Université du Québec
Institut national de la recherche scientifique
Centre Énergie, Matériaux et Télécommunications

NETWORK PLANNING AND RESOURCE MANAGEMENT FOR UAV-BASED WIRELESS NETWORKS

Par

MINH DAT NGUYEN

Thèse présentée pour l'obtention du grade de
Doctorat en philosophie, Ph.D.
en télécommunications

Jury d'évaluation

Examinateur externe	Prof. Wessam Ajib *Université du Québec à Montréal*
	Prof. Wei-Ping Zhu *Concordia University*
Examinateur interne	Prof. Jean-Charles Grégoire *INRS-ÉMT*
Codirecteur de recherche	Prof. André Girard *INRS-ÉMT*
Directeur de recherche	Prof. Long Bao Le *INRS-ÉMT*

Acknowledgments

Firstly, I am immensely thankful to my supervisor, Professor Long Bao Le, for his gentle encouragement and instruction in my doctoral study at INRS-ÉMT, University of Québec. He has always inspired me in my research. Without his supervision and motivation in the direction of the research, this research could not have been completed. Thanks to his invaluable support and guidance, my Ph.D. is fruitful not only in terms of results but also in terms of the skills that I have gained during my study. I would like to affirm my profound respect and gratitude to my co-supervisor, Professor André Girard of INRS-ÉMT. I sincerely appreciate him for spending his precious time, patiently mentoring, professional consultation, and judgment, as well as inspired recommendations that are very helpful for accomplishing this research.

I would like to extend my gratefulness to other members of my Ph.D. committee - Professor Jean-Charles Grégoire of INRS-ÉMT, University of Québec who has regularly reviewed and constructively commented on the progress of my doctoral study. I would also like to thank Professor Wessam Ajib of Université du Québec à Montréal and Professor Wei-Ping Zhu of Concordia University for serving as the external examiners for my Ph.D. dissertation.

I would like to express gratitude to all my colleagues at the Networks and Cyber Physical Systems Lab (NECPHY-Lab) of INRS-ÉMT, University of Québec, as well as my friends in Montreal: Vu Ha, Tuong Hoang, Thinh Tran, Ti Nguyen, Dung Le, Hoang Vu, Tung Phan, Tri Nguyen, Binh Truong, Vu Truong for their recommendations in research and for the memorable life moments during my time at INRS.

Finally, I want to say the deepest thanks to my family for their unconditional love and continuous support. Based on the constant encouragement from my family and essential guidance from my father, I can follow my dream of doing research and chasing my life's goals. Because of my family, I could continue my study under the pressure of research and the stress of the unexpected situation and finally achieved an academic degree in Canada. I would like to consecrate my work to my family for their endless love, support, and encouragement throughout my life. I thank you all, and I hope my accomplishments will make you proud.

Abstract

Next-generation wireless networks will enable to support applications in various domains including smart factories, intelligent transportation, e-health, and more. Therefore, future wireless communications are expected to provide higher capacity and much lower latency and offer excellent stability, ubiquitous communications, and connectivity to billions of devices. However, the deployment of terrestrial infrastructure faces challenges in various practical scenarios, such as communications to serve temporary events and emergencies like natural disasters and fast service recovery. Toward this end, several promising technologies have been under consideration, including satellite communications, unmanned aerial vehicle (UAV) communications, intelligent reflecting surface (IRS), and mobile edge computing (MEC). The overall objective of this Ph.D. research is to develop network planning and resource management for UAV-based wireless networks. Our research has resulted in three major research contributions, which are presented in three corresponding main chapters of this dissertation.

First, we study the trajectory control, sub-channel assignment, and user association design for UAVs-based wireless networks, which is presented in Chapter 5. In particular, we propose a method to optimize the max-min average rate subject to data demand constraints of ground users (GUs) where spectrum reuse and co-channel interference management are considered. The mathematical model is a mixed integer non-linear optimization problem which we solve by using the alternating optimization approach where we iteratively optimize the user association, sub-channel assignment, and UAV trajectory control until convergence. For the sub-channel assignment sub-problem, we propose an iterative sub-channel assignment (ISA) algorithm to obtain an efficient solution. Moreover, the successive convex approximation (SCA) is used to convexify and solve the non-convex UAV trajectory control sub-problem.

Second, we design an UAV-based wireless network with wireless access and backhaul links leveraging an IRS, which is covered in Chapter 6. Particularly, this design aims to maximize the sum rate achieved by GUs through optimizing the UAV placement, IRS phase shifts, and sub-channel assignments considering the wireless backhaul capacity constraint. To tackle the underlying mixed integer non-linear optimization problem (MINLP), we first derive the closed-form IRS phase shift solution; we then optimize the sub-channel assignment and UAV placement by using the alternating optimization method. Specifically, we propose an iterative sub-channel assignment method to efficiently utilize the bandwidth and balance bandwidth allocation for wireless access and backhaul links while maintaining the backhaul capacity constraint. Moreover, we employ the successive convex approximation (SCA) method to solve the UAV placement optimization sub-problem.

Finally, we study the computation offloading problem in space-air-ground integrated networks (SAGIN), where joint optimization of partial computation offloading, UAV trajectory control, user scheduling, computation, resource allocation, and admission control is performed. The research outcomes of this study are presented in Chapter 7. Specifically, the considered SAGIN employs multiple UAV-mounted edge servers with controllable UAV trajectory and a cloud sever which can be reached by GUs via multi-hop low-earth-orbit (LEO) satellite communications. This design aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of underlying computation tasks. To tackle the underlying non-convex mixed integer non-linear optimization problem, we use the alternating optimization approach where we iteratively solve four sub-problems, namely user scheduling, partial offloading control and bit allocation over time slots, computation resource and bandwidth allocation, and multi-UAV trajectory control until convergence. Moreover, feasibility verification and admission control strategies are proposed to handle overloaded network scenarios. Furthermore, the successive convex approximation (SCA) method is employed to convexify and solve the non-convex computation resource and bandwidth allocation and UAV trajectory control sub-problems.

For all proposed designs and algorithms, we provide extensive analytical and numerical studies which illustrate their achievable performances as the values of different key parameters vary. The numerical studies also demonstrate the efficacy of our proposed algorithms and their significant performance gains versus the state-of-the-art designs.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

3GPP	3rd generation partnership project
5G	Fifth-generation mobile communications technology
5G NR	5G new radio
6G	Sixth-generation mobile communications technology
AC	Admission control
AF	Amplify-and-forward
BS	Base station
FDMA	Frequency division multiple access
DC	Different of two convex functions
GEO	Geostationary earth orbit
GU	Ground user
IoT	Internet of things
IoV	Internet of vehicles
IRS	Intelligent reflecting surface
ISL	Inter-satellite link
LEO	Low earth orbit
LoS	Line-of-sight
LP	Linear programming

LTE	Long-term evolution
MEO	Medium earth orbit
MEC	Mobile edge computing
MINLP	Mixed integer nonlinear programming
MIMO	Multiple-input multiple-output
mmWave	Millimeter wave
NLoS	Non-line-of-sight
NOMA	Non-orthogonal multiple access
NP-Hard	Non-deterministic polynomial-time hard
NT, NTN	Non-terrestrial, non-terrestrial network
OFDM	Orthogonal frequency division multiplexing
OFDMA	Orthogonal frequency division multiple access
QoS	Quality-of-service
RIS	Reconfigurable intelligent surface
RSS	Received signal strength
SAGIN	Space-air-ground integrated networks
SCA	Sucessive convex approximation
SINR	Signal to interference plus noise ratio
SNR	Signal to noise ratio
TDMA	Time division multiple access
TN	Terrestrial network
UAV	Unmanned aerial vehicle

Chapter 1

Extended Summary

1.1 Background and Motivations

Next-generation wireless networks will enable to support applications in various domains including smart factories, intelligent transportation, e-health, and more [3, 4]. The proliferation of many human and Internet of Things (IoT) applications have led to a mobile traffic explosion. Therefore, future wireless communications are expected to provide higher capacity and much lower latency, offer enhanced stability, ubiquitous communications, and connectivity to billions of devices [5–8]. However, the deployment of terrestrial infrastructure faces challenges in various practical scenarios, such as communications to serve temporary events and emergencies like natural disasters and fast service recovery [9–11].

Toward this end, several promising technologies have been under consideration, including satellite communications, unmanned aerial vehicle (UAV) communications, intelligent reflecting surface (IRS), and mobile edge computing (MEC) [12, 13]. In particular, UAV communications have emerged as a potential solution to overcome the limitations of current infrastructure, offering wider coverage, higher resilience, and availability, and improving user's quality of service (QoS) due to their superior attributes such as mobility, flexibility, and adaptive altitude [14, 15]. Besides, the IRS-assisted UAV communications have attracted extensive attention because they can significantly enhance the communication quality. In this system, UAV communicates with ground users (GUs) and IRS can reflect the dissipated signals from the UAV, improving the UAV-GU communications quality [16–18]. Moreover, space-air-ground integrated networks (SAGIN) have emerged as promising architecture to provide high-quality and ubiquitous communications by leveraging the complementary strengths of space, air, and ground networks and enabling the technologies such as edge computing [19–21].

Firstly, there has been strong interest in providing wireless coverage in the three-dimensional (3D) space and leveraging different flying platforms to enhance wireless connectivity and/or the performance of the terrestrial wireless networks [4,10,22,23]. UAV communications can provide low-cost solutions for various communications scenarios, e.g., wireless areas with limited infrastructure or high traffic demand. Moreover, the UAV-based wireless networks can provide extra degrees of freedom to optimize the underlying wireless network to enhance the coverage, throughput, and energy efficiency thanks to unique UAV's attributes such as mobility, flexibility, and controllable altitude.

Secondly, intelligent reflecting surface (IRS) or reconfigurable intelligent surface (RIS) is a promising paradigm which can substantially improve the spectral and energy efficiency of wireless networks by constructing favorable communication channels via tuning massive low-cost passive reflecting elements [24]. In essence, an IRS consists of a large number of low-cost passive elements, where each element can be adjusted with an independent phase shift to reflect the electromagnetic incident signals, to be added coherently at GUs. IRS can be flexibly deployed on various structures, such as building facades, roadside billboards, and indoor walls [25]. IRS-assisted wireless communications can be realized by deploying the IRS between the BS or aerial BS and mobile users to enhance the received signal power [26–30].

Thirdly, space-air-ground integrated networks (SAGIN) have emerged as an effective means of providing high quality and ubiquitous communications by leveraging the complementary strengths of space, air, and ground networks segments [31–33]. On the one hand, in the space network of SAGIN, geostationary earth orbit (GEO) satellites, medium earth orbit (MEO) satellites, and low earth orbit (LEO) satellites are the main components [34]. LEO satellites are liable to form networks by inter-satellite links (ISLs), which guarantee lower propagation delay, high communication rates, and seamless communication services for wide geographical areas [35,36]. On the other hand, in the air network of SAGIN, there is a mobile aerial system that uses aircraft as carriers for information acquisition, transmission, and processing. The UAVs, airships, and balloons are the main infrastructures making up the high and low-altitude platforms (HAPs & LAPs) which can provide broadband wireless communications complementing the terrestrial networks [31, 37]. Meanwhile, in the ground network of SAGIN, the network mainly consists of terrestrial communication systems such as cellular networks, mobile ad hoc networks, wireless local area networks, and so on [31]. A UAV-assisted MEC system enables efficient support for computation-extensive mobile applications thanks to controllable UAVs' trajectories, extensive coverage and additional computation capability.

In this dissertation, our main objective is to study network planning and resource management for UAV-based wireless networks. In more precise contexts, the results of this dissertation could be useful to address some long-range planning problems over a horizon of a year or more. Real-time implementation issues are outside the scope of this dissertation. Besides, the models are deterministic optimization problems where all input data are known, e.g., horizontal coordinates of UAVs and GUs. The GUs can be viewed as aggregates of traffic sources over a small region. The traffic demands are also averages of the demand over the planning horizon. Moreover, the proposed models and designs can provide answer some questions that are relevant in this context. Examples are how many UAVs should the network provider buy, whether or not the UAVs should be fixed or moving, how many IRS to buy and where they should be installed, whether a cloud architecture is worth it and where it should be located, etc. Furthermore, the more real-time issues such as GU-UAV association, UAV trajectory control, or computation splitting and offloading can also be used as guidelines for the real-time algorithms. Specifically, the research contributions of this dissertation are summarized in the next sections.

1.2 Research Contributions

In this dissertation, our main objective is to develop network planning and resource management in UAV communications for future wireless networks. In particular, our work focuses on three aspects. The first aspect is integrated UAV trajectory control and resource allocation for UAV-based wireless networks with co-channel interference management. In the second aspect, we study UAV placement and resource allocation for intelligent reflecting surface-assisted UAV-based wireless networks. Finally, we study integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control in SAGIN with multi-hop LEO satellite communications. The following sections describe the main contributions of this dissertation.

1.2.1 Integrated UAV Trajectory Control and Resource Allocation for UAV-Based Wireless Networks with Co-channel Interference Management

In this contribution, we study the joint UAV-GU association, resource allocation, and UAV trajectory control for UAV-based wireless networks with spectrum reuse and interference management. The main contributions can be summarized as follows:

- We formulate the joint UAV-GU association, UAV trajectory control, and non-orthogonal sub-channel assignment problem for UAV-based wireless networks. We maximize the minimum average rate of all GUs considering constraints on data transmission demands of individual GUs.
- We solve the underlying mixed-integer non-linear optimization problem (MINLP) problem using the alternating optimization approach. We solve the UAV-GU association, sub-channel assignment, and UAV trajectory control sub-problems separately in each iteration until convergence. We develop an iterative sub-channel assignment (ISA) algorithm to tackle the sub-channel assignment sub-problem. Given the UAV-GU association and sub-channel assignment solutions, the UAV trajectory control sub-problem is a difficult non-convex problem. We propose to use the successive convex approximation (SCA) technique to convexify and solve this sub-problem. We then present a short complexity analysis of the proposed algorithm.
- Extensive numerical results are presented to show the performance of our algorithm. Specifically, we compare the network performance when the proposed ISA sub-channel algorithm and a baseline heuristic sub-channel assignment with interference management (SAIM) algorithm are used to solve the joint problem. We also study the impacts of different parameters and the importance of trajectory control on the achievable performance. Finally, we illustrate the convergence of the algorithm.

1.2.1.1 System Model

We consider a network where a set of UAVs denoted as $\mathcal{M} = \{1, ..., M\}$, provides wireless connectivity for a set of GUs, denoted as $\mathcal{K} = \{1, ..., K\}$. We assume that each GU needs to receive a specific amount of data from UAVs in the downlink direction. This can be the case in many practical scenarios, e.g., GUs want to receive video files from the UAV such as specific scenes of a football match.

Because the UAVs are flying at a relatively high altitude, we assume that all communications, be it UAV-to-BS or UAV-to-GU, are dominated by line-of-sight (LoS) propagation. The UAVs are assumed to be connected to the core network wirelessly through one cellular BS where the UAV-BS links are assumed to have a sufficiently large capacity i.e., by using mmWave communications. We assume that the UAVs fly at a fixed altitude H over a flight period of $T > 0$ seconds. The flight period is divided into N time slots where the set of time slots is denoted as $\mathcal{N} = \{1, ..., N\}$. At any time slot during the flight period T, each UAV can communicate with multiple GUs at the same time using orthogonal frequency division multiple access (OFDMA) technique. The GUs are assumed to be located on the ground at zero altitude with fixed horizontal coordinates $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. Moreover, the horizontal coordinate of UAV m in time slot n is denoted as $\mathbf{q}_m[n] = (x_m^{\mathsf{d}}[n], y_m^{\mathsf{d}}[n])$. We assume that each UAV m must come back to its initial position at the end of the flight period and the slot interval $\Delta t = T/N$ is set sufficiently small so that each UAV just flies a small distance during each time slot even at the maximum speed $V_{\max}$.

Let C be the number of sub-channels available to support the wireless access links between UAVs and GUs. We denote the total transmit power of each UAV as $P_{\max} \geq 0$. We assume that the uniform power allocation is used by each UAV i.e., the transmit power on each sub-channel is equal the total transmit power $P_{\max}$ divided by the total sub-channels used for downlink communications and is given by $p = P_{\max}/C$. We define the binary UAV-GU association decision variable $\omega_{k,m}[n]$ which is equal to 1 if GU k is served by UAV m in time slot n and equal to 0, otherwise. In addition to the UAV assignment, let W (MHz) denote the bandwidth of each sub-channel and $\mathcal{C} = \{1, ..., C\}$ denote the set of sub-channels. The sub-channel assignment variables are defined as $\theta_{k,c}[n]$ which are equal to 1 if sub-channel c is assigned to GU k in time slot n and equal to 0, otherwise.

Recall that we have assumed that the communication links from UAVs to GUs are dominated by the LoS propagation where the channel quality is mostly dependent on the UAV-GU distance. In time slot n, the distance between UAV m and GU k can be calculated as $d_{k,m}[n]=\sqrt{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$. Then, the channel power gain from UAV m to GU k in time slot n on sub-channel c is assumed to follow the free-space path loss model and it can be expressed as $g_{k,m}[n] = \rho_0 d_{k,m}^{-2}[n] = \frac{\rho_0}{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$, where ρ_0 presents the channel power gain at the reference distance of 1 m. The received signal to interference plus noise ratio (SINR) at GU k on sub-channel c can be calculated as

$$\gamma_{k,m,c}[n] = \frac{p g_{k,m}[n]}{\sum_{j=1,j\neq m}^{M} \sum_{z=1,z\neq k}^{K} \omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2}, \tag{1.1}$$

where σ^2 is the power of the additive white Gaussian noise (AWGN) at the receiver. The term $\sum_{j=1,j\neq m}^{M} \sum_{z=1,z\neq k}^{K} \omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n]$ represents the interference at GU k on the sub-channel c due to the transmissions of other UAVs in time slot n on this sub-channel. The achievable rate of GU k served by UAV m in time slot n on the sub-channel c, denoted by $R_{k,m,c}[n]$ in bits/second (bps), can then be expressed as

$$R_{k,m,c}[n] = \omega_{k,m}[n]\theta_{k,c}[n] W \log_2\left(1 + \gamma_{k,m,c}[n]\right). \tag{1.2}$$

Therefore, the total rate achieved by GU k in time slot n, denoted by $R_k[n]$, can be written as

$$R_k[n] = \sum_{m=1}^{M} \sum_{c=1}^{C} R_{k,m,c}[n]. \tag{1.3}$$

As a result, the average rate per slot of GU k over N time slots can be expressed as

$$\bar{R}_k = \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{c=1}^{C} \omega_{k,m}[n]\theta_{k,c}[n] W \log_2\left(1 + \gamma_{k,m,c}[n]\right). \tag{1.4}$$

1.2.1.2 Problem Formulation

For convenience, we gather different decision variables as $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$, $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$ and $\mathbf{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$. Our design goal is to maximize the minimum average rate achieved by all GUs by jointly optimizing the user association i.e., $\mathbf{\Omega}$, sub-channel assignment i.e., $\mathbf{\Theta}$, and UAV trajectory i.e., $\mathbf{Q}$ over all time slots of the flight period.

The average rate $\bar{R}_k$ in (1.4) is a non-linear function with respect to three decision variables $\mathbf{\Omega}, \mathbf{\Theta}$, and $\mathbf{Q}$. Instead of performing the max-min optimization of this non-linear function, we introduce the function $\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right) = \min_{k\in\mathcal{K}} \bar{R}_k$ as the minimum average rate of all GUs. Then, our optimization problem becomes equivalent to maximizing $\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right)$, which is more tractable. Moreover, we assume that GU $k, \forall k \in \mathcal{K}$, has the minimum data transmission demand of D_k^{min}, which must be received in the downlink direction over the UAV flight period. Then, the joint

UAV-GU association, sub-channel assignment, and UAV trajectory control optimization problem to maximize the minimum average rate over all GUs can be formulated as

$$\textbf{(P1):} \max_{\eta,\boldsymbol{\Omega},\boldsymbol{\Theta},\mathbf{Q}} \ \eta \tag{1.5}$$

$$\text{s.t. } \bar{R}_k \geq \eta,\ \forall k, \tag{1.5a}$$

$$\sum_{n=1}^{N} \Delta t R_k[n] \geq D_k^{\mathsf{min}},\ \forall k, \tag{1.5b}$$

$$\|r_0 - \mathbf{q}_m[n]\| \leq R_0,\ \forall m,n, \tag{1.5c}$$

$$\sum_{m=1}^{M} \omega_{k,m}[n] = 1,\ \forall k,n, \tag{1.5d}$$

$$\sum_{k=1}^{K} \omega_{k,m}[n]\theta_{k,c}[n] \leq 1, \forall m,n,c, \tag{1.5e}$$

$$\sum_{c=1}^{C} \theta_{k,c}[n] \geq 1, \forall k,n, \tag{1.5f}$$

$$\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m, \tag{1.5g}$$

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \leq S_{\mathsf{max}}^2,\ n{=}1,...,N{-}1, \tag{1.5h}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d_{\mathsf{min}}^2,\ \forall n,m,j \neq m, \tag{1.5i}$$

$$\omega_{k,m}[n] \in \{0,1\}, \forall k,m,n, \tag{1.5j}$$

$$\theta_{k,c}[n] \in \{0,1\}, \forall k,c,n, \tag{1.5k}$$

where R_0 represents the radius of the network area centered at r_0. Constraints (1.5b) capture the required data transmission demand for each GU over the flight period of T seconds, while constraints (1.5c) restrict the trajectories of all UAVs inside the desired network area. Moreover, (1.5d)-(1.5e) present the UAV-GU association constraints, (1.5e)-(1.5f) capture constraints on the sub-channel assignment, and (1.5g)-(1.5i) represent constraints on the UAVs' trajectories. It can be seen that the constraints (1.5a), (1.5b), and (1.5i) are non-linear and integer decision variables are involved in (1.5j) and (1.5k) for the UAV-GU association and sub-channel assignment, respectively. Hence, problem (1.5) is a mixed integer non-linear optimization problem (MINLP), which is difficult to solve optimally.

1.2.1.3 Proposed Algorithm

We adopt the alternating optimization approach to solve problem (1.5) where we iteratively optimize each set of variables given the values of other variables in the corresponding sub-problems until convergence. We describe how to solve these different sub-problems in the following.

a) UAV-GU Association Given Sub-channel Assignment and UAV Trajectory Control

For the given sub-channel assignment $\mathbf{\Theta}$ and UAV trajectory $\mathbf{Q}$, the problem of optimizing the UAV-GU association $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$ to achieve the max-min average rate over all GUs is still a integer non-linear optimization problem. To make the problem more tractable, we relax the integer decision variables in $\mathbf{\Omega}$ into continuous decision variables, which yields the following problem

$$\textbf{(P1.1):} \max_{\eta,\mathbf{\Omega}} \quad \eta \tag{1.6}$$

$$\text{s.t.} \quad 0 \leq \omega_{k,m}[n] \leq 1, \forall k, m, n, \tag{1.6a}$$

$$\text{constraints (1.5a), (1.5b), (1.5d), (1.5e).}$$

Even with this relaxation, problem (1.6) is still a non-convex optimization problem due to the non-convex constraints (1.5a) and (1.5b). To this end, $R_{k,m,c}[n]$, in constraints (1.5a) and (1.5b), can be re-written as

$$\begin{aligned} R_{k,m,c}[n] &= \omega_{k,m}[n]\theta_{k,c}[n]W \log_2\left(1 + \frac{pg_{k,m}[n]}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n] + \sigma^2}\right) \\ &\geq \omega_{k,m}[n]\theta_{k,c}[n]W R^{\mathsf{A}}_{k,m,c}[n], \end{aligned} \tag{1.7}$$

where

$$R^{\mathsf{A}}_{k,m,c}[n] \leq \log_2\left(1+\frac{pg_{k,m}[n]}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2}\right). \tag{1.8}$$

By introducing auxiliary variables $\mathbf{R}^{\mathsf{A}} = \{R^{\mathsf{A}}_{k,m,c}[n], \forall k, m, c, n\}$, and based on the first-order Taylor expansion at the given points $\omega^{r}_{k,m}[n]$ and $R^{\mathsf{A},r}_{k,m,c}[n]$ in the r-th iteration of the approximation process, we can obtain the following inequality

$$\omega_{k,m}[n]R^{\mathsf{A}}_{k,m,c}[n] \geq \frac{1}{4}\Big[-\Big(\omega^r_{k,m}[n]+R^{\mathsf{A},r}_{k,m,c}[n]\Big)^2+2\Big(\omega^r_{k,m}[n]+R^{\mathsf{A},r}_{k,m,c}[n]\Big)\Big(\omega_{k,m}[n]+R^{\mathsf{A}}_{k,m,c}[n]\Big)- \Big(\omega_{k,m}[n]-R^{\mathsf{A}}_{k,m,c}[n]\Big)^2\Big] \triangleq R^{\mathsf{Alb},r}_{k,m,c}[n]. \tag{1.9}$$

Moreover, the right-hand side (RHS) of constraints (1.8) is convex with respect to $\omega_{z,j}[n]$. Thus, by applying the first-order Taylor expansion at the given points $\omega^r_{z,j}[n]$, we can obtain the lower bound $R^{\mathsf{AA},r}_{k,m,c}[n]$ as in (1.10).

$$\log_2\left(1+\frac{pg_{k,m}[n]}{\sum_{j\neq m}\sum_{z\neq k}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2}\right) \geq \log_2\left(1+\frac{pg_{k,m}[n]}{\sum_{j\neq m}\sum_{z\neq k}\omega^r_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2}\right)- \sum_{j\neq m}\sum_{z\neq k}A_{z,j,k,m,c}[n]\big(\omega_{z,j}[n]-\omega^r_{z,j}[n]\big) \triangleq R^{\mathsf{AA},r}_{k,m,c}[n], \tag{1.10}$$

where

$$A_{z,j,k,m,c}[n]=\frac{\omega^r_{z,j}[n]\theta_{z,c}[n]p^2g_{k,j}[n]g_{k,m}[n]\log_2(e)}{\Big(\sum_{j\neq m}\sum_{z\neq k}\omega^r_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2\Big)\Big(\sum_{j\neq m}\sum_{z\neq k}\omega^r_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2+pg_{k,m}[n]\Big)}.$$

Using the approximations above, problem (1.6) can be approximated by the following problem:

$$\textbf{(P1.1"):} \max_{\eta^r_{\mathsf{a}},\mathbf{\Omega},\mathbf{R}^{\mathsf{A}}} \ \eta^r_{\mathsf{a}} \tag{1.11}$$

$$\text{s.t.} \ \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}R^{\mathsf{Alb},r}_{k,m,c}[n]\theta_{k,c}[n]W \geq \eta^r_{\mathsf{a}}, \forall k, \tag{1.11a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\Delta t R^{\mathsf{Alb},r}_{k,m,c}[n]\theta_{k,c}[n]W \geq D^{\mathsf{min}}_k, \forall k, \tag{1.11b}$$

$$R^{\mathsf{A}}_{k,m,c}[n] \leq R^{\mathsf{AA},r}_{k,m,c}[n], \forall k,m,c,n, \tag{1.11c}$$

$$\text{constraints (1.5d), (1.5e), (1.6a).}$$

It can be seen that all constraints are linear. Hence, problem (1.11) is a standard convex optimization problem which can be solved efficiently by any convex optimization solvers such as CVX-Mosek [38]. Detailed description of our proposed algorithm to solve the UAV-GU association problem is given in Algorithm 1.1. In the solution obtained by Algorithm 1.1, if the UAV-GU association variables $\omega_{k,m}[n]$ are all binary, then the relaxation is tight and

Algorithm 1.1. SCA-based Algorithm to Solve (1.6)

Initialization: Set $r := 0$, generate an initial point $\left(\mathbf{\Omega}^{0}, \mathbf{R}^{\mathsf{A},0}\right)$ of (1.11);
repeat
  $r := r + 1$;
  Solve (1.11) to obtain optimal values $\left(\mathbf{\Omega}^{*}, \mathbf{R}^{\mathsf{A},*}\right)$;
  Update $\left(\mathbf{\Omega}^{r}, \mathbf{R}^{\mathsf{A},r}\right) := \left(\mathbf{\Omega}^{*}, \mathbf{R}^{\mathsf{A},*}\right)$;
until Convergence
Output $\eta_{\mathsf{a}}^{*}, \mathbf{\Omega}^{*}, \mathbf{R}^{\mathsf{A},*}$.

the obtained solution is also a feasible solution of problem **(P1)**. Otherwise, the UAV-GU association solution needs to be recovered by rounding it to the nearest integer of 0 or 1. Furthermore, since constraints (1.5d) and (1.5e) are met with equalities in the solution of (1.11), a binary solution can be recovered.

b) Sub-channel Assignment Given UAV-GU Association and UAV Trajectory

For the given UAV-GU association and UAV trajectory $\{\mathbf{\Omega}, \mathbf{Q}\}$, we optimize the sub-channel assignment $\mathbf{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$ to achieve the max-min average rate among all GUs. This problem can be expressed as follows:

$$\textbf{(P1.2):} \max_{\eta, \mathbf{\Theta}} \ \eta \tag{1.12}$$
$$\text{s.t.} \quad \text{constraints } (1.5\text{a}), (1.5\text{b}), (1.5\text{e}), (1.5\text{f}), (1.5\text{k}).$$

We propose a heuristic but efficient algorithm for sub-channel assignments. Recall that our design objective is to maximize the minimum average rate among all GUs and satisfy the data transmission demands of individual GUs, i.e., $D_k^{\mathsf{min}}, \forall k \in \mathcal{K}$. Hence, in the first phase, we perform sub-channel assignments for each GU to not only improve the design objective, but also ensure the constraints on data transmission demands of all GUs be satisfied. Specifically, we search a sub-channel assignment for each GU k associated with UAV m in a certain time slot n to achieve higher and maximum increase in the average rate of GU k and ensure the minimum average rate of the system is not decreasing in each assignment step.

After the required data transmission demands of all GUs are satisfied, the algorithm enters an iterative sub-channel assignment loop where in each iteration, it searches the GU with the minimum average rate and finds the best sub-channel assignment achieving the highest and better average rate for the underlying GU while improving the minimum average rate of the system. In fact, the method to determine the best sub-channel assignment solution in this

Algorithm 1.2. Iterative Sub-channel Assignment (ISA) Algorithm

Require: M UAVs, K GUs, C sub-channels;
1: ***Given*:** UAV-GU association, UAV trajectory control;
Ensure: Max-min average rate ($\bar{R}_k$), η;
2: $k = 1$;
3: **while** $k \le K$ **do**
4:   **repeat**
5:     Calculate the minimum average rate of the system: minrate $= \min_{k\in\mathcal{K}}\{\bar{R}_k\}$;
6:     Given GU k, identify all UAV and time slot pairs $\{m, n\}$ with $\{\omega_{k,m}[n] = 1\}$;
7:     Given GU k and each pair $\{m, n\}$ identified in step 6, find the sub-channel c for assignment to achieve the highest and better average rate for GU k;
8:     Compare all potential sub-channel assignments for different pairs $\{m, n\}$ found in step 7, realize the best sub-channel assignment if it can improve the minimum average rate of the system, i.e., we calculate rate $= \min_{k\in\mathcal{K}}\{\bar{R}_k\}$ and the new sub-channel assignment must satisfy minrate $<$ rate;
9:   **until** $\sum_n D_k[n] \ge D_k^{\mathsf{min}}$
10:   $k \leftarrow k + 1$;
11: **end while**
12: **repeat**
13:   Find GU $k = \text{argmin}_{k\in\mathcal{K}}\{\bar{R}_k\}$;
14:   Calculate the minimum average rate of the system: minrate$^* = \min_{k\in\mathcal{K}}\{\bar{R}_k\}$;
15:   Given GU k, identify all UAV and time slot pairs $\{m, n\}$ with $\{\omega_{k,m}[n] = 1\}$;
16:   Given GU k and each pair $\{m, n\}$ identified in step 15, find the sub-channel c for assignment to achieve the highest and better average rate for GU k;
17:   Compare all potential sub-channel assignments for different pairs $\{m, n\}$ found in step 16, realize the best sub-channel assignment if it can improve the minimum average rate of the system, i.e., we calculate rate $= \min_{k\in\mathcal{K}}\{\bar{R}_k\}$ and the new sub-channel assignment must satisfy minrate$^* <$ rate;
18:   Update minrate* = rate;
19: **until** Convergence
20: Update $\eta^* \leftarrow$ minrate*;
21: Return η^*, Θ^*.

loop is similar to that in the previous phase. The algorithm terminates when the minimum average rate of all GUs cannot be improved further. Details of the proposed algorithm called *"Iterative Sub-channel Assignment (ISA) Algorithm"* are given in Algorithm 1.2.

c) UAV Trajectory Control Given UAV-GU Association and Sub-channel Assignment

Given the UAV-GU association and sub-channel assignment $\{\mathbf{\Omega}, \mathbf{\Theta}\}$, the problem optimizing the UAV trajectory control $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$ to achieve the max-min average rate over all GUs can be written as follows:

$$\textbf{(P1.3):} \max_{\eta,\mathbf{Q}} \quad \eta \tag{1.13}$$
$$\text{s.t.} \quad \text{constraints (1.5a), (1.5b), (1.5c), (1.5g), (1.5h), (1.5i).}$$

Algorithm 1.3. SCA-based Algorithm to Solve (1.13)

1: **Initialization:** Set $r := 0$, generate an initial point $\left(\mathbf{Q}^0, \mathbf{S}^0, \mathbf{R}^0\right)$ of (1.14);
2: **repeat**
3: $r := r + 1$;
4: Solve (1.14) to obtain optimal values $(\mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*)$;
5: Update $(\mathbf{Q}^r, \mathbf{S}^r, \mathbf{R}^r) := (\mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*)$;
6: **until** Convergence
7: **Output** $\eta^*_{\mathsf{trj}}, \mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*$.

This problem is a non-convex optimization problem due to the non-convex constraints (1.5a), (1.5b) and (1.5i). Therefore, it is difficult to solve this problem optimally. We design an algorithm with three main steps to solve this problem as follows. In Step 1, we introduce some auxiliary variables and transform problem (1.13) into an equivalent form. Then, we approximately convexify the corresponding problem in Step 2. Finally, we use a convex optimization solver to solve the obtained convex problem in Step 3. Therefore, problem (1.13) can be approximated as

$$\textbf{(P1.3"):} \max_{\eta^r_{\mathsf{trj}}, \mathbf{Q}, \mathbf{S}, \mathbf{R}} \quad \eta^r_{\mathsf{trj}} \tag{1.14}$$

$$\text{s.t. } \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]\right) \geq \eta^r_{\mathsf{trj}},\ \forall k, \tag{1.14a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\Delta t\omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]\right) \geq D^{\mathsf{min}}_k,\ \forall k, \tag{1.14b}$$

$$S_{k,m}[n] \leq \|\mathbf{q}^r_m[n] - \mathbf{r}^{\mathsf{u}}_k\|^2 + 2\left(\mathbf{q}^r_m[n] - \mathbf{r}^{\mathsf{u}}_k\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}^r_m[n]\right), \forall k, m, n, \tag{1.14c}$$

$$d^2_{\mathsf{min}} \leq -\left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right), \forall j \neq m, n, \tag{1.14d}$$

$$\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0 - R^{\mathsf{Ab}}_{z,j,c,k,m}[n]H^2 \geq R^{\mathsf{App},r}[n], \forall n, \tag{1.14e}$$

$$(1.5\text{c}), (1.5\text{g}), (1.5\text{h}),$$

where $\mathbf{R} = \{R^{\mathsf{Ab}}_{z,j,c,k,m}[n], \forall k, m, z, j, c, n\}$ and $\mathbf{S} = \{S_{k,m}[n], \forall k, m, n\}$.

Hence, problem (1.14) is a standard convex optimization problem which can be solved efficiently by any convex optimization solvers such as CVX-Mosek [38]. Detailed description of our proposed algorithm to solve the UAV trajectory control optimization problem is given in Algorithm 1.3.

d) Integrated UAV-GU Association, Sub-channel Assignment and UAV Trajectory Control

Algorithm 1.4. Integrated UAV-GU Association, Sub-channel Assignment and UAV Trajectory Control

Require: M UAVs, K GUs, C sub-channels and T;
Ensure: Max-min average rate $(\bar{R}_k)$, η; Let $r = 1$;
1: **repeat**
2: Optimize the UAV-GU association given the sub-channel assignment and UAVs' trajectories by solving sub-problem using Algorithm 1.1 to obtain $\mathbf{\Omega}^r$;
3: Optimize the sub-channel assignment given the UAV-GU association and UAVs' trajectories by solving sub-problem using Algorithm 1.2 to obtain $\mathbf{\Theta}^r$;
4: Optimize the UAVs' trajectories given the UAV-GU association and sub-channel assignment using Algorithm 1.3 to obtain $\mathbf{Q}^r$;
5: Update $r = r + 1$;
6: **until** Convergence
7: Return $\eta^*, \mathbf{\Omega}^*, \mathbf{\Theta}^*, \mathbf{Q}^*$.;

Using the results presented in Sections **a)**, **b)** and **c)**, our proposed algorithm based on the alternating optimization method is described in Algorithm 1.4. The convergence of this algorithm is stated in the following proposition.

Proposition 1.1. The proposed Algorithm 1.4 creates a sequence of feasible solutions where the objective value monotonically increases over iterations. As a result, the algorithm converges to a feasible solution.

1.2.1.4 Numerical Results

In this section, we evaluate the performance of the proposed algorithm. The parameter setting for the simulations is similar to that in [39–41] and are summarized in Table 5.3. We consider a circular network area with radius $R_0 = 500$m with two or more clusters i.e., hotspots, of GUs. The radius of each circular cluster area is $r_{\mathsf{c}} = 200$m and different clusters are placed far enough apart not to overlap. The distance between two neighboring clusters' centers is set to satisfy the constraint $D^0 \geq d_{\mathsf{min}} + 2 \times r_{\mathsf{c}}$(m). The altitude of all UAVs is assumed to be fixed at $H = 100$m. Moreover, the required transmission data demand for each GU k (D_k^{min}) is set according to the size of short videos, e.g., video files with the resolution of 30 frames per second (fps) [42]).

We first evaluate the performance of the proposed ISA algorithm described in Alg. 1.2 and the heuristic sub-channel assignment with interference management (SAIM) algorithm shown in Fig. 5.3. Specifically, the max-min average rates due to different schemes are shown in Fig. 1.1 for the network with 2 UAVs, 40 sub-channels, UAV's flight period $T = 20$s, and the maximum velocity of UAVs $V_{\mathsf{max}} = 40$ (m/s). We see that the proposed design with ISA algorithm, optimized

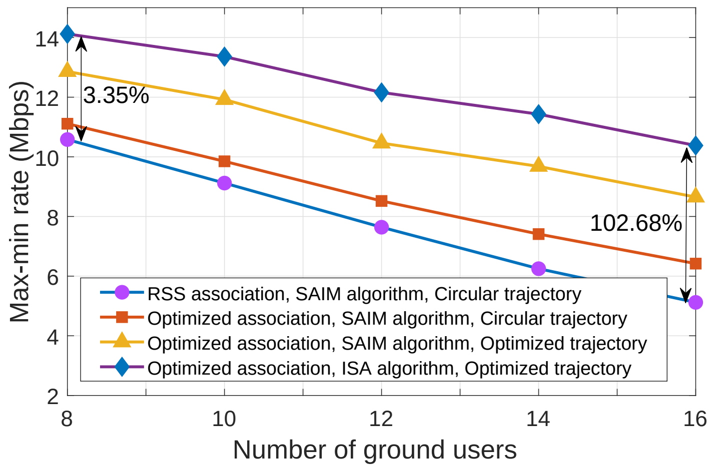

Figure 1.1 – Performance comparison of different schemes with 2 UAVs and $T = 20s$.

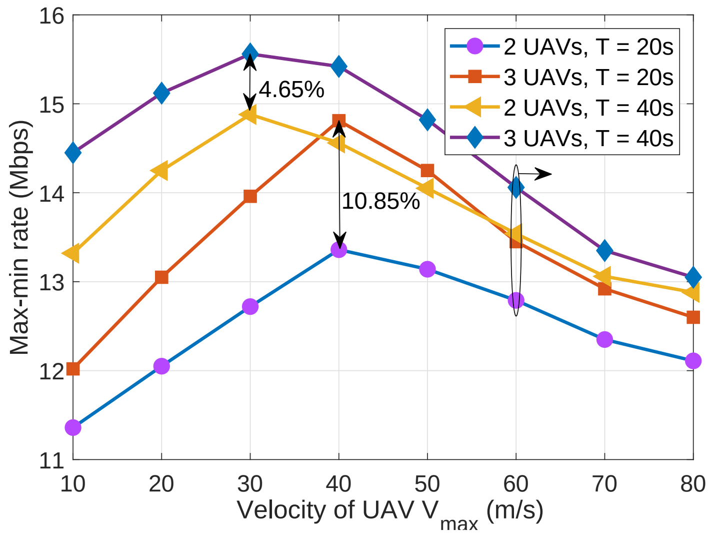

Figure 1.2 – Max-min rate under different velocity of UAV V_{max}.

UAV-GU association and trajectory control achieves the highest max-min average rate among the considered schemes. In addition, the rate gaps between the proposed ISA algorithm and other schemes increase when the number of GUs increases. For a given number of sub-channels, more sub-channels are likely to be reused by different UAVs to meet the GUs' data transmission demands when the number of GUs increases and this will likely lead to stronger co-channel interference. The results in Fig. 1.1 imply that the proposed ISA algorithm can effectively manage interference and resources.

We study the impact of the maximum UAV's velocity V_{max} on the max-min average rate in Fig. 1.2 for scenarios with 2 UAVs and 3 UAVs, 10 GUs, and 40 sub-channels where V_{max} varies in the range of $10-80$ (m/s). It can be seen that the peaks of the max-min average rate are achieved at the maximum UAV's velocity of 40 (m/s) and 30 (m/s) for $T = [20, 40]$s, respectively. Moreover, the rate gains at the peak rates for the 3-UAV setting versus the 2-UAV setting are 4.65% and 10.85% for $V_{\mathsf{max}} = [30, 40]$ (m/s) and $T = [20, 40]$s, respectively. However, this rate gain tends to decrease with the higher maximum velocity of UAVs. In fact, with the restricted network area of radius, $d_m^{\mathsf{init}}, (\forall m)$ given in Eq. (5.39), the velocity of UAVs strongly impacts the initial and the optimized trajectories of UAVs. This is because when the UAVs fly faster, the inter-UAV distances can become smaller in larger portions of the flight and the co-channel interference would be stronger, especially with a large number of UAVs. Specifically, the max-min average rate with $V_{\mathsf{max}} \geq 60$ (m/s) in the 3-UAV deployment and $T = 20$s is smaller than that in the 2-UAV scenario with $T = 40$s.

1.2.2 UAV Placement and Resource Allocation for Intelligent Reflecting Surface Assisted UAV-Based Wireless Networks

To the best of our knowledge, none of the existing work has studied the multi-carrier IRS-assisted UAV-based wireless network taking into account the constrained capacity of wireless backhauls. To fill this research gap, we study the joint optimization of UAV placement, IRS phase shifts, and sub-channel assignments for wireless access and backhaul links where our design objective is to maximize the sum rate achieved by GUs.

