] := P_n^{\max}$.
5: **end for**
6: Each LSat uniformly allocates BW to linked BSs.
7: Each $\mathsf{UE}_k \in \mathcal{K}_{\mathtt{D}}$ selects BS_n with the best channel gain.
8: Each BS assigns SCs to UEs in descending order of channel gain.
9: Build matrix $\boldsymbol{\alpha}$. Set $\bar{p}_{\mathsf{UE}}^{\max} = p_{\mathsf{UE}}^{\max}$.
10: **for** $n = 1 \rightarrow N$ **do**
11: **while** $R_n^{\mathsf{BS}} < R_n^{\mathsf{total}}$ or $|R_n^{\mathsf{BS}} - R_n^{\mathsf{total}}| > \epsilon_{\mathsf{rate}}$. **do**
12: Utilize water-filling algorithm for each UE linked with BS_n to find power allocation using $\bar{p}_{\mathsf{UE}}^{\max}$ as the maximum power.
13: Adjust $\bar{p}_{\mathsf{UE}}^{\max}$ by bisection search method based on R_n^{total}.
14: **end while**
15: **end for**
16: Calculate $d_k[t], \forall k$ based on (2.18).
17: $\forall k \in \mathcal{K}_{\mathtt{D}}$, remove UE_k with $d_l[t] = 0$: $\mathcal{K}_{\mathtt{D}} = \mathcal{K}_{\mathtt{D}} - \{k\}$.
18: Set $t = t + 1$.
19: **until** $\mathcal{K}_{\mathtt{D}} = \emptyset$.
20: **Output:** The solution $(\mathbf{p}, \mathbf{P}, \mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\mu})$.

one LSat; (iv) ignoring the inter-satellite interference, then utilizing the water-filling method to allocate power transmission for BSs.

2.4.3 Complexity Analysis

Greedy Algorithm

Regarding Algorithm 3, the FOR loop given in **Steps 3-5**, the BW allocation procedure in **Step 6**, and BS-UE association operation in **Step 7** demand a complexity of $O(2N + K)$. Then, the complexity of SC assignment process in **Step 8** is $O(NS)$. For the power allocation work given in **Steps 10-15**, one requires $-\log_2(\epsilon_{\mathsf{pow}})$ and $-\log_2(\epsilon_{\mathsf{rate}})$ iterations to find $\bar{p}_{\mathsf{UE}}^{\max}$ level and water level by bisection search method, respectively. Additionally, the complexities of obtaining UE power control and calculating the data remaining are $O(NK\bar{S})$ and $O(K)$ computations, respectively. In summary, the complexity of the greedy algorithm can be expressed as

$$X_{\mathsf{Gre.Alg.}} = O\Big(T^{\mathsf{Gre}} NK\bar{S} \log_2(\epsilon_{\mathsf{rate}}) \log_2(\epsilon_{\mathsf{pow}})\Big), \tag{2.42}$$

where T^{Gre} is the number of TS required for uploading all UE data of the greedy algorithm.

Di Algorithm

In Di Algorithm, the most computation complexity is due to the three-dimension matching process. In particular, a three-dimension matching approach is utilized for the BS-UE-SC association repeatedly to let each UE can access multiple SCs from one BS, this process demands a complexity of $\mathcal{O}(NK^2S^2\bar{S})$. For the BS-LSat transmission, the three-dimension matching approach is applied again for the LSat-BS-SC association to let each BS can connect to one LSat over at most $\bar{S}_{\mathsf{BS}}$ SCs. The complexity of this process can be expressed as $\mathcal{O}(MN^2S^2_{\mathtt{LEO}}\bar{S}_{\mathtt{BS}})$, where $S_{\mathtt{LEO}}$ is the number SCs in the LSat-BS transmission. In summary, the complexity in terms of big-O of Di Algorithm can be expressed as

$$X_{\mathsf{Di.Alg.}} = \mathcal{O}\Big(N_{\mathsf{out}}(NK^2S^2\bar{S} + MN^2S^2_{\mathsf{LEO}}\bar{S}_{\mathsf{BS}})\Big), \tag{2.43}$$

where N_{out} is the number of outer iterations for convergence.

Proposed Algorithms

Each proposed algorithm consists of two loops: the outer loop corresponds to the number of TSs of transmission required for completing the upload of all UEs data while the inner loop iteratively solves the optimization problems until the solution for each TS is obtained. To implement each algorithm, the successive convex problems (2.27) and (2.38) should be transformed into the conic form which is accepted by the convex solvers, e.g., MOSEK [83,84], resulting in requires extra constraints and auxiliary variables. Regarding the centralized algorithm, one needs to deal with problem (2.27) several times. According to [84], this problem can be transformed into a conic problem which consists of $x_1 = NKS + NS + NK + 4N + 2K + M_t$ linear constraints, $y_1 = N^2K^2S + 2NKS + M_tN$ exponential cone constraints, and $z_1 = N^2K^2S+4NKS+NK+3M_tN+N$ variables. Based on [84,85], to obtain ϵ-accuracy solution for problem (2.27), the connect-based solving procedure requires a number of $\mathcal{O}\Big(3\log(1/\epsilon)\Big)$ sub-steps. Therefore, the complexity of the centralized algorithm can be expressed as

$$X_{\mathsf{Cen}} = \mathcal{O}\Big(9T^{\mathsf{Cen}}N_{\mathsf{iter}}\log(1/\epsilon)\left((2(z_1+1)+x_1)^3 + y_1\right)\Big), \tag{2.44}$$

where N_{iter} is the number of inner-loop iterations, T^{Cen} is the required number of TSs due to the centralized algorithm.

Table 2.2: Simulation Parameters

Parameter	Value	Parameter	Value
Number of BS clusters	6	LSat BW, W^{LEO}	20 MHz
Number of UEs, K	60	Orbit type	Walker-star
Number of BSs, N	18	Number of BSs per cluster	3
LSat altitude	600 km	LSat velocity	7.57 km/s
Number of satellite orbits	2	LSat antenna gain, G^{LEO}	37.1 dBi
UE data demand, $\bar{D}$	5 Mbits	SC BW, W_{SC}	720 kHz ($\mu = 2$)
BS antenna gain, G^{BS}	32.8 dBi	BS-LEO operation frequency, f_c	30 GHz
Number of SCs, S	8	Max. number of SCs at UE, $\bar{S}$	4
TS duration, T_S	30 ms	Max. UE transmit power, $p_{\mathsf{UE}}^{\max}$	26 dBm
Rician κ factor, κ^{TN}	5 dB	Max. BS transmit power, $P_{\mathsf{BS}}^{\max}$	14 dBW
Radius of TN cell	200 m	Considered area	(7km $\times$ 5km)
Satellite 3dB beam diameter	12.35 km	Noise power density	-174 dBm/Hz

For the decentralized algorithm, it is equivalent to a conic problem with $x_2 = N_c K_c S + N_c S + N_c K_c + 4N_c + 2K_c + M_t$ linear constraints, $y_2 = N_c^2 K_c^2 S + 2N_c K_c S + M_t N_c$ conic constraints, and $z_2 = N_c^2 K_c^2 S + 4N_c K_c S + N_c K_c + 3M_t N_c + N_c$ variables. Therefore, its complexity can be written as

$$X_{\mathsf{Dec}} = O\Big(9T^{\mathsf{Dec}} N_{\mathsf{iter}} N_{\mathsf{Lag}} \log(1/\epsilon) \Big(\big(2(z_2+1) + x_2\big)^3 + y_2 \Big) \Big), \tag{2.45}$$

where N_{Lag} is the number of Lagrangian iterations, T^{Dec} is the required number of TSs due to the decentralized algorithm. Here, both algorithms have polynomial time complexities while the decentralized approach has a lower complexity.

2.5 Numerical Results

2.5.1 Setting Parameters

This section presents the simulation results to evaluate the performance of the proposed solutions and compare them with reference algorithms. The simulations were performed over an area of 7×5 km^2, with the center geographical coordinate set at $(\varphi_0, \theta_0) = (40°\text{N}, 20°\text{E})$. Within this area, terrestrial clusters were uniformly deployed and CNC is assumed to be located at the center of the area. The rain attenuation $\varrho^{\mathsf{c}}_{m,n}[t]$ is assumed to follow "log-normal distribution" with a mean of 2.6 dB and a variance of 1.63 dB [86], while the cloud attenuation $\varrho^{\mathsf{r}}_{m,n}[t]$ is set as detailed in [87]. The Doppler shift, ν^t_m, is calculated based on LSat velocity and position as described in [88]. Other key parameters are listed in Table 2.2. To ease the readability, two proposed algorithms, i.e., Algorithm 1 and Algorithm 2 are referred to as *"Cen-Alg"* and *"Dec-Alg"*, respectively. The Greedy Algorithm and Di Algorithm are also implemented for comparison purposes.

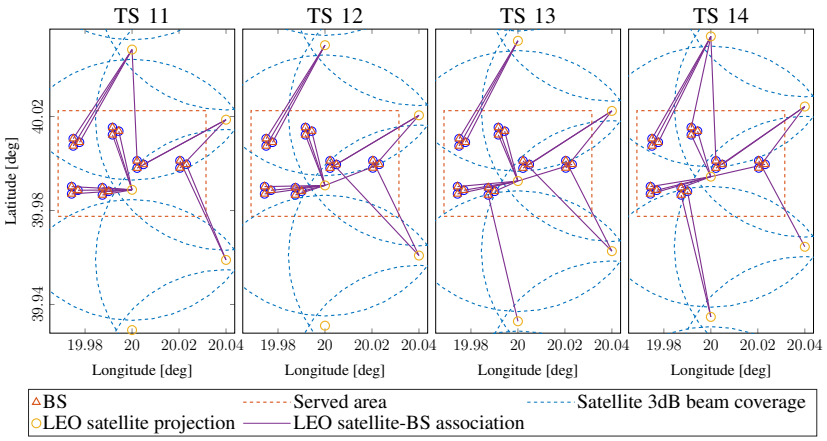


Figure 2.5: The LSat-BS association result over TSs.

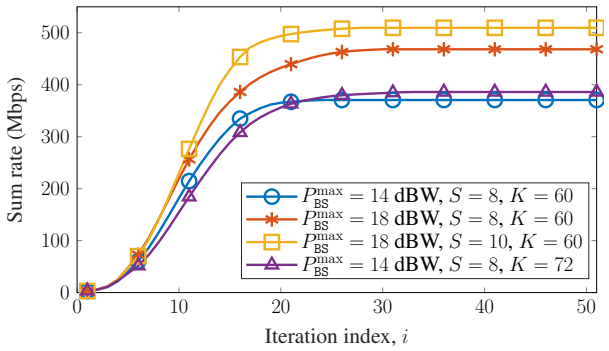


Figure 2.6: Sum-rate of the centralized algorithm in TS 1 over iterations.

2.5.2 Numerical Results

Fig. 2.5 displays a simulated topology along with the LSat-BS association results, as obtained from the proposed *Cen-Alg* during the duration from TS 11 to TS 14, wherein the satellite 3dB beam coverage is calculated based on beam pattern formulation given in Section 2.2.2 and parameters in Table 2.2. It can be observed that BSs in the service area tend to prioritize connection with the nearest LSat due to the strong channel gain (the middle LSat in Fig. 2.5). However, due to limited backhaul capacity, several BSs connect to distant LSats. Moreover, it can be seen that LEO switching does not occur too frequently. Specifically, in this duration, there are a few BSs that need to switch their LEO association.

Fig. 2.6 illustrates the SR achieved by the proposed *Cen-Alg* over iterations in TS 1 for different settings. As can be observed, in all settings, the achieved SR increases and saturates after just a few tens of iterations, that has confirmed the convergence of our proposed *Cen-Alg*. For instance, with the settings of $(P_{\mathrm{BS}}^{\max}, S, K) = (14, 8, 60)$ and $(P_{\mathrm{BS}}^{\max}, S, K) = (18, 8, 60)$, our proposed *Cen-Alg* requires approximately 23 and 26 iterations for convergence, respectively.

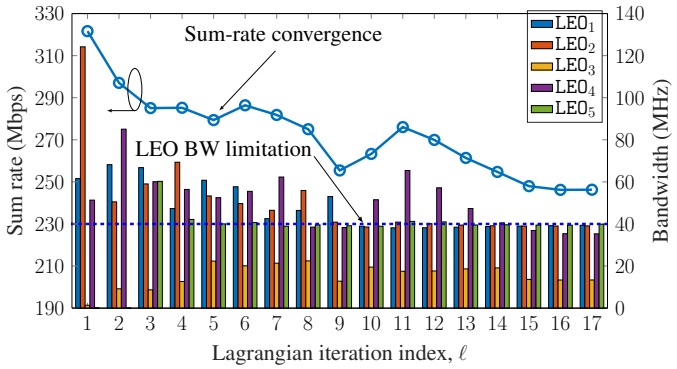


Figure 2.7: Sum-rate and BW of proposed *Dec-Alg* in TS 1 over iterations.

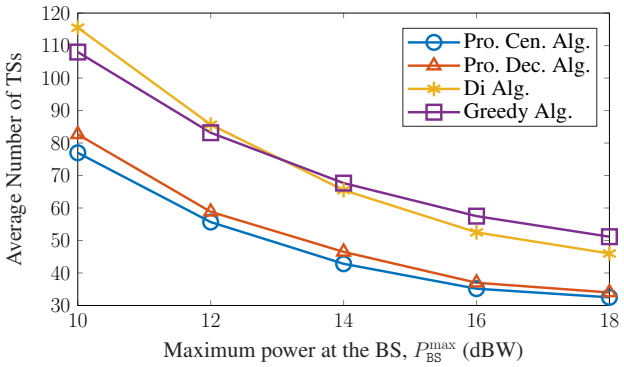


Figure 2.8: Average TS number vs. BS max power.

Even with a larger setting, $(P_{\mathrm{BS}}^{\max}, S, K) = (18, 8, 60)$ and $(P_{\mathrm{BS}}^{\max}, S, K) = (14, 8, 72)$, a similar number of iterations for convergence is required, i.e., about 28 and 27 iterations, respectively.

Fig. 2.7 depicts the SR convergence and the allocated BW of each LSat over Lagrangian iterations. In this approach, at each Lagrangian iteration, each cluster solves a sub-problem similar to the problem addressed in the centralized approach, but on a smaller scale. Therefore, the SR convergence of the *Dec-Alg* at each cluster can be achieved with a lower number of iterations in comparison to that of the proposed *Cen-Alg*. In this instance, the proposed *Dec-Alg* requires approximately 15 Lagrangian iterations for convergence. Intriguingly, during the initial iterations, the algorithm delivers superior SR performance, but there are some LSats where the total allocated BW exceeds the maximum BW at the LSat (indicated by the threshold dashed line). This phenomenon occurs due to the relaxation of the BW constraint at each LSat. However, thanks to the adjustment of the Lagrangian multiplier $\boldsymbol{\tau}$ based on (2.40), the BW violation decreases in later Lagrangian iterations, and the BW allocation constraint is fulfilled upon achieving convergence.

To assess the effect of backhaul link capacity, Fig. 2.8 depicts the average numbers of TSs

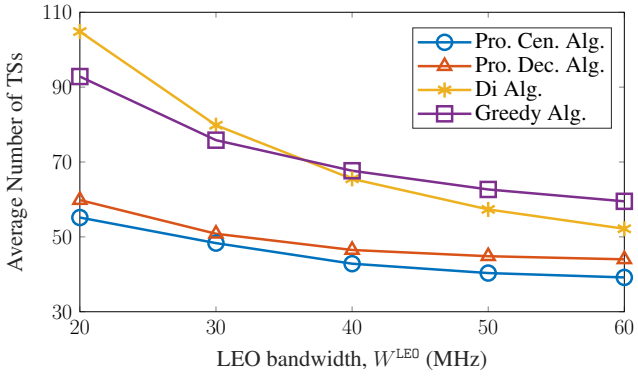


Figure 2.9: Average TS number vs. LEO BW.

required to meet UE demand according to the four algorithms versus different maximum BS power levels. Overall, the results of all algorithms share the same trend that the average TS number decreases as the maximum BS transmission power increases. As anticipated, both proposed algorithms outperform the benchmarks with considerable performance gaps. For instance, at $P_{\text{BS}}^{\max} = 12$ dBW, the average numbers of TSs required to meet UE demand for the two proposed algorithms are approximately 55.7 and 58.8 TSs. In contrast, the Di Algorithm and the Greedy Algorithm require substantially more TSs, specifically around 85.6 and 83.2 TSs, respectively. Across all settings of $P_{\text{BS}}^{\max}$, the gap between the proposed algorithm and the benchmark lines seems to be unchanged, i.e., approximately 30 TSs. Particularly noteworthy is the minor performance gap between the *Cen-Alg* and *Dec-Alg*, roundly about 3 TSs, even though the latter algorithm is executed separately at the LCs.

To further evaluate the influence of the backhaul link capacity, Fig. 2.9 illustrates the performance of the four algorithms in terms of the average number of TSs versus the maximum backhaul BW which can be provided by the LSats. The trend of the results in this figure is similar to that shown in Fig. 2.8, where the average number of TSs decreases as the backhaul capacity budget increases. As can be seen, the two proposed algorithms deliver superior performance in terms of transmission time compared to the Di Algorithm and the Greedy Algorithm. For instance, at $W^{\text{LEO}} = 40$ MHz, the proposed *Cen-Alg* and *Dec-Alg* require approximately 42.8 and 46.5 TSs to fulfill UE demand, while the benchmark algorithms require about 65.5 and 67.6 TSs, respectively. Contrary to the results in Fig. 2.8, the performance gap between the proposed algorithms and the two benchmark algorithms is larger with a smaller W^{LEO}. This phenomenon demonstrates the efficiency and capability of the two proposed algorithms to exploit LSat BW compared to the benchmarks.

Subsequently, Fig. 2.10 illustrates the average number of TSs as a function of the UE

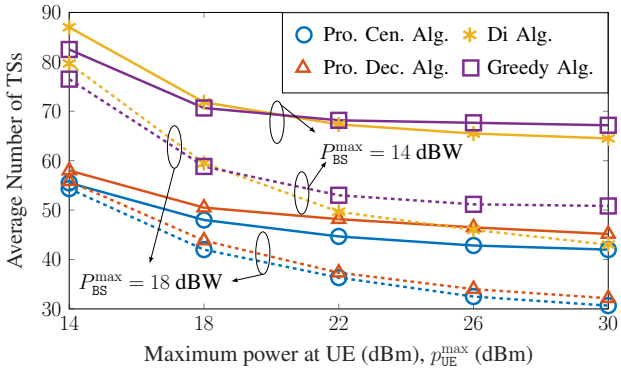


Figure 2.10: The number of required TSs vs. the maximum power at UE.

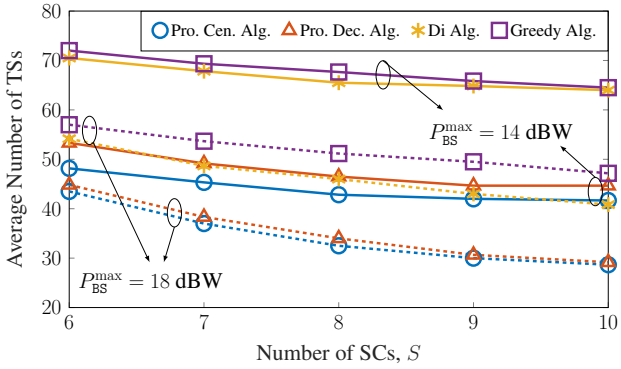


Figure 2.11: The number of TSs versus the number of SCs.

transmission power budget. Again, the two proposed algorithms consistently provide superior solutions with a significant performance gap compared to the benchmarks. With $P_{\text{BS}}^{\max} = 14$ dBW, the Di Algorithm and the Greedy Algorithm require about 71.5 and 70.6 TSs, respectively, while the proposed algorithms require only approximately 48 and 50.5 TSs at $p_{\text{UE}}^{\max} = 18$ dBm. These numbers of TSs are reduced to about $64.5, 67.2, 42$ and 45.2 TSs at $p_{\text{UE}}^{\max} = 30$ dBm, respectively. Interestingly, the numbers of TSs corresponding to the four schemes decrease quickly when $p_{\text{UE}}^{\max}$ varies from 14 dBm to 18 dBm, but saturate when $p_{\text{UE}}^{\max}$ is set to a higher value. This indicates that the performance seems to be bottlenecked by the backhaul link capacity. For a clearer assessment, we examine the setting with $P_{\text{BS}}^{\max} = 18$ dBW. As expected, the number of TSs for the four algorithms decreases significantly as $p_{\text{UE}}^{\max}$ increases. When $p_{\text{UE}}^{\max}$ is set in the range from 22 dBm to 30 dBm, the performance of the Di Algorithm improves rapidly compared to that of the Greedy Algorithm. However, the performance gap remains substantial compared to the two proposed algorithms. In particular, the average numbers of TSs for the benchmarks are about 43 and 50.8 TSs, respectively, whereas those for the proposed algorithms are only 30.6 and 32.1 TSs at $p_{\text{UE}}^{\max} = 30$ dBm.

Next, Fig. 2.11 illustrates the impact of the number of SCs on the system performance in

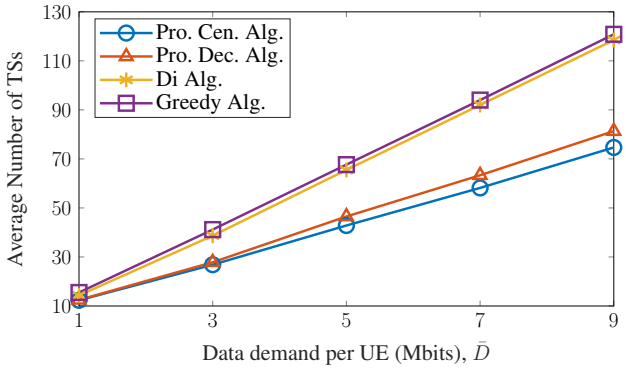


Figure 2.12: Required-TS number vs. UE data demand.

terms of the complete time of UE data upload. With $P_{\mathsf{BS}}^{\max} = 14$ dBW, the benchmark algorithms require a similar number of TSs across all different settings of S, and this transmission time is much larger than those of the two proposed algorithms. Specifically, the proposed algorithms can reduce the requirement by approximately 23 and 21.5 TSs for all values of S in this simulation. Conversely, it can be seen that the average number of TSs achieved by all four schemes decreases slowly as the number of SCs increases. Especially, when S varies from 8 to 10, the changes in the average number of TSs for the proposed *Cen-Alg* and *Dec-Alg*, the Di Algorithm, and the Greedy Algorithm are only about $1.2, 1.9, 1.5$, and 3 TSs, respectively. Similar to the result shown in Fig. 2.10, this indicates the presence of a backhaul link bottleneck. When we increase $P_{\mathsf{BS}}^{\max}$ to 18 dBW, as expected, the number of SCs has shown a more significant impact on the number of TSs. In this case, the Di Algorithm shows better transmission time performance compared to that of the Greedy Algorithm. However, the performance gap between these benchmarks and the two proposed algorithms remains notable. In particular, in the setting of 10 SCs available at the BSs, the Di Algorithm and the Greedy Algorithm require about 40.8 and 47.1 TSs, respectively, while the proposed algorithms require only about 28.6 and 29.1 TSs.

Additionally, Fig. 2.12 presents the TS numbers of four schemes as the functions of transmission data-amount demanding by each UE, $\bar{D}$. For all considered algorithms, the required TS numbers of all schemes appear to increase linearly as the data demand per UE escalates. Notably, the gaps between the proposed algorithms and the benchmarks are proportionally linear to $\bar{D}$. For instance, at $\bar{D} = 3$ Mbits, the required-TS gaps from the *Cen-Alg* and *Dec-Alg* to the Di Algorithm are about 11.8 and 11.2 TSs, respectively. Conversely, at $\bar{D} = 9$ Mbits, these gaps grow to about 43.9 and 37.2 TSs. Again, this figure confirms the effectiveness of the two proposed algorithms in comparison to the two benchmark algorithms when the data

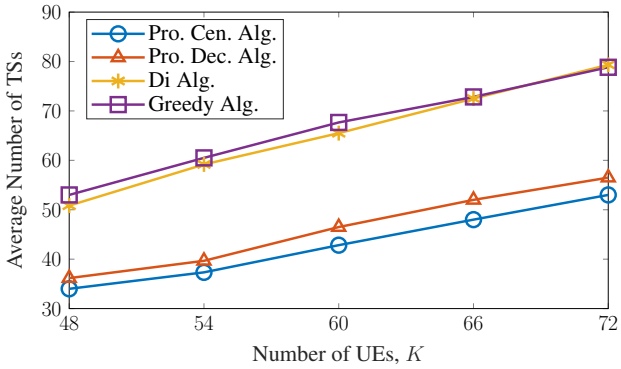
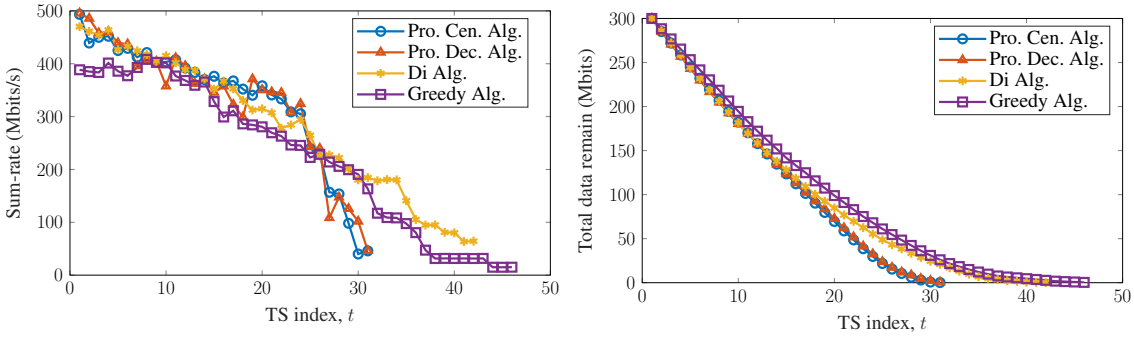


Figure 2.13: Required-TS number vs. UE number.



(a) Sum-rate over TSs.

(b) Total UE data remain over TSs.

Figure 2.14: Sum-rate system and data remain versus TS of a realization with $P_{\mathsf{BS}}^{\max} = 18$ dBW.

demand of the UE increases.

Fig. 2.13 illustrates the impact of variations in the number of UEs on transmission time performance. With fixed resources, it is observed that an increased number of UEs necessitates more TSs to meet the data demand for all UEs. As expected, the solutions provided by the two proposed algorithms surpass those of both the Di Algorithm and the Greedy Algorithm in terms of transmission time across all instances of K. Notably, the transmission time performance gaps between the two proposed algorithms and the two benchmark algorithms are significant and seem to remain relatively constant with different values of K, i.e., about 24 and 20 TSs from the *Cen-Alg* and *Dec-Alg* to the Di Algorithm, respectively. In addition to showcasing the superior performance of the two proposed algorithms in terms of minimizing transmission time, this figure also demonstrates the adaptability of our proposed algorithms with varying numbers of UEs in the network.

Next, Fig. 2.14a and 2.14b respectively present the achieved SR and total remaining UE data over TSs for the four considered algorithms for a particular realization. In almost all TS, Greedy Algorithm achieves the lowest SR performance, thus it requires the largest number

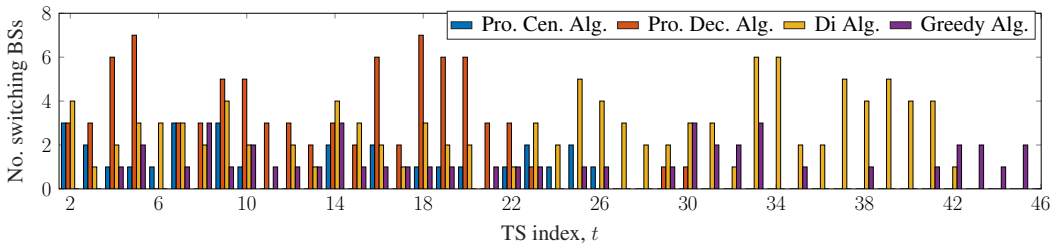


Figure 2.15: The number of switching BSs over TSs.

of TSs to complete UE demand, i.e., 46 TSs. Regarding the other three algorithms, the differences among the SR achieved by these schemes are not significant, with similar totals of remaining data, in the duration from TS 1 to TS 15. In the subsequent TSs, the two proposed algorithms demonstrate superior SR performance; hence, they are able to fulfill UE data demand in fewer numbers of TSs compared to the benchmarks. Specifically, the outcomes of the proposed algorithms are 30 and 31 TSs, respectively, while the Di Algorithm requires 42 TSs. The reason for this phenomenon can be expressed as follows. After a certain TS, although the total remaining data for all algorithms is similar, the distribution of remaining data among UEs may differ. Therefore, adjusting the weight for the SR objective and accounting for capacity balancing in the two proposed algorithms can enhance SR performance. These figures further clarify the effectiveness of the two proposed algorithms.

Finally, Fig. 2.15 shows the number of BSs that have to switch their associated LEO over time. Herein, such BSs are so-called the switching BSs. As can be observed, the number of switching BSs due to Greedy Algorithm implementation is the lowest in almost all TSs due to the simple channel-gain-based BS-LEO association. While the proposed *Cen-Alg* returns the number of switching BSs lower than the other two remaining schemes, i.e., ranging from 1 to 3 BSs in nearly all TSs. Interestingly, the outcomes of the decentralized approach vary with large variance over the duration from TS 1 to TS 22 while Di Algorithm shows significant fluctuation in the following TSs. Additionally, the maximum number of switching BSs of these two algorithms are 7 and 6. Herein, the *Cen-Alg* has fewer switching BSs compared to the decentralized one but the returns of the latter are still acceptable, as the maximum number of HO BSs is 7 out of the total of 18 BSs. This can be interpreted as a trade-off between the two proposed algorithms in terms of workload offloading for the CNC.

2.5.3 Numerical-Result Discussion

The Superiority of Proposed Algorithms Compared to Benchmarks

In all considered scenarios, our two proposed algorithms demonstrate superior performance in terms of transmission time and achievable rates when compared to two baseline algorithms: the Greedy Algorithm and the Di Algorithm. The advantages of our proposed algorithms over these baselines are outlined as follows:

- Greedy Algorithm: establishing BS-UE and BS-LSat associations and SC assignments based solely on channel gain. The UE power allocation is performed using the water-filling algorithm, which does not account for interference, potentially leading to low performance.
- Di Algorithm: In this approach, UE transmit power is fixed. The resource management task is split into two separate problems for TN and NTN. As a result, this algorithm does not efficiently address interference management and two-tier user association issues.
- Our Proposed Algorithms: Our algorithms perform a comprehensive optimization of two-tier user associations, SC assignment, power control, and bandwidth allocation, all under the backhaul-link limitations. Moreover, by prioritizing users based on their remaining demand and incorporating UE fairness, our algorithms achieve a notable improvement in transmission time performance.

Comparable Performance of Decentralized and Centralized Algorithms

Both the *Cen-Alg* and *Dec-Alg* are developed using similar technical methods, resulting in the *Dec-Alg* having a performance comparable to that of the *Cen-Alg*. Specifically, in the decentralized approach, problem (2.25) from the centralized framework is divided into sub-problems $\mathcal{P}_{\mathsf{l,c}}^{\mathsf{clus}}$ for clusters as delineated in (2.36). Notably, each sub-problem $\mathcal{P}_{\mathsf{l,c}}^{\mathsf{clus}}$ shares a similar structure with (2.25). Consequently, we employ the same technical methods, namely compressed-sensing and SCALE methods, to solve them. The primary distinction between the centralized and decentralized solutions lies in the latter's requirement for updating the Lagrangian multiplier $\boldsymbol{\tau}$. This challenge is efficiently addressed using the sub-gradient descent method. Therefore, the *Dec-Alg* can obtain comparable performance to that of the *Cen-Alg*.

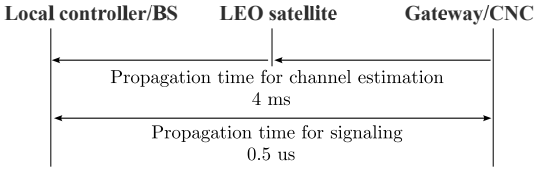


Figure 2.16: Time consumption of propagation and signaling procedures.

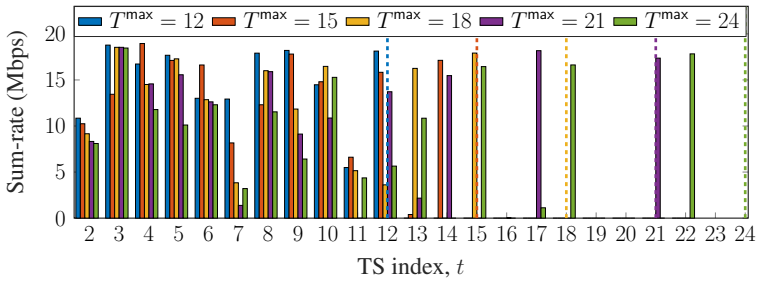


Figure 2.17: Sum-rate of a delay-sensitive UE versus TS of a realization.

Practical Implementation Analysis

The studied network consists of two separate signal transmission hops, i.e., BS-UE and BS-LSat, which are processed on different frequency bands. Consequently, the synchronization procedures for these two hops can be conducted independently. For the BS-UE hop, we assume adherence to the 5G-NR standard. For the BS-LSat hop, the time and frequency synchronization can be conducted at each BS thanks to pre-compensation as outlined in the standard [8]. It is noted that both proposed algorithms are executed at the ground controllers without frequent signaling between LSats and ground segments. The eligible time consumption is the propagation time for channel estimation between BS and LSat. Hence, this time consumption is approximately 4 ms which is the round-trip-time between the LSat and the ground segment calculated based on the LSat altitude of 600 km.

Next, as per the simulation settings of the simulation area and CNC location, the distance between CNC and LC is about 4.3 km, thus the transmission time between CNC and LC $T_{\mathsf{CNC,LC}} \leq 1.43 \times 10^{-8}$ s. On the other hand, it can be seen that the *Dec-Alg* requires about 15 Lagrangian iterations for convergence as per Fig. 2.7. Based on the data exchange procedure described in Fig. 2.4, the time consumption due to the signaling procedure can be calculated as $T_{\mathsf{sig}} \approx 0.5$ us. Summarily, the time consumption due to the propagation and data exchange is about 4.0005 ms, thus it is feasible to consider the TS duration of 30 ms. Therefore, this further demonstrates the practicality of the *Dec-Alg*. Specifically, the time consumption due to propagation and signaling procedure is described as per Fig. 2.16.

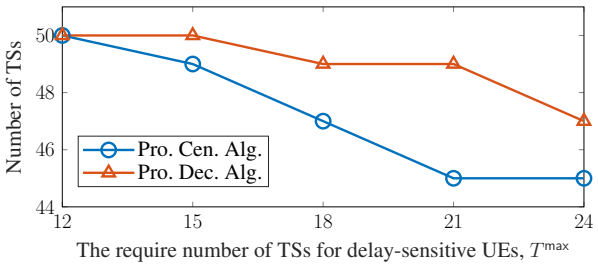


Figure 2.18: Number of TSs versus $T_{\max}$ of a realization.

Impact of Temporary Link Disruption

Regarding one user's data transmission during a specific time window, temporary link disruptions can occur when no system resources are allocated for data transmission in one or multiple time slots (TSs). Such disruptions can cause latency and delays, which are particularly critical for users with stringent delay requirements, referred to as delay-sensitive UEs (DUEs). Let $\mathcal{K}_{\text{sen}}$ be the DUE set. Regarding the delay requirements of the DUEs, we consider the following additional constraints,

$$(C_{\text{DI}}): \ t_k \leq T_k^{\max}, \forall k \in \mathcal{K}_{\text{sen}}, \text{ and } d_k(t_k) = 0, \forall k, \tag{2.46}$$

wherein t_k indicates the TS in that the $\texttt{DUE}_k$ can finish transmitting all data while $T_k^{\max}$ stands for the delay tolerance of $\texttt{DUE}_k$. To satisfy this constraint, we proposed the weight update approach instead of (2.21) as

$$\omega_k[t] = T_k^{\max}/(\max(0, T_k^{\max} - t) + \epsilon) d_k[t], \forall k \in \mathcal{K}_{\text{sen}}. \tag{2.47}$$

Based on this, the DUEs have higher priority for resource allocation at the TS closed to their TS tolerance, i.e., $T_k^{\max}$.

To evaluate the effectiveness of this updated approach in meeting the delay requirements for DUEs, we consider an additional simulation scenario in which 20% of all UEs are DUEs and their delay-tolerance for the DUEs varies from 12 to 24 TSs. In Fig. 2.17, one depicts the SR of a DUE with the different delay requirements $T_k^{\max}$, wherein the dashed lines indicate the delay requirements. It can be seen that the DUE transmits data continuously in almost all cases. Especially, the delay requirement is satisfied in all cases for this UE. However, the effort to satisfy the delay requirement impacts the overall system performance. Particularly, Fig. 2.18 shows the TS number to offload all UEs data for the various delay requirements

$T_k^{\max}$. It can be observed that the required TS number increases due to the stricter delay requirements. This occurs because when higher priorities are set for the DUEs, the lower system resources are allocated to other users, leading to the degradation in overall system performance. This interestingly presents the trade-off between satisfying delay requirements and maintaining overall system performance.

2.6 Summary

This work studied the novel joint two-tier association and RA for ISTNs over time regarding the mobility of satellites in the LEO constellation. An optimization problem focusing on minimizing transmission time under UE demand and backhaul link constraints was then considered. To address this problem, an iterative *Cen-Alg* implementable at the CNC is proposed. Moreover, with the objective of offloading computations from the central node, we proposed *Dec-Alg* capable of parallel execution at local network controllers efficiently. Through a range of numerical evaluations, both proposed algorithms demonstrated superior performance in terms of transmission time minimization and practicality, compared with two benchmarks. Additionally, the simulation results highlighted the proposed algorithms' effectiveness in terms of capacity balancing, SR performance, and adaptation ability, further validating their advantages and potential utility in real-world applications.

This study offers avenues for further extension, particularly in addressing the challenges of handover and scalability. Particularly, the handover between BSs and LSats, which consumes time for connection establishment and reduces system performance, should be minimized. These handovers need to be carefully considered in the optimization problem to avoid frequent occurrences. In terms of scalability, the controller link limitations at the CNC could lead to overload when scaling up the system. Therefore, scheduling for data exchange must be carefully planned for practical implementation. Additionally, the application of machine learning and deep learning-based mechanisms offers promising prospects for enhancing user association and RA in ISTNs, especially utilizing learning-based techniques in the MINLP problems. These emerging techniques, as discussed in [89–91], may facilitate comprehensive exploration in future research endeavors.

Chapter 3

Enhanced Throughput and Seamless Handover Solutions for 5G-ISTNs

THIS chapter investigates downlink transmission in 5G-ISTNs supporting automotive UEs in urban environments, where BSs and LSats cooperate to serve moving UEs over shared C-band frequency carriers. Urban settings, characterized by dense obstructions, together with UE mobility, and the dynamic movement and coverage of LSats pose significant challenges to user association and resource allocation. To address these challenges, we formulate a multi-objective optimization problem designed to improve both throughput and seamless HO. Particularly, the formulated problem balances sum-rate (SR) maximization and connection change (CC) minimization through a weighted trade-off by jointly optimizing power allocation and BS-UE/LSat-UE associations over a given time window. This is a mixed-integer and non-convex problem which is inherently difficult to solve. To solve this problem efficiently, we propose an iterative algorithm based on the Successive Convex Approximation (SCA) technique. Furthermore, we introduce a practical prediction-based algorithm capable of providing efficient solutions in real-world implementations. Especially, the simulations use a *realistic 3D map of London* and UE routes obtained from the Google Navigator application to ensure practical examination. Thanks to these realistic data, the simulation results can show valuable insights into the link budget assessment in urban areas due to the impact of buildings on transmission links under the blockage, reflection, and diffraction effects. Furthermore, the numerical results demonstrate the effectiveness of our proposed algorithms in terms of SR and the CC-number compared to the greedy and benchmark algorithms.

3.1 Introduction

In recent years, wireless communication networks have rapidly evolved to meet increasing traffic demands and connectivity needs. Consequently, terrestrial networks (TNs) have been swiftly deployed to accommodate this growth. However, TNs often struggle to maintain the connectivity and quality of service for a vast number of devices. Therefore, the next-generation network aims to provide extensive and seamless connectivity, addressing the high traffic demands and service requirements of new applications [2–4]. Especially, these requirements become more crucial in urban environments that generally have high traffic and user demand. A straightforward solution is the dense deployment of base stations (BSs) to enhance network coverage and capacity. However, this approach can be challenging and costly to implement. Additionally, satellite communication (SatCom) emerges as a promising solution to expand network coverage and capacity [55]. Particularly, due to the ubiquitous and high capacity, SatCom can provide complementary coverage, aid in offloading for TNs, and jointly serve with TNs. This approach is supported by the United States (US) Federal Communications Commission (FCC), which has proposed complementary coverage for TNs from space segments [23]. Although satellite coverage is typically seen as the solution for areas lacking TN connectivity, using SatCom in urban environments, rather than relying on dense base station (BS) deployment, presents a cost-effective alternative. This approach not only enhances coverage but also helps traffic offloading, addressing issues of susceptibility to blockage and congestion events, respectively. While SatCom was previously expensive, costs have now decreased, especially for Low Earth Orbit satellites (LSats). Hence, reasonable manufacturing and launching costs of LSats led to many interests in developing LEO constellations such as Starlink and OneWeb. Compared to geostationary and medium-earth orbit satellites, LSats offer higher channel gain and lower latency due to their lower altitude, making the integration of satellite and TN systems more feasible [49].

Traditionally, TNs and SatCom Networks (SatNets) have operated in distinct frequency bands (RFBs) due to spectrum regulation. However, to meet the growing demand for connectivity, both networks have expanded their spectrum utilization, resulting in the development of coexistence radio-access systems. For instance, many countries have extended the TN radio frequency (RF) spectrum into the millimeter-wave frequency and C-band (3.4 – 4.2 GHz). Concurrently, significant efforts from academia and industry aim to lower the RFBs

of Non-Terrestrial Networks (NTNs), including SatNets, from Frequency Range 2 (FR2) to Frequency Range 1 (FR1), incorporating bands such as the S-band and L-band [25], which can offer better penetration through precipitation and a longer propagation distance compared to higher RFBs. Following this trend, Mediatek has proposed scenarios in which the TN spectrum is shared with NTNs under complementary management [22]. Notably, various coexistence scenarios between TNs and NTNs have been investigated by 3GPP [9].