To solve the underlying mixed integer nonlinear program (MINLP), we first derive the closed-form IRS phase shift solution and then optimize the sub-channel assignment and UAV placement in an iterative manner by using the alternating optimization method. The sets of sub-channels assigned for the access and backhaul links are iteratively updated to efficiently use the available bandwidth while maintaining the backhaul capacity constraint. Moreover, we use the successive convex approximation (SCA) technique to solve the UAV placement sub-problem. Numerical results are presented to study the impacts of different parameters on the achieved sum rate.

1.2.2.1 System Model

We consider downlink communications between a UAV and a set of GUs in an IRS-assisted wireless network with the backhaul link between the UAV and a base station (BS). We define $\mathcal{K}$ as the set of GUs, i.e., $\mathcal{K} = \{1, ..., K\}$, located on the ground at fixed horizontal coordinates $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$.

We assume that the UAV is placed at the altitude H with the horizontal coordinate $\mathbf{q} = (x^{\mathsf{d}}, y^{\mathsf{d}})$. The UAV acts as an airborne BS connected to the core network wirelessly through a cellular BS which is placed at the coordinate $\mathbf{r}^{\mathsf{b}} = (x^{\mathsf{b}}, y^{\mathsf{b}})$ and a fixed altitude H^{b}.

We assume that a single IRS is installed on the surface of a building wall at the altitude H^{i} and horizontal coordinate $\mathbf{w}^{\mathsf{i}} = (x^{\mathsf{i}}, y^{\mathsf{i}})$. The IRS is made up of $I_r \times I_c$ passive reflection elements units installed as a uniform planar array (UPA) with I_c and I_r elements on each column and each row, respectively. The distance between any two adjacent elements of the IRS is denoted by d. The phase shift matrix of the IRS is denoted by $\mathbf{\Phi} = \mathrm{diag}\left\{e^{j\phi_{1,1}}, \ldots, e^{j\phi_{i_r,i_c}}, \ldots, e^{j\phi_{I_r,I_c}}\right\} \in \mathbb{C}^{I_r \times I_c}$, where $\phi_{i_r,i_c} \in [0, 2\pi), \forall i_r = 1, \ldots, I_r$, and $i_c = 1, \ldots, I_c$.

We assume that orthogonal frequency-division multiple access (OFDMA) is useed for both wireless access and backhaul links where $\mathcal{C} = \{1, ..., C\}$ denotes the set of available sub-channels and the bandwidth of each sub-channel is W (Hz). Let $\psi^{\mathsf{A}}_{k,c}$ denote sub-channel assignment variables for the access links between the UAV and K GUs, where $\psi^{\mathsf{A}}_{k,c} = 1$, if sub-channel c is assigned for GU k and $\psi^{\mathsf{A}}_{k,c} = 0$, otherwise. Similary, we define $\psi^{\mathsf{B}}_{0,c}$ as sub-channel assignment variables for the backhaul link, where $\psi^{\mathsf{B}}_{0,c} = 1$, if sub-channel c is assigned for the backhaul link and $\psi^{\mathsf{B}}_{0,c} = 0$, otherwise.

We assume that all BS-UAV, UAV-IRS, and UAV-GU communication links are dominated by the LoS propagation while communications channels between the IRS and GUs experience Rayleigh channel fading due to blockages. Hence, the distances among BS, UAV, IRS, and GUs can be calculated based on their coordinates as $d^{\mathsf{BU}} = \sqrt{\|\mathbf{r}^{\mathsf{b}} - \mathbf{q}\|^2 + (H^{\mathsf{b}} - H)^2}$, $d^{\mathsf{UI}} = \sqrt{\|\mathbf{q} - \mathbf{w}^{\mathsf{i}}\|^2 + (H - H^{\mathsf{i}})^2}$, $d^{\mathsf{UG}}_k = \sqrt{\|\mathbf{q} - \mathbf{r}^{\mathsf{u}}_k\|^2 + H^2}, \forall k$, $d^{\mathsf{IG}}_k = \sqrt{\|\mathbf{w}^{\mathsf{i}} - \mathbf{r}^{\mathsf{u}}_k\|^2 + (H^{\mathsf{i}})^2}, \forall k$, corresponding to the distances from BS to UAV, UAV to IRS, UAV to GU k, and IRS to GU k, respectively.

As discussed in [43], the received signal at GU k due to the communications from the UAV is given by $y_k = \sqrt{p}\left((\mathbf{h}^{\mathsf{IG}}_k)^H \mathbf{\Phi} \mathbf{h}^{\mathsf{UI}} + h^{\mathsf{UG}}_k\right) x_k + n^{\mathsf{G}}$, where x_k represents the transmitted symbol from the UAV, which satisfies $\mathbb{E}(|x_k|^2) = 1$, and p denotes the transmit power of the UAV for GU k on each sub-channel, i.e., $p = P_{\max}/C$ assuming uniform power allocation where $P_{\max}$ is the total transmit power of UAV, and n^{G} denotes the additive white Gaussian noise (AWGN) at GU, with zero mean and variance σ^2. Also, let $h^{\mathsf{UG}}_k, \mathbf{h}^{\mathsf{UI}}$, and $\mathbf{h}^{\mathsf{IG}}_k$ denote the channel coefficients of the links between UAV and GU k, UAV and IRS, IRS and GU k, respectively, which are expressed as

$h_k^{\mathsf{UG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{UG}})^2}}, \forall k$, and

$$\mathbf{h}^{\mathsf{UI}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{UI}})^2}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}\right]^H \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}\right]^H, \tag{1.15}$$

$$\mathbf{h}_k^{\mathsf{IG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{IG}})^\kappa}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}\right]^H \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}\right]^H \times \alpha^{\mathsf{IG}}, \forall k, \tag{1.16}$$

where β_0 denotes the channel gain at the reference distance of 1 meter, κ is the path loss exponent, λ is the wavelength of the carrier wave, and α^{IG} is the random scattering components modeled by a circularly symmetric complex Gaussian random variable with zero mean and unit variance. In addition, $(\theta^{\mathsf{UI}}, \xi^{\mathsf{UI}})$ and $(\theta_k^{\mathsf{IG}}, \xi_k^{\mathsf{IG}})$ represent the vertical and horizontal angle-of-departures from the UAV to the IRS and from the IRS to GU k, respectively, which can be calculated from $\sin\theta^{\mathsf{UI}} = \frac{|H-H^{\mathsf{i}}|}{d^{\mathsf{UI}}}$, $\sin\xi^{\mathsf{UI}} = \frac{|x^{\mathsf{i}}-x^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\cos\xi^{\mathsf{UI}} = \frac{|y^{\mathsf{i}}-y^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\sin\theta_k^{\mathsf{IG}} = \frac{H^{\mathsf{i}}}{d_k^{\mathsf{IG}}}$, $\sin\xi_k^{\mathsf{IG}} = \frac{|x^{\mathsf{i}}-x_k^{\mathsf{u}}|}{\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}$, and $\cos\xi_k^{\mathsf{IG}} = \frac{|y^{\mathsf{i}}-y_k^{\mathsf{u}}|}{\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}$, $\forall k \in \mathcal{K}$.

As presented in [43], the achievable rate for GU k served by the UAV on sub-channel c can be expressed as

$$R_{k,c}^{\mathsf{A}} = \psi_{k,c}^{\mathsf{A}} W \log_2\left(1+\frac{p}{\sigma^2}\left|\frac{\sqrt{\beta_0}}{d_k^{\mathsf{UG}}}+\frac{\beta_0 f_k|\alpha^{\mathsf{IG}}|}{(d_k^{\mathsf{IG}})^{\kappa/2}d^{\mathsf{UI}}}\right|^2\right), \tag{1.17}$$

where $f_k = \sum_{i_c=1}^{I_c}\sum_{i_r=1}^{I_r} e^{j\left(F_k^{i_r,i_c}+\phi_{i_r,i_c}\right)}, \forall k$, and $F_k^{i_r,i_c} = -\frac{2\pi d}{\lambda}((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}})+(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}))-\arg(\alpha^{\mathsf{IG}})$.

Also, the achievable rate of the backhaul link on sub-channel c can be expressed as

$$R_{0,c}^{\mathsf{B}} = \psi_{0,c}^{\mathsf{B}} W \log_2\left(1+\frac{p_0\beta_0}{(d^{\mathsf{BU}})^2\sigma^2}\right), \tag{1.18}$$

where p_0 denotes the transmit power of the cellular BS.

Moreover, to maintain the good end-to-end performance, the total data rate of all access links from the UAV to all the GUs should not exceed the backhaul rate. This constraint can be described as $\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}} R^{\mathsf{B}}_{0,c}$.

Let $\mathbf{\Psi} = \{\psi^{\mathsf{A}}_{k,c}, \psi^{\mathsf{B}}_{0,c}, \forall k, c\}$, $\mathbf{\Phi}$, and $\mathbf{Q} = \{\mathbf{q}\}$ denote vectors of all decision variables for sub-channel assignment, IRS phase shifts, and UAV placement, respectively. We want to maximize the sum rate of all GUs by optimizing all variables $\mathbf{\Psi}$, $\mathbf{\Phi}$, and $\mathbf{Q}$. This design problem can be formulated as

$$\textbf{(P2):} \max_{\mathbf{\Psi},\mathbf{\Phi},\mathbf{Q}} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} R^{\mathsf{A}}_{k,c} \tag{1.19}$$

$$\text{s.t.} \sum_{c\in\mathcal{C}} \psi^{\mathsf{A}}_{k,c} \geq 1, \forall k, \tag{1.19a}$$

$$\sum_{k\in\mathcal{K}} \psi^{\mathsf{A}}_{k,c} + \psi^{\mathsf{B}}_{0,c} \leq 1, \forall c, \tag{1.19b}$$

$$\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}} R^{\mathsf{B}}_{0,c}, \tag{1.19c}$$

$$\psi^{\mathsf{A}}_{k,c}, \psi^{\mathsf{B}}_{0,c} \in \{0,1\}, \forall k, c, \tag{1.19d}$$

$$\phi_{i_r,i_c} \in [0, 2\pi), \forall i_r = 1, \ldots, I_r; \forall i_c = 1, \ldots, I_c. \tag{1.19e}$$

Because of the non-convex constraint (1.19c) and integer variables in (1.19d), problem (1.19) is a non-convex mixed integer nonlinear optimization program (MINLP), which is difficult to solve. One might argue that adding constraint (1.19c) is a trivial modification of the previous models. While this is certainly true as far as writing the mathematical model, this constraint is not convex and thus makes the design of an efficient solution algorithm much more complicated. In the following, we describe the details of our proposed algorithm.

1.2.2.2 Proposed Algorithm

To solve problem **(P2)** we first derive the closed-form phase shift solution and then optimize the sub-channel assignment and UAV placement iteratively. Let $\mathcal{C}^{\mathsf{A}}$ and $\mathcal{C}^{\mathsf{B}}$ be the sets of sub-channels assigned for access and backhaul links, respectively where $\mathcal{C} = \mathcal{C}^{\mathsf{A}} \cup \mathcal{C}^{\mathsf{B}}$. Initially, the number of sub-channels allocated in $\mathcal{C}^{\mathsf{A}}$ is equal to the number of GUs K to ensure each GU is assigned at least one sub-channel and all remaining sub-channels are allocated to $\mathcal{C}^{\mathsf{B}}$. Then, the sets $\mathcal{C}^{\mathsf{A}}$ and $\mathcal{C}^{\mathsf{B}}$

are updated by taking sub-channels from $\mathcal{C}^{\mathsf{B}}$ and re-allocating to $\mathcal{C}^{\mathsf{A}}$ and the IRS phase shifts and UAV placement are optimized accordingly while maintaining the backhaul capacity constraint.

To obtain the maximum access rate $R^{\mathsf{A}}_{k,c}$ given in (1.17) and hence the sum rate, i.e., the objective function, the IRS phase shift $\mathbf{\Phi}^*$ must be aligned with the phases of channel coefficients. Such optimal IRS phase shifts, which result in $f_k^* = I_c I_r$, can be expressed as

$$\begin{aligned}\phi^*_{i_r,i_c} \quad = \quad & \frac{2\pi d}{\lambda}\Big((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}) \\ & +(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}})\Big) + \arg(\alpha^{\mathsf{IG}}). \qquad (1.20)\end{aligned}$$

Substituting this IRS phase shifts into problem **(P2)** still results in a non-convex MINLP problem. Thus, we use the alternating optimization approach to tackle this problem where we iteratively optimize each set of optimization variables given the values of other variables until convergence.

a) Optimization of Sub-channel Assignment

For given $\mathbf{\Phi}$ and $\mathbf{Q}$, the sub-problem to optimize the sub-channel assignment $\mathbf{\Psi}$ can be stated as

$$\textbf{(P2.1):} \max_{\mathbf{\Psi}} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{A}}_{k,c} \qquad (1.21)$$

$$\text{s.t.} \quad \sum_{c\in\mathcal{C}^{\mathsf{A}}} \psi^{\mathsf{A}}_{k,c} \geq 1, \forall k, \qquad (1.21a)$$

$$\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}^{\mathsf{B}}} R^{\mathsf{B}}_{0,c}, \qquad (1.21b)$$

$$\text{constraints} \;\; (1.19b), (1.19d).$$

This is a standard mixed integer linear program (MILP), which can be solved efficiently by using the CVX-Mosek solver [38].

b) Optimization of UAV Placement

For given $\mathbf{\Psi}$ and $\mathbf{\Phi}$, the sub-problem to optimize the UAV placement $\mathbf{Q}$ is non-convex. To solve this problem, we first introduce some auxiliary variables and then solve the transformed problem by using the SCA method. Specifically, we introduce variables $\nu_k \geq \|\mathbf{q}-\mathbf{r}^{\mathsf{u}}_k\|^2 + H^2, \forall k$,

$\mu \geq \|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2+(H-H^{\mathsf{i}})^2$, and $\epsilon \geq \left\|\mathbf{r}^{\mathsf{b}}-\mathbf{q}\right\|^2+(H^{\mathsf{b}}-H)^2$. From (1.17) and (1.18), we have

$$R_{k,c}^{\mathsf{Aq}} = \psi_{k,c}^{\mathsf{A}} W \log_2\left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k}+\frac{Y_k^2}{\mu}+\frac{2X_kY_k}{\nu_k^{1/2}\mu^{1/2}}\right)\right) \leq R_{k,c}^{\mathsf{A}}, \tag{1.22}$$

$$R_{0,c}^{\mathsf{Bq}} = \psi_{0,c}^{\mathsf{B}} W \log_2\left(1+\frac{Z}{\epsilon}\right) \leq R_{0,c}^{\mathsf{B}}, \tag{1.23}$$

where $X_k = \sqrt{\beta_0}$, $Y_k = \beta_0 f_k^* |\alpha^{\mathsf{IG}}|(d_k^{\mathsf{IG}})^{-\kappa/2}$, $Z = p_0\beta_0/\sigma^2$, in which $f_k^* = I_c I_r$ is a solution given by the IRS phase shifts expressed in (1.20). It can be verified that $R_{k,c}^{\mathsf{Aq}}$ is a convex function with respect to ν_k and μ and it can be lower-bounded by its first-order Taylor expansion at r-th iteration in the approximation process as follows:

$$R_{k,c}^{\mathsf{Aq}} \geq \psi_{k,c}^{\mathsf{A}} W \log_2 D^r + \frac{L^r}{D^r}(\nu_k-\nu_k^r)+\frac{S^r}{D^r}(\mu-\mu^r) \triangleq R_{k,c}^{\mathsf{Aqlb}}, \tag{1.24}$$

where

$$\begin{aligned}
D^r &= \left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k^r}+\frac{Y_k^2}{\mu^r}+\frac{2X_kY_k}{\nu_k^{1/2,r}\mu^{1/2,r}}\right)\right), \\
L^r &= -\psi_{k,c}^{\mathsf{A}} W \log_2(e)\left(\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k^{2,r}}+\frac{X_kY_k}{\nu_k^{3/2,r}\mu^{1/2,r}}\right)\right), \\
S^r &= -\psi_{k,c}^{\mathsf{A}} W \log_2(e)\left(\frac{p}{\sigma^2}\left(\frac{Y_k^2}{\mu^{2,r}}+\frac{X_kY_k}{\nu_k^{1/2,r}\mu^{3/2,r}}\right)\right).
\end{aligned}$$

Similarly, since $R_{0,c}^{\mathsf{Bq}}$ is convex with respect to ϵ, by applying the first-order Taylor expansion at the given point ϵ^r, it can be lower-bounded as

$$R_{0,c}^{\mathsf{Bq}} \geq \psi_{0,c}^{\mathsf{B}} W \log_2\left(1+\frac{Z}{\epsilon^r}\right) - \psi_{0,c}^{\mathsf{B}} W \frac{\log_2(e)Z}{\epsilon^r(\epsilon^r+Z)}(\epsilon-\epsilon^r) \triangleq R_{0,c}^{\mathsf{Bqlb}}, \tag{1.25}$$

Moreover, the upper-bound of the access rate given in (1.17) can be expressed by introducing auxiliary variables $\alpha_k \leq \|\mathbf{q}-\mathbf{r}_k^{\mathsf{u}}\|^2+H^2, \forall k$, $\gamma \leq \|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2+(H-H^{\mathsf{i}})^2$ and we have

$$R_{k,c}^{\mathsf{Aub}} = \psi_{k,c}^{\mathsf{A}} W \log_2\left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\alpha_k}+\frac{Y_k^2}{\gamma}+\frac{2X_kY_k}{\alpha_k^{1/2}\gamma^{1/2}}\right)\right) \geq R_{k,c}^{\mathsf{A}}. \tag{1.26}$$

Algorithm 1.5. Joint Algorithm for Sub-channel Assignment, IRS Phase Shifts, and UAV Placement

Initialization: $\mathcal{C}^{\mathsf{A}}, \mathcal{C}^{\mathsf{B}}, \mathbf{Q}^0, \mathbf{\Phi}^0, \mathbf{\Psi}^0$, $S^* = 10^2, S = S_1 = 0$, $t = 0$;
repeat
    $S^* = S$ and $t = t + 1$;
    Take a sub-channel c from $\mathcal{C}^{\mathsf{B}}$; update $\mathcal{C}^{\mathsf{A}} = \mathcal{C}^{\mathsf{A}} \cup \{c\}$ and $\mathcal{C}^{\mathsf{B}} \setminus \{c\}$;
    repeat
        Given $\mathbf{\Phi}^{r,*}$ in (1.20);
        Solve **(P2.1)** iteratively until convergence to obtain $\mathbf{\Psi}^r$;
        Solve **(P2.2)** iteratively until convergence to obtain $\mathbf{Q}^r$;
    until Convergence
    if Obtain a feasible solution with maximum sum rate S_1 and $S<S_1$ **then**
        Update $S = S_1$ and $\{\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*\} = \{\mathbf{\Psi}^t, \mathbf{Q}^t, \mathbf{\Phi}^t\}$;
    else
        Update $\mathcal{C}^{\mathsf{B}} = \mathcal{C}^{\mathsf{B}} \cup \{c\}$ and $\mathcal{C}^{\mathsf{A}} \setminus \{c\}$;
        Update $\{\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*\} = \{\mathbf{\Psi}^{t-1}, \mathbf{Q}^{t-1}, \mathbf{\Phi}^{t-1}\}$;
    end if
until $|S - S^*| < 10^{-6}$
Return $\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*$;

Therefore, the UAV placement optimization problem can be approximated by

$$\textbf{(P2.2):} \quad \max_{\mathbf{Q}, \nu_k, \mu, \epsilon, \alpha_k, \gamma} \sum_{k \in \mathcal{K}} \sum_{c \in \mathcal{C}^{\mathsf{A}}} R_{k,c}^{\mathsf{Aqlb}} \tag{1.27}$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}} \sum_{c \in \mathcal{C}^{\mathsf{A}}} R_{k,c}^{\mathsf{Aub}} - \sum_{c \in \mathcal{C}^{\mathsf{B}}} R_{0,c}^{\mathsf{Bqlb}} \leq 0, \tag{1.27a}$$

$$\nu_k \geq \|\mathbf{q} - \mathbf{r}_k^{\mathsf{u}}\|^2 + H^2, \forall k; \ \mu \geq \left\|\mathbf{q} - \mathbf{w}^{\mathsf{i}}\right\|^2 + (H - H^{\mathsf{i}})^2, \tag{1.27b}$$

$$\alpha_k \leq \|\mathbf{q}^r - \mathbf{r}_k^{\mathsf{u}}\|^2 + 2\left(\mathbf{q}^r - \mathbf{r}_k^{\mathsf{u}}\right)^T \left(\mathbf{q} - \mathbf{q}^r\right) + H^2, \forall k, \tag{1.27c}$$

$$\gamma \leq \left\|\mathbf{q}^r - \mathbf{w}^{\mathsf{i}}\right\|^2 + 2\left(\mathbf{q}^r - \mathbf{w}^{\mathsf{i}}\right)^T \left(\mathbf{q} - \mathbf{q}^r\right) + (H - H^{\mathsf{i}})^2, \tag{1.27d}$$

$$\epsilon \geq \left\|\mathbf{r}^{\mathsf{b}} - \mathbf{q}\right\|^2 + (H^{\mathsf{b}} - H)^2. \tag{1.27e}$$

This is a convex problem, which can be solved efficiently by using the CVX-Mosek solver [38].

Solutions of these sub-problems are used in our proposed algorithm which is described in Algorithm 1.5.

1.2.2.3 Numerical Results

We consider a rectangular network area with size $1000 \times 1000 (\mathrm{m}^2)$. The altitude of the UAV is fixed at $H = 120$m and the BS is located at $(0, 0, 20)$m. In addition, the IRS is fixed at $(500, 500, 50)$m and GUs are placed inside circular clusters with a radius of $r_c = 200$m. We initially locate the UAV at the center of the GUs' cluster. The remaining parameters are set as $p_0 = 33$dBm, $P_{\mathsf{max}} = 30$dBm,

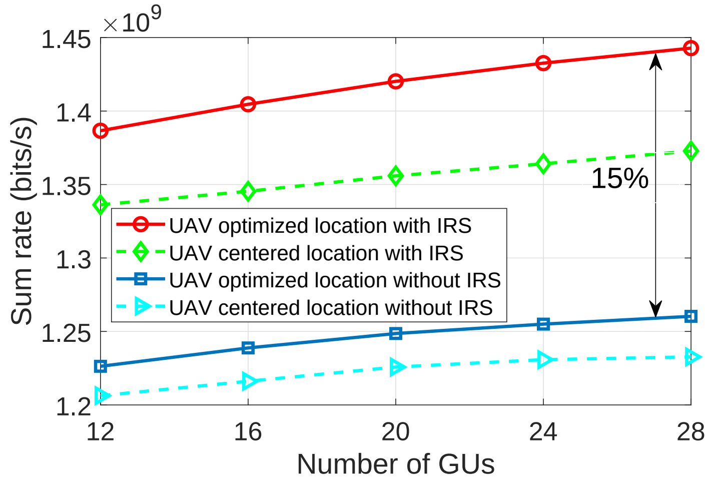

Figure 1.3 – Sum rate for different number of GUs.

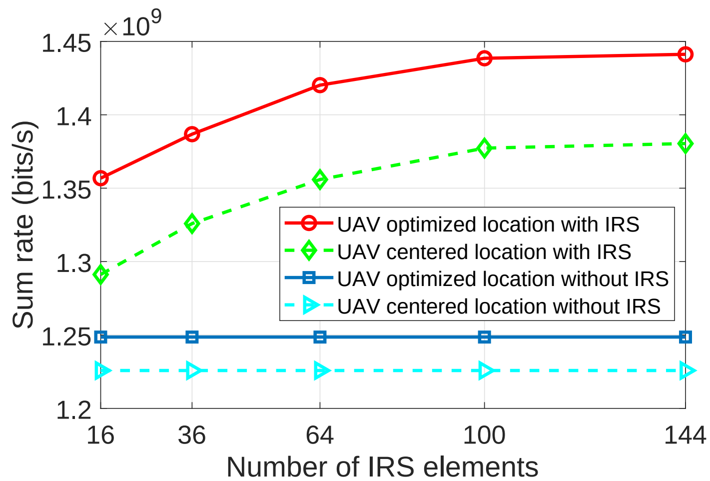

Figure 1.4 – Sum rate for different number of IRS elements.

$W = 1$MHz, $\sigma^2 = -110$dBm, $f_c = 2.5$GHz, $d = \lambda/2$, and $\kappa = 2$. Square IRSs with $I_r = I_c$ will be considered where the number of IRS elements is denoted by $I = I_r I_c$.

Fig. 1.3 – Fig. 1.4 show the sum rate achieved by the proposed algorithm, i.e., Alg. 1.5, and compared with the case where the UAV is placed at the cluster's center, which are indicated as "UAV optimized location" and "UAV centered location", respectively. Fig. 1.3 shows the sum rate for different number of GUs with $C = 60$ and $I = 64$. It can be seen that the sum rate slightly increases with increasing number of GUs and the difference in achieved sum rate between the optimized and centered location of UAV becomes larger as the number of GUs increases. The rate gain due to the proposed algorithm with and without leveraging the IRS is about 15%. Fig. 1.4 illustrates the sum rate for different number of IRS elements with 20 GUs and $C = 60$. In fact,

larger numbers of sub-channels or IRS elements lead to higher system diversity, which improves the achieved sum rate. These results confirm the effectiveness of the proposed algorithm in optimizing the UAV placement and sub-channel assignment in the IRS-assisted UAV communications.

1.2.3 Integrated Computation Offloading, UAV Trajectory Control, User Scheduling, Resource Allocation, and Admission Control in SAGIN

In this contribution, we investigate the integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control for SAGIN with multi-hop satellite communications. The main contributions can be summarized as follows:

- We study partial computation offloading in SAGIN where fractions of computation tasks from GUs are processed locally and/or offloaded and processed at the UAV-mounted edge servers and cloud server leveraging multi-hop LEO satellite communications. We formulate an optimization problem that aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of underlying computation tasks by jointly optimizing the user scheduling, partial offloading control and bit allocation over time, computation resource and bandwidth allocation, and UAV trajectory control.
- The alternating optimization approach is used to solve the underlying non-convex mixed-integer non-linear optimization problem (MINLP). Moreover, the successive convex approximation (SCA) method is used to solve the computation resource and bandwidth allocation and UAV trajectory control sub-problems.
- We propose efficient strategies for feasibility verification and admission control in the overloaded network scenarios. Specifically, an iterative algorithm is proposed to solve the feasibility verification problem and an efficient user removal strategy is developed for admission control while satisfying all GUs' and system constraints.
- Numerical results are presented to show the impacts of different parameters including the hop count in the multi-hop satellite communications, number of GUs, bandwidth, and computation task size on the achievable performance and the gains due to optimizing the UAV trajectory control, user scheduling, resource allocation, and computation offloading. Moreover, the admission ratio of GUs that are actually served in the different scenarios is presented.

1.2.3.1 System Model

We consider the computation offloading design in the SAGIN-based edge-cloud system as shown in Fig. 7.1, where the terrestrial network is made up of K GUs located on the ground, the aerial network layer employs M UAVs, and the space network layer relies on LEO satellites for connections to a distant cloud server. We denote the sets of satellites, UAVs, and GUs as $\mathcal{S} = \{1, ..., S\}$, $\mathcal{M} = \{1, ..., M\}$, and $\mathcal{K} = \{1, ..., K\}$, respectively.

We assume that partial computation offloading is used for a computation task of each GU. Specifically, each GU partitions its computation task into three sub-tasks where the first sub-task is processed locally and the other two sub-tasks are offloaded and processed at the UAV-mounted edge server and the cloud server, respectively. Moreover, the data related to the second sub-task must be transmitted from the associated GU to the connected UAV while the data related to the third sub-task must be transmitted from the GU to the cloud server via a multi-hop satellite communication path.

All GUs located on the ground at zero altitude are assumed to have fixed horizontal coordinates of $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. Besides, we assume that the UAVs fly at a fixed altitude H over a flight period of $T > 0$ seconds. We divide the flight period into N time slots where the set of time slots is denoted as $\mathcal{N} = \{1, ..., N\}$. Moreover, we assume that uplink communications from multiple GUs to their associated UAVs use the frequency division multiple access (FDMA). Specifically, let W denote the total bandwidth available to support uplink communications from GUs to UAVs. We assume that the available bandwidth is partitioned into orthogonal sub-bands each of which is allocated to one corresponding UAV to serve its associated GU. We denote the bandwidth allocated for UAV m as W_m^{u} then we have $\sum_{m \in \mathcal{M}} W_m^{\mathsf{u}} = W$. We also assume that the associations between GUs and UAVs and between GUs and satellites are fixed during the computation offloading process. Furthermore, we assume that the data size corresponding to the computation results is much smaller than that of the offloading data so that we can neglect the download time of the computation results in the offloading process. For ease of reference, the list of key notations is given in Table 7.2.

a) Computation Task Model

We assume that each GU k has one delay-constrained computation task represented by $U_k = (f_k, s_k, c_k, T_k^{\mathsf{max}})$, where f_k denotes the computation demand expressed by the number of central process unit (CPU) cycles per second (CPU cycles/second), s_k (bits) represents the size of input raw data, c_k (CPU cycles/bit) denotes the computation resource required for 1-bit input data, and T_k^{max} (seconds) describes the maximum tolerable latency of computation task U_k.

We assume that each GU's computation task is partitioned into three sub-tasks that are processed in parallel at the GU, the UAV-mounted edge server, and the cloud server reached via the multi-hop LEO satellite communication as considered in [44, 45]. Then, the task processing time for GU k can be expressed as

$$T_k = \max\left\{T_k^{\mathsf{lo}}, T_k^{\mathsf{ed}}, T_k^{\mathsf{cl}}\right\}, \tag{1.28}$$

where T_k^{lo}, T_k^{ed}, and T_k^{cl} represent the total data transmission and task execution time at the GU, UAV-mounted edge server, and cloud server, respectively. Specifically, T_k^{ed} includes both the data transmission time from GU k to the associated UAV and the execution time of the sub-task from GU k at the associated UAV. We will describe in more detail how to calculate this execution time later. Hence, the delay constraint for GU k can be expressed as $T_k \leq T_k^{\mathsf{max}}$.

To model the task partitioning for GU k, we introduce variables λ_k^{lo} and λ_k^{ed}, $(0 \leq \lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}} \leq 1)$ that represent the fractions of input data to be processed locally at GU k and to be offloaded and processed at the UAV-mounted edge server, respectively. Hence, $(1-\lambda_k^{\mathsf{lo}}-\lambda_k^{\mathsf{ed}})$ represents the fraction of input data from GU k to be offloaded and processed at the distant cloud server.

b) UAV Trajectory Control

The horizontal coordinates of UAV m in time slot n are denoted as $\mathbf{q}_m[n] = (x_m^{\mathsf{d}}[n], y_m^{\mathsf{d}}[n])$. We assume that each UAV must come back to its initial position at the end of the flight period, i.e., $\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m \in \mathcal{M}$. In addition, the slot interval $\Delta t = T/N$ is chosen to be sufficiently small so that the UAVs' locations are within a bounded small neighborhood in each time slot even at the maximum flight speed $V_{\max}$ in meter/second (m/s).

c) User Scheduling

Let $\phi_{k,m}^{\mathsf{u}}[n]$ denote binary decision variables for the association between the GUs and UAVs over flight period T, where $\phi_{k,m}^{\mathsf{u}}[n] = 1$ if GU k is served by UAV m in time slot n and $\phi_{k,m}^{\mathsf{u}}[n] = 0$, otherwise. The first requirement for the association is that each GU can offload its computation sub-task to at most one UAV in each time slot, i.e., $\sum_{m\in\mathcal{M}} \phi_{k,m}^{\mathsf{u}}[n] \leq 1$.

We assume that each GU k is initially associated with the UAV providing the highest average received signal strength (RSS), i.e., $\phi^{\mathsf{u}}_{k,m}[n] = 1$ with $m = \arg\max_{k}(RSS_{k,m}[n])$, where $RSS_{k,m}[n](\text{dBm}) = P^{\mathsf{u}}_k(\text{dBm}) - g_{k,m}[n](\text{dBm})$ with P^{u}_k denotes the transmit power of GU k to its associated UAV and $g_{k,m}[n]$ stands for channel power gain from GU k to UAV m. To satisfy the delay constraint of each GU, the number of *consecutive time slots* required to completely process the computation task of GU k can be denoted as $N_k = \lceil T^{\mathsf{max}}_k/\Delta t \rceil$, where $\lceil . \rceil$ denotes the round-up operation. We now introduce binary user scheduling variables $\theta_k[n]$, where $\theta_k[n] = 1$ if GU k is scheduled to transmit to its associated UAV in time slot n and $\theta_k[n] = 0$, otherwise. We need to impose the following constraints on the user scheduling decisions:

$$\sum_{m\in\mathcal{M}} \sum_{t=0}^{N_k-1} \theta_k[n+t]\phi^{\mathsf{u}}_{k,m}[n+t] = N_k, \forall k, n \in \{1, ..., N-N_k\}. \tag{1.29}$$

d) Computing Models

- **Local Computing Model**:

  The local task execution time at GU k can be expressed as

  $$T^{\mathsf{lo}}_k = \frac{\lambda^{\mathsf{lo}}_k s_k c_k}{f_k}. \tag{1.30}$$

  The delay constraint imposed to the local processing can be expressed as $T^{\mathsf{lo}}_k \leq T^{\mathsf{max}}_k$. The energy consumption due to local task execution can be calculated as

  $$E^{\mathsf{lo}}_k = \kappa\lambda^{\mathsf{lo}}_k s_k c_k (f_k)^2, \tag{1.31}$$

  where κ is the effective switched capacitance depending on the chip architecture [46].

- **UAV-Mounted Edge Computing Model**:

  For the partitioned sub-tasks offloaded to the UAVs, let $l^{\mathsf{u}}_k[n]$ denote the number of offloading bits from GU k to the associated UAV over time slot n. Besides, let us denote the computing resource of UAV m allocated to handle the sub-task offloaded from GU k in time slot n by $f^{\mathsf{u}}_k[n]$ (CPU cycles/second).

Hence, the total energy consumption at the associated UAVs to process the offloading sub-task from GU k can be calculated as

$$E_k^{\mathsf{ed}} = \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\big(l_k^{\mathsf{u}}[n]c_k e^{\mathsf{ed}}\big). \tag{1.32}$$

In addition, we assume that the communication links from the GUs to UAVs are dominated by the line-of-sight (LoS) propagation where the channel quality is mostly dependent on the UAV-GU distance. The distance between GU k and UAV m in time slot n can be calculated as $d_{k,m}[n]{=}\sqrt{H^2{+}\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$. Moreover, the channel power gain from GU k to UAV m in time slot n is assumed to follow the free-space path loss model, which can be expressed as $g_{k,m}[n]{=}\rho_0(d_{k,m}[n])^{-2}{=}\frac{\rho_0}{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$, where ρ_0 presents the channel power gain at the reference distance of 1 m. Hence, the achievable rate of the uplink transmission from GU k to the associated UAV m in time slot n, denoted by $R_{k,m}^{\mathsf{u}}[n]$ in bits/second (bps), can be expressed as

$$R_{k,m}^{\mathsf{u}}[n] = \beta_k^{\mathsf{u}}[n]\log_2\left(1+\frac{P_k^{\mathsf{u}}g_{k,m}[n]}{\beta_k^{\mathsf{u}}[n]\sigma^2}\right), \tag{1.33}$$

where $\beta_k^{\mathsf{u}}[n]$ and P_k^{u} represent the bandwidth allocated to GU k in time slot n and the transmit power of GU k for its uplink transmission, respectively, and σ^2 denotes the power density of the additive white Gaussian noise (AWGN) at the receiver.

Moreover, we assume that the partial task from GU k is offloaded and processed completely at each associated UAV in each time slot. Then we have following constraints

$$T_{k,m}^{\mathsf{ed}}[n]{=}\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}{+}\frac{l_k^{\mathsf{u}}[n]}{R_{k,m}^{\mathsf{u}}[n]}\right) \le \Delta t, \forall k,m,n. \tag{1.34}$$

Then, the total processing time at the UAVs to serve GU k can be written as

$$T_k^{\mathsf{ed}} = \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n]T_{k,m}^{\mathsf{ed}}[n]. \tag{1.35}$$

Furthermore, each UAV consumes some energy during its hovering time. Specifically, the flying energy consumption of UAV m can be expressed as $E_m^{\mathsf{edf}} = P_m^{\mathsf{f}}T$, where P_m^{f} denotes the flying power of UAV m.

- **Satellite Cloud Computing Model**:

We omit the processing time at the cloud server and we also ignore the cloud energy consumption involved in computation task execution and transmission of the computation results from the cloud server to GUs. In a recent work [2], an algorithm to determine the number of hops, i.e., the number of inter-satellite links (ISLs), and the corresponding satellites to establish the multi-hop communication path between two locations on the ground was proposed, i.e., see Algorithm 1 of [2]. By using this algorithm, the number of hops between the first and the last satellites connecting the considered terrestrial network area and the cloud server can be determined as L. Hence, the total data processing time and the propagation time from GU k to the cloud server can be calculated as

$$T_k^{\mathsf{cl}} = (1-\lambda_k^{\mathsf{lo}}-\lambda_k^{\mathsf{ed}})s_k\Big(\frac{1}{R_k^{\mathsf{s}}}+\sum_{i=1}^{L}(\frac{1}{R_i^{\mathsf{ss}}})+\frac{1}{R^{\mathsf{cl}}}\Big)+T_k^{\mathsf{prop}}, \tag{1.36}$$

where $R_k^{\mathsf{s}}, R_i^{\mathsf{ss}}, R^{\mathsf{cl}}$ stand for the transmission rates between the GU k and the first satellite, between the satellites in the i-th hop, and between the last satellite and the cloud server, respectively. Here, T_k^{prop} represents the total propagation delay from GU k to the first satellite, between satellites over the L ISLs, and from the last satellite to the cloud server. Moreover, the energy consumption of GU k for transmitting the data related to the offloaded sub-task to the first satellite can be calculated as

$$E_k^{\mathsf{s}} = \frac{(1-\lambda_k^{\mathsf{lo}}-\lambda_k^{\mathsf{ed}})s_k P_k^{\mathsf{s}}}{R_k^{\mathsf{s}}}, \tag{1.37}$$

where P_k^{s} represents the transmission power of GU k to the satellite.

1.2.3.2 Problem Formulation

In this work, we are interested in minimizing the weighted energy consumption of all GUs and UAVs for all involved computation tasks, which can be expressed as

$$\begin{aligned} E^{\mathsf{sum}} \;\; = \;\; & \alpha_1\Big(\sum_{k\in\mathcal{K}}E_k^{\mathsf{ed}}+\sum_{m\in\mathcal{M}}P_m^{\mathsf{f}}T\Big) \\ & +\alpha_2\sum_{k\in\mathcal{K}}\Big(E_k^{\mathsf{lo}}+\sum_{m\in\mathcal{M}}\sum_{n\in\mathcal{N}}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]E_{k,m}^{\mathsf{edt}}[n]+E_k^{\mathsf{s}}\Big), \end{aligned} \tag{1.38}$$

where $\alpha_1, \alpha_2 \in [0,1]$ represent the weight factors of the energy consumption of the UAVs and GUs, respectively, which strike to balance the energy consumption between the UAVs and GUs.

For convenience, we gather different decision variables and define the corresponding groups of variables as follows: user scheduling $\mathbf{\Theta} = \{\theta_k[n], \forall k, n\}$, partial offloading control $\mathbf{\Lambda} = \{\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, \forall k\}$, bit allocation $\mathbf{L} = \{l_k^{\mathsf{u}}[n], \forall k, n\}$, bandwidth allocation $\boldsymbol{\beta} = \{\beta_k^{\mathsf{u}}[n], \forall k, n\}$, computation resource allocation $\mathbf{F} = \{f_k^{\mathsf{u}}[n], \forall k, n\}$, and UAV trajectory control $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$. Our design aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of individual computation tasks. The optimization problem can be formulated as

$$\textbf{(P3):} \min_{\mathbf{\Theta},\mathbf{\Lambda},\mathbf{L},\boldsymbol{\beta},\mathbf{F},\mathbf{Q}} \quad E^{\mathsf{sum}} \tag{1.39}$$

$$\text{s.t.} \quad T_k^{\mathsf{lo}} \le T_k^{\mathsf{max}}, \forall k, \tag{1.39a}$$

$$T_k^{\mathsf{cl}} \le T_k^{\mathsf{max}}, \forall k, \tag{1.39b}$$

$$\sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] \le T_k^{\mathsf{max}}, \forall k, \tag{1.39c}$$

$$\theta_k[n] T_{k,m}^{\mathsf{ed}}[n] \le \Delta t, \forall k, m, n, \tag{1.39d}$$

$$\sum_{m\in\mathcal{M}} \sum_{t=0}^{N_k-1} \theta_k[n+t] \phi_{k,m}^{\mathsf{u}}[n+t] = N_k, \forall k, n \in \{1, ..., N-N_k\}, \tag{1.39e}$$

$$\sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] l_k^{\mathsf{u}}[n] = \lambda_k^{\mathsf{ed}} s_k, \forall k, \tag{1.39f}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] \beta_k^{\mathsf{u}}[n] \le W_m^{\mathsf{u}}, \forall m, n, \tag{1.39g}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] f_k^{\mathsf{u}}[n] \le F_m^{\mathsf{max}}, \forall m, n, \tag{1.39h}$$

$$\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m, \tag{1.39i}$$

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \le D_{\max}^2, \forall m, n = 1, ..., N-1, \tag{1.39j}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \ge d_{\min}^2, \ \forall n, m, j \ne m, \tag{1.39k}$$

$$\theta_k[n] \in \{0, 1\}, \forall k, n, \tag{1.39l}$$

$$0 \le \lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, 1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}} \le 1, \forall k, \tag{1.39m}$$

$$\beta_k^{\mathsf{u}}[n], f_k^{\mathsf{u}}[n], l_k^{\mathsf{u}}[n] \ge 0, \forall k, n, \tag{1.39n}$$

where constraints (1.39a)-(1.39d) capture the delay requirements for the GUs. Constraints (1.39e) and (1.39l) describe the binary user scheduling constraints for the GUs served by the associated

UAVs. Constraints (1.39g) capture the bandwidth allocation for transmission between the GUs and UAVs while constraints (1.39h) present the UAVs' computation constraints where F_m^{max} denotes the maximum computation resource of UAV m. It can be seen that the objective and constraint functions (1.39a)-(1.39d) are non-linear and integer decision variables are involved in (1.39l) for the user scheduling. Hence, problem (1.39) is a non-convex mixed integer non-linear optimization problem (MINLP), which is difficult to solve optimally.

1.2.3.3 Proposed Algorithm

In the section, we develop an algorithm to solve the formulated problem when it is feasible. Specifically, we adopt the alternating optimization approach to solve problem (1.39) where we iteratively optimize each set of variables given the values of other variables in the corresponding sub-problems until convergence. We describe how to solve different sub-problems in the following.

a) Optimization of User Scheduling

Given $\{\mathbf{L}, \mathbf{\Lambda}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, the user scheduling sub-problem to optimize $\mathbf{\Theta}$ can be formulated as

$$\begin{aligned} \textbf{(P3.1):} \min_{\mathbf{\Theta}} \;\; & E^{\mathsf{sum}} && (1.40) \\ & \text{s.t. constraints } (1.39\text{c}) - (1.39\text{h}), (1.39\text{l}). \end{aligned}$$

It can be verified that problem (1.40) is a standard mixed integer linear program (MILP), which can be solved efficiently by using the CVX-Gurobi solver [38].

b) Optimization of Partial Offloading Control and Bit Allocation Over Time Slots

Given $\{\mathbf{\Theta}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, the sub-problem optimizing the partial offloading control and bit allocation $\{\mathbf{\Lambda}, \mathbf{L}\}$ can be formulated as

$$\begin{aligned} \textbf{(P3.2):} \; \min_{\mathbf{\Lambda}, \mathbf{L}} \;\; & E^{\mathsf{sum}} && (1.41) \\ \text{s.t.} \;\; & (1.39\text{a}) - (1.39\text{d}), (1.39\text{f}), (1.39\text{m}), (1.39\text{n}). \end{aligned}$$

It can be verified that problem (1.41) is a linear problem (LP), it can be solved by using the CVX-Gurobi solver [38].

c) Optimization of Computation Resource and Bandwidth Allocation

Given $\{\mathbf{\Theta}, \mathbf{\Lambda}, \mathbf{L}, \mathbf{Q}\}$, the sub-problem optimizing the computation resource and bandwidth allocation $\{\mathbf{F}, \boldsymbol{\beta}\}$ can be stated as

$$\textbf{(P3.3):} \min_{\mathbf{F},\boldsymbol{\beta}} \ \alpha_2\bigg(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{R^{\mathsf{u}}_{k,m}[n]}\bigg)+E^{\mathsf{sum1}} \tag{1.42}$$

$$\text{s.t. constraints (1.39c), (1.39d), (1.39g), (1.39h), (1.39n),}$$

where

$$\begin{aligned} E^{\mathsf{sum1}} = \alpha_2\bigg(\sum_{k\in\mathcal{K}}(\kappa\lambda^{\mathsf{lo}}_k s_k c_k(f_k)^2+\frac{(1-\lambda^{\mathsf{lo}}_k-\lambda^{\mathsf{ed}}_k)s_kP^{\mathsf{s}}_k}{R^{\mathsf{s}}_k})\bigg)+ \\ \alpha_1\bigg(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n](l^{\mathsf{u}}_k[n]c_ke^{\mathsf{ed}})+\sum_m P^{\mathsf{f}}_mT\bigg). \end{aligned} \tag{1.43}$$

We first introduce auxiliary variables

$$\xi_{k,m}[n] = R^{\mathsf{u}}_{k,m}[n] = \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u}}_k[n]}\right), \tag{1.44}$$

where $B_{k,m}[n]=\frac{P^{\mathsf{u}}_kg_{k,m}[n]}{\sigma^2}$.

It can be verified that $\beta^{\mathsf{u}}_k[n]\log_2(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u}}_k[n]})$ is a concave function with respect to $\beta^{\mathsf{u}}_k[n]$. Using the successive convex approximation (SCA) method, the upper-bound for this concave function by using the first-order Taylor expansion at the given point $\beta^{\mathsf{u},r}_k[n]$ in the r-th iteration of the approximation process can be derived as

$$\begin{aligned} \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u}}_k[n]}\right) \le \beta^{\mathsf{u},r}_k[n]\log_2\left(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u},r}_k[n]}\right)+ \\ \left(\log_2(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u},r}_k[n]})-\frac{\log_2(e)B_{k,m}[n]}{B_{k,m}[n]+\beta^{\mathsf{u},r}_k[n]}\right)(\beta^{\mathsf{u}}_k[n]-\beta^{\mathsf{u},r}_k[n]) \triangleq R^{\mathsf{ub}}_{k,m}[n]. \end{aligned} \tag{1.45}$$

Then, problem (1.42) can be approximated by the following problem:

$$\begin{aligned}
\textbf{(P3.3.2):} \min_{\mathbf{F},\boldsymbol{\beta},\boldsymbol{\Xi}} \quad & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{\xi_{k,m}[n]}\Big) \\
& + E^{\mathsf{sum1}} && (1.46)\\
\text{s.t.} \quad & \sum_{m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\Big(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\xi_{k,m}[n]}\Big)\leq T^{\mathsf{max}}_k, \forall k, && (1.46a)\\
& \theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\Big(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\xi_{k,m}[n]}\Big) \leq \Delta t, \forall k,m,n, && (1.46b)\\
& \xi_{k,m}[n] \leq R^{\mathsf{ub}}_{k,m}[n], \forall k,m,n, && (1.46c)\\
& \text{constraints (1.39g), (1.39h), (1.39n),}
\end{aligned}$$

where $\boldsymbol{\Xi} = \{\xi_{k,m}[n], \forall k,m,n\}$.

Since $\frac{1}{f^{\mathsf{u}}_k[n]}$ and $\frac{1}{\xi_{k,m}[n]}$ are convex functions with respect to $f^{\mathsf{u}}_k[n]$ and $\xi_{k,m}[n]$, respectively, it can be seen that the objective function is convex and all constraints are linear. Hence, problem (1.46) is a convex problem, which can be solved effectively by using the CVX-Gurobi solver [38].

d) Optimization of Multi-UAV Trajectory

Given $\{\boldsymbol{\Theta},\boldsymbol{\Lambda},\mathbf{L},\mathbf{F},\boldsymbol{\beta}\}$, the sub-problem optimizing multi-UAV trajectory control variables $\mathbf{Q}$ can be formulated as

$$\begin{aligned}
\textbf{(P3.4):} \min_{\mathbf{Q}} \quad & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{R^{\mathsf{u}}_{k,m}[n]}\Big) + E^{\mathsf{sum1}} && (1.47)\\
\text{s.t.} \quad & \text{constraints (1.39c), (1.39d), (1.39i), (1.39j), (1.39k).}
\end{aligned}$$

To approximate this problem, we introduce auxiliary variables $\gamma_{k,m}[n] = R^{\mathsf{u}}_{k,m}[n]$ and $S_{k,m}[n] \leq H^2 + \|\mathbf{q}_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2$ and we have

$$\begin{aligned}
\gamma_{k,m}[n] \quad &= \quad \beta^{\mathsf{u}}_k[n]\log_2\Big(1+\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2(H^2+\|\mathbf{q}_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2)}\Big)\\
&\leq \quad \beta^{\mathsf{u}}_k[n]\log_2\Big(1+\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2 S_{k,m}[n]}\Big). && (1.48)
\end{aligned}$$

It can be verified that $\beta_k^{\mathsf{u}}[n]\log_2\left(1+\frac{R_k[n]}{S_{k,m}[n]}\right)$ is a convex function with respect to $S_{k,m}[n]$, where $R_k[n]{=}\frac{P_k^{\mathsf{u}}\rho_0}{\beta_k^{\mathsf{u}}[n]\sigma^2}$. By applying the SCA method, the lower-bound for the right hand side (RHS) of (1.48) derived by using the first-order Taylor expansion at the given point $S_{k,m}^r[n]$ in the r-th iteration of the approximation process can be expressed as

$$\beta_k^{\mathsf{u}}[n]\log_2\left(1{+}\frac{R_k[n]}{S_{k,m}[n]}\right) \geq \beta_k^{\mathsf{u}}[n]\Bigg(\log_2\left(S_{k,m}[n]{+}R_k[n]\right){-} \log_2(S_{k,m}^r[n]){-}\frac{\log_2(e)}{S_{k,m}^r[n]}(S_{k,m}[n]{-}S_{k,m}^r[n])\Bigg) \triangleq R_{k,m}^{\mathsf{lb}}[n]. \tag{1.49}$$

Therefore, the optimization problem (1.47) can be approximated by the following problem:

$$\textbf{(P3.4.2):} \min_{\mathbf{Q},\boldsymbol{\Gamma},\mathbf{S}} \ \alpha_2\Bigg(\sum_{k,m,n}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]l_k^{\mathsf{u}}[n]P_k^{\mathsf{u}}\frac{1}{\gamma_{k,m}[n]}\Bigg) + E^{\mathsf{sum1}} \tag{1.50}$$

$$\text{s.t.} \sum_{m,n}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}{+}\frac{l_k^{\mathsf{u}}[n]}{\gamma_{k,m}[n]}\right) \leq T_k^{\mathsf{max}}, \forall k, \tag{1.50a}$$

$$\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}{+}\frac{l_k^{\mathsf{u}}[n]}{\gamma_{k,m}[n]}\right) \leq \Delta t, \forall k,m,n, \tag{1.50b}$$

$$\gamma_{k,m}[n] \leq R_{k,m}^{\mathsf{lb}}[n], \forall k,m,n, \tag{1.50c}$$

$$S_{k,m}[n] \leq \|\mathbf{q}_m^r[n]{-}\mathbf{r}_k^{\mathsf{u}}\|^2 + 2\left(\mathbf{q}_m^r[n]{-}\mathbf{r}_k^{\mathsf{u}}\right)^T\left(\mathbf{q}_m[n]{-}\mathbf{q}_m^r[n]\right) + H^2, \forall k,m,n, \tag{1.50d}$$

$$d_{\min}^2 \leq -\left\|\mathbf{q}_m^r[n]{-}\mathbf{q}_j^r[n]\right\|^2 + 2\left(\mathbf{q}_m^r[n]{-}\mathbf{q}_j^r[n]\right)^T\left(\mathbf{q}_m[n]{-}\mathbf{q}_j[n]\right), \forall j{\neq}m, n, \tag{1.50e}$$

$$\text{constraints (1.39i), (1.39j),}$$

where $\boldsymbol{\Gamma}{=}\{\gamma_{k,m}[n], \forall k,m,n\}$, $\mathbf{S} = \{S_{k,m}[n], \forall k,m,n\}$.

Since $\frac{1}{\gamma_{k,m}[n]}$ is a convex function with respect to $\gamma_{k,m}[n]$, the objective function is convex. In addition, all constraints are linear. Therefore, problem (1.50) is a convex problem, which can be solved effectively by using the CVX-Gurobi solver [38].

e) Integrated User Scheduling, Partial Offloading Control, Computation Resource, Bandwidth Allocation, and Multi-UAV Trajectory Control Algorithm

Using the results above, we can develop an integrated algorithm based on the alternating optimization method as described in Algorithm 1.6. The convergence of this algorithm is stated in the following proposition.

Algorithm 1.6. Integrated User Scheduling, Partial Offloading, Computation, Bandwidth Allocation, and Multi-UAV Trajectory Control Algorithm

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, and locations of GUs, satellites and cloud server;
1: **Initialization:** $\mathbf{L}^0, \mathbf{\Lambda}^0, \mathbf{F}^0, \beta^0, \mathbf{Q}^0$;
Ensure: Min weighted energy consumption (E^{sum}); Let $r = 1$;
2: **repeat**
3: Solve sub-problem (1.40) to obtain $\mathbf{\Theta}^r$;
4: Solve sub-problem (1.41) to obtain $\mathbf{L}^r$ and $\mathbf{\Lambda}^r$;
5: Solve sub-problem (1.46) to obtain $\boldsymbol{\beta}^r$ and $\mathbf{F}^r$;
6: Solve sub-problem (1.50) to obtain $\mathbf{Q}^r$;
7: Update $r = r + 1$;
8: **until** Convergence
9: Return $E^{\mathsf{sum},*}, \mathbf{\Theta}^*, \mathbf{L}^*, \mathbf{\Lambda}^*, \mathbf{F}^*, \beta^*, \mathbf{Q}^*$.

Proposition 1.2. The proposed Algorithm 1.6 creates a sequence of feasible solutions where the objective value monotonically decreases over iterations. As a result, the algorithm converges to a feasible solution.

1.2.3.4 Joint Admission Control and Network Management Design

If problem **(P3)** is feasible then Algorithm 1.6 converges to a feasible solution. However, problem **(P3)** can be infeasible in certain overloaded scenarios. To this end, we develop an algorithm to verify the feasibility of problem **(P3)** and propose a joint admission control and network management algorithm to tackle problem **(P3)** in a generic scenario where this problem can be feasible or infeasible.

a) Feasibility Verification

We address the feasibility verification for problem **(P3)** in this section. We introduce a new variable δ and use it to all inequality constraints of problem **(P3)** and consider a related optimization problem aiming to minimize δ. This feasibility verification problem can be formulated as

$$
\begin{aligned}
\textbf{(P3')}: \quad & \min_{\boldsymbol{\Theta},\boldsymbol{\Lambda},\mathbf{L},\boldsymbol{\beta},\mathbf{F},\mathbf{Q},\delta} \ \delta && (1.51)\\
\text{s.t. } & T_k^{\mathsf{lo}} - T_k^{\mathsf{max}} - \delta \le 0, \forall k, && (1.51a)\\
& T_k^{\mathsf{cl}} - T_k^{\mathsf{max}} - \delta \le 0, \forall k, && (1.51b)\\
& \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] - T_k^{\mathsf{max}} - \delta \le 0, \forall k, && (1.51c)\\
& \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] - \Delta t - \delta \le 0, \forall k, m, n, && (1.51d)\\
& \sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] \beta_k^{\mathsf{u}}[n] - W_m^{\mathsf{u}} - \delta \le 0, \forall m, n, && (1.51e)\\
& \sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] f_k^{\mathsf{u}}[n] - F_m^{\mathsf{max}} - \delta \le 0, \forall m, n, && (1.51f)\\
& \|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 - D_{\max}^2 - \delta \le 0, \forall m, n = 1, ..., N-1, && (1.51g)\\
& d_{\min}^2 - \|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 - \delta \le 0, \ \forall n, m, j \ne m, && (1.51h)\\
& \text{constraints } (1.39e), (1.39f), (1.39i), (1.39l), (1.39m), (1.39n).
\end{aligned}
$$

Note that this problem is feasible and there exists an optimal value of δ that can be used to determine the feasibility of problem **(P3)** as follows. Specifically, problem **(P3)** is feasible, i.e., all constraints are satisfied, if $\delta \le 0$ and it is infeasible, otherwise. However, problem **(P3')** is also a mixed integer non-linear optimization problem (MINLP), which is difficult to solve optimally. Using the similar alternation optimization approach discussed in Section 1.2.3.3, after each sub-problem, we can obtain the values of δ that would be checked to verify the problem is feasible or not. The summary of the feasibility verification algorithm is described in the following.

- **Feasibility Verification Algorithm:**

  Summary of the feasibility verification algorithm is given in Algorithm 1.7. Initially, we set *feasibility* = *true*, and initialize all variables $\boldsymbol{\omega}^0 = \{\boldsymbol{\Theta}^0, \mathbf{L}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0\}$. Then, we apply the alternating optimization method and iteratively solve each set of variables given the values of other variables until convergence to a stable value of δ^* as described from step 1 to step 19. Specifically, after solving each sub-problem in the r-th iteration, we check the obtained objective value δ^r as follows: if $\delta^r > 0$, we return this value and $\boldsymbol{\omega}^r$ and then break the *"repeat-until"* loop. Otherwise, if $\delta^r \le 0$, we continue solving next

Algorithm 1.7. Feasibility Verification Algorithm

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, and locations of GUs, satellites and cloud server;
Ensure: Min δ; Let $r = 1$; *feasibility* = *true*, and $\boldsymbol{\omega}^0 = \{\boldsymbol{\Theta}^0, \mathbf{L}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \beta^0, \mathbf{Q}^0\}$;
1: **repeat**
2: Solve sub-problem (7.36) to obtain δ^r and $\boldsymbol{\Theta}^r$;
3: **if** $\delta^r > 0$ **then**
4: Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^{r-1}, \boldsymbol{\Lambda}^{r-1}, \mathbf{F}^{r-1}, \beta^{r-1}, \mathbf{Q}^{r-1}\}$,
Break **repeat** loop;
5: **end if**
6: Solve sub-problem (7.37) to obtain δ^r, $\mathbf{L}^r$, and $\boldsymbol{\Lambda}^r$;
7: **if** $\delta^r > 0$ **then**
8: Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^{r-1}, \beta^{r-1}, \mathbf{Q}^{r-1}\}$,
Break **repeat** loop;
9: **end if**
10: Solve sub-problem (7.38) to obtain δ^r, β^r, and $\mathbf{F}^r$;
11: **if** $\delta^r > 0$ **then**
12: Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \beta^r, \mathbf{Q}^{r-1}\}$,
Break **repeat** loop;
13: **end if**
14: Solve sub-problem (7.39) to obtain δ^r and $\mathbf{Q}^r$;
15: **if** $\delta^r > 0$ **then**
16: Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \beta^r, \mathbf{Q}^r\}$,
Break **repeat** loop;
17: **end if**
18: Update $r = r + 1$;
19: **until** Convergence δ^*
20: **if** $\delta^* \leq 0$ or $\delta^r \leq 0$ **then**
21: *feasibility* = *true*;
22: **else**
23: *feasibility* = *false*;
24: **end if**
25: **Output** *feasibility result* and

$$\boldsymbol{\omega}^* \in \Big\{ \{\boldsymbol{\Theta}^r, \mathbf{L}^{r-1}, \boldsymbol{\Lambda}^{r-1}, \mathbf{F}^{r-1}, \beta^{r-1}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^{r-1}, \beta^{r-1}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \beta^r, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \beta^r, \mathbf{Q}^r\} \Big\};$$

sub-problem. Steps 20 to 24 check the obtained value of δ^* and output the *feasibility result* in step 25. If *feasibility* = *true*, all constraints of problem **(P3)** are satisfied and we can solve the considered optimization problem to obtain a feasible solution. Otherwise, if *feasibility* = *false*, problem **(P3)** is infeasible, i.e., certain constraints of problem **(P3)** cannot be satisfied. The outputs of Algorithm 1.7 are *feasibility* results and $\boldsymbol{\omega}^*$.

b) Admission Control and Network Management Algorithm

For problem **(P3)**, the maximum delay constraints and the constraints requiring partial tasks from GUs be offloaded and processed completely at the associated UAVs in each time slot are challenging ones to satisfy. We propose a user removal strategy that iteratively removes in each removal step one "worst" GU that requires the largest amount of resource to satisfy its

Algorithm 1.8. Joint Admission Control and Network Management Algorithm

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, and locations of GUs, satellites and cloud server;
Initialization: $\mathbf{L}^0, \mathbf{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0, \mathcal{K}^{\mathsf{ac}} = \mathcal{K}$;
Ensure: Min weighted energy consumption (E^{sum});
Find number of hop-count L by running the Alg. 1 in [2];
Determine the bandwidth W_m^{u}, i.e., $\sum_{m \in \mathcal{M}} W_m^{\mathsf{u}} \leq W$;
Let *feasibility* = *true*;
repeat
    Run feasibility verification Algorithm 1.7;
    if *feasibility* = *true* **then**
        Run Algorithm 1.6
        Break **repeat** loop;
    else
        Given $\boldsymbol{\omega}^*$ obtained from Algorithm 1.7, calculate T_k based on (1.28);
        Find the worst GU $k = \text{argmax}_{k \in \mathcal{K}} \left(T_k / T_k^{\mathsf{max}} \right)$;
        Assign $\mathcal{K}^{\mathsf{ac}} = \mathcal{K} \setminus \{k\}$;
        Update $\mathcal{K} \leftarrow \mathcal{K}^{\mathsf{ac}}$;
    end if
until $\mathcal{K} = \emptyset$
Return $E^{\mathsf{sum},*}, \mathcal{K}^{\mathsf{ac}}, \mathbf{\Theta}^*, \mathbf{L}^*, \mathbf{\Lambda}^*, \mathbf{F}^*, \boldsymbol{\beta}^*, \mathbf{Q}^*$.

stringent delay constraint. Specifically, given the output of the Algorithm 1.7, the total data transmission, propagation, and task processing time for each GU could be calculated as in (1.28). Then, we find the "worst" GU k that achieves the maximum value of T_k/T_k^{max}, remove it and update the set of remaining GUs $\mathcal{K}^{\mathsf{ac}}$ accordingly, i.e., removing the identified GU k from the set $\mathcal{K}$. We propose a joint admission control and network management algorithm to solve problem **(P3)** to achieve the minimum weighted energy consumption of the GUs and UAVs as in Algorithm 1.8.

c) Algorithm Initialization

- **Initial Circular UAV Trajectory:**

  The circular UAVs trajectories to serve groups of GUs are considered and initialized as in section 7.6.4.1.

- **Initial Partial Offloading Control, Computation Resource, and Bandwidth Allocation Variables:**

  The task size values s_k are set randomly in range of $[1, 10]$Mbits and the values of the maximum tolerable delay T_k^{max} are also set randomly in range of $[1, 3]$(seconds). Moreover, the initial values of partial offloading control variables are randomly generated in $\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}} \in [0, 0.5]$, and a uniform allocation of bit, computation resource, and bandwidth is applied, i.e., $l_k^{\mathsf{u}}[n] = \lambda_k^{\mathsf{ed}} s_k / N_k$, $f_k^{\mathsf{u}}[n] = M F_m^{\mathsf{max}} / K$, and $\beta_k^{\mathsf{u}}[n] = W/K$, respectively.

To investigate the effectiveness of the proposed algorithms, we consider the following baselines. In an "early scheduling" baseline, all GUs are scheduled continuously from the first time slot of the UAV flight period. In the second baseline, called "baseline edge", we initially set circular UAVs' trajectories to serve the corresponding groups of GUs, the values of partial offloading control variables are randomly set and a uniform allocation of bit, computation resource, and bandwidth is applied as described above. For comparison, the "optimized edge" strategy represents our proposed design where all variables are optimized.

1.2.3.5 Numerical Results

We consider different scenarios in which a cloud server is far away from a considered network area. For particular, the group of GUs is located in Montreal (45.50°N, 73.56°W) while the cloud server is located in Vancouver (49.28°N, 123.12°W). By running the Alg. 1 in [2], we can determine the number of satellite hops $L = 4$. The parameters for our simulations are set similarly to those in [44, 46–48] and the chosen values of key parameters are summarized in Table 7.3.

Fig. 1.5 illustrates the weighted sum of energy for different number of GUs, 2 UAVs, $W = 10$ MHz, $L = 4$, and $T = [10, 15]$s. It can be seen that the weighted sum of energy becomes higher with larger number of GUs and the proposed algorithm achieves the smallest weighted sum of energy compared to those due to other baselines in both scenarios with $T = [10, 15]$s. For 18 GUs, the weighted sum of energy can be reduced by 18.05% and 9.64% compared to the corresponding values due to the "early scheduling" and "baseline edge" baselines with $T = 10$s and $T = 15$s, respectively. Fig. 1.6 illustrates the computation load distribution over network layers. This figure shows that larger task size values lead to less computation load distributed at the GUs while larger satellite hop counts result in higher computation load to be processed at the edge servers.

To evaluate the performance achieved by the proposed admission control design, we define an admission ratio as the ratio between the number of actual GUs served to the total number of GUs, i.e., $\frac{|\mathcal{K}^{\mathsf{ac}}|}{|\mathcal{K}_0|}$, where $\mathcal{K}_0$ denotes the set of original GUs and $\mathcal{K}^{\mathsf{ac}}$ represents the set of GUs admitted for which a feasible solution can be found by the proposed algorithm. Fig. 1.7 shows the admission ratio for different number of GUs for the networks with 2 and 3 UAVs, $W = 10$ MHz, $L = 4$, and $T = [10, 15]$s. It can be seen that the admission ratio decreases as the number of GUs increases. This

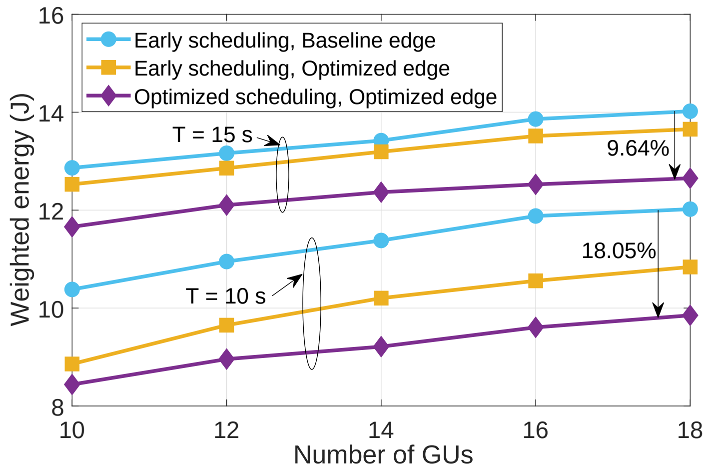

Figure 1.5 – Weighted sum of energy for different number of GUs.

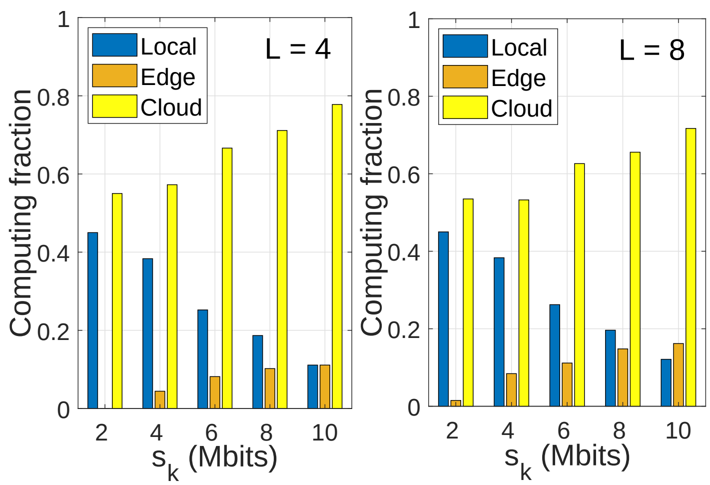

Figure 1.6 – Computation distribution for different task size values.

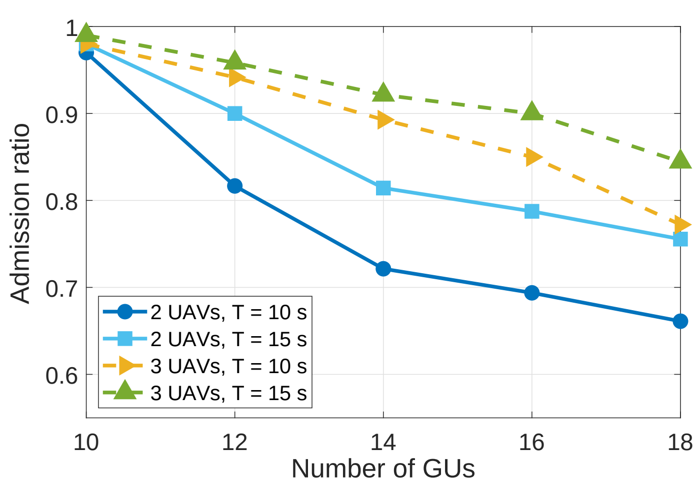

Figure 1.7 – Admission ratio for different number of GUs.

is because given the fixed radio and computation resources, the number of GUs that the network can support is limited. Hence, a larger number of GUs would be removed from the system as the number of GUs increases resulting in a decreasing admission ratio. It can also be seen that the difference in the admission ratios for the two scenarios with $T = 10$s and $T = 15$s is larger for 2 UAVs compared to that for 3 UAVs. In fact, for the network setting with a larger number of UAVs, i.e., edge servers, and larger UAV flight period T, the network can be covered better; therefore, a larger number of GUs can be served.

1.3 Concluding Remarks

In this doctoral dissertation, we have developed novel various network architectures and efficient resource allocation algorithms for UAV-based wireless networks. Specifically, we made three important research contributions. First, we study integrated UAV trajectory control and resource allocation for UAV-based wireless networks with co-channel interference management. Second, we consider UAV placement and resource allocation for intelligent reflecting surface assisted UAV-based wireless networks. Third, we study integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control in SAGIN with multi-hop LEO satellite wireless communications.

Chapter 2

Résumé Long

Ce chapitre est le résumée en français de la thèse intitulée:

"Planification de réseau et gestion de ressources pour réseaux sans fil basés sur UAV"

2.1 Contexte et motivation

Les réseaux sans fil de nouvelle génération permettront de prendre en charge des applications dans divers domaines, notamment les usines intelligentes, les transports intelligents, la santé en ligne, et plus encore [3, 4]. Les prolifération de nombreuses applications humaines et l'Internet des Objets (IoT) ont entraîné une explosion du trafic mobile. Par conséquent, les futures communications sans fil devront fournir une capacité plus élevée et une latence beaucoup plus faible, offrir une excellente stabilité, des communications omniprésentes et une connectivité à des milliards d'appareils [5–8]. Le déploiement de l'infrastructure terrestre, cependant, est confronté à des défis dans divers scénarios comme les communications pour répondre aux événements temporaires et aux urgences comme les catastrophes naturelles et la reprise rapide des services [9–11].

À cette fin, un certain nombre de technologies prometteuses ont été envisagées, notamment les communications par satellite, les communications par véhicules aériens sans pilote (UAV), la surface réfléchissante intelligente (IRS) et l'informatique mobile de pointe (MEC) [12, 13]. En particulier, les communications UAV sont apparues comme une solution potentielle pour surmonter les limites de l'infrastructure actuelle, offrant une couverture plus large, une résilience et une disponibilité plus élevées, et améliorant la qualité de service (QoS) de l'utilisateur en raison de leurs caractéristiques supérieures, comme la mobilité, la flexibilité et leur capacité à adapter leur altitude [14, 15]. En outre, le système de communication UAV assisté par IRS a attiré une attention considérable pour améliorer l'environnement de propagation et la qualité de la communication. Dans ce système, l'UAV communique avec les utilisateurs au sol (GU) tout au long de sa trajectoire, et l'IRS peut refléter les signaux émis par de l'UAV, ce qui améliore la grande flexibilité de l'UAV pour optimiser sa trajectoire [16–18]. De plus, les réseaux intégrés espace-air-sol (SAGIN) sont apparus comme une architecture prometteuse pour fournir des communications de haute qualité et omniprésentes en tirant parti des atouts complémentaires des réseaux spatiaux, aériens et terrestres et en permettant des technologies telles que l'informatique de pointe [19–21].

En premier lieu, il y a eu un vif intérêt à fournir une couverture sans fil entre des usagers au-dessus du sol (3D) et à tirer parti de différentes plates-formes de vol pour améliorer la connectivité sans fil et/ou les performances des réseaux sans fil terrestres [4, 10, 22, 23]. Les plates-formes de communication UAV peuvent fournir des solutions à faible coût pour divers scénarios de communication, par exemple, les zones sans fil blanches, avec une infrastructure limitée et une forte demande de trafic. Ainsi, les réseaux sans fil basés sur les UAV offrent des degrés de liberté supplémentaires pour optimiser le réseau sans fil sous-jacent afin d'améliorer la couverture, le débit et l'efficacité énergétique grâce aux attributs uniques des UAV tels que leur mobilité, leur flexibilité et le contrôle d'altitude.

En deuxième lieu, la surface réfléchissante intelligente (IRS) ou la surface intelligente reconfigurable (RIS) est un nouveau paradigme prometteur pour améliorer considérablement l'efficacité spectrale et énergétique des réseaux sans fil, en construisant des canaux de communication favorables via le réglage d'éléments réfléchissants passifs en grand quantité et à faible coût [24], où chaque élément peut être ajusté avec un déphasage indépendant pour refléter les signaux électromagnétiques incidents. Ceux-ci s'ajoutent alors de manière cohérente aux GU. L'IRS peut être déployé sur diverses structures, comme les façades de bâtiments, les panneaux d'affichage en bordure de route et les murs intérieurs [25]. Actuellement, les communications sans fil assistées par IRS utilisent généralement la surface entre la BS ou la BS aérienne et les utilisateurs mobiles pour améliorer la puissance du signal reçu [26–30].

Subsidiairement, les réseaux intégrés espace-air-sol (SAGIN) sont apparus comme un moyen efficace de fournir des communications de haute qualité et omniprésentes en tirant parti des atouts complémentaires des segments de réseaux spatiaux, aériens et terrestres [31–33]. D'une part, dans le réseau SAGIN, les satellites en orbite terrestre géostationnaire (GEO), les satellites en orbite terrestre moyenne (MEO) et les satellites en orbite terrestre basse (LEO) en sont les principaux composants [34]. Les satellites LEO sont susceptibles de former des réseaux par liaisons inter-satellites (ISL), qui garantissent un délai de propagation plus faible, des débits de communication élevés et des services de communication transparents pour de vastes zones géographiques [35, 36]. D'autre part, dans le réseau aérien de SAGIN, il existe un système mobile aérien aéroportés pour l'acquisition, la transmission et le traitement des informations. Les UAV, les dirigeables et les ballons sont les principales infrastructures composant les plates-formes haute et basse altitude (HAP & LAP) qui peuvent fournir des communications sans fil à large bande en complément des réseaux terrestres [31, 37]. Le réseau terrestre se compose alors principalement de systèmes de communication tels que le réseaux cellulaires, les réseaux mobiles ad hoc, les réseaux locaux sans fil, etc. [31]. Un système MEC assisté par UAV permet une prise en charge efficace des applications mobiles gourmandes en calcul grâce aux trajectoires contrôlables des UAV, et offre une couverture étendue et une capacité de calcul supplémentaire.

Dans cette thèse, notre objectif principal est d'étudier la planification du réseau et la gestion des ressources pour les réseaux sans fil basés sur des UAV. Dans des contextes plus précis, les résultats de cette thèse pourraient être utiles pour traiter certains problèmes de planification à long terme sur un horizon d'un an ou plus. Les problèmes d'implémentation en temps réel sortent du cadre de cette thèse. Les modèles sont des problèmes d'optimisation déterministes où toutes les données d'entrée sont connues, par exemple, les coordonnées des UAV et des GU. Les GU peuvent être vues comme des agrégats de sources de trafic sur une petite région. Les demandes de trafic sont aussi des moyennes de la demande sur l'horizon de planification. De plus, les modèles et conceptions proposés peuvent fournir des réponses à certaines questions pertinentes dans ce contexte. Par exemple, combien de l'UAV le fournisseur de réseau doit-il acheter, les UAV devraient-ils être fixes ou mobiles, combien de l'IRS faut-il en acheter et où devraient-ils être installés ou encore est-ce qu'une architecture cloud en vaut la peine, etc. En outre, les problèmes plus en temps réel tels que l'association GU-UAV, le contrôle de la trajectoire de l'UAV ou la fractionnement et le déchargement des calculs peuvent également être utilisés comme guides pour les algorithmes en temps réel. Spécifiquement, les contributions de cette thèse sont résumées dans la section suivante.

2.2 Contributions à la recherche

Dans cette thèse, notre objectif principal est de développer la planification du réseau et la gestion des ressources dans les communications UAV pour les futurs réseaux sans fil. En particulier, nos travaux portent sur trois aspects. Le premier est le contrôle intégré de la trajectoire des UAV et l'allocation des ressources pour les réseaux sans fil basés sur les UAV avec gestion des interférences dans le même canal. Dans le deuxième, nous étudions le placement de l'UAV et l'allocation de ressources pour les réseaux sans fil basés sur des UAV avec surface réfléchissante. Enfin, nous étudions le délestage des calculs, le contrôle de trajectoire UAV, la planification des utilisateurs, l'allocation des ressources et le contrôle d'admission dans SAGIN avec des communications par satellite LEO multi-sauts. Cette section présente un résumé des contributions de cette thèse.

2.2.1 Contrôle de trajectoire UAV intégré et allocation de ressources pour les réseaux sans fil basés sur UAV avec gestion des interférences intra-canal

Dans cette contribution, nous étudions l'association conjointe UAV-GU, l'allocation des ressources et le contrôle de la trajectoire des UAV pour les réseaux sans fil basés sur UAV avec réutilisation du spectre et gestion des interférences. Les principales contributions peuvent être résumées comme suit:

- Nous formulons l'association conjointe UAV-GU, le contrôle de trajectoire UAV et le problème d'attribution de sous-canal non orthogonal pour les réseaux sans fil basés sur UAV. Nous maximisons le débit moyen minimum de tous les GU en tenant compte des contraintes sur les demandes de transmission de données des GU individuels.
- Nous résolvons le problème sous-jacent d'optimisation non linéaire en nombres entiers mixte (MINLP) en utilisant l'approche d'optimisation alternée. Nous résolvons les sous-problèmes d'association UAV-GU, d'attribution de sous-canal et de contrôle de trajectoire UAV séparément à chaque itération jusqu'à convergence. Nous développons un algorithme itératif

d'affectation de sous-canal (ISA). Compte tenu de l'association UAV-GU et des solutions d'affectation des sous-canaux, le sous-problème de contrôle de trajectoire UAV est un problème non convexe difficile. Nous proposons d'utiliser la technique d'approximation convexe successive (SCA) pour convexifier et résoudre ce sous-problème. Nous présentons ensuite une courte analyse de la complexité de l'algorithme proposé.

- De nombreux résultats numériques sont présentés pour montrer les performances de notre algorithme. Spécifiquement, nous comparons les performances du réseau lorsque l'algorithme de sous-canal ISA proposé et un algorithme d'affectation de sous-canal heuristique de base avec gestion des interférences (SAIM) sont utilisés pour résoudre le problème commun. Nous étudions également les impacts de différents paramètres et l'importance du contrôle de trajectoire sur les performances atteignables. Enfin, nous illustrons la convergence de l'algorithme.

2.2.1.1 Modèle de système

Nous considérons un réseau où un ensemble d'UAV noté $\mathcal{M} = \{1, ..., M\}$, fournit une connectivité sans fil pour un ensemble de GU, noté $\mathcal{K} = \{1, ..., K\}$. Nous supposons que chaque GU doit recevoir une quantité spécifique de données des UAV sur la liaison descendante. Cela peut être le cas dans de nombreux scénarios pratiques, par exemple, les GU souhaitent recevoir des fichiers vidéo de l'UAV tels que des scènes spécifiques d'un match de football.

Étant donné que les UAV volent à une altitude relativement élevée, nous supposons que toutes les communications, qu'il s'agisse de l'UAV vers BS ou de l'UAV vers GU, sont dominées par la propagation en visibilité directe (LoS). Les UAV sont supposés être connectés au réseau central sans fil via une BS cellulaire où les liaisons UAV-BS sont supposées avoir une capacité suffisamment grande, c'est-à-dire en utilisant des communications mmWave. Nous supposons que les UAV volent à une altitude fixe H sur une période de vol de $T > 0$ secondes. La période de vol est divisée en tranches de temps N où l'ensemble de tranches de temps est noté $\mathcal{N} = \{1, ..., N\}$. Pendant tout créneau horaire pendant la période de vol T, chaque UAV peut communiquer avec plusieurs GU en même temps en utilisant l'OFDMA. Les GU sont supposés être situés au sol avec les coordonnées $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. De plus, la coordonnée de l'UAV m dans la tranche de temps n est notée $\mathbf{q}_m[n] = (x_m^{\mathsf{d}}[n], y_m^{\mathsf{d}}[n])$. Nous supposons que chaque UAV m doit revenir à sa position initiale à la fin de la période de vol, et que l'intervalle de créneaux $\Delta t = T/N$ est défini suffisamment petit pour que chaque UAV se déplace sur une petite distance pendant chaque créneau horaire même à la vitesse maximale $V_{\max}$.

Soit C le nombre de sous-canaux disponibles pour prendre en charge les liaisons d'accès sans fil entre les UAV et les GU. Nous désignons la puissance d'émission totale de chaque UAV par $P_{\max} \geq 0$. Nous supposons que l'allocation de puissance uniforme est utilisée par chaque UAV, c'est-à-dire que la puissance d'émission sur chaque sous-canal est égale à la puissance d'émission totale $P_{\max}$ divisée par le nombre total de sous-canaux utilisés pour les communications en liaison descendante et est donnée par $p = P_{\max}/C$. Nous définissons la variable binaire de décision d'association UAV-GU $\omega_{k,m}[n]$ qui est égale à 1 si le GU k est desservi par l'UAV m dans le créneau horaire n et égal à 0, sinon. En plus de l'affectation UAV, W (MHz) dénote la bande passante de chaque sous-canal et $\mathcal{C} = \{1, ..., C\}$ dénote l'ensemble de sous-canaux. En outre, nous devons décider de l'ensemble de sous-canaux à attribuer à chaque GU. Les variables d'affectation de sous-canal sont définies comme $\theta_{k,c}[n]$ qui sont égales à 1 si le sous-canal c est affecté à GU k dans le créneau horaire n et égales à 0, sinon.

Rappelons que nous avons supposé que les liaisons de communication des UAV aux GU sont dominées par la propagation LoS, où la qualité du canal dépend principalement de la distance UAV-GU. Dans le créneau horaire n, la distance entre UAV m et GU k peut être calculée comme $d_{k,m}[n] = \sqrt{H^2 + \left\|\mathbf{q}_m[n] - \mathbf{r}_k^{\mathsf{u}}\right\|^2}$. Le gain de puissance du canal de l'UAV m à GU k dans l'intervalle de temps n sur le sous-canal c suit un modèle de propagation en espace libre et peut être exprimé comme $g_{k,m}[n] = \rho_0 d_{k,m}^{-2}[n] = \frac{\rho_0}{H^2 + \left\|\mathbf{q}_m[n] - \mathbf{r}_k^{\mathsf{u}}\right\|^2}$, où ρ_0 présente le gain de puissance du canal à la distance de référence de 1 m. Le rapport signal/brouillage plus bruit reçu (SINR) à GU k sur le sous-canal c peut être calculé comme suit

$$\gamma_{k,m,c}[n] = \frac{p g_{k,m}[n]}{\sum_{j=1, j\neq m}^{M} \sum_{z=1, z\neq k}^{K} \omega_{z,j}[n] \theta_{z,c}[n] p g_{k,j}[n] + \sigma^2}, \tag{2.1}$$

où σ^2 est la puissance du bruit blanc gaussien additif (AWGN) au récepteur. Le terme $\sum_{j=1,j\neq m}^{M} \sum_{z=1,z\neq k}^{K} \omega_{z,j}[n] \theta_{z,c}[n] p g_{k,j}[n]$ représente l'interférence perçue par le GU k sur le sous-canal c en raison des transmissions d'autres UAV dans le créneau horaire n sur ce sous-canal. Le débit réalisable du GU k desservi par l'UAV m dans le créneau horaire n sur le sous-canal c, noté

$R_{k,m,c}[n]$ en bits/seconde (bps), peut alors être exprimé comme

$$R_{k,m,c}[n] = \omega_{k,m}[n]\theta_{k,c}[n]W \log_2 \left(1 + \gamma_{k,m,c}[n]\right). \tag{2.2}$$

Par conséquent, le débit total atteint par GU k dans l'intervalle de temps n, noté $R_k[n]$, peut être écrit comme

$$R_k[n] = \sum_{m=1}^{M} \sum_{c=1}^{C} R_{k,m,c}[n]. \tag{2.3}$$

En conséquence, le débit moyen par créneau du GU k sur N créneaux horaires peut être exprimé comme suit :

$$\bar{R}_k = \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{c=1}^{C} \omega_{k,m}[n]\theta_{k,c}[n]W \log_2 \left(1 + \gamma_{k,m,c}[n]\right). \tag{2.4}$$

2.2.1.2 Formulation du problème

Pour plus de commodité, nous rassemblons différentes variables dans les ensembles $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$, $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$ et $\mathbf{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$. Notre objectif de conception est de maximiser le débit moyen minimum atteint par toutes les unités GU en optimisant conjointement l'association d'utilisateurs, c'est-à-dire $\mathbf{\Omega}$, l'affectation des sous-canaux, c'est-à-dire $\mathbf{\Theta}$, et la trajectoire UAV, c'est-à-dire $\mathbf{Q}$ sur tous les créneaux horaires de la période de vol.