In light of the evolving trend toward Integrated Satellite-Terrestrial Networks (ISTNs), sharing the same RFBs between TN and NTN systems is considered a promising solution to enhance spectrum efficiency and increase overall network capacity. Based on the agreement between terrestrial and satellite operators, BSs and LSats can work together in ISTNs to provide Internet connectivity. This approach is further supported by the FCC, which has proposed the co-primary use of mobile and satellite services within the same RFBs [23]. Furthermore, direct-to-device connectivity has attracted significant attention and is supported by the European Space Agency [92]. Hence, this approach offers the advantage of improving seamless connectivity by allowing UEs to transit smoothly between TN and NTN systems [9,23]. Moreover, the C-band has emerged as a critical RFB for future networks. Historically reserved for Fixed Satellite Services (FSS), the C-band is now attracting interest from TN operators due to its wide coverage and high capacity. In the US, the FCC has implemented a plan to repurpose 280 MHz of C-band spectrum from FSS downlink (DL) to 5G, requiring satellite companies to vacate this portion by December 2023 [27]. In Europe, the Radio Spectrum Policy Group has identified three key bands for 5G deployment: 700 MHz for broad coverage, 3.4 GHz for a balance between coverage and capacity, and 26 GHz for high-speed data rates, with the 3.4 GHz C-band designated as the primary band for implementation. Additionally, 3GPP has defined three specific C-band frequencies ($n77$, $n78$, and $n79$) for 5G-NR [24], underscoring the C-band's pivotal role in the future of ISTNs.

Although LSats in ISTNs offer substantial benefits, such as enhanced capacity, improved coverage, and seamless connectivity, their inherent movement introduces significant dynamic challenges. The extensive coverage provided by satellites also introduces heterogeneity in ISTNs in terms of propagation environments due to the diverse characteristics of the TN areas (urban, suburban, rural) and even environmental differences within a single urban area [30]. This environmental diversity necessitates accounting for the varying impact of the environment within the ISTN's coverage in system design. Additionally, UE mobility, particularly the

movement of vehicles in urban areas, further complicates the network dynamics [31]. Although utilizing the same RFB in ISTNs, such as the C-band for DL as per the current spectrum regulation, can enhance spectrum efficiency, increase system capacity, and support seamless connectivity, it also introduces critical cross-interference issues between TN and NTN systems. Consequently, system design must carefully address seamless handovers, cross-system interference, and resource allocation, factoring in the unique characteristics of urban environments and realistic UE mobility to ensure these challenges are adequately reflected in practical deployments.

3.1.1 Related Works

Recently, the analysis of interference in TN-NTN coexistence and spectrum sharing systems has been investigated in many works [9, 28, 29, 93, 94]. In [9], the 3GPP has considered various TN-NTN coexistence scenarios wherein TN and NTN systems operate in adjacent bands in the S-band. Subsequently, the throughput loss due to adjacent channel interference was investigated. In [28, 29], the coexistence systems in C-band were studied. Particularly, the authors in [28] analyzed the impact of TNs on DL FSS systems wherein the aggregated interference of 5G BSs at the ground FSS receiver was investigated. Subsequently, the switch-off and power-back-off schemes are proposed to reduce the interference from critical BSs. The authors in [29] conducted the measurement of the interference power caused by TN BSs to the FSS receiver. Subsequently, the impact of various experimental deployment scenarios of the tuner and filter on the interference has been carried out. Additionally, in [93, 94], the performance analysis is studied based on the stochastic geometry. In particular, the authors in [93] studied different coexistence scenarios in the S-band, where the system performances in terms of coverage and rate probability were analyzed. In [94], the authors studied a system-level performance analysis for the sharing systems, where the ergodic spectrum efficiency of LEO-SatNets is provided when the SatNet spectrum is shared with TNs. However, in these works, the impact of complex environments, such as in urban areas, is not considered. Furthermore, these works focus on snap-shot model analysis without considering the dynamics of systems.

Besides, the resource allocation in spectrum-sharing systems has been studied in many works [16, 17, 45–48]. In [16, 45–47], the snap-shot-based spectrum-sharing scenarios for ISTNs have been studied. In particular, the authors in [45] studied cooperative and competitive

scenarios between TN and SatNet systems, where the second-price auction mechanism is designed to achieve equilibrium between channel sharing from TNs and offloading ability from SatNets. In [46], the authors studied the NOMA ISTNs where one satellite provides the backhaul link for BSs and the access link for UEs in the C-band. The system bandwidth (BW) is shared between the backhaul and access links. Subsequently, the user association (UA), BW allocation, and power allocation (PA) are optimized under the QoS and backhaul link constraints. In [47], the author examined various TN-NTN spectrum-sharing scenarios. Subsequently, the interference-aware and group-sharing mechanisms are studied for NTN UL transmission in the reverse paring spectrum sharing scenario, which aims to group and assign BSs and UEs to resource blocks. The authors in [16] studied the rate splitting multiple access (RSMA)-based spectrum sharing systems, wherein the SR maximization problems in RSMA systems are considered for underlay and overlay spectrum sharing scenarios. However, the snap-shot-based models in these works can not consider the dynamics of networks, especially the LEO-SatNets. Furthermore, these studies are limited to single-satellite scenarios.

Subsequently, the time-varying systems have been studied in [17,48]. In [48], author proposed a dynamic spectrum-sharing mechanism in which bandwidth allocation for each network is determined based on its load, with coordination managed by a spectrum management server. The authors in [17] investigated beam management and scheduling in coexistence networks, focusing on LSat transmissions to satellite cells. In this scenario, TNs and SatNets operate in separate RFBs, with the TN band temporarily shared with SatNets. To address this, the study proposed sequential solutions for handover, beam selection, and spectrum sharing. However, the analysis was limited to the system level, without considering BS/LSat-to-UE transmissions, which may pose challenges in accounting for UE mobility. Furthermore, existing studies on coexistence and spectrum-sharing systems predominantly rely on statistical channel models, which limit their ability to accurately capture environmental characteristics in the system model. Although the TN-based UA/PA problems have been extensively studied in literature such as [95,96], the joint optimization of BS/LSat-UA/PA in co-primary spectrum-sharing ISTNs, incorporating a realistic channel model based on an actual three-dimensional (3D) urban map, remains largely unexplored. This chapter aims to address this research gap.

3.1.2 Research Contributions

In this work, we study a time-window (TW)-based UA and PA mechanism for ISTNs in urban environments wherein TNs and SatNets use the same RFB in the C-band. The main contributions of this work can be summarized as follows.

- We analyze time-varying systems and transmission models for ISTNs operating in the same RFB. Particularly, the utilized channel and antenna models are especially compatible with the ray-tracing (RayT) mechanism and the actual 3D maps, which can apply to various areas thanks to the 3D map availability in the *OpenStreetMap* resources [97]. Based on which we formulate a multi-objective optimization problem that aims to maximize the sum-rate (SR) while minimizing the number of connection state changes (CCs), with joint optimization of BS/LSat-UA and PA. This problem is categorized as mixed-integer non-convex programming, which presents challenges for direct solutions.

- To address this problem, we propose two iterative algorithms using the successive-convex-approximation (SCA) method. The first algorithm can provide the solution over the entire TW but requires knowledge of the channel gain for the entire TW, which can be challenging in practice. For practical implementation, we further introduce the second prediction-based iterative algorithm. Specifically, we develop an RayT-based mechanism to sequentially predict the channel gain, enabling the solution of consecutive sub-TW problems.

- For practical simulations, we utilize a realistic 3D map of London to assess the impact of the urban environment on ISTN channels and performance. Additionally, UE mobility and network load are modeled based on actual routes from Google Navigator and population data. The simulation results provide valuable insights into the impact of the urban environment with the blockage, reflection, and diffraction effects on the time-varying link budget. Moreover, the numerical results demonstrate the superiority of our proposed algorithms over benchmark schemes in solving the joint SR-CC optimization problem. Furthermore, the small gap in performance between the two proposed algorithms highlights the effectiveness of the prediction-based approach.

The preliminary result of this work were presented in [31] wherein only the UA for SR maximization problem was studied. The rest of this chapter is organized as follows. Section 3.2

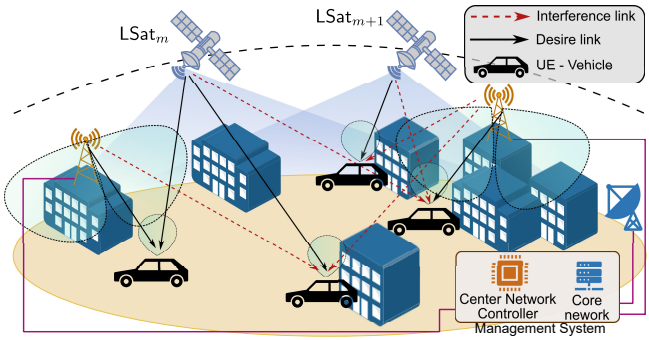


Figure 3.1: System model.

presents the system model and the problem formulation. Section 3.3 and 3.4 show the proposed algorithms, the benchmark algorithms, and the complexity analysis. Section 3.5 and 3.6 provide the simulation and numerical results and the conclusion, respectively.

3.2 System Model

In this work, we study ISTNs as per Fig. 3.1, where LSats and BSs jointly serve multiple automotive UEs. Specifically, we focus on urban environments where the impact of buildings is considered, and the UEs represent moving vehicles on city streets. According to the current spectrum regulation discussed above with the coexistence and overlap of TN and NTN DL transmission in the C-band, we study the DL operation of ISTNs wherein we assume that TNs and NTNs operate in the same RFB. Let $\mathcal{N} \triangleq \{1, \ldots, N\}$, $\mathcal{M} \triangleq \{1, \ldots, M\}$ and $\mathcal{K} \triangleq \{1, \ldots, K\}$ denote the sets of BSs, LSats, and UEs, respectively. In addition, let $\mathsf{BS_n}$, $\mathsf{LSat_m}$ and $\mathsf{UE_k}$ be the $n-$th BS, $m-$th LSat, and $k-$th UE, respectively. To account for other UEs in the system, we also assume the presence of terrestrial UEs (TUEs) and satellite UEs (SUEs), served by BSs and LSats, respectively. Additionally, we employ a window-based model consisting of N_{TS} time slots (TSs), where the duration of each TS is T_{S}. During this TW, each automotive vehicle UE is moving in its predetermined route. Subsequently, based on the satellite orbit and UE movement, we assume that the time alignment at LSats is perfect. Assume that each LSat operates in a quasi-fixed-beam mode and LSats point their beam towards the designated serving area [8]. Additionally, one assumes that BSs and satellite gateways are connected to a management system, wherein each node can exchange data with the core network and the center network controller (CNC) can manage the system resource.

3.2.1 Tx-UE Channel Model

This subsection presents the channel model of BS-UE and LSat-UE links which follows the coexistence of terrestrial and satellite system scenarios described in 3GPP TR 38863 [9]. In particular, for simplification, let $h_{n,k,t}$ and $g_{m,k,t}$ be the effective channel gain coefficients of the $\mathsf{BS}_n - \mathsf{UE}_k$ and $\mathsf{LSat}_m - \mathsf{UE}_k$ links at TS t which consist of the path-loss and the radiation beam pattern created by the terminal antennas, respectively. Subsequently, the channel gain is calculated by the inverse of the corresponding path loss (PL).

Path-loss Model

Based on the multi-path channel model [30, 68], let Tx represent either BS or LSat, the equivalent PL of $\mathsf{Tx}_x - \mathsf{UE}_k$, $x \in \{m, n\}$ link at TS t can be modeled as

$$L_{x,k,t}^{\mathsf{Tx-UE}} = -\log_{10}\left(\left|\sum_{i=0}^{N_{\mathsf{ray}}} \frac{G_{\mathsf{r}}(\varphi_{x,k,i}^{\mathsf{a},t}, \theta_{x,k,i}^{\mathsf{a},t}) G_{\mathsf{t}}(\varphi_{x,k,i}^{\mathsf{d},t}, \theta_{x,k,i}^{\mathsf{d},t})}{L_{x,k,i}^{\mathsf{pro},t} L_{x,k}^{\mathsf{B},t}} e^{-j\phi_{x,k,i}^{t} + j2\pi t T_{\mathsf{S}} f_{x,k,i}^{\mathsf{D},t}}\right|^{-1}\right), \tag{3.1}$$

wherein N_{ray} is the number of propagation rays from the Tx to an UE. $L_{x,k,i}^{\mathsf{pro},t}$, $\phi_{x,k,i}^{t}$, and $f_{x,k,i}^{\mathsf{d},t}$ are the propagation loss, phase delay, and Doppler shift of ray i from Tx_x to UE_k at TS t. For BS-UE links, we assume the Doppler shift is negligible. For LSat-UE links, due to the very high altitude compared to the distance between UEs and scatterers on the ground, one assumes that the Doppler shifts of all rays are identical, i.e., $f_{m,k,i}^{\mathsf{D},t} = f_{m,k}^{\mathsf{D},t}, \forall m, k, t$ [68]. Additionally, thanks to the satellite orbit and UE position knowledge, one assumes that the Doppler shift can be compensated at the LEO satellite payload [8, 49, 98]. Besides, $G_{\mathsf{r}}(\cdot)$ and $G_{\mathsf{t}}(\cdot)$ are the antenna gain patterns of the UE and the Tx. $(\varphi_{x,k,i}^{\mathsf{a},t}, \theta_{x,k,i}^{\mathsf{a},t})$ and $(\varphi_{x,k,i}^{\mathsf{d},t}, \theta_{x,k,i}^{\mathsf{d},t})$ are the elevation and azimuth angles of arrive at UE_k and departure at Tx_x corresponding to ray i at TS t. Especially, ray i, $i \geq 1$, is the LoS ray or the ray reaching UE by reflection or diffraction on obstacles in the environment. In case existing obstacles, ray $i = 0$ is the ray propagating through the wall to reach the UE, the channel gain loss of this ray can be modeled as twice of the wall-loss as [99]

$$L_{x,k,0}^{\mathsf{pro},t} = L_{x,k}^{\mathsf{FS},t} + 2\Big(5 - 10\log\left(0.7 \cdot 10^{\frac{-L_{\mathsf{IRRglass}}}{10}} + 0.3 \cdot 10^{\frac{-L_{\mathsf{concrete}}}{10}}\right)\Big), \tag{3.2}$$

$L_{x,k}^{\mathsf{FS},t}$ is the free-space PL, L_{IRRglass} and L_{concrete} are the penetration loss metrics of IRR glass and concrete materials, respectively. In addition, $L_{x,k}^{\mathsf{B},t}$ is the additional basic loss. For

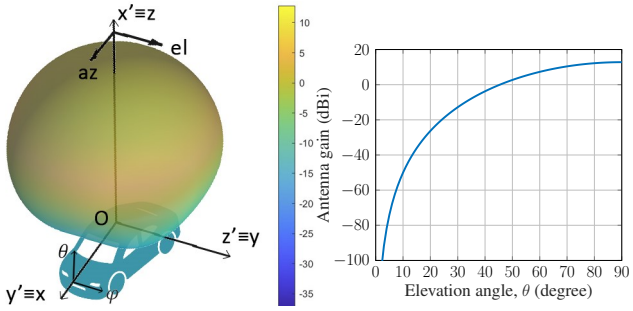


Figure 3.2: The gain pattern of UE's antenna.

$\mathsf{BS}_x - \mathsf{UE}_k$ link, since path-loss and shadow fading is accounted in $L^{\mathsf{pro},i}_{x,k,i}$ [99], $L^{\mathsf{B},t}_{x,k} = 1$. For $\mathsf{LSat}_x - \mathsf{UE}_k$ link, $L^{\mathsf{B},t}_{x,k}$ is computed by atmospheric absorption and attenuation of rain and clouds [65].

Antenna Model

In this work, we assume that the vehicle UE is equipped with a low-cost patch antenna on the rooftop with no beamforming capabilities, i.e. pointing to the Zenith (as shown in Fig. 3.2). Additionally, its radiation pattern can be modeled by employing a cosine pattern as

$$\bar{G}_{\mathsf{veh}}(az, el) = \cos^{2\bar{m}}(az)\cos^{2\bar{n}}(el), \tag{3.3}$$

where $\bar{m}$ and $\bar{n}$ are the order of the. az and el are the azimuth and elevation angles in the $\mathsf{Ox'y'z'}$ coordinate system in which its $\mathsf{Oy'z'}$ plane is mounted on the vehicle's rooftop. Besides, azimuth and elevation angles φ and θ in $G_{\mathsf{r}}(\varphi, \theta)$ in (3.1) are measured in the Oxyz coordinate system where its Oxy plane mounted on the vehicle's rooftop. Hence, it is noted that angles az and el can be computed from angles φ and θ by the coordinate rotation operation. For instance, the directivity gain pattern of the patch antenna in the Oxyz and $\mathsf{Ox'y'z'}$ coordinate systems is illustrated in Fig. 3.2. Regarding the LSat's antenna, its antenna gain pattern is modeled as [8].

$$G_{\mathsf{Sat}}(\theta) = \begin{cases} G^{\mathsf{max}}_{\mathsf{Sat}}, & \theta = 0, \\ G^{\mathsf{max}}_{\mathsf{Sat}} 4 \left| \frac{J_1(ka\sin\theta)}{ka\sin\theta} \right|^2, & \theta \neq 0, \end{cases} \tag{3.4}$$

wherein $G^{\mathsf{max}}_{\mathsf{S}}$ is the maximum LSat antenna gain, $J_1(\cdot)$ is the Bessel function of the first kind and first order, $k = 2\pi f_c/c$. In addition, a, f_c and c are the antenna aperture radius, operation frequency, and light speed, respectively. Regarding the BS, for simplicity, we employ the non-active antenna system pattern for the BS which is used in the coexistence scenarios examined

in 3GPP TR 38.863 [9]. Particularly, the beam pattern created by BS's antennas is modeled as

$$G_{\mathsf{BS}}(\theta,\varphi) = -\min\{-(A_{\mathsf{E,V}}(\theta) + A_{\mathsf{E,H}}(\varphi)), A_{\mathsf{m}}\} + G_{\mathsf{BS}}^{\mathsf{max}}, \tag{3.5}$$

where $A_{\mathsf{E,V}}(\theta)$ and $A_{\mathsf{E,H}}(\varphi)$ are the vertical and horizontal gain patterns which are expressed in [9], $G_{\mathsf{BS}}^{\mathsf{max}}$ is the maximum gain the the BS's antenna. Besides, a downtill angle of 10° is applied to the BS antenna in urban areas [9].

3.2.2 BS/LSat-UE Association Model

To indicate the BS-UA, we first introduce binary variable $\boldsymbol{\alpha} \triangleq [\alpha_{n,k,t}]_{\forall(n,k,t)}$ as

$$\alpha_{n,k,t} = \begin{cases} 1, & \text{if } \mathsf{UE}_k \text{ connects to } \mathsf{BS}_n \text{ at TS } t, \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

In addition, one assumes that each UE can connect to at most one BS at each TS, which is ensured as

$$(C1): \quad \sum_{\forall n} \alpha_{n,k,t} \leq 1, \quad \forall(k,t). \tag{3.7}$$

Regarding the available BS-UA at each TS, assuming that BS_n can serve at most ψ_n^{B} UEs at each TS. Then, the connection limitation at each BS can be cast by

$$(C2): \quad \sum_{\forall k} \alpha_{n,k,t} \leq \psi_n^{\mathsf{B}} - \eta_{n,t}^{\mathsf{B}}, \quad \forall(n,t), \tag{3.8}$$

wherein $\eta_{n,t}^{\mathsf{B}}$ is the number of TUEs[1] served by BS_n at TS t. Regarding the LSat-UA, one can see that UEs can connect to the LSat if it is in the field of view (FoV) of UEs. Let $\Theta_t = \{\forall(m,k) | \theta_{m,k,t} \geq \tilde{\theta}\}$ be the set of LSat-UE pairs satisfying the FoV criteria above at TS t, wherein $\theta_{m,k,t}$ is the elevation angle at UE_k toward LSat_m at TS t and $\tilde{\theta}$ is the minimum elevation angle requirement. Next, we introduce the LSat-UA variable $\boldsymbol{\beta} \triangleq [\beta_{m,k,t}]_{\forall(m,k,t)}$ corresponding to TS t as

$$\beta_{m,k,t} = \begin{cases} 1, & \text{if } (m,k) \in \Theta_t \text{ and } \mathsf{UE}_k \text{ connects to } \mathsf{LSat}_m, \\ 0, & \text{otherwise.} \end{cases} \tag{3.9}$$

[1]The presence of TUEs is used to emulate the congestion caused by other UEs at BSs.

Assuming that each UE can connect to at most one LSat at each TS while each LSat_m can serve at most ψ_m^{S} UEs at each TS. These yield the following constraints

$$(C3): \quad \sum_{\forall m} \beta_{m,k,t} \leq 1, \quad \forall (k,t), \tag{3.10}$$

$$(C4): \quad \sum_{\forall k} \beta_{m,k,t} \leq \psi_m^{\mathsf{S}} - \eta_{m,t}^{\mathsf{S}}, \quad \forall (m,t), \tag{3.11}$$

where $\eta_{m,t}^{\mathsf{S}}$ is the number of SUEs[2] served by LSat_m at TS t.

Regarding UE experience, to ensure the quality of experience (QoE) in terms of seamless connectivity and to guarantee that all UEs will be served, we assume that each UE has at least one connection at each TS, which yields the following constraint.

$$(C5): \quad \sum_{\forall n} \alpha_{n,k,t} + \sum_{\forall m} \beta_{m,k,t} \geq 1, \quad \forall (k,t). \tag{3.12}$$

Additionally, due to system dynamics, UEs must adapt their association with BSs and LSats to enhance throughput performance. Particularly, UEs may carry out the HO, establish an additional connection, or remove the existing connection. Based on the UA decision, the connection state transition procedure can be executed accordingly to the system standard [100]. Thanks to the UA, the number of connection state changes (CCs) is quantified by evaluating the difference in UA decisions between consecutive TSs as follows

$$f^{\mathsf{CC}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{\forall t>1} \Big(\sum_{\forall (n,k)} |\alpha_{n,k,t} - \alpha_{n,k,t-1}| + \sum_{\forall (m,k)} |\beta_{m,k,t} - \beta_{m,k,t-1}| \Big). \tag{3.13}$$

3.2.3 Transmission Model

This subsection describes the transmission from BSs and LSats to UEs. Since TNs and NTNs operate in the same RFB, the transmission involves intra-system interference (IaSI) and inter-system interference (ISyI). Depending on the standard in each system, based on the employed resource allocation and multiple access technique, we assume that BSs are coordinated to mitigate the TN IaSI such as in [77, 95, 96] and LSats are coordinated for that in NTNs such as in [101, 102]. However, due to the long distance between LSats and ground segments, TNs and NTNs are not coordinated to mitigate ISyI. Hence, for simplification, we assume that IaSI is negligible and only ISyI between TNs and NTNs is considered. Consequently, if UE_k connects to BS_n, the corresponding signal-to-interference-plus-noise-ratio (SINR) can be

[2] The presence of SUEs is used to emulate the congestion caused by other UEs at LSats.

expressed as

$$\gamma^{\mathsf{B}}_{n,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) = \frac{\alpha_{n,k,t}P^{\mathsf{B}}_{n,k,t}h_{n,k,t}}{\Xi_{k,t}(\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{S}})+\sigma_k^2}, \tag{3.14}$$

where σ_k^2 is the additive white Gaussian noise power and $\Xi_{k,t}(\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{S}}) = \sum_{\forall m}(\sum_{\forall k'}\beta_{m,k',t}P^{\mathsf{S}}_{m,k',t}+\eta^{\mathsf{S}}_{m,t}\bar{P}^{\mathsf{S}}_{m,t})g_{m,k,t}$ is ISyI caused by LSats at UE_k. Specifically, the first summation and the second term in the brackets are the total transmit power of LSat_m to K UEs and its own $\eta^{\mathsf{S}}_{m,t}$ SUEs, respectively. Particularly, $\boldsymbol{P}^{\mathsf{B}} \triangleq [P^{\mathsf{B}}_{n,k,t}]_{\forall(n,k,t)}$, $\boldsymbol{P}^{\mathsf{S}} \triangleq [P^{\mathsf{S}}_{m,k,t}]_{\forall(m,k,t)}$, $P^{\mathsf{B}}_{n,k,t}$ and $P^{\mathsf{S}}_{m,k,t}$ are the transmit powers of BS_n and LSat_m to UE_k at TS t, respectively. Additionally, $\bar{P}^{\mathsf{S}}_{m,t}$ is the transmit power of LSat_m to each of its $\eta^{\mathsf{S}}_{m,t}$ SUEs which is uniformly allocated as

$$\bar{P}^{\mathsf{S}}_{m,t} = P^{\mathsf{S,max}}_m/\min\{\psi^{\mathsf{S}}_m,\eta^{\mathsf{S}}_{m,t}+K\},\forall(m,t), \tag{3.15}$$

wherein $P^{\mathsf{S,max}}_m$ is the maximum power budget of LSat_m and $\min\{\psi^{\mathsf{S}}_m,\eta^{\mathsf{S}}_{m,t}+K\}$ indicates the maximum number of UEs which can connect to LSat_m at TS t. Subsequently, the throughput from BS_n to UE_k at TS t can be computed as

$$R^{\mathsf{B}}_{n,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) = \log\left(1+\gamma^{\mathsf{B}}_{n,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}})\right). \tag{3.16}$$

Regarding the LSat-UE connection, if UE_k connects to LSat_m, the corresponding SINR can be expressed as

$$\gamma^{\mathsf{S}}_{m,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) = \frac{\beta_{m,k,t}P^{\mathsf{S}}_{m,k,t}g_{m,k,t}}{\Phi_{k,t}(\boldsymbol{\alpha},\boldsymbol{P}^{\mathsf{B}})+\sigma_k^2}, \tag{3.17}$$

where $\Phi_{k,t}(\boldsymbol{\alpha},\boldsymbol{P}^{\mathsf{B}}) = \sum_{\forall n}\left(\sum_{\forall k'}\alpha_{n,k',t}P^{\mathsf{B}}_{n,k',t}+\eta^{\mathsf{B}}_{n,t}\bar{P}^{\mathsf{B}}_{n,t}\right)h_{n,k,t}$ is the ISyI caused by BSs, the first summation and the second term inside brackets are the total transmit power of BS_n to K UEs and its own $\eta^{\mathsf{B}}_{n,t}$ TUEs, respectively. Particularly, $\bar{P}^{\mathsf{B}}_{n,t}$ is the transmit power from BS_n to each of its $\eta^{\mathsf{B}}_{n,t}$ TUEs which is uniformly allocated as

$$\bar{P}^{\mathsf{B}}_{n,t} = P^{\mathsf{B,max}}_n/\min\{\psi^{\mathsf{B}}_n,\eta^{\mathsf{B}}_{n,t}+K\},\forall(n,t), \tag{3.18}$$

$P^{\mathsf{B,max}}_n$ is the maximum power budget of BS_n and $\min\{\psi^{\mathsf{B}}_n,\eta^{\mathsf{B}}_{n,t}+K\}$ is the maximum number of UEs which can connect to BS_n at TS t. Subsequently, the transmission throughput from

LSat_m to UE_k can be formulated as

$$R^{\mathsf{S}}_{m,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) = \log\left(1+\gamma^{\mathsf{S}}_{m,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}})\right). \tag{3.19}$$

Hence, the aggregated transmission throughput of UE_k at TS t can be expressed as

$$R_{k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) = \sum_{n} R^{\mathsf{B}}_{n,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) + \sum_{m} R^{\mathsf{S}}_{m,k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}). \tag{3.20}$$

Regarding the QoS of UEs, the rate threshold constraint should be considered. However, ensuring this constraint at each TS may be infeasible since the transmission links can be blocked by buildings in urban environments at certain TSs. Hence, one can guarantee the QoS of UEs by considering the minimum average rate constraint at consecutive periods as

$$(C6): \quad \frac{1}{N^{\mathsf{QoS}}_{\mathsf{TS}}} \sum_{\forall t \in \mathcal{T}_q} R_{k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) \geq \bar{R}_k, \forall k, q > 0,$$

wherein $\mathcal{T}_q = \{(q-1)N^{\mathsf{QoS}}_{\mathsf{TS}}+1,\ldots,qN^{\mathsf{QoS}}_{\mathsf{TS}}\}$ is the set of TSs in period q, each period consists of $N^{\mathsf{QoS}}_{\mathsf{TS}}$ TSs, and $\bar{R}_k$ is the average rate threshold of UE_k at each period.

3.2.4 Problem Formulation

In this subsection, we formulate a multi-objective optimization problem aiming to maximize the average system SR and minimize the average number of CCs per TS, wherein the UA and PA are jointly optimized.

Assuming that the deployment of BSs and LSats is sufficiently dense, UEs may have multiple connection options at each TS. However, due to the limitation of radio resources, BSs and LSats can overload at certain times depending on the number of connected devices. In addition, the system is highly dynamic due to UE mobility, LSat movement, and the complexity of urban environments, leading to frequent variations in link availability and quality over time.

Therefore, to ensure connectivity under such dynamics, UEs must appropriately adapt their associations with BSs and LSats through HO, establishing new connections, or releasing existing ones. Concurrently, due to the limitation of BS/LSat's transmit power and the impact of cross-system interference, an appropriate PA is essential.

Besides, each connection state transition incurs signaling overhead between UEs and the

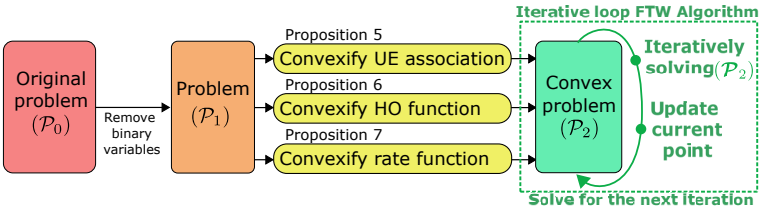


Figure 3.3: Flowchart of the full-time-window algorithm

corresponding BS/LSat. Furthermore, unnecessary state transitions may arise, which are undesirable and should be minimized. Therefore, in addition to enhancing throughput and enabling seamless HO, the optimization goal must also aim to minimize the number of CCs. Particularly, the optimization is mathematically formulated as

$$
\begin{aligned}
(\mathcal{P}_0): \quad & \max_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}} \frac{\rho}{N_{\mathsf{TS}}} \sum_{k,t} R_{k,t}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) - \frac{1-\rho}{N_{\mathsf{TS}}} f^{\mathsf{CC}}(\boldsymbol{\alpha},\boldsymbol{\beta}) \\
& \text{s.t.} \quad \text{constraints } (C1)-(C6), \\
(C0): \quad & \alpha_{n,k,t}, \beta_{m,k,t} \in \{0,1\}, \forall (m,n,k,t), \\
(C7): \quad & \eta^{\mathsf{B}}_{n,t} \bar{P}^{\mathsf{B}}_{n,t} + \sum_{\forall k} \alpha_{n,k,t} P^{\mathsf{B}}_{n,k,t} \le P^{\mathsf{B,max}}_{n}, \forall (n,t), \\
(C8): \quad & \eta^{\mathsf{S}}_{m,t} \bar{P}^{\mathsf{S}}_{m,t} + \sum_{\forall k} \beta_{m,k,t} P^{\mathsf{S}}_{m,k,t} \le P^{\mathsf{S,max}}_{m}, \forall (m,t),
\end{aligned}
$$

wherein $\rho \in (0,1)$ is the multi-objective factor, a higher ρ leads to higher priority of SR maximization. In addition, constraints $(C7)$ and $(C8)$ ensure that the transmit power does not exceed the BS and LSat power budgets, respectively. The outcome of solving this problem can provide solutions for UA and PA over time, helping to improve system performance. However, due to the non-convexity of the rate functions, and the coupling between binary and continuous variables, this problem belongs to mixed-integer non-linear programming which is difficult to solve directly.

3.3 Proposed Solutions

This section proposed two algorithms for the full time window and for the prediction time window. The developing procedure of the first algorithm is summarized in Fig. 3.3.

3.3.1 Problem Transformation

Analyzing the $(\mathcal{P}_0)$ structure and its optimization variables, we obtain the following relationships, each TS t

- If UE_k connects to BS_n, $P^{\mathsf{B}}_{n,k,t} > 0$, and $P^{\mathsf{B}}_{n,k,t} = 0$ if UE_k does not connect to BS_n.
- If UE_k connects to LSat_m, $P^{\mathsf{S}}_{m,k,t} > 0$, and $P^{\mathsf{S}}_{m,k,t} = 0$ if UE_k is not served by LSat_m.

Based on these relations, the binary variable can be represented by the ℓ_0-norm components of continuous ones as

$$\alpha_{n,k,t} = \|P^{\mathsf{B}}_{n,k,t}\|_0, \forall(n,k,t), \tag{3.20a}$$

$$\beta_{m,k,t} = \|P^{\mathsf{S}}_{m,k,t}\|_0, \forall(m,k,t). \tag{3.20b}$$

One notes that representations in (3.21) are valid for $\forall P^{\mathsf{B}}_{n,k,t} \geq 0$ and $\forall P^{\mathsf{S}}_{m,k,t} \geq 0$. Additionally, due to the FoV criteria for LSat-UA in (3.9), the LSat's power control elements are set as $P^{\mathsf{S}}_{m,k,t} = 0$, $\forall(m,k) \notin \Theta_t$. By employing the relations in (3.21), the binary terms in rate and SINR functions can be omitted. Subsequently, problem $\mathcal{P}_0$ can be rewritten as

$$\begin{aligned}
(\mathcal{P}_1): \max_{\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}},\boldsymbol{\lambda}^{\mathsf{B}},\boldsymbol{\lambda}^{\mathsf{S}}} \quad & \frac{\rho}{N_{\mathsf{TS}}}\tilde{\lambda} - \frac{1-\rho}{N_{\mathsf{TS}}}\bar{f}^{\mathsf{CC}}(\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) \\
\text{s.t.} \quad (\bar{C}1): \quad & \sum_{\forall n}\|P^{\mathsf{B}}_{n,k,t}\|_0 \leq 1, \forall(k,t), \\
(\bar{C}2): \quad & \sum_{\forall k}\|P^{\mathsf{B}}_{n,k,t}\|_0 \leq \psi^{\mathsf{B}}_n - \eta^{\mathsf{B}}_{n,t}, \forall(n,t), \\
(\bar{C}3): \quad & \sum_{\forall m}\|P^{\mathsf{S}}_{m,k,t}\|_0 \leq 1, \forall(k,t), \\
(\bar{C}4): \quad & \sum_{\forall k}\|P^{\mathsf{S}}_{m,k,t}\|_0 \leq \psi^{\mathsf{S}}_m - \eta^{\mathsf{S}}_{m,t}, \forall(m,t), \\
(\bar{C}5): \quad & \sum_{\forall n}\|P^{\mathsf{B}}_{n,k,t}\|_0 + \sum_{\forall m}\|P^{\mathsf{S}}_{m,k,t}\|_0 \geq 1, \forall(k,t), \\
(\tilde{C}6): \quad & \frac{1}{N^{\mathsf{QoS}}_{\mathsf{TS}}}\sum_{\forall t\in\mathcal{T}_q}\Big(\sum_{\forall n}\lambda^{\mathsf{B}}_{n,k,t} + \sum_{\forall m}\lambda^{\mathsf{S}}_{m,k,t}\Big) \geq \bar{R}_k, \forall(k,q), \\
(\tilde{C}7): \quad & \eta^{\mathsf{B}}_{n,t}\bar{P}^{\mathsf{B}}_{n,t} + \sum_{\forall k}P^{\mathsf{B}}_{n,k,t} \leq P^{\mathsf{B,max}}_n, \forall(n,t), \\
(\tilde{C}8): \quad & \eta^{\mathsf{S}}_{m,t}\bar{P}^{\mathsf{S}}_{m,t} + \sum_{\forall k}P^{\mathsf{S}}_{m,k,t} \leq P^{\mathsf{S,max}}_m, \forall(m,t), \\
(C9): \quad & \lambda^{\mathsf{B}}_{n,k,t} \leq R^{\mathsf{B}}_{n,k,t}(\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}), \quad \forall(n,k,t), \\
(C10): \quad & \lambda^{\mathsf{S}}_{m,k,t} \leq R^{\mathsf{S}}_{m,k,t}(\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}), \quad \forall(m,k,t),
\end{aligned}$$

where $\bar{f}^{\mathsf{CC}}(\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}})$ is rewritten from $f^{\mathsf{CC}}(\boldsymbol{\alpha},\boldsymbol{\beta})$ as

$$\bar{f}^{\mathsf{CC}}(\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}}) = \sum_{\forall t>1}\left(\sum_{\forall(n,k)}\Big|\|P^{\mathsf{B}}_{n,k,t}\|_0 - \|P^{\mathsf{B}}_{n,k,t-1}\|_0\Big| + \sum_{\forall(m,k)}\Big|\|P^{\mathsf{S}}_{m,k,t}\|_0 - \|P^{\mathsf{S}}_{m,k,t-1}\|_0\Big|\right). \tag{3.21}$$

In addition, $\tilde{\lambda} = \sum_{n,k,t} \lambda^{\mathsf{B}}_{n,k,t} + \sum_{\forall m,k,t} \lambda^{\mathsf{S}}_{m,k,t}$, and $\boldsymbol{\lambda}^{\mathsf{B}} \triangleq [\lambda^{\mathsf{B}}_{n,k,t}]_{\forall(n,k,t)}$ and $\boldsymbol{\lambda}^{\mathsf{S}} \triangleq [\lambda^{\mathsf{S}}_{m,k,t}]_{\forall(m,k,t)}$ are the slack variables which are considered the lower bounds of the rate functions as in constraints $(C9)$ and $(C10)$, respectively. It can be seen that problem $(\mathcal{P}_1)$ is still difficult to solve due to the nonconvex rate constraint $(C9)-(C10)$ and the sparse ℓ_0-norm terms. The sparsity in ℓ_0-norm terms and the non-convexity in rate constraints will be addressed by appropriate approaches in the following sub-sections.

3.3.2 Proposed Iterative Solution

Approximate ℓ_0-norm components

To address the sparsity in problem $(\mathcal{P}_1)$, the ℓ_0-norm components can be approximated by a concave function as

$$\|P^{\mathsf{B}}_{n,k,t}\|_0 \approx f_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}), \quad \forall(n,k,t), \tag{3.21a}$$

$$\|P^{\mathsf{S}}_{m,k,t}\|_0 \approx f_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}), \quad \forall(m,k,t), \tag{3.21b}$$

in which $f_{\mathsf{apx}}(x) \triangleq 1-e^{-\zeta x},\ x \geq 0$ is the approximate concave function, ζ is a positive number used to control the smoothness of the approximated function [77]. Thanks to (3.22), $(\bar{C}5)$ can be transformed into the convex constraint as

$$(\tilde{C}5): \quad \sum_{\forall n} f_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}) + \sum_{\forall m} f_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}) \geq 1, \quad \forall(k,t).$$

Proposition 5. *Constraints $(\bar{C}1)-(\bar{C}4)$ can be approximated respectively at iteration $(i+1)$ as*

$$(\tilde{C}1): \quad \sum_{\forall n} \tilde{f}^{(i)}_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}) \leq 1, \forall(k,t),$$

$$(\tilde{C}2): \quad \sum_{\forall k} \tilde{f}^{(i)}_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}) \leq \psi^{\mathsf{B}}_n - \eta^{\mathsf{B}}_{n,t}, \forall(n,t),$$

$$(\tilde{C}3): \quad \sum_{\forall m} \tilde{f}^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}) \leq 1, \forall(k,t),$$

$$(\tilde{C}4): \quad \sum_{\forall k} \tilde{f}^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}) \leq \psi^{\mathsf{S}}_m - \eta^{\mathsf{S}}_{m,t}, \forall(m,t),$$

wherein $\tilde{f}^{(i)}_{\mathsf{apx}}(x) \triangleq \zeta \exp(-\zeta x^{(i)})(x - x^{(i)}) - \exp(-\zeta x^{(i)}) + 1$ and $x^{(i)}$ is the feasible point at iteration i.

Proof: Thanks to (3.22), the ℓ_0-norm components in constraints $(\bar{C}1) - (\bar{C}4)$ can be approximated by $f_{\mathsf{apx}}(\cdot)$. However, the approximated constraints are nonconvex due to the concavity of $f_{\mathsf{apx}}(\cdot)$. Subsequently, by utilizing the first-order approximation, an upper bound of $f_{\mathsf{apx}}(x)$ around $x^{(i)}$ is obtained as

$$f_{\mathsf{apx}}(x) \leq \nabla f_{\mathsf{apx}}(x)|_{x=x^{(i)}}(x - x^{(i)}) + f_{\mathsf{apx}}(x^{(i)}) := \tilde{f}^{(i)}_{\mathsf{apx}}(x). \tag{3.22}$$

By employing (3.22), constraints $(\bar{C}1) - (\bar{C}4)$ are convexified as $(\tilde{C}1) - (\tilde{C}4)$, respectively. ■

Proposition 6. *An upper bound of the CC-number function can be expressed as*

$$\bar{f}^{\mathsf{CC}}(\boldsymbol{P}^{\mathsf{B}}, \boldsymbol{P}^{\mathsf{S}}) \leq \sum_{\forall t>1} \left(\sum_{\forall (n,k)} (a^{\mathsf{up}}_{n,k,t} - a^{\mathsf{low}}_{n,k,t}) + \sum_{\forall (m,k)} (b^{\mathsf{up}}_{m,k,t} - b^{\mathsf{low}}_{m,k,t}) \right) := \tilde{f}^{\mathsf{CC},(i)}(\boldsymbol{a}, \boldsymbol{b}), \tag{3.23}$$

where $\boldsymbol{a} = [a^{\mathsf{low}}_{n,k,t}, a^{\mathsf{up}}_{n,k,t}]_{\forall (n,k,t)}$ *and* $\boldsymbol{b} = [b^{\mathsf{low}}_{m,k,t}, b^{\mathsf{up}}_{m,k,t}]_{\forall (m,k,t)}$ *are the new slack variables satisfying constraints* $\forall (m,n,k,t)$*:*

$$\begin{aligned}
&(\tilde{C}11.1): a^{\mathsf{low}}_{n,k,t} \leq f_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}),\ a^{\mathsf{low}}_{n,k,t} \leq f_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t-1}),\\
&(\tilde{C}11.2): a^{\mathsf{up}}_{n,k,t} \geq f^{(i)}_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}),\ a^{\mathsf{up}}_{n,k,t} \geq f^{(i)}_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t-1}),\\
&(\tilde{C}12.1): b^{\mathsf{low}}_{m,k,t} \leq f_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}),\ b^{\mathsf{low}}_{m,k,t} \leq f_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t-1}),\\
&(\tilde{C}12.2): b^{\mathsf{up}}_{m,k,t} \geq f^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}),\ b^{\mathsf{up}}_{m,k,t} \geq f^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t-1}).
\end{aligned}$$

Proof: Consider function $f_{\mathsf{abs}}(x,y) = |\|x\|_0 - \|y\|_0|$, $x, y \geq 0$ which has the form of components in the CC-number function. $f_{\mathsf{abs}}(x,y)$ can be represented as $f_{\mathsf{abs}}(x,y) = \max\{\|x\|_0, \|y\|_0\} - \min\{\|x\|_0, \|y\|_0\}$. Let introduce slack variables u^{low} and u^{up} satisfying

$$u^{\mathsf{low}} \leq \|x\|_0,\ u^{\mathsf{low}} \leq \|y\|_0, \tag{3.23a}$$

$$u^{\mathsf{up}} \geq \|x\|_0,\ u^{\mathsf{up}} \geq \|y\|_0, \tag{3.23b}$$

an upper bound of function $f_{\mathsf{abs}}(x,y)$ can be obtained as

$$f_{\mathsf{abs}}(x,y) \leq u^{\mathsf{up}} - u^{\mathsf{low}}. \tag{3.24}$$

Additionally, thanks to the approximations in (3.22) and (3.22), constraint (3.24) is approxi-

mated as

$$u^{\mathsf{low}} \le f_{\mathsf{apx}}(x),\ u^{\mathsf{low}} \le f_{\mathsf{apx}}(y), \tag{3.24a}$$

$$u^{\mathsf{up}} \ge f^{(i)}_{\mathsf{apx}}(x),\ u^{\mathsf{up}} \ge f^{(i)}_{\mathsf{apx}}(y). \tag{3.24b}$$

Constraint (3.25) is convex since $f_{\mathsf{apx}}(\cdot)$ is concave and $f^{(i)}_{\mathsf{apx}}(\cdot)$ is affine. Subsequently, substituting this approximation for each absolute term in the CC-number function $\bar{f}^{\mathsf{CC}}(\boldsymbol{P}^{\mathsf{B}}, \boldsymbol{P}^{\mathsf{S}})$ setting $x = P^{\mathsf{B}}_{n,k,t}$, $y = P^{\mathsf{B}}_{n,k,t-1}$, $u^{\mathsf{low}} = a^{\mathsf{low}}_{n,k,t}$, and $u^{\mathsf{up}} = a^{\mathsf{up}}_{n,k,t}$ for the term $|\|P^{\mathsf{B}}_{n,k,t}\|_0 - \|P^{\mathsf{B}}_{n,k,t-1}\|_0|$, and $x = P^{\mathsf{S}}_{m,k,t}$, $y = P^{\mathsf{S}}_{m,k,t-1}$, $u^{\mathsf{low}} = b^{\mathsf{low}}_{m,k,t}$, and $u^{\mathsf{up}} = b^{\mathsf{up}}_{m,k,t}$ for term $|\|P^{\mathsf{S}}_{m,k,t}\|_0 - \|P^{\mathsf{S}}_{m,k,t-1}\|_0|$, we obtain the upper bound $\tilde{f}^{\mathsf{CC},(\mathrm{i})}(\boldsymbol{a}, \boldsymbol{b})$ of the CC-number function $\bar{f}^{\mathsf{CC}}(\boldsymbol{P}^{\mathsf{B}}, \boldsymbol{P}^{\mathsf{S}})$ as in (3.23) and convex constraints $(\tilde{C}11), (\tilde{C}12)$. ■

Convexify rate constraints

Rate constraints $(C9)$ and $(C10)$ can be convexified by the following proposition.

Proposition 7. *Constraints* $(C9), (C10)$ *can be convexified as*

$$(\tilde{C}9.1) : \ln(P^{\mathsf{B}}_{n,k,t} h_{n,k,t} + \Xi_{k,t}(\boldsymbol{P}^{\mathsf{S}}) + \sigma_k^2) \ge \lambda^{\mathsf{B}}_{n,k,t} + \mu^{\mathsf{B}}_{k,t},\ \forall (n,k,t),$$
$$(\tilde{C}9.2) : \Xi_{k,t}(\boldsymbol{P}^{\mathsf{S}}) + \sigma_k^2 \le \exp(\mu^{\mathsf{B},(i)}_{k,t})(\mu^{\mathsf{B}}_{k,t} - \mu^{\mathsf{B},(i)}_{k,t} + 1),\ \forall (k,t),$$
$$(\tilde{C}10.1) : \ln(P^{\mathsf{S}}_{m,k,t} g_{m,k,t} + \Phi_{k,t}(\boldsymbol{P}^{\mathsf{B}}) + \sigma_k^2) \ge \lambda^{\mathsf{S}}_{m,k,t} + \mu^{\mathsf{S}}_{k,t}, \forall (m,k,t),$$
$$(\tilde{C}10.2) : \Phi_{k,t}(\boldsymbol{P}^{\mathsf{B}}) + \sigma_k^2 \le \exp(\mu^{\mathsf{S},(i)}_{k,t})(\mu^{\mathsf{S}}_{k,t} - \mu^{\mathsf{S},(i)}_{k,t} + 1),\ \forall (k,t),$$

Proof: Let consider function $f_{\mathsf{rate}}(x,y) = \ln(1 + \frac{x}{y+c})$, $x, y \ge 0, c > 0$, and constraint $f_{\mathsf{rate}}(x,y) \ge z$. By expressing $f_{\mathsf{rate}}(x,y) = \ln(x+y+c) - \ln(y+c)$, constraint $f_{\mathsf{rate}}(x,y) \ge z$ can be decomposed as

$$\begin{cases} \ln(x+y+c) \ge z + u, & \text{(3.25a)} \\ \ln(y+c) \le u, & \text{(3.25b)} \end{cases}$$

wherein z is a slack variable. By taking the exponent operator on both sides of (3.25b), it is transformed into

$$y + c \le \exp(u). \tag{3.26}$$

It can be seen that (3.26) is still non-convex. To convexify (3.26), we employ the first-order

approximation to the exponential term and obtain the following convex constraint

$$\begin{cases} \ln(x+y+c) \geq z+u, & \text{(3.27a)} \\ y+c \leq \exp(u^{(i)})(u-u^{(i)}+1), & \text{(3.27b)} \end{cases}$$

wherein $u^{(i)}$ is the feasible point of u at iteration i.

By setting $x = P^{\mathsf{B}}_{n,k,t} h_{n,k,t}$, $y = \sum_{\forall m} \sum_{\forall k'} P^{\mathsf{S}}_{m,k',t} g_{m,k,t}$, $c = \sum_{\forall m} \eta^{\mathsf{S}}_{m,t} \bar{P}^{\mathsf{S}}_{m,t} g_{m,k,t} + \sigma^2_k$, $z = \lambda^{\mathsf{B}}_{n,k,t}$, and $u = \mu^{\mathsf{B}}_{k,t}$, we obtained the convex approximated form of constraint $(C8)$ as constraints $(\tilde{C}8.1) - (\tilde{C}8.2)$. Subsequently, by setting $x = P^{\mathsf{S}}_{m,k,t} g_{m,k,t}$, $y = \sum_{\forall n} \sum_{\forall k'} P^{\mathsf{B}}_{n,k',t} h_{n,k,t}$, $c = \sum_{\forall n} \eta^{\mathsf{B}}_{n,t} \bar{P}^{\mathsf{B}}_{n,t} h_{n,k,t} + \sigma^2_k$, $z = \lambda^{\mathsf{S}}_{m,k,t}$, and $u = \mu^{\mathsf{S}}_{k,t}$, the convex approximated form of constraint $(C9)$ is derived as $(\tilde{C}9.1) - (\tilde{C}9.2)$. ■

Thanks to Propositions 5, 6, and 7, problem $(\mathcal{P}_1)$ can be transformed into a successive convex problem at iteration i as

$$(\mathcal{P}_2): \max_{\substack{\boldsymbol{P}^{\mathsf{B}}, \boldsymbol{P}^{\mathsf{S}}, \boldsymbol{\lambda}^{\mathsf{B}}, \boldsymbol{\lambda}^{\mathsf{S}}, \\ \boldsymbol{\mu}^{\mathsf{B}}, \boldsymbol{\mu}^{\mathsf{S}}, \boldsymbol{a}, \boldsymbol{b}}} \quad \frac{\rho}{N_{\mathsf{TS}}} \tilde{\lambda} - \frac{1-\rho}{N_{\mathsf{TS}}} \tilde{f}^{\mathsf{CC},(i)}(\boldsymbol{a}, \boldsymbol{b})$$
$$\text{s.t.} \quad (\tilde{C}1) - (\tilde{C}12),$$

with $\boldsymbol{\mu}^{\mathsf{B}} = [\mu^{\mathsf{B}}_{n,k,t}]_{\forall (n,k,t)}$ and $\boldsymbol{\mu}^{\mathsf{S}} = [\mu^{\mathsf{S}}_{m,k,t}]_{\forall (m,k,t)}$. By iteratively solving problem $(\mathcal{P}_2)$, we obtain the PA solution for problem $(\mathcal{P}_0)$. In addition, the UA solution can be retrieved as [77]

$$\alpha_{n,k,t} = \begin{cases} 1, \text{ if } P^{\mathsf{B}}_{n,k,t} \geq \epsilon, \\ 0, \text{ otherwise }, \end{cases}, \quad \beta_{m,k,t} = \begin{cases} 1, \text{ if } P^{\mathsf{S}}_{m,k,t} \geq \epsilon, \\ 0, \text{ otherwise }, \end{cases} \tag{3.28}$$

wherein ϵ is a small positive number. The proposed iterative algorithm for full time-window-based (FTW), so called FTW algorithm, is summarized in Algorithm 4. The convergence of this algorithm is discussed in the following proposition.

Proposition 8. *Algorithm 4 converges to a local optimal solution of problem* $(\mathcal{P}_0)$.

Proof: First, one needs to prove the convergence of Algorithm 4. Let $\Omega \triangleq (\boldsymbol{P}^{\mathsf{B}}, \boldsymbol{P}^{\mathsf{S}}, \boldsymbol{\lambda}^{\mathsf{B}}, \boldsymbol{\lambda}^{\mathsf{S}}, \boldsymbol{\mu}^{\mathsf{B}}, \boldsymbol{\mu}^{\mathsf{S}}, \boldsymbol{a}, \boldsymbol{b}$ be the variable tuple, and $\Omega^{(i)}$, $\Omega^{\star(i)}$, and $\mathcal{F}^{(i)}$ be the feasible point, optimal solution, and feasible set of $(\mathcal{P}_2)$ at iteration i. In addition, let $F^{(i)}(\Omega)$ denote the objective value of $(\mathcal{P}_2)$ with solution Ω at iteration i and feasible point $\Omega^{(i)}$. Since $(\mathcal{P}_2)$ is derived by employing the SCA technique, it satisfies the properties of SCA-based problem [103]. In particular, we have $\Omega^{\star(i)} \in \mathcal{F}^{(i)} \cap \mathcal{F}^{(i+1)}$. This leads to $F^{(i+1)}(\Omega^{\star(i+1)}) \geq F^{(i)}(\Omega^{\star(i)})$. Hence, the objective

Algorithm 4 FTW Algorithm for Solving Problem $(\mathcal{P}_0)$

1: Set $i=0$ and generate an initial point $(\boldsymbol{P}^{\mathsf{B},(0)}, \boldsymbol{P}^{\mathsf{S},(0)}, \boldsymbol{\mu}^{\mathsf{B},(0)}, \boldsymbol{\mu}^{\mathsf{S},(0)})$ by solving problem (3.29).
2: **repeat**
3: Solve problem $(\mathcal{P}_2)$ to obtain $(\boldsymbol{P}^{\mathsf{B},\star}, \boldsymbol{P}^{\mathsf{S},\star}, \boldsymbol{\mu}^{\mathsf{B},\star}, \boldsymbol{\mu}^{\mathsf{S},\star})$.
4: Update $(\boldsymbol{P}^{\mathsf{B},(i)}, \boldsymbol{P}^{\mathsf{S},(i)}, \boldsymbol{\mu}^{\mathsf{B},(i)}, \boldsymbol{\mu}^{\mathsf{S},(i)}) := (\boldsymbol{P}^{\mathsf{B},\star}, \boldsymbol{P}^{\mathsf{S},\star}, \boldsymbol{\mu}^{\mathsf{B},\star}, \boldsymbol{\mu}^{\mathsf{S},\star})$.
5: Set $i = i+1$.
6: **until** Convergence
7: Recovery UA solutions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ by (3.28).
8: **Output:** The solution for problem $(\mathcal{P}_0)$ $(\boldsymbol{\alpha}^\star, \boldsymbol{\beta}^\star, \boldsymbol{P}^{\mathsf{B},\star}, \boldsymbol{P}^{\mathsf{S},\star})$.

value sequence is non-decreasing. Additionally, the transmit power is bounded by constraints $(\bar{C}7)-(\bar{C}8)$, which prove the convergence of Algorithm 4. Subsequently, with $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ recovered from $(\boldsymbol{P}^{\mathsf{B}}, \boldsymbol{P}^{\mathsf{S}})$ by (3.28), one can see that the feasible set of $(\mathcal{P}_2)$ is subset of that of $(\mathcal{P}_0)$. Therefore, Algorithm 4 converges to a local solution of $(\mathcal{P}_0)$. ■

To solve $(\mathcal{P}_2)$, one requires an initial point $(\boldsymbol{P}^{\mathsf{B},(0)}, \boldsymbol{P}^{\mathsf{S},(0)}, \boldsymbol{\mu}^{\mathsf{B},(0)}, \boldsymbol{\mu}^{\mathsf{S},(0)})$ satisfying $(\tilde{C}1)-(\tilde{C}12)$. To obtain an initial point, we can relax and penalize constraints $(\tilde{C}5)$, $(\tilde{C}6)$ [104], and solve the following problem

$$\min_{\substack{\boldsymbol{P}^{\mathsf{B}},\boldsymbol{P}^{\mathsf{S}},\boldsymbol{\lambda}^{\mathsf{B}},\boldsymbol{\lambda}^{\mathsf{S}},\\ \boldsymbol{\mu}^{\mathsf{B}},\boldsymbol{\mu}^{\mathsf{S}},\boldsymbol{\tau},\boldsymbol{\chi}}} \quad \sum \tau_{k,t} + \sum \chi_{k,q} \tag{3.28a}$$

$$\text{s.t.} \quad (\tilde{C}1)-(\tilde{C}4), (\tilde{C}7)-(\tilde{C}10),$$

$$\sum_{\forall n} f_{\mathsf{apx}}(P^{\mathsf{B}}_{n,k,t}) + \sum_{\forall m} f_{\mathsf{apx}}(P^{\mathsf{S}}_{m,k,t}) + \tau_{k,t} \geq 1, \quad \forall (k,t), \tag{3.28b}$$

$$\frac{1}{N^{\mathsf{QoS}}_{\mathsf{TS}}} \sum_{\forall t \in \mathcal{T}_q} \Big(\sum_{\forall n} \lambda^{\mathsf{B}}_{n,k,t} + \sum_{\forall m} \lambda^{\mathsf{S}}_{m,k,t}\Big) + \chi_{k,q} \geq \bar{R}_k, \quad \forall (k,q), \tag{3.28c}$$

$$\tau_{k,t} \geq 0, \chi_{k,q} \geq 0, \quad \forall (k,t,q), \tag{3.28d}$$

where $\boldsymbol{\tau} = \{\tau_{k,t}\}_{\forall (k,t)}$ and $\boldsymbol{\chi} = \{\chi_{k,q}\}_{\forall (k,q)}$ are slack variables. An initial point can be found by solving problem (3.29) until convergence.

3.3.3 Prediction-based Solution

The implementation of the proposed FTW algorithm above requires the channel gain information of all TSs in the considered time window which is difficult to achieve in practice if the number of TSs is sufficiently large. Hence, in this sub-section, we further propose a prediction-based solution. The overall approach of this solution is described in Fig. 3.4. Particularly, the entire TW is divided into multiple sub-TWs with the size of $N^{\mathsf{pred}}_{\mathsf{TS}}$, yielding $N^{\mathsf{pred}}_{\mathsf{TW}} = \lceil N_{\mathsf{TS}}/N^{\mathsf{pred}}_{\mathsf{TS}} \rceil$ sub-TWs. Subsequently, the channel gain information of each sub-

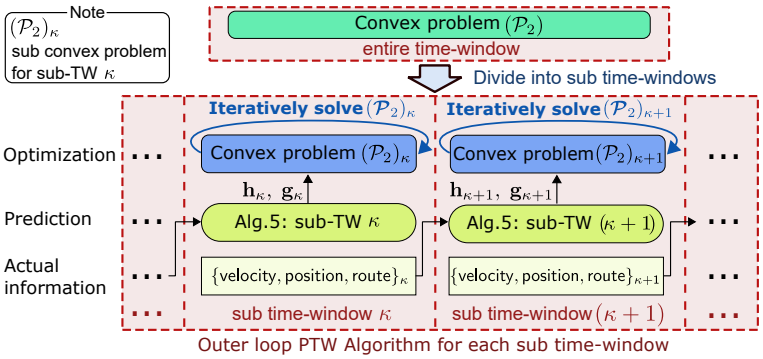


Figure 3.4: Flowchart of the prediction-time-window algorithm

TW is predicted based on the UE's information in the previous one. For convenience, let $\mathcal{T}_\kappa^{\text{pred}} = \{(\kappa-1)N_{\text{TS}}^{\text{pred}}+1,\ldots,\kappa N_{\text{TS}}^{\text{pred}}\}$ be the set of TSs in sub-TW κ.

Supposing that the satellite orbit information, BS's position, and 3D map knowledge are available, one can further assume that the channel gain of LSat/BS-UE links can be identified by employing the RayT mechanism for the given UE's position. However, to obtain the channel gain information in a sub-TW while avoiding the frequent update of UE's position at each TS, one necessitates predicting the UE position based on UE's information in the past. It is worth noting that UE is the vehicle, for the given UE route which can be obtained by the navigator application, e.g., Google Navigator, the UE position can be predicted as follows. Regarding the UE mobility, it can be seen that the changing of the UE's velocity and trajectory in each sub-TW is negligible for a sufficiently small window size. Hence, let $v_{k,t}$ and $\hat{v}_{k,t}$ be the actual and predicted velocities of UE_k at TS t, UE_k's velocity in sub-TW κ is predicted as

$$\hat{v}_{k,t} = \frac{1}{N_{\text{TS}}^{\text{pred}}} \sum_{\forall t' \in \mathcal{T}_{\kappa-1}^{\text{pred}}} v_{k,t'}, \quad \forall t \in \mathcal{T}_\kappa^{\text{pred}}, \forall k. \tag{3.29}$$

Subsequently, let $d_{k,t}$ be the distance from the starting point to the UE_k's position at TS t along the given route, the predicted distance $\hat{d}_{k,t}$ in sub-TW κ−th can be identified as

$$\hat{d}_{k,t} = d_{k,t'} + T_{\text{S}}\hat{v}_{k,t}, \quad \forall t \in \mathcal{T}_\kappa^{\text{pred}}, \tag{3.30}$$

where $t' = (\kappa-1)N_{\text{TS}}^{\text{pred}}$ is the TS before sub-TW κ. Based on the predicted UE position, the channel gain information can be identified by the RayT mechanism as the above assumption. The channel gain prediction procedure is summarized as in Algorithm 5. Given the channel

gain in sub-TW κ, the optimization problem in sub-TW κ is revised from problem $(\mathcal{P}_2)$ as

$$(\mathcal{P}_2)_\kappa : \max_{\substack{\boldsymbol{P}_\kappa^{\mathsf{B}}, \boldsymbol{P}_\kappa^{\mathsf{S}}, \boldsymbol{\lambda}_\kappa^{\mathsf{B}}, \boldsymbol{\lambda}_\kappa^{\mathsf{S}}, \\ \boldsymbol{\mu}_\kappa^{\mathsf{B}}, \boldsymbol{\mu}_\kappa^{\mathsf{S}}, \boldsymbol{a}_\kappa, \boldsymbol{b}_\kappa}} \quad \frac{\rho}{N_{\mathsf{TS}}^{\mathsf{pred}}} \tilde{\lambda}_\kappa - \frac{1-\rho}{N_{\mathsf{TS}}^{\mathsf{pred}}} \tilde{f}^{\mathsf{CC},(i)}(\boldsymbol{P}_\kappa^{\mathsf{B}}, \boldsymbol{P}_\kappa^{\mathsf{S}})$$
$$\text{s.t.} \quad (\tilde{C}1)_\kappa - (\tilde{C}12)_\kappa,$$

wherein the added low suffix $(\cdot)_\kappa$ in the constraint and variable notations indicates the adaptation of TS range which is limited within sub-TW κ. For simplification, we assume that sub-TW size $N_{\mathsf{TS}}^{\mathsf{pred}}$ is equal to integer times $N_{\mathsf{TS}}^{\mathsf{QoS}}$. To obtain the solution for problem $(\mathcal{P}_0)$, channel gain prediction is executed and problem $(\mathcal{P}_2)_\kappa$ is solved for each sub-TW. Particularly, at the end of sub-TW $(\kappa-1)$, channel gain for sub-TW κ is predicted and problem $(\mathcal{P}_2)_\kappa$ is solved. The proposed prediction-based algorithm, so-called prediction-based time-window (PTW) algorithm, is described in Algorithm 6. Its convergence is discussed in the following proposition.

Proposition 9. *Algorithm 6 is guaranteed for convergence to a local optimal solution of problem* $(\mathcal{P}_0)$.

Proof: To proof the convergence of Algorithm 6, one needs to proof the independence among problems $(\mathcal{P}_2)_\kappa$ of different sub-TWs and the convergence in solving each problem $(\mathcal{P}_2)_\kappa$. Let $\Omega_\kappa = (\boldsymbol{P}_\kappa^{\mathsf{B}}, \boldsymbol{P}_\kappa^{\mathsf{S}}, \boldsymbol{\lambda}_\kappa^{\mathsf{B}}, \boldsymbol{\lambda}_\kappa^{\mathsf{S}}, \boldsymbol{\mu}_\kappa^{\mathsf{B}}, \boldsymbol{\mu}_\kappa^{\mathsf{S}}, \boldsymbol{a}_\kappa, \boldsymbol{b}_\kappa)$ be the variable tuple of problem $(\mathcal{P}_2)_\kappa$. It can be seen that there is no coupling between variable tuples of problems $(\mathcal{P}_2)_\kappa$ and $(\mathcal{P}_2)_{\kappa'}, \forall \kappa \neq \kappa'$, i.e., $\Omega_\kappa \cap \Omega_{\kappa'} = \varnothing$. Therefore, problem $(\mathcal{P}_2)_\kappa$ for each sub-TW κ can be solved independently. In addition, problem $(\mathcal{P}_2)_\kappa$ has the same structure with problem $(\mathcal{P}_2)$. Hence, solving each problem $(\mathcal{P}_2)_\kappa$ as in steps 4 of Algorithm 6 is guaranteed to converge, and the aggregated solution for all sub-TWs is the local optimal solution of problem $(\mathcal{P}_0)$ which can be proved similarly to the proof of Proposition 8. ■

Remark 6. *The proposed FTW and PTW algorithms are designed based on the average joint max-sum-rate and min-CC-number problems* $(\mathcal{P}_\epsilon)$ *and* $(\mathcal{P}_2)_\kappa$*, however, they can easily adapt to the average joint max-sum-rate and min-CC-number objective with simple modifications. Particularly, let* λ_t^{min} *be the slack variable indicating the minimum UE rate at TS* t*. Indeed, by the following modifications:*

- *replacing* $\tilde{\lambda}$ *by* λ_t^{min} *in the objective function,*

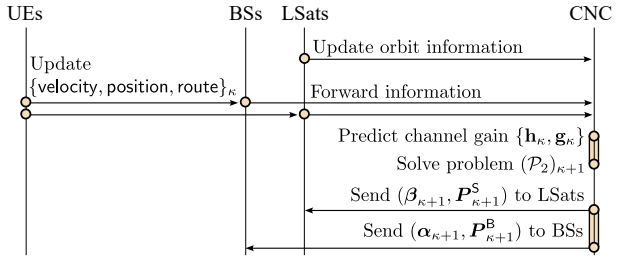


Figure 3.5: Information exchange in the prediction-based algorithm

- *replacing the minimum rate requirement* $\bar{R}_k$ *by* $\lambda_t^{\min}$ *in constraint* $(\tilde{C}6)$ *and* $(\tilde{C}6)_\kappa$ *of problems* $(\mathcal{P}_2)$ *and* $(\mathcal{P}_2)_\kappa$*, respectively,*

the proposed FTW and PTW algorithms can be directly applied to solve average joint max-min-rate and min-CC-number problems.

3.3.4 Prediction-based Algorithm Implementation

To clarify the practical implementation, Fig. 3.5 depicts the update and execution procedures of the prediction-based algorithm. In particular, at the end of each sub-TW κ, UEs update their velocity, position, and route via the connected BS/LSat while LSats update their orbit information to the CNC. Subsequently, the CNC executes Algorithm 5 to predict the channel gain for the next sub-TW $(\kappa+1)$. Based on which the CNC executes problem $(\mathcal{P}_2)_{\kappa+1}$ to find the solution for the next sub-TW $(\kappa+1)$. Since the update and execution procedures occur only once in each sub-TW, this approach reduces the signaling amount between nodes and the computation frequency at CNC. Additionally, depending on the stability of satellite orbits, LSats can update information less frequently instead of at every sub-TW which further reduces the data exchange frequency over the long propagation link between LSats and ground segments. Furthermore, the implementation analysis with the specific parameter setting is provided in the simulation results section.

Remark 7. *Thanks to updating the UE route at each sub-TW, the PTW algorithm can adapt to the changes of the UE route in practice which avoids the outdated route information provided at the beginning time.*

Remark 8. *In practical implementation, the resource allocation task should be completed within a finite time horizon, which may vary depending on the practical systems. For simplification, we assume that the computation capacity of the CNC is adequate and the time horizon for the resource allocation task is extensive enough.*

Algorithm 5 CHANNEL GAIN PREDICTION

1: **Input:** Satellite orbit, BS's positions, 3D map data, UE's routes, sub-TW index κ, and UE's velocity and position in the previous sub-TWs.
2: Predict $\hat{v}_{k,t}$ and $\hat{d}_{k,t}$ of UEs in sub-TW κ by (3.29) and (3.30).
3: Identify UE positions from $\hat{d}_{k,t}$ with respect to given UE's routes.
4: Identify channel gain from UE position by RayT mechanism.
5: **Output:** The channel gain information of UEs in sub-TW κ.

Algorithm 6 PTW ALGORITHM FOR SOLVING PROBLEM $(\mathcal{P}_0)$

1: Set $\kappa = 1$.
2: **while** $\kappa \leq N_{\mathsf{TW}}^{\mathsf{pred}}$ **do**
3: Predict the channel gain in sub-TW κ by Algorithm 5.
4: Solve problem $(\mathcal{P}_2)_\kappa$ by Algorithm 4 with adaptation for sub-TW κ.
5: Set $\kappa = \kappa + 1$.
6: **end while**
7: Recovery UA solutions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ by (3.28).
8: **Output:** The solution $(\boldsymbol{\alpha}^\star, \boldsymbol{\beta}^\star, \boldsymbol{P}^{\mathsf{B},\star}, \boldsymbol{P}^{\mathsf{S},\star})$.

3.4 Benchmarks and Complexity Analysis

3.4.1 Greedy Algorithm

In this benchmark algorithm, the UA is selected based on the channel gain while the transmit power of each BS/LSat is allocated uniformly. For BS-UE connection, at each TS, the $\mathsf{BS_n}$ – $\mathsf{UE_k}$ link is chosen based on the descending channel gain order if: 1) $\mathsf{UE_k}$ is still not connected to any BSs to ensure $(C1)$ and 2) the total number of connected and selected UEs at $\mathsf{BS_n}$ is still less than Ψ_n^{B} to ensure $(C2)$. The same procedure is conducted for the LSat-SUE connection such that constraints $(C3)$ and $(C4)$ are satisfied. Since only the channel gain-based UA is considered, the QoS constraint $(C6)$ is not guaranteed. The Greedy Algorithm is described in Algorithm 7. For convenience, the MATLAB notation is used to perform the matrix index.

3.4.2 Other Benchmark Algorithm

For comparison purposes, we introduce a benchmark algorithm based on our previous work [31]. In [31], we studied a similar system model to that in this work wherein the UA is optimized with a fixed transmit power. However, in [31] the objective is only SR maximization. To adapt the algorithm proposed in [31] to this work, we aim to modify this algorithm as follows. To consider the multi-objective in problem $(\mathcal{P}_0)$, the CC minimization objective $f^{\mathsf{CC}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is added, wherein multi-objective factors ρ and $(1-\rho)$ are added into SR and CC objectives. Subsequently, proposition 6 is applied to $f^{\mathsf{CC}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$, where components $|\alpha_{n,k,t} - \alpha_{n,k,t-1}|$ and $|\beta_{m,k,t} - \beta_{m,k,t-1}|$ are convexified similar to the procedure applied to $|\|\boldsymbol{P}_{n,k,t}^{\mathsf{B}}\|_0 - \|\boldsymbol{P}_{n,k,t-1}^{\mathsf{B}}\|_0|$

Algorithm 7 GREEDY ALGORITHM

Input: Channel gain $\mathbf{h} = [h_{n,k,t}]_{\forall(n,k,t)}$ and $\mathbf{g} = [g_{m,k,t}]_{\forall(m,k,t)}$, and parameters $\{\psi_n^{\mathsf{B}}, \psi_m^{\mathsf{S}}, \eta_{n,t}^{\mathsf{B}}, \eta_{m,t}^{\mathsf{S}}\}_{\forall(m,n,t)}$.
Initialize: Zero matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$
for Each TS t **do**
  while $\mathbf{h}(:,:,t) \neq \mathbf{0}$ **do**
    Find index $(\hat{n}, \hat{k})$ satisfying $\mathbf{h}(\hat{n}, \hat{k}, t) = \mathsf{max}(\mathbf{h}(:,:,t))$.
    if $\sum_{\forall k} \alpha_{\hat{n},k,t} \leq \psi_{\hat{n}}^{\mathsf{B}} - \eta_{\hat{n},t}^{\mathsf{B}}$
    then Update $\alpha_{\hat{n},\hat{k},t} = 1$ and $\mathbf{h}(:,\hat{k},t) = \mathbf{0}$
    else Update $\mathbf{h}(\hat{n},:,t) := \mathbf{0}$
  end while
  while $\mathbf{g}(:,:,t) \neq \mathbf{0}$ **do**
    Find index $(\hat{m}, \hat{k})$ satisfying $\mathbf{g}(\hat{m}, \hat{k}, t) = \mathsf{max}(\mathbf{g}(:,:,t))$.
    if $\sum_{\forall k} \beta_{\hat{m},k,t} \leq \psi_{\hat{m}}^{\mathsf{S}} - \eta_{\hat{m},t}^{\mathsf{S}}$
    then Update $\beta_{\hat{m},\hat{k},t} = 1$ and $\mathbf{g}(:,\hat{k},t) = \mathbf{0}$
    else Update $\mathbf{g}(\hat{m},:,t) := \mathbf{0}$
  end while
end for
Output: UA solution $\boldsymbol{\alpha}, \boldsymbol{\beta}$.

and $|\|\boldsymbol{P}_{m,k,t}^{\mathsf{S}}\|_0 - \|\boldsymbol{P}_{m,k,t-1}^{\mathsf{S}}\|_0|$, respectively. Additionally, the transmit power at each BS/LSat is fixed and uniformly allocated. This modified algorithm is named the full-time-window UA (FWUA) algorithm.

3.4.3 Complexity Analysis

This section analyzes the complexity of the proposed algorithms and benchmark algorithms described above regarding big-O notation. One notes that FTW, PTW, and FWUA algorithms involve solving convex problems. Assuming that the interior method is employed to solve the convex problems, the complexity in terms of big-O is $\mathcal{O}(c^{1/2}(c+v)v^2)$ where c and v are the numbers of inequality constraints and variables, respectively [105].

FTW Algorithm

As described in Algorithm 4, the FTW Algorithm includes one loop in **Steps 2-6**, wherein problem $(\mathcal{P}_2)$ is solved iteratively. Since problem $(\mathcal{P}_2)$ consists of $v_{\mathsf{FTW}} = 3NKN_{\mathsf{TS}} + 3MKN_{\mathsf{TS}} + NN_{\mathsf{TS}} + MN_{\mathsf{TS}}$ variables and $c_{\mathsf{FTW}} = 4NKN_{\mathsf{TS}} + 4MKN_{\mathsf{TS}} + 2NN_{\mathsf{TS}} + 2MN_{\mathsf{TS}} + 5KN_{\mathsf{TS}} + KN_{\mathsf{TS}}/N_{\mathsf{TS}}^{\mathsf{QoS}}$ inequality constraints, the complexity of the FTW Algorithm can be expressed as

$$X_{\mathsf{FTW}} = \mathcal{O}(N_{\mathsf{iter}} c_{\mathsf{FTW}}^{1/2}(c_{\mathsf{FTW}} + v_{\mathsf{FTW}})v_{\mathsf{FTW}}^3), \tag{3.31}$$

wherein N_{iter} is the required number iteration for convergence.

PTW Algorithm

Regarding Algorithm 6, the PTW Algorithm consists of two loops: the outer loop corresponds to the number of sub-TWs $N_{\mathsf{TW}}^{\mathsf{pred}}$ and the inner loop relates to solving problem $(\mathcal{P}_2)_\kappa$ iteratively. Since problem $(\mathcal{P}_2)_\kappa$ consists of $v_{\mathsf{PTW}} = 3NKN_{\mathsf{TS}}^{\mathsf{pred}} + 3MKN_{\mathsf{TS}}^{\mathsf{pred}} + NN_{\mathsf{TS}}^{\mathsf{pred}} + MN_{\mathsf{TS}}^{\mathsf{pred}}$ variables and $c_{\mathsf{PTW}} = 4NKN_{\mathsf{TS}}^{\mathsf{pred}} + 4MKN_{\mathsf{TS}}^{\mathsf{pred}} + 2NN_{\mathsf{TS}}^{\mathsf{pred}} + 2MN_{\mathsf{TS}}^{\mathsf{pred}} + 5KN_{\mathsf{TS}}^{\mathsf{pred}} + KN_{\mathsf{TS}}^{\mathsf{pred}}/N_{\mathsf{TS}}^{\mathsf{QoS}}$ inequality constraints, the complexity of the PTW Algorithm can be expressed as

$$X_{\mathsf{PTW}} = \mathcal{O}(N_{\mathsf{iter}} N_{\mathsf{TW}}^{\mathsf{pred}} c_{\mathsf{PTW}}^{1/2} (c_{\mathsf{PTW}} + v_{\mathsf{PTW}}) v_{\mathsf{PTW}}^3). \tag{3.32}$$

Greedy Algorithm

Regarding Algorithm 7, the Greedy Algorithm consists of one outer-loop with N_{TS} loops and two inner-loops for BS-UE and LSat-UA, respectively. In the first inner-loop in **Step 4-8**, one needs to execute K times until $\mathbf{h}(:,:,t) = \mathbf{0}$ in worst case, wherein finding and updating in **Step 5-8** demand a complexity of $\mathcal{O}(NK + N)$ for each time. The computation for each time in the second inner loop is identified similarly. In summary, the complexity of the Greedy Algorithm can be given as

$$X_{\mathsf{Gre.Alg.}} = \mathcal{O}(N_{\mathsf{TS}}(NK^2 + NK + MK^2 + MK)). \tag{3.33}$$

FWUA Algorithm

In the FWUA Algorithm, a convex problem for UA is solved iteratively until convergence. The convex problem associated with the FTWUA Algorithm consists of $v_{\mathsf{FWUA}} = 3NKN_{\mathsf{TS}} + 3MKN_{\mathsf{TS}} + NN_{\mathsf{TS}} + MN_{\mathsf{TS}}$ variables and $c_{\mathsf{FWUA}} = 5NKN_{\mathsf{TS}} + 5MKN_{\mathsf{TS}} + NN_{\mathsf{TS}} + MN_{\mathsf{TS}} + 5KN_{\mathsf{TS}} + KN_{\mathsf{TS}}/N_{\mathsf{TS}}^{\mathsf{QoS}}$ inequality constraints. Therefore, the complexity of the FTWUA Algorithm can be expressed as

$$X_{\mathsf{FWUA}} = \mathcal{O}(N_{\mathsf{iter}} c_{\mathsf{FWUA}}^{1/2} (c_{\mathsf{FWUA}} + v_{\mathsf{FWUA}}) v_{\mathsf{FWUA}}^3). \tag{3.34}$$

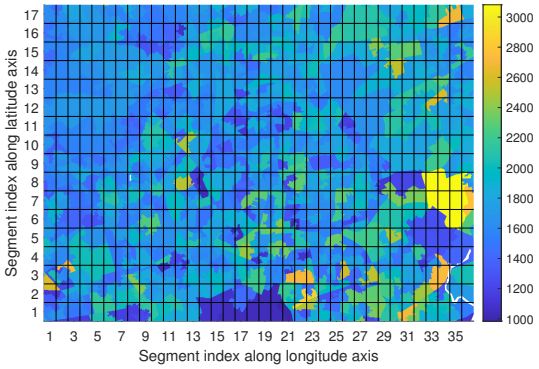


Figure 3.6: Population distribution in the entire examined area, the total population in each colored area is indicated by its color.

3.5 Numerical Results

3.5.1 Simulation Setup

The simulation is conducted in urban environments on an area in London city limited by latitude and longitude ranges of $[51.5115°\text{N}, 51.5965°\text{N}]$ and $[0.1772°\text{W}, 0.0028°\text{E}]$, respectively. Subsequently, the map is divided into a grid of segments where the size of each segment is $(0.005° \times 0.005°)$ in latitude and longitude, respectively. To evaluate the link budget across map locations, the receiver grid size in each map segment is (100×100) points. It is worth noting that UEs are the vehicles, thus receiver points are deployed outdoors. Regarding the transmitters, one assumes that one BS is deployed in each map segment at the highest building. In addition, one LEO constellation at 500 km is considered. To reduce the shadow area caused by high-buildings when LSats at low elevation angle and improve the coverage, one assumes that there are 2 LSats which can be used in the FoV at each instance time. Regarding UE mobility, the UE's routes are obtained by the navigator data in the Google Map application. Additionally, the number of TUEs at BS_n $\eta^{\mathsf{B}}_{n,t}$ and SUEs at LSat_m $\eta^{\mathsf{S}}_{m,t}$ are assumed to be Poisson random variables with means $\bar{\eta}^{\mathsf{B}}_n$ and $\bar{\eta}^{\mathsf{S}}_m$. Furthermore, to consider the heterogeneity of UE number in the considered area, the means number of TUEs at BSs $\bar{\eta}^{\mathsf{B}}_n$ are set proportionally to the population distribution in the area which is extracted from London's age data set [106] and depicted in Fig. 3.6. Particularly, the population in each map segment is computed based on the dataset, then $\bar{\eta}^{\mathsf{B}}_n$ for each map segment is scaled proportionally to its population so that $\max_n\{\bar{\eta}^{\mathsf{B}}_n\} = \bar{\eta}^{\mathsf{B}}_{\max}$. The key simulation parameters are described in Table. 3.1 wherein parameter set 1 for co-existence study in 3GPP TR 38-863

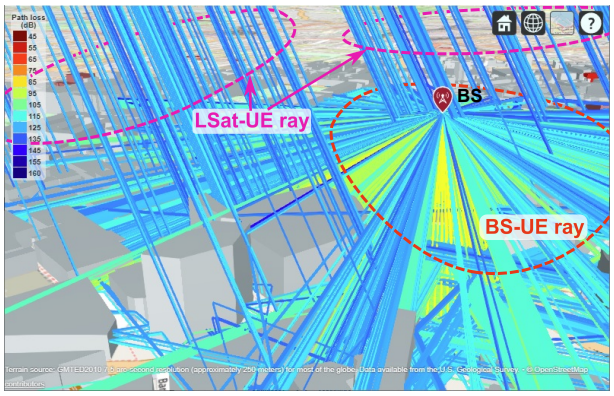


Figure 3.7: Raytracing result of BS/LSat-UE links.