Le débit moyen $\bar{R}_k$ in (2.4) est une fonction non linéaire par rapport de trois variables de décision $\mathbf{\Omega}, \mathbf{\Theta}$ et $\mathbf{Q}$. Au lieu d'effectuer l'optimisation max-min de cette fonction non linéaire, nous introduisons la fonction $\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right) = \min_{k \in \mathcal{K}} \bar{R}_k$ comme le débit moyen minimum de tous les GU. Le problème d'optimisation devient équivalent à la maximisation $\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right)$, ce qui est plus maniable. De plus, nous supposons que chaque utilisateur k doit recevoir une quantité minimale de données D_k^{min} sur la liaison descendante pendant la période de vol de l'UAV. Ensuite, le problème conjoint d'association UAV-GU, d'attribution de sous-canal et d'optimisation du contrôle de trajectoire UAV pour maximiser le débit moyen minimum sur tous les GU peut être formulé comme suit

$$\textbf{(P1):} \quad \max_{\eta,\mathbf{\Omega},\mathbf{\Theta},\mathbf{Q}} \ \eta \tag{2.5}$$

$$\text{s.t. } \bar{R}_k \geq \eta, \ \forall k, \tag{2.5a}$$

$$\sum_{n=1}^{N} \Delta t R_k[n] \geq D_k^{\mathsf{min}}, \ \forall k, \tag{2.5b}$$

$$\|r_0 - \mathbf{q}_m[n]\| \leq R_0, \ \forall m, n, \tag{2.5c}$$

$$\sum_{m=1}^{M} \omega_{k,m}[n] = 1, \ \forall k, n, \tag{2.5d}$$

$$\sum_{k=1}^{K} \omega_{k,m}[n]\theta_{k,c}[n] \leq 1, \forall m, n, c, \tag{2.5e}$$

$$\sum_{c=1}^{C} \theta_{k,c}[n] \geq 1, \forall k, n, \tag{2.5f}$$

$$\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m, \tag{2.5g}$$

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \leq S_{\mathsf{max}}^2, \ n{=}1, ..., N{-}1, \tag{2.5h}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d_{\mathsf{min}}^2, \ \forall n, m, j \neq m, \tag{2.5i}$$

$$\omega_{k,m}[n] \in \{0, 1\}, \forall k, m, n, \tag{2.5j}$$

$$\theta_{k,c}[n] \in \{0, 1\}, \forall k, c, n, \tag{2.5k}$$

où R_0 représente le rayon de la zone du réseau centrée sur r_0. Les contraintes (2.5b) capturent la demande de transmission de données requise pour chaque GU sur la période de vol de T secondes, tandis que les contraintes (2.5c) restreignent les trajectoires de tous les UAV à l'intérieur de la zone de réseau souhaitée. De plus, les contraintes (2.5d)-(2.5e) représentent les contraintes d'association UAV-GU, (2.5e)-(2.5f), les contraintes sur l'affectation des sous-canaux, et (2.5g)-(2.5i), les contraintes sur les trajectoires des UAV. On peut voir que les contraintes (2.5a), (2.5b) et (2.5i) sont des variables de décision non linéaires entières et sont impliquées dans (2.5j) et (2.5k) pour l'association UAV-GU et l'attribution de sous-canal. Par conséquent, le problème (2.5) est un problème d'optimisation non linéaire en entier mixte (MINLP), difficile à résoudre de manière optimale.

2.2.1.3 Algorithmes proposés

Nous adoptons l'approche d'optimisation alternée pour résoudre le problème (2.5) où nous optimisons itérativement chaque ensemble de variables en fonction des valeurs des autres variables dans les sous-problèmes correspondants jusqu'à convergence. Nous décrivons comment résoudre ces différents sous-problèmes dans ce qui suit.

a) Association UAV-GU étant donné l'attribution de sous-canaux et le contrôle de trajectoire UAV

Pour l'affectation de sous-canal donnée $\mathbf{\Theta}$ et la trajectoire UAV $\mathbf{Q}$, le problème d'optimisation de l'association UAV-GU $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$ pour atteindre le débit moyen max-min sur tous les GU reste un problème d'optimisation non linéaire entier. Pour rendre le problème plus facile à résoudre, nous relaxons les variables de décision entières dans $\mathbf{\Omega}$ en variables de décision continues, ce qui donne le problème suivant

$$\textbf{(P1.1):} \max_{\eta, \mathbf{\Omega}} \ \eta \tag{2.6}$$

$$\text{s.t.} \quad 0 \leq \omega_{k,m}[n] \leq 1, \forall k, m, n, \tag{2.6a}$$

$$\text{contraintes } (2.5\text{a}), (2.5\text{b}), (2.5\text{d}), (2.5\text{e}).$$

Même avec cette relaxation, le problème (2.6) est toujours un problème d'optimisation non convexe en raison des contraintes non convexes (2.5a) et (2.5b). On peut récrire $R_{k,m,c}[n]$ dans les contraintes (2.5a) et (2.5b) sous la forme

$$\begin{aligned} R_{k,m,c}[n] &= \omega_{k,m}[n]\theta_{k,c}[n] W \log_2\left(1 + \frac{pg_{k,m}[n]}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n] + \sigma^2}\right) \\ &\geq \omega_{k,m}[n]\theta_{k,c}[n] W R^{\mathsf{A}}_{k,m,c}[n], \end{aligned} \tag{2.7}$$

où

$$R^{\mathsf{A}}_{k,m,c}[n] \leq \log_2\left(1 + \frac{pg_{k,m}[n]}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n] + \sigma^2}\right). \tag{2.8}$$

En introduisant des variables auxiliaires $\mathbf{R}^{\mathsf{A}} = \{R^{\mathsf{A}}_{k,m,c}[n], \forall k, m, c, n\}$, et basé sur le développement de Taylor du premier ordre aux points donnés $\omega^{r}_{k,m}[n]$ et $R^{\mathsf{A},r}_{k,m,c}[n]$ à la r-ième itération du processus d'approximation, on peut obtenir l'inégalité suivante

$$\omega_{k,m}[n]R^{\mathsf{A}}_{k,m,c}[n] \geq \frac{1}{4}\Big[-\Big(\omega^{r}_{k,m}[n]+R^{\mathsf{A},r}_{k,m,c}[n]\Big)^{2}+2\Big(\omega^{r}_{k,m}[n]+R^{\mathsf{A},r}_{k,m,c}[n]\Big)\Big(\omega_{k,m}[n]+R^{\mathsf{A}}_{k,m,c}[n]\Big)- \Big(\omega_{k,m}[n]-R^{\mathsf{A}}_{k,m,c}[n]\Big)^{2}\Big] \triangleq R^{\mathsf{Alb},r}_{k,m,c}[n]. \tag{2.9}$$

De plus, le membre de droite des contraintes (2.8) est convexe par rapport à $\omega_{z,j}[n]$. Ainsi, appliquant le développement de Taylor du premier ordre aux points donnés $\omega^{r}_{z,j}[n]$, on peut obtenir la borne inférieure $R^{\mathsf{AA},r}_{k,m,c}[n]$ comme dans (2.10).

$$\log_2\left(1+\frac{pg_{k,m}[n]}{\sum_{j\neq m}\sum_{z\neq k}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2}\right) \geq \log_2\left(1+\frac{pg_{k,m}[n]}{\sum_{j\neq m}\sum_{z\neq k}\omega^{r}_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2}\right)- \sum_{j\neq m}\sum_{z\neq k}A_{z,j,k,m,c}[n]\big(\omega_{z,j}[n]-\omega^{r}_{z,j}[n]\big) \triangleq R^{\mathsf{AA},r}_{k,m,c}[n], \tag{2.10}$$

où

$$A_{z,j,k,m,c}[n]=\frac{\omega^{r}_{z,j}[n]\theta_{z,c}[n]p^2g_{k,j}[n]g_{k,m}[n]\log_2(e)}{\Big(\sum_{j\neq m}\sum_{z\neq k}\omega^{r}_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2\Big)\Big(\sum_{j\neq m}\sum_{z\neq k}\omega^{r}_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2+pg_{k,m}[n]\Big)}.$$

Avec ces approximations, le problème (2.6) peut être approximé par le problème suivant :

$$\textbf{(P1.1”):} \max_{\eta^{r}_{\mathsf{a}},\mathbf{\Omega},\mathbf{R}^{\mathsf{A}}} \quad \eta^{r}_{\mathsf{a}} \tag{2.11}$$

$$\text{s.t.} \quad \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}R^{\mathsf{Alb},r}_{k,m,c}[n]\theta_{k,c}[n]W \geq \eta^{r}_{\mathsf{a}}, \forall k, \tag{2.11a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\Delta t R^{\mathsf{Alb},r}_{k,m,c}[n]\theta_{k,c}[n]W \geq D^{\min}_{k}, \forall k, \tag{2.11b}$$

$$R^{\mathsf{A}}_{k,m,c}[n] \leq R^{\mathsf{AA},r}_{k,m,c}[n], \forall k,m,c,n, \tag{2.11c}$$

$$\text{constraints (2.5d), (2.5e), (2.6a).}$$

On peut voir que toutes les contraintes sont linéaires. Par conséquent, le problème (2.11) est un problème d'optimisation convexe standard qui peut être résolu efficacement par n'importe quel outil d'optimisation convexe tel que CVX-Mosek [38]. Une description détaillée de notre algorithme proposé pour résoudre le problème d'association UAV-GU est donnée dans l'algorithme 2.1. Dans la solution obtenue par l'algorithme 2.1, si les variables d'association UAV-GU $\omega_{k,m}[n]$ sont toutes binaires, alors la relaxation est serrée et la solution obtenue est

Algorithm 2.1. Algorithme basé sur SCA pour résoudre (2.6)

Initialisation: Set $r := 0$, générer un point initial $\left(\mathbf{\Omega}^0, \mathbf{R}^{\mathsf{A},0}\right)$ de (2.11);
repeat
  $r := r + 1$;
  Résoudre (2.11) pour obtenir les valeurs optimales $\left(\mathbf{\Omega}^*, \mathbf{R}^{\mathsf{A},*}\right)$;
  Mise à jour $\left(\mathbf{\Omega}^r, \mathbf{R}^{\mathsf{A},r}\right) := \left(\mathbf{\Omega}^*, \mathbf{R}^{\mathsf{A},*}\right)$;
until Convergence
Production $\eta_{\mathsf{a}}^*, \mathbf{\Omega}^*, \mathbf{R}^{\mathsf{A},*}$.

aussi une solution réalisable du problème **(P1)**. Sinon, la solution d'association UAV-GU doit être récupérée en l'arrondissant à l'entier le plus proche de 0 ou 1. De plus, puisque les contraintes (2.5d) et (2.5e) sont rencontrées avec des égalités dans la solution de (2.11), une solution binaire peut être récupérée.

b) Attribution de sous-canal en fonction de l'association UAV-GU et de la trajectoire UAV

Etant données l'association UAV-GU et la trajectoire de l'UAV $\{\mathbf{\Omega}, \mathbf{Q}\}$, nous optimisons l'affectation de sous-canal $\mathbf{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$ pour atteindre le débit moyen max-min sur l'ensemble de tous les GU. Ce problème peut être exprimé comme suit:

$$\begin{aligned} \textbf{(P1.2):} \max_{\eta, \mathbf{\Theta}} \quad & \eta \\ \text{s.t.} \quad & \text{contraintes } (2.5\text{a}), (2.5\text{b}), (2.5\text{e}), (2.5\text{f}), (2.5\text{k}). \end{aligned} \tag{2.12}$$

Nous proposons un algorithme heuristique mais efficace pour les affectations de sous-canaux. Rappelons que notre objectif est de maximiser le débit moyen minimum parmi tous les GU et de satisfaire les demandes de transmission de données des GU individuels $D_k^{\mathsf{min}}, \forall k \in \mathcal{K}$. Par conséquent, dans la première phase, nous affectons affectons des sous-canaux à chaque GU, non seulement pour améliorer l'objectif, mais également pour garantir que les contraintes sur les demandes de transmission de données de tous les GU soient satisfaites. Spécifiquement, nous recherchons une affectation de sous-canal pour chaque GU k associée à UAV m dans un certain créneau horaire n pour obtenir une augmentation supérieure et maximale du débit moyen du GU k et garantir que le débit moyen minimum du système ne décroît pas à chaque pas d'affectation.

Une fois que les demandes de transmission de données requises de tous les GU sont satisfaites, l'algorithme entre dans une boucle d'attribution de sous-canal itérative où, à chaque itération,

Algorithm 2.2. Algorithme itératif d'affectation de sous-canaux (ISA)

Require: M UAV, K GU, C sous-canaux;
1: ***Donné*:** Association UAV-GU, contrôle de trajectoire UAV ;
Ensure: Débit moyen max-min $(\bar{R}_k)$, η;
2: $k = 1$;
3: **while** $k \leq K$ **do**
4: **repeat**
5: Calculer le débit moyen minimum du système : $\text{minrate} = \min_{k \in \mathcal{K}}\{\bar{R}_k\}$;
6: Étant donné GU k, identifier toutes les paires UAV et créneaux horaires $\{m, n\}$ avec $\{\omega_{k,m}[n] = 1\}$;
7: Étant donné GU k et chaque paire $\{m, n\}$ identifiée à l'étape 6, trouver le sous-canal c pour l'affectation afin d'obtenir le débit moyen le plus élevé et le meilleur pour GU k ;
8: Comparer toutes les affectations de sous-canal potentielles pour différentes paires $\{m, n\}$ trouvées à l'étape 7, choisir la meilleure affectation de sous-canal si elle peut améliorer le débit moyen minimum du système (c'est-à-dire que nous calculons $\text{rate} = \min_{k \in \mathcal{K}}\{\bar{R}_k\}$ et la nouvelle affectation de sous-canal doit satisfaire minrate < rate);
9: **until** $\sum_n D_k[n] \geq D_k^{\mathsf{min}}$
10: $k \leftarrow k + 1$;
11: **end while**
12: **repeat**
13: Trouver GU $k = \text{argmin}_{k \in \mathcal{K}}\{\bar{R}_k\}$;
14: Calculer le débit moyen minimum du système : $\text{minrate}^* = \min_{k \in \mathcal{K}}\{\bar{R}_k\}$;
15: Étant donné GU k, identifier toutes les paires UAV et créneaux horaires $\{m, n\}$ avec $\{\omega_{k,m}[n] = 1\}$;
16: Étant donné GU k et chaque paire $\{m, n\}$ identifiée à l'étape 15, trouver le sous-canal c pour l'affectation afin d'obtenir le débit moyen le plus élevé et le meilleur pour GU k ;
17: Comparer toutes les affectations de sous-canal potentielles pour différentes paires $\{m, n\}$ trouvées à l'étape 16, choisir la meilleure affectation de sous-canal si elle peut améliorer le débit moyen minimum du système (c'est-à-dire que nous calculons $\text{rate} = \min_{k \in \mathcal{K}}\{\bar{R}_k\}$ et la nouvelle affectation de sous-canal doit satisfaire $\text{minrate}^* < \text{rate}$);
18: Mise à jour $\text{minrate}^* = \text{rate}$;
19: **until** Convergence
20: Mise à jour $\eta^* \leftarrow \text{minrate}^*$;
21: Retour η^*, Θ^*.

il recherche le GU avec le débit moyen minimum et trouve la meilleure attribution de sous-canal obtenant la moyenne la plus élevée et le meilleur débit pour l'GU sous-jacente tout en améliorant le débit moyen minimum du système. En fait, la méthode pour déterminer la meilleure solution d'affectation de sous-canal dans cette boucle est similaire à celle de la phase précédente. L'algorithme se termine lorsque le débit moyen minimum de tous les GU ne peut plus être amélioré. Les détails de l'algorithme proposé appelé *"Algorithme itératif d'affectation de sous-canaux (ISA)"* sont donnés dans l'algorithme 2.2.

c) Contrôle de trajectoire UAV étant donné l'association UAV-GU et l'attribution de sous-canal

Étant donné l'association UAV-GU et l'affectation de sous-canal $\{\mathbf{\Omega}, \mathbf{\Theta}\}$, le problème d'optimisation du contrôle de trajectoire UAV $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$ pour atteindre le débit moyen max-min sur tous les GU peut s'écrire comme suit :

$$\textbf{(P1.3):} \max_{\eta,\mathbf{Q}} \quad \eta \tag{2.13}$$
$$\text{s.t.} \quad \text{contraintes (2.5a), (2.5b), (2.5c), (2.5g), (2.5h), (2.5i).}$$

Ce problème est un problème d'optimisation non convexe dû aux contraintes non convexes (2.5a), (2.5b) et (2.5i). Il est donc difficile de résoudre ce problème de manière optimale. Nous concevons un algorithme avec trois étapes principales pour résoudre ce problème. À l'étape 1, nous introduisons quelques variables auxiliaires et transformons le problème (2.13) en une forme équivalente. Ensuite, nous convexifions approximativement le problème correspondant à l'étape 2. Enfin, nous utilisons un outil d'optimisation convexe pour résoudre le problème convexe obtenu à l'étape 3. Par conséquent, le problème (2.13) peut être approximé comme

$$\textbf{(P1.3"):} \max_{\eta^r_{\mathsf{trj}},\mathbf{Q},\mathbf{S},\mathbf{R}} \quad \eta^r_{\mathsf{trj}} \tag{2.14}$$
$$\text{s.t.} \ \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]\right) \geq \eta^r_{\mathsf{trj}},\ \forall k, \tag{2.14a}$$
$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\Delta t\omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]\right) \geq D^{\mathsf{min}}_k,\ \forall k, \tag{2.14b}$$
$$S_{k,m}[n] \leq \|\mathbf{q}^r_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{r}^{\mathsf{u}}_k\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}^r_m[n]\right), \forall k,m,n, \tag{2.14c}$$
$$d^2_{\mathsf{min}} \leq -\left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right), \forall j\neq m,n, \tag{2.14d}$$
$$\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0 - R^{\mathsf{Ab}}_{z,j,c,k,m}[n]H^2 \geq R^{\mathsf{App},r}[n], \forall n, \tag{2.14e}$$
$$\text{(2.5c), (2.5g), (2.5h),}$$

où $\mathbf{R} = \{R^{\mathsf{Ab}}_{z,j,c,k,m}[n], \forall k,m,z,j,c,n\}$ et $\mathbf{S} = \{S_{k,m}[n], \forall k,m,n\}$.

Le problème (2.14) est un problème d'optimisation convexe standard qui peut être résolu efficacement par n'importe quel outil d'optimisation convexe tel que CVX-Mosek [38]. Une description détaillée est donnée dans l'algorithme 2.3.

d) Association UAV-GU intégrée, affectation de sous-canaux et contrôle de trajectoire UAV

Algorithm 2.3. Algorithme basé sur SCA pour résoudre (2.13)

Initialisation: Fixer $r := 0$, générer un point initial $\left(\mathbf{Q}^0, \mathbf{S}^0, \mathbf{R}^0\right)$ of (2.14);
repeat
    $r := r + 1$;
    Résoudre (2.14) pour obtenir les valeurs optimales $(\mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*)$;
    Mise à jour $(\mathbf{Q}^r, \mathbf{S}^r, \mathbf{R}^r) := (\mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*)$;
until Convergence
Production $\eta^*_{\mathsf{trj}}, \mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*$.

Algorithm 2.4. Association UAV-GU intégrée, affectation de sous-canaux et contrôle de trajectoire UAV

Require: M UAV, K GU, C sous-canaux et T;
Ensure: débit moyen max-min $(\bar{R}_k)$, η; Let $r = 1$;
repeat
    Optimiser l'association UAV-GU compte tenu de l'affectation des sous-canaux et des trajectoires des UAV en résolvant le sous-problème à l'aide de l'algorithme 2.1 pour obtenir $\mathbf{\Omega}^r$;
    Optimiser l'affectation des sous-canaux compte tenu de l'association UAV-GU et des trajectoires des UAV en résolvant le sous-problème à l'aide de l'algorithme 2.2 pour obtenir $\mathbf{\Theta}^r$;
    Optimiser les trajectoires des UAV compte tenu de l'association UAV-GU et de l'affectation des sous-canaux à l'aide de l'algorithme 2.3 pour obtenir $\mathbf{Q}^r$;
    Mise à jour $r = r + 1$;
until Convergence
Retour $\eta^*, \mathbf{\Omega}^*, \mathbf{\Theta}^*, \mathbf{Q}^*$.;

En utilisant les résultats présentés dans les sections **a)**, **b)** et **c)**, notre algorithme basé sur la méthode d'optimisation alternée est décrit dans l'algorithme 2.4. La convergence de cet algorithme est énoncée dans la proposition suivante.

Proposition 2.1. L'algorithme 2.4 crée une séquence de solutions réalisables où la valeur fonction objectif augmente de manière monotone au fil des itérations. En conséquence, l'algorithme converge vers une solution réalisable.

2.2.1.4 Résultats numériques

Dans cette section, nous évaluons les performances de l'algorithme proposé. Le paramétrage des simulations est similaire à celui de [39–41] et est résumé dans le tableau 5.3. Nous considérons une zone de réseau circulaire avec un rayon $R_0 = 500$m avec deux clusters ou plus de GU. Le rayon de chaque zone de cluster circulaire est de $r_{\mathsf{c}} = 200$m et les différents clusters sont suffisamment éloignés pour ne pas se chevaucher. La distance entre les centres de deux clusters voisins est définie pour satisfaire la contrainte suivante $D^{\mathsf{0}} \geq d_{\mathsf{min}} + 2 \times r_{\mathsf{c}}(\mathrm{m})$. L'altitude de tous les UAV est supposée être fixée à $H = 100$m. De plus, la demande de données de transmission requise pour chaque GU k

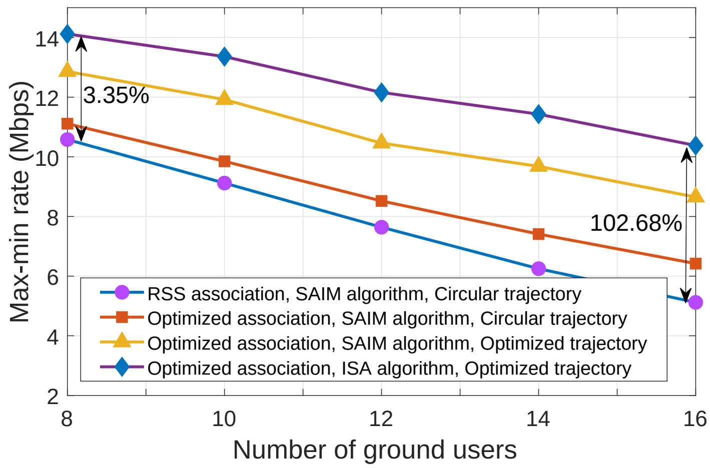

Figure 2.1 – Comparaison des performances des différents schémas avec 2 UAV et $T = 20s$.

(D_k^{min}) est définie en fonction de la taille de courtes transmissions vidéo, par exemple, des fichiers vidéo avec une résolution de 30 images par seconde (fps) [42]).

Nous évaluons d'abord les performances de l'algorithme ISA proposé décrit dans l'algorithme 2.2 et de l'algorithme SAIM illustré dans la figure 5.3. Spécifiquement, les débit moyens max-min dus à différentes méthodes de conception sont illustrés à la figure 2.1 pour le réseau avec 2 UAV, 40 sous-canaux, la période de vol de l'UAV $T = 20$s et la vitesse maximale des UAV $V_{\mathsf{max}} = 40$ (m/s). Nous voyons que la conception proposée avec l'algorithme ISA, l'association UAV-GU optimisée et le contrôle de trajectoire, atteint le débit moyen max-min le plus élevé parmi les méthodes de conception. De plus, les écarts de débit entre l'algorithme ISA proposé et d'autres méthodes de conception augmentent lorsque le nombre de GU augmente. Pour un nombre donné de sous-canaux, davantage de sous-canaux sont susceptibles d'être réutilisés par différents UAV pour répondre aux demandes de transmission de données des GU lorsque le nombre de GU augmente, ce qui entraînera probablement une interférence co-canal plus forte. Les résultats de la figure 2.1 impliquent que l'algorithme ISA proposé peut gérer efficacement les interférences et les ressources.

Nous étudions l'impact de la vitesse maximale de l'UAV V_{max} sur le débit moyen max-min dans la figure 2.2 pour les scénarios avec 2 UAV et 3 UAV, 10 GU et 40 sous-canaux où V_{max} varie entre 10 et 80 (m/s). On peut voir que les pics du débit moyen max-min sont atteints à la vitesse maximale de l'UAV de 40 (m/s) et 30 (m/s) pour $T = [20, 40]$s. De plus, les gains de débit aux débit de pointe pour la configuration avec 3 UAV par rapport à la configuration avec 2 UAV sont de 4.65 % et 10.85 % pour $V_{\mathsf{max}} = [30, 40]$ (m/s) et $T = [20, 40]$s. Cependant, ce gain de vitesse a

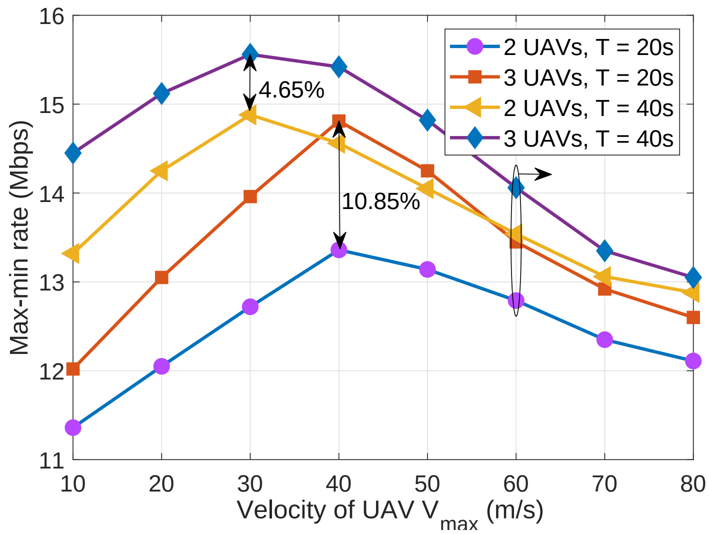

Figure 2.2 – Débit max-min sous différentes vitesses de l'UAV V_{max}.

tendance à diminuer avec la vitesse maximale plus élevée des UAV. En fait, avec la zone de réseau restreinte de rayon, $d_m^{\mathsf{init}}, (\forall m)$ donnée dans Eq. (5.39), la vélocité des UAV impacte fortement les trajectoires initiales et optimisées des UAV. En effet, lorsque les UAV volent plus vite, les distances inter-UAV peuvent devenir plus petites dans de plus grandes portions du vol et l'interférence dans le même canal serait plus forte, en particulier avec un grand nombre d'UAV. Plus précisément, le débit moyen max-min avec $V_{\mathsf{max}} \geq 60$ (m/s) dans le déploiement à 3 UAV et $T = 20$s est inférieur à celui dans le scénario à 2 UAV avec $T = 40$s.

2.2.2 Placement de l'UAV et allocation de ressources pour les réseaux sans fil basés sur des UAV avec surfaces réfléchissantes intelligentes

Au meilleur de notre connaissance, aucun des travaux existants n'a étudié les contraintes des réseaux sans fil assistés par IRS et basés sur des UAV multiporteurs. Pour combler cette lacune, nous étudions l'optimisation conjointe du placement des UAV, des déphasages IRS et des affectations de sous-canaux pour l'accès sans fil et les liaisons terrestres, où notre objectif est de maximiser débit total atteint par les GU.

Pour résoudre le programme non linéaire à entiers mixtes sous-jacent (MINLP), nous calculons une solution analytique au problème de déphasage IRS, puis nous optimisons l'affectation des sous-canaux et le placement des UAV de manière itérative en utilisant la méthode d'optimisation alternée. Les ensembles de sous-canaux attribués aux liaisons d'accès et de liaison sont mis à jour de manière itérative pour utiliser efficacement la bande passante disponible tout en maintenant la contrainte de capacité de liaison. De plus, nous utilisons la technique d'approximation convexe successive (SCA) pour résoudre le sous-problème de placement de l'UAV. Des résultats numériques sont présentés pour étudier les impacts de différents paramètres sur le débit total.

2.2.2.1 Modèle de système

Nous considérons les communications en liaison descendante entre un UAV et un ensemble du GU dans un réseau sans fil assisté par IRS avec une liaison entre l'UAV et une station de base (BS). Nous définissons $\mathcal{K}$ comme l'ensemble de GU, c'est-à-dire $\mathcal{K} = \{1, ..., K\}$, situés au sol aux points fixes $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. Nous supposons que l'UAV est placé à l'altitude H avec les coordonnées $\mathbf{q} = (x^{\mathsf{d}}, y^{\mathsf{d}})$. L'UAV agit comme une station de base aéroportée connectée au réseau central sans fil via une station de base cellulaire qui est placée au point $\mathbf{r}^{\mathsf{b}} = (x^{\mathsf{b}}, y^{\mathsf{b}})$ et une altitude fixe H^{b}.

Nous supposons qu'un seul IRS est installé sur la surface d'un mur de bâtiment à l'altitude H^{i} et au point $\mathbf{w}^{\mathsf{i}} = (x^{\mathsf{i}}, y^{\mathsf{i}})$. L'IRS est composé d'éléments de réflexion passive $I_r \times I_c$ installés en réseau planaire uniforme (UPA) avec des éléments I_c et I_r sur chaque colonne et chaque ligne. De plus, la distance entre deux éléments adjacents de l'IRS est notée d. La matrice de déphasage de l'IRS est $\mathbf{\Phi} = \mathrm{diag}\left\{e^{j\phi_{1,1}}, \ldots, e^{j\phi_{i_r,i_c}}, \ldots, e^{j\phi_{I_r,I_c}}\right\} \in \mathbb{C}^{I_r \times I_c}$, où $\phi_{i_r,i_c} \in [0, 2\pi), \forall i_r = 1, \ldots, I_r$, et $i_c = 1, \ldots, I_c$.

Nous supposons que l'accès multiple par répartition orthogonale de la fréquence (OFDMA) est utilisé à la fois pour l'accès sans fil et les liaisons terrestres où $\mathcal{C} = \{1, ..., C\}$ désigne l'ensemble des sous-canaux disponibles et la bande passante de chaque sous-canal est de W (Hz). Soit $\psi_{k,c}^{\mathsf{A}}$ les variables d'affectation de sous-canaux pour les liaisons d'accès entre l'UAV et les K GU, où $\psi_{k,c}^{\mathsf{A}} = 1$, si le sous-canal c est assigné au GU k et $\psi_{k,c}^{\mathsf{A}} = 0$, sinon. De même, nous définissons $\psi_{0,c}^{\mathsf{B}}$ comme variables d'affectation de sous-canal pour la liaison terrestre, où $\psi_{0,c}^{\mathsf{B}} = 1$, si le sous-canal c est affecté à la liaison backhaul et $\psi_{0,c}^{\mathsf{B}} = 0$, sinon.

Nous supposons que toutes les liaisons BS-UAV, UAV-IRS et UAV-GU sont dominées par la propagation LoS tandis que les canaux de communication entre l'IRS et les GU subissent un évanouissement de canal de type Rayleigh en raison de blocages. Par conséquent, les distances entre BS, UAV, IRS et GU peuvent être calculées en fonction de leurs coordonnées comme $d^{\mathsf{BU}} = \sqrt{\|\mathbf{r}^{\mathsf{b}} - \mathbf{q}\|^2 + (H^{\mathsf{b}} - H)^2}$, $d^{\mathsf{UI}} = \sqrt{\|\mathbf{q} - \mathbf{w}^{\mathsf{i}}\|^2 + (H - H^{\mathsf{i}})^2}$, $d_k^{\mathsf{UG}} = \sqrt{\|\mathbf{q} - \mathbf{r}_k^{\mathsf{u}}\|^2 + H^2}, \forall k$, $d_k^{\mathsf{IG}} =$

$\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2+(H^{\mathsf{i}})^2}, \forall k$, correspondant aux distances de BS à UAV, UAV à IRS, UAV à GU k et IRS à GU k.

Selon [43], le signal reçu par le GU k en provenance de l'UAV est donné par $y_k = \sqrt{p}\left((\mathbf{h}_k^{\mathsf{IG}})^H \mathbf{\Phi} \mathbf{h}^{\mathsf{UI}} + h_k^{\mathsf{UG}}\right) x_k + n^{\mathsf{G}}$, où x_k représente le symbole transmis par l'UAV, qui satisfait $\mathbb{E}(|x_k|^2) = 1$, et p indique la puissance d'émission de l'UAV au GU k sur chaque sous- canal, (c'est-à-dire $p = P_{\max}/C$ en supposant une allocation de puissance uniforme où $P_{\max}$ est la puissance d'émission totale de l'UAV), et n^{G} désigne le bruit gaussien blanc additif reçu par le GU, avec une moyenne nulle et une variance σ^2. De plus, $h_k^{\mathsf{UG}}, \mathbf{h}^{\mathsf{UI}}$ et $\mathbf{h}_k^{\mathsf{IG}}$ désignent les coefficients de canal des liens entre UAV et GU k, UAV et IRS, IRS et GU k, qui sont exprimés par $h_k^{\mathsf{UG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{UG}})^2}}, \forall k$, et

$$\begin{aligned} \mathbf{h}^{\mathsf{UI}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{UI}})^2}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}\right]^H \\ \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}\right]^H, \end{aligned} \tag{2.15}$$

$$\begin{aligned} \mathbf{h}_k^{\mathsf{IG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{IG}})^\kappa}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}\right]^H \\ \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}\right]^H \times \alpha^{\mathsf{IG}}, \forall k, \end{aligned} \tag{2.16}$$

où β_0 désigne le gain du canal à la distance de référence de 1 mètre, κ est l'exposant de perte de trajet, λ est la longueur de l'onde porteuse et α^{IG} est la composante de diffusion aléatoire, modélisée par une variable aléatoire gaussienne complexe à symétrie circulaire avec une moyenne nulle et une variance unitaire. De plus, $(\theta^{\mathsf{UI}}, \xi^{\mathsf{UI}})$ et $(\theta_k^{\mathsf{IG}}, \xi_k^{\mathsf{IG}})$ représentent les angles de départ vertical et horizontal de l'UAV vers l'IRS et de l'IRS vers GU k donnés par $\sin\theta^{\mathsf{UI}} = \frac{|H-H^{\mathsf{i}}|}{d^{\mathsf{UI}}}$, $\sin\xi^{\mathsf{UI}} = \frac{|x^{\mathsf{i}}-x^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\cos\xi^{\mathsf{UI}} = \frac{|y^{\mathsf{i}}-y^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\sin\theta_k^{\mathsf{IG}} = \frac{H^{\mathsf{i}}}{d_k^{\mathsf{IG}}}$, $\sin\xi_k^{\mathsf{IG}} = \frac{|x^{\mathsf{i}}-x_k^{\mathsf{u}}|}{\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}$, et $\cos\xi_k^{\mathsf{IG}} = \frac{|y^{\mathsf{i}}-y_k^{\mathsf{u}}|}{\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}$, $\forall k \in \mathcal{K}$.

Selon [43], le débit réalisable par le GU k desservi par l'UAV sur le sous-canal c peut être exprimé comme suit :

$$R_{k,c}^{\mathsf{A}} = \psi_{k,c}^{\mathsf{A}} W \log_2\left(1+\frac{p}{\sigma^2}\left|\frac{\sqrt{\beta_0}}{d_k^{\mathsf{UG}}}+\frac{\beta_0 f_k|\alpha^{\mathsf{IG}}|}{(d_k^{\mathsf{IG}})^{\kappa/2}d^{\mathsf{UI}}}\right|^2\right), \tag{2.17}$$

où $f_k = \sum_{i_c=1}^{I_c}\sum_{i_r=1}^{I_r} e^{j\left(F_k^{i_r,i_c}+\phi_{i_r,i_c}\right)}, \forall k$, et $F_k^{i_r,i_c} = -\frac{2\pi d}{\lambda}((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}} + \sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}) + (i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}} + \sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}})) - \arg(\alpha^{\mathsf{IG}})$.

Par ailleurs, le débit réalisable de la liaison backhaul sur le sous-canal c peut être exprimé comme

$$R_{0,c}^{\mathsf{B}} = \psi_{0,c}^{\mathsf{B}} W \log_2 \left(1 + \frac{p_0 \beta_0}{(d^{\mathsf{BU}})^2 \sigma^2}\right), \tag{2.18}$$

où p_0 désigne la puissance d'émission de la BS cellulaire.

De plus, pour maintenir de bonnes performances de bout en bout, le débit de données total de toutes les liaisons d'accès de l'UAV à tous les GU ne doit pas dépasser le débit de liaison. Cette contrainte peut être décrite comme $\sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}} R_{k,c}^{\mathsf{A}} \leq \sum_{c\in\mathcal{C}} R_{0,c}^{\mathsf{B}}$.

Soit $\mathbf{\Psi} = \{\psi_{k,c}^{\mathsf{A}}, \psi_{0,c}^{\mathsf{B}}, \forall k, c\}$, $\mathbf{\Phi}$ et $\mathbf{Q} = \{\mathbf{q}\}$ les vecteurs de toutes les variables de décision pour l'affectation des sous-canaux, les déphasages IRS et le placement des UAV. Nous voulons maximiser le débit total de tous les GU en optimisant toutes les variables $\mathbf{\Psi}$, $\mathbf{\Phi}$ et $\mathbf{Q}$. Ce problème de conception peut être formulé comme

$$\textbf{(P2):} \max_{\mathbf{\Psi},\mathbf{\Phi},\mathbf{Q}} \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}} R_{k,c}^{\mathsf{A}} \tag{2.19}$$

$$\text{s.t.} \sum_{c\in\mathcal{C}} \psi_{k,c}^{\mathsf{A}} \geq 1, \forall k, \tag{2.19a}$$

$$\sum_{k\in\mathcal{K}} \psi_{k,c}^{\mathsf{A}} + \psi_{0,c}^{\mathsf{B}} \leq 1, \forall c, \tag{2.19b}$$

$$\sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}} R_{k,c}^{\mathsf{A}} \leq \sum_{c\in\mathcal{C}} R_{0,c}^{\mathsf{B}}, \tag{2.19c}$$

$$\psi_{k,c}^{\mathsf{A}}, \psi_{0,c}^{\mathsf{B}} \in \{0, 1\}, \forall k, c, \tag{2.19d}$$

$$\phi_{i_r,i_c} \in [0, 2\pi), \forall i_r = 1, \ldots, I_r; \forall i_c = 1, \ldots, I_c. \tag{2.19e}$$

À cause de la contrainte non convexe (2.19c) et des variables entières dans (2.19d), le problème (2.19) est un programme d'optimisation non linéaire mixte à nombre entier non convexe (MINLP), difficile à résoudre. On pourrait dire que l'ajout de la contrainte (2.19c) est une modification triviale des modèles précédents. Bien que cela soit certainement vrai pour l'écriture du modèle mathématique, cette contrainte n'est pas convexe et rend donc la conception d'un algorithme de résolution efficace beaucoup plus compliquée. Dans ce qui suit, nous décrivons les détails de notre algorithme.

2.2.2.2 Algorithmes proposés

Pour résoudre le problème **(P2)**, nous dérivons d'abord une solution analytique de déphasage, puis nous optimisons l'affectation des sous-canaux et le placement des UAV de manière itérative. Soient $\mathcal{C}^{\mathsf{A}}$ et $\mathcal{C}^{\mathsf{B}}$ les ensembles de sous-canaux affectés aux liaisons d'accès et de liaison, où $\mathcal{C} = \mathcal{C}^{\mathsf{A}} \cup \mathcal{C}^{\mathsf{B}}$. Initialement, le nombre de sous-canaux alloués dans $\mathcal{C}^{\mathsf{A}}$ est égal au nombre de GU K pour s'assurer que chaque GU se voit attribuer au moins un sous-canal et tous les sous-canaux restants sont alloués à $\mathcal{C}^{\mathsf{B}}$. Ensuite, les ensembles $\mathcal{C}^{\mathsf{A}}$ et $\mathcal{C}^{\mathsf{B}}$ sont mis à jour en prenant des sous-canaux de $\mathcal{C}^{\mathsf{B}}$ en les réallouant à $\mathcal{C}^{\mathsf{A}}$. Les décalages de phase IRS et le placement des UAV sont optimisés en conséquence tout en maintenant la contrainte de capacité de liaison.