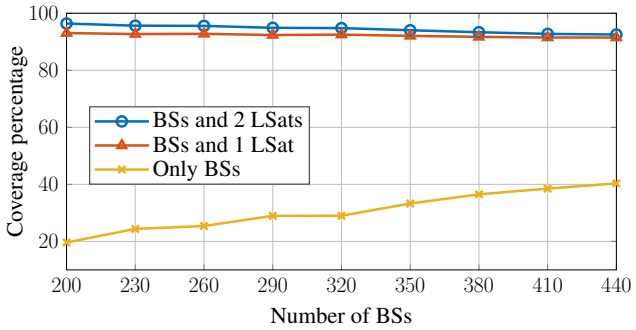
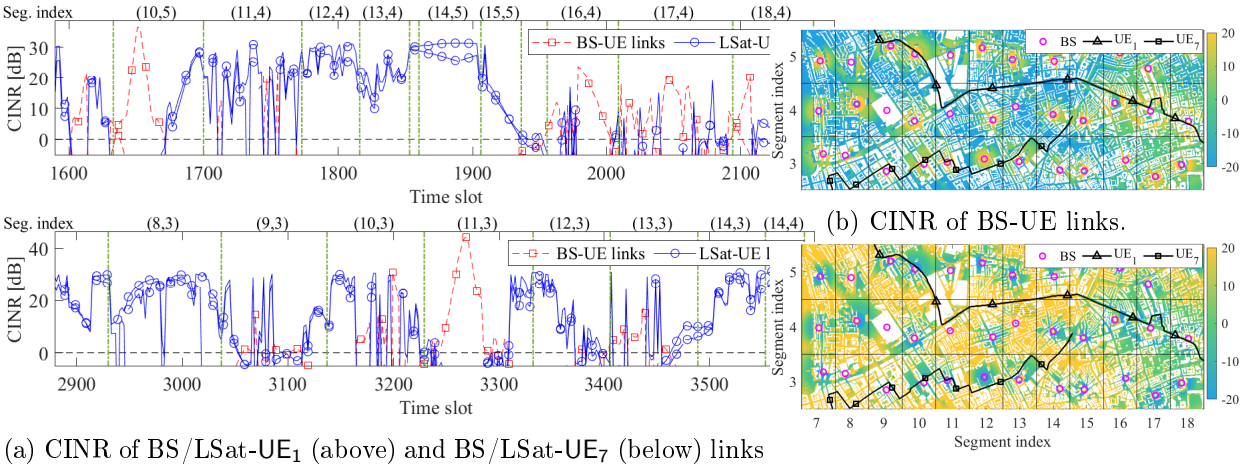


Figure 3.8: The coverage percentage (CINR $\geq$ 3 dB) versus the number of deployed BSs.

is utilized. In practical systems, if the power flux density limit is explicit, the maximum LSat transmit power could be adjusted appropriately. The simulations are carried out on a computer equipped with an *Intel Xeon E5 CPU @ 2.4 GHz* and *128 GB RAM*, wherein the iterative algorithms are implemented by using the MATLAB programming environment, leveraging the CVX modeling framework and MOSEK solver. For an intuitive view of complicated environments in urban areas, we first investigate the link budget assessment in terms of the CINR of BS/LSat-UE links with practical data. Subsequently, the numerical results are carried out to evaluate the performance in terms of SR and CC number optimization of the considered algorithms.

3.5.2 Link Budget Simulation Results

This section carries out the simulation to assess the link budget in terms of the CINR of BS/LSat-UE links in urban environments. The CINR of links can be calculated as described in [107]. Additionally, based on the outcomes of RayT simulation, i.e., propagation loss, phase



(a) CINR of BS/LSat-UE_1 (above) and BS/LSat-UE_7 (below) links over TSs.

(b) CINR of BS-UE links.

(c) CINR of LSat-UE links.

Figure 3.9: CINR of BS/LSat-UE_1 and BS/LSat-UE_7 links over TSs corresponding to the CINR heatmap in a small area.

delay, and AoA and AoD of rays, the equivalent PL of each link is computed by (3.1). To be more intuitive, the RayT simulation of the BS-UE and LSat-UE links is shown in Fig. 3.7, where the elevation angle at the receiver points to LSat is 60°. This shows the impact of building and the complicated propagation of rays from BS/LSat to the receiver points in urban environments, which is difficult to determine mathematically.

To evaluate the effectiveness of ISTNs in improving coverage, Fig. 3.8 illustrates the percentage of coverage area with CINR ≥ 3 dB of different deployment options versus the number of BSs. Obviously, increasing the number of BSs improves coverage. However, based on the trend, one may require to deploy a large number of BSs to achieve more than 90% coverage. In contrast, using integrated BSs with 1 or 2 LSats can significantly increase the coverage percentage, i.e., more than 90% of the considered area. This also demonstrates the cost-effectiveness of ISTNs in improving coverage and facilitating seamless connectivity.

Fig. 3.9 shows the CINR heatmap of BS/LSat-UE links and the time-varying CINR of two UEs corresponding to their route in a considered area. Regarding the CINR heatmap, one can see the heterogeneity of both BS-UE and LSat-UE links. In particular, the CINR of BS-UE links is improved in areas with dense buildings and dense BS deployment. This is due to the blockage of the high and dense building to the LSat-UE links while the BS-UE links against the impact of the building by dense deployment of BSs as depicted in Fig. 3.7. Additionally, the CINR of LSat-UE links increases in the area with fewer buildings. Interestingly, despite the significant propagation loss due to the long distance of the LSat-UE links compared to BS-UE ones, its CINR is higher than that of the BS-UE link in certain areas. This phenomenon

Table 3.1: Simulation Parameters

Parameter	Value
Number of TSs, N_{TS}	3600 TSs
Sub-TW size, $N_{\mathsf{TS}}^{\mathsf{pred}}$	90 TSs
TS duration, T_{S}	0.5 s
Operation frequency, f_c	3.4 GHz
System bandwidth, B	20 MHz
Latitude limitation of area	$[51.5115°\text{N}, 51.5965°\text{N}]$
Longitude limitation of area	$[0.1772°\text{W}, 0.0028°\text{E}]$
Size of one map segment	$0.005° \times 0.005°$
Receiver point grid size in one segment	100×100 points
Number of BSs	467
BS antenna parameters	[9]
LEO constellation	500 km
LEO inclination angle	53°
Minimum elevation angle in FoV $\tilde{\theta}$	60°
Number of LSats at each time	2
LSat antenna aperture, a	1 m [8, 9]
Maximum LSat antenna gain, $G_{\mathsf{Sat}}^{\mathsf{max}}$	30 dBi [8, 9]
Number of UEs	12
UE antenna height	1 m
Maximum UE antenna gain	12.8 dBi
BS's power budget, $P_n^{\mathsf{B,max}}$	42 dBm
LSat's power budget, $P_m^{\mathsf{S,max}}$	14 dBW
UE receiver noise figure, G_f	1.2 dB
UE antenna temperature, T_a	150 K
Penetration loss of IRR glass, L_{IRRglass}	$23 + 0.3f$[GHz] [99]
Penetration loss of concrete, L_{concrete}	$5 + 4f$[GHz] [99]
Maximum number of served UEs at BSs, ψ_n^{B}	20
Maximum number of served UEs at LSats, ψ_m^{S}	100
Maximum mean TUE number at BSs, $\bar{\eta}_{\max}^{\mathsf{B}}$	18
Mean SUE number at LSats, $\bar{\eta}_m^{\mathsf{S}}$	95
Multi-objective factor, ρ	0.9
Smooth parameter of the approx. function, ζ	10

is due to the impact of the antenna gain pattern at the UE which is equipped with the patch antenna. In particular, the elevation angle at the UE towards the LSat can be seen as constant across the map due to the long UE-LSat distance as in Fig. 3.7. This elevation angle is in a range of about $60° - 90°$, which results in an antenna gain at UE of about $7.3 - 12.8$ dBi as shown in Fig. 3.2. However, except for locations near BSs, the BS-UE links are transmitted almost in the horizontal plane with a low elevation angle, i.e., about $20° - 30°$, leading to the low antenna gain at UE, i.e., about -26.3 dBi to -12.8 dBi. For instance, at locations wherein the elevation angles of LSat-UE and BS-UE links are 60° and 30°, the gap in antenna gain between two links is about 20.1 dBi. Therefore, in areas where there exists a significant

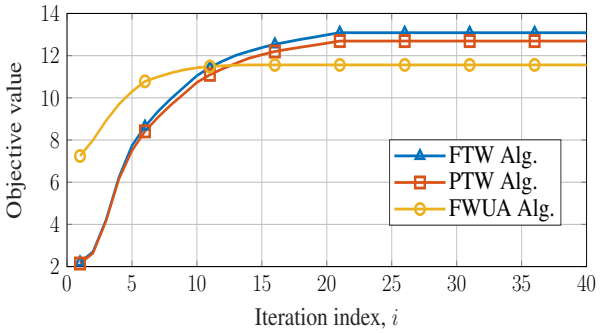


Figure 3.10: Convergence of the objective function.

antenna gain disparity, the antenna gain can compensate the distance gap between the LSat-UE and BS-UE links, allowing the LSat-UE links to outperform the BS-UE links in terms of CINR. Additionally, regarding time-varying CINR, one can see the correlation between the UE routes on the CINR heatmap and their CINR over TSs. This change in both BS/LSat-UE links is due to the combination of UE mobility, LSat movement, and building influence. Depending on the UE location, the quality of BS/LSat-UE links is different, especially the LSat-UE links which are time-dependence due to LSat movement. In particular, the CINR can be changed with different positive magnitudes in a range of about $5 - 30$ dB or even be negative, i.e., the link is blocked. This emphasizes the necessity of PA and UA optimization as formulated in problem $(\mathcal{P}_0)$.

3.5.3 Numerical Results

This section discusses the numerical results of the proposed and benchmark algorithms in solving the SR-CC problem. Fig. 3.10 shows the convergence of the objective function which consists of the average SR and the CC number of the FTW, PTW, and FTUA algorithms versus iterations. All considered algorithms require only a few tens iterations for convergence. Due to the optimization of only UA, the FWUA algorithm requires the smallest number of iterations for convergence among three, about 14 iterations. Additionally, the proposed FTW and PTW algorithms need a slightly higher number of iterations, i.e., about 23 and 24 iterations, respectively. In practical implementation, if the running time of the proposed algorithms is too long, one can stop the execution before convergence to obtain a good enough solution.

Fig. 3.11 depicts the Pareto curves of the average SR and CC number per TS of the considered algorithms versus multi-objective factor ρ. One can see that the Pareto curves of

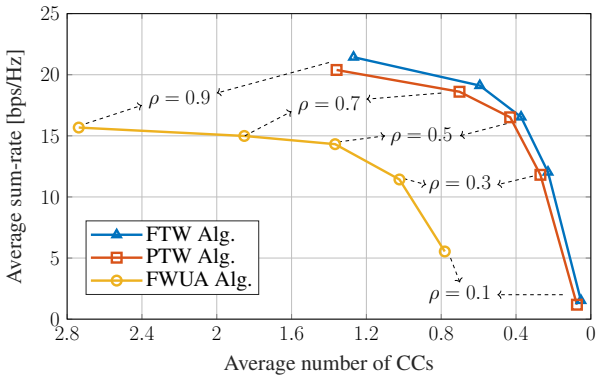


Figure 3.11: Pareto front of the average SR and CC number vs multi-objective factor ρ.

the two proposed algorithms are close and outperform that of the FWUA algorithm in terms of the SR and the CC number. Particularly, at $\rho = 0.9$, $\rho = 0.7$, and $\rho = 0.5$, both SR and the CC number results provided by the two proposed algorithms are better than those by the FWUA algorithm. At $\rho = 0.3$ and $\rho = 0.1$, the FWUA algorithm can provide a better SR compared to the two proposed algorithms, however, the CC number outcome is worse. Obviously, a higher ρ increases the priority of SR maximization and decreases that of CC number minimization. However, for the FTW and PTW algorithms, the Pareto curve increases quickly from $\rho = 0.1$ to $\rho = 0.7$, i.e., SR rapidly raises and the CC number slowly increases, whereas it slowly increases from $\rho = 0.7$ to $\rho = 0.9$. Particularly, SR increases about 10.6, 5.9, and 2.6 bps/Hz and the average CC number increases about 0.15, 0.23, and 0.35 when increasing $\rho = 0.1 \rightarrow 0.3$, $\rho = 0.3 \rightarrow 0.5$, and $\rho = 0.5 \rightarrow 0.7$, whereas the average SR and the CC number increase about 2.27 bps/Hz and 0.6 when increasing $\rho = 0.7 \rightarrow 0.9$, respectively. Based on this observation, we choose $\rho = 0.7$ to balance between the average SR and CC number.

Fig. 3.12 shows the statistic channel prediction error in terms of the mean absolute percentage error (MAPE). Obviously, a larger prediction window leads to a higher channel prediction error. Particularly, the MAPE values are about 0.07 and 0.14 at $N_{\mathsf{TS}}^{\mathsf{pred}} = 60$ TSs and $N_{\mathsf{TS}}^{\mathsf{pred}} = 180$ TSs, respectively. For more information on the impact of prediction window size, Fig. 3.13 illustrates the performance of the PTW Algorithm in terms of the average SR and the CC number with the different sizes of prediction sub-TW $N_{\mathsf{TS}}^{\mathsf{pred}}$ and those of the FTW and FWUA algorithms as the baselines. Regarding the CC, since the sub-TW is optimized independently, a longer prediction window size leads to a lower number of sub-TW, thus the CC number decreases. However, a larger window size results in less accuracy in channel gain

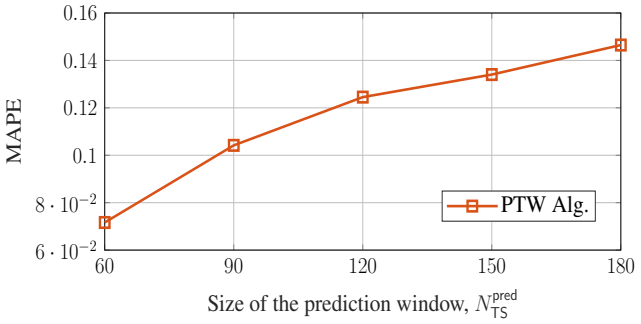


Figure 3.12: Mean absolute percentage error (MAPE) of Alg. 5 vs. prediction sub-TW size.

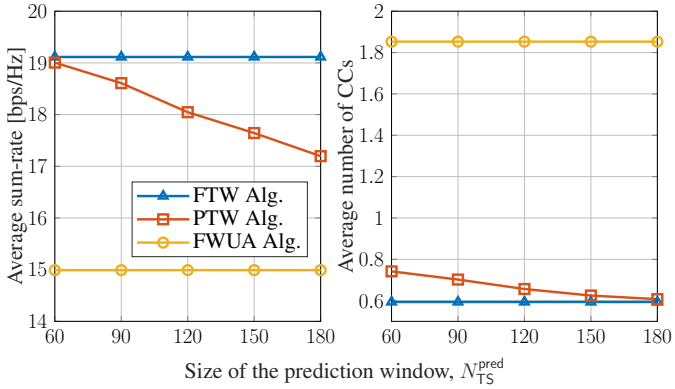


Figure 3.13: Sum-rate and CC number of Alg. 6 vs. prediction sub-TW size.

prediction as MAPE increases, leading to degradation of SR performance. Additionally, the UA decision is made based on the predicted channel, the higher prediction error may lead to more unsuitable UA decisions with the actual channel, which also contributes to the SR degradation. For instance, the SR degradation is about 2 bps/Hz and the average number of CCs decreases about 0.15 when raising $N_{\mathrm{TS}}^{\mathrm{pred}}$ from 60 to 180. This can be considered as the trade-off between SR and CC performances. However, these performances can be balanced and comparable with those of the FTW Algorithm by choosing an appropriate prediction sub-TW size.

Fig. 3.14 depicts the impact of the BS's power budget on the average SR and number of CCs. Regarding the CC results, one can see that the change in BS's power budget does not significantly affect the average number of CCs. However, two proposed FTW and PTW Algorithms outperform the two other benchmarks in terms of average CC number. In particular, the average CC number outcome of the two proposed algorithms is about 0.65 whereas those

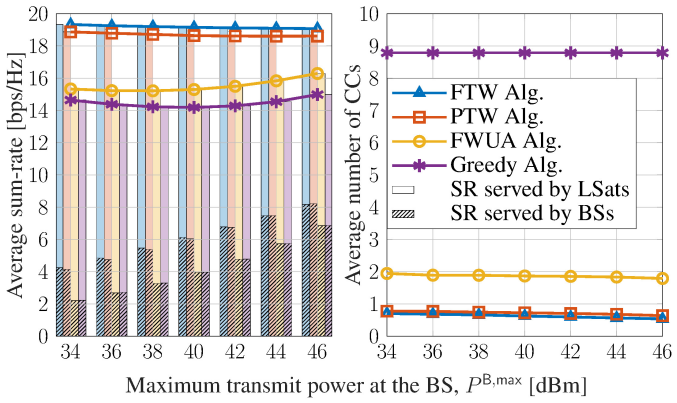


Figure 3.14: Sum-rate and CC number vs. BS maximum transmit power.

of the FWUA and Greedy Algorithms are about 1.9 and 8.7, respectively. Regarding the SR, the trends of the FWUA and Greedy Algorithms first decrease and then increase as the BS's power budget increases. This phenomenon arises because interference from BSs to LSat-UE links increases with the BS's power budget, becoming significant at particular BS power levels. Subsequently, at a sufficiently high BS power, the SR offered by BS-UE links dominates that by LSat-UE links, which leads to an increase in system SR. Additionally, increasing the BS's power budget can improve the BS-UE links. This is shown by the increase in the ratio of the SR contributed by BSs with the BS's power budget. Regarding the two proposed algorithms, the ratio is about 22% at $P^{\mathsf{B,max}} = 34$ dBm and 43% at $P^{\mathsf{B,max}} = 46$ dBm. Besides, due to the flexibility of PA in the FTW and PTW algorithms, the two proposed algorithms offer a significant SR improvement. Especially, the SR gap between the two lines is not significant, i.e., about 0.5 bps/Hz. For instance, SR outcomes of the FTW and PTW algorithms are about 19.1 and 18.6 bps/Hz at $P^{\mathsf{B,max}} = 40$ dBm, respectively.

For more insight into the impact of BS's power budget, Fig. 3.15 displays the average number of UEs served by BSs and LSats with the different $P^{\mathsf{B,max}}$. Regarding the two proposed algorithms, one can see that the average number of UEs served in BSs increases as $P^{\mathsf{B,max}}$ increases. For instance, that is about 3.8 and 4.7 UEs at $P^{\mathsf{B,max}} = 34$ dBm and $P^{\mathsf{B,max}} = 46$ dBm, respectively. Additionally, that number at LSats decreases as $P^{\mathsf{BS,max}}$ increases, which is about 8.1 and 7.3 UEs at $P^{\mathsf{B,max}} = 34$ dBm and $P^{\mathsf{B,max}} = 46$ dBm, respectively. This indicates that UEs priories BSs's connections to achieve a higher SR as the BS's power budget increases. Furthermore, the average numbers of served UEs at BSs and LSats provided by the proposed FTW and PTW Algorithms are less than that by the greedy algorithm. This emphasizes the effectiveness of the proposed algorithms wherein they require less number of connections

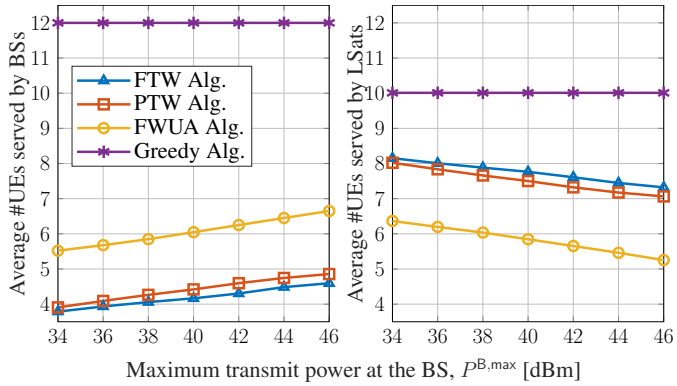


Figure 3.15: Served-UE number vs BS maximum transmit power.

while providing better SR compared to the two benchmark ones, i.e., the FWUA and greedy algorithms.

Fig. 3.16 shows the average SR and number of CCs versus the LSat's power budget. Regarding the CC number, this is similar to that in Fig. 3.14, wherein its change is not significant as different $P^{\mathsf{S,max}}$. Regarding the SR performance, the trends of the four considered algorithms are similar. In particular, as the power budget at LSats increases, the SR first drops due to the increase of interference from LSats to BS-UE links. Subsequently, the SR rises at a sufficiently high LSat power budget due to the domination quality of LSat-UE links compared to that of BS-UE links. For instance, at $P^{\mathsf{S,max}} = (6, 12, 16)$ dBW, the SR provided by the FWT and PTW Algorithms is approximately $(18.8, 18.4, 20.4)$ bps/Hz and $(18.4, 17.9, 19.8)$ bps/Hz, respectively. Furthermore, the SR contribution ratio of LSats increases with the LSat's power budget due to the LSat-UE link's improvement. Particularly, regarding the two proposed algorithms, LSats contributes about 24% and 74% of SR at $P^{\mathsf{S,max}} = 6$ dBW and $P^{\mathsf{S,max}} = 16$ dBW, respectively. Compared to SR at different $P^{\mathsf{BS,max}}$ in Fig. 3.14, the SR degradation in this figure shows a more critical impact of interference from LSats to BS-UE links, even with PA of the two proposed algorithms. This is explained by the impact of interference from LSats on BS-UE links in a vast area, whereas BSs cause interference on LSat-UE links locally due to the isolation by high buildings.

For more insight into the impact of LSat's power budget, Fig. 3.17 shows the average number of served UEs at BSs and at LSats with the different cases of $P^{\mathsf{S,max}}$. In the two proposed algorithms, the UEs reduce the connections to the BSs and increase those at LSats as the maximum LSat power increases since the SR offered by LSats is improved. Particularly, the average numbers of UEs served at BSs and LSats are about 5.7 and 6.3 at $P^{\mathsf{S,max}} = 6$

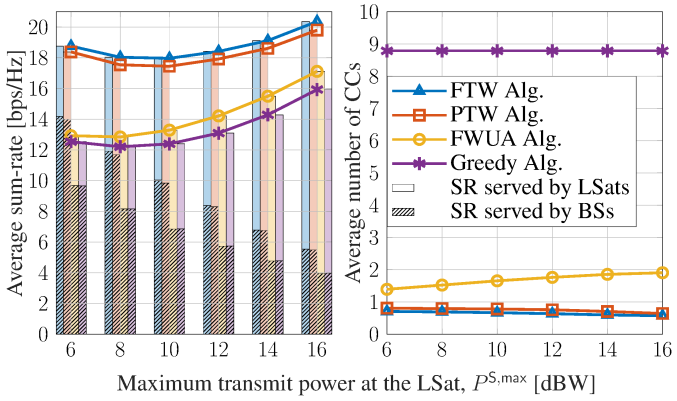


Figure 3.16: Sum-rate and the number of CCs versus the maximum transmit power at the LSat.

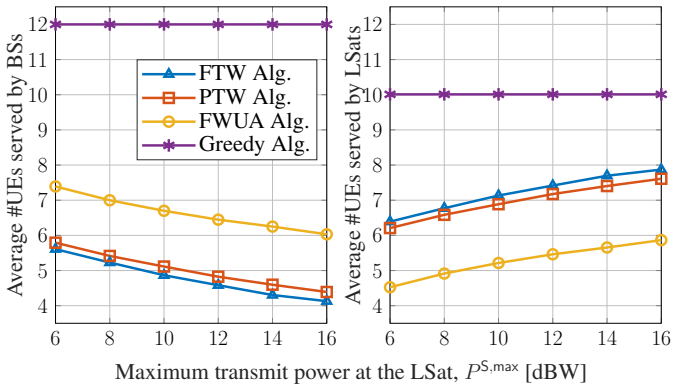


Figure 3.17: Average served-UE number vs LSat max transmit power.

dBW, and about 4.2 and 7.8 at $P^{\mathsf{S,max}} = 16$ dBW. Furthermore, compared to the benchmark algorithms, this figure again shows the effectiveness of the two proposed algorithms in terms of SR and connection number wherein their solutions have similar or fewer numbers of connections but better SR performance.

Subsequently, Fig. 3.18 illustrates the average SR and the number of CCs versus the different rate requirement $\bar{R}_k$. Regarding the FWUA algorithm, the SR performance changes insignificantly with different rate requirements. However, to satisfy the QoS constraint, the number of CCs increases as $\bar{R}_k$ increases. For the two proposed algorithms, as $\bar{R}_k$ increases, the SR outcomes decrease whereas the numbers of CCs increase. One can see that the increase in the rate requirement leads to a smaller feasible set of the SR-CC problem, which results in degradation in the SR and CC performance. Particularly, at $\bar{R}_k = 0.2$ bps/Hz, the average SR results of the FTW and PTW algorithms are about 20.12 bps/Hz and 19.45 bps/Hz, and the CC number is about 0.9; at $\bar{R}_k = 1$ bps/Hz, the those SR results are about 18.8 bps/Hz and 18.1 bps/Hz, and the CC number is about 1.7, respectively. However, at all considered rate requirement levels, the two proposed algorithms always outperform the FWUA and greedy

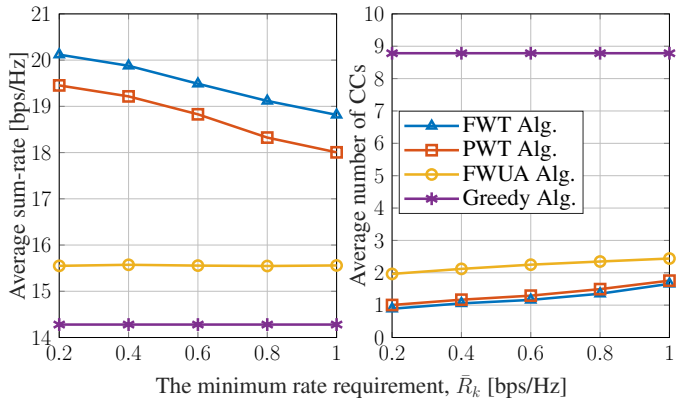


Figure 3.18: Sum-rate and CC number versus the minimum rate requirement.

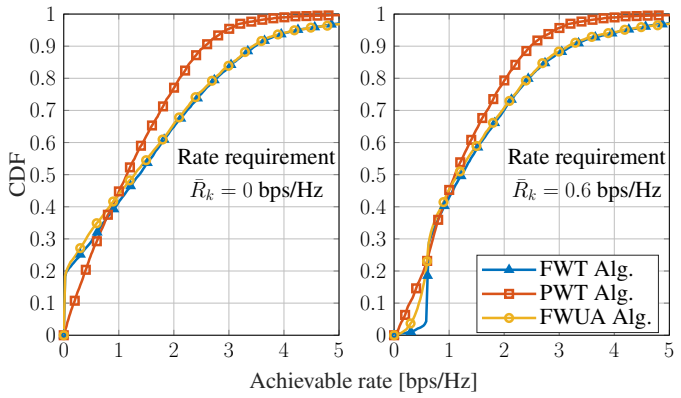


Figure 3.19: The cumulative distribution function of the achievable rate at the minimum rate requirements of $\bar{R}_k = 0$ bps/Hz and $\bar{R}_k = 0.6$ bps/Hz.

algorithms in terms of SR and the CC-number.

For more information on the impact of QoS requirements, Fig. 3.19 shows the cumulative distribution function (CDF) of the achievable rate with rate requirement $\bar{R}_k = 0$ bps/Hz and $\bar{R}_k = 0.6$ bps/Hz. Regarding the FWUA algorithm, the CDF of achievable rates at $\bar{R}_k = 0$ bps/Hz and $\bar{R}_k = 0.6$ bps/Hz are similar, which further clarifies the near constancy of SR in Fig. 3.18. In contrast, the CDFs of the rate outcomes of the FTW and PTW algorithms change significantly at $\bar{R}_k = 0$ bps/Hz and $\bar{R}_k = 0.6$ bps/Hz. Particularly, compared to the CDFs at $\bar{R}_k = 0$ bps/Hz, those at $\bar{R}_k = 0.6$ bps/Hz are forced to satisfy the minimum rate requirement. However, due to the channel prediction error in the PWT algorithm, one can see the degradation of the achievable rate, especially at lower than $\bar{R}_k = 0.6$ bps/Hz. However, the probabilities of rate outcomes of the FTW and PTW algorithms lower than $\bar{R}_k = 0.6$ bps/Hz are relatively small, i.e., about 0.03 and 0.2, respectively.

To assess computational efficiency, the average running time of the PTW algorithm with

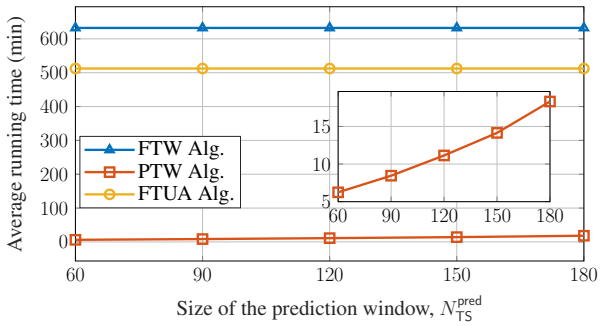


Figure 3.20: Average running time of examined iterative algorithms.

varying prediction window sizes $N_{\mathsf{TS}}^{\mathsf{pred}}$ and those of the benchmark, i.e., the FTW and FTUA algorithms, are shown in Fig. 3.20. Since the FTW and FTUA algorithms solve the problem over the entire time window, they incur an extremely high time for execution, i.e., about 632.3 and 512.4 minutes. In contrast, the PTW algorithm exhibits much lower running time. While its running time increases slightly with larger prediction window size, due to the increasing size of problem $(\mathcal{P}_{\in})_{\leq}$, it remains substantially more efficient than the FTW and FTUA ones. Specifically, for $N_{\mathsf{TS}}^{\mathsf{pred}} = (60, 90, 120, 150, 180)$ TSs, the average running time of the PTW algorithm is approximately 6.2, 8.4, 11.1, 14.2, and 18.31 minutes, respectively. Notably, the proposed PTW algorithm can achieve performance comparable to the FTW algorithm and even outperforms the FTUA one in terms of max-SR and min-CC-number, while requiring only a fraction of the running time, i.e, approximately $1\% - 3.5\%$ of the benchmarks' running time.

3.5.4 Practical Implementation Analysis

According to the simulation setting, assuming that the CNC is deployed at the center of the examined area, the distance between the CNC and BSs is about 7.8 km leading to the propagation time between CNC and BSs of $T_{\mathsf{CNC,BS}} \leq 0.03$ µs. Besides, due to the satellite orbit altitude of 500 km, the propagation time between the CNC and LSats is about $T_{\mathsf{CNC,LSat}} \approx 1.7$ ms. The time consumption due to propagation and signaling is depicted in Fig. 3.21. In the best case where UEs can update their status via BSs, this time consumption is about $T_{\mathsf{best}} = 2 \times T_{\mathsf{CNC,LSat}} = 3.4$ ms, while in the worst case where existing UEs which need to update their status via LSats this time is about $T_{\mathsf{worst}} = 3 \times T_{\mathsf{CNC,LSat}} = 5.1$ ms. Additionally, for TS duration of $T_{\mathsf{S}} = 0.5$ s and the sub-TW size of $N_{\mathsf{TS}}^{\mathsf{pred}} = 90$ TSs as in simulation setting, the sub-TW duration is $T_{\mathsf{sub-TW}} = 45$ s. One can see that the time consumption

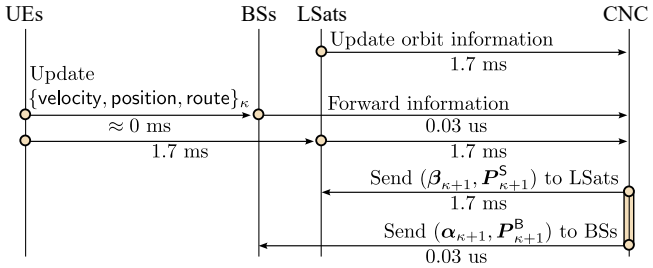


Figure 3.21: Propagation and signaling time consumption.

for propagation and signaling in the worst case is negligible compared to sub-TW duration. Hence, this further shows the practicality of the proposed PTW algorithm.

Besides, the segmentation of the map in the considered area can reduce the complexity of the prediction mechanism which facilitates scalability. Particularly, based on the outcome of the UE position prediction step, for each UE, one can determine the map segment corresponding to its position in a sub-TW. Based on which the RayT mechanism can execute for the identified map segment rather than the entire considered map. Furthermore, these steps can be executed in a parallel manner for UEs which further facilitates scalability.

3.6 Summary

This work studied the joint UA and PA for the practical co-primary spectrum sharing ISTNs in urban environments. Subsequently, the multi-objective of SR maximization and the number of CCs minimization under the resource limitation constraints is formulated. To solve this problem, we proposed two algorithms: 1) FTW: an iterative algorithm for the entire time-window and 2) PTW: a prediction-based algorithm for the implementation purpose that sequentially predicts the channel gain and solves the problem for sub-time-windows. For practice evaluation, simulation with the actual 3D map and UE mobility is performed. The simulation results provide insight into the link budget assessment in urban environments. Additionally, numerical results show the effectiveness of our two proposed algorithms compared to the greedy algorithm and the benchmark one (FTUA) in terms of both the system SR and the CC number.

The approach of using RayT, 3D map, and realistic mobility can be used as a framework for future works. Moreover, this study provides opportunities for further extensions, particularly in improving the channel model as well as addressing prediction error and intra-system-interference issues. Specifically, for practical applications, random impacts in the

actual environment can be characterised in the channel model by incorporating corresponding statistical components. Additionally, the prediction mechanism in the proposed PTW algorithm could be improved to reduce prediction error and SR degradation. For more in-depth analysis, the potential intra-system-interference issues should be considered in future research.

Chapter 4

Digital-Twin-based Resource Management for ISTNs

THE explosive growth in wireless service demand has prompted the evolution of ISTNs to overcome the limitations of traditional TNs in terms of coverage, spectrum efficiency, and deployment cost. Leveraging LEOsats and dynamic spectrum sharing (DSS), ISTNs offer promising solutions but face significant challenges due to diverse terrestrial environments, user and satellite mobility, and long propagation LEO-to-ground distance. To address these challenges, digit-twin (DT) has emerged as a promising technology to offer virtual replicas of real-world systems, facilitating prediction for resource management. In this work, we study a time-window-based DT-aided DSS framework for ISTNs, enabling joint long-term and short-term resource decisions to reduce system congestion. Based on that, two optimization problems are formulated, which aim to optimize resource management using DT information and to refine obtained solutions with actual real-time information, respectively. To efficiently solve these problems, we proposed algorithms using compressed-sensing-based and successive convex approximation techniques. Simulation results using actual traffic data and the London 3D map demonstrate the superiority in terms of congestion minimization of our proposed algorithms compared to benchmarks. Additionally, it shows the adaptation ability and practical feasibility of our proposed solutions.

4.1 Introduction

In recent years, the explosive growth in mobile traffic and emerging service demands, such as seamless and ubiquitous connectivity, have driven the evolution of wireless communication. While next-generation (NextG) networks are envisioned to meet these requirements [3, 21], relying solely on denser terrestrial networks (TNs) deployment faces significant challenges, including inefficient spectrum utilization in sparsely populated areas and high infrastructure costs. To overcome these limitations, non-terrestrial networks (NTNs), notably satellite networks (SatNet), can offer complementary coverage and traffic offloading. Leveraging both TN and SatNet capabilities, integrated satellite-terrestrial networks (ISTNs) have attracted much attention as a viable architecture to meet NextG requirements, advancing the goal of global, uninterrupted coverage. Backed by regulatory and institutional bodies such as the U.S. Federal Communications Commission (FCC) and the European Space Agency (ESA), ISTNs are recognized as enablers of direct-to-device (D2D) connectivity and TN support [23, 92]. Among the SatNet options, low Earth orbit (LEO) satellites (LSats) offer lower latency and higher channel gain, making them especially suited for tight TN-NTN integration and native support of NTNs in 6G [21].

To realize ubiquitous connectivity in 6G with ISTNs, efficient use of radio spectrum across TNs and SatNets becomes critical. Traditionally, these two networks have operated in separate radio frequency bands (RFBs), with TNs utilizing lower bands and SatNet relying on higher ones. However, due to limited spectrum availability, this static allocation strategy poses limitations, especially in densely deployed or heterogeneous traffic environments. To address these constraints, dynamic spectrum sharing (DSS) between TNs and SatNets appears as an important strategy within ISTNs to improve spectral efficiency. Recent proposals, such as MediaTek's spectrum reuse framework, highlight the potential of allowing SatNet to opportunistically access terrestrial spectrum [22]. This vision is further reinforced by regulatory bodies like the FCC and ESA, which support co-primary spectrum use between TN and NTN systems, particularly for D2D connectivity [23, 92]. Nevertheless, enabling spectrum sharing in ISTNs introduces critical challenges, especially inter-system interference (ISyI) between coexisting TN and SatNet transmissions.

Effectively managing ISyI is essential to realize the full potential of DSS in 6G-ISTN architectures. This, however, demands advanced technologies to address key challenges stemming from the wide coverage areas, high mobility of LSats, and the long propagation LSat-to-ground

(Sat2G) distance. Moreover, LSats often serve regions with diverse terrestrial conditions, especially in urban environments, where dense buildings introduce reflection, diffraction, and signal blockage, further complicating interference patterns. In addition, the dynamic movement of both LSats and users leads to rapidly varying link conditions, while the long Sat2G increases the overhead for signaling and real-time channel estimation. To address these challenges, digital-twin (DT) has emerged as a promising technology to offer virtual replicas of real-world systems, which enables capturing and emulating the evolution of the physical environments, network dynamics, and network entities in real time [108, 109] through continuous monitoring. In the context of ISTNs, utilizing DT to replicate traffic demand and environmental conditions [110] can significantly enhance resource management (RM), support proactive interference mitigation, and reduce Sat2G control signaling by using DT information. In this paper, we leverage DT to address the key challenges of ISyI, DSS coordination, and RM in practical ISTN deployments.

4.1.1 Related Works

The interference analysis in TN-NTN systems has been studied in the literature [9, 28, 93, 94]. Specifically, the TN-NTN coexistence scenarios were investigated in the S-band [9] and in the C-band [28]. The stochastic-geometry-based analyses were conducted in [93] to evaluate system coverage and rate probabilities, and in [94] to assess ergodic spectrum efficiency. These works, however, have not considered RM and relied on statistical channel models that struggle to reflect actual environments. Additionally, these studies evidenced that uncoordinated DSS in coexisting TN and NTN generates harmful ISyI, leading to throughput degradation.

Spectrum sharing (SS) designs in ISTNs have been considered in literature [16, 17, 45–48, 52]. In [16, 45–47], snapshot-based SS solutions for ISTNs were considered. In particular, [45] designed the second-price auction mechanism to achieve an equilibrium SS solution for TNs and SatNet-based traffic offloading. [46] considered a backhaul and access SS framework, where one satellite served both base stations (BSs) and users (UEs). Limited backhaul link capacity, the UE association, allocates bandwidth (BW) assignment, and power control were addressed through the corresponding sub-problems. [47] and [16] studied the RA problems for throughput maximization in reverse SS scenarios [47] and in underlay/overlay sharing cases [16]. However, these snapshot-based works have not considered ubiquitous and seamless connectivity and struggle to capture the dynamic behavior of ISTNs.

The time-window-based SS systems have been examined in [17, 48, 52]. Specifically, [48]

studied DSS for ISTNs, where a centralized coordinator BW to each network based on traffic load. In [17], the TN spectrum was temporarily shared with SatNet. The study focused on beam management, scheduling, and DSS for only SatNet, while TNs were considered under interference constraints. Both [17, 48] have not considered interference and service requirements at the user side. Additionally, their designs mainly rely on the statistical channel model. Filling these gaps, our previous work [52] has utilized 3D map and the ray-tracing (RayT) mechanism to examine the seamless connectivity and coverage enhancement in urban areas. However, [52] did not consider BW and RB allocation.

Recently, the DT technology has been utilized in RRM with limitations [111–114]. Particularly, these DT models are predominantly used to replicate the computation information [111, 114], traffic demand, [112, 113], and UE positions [114]. In addition, the channel state information (CSI), crucial for wireless RRM, should be replicated in DT models using the 3D map and RayT mechanism [110]. To the best of our knowledge, the joint design of BW allocation, traffic steering, AP/LSat-UE association, RB assignment, and power control in DSS ISTNs with DT-aided, incorporating a realistic channel model based on actual 3D maps, has not been thoroughly investigated. This motivates us to fill this research gap in this work.

4.1.2 Research Contributions

According to the discussed research gap as well as D2D scenarios driven by ESA and FCC, this work studies novel DT-aided DSS mechanisms for ISTNs supporting D2D connectivity, where TNs and SatNets share the same 5G-NR RFBs. Regarding uneven traffic demand among services and systems, the work aims to minimize system congestion by improving spectrum usage through DSS and RRM under service and system constraints. This work's main contributions can be summarized as follows.

- We investigate the 5G-NR transmission model for ISTNs under a practical system architecture and develop a novel DT model to replicate key components of the network. Beyond replicating traffic demand and terminal positions, as done in prior works, our model incorporates realistic terminal antenna patterns, an actual 3D map of the target area, and a ray-tracing mechanism to accurately emulate LSat/TN access point (TAP)-to-UE CSIs.

- We formulate two resource-allocation (RA) optimization problems, "*joint-RA*" and "*re-*

finement", which respectively rely on DT-predicted information and real-time system feedback, such as CSIs and traffic arrival. The joint-RA problem leverages DT-predicted information to optimize both long-term RA decisions (i.e., traffic steering and BW allocation across services and TN/SatNet systems) and preliminary estimated short-term RA decisions (i.e., TAP/LSat–UE association, RB assignment, and power control). The refinement problem adjusts the TN short-term decisions based on real-time feedback to align the DT-optimized strategy with current environmental conditions. Both problems are challenging due to non-convex constraints; notably, the joint-RA problem is formulated as a mixed-integer non-linear program (MINLP) and relies on predicted information.

- By leveraging the DT model for predictive information and applying compressed sensing and successive convex approximation (SCA) techniques, we propose two efficient solution methods, named DT-based joint RA and Real-time refinement algorithms (*DT-JointRA* and *RT-Refine*), that address the problems in alignment with the system architecture.
- For practical evaluation, we utilize a realistic 3D map of London, mobility of UEs such as vehicle UEs extracted from UE route in Google Navigator, and actual data traffic datasets. Numerical results demonstrate that our proposed algorithms significantly outperform greedy and heuristic benchmarks in minimizing queue lengths. Moreover, with DT-assisted prediction and pre-optimization in DT-JointRA, the RT-Refine scheme achieves fast convergence within a few iterations, highlighting the practical feasibility and efficiency of the proposed solutions.

The rest of this chapter is organized as follows. Section 4.2 presents the system model and problem formulation. Section 4.4 and 4.5 describe the proposed solutions, benchmarks, and complexity analysis, respectively. Section 4.6 and 4.7 provide numerical results and the conclusion.

4.2 5G-NR and Digital Twin System Model

This chapter considers a 5G-NR-based downlink transmission in an ISTN consisting of N TAPs and one LSat jointly serving K UEs, as depicted in Fig. 4.1. It is assumed that SatNet employs the FDD strategy to manage downlink (DL) and uplink (UL) transmissions in distinct RFBs, whereas TN operates under the TDD mechanism for allocating radio resources between

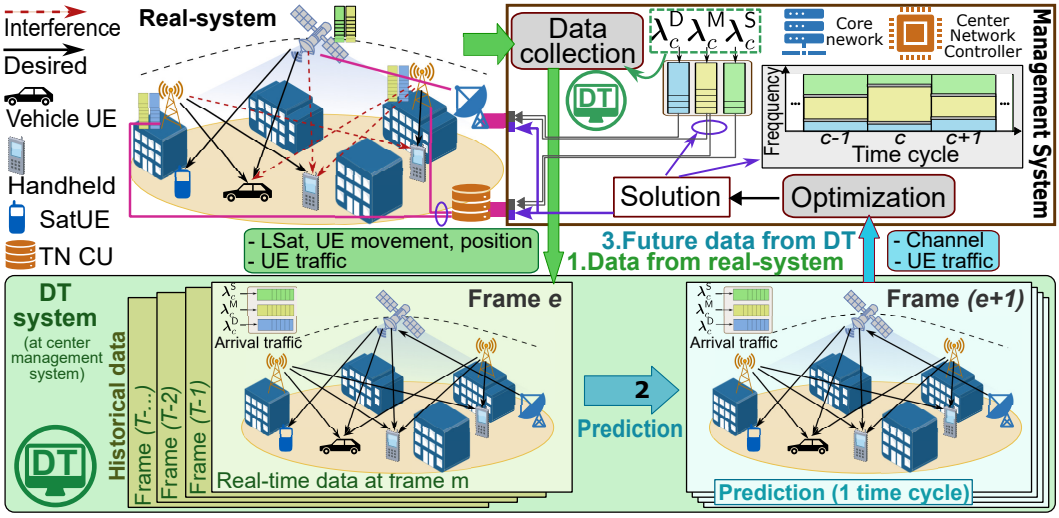


Figure 4.1: Digital-twin-aided ISTNs.

DL and UL in the same RFB [115]. In this context, we focus on utilizing the RFB, denoted $BW^{\text{total}} = [f^{\min}, f^{\max}]$ (Hz), which is assumedly accessible to TN for both DL and UL with TDD-mode management as well as the SatNet DL transmission.

4.2.1 Services and DT-based Network Management

The system supports three types of UEs, each associated with a specific communication service: (i) Delay-sensitive (D), such as uRLLC, exclusively served by TAPs; (ii) SatCom (S), provided solely by the LSat; and (iii) Multinet (M) supported by both TAPs and LSat. The sets of UEs associated with these three services are denoted as $\mathcal{K}_{\mathsf{D}}$, $\mathcal{K}_{\mathsf{S}}$, and $\mathcal{K}_{\mathsf{M}}$, respectively. The total BW is dynamically divided into three BW parts (BWPs) for these three services. Each BWP employs a 5G-NR numerology, i.e., μ_{D}, μ_{M}, and μ_{S}, tailored to its respective latency perspective. Let $\mathcal{K} = \mathcal{K}_{\mathsf{D}} \cup \mathcal{K}_{\mathsf{S}} \cup \mathcal{K}_{\mathsf{M}}$, $\mathcal{L}$, $\mathcal{S}$, $L = |\mathcal{L}|$, $K_{\mathsf{D}} = |\mathcal{K}_{\mathsf{D}}|$, $K_{\mathsf{S}} = |\mathcal{K}_{\mathsf{S}}|$, and $K_{\mathsf{S}} = |\mathcal{K}_{\mathsf{M}}|$ be the sets of all UEs, TAPs, services, and numbers of TAPs and UEs. Hereafter, we denote AP_ℓ and UE_k the ℓ-th TAP and k-th UE.

This chapter aims to jointly optimize *"long-term decisions"*, i.e., traffic steering from the core network (CN) to the LSat and TAPs, BW allocation (BWA) between TN and SatNet systems and across BWPs, as well as *"short-term decisions"*, i.e., UE association, RB assignment, and power control. To facilitate such optimization, we assume the availability of a DT model with the virtual representation of network components to enable the prediction of network information. *The DT model will be presented in Section 4.2.4.* The proposed optimization framework operates over multiple time scales, including cycle, frame, and sub-frame (SF), which will be detailed later.

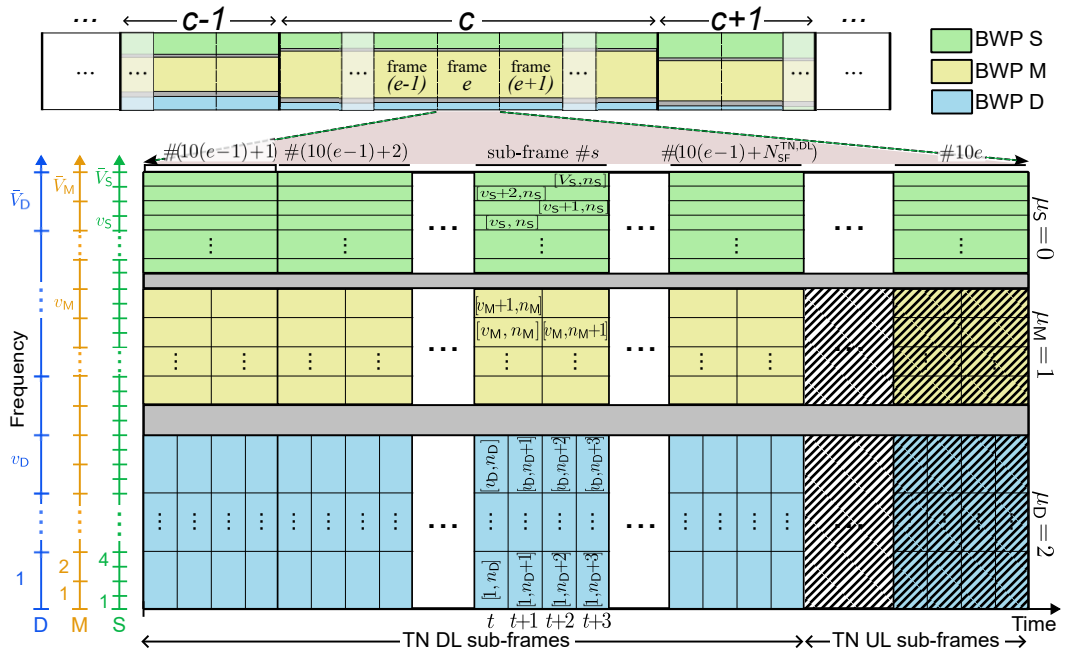


Figure 4.2: Resource block grid in cycle c.

This chapter considers a time window of N_{Cy} cycles. At the end of each cycle, the DT model updates its state and provides predicted system-level information of the subsequent cycle to the central network management system (NMS). Leveraging this foresight, the NMS performs joint resource management optimization, generating long-term decisions along with preliminary estimates of short-term ones. This stage is referred to as the **"DT-based Joint RA Stage"**. Each cycle is composed of multiple frames, with each frame consisting of 10 sub-frames (SFs). We assume that traffic rate information for D UEs can be updated at the SF level [116]. Therefore, *"short-term decisions"* for the TN can be further refined at the SF granularity by the TN center unit (TCU). This SF-level re-optimization process constitutes the **"RT-Refine Stage"**.

4.2.2 5G-NR Standard-based Setting

Frame and Sub-frame Setting

According to 5G-NR standard, each frame duration is 10 ms ($T_{\mathsf{TF}} = 10$ ms); hence, duration of one SF is 1 ms ($T_{\mathsf{SF}} = 1$ ms). Assuming that each cycle c contains N_{TF} frames, the number of SFs in one cycle is $N_{\mathsf{SF}} = 10N_{\mathsf{TF}}$. Regarding the TDD transmission mode for TN, we assume that TN uses $N_{\mathsf{SF}}^{\mathsf{TN,DL}}$ beginning SFs for DL and $(10 - N_{\mathsf{SF}}^{\mathsf{TN,DL}})$ ending SFs for UL.

Numerology Setting

In this chapter, numerologies used in BWPs are set as $\mu_{\mathsf{S}} = 0, \mu_{\mathsf{M}} = 1$ and $\mu_{\mathsf{D}} = 2$, reflecting the increasing tolerance to delay across the three service types [117, 118]. Based on this, the RB grid is illustrated in Fig. 4.2.

Time Domain Hereafter, we define the RB duration for service D as a *time slot* (TS) - the time unit. Based on that, the time indices RBs of service x n_{x} can be defined as

$$n_{\mathsf{D}} = t, \quad n_{\mathsf{M}} = \lceil t/2 \rceil, \quad n_{\mathsf{S}} = \lceil t/4 \rceil, \tag{4.1}$$

respectively, where $\lceil \cdot \rceil$ is the ceil function. Furthermore, the RB duration of service x can be given as $T_{\mathsf{x}} = 2^{-\mu_{\mathsf{x}}}$ ms; hence, there are $N_{\mathsf{x}} = 2^{\mu_{\mathsf{x}}}$ RBs over one sub-channel due to service x in within a SF. Then, the time indices of cycles (c), frames (m), sub-frames (s) can be counted based on N_{D} as $c = \lceil t/(N_{\mathsf{D}} N_{\mathsf{SF}}) \rceil, m = \lceil t/(10 N_{\mathsf{D}}) \rceil, s = \lceil t/N_{\mathsf{D}} \rceil$. In addition, the sets of RB indices in the time domain corresponding to service x within cycle c, frame e, and SF s can be defined as $\mathcal{T}_{\mathsf{x}}^{\mathsf{Cy},c} \triangleq \{(c-1)N_{\mathsf{x}} N_{\mathsf{SF}} + 1, \ldots, cN_{\mathsf{x}} N_{\mathsf{SF}}\}$, $\mathcal{T}_{\mathsf{x}}^{\mathsf{TF},e} \triangleq \{(e-1)10N_{\mathsf{x}} + 1, \ldots, e10N_{\mathsf{x}}\}$, and $\mathcal{T}_{\mathsf{x}}^{\mathsf{SF},s} \triangleq \{(s-1)N_{\mathsf{x}} + 1, \ldots, sN_{\mathsf{x}}\}$.

Frequency Domain Let $BW_{\mathsf{x},c} = [f_{\mathsf{x},c}^{\mathsf{min}}, f_{\mathsf{x},c}^{\mathsf{max}}]$ be the frequency ranges of BWPs sliced for service x in cycle c. For simplicity, we set $f_{\mathsf{D},c}^{\mathsf{max}} \leq f_{\mathsf{M},c}^{\mathsf{min}}$ and $f_{\mathsf{M},c}^{\mathsf{max}} \leq f_{\mathsf{S},c}^{\mathsf{min}}$ for all c. Then, the BW of service x in cycle c is given as $W_{\mathsf{x},c} = f_{\mathsf{x},c}^{\mathsf{max}} - f_{\mathsf{x},c}^{\mathsf{min}}$. The BWPs are divided into sub-channels (SCs) due to various numerologies. For service x, the RB channel spacing is given as $w_{\mathsf{x}} = 2^{\mu_{\mathsf{x}}} \times 180$ kHz. Let $W^{\mathsf{tot}} = f^{\mathsf{max}} - f^{\mathsf{min}}$ (Hz) denote the total BW. Then, the maximum SC number that can be assigned to the service x is $\bar{V}_{\mathsf{x}} = \lfloor W^{\mathsf{tot}}/w_{\mathsf{x}} \rfloor$. We further denote v_{x} indices of the SCs in BWP x, and $\mathcal{V}_{\mathsf{x}} \triangleq \{1, \ldots, \bar{V}_{\mathsf{x}}\}$, i.e., $v_{\mathsf{x}} \in \mathcal{V}_{\mathsf{x}}$. Once, the BWP of service x is defined, the SCs of $\mathcal{V}_{\mathsf{x}}$ located in the corresponding frequency range will be activated for service-x transmission. Hereafter, the RBs in BWP x are identified based on both time and frequency indices, i.e., $[v_{\mathsf{x}}, n_{\mathsf{x}}]$.

Following 5G standard, guard bands (GBs) between BWPs are considered. Here, the $\mathsf{D}-\mathsf{M}$ separating GB BW should be half of w_{D}, i.e., $W_{\mathsf{G1}} = f_{\mathsf{M},c}^{\mathsf{min}} - f_{\mathsf{D},c}^{\mathsf{max}} = w_{\mathsf{D}}/2$, and the $\mathsf{M}-\mathsf{S}$ separating GB BW is half of w_{M}, $W_{\mathsf{G2}} = f_{\mathsf{S},c}^{\mathsf{min}} - f_{\mathsf{M},c}^{\mathsf{max}} = w_{\mathsf{M}}/2$.

Remark 9. *Denote* $v_{\mathsf{D},c}^{\mathsf{max}}$ *and* $v_{\mathsf{M},c}^{\mathsf{max}}$ *the maximum indices of activated SCs of service* D *and* M *in cycle* c. *With respect to GB, we can obtain* $f_{\mathsf{M},c}^{\mathsf{min}} = f^{\mathsf{min}} + (2v_{\mathsf{D},c}^{\mathsf{max}} + 1)w_{\mathsf{M}} = f^{\mathsf{min}} + (4v_{\mathsf{D},c}^{\mathsf{max}} + 2)w_{\mathsf{S}}$

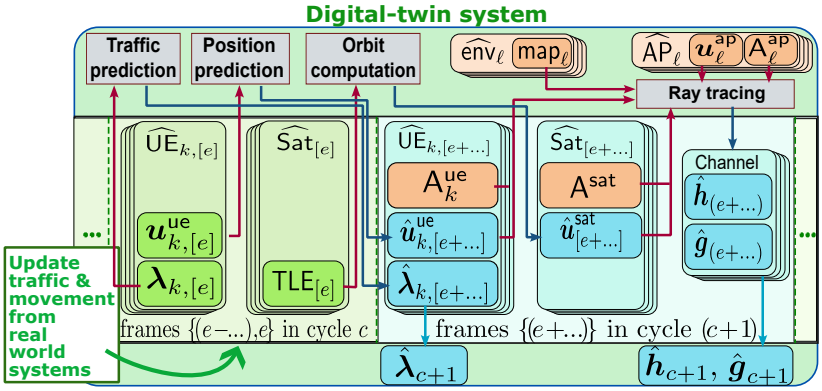


Figure 4.3: Summary of the DT-system.

and $f_{\mathsf{S},c}^{\min} = f^{\min} + (2v_{\mathsf{M},c}^{\max} + 1)w_{\mathsf{S}}$.

Let $\boldsymbol{b}_c \triangleq \{b_{v_{\mathsf{x}},c}^{\mathsf{x}} | \forall v_{\mathsf{x}} \in \mathcal{V}_{\mathsf{x}}, \mathsf{x} \in \mathcal{S}\}$ where $b_{v_{\mathsf{x}},c}^{\mathsf{x}} = 1$ if SC v_{x} is activated in cycle c and $b_{v_{\mathsf{x}},c}^{\mathsf{x}} = 0$ otherwise. Remark 9 yields,

$$(C1): \quad b_{v_{\mathsf{D}},c}^{\mathsf{D}} + \sum_{i=1}^{2v_{\mathsf{D}}+1} b_{i,c}^{\mathsf{M}} + \sum_{j=1}^{4v_{\mathsf{D}}+2} b_{j,c}^{\mathsf{s}} \leq 1, \ \forall (v_{\mathsf{D}}, c),$$
$$(C2): \quad b_{v_{\mathsf{M}},c}^{\mathsf{M}} + \sum_{j=1}^{2v_{\mathsf{M}}+1} b_{j,c}^{\mathsf{S}} \leq 1, \ \forall (v_{\mathsf{M}}, c).$$

Additionally, the system BW constraint yields

$$(C3): \quad \sum_{\mathsf{x} \in \mathcal{S}} \sum_{\forall v_{\mathsf{x}} \in \mathcal{V}_{\mathsf{x}}} b_{v_{\mathsf{x}},c}^{\mathsf{x}} w_{\mathsf{x}} + W_{\mathsf{G1}} + W_{\mathsf{G2}} \leq W^{\mathsf{tot}}, \ \forall c.$$

4.2.3 Channel Model

Let $\tilde{h}_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}$ and $\tilde{g}_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]}$ be the channel coefficients of $\mathsf{AP}_\ell - \mathsf{UE}_k$ and $\mathsf{LSat} - \mathsf{UE}_k$ links over $\mathsf{RB}_{[v_{\mathsf{x}},n_{\mathsf{x}}]}$, respectively. Omitting $[v_{\mathsf{x}}, n_{\mathsf{x}}]$, the channel coefficients are modeled as

$$\tilde{h}_{\ell,k} = \sqrt{\mathsf{PL}_{\ell,k}} \left(\sqrt{\rho_{\ell,k}} \tilde{h}_{\ell,k}^{\mathsf{ls}} + \sqrt{1 - \rho_{\ell,k}} \tilde{h}_{\ell,k}^{\mathsf{nl}} \right), \tag{4.2}$$
$$\tilde{g}_k = \sqrt{\mathsf{PL}_{0,k}} \left(\sqrt{\rho_{0,k}} \tilde{g}_k^{\mathsf{ls}} + \sqrt{1 - \rho_{0,k}} \tilde{g}_k^{\mathsf{nl}} \right), \tag{4.3}$$

based on the Rician model where $\rho_{\ell,k} = \tilde{K}_{\ell,k}/(\tilde{K}_{\ell,k} + 1)$; $\tilde{K}_{\ell,k}$, $\mathsf{PL}_{\ell,k}$, $\tilde{h}_{\ell,k}$, $\tilde{h}_{\ell,k}$, $\tilde{K}_{0,k}$, $\mathsf{PL}_{0,k}$, $\tilde{g}_k^{\mathsf{ls}}$, and $\tilde{g}_k^{\mathsf{nl}}$ are the K-factor, path-loss, the corresponding line-of-sight (LoS) and non-line-of-sight (NLoS) components, respectively. In this work, we assume that TAPs can estimate perfectly CSIs for their served UEs in each frame thanks to the pilot sent in TN uplink SFs [119]. The channel gain in cycle c and in frame e are denoted by $\{\boldsymbol{h}_c, \boldsymbol{g}_c\}$ and $\{\boldsymbol{h}_e, \boldsymbol{g}_e\}$.

4.2.4 Digital-Twin Model

The DT system is deployed in NMS to replicate the environment and network components. As illustrated in Fig. 4.3, the DT model comprises a virtual system and computational functionalities. The virtual system consists of *static components*, including environmental features, TAP positions, and terminal antenna characteristics, as well as *dynamic components*, such as UE mobility, traffic demand, LSat movement, and CSIs. The DT model operation over each cycle is depicted in Fig. 4.1. Specifically, at the end of each cycle, the real mobility and traffic information are updated to the DT model. Using this information, computational functionalities are leveraged to emulate the real system's evolution in the next cycle and reflect it in the virtual one. The resulting predictions are then used in resource management for the subsequent cycle.

Geographical Environment

The DT system integrates 3D maps of the targeted areas. In cell ℓ, the corresponding environment in DT system is represented by

$$\widehat{\mathsf{env}}_\ell = \{\mathsf{map}_\ell\}, \quad \widehat{\mathsf{AP}}_\ell = \{\mathsf{A}_\ell^{\mathsf{ap}}, \boldsymbol{u}_\ell^{\mathsf{ap}}\}, \quad \forall \ell \in \mathcal{L}, \tag{4.4}$$

wherein map_ℓ indicates the 3D map of the area in cell n, $\mathsf{A}_\ell^{\mathsf{ap}}$ and $\boldsymbol{u}_\ell^{\mathsf{ap}}$ are the antenna properties and position of AP_ℓ.

UE Information

Virtual $\mathsf{UE_k}$ in frame e is modeled as

$$\widehat{\mathsf{UE}}_{k,[e]}^{\mathsf{D}} = \{\mathsf{A}_k^{\mathsf{ue}}, \hat{\boldsymbol{u}}_{k,[e]}^{\mathsf{ue}}, \hat{\lambda}_{k,[s]}^{\mathsf{SF,D}}|_{\forall s \in \mathcal{T}_{\mathsf{SF}}^{\mathsf{TF},e}}\}, \ \forall k \in \mathcal{K}_{\mathsf{D}}, \tag{4.4a}$$

$$\widehat{\mathsf{UE}}_{k,[e]}^{\mathsf{x}} = \{\mathsf{A}_k^{\mathsf{ue}}, \hat{\boldsymbol{u}}_{k,[e]}^{\mathsf{ue}}, \hat{\lambda}_{k,[e]}^{\mathsf{TF,x}}\}, \ \forall k \in \mathcal{K}_{\mathsf{x}}, \mathsf{x} \in \{\mathsf{M}, \mathsf{S}\}, \tag{4.4b}$$

wherein $\mathcal{T}_{\mathsf{SF}}^{\mathsf{TF},e}$ is SF set in frame e, $\mathsf{A}_k^{\mathsf{ue}}$, $\hat{\boldsymbol{u}}_{k,[e]}^{\mathsf{ue}}$, $\hat{\lambda}_{k,[e]}^{\mathsf{TF,x}}$, and $\hat{\lambda}_{k,[s]}^{\mathsf{SF,D}}$ are the virtual antenna properties, position and traffic rate of UE_k in frame e and SF s, respectively. *The UE traffic rates will be described in Section 4.3.5.* This UE information in the upcoming cycle is obtained based on prediction techniques and updated information from the real environment.

LSat Model

The LSat DT at frame e is modeled as

$$\widehat{\mathsf{Sat}}_{[e]} = \{\mathsf{A}^{\mathsf{sat}}, \hat{\boldsymbol{u}}^{\mathsf{sat}}_{[e]}, \mathsf{TLE}_{[e]}\}, \tag{4.5}$$

where $\mathsf{A}^{\mathsf{sat}}$, $\hat{\boldsymbol{u}}^{\mathsf{sat}}_{[e]}$, and $\mathsf{TLE}_{[e]}$ are the antenna properties, predicted position, and two-line-element (TLE) data of LSat in frame e, respectively. Due to the orbit stability and the frequent update framework, we assume that the LSat position is predicted with a negligible error, that is, $\hat{\boldsymbol{u}}^{\mathsf{sat}}_{[e]} \approx \boldsymbol{u}^{\mathsf{sat}}_{[e]}$.

Channel Prediction

Employing the RayT tool as in Fig. 4.3, the DT channel model is identified as [52]

$$\mathsf{chan}_{[e]} = \mathsf{RT}(\widehat{\mathsf{env}}_\ell, \widehat{\mathsf{AP}}_\ell, \widehat{\mathsf{Sat}}_{[e]}, \widehat{\mathsf{UE}}_{k,[e]}), \tag{4.6}$$

where $\widehat{\mathsf{UE}}_{k,[e]} = \{\widehat{\mathsf{UE}}^{\mathsf{x}}_{k,[e]}\}_{\forall \mathsf{x}}$, $\mathsf{chan}_{[e]} = \{\tilde{h}^{\mathsf{ls},[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}, \bar{h}^{\mathsf{nl},[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}, \mathsf{PL}^{[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}, \tilde{K}^{[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}, \tilde{g}^{\mathsf{ls},[\nu_\mathsf{x},n_\mathsf{x}]}_{k}, \bar{g}^{\mathsf{nl},[\nu_\mathsf{x},n_\mathsf{x}]}_{k}, \mathsf{PL}^{[\nu_\mathsf{x},n_\mathsf{x}]}_{0,k}, \tilde{K}^{[\nu_\mathsf{x},n}_{0,k}$ with $\forall(\ell,k), \forall n_\mathsf{x} \in \mathcal{T}^{\mathsf{TF},e}_{\mathsf{x}}$. In virtual environment, $\bar{h}^{\mathsf{nl},[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}$ and $\bar{g}^{\mathsf{nl},[\nu_\mathsf{x},n_\mathsf{x}]}_{k}$ are introduced to represent the NLoS components of DT channels. Those are different with $\tilde{h}^{\mathsf{nl},[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}$ and $\tilde{g}^{\mathsf{nl},[\nu_\mathsf{x},n_\mathsf{x}]}_{k}$ of real system channels, primarily due to the absence of information in the 3D map. The relationship between real and virtual NLoS components is modeled as

$$\tilde{h}^{\mathsf{nl}}_{\ell,k} = \sqrt{\xi}\bar{h}^{\mathsf{nl}}_{\ell,k} + \sqrt{(1-\xi)}\delta_{\ell,k}, \tag{4.6a}$$

$$\tilde{g}^{\mathsf{nl}}_{k} = \sqrt{\xi}\bar{g}^{\mathsf{nl}}_{k} + \sqrt{(1-\xi)}\delta_{0,k}, \tag{4.6b}$$

where $\delta_{\ell,k}$ and $\delta_{0,k}$ indicate errors caused by the absence of map information, which are assumed as complex normal random variables. $\xi \in (0,1)$ is the correlation factor.

4.3 DT-based Optimization Problem Formulation

4.3.1 User Association

Let $\boldsymbol{\alpha} = [\alpha^{[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k}]$ be the binary association variable for $\mathsf{x} \in \{\mathsf{D},\mathsf{M}\}$ where $\alpha^{[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k} = 1$ if AP_ℓ served $\mathsf{UE_k}$ over $\mathsf{RB}_{[\nu_\mathsf{x},\mathsf{n_x}]}$ for DL and $\alpha^{[\nu_\mathsf{x},n_\mathsf{x}]}_{\ell,k} = 0$ otherwise. Note that the association with UE^{x} over SC ν_x in BWP x is allowed only if this SC is activated, which is ensured as

$$(C4): \quad \alpha_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]} \leq b_{v_{\mathsf{x}},c}^{\mathsf{x}}, \forall (c,\ell,k,[v_{\mathsf{x}},n_{\mathsf{x}}]), \mathsf{x} \in \{\mathsf{D},\mathsf{M}\}.$$

In addition, one assumes that each RB of each AP can be assigned to at most one UE, which yields

$$(C5): \quad \sum\nolimits_{\forall k} \alpha_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]} \leq 1, \ \forall (c,\ell,[v_{\mathsf{x}},n_{\mathsf{x}}]), \mathsf{x} \in \{\mathsf{D},\mathsf{M}\}.$$

While each UE^{D} can be served by multiple APs at each TS via different RBs, which is ensured as

$$(C6): \ \sum\nolimits_{\forall \ell} \alpha_{\ell,k}^{[v_{\mathsf{D}},n_{\mathsf{D}}]} \leq 1, \ \forall (c,k,[v_{\mathsf{D}},n_{\mathsf{D}}]).$$

For LSat-UE association, we introduce a variable $\boldsymbol{\beta} = [\beta_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]}]$ for $\mathsf{x} \in \{\mathsf{M},\mathsf{S}\}$ where $\beta_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]} = 1$ if $\mathsf{UE_k}$ is served by LSat over $\mathsf{RB}[v_{\mathsf{x}},n_{\mathsf{x}}]$, $\beta_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]} = 0$ otherwise. Similar to the AP-UE association, we have

$$(C7): \ \beta_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]} \leq b_{v_{\mathsf{x}},c}^{\mathsf{x}}, \forall (c,k,[v_{\mathsf{x}},n_{\mathsf{x}}],$$
$$(C8): \ \sum\nolimits_{\forall k} \beta_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]} \leq 1, \ \forall (c,[v_{\mathsf{x}},n_{\mathsf{x}}]), \mathsf{x} \in \{\mathsf{M},\mathsf{S}\},$$

Additionally, we assume that UEs using MS services can be served by both APs and LSat at the same time via different RBs in BWP M, which yields the constraint

$$(C9): \ \sum\nolimits_{\forall \ell} \alpha_{\ell,k}^{[v_{\mathsf{M}},n_{\mathsf{M}}]} + \beta_k^{[v_{\mathsf{M}},n_{\mathsf{M}}]} \leq 1, \ \forall (c,k,[v_{\mathsf{M}},n_{\mathsf{M}}]).$$

4.3.2 Service-D Transmission

Assuming that $\mathsf{UE}_k^{\mathsf{D}}$ is served by AP_ℓ over $\mathsf{RB}_{[v_{\mathsf{D}},n_{\mathsf{D}}]}$, the received signal $y_k^{[v_{\mathsf{D}},n_{\mathsf{D}}]}$ and its components are expressed as

$$y_k^{\mathsf{D},[v_{\mathsf{D}},n_{\mathsf{D}}]} = \tilde{y}_k^{\mathsf{TN},[v_{\mathsf{D}},n_{\mathsf{D}}]} + \varsigma_k^{[v_{\mathsf{D}},n_{\mathsf{D}}]}, \tag{4.7}$$

where $\tilde{y}_k^{\mathsf{TN},[v_{\mathsf{x}},n_{\mathsf{x}}]} = \Sigma_{\forall (i,j)} \sqrt{\alpha_{i,j}^{[v_{\mathsf{x}},n_{\mathsf{x}}]} p_{i,j}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}} \tilde{h}_{i,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]} x_{i,j}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}$; $p_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}$ and $x_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}$ are the transmit power and transmission symbol from AP_ℓ to UE_k over $\mathsf{RB}_{[v_{\mathsf{x}},n_{\mathsf{x}}]}$. $\varsigma_k^{[v_{\mathsf{x}},n_{\mathsf{x}}]} \sim \mathcal{CN}(0,\sigma_{\mathsf{x},\mathsf{k}}^2)$ is the AGWN at UE_k. Hence, the SINR of UE_k is expressed as

$$\gamma_{\ell,k}^{\mathsf{D},[v_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c) = \frac{\alpha_{\ell,k}^{[v_{\mathsf{D}},n_{\mathsf{D}}]} p_{\ell,k}^{[v_{\mathsf{D}},n_{\mathsf{D}}]} h_{\ell,k}^{[v_{\mathsf{D}},n_{\mathsf{D}}]}}{\Psi_{\ell,k}^{[v_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c) + \sigma_{\mathsf{D},k}^2}, \tag{4.8}$$

where $\Psi_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)$ is the ICI power defined as

$$\Psi_{\ell,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)=\sum\nolimits_{\forall i\neq\ell,\forall j}\alpha_{i,j}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}p_{i,j}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}h_{i,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}, \tag{4.9}$$

with $(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)\triangleq\{(p_{\ell,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]},\alpha_{\ell,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]})|_{\forall(\ell,k,[\nu_{\mathsf{x}},n_{\mathsf{x}}]),\mathsf{x}\in\{\mathsf{D},\mathsf{M}\}}\}$ and $h_{\ell,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}=|\tilde{h}_{\ell,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}|^2$. Due to the delay-sensitive requirement, the short packet framework is used to model the transmission of D service. Hence, the aggregated achievable rate of $\mathsf{UE}_k^{\mathsf{D}}$ served by AP_ℓ at SF s can be expressed as [120]

$$R_{\ell,k,[s]}^{\mathsf{SF},\mathsf{D}}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)=w_{\mathsf{D}}\left[\sum_{\forall(\nu_{\mathsf{D}},n_{\mathsf{D}})}\log_2\left(1+\gamma_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)\right)-\frac{1}{\ln(2)}\frac{\alpha_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}\sqrt{V_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}}Q^{-1}(P_\epsilon)}{\sqrt{\sum_{\forall(\nu_{\mathsf{D}},n_{\mathsf{D}})}\alpha_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}T_{\mathsf{d}}w_{\mathsf{d}}}}\right], \tag{4.10}$$

where $V_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}=1-(1+\gamma_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c))^{-2}$, $Q^{-1}(\cdot)$ and P_ϵ are the channel dispersion, the inverse of the Q-function, and the error probability. One can see that the channel dispersion can be approximated as $V_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}\approx 1$ for a sufficiently high $\gamma_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)\geq\gamma_0^{\mathsf{D}}$ with $\gamma_0^{\mathsf{D}}\geq 5$ dB [121]. Regarding this approximation, we consider the following constraint

$$(C10):\quad \gamma_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)\geq\alpha_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}\gamma_0^{\mathsf{D}},\ \forall(\ell,k,\nu_{\mathsf{D}},n_{\mathsf{D}}).$$

Subsequently, $R_{\ell,k,[s]}^{\mathsf{SF},\mathsf{D}}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)$ can be rewritten as

$$R_{\ell,k,[s]}^{\mathsf{SF},\mathsf{D}}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)=w_{\mathsf{D}}\sum_{\forall(\nu_{\mathsf{D}},n_{\mathsf{D}})}\log_2\left(1+\gamma_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{p}_c,\boldsymbol{\alpha}_c)\right)-\chi_{\mathsf{D}}\sqrt{\sum_{\forall(\nu_{\mathsf{D}},n_{\mathsf{D}})}\alpha_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}}, \tag{4.11}$$

where $\chi_{\mathsf{D}}=\sqrt{Vw_{\mathsf{D}}/T_{\mathsf{D}}}Q^{-1}(P_\epsilon)/\ln(2)$ and $V\approx V_{\ell,k}^{\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}\approx 1$.

4.3.3 Service-M Transmission

In this BWP, UEs can be served by APs and the LSat. The received signal $y_k^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ at $\mathsf{UE}_k^{\mathsf{M}}$ over $\mathsf{RB}_{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ and its components is expressed as

$$y_k^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}=\tilde{y}_k^{\mathsf{TN},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}+\tilde{y}_k^{\mathsf{Sat},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}+\varsigma_k^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}, \tag{4.12}$$

$$\tilde{y}_k^{\mathsf{Sat},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}\triangleq\sum\nolimits_{\forall j}\sqrt{\beta_j^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}p_{0,j}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}}\tilde{g}_k^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}x_j^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}, \tag{4.13}$$

where $p_{0,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}$ is the transmit power from the LSat to UE_k over $\mathsf{RB}_{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}$, and $g_k^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}=|\tilde{g}_k^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]}|^2$. For brevity, let's define $(\boldsymbol{p}_{0,c},\boldsymbol{\beta}_c)\triangleq\{(p_{0,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]},\beta_k^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]})|_{\forall(k,\nu_{\mathsf{x}},n_{\mathsf{x}}),\mathsf{x}\in\{\mathsf{M},\mathsf{S}\}}\}$ and $\boldsymbol{P}_c\triangleq\{\boldsymbol{p}_c,\boldsymbol{p}_{0,c}\}$. Hereafter, argument $\boldsymbol{p}_c$ in defined functions is replaced appropriately by $\boldsymbol{P}_c$.

TN-served UEs

Assuming that $\mathsf{UE}_k^{\mathsf{M}}$ is served by AP_ℓ, the corresponding SINR is expressed as

$$\gamma_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)=\frac{\alpha_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}p_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}h_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}}{\Psi_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c)+\Theta_k^{\mathsf{TN},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c)+\sigma_{\mathsf{M},k}^2},\tag{4.14}$$

where $\Psi_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ is defined in (4.9) and $\Theta_k^{\mathsf{TN},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c)$ is the inter-system-interference (ISyI) caused by the LSat to $\mathsf{UE}_k^{\mathsf{M}}$ served by TN over $\mathsf{RB}_{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ which is given as

$$\Theta_k^{\mathsf{TN},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c)=\textstyle\sum_{\forall j}\beta_j^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}p_{0,j}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}g_k^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]},\tag{4.15}$$

The corresponding achievable rate over $\mathsf{RB}_{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ and the aggregated rate at RB time n_{M} are expressed as

$$R_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)=w_{\mathsf{M}}\log_2\left(1+\gamma_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)\right),\tag{4.15a}$$

$$R_{\ell,k}^{\mathsf{M},[n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)=\textstyle\sum_{\forall\nu_{\mathsf{M}}\in\mathcal{V}_{\mathsf{M}}}R_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c).\tag{4.15b}$$

UE served by SatNet

Assuming that $\mathsf{UE}_k^{\mathsf{M}}$ is served by the LSat over $\mathsf{RB}_{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$, the corresponding SINR is given as

$$\gamma_{0,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)=\frac{\beta_k^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}p_{0,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}g_k^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}}{\Theta_k^{\mathsf{Sat},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c)+\sigma_{\mathsf{M},k}^2},\tag{4.16}$$

where $\Theta_k^{\mathsf{Sat},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c)$ is the ISyI caused by TAPs to $\mathsf{UE}_k^{\mathsf{M}}$ served by the LSat which is defined as

$$\Theta_k^{\mathsf{Sat},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c)=\textstyle\sum_{\forall(i,j)}\alpha_{i,j}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}p_{i,j}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}h_{i,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}.\tag{4.17}$$

Therefore, the achievable rate over $\mathsf{RB}_{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ and the aggregated rate at RB time n_{M} of $\mathsf{UE}_k^{\mathsf{M}}$ served by the LSat can be respectively expressed as

$$R_{0,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)=w_{\mathsf{M}}\log_2\left(1+\gamma_{0,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)\right),\tag{4.17a}$$

$$R_{0,k}^{\mathsf{M},[n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c)=\textstyle\sum_{\forall\nu_{\mathsf{M}}\in\mathcal{V}_{\mathsf{M}}}R_{0,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c).\tag{4.17b}$$

4.3.4 Service-S Transmission

In this BWP, if $\mathsf{UE}_k^{\mathsf{S}}$ is served by LSat over $\mathsf{RB}_{[\nu_{\mathsf{S}},n_{\mathsf{S}}]}$, the corresponding SNR is given by

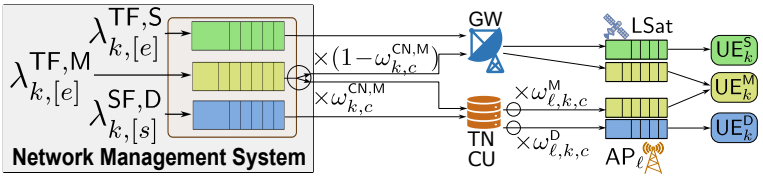


Figure 4.4: Traffic steering model.

$$\gamma_{0,k}^{\mathsf{S},[v_\mathsf{S},n_\mathsf{S}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c) = (\beta_k^{[v_\mathsf{S},n_\mathsf{S}]} p_{0,k}^{[v_\mathsf{S},n_\mathsf{S}]} g_k^{[v_\mathsf{S},n_\mathsf{S}]})/\sigma_{\mathsf{S},k}^2, \tag{4.18}$$

Therefore, the achievable rate over $\mathsf{RB}_{[v_\mathsf{S},n_\mathsf{S}]}$ and the aggregated rate at RB time n_S of UE_k^S can be expressed as

$$R_{0,k}^{\mathsf{S},[v_\mathsf{S},n_\mathsf{S}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c) = w_\mathsf{S} \log_2\left(1+\gamma_{0,k}^{\mathsf{S},[v_\mathsf{S},n_\mathsf{S}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c)\right), \tag{4.18a}$$

$$R_{0,k}^{\mathsf{S},[n_\mathsf{S}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c) = \sum\nolimits_{\forall v_\mathsf{S}\in\mathcal{V}_\mathsf{S}} R_{0,k}^{\mathsf{S},[v_\mathsf{S},n_\mathsf{S}]}(\boldsymbol{P}_c,\boldsymbol{\beta}_c). \tag{4.18b}$$

Regarding the TAP/LSat power budget, one must satisfy

$$(C11): \quad \sum_{\mathsf{x}\in\{\mathsf{D},\mathsf{M}\}} \sum_{\forall k} \sum_{\forall v_\mathsf{x}\in\mathcal{V}_\mathsf{x}} p_{\ell,k}^{[v_\mathsf{x},n_\mathsf{x}]} \le p_{\mathsf{AP},\ell}^{\max}, \quad \forall(\ell,t),$$

$$(C12): \quad \sum_{\mathsf{x}\in\{\mathsf{M},\mathsf{S}\}} \sum_{\forall k} \sum_{\forall v_\mathsf{x}\in\mathcal{V}_\mathsf{x}} p_{0,k}^{[v_\mathsf{x},n_\mathsf{x}]} \le p_{\mathsf{LSat}}^{\max}, \quad \forall t.$$

4.3.5 Traffic and Queuing Model

Regarding latency tolerance, different time scales are applied to serve UEs associated with different services. Specifically, the arrival data of M and S services are buffered and scheduled for transmission in the following frame, whereas the data stream of D services is processed at the SF time scale to meet stringent latency requirements. Assuming K traffic flows arriving the CN due to K UEs. Denote $\boldsymbol{\lambda}_{[e]}^{\mathsf{TF,M}} = \{\lambda_{k,[e]}^{\mathsf{TF,M}}|_{\forall k\in\mathcal{K}_\mathsf{M}}\}$, $\boldsymbol{\lambda}_{[e]}^{\mathsf{TF,S}} = \{\lambda_{k,[e]}^{\mathsf{TF,S}}|_{\forall k\in\mathcal{K}_\mathsf{S}}\}$, and $\boldsymbol{\lambda}_{[s]}^{\mathsf{SF,D}} = \{\lambda_{k,[s]}^{\mathsf{SF,D}}|_{\forall k\in\mathcal{K}_\mathsf{D}}\}$, where $\lambda_{k,[e]}^{\mathsf{TF,M}}$, $\lambda_{i,[e]}^{\mathsf{TF,S}}$, and $\lambda_{j,[s]}^{\mathsf{SF,D}}$ are the data arrival rates of UE_k^M, UE_i^S at frame e, and of UE_j^D at SF s, respectively. Additionally, the traffic rates in cycle c and frame e are denoted by $\boldsymbol{\lambda}_c$ and $\boldsymbol{\lambda}_e$, respectively.

Traffic Steering

As depicted in Fig. 4.4, at the CN, the D and S traffic flows are routed to the TNs and SatNet, respectively, while M traffic is split and dynamically steered to both domains. Subsequently, at the TCU, the D and M flows are further directed to TAPs due to a TN steering scheduler.