Pour obtenir le débit d'accès maximal $R^{\mathsf{A}}_{k,c}$ donné dans (2.17) et donc le débit total (c'est-à-dire la fonction objectif), les déphasages IRS $\mathbf{\Phi}^*$ doivent être alignés avec les phases des coefficients du canal. De tels déphasages IRS optimaux, qui se traduisent par $f_k^* = I_c I_r$, peuvent être exprimés comme

$$\begin{aligned} \phi^*_{i_r,i_c} \quad = \quad & \frac{2\pi d}{\lambda}\Big((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}) \\ & +(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}})\Big) + \arg(\alpha^{\mathsf{IG}}). \end{aligned} \tag{2.20}$$

La substitution de ces déphasages IRS dans le problème **(P2)** entraîne toujours un problème MINLP non convexe. Ainsi, nous utilisons l'approche d'optimisation alternée pour résoudre ce problème où nous optimisons de manière itérative chaque ensemble de variables d'optimisation compte tenu des valeurs des autres variables jusqu'à la convergence.

a) Optimisation de l'affectation des sous-canaux

Pour $\mathbf{\Phi}$ et $\mathbf{Q}$ donnés, le sous-problème pour optimiser l'affectation de sous-canal $\mathbf{\Psi}$ peut être énoncé comme suit

$$\textbf{(P2.1):} \max_{\mathbf{\Psi}} \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{A}}_{k,c} \tag{2.21}$$

$$\text{s.t.} \quad \sum_{c\in\mathcal{C}^{\mathsf{A}}} \psi^{\mathsf{A}}_{k,c} \geq 1, \forall k, \tag{2.21a}$$

$$\sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}^{\mathsf{B}}} R^{\mathsf{B}}_{0,c}, \tag{2.21b}$$

$$\text{contraintes} \quad (2.19\text{b}), (2.19\text{d}).$$

Il s'agit d'un programme linéaire mixte en nombres entiers (MILP) standard, qui peut être résolu efficacement en utilisant le outil CVX-Mosek [38].

b) Optimisation du placement de l'UAV

Pour $\mathbf{\Psi}$ et $\mathbf{\Phi}$ donnés, le sous-problème pour optimiser le placement de l'UAV $\mathbf{Q}$ est non convexe. Pour résoudre ce problème, nous introduisons d'abord quelques variables auxiliaires, puis résolvons le problème transformé en utilisant la méthode SCA. Spécifiquement, nous introduisons des variables $\nu_k \geq \|\mathbf{q}-\mathbf{r}^{\mathsf{u}}_k\|^2 + H^2, \forall k,$ $\mu \geq \left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2 + (H-H^{\mathsf{i}})^2$, et $\epsilon \geq \left\|\mathbf{r}^{\mathsf{b}}-\mathbf{q}\right\|^2 + (H^{\mathsf{b}}-H)^2$. À partir de (2.17) et (2.18), nous avons

$$R^{\mathsf{Aq}}_{k,c} = \psi^{\mathsf{A}}_{k,c} W \log_2 \left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k}+\frac{Y_k^2}{\mu}+\frac{2X_kY_k}{\nu_k^{1/2}\mu^{1/2}}\right)\right) \leq R^{\mathsf{A}}_{k,c}, \tag{2.22}$$

$$R^{\mathsf{Bq}}_{0,c} = \psi^{\mathsf{B}}_{0,c} W \log_2 \left(1+\frac{Z}{\epsilon}\right) \leq R^{\mathsf{B}}_{0,c}, \tag{2.23}$$

où $X_k = \sqrt{\beta_0}$, $Y_k = \beta_0 f_k^* |\alpha^{\mathsf{IG}}| (d_k^{\mathsf{IG}})^{-\kappa/2}$, $Z = p_0\beta_0/\sigma^2$, dans lequel $f_k^* = I_c I_r$ est une solution donnée par les déphasages IRS exprimés en (2.20). On peut vérifier que $R^{\mathsf{Aq}}_{k,c}$ est une fonction convexe par rapport à ν_k et μ et peut être minorée par son développement de Taylor du premier ordre à r-ième itération dans le processus d'approximation comme suit :

$$R^{\mathsf{Aq}}_{k,c} \geq \psi^{\mathsf{A}}_{k,c} W \log_2 D^r + \frac{L^r}{D^r}(\nu_k - \nu_k^r) + \frac{S^r}{D^r}(\mu - \mu^r) \triangleq R^{\mathsf{Aqlb}}_{k,c}, \tag{2.24}$$

où

$$\begin{aligned} D^r &= \left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k^r}+\frac{Y_k^2}{\mu^r}+\frac{2X_kY_k}{\nu_k^{1/2,r}\mu^{1/2,r}}\right)\right), \\ L^r &= -\psi_{k,c}^{\mathsf{A}}W\log_2(e)\left(\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k^{2,r}}+\frac{X_kY_k}{\nu_k^{3/2,r}\mu^{1/2,r}}\right)\right), \\ S^r &= -\psi_{k,c}^{\mathsf{A}}W\log_2(e)\left(\frac{p}{\sigma^2}\left(\frac{Y_k^2}{\mu^{2,r}}+\frac{X_kY_k}{\nu_k^{1/2,r}\mu^{3/2,r}}\right)\right). \end{aligned}$$

De même, puisque $R_{0,c}^{\mathsf{Bq}}$ est une fonction convexe de ϵ, on peut utiliser le premier terme du développement de Taylor comme borne inférieure

$$R_{0,c}^{\mathsf{Bq}} \geq \psi_{0,c}^{\mathsf{B}}W\log_2\left(1+\frac{Z}{\epsilon^r}\right) - \psi_{0,c}^{\mathsf{B}}W\frac{\log_2(e)Z}{\epsilon^r(\epsilon^r+Z)}(\epsilon-\epsilon^r) \triangleq R_{0,c}^{\mathsf{Bqlb}}, \tag{2.25}$$

De plus, la borne supérieure du débit d'accès donnée dans (2.17) peut être exprimée en introduisant des variables auxiliaires $\alpha_k \leq \|\mathbf{q}-\mathbf{r}_k^{\mathsf{u}}\|^2+H^2, \forall k,\ \gamma \leq \left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2+(H-H^{\mathsf{i}})^2$ et nous avons

$$R_{k,c}^{\mathsf{Aub}} = \psi_{k,c}^{\mathsf{A}}W\log_2\left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\alpha_k}+\frac{Y_k^2}{\gamma}+\frac{2X_kY_k}{\alpha_k^{1/2}\gamma^{1/2}}\right)\right) \geq R_{k,c}^{\mathsf{A}}. \tag{2.26}$$

Par conséquent, le problème d'optimisation du placement des UAV peut être approximé par

$$\textbf{(P2.2):} \max_{\mathbf{Q},\nu_k,\mu,\epsilon,\alpha_k,\gamma} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}^{\mathsf{A}}} R_{k,c}^{\mathsf{Aqlb}} \tag{2.27}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}^{\mathsf{A}}} R_{k,c}^{\mathsf{Aub}} - \sum_{c\in\mathcal{C}^{\mathsf{B}}} R_{0,c}^{\mathsf{Bqlb}} \leq 0, \tag{2.27a}$$

$$\nu_k \geq \|\mathbf{q}-\mathbf{r}_k^{\mathsf{u}}\|^2+H^2, \forall k;\ \mu \geq \left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2+(H-H^{\mathsf{i}})^2, \tag{2.27b}$$

$$\alpha_k \leq \|\mathbf{q}^r-\mathbf{r}_k^{\mathsf{u}}\|^2+2\left(\mathbf{q}^r-\mathbf{r}_k^{\mathsf{u}}\right)^T(\mathbf{q}-\mathbf{q}^r)+H^2, \forall k, \tag{2.27c}$$

$$\gamma \leq \left\|\mathbf{q}^r-\mathbf{w}^{\mathsf{i}}\right\|^2+2\left(\mathbf{q}^r-\mathbf{w}^{\mathsf{i}}\right)^T(\mathbf{q}-\mathbf{q}^r)+(H-H^{\mathsf{i}})^2, \tag{2.27d}$$

$$\epsilon \geq \left\|\mathbf{r}^{\mathsf{b}}-\mathbf{q}\right\|^2+(H^{\mathsf{b}}-H)^2. \tag{2.27e}$$

Il s'agit d'un problème convexe, qui peut être résolu efficacement en utilisant le outil CVX-Mosek [38]. Les solutions de ces sous-problèmes sont utilisées dans notre algorithme qui est décrit dans l'algorithme 2.5.

Algorithm 2.5. Algorithme conjoint pour l'attribution de sous-canaux, les déphasages IRS et le placement de l'UAV

1: Initialisation: $\mathcal{C}^{\mathsf{A}}, \mathcal{C}^{\mathsf{B}}, \mathbf{Q}^0, \mathbf{\Phi}^0, \mathbf{\Psi}^0$, $S^* = 10^2, S = S_1 = 0$, $t = 0$;
2: **repeat**
3: $S^* = S$ et $t = t + 1$;
4: Choisir un sous-canal c de $\mathcal{C}^{\mathsf{B}}$; mettre à jour $\mathcal{C}^{\mathsf{A}} = \mathcal{C}^{\mathsf{A}} \cup \{c\}$ et $\mathcal{C}^{\mathsf{B}} \setminus \{c\}$;
5: **repeat**
6: Étant donné $\mathbf{\Phi}^{r,*}$ dans (2.20) ;
7: Résoudre **(P2.1)** itérativement jusqu'à convergence pour obtenir $\mathbf{\Psi}^r$;
8: Résoudre **(P2.2)** itérativement jusqu'à convergence pour obtenir $\mathbf{Q}^r$;
9: **until** Convergence
10: **if** Obtenir une solution réalisable avec un débit total maximum S_1 et $S<S_1$ **then**
11: Mise à jour $S = S_1$ et $\{\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*\} = \{\mathbf{\Psi}^t, \mathbf{Q}^t, \mathbf{\Phi}^t\}$;
12: **else**
13: Mise à jour $\mathcal{C}^{\mathsf{B}} = \mathcal{C}^{\mathsf{B}} \cup \{c\}$ et $\mathcal{C}^{\mathsf{A}} \setminus \{c\}$;
14: Mise à jour $\{\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*\} = \{\mathbf{\Psi}^{t-1}, \mathbf{Q}^{t-1}, \mathbf{\Phi}^{t-1}\}$;
15: **end if**
16: **until** $|S - S^*| < 10^{-6}$
17: Retour $\mathbf{\Psi}^*, \mathbf{Q}^*$, $\mathbf{\Phi}^*$;

2.2.2.3 Résultats numériques

Nous considérons une zone de réseau rectangulaire de taille $1000 \times 1000 (\mathrm{m}^2)$. L'altitude de l'UAV est fixée à $H = 120$m et la BS est située au point $(0, 0, 20)$m. De plus, l'IRS est fixé à $(500, 500, 50)$m et les GU sont regroupé en régions circulaires de rayon $r_c = 200$m. Nous plaçons initialement l'UAV au centre de chaque groupe. Les paramètres restants sont définis comme $p_0 = 33$dBm, $P_{\mathsf{max}} = 30$dBm, $W = 1$MHz, $\sigma^2 = -110$dBm, $f_c = 2.5$ GHz, $d = \lambda/2$ et $\kappa = 2$. Les IRS carrés avec $I_r = I_c$ seront considérés où le nombre d'éléments IRS est indiqué par $I = I_r I_c$.

Les figures 2.3 et 2.4 montrent le débit total atteint par l'algorithme (c'est-à-dire l'algorithme 2.5) dans le cas où l'UAV est placé au centre du cluster, qui sont indiqués comme "emplacement optimisé UAV" et "emplacement centré UAV". La figure 2.3 montre le débit total pour différents nombres de GU avec $C = 60$ et $I = 64$. On peut voir que le débit total augmente légèrement avec l'augmentation du nombre de GU et que la différence entre l'emplacement optimisé et centré de l'UAV devient plus grande à mesure que le nombre de GU augmente. Le gain de débit dû à l'algorithme proposé avec et sans l'IRS est d'environ 15 %. La figure 2.4 illustre le débit total pour différents nombres d'éléments IRS avec 20 GU et $C = 60$. En fait, un plus grand nombre de sous-canaux ou d'éléments IRS conduit à une plus grande diversité de système, ce qui améliore le débit total obtenu. Ces résultats confirment l'efficacité de l'algorithme proposé pour optimiser le placement des UAV et l'affectation des sous-canaux dans les communications UAV assistées par IRS.

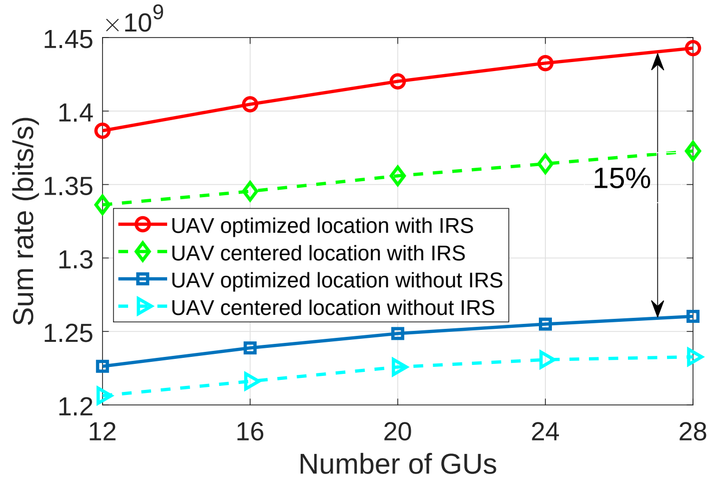

Figure 2.3 – Débit total pour différents nombres de GU.

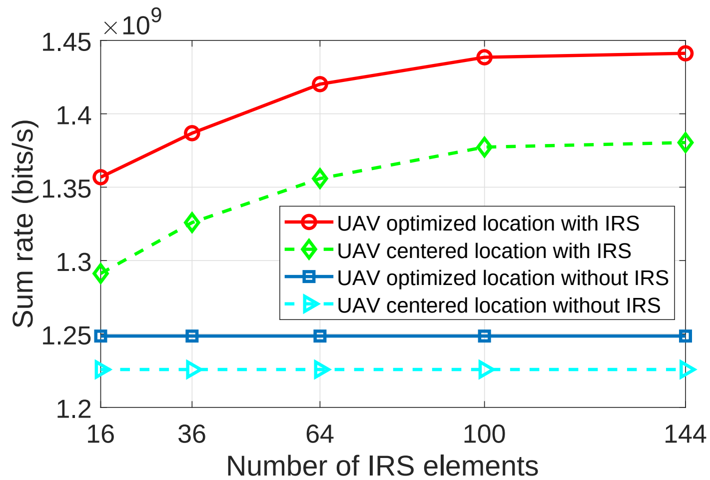

Figure 2.4 – Débit total pour différents nombres d'éléments IRS.

2.2.3 Délestage des calculs intégré, contrôle de trajectoire UAV, planification des utilisateurs, allocation de ressources et contrôle d'admission dans SAGIN

Dans cette partie, nous étudions le délestage de calcul intégré, le contrôle de la trajectoire UAV, la planification des utilisateurs, l'allocation des ressources et le contrôle d'admission pour les SAGIN avec des communications par satellite multi-sauts. Les principales contributions peuvent être résumées comme suit:

- Nous étudions le délestage partiel des calculs dans SAGIN où une partie des tâches de calcul des GU sont traitées localement et/ou par délestage, sur les serveurs de bordure montés sur

UAV ou un serveur cloud, en exploitant par satellite LEO multi-sauts. Nous formulons un problème d'optimisation qui vise à minimiser la consommation d'énergie pondérée des GU et des UAV tout en satisfaisant les contraintes de délai maximal des tâches de calcul sous-jacentes.

- La méthode d'optimisation alternée est utilisée pour résoudre le problème d'optimisation non linéaire à nombre entier mixte non convexe sous-jacent (MINLP). De plus, la méthode d'approximation convexe successive (SCA) est utilisée pour résoudre les sous-problèmes d'allocation de ressources de calcul et de bande passante et de contrôle de trajectoire UAV.

- Nous proposons des stratégies efficaces pour la vérification de faisabilité et le contrôle d'admission dans des scénarios de réseau surchargé. Plus précisément, un algorithme itératif est proposé pour résoudre le problème de vérification de faisabilité et une stratégie efficace de suppression d'utilisateurs est développée pour le contrôle d'admission tout en satisfaisant toutes les contraintes des GU et du système.

- Des résultats numériques sont présentés pour montrer les impacts de différents paramètres, y compris le nombre de sauts dans les communications par satellite multi-sauts, le nombre de GU, la bande passante et la taille de la tâche de calcul sur les performances réalisables et les gains dus à l'optimisation du contrôle de trajectoire UAV, la planification des utilisateurs, l'allocation des ressources et le délestage des calculs. De plus, le débit d'admission des GU qui sont effectivement desservies dans les différents scénarios est présenté.

2.2.3.1 Modèle de système

Nous considérons la conception de délestage des calculs dans le système de cloud de périphérie basé sur SAGIN, illustré à la figure 7.1, où le réseau terrestre comprend K GU situés au sol, la couche du réseau aérien utilise M UAV et la couche du réseau spatial repose sur des satellites LEO pour les connexions à un serveur cloud distant. Nous désignons les ensembles de satellites, UAV et GU par $\mathcal{S} = \{1, ..., S\}$, $\mathcal{M} = \{1, ..., M\}$, et $\mathcal{K} = \{1, ..., K\}$.

Nous supposons que le délestage partiel du calcul est utilisé pour une tâche de calcul de chaque GU. Spécifiquement, chaque GU partitionne sa tâche de calcul en trois sous-tâches où la première est traitée localement et les deux autres sont traitées sur le serveur périphérique monté sur UAV et le serveur cloud. De plus, les données liées à la deuxième sous-tâche doivent être transmises du GU associé à l'UAV de calcul, tandis que les données liées à la troisième sous-tâche doivent être transmises du GU au serveur cloud via une communication par satellite multi-sauts.

Tous les GU situés au sol au point $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. De plus, nous supposons que les UAV volent à une altitude fixe H sur une période de vol de $T > 0$ secondes. Nous divisons la période de vol en N tranches où l'ensemble de tranches de temps est noté $\mathcal{N} = \{1, ..., N\}$. De plus, nous supposons que les communications en liaison montante des plusieurs GU vers leurs UAV associés utilisent l'accès multiple par répartition en fréquence (FDMA). Dénotons W la bande passante totale disponible pour prendre en charge les communications de liaison montante des GU aux UAV. Nous supposons que la bande passante disponible est partitionnée en sous-bandes orthogonales dont chacune est allouée à un UAV pour desservir son GU associé. Nous désignons la bande passante allouée à l'UAV m par W_m^{u} où $\sum_{m \in \mathcal{M}} W_m^{\mathsf{u}} = W$. Nous supposons également que les associations entre les GU et les UAV et entre les GU et les satellites sont fixées pendant le processus de délestage des calculs. De plus, nous supposons que la taille des données correspondant aux résultats de calcul est beaucoup plus petite que celle des données de délestage, de sorte que nous pouvons négliger le temps de téléchargement des résultats de calcul dans le processus de délestage. Pour plus de facilité, la liste des principales notations du document est donnée dans le tableau 7.2.

a) Modèle de tâche de calcul

Nous supposons que chaque GU k a une tâche de calcul sous une contrainte de délai représentée par $U_k = (f_k, s_k, c_k, T_k^{\mathsf{max}})$, où f_k désigne la demande de calcul exprimée par le nombre de cycles de l'unité centrale de traitement (CPU) par seconde (cycles CPU/seconde), s_k (bits) représente la taille des données brutes d'entrée, c_k (cycles CPU/ bit) désigne la ressource de calcul requise pour les données d'entrée de 1 bit, et T_k^{max} (secondes) décrit la latence maximale tolérable de la tâche de calcul U_k.

Nous supposons que la tâche de calcul de chaque GU est divisée en trois sous-tâches qui sont traitées en parallèle par le GU, le serveur périphérique monté sur UAV et le serveur cloud atteint via la communication par satellite LEO multi-sauts, comme indiqué dans [44, 45].

Le temps de traitement de la tâche au GU k peut être exprimé comme suit

$$T_k = \max\left\{T_k^{\mathsf{lo}}, T_k^{\mathsf{ed}}, T_k^{\mathsf{cl}}\right\}, \tag{2.28}$$

où T_k^{lo}, T_k^{ed} et T_k^{cl} représentent le temps total de transmission de données et d'exécution de la tâche par le GU, le serveur périphérique monté sur UAV, et le serveur cloud. Spécifiquement, T_k^{ed} inclut à la fois le temps de transmission des données du GU k vers l'UAV associé et le temps d'exécution de la sous-tâche par l'UAV. Nous décrirons plus en détail comment calculer ce temps d'exécution ultérieurement. Par conséquent, la contrainte de délai au GU k peut être exprimée sous la forme $T_k \leq T_k^{\mathsf{max}}$.

Pour modéliser le partitionnement des tâches au GU k, nous introduisons les variables λ_k^{lo} et λ_k^{ed}, $(0 \leq \lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}} \leq 1)$, qui représentent les fractions de données d'entrée à traiter localement par le GU k et à traiter sur le serveur périphérique. Par conséquent, $(1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}})$ représente la fraction des données d'entrée du GU k à traiter sur le serveur cloud distant.

b) Contrôle de trajectoire UAV

Les coordonnées de l'UAV m dans le créneau horaire n sont notées $\mathbf{q}_m[n] = (x_m^{\mathsf{d}}[n], y_m^{\mathsf{d}}[n])$. Nous supposons que chaque UAV doit revenir à sa position initiale à la fin de la période de vol, c'est-à-dire $\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m \in \mathcal{M}$. De plus, la taille des créneaux $\Delta t = T/N$ est choisi suffisamment petite pour que les UAV se trouvent dans un petit voisinage dans chaque créneau temporel, même à la vitesse de vol maximale $V_{\max}$.

c) Planification des utilisateurs

Soit $\phi_{k,m}^{\mathsf{u}}[n]$ les variables de décision binaires pour l'association entre les GU et les UAV sur la période de vol T, où $\phi_{k,m}^{\mathsf{u}}[n] = 1$ si le GU k est desservi par l'UAV m dans le créneau horaire n et $\phi_{k,m}^{\mathsf{u}}[n] = 0$, sinon. La première exigence pour l'association est que chaque GU puisse délester sa sous-tâche de calcul sur au plus un UAV dans chaque tranche de temps, c'est-à-dire $\sum_{m \in \mathcal{M}} \phi_{k,m}^{\mathsf{u}}[n] \leq 1$. Nous supposons que chaque GU k est initialement associé à l'UAV fournissant la puissance moyenne du signal reçu (RSS) la plus élevée, c'est-à-dire $\phi_{k,m}^{\mathsf{u}}[n] = 1$ avec $m = \arg\max_k(RSS_{k,m}[n])$, où $RSS_{k,m}[n](\text{dBm}) = P_k^{\mathsf{u}}(\text{dBm}) - g_{k,m}[n](\text{dBm})$ où P_k^{u} indique la puissance d'émission du GU k vers son UAV associé et $g_{k,m}[n]$ représente le gain de puissance du canal du GU k à UAV m. Pour satisfaire la contrainte de délai de chaque GU, le nombre *d'intervalles de temps consécutifs* requis pour traiter complètement la tâche de calcul du GU k peut être noté $N_k = \lceil T_k^{\mathsf{max}}/\Delta t \rceil$, où $\lceil . \rceil$ désigne l'opération arrondi à l'entier supérieur. Nous introduisons maintenant les variables binaires pour la planification des utilisateurs $\theta_k[n]$, où $\theta_k[n] = 1$ si le GU k est programmé pour transmettre à son UAV associé dans le créneau horaire n et $\theta_k[n] = 0$, sinon. Nous devons imposer les contraintes suivantes aux décisions de planification des utilisateurs :

$$\sum_{m\in\mathcal{M}}\sum_{t=0}^{N_k-1}\theta_k[n+t]\phi^{\mathsf{u}}_{k,m}[n+t]=N_k,\forall k,n\in\{1,...,N-N_k\}.\tag{2.29}$$

d) Modèles informatiques

- **Modèle informatique local**:

  Le temps d'exécution de la tâche locale par le GU k peut être exprimé comme

$$T_k^{\mathsf{lo}}=\frac{\lambda_k^{\mathsf{lo}}s_kc_k}{f_k}.\tag{2.30}$$

  La contrainte de délai imposée au traitement local peut être exprimée sous la forme $T_k^{\mathsf{lo}}\leq T_k^{\mathsf{max}}$. La consommation d'énergie due à l'exécution des tâches locales peut être calculée comme

$$E_k^{\mathsf{lo}}=\kappa\lambda_k^{\mathsf{lo}}s_kc_k(f_k)^2,\tag{2.31}$$

  où κ est la capacité commutée effective en fonction de l'architecture du processeur [46].

- **Modèle edge computing monté sur UAV**:

  Pour les sous-tâches délestées vers les UAV, nous dénotons par $l_k^{\mathsf{u}}[n]$ le nombre de bits transmis du GU k vers l'UAV associé sur l'intervalle de temps n. En outre, notons la ressource de calcul de l'UAV m allouée pour gérer la sous-tâche du GU k dans le créneau horaire n par $f_k^{\mathsf{u}}[n]$ (cycles CPU/seconde).

  Par conséquent, la consommation totale d'énergie des UAV associés pour traiter la sous-tâche du GU k peut être calculée comme suit :

$$E_k^{\mathsf{ed}}=\sum_{m\in\mathcal{M}}\sum_{n\in\mathcal{N}}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(l_k^{\mathsf{u}}[n]c_ke^{\mathsf{ed}}\right).\tag{2.32}$$

  De plus, nous supposons que les liaisons de communication entre les GU et les UAV sont dominées par la propagation en visibilité directe (LoS) où la qualité du canal dépend principalement de la distance UAV-GU. La distance entre GU k et UAV m dans le créneau horaire n peut être calculée comme suit : $d_{k,m}[n]=\sqrt{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$. De plus, le gain de puissance du canal du GU k à UAV m dans l'intervalle de temps n est supposé suivre le modèle de perte de trajet en espace libre, qui peut être exprimé par $g_{k,m}[n]=\rho_0(d_{k,m}[n])^{-2}=\frac{\rho_0}{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$, où ρ_0 présente le gain de puissance du canal à la distance de référence de 1 m. Par conséquent, le débit réalisable de la transmission montante du GU k à l'UAV associé m dans l'intervalle de temps n, noté $R_{k,m}^{\mathsf{u}}[n]$ en bits/seconde (bps), peut être exprimé comme

$$R_{k,m}^{\mathsf{u}}[n] = \beta_k^{\mathsf{u}}[n] \log_2 \left(1 + \frac{P_k^{\mathsf{u}} g_{k,m}[n]}{\beta_k^{\mathsf{u}}[n]\sigma^2}\right), \tag{2.33}$$

où $\beta_k^{\mathsf{u}}[n]$ et P_k^{u} représentent la bande passante allouée au GU k dans le créneau horaire n et la transmission puissance du GU k pour sa transmission en liaison montante et σ^2 désigne la densité de puissance du bruit gaussien blanc additif (AWGN) du récepteur.

De plus, nous supposons que la tâche partielle du GU k est déchargée et traitée complètement à chaque UAV associé dans chaque tranche de temps. On a alors les contraintes suivantes

$$T_{k,m}^{\mathsf{ed}}[n]=\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}+\frac{l_k^{\mathsf{u}}[n]}{R_{k,m}^{\mathsf{u}}[n]}\right) \leq \Delta t, \forall k, m, n. \tag{2.34}$$

Le temps de traitement total des UAV pour desservir le GU k peut être écrit comme

$$T_k^{\mathsf{ed}} = \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n]. \tag{2.35}$$

De plus, chaque UAV consomme d'énergie pendant son vol stationnaire. La consommation d'énergie en vol de l'UAV m peut être exprimée comme $E_m^{\mathsf{edf}} = P_m^{\mathsf{f}} T$, où P_m^{f} indique la puissance de vol de l'UAV m.

- **Modèle de cloud computing par satellite**:

  Nous omettons le temps de traitement du serveur cloud et nous ignorons également la consommation d'énergie du cloud impliquée dans l'exécution des tâches de calcul et la transmission des résultats de calcul du serveur cloud aux GU. Dans un travail récent [2], un algorithme pour déterminer le nombre de liaisons inter-satellites (ISL) et les satellites correspondants pour établir le chemin de communication multi-sauts entre deux emplacements sur le terrain a été proposé (voir l'algorithme 1 de [2]). En utilisant cet algorithme, le nombre de sauts entre le premier et le dernier satellite reliant la zone de réseau terrestre considérée et le serveur cloud peut être déterminée comme L. Par conséquent, le temps total de traitement des données et le temps de propagation du GU

k au serveur cloud peuvent être calculés comme suit

$$T_k^{\mathsf{cl}} = (1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}})s_k\bigg(\frac{1}{R_k^{\mathsf{s}}} + \sum_{i=1}^{L}(\frac{1}{R_i^{\mathsf{ss}}}) + \frac{1}{R^{\mathsf{cl}}}\bigg) + T_k^{\mathsf{prop}}, \tag{2.36}$$

où $R_k^{\mathsf{s}}, R_i^{\mathsf{ss}}, R^{\mathsf{cl}}$ représentent les débits de transmission entre le GU k et le premier satellite, entre les satellites dans le i-ième saut, et entre le dernier satellite et le serveur cloud. Ici, T_k^{prop} représente le délai de propagation total du GU k jusqu'au serveur cloud. De plus, la consommation d'énergie du GU k pour transmettre les données liées à la sous-tâche déchargée vers le premier satellite peut être calculée comme suit :

$$E_k^{\mathsf{s}} = \frac{(1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}})s_k P_k^{\mathsf{s}}}{R_k^{\mathsf{s}}}, \tag{2.37}$$

où P_k^{s} représente la puissance d'émission du GU k vers le satellite.

2.2.3.2 Formulation du problème

Dans ce travail, nous intéressons à la minimisation de la consommation d'énergie pondérée de tous les GU et UAV pour toutes les tâches de calcul impliquées, qui peut être exprimée comme

$$\begin{aligned} E^{\mathsf{sum}} \quad &= \quad \alpha_1\bigg(\sum_{k\in\mathcal{K}} E_k^{\mathsf{ed}} + \sum_{m\in\mathcal{M}} P_m^{\mathsf{f}} T\bigg) \\ &+ \alpha_2 \sum_{k\in\mathcal{K}} \bigg(E_k^{\mathsf{lo}} + \sum_{m\in\mathcal{M}}\sum_{n\in\mathcal{N}} \theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]E_{k,m}^{\mathsf{edt}}[n] + E_k^{\mathsf{s}}\bigg), \end{aligned} \tag{2.38}$$

où $\alpha_1, \alpha_2 \in [0, 1]$ représentent les facteurs de pondération pour équilibrer la consommation d'énergie entre les UAV et les GU.

Pour plus de commodité, nous rassemblons différentes variables de décision et définissons les groupes de variables correspondants comme suit: planification de l'utilisateur $\mathbf{\Theta} = \{\theta_k[n], \forall k, n\}$, contrôle de délestage partiel $\mathbf{\Lambda} = \{\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, \forall k\}$, allocation de bits $\mathbf{L} = \{l_k^{\mathsf{u}}[n], \forall k, n\}$, allocation de bande passante $\boldsymbol{\beta} = \{\beta_k^{\mathsf{u}}[n], \forall k, n\}$, allocation de ressources de calcul $\mathbf{F} = \{f_k^{\mathsf{u}}[n], \forall k, n\}$, et contrôle de trajectoire UAV $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$. Notre conception vise à minimiser la consommation d'énergie pondérée des GU et des UAV tout en satisfaisant les contraintes de délai maximum des tâches de calcul individuels. Le problème d'optimisation peut être formulé comme

$$\textbf{(P3):} \min_{\mathbf{\Theta},\mathbf{\Lambda},\mathbf{L},\boldsymbol{\beta},\mathbf{F},\mathbf{Q}} \quad E^{\mathsf{sum}} \tag{2.39}$$

$$\text{s.t. } T_k^{\mathsf{lo}} \leq T_k^{\mathsf{max}}, \forall k, \tag{2.39a}$$

$$T_k^{\mathsf{cl}} \leq T_k^{\mathsf{max}}, \forall k, \tag{2.39b}$$

$$\sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] \leq T_k^{\mathsf{max}}, \forall k, \tag{2.39c}$$

$$\theta_k[n] T_{k,m}^{\mathsf{ed}}[n] \leq \Delta t, \forall k, m, n, \tag{2.39d}$$

$$\sum_{m\in\mathcal{M}} \sum_{t=0}^{N_k-1} \theta_k[n+t]\phi_{k,m}^{\mathsf{u}}[n+t]{=}N_k, \forall k, n \in \{1,...,N{-}N_k\}, \tag{2.39e}$$

$$\sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n]\phi_{k,m}^{\mathsf{u}}[n] l_k^{\mathsf{u}}[n] = \lambda_k^{\mathsf{ed}} s_k, \forall k, \tag{2.39f}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\beta_k^{\mathsf{u}}[n] \leq W_m^{\mathsf{u}}, \forall m, n, \tag{2.39g}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n]\phi_{k,m}^{\mathsf{u}}[n] f_k^{\mathsf{u}}[n] \leq F_m^{\mathsf{max}}, \forall m, n, \tag{2.39h}$$

$$\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m, \tag{2.39i}$$

$$\|\mathbf{q}_m[n{+}1] - \mathbf{q}_m[n]\|^2 \leq D_{\max}^2, \forall m, n = 1, ..., N{-}1, \tag{2.39j}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d_{\min}^2, \ \forall n, m, j \neq m, \tag{2.39k}$$

$$\theta_k[n] \in \{0,1\}, \forall k, n, \tag{2.39l}$$

$$0 \leq \lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, 1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}} \leq 1, \forall k, \tag{2.39m}$$

$$\beta_k^{\mathsf{u}}[n], f_k^{\mathsf{u}}[n], l_k^{\mathsf{u}}[n] \geq 0, \forall k, n, \tag{2.39n}$$

où les contraintes (2.39a)-(2.39d) capturent les exigences de délai pour les GU. Les contraintes (2.39e) et (2.39l) décrivent les contraintes binaires d'ordonnancement des utilisateurs pour les GU desservies par les UAV associés. Les contraintes (2.39g) capturent l'allocation de bande passante pour la transmission entre les GU et les UAV tandis que les contraintes (2.39h) présentent les contraintes de calcul des UAV où F_m^{max} désigne la ressource de calcul maximale de l'UAV m. On peut voir que les fonctions objectif et de contrainte (2.39a)-(2.39d) sont non linéaires et que des variables de décision entières sont impliquées dans (2.39l) pour la planification de l'utilisateur. Par conséquent, le problème (2.39) est un problème d'optimisation non linéaire mixte à nombre entier non convexe (MINLP), difficile à résoudre de manière optimale.

2.2.3.3 Algorithmes proposés

Dans cette section, nous développons un algorithme pour résoudre le problème (2.39) en supposant qu'il existe une solution réalisable. Nous adoptons l'approche d'optimisation alternée pour résoudre le problème (2.39) où nous optimisons itérativement chaque ensemble de variables en fonction des valeurs des autres variables dans les sous-problèmes correspondants jusqu'à convergence. Nous décrivons comment résoudre différents sous-problèmes dans ce qui suit.

a) Optimisation de la planification des utilisateurs

Étant donné $\{\mathbf{L}, \mathbf{\Lambda}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, le sous-problème de planification de l'utilisateur pour optimiser $\mathbf{\Theta}$ peut être formulé comme

$$\textbf{(P3.1):} \min_{\mathbf{\Theta}} \; E^{\mathsf{sum}} \tag{2.40}$$
$$\text{s.t. contraintes } (2.39\text{c}) - (2.39\text{h}), (2.39\text{l}).$$

On peut vérifier que le problème (2.40) est un programme linéaire mixte en nombres entiers (MILP) standard, qui peut être résolu efficacement en utilisant le outil CVX-Gurobi [38].

b) Optimisation du contrôle de délestage partiel et de l'allocation de bits sur des intervalles de temps

Étant donné $\{\mathbf{\Theta}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, le sous-problème optimisant le contrôle de délestage partiel et l'allocation de bits $\{\mathbf{\Lambda}, \mathbf{L}\}$ peut être formulé comme

$$\textbf{(P3.2):} \min_{\mathbf{\Lambda}, \mathbf{L}} \; E^{\mathsf{sum}} \tag{2.41}$$
$$\text{s.t. } (2.39\text{a}) - (2.39\text{d}), (2.39\text{f}), (2.39\text{m}), (2.39\text{n}).$$

On peut vérifier que le problème (2.41) est un problème linéaire (LP), il peut être résolu en utilisant le outil CVX-Gurobi [38].

c) Optimisation des ressources de calcul et de l'allocation de bande passante

Étant donné $\{\mathbf{\Theta}, \mathbf{\Lambda}, \mathbf{L}, \mathbf{Q}\}$, le sous-problème optimisant la ressource de calcul et l'allocation de bande passante $\{\mathbf{F}, \boldsymbol{\beta}\}$ peut être défini comme

$$\begin{aligned}
\textbf{(P3.3):}\ \min_{\mathbf{F},\boldsymbol{\beta}}\ & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{R^{\mathsf{u}}_{k,m}[n]}\Big) \\
& + E^{\mathsf{sum1}} \qquad (2.42)\\
\text{s.t.}\ & \text{contraintes } (2.39\text{c}), (2.39\text{d}), (2.39\text{g}), (2.39\text{h}), (2.39\text{n}),
\end{aligned}$$

où

$$\begin{aligned}
E^{\mathsf{sum1}} = \alpha_2\Big(&\sum_{k\in\mathcal{K}}(\kappa\lambda^{\mathsf{lo}}_k s_k c_k(f_k)^2+\frac{(1-\lambda^{\mathsf{lo}}_k-\lambda^{\mathsf{ed}}_k)s_k P^{\mathsf{s}}_k}{R^{\mathsf{s}}_k})\Big)+ \\
&\alpha_1\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n](l^{\mathsf{u}}_k[n]c_k e^{\mathsf{ed}})+\sum_m P^{\mathsf{f}}_m T\Big). \qquad (2.43)
\end{aligned}$$

Nous introduisons d'abord les variables auxiliaires

$$\xi_{k,m}[n] = R^{\mathsf{u}}_{k,m}[n] = \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u}}_k[n]}\right), \qquad (2.44)$$

où $B_{k,m}[n]{=}\frac{P^{\mathsf{u}}_k g_{k,m}[n]}{\sigma^2}$.

On peut vérifier que $\beta^{\mathsf{u}}_k[n]\log_2(1{+}\frac{B_{k,m}[n]}{\beta^{sfu}_k[n]})$ est une fonction concave de $\beta^{\mathsf{u}}_k[n]$. En utilisant la méthode d'approximation convexe successive (SCA), on peut calculer une borne supérieure de cette fonction concave en utilisant le développement de Taylor du premier ordre au point donné $\beta^{\mathsf{u},r}_k[n]$ dans la r-ième itération du processus d'approximation

$$\begin{aligned}
\beta^{\mathsf{u}}_k[n]\log_2\left(1{+}\frac{B_{k,m}[n]}{\beta^{\mathsf{u}}_k[n]}\right) \le\ & \beta^{\mathsf{u},r}_k[n]\log_2\left(1{+}\frac{B_{k,m}[n]}{\beta^{\mathsf{u},r}_k[n]}\right)+ \\
&\left(\log_2(1{+}\frac{B_{k,m}[n]}{\beta^{\mathsf{u},r}_k[n]})-\frac{\log_2(e)B_{k,m}[n]}{B_{k,m}[n]+\beta^{\mathsf{u},r}_k[n]}\right)(\beta^{\mathsf{u}}_k[n]-\beta^{\mathsf{u},r}_k[n]) \\
&\triangleq R^{\mathsf{ub}}_{k,m}[n]. \qquad (2.45)
\end{aligned}$$

Ensuite, le problème (2.42) peut être approximé par le problème suivant :

$$\begin{aligned}\textbf{(P3.3.2):} \min_{\mathbf{F},\boldsymbol{\beta},\boldsymbol{\Xi}} \quad & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{\xi_{k,m}[n]}\Big) \\ & + E^{\mathsf{sum1}} \qquad (2.46)\\ \text{s.t.} \quad & \sum_{m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\Big(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\xi_{k,m}[n]}\Big)\leq T^{\mathsf{max}}_k, \forall k, \qquad (2.46a)\\ & \theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\Big(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\xi_{k,m}[n]}\Big) \leq \Delta t, \forall k,m,n, \qquad (2.46b)\\ & \xi_{k,m}[n] \leq R^{\mathsf{ub}}_{k,m}[n], \forall k,m,n, \qquad (2.46c)\\ & \text{contraintes } (2.39g), (2.39h), (2.39n),\end{aligned}$$

où $\boldsymbol{\Xi} = \{\xi_{k,m}[n], \forall k, m, n\}$.

Puisque $\frac{1}{f^{\mathsf{u}}_k[n]}$ et $\frac{1}{\xi_{k,m}[n]}$ sont des fonctions convexes par rapport à $f^{\mathsf{u}}_k[n]$ et $\xi_{k,m}[n]$, on peut voir que la fonction objectif est convexe et que toutes les contraintes sont linéaires. Par conséquent, le problème (2.46) est un problème convexe, qui peut être résolu efficacement en utilisant le outil CVX-Gurobi [38].

d) Optimisation de la trajectoire multi-UAV

Étant donné $\{\boldsymbol{\Theta}, \boldsymbol{\Lambda}, \mathbf{L}, \mathbf{F}, \boldsymbol{\beta}\}$, le sous-problème optimisant le contrôle de trajectoire multi-UAV représentées par les variables $\mathbf{Q}$ peuvent être formulé comme

$$\begin{aligned}\textbf{(P3.4):} \min_{\mathbf{Q}} \quad & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{R^{\mathsf{u}}_{k,m}[n]}\Big) \\ & + E^{\mathsf{sum1}} \qquad (2.47)\\ \text{s.t.} \quad & \text{contraintes } (2.39c), (2.39d), (2.39i), (2.39j), (2.39k).\end{aligned}$$

Pour approximer ce problème, nous introduisons des variables auxiliaires $\gamma_{k,m}[n] = R^{\mathsf{u}}_{k,m}[n]$ et $S_{k,m}[n] \leq H^2 + \|\mathbf{q}_m[n] - \mathbf{r}^{\mathsf{u}}_k\|^2$ et on a

$$\begin{aligned}\gamma_{k,m}[n] &= \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2(H^2+\|\mathbf{q}_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2)}\right)\\ &\leq \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2 S_{k,m}[n]}\right). \qquad (2.48)\end{aligned}$$

On peut vérifier que $\beta_k^{\mathsf{u}}[n]\log_2\left(1+\frac{R_k[n]}{S_{k,m}[n]}\right)$ est une fonction convexe par rapport à $S_{k,m}[n]$, où $R_k[n]{=}\frac{P_k^{\mathsf{u}}\rho_0}{\beta_k^{\mathsf{u}}[n]\sigma^2}$. Afin de pouvoir utiliser la méthode SCA, on peut calculer une borne inférieure (2.48) en utilisant le développement de Taylor du premier ordre au point donné $S_{k,m}^r[n]$ dans la r-ième itération du processus d'approximation. On trouve

$$\beta_k^{\mathsf{u}}[n]\log_2\left(1{+}\frac{R_k[n]}{S_{k,m}[n]}\right)\geq\beta_k^{\mathsf{u}}[n]\Bigg(\log_2\left(S_{k,m}[n]{+}R_k[n]\right)\\ -\log_2(S_{k,m}^r[n]){-}\frac{\log_2(e)}{S_{k,m}^r[n]}(S_{k,m}[n]{-}S_{k,m}^r[n])\Bigg)\triangleq R_{k,m}^{\mathsf{lb}}[n]. \tag{2.49}$$

Par conséquent, le problème d'optimisation (2.47) peut être approximé par le problème suivant:

$$\textbf{(P3.4.2):}\ \min_{\mathbf{Q},\mathbf{\Gamma},\mathbf{S}}\ \alpha_2\Bigg(\sum_{k,m,n}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]l_k^{\mathsf{u}}[n]P_k^{\mathsf{u}}\frac{1}{\gamma_{k,m}[n]}\Bigg)\\ +E^{\mathsf{sum1}} \tag{2.50}$$

$$\text{s.t.}\sum_{m,n}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}{+}\frac{l_k^{\mathsf{u}}[n]}{\gamma_{k,m}[n]}\right)\leq T_k^{\mathsf{max}},\forall k, \tag{2.50a}$$

$$\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}{+}\frac{l_k^{\mathsf{u}}[n]}{\gamma_{k,m}[n]}\right)\leq\Delta t,\forall k,m,n, \tag{2.50b}$$

$$\gamma_{k,m}[n]\leq R_{k,m}^{\mathsf{lb}}[n],\forall k,m,n, \tag{2.50c}$$

$$S_{k,m}[n]\leq\|\mathbf{q}_m^r[n]{-}\mathbf{r}_k^{\mathsf{u}}\|^2+2\left(\mathbf{q}_m^r[n]{-}\mathbf{r}_k^{\mathsf{u}}\right)^T\left(\mathbf{q}_m[n]{-}\mathbf{q}_m^r[n]\right)+H^2,\forall k,m,n, \tag{2.50d}$$

$$d_{\min}^2\leq-\left\|\mathbf{q}_m^r[n]{-}\mathbf{q}_j^r[n]\right\|^2+2\left(\mathbf{q}_m^r[n]{-}\mathbf{q}_j^r[n]\right)^T\left(\mathbf{q}_m[n]{-}\mathbf{q}_j[n]\right),\forall j{\neq}m,n, \tag{2.50e}$$

$$\text{constraints (2.39i), (2.39j),}$$

où $\mathbf{\Gamma}{=}\{\gamma_{k,m}[n],\forall k,m,n\}$, $\mathbf{S}=\{S_{k,m}[n],\forall k,m,n\}$.