For M-traffic steering, we introduce variables $\boldsymbol{\omega}_c^{\mathsf{CN,M}} = [\omega_{k,c}^{\mathsf{CN,M}}], \forall k \in \mathcal{K}_\mathsf{M}$, where $\omega_{k,c}^{\mathsf{CN,M}} \in$

$[0,1]$ and $(1-\omega_{k,c}^{\mathsf{CN,M}})$ are the portions of $\mathsf{UE}_k^{\mathsf{M}}$'s traffic which are steered to TN and SatNet in cycle c, respectively. In the TN, let $\boldsymbol{\omega}_c^{\mathsf{x}} = [\omega_{\ell,k,c}^{\mathsf{x}}], \forall(\ell,k) \in (\mathcal{L}\times\mathcal{K}_{\mathsf{x}}), \mathsf{x}\in\{\mathsf{D},\mathsf{M}\}$ be the TN steering variables, wherein $\omega_{\ell,k,c}^{\mathsf{x}} \in [0,1]$ indicates the flow-split portion of service $\mathsf{x}\in\{\mathsf{D},\mathsf{M}\}$ from TCU to AP_ℓ in cycle c. Besides, the integrity of traffic flows is ensured by

$$(C13): \quad \sum_{\forall \ell\in\mathcal{L}} \omega_{\ell,k,c}^{\mathsf{x}} = 1, \quad \forall(c,k), \mathsf{x}\in\{\mathsf{D},\mathsf{M}\}.$$

Subsequently, the data arrival of D service in SF s, and that of M and S services in frame e at AP_ℓ/LSat corresponding to UEs are respectively expressed as

$$\lambda_{\ell,k,[s]}^{\mathsf{SF,D}} = \omega_{\ell,k,c}^{\mathsf{D}} \lambda_{k,[s]}^{\mathsf{SF,D}}, \ \forall(\ell,k,s) \in \mathcal{L}\times\mathcal{K}_{\mathsf{D}}\times\mathcal{T}_{\mathsf{SF}}^{\mathsf{Cy},c}, \forall c \tag{4.18a}$$

$$\lambda_{\ell,k,[e]}^{\mathsf{TF,M}} = \omega_{k,c}^{\mathsf{CN,M}} \omega_{\ell,k,c}^{\mathsf{M}} \lambda_{k,[e]}^{\mathsf{TF,M}}, \ \forall(\ell,k,e) \in \mathcal{L}\times\mathcal{K}_{\mathsf{M}}\times\mathcal{T}_{\mathsf{TF}}^{\mathsf{Cy},c}, \forall c, \tag{4.18b}$$

$$\lambda_{0,k,[e]}^{\mathsf{TF,M}} = (1-\omega_k^{\mathsf{CN,M}}) \lambda_{k,[e]}^{\mathsf{TF,M}}, \ \forall(k,e) \in \mathcal{K}_{\mathsf{M}}\times\mathcal{T}_{\mathsf{TF}}^{\mathsf{Cy},c}, \forall c, \tag{4.18c}$$

$$\lambda_{0,k,[e]}^{\mathsf{TF,S}} = \lambda_{k,[e]}^{\mathsf{TF,S}}, \ \forall k \in \mathcal{K}_{\mathsf{S}}, \forall e. \tag{4.18d}$$

where $\mathcal{T}_{\mathsf{TF}}^{\mathsf{Cy},c}$ and $\mathcal{T}_{\mathsf{SF}}^{\mathsf{Cy},c}$ are frame and SF sets in cycle c. For D service, the SF time-scale data-processing requirement is cast by the following constraint,

$$(C14): \quad T_{\mathsf{D}} R_{\ell,k,[s]}^{\mathsf{SF,D}} \geq \lambda_{\ell,k,[s]}^{\mathsf{SF,D}}, \forall \ell, \forall k \in \mathcal{K}_{\mathsf{D}}, \forall s.$$

Service queues

Assume that TAPs and LSat are equipped with service-specific buffers. Each TAP/LSat is assumed to maintain separate queues for the UE data flows [116]. Subsequently, let $q_{\ell,k}^{\mathsf{M},[n_{\mathsf{M}}]}$, $q_{0,k}^{\mathsf{M},[n_{\mathsf{M}}]}$, and $q_{0,k}^{\mathsf{S},[n_{\mathsf{S}}]}$ be the queue lengths at AP_ℓ/LSat of flow k of services M,S at each RB time, the queue length evolution over time is expressed as

$$q_{\ell,k}^{\mathsf{M},[n_{\mathsf{M}}+1]} = \left[q_{\ell,k}^{\mathsf{M},[n_{\mathsf{M}}]} + \lambda_{\ell,k}^{\mathsf{M},[n_{\mathsf{M}}]} - T_{\mathsf{M}} R_{\ell,k}^{\mathsf{M},[n_{\mathsf{M}}]}\right]^+, \forall \ell, \forall k \in \mathcal{K}_{\mathsf{M}}, \tag{4.18a}$$

$$q_{0,k}^{\mathsf{x},[n_{\mathsf{x}}+1]} = \left[q_{0,k}^{\mathsf{x},[n_{\mathsf{x}}]} + \lambda_{0,k}^{\mathsf{x},[n_{\mathsf{x}}]} - T_{\mathsf{x}} R_{0,k}^{\mathsf{x},[n_{\mathsf{x}}]}\right]^+, \forall k \in \mathcal{K}_{\mathsf{x}}, \forall \mathsf{x}\in\{\mathsf{M},\mathsf{S}\}, \tag{4.18b}$$

where $[\cdot]^+ = \max\{0, \cdot\}$; $\lambda_{\ell,k}^{\mathsf{M},[n_\mathsf{M}]}$, $\lambda_{0,k}^{\mathsf{M},[n_\mathsf{M}]}$, and $\lambda_{0,k}^{\mathsf{S},[n_\mathsf{S}]}$ are intermediate data arrival notations with setting

$$\begin{cases} \lambda_{\ell,k}^{\mathsf{M},[n_\mathsf{M}]} = \lambda_{\ell,k,[e]}^{\mathsf{TF},\mathsf{M}},\ \lambda_{0,k}^{\mathsf{x},[n_\mathsf{x}]} = \lambda_{0,k,[e]}^{\mathsf{TF},\mathsf{x}} \text{ if } \mathsf{mod}(n_\mathsf{x}, 10N_\mathsf{x}) = 1, \\ \lambda_{\ell,k}^{\mathsf{M},[n_\mathsf{M}]} = 0,\ \lambda_{0,k}^{\mathsf{x},[n_\mathsf{x}]} = 0 \text{ otherwise}, \end{cases} \tag{4.19}$$

where $\{n_\mathsf{x}|_{\mathsf{mod}(n_\mathsf{x},10N_\mathsf{x})=1}\}$ indicates first RB time of each frame. To maintain stability, the total queue length of each service $\mathsf{x} \in \{\mathsf{M}, \mathsf{S}\}$ at AP_ℓ/LSat must satisfy [116]

$$\begin{aligned} (C15):\ & q_\ell^{\mathsf{M},[n_\mathsf{M}]} = \textstyle\sum_{\forall k \in \mathcal{K}_\mathsf{m}} q_{\ell,k}^{\mathsf{M},[n_\mathsf{M}]} \le q_\ell^{\mathsf{M},\mathsf{max}},\ \forall(\ell, n_\mathsf{M}), \\ (C16):\ & q_0^{\mathsf{x},[n_\mathsf{x}]} = \textstyle\sum_{\forall k \in \mathcal{K}_\mathsf{x}} q_{0,k}^{\mathsf{x},[n_\mathsf{x}]} \le q_0^{\mathsf{x},\mathsf{max}},\ \forall n_\mathsf{x}, \mathsf{x} \in \{\mathsf{M}, \mathsf{S}\}, \end{aligned}$$

where $q_\ell^{\mathsf{M},\mathsf{max}}$ and $q_0^{\mathsf{x},\mathsf{max}}$ are the maximum queue-length of AP_ℓ and LSat for corresponding services, respectively. For convenience, let's denote $\boldsymbol{q}_c^{\mathsf{TN}} = \{q_{\ell,k}^{\mathsf{M},[n_\mathsf{M}]}|_{\forall \ell, \forall k, \forall n_\mathsf{M} \in \mathcal{T}_\mathsf{M}^{\mathsf{Cy},c}}\}$, $\boldsymbol{q}_c^{\mathsf{Sat}} = \{q_{0,k}^{\mathsf{x},[n_\mathsf{x}]}|_{\forall k, \forall n_\mathsf{x} \in \mathcal{T}_\mathsf{x}^{\mathsf{Cy},c}, \forall \mathsf{x} \in \{\mathsf{M},\mathsf{S}\}}\}$, and $\boldsymbol{q}_c \triangleq \{\boldsymbol{q}_c^{\mathsf{TN}}, \boldsymbol{q}_c^{\mathsf{Sat}}\}$.

Remark 10. *The arrival data packet of* D *services is completely served within the next SF, as ensured by constraint* $(C14)$*. Hence, the buffer of* D *services is cleared after each SF, and the QL of* D *services is not considered.*

4.3.6 Problem Formulation

Typically, the uneven traffic demand of services and the inefficient spectrum utilization may lead to congestion in systems. Hence, this work aims to minimize the system congestion, i.e., the mean system QL, by optimizing the BWA, traffic split decision, AP-UE and LSat-UE associations, RB assignment, and power control. As a result, the utility function is the total average QL per RB time of M and S services at TAPs and LSats in each cycle c defined as

$$f_{\mathsf{obj}}(\boldsymbol{q}_c) = \frac{1}{N_\mathsf{M} N_\mathsf{SF}} \sum_{\forall n_\mathsf{M} \in \mathcal{T}_\mathsf{M}^{\mathsf{Cy},c}, \forall \ell} q_\ell^{\mathsf{M},[n_\mathsf{M}]} + \sum_{\mathsf{x} \in \{\mathsf{M},\mathsf{S}\}} \frac{1}{N_\mathsf{x} N_\mathsf{SF}} \sum_{\forall n_\mathsf{x} \in \mathcal{T}_\mathsf{x}^{\mathsf{Cy},c}} q_0^{\mathsf{x},[n_\mathsf{x}]}. \tag{4.20}$$

Subsequently, two problems, "DT-JointRA" and "RT-Refine", will be studied. The former optimizes systems for each cycle c and the latter optimizes TN short-term decisions for each SF, which are solved at the CNC and TCU, respectively.

DT-JointRA Problem

For cycle c, we consider

$$(\mathcal{P}_0)_c: \min_{\boldsymbol{b}_c,\boldsymbol{\omega}_c,\boldsymbol{P}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c,\boldsymbol{q}_c} f_{\mathsf{obj}}(\boldsymbol{q}_c) \text{ s.t. constraints } (C1)-(C16),$$
$$(C0): b^{\mathsf{x}}_{v_{\mathsf{x}},c}, \alpha^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{\ell,k}, \beta^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{k} \in \{0,1\}, \forall(\ell,k,v_{\mathsf{x}},n_{\mathsf{x}},c).$$

RT-Refine Problem

Regarding implementation aspects, problem $(\mathcal{P}_0)_c$ requires future information $\{\boldsymbol{h}_c, \boldsymbol{g}_c, \boldsymbol{\lambda}_c\}$ in cycle c which is challenging in practice. Assuming that the predicted information $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c, \hat{\boldsymbol{\lambda}}_c\}$ is used instead of $\{\boldsymbol{h}_c, \boldsymbol{g}_c, \boldsymbol{\lambda}_c\}$, the corresponding solution should be adjusted appropriately with actual information. Hence, this section formulates the refinement problem to adjust the obtained solution at each SF by re-optimizing TN short-term decisions. However, due to the long propagation distance in SatNet, the channel estimation and RA in SatNet at each SF are challenging. Hence, the given RA in SatNet can remain unchanged.

Let's call UEs served by TAPs and LSat as TN UEs and SatNet UEs. Assuming that TAPs only estimate TN UE CSIs, the predicted CSIs of SatNet UEs should be used. For a given SatCom RA, the channel uncertainty of M UEs served by LSat leads to errors in the ISyI power $\Theta_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ and $\Theta_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]}$. Hence, we impose the interference margin coefficient $\kappa \geq 1$ and use $\hat{\Theta}_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]} = \kappa\Theta_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ and $\hat{\Theta}_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]} = \kappa\Theta_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ for refinement problem.

Let $\boldsymbol{h}_e^*$ be the channel gain matrix in frame e where CSIs of TN UEs are actual while those of SatNet UEs are extracted from $\hat{\boldsymbol{h}}_e$. Based on channel gains $\{\boldsymbol{h}_e^*, \hat{\boldsymbol{g}}_e\}$ and actual traffic $\boldsymbol{\lambda}_e$ in frame e, for a fixed $\{\boldsymbol{b}_c, \boldsymbol{\omega}_c, \boldsymbol{p}_{0,c}, \boldsymbol{\beta}_c\}$, the refinement problem for TN short-term decisions at SF s is formulated as

$$(\mathcal{P}_0^{\mathsf{refi}})_s: \min_{\boldsymbol{p}[s],\boldsymbol{\alpha}[s],\boldsymbol{q}[s]} f_{\mathsf{obj}}(\boldsymbol{q}[s])$$
$$\text{s.t. } (C4)_s-(C6)_s, (C9)_s, (C10)_s, (C11)_s, (C14)_s, (C15)'_s, (C16)'_s,$$

where index s used in $\boldsymbol{x}[s]$ and $(Cx)_s$ indicates time index adaptation to SF s for variable $\boldsymbol{x}$ and constraint (Cx), $(C15)'$ and $(C16)'$ are the constraint revised from $(C15)$ and $(C16)$ in which terms $\hat{\Theta}_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ and $\hat{\Theta}_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ are used instead of $\Theta_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ and $\Theta_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ in rate functions, respectively.

Challenges in solving $(\mathcal{P}_0)_c$ *and* $(\mathcal{P}_0^{\mathsf{refi}})_s$: these problems contain the non-convex constraints, SINR and rate functions. Especially, problem $(\mathcal{P}_0)_c$ further consists of both contin-

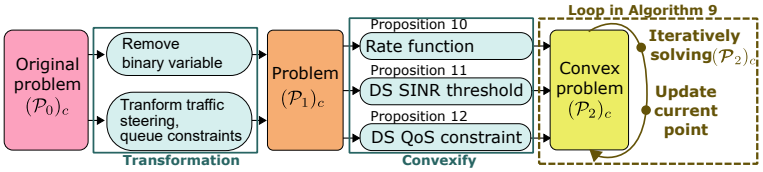


Figure 4.5: Developing workflow.

uous and binary variables. Hence, $(\mathcal{P}_0^{\text{refi}})_s$ and $(\mathcal{P}_0)_c$ are non-convex and MINLP problems. Additionally, the unknown channel and traffic information in $(\mathcal{P}_0)_c$ makes it more complicated.

Overall workflow

Let $\{\hat{\boldsymbol{x}}_c\}$ and $\{\boldsymbol{x}_c\}$ with $\boldsymbol{x} \in \{\boldsymbol{h}, \boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{u}^{\text{ue}}, \boldsymbol{u}^{\text{sat}}, \text{TLE}\}$ be the DT-virtual and real information vector in cycle c, the overall workflow is summarized as

① Based on the updated UE position $\hat{\boldsymbol{u}}_{c-1}^{\text{ue}}$, orbit information TLE_{c-1}, and traffic $\boldsymbol{\lambda}_{c-1}$ in cycle $(c-1)$, the DT model first predicts $\hat{\boldsymbol{\lambda}}_c$, $\{\hat{\boldsymbol{u}}_c^{\text{ue}}, \hat{\boldsymbol{u}}_c^{\text{sat}}\}$, and then $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c\}$ in cycle c.

② Based on $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c, \hat{\boldsymbol{\lambda}}_c\}$, problem $(\mathcal{P}_0)_c$ is solved at the CNC to obtain long-term decisions and preliminary estimates of short-term ones $\{\boldsymbol{b}_c, \boldsymbol{\omega}_c, \boldsymbol{P}_c, \boldsymbol{\alpha}_c, \boldsymbol{\beta}_c\}$ for cycle c.

③ For each frame e in cycle c, using channel gain $\boldsymbol{h}_e^*$, arrival data rate $\boldsymbol{\lambda}_e$ in the real system, and given solution $\{\boldsymbol{b}_c, \boldsymbol{\omega}_c, \boldsymbol{P}_c, \boldsymbol{\alpha}_c, \boldsymbol{\beta}_c\}$, the initial power control, UE association, and RB assignment $\{\boldsymbol{p}_c, \boldsymbol{\alpha}_c\}$ in the TN, which are outcomes of step 2, are refined by solving problem $(\mathcal{P}_0^{\text{refi}})_s$ at TCU at each SF s in frame e.

4.4 Proposed Solutions

This section first proposes the solution for problem $(\mathcal{P}_0)_c$. Subsequently, the solution for the problem $(\mathcal{P}_0^{\text{refi}})_s$ is proposed to refine TN short-term decisions with actual information. The solution workflows are summarized in Figs. 4.5 and 4.6.

4.4.1 DT-based Prediction

At the beginning of every cycle, DT model needs the updated UEs position and LSat orbit information for channel prediction. Assuming that the cycle duration is sufficiently small, the changes in the movement trend of UEs and velocity are negligible. Hence, with a given updated movement of route and velocity in cycle $(c-1)$, the UE position in cycle c can be estimated by the approach given in [52]. Subsequently, channel components in cycle c are

Algorithm 8 DT-BASED PREDICTION

1: **Input:** Actual information $\{\boldsymbol{\lambda}_{c-1}, \boldsymbol{u}^{\mathsf{ue}}_{c-1}, \mathsf{TLE}_{c-1}\}$ in cycle $(c-1)$.
2: Predict $\{\hat{\boldsymbol{u}}^{\mathsf{ue}}_{c}, \hat{\boldsymbol{u}}^{\mathsf{sat}}_{c}\}$ and construct $\{\widehat{\mathsf{UE}}[e], \widehat{\mathsf{Sat}}[e] | \forall e \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{TF}}\}$,
3: Predict $\{\mathsf{chan}_c, \hat{\boldsymbol{\lambda}}_c\}$ by (4.6), (4.21).
4: Construct $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c\}$ by using $\mathsf{chan}_{\mathsf{c}}$ with $\xi = 1$.
5: **Output:** Predicted information $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c, \hat{\boldsymbol{\lambda}}_c\}$.

identified by (4.6). Based on output $\mathsf{chan}_c \triangleq \{\mathsf{chan}_{[e]} | \forall e \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{TF}}\}$, channel gain $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c\}$ used for problem $(\mathcal{P}_0)_c$ is computed with DT channel coefficient $\xi = 1$.

Regarding UE traffic, one can assume that UEs continue to use specific services over a given time duration. As a result, their traffic pattern and arrival traffic remain relatively stable. Therefore, the arrival data in cycle c can be estimated by the average in cycle $(c-1)$ as

$$\hat{\lambda}^{\mathsf{SF,D}}_{k,[s]} = \frac{1}{N_{\mathsf{SF}}} \sum\nolimits_{s' \in \mathcal{T}^{\mathsf{Cy},c-1}_{\mathsf{SF}}} \lambda^{\mathsf{SF,D}}_{k,[s']}, \ \forall s \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{SF}}, \tag{4.20a}$$

$$\hat{\lambda}^{\mathsf{TF},\mathsf{x}}_{k,[e]} = \frac{1}{N_{\mathsf{TF}}} \sum\nolimits_{e' \in \mathcal{T}^{\mathsf{Cy},c-1}_{\mathsf{TF}}} \lambda^{\mathsf{TF},\mathsf{x}}_{k,[e']}, \ \forall e \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{TF}}, \mathsf{x} \in \{\mathsf{M}, \mathsf{S}\}. \tag{4.20b}$$

The DT-based prediction mechanism is summarized in Alg. 8.

4.4.2 DT-JointRA Solution

Compressed-Sensing-Based Relaxation

Considering the system model and the optimization problem $(\mathcal{P}_0)_c$, one can see the relationship between the binary variable $\{\boldsymbol{\alpha}_c, \boldsymbol{\beta}_c\}$ and the continuous variable $\boldsymbol{P}_c$ as follows. Consider $\mathsf{RB}_{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}$:

- If AP_ℓ serves $\mathsf{UE}_k \Rightarrow \alpha^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{\ell,k} = 1$, $p^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{\ell,k} > 0$;
  if AP_ℓ does not serve UE_k $\alpha^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{\ell,k} = 0, p^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{\ell,k} = 0$.
- If the LSat serves $\mathsf{UE}_k \Rightarrow \beta^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{k} = 1$, $p^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{0,k} > 0$;
  if the LSat does not serve UE_k $\beta^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{k} = 0, p^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{0,k} = 0$.

Based on this relationship, the binary variables $\boldsymbol{\alpha}_c$ and $\boldsymbol{\beta}_c$ can be respectively represented by the continuous one $\boldsymbol{P}_c$ as

$$\alpha^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{\ell,k} = \|p^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{\ell,k}\|_0, \quad \beta^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{k} = \|p^{[\nu_{\mathsf{x}}, n_{\mathsf{x}}]}_{0,k}\|_0. \tag{4.21}$$

Besides, consider $(C4), (C7)$, one can see that SC v_{x} is activated in cycle c, i.e., $b^{\mathsf{x}}_{v_{\mathsf{x}},c} = 1$, only if it is used by at least one TAP/LSat-UE link at any RB time $n_{\mathsf{x}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{x}}$, i.e., $\alpha^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{\ell,k} = 1$ or $\beta^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{k} = 1$. Combining (4.21), $\boldsymbol{b}_c$ is rewritten as

$$b^{\mathsf{D}}_{v_{\mathsf{D}},c} = \| \sum_{\forall(\ell,k),\forall n_{\mathsf{D}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{D}}} \alpha^{[v_{\mathsf{D}},n_{\mathsf{D}}]}_{\ell,k} \|_0 = \| p^{\mathsf{D}}_{v_{\mathsf{D}},c} \|_0, \ \forall (v_{\mathsf{D}}, c), \tag{4.21a}$$

$$b^{\mathsf{M}}_{v_{\mathsf{M}},c} = \| \sum_{\forall(\ell,k),\forall n_{\mathsf{M}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{M}}} \alpha^{[v_{\mathsf{M}},n_{\mathsf{M}}]}_{\ell,k} + \sum_{\forall k,\forall n_{\mathsf{M}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{M}}} \beta^{[v_{\mathsf{M}},n_{\mathsf{M}}]}_{k} \|_0 = \| P^{\mathsf{M}}_{v_{\mathsf{M}},c} \|_0, \ \forall (v_{\mathsf{M}}, c), \tag{4.21b}$$

$$b^{\mathsf{S}}_{v_{\mathsf{S}},c} = \| \sum\nolimits_{\forall k,\forall n_{\mathsf{S}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{S}}} \beta^{[v_{\mathsf{S}},n_{\mathsf{S}}]}_{k} \|_0 = \| p^{\mathsf{S}}_{0,v_{\mathsf{S}},c} \|_0, \ \forall (v_{\mathsf{S}}, c). \tag{4.21c}$$

where $P^{\mathsf{M}}_{v_{\mathsf{M}},c} = p^{\mathsf{M}}_{v_{\mathsf{M}},c} + p^{\mathsf{M}}_{0,v_{\mathsf{M}},c}$, $p^{\mathsf{x}}_{v_{\mathsf{x}},c} = \sum_{\forall(\ell,k),\forall n_{\mathsf{x}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{x}}} p^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{\ell,k}$ and $P^{\mathsf{x}}_{v_{\mathsf{x}},c} = \sum_{\forall k,\forall n_{\mathsf{x}} \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{x}}} p^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{0,k}$. Thanks to relationship (4.22), constraints $(C4)$,$(C7)$ can be omitted.

Exploiting (4.21), the binary components can be replaced by the corresponding ℓ_0-norm components. Moreover, the binary variables in the production components and the binary arguments of SINR/rate functions can be omitted. Specifically, $R^{\mathsf{SF,D}}_{\ell,k,[s]}(\boldsymbol{P}_c, \boldsymbol{\alpha}_c)$ is reformulated as

$$R^{\mathsf{SF,D}}_{\ell,k,[s]}(\boldsymbol{P}_c) = w_{\mathsf{D}} \sum_{\forall(v_{\mathsf{D}},n_{\mathsf{D}})} \log_2\left(1 + \gamma^{\mathsf{D},[v_{\mathsf{D}},n_{\mathsf{D}}]}_{\ell,k}(\boldsymbol{P}_c)\right) - \chi_{\mathsf{D}} \sqrt{\sum_{\forall(v_{\mathsf{D}},n_{\mathsf{D}})} \| p^{[v_{\mathsf{D}},n_{\mathsf{D}}]}_{\ell,k} \|_0}, \tag{4.22}$$

Subsequently, the ℓ_0-norm component $\|x\|_0, \forall x \geq 0$, can be approximated by concave function $\mathcal{F}_{\mathsf{apx}}(x) \triangleq 1 - e^{-x/\epsilon}$, $0 < \epsilon \ll 1$, as $\|x\|_0 \approx \mathcal{F}_{\mathsf{apx}}(x)$. Let $\mathcal{F}^{(i)}_{\mathsf{apx}}(x)$ be an upper bound of $\mathcal{F}_{\mathsf{apx}}(x)$ at feasible point $(x^{(i)})$, the ℓ_0-norm components in (4.21) and (4.22) can be approximated at iteration i as

$$\|\tau\|_0 \approx \mathcal{F}_{\mathsf{apx}}(\tau) \leq \mathcal{F}^{(i)}_{\mathsf{apx}}(\tau), \ \tau \in \{p^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{\ell,k}, p^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{0,k}, p^{\mathsf{x}}_{v_{\mathsf{x}},c}, p^{\mathsf{x}}_{0,v_{\mathsf{x}},c}, P^{\mathsf{x}}_{v_{\mathsf{x}},c}\}, \tag{4.23}$$

where upper bound $\mathcal{F}^{(i)}_{\mathsf{apx}}(x)$ can be obtained based on [52] as

$$\mathcal{F}_{\mathsf{apx}}(x) \leq \mathcal{F}^{(i)}_{\mathsf{apx}}(x) \triangleq 1/\epsilon \exp(-x^{(i)}/\epsilon)(x - x^{(i)} - \epsilon) + 1 \tag{4.24}$$

Utilize (4.21), (4.22), and (4.23), by replacing binary component $b^{\mathsf{x}}_{v_{\mathsf{x}},c}$, $\alpha^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{\ell,k}$, and $\beta^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{k}$ by the corresponding approximated terms $\mathcal{F}^{(i)}_{\mathsf{apx}}(\cdot)$ at iteration i, constraints $(C1)$–$(C3)$, $(C5)$, $(C6)$, $(C8)$–$(C9)$ are directly transformed into convex ones and $(C10)$ is approximated by a non-convex

constraint as

$$(\tilde{C}1): \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{\mathsf{D}}_{v_{\mathsf{D}},c}) + \sum_{i=1}^{2v_{\mathsf{D}}+1} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{\mathsf{M}}_{i,c} + P^{\mathsf{M}}_{i,c}) + \sum_{j=1}^{4v_{\mathsf{D}}+2} \mathcal{F}^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{j,c}) \leq 1,\ \forall (v_{\mathsf{D}}, c),$$

$$(\tilde{C}2): \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{\mathsf{M}}_{f,c} + P^{\mathsf{M}}_{f,c}) + \sum_{j=1}^{\bar{Q}^{\mathsf{M}}_{\mathsf{S}}(f+0.5)} \mathcal{F}^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{j,c}) \leq 1,\ \forall (v_{\mathsf{M}}, c),$$

$$(\tilde{C}3): w_{\mathsf{D}} \sum_{\forall f \in \mathcal{V}_{\mathsf{D}}} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{\mathsf{D}}_{f,c}) + w_{\mathsf{M}} \sum_{\forall f \in \mathcal{V}_{\mathsf{M}}} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{\mathsf{M}}_{f,c} + P^{\mathsf{M}}_{f,c}) + w_{\mathsf{S}} \sum_{\forall f \in \mathcal{V}_{\mathsf{S}}} \mathcal{F}^{(i)}_{\mathsf{apx}}(P^{\mathsf{S}}_{f,c}) + B_{\mathsf{G1}} + B_{\mathsf{G2}} \leq B,\ \forall c,$$

$$(\tilde{C}5): \sum_{\forall k} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{\ell,k}) \leq 1, \quad \forall \ell, \forall (v_{\mathsf{x}}, n_{\mathsf{x}}) \in \mathcal{B}^{\mathsf{x}}_c, \mathsf{x} \in \{\mathsf{D}, \mathsf{M}\},$$

$$(\tilde{C}6): \sum_{\forall \ell} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{[v_{\mathsf{D}},n_{\mathsf{D}}]}_{\ell,k}) \leq 1, \quad \forall k \in \mathcal{K}_{\mathsf{D}}, \forall (v_{\mathsf{D}}, n_{\mathsf{D}}) \in \mathcal{B}^{\mathsf{D}}_c,$$

$$(\tilde{C}8): \sum_{\forall k} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{[v_{\mathsf{x}},n_{\mathsf{x}}]}_{0,k}) \leq 1, \quad \forall (v_{\mathsf{x}}, n_{\mathsf{x}}) \in \mathcal{B}^{\mathsf{x}}_c, \mathsf{x} \in \{\mathsf{M}, \mathsf{S}\},$$

$$(\tilde{C}9): \sum_{\forall \ell} \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{[v_{\mathsf{M}},n_{\mathsf{M}}]}_{\ell,k}) + \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{[v_{\mathsf{M}},n_{\mathsf{M}}]}_{0,k}) \leq 1, \forall k \in \mathcal{K}_{\mathsf{M}}, \forall (v_{\mathsf{M}}, n_{\mathsf{M}}) \in \mathcal{B}^{\mathsf{M}}_c,$$

$$(\bar{C}10): \gamma^{\mathsf{D},[v_{\mathsf{D}},n_{\mathsf{D}}]}_{\ell,k}(\boldsymbol{P}_c) \geq \mathcal{F}^{(i)}_{\mathsf{apx}}(p^{[v_{\mathsf{D}},n_{\mathsf{D}}]}_{\ell,k})\gamma^{\mathsf{D}}_0,\ \forall (\ell, k, v_{\mathsf{D}}, n_{\mathsf{D}}).$$

Transform Traffic and Queue Constraints

$(C14)-(C16)$ are non-convex due to non-convex rate functions, coupling between traffic steering variables, and queue length dependency. First, we introduce variable $\bar{\boldsymbol{\omega}} = \{\bar{\omega}^{\mathsf{x}}_{\ell,k,c} | \forall (\ell, k, c, \mathsf{x})\}$ and use it instead of $\omega^{\mathsf{CN,M}}_{k,c}$ and $\omega^{\mathsf{x}}_{\ell,k,c}$ with the relationship

$$\bar{\omega}^{\mathsf{D}}_{\ell,k,c} = \omega^{\mathsf{D}}_{\ell,k,c} \quad \text{and} \quad \bar{\omega}^{\mathsf{M}}_{\ell,k,c} = \omega^{\mathsf{CN,M}}_{k,c}\omega^{\mathsf{M}}_{\ell,k,c}, \tag{4.25}$$

Hence, constraint $(C13)$ is rewritten for the DS service as

$$(\tilde{C}13): \quad \sum_{\forall \ell \in \mathcal{L}} \bar{\omega}^{\mathsf{D}}_{\ell,k,c} = 1, \forall k \in \mathcal{K}_{\mathsf{D}}, \forall c.$$

The traffic arrival rates (4.18a)-(4.18c) are rewritten as

$$\lambda^{\mathsf{SF,D}}_{\ell,k,[s]} = \bar{\omega}^{\mathsf{D}}_{\ell,k,c}\lambda^{\mathsf{SF,D}}_{k,[s]},\ \ \forall \ell, \forall k \in \mathcal{K}_{\mathsf{D}}, \forall s \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{SF}}, \forall c, \tag{4.25a}$$

$$\lambda^{\mathsf{TF,M}}_{\ell,k,[e]} = \bar{\omega}^{\mathsf{M}}_{\ell,k,c}\lambda^{\mathsf{TF,M}}_{k,[e]}, \quad \forall \ell, \forall k \in \mathcal{K}_{\mathsf{M}},\ \ e \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{TF}}, \forall c, \tag{4.25b}$$

$$\lambda^{\mathsf{TF,M}}_{0,k,[e]} = (1 - \sum_{\forall \ell} \bar{\omega}^{\mathsf{M}}_{\ell,k,c})\lambda^{\mathsf{TF,M}}_{k,[e]},\ \ \forall k \in \mathcal{K}_{\mathsf{M}}, \forall e \in \mathcal{T}^{\mathsf{Cy},c}_{\mathsf{TF}}, \forall c. \tag{4.25c}$$

Remark 11. *The integrity of MS traffic is ensured thanks to representation* (4.25), (4.26). *Particularly, based on* (4.25), (4.18c) *is rewritten by* (4.25c) *due to* $\sum_{\forall \ell} \bar{\omega}^{\mathsf{M}}_{\ell,k,c} = \omega^{\mathsf{CN,M}}_{k,c}$. *Since* $\sum_{\forall \ell} \lambda^{\mathsf{TF,M}}_{\ell,k,[e]} + \lambda^{\mathsf{TF,M}}_{0,k,[e]} = \lambda^{\mathsf{TF,M}}_{k,[e]}$, *the integrity is ensured.*

Subsequently, we introduce slack variables $\bar{\boldsymbol{q}}_c = \{\bar{q}^{\mathsf{x},[n_\mathsf{x}]}_{\ell,k}, \bar{q}^{\mathsf{x},[n_\mathsf{x}]}_{0,k} \ |\forall(\ell,k),\ \forall n_\mathsf{x} \in \mathcal{T}^{\mathsf{Cy},c}_\mathsf{x}, \forall \mathsf{x} \in \{\mathsf{M},\mathsf{S}\}\}$ as the queue length upper bound and $\boldsymbol{r}_c = \{r^{\mathsf{x},[v_\mathsf{x},n_\mathsf{x}]}_{\ell,k}, r^{\mathsf{x},[v_\mathsf{x},n_\mathsf{x}]}_{0,k}, r^{\mathsf{SF,D}}_{\ell,k,[s]} | \forall(\ell,k,s,v_\mathsf{x},n_\mathsf{x}), \forall \mathsf{x} \in \mathcal{S}\}$ as the rate's lower bound; and transform (4.19) and $(C14)-(C16)$ as

$$
\begin{aligned}
&(\tilde{C}14): T_\mathsf{D} r^{\mathsf{SF,D}}_{\ell,k,[s]} \geq \lambda^{\mathsf{SF,D}}_{\ell,k,[s]}, \forall \ell, \forall k \in \mathcal{K}_\mathsf{d}, \forall s,\\
&(\tilde{C}15a): \bar{q}^{\mathsf{M},[n_\mathsf{M}]}_{\ell,k} + \lambda^{\mathsf{m},[n_\mathsf{M}]}_{\ell,k} - T_\mathsf{M} r^{\mathsf{M},[n_\mathsf{M}]}_{\ell,k} \leq \bar{q}^{\mathsf{M},[n_\mathsf{M}+1]}_{\ell,k}, \forall(\ell,k,n_\mathsf{M}),\\
&(\tilde{C}15b): 0 \leq \bar{q}^{\mathsf{M},[n_\mathsf{M}]}_{\ell,k}, \textstyle\sum_{\forall k \in \mathcal{K}_\mathsf{M}} \bar{q}^{\mathsf{M},[n_\mathsf{M}]}_{\ell,k} \leq q^{\mathsf{M,max}}_{\ell}, \forall(\ell,n_\mathsf{M}),\\
&(\tilde{C}16a): \bar{q}^{\mathsf{x},[n_\mathsf{x}]}_{0,k} + \lambda^{\mathsf{x},[n_\mathsf{x}]}_{0,k} - T_\mathsf{x} r^{\mathsf{x},[n_\mathsf{x}]}_{0,k} \leq \bar{q}^{\mathsf{x},[n_\mathsf{x}+1]}_{0,k}, \forall(k,n_\mathsf{x}), \forall \mathsf{x} \in \{\mathsf{M},\mathsf{S}\},\\
&(\tilde{C}16b): 0 \leq \bar{q}^{\mathsf{x},[n_\mathsf{x}]}_{0,k}, \textstyle\sum_{\forall k \in \mathcal{K}_\mathsf{x}} \bar{q}^{\mathsf{x},[n_\mathsf{x}]}_{0,k} \leq q^{\mathsf{x,max}}_{0},\ \forall n_\mathsf{x}, \forall \mathsf{x} \in \{\mathsf{M},\mathsf{S}\},
\end{aligned}
$$

where $r^{\mathsf{M},[n_\mathsf{M}]}_{\ell,k} = \sum_{\forall v_\mathsf{M}} r^{\mathsf{M},[v_\mathsf{M},n_\mathsf{M}]}_{\ell,k}$, $r^{\mathsf{x},[n_\mathsf{x}]}_{0,k} = \sum_{\forall v_\mathsf{x}} r^{\mathsf{x},[v_\mathsf{x},n_\mathsf{x}]}_{0,k}$ while $r^{\mathsf{M},[v_\mathsf{M},n_\mathsf{M}]}_{\ell,k}$, $r^{\mathsf{SF,D}}_{\ell,k,[s]}$, and $r^{\mathsf{x},[v_\mathsf{x},n_\mathsf{x}]}_{0,k}$ satisfy constraints

$$
\begin{aligned}
&(C17.1): R^{\mathsf{M},[v_\mathsf{M},n_\mathsf{M}]}_{\ell,k}(\boldsymbol{P}_c) \geq r^{\mathsf{M},[v_\mathsf{M},n_\mathsf{M}]}_{\ell,k},\ \forall(\ell,v_\mathsf{M},n_\mathsf{M}), \forall k \in \mathcal{K}_\mathsf{M},\\
&(C17.2): R^{\mathsf{SF,D}}_{\ell,k,[s]}(\boldsymbol{P}_c) \geq r^{\mathsf{SF,D}}_{\ell,k,[s]},\ \forall \ell, \forall k \in \mathcal{K}_\mathsf{D}, \forall s,\\
&(C18): r^{\mathsf{x},[v_\mathsf{x},n_\mathsf{x}]}_{0,k}(\boldsymbol{P}_c) \geq r^{\mathsf{x},[v_\mathsf{x},n_\mathsf{x}]}_{0,k},\ \forall k \in \mathcal{K}_\mathsf{x}, \forall(v_\mathsf{x},n_\mathsf{x}), \forall \mathsf{x} \in \{\mathsf{M},\mathsf{S}\},
\end{aligned}
$$

Therefore, problem $(\mathcal{P}_0)_c$ can be rewritten as

$$
\begin{aligned}
(\mathcal{P}_1)_c: \min_{\bar{\boldsymbol{\omega}}_c, \boldsymbol{P}_c, \bar{\boldsymbol{q}}_c, \boldsymbol{r}_c}\ & f_{\mathsf{obj}}(\bar{\boldsymbol{q}}_c) \text{ s.t. } (\tilde{C}1)-(\tilde{C}3), (\tilde{C}5), (\tilde{C}6), (\tilde{C}8), (\tilde{C}9),\\
& (\bar{C}10), (C11), (C12), (\tilde{C}14)-(\tilde{C}16), (C17), (C18).
\end{aligned}
$$

Obviously, problem $(\mathcal{P}_1)$ is still non-convex due to the non-convexity of SINR and rate constraints $(\bar{C}10), (C17), (C18)$.

Convexify SINR and Rate Constraints

These constraints can be convexified by the following propositions.

Proposition 10. *Constraints* $(C17.1), (C18)$ *are convexified as*

$$
\begin{aligned}
&(\tilde{C}17a): w_{\mathsf{M}} F_{\ell,k}^{\mathsf{R},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) \geq r_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]},\\
&(\tilde{C}17b): \Psi_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c) + \Theta_{k}^{\mathsf{TN},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c) + \sigma_{\mathsf{M},k}^2 \leq F_{\mathsf{exp}}^{(i)}(\eta_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}),\\
&(\tilde{C}18a): w_{\mathsf{M}} F_{0,k}^{\mathsf{R},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) \geq r_{0,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]},\\
&(\tilde{C}18b): \Theta_{k}^{\mathsf{Sat},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c) + \sigma_{\mathsf{M},k}^2 \leq \mathcal{F}_{\mathsf{exp}}^{(i)}(\eta_{0,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}),\\
&(\tilde{C}18c): R_{0,k}^{\mathsf{S},[\nu_{\mathsf{S}},n_{\mathsf{S}}]}(\boldsymbol{P}_c) \geq R_{0,k}^{\mathsf{S},[\nu_{\mathsf{S}},n_{\mathsf{S}}]},
\end{aligned}
$$

with $\mathcal{F}_{\mathsf{exp}}^{(i)}(u) \triangleq \exp(u^{(i)})(u-u^{(i)}+1)$, *and*

$$
\begin{aligned}
F_{\ell,k}^{\mathsf{R},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) &\triangleq \log_2(p_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]} h_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}\\
&\quad + \Psi_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c) + \Theta_{k}^{\mathsf{TN},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c) + \sigma_{\mathsf{M},k}^2) - \eta_{\ell,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}/\ln(2), \qquad (4.26)\\
F_{0,k}^{\mathsf{R},\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) &\triangleq \log_2\left(p_{0,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]} g_{k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]} + \Theta_{k}^{\mathsf{Sat},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}(\boldsymbol{P}_c) + \sigma_{\mathsf{M},k}^2\right) - \eta_{0,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}/\ln(2), \qquad (4.27)
\end{aligned}
$$

and $\boldsymbol{\eta}_c = \{\eta_{0,k}^{[\nu_{\mathsf{M}},n_{\mathsf{M}}]}, \eta_{\ell,k}^{[\nu_{\mathsf{x}},n_{\mathsf{x}}]} | \forall(\ell,k,\nu_{\mathsf{x}},n_{\mathsf{x}}), \mathsf{x} \in \{\mathsf{D},\mathsf{M}\}\}$ *is a slack variable.*

Proof: Consider function $f_{\mathsf{r}}(x,y) = \log_2(1+\frac{x}{y+a}), x,y \geq 0, a > 0$ which has the form of the rate function, constraint $f_{\mathsf{r}}(x,y) \geq z,\ z \geq 0$ is approximated as [31]

$$
\begin{cases}
\log_2(x+y+a) \geq z + u/\ln(2),\\
y + a \leq f_{\mathsf{exp}}^{(i)}(u) \triangleq \exp(u^{(i)})(u-u^{(i)}+1),
\end{cases} \qquad (4.28)
$$

where u and $u^{(i)}$ are the slack variable and its feasible point at iteration i, respectively. Apply approximation in (4.28) to log component in rate functions $R_{\ell,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ and $R_{0,k}^{\mathsf{M},[\nu_{\mathsf{M}},n_{\mathsf{M}}]}$ with setting x, y, a, and u by the desired received power, the total interference terms, noise power, and the corresponding η, constraints $(C17.1)$ and $(C18)$ are convexified by $(\tilde{C}17a)$, $(\tilde{C}17b)$ and $(\tilde{C}18)$, respectively. Specifically, $R_{0,k}^{\mathsf{S},[\nu_{\mathsf{S}},n_{\mathsf{S}}]}(\boldsymbol{P}_c)$ is concave, hence, constraint $(\tilde{C}20c)$ is convex. ■

The SINR constraint $(\bar{C}11)$ and rate constraint $(C17.2)$ for DS services can be convexified by the following proposition.