Puisque $\frac{1}{\gamma_{k,m}[n]}$ est une fonction convexe par rapport à $\gamma_{k,m}[n]$, la fonction objectif est convexe. De plus, toutes les contraintes sont linéaires. Par conséquent, le problème (2.50) est un problème convexe, qui peut être résolu efficacement en utilisant le outil CVX-Gurobi [38].

e) **Algorithme de planification intégrée des utilisateurs, contrôle de délestage partiel, ressource de calcul, allocation de bande passante, et contrôle de trajectoire multi-UAV**

Algorithm 2.6. Algorithme de planification intégrée des utilisateurs, délestage partiel, calcul, allocation de bande passante, et contrôle de trajectoire multi-UAV

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, et les emplacements des GU, des satellites et du serveur cloud;
1: **Initialisation:** $\mathbf{L}^0, \mathbf{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0$;
Ensure: Consommation d'énergie pondérée minimale (E^{sum}); Let $r = 1$;
2: **repeat**
3: Résoudre le sous-problème (2.40) pour obtenir $\mathbf{\Theta}^r$;
4: Résoudre le sous-problème (2.41) pour obtenir $\mathbf{L}^r$ et $\mathbf{\Lambda}^r$;
5: Résoudre le sous-problème (2.46) pour obtenir $\boldsymbol{\beta}^r$ et $\mathbf{F}^r$;
6: Résoudre le sous-problème (2.50) pour obtenir $\mathbf{Q}^r$;
7: Mise à jour $r = r + 1$;
8: **until** Convergence
9: Retour $E^{\mathsf{sum},*}, \mathbf{\Theta}^*, \mathbf{L}^*, \mathbf{\Lambda}^*, \mathbf{F}^*, \boldsymbol{\beta}^*, \mathbf{Q}^*$.

En utilisant les résultats ci-dessus, nous pouvons développer un algorithme intégré basé sur la méthode d'optimisation alternée décrite dans l'algorithme 2.6. La convergence de cet algorithme est énoncée dans la proposition suivante.

Proposition 2.2. L'algorithme 2.6 crée une séquence de solutions réalisables où la valeur fonction objectif diminue de manière monotone au fil des itérations. En conséquence, l'algorithme converge vers une solution réalisable.

2.2.3.4 Conception conjointe du contrôle d'admission et de la gestion du réseau

Si le problème **(P3)** est faisable alors l'algorithme 2.6 converge vers une solution faisable. Cependant, le problème **(P3)** peut être irréalisable dans certains scénarios surchargés. À cette fin, nous développons un algorithme pour vérifier la faisabilité du problème **(P3)** et proposons un algorithme conjoint de contrôle d'admission et de gestion de réseau pour résoudre le problème **(P3)** dans un scénario générique où ce problème peut être réalisable ou non.

a) Vérification de faisabilité

Nous abordons la vérification de faisabilité pour le problème **(P3)** dans cette section. Nous introduisons une nouvelle variable δ et l'utilisons pour toutes les contraintes d'inégalité du problème **(P3)** et considérons un problème d'optimisation formulé comme

$$\textbf{(P3'):} \min_{\boldsymbol{\Theta},\boldsymbol{\Lambda},\mathbf{L},\boldsymbol{\beta},\mathbf{F},\mathbf{Q},\delta} \delta \tag{2.51}$$

$$\text{s.t. } T_k^{\mathsf{lo}} - T_k^{\mathsf{max}} - \delta \leq 0, \forall k, \tag{2.51a}$$

$$T_k^{\mathsf{cl}} - T_k^{\mathsf{max}} - \delta \leq 0, \forall k, \tag{2.51b}$$

$$\sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] - T_k^{\mathsf{max}} - \delta \leq 0, \forall k, \tag{2.51c}$$

$$\theta_k[n] T_{k,m}^{\mathsf{ed}}[n] - \Delta t - \delta \leq 0, \forall k, m, n, \tag{2.51d}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] \beta_k^{\mathsf{u}}[n] - W_m^{\mathsf{u}} - \delta \leq 0, \forall m, n, \tag{2.51e}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] f_k^{\mathsf{u}}[n] - F_m^{\mathsf{max}} - \delta \leq 0, \forall m, n, \tag{2.51f}$$

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 - D_{\max}^2 - \delta \leq 0, \forall m, n = 1, ..., N-1, \tag{2.51g}$$

$$d_{\min}^2 - \|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 - \delta \leq 0, \ \forall n, m, j \neq m, \tag{2.51h}$$

$$\text{contraintes (2.39e), (2.39f), (2.39i), (2.39l), (2.39m), (2.39n).}$$

Notez que ce problème est réalisable et qu'il existe une valeur optimale de δ qui peut être utilisée pour déterminer la faisabilité du problème **(P3)** comme suit. Spécifiquement, le problème **(P3)** est réalisable (c'est-à-dire que toutes les contraintes sont satisfaites) si $\delta \leq 0$ et il est irréalisable, sinon. Cependant, le problème **(P3')** est aussi un problème d'optimisation non linéaire mixte en nombres entiers (MINLP), difficile à résoudre de manière optimale. En utilisant l'approche d'optimisation d'alternance similaire décrite dans la section 2.2.3.3, après chaque sous-problème, nous pouvons obtenir les valeurs de δ qui seraient requises pour vérifier si le problème est faisable ou non. Le résumé de l'algorithme de vérification de faisabilité est décrit dans ce qui suit.

- **Algorithme de vérification de faisabilité:**

  Un résumé de l'algorithme de vérification de faisabilité est donné dans l'algorithme 2.7. Initialement, nous définissons *feasibility* = *true* et initialisons toutes les variables $\boldsymbol{\omega}^0 = \{\boldsymbol{\Theta}^0, \mathbf{L}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0\}$. Ensuite, nous appliquons la méthode d'optimisation alternée et résolvons itérativement chaque ensemble de variables en fonction des valeurs des autres variables jusqu'à convergence vers une valeur stable de δ^* comme décrit de l'étape 1 à l'étape 19. Spécifiquement, après avoir résolu chaque sous-problème dans la r-ième itéra-

Algorithm 2.7. Algorithme de vérification de faisabilité

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, et les emplacements des GU, des satellites et du serveur cloud ;
Ensure: Min δ; Let $r = 1$; *feasibility* = *true*, et $\boldsymbol{\omega}^0 = \{\boldsymbol{\Theta}^0, \mathbf{L}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0\}$;

1: **repeat**
2: Résoudre le sous-problème (7.36) pour obtenir δ^r et $\boldsymbol{\Theta}^r$;
3: **if** $\delta^r > 0$ **then**
4: Retour δ^r et $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^{r-1}, \boldsymbol{\Lambda}^{r-1}, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}$,
Sortie de boucle **repeat**;
5: **end if**
6: Résoudre le sous-problème (7.37) pour obtenir δ^r, $\mathbf{L}^r$, et $\boldsymbol{\Lambda}^r$;
7: **if** $\delta^r > 0$ **then**
8: Retour δ^r et $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^{r}, \boldsymbol{\Lambda}^{r}, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}$,
Sortie de boucle **repeat**;
9: **end if**
10: Résoudre le sous-problème (7.38) pour obtenir δ^r, $\boldsymbol{\beta}^r$, et $\mathbf{F}^r$;
11: **if** $\delta^r > 0$ **then**
12: Retour δ^r et $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^{r}, \boldsymbol{\Lambda}^{r}, \mathbf{F}^{r}, \boldsymbol{\beta}^{r}, \mathbf{Q}^{r-1}\}$,
Sortie de boucle **repeat**;
13: **end if**
14: Résoudre le sous-problème (7.39) pour obtenir δ^r et $\mathbf{Q}^r$;
15: **if** $\delta^r > 0$ **then**
16: Retour δ^r et $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^{r}, \boldsymbol{\Lambda}^{r}, \mathbf{F}^{r}, \boldsymbol{\beta}^{r}, \mathbf{Q}^{r}\}$,
Sortie de boucle **repeat**;
17: **end if**
18: Mise à jour $r = r + 1$;
19: **until** Convergence δ^*
20: **if** $\delta^* \leq 0$ or $\delta^r \leq 0$ **then**
21: *feasibility* = *true*;
22: **else**
23: *feasibility* = *false*;
24: **end if**
25: **Production** *résultat de feasibility* et
$$\boldsymbol{\omega}^* \in \Big\{\{\boldsymbol{\Theta}^r, \mathbf{L}^{r-1}, \boldsymbol{\Lambda}^{r-1}, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^{r}, \boldsymbol{\Lambda}^{r}, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^{r}, \boldsymbol{\Lambda}^{r}, \mathbf{F}^{r}, \boldsymbol{\beta}^{r}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^{r}, \boldsymbol{\Lambda}^{r}, \mathbf{F}^{r}, \boldsymbol{\beta}^{r}, \mathbf{Q}^{r}\}\Big\};$$

tion, nous vérifions la valeur objectif obtenu δ^r comme suit : si $\delta^r > 0$, nous renvoyons cette valeur et $\boldsymbol{\omega}^r$ et puis sortie de boucle *"repeat-until"*. Sinon, si $\delta^r \leq 0$, on continue à résoudre le sous-problème suivant. Les étapes 20 à 24 vérifient la valeur obtenue de δ^* et génèrent le *résultat de feasibility* à l'étape 25. Si *feasibility* = *true*, toutes les contraintes du problème **(P3)** sont satisfaites et on peut résoudre le problème d'optimisation considéré pour obtenir une solution réalisable. Sinon, si *feasibility* = *false*, le problème **(P3)** est irréalisable (c'est-à-dire que certaines contraintes du problème **(P3)** ne peuvent pas être satisfaites). Les sorties de l'algorithme 2.7 sont des résultats *feasibility* et $\boldsymbol{\omega}^*$.

b) Algorithme de contrôle d'admission et de gestion de réseau

Algorithm 2.8. Algorithme conjoint de contrôle d'admission et de gestion de réseau

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, et les emplacements des GU, des satellites et du serveur cloud ;
Initialisation: $\mathbf{L}^0, \mathbf{\Lambda}^0, \mathbf{F}^0, \beta^0, \mathbf{Q}^0, \mathcal{K}^{\mathsf{ac}} = \mathcal{K}$;
Ensure: Consommation d'énergie pondérée minimale (E^{sum});
Trouver le nombre de sauts L par en exécutant l'algorithme 1 dans [2];
Déterminer la bande passanteh W^{u}_m, i.e., $\sum_{m\in\mathcal{M}} W^{\mathsf{u}}_m \leq W$;
Let *feasibility* = *true*;
repeat
  Exécuter la vérification de faisabilité l'algorithme 2.7;
  if *feasibility* = *true* **then**
    Exécuter l'algorithme 2.6
    Sortie de boucle **repeat**;
  else
    Donné $\boldsymbol{\omega}^*$ obtenu à partir de l'algorithme 2.7, calculer T_k basé sur (2.28);
    Trouver le pire GU $k = \text{argmax}_{k\in\mathcal{K}} \left(T_k/T_k^{\mathsf{max}}\right)$;
    Assigner $\mathcal{K}^{\mathsf{ac}} = \mathcal{K} \setminus \{k\}$;
    Mise à jour $\mathcal{K} \leftarrow \mathcal{K}^{\mathsf{ac}}$;
  end if
until $\mathcal{K} = \emptyset$
Retour $E^{\mathsf{sum},*}, \mathcal{K}^{\mathsf{ac}}, \mathbf{\Theta}^*, \mathbf{L}^*, \mathbf{\Lambda}^*, \mathbf{F}^*, \beta^*, \mathbf{Q}^*$.

Pour le problème **(P3)**, les contraintes de délai maximum et les contraintes nécessitant que les tâches partielles des GU soient traitées complètement par les UAV associés dans chaque tranche de temps sont difficiles à satisfaire. Nous proposons une stratégie qui supprime de manière itérative à chaque étape le pire GU, c'est-à-dire celui qui nécessite la plus grande quantité de ressources pour satisfaire sa contrainte de délai. Spécifiquement, étant donné la sortie de l'algorithme 2.7, le temps total de transmission, de propagation et de traitement des tâches pour chaque GU pourrait être calculé comme dans (2.28). Ensuite, nous trouvons le GU k qui atteint la valeur maximale de T_k/T_k^{max}, le supprimons et mettons à jour l'ensemble de GU restants $\mathcal{K}^{\mathsf{ac}}$ en supprimant ce GU de l'ensemble $\mathcal{K}$. Nous proposons un algorithme conjoint de contrôle d'admission et de gestion de réseau pour résoudre le problème **(P3)** afin d'atteindre la consommation d'énergie pondérée minimale des GU et des UAV comme dans l'algorithme 2.8.

c) Initialisation de l'algorithme

- **Trajectoire circulaire initiale des UAV:**

  Les trajectoires circulaires des UAV pour desservir des groupes de GU sont considérées et initialisées comme dans la section 7.6.4.1.

- **Variables initiales de contrôle de délestage partiel, de ressource de calcul et d'allocation de bande passante:**

Les valeurs de la taille des tâches s_k sont définies de manière aléatoire dans une plage de $[1, 10]$Mbits et les valeurs du délai maximal tolérable T_k^{max} sont également définies de manière aléatoire dans une plage de $[1, 3]$(secondes). De plus, les valeurs initiales des variables de contrôle de délestage partiel sont générées aléatoirement dans $\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}} \in [0, 0.5]$, et nous utilisons une allocation uniforme de bit, de ressource de calcul, et de bande passante $l_k^{\mathsf{u}}[n] = \lambda_k^{\mathsf{ed}} s_k/N_k$, $f_k^{\mathsf{u}}[n] = MF_m^{\mathsf{max}}/K$, et $\beta_k^{\mathsf{u}}[n] = W/K$.

Pour étudier l'efficacité des algorithmes proposés, nous considérons le cas de référence, “early scheduling”, où tous les GU sont programmés en continu à partir du premier créneau horaire de la période de vol de l'UAV. Dans le cas de référence, appelée “baseline edge”, nous avons initialement défini des trajectoires circulaires des UAV pour desservir les groupes correspondants de GU, les valeurs des variables de commande de délestage partiel sont définies de manière aléatoire avec l'allocation uniforme de bit, de ressource de calcul et de bande passante décrite plus haut. À titre de comparaison, la stratégie “optimized edge” représente notre conception où toutes les variables sont optimisées.

2.2.3.5 Résultats numériques

Nous considérons différents scénarios dans lesquels un serveur cloud est loin du réseau: les GU sont situés à Montréal (45.50 °N, 73.56 °O) tandis que le serveur cloud est situé à Vancouver (49.28 °N, 123.12 °O). En exécutant l'algorithme 1 dans [2], nous pouvons déterminer le nombre de sauts de satellite $L = 4$. Les paramètres de nos simulations sont définis de manière similaire à ceux de [44, 46–48] et les valeurs choisies des paramètres clés sont résumées dans tableau 7.3.

La figure 2.5 illustre la somme pondérée d'énergie pour différents nombres de GU, 2 UAV, $W = 10$ MHz, $L = 4$, et $T = [10, 15]$s. On peut voir que la somme pondérée d'énergie devient plus élevée avec un plus grand nombre de GU et l'algorithme proposé atteint la plus petite somme pondérée d'énergie par rapport à celles des cas de référence dans les deux scénarios avec $T = [10, 15]$s. Pour 18 GU, la somme pondérée d'énergie peut être réduite de 18.05% et de 9.64% par rapport aux valeurs correspondantes du “early scheduling” et de la “baseline edge” avec $T = 10$s et $T = 15$s. La figure 2.6 illustre la répartition de la charge de calcul sur les couches du réseau. Cette figure montre que des tâches de grande taille entraînent une charge de calcul inférieure répartie sur les GU, tandis

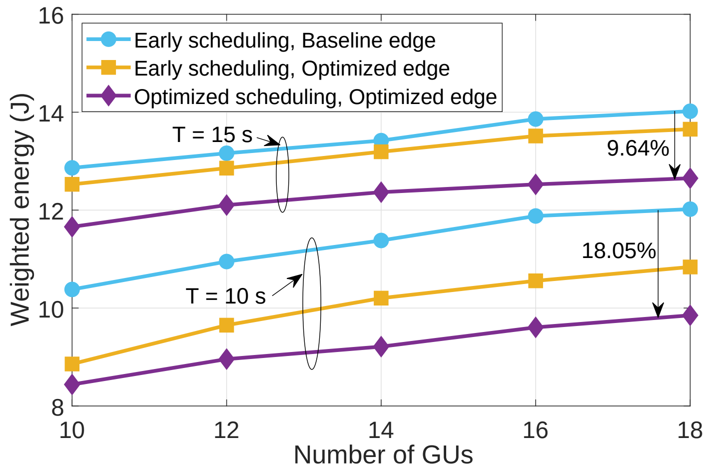

Figure 2.5 – Somme pondérée d'énergie pour différents nombres de GU.

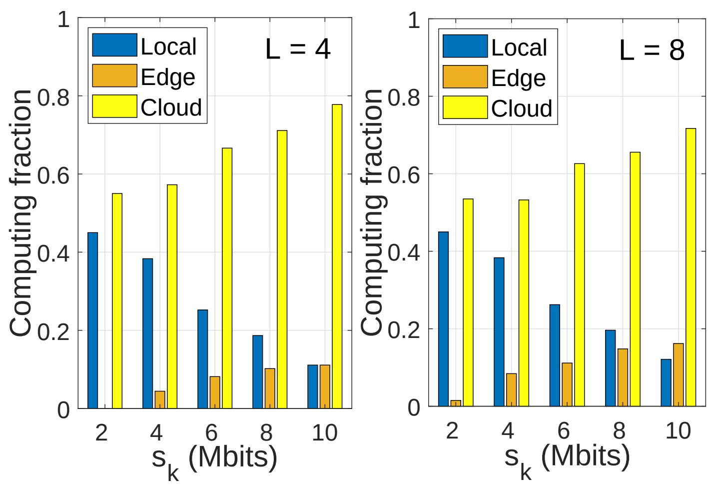

Figure 2.6 – Répartition de calcul pour différentes valeurs de taille de tâche.

qu'un nombre de sauts satellite plus important entraîne une charge de calcul plus élevée à traiter sur les serveurs périphériques.

Pour évaluer les performances obtenues par la conception de contrôle d'admission proposée, nous définissons un débit d'admission comme le rapport entre le nombre de GU réelles desservies et le nombre total de GU, c'est à dire $\frac{|\mathcal{K}^{\mathsf{ac}}|}{|\mathcal{K}_0|}$, où $\mathcal{K}_0$ désigne l'ensemble de GU d'origine et $\mathcal{K}^{\mathsf{ac}}$ représente l'ensemble de GU admises pour lesquelles une solution réalisable peut être trouvée. La figure 2.7 montre le débit d'admission pour différents nombre de GU pour les réseaux avec 2 et 3 UAV, $W = 10$ MHz, $L = 4$, et $T = [10, 15]$s. On constate que le débit d'admission diminue à mesure que le nombre de GU augmente. En effet, compte tenu des ressources radio et de calcul fixes, le nombre de GU que le réseau peut supporter est limité. Par conséquent, un plus grand nombre de GU serait retiré du système à mesure que le nombre de GU augmenterait, ce qui entraînerait une diminution du

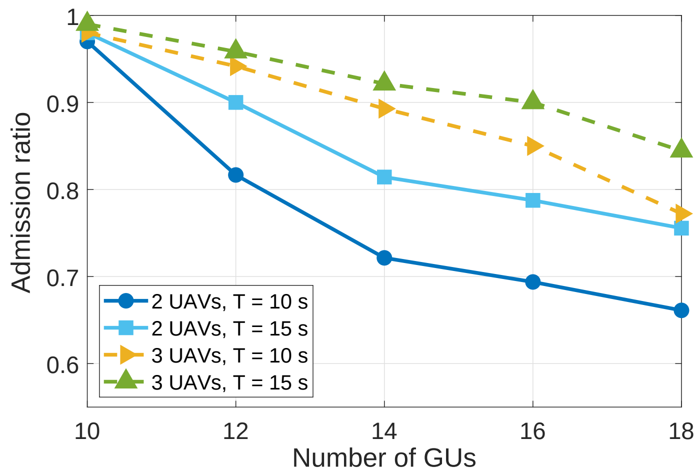

Figure 2.7 – Débit d'admission pour différents nombre de GU.

débit d'admission. On peut également voir que la différence dans les ratios d'admission pour les deux scénarios avec $T = 10$s et $T = 15$s est plus grande pour le cas de 2 UAV que pour 3. En fait, pour la configuration du réseau avec un plus grand nombre d'UAV (c'est-à-dire des serveurs de périphérie) et une plus grande période de vol de l'UAV T, le réseau peut être mieux couvert; par conséquent, un plus grand nombre de GU peut être desservi.

2.3 Remarques finales

Dans cette thèse de doctorat, nous avons développé diverses de nouvelles techniques de planification et de gestion des réseaux UAV pour les futurs réseaux sans fil. Spécifiquement, nous avons apporté trois contributions importantes à la recherche. Tout d'abord, nous étudions contrôle de trajectoire UAV intégré et l'allocation de ressources pour les réseaux sans fil basés sur UAV avec gestion des interférences dans le même canal. Deuxièmement, nous considérons le placement de l'UAV et l'allocation de ressources pour les réseaux sans fil basés sur des UAV avec surfaces réfléchissantes intelligentes. Troisièmement, nous étudions délestage des calculs, le contrôle de trajectoire UAV, la planification des utilisateurs, l'allocation de ressources et le contrôle d'admission dans SAGIN.

Chapter 3

Introduction

3.1 Background and Motivations

Next-generation wireless networks will enable to support applications in various domains including smart factories, intelligent transportation, e-health, and more [3, 4]. The proliferation of many human and Internet of Things (IoT) applications have led to a mobile traffic explosion. In fact, it is predicted by Ericsson that total mobile traffic volume can reach 131 exabytes per month by the end of 2024 [49]. Moreover, the recent forecast shows that billions of wireless devices, from low-cost IoT devices and wearables to virtual/augmented/mixed reality devices and smart vehicles, will be connected to wireless networks over the next few years [50, 51]. Therefore, future wireless communications are expected to provide higher capacity and much lower latency, offer enhanced stability, ubiquitous communications, and connectivity to billions of devices [5–8]. However, the deployment of terrestrial infrastructure faces challenges in various practical scenarios, such as communications to serve temporary events and emergencies like natural disasters and fast service recovery [9–11].

Toward this end, several promising technologies have been under consideration, including satellite communications, unmanned aerial vehicle (UAV) communications, intelligent reflecting surface (IRS), and mobile edge computing (MEC) [12, 13]. In particular, UAV communications have emerged as a potential solution to overcome the limitations of current infrastructure, offering wider coverage, higher resilience, and availability, and improving user's quality of service (QoS) due to their superior attributes such as mobility, flexibility, and adaptive altitude [14, 15]. Besides, the

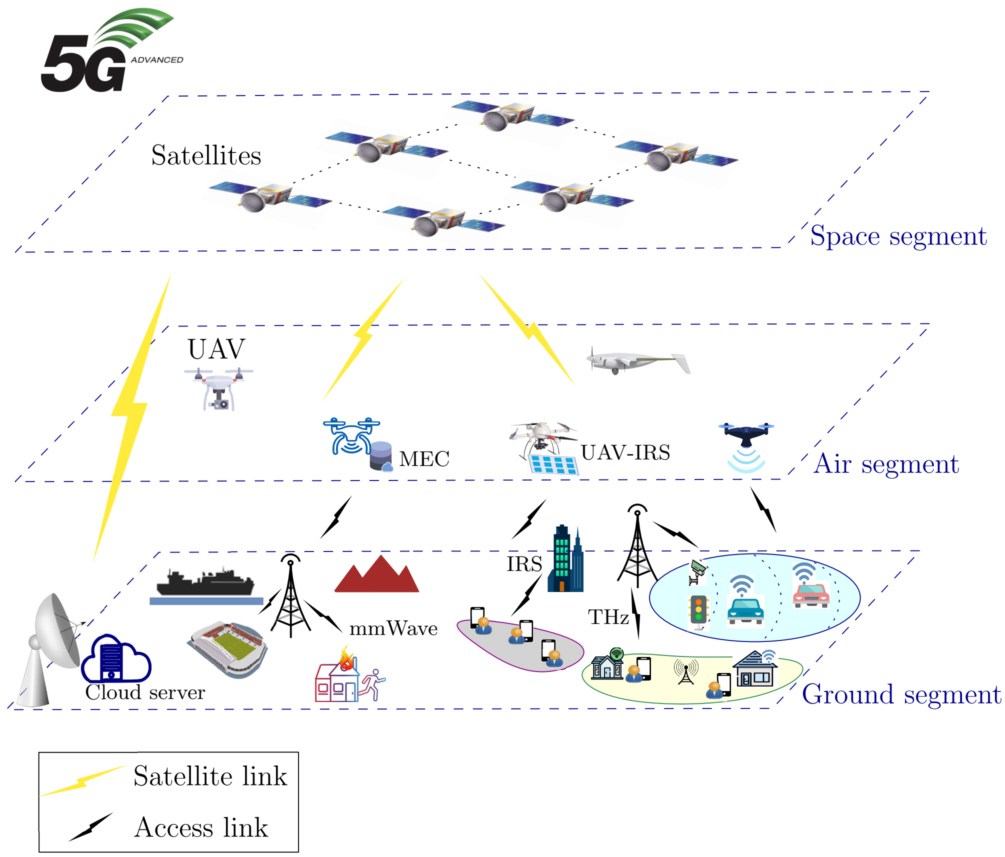

Figure 3.1 – Next-generation wireless networks.

IRS-assisted UAV communications have attracted extensive attention because they can significantly enhance the communication quality. In this system, UAV communicates with ground users (GUs) and IRS can reflect the dissipated signals from the UAV, improving the UAV-GU communications quality [16–18]. Moreover, space-air-ground integrated networks (SAGIN) have emerged as promising architecture to provide high-quality and ubiquitous communications by leveraging the complementary strengths of space, air, and ground networks and enabling the technologies such as edge computing [19–21].

One typical network architecture for the fifth-generation (5G) and beyond wireless system is shown in Fig. 3.1, which employs various enabling technologies. Besides, wireless applications divided into different relevant groups are described in Fig. 3.2. In the following, we discuss these enabling technologies in more details.

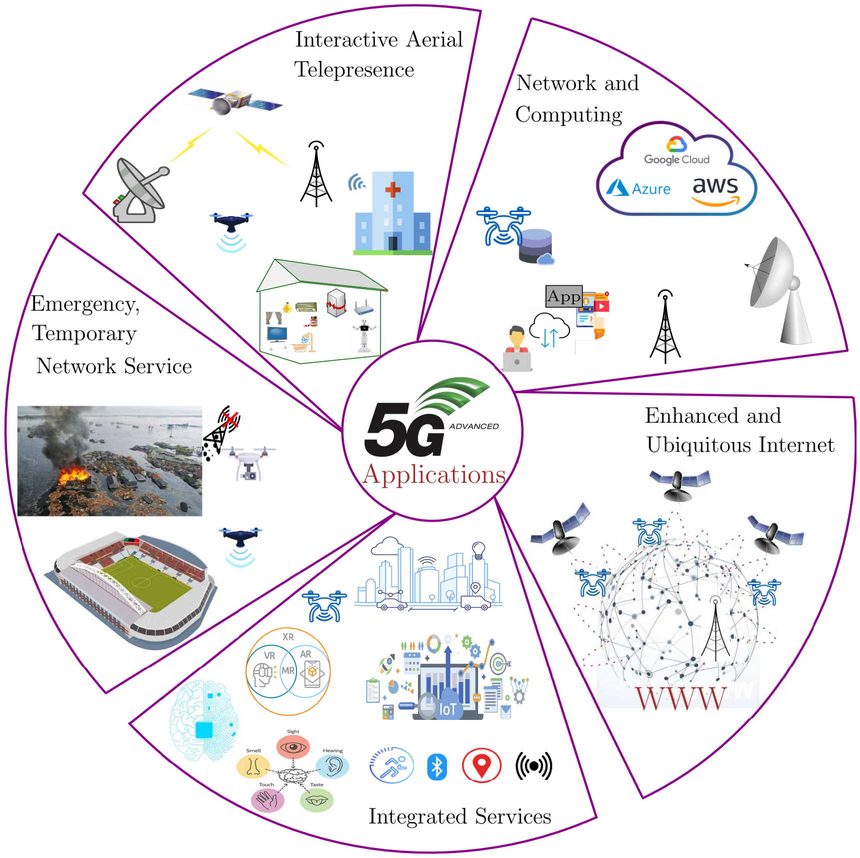

Figure 3.2 – Applications in next-generation wireless networks.

3.1.1 UAV Communications

There has been strong interest in providing wireless coverage in the three-dimensional (3D) space and leveraging different flying platforms to enhance wireless connectivity and/or the performance of the terrestrial wireless networks [4, 10, 22, 23]. UAV communications can provide low-cost solutions for various communications scenarios, e.g., wireless areas with limited infrastructure or high traffic demand. Moreover, the UAV-based wireless networks can provide extra degrees of freedom to optimize the underlying wireless network to enhance the coverage, throughput, and energy efficiency thanks to unique UAV's attributes such as mobility, flexibility, and controllable altitude. With appropriate deployment, UAV communications can provide favorable line-of-sight (LoS) communications with GUs [52]. UAV communications can also be leveraged to support various Internet of Things (IoT) applications such as data dissemination, or data collection [53]. Therefore, UAV-based wireless networks are expected to play an important role in 5G and beyond wireless systems [54].

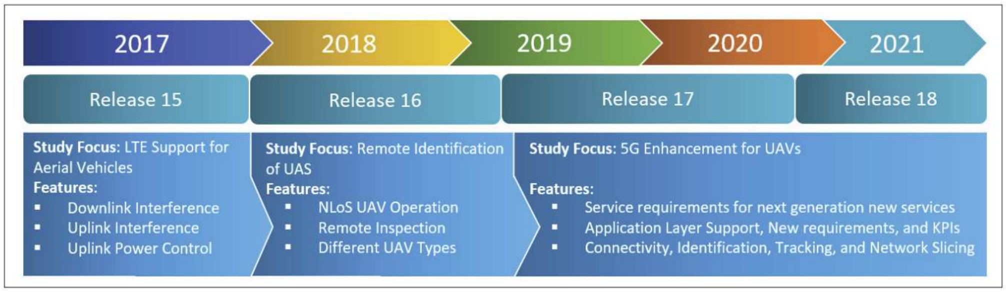

Figure 3.3 – High-level timeline for 3GPP Releases for UAVs [1].

Third Generation Partnership Project (3GPP) considers UAV communication as an essential technology for the next-generation cellular system. Specifically, the suggested timeline for 3GPP releases for UAVs was discussed in Fig. 3.3. In 2017, 3GPP conducted several studies and issued Rel-15 to acknowledge LTE-empowered UAVs [55]. These studies focused on UAV traffic requirements, channel modeling for air-to-ground propagation channels, and utilization of current cellular networks to support UAV communications. In addition, these studies introduced the signaling for subscription-based aerial user identification, reporting UAV height, location, speed, and flight path, and measurement data to address aerial interference for certain density of low-altitude UAVs. In subsequent releases [56–59], 3GPP addressed application layer support and security for connected UAVs, also defined the service interactions between UAVs and the UAV traffic management (UTM) system. As 5G use-cases evolve, Rel-18 will consider 5G NR support for devices onboard aerial vehicles, studying additional triggers for conditional handover, base station uptilting, and signaling to indicate UAV beamforming capabilities, among other enhancements [60].

Generally, UAVs can act as mobile users, relays, or flying base stations (BSs), which can be leveraged to enhance the coverage and capacity of wireless networks. There has been a great deal of research on these UAV communications scenarios in recent years. In particular, research on data collection for wireless nodes in wireless sensor networks (WSNs) and Internet of Things (IoT) leveraging the UAV communication has been an active research topic [61–64]. There has also been much work for UAV-enabled wireless networks in which UAVs act as relays [65–69]. Finally, UAVs can serve as aerial BSs to provide on-the-fly communications and enhance the performance of the terrestrial wireless networks [70, 71]. In fact, UAV placement and trajectory control optimization have been studied in [41, 72–77]. Various studies [78, 79] have shown that UAV communications can be employed to improve system performances such as coverage, throughput, and energy efficiency. Specifically, UAVs's flying trajectories can be efficiently controlled and optimized to provide reliable and line-of-sight (LoS) wireless connectivity for ground users. Moreover, a sufficiently large number of UAVs should be deployed and properly controlled to cope with dynamic traffic [80]. Joint optimization of UAVs' trajectories and resource allocation can potentially support more users with limited radio resources, which is highly desirable in various application scenarios such as massive connectivity and data dissemination for Internet of Things (IoT).

The number of IoT devices has been increasing rapidly in recent years [51]. Energy-efficient design for IoT networks is very important because it helps elongate the lifetime of IoT devices and networks [81]. Energy-efficient deployment of fixed base stations can be challenging in certain scenarios such as remote areas with limited infrastructure or disaster response and recovery. Some recent work [82–84] show that UAV communications can provide promising solutions to improve the energy efficiency of IoT wireless networks. Succinctly, efficient design of UAVs' trajectories or placement can leverage the LoS communications between the IoT devices and UAVs, which significantly improve the energy efficiency of the IoT devices and networks. However, joint design of resource allocation and multi-UAV trajectories control still requires much further research.

3.1.2 Intelligent Reflecting Surface Assisted Wireless Communications

Intelligent reflecting surface (IRS) or reconfigurable intelligent surface (RIS) is a promising paradigm which can substantially improve the spectral and energy efficiency of wireless networks by constructing favorable communication channels via tuning massive low-cost passive reflecting elements [24]. In essence, an IRS consists of a large number of low-cost passive elements, where each element can be adjusted with an independent phase shift to reflect the electromagnetic incident signals, to be added coherently at GUs. In particular, each of these elements can be reconfigured via amplitudes and phase shifts, thus modifying the propagation of incident signals. By optimizing the reflection coefficients of the IRS, reflected signals can be combined coherently with the non-reflecting signal to enhance the desired signal strength or destructively to suppress interference [85]. More detailed studies about physics, propagation, and path loss models of the IRS were considered in [26, 86–88]. The 3GPP has had several discussions about the IRS in the Rel-18 workshops [RWS-210247, 0300, 0306, 0390, 0361, 0465] [89]. Nevertheless, the IRS technology still needs to be further studied before it can be adopted for future wireless networks such as 6G and beyond.

IRS can be flexibly deployed on various structures, such as building facades, roadside billboards, and indoor walls [25]. IRS-assisted wireless communications can be realized by deploying the IRS between the BS or aerial BS and mobile users to enhance the received signal power [26–30]. In some applications, it is also helpful to enhance the incoming signal from the BS to the distant users using IRS, especially when the users are in the non-line-of-sight (NLoS) region. However, the network coverage enhanced by the IRS technology is limited by the fact that the surface is composed of passive elements; this is the key difference between IRS and active amplify-and-forward (AF) relays. Nevertheless, IRS is effective for indoor applications, wherein it is mounted on the walls of buildings to provide an additional link between the BS and indoor mobile users that may be inaccessible via LoS communications provided by conventional network structures. Furthermore, deployment of IRS-assisted UAV-based wireless networks to improve the performance of the network is a new research direction attracting growing research attention.

3.1.3 Space-Air-Ground Integrated Networks

Space-air-ground integrated networks (SAGIN) have emerged as an effective means of providing high quality and ubiquitous communications by leveraging the complementary strengths of space, air, and ground networks segments [31–33]. On the one hand, in the space network of SAGIN, geostationary earth orbit (GEO) satellites, medium earth orbit (MEO) satellites, and low earth orbit (LEO) satellites are the main components [34]. LEO satellites are liable to form networks by inter-satellite links (ISLs), which guarantee lower propagation delay, high communication rates, and seamless communication services for wide geographical areas [35, 36]. A notable example of LEO satellite network is Starlink of SpaceX, which plans to launch more than 12, 000 LEO satellites to offer global seamless service for terrestrial users, where the lower group (7, 518 satellites) will operate at 340 km altitude, while the higher group (4, 425 satellites) will operate at 1, 100 km altitude [19, 90]. Satellite communications are, therefore, vital in providing communications for areas where ground base stations are not available or damaged by natural disasters [47, 91, 92].

On the other hand, in the air network of SAGIN, there is a mobile aerial system that uses aircraft as carriers for information acquisition, transmission, and processing. The UAVs, airships, and balloons are the main infrastructures making up the high and low-altitude platforms (HAPs & LAPs) which can provide broadband wireless communications complementing the terrestrial networks [31, 37]. Compared with the terrestrial network employing base stations, the air network has the features of low cost, easy deployment, and large coverage offering wireless access services on a regional basis. Meanwhile, in the ground network of SAGIN, the network mainly consists of terrestrial communication systems such as cellular networks, mobile ad hoc networks, wireless local area networks, and so on [31]. Specifically, the cellular/mobile network has evolved to advanced wireless architectures such as 5G and beyond. However, deployment of many emerging delay-sensitive and computation-extensive applications on wireless devices is still challenging due to their limited energy and computing resources. To tackle this challenge, mobile edge computing (MEC) is a promising approach, which allows mobile users to offload their computation-intensive tasks to the edge servers [93, 94]. In addition, a UAV-assisted MEC system enables efficient support for computation-extensive mobile applications thanks to controllable UAVs' trajectories, extensive coverage and additional computation capability. Therefore, the SAGIN system is very effective in providing higher quality, lower latency, and better communication reliability and stability.

In fact, 3GPP has developed a set of specifications for non-terrestrial networks (NTN), which include unmanned aerial vehicles (UAVs), high altitude platforms (HAPs), and satellite networks that are traditionally used for certain applications such as disaster management, navigation, television broadcasting, and remote sensing [3, 95, 96]. Specifically, the studies on satellite access have been presented in [97–99], which identified use-cases for the service provisioning considering the integration of 5G satellite-based access components and specified enhancements for radio frequency and physical layer, protocols, radio resource management, and frequency bands. In the end, these studies identified new services, corresponding requirements, and suitable architectures. Moreover, the 3GPP Rel-18 specifications will enhance 5G NR NTN operation by studying the enablers for NR-based satellite access in bands above 10 GHz to serve fixed and moving platforms, e.g., aircraft, UAVs, as well as building-mounted devices, e.g., businesses and premises. The aim of these efforts is to further optimize satellite access performance, address new bands with their specific regulatory requirements, and support new capabilities and services as the evolution of 5G continues [100].

The overall objective of this dissertation is to study network planning and resource management leveraging UAV communications for future wireless networks. Toward this end, we consider three fundamental design aspects, namely joint UAV deployment and resource allocation, IRS-assisted UAV-based wireless networks, and computation offloading and resource allocation in SAGIN. In the following, we discuss research challenges and perform literature survey related to these research directions. Then we summarize key contributions of this dissertation.

3.2 Research Challenges

For complicated next-generation wireless systems exploiting UAV communications, efficient and sophisticated resource management design is needed to enable robust and efficient coexistence between UAVs and cellular networks. However, various research challenges arise as one considers designing efficient communication and computation management strategies. We discuss some of these significant challenges in the following subsections.

3.2.1 Joint UAV Deployment and Resource Allocation

The design of effective UAV-enabled communications networks is quite challenging [9]. First, channel modeling for different UAV communication scenarios is a major research challenge. Specifically, communication channels for air-to-ground communications between UAVs and ground users must be appropriately modeled considering possible line-of-sight (LoS) and non-line-of-sight (NLoS) propagation conditions. Second, efficient deployment of UAVs in three-dimensional (3D) space or effective control of UAVs' trajectories significantly impact communications performance, such as UAV's flight time, energy consumption, and GUs' quality of service (QoS). Finally, the development of resource allocation algorithms that can efficiently manage and assign various network resources, including communication bandwidth and transmit power for users, is of critical importance for UAV-enabled communications networks. Therefore, joint UAV deployment, i.e., UAV placement or trajectory control, and resource allocation are challenging research problems which are discussed in the following.

The first challenge is related to the system design. In particular, spectrum reuse to support communications between multiple UAVs and GUs is needed in practice to enhance the spectral efficiency and network performance; however, efficient co-channel interference management techniques must be developed. Consequently, an efficient design must properly capture these key aspects of the joint UAV deployment and resource allocation system.

The second challenge concerns the quality of service (QoS) of GUs. Many practical application scenarios, such as data collection and information sharing require to guarantee data transmission demand constraints of individual GUs. Therefore, proper system and QoS modes are needed to capture these constraints to achieve reliable designs meeting the required QoS constraints.

The third challenge is related to the underlying optimization arisen in the design. In fact, the joint optimization of UAV deployment and resource allocation often results in optimization problems in the form of mixed integer nonlinear programming (MINLP). In general, it is non-trivial to develop an efficient algorithm to solve these MINLP problems optimally. To tackle this challenge, the alternating optimization approach where the variable sets is decomposed into smaller sets, e.g., user association, sub-channel assignment, UAV trajectory control, can be adopted to effectively solve the complex optimization problem. Specifically, each set of variables is iteratively optimized given the values of other variables in the corresponding sub-problems until convergence. Moreover, in some cases, the sub-problems are still non-convex, so successive convex approximation (SCA) or difference of convex functions (DC) methods could be applied to convexify the sub-problems and the resulting convex problems can be solved effectively by any convex optimization solvers.

3.2.2 IRS-assisted UAV-based Wireless Networks

There are many challenges associated with the aerial deployment of IRSs [101]. The first challenge is related to the deployment to IRSs for flying assets, where the main objectives are low-mass, low-power, and deployment and communication flexibility. Another challenge is the implementation of effective controllers for the surface configuration, given that the channel might change rapidly and significantly while the propagation distance/delay between the flying BS and the surface would be considerable. Moreover, IRS-assisted UAV-based wireless networks raise various research challenges.

Firstly, design of the multi-carrier IRS-assisted UAV-based wireless network considering limited capacity of wireless backhauls has been quite under-explored in the literature.. Therefore, the study on joint optimization of UAV placement, IRS phase shifts, and sub-channel assignments for wireless access and backhaul links still needs further research efforts.

Secondly, the complex IRS phase shifts, i.e., including real and imaginary parts, generally make the design and formulated problem very challenging. Therefore, to handle these difficulties, the IRS phase shifts are usually aligned with the phases of the channel coefficients.

Thirdly, for certain design objectives such as maximization of the sum rate of all GUs, the resulting optimization problems are usually in form of non-convex MINLP, due to integer variables of association between UAVs and GUs or sub-channel assignment between access and backhaul links, and non-convex objective function and constraints. Therefore, it is difficult to solve directly and optimally the underlying optimization problem. To this end, the iterative algorithm and the SCA method can be used to tackle these optimization problems.

3.2.3 Computation Offloading and Resource Allocation in Space-Air-Ground Integrated Networks

Space-air-ground integrated networks (SAGIN) have attracted great attention from academia and industry in recent years. For instance, some organizations have been starting the projects on SAGIN such as the Global Information Grid (GIG) [102, 103], Oneweb [104], and SpaceX [105]. However, engineering the SAGIN requires to solve many challenges such as efficient physical-layer transmission techniques, efficient routing mechanism to enable packet transmission across different network components, efficient integration and management of the heterogeneous resources, and network security [106]. The resource allocation, traffic offloading, and service coordination/network function virtualization in SAGIN have been studied [107–113]. In addition, the QoS, performance and outage performance analysis in SAGIN were investigated in [114–117]. Moreover, the serious security threats in data storage, transmission and sharing of space–air–ground integrated vehicular network were discussed in [118], while SAGIN design empowered by the blockchain has been studied in [119–122], and artificial intelligence (AI), federated learning, and deep learning in SAGIN have been studied in [123–128]. Nevertheless, the computation offloading and resource allocation for SAGIN still require much further research to deal with various technical challenges.

Firstly, none of the previous work has studied computation offloading for the hybrid edge-cloud SAGIN considering user scheduling design over the UAV flight period and multi-hop communications in the satellite network segment. This design problem, however, reflects a practical scenario where a cloud server is deployed far away from the considered terrestrial network area. Therefore, it is essential to investigate the integrated computation offloading, UAV trajectory control, user scheduling, and resource allocation for SAGIN with multi-hop LEO satellite communications.

Secondly, a UAV-assisted MEC system is especially crucial for practical scenarios with limited or no communications infrastructures [129, 130], where UAVs can be aerial users with computation tasks to be executed, or relay nodes to assist ground users in executing/offloading computation tasks, or MEC servers for executing computation tasks. Particularly, when UAVs are equipped with MEC servers then ground users can offload some of their computation-intensive tasks for execution at the UAV-mounted servers. This is especially important for certain practical scenarios where there is no the terrestrial MEC network, such as natural disasters. Generally, a UAV with sufficient energy and computing resources offers some advantages compared to the conventional MEC system where servers are deployed at fixed BSs. Nevertheless, there are several research challenges one must tackle to realize the benefits offered by UAV-assisted MEC systems.

Thirdly, efficient utilization of edge computing resources in SAGIN requires further research to address various open challenges. First, computing delay and bandwidth constraints must be taken into consideration in the design. Second, UAVs and GUs typically have limited energy; therefore, energy-efficient design in SAGIN is an important research issue. Finally, many emerging IoT applications have complex design requirements and functionalities such as demanding data transmission, e.g., video downloads, data processing and analysis, e.g., video analysis and speech recognition, and content caching. Therefore, low-latency computation task processing and efficient resource management are needed to enable edge computing-based applications in SAGIN.

3.3 Literature Review

In the following, we present the survey of the existing literature on the research issues considered in this dissertation. First, we describe the existing work on UAV deployment and resource allocation in section 3.3.1. Then, we survey recent researches on IRS-assisted UAV-based wireless networks in section 3.3.2. Finally, we discuss the research work on computation offloading and resource allocation in SAGIN in section 3.3.3.

3.3.1 Joint UAV Deployment and Resource Allocation

In this section, we provide a brief review of existing research on UAV deployment and resource allocation for UAV-based wireless networks, where UAVs act as flying BSs to provide wireless connectivity to GUs. More information about other use cases and application scenarios can be found in [131–133]. Our following survey discusses existing work where we pay special attention to the data transmission demand constraints and spectrum reuse with interference management, because these design aspects could significantly impact the achievable performance and required QoS of UAV-based wireless networks.

On the one hand, the paper [134] tackled the joint optimization of the power, continuous bandwidth assignment, and 3D UAV's trajectory where its design objective is to minimize the total UAV's energy consumption. This work, however, did not consider data transmission demand constraints and co-channel interference. A block coordinate descent algorithm was used to iteratively optimize the resource allocation and UAV's trajectory control. The authors in [135] addressed the joint design of user scheduling, transmit power, continuous bandwidth assignment, and UAV's trajectory control in the 3D space to maximize the system energy efficiency. However, this work did not consider co-channel interference and an iterative algorithm using the Dinkelbach and block coordinate descent techniques was proposed to solve the underlying problem. The authors in [136] also studied the UAV's trajectory and continuous bandwidth assignment without considering co-channel interference where its design goal was to maximize the minimum average rate of GUs using an alternating optimization technique.

On the other hand, the work [39, 137–140] mainly studied the sub-channel assignment and UAV trajectory control. An exception is [137] where the authors discussed the sub-channel assignment considering co-channel interference and UAV velocity control with a known trajectory to maximize the uplink sum rate through an iterative algorithm. The authors of [138] studied the network setting with only two UAVs, i.e., transmitter and jammer, and they considered maximizing the system energy efficiency by jointly optimizing the transmit power, sub-channel assignment, and UAV trajectory control using an alternating optimization algorithm and relaxation of the binary sub-channel assignment decision variables. Moreover, the designs in [39, 139, 140] aim to maximize the minimum average rate of GUs. The co-channel interference was considered only in [140]. The sub-channel assignment was investigated in [39] for a network where the UAV employs orthogonal frequency-division multiple access (OFDMA). Bandwidth, power allocation, and UAV's trajectory control were jointly optimized using the block coordinate descent method. A backhaul-aware design to maximize the minimum average rate for GUs was tackled in [139, 140] using an alternating optimization approach to solve the joint problem of sub-channel assignment and UAV's trajectory control.

Even though there have been some existing studies on the joint sub-channel assignment and UAV's trajectory design considering spectrum reuse and co-channel interference management, this research direction remains under-explored in the UAV communication literature.

3.3.2 IRS-assisted UAV-based Wireless Networks

Intelligent reflecting surfaces (IRS) assisted UAV-based wireless network can play an important role in the 5G and beyond wireless system [26, 141]. Specifically, the IRSs can be installed on UAVs or deployed on the facades of the buildings. On the one hand, such terrestrial deployment may be hindered by appropriate site selection, service access which can be limited to only half of the space, and scattering in undesired direction in urban areas [142]. To address these shortcomings of the terrestrial IRS, aerial deployment of IRSs on UAVs has been recently explored [143–147], where the designs of the UAV placements and beamforming matrics were investigated to maximize the energy efficiency or maximize data rate of the system. However, the enabling techniques on UAVs and energy limitations of UAVs make this design very challenging.

On the other hand, most of the existing work on IRS deployment have focused on the terrestrial deployment of the IRSs such as on the facades of the buildings. In particular, the optimization of UAV placement, number of IRS elements, and phase shifts of IRS was studied in [148–150]. The single UAV and IRS were considered to improve the strength of received signals at GUs in [148] and maximize secrecy rate in [149]. The design of multi-UAV placement and non-orthogonal multiple access (NOMA) to maximize sum rate of the system was considered in [150]. Most of these designs result in the non-convex optimization problems due to the non-convex constraints of IRS phase shifts and UAV placements. Hence, an alternating optimization approach can be employed to solve the underlying problems and obtain a feasible solution.

Furthermore, the optimization of UAV trajectory, transmit power control, and IRS phase shifts for a single IRS deployed on a building wall was investigated in [27, 43, 151–156]. The design in [151] aims to minimize the total power consumption, while maximization of the secrecy rate and the sum rate was considered in [43, 152] and [27, 153], respectively. In addition, the sub-channel assignment design was addressed in [153]. Moreover, the design in [154] maximizes the minimum achievable rate of all GUs, while the authors in [155] considered maximizing the received signal power at one GU. The authors in [156] investigated two scenarios of UAV trajectory control and IRS phase shifts to minimize the weighted and maximum bit error rate (BER) among all IRSs. In these papers, different iterative algorithms have been developed and the IRS phase shifts are aligned with the phases of the channel coefficients.

However, the design of multi-carrier IRS-assisted UAV-based wireless network taking into account the constrained capacity of wireless backhauls has not been addressed in the aforementioned existing work.

3.3.3 Computation Offloading and Resource Allocation in SAGIN

Computation offloading and resource allocation in the SAGIN have attracted great attention [126, 157–159] where computation tasks can be offloaded from GUs to UAVs and satellites to save GUs' energy and/or improve computation latency. In addition, the authors in [160] considered a simple SAGIN setting with a single LEO satellite providing cloud computing capability and a UAV-mounted MEC server providing computing resources near the GUs where the design aims to maximize the long-term time-averaged total system computation rate by optimizing the computing resource, power allocation, and UAV trajectory control. Optimization of user association between the GUs and multiple UAV-mounted edge servers was studied in [161–164]. However, these papers only optimized the UAV placement or assumed that the UAVs' trajectories are pre-determined.

Optimization of UAV trajectory control and user association for SAGIN has been performed considering different design objectives as functions of throughput/capacity and energy consumption [48, 165, 166]. Specifically, the SAGIN setting with one LEO satellite and multiple UAVs was considered in [165] where the design aims to maximize the system capacity by jointly optimizing user association, power control, and UAV trajectory. Meanwhile, the SAGIN with one satellite and multiple UAVs was studied in [48] with the objective of minimizing the weighted energy consumption via joint optimization of device association, resource partitioning, bit allocation, and UAV trajectory control. Moreover, in [166], the authors considered the downlink communications for SAGIN with a single satellite, multiple UAVs, and base station with user terminals where the design aims to maximize the average throughput among GUs by jointly optimizing user association, power control, and UAV trajectory. However, the above existing studies have not considered the maximum delay constraints in the computation offloading design and they only tackled the binary task assignment or task partitioning between the local device and the MEC or cloud servers.

The general computation offloading designs with the parallel task execution at local devices, MEC, and/or cloud servers were investigated in [44, 45, 167]. Specifically, minimization of the maximum delay experienced by different GUs by jointly optimizing UAV-device association, task assignment, power control, bandwidth allocation, computation resource, and UAV placement was studied in [44]. The authors in [45] considered a multi-user MEC system and optimized user association and task partitioning to achieve minimum average latency for all GUs where independent and dependent sub-tasks are explored. Moreover, a SAGIN setting with a single satellite, a single UAV, and multiple small cells was studied in [167] whose design aims to maximize the sum rate of the small cells by jointly optimizing the user association, sub-channel, and power allocation considering the maximum delay constraint.

Several key design aspects were not addressed satisfactorily in the aforementioned existing work. First, partial offloading for efficient computation load balancing among GUs, edge and cloud servers considering radio resource allocation, UAV trajectory control, and multi-hop satellite communications has not been studied in the literature. Second, the maximum delay constraints imposed by underlying computation tasks may not be achieved due to limited radio and computation resources. To this end, admission control is a critical issue which has not been addressed in the SAGIN context.

3.4 Research Objectives and Contributions

The general objective of my Ph.D. research is to develop novel network architectures and efficient resource allocation algorithms for UAV-based wireless networks which contribute to enable efficient integration of terrestrial and non-terrestrial networks. Specifically, our main contributions, which are highlighted in Fig. 3.4, can be described as follows.

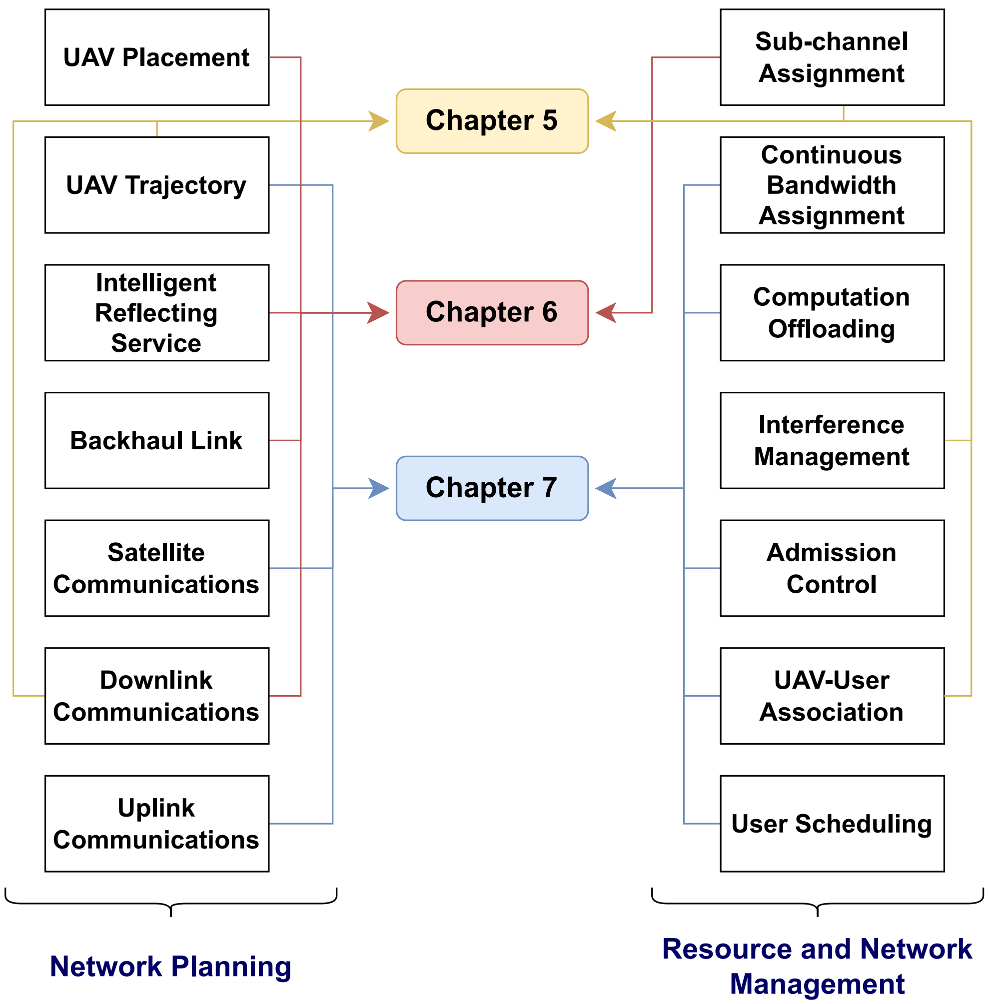

Figure 3.4 – Research contributions.

1. *Integrated UAV trajectory control and resource allocation for UAV-based wireless networks with co-channel interference management*:

   We study the trajectory control, sub-channel assignment, and user association design for unmanned aerial vehicles (UAVs)-based wireless networks. Our design optimizes the max-min average rate subject to data demand constraints of GUs where spectrum reuse and co-channel interference management are considered. The considered optimization leads to a mixed integer non-linear optimization problem which is solved by using the alternating optimization approach where we iteratively optimize the user association, sub-channel assignment, and UAV trajectory control until convergence. For the sub-channel assignment sub-problem, we propose an iterative sub-channel assignment (ISA) algorithm to obtain an efficient solution. Moreover, the successive convex approximation (SCA) is used to convexify and solve the non-convex UAV trajectory control sub-problem. Via extensive numerical studies, we illustrate the effectiveness of our proposed design considering different UAV flight periods and number of sub-channels and GUs compared with a simple heuristic.

2. *UAV placement and resource allocation for intelligent reflecting surface assisted UAV-based wireless networks*:

   We design a UAV-based wireless network with wireless access and backhaul links leveraging an intelligent reflecting surface (IRS). This design aims to maximize the sum rate achieved by GUs through optimizing the UAV placement, IRS phase shifts, and sub-channel assignments considering the wireless backhaul capacity constraint. To tackle the underlying mixed integer non-linear optimization problem (MINLP), we first derive the closed-form IRS phase shift solution; we then optimize the sub-channel assignment and UAV placement by using the alternating optimization method. Specifically, we propose an iterative sub-channel assignment strategy to efficiently utilize the bandwidth and balance bandwidth allocation for wireless access and backhaul links while maintaining the backhaul capacity constraint. Moreover, we employ the successive convex approximation (SCA) method to solve the UAV placement optimization sub-problem. We show the effectiveness of our proposed design via extensive numerical studies.

3. *Integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control in SAGIN*:

   We study the computation offloading problem in SAGIN, where joint optimization of partial computation offloading, UAV trajectory control, user scheduling, computation, resource allocation, and admission control is performed. Specifically, the considered SAGIN employs multiple UAV-mounted edge servers with controllable UAV trajectories and a cloud sever which can be reached by GUs via multi-hop LEO satellite communications. This design aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of underlying computation tasks. To tackle the underlying non-convex mixed integer non-linear optimization problem, we use the alternating optimization approach where we iteratively solve four sub-problems, namely user scheduling, partial offloading control and bit allocation over time slots, computation resource and bandwidth allocation, and multi-UAV trajectory control until convergence. Moreover, feasibility verification and admission control strategies are proposed to handle overloaded network scenarios. Furthermore, the successive convex approximation (SCA) method is employed to convexify and solve the non-convex computation resource and bandwidth allocation and UAV trajectory control sub-problems. Via extensive numerical studies, we illustrate the effectiveness of our proposed design with respect to several baselines.

3.5 Dissertation Outline

The remaining of this dissertation is organized as follows. Chapter 4 reviews some fundamental background including mathematical optimization, UAV communications, IRS-assisted UAV-based wireless network, and SAGIN. Chapter 5 covers our study about integrated UAV trajectory control and resource allocation for UAV-based wireless networks with co-channel interference management. Then, we study the UAV placement and resource allocation for intelligent reflecting surface assisted UAV-based wireless networks in Chapter 6. Finally, we present our study integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control in SAGIN in Chapter 7. The main contributions of the dissertation and some potential direction for future research are discussed in Chapter 8.

Chapter 4

Background

In this chapter, we present some fundamentals of mathematical optimization, UAV communications, IRS-assisted UAV communications, computation task model, and satellite communications. Particularly, basic concepts of mathematical optimization and some popular techniques to solve optimization problems are introduced in Section 4.1 while we present working principles of UAV communications and IRS-assisted UAV-based wireless networks in Section 4.2 and Section 4.3, respectively. Section 4.4 presents the computation task models and Section 4.5 describes the satellite communications.

4.1 Mathematical Optimization

4.1.1 Fundamental Concepts

A mathematical optimization problem can be written in the following form [168]

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_0(\mathbf{x}), \\ \text{subject to} \quad & f_i(\mathbf{x}) \leq 0, \quad i = 1, 2, ..., m, \end{aligned} \tag{4.1}$$

where the vector $\mathbf{x} \in \mathbb{R}^n$ is the *optimization variable*, and $f_0(\mathbf{x})$ is the *objective function*. The inequalities $f_i(\mathbf{x}) \leq 0,\ i = 1, 2, ..., m$ are the *constraints* of the problem. Let $\mathcal{D}$ be the intersection of the domains of $f_i(\mathbf{x}),\ i = 0, 1, ..., m$, the *feasible set* of the problem is the set of all $\mathbf{x} \in \mathcal{D}$ that satisfies all these m constraints. A vector $\mathbf{x}^*$ is called *optimal*, or *optimal solution*, of the problem if $f_0(\mathbf{x}^*)$ achieves the smallest value among all values of $f_0(\mathbf{x})$ where $\mathbf{x}$ belongs to the feasible set, and the value of $f_0(\mathbf{x}^*)$ is called *optimal value*. If the feasible set is empty, the problem is *infeasible*. Conventionally, the optimal value of the problem is $+\infty$ if the problem is infeasible [168]. Hereafter, the term 'subject to' is written as 's.t.'.

4.1.2 Convex Optimization

Among many classes of optimization problems, convex optimization problems are of particular interest. First, it is a fundamental property of convex optimization problems that any locally optimal point is also globally optimal. Hence, compared to a generic optimization problem, it is generally easier to solve a convex optimization problem as one only needs to find a local optimal solution. Second, many sub-classes of convex optimization problems are well-studied, and the technologies to solve most of problems in these sub-classes are mature and can be deployed in many applications. In the following sections, some fundamentals of convex optimization are briefly introduced.

Convex set: A set S is convex if for any vectors $\mathbf{x}, \mathbf{y} \in S$, the following holds for any value of θ where $\theta \in [0, 1]$

$$\theta\mathbf{x} + (1-\theta)\mathbf{y} \in S. \tag{4.2}$$

Convex function: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain (denoted as $\mathcal{D}$) is a convex set and the following inequality holds for any $\mathbf{x}, \mathbf{y} \in \mathcal{D}$ and any $\theta \in [0, 1]$

$$f\left(\theta\mathbf{x} + (1-\theta)\mathbf{y}\right) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}). \tag{4.3}$$

Convex optimization problem: An optimization problem is convex if it can be written in the following form

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) \leq 0, \quad i = 1, ..., m, \\ & h_j(\mathbf{x}) = 0, \quad j = 1, ..., p, \end{aligned} \tag{4.4}$$

where the functions $f_i,\ i = 0, 1, ..., m$, are convex and the functions $h_j,\ j = 1, ..., p$ are linear. Convex optimization problems can be called ‘convex problems’ for short.

4.1.3 Methods to Solve Optimization Problems

It is known that there is no efficient methods that can find a global solution of a generic nonconvex optimization problem in polynomial time [168]. However, various methods have been developed to solve some classes of convex optimization problems with desired accuracy and in polynomial time with respect to the problem dimensions [169, 170]. Particularly, the interior-point (or barrier) methods are currently considered the most powerful algorithms for large-scale problems. The interior-point methods are already used in many optimization solvers. For brevity, we do not intend to go into details about these methods. Interested readers are encouraged to read [168, 171] where many rigorous definitions of mathematical optimization and solving techniques are presented in details. Furthermore, since available solvers already do a decent job in solving popularly encountered convex problems, we use them to solve convex problems that arise in our research, rather than developing specific numerical methods. In particular, we use the CVX [38] on MATLAB to solve convex problems in our research where the underlying solver is Mosek academic version [172]. That being said, several formulated problems arising in our studies are nonconvex, so techniques other than interior-point methods have to be devised to tackle them. In this dissertation, two main techniques that we use to find sub-optimal solutions of the formulated nonconvex problems are Successive Convex Approximation (SCA) and Difference of Convex Functions (DC) Programming where a nonconvex problem is solved sub-optimally by iteratively solving a series of convex optimization problems. Details of these techniques are presented in the following.

4.1.3.1 Successive Convex Approximation

There are approaches to tackle the nonconvexity of optimization problems. We will describe one popular approach in this regard which is the Successive Convex Approximation (SCA) method. In the SCA method, nonconvex functions are approximated by convex functions and the resulting approximated convex problem is solved iteratively until convergence. In this iterative process, the solution obtained in each iteration is used in the new approximations of the objective and constraint functions in the next iteration.

Consider the following nonconvex optimization problem

$$\begin{aligned} \mathcal{P}: \quad & \underset{\boldsymbol{x}}{\text{minimize}} \quad f_0(\boldsymbol{x}), \\ & \text{s.t.} \quad f_i(\boldsymbol{x}) \leq 0, \quad i = 1, 2, ..., m. \end{aligned} \tag{4.5}$$

Assume that the local point $\boldsymbol{x}^r$ is given, in iteration $r+1$, the SCA method approximates functions $f_i(\boldsymbol{x})$ by $\tilde{f}_i(\boldsymbol{x}|\boldsymbol{x}^r)$ and solves the following approximated optimization problem

$$\begin{aligned} \mathcal{P}^{r+1}: \quad & \underset{\boldsymbol{x}}{\text{minimize}} \quad \tilde{f}_0(\boldsymbol{x}|\boldsymbol{x}^r), \\ & \text{s.t.} \quad \tilde{f}_i(\boldsymbol{x}|\boldsymbol{x}^r) \leq 0, \quad i = 1, 2, ..., m, \end{aligned} \tag{4.6}$$

where the following conditions hold [173]

$$\text{Upper-bound:} \quad \tilde{f}_i(\boldsymbol{x}|\boldsymbol{x}^r) \geq f_i(\boldsymbol{x}), \tag{4.7a}$$

$$\text{Function value consistency:} \quad \tilde{f}_i(\boldsymbol{x}^r|\boldsymbol{x}^r) = f_i(\boldsymbol{x}^r), \tag{4.7b}$$

$$\text{Gradient consistency:} \quad \nabla\tilde{f}_i(\boldsymbol{x}^r|\boldsymbol{x}^r) = \nabla f_i(\boldsymbol{x}^r), \tag{4.7c}$$

$$\text{Convexity:} \quad \tilde{f}_i(\boldsymbol{x}|\boldsymbol{x}^r) \text{ is convex with respect to } \boldsymbol{x}. \tag{4.7d}$$

These conditions guarantee that in each iteration, an original function is approximated by a upper-bound whose first order derivative is equal to that of the original function. The typical SCA based algorithm to solve problem $\mathcal{P}$ is described in Algorithm 4.1.

Algorithm 4.1. Typical SCA algorithm

```
Initiate x by a feasible x^0 and set r = 0
while 1 do
    Solve problem P^{r+1}, update x^r by the obtained solution.
    if Convergence condition is met then
        Break the loop.
    else
        Let r = r + 1.
    end if
end while
End of algorithm.
```

4.1.3.2 Difference of Convex Functions (DC) Programming

DC programming can be used to solve a particular family of non-convex problems [174]. In the wireless domain, the transmission rate in many practical scenarios has the DC form which is a feature of DC programming problem (DCP). There are several popular techniques such as branch-and-bound and cutting planes algorithms to solve optimization problems, but in general, they are inefficient. One desirable aspect of DCP is generally possible to build the approximated function satisfying three conditions in (4.7) and thus one can obtain the local optimal solutions [175]. Further details about the DC approach are briefly presented as follows.

DC function: A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a DC function if there exists *convex* functions $g, h : \mathbb{R}^n \mapsto \mathbb{R}$ such that f can be expressed as the difference between g and h as

$$f(\boldsymbol{x}) = g(\boldsymbol{x}) - h(\boldsymbol{x}). \tag{4.8}$$

DCP: A problem is a DCP if it has the following form

$$\min_{\boldsymbol{x}} \quad f_0(\boldsymbol{x}) \tag{4.9}$$

$$\text{s.t.} \quad f_i(\boldsymbol{x}) \leq 0, \; i = 1, \ldots, m, \tag{4.9a}$$

where the function $f_i : \mathbb{R}^n \mapsto \mathbb{R}$ is a differentiable DC function for $i = 0, 1, \ldots, m$.

Approximated DCP: Using the first Taylor approximation, the convex term $h_i(\boldsymbol{x})$ in $f_i(\boldsymbol{x})$ can be approximated by its lower-bound at iteration r in the approximation process as follows:

$$h_i(\boldsymbol{x}) \geq \tilde{h}_i(\boldsymbol{x}, \boldsymbol{x}^r) = h_i(\boldsymbol{x}^r) + \nabla h_i^T(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{x}^r). \tag{4.10}$$

Then, we have $f_i(\boldsymbol{x}) = g_i(\boldsymbol{x}) - h_i(\boldsymbol{x}) \leq g_i(\boldsymbol{x}) - \tilde{h}_i(\boldsymbol{x}, \boldsymbol{x}^r) = \tilde{f}_i(\boldsymbol{x}, \boldsymbol{x}^r)$. Using these upper-bound functions of DC functions, the approximated DCP at iteration $r + 1$ which is a standard convex problem, can be expressed as

$$\begin{aligned} \min_{\boldsymbol{x},\eta} \quad & \eta && (4.11) \\ \text{s.t.} \quad & \tilde{f}_0(\boldsymbol{x}, \boldsymbol{x}^r) - \eta \leq 0, && (4.11a) \\ & \tilde{f}_i(\boldsymbol{x}, \boldsymbol{x}^r) \leq 0, \ i = 1, \ldots, m. && (4.11b) \end{aligned}$$

4.2 UAV Communications and Networks

In this section, we introduce some fundamental concepts of UAV communications and UAV-based wireless networks. In particular, we present widely adopted channel models in section 4.2.1 and typical optimization formulations in section 4.2.2.

4.2.1 Channel Models in UAV-based Wireless Networks

Most studies on channel modeling for air-to-ground and ground-to-air communications in UAV-based wireless networks are fairly recent compared to those on channel modeling for the conventional cellular networks. One popular air-to-ground channel model considers a binomial random event in which the LoS and non-line-of-sight (NLoS) communications between the UAV and the ground users occur with certain probabilities [176]. Specifically, these LoS/NLoS probabilities depend on the elevation angle, types of communications environment, e.g., urban, sub-urban, and rural, and the relative locations of the UAV and users. We denote the 2-D coordinate of the considered ground user as $\mathbf{u}$ and the 2-D coordinate of the UAV as $\mathbf{q}$. For convenience, we assume that the ground user has altitude of 0 meters and the UAV has altitude of h meters. Then, the LoS and NLoS probabilities can be calculated as follows:

$$\begin{aligned} P_{LoS} &= \alpha \left(\frac{180}{\pi} \theta - 15 \right)^{\gamma}, \\ P_{NLoS} &= 1 - P_{LoS}, \end{aligned} \tag{4.12}$$

where α and γ are the constants which depend on the communications environment, and θ is the elevation angle that can be calculated as

$$\theta = \tan^{-1}\left(\frac{h}{\|\mathbf{u}-\mathbf{q}\|}\right). \tag{4.13}$$

Moreover, the shadow fading components for the LoS and NLoS links which are denoted as ξ_{LoS} and ξ_{NLoS}, respectively, are assumed to follow a log-normal distribution, which means their logarithms are normal distribution with mean 0 and variances defined as follows:

$$\begin{aligned} \sigma_{LoS} &= k_{LoS}\exp\left(-g_{LoS}\theta\right), \\ \sigma_{NLoS} &= k_{NLoS}\exp\left(-g_{NLoS}\theta\right), \end{aligned} \tag{4.14}$$

where $k_{LoS}, g_{LoS}, k_{NLoS}$ and g_{NLoS} are positive constants depending on the communications environment. Then, the channel power gains (τ_{LoS}, τ_{NLoS}) can be computed for the LoS and NLoS communications scenarios as follows [176]:

$$\begin{aligned} \tau_{LoS} &= \frac{\zeta}{\xi_{LoS}}\left(\frac{4\pi f_c(h^2+\|\mathbf{u}-\mathbf{q}\|^2)^{1/2}}{c}\right)^{-\kappa}, \\ \tau_{NLoS} &= \frac{\zeta}{\xi_{NLoS}}\left(\frac{4\pi f_c(h^2+\|\mathbf{u}-\mathbf{q}\|^2)^{1/2}}{c}\right)^{-\kappa}, \end{aligned} \tag{4.15}$$

where ζ is the constant that accounts for the antenna gain, κ is the free-space path loss exponent, f_c and c are the carrier frequency and the speed of light, respectively.

Even though the above probabilistic channel model accounts for several factors that affect the channel power gain, it can be difficult to estimate the involved constants for different types of communications environment. Moreover, in the multi-UAV based wireless networks, it is natural to associate users with their closest UAVs [70], which would increase the elevation angle θ. Besides, the high elevation angle significantly reduces the effect of shadowing (from (4.14)) and increases the LoS probability (from (4.12)). Therefore, in multi-UAV deployments with high elevation angles, the air-to-ground channel can be considered effectively LoS. As a result, many research studies in UAV communications have adopted a simplified channel model in which the air-to-ground channel power gain can be expressed as follows:

$$\tau = \mu(h^2+\|\mathbf{u}-\mathbf{q}\|^2)^{-\kappa/2}, \tag{4.16}$$

where μ is the channel power gain at a reference distance. In fact, μ accounts for the antenna gain ζ and other constants in (4.15).

Despite its simplicity, the channel model in (4.16) and its results can serve as benchmarks and help gain insights about efficient network designs and achievable performance.

4.2.2 Design Optimization in UAV-based Wireless Networks

In this section, we discuss a generic design optimization formulation in UAV-based wireless networks where there are M UAVs communicating with K users in the downlink direction. The service period is divided into N time slots ($n = 1, 2, ..., N$), each having equal length of Δt seconds. The slot length Δt should be chosen appropriately such that network conditions stay approximately the same during each time slot. We consider a communication system in which users and UAVs communicate via orthogonal resource blocks.[1] Let the channel power gain, bandwidth, and transmit power allocated for user k associated with UAV m in time slot n be $g_{k,m}[n]$, $b_{k,m}[n]$, and $p_{k,m}[n]$, respectively. The achievable data rate of user k associated with UAV m at time slot n, denoted as $r_{k,m}[n]$, can be expressed as follows:

$$r_{k,m}[n] = b_{k,m}[n] \log\left(1 + \frac{g_{k,m}[n] p_{k,m}[n]}{\sigma^2 b_{k,m}[n]}\right), \tag{4.17}$$

where σ^2 denotes the white noise power density (W/Hz).

The total transmit power of each UAV, and bandwidth of the system, denoted as B, are limited. Hence, we have the following constraints:

$$\sum_{k=1}^{K} p_{k,m}[n] \leq P_{\mathsf{max}}, \quad \forall m, n, \tag{4.18a}$$

$$\sum_{k=1}^{K} \sum_{n=1}^{N} b_{k,m}[n] \leq B, \quad \forall n, \tag{4.18b}$$

where P_{max} is the maximum total power of each UAV.

[1]Note that co-channel interference among user-UAV communications can occur and it has been considered in some existing work where non-orthogonal resource blocks are allocated for nearby communication links/users. However, for the introductory purpose of this section, we choose to present a design optimization for an interference-free scenario.

Let the 3-D coordinate of UAV m in time slot n be $\mathbf{q}_m[n] = (x_n[n], y_n[n], z_n[n])$, the following constraints are usually considered for the UAV's trajectory control:

$$\|\mathbf{q}_m[n] - \mathbf{q}_m[n-1]\| \leq \Delta t V_{max}, \quad \forall m, n = 1, ..., N-1, \tag{4.19a}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\| \geq d_{\mathsf{min}}, \quad \forall j \neq m, n, \tag{4.19b}$$

$$\mathbf{q}_m[1] = \mathbf{q}_{start}, \quad \forall m, \tag{4.19c}$$

$$\mathbf{q}_m[N] = \mathbf{q}_{end}, \quad \forall m, \tag{4.19d}$$

where V_{max} is the maximum speed of a UAV, d_{min} is the safety distance between any two UAVs to avoid collision, and $\mathbf{q}_{start}$ and $\mathbf{q}_{end}$ capture the starting and final positions on the trajectory of a UAV, respectively.

A generic resource allocation optimization problem in UAV-based wireless networks can be stated as follows:

$$\begin{aligned}(\mathbf{P}): \min_{\{\mathbf{B}[n],\mathbf{P}[n],\mathbf{Q}[n]\}} \quad & f(\mathbf{B}[n], \mathbf{P}[n], \mathbf{Q}[n]) \\ \text{s.t.} \quad & \text{constraints related to } r_{k,m}[n] \text{ in (4.17)}, \\ & \text{constraints (4.18a), (4.18b), (4.19a), (4.19b), (4.19c), (4.19d)},\end{aligned} \tag{4.20}$$

where $f(\mathbf{B}[n], \mathbf{P}[n], \mathbf{Q}[n])$ is the objective function that depends on the optimization variables, $\mathbf{B}[n], \mathbf{P}[n], \mathbf{Q}[n]$ are the matrices of resource allocation variables, power allocation varibales, and the matrix of UAVs' coordinates in time slot n, respectively, and $\{\mathbf{B}[n], \mathbf{P}[n], \mathbf{Q}[n]\}$ denotes the set of all optimization variables for all values of n. In the single-UAV setting, we have $M = 1$ and constraints (4.19b) are omitted. Furthermore, there may be more variables and constraints in optimization problems for more complicated designs and network settings.

Optimization problems formulated in UAV-based wireless networks similar to (4.20) are usually complicated and non-convex which involves both resource allocation and spatial variables to be optimized in each time slot. Solving these types of optimization problems usually requires sophisticated algorithms which usually exploit certain structure of the underlying problem. Beside analysis, modeling, and formulations, developing algorithms to solve different design optimization problems is the core part of many existing studies in the literature.

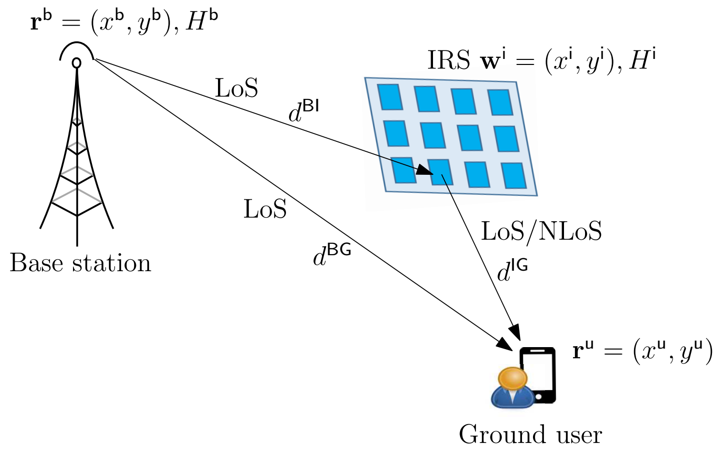

Figure 4.1 – General model of IRS-aided communications.

4.3 IRS-assisted UAV Communications

In this section, we first describe the IRS-aided communications in section 4.3.1 and then the model of IRS-assisted UAV communications will be presented in section 4.3.2.

4.3.1 IRS-aided Communications

We consider a general model of the IRS-aided communications in Fig. 4.1. The channel between the base station (BS) and IRS, BS and GU are dominated by LoS propagation due to higher altitude, while the channel between the IRS and GU can be dominated by LoS or NLoS propagation due to blockages. Hence, the equivalent channel h from the BS to GU can be expressed as

$$h = (\mathbf{h}^{\mathsf{IG}})^H \mathbf{\Phi} \mathbf{h}^{\mathsf{BI}} + h^{\mathsf{BG}}, \tag{4.21}$$

where $h^{\mathsf{BG}}, \mathbf{h}^{\mathsf{BI}}$, and $\mathbf{h}_k^{\mathsf{IG}}$ denote the channel coefficients of the links between the BS and GU, BS and IRS, IRS and GU k, respectively. In addition, $\mathbf{\Phi}$ denotes the phase shift matrix of the IRS [43,153].

We have $h^{\mathsf{BG}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{BG}})^2}}$, where β_0 denotes the channel gain at the reference distance of 1 meter and d^{BG} represents the distance from BS to GU. Meanwhile, the expression of other functions depends on what type of array elements is installed for IRS, i.e., uniform linear array (ULA) or uniform planar array (UPA). We will discuss these two cases in the following.

4.3.1.1 Uniform Linear Array

For the IRS made up of a square array of I elements, individual reflecting signals from the UAV to GUs are adaptively assisted by an IRS controller that tunes the phase shifters of the elements accordingly where the phase shift matrix can be written as $\mathbf{\Phi} = \operatorname{diag}\left\{e^{j\phi_1}, \ldots, e^{j\phi_I}\right\} \in \mathbb{C}^{I\times 1}$, where $\phi_i \in [0, 2\pi)\,, \forall i = 1, \ldots, I$. Besides, the distance between any two adjacent elements of the IRS is denoted by d. Then we have

$$\mathbf{h}^{\mathsf{BI}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{BI}})^2}} \left[1, \ldots, e^{-j\frac{2\pi}{\lambda}(\sqrt{I}-1)d\theta^{\mathsf{BI}}}\right]^T, \tag{4.22}$$

$$\mathbf{h}^{\mathsf{IG}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{IG}})^\kappa}} \left[1, \ldots, e^{-j\frac{2\pi}{\lambda}(\sqrt{I}-1)d\theta^{\mathsf{IG}}}\right]^T \alpha^{\mathsf{IG}}, \tag{4.23}$$

where κ is the path loss exponent, λ is the wavelength of the carrier wave, α^{IG} is the random scattering components modeled by a circularly symmetric complex Gaussian random variable with zero mean and unit variance, and d^{BI} and d^{IG} are the distance from BS to IRS, IRS to GU, respectively. In addition, θ^{BI} and θ^{IG} denote the cosine of the angle of arrival of the signal from BS to IRS and the cosine of the angle of departure of the signal from IRS to GU, respectively, which can be calculated as $\theta^{\mathsf{BI}} = \frac{x^{\mathsf{i}}-x^{\mathsf{b}}}{d^{\mathsf{BI}}}$ and $\theta^{\mathsf{IG}} = \frac{x^{\mathsf{u}}-x^{\mathsf{i}}}{d^{\mathsf{IG}}}$.