Proposition 11. *Constraint* $(\bar{C}10)$ *can be convexified as*

$$
(\tilde{C}10a): F_{\ell,k}^{\mathsf{R},\mathsf{D},[\nu_{\mathsf{D}},n_{\mathsf{D}}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) \geq \mathcal{F}_{\mathsf{apx}}^{(i)}(p_{\ell,k}^{[\nu_{\mathsf{D}},n_{\mathsf{D}}]}) \log_2(1+\gamma_0^{\mathsf{D}}),
$$

$$(\tilde{C}10b): \Psi_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}(\boldsymbol{P}_c) + \sigma_{\mathsf{D},k}^2 \leq \mathcal{F}_{\mathsf{exp}}^{(i)}(\eta_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}),$$

with $F_{\ell,k}^{\mathsf{R},\mathsf{D},[v_\mathsf{D},n_\mathsf{D}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) = \log_2\left(p_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}h_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]} + \Psi_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}(\boldsymbol{P}_c) + \sigma_{\mathsf{D},k}^2\right) - \eta_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}/\ln(2)$.

Proof: Constraint $(\bar{C}10)$ is equivalently transformed as

$$\log_2(1+\gamma_{\ell,k}^{\mathsf{D},[v_\mathsf{D},n_\mathsf{D}]}(\boldsymbol{P}_c)) \geq \mathcal{F}_{\mathsf{apx}}^{(i)}(p_{\ell,k}^{\mathsf{D},[v_\mathsf{D},n_\mathsf{D}]})\log_2(1+\gamma_0^\mathsf{D}), \quad \forall(\ell,k,v_\mathsf{D},n_\mathsf{D}). \tag{4.29}$$

Apply approximation (4.28) to the log component by setting $x = p_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}h_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}$, $y = \Psi_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}$, $a = \sigma_{\mathsf{D},k}^2$, and $u = \eta_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}$, we obtain convex constraint $(\tilde{C}10a)$ and $(\tilde{C}10b)$. ■

Proposition 12. *Constraint* $(C17.2)$ *can be convexified as*

$$(\tilde{C}17c): w_\mathsf{D}\sum_{\forall(v_\mathsf{D},n_\mathsf{D})} F_{\ell,k}^{\mathsf{R},\mathsf{D},[v_\mathsf{D},n_\mathsf{D}]}(\boldsymbol{P}_c,\boldsymbol{\eta}_c) - \chi_\mathsf{D}\mathcal{F}_{\mathsf{sqrt}}^{(i)}(\zeta_{\ell,k,[s]}) \geq r_{\ell,k,[s]}^{\mathsf{SF},\mathsf{D}},$$
$$(\tilde{C}17d): \zeta_{\ell,k,[s]} \geq \sum\nolimits_{\forall(v_\mathsf{x},n_\mathsf{x})}\mathcal{F}_{\mathsf{apx}}^{(i)}(p_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}),$$

where $\mathcal{F}_{\mathsf{sqrt}}^{(i)}(x) \triangleq \frac{x}{2\sqrt{x^{(i)}}} + \frac{\sqrt{x^{(i)}}}{2}$ *and* $\boldsymbol{\zeta}_c \triangleq \{\zeta_{\ell,k,[s]}|\forall\ell,\forall k\in\mathcal{K}_\mathsf{D},\forall s\in\mathcal{T}_\mathsf{SF}^{\mathsf{Cy},c}\}$ *is a slack variable.*

Proof: Let $\zeta_{\ell,k,[s]}$ be the upper bound of the RB number allocated to $\mathsf{AP}_\ell - \mathsf{UE}_k^\mathsf{D}$ link in SF s, constraint $(C17.2)$ is rewritten as

$$w_\mathsf{D}\sum_{\forall(v_\mathsf{D},n_\mathsf{D})}\log_2\left(1+\gamma_{\ell,k}^{\mathsf{D},[v_\mathsf{D},n_\mathsf{D}]}(\boldsymbol{P}_c)\right) - \chi_\mathsf{D}\sqrt{\zeta_{\ell,k,[s]}} \geq r_{\ell,k,[s]}^{\mathsf{SF},\mathsf{D}}, \tag{4.29a}$$

$$\sum_{\forall(v_\mathsf{D},n_\mathsf{D})}\|p_{\ell,k}^{[v_\mathsf{D},n_\mathsf{D}]}\|_0 \leq \zeta_{\ell,k,[s]}. \tag{4.29b}$$

Apply approximation (4.23) to ℓ_0-norm components, (4.29b) is convexified as $(\tilde{C}17d)$. In (4.29a), the log component can be convexified as in proposition 12 while the square root component is convexified as follows. Apply SCA technique, function $\mathcal{F}_{\mathsf{sqrt}}(x) = \sqrt{x}$ is approximated at iteration i as

$$\mathcal{F}_{\mathsf{sqrt}}(x) \leq \frac{x}{2\sqrt{x^{(i)}}} + \frac{\sqrt{x^{(i)}}}{/}2 := \mathcal{F}_{\mathsf{sqrt}}^{(i)}(x). \tag{4.30}$$

Use approximation (4.30) for `sqrt` term in (4.29a), we obtain the convex form as $(\tilde{C}17c)$. ■

Thanks to proposition 10, 11, and 12 problem $(\mathcal{P}_1)_c$ is transformed into the iterative convex problem $(\mathcal{P}_2)_c$ as

$$(\mathcal{P}_2)_c: \min_{\boldsymbol{\omega}_c,\boldsymbol{P}_c,\bar{\boldsymbol{q}}_c,\boldsymbol{r}_c,\boldsymbol{\eta}_c,\boldsymbol{\zeta}_c} f_{\mathsf{obj}}(\bar{\boldsymbol{q}}_c)$$

$$\text{s.t.} \qquad (\tilde{C}1)-(\tilde{C}3),(\tilde{C}5),(\tilde{C}6),(\tilde{C}8)-(\tilde{C}10),(C11),(C12),(\tilde{C}13)-(\tilde{C}18),$$

Using outcome of solving $(\mathcal{P}_2)_c$, $\boldsymbol{\alpha}_c$ and $\boldsymbol{\beta}_c$ are recovered as

$$\alpha_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}=1 \text{ if } p_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}\geq\epsilon;\ \alpha_{\ell,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}=0 \text{ otherwise}, \tag{4.30a}$$

$$\beta_{k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}=1 \text{ if } p_{0,k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}\geq\epsilon;\ \beta_{k}^{[v_{\mathsf{x}},n_{\mathsf{x}}]}=0 \text{ otherwise}. \tag{4.30b}$$

The BWA variable $\boldsymbol{b}_c$ is recovered by (4.22). The proposed solution *"DT-JointRA"* to solve problem $(\mathcal{P}_0)_c$ using predicted information $\{\hat{\boldsymbol{h}}_c,\hat{\boldsymbol{g}}_c,\hat{\boldsymbol{\lambda}}_c\}$ is summarized in Alg. 9.

Proposition 13. *Alg. 9 is guaranteed to converge to a local optimal solution of problem* $(\mathcal{P}_0)_c$.

Proof: Let $\boldsymbol{\Omega}=(\boldsymbol{\omega}_c,\boldsymbol{P}_c,\bar{\boldsymbol{q}}_c,\boldsymbol{r}_c,\boldsymbol{\eta}_c,\boldsymbol{\zeta}_c)$, $\boldsymbol{\Omega}_c^{(i)}$, $\boldsymbol{\Omega}_c^{\star(i)}$, $\mathcal{F}_c^{(i)}$, and $F_c^{(i)}$ be the variable tuple, feasible point, optimal solution, feasible set, and optimal objective value of $(\mathcal{P}_2)_c$ at point $\boldsymbol{\Omega}_c^{(i)}$ at iteration i. Since $(\mathcal{P}_2)_c$ is derived by using the SCA technique [103], the following properties hold. $\boldsymbol{\Omega}_c^{\star(i)}\in\mathcal{F}^{(i)}\cap\mathcal{F}^{(i+1)}$, leading to $F_c^{(i+1)}\leq F_c^{(i)}$. Hence, by iteratively solve $(\mathcal{P}_2)_c$ in Alg. 9, the generated objective value sequence is non-increasing. Additionally, since the transmit power is limited by $(C11),(C12)$, convergence of Alg. 9 is ensured. Subsequently, with recovered binary variable $\{\boldsymbol{b}_c,\boldsymbol{\alpha}_c,\boldsymbol{\beta}_c\}$, the feasible set of $(\mathcal{P}_2)_c$ is a subset of that of $(\mathcal{P}_0)_c$. Hence, Alg. 9 converges to a local optimal solution of $(\mathcal{P}_0)_c$. ■

4.4.3 RT-Refine Solution

Problems $(\mathcal{P}_0^{\mathsf{refi}})_s$ and $(\mathcal{P}_0)_c$ share a similar structure with certain omitted constraints and variables in $(\mathcal{P}_0^{\mathsf{refi}})_s$. Hence, by using approximation steps for $(\mathcal{P}_0)_c$, $(\mathcal{P}_0^{\mathsf{refi}})_s$ is transformed into iterative convex problem $(\mathcal{P}_1^{\mathsf{refi}})_s$ as

$$\begin{aligned}(\mathcal{P}_1^{\mathsf{refi}})_s:\ \min_{\boldsymbol{p}[s],\boldsymbol{r}[s],\boldsymbol{\eta}[s],\bar{\boldsymbol{q}}[s]}\ & f_{\mathsf{obj}}(\bar{\boldsymbol{q}}[s])\\ \text{s.t.}\quad & (\tilde{C}5)_s,(\tilde{C}6)_s,(\tilde{C}9)_s,(\tilde{C}10)_s,(C11)_s,\\ & (\tilde{C}14)_s,(\tilde{C}15)_s,(\tilde{C}16)_s,(\tilde{C}17)'_s,(\tilde{C}18a)'_s,(\tilde{C}18b)'_s,\end{aligned}$$

where $(\tilde{C}17)'$, $(\tilde{C}18a)'$, and $(\tilde{C}18b)'$ are the convex constraint revised from $(\tilde{C}17)$, $(\tilde{C}18a)$, and $(\tilde{C}18b)$ using terms $\hat{\Theta}_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ and $\hat{\Theta}_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ instead of $\Theta_k^{\mathsf{TN},[v_{\mathsf{M}},n_{\mathsf{M}}]}$ and $\Theta_k^{\mathsf{Sat},[v_{\mathsf{M}},n_{\mathsf{M}}]}$, respectively. It is worth noting that information $\{\boldsymbol{h}_e^*,\hat{\boldsymbol{g}}_e,\boldsymbol{\lambda}_e\}$ is used to solve $(\mathcal{P}_1^{\mathsf{refi}})_s$. The overall solution of RT-Refine Algorithm, is described in Alg. 10, with the workflow illustrated

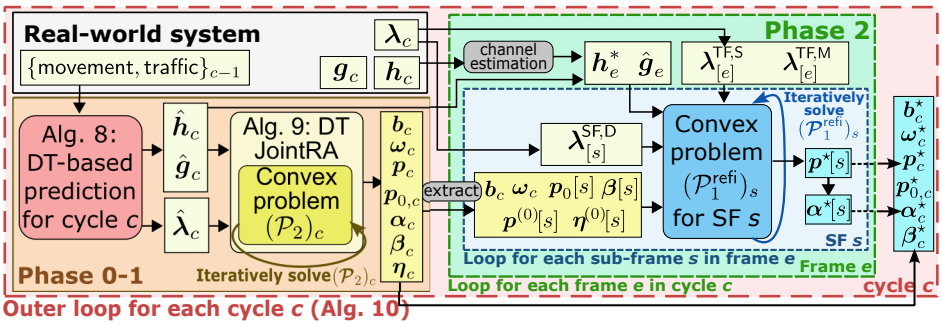


Figure 4.6: Flowchart of the proposed DT-based algorithm.

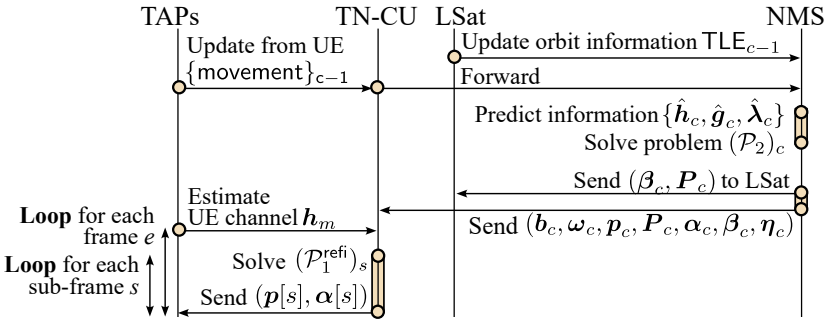


Figure 4.7: Data exchange in DT-based algorithm.

in Fig. 4.6.

Proposition 14. *Alg. 10 converges to a local optimal solution of problem sequence* $\{(\mathcal{P}_0^{\text{refi}})_s\}_{\forall s}$.

Proof: In each loop for cycle c, Alg. 10 solves $(\mathcal{P}_2)_c$ and $(\mathcal{P}_1^{\text{refi}})_s$ in phase 1 and 2, respectively. While convergence properties of solving $(\mathcal{P}_2)_c$ is guarantee in proposition 13, similarly, solving $(\mathcal{P}_1^{\text{refi}})_s$ in steps $10-12$ is guaranteed to converge since $(\mathcal{P}_1^{\text{refi}})_s$ is a SCA-based problem [103]. Besides, for each loop of SF s in steps $8-16$, the feasible set of $(\mathcal{P}_1^{\text{refi}})_s$ is a subset of that of corresponding $(\mathcal{P}_0^{\text{refi}})_s$. Hence, Alg. 10 converges to a local optimal solution of $\{(\mathcal{P}_0^{\text{refi}})_s\}_{\forall s}$. ■

4.4.4 RT-Refine Algorithm Implementation

Regarding implementation, Fig. 4.7 shows execution steps and data exchange among TAPs, TCU, LSat, and NMS in the DT-based algorithm. The procedure is summarized as

- At each cycle, TAPs and LSat send the updated UE movement and orbit information to NMS. Based on which, the necessary information is predicted and Alg. 9 is executed to solve $(\mathcal{P}_2)_c$ for upcoming cycle c.

- Once $(\mathcal{P}_2)_c$ is solved, all outcomes are sent to TCU for the next phase, while only power

Algorithm 9 DT-based Joint-RA Algorithm (DT-JointRA)

Input: Predicted $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c, \hat{\boldsymbol{\lambda}}_c\}$ from DT.
Set $i = 1$ and generate an initial point $(\boldsymbol{P}_c^{(0)}, \boldsymbol{P}_c^{(0)}, \boldsymbol{\eta}_c^{(0)}, \boldsymbol{\zeta}_c^{(0)})$.
repeat
    Solve problem $(\mathcal{P}_2)_c$ to obtain $(\boldsymbol{P}_c^{\star}, \boldsymbol{\eta}_c^{\star}, \boldsymbol{\zeta}_c^{\star})$.
    Update $(\boldsymbol{P}_c^{(i)}, \boldsymbol{\eta}_c^{(i)}, \boldsymbol{\zeta}_c^{(i)}) = (\boldsymbol{P}_c^{\star}, \boldsymbol{\eta}_c^{\star}, \boldsymbol{\zeta}_c^{\star})$ and $i := i + 1$.
until Convergence
Recovery $\boldsymbol{\omega}_c^{\star}$ and binary solutions $\boldsymbol{b}_c^{\star}$, $\boldsymbol{\alpha}_c^{\star}$, and $\boldsymbol{\beta}_c^{\star}$ by (4.25), (4.22), and (4.31).
Output: Solution $\{\boldsymbol{b}_c^{\star}, \boldsymbol{\omega}_c^{\star}, \boldsymbol{P}_c^{\star}, \boldsymbol{P}_c^{\star}, \boldsymbol{\alpha}_c^{\star}, \boldsymbol{\beta}_c^{\star}, \boldsymbol{\eta}_c^{\star}\}$ for cycle c.

Algorithm 10 Real-time-Refine Algorithm (RT-Refine)

Initialize: Actual information $\{\boldsymbol{\lambda}_0, \boldsymbol{u}_0^{\mathsf{ue}}, \mathsf{TLE}_0\}$ in cycle 0.
for Each cycle c, $c = 1 \rightarrow N_{\mathsf{Cy}}$ **do**
    Phase 0: Predict $\{\hat{\boldsymbol{h}}_c, \hat{\boldsymbol{g}}_c, \hat{\boldsymbol{\lambda}}_c\}$ by Alg. 8.
    Phase 1: Execute Alg. 9 for solution $\{\boldsymbol{b}_c^{\star}, \boldsymbol{\omega}_c^{\star}, \boldsymbol{P}_c^{\star}, \boldsymbol{\alpha}_c^{\star}, \boldsymbol{\beta}_c^{\star}, \boldsymbol{\eta}_c^{\star}\}$ for cycle c..
    Phase 2: Re-optimize using $\boldsymbol{h}_e^*$,$\hat{\boldsymbol{g}}_e$,$\boldsymbol{\lambda}_e$ and initial point from output in **step 4**.
    for Each frame e in cycle c **do**
        Collect $\{\boldsymbol{\lambda}_e^{\mathsf{TF},\mathsf{x}}\}, \mathsf{x} \in \{\mathsf{M}, \mathsf{S}\}$. Estimate channel of TN UEs and build $\boldsymbol{h}_e^*$.
        for Each SF s in frame e **do**
            Collect $\boldsymbol{\lambda}_s^{\mathsf{SF},\mathsf{D}}$, set i=0 and extract initial point from solution in **step 4**.
            repeat
                Solve problem $(\mathcal{P}_1^{\mathsf{refi}})_s$ to obtain $(\boldsymbol{p}^{\star}[s], \boldsymbol{\eta}^{\star}[s])$.
                Update $(\boldsymbol{p}^{(i)}[s], \boldsymbol{\eta}^{(i)}[s]) = (\boldsymbol{p}^{\star}[s], \boldsymbol{\eta}^{\star}[s])$ and $i := i + 1$.
            until Convergence
            Recover binary variable $\boldsymbol{\alpha}^{\star}[s]$ by (4.31).
            Output SF s**:** The adjusted solution $(\boldsymbol{p}^{\star}[s], \boldsymbol{\alpha}^{\star}[s])$.
        end for
    end for
    Output cycle c**:** Solution $\{\boldsymbol{b}_c^{\star}, \boldsymbol{\omega}_c^{\star}, \boldsymbol{P}_c^{\star}, \boldsymbol{P}_c^{\star}, \boldsymbol{\alpha}_c^{\star}, \boldsymbol{\beta}_c^{\star}\}$ with adjusted $(\boldsymbol{P}_c^{\star}, \boldsymbol{\alpha}_c^{\star})$.
end for
Output: Solution for N_{cy} cycles.

control and UE/RB assignment solutions for LSat-UE links are sent to LSat.

- For the given solution sent from NMS, TCU re-optimizes TN short-term decision each SF s. Subsequently, solutions are sent to TAPs.

One can see that, at each cycle, there is only one round-trip communication between LSat and ground segments, i.e., LSat updates orbit information and then NMS sends transmission solutions. Additionally, depending on the stability and accuracy of orbit information (TLE data), LSat can update less frequently, which further reduces costly signaling. The complexity aspect is further discussed in the following section.

4.5 Benchmarks and Complexity Analysis

4.5.1 Greedy Algorithm

The **Greedy Algorithm** is summarized in Alg. 11. The number of used DS SCs is fixed to $N_{\mathsf{SC}}^{\mathsf{D}}$ while BW for MS and SS services is uniformly allocated. The AP/LSat-UE association and

Algorithm 11 GREEDY ALGORITHM

Input: Channel $\boldsymbol{h}_c = [h^{\mathsf{x}}(\ell, k, \nu_{\mathsf{x}}, n_{\mathsf{x}})]$, $\boldsymbol{g}_c = [g^{\mathsf{x}}(k, \nu_{\mathsf{x}}, n_{\mathsf{x}})]$, $\forall(\ell, k, \nu_{\mathsf{x}}, n_{\mathsf{x}}, \mathsf{x}, c)$.
Initialize: Zero matrices $\boldsymbol{\alpha}_c$, $\boldsymbol{\beta}_c$, and $\boldsymbol{b}_c$ with $\forall c$.
for Each cycle c **do**
  Use $N_{\mathsf{sb}}^{\mathsf{d}}$ DS SCs, uniformly allocate BW for services $\mathsf{x} \in \{\mathsf{m}, \mathsf{s}\}$. Build $\boldsymbol{b}_c$.
  for Each frame e **do**
    Extract $\{\boldsymbol{h}_e, \boldsymbol{g}_e, \boldsymbol{\alpha}_e, \boldsymbol{\beta}_e\}$ for frame e. Set elements of non-using SCs to zero.
    for Each UE k service x **do**
      while $\boldsymbol{h}_e^{\mathsf{x}}(:, k, :, :) \neq \mathbf{0}$ **do**
        Find index $(\hat{\ell}, \hat{\nu}, \hat{n})$ where $\boldsymbol{h}_e^{\mathsf{x}}(\hat{\ell}, k, \hat{\nu}, \hat{n}) = \mathsf{max}(\boldsymbol{h}_e^{\mathsf{x}}(:, k, :, :))$.
        Set $\boldsymbol{\alpha}_e^{\mathsf{x}}(\hat{\ell}, k, \hat{\nu}, \hat{n}) = 1$, $\boldsymbol{h}_e^{\mathsf{x}}(:, k, \hat{\nu}, \hat{n}) = 0$ and $\boldsymbol{h}_e^{\mathsf{x}}(\hat{\ell}, :, \hat{\nu}, \hat{n}) = 0$.
      end while
      while $\boldsymbol{g}_e^{\mathsf{x}}(k, :, :) \neq \mathbf{0}$ **do**
        Find index $(\hat{\nu}, \hat{n})$ satisfying $\boldsymbol{g}_e^{\mathsf{x}}(k, \hat{\nu}, \hat{n}) = \mathsf{max}(\boldsymbol{g}_e^{\mathsf{x}}(k, :, :))$.
        Set $\boldsymbol{\beta}_e^{\mathsf{x}}(k, \hat{\nu}, \hat{n}) = 1$ and $\boldsymbol{g}_e^{\mathsf{x}}(:, \hat{\nu}, \hat{n}) = 0$.
      end while
    end for
  end for
  Assign $\boldsymbol{\alpha}_e, \boldsymbol{\beta}_e$ to $\boldsymbol{\alpha}_c, \boldsymbol{\beta}_c$
  Power control $\{\boldsymbol{P}_c\}$**:** using water-filling scheme by ignoring interference.
  Traffic steering $\boldsymbol{\omega}_c$**:** set proportional to UE rate.
end for
Output: Solution $\{\boldsymbol{b}_c, \boldsymbol{\omega}_c, \boldsymbol{P}_c, \boldsymbol{\alpha}_c, \boldsymbol{\beta}_c\}$, $\forall c$ for N_{Cy} cycles.

RB assignment are selected based on channel gain matrices in cycle c. Afterward, the water-filling algorithm is utilized for power allocation by ignoring interference with ϵ_{pow} tolerance for water level searching. The traffic steering decision is set proportional to the UE rate. One notes that this algorithm does not guarantee DS SINR/rate and queue length constraints $(C10), (C14) - (C16)$.

4.5.2 Other Benchmark

A **Heuristic Algorithm** is further proposed by updating the greedy algorithm above, where the BWA is adopted from [116]. In particular, the approach is developed from Alg. 11 as: **step 1)** traffic $\boldsymbol{\lambda}_c$ is added to input; **step 4)** BW for the MS and SS services is set proportional to the remaining queue length in cycle $(c-1)$; **step 19)** after power allocation, the remaining queue length is computed based on the rate of UE and traffic.

4.5.3 Complexity Analysis

This section analyzes the complexity of proposed and benchmark algorithms in terms of big-O notation. To solve convex problems in the DT-JointRA and RT-Refine algorithms, we assume that the interior-point method is utilized. Hence, the complexity to solve one convex problem is $\mathcal{O}(a^2(a+b)b^{1/2})$ where a and b are the number of variables and inequality constraints, respectively.

DT-JointRA Algorithm

Regarding Alg. 9, the DT-JointRA algorithm operates over a single cycle and consists of one iterative loop, where $(\mathcal{P}_2)_c$ is solved iteratively in **Step 3-6**. Problem $(\mathcal{P}_2)_c$ includes $a_{\mathsf{DTJRA}} = N_{\mathsf{SF}}(L(K_{\mathsf{D}}(8\bar{V}_{\mathsf{D}}+2)+K_{\mathsf{M}}(6\bar{V}_{\mathsf{M}}+2))+K_{\mathsf{M}}(6\bar{V}_{\mathsf{M}}+2)+K_{\mathsf{S}}(3\bar{V}_{\mathsf{S}}+1))+L(\bar{V}_{\mathsf{D}}+\bar{V}_{\mathsf{M}})$ variables and $b_{\mathsf{DTJRA}} = N_{\mathsf{SF}}(L(K_{\mathsf{D}}(8\bar{V}_{\mathsf{D}}+3)+K_{\mathsf{M}}(\bar{V}_{\mathsf{M}}+1)+2(2\bar{V}_{\mathsf{D}}+\bar{V}_{\mathsf{M}}+3))+2(2K_{\mathsf{D}}\bar{V}_{\mathsf{D}}+K_{\mathsf{M}}(3\bar{V}_{\mathsf{M}}+2))+2\bar{V}_{\mathsf{M}}+K_{\mathsf{S}}(\bar{V}_{\mathsf{S}}+2)+\bar{V}_{\mathsf{S}}+5)+\bar{V}_{\mathsf{D}}+\bar{V}_{\mathsf{M}}+1$ inequality constraints. The DT-JointRA algorithm complexity over N_{Cy} cycles is expressed as

$$X_{\mathsf{DTJRA}} = O(N_{\mathsf{Cy}}N_{\mathsf{iter}}a^2_{\mathsf{DTJRA}}(a_{\mathsf{DTJRA}}+b_{\mathsf{DTJRA}})b^{1/2}_{\mathsf{DTJRA}}), \tag{4.31}$$

where N_{iter} is the number of iterations for convergence.

RT-Refine Algorithm

As described in Alg. 10, the RT-Refine algorithm consists of four nested loops: 1) loop for N_{cy} cycles, which involves solving $(\mathcal{P}_2)_c$ in **step 4**, 2) loop for N_{TF} frames in each cycle, 3) loop for 10 SFs in each frame, 4) iteratively solving problem $(\mathcal{P}_1^{\mathsf{refi}})_s$. Since $(\mathcal{P}_1^{\mathsf{refi}})_s$ consists of $a_{\mathsf{RTRefi}} = L(K_{\mathsf{D}}(8\bar{V}_{\mathsf{D}}+2)+K_{\mathsf{M}}(6\bar{V}_{\mathsf{M}}+2))+K_{\mathsf{M}}(4\bar{V}_{\mathsf{M}}+2)$ variables and $b_{\mathsf{RTRefi}} = L(K_{\mathsf{D}}(8\bar{V}_{\mathsf{D}}+3)+K_{\mathsf{M}}(\bar{V}_{\mathsf{M}}+1)+2(2\bar{V}_{\mathsf{D}}+\bar{V}_{\mathsf{M}}+3))+2(2K_{\mathsf{D}}\bar{V}_{\mathsf{D}}+K_{\mathsf{M}}(3\bar{V}_{\mathsf{M}}+2)+1)$ inequality constraints, and the complexity for solving $(\mathcal{P}_2)_c$ in N_{Cy} cycles is given in (4.31), the complexity of the RT-Refine algorithm is given as

$$X_{\mathsf{RTRefi}} = X_{\mathsf{DTJRA}} + O(N_{\mathsf{Cy}}N_{\mathsf{SF}}\bar{N}_{\mathsf{iter}}a^2_{\mathsf{RTRefi}}(a_{\mathsf{RTRefi}}+b_{\mathsf{RTRefi}})b^{1/2}_{\mathsf{RTRefi}}), \tag{4.32}$$

where $\bar{N}_{\mathsf{iter}}$ is the number of required iterations for convergence in solving $(\mathcal{P}_1^{\mathsf{refi}})_s$. In (4.32), one notes that the first term is due to execution at the CNC with predicted information while the second term corresponds to execution at TCU with real-time feedback from actual systems at each SF. (4.32) shows that the complexity for executing each SF at TCU is significantly lower than that for one cycle at CNC, i.e., about $(N_{\mathsf{TF}})^{3.5}$ times in terms of big-O notation. This further emphasizes the practice of the proposed solution.

Table 4.1: Queue length and DS traffic remaining outcomes of the Heuristic algorithm versus $N_{\mathsf{SC}}^{\mathsf{D}}$.

Num. SCs for DS services, $N_{\mathsf{SC}}^{\mathsf{D}}$	1	2	3	4	5
Mean queue length (MB)	20.12	23.12	27.08	31.55	37.29
Remaining DS traffic (%)	22.2	19.5	18.4	17.8	17.4

4.5.4 Greedy and Heuristic Algorithms

The greedy and heuristic algorithms share a similar structure with minor differences in **Step 4** and **19**. Hence, these algorithms have the same complexity in terms of big-O notation. Particularly, Alg. 11 consists of nested loops. For each cycle loop in **Steps 3-21**, there are two outer-nested-loops in **Step 5-17** for each frame and UE, and two inner-loops for AP-UE and LSat-UE associations in **Step 8-11** and **Step 12-15**, respectively. **Step 5-17** for UE association demand a complexity of $\mathcal{O}(X_{\mathsf{Assoc}})$ with $X_{\mathsf{Assoc}} = \mathcal{O}(N_{\mathsf{TF}}(40LK_{\mathsf{D}}\bar{V}_{\mathsf{D}} + 20(L+1)K_{\mathsf{M}}\bar{V}_{\mathsf{M}} + 10K_{\mathsf{S}}\bar{V}_{\mathsf{S}}))$. In **Step 19**, the water searching for water-filling algorithm for N TAPs and 1 LSat demand a complexity of $\mathcal{O}(-\log_2(\epsilon_{\mathsf{pow}})(N+1))$. Hence, the complexity of Alg. 11 is expressed as

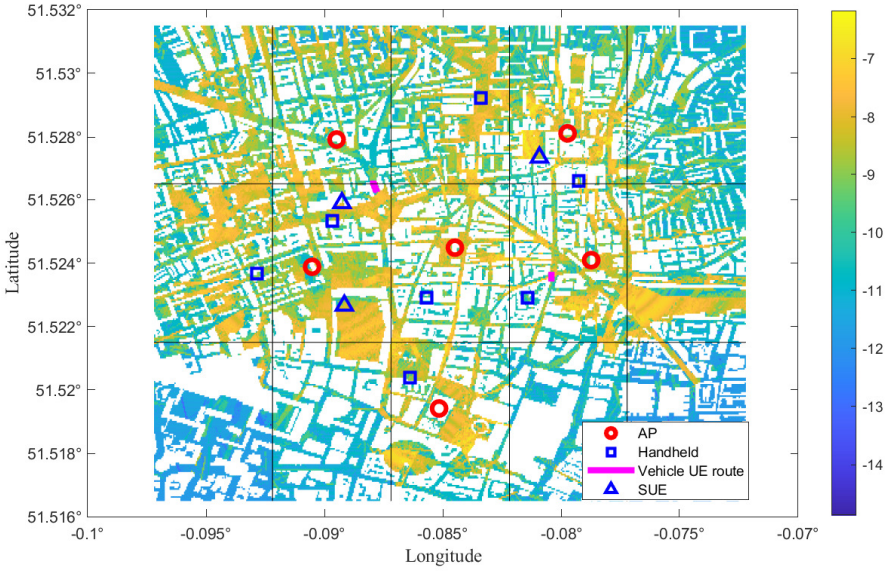
$$\mathsf{X}_{\mathsf{Grd-Heu}} = \mathcal{O}(N_{\mathsf{Cy}}(X_{\mathsf{Assoc}} - N\log_2(\epsilon_{\mathsf{pow}}))). \tag{4.33}$$

4.6 Numerical Results

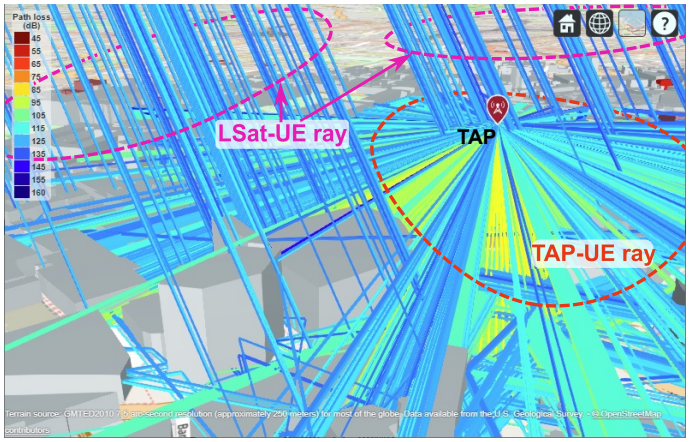
4.6.1 Simulation Setup

The simulation is conducted with 3D map in an area of 3km^2 at (51.524°N, 0.085°W) in London city. For simulation data, the vehicular UE route and LSat orbit are taken from Google Navigator and Starlink LEO TLE, the arrival traffic is taken from an actual dataset [122]. Environment channels are generated using the DT model with coefficient $\xi = 0.5$.

The key simulation parameters are: system BW $W^{\mathsf{tot}} = 15$ MHz, operation frequency $f_c = 3.4$ GHz, LSat altitude 500 km, antenna parameters of AP, LSat, SUE as in [8, 9], vehicle UE antenna and RayT parameters as in [31,52], number of cycles $N_{\mathsf{Cy}} = 20$, number of frames/cycle $N_{\mathsf{TF}} = 5$, numbers of TAPs, UEs $(N, K_{\mathsf{D}}, K_{\mathsf{M}}, K_{\mathsf{S}}) = (6, 4, 5, 3)$, power budget at AP and LSat $(p_{\mathsf{AP}}^{\mathsf{max}}, p_{\mathsf{Lsat}}^{\mathsf{max}}) = (34, 36)$ dBm, maximum queue length $q_{\ell}^{\mathsf{m,max}} = q_0^{\mathsf{x,max}} = 2$ MB, interference margin for the RT-Refine algorithm $\kappa = 1.1$. For intuition, the simulation scenario with UE position and channel gain heatmap, and RayT result are shown in Fig. 4.8. For benchmarks, a **full information algorithm (FIA)** is further examined where DT-JointRA algorithm is executed with full actual information. In addition, for heuristic algorithms, mean



(a) Simulation scenario and AP channel gain heatmap (log10 scale).



(b) Ray-tracing result in the examined area.

Figure 4.8: Simulation scenario, TN channel heatmap (log10 scale), and RayT.

queue length and remaining DS traffic versus $N_{\mathsf{SC}}^{\mathsf{D}}$ are shown in Table. 4.1. Since the remaining DS traffic decreases slowly while the mean queue length quickly increases as $N_{\mathsf{SC}}^{\mathsf{D}}$ increases. Based on which, $N_{\mathsf{SC}}^{\mathsf{D}} = 1$ is selected.

Fig. 4.9 shows the convergence rate of the proposed algorithms in different cases. It is worth noting that the DT-JointRA algorithm is encompassed in the RT-Refine one in phase 1 at step 4. Since the DT-JointRA algorithm and FIA share the same structure with only differences in inputs, their convergence rates are similar. Particularly, the DT-JointRA algorithm and FIA require only about 25 iterations for convergence in both cases $p_{\mathsf{AP}}^{\mathsf{max}} = 30$

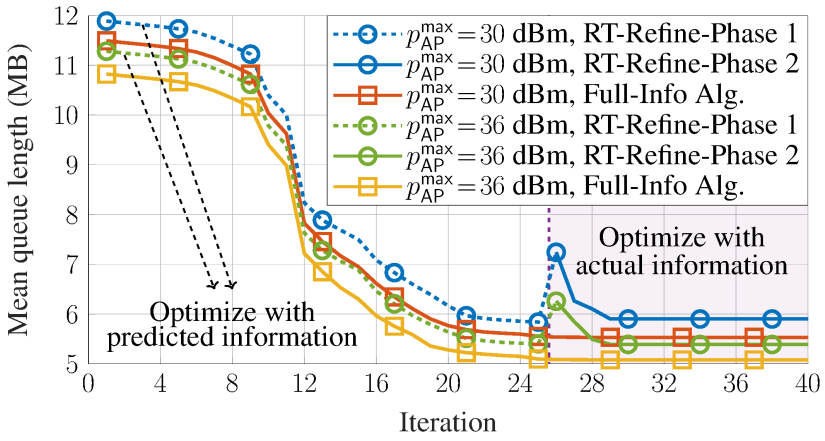


Figure 4.9: Convergence rate of the proposed algorithms.

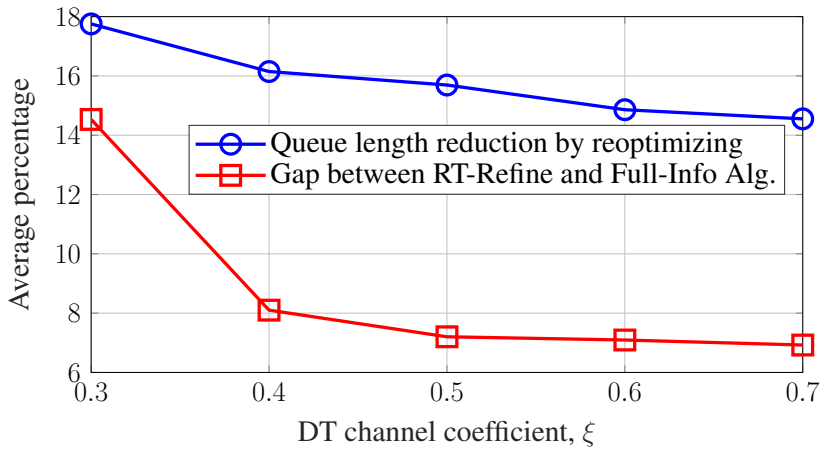


Figure 4.10: Queue length versus DT channel coefficient, ξ.

dBm and $p_{\mathsf{AP}}^{\max} = 36$ dBm. Although in the RT-Refine algorithm, phase 2 is executed right after phase 1 (i.e., DT-JointRA), it is bumpy at the beginning of phase 2. This phenomenon is due to the changes in inputs, i.e., replacing $\{\hat{\boldsymbol{h}}_t, \hat{\boldsymbol{\lambda}}_t\}$ by $\{\boldsymbol{h}_t^*, \boldsymbol{\lambda}_t\}$. However, phase 2 requires only about 3 iterations for convergence. This emphasizes the practical implementation aspect of the proposed RT-Refine algorithm.

Regarding the impact of DT channel coefficient ξ on the performance of the proposed RT-Refine algorithm, Fig. 4.10 illustrates the queue length improvement of re-optimization in phase 2 of the RT-Refine algorithm and the queue length gap between the RT-Refine algorithm and FIA. Based on the channel model in Section 4.2.3, one can see that a larger ξ results in a smaller difference between the emulated real environment and its DT reduces. Hence, both lines decrease as ξ increases. Particularly, at $\xi = 0.3$, $\xi = 0.5$, and $\xi = 0.7$, re-optimizing in phase 2 can reduce the queue length by about 17.7%, 15.7%, and 14.5% compared to DT-JointRA outcomes, and the percentage gaps between the RT-Refine algorithm and FIA are

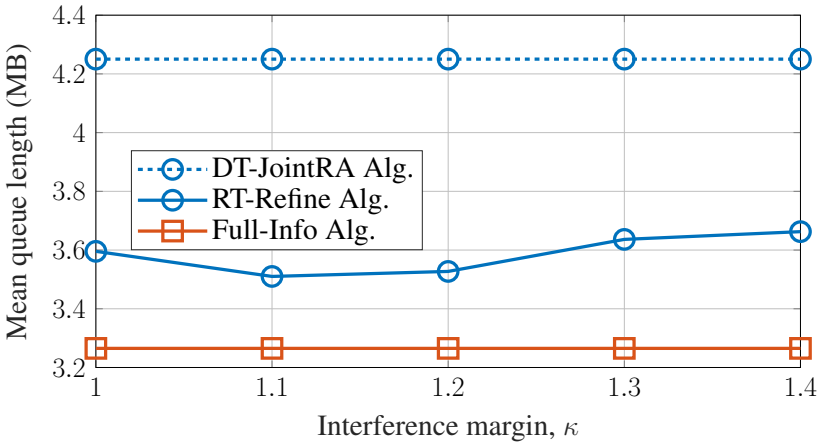


Figure 4.11: Queue length versus margin interference.

about 14.5%, 7.3%, and 6.9% compared to FIA outcomes, respectively.

Fig. 4.11 depicts the mean queue length versus interference margin κ in the RT-Refine algorithm. One can see that the mean queue length outcome first decreases and then increases as κ increases. Particularly, the mean queue length at $\kappa = (1, 1.1, 1.2, 1.3)$ is about $(3.59, 3.51, 3.53, 3.63)$ MB. This phenomenon can be explained as follows. First, one notes that in phase 2 of the RT-Refine algorithm, only actual channels from TAPs to their own served UEs are used while other channels are kept as predicted ones. For $\kappa = 1$, the prediction error of interference channels is not aware. In contrast, for $\kappa > 1$, this error is taken into account and TAPs reduce their transmit power to protect UEs served by LSat. Hence, at higher κ, e.g., at $\kappa = 1.1$ and $\kappa = 1.2$, mean queue length decreases. However, for sufficiently high κ, the prediction error is overestimated, leading the TAPs to reduce transmit power more than necessary, which results in rate degradation and an increase in queue length. By appropriately choosing this margin, i.e., $\kappa = 1.1$, re-optimization in phase 2 can reduce about 0.75 MB in mean queue length, and the gap between the RT-Refine and FIA lines is only about 0.25 MB. This result demonstrates the effectiveness of considering this margin against interference channel prediction errors while avoiding costly channel estimation. Based on this result, margin $\kappa = 1.1$ is selected for other simulations.

Fig. 4.12 shows the mean queue length outcomes of proposed algorithms and benchmarks versus the AP power budget. Regarding the greedy and heuristic benchmarks, the mean queue length gap between two lines is about 1.12 MB. This improvement gap in the heuristic algorithm is due to the adaptive BWA based on the remaining queue length. However, compared to the proposed algorithms, since benchmarks ignore interference, the queue length

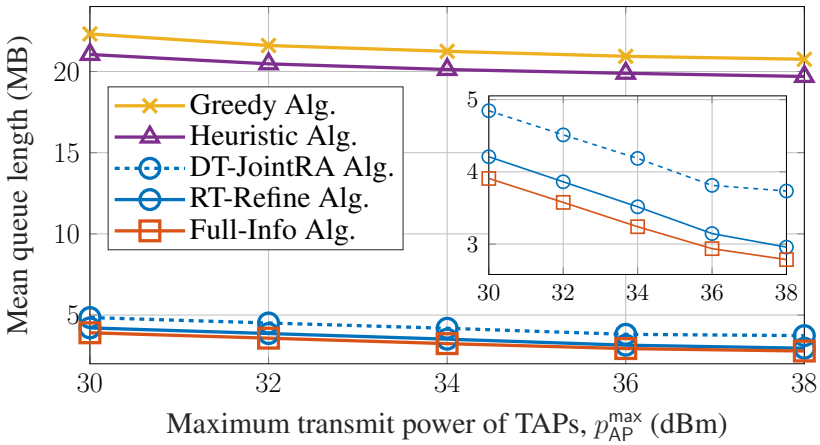


Figure 4.12: Queue length versus AP power budget.

performance gap between the proposed and benchmark algorithms is significant, i.e., about 17 MB. Additionally, the proposed algorithms can satisfy DS requirements in all considered cases while unserved DS traffic is about 22.2% in benchmarks as shown in Table. 4.1. Regarding proposed algorithms, the gap between the DT-JointRA and RT-Refine lines, which indicates improvement by re-optimizing in RT-Refine-phase 2, is about 0.7 MB. Beside, compared to the proposed benchmark FIA which uses full actual information for optimization, the gap between the RT-Refine and FIA lines is only about 0.27 MB. Hence, in addition to superiority compared to two benchmarks, the proposed RT-Refine scheme shows the effectiveness of re-optimizing AP power control to adjust solutions with the actual information at all considered cases of $p_{\mathsf{AP}}^{\max}$. Additionally, a similar outcome trend of the proposed algorithms can be recognized. Particularly, with AP power budget $p_{\mathsf{AP}}^{\max} = 30 \rightarrow 36$ dBm, the mean queue length decreases quickly, and with $p_{\mathsf{AP}}^{\max} = 36 \rightarrow 38$ dBm, their decrease trend slightly reduces. This phenomenon indicates that with $p_{\mathsf{AP}}^{\max} = 36$ or above, the AP power budget starts to be redundant for queue length minimization.

Fig. 4.13 depicts the mean queue length performance of the examined schemes with different LSat power budgets. Regarding the greedy and heuristic benchmarks, one can see that the mean queue length outcomes first decrease and then increase. This is due to the impact of interference caused by LSat in the whole coverage area is not addressed in the power control phase. Similar to results in Fig.4.12, the adaptive BWA in the heuristic scheme brings the improvement in terms of queue length compared to the greedy one. However, compared to the proposed algorithms, the queue length outcome gap is still significant. Regarding the proposed algorithms, a similar trend compared to results in Fig 4.12 can be recognized. Particularly, the

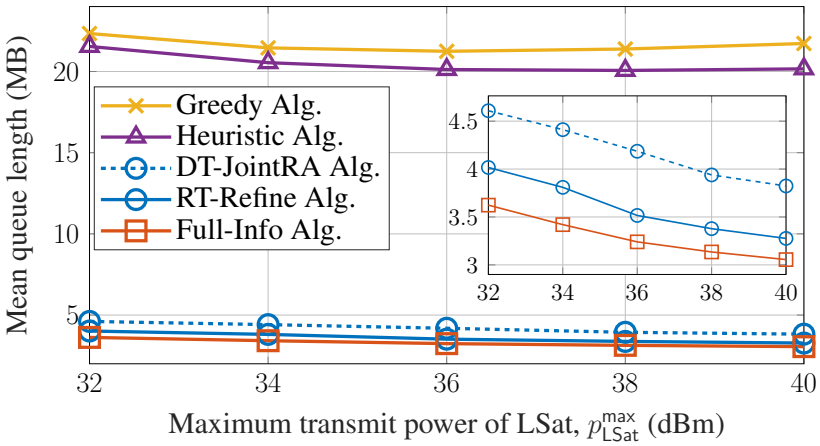


Figure 4.13: Queue length versus LSat power budget.

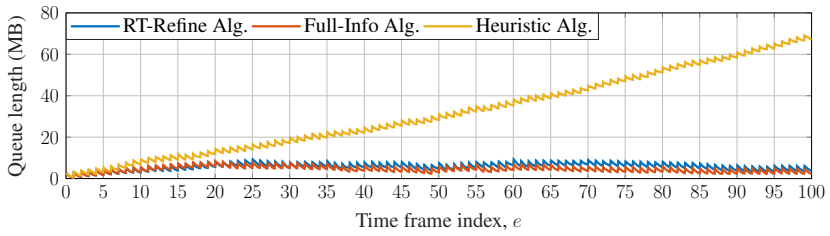
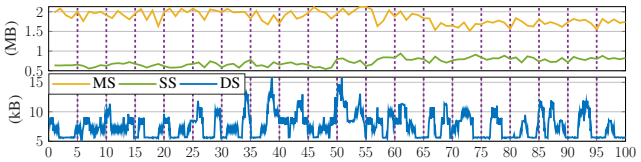


Figure 4.14: Queue length over time of examined algorithms.

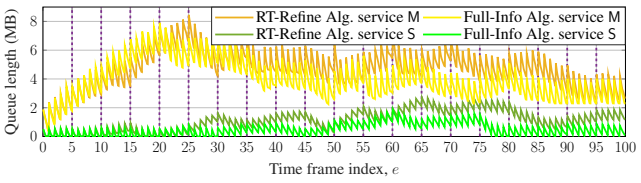
queue length outcomes linearly decrease as LSat power budget $p_{\mathsf{LSat}}^{\max}$ increases, while the decreasing trend slightly reduces with $p_{\mathsf{LSat}}^{\max} = 36 \rightarrow 40$ dBm. This slight degradation indicates that the LSat power budget starts to be redundant for queue length minimization.

For an insight into the queue length minimization performance, Fig. 4.14 shows the queue length of FIA and heuristic benchmarks and the RT-Refine algorithm over time (TFs). One can see that the queue length outcome of the heuristic benchmark increases linearly over frames while that of proposed algorithms remains below about 10 MB over frames. This result demonstrates the superiority of the proposed solutions compared to benchmarks in terms of queue length stability.

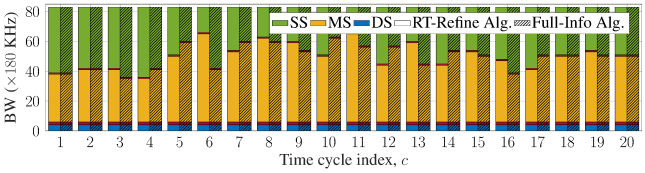
For more insight into the queue length over time, Fig. 4.15 shows the arrival data of services, queue length and BWA outcomes over time of the RT-Refine algorithm and FIA. Although the arrival data of DS services is quite small, i.e., only about $5-10$ kB/frame, it is not completely served by using benchmarks with multiple SCs for DS services as Table 4.1. However, the proposed algorithms can satisfy DS requirements by allocating only 1 SC for BWP D. Besides, due to the larger varying total arrival data size of MS and SS services, the BWAs for these services change over cycles for queue length minimization. One can see that, at each cycle, the BW is allocated adaptively based on the remaining queue length in



(a) Arrival data of services.



(b) Queue length over time.



(c) BW allocation over time.

Figure 4.15: Arrival data, queue length, and BWA over time.

the previous cycle and the arrival data. Additionally, compared with the FIA benchmark using full actual information, the RT-Refine algorithm, which is proposed for implementation purposes, can provide a comparable performance in terms of queue length minimization over time.

4.7 Summary

This work studied the novel DT-aided joint spectrum sharing, traffic steering, and resource allocation for ISTNs over time. The optimization problems aim to minimize the system congestion under system and service constraints. To assist RRM, a digital-twin system is developed for information prediction. Based on this, we proposed a DT-JointRA algorithm implementable at the CNC to address the optimization problem. For practical implementation, a re-optimization phase is built on top of the DT-JointRA algorithm to refine its solution with the information in real-world systems. Through numerical results with the actual 3D map and traffic information, the proposed algorithms show the superiority in terms of congestion minimization, compared to benchmarks. Furthermore, the simulation results emphasize the practicality and effectiveness of the proposed algorithms in terms of adaptation ability.

Chapter 5

Conclusions and Future Work

THIS thesis addressed several spectrum management and resource allocation (RA) challenges in integrated satellite-terrestrial networks (ISTNs). In summary, three key problems were identified, each representing a critical issue in the deployment and operation of ISTNs. These problems were carefully analyzed and addressed in the subsequent chapters. The main conclusions drawn from the contributions of this thesis, along with directions for future work, are outlined below.

5.1 Main Conclusions

Motivated by practical ISTN scenarios and spectrum regulations promoted by standardization bodies such as 3GPP, ESA, and the U.S. FCC as well as the associated technical challenges, this thesis has identified and comprehensively investigated three critical problems in Integrated Satellite-Terrestrial Networks (ISTNs). Particularly, a comprehensive overview of ISTNs, emphasizing how emerging service and increasing data demands have driven the integration of terrestrial networks (TNs) and satellite networks (SatNets), has been presented in Chapter 1. Building upon this background with associated technique issues and a review of existing literature, three key research problems were formulated:

- Problem 1: Backhaul-Link BW Allocation and Two-Tier UE Association.
- Problem 2: Throughput Enhancement and Seamless Handover in C-Band 5G-ISTNs.
- Problem 3: Resource Management in Dynamic Spectrum Sharing ISTNs.

These problems were systematically addressed in Chapters 2, 3, and 4, respectively.

Problem 1 was investigated in Chapter 2 through a novel time-window-based joint two-tier association and RA framework for ISTNs. Given the objective of transmission time minimization, two algorithms-one centralized and one decentralized-were developed to jointly optimize backhaul BW allocation, UE-BS-LSat association, RB assignment, and power control. Especially, the decentralized algorithm was designed for implementation purposes enabling parallel execution among local controllers. Extensive simulations demonstrated that:

- Both proposed algorithms significantly outperformed baseline benchmarks in minimizing transmission time, achieving a reduction of approximately $30-50\%$ compared to the benchmark schemes.
- The solutions provided enhanced load balancing, spectral efficiency, and adaptability.
- The decentralized algorithm showed strong potential for practical deployment in distributed network scenarios.

Problem 2 was addressed in Chapter 3 by developing a joint user association and power allocation optimization framework for co-primary spectrum sharing ISTNs in urban environments. Aiming to simultaneously maximize system sum-rate (SR) and minimize connection changes (CCs), algorithms were proposed: FTW-Alg and PTW-Alg, which solve the problem in the full time-window and subsequent sub-time-windows. Especially, PTW-Alg was designed for implementation purposes with incorporation of prediction mechanisms. Evaluation and analysis using a 3D map and realistic UE mobility scenarios demonstrated that:

- The proposed algorithms significantly outperformed benchmarks in both SR enhancement and CC minimization, achieving an increase of approximately $20-50\%$ on SR and reduction of about $65-90\%$ on CCs.
- The solution demonstrated the feasibility of co-primary spectrum usage between TNs and SatNets through effective interference mitigation mechanisms.
- The prediction-based time-window (PTW) algorithm showed strong practicality, requiring only a single LSat position update per sub-time-window, thereby minimizing information exchange between LEO satellites and ground segments.
- The framework leveraging Ray tracing, 3D map data, and mobility modeling presented a valuable foundation for future ISTN research.

Problem 3 was tackled in Chapter 4 by developing a digital-twin (DT) model and a DT-aided framework for joint spectrum sharing, traffic steering, and resource allocation in ISTNs. Given the objective of minimizing system congestion, two tailored algorithms were designed:

- **DT-JointRA Alg:** a resource management algorithm leveraging DT-predicted system information for joint optimization of cycle time level.
- **RT-Refine Alg:** a refinement algorithm that updates DT-JointRA decisions using real-time system feedback to enhance accuracy at sub-frame time level.

Extensive numerical evaluations, including simulations with real 3D urban maps and traffic data, demonstrated that:

- The proposed algorithms significantly outperformed benchmark schemes in reducing system congestion. The DT-JointRA algorithm achieved a queue length reduction of approximately 78−82% compared to benchmarks, while the RT-Refine algorithm further improved performance by an additional 16%.
- Two algorithms exhibited the practicality and effectiveness in terms of adaptation ability with varying traffic demand.
- The DT model effectively supported resource management by reducing LSat–ground segment information exchange and addressing challenges in LSat–UE CSI estimation.
- The developed DT model–incorporating ray-tracing, 3D map data, user mobility, and traffic patterns–emulated key network components and predicted system dynamics, offering a valuable foundation for future ISTN research.

5.2 Future Work

While this thesis has addressed critical problems in ISTNs, several directions remain for future exploration. The possible extensions and open research topics can be listed in subsequent sections.

5.2.1 Possible Extensions

- **Performance loss due to connection state changes:** In the second work, the goal is to maximize sum-rate (SR) while minimizing the number of connection state changes

(CCs). However, the direct impact of CCs on SR performance has not been explicitly analyzed. According to connection state transition procedure in 3GPP 38.300 [100], frequent CCs may disrupt data transmission due to delays in releasing and establishing connections, effectively reducing transmission duration. A more comprehensive performance model incorporating SR degradation due to CCs would allow for optimizing the effective SR and improving the trade-off between connectivity stability and throughput.

- **Heterogeneity BW slicing among TN cells:** The third study assumed uniform (homogeneous) bandwidth slicing across terrestrial network (TN) cells. However, in practical large-scale deployments, the traffic pattern is spatially heterogeneous. Allowing BW slicing to vary among TN cells could further enhance spectrum efficiency. This design, however, introduces inter-numerology interference across neighboring cells [123, 124], which would require sophisticated coordination or interference mitigation strategies in future works.

- **Multi-Satellite in DSS-Enabled ISTNs:** In the third work, the model is limited to a single LEO satellite. Extending the framework to support multiple satellites operating concurrently would enable load balancing, increase SatCom capacity, and improve coverage [49, 52, 101]. Such multi-satellite scenarios would introduce new challenges in satellite coordination, resource partitioning, and inter-satellite communication.

- **DFT-based Beamforming and Precoding Technologies in SatCom Systems:** Advanced physical-layer technologies, such as DFT-based beamforming and precoding in SatCom systems, remain underexplored in this thesis. Particularly, multi-beam technology in SatCom can enhance spatial multiplexing and total system capacity. Given the limited computational resources onboard satellites, DFT-based beamforming offers a promising and lightweight solution [101, 125]. Based on which, precoding technology can further improve spectral efficiency and reduce interference. Future work should explore the integration of these approaches into SatCom system architectures, accounting for satellite-specific constraints and mobility. Additionally, the challenge of acquiring accurate channel state information (CSI), especially given long propagation delays in satellite links, must be addressed to enable practical precoding implementations.

- **Machine Learning and Deep Learning for Decision Making:** The integration of machine learning (ML) and deep learning (DL) techniques into ISTNs can enable

adaptive, data-driven decision-making across various layers of the network [116, 126]. Promising applications include traffic demand prediction, user behavior modeling, and dynamic resource allocation. In particular, for the second and third problems addressed in this thesis, time-series forecasting models can be leveraged to predict UE mobility based on historical trajectory data, thereby enhancing association and handover strategies. Moreover, in the digital-twin (DT) framework of the third work, replacing heuristic methods with ML/DL-based models for traffic demand prediction can significantly improve accuracy and responsiveness to spatiotemporal variations in traffic patterns.

5.2.2 Possible Open Research Direction

- **Open-RAN Integration in ISTNs:** Open Radio Access Network (O-RAN) architecture promotes interoperability, flexibility, and vendor diversity through the disaggregation of network functions and RAN intelligent control (RIC) loops [127]. Its application in ISTNs presents a promising direction, enabling more modular and software-defined architectures. Particularly, the design of network and resource decision making through RIC loops [116], considering latency management as well as synchronization between space and ground segments, could be a promising research direction.

- **Support for Both Unicast and Multicast Services in ISTNs:** SatCom systems are particularly well-suited for content broadcasting and multicast services due to their wide coverage [128,129]. In addition to unicast services, future work could explore ISTN scenarios that support both unicast and multicast transmissions, especially in scenarios involving high-demand media content distribution. In such scenarios, the network and resource management design, considering different buffers dedicated to service types, should be studied to effectively handle different services and quality-of-service (QoS) requirements.

- **Content caching at LEO satellites in ISTNs:** Deploying intelligent caching strategies at LEO satellites could significantly reduce feeder link load, improve latency performance, and offload traffic from TNs for popular content [130,131]. Considering content-aware caching and delivery mechanisms in network and resource management designs represents a valuable research direction. Future work can focus on designing intelligent caching mechanisms that consider dynamic networks, content popularity, and coordination with terrestrial caches for optimized content delivery.

Bibliography

[1] Ericsson, "Ericsson mobility report," 2025.

[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, March 2020.

[3] C.-X. Wang *et al.*, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys Tut.*, vol. 25, no. 2, pp. 905–974, Secondquarter 2023.

[4] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, July 2021.

[5] M. A. Jamshed, A. Kaushik, S. Manzoor, M. Z. Shakir, J. Seong, M. Toka, W. Shin, and M. Schellmann, "A tutorial on non-terrestrial networks: Towards global and ubiquitous 6g connectivity," *Foundations and Trends® in Networking*, vol. 14, no. 3, pp. 160–253, 2025. [Online]. Available: http://dx.doi.org/10.1561/1300000072

[6] M. A. Jamshed *et al.*, "Non-terrestrial networks for 6g: Integrated, intelligent and ubiquitous connectivity," 2024. [Online]. Available: https://arxiv.org/abs/2407.02184

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

[8] 3GPP, "Solutions for NR to support non-terrestrial networks," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.821, March 2023, version 16.2.0.

[9] ——, "Solutions for NR to support non-terrestrial networks (NTN): Non-terrestrial networks (NTN) related RF and co-existence aspects," Technical Specification (TS) 38.863, Dec. 2023, 18.0.0.

[10] J. Li, K. Xue, D. S. L. Wei, J. Liu, and Y. Zhang, "Energy efficiency and traffic offloading optimization in integrated satellite/terrestrial radio access networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2367–2381, April 2020.

[11] Z. Gao, A. Liu, C. Han, and X. Liang, "Max completion time optimization for internet of things in LEO satellite-terrestrial integrated networks," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9981–9994, 2021.

[12] B. Di, H. Zhang, L. Song, Y. Li, and G. Y. Li, "Ultra-dense LEO: Integrating terrestrial-satellite networks into 5G and beyond for data offloading," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 47–62, 2019.

[13] R. Deng, B. Di, S. Chen, S. Sun, and L. Song, "Ultra-dense LEO satellite offloading for terrestrial networks: How much to pay the satellite operator?" *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6240–6254, Oct. 2020.

[14] R. Deng, B. Di, H. Zhang, and L. Song, "Ultra-dense LEO satellite constellation design for global coverage in terrestrial-satellite networks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Dec. 2020, pp. 1–6.

[15] R. Deng, B. Di, H. Zhang, L. Kuang, and L. Song, "Ultra-dense LEO satellite constellations: How many LEO satellites do we need?" *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4843–4857, Aug. 2021.

[16] Z. Li, S. Han, M. Peng, C. Li, and W. Meng, "Dynamic multiple access based on RSMA and spectrum sharing for integrated satellite-terrestrial networks," *IEEE Transactions on Wireless Communications*, vol. 23, no. 6, pp. 5393–5408, 2024.

[17] J. Zhu, Y. Sun, and M. Peng, "Beam management in low earth orbit satellite communication with handover frequency control and satellite-terrestrial spectrum sharing," *IEEE Transactions on Communications*, pp. 1–1, 2024.

[18] F. Rinaldi, H.-L. Maattanen, J. Torsner, S. Pizzi, S. Andreev, A. Iera, Y. Koucheryavy, and G. Araniti, "Non-Terrestrial Networks in 5G & Beyond: A Survey," *IEEE Access*, vol. 8, pp. 165 178–165 200, 2020.

[19] 3GPP, "Study on using satellite access in 5G," 3rd Generation Partnership Project (3GPP), Technical report (TR) 22.822, June 2018, version 16.0.0.

[20] ——, "Seamless integration of satellite and/or HAPS systems into 5G and related architecture options," 3rd Generation Partnership Project (3GPP), Technical report (TR) 1.1.1, June 2020, version 19.0.0.

[21] ETSI, "Non-terrestrial networks, a native component of 6g: Vision on non-terrestrial networks in 6g system (or imt-2030)," 2024.

[22] 3GPP, "Study on complementary reuse of terrestrial spectrum in satellite deployments," Tdoc RWS-230110, June 2023, TSG RAN R19 Workshop.

[23] FCC, "Single network future: Supplemental coverage from space," 2023. [Online]. Available: https://www.fcc.gov/document/fcc-advances-supplemental-coverage-space-framework-0

[24] 3GPP, "NR; Base Station (BS) radio transmission and reception," 3rd Generation Partnership Project (3GPP), Technical specification (TS) 38.104, March 2024, version 18.5.0.

[25] ——, "NR; Part 5: Satellite access Radio Frequency (RF) and performance requirements," 3rd Generation Partnership Project (3GPP), Technical specification (TS) 38.104, June 2024, version 18.6.0.

[26] RSPG, "Opinion on 5G developments and possible implications for 6G spectrum needs and guidance on the rollout of future wireless broadband networks," 2023.

[27] FCC, "C-Band spectrum will be available for 5G services on accelerated basis," 2020. [Online]. Available: https://www.fcc.gov/document/c-band-spectrum-will-be-available-5g-services-accelerated-basis

[28] E. Lagunas, C. G. Tsinos, S. K. Sharma, and S. Chatzinotas, "5G cellular and fixed satellite service spectrum coexistence in C-band," *IEEE Access*, vol. 8, pp. 72 078–72 094, 2020.

[29] A. Al-Jumaily *et al.*, "Evaluation of 5G coexistence and interference signals in the C-band satellite earth station," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6189–6200, June 2022.

[30] H. Nguyen-Kha, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "An experimental study of C-band channel model in integrated LEO satellite and terrestrial systems," in *Proc. MeditCom 2024*, 2024.

[31] ——, "Seamless 5g automotive connectivity with integrated satellite terrestrial networks in c-band," in *2024 IEEE 100th VTC-Fall*, 2024, pp. 1–5.

[32] SES, "D2.5 report on regulatory requirements: Spectrum policies and regulatory analysis," Tech. Rep., 2024.

[33] M. V. Katwe, A. Kaushik, K. Singh, M. D. Renzo, S. Sun, D. Lee, A. G. Armada, Y. C. Eldar, O. A. Dobre, and T. S. Rappaport, "Cmwave and sub-thz: Key radio enablers and complementary spectrum for 6g," 2024. [Online]. Available: https://arxiv.org/abs/2406.18391

[34] S. Kang *et al.*, "Cellular wireless networks in the upper mid-band," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 2058–2075, 2024.

[35] M. Ghosh, "Sharing in the 12 GHz band," *IEEE Wireless Communications*, vol. 30, no. 3, pp. 10–11, 2023.

[36] 3GPP, "Study on the 7 to 24 GHz Frequency Range for NR," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.820, Sept. 2021, version 16.1.0.

[37] FCC, "A preliminary view of spectrum bands in the 7.125-24 ghz range; and a summary of spectrum sharing frameworks," 2023.

[38] ——, "Notice of proposed rulemaking, expanding flexible use in the 12.2-12.7 ghz band," 2021.

[39] ——, "Notice of inquiry and order, expanding use of the 12.7-13.25 ghz band for mobile broadband or other expanded use;," 2022.

[40] J. Du, C. Jiang, A. Benslimane, S. Guo, and Y. Ren, "Sdn-based resource allocation in edge and cloud computing systems: An evolutionary stackelberg differential game

approach," *IEEE/ACM Transactions on Networking*, vol. 30, no. 4, pp. 1613–1628, Aug. 2022.

[41] S. Chan, H. Lee, S. Kim, and D. Oh, "Intelligent low complexity resource allocation method for integrated satellite-terrestrial systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 5, pp. 1087–1091, 2022.

[42] L. Zhang, H. Zhang, C. Guo, H. Xu, L. Song, and Z. Han, "Satellite-aerial integrated computing in disasters: User association and offloading decision," in *ICC 2020 - 2020 IEEE Inter. Conf. Commun. (ICC)*, June 2020, pp. 554–559.

[43] Y. Zhang, H. Zhang, H. Zhou, and W. Li, "Interference cooperation based resource allocation in NOMA terrestrial-satellite networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2021, pp. 01–05.

[44] C.-Q. Dai, J. Luo, S. Fu, J. Wu, and Q. Chen, "Dynamic user association for resilient backhauling in satellite-terrestrial integrated networks," *IEEE Systems Journal*, vol. 14, no. 4, pp. 5025–5036, Dec. 2020.

[45] J. Du, C. Jiang, H. Zhang, Y. Ren, and M. Guizani, "Auction design and analysis for SDN-based traffic offloading in hybrid satellite-terrestrial networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 10, pp. 2202–2217, Oct. 2018.

[46] Y. Zhang, H. Zhang, H. Zhou, K. Long, and G. K. Karagiannidis, "Resource allocation in terrestrial-satellite-based next generation multiple access networks with interference cooperation," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1210–1221, 2022.

[47] H.-W. Lee, C.-C. Chen, S. Liao, A. Medles, D. Lin, I.-K. Fu, and H.-Y. Wei, "Interference mitigation for reverse spectrum sharing in B5G/6G satellite-terrestrial networks," *IEEE Trans. Veh. Technol.*, pp. 1–16, 2023.

[48] H. Martikainen, M. Majamaa, and J. Puttonen, "Coordinated dynamic spectrum sharing between terrestrial and non-terrestrial networks in 5g and beyond," in *IEEE 24th Int. Symposium WoWMoM*, June 2023, pp. 419–424.

[49] H. Nguyen-Kha, V. Nguyen Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Joint two-tier user association and resource management for integrated satellite-terrestrial net-

works," *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 16 648–16 665, 2024.

[50] H. Nguyen-Kha, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Two-tier user association and resource allocation design for integrated satellite-terrestrial networks," in *Proc. IEEE International Conf. Commun. Workshop*, 2023.

[51] ——, "LEO-to-user assignment and resource allocation for uplink transmit power minimization," in *Proc. WSA & SCC 2023*, 2023.

[52] H. Nguyen-Kha, V. N. Ha, E. lagunas, S. Chatzinotas, and J. Grotz, "Enhanced throughput and seamless handover solutions for urban 5G-vehicle C-band integrated satellite-terrestrial networks," *IEEE Transactions on Communications*, 2025.

[53] H. Nguyen-Kha, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Digital-twin-aided dynamic spectrum sharing and resource management in integrated satellite-terrestrial networks," *IEEE JSAC (submitted)*, 2025.

[54] H. Nguyen-Kha, V. N. Ha, T. T. Nguyen, E. Lagunas, S. Chatzinotas, and J. Grotz, "DT-aided resource management in integrated satellite-terrestrial networks," in *GLOBECOM 2025 - 2025 IEEE Global Commun. Conf.*, 2025.

[55] V. N. Ha, E. Lagunas, T. S. Abdu, H. Chaker, S. Chatzinotas, and J. Grotz, "Large-scale beam placement and resource allocation design for MEO-constellation SATCOM," in *ICC Workshop - 6GSatComNet*, 2023.

[56] T. S. Abdu, E. Lagunas, V. N. Ha, J. Grotz, S. Kisseleff, and S. Chatzinotas, "Demand-aware flexible handover strategy for LEO constellation," in *ICC Workshop - 6GSatComNet*, 2023.

[57] C. Handforth, H. Croxson, and G. Cruz, "Closing the coverage gap: How innovation can drive rural connectivity," 2019, GSMA, London, U.K., Rep.

[58] X. Zhu and C. Jiang, "Integrated satellite-terrestrial networks toward 6G: Architectures, applications, and challenges," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 437–461, Jan. 2022.

[59] B. Tezergil and E. Onur, "Wireless backhaul in 5G and beyond: Issues, challenges and opportunities," *IEEE Commun. Surveys Tut.*, vol. 24, no. 4, pp. 2579–2632, Fourth quarter 2022.

[60] X. Lin, S. Cioni, G. Charbit, N. Chuberre, S. Hellsten, and J.-F. Boutillon, "On the path to 6G: Embracing the next wave of low earth orbit satellite access," *IEEE Commun. Magazine*, vol. 59, no. 12, pp. 36–42, Dec. 2021.

[61] G. Geraci, D. López-Pérez, M. Benzaghta, and S. Chatzinotas, "Integrating terrestrial and non-terrestrial networks: 3D opportunities and challenges," *IEEE Commun. Mag.*, vol. 61, no. 4, pp. 42–48, April 2023.

[62] N. U. Hassan, C. Huang, C. Yuen, A. Ahmad, and Y. Zhang, "Dense small satellite networks for modern terrestrial communication systems: Benefits, infrastructure, and technologies," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 96–103, Oct. 2020.

[63] R. M. Calvo, T. de Cola, J. Poliak, L. Macrì, A. Papa, S. Ayvasik, E. Babaians, and W. Kellerer, "Optical feeder links for future very high-throughput satellite systems in B5G networks," in *European Conf. Optical Commun.*, 2020, pp. 1–4.

[64] Z. Xiao, J. Yang, T. Mao, C. Xu, R. Zhang, Z. Han, and X.-G. Xia, "LEO satellite access network (LEO-SAN) towards 6G: Challenges and approaches," *IEEE Wireless Communications*, pp. 1–8, 2022.

[65] 3GPP, "Study on New Radio (NR) to support non-terrestrial networks," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.811, Sept 2020, version 15.4.0.

[66] ——, "Study on satellite access phase 3," 3rd Generation Partnership Project (3GPP), Technical report (TR) 22.865, June 2023, version 19.0.0.

[67] C. Caini, H. Cruickshank, S. Farrell, and M. Marchese, "Delay- and disruption-tolerant networking (dtn): An alternative solution for future satellite networking applications," *Proceedings of the IEEE*, vol. 99, no. 11, pp. 1980–1997, Nov 2011.

[68] L. You, X. Qiang, K.-X. Li, C. G. Tsinos, W. Wang, X. Gao, and B. Ottersten, "Hybrid analog/digital precoding for downlink massive MIMO LEO satellite communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 5962–5976, Aug. 2022.

[69] V. N. Ha, T. T. Nguyen, E. Lagunas, J. C. Merlano Duncan, and S. Chatzinotas, "GEO payload power minimization: Joint precoding and beam hopping design," in *GLOBECOM 2022 - 2022 IEEE Global Commun. Conf.*, 2022, pp. 6445–6450.

[70] A. Alimohammad, S. Fard, B. Cockburn, and C. Schlegel, "Compact rayleigh and rician fading simulator based on random walk processes," *IET Commun.*, vol. 3, no. 8, pp. 1333–1342, Aug. 2009.

[71] S. Haggui, F. Rouissi, Y. Mlayeh, and F. Tlili, "A compact architecture of a mobile-to-mobile fading channel emulator based on random walk process," in *ICMCS'16*, Sep. 2016, pp. 564–569.

[72] V. N. Ha, T. T. Nguyen, L. B. Le, and J.-F. Frigon, "Admission control and network slicing for multi-numerology 5G wireless networks," *IEEE Networking Letters*, vol. 2, no. 1, pp. 5–9, March 2020.

[73] T. T. Nguyen, V. N. Ha, and L. B. Le, "Wireless scheduling for heterogeneous services with mixed numerology in 5g wireless networks," *IEEE Communications Letters*, vol. 24, no. 2, pp. 410–413, 2020.

[74] F. She, H. Luo, W. Chen, and X. Wang, "Joint queue control and user scheduling in MIMO broadcast channel under zero-forcing multiplexing," in *2008 IEEE Inter. Conf. Commun. (ICC)*, May 2008, pp. 275–279.

[75] H. Viswanathan and K. Kumaran, "Rate scheduling in multiple antenna downlink wireless systems," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 645–655, Apr. 2005.

[76] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted L1 minimization," *J. Fourier Analysis and Applications*, vol. 14, pp. 877–905, Dec 2008.

[77] V. N. Ha, L. B. Le, and N.-D. Dao, "Coordinated multipoint transmission design for cloud-RANs with limited fronthaul capacity constraints," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7432–7447, 2016.

[78] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *2006 IEEE Int. Conf. Commun.*, vol. 7, June 2006, pp. 3270–3275.

[79] M. Monteiro, N. Lindqvist, and A. Klautau, "Spectrum balancing algorithms for power minimization in DSL networks," in *2009 IEEE Int. Conf. Commun.*, June 2009, pp. 1–5.

[80] J. Li, L. Zhang, K. Xue, Y. Fang, and Q. Sun, "Secure transmission by leveraging multiple intelligent reflecting surfaces in miso systems," *IEEE Trans. Mobile Computing*, vol. 22, no. 4, pp. 2387–2401, April 2023.

[81] S. Boyd and J. Park, "Subgradient methods," *lecture notes of EE364b, Stanford University, Spring Quarter*, 2014.

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

[83] J. Dahl and E. Andersen, "A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization," *Mathematical Programming*, vol. 194, 03 2021.

[84] A. Serrano, "Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone," Thesis, Stanford University, 2015. [Online]. Available: http://purl.stanford.edu/sn367tt9726

[85] A. Skajaa and Y. Ye, "A homogeneous interior-point algorithm for nonsymmetric convex conic optimization," *Mathematical Programming*, vol. 150, 05 2014.

[86] G. Zheng, S. Chatzinotas, and B. Ottersten, "Generic optimization of linear precoding in multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2308–2320, June 2012.

[87] E. Salonen and S. Uppala, "New prediction method of cloud attenuation," *Electronics Letters*, vol. 27, no. 12, pp. 1106–1108, June 1991.

[88] I. Ali, N. Al-Dhahir, and J. Hershey, "Doppler characterization for LEO satellites," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 309–313, 1998.

[89] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine learning for 6g wireless networks: Carrying forward enhanced bandwidth, massive access, and ultrareliable/low-latency service," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 122–134, Dec. 2020.

[90] T. M. Kebedew *et al.*, "Reinforcement learning for qoe-oriented flexible bandwidth allocation in satellite communication," in *IEEE Globecom Workshops on Cellular UAV and Satellite Communications*, 2023.

[91] G. Fontanesi *et al.*, "Artificial intelligence for satellite communication and non-terrestrial networks: A survey," 2023.

[92] ESA, "Conference: Direct-to-device connectivity: An opportunity for europe," 2024. [Online]. Available: https://connectivity.esa.int/directtodevice-connectivity-opportunity-europe

[93] N. Okati, A. N. Barreto, L. U. Garcia, and J. Wigard, "Co-existence of terrestrial and non-terrestrial networks in S-band," 2024. [Online]. Available: http://arxiv.org/abs/2401.08453

[94] D. Kim, J. Park, J. Choi, and N. Lee, "Spectrum sharing between low earth orbit satellite and terrestrial networks: A stochastic geometry perspective analysis," 2024. [Online]. Available: http://arxiv.org/abs/2408.12145

[95] L. F. Abanto-Leon, A. Asadi, A. Garcia-Saavedra, G. H. Sim, and M. Hollick, "Radiorchestra: Proactive management of millimeter-wave self-backhauled small cells via joint optimization of beamforming, user association, rate selection, and admission control," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 153–173, 2023.

[96] H. H. M. Tam, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, "Joint load balancing and interference management for small-cell heterogeneous networks with limited backhaul capacity," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 872–884, 2017.

[97] OpenStreetMap. [Online]. Available: www.openstreetmap.org

[98] B.-H. Yeh, J.-M. Wu, and R. Y. Chang, "Efficient doppler compensation for LEO satellite downlink OFDMA systems," *IEEE Transactions on Vehicular Technology*, 2024.

[99] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.901, Dec. 2019, 16.0.1.

[100] ——, "NR and NG-RAN overall description; stage-2," 3rd Generation Partnership Project (3GPP), Technical specification (TS) 38.300, March 2024, version 18.1.0.

[101] V. N. Ha *et al.*, "User-centric beam selection and precoding design for coordinated multiple-satellite systems," in *IEEE PIMRC*, 2024.

[102] H. Xv, Y. Sun, Y. Zhao, M. Peng, and S. Zhang, "Joint beam scheduling and beamforming design for cooperative positioning in multi-beam leo satellite networks," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 4, pp. 5276–5287, 2024.

[103] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *J. Global Optimization*, vol. 47, pp. 29–51, May 2010.

[104] K.-H. Nguyen, H. V. Nguyen, M. T. P. Le, L. Sanguinetti, and O.-S. Shin, "On the energy efficiency maximization of NOMA-aided downlink networks with dynamic user pairing," *IEEE Access*, vol. 10, pp. 35 131–35 145, 2022.

[105] A. Nemirovski, "Interior point polynomial time methods in convex programming," lecture notes 42.16 (2004): 3215-3224. [Online]. Available: https://www2.isye.gatech.edu/~nemirovs/Lect_IPM.pdf

[106] T. G. L. Authority, "London's age profile," 2021. [Online]. Available: https://mangomap.com/demographics/maps/50068/age-distribution

[107] A. Guidotti, A. Vanelli-Coralli, A. Mengali, and S. Cioni, "Non-terrestrial networks: Link budget analysis," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–7.

[108] S. Communications, "Simplifying 5G with a network digital twin," 2019, white paper.

[109] H. X. Nguyen, R. Trestian, D. To, and M. Tatipamula, "Digital twin for 5g and beyond," *IEEE Commun. Mag.*, vol. 59, no. 2, pp. 10–15, 2021.

[110] M. Zhu, L. Cazzella, F. Linsalata, M. Magarini, M. Matteucci, and U. Spagnolini, "Toward real-time digital twins of em environments: Computational benchmark for ray launching software," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 6291–6302, 2024.

[111] D. Van Huynh *et al.*, "Urllc edge networks with joint optimal user association, task offloading and resource allocation: A digital twin approach," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7669–7682, 2022.

[112] C. Hu *et al.*, "Digital twin-assisted real-time traffic data prediction method for 5g-enabled internet of vehicles," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2811–2819, 2022.

[113] W. Sun, P. Wang, N. Xu, G. Wang, and Y. Zhang, "Dynamic digital twin and distributed incentives for resource allocation in aerial-assisted internet of vehicles," *IEEE IoT J.*, vol. 9, no. 8, pp. 5839–5852, 2022.

[114] Y. Gong, Y. Wei, Z. Feng, F. R. Yu, and Y. Zhang, "Resource allocation for integrated sensing and communication in digital twin enabled internet of vehicles," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 4510–4524, 2023.

[115] Chee, Tommy and Emiliani, Luis and Politis, Christos and Combelles,Laurent , "D2.5 report on regulatory requirements: Spectrum policies and regulatory analysis," 2024, 6G-NTN (6G Non Terrestrial Network) project report D2.5.

[116] F. Kavehmadavani, V.-D. Nguyen, T. X. Vu, and S. Chatzinotas, "Intelligent traffic steering in beyond 5g open ran based on lstm traffic prediction," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 7727–7742, 2023.

[117] A. B. Kihero, M. S. J. Solaija, and H. Arslan, "Inter-numerology interference for beyond 5g," *IEEE Access*, vol. 7, 2019.

[118] P. K. Korrai, E. Lagunas, A. Bandi, S. K. Sharma, and S. Chatzinotas, "Joint power and resource block allocation for mixed-numerology-based 5g downlink under imperfect csi," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1583–1601, 2020.

[119] 3GPP, "NR; physical channels and modulation," Technical specification (TS) 38.211, Dec. 2023, version 18.1.0.

[120] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel Coding Rate in the Finite Block-length Regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[121] S. Schiessl, J. Gross, and H. Al-Zubaidy, "Delay analysis for wireless fading channels with finite blocklength channel coding," in *Proc. 18th ACM MSWiM*, Nov 2015, pp. 13–22.

[122] L. Bonati, R. Shirkhani, C. Fiandrino, S. Maxenti, S. D'Oro, M. Polese, and T. Melodia, "Twinning commercial network traces on experimental open ran platforms," in *Proc. MobiCom*, Dec 2024, pp. 1914–1921.

[123] T. V. S. Sreedhar and N. B. Mehta, "Inter-numerology interference in mixed numerology ofdm systems in time-varying fading channels with phase noise," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5473–5485, 2023.

[124] J. Mao, L. Zhang, P. Xiao, and K. Nikitopoulos, "Interference Analysis and Power Allocation in the Presence of Mixed Numerologies," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5188–5203, Aug. 2020.

[125] V. N. Ha, Z. Abdullah, G. Eappen, J. C. M. Duncan, R. Palisetty, J. L. G. Rios, W. A. Martins, H.-F. Chou, J. A. Vasquez, L. M. Garces-Socarras, H. Chaker, and S. Chatzinotas, "Joint Linear Precoding and DFT Beamforming Design for Massive MIMO Satellite Communication," in *IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, pp. 1121–1126.

[126] 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*, vol. 23, no. 8, pp. 10 091–10 103, Aug. 2024.

[127] R. Campana, C. Amatetti, and A. Vanelli-Coralli, "O-RAN based non-terrestrial networks: Trends and challenges," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Jun. 2023, pp. 264–269, iSSN: 2575-4912.

[128] Z. Li, S. Wang, S. Han, and C. Li, "Non-orthogonal broadcast and unicast joint transmission for multibeam satellite system," *IEEE Transactions on Broadcasting*, vol. 69, no. 3, pp. 647–660, 2023.

[129] T. Shi, S. Sun, S. Kang, W. Sun, and R. Liu, "Robust multi-cast beamforming optimization on real spectral efficiency in satellite systems," *IEEE Transactions on Vehicular Technology*, 2024.

[130] S. Bhandari, T. X. Vu, and S. Chatzinotas, "User-centric flexible resource management framework for leo satellites with fully regenerative payload," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 5, pp. 1246–1261, 2024.

[131] H. G. Abreha, I. Maity, H. Chougrani, C. Politis, and S. Chatzinotas, "On-board content caching with dynamic cache reconfiguration in multi-layer satellite edge networks," *IEEE Open Journal of the Communications Society*, pp. 1–1, 2025.