4.3.1.2 Uniform Planar Array

For the IRS made up of $I_r \times I_c$ passive reflection elements units installed as a uniform planar array (UPA) with I_c and I_r elements on each column and each row, respectively, the phase shift matrix can be expressed as $\mathbf{\Phi} = \operatorname{diag}\left\{e^{j\phi_{1,1}}, \ldots, e^{j\phi_{i_r,i_c}}, \ldots, e^{j\phi_{I_r,I_c}}\right\} \in \mathbb{C}^{I_r\times I_c}$, where $\phi_{i_r,i_c} \in [0, 2\pi)\,, \forall i_r = 1, \ldots, I_r$, and $i_c = 1, \ldots, I_c$. Besides, the distance between any two adjacent elements of the IRS is denoted by d. Then we have

$$\begin{aligned}\mathbf{h}^{\mathsf{BI}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{BI}})^2}} \times &\left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{BI}}\cos\xi^{\mathsf{BI}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta^{\mathsf{BI}}\cos\xi^{\mathsf{BI}}}\right]^H \\ &\otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{BI}}\sin\xi^{\mathsf{BI}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta^{\mathsf{BI}}\sin\xi^{\mathsf{BI}}}\right]^H,\end{aligned} \tag{4.24}$$

$$\begin{aligned}\mathbf{h}^{\mathsf{IG}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{IG}})^\kappa}} \times &\left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{IG}}\cos\xi^{\mathsf{IG}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta^{\mathsf{IG}}\cos\xi^{\mathsf{IG}}}\right]^H \\ &\otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{IG}}\sin\xi^{\mathsf{IG}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta^{\mathsf{IG}}\sin\xi^{\mathsf{IG}}}\right]^H \times \alpha^{\mathsf{IG}},\end{aligned} \tag{4.25}$$

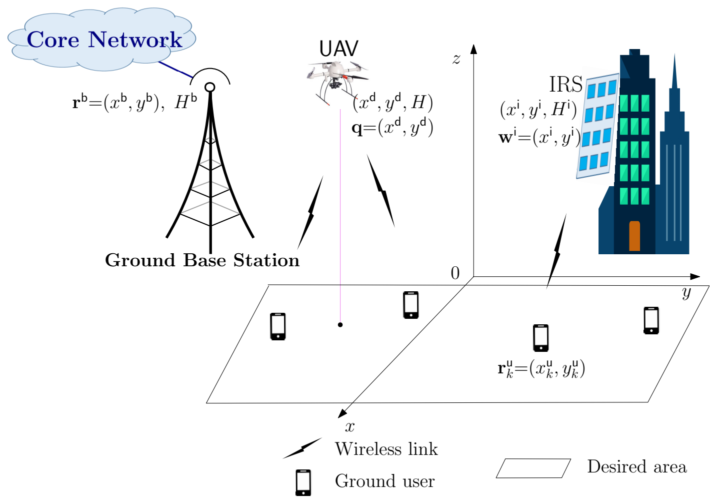

Figure 4.2 – System model IRS-assisted UAV-based wireless networks.

where κ is the path loss exponent, λ is the wavelength of the carrier wave, α^{IG} is the random scattering components modeled by a circularly symmetric complex Gaussian random variable with zero mean and unit variance, and d^{BI} and d^{IG} are the distance from BS to IRS, IRS to GU, respectively. In addition, $(\theta^{\mathsf{UI}}, \xi^{\mathsf{UI}})$ and $(\theta^{\mathsf{IG}}, \xi^{\mathsf{IG}})$ represent the vertical and horizontal angles of departure from the BS to the IRS and from the IRS to GU k, respectively, which can be calculated from $\sin\theta^{\mathsf{BI}} = \frac{|H^{\mathsf{b}}-H^{\mathsf{i}}|}{d^{\mathsf{BI}}}$, $\sin\xi^{\mathsf{BI}} = \frac{|x^{\mathsf{i}}-x^{\mathsf{b}}|}{\sqrt{\|\mathbf{r}^{\mathsf{b}}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\cos\xi^{\mathsf{BI}} = \frac{|y^{\mathsf{i}}-y^{\mathsf{b}}|}{\sqrt{\|\mathbf{r}^{\mathsf{b}}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\sin\theta^{\mathsf{IG}} = \frac{H^{\mathsf{i}}}{d^{\mathsf{IG}}}$, $\sin\xi^{\mathsf{IG}} = \frac{|x^{\mathsf{i}}-x^{\mathsf{u}}|}{\sqrt{\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}^{\mathsf{u}}\|^2}}$, and $\cos\xi^{\mathsf{IG}} = \frac{|y^{\mathsf{i}}-y^{\mathsf{u}}_k|}{\sqrt{\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}^{\mathsf{u}}\|^2}}$.

In the following section, we will discuss more detail the channel model with the received signals that could be improved by employing the IRS.

4.3.2 IRS-assisted UAV Communications Model

In this section, we describe in detail the IRS-assisted UAV communications model. Specially, we will discuss how the data rate could be improved. In the literature survey, the IRS model with ULA was considered in [27, 43], while the IRS with UPA was studied in [151, 153]. To make the design more general and reliable, we describe the system with UPA for the IRS where these results will be studied in Chapter 6.

The general system model is illustrated in Fig. 4.2, in which we assume that all BS-UAV, UAV-IRS, and UAV-GU communication links are dominated by the LoS propagation while communications channels between the IRS and GUs experience Rayleigh channel fading due to blockages. The IRS is made up of $I_r \times I_c$ passive reflection elements units installed as a uniform planar array (UPA) with I_c and I_r elements on each column and each row, respectively. Besides, the distance between any two adjacent elements of the IRS is denoted by d. The phase shift matrix of the IRS is denoted by $\mathbf{\Phi} = \mathrm{diag}\left\{e^{j\phi_{1,1}}, \ldots, e^{j\phi_{i_r,i_c}}, \ldots, e^{j\phi_{I_r,I_c}}\right\} \in \mathbb{C}^{I_r \times I_c}$, where $\phi_{i_r,i_c} \in [0, 2\pi), \forall i_r = 1, \ldots, I_r$, and $i_c = 1, \ldots, I_c$.

The received signal at GU k due to the communications from the UAV is given by

$$y_k = \sqrt{p}\left((\mathbf{h}_k^{\mathsf{IG}})^H \mathbf{\Phi} \mathbf{h}^{\mathsf{UI}} + h_k^{\mathsf{UG}}\right) x_k + n^{\mathsf{G}}, \tag{4.26}$$

where x_k represents the transmitted symbol from the UAV, which satisfies $\mathbb{E}(|x_k|^2) = 1$, and p denotes the transmit power of the UAV for GU k on each channel, and n^{G} denotes the additive white Gaussian noise (AWGN) at GU, with zero mean and variance σ^2. Besides, $h_k^{\mathsf{UG}}, \mathbf{h}^{\mathsf{UI}}$, and $\mathbf{h}_k^{\mathsf{IG}}$ denote the channel coefficients of the links between UAV and GU k, UAV and IRS, IRS and GU k, respectively, which are expressed as $h_k^{\mathsf{UG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{UG}})^2}}, \forall k$, and

$$\begin{aligned} \mathbf{h}^{\mathsf{UI}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{UI}})^2}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}\right]^H \\ \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}\right]^H, \end{aligned} \tag{4.27}$$

$$\begin{aligned} \mathbf{h}_k^{\mathsf{IG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{IG}})^\kappa}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}\right]^H \\ \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}, \ldots, e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}\right]^H \times \alpha^{\mathsf{IG}}, \forall k, \end{aligned} \tag{4.28}$$

where β_0 denotes the channel gain at the reference distance of 1 meter, κ is the path loss exponent, λ is the wavelength of the carrier wave, and α^{IG} is the random scattering components modeled by a circularly symmetric complex Gaussian random variable with zero mean and unit variance. In addition, $(\theta^{\mathsf{UI}}, \xi^{\mathsf{UI}})$ and $(\theta_k^{\mathsf{IG}}, \xi_k^{\mathsf{IG}})$ represent the vertical and horizontal angle-of-departures from the UAV to the IRS and from the IRS to GU k, respectively, which can be calculated from $\sin\theta^{\mathsf{UI}} = \frac{|H-H^{\mathsf{i}}|}{d^{\mathsf{UI}}}$, $\sin\xi^{\mathsf{UI}} = \frac{|x^{\mathsf{i}}-x^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\cos\xi^{\mathsf{UI}} = \frac{|y^{\mathsf{i}}-y^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\sin\theta_k^{\mathsf{IG}} = \frac{H^{\mathsf{i}}}{d_k^{\mathsf{IG}}}$, $\sin\xi_k^{\mathsf{IG}} = \frac{|x^{\mathsf{i}}-x_k^{\mathsf{u}}|}{\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}$, and $\cos\xi_k^{\mathsf{IG}} = \frac{|y^{\mathsf{i}}-y_k^{\mathsf{u}}|}{\sqrt{\left\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}$, $\forall k \in \mathcal{K}$.

Then, we can obtain

$$(\mathbf{h}_k^{\mathsf{IG}})^H \boldsymbol{\Phi} \mathbf{h}^{\mathsf{UI}} = \frac{\beta_0}{d^{\mathsf{UI}}(d_k^{\mathsf{IG}})^{\kappa/2}} \sum_{i_c=1}^{I_c} \sum_{i_r=1}^{I_r} e^{-j\left(\frac{2\pi d}{\lambda}\left((i_r-1)\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+(i_c-1)\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}\right)+\arg(\alpha^{\mathsf{IG}})\right)} e^{j\phi_{i_r,i_c}}$$
$$\times\, e^{-j\frac{2\pi d}{\lambda}\left((i_r-1)\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}+(i_c-1)\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}\right)}|\alpha^{\mathsf{IG}}|$$
$$= \frac{\beta_0}{d^{\mathsf{UI}}(d_k^{\mathsf{IG}})^{\kappa/2}}$$
$$\sum_{i_c=1}^{I_c} \sum_{i_r=1}^{I_r} e^{-j\left(\frac{2\pi d}{\lambda}\left((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}})+(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}})\right)+\arg(\alpha^{\mathsf{IG}})\right)} e^{j\phi_{i_r,i_c}}|\alpha^{\mathsf{IG}}| \tag{4.29}$$

Therefore, the achievable rate for GU k served by the UAV can be expressed as

$$R_k^{\mathsf{A}} = W\log_2\left(1+\frac{p_k}{\sigma^2}\left|\frac{\sqrt{\beta_0}}{d_k^{\mathsf{UG}}}+\frac{\beta_0 f_k|\alpha^{\mathsf{IG}}|}{(d_k^{\mathsf{IG}})^{\kappa/2}d^{\mathsf{UI}}}\right|^2\right), \tag{4.30}$$

where W is the bandwidth of channel and

$$\begin{aligned} f_k &= \sum_{i_c=1}^{I_c}\sum_{i_r=1}^{I_r} e^{j\left(-\frac{2\pi d}{\lambda}\left((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}})+(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}})\right)-\arg(\alpha^{\mathsf{IG}})+\phi_{i_r,i_c}\right)}, \\ &= \sum_{i_c=1}^{I_c}\sum_{i_r=1}^{I_r} e^{j\left(F_k^{i_r,i_c}+\phi_{i_r,i_c}\right)}, \forall k, \end{aligned} \tag{4.31}$$

in which

$$F_k^{i_r,i_c} = -\frac{2\pi d}{\lambda}((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}})+(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}))-\arg(\alpha^{\mathsf{IG}}). \tag{4.32}$$

4.4 Computation Task Model

In a mobile edge computing (MEC) system, computation resources of different capacity are available at mobile devices, cloudlets, and cloud. Therefore, efficient distribution and processing of the computation workloads from different wireless applications play a vital role in designing MEC systems. A general model to describe computation tasks or workload is important as one wants to achieve the broad applicability over different practical applications and mathematical tractability. Various

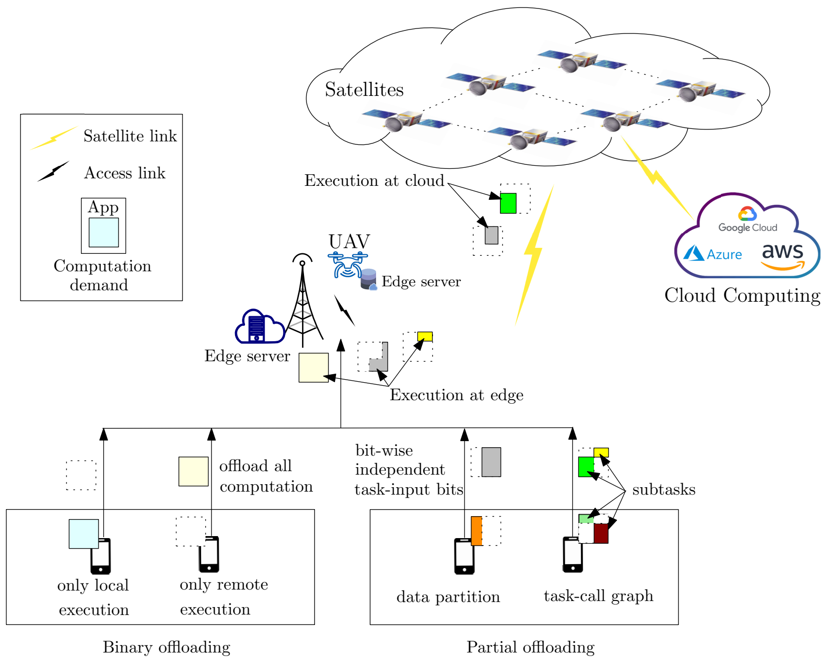

Figure 4.3 – Computation task models.

factors such as context awareness and generality, simplicity, and tractability must be considered in such the model, which must capture the essentials of an application and offer meaningful insights for engineering practice, from the MEC design perspective. In particular, the computation workload of a specific application can be partitioned into sub-tasks in certain practical scenarios and applications [177–179]. Accordingly, the two offloading mechanisms which are binary and partial offloading as shown in Fig. 4.3, have been used in the management of the computation workloads and the offloading decision making process. This section briefly introduces the task models for binary and partial offloading.

4.4.1 Task Model for Binary Offloading

The binary offloading model can be used for computation demand from a compact task that cannot be partitioned and must be executed entirely either at the end device or in a remote cloud or edge server [180, 181].

Such a task can be parameterized by the task input-data size (in bits), the computation workload (CPU cycles), the completion deadline (seconds), and the task output-data size (in bits). These parameters are related to the nature of the applications and can be estimated through task profilers

[182]. In the existing literature, the relation between the computation workload and the task input-data size is captured by using the probabilistic and deterministic relational models. In particular, the number of CPU cycles needed to execute 1-bit of task input data can be modeled as a random variable in case of probabilistic case [181], and as a fixed relation in case of deterministic case [183].

4.4.2 Task Models for Partial Offloading

In practice, many mobile applications have computation demand captured by multiple procedures/tasks. For example, the action recognition application for videos can be decomposed into two main tasks, the first one for capturing the spacial information and the second one for analyzing the temporal information [184]. The partial offloading can be applied in this scenario where a part of each arrival computation demand is locally processed at end devices and the remaining part is remotely executed at cloud/edge servers.

One simple task model for partial offloading is the data-partition model, where the task-input bits are bit-wise independent and they can be divided into different sizes and executed by different entities in MEC systems, e.g., parallel execution at the mobiles and MEC server [185]. Another task model for partial offloading is the task-call graph, which can capture the number of CPU cycles and inter-dependency among different procedures/subtasks in an application [186]. For the task-call graph model, the computation workload of a particular application is captured by subtasks where each subtask can be parameterized by the task input-data size (in bits), the computation workload (CPU cycles), the completion deadline (seconds), and the task output-data size (in bits).

4.5 Satellite Communications

In this section, we will describe the integrated satellite and terrestrial network and inter-satellite communications.

4.5.1 Integrated Satellite and Terrestrial Network

An integrated satellite and terrestrial network has attracted much attention from academia to industry [47, 48, 187, 188]. By using advanced communication such as multiple input multiple output

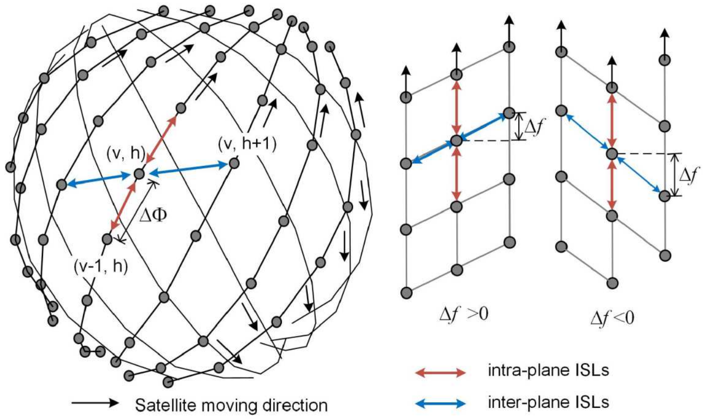

Figure 4.4 – Constellation topology and ISLs [2].

(MIMO) communications, GUs on the ground can communicate directly with the satellites. The design performed in our dissertation captures the GU-satellite communications via the data transmission rate, transmission and propagation time between the GUs and satellites. The values of the related parameters are set similarly to those in [47, 48].

4.5.2 Inter-satellite Communications

In a recent work [2], an algorithm to determine the number of hops, i.e., the number of inter-satellite links (ISLs), and the corresponding satellites to establish the multi-hop communication path between two locations on the ground was proposed, i.e., see Algorithm 1 of [2]. In this section, we would like to describe a general model of multi-hop satellite communications and the hop-count estimation algorithm of [2] in more detail and these results will be used in Chapter. 7.

A satellite constellation and ISLs are illustrated in Fig. 4.4. In particular, the Walker-Delta constellations are considered on [189], where these constellations consist of $N_P \times M_P$ satellites, where N_P is the number of orbit planes and M_P is the number of satellites per plane. All the orbits have the same inclination α and are equally spaced along the equator. The difference of the right ascension of ascending node between adjacent planes is $\Delta\Omega = 2\pi/N_P$. Besides, M_P satellites are evenly distributed in each plane with the phase difference between adjacent satellites equals $\Delta\Phi = 2\pi/M_P$. Moreover, the phase offset between satellites in adjacent planes is given by $\Delta f = 2\pi F/(N_P M_P)$, where F denotes a phasing factor. Furthermore, each satellite is assigned a two-dimension logical index (v, h), denoting the v-th satellite in the h-th orbit plane [2]).

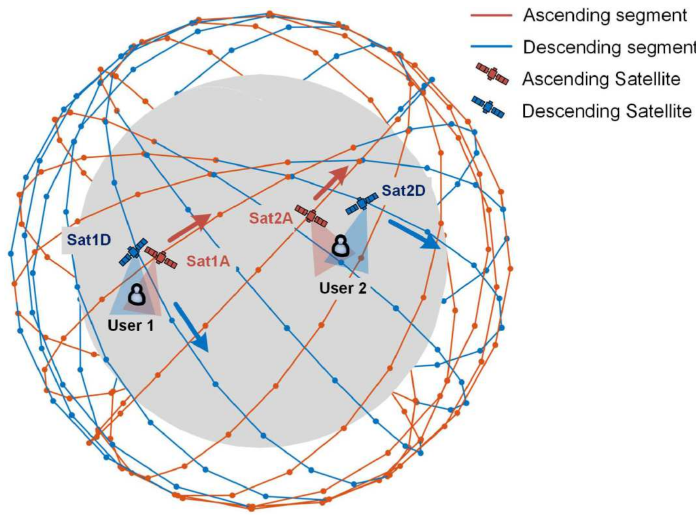

Figure 4.5 – Ascending and descending satellites [2].

A general model is derived to estimate the required minimum ISL hop-count between two users as illustrated in Fig. 4.5. Then, the model is applied to different path modes to derive the hop-count. Moreover, based on the flying direction, all the satellites can be classified into two types: ascending (A) satellites and descending (D) satellites [190]. Hence, the basic idea is to establish a general model to estimate the hop-count based on the ground projection of the satellites and ISL connection mode. Then, considering the access satellite type, variables in the general model are further specified according to four potential path modes, which depends on whether the access satellite is ascending or descending. The final hop-count is the minimum of the four path modes, i.e., A2A, A2D, D2A, D2D.

4.5.2.1 Satellite Ground Location

The sub-satellite point (SSP) of a generic satellite on the ground is represented by its latitude φ and longitude λ. At the generic time t, the SSP location can be determined by

$$\varphi = \arcsin(\sin\alpha \sin u), \tag{4.33}$$

$$\lambda = \zeta(u) + L_0 - \omega_e t, \tag{4.34}$$

$$\zeta(u) = \begin{cases} \arctan(\cos\alpha \tan u), & \text{ascending segment} \\ \arctan(\cos\alpha \tan u)+\pi, & \text{descending segment} \end{cases} \tag{4.35}$$

where α is the orbit inclination and $u \in [-\pi, \pi]$ is the satellite phase angle from its ascending node, which describes the satellite position in orbit. When $u \in [-\pi/2, \pi/2]$, the satellite is in the ascending segment and flies towards the northeast, while $u \in [-\pi, -\pi/2) \cup (\pi/2, \pi]$ means the descending segment towards the southeast. In addition, $\zeta(u)$ denotes the longitude difference from the satellite to its ascending node, which varies with the satellite phase. L_0 represents the initial longitude of the orbit ascending node, which is an absolute parameter determining the orbit plane position, and ω_e denotes earth rotation speed.

4.5.2.2 System Design

Consider two users whose locations are specified as follows: User 1 (φ_1, λ_1) and User 2 (φ_2, λ_2), where $\varphi_1, \varphi_2 \in [-\alpha, \alpha]$ and $\lambda_1, \lambda_2 \in [-\pi, \pi]$. As assumption, Sat 1 and Sat 2 are also at (φ_1, λ_1) and (φ_2, λ_2). Let Sat 1 be the one on the west, then the longitude difference is $\Delta\lambda = \lambda_2 - \lambda_1, \Delta\lambda \in [0, \pi]$. The difference of the right ascension of ascending node between Sat 1 and Sat 2 is

$$\Delta L_0 = L_{0,2} - L_{0,1} = H_h \Delta\Omega. \tag{4.36}$$

Besides, we can obtain another expression for ΔL_0 as

$$\Delta L_0 = \Delta\lambda + \zeta(u_1) - \zeta(u_2). \tag{4.37}$$

Then, the inter-plane hop-count H_h can be calculated by

$$H_h = \text{Round}\left[\frac{\Delta L_0}{\Delta\Omega}\right] = \text{Round}\left[\frac{\Delta\lambda + \zeta(u_1) - \zeta(u_2)}{\Delta\Omega}\right], \tag{4.38}$$

where Round$[x]$ represents the standard rounding function that returns the integer closet to x.

Since each intra-plane and inter-plane relay respectively add $\Delta\Phi$ and Δf to satellite phase angle, the phase angle difference between Sat 1 and Sat 2 is

$$\Delta u = u_2 - u_1 = H_v \Delta\Phi + H_h \Delta f, \tag{4.39}$$

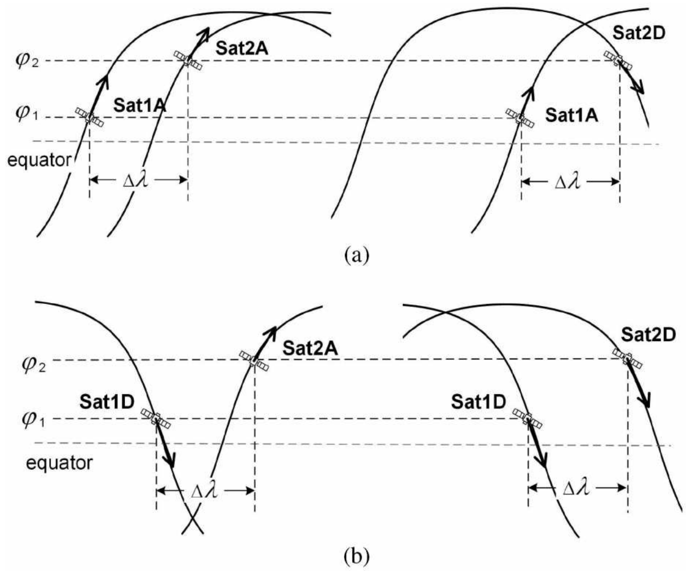

Figure 4.6 – Illustration of different path modes using satellite ground tracks [2].

Table 4.1 – Specified values of u and $\zeta(u)$ in different path modes [2]

Path mode	A2A	A2D	D2A	D2D
u_1	$u_1 = \arcsin\frac{\sin\varphi_1}{\sin\alpha}$	$u_1 = \arcsin\frac{\sin\varphi_1}{\sin\alpha}$	$u_1 = \frac{\varphi_1}{\lvert\varphi_1\rvert}\pi - \arcsin\frac{\sin\varphi_1}{\sin\alpha}$	$u_1 = \frac{\varphi_1}{\lvert\varphi_1\rvert}\pi - \arcsin\frac{\sin\varphi_1}{\sin\alpha}$
u_2	$u_2 = \arcsin\frac{\sin\varphi_2}{\sin\alpha}$	$u_2 = \frac{\varphi_2}{\lvert\varphi_2\rvert}\pi - \arcsin\frac{\sin\varphi_2}{\sin\alpha}$	$u_2 = \arcsin\frac{\sin\varphi_2}{\sin\alpha}$	$u_2 = \frac{\varphi_2}{\lvert\varphi_2\rvert}\pi - \arcsin\frac{\sin\varphi_2}{\sin\alpha}$
$\zeta(u_1)$	$\arctan(\cos\alpha\tan u_1)$	$\arctan(\cos\alpha\tan u_1)$	$\pi + \arctan(\cos\alpha\tan u_1)$	$\pi + \arctan(\cos\alpha\tan u_1)$
$\zeta(u_2)$	$\arctan(\cos\alpha\tan u_2)$	$\pi + \arctan(\cos\alpha\tan u_2)$	$\arctan(\cos\alpha\tan u_2)$	$\pi + \arctan(\cos\alpha\tan u_2)$

where u_1 and u_2 satisfy

$$\sin u = \sin\varphi / \sin\alpha. \tag{4.40}$$

Hence, the intra-plane hop-count can be calculated as

$$H_v = \text{Round}\left[\frac{\Delta u - H_h \Delta f}{\Delta\Phi}\right]. \tag{4.41}$$

Therefore, the total hop-count between two access satellites is

$$H = |H_h| + |H_v|. \tag{4.42}$$

Fig. 4.6 illustrates a specific topology of different path modes using satellite ground tracks [2]. Specifically, Sat1A and Sat2A are on the ascending orbit segment, while Sat1D and Sat2D are on the descending segment. (a) The access satellite of User 2 can be Sat2A or Sat2D, corresponding

Algorithm 4.2. Hop-Count Estimation Between Two Ground Users [2].

Input: $\varphi_1, \varphi_2, \Delta\lambda$
Output: H
for X2X **in** A2A, A2D, D2A, A2D **do**
  Specify $u_1, u_2, \zeta(u_1)$ and $\zeta(u_2)$ in Table 4.1
  $\Delta L_0 \leftarrow \Delta\lambda + \zeta(u_1) - \zeta(u_2)$
  if $|\Delta L_0| > \pi$ **then**
    $\overline{\Delta L_0} \leftarrow \mod(\Delta L_0 + \pi, 2\pi) - \pi$
  end if
  $H_h \leftarrow \text{Round}(\overline{\Delta L_0}/\Delta\Omega)$
  $\Delta U \leftarrow u_2 - u_1 - H_h \Delta f$
  if $|\Delta U| > \pi$ **then**
    $\overline{\Delta U} \leftarrow \mod(\Delta U + \pi, 2\pi) - \pi$
  end if
  $H_v \leftarrow \text{Round}(\overline{\Delta U}/\Delta\Phi)$
  $H^{\mathsf{X2X}} = |H_h| + |H_v|$
end for
$H \leftarrow \min(H^{\mathsf{A2A}}, H^{\mathsf{A2D}}, H^{\mathsf{D2A}}, H^{\mathsf{D2D}})$
Return H.

to A2A or A2D path mode, respectively. (b) The access satellite of User 1 is assumed at the descending segment, corresponding to D2A and D2D path mode. Furthermore, the satellite phase angle u from its ascending node and longitude difference from the satellite to its ascending node $\zeta(u)$ for four path modes are specified in Table. 4.1. Finally, the overall algorithm to estimate the number hop-count between two ground users (or two satellites) is summarized in Algorithm 4.2. This algorithm considers the access satellite types and variables in the general model corresponding to four potential path modes and the final hop-count is the minimum of the four path modes $H = \min(H^{\mathsf{A2A}}, H^{\mathsf{A2D}}, H^{\mathsf{D2A}}, H^{\mathsf{D2D}})$.

Chapter 5

Integrated UAV Trajectory Control and Resource Allocation for UAV-Based Wireless Networks with Co-channel Interference Management

The content of this chapter was published in the following paper:

Minh Dat Nguyen, Long Bao Le, and André Girard , "Integrated UAV Trajectory Control and Resource Allocation for UAV-Based Wireless Networks With Co-Channel Interference Management," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12754–12769, Jul. 2022.

5.1 Abstract

In this chapter, we study the trajectory control, sub-channel assignment, and user association design for unmanned aerial vehicles (UAVs)-based wireless networks. We propose a method to optimize the max-min average rate subject to data demand constraints of ground users (GUs) where spectrum reuse and co-channel interference management are considered. The mathematical model is a mixed integer non-linear optimization problem which we solve by using the alternating optimiza-

tion approach where we iteratively optimize the user association, sub-channel assignment, and UAV trajectory control until convergence. For the sub-channel assignment sub-problem, we propose an iterative sub-channel assignment (ISA) algorithm to obtain an efficient solution. Moreover, the successive convex approximation (SCA) is used to convexify and solve the non-convex UAV trajectory control sub-problem. Via extensive numerical studies, we illustrate the effectiveness of our proposed design considering different UAV flight periods and number of sub-channels and GUs as compared with a simple heuristic.

5.2 Introduction

Next-generation wireless communications networks are expected to provide much higher capacity, lower latency, better communication reliability and stability for billions of devices anywhere and anytime [5, 191]. However, deployment of an efficient fixed terrestrial wireless infrastructure can be quite challenging in certain scenarios, e.g., emergency situations such as natural disasters and fast service recovery. To this end, unmanned aerial vehicle (UAV) communications can overcome certain limitations of a fixed wireless communications infrastructure where UAV communications can improve the coverage, users' quality of service (QoS), and the communication resilience and availability thanks to their attributes such as mobility, flexibility, and controllable altitude [9, 14, 70, 192].

The design of effective UAV-enabled communications networks, however, is quite challenging [9]. First, channel modeling for different UAV communication is a major research challenge. Specifically, communication channels for air-to-ground (A2G) communications between UAVs and ground users must be appropriately modeled considering possible line-of-sight (LoS) and non-line-of-sight (NLoS) propagation conditions. Second, efficient deployment of UAVs in three-dimensional (3D) space or effective control of UAVs' trajectories significantly impacts communications performance such as UAV's flight time, energy consumption, and GUs' QoS. Finally, development of resource allocation algorithms that can efficiently manage and assign various types of network resources, including communication bandwidth and transmit power for users, is of critical importance for UAV-enabled communications networks.

Generally, UAVs can act as mobile users, relays, or flying base stations (BSs) to enhance the coverage and capacity of wireless networks. There has been a great deal of research on these UAV communications scenarios in recent years. In particular, research on data collection for wireless nodes in wireless sensor networks (WSNs) and Internet of Things (IoT) leveraging the UAV communication has been an active research topic [61–64]. In [61], a cellular-enabled UAV communication setting with given UAV's initial and final locations of a single UAV was considered where the design goal was to minimize the UAV's mission completion time by optimizing its trajectory. The optimization of the UAV's trajectory was also studied in [62] for a WSN where one UAV is used as a mobile data collector to minimize the maximum energy consumption of all sensor nodes. Similarly, the authors of [63] proposed an energy-efficient framework for a WSN using a flying UAV whose stopping positions were optimized for efficient data collection. Finally, placement optimization for multiple UAVs was studied to achieve low overhead and high sensor search accuracy in [64].

There have also been much work for UAV-enabled wireless networks in which UAVs act as relays [65–69]. Specifically, two-dimensional placement or trajectory optimization of UAVs have been studied where UAVs are mostly assumed to stay at a fixed altitude. The joint power and UAV's trajectory optimization to maximize the end-to-end throughput from a source to its destination for the UAV-based relay network was also studied in [65], For the network setting with multiple UAVs, the authors of [66] considered the joint optimization of the power, bandwidth allocation, and UAVs' trajectories to maximize the average end-to-end throughput. In [67], the authors studied a UAV-to-ground secure communication system with a single UAV relay at a fixed altitude and multiple eavesdroppers. Here, the design objective is to maximize the minimum secrecy rate by jointly optimizing the transmit power and location of the UAV. Moreover, by using a UAV full-duplex relay at a fixed altitude, the authors of [68] jointly optimized the transmit power and UAV's trajectory to achieve the maximum destination node's throughput considering spectrum sharing with terrestrial device-to-device (D2D) communications. The authors of [69] considered the 3D placement of a single UAV acting as a relay for a network with multiple communication pairs of source and destination nodes on the ground. Here, the design objective was to maximize the network throughput by jointly optimizing the transmit power, bandwidth allocation, and UAV placement.

Finally, UAVs can serve as aerial BSs to provide on-the-fly communications and enhance the performance of the terrestrial wireless networks. In fact, UAV placement and trajectory control optimization have been studied in [41, 72–77]. For the UAV placement, the authors of [72] proposed an analytical framework to optimize the UAV's altitude providing maximum coverage. The authors of [73] optimized the required number of UAVs and their positions to provide the best wireless coverage for a group of GUs. The joint bandwidth and power allocation for multi-hop UAV based downlink communications was studied in [74] where orthogonal bandwidth allocation for both access and backhaul links was considered. In our previous work [75], we considered the joint placement for stationary UAVs and non-orthogonal bandwidth allocation for wireless access links. In [76], the joint optimization of power allocation and trajectory control for the frequency-division multiple access (FDMA) UAV-based wireless network was studied. In [77], the joint optimization of user association, power allocation, and UAV's trajectory control for the wireless network with multiple UAVs was considered assuming no co-channel interference. Moreover, the authors in [41] jointly optimized the scheduling, user association, power allocation, and UAV's trajectory control for the time-division multiple access (TDMA) based wireless network.

While the papers mentioned above have considered the optimization of the UAV placement or UAV trajectory control and resource allocation in different network settings, there are still various research issues deserving more in-depth studies. In particular, spectrum reuse to support communications between multiple UAVs and GUs is needed in practice to enhance the spectrum efficiency and network performance; however, efficient co-channel interference management techniques must be developed. Moreover, many practical application scenarios such as data collection, information sharing require to guarantee data transmission demand constraints of individual GUs. To fill these research gaps, we study in this chapter the joint UAV-GU association, resource allocation, and UAV trajectory control for UAV-based wireless networks with spectrum reuse and interference management. The main contributions can be summarized as follows:

- We formulate the joint UAV-GU association, UAV trajectory control, and non-orthogonal sub-channel assignment problem for UAV-based wireless networks. We maximize the minimum average rate of all GUs considering constraints on data transmission demands of individual GUs.
- We solve the underlying mixed-integer non-linear optimization problem (MINLP) problem using the alternating optimization approach. We solve the UAV-GU association, sub-channel assignment, and UAV trajectory control sub-problems separately in each iteration until convergence. We develop an iterative sub-channel assignment (ISA) algorithm to tackle the sub-

channel assignment sub-problem. Given the UAV-GU association and sub-channel assignment solutions, the UAV trajectory control sub-problem is a difficult non-convex problem. We propose to use the successive convex approximation (SCA) technique to convexify and solve this sub-problem. We then present a short complexity analysis of the proposed algorithm.

- Extensive numerical results are presented to show the performance of our algorithm. Specifically, we compare the network performance when the proposed ISA sub-channel algorithm and a baseline heuristic sub-channel assignment with interference management (SAIM) algorithm are used to solve the joint problem. We also study the impacts of different parameters and the importance of trajectory control on the achievable performance. Finally, we illustrate the convergence of the algorithm.

The remainder of this chapter is organized as follows. In Section 5.3, we discuss the related work on joint optimization of UAV trajectory control and bandwidth allocation. Section 5.4 presents the system model and problem formulation. In Section 5.5, we describe how we solve the three sub-problems and provide the convergence and complexity analysis of the proposed algorithm. Section 5.6 presents the numerical results for performance evaluations of the algorithm. Finally, Section 5.7 concludes the chapter.

5.3 Related Work

A summary of recent work on joint UAV trajectory control and bandwidth allocation is given in Table 5.1. In fact, data transmission demand constraints and spectrum reuse with interference management have a significant impact on the achievable performance of UAV-based wireless networks; however, taking the these aspects into account makes the design very challenging. Therefore, we include these design aspects in addition to others for related work in Table 5.1. This table confirms that our current work considers all key design aspects and provides fairness for GUs by maximizing the minimum average rate of the GUs so that our work presents a more extensive design framework for UAV-based wireless networks compared to the existing literature.

Table 5.1 – Related work on UAV trajectory and bandwidth allocation for UAV-based wireless networks

Ref.	Type	Objective	Data Demand Constraint	Trajectory Optimization	Bandwidth Allocation	Spectrum Reuse
[134]	Single UAV	Minimize UAV energy consumption	No	Yes	Continuous bandwidth assignment	No
[135]	Single UAV	Maximize system energy efficiency	Yes	Yes	Continuous bandwidth assignment	No
[136]	Single UAV	Maximize minimum average rate of delay-tolerant users	Yes	Yes	Continuous bandwidth assignment	No
[137]	Multiple UAVs	Maximize uplink sum-rate	Yes	No	Sub-channel assignment	Yes
[138]	Multiple UAVs	Maximize system energy efficiency	Yes	Yes	Sub-channel assignment	Yes
[39]	Single UAV	Maximize minimum average rate of GUs	No	Yes	Sub-channel assignment	No
[139]	Single UAV	Maximize minimum average rate of GUs	Yes	Yes	Sub-channel assignment	No
[140]	Multiple UAVs	Maximize minimum average rate of GUs	Yes	Yes	Sub-channel assignment	Yes
This Work	Multiple UAVs	Maximize minimum average rate of GUs	Yes	Yes	Sub-channel assignment	Yes

On the one hand, the paper [134] jointly optimized the power, continuous bandwidth assignment, and 3D UAV's trajectory where the objective is to minimize the total UAV's energy consumption. This work did not consider data transmission demand constraints and co-channel interference. A block coordinate descent algorithm was used to iteratively optimize the resource allocation and UAV's trajectory control. The authors in [135] addressed the joint design of user scheduling, transmit power, continuous bandwidth assignment, and UAV's trajectory control in the 3D space to maximize the system energy efficiency. However, this work did not consider co-channel interference and an iterative algorithm using the Dinkelbach and block coordinate descent techniques was proposed to solve the problem. The authors in [136] also studied the UAV's trajectory and continuous bandwidth assignment without considering co-channel interference. The design goal was to maximize the minimum average rate of GUs using an alternating optimization technique.

On the other hand, the work [39, 137–140] mainly studied the sub-channel assignment and UAV trajectory control. An exception is [137] where the authors discussed the sub-channel assignment while considering co-channel interference and UAV velocity control with a known trajectory to maximize the uplink sum rate through an iterative algorithm. The authors of [138] studied the network setting with only two UAVs i.e., transmitter and jammer, to maximize the system energy efficiency by jointly optimizing the transmit power, sub-channel assignment, and UAV trajectory control using an alternating optimization algorithm with a relaxation of the binary sub-channel assignment decision variables. Moreover, the design in [39, 139, 140] was to maximize the minimum average rate of GUs. The co-channel interference was considered only in [140]. The sub-channel assignment was investigated in [39] for a network where the UAV used orthogonal frequency-division multiple access (OFDMA). Bandwidth, power allocation, and UAV's trajectory control were jointly optimized using the block coordinate descent method. A backhaul-aware design to maximize the minimum average rate for GUs was proposed in [139,140] using an alternating optimization approach to solve the joint problem of sub-channel assignment and UAV's trajectory control.

Even though there has been some limited work on the joint sub-channel assignment and UAV's trajectory design considering spectrum reuse and co-channel interference management, this research direction remains under-explored in the UAV communication literature. In our preliminary work [140], we developed a heuristic algorithm for the UAV trajectory control and sub-channel assignment problem. The present work makes several significant extensions of this conference work. Specifically, we solve three sub-problems, namely the UAV-GU association, sub-channel assignment and UAV trajectory control, and develop an integrated algorithm to solve the joint optimization problem of UAV-GU association, sub-channel assignment and UAV trajectory control. Moreover, we give a complexity analysis and prove the convergence of the integrated algorithm. Finally, much more extensive numerical results are presented in this chapter compared to those in the conference paper to demonstrate the efficiency and desirable performance of the proposed algorithm.

5.4 System Model

We consider a network where a set of UAVs denoted as $\mathcal{M} = \{1, ..., M\}$, provides wireless connectivity for a set of GUs, denoted as $\mathcal{K} = \{1, ..., K\}$ as shown in Fig. 5.1.

We assume that each GU needs to receive a specific amount of data from UAVs in the downlink direction. This can be the case in many practical scenarios, e.g., GUs want to receive video files from the UAV such as specific scenes of a football match. Because the UAVs are flying at a relatively high altitude, we assume that all communications, be it UAV-to-BS or UAV-to-GU, are dominated by line-of-sight (LoS) propagation.

The UAVs are assumed to be connected to the core network wirelessly through one cellular BS where the UAV-BS links are assumed to have sufficiently large capacity i.e., by using mmWave communications. The assumption of LoS propagation for these channels is necessary because mmWave communications are very sensitive to blockage, which degrades the communication rate and relia-

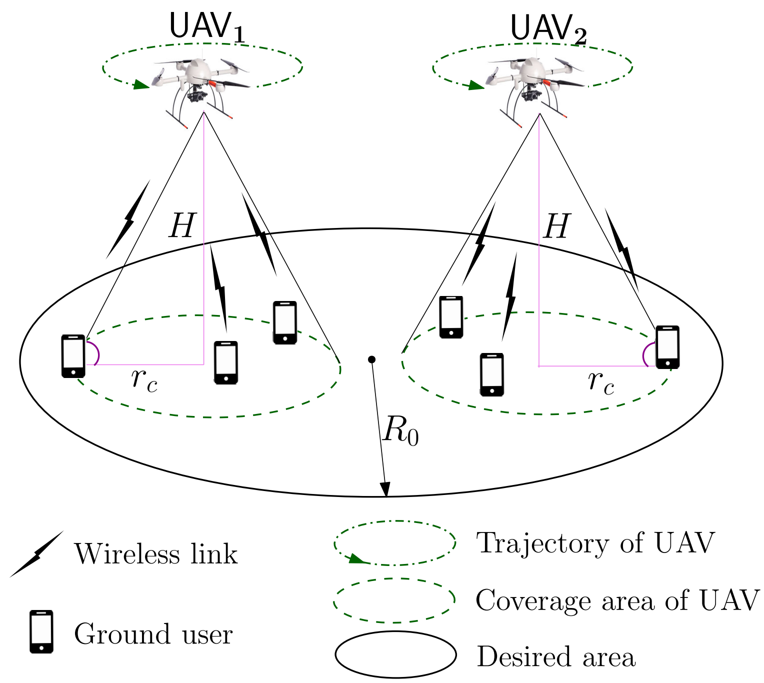

Figure 5.1 – UAV based wireless network.

bility significantly. Also, the large bandwidth available at mmWave bands enables us to achieve the high capacity required by the backhaul links [193].

The UAV-GU channels, on the other hand, don't use mmWave communications since there are still many unresolved issues on the design of the hardware and physical layer for the transceivers deployed on UAV-based mmWave communications. For instance, more work is needed on highly directional antennas with efficient beamforming training and tracking to take into account the UAV movement and the channel Doppler effect. This is particularly difficult since the UAV's position and GUs discovery are tightly coupled [9]. For these reasons, we assume that only omni-directional antennas are installed on the UAVs to enable the low-complexity transceivers needed to achieve omnipresent coverage and dynamic GUs connections.

We assume that the UAVs fly at a fixed altitude H over a flight period of $T > 0$ seconds. The flight period is divided into N time slots where the set of time slots is denoted as $\mathcal{N} = \{1, ..., N\}$. At any time slot during the flight period T, each UAV can communicate with multiple GUs at the same time using OFDMA. The GUs are assumed to be located on the ground at zero altitude with fixed horizontal coordinates $\mathbf{r}_k^{\mathtt{u}} = (x_k^{\mathtt{u}}, y_k^{\mathtt{u}}), \forall k \in \mathcal{K}$.

Table 5.2 – Key notations

Known Parameters	
C	Number of sub-channels
$\mathcal{C}$	Set of sub-channels
d_{min}	Minimum inter-UAV distance
D_k^{min}	Minimum data transmission demand of GU k
H	Fixed altitude of UAVs
K	Number of ground users (GUs)
$\mathcal{K}$	Set of GUs
M	Number of UAVs
$\mathcal{M}$	Set of UAVs
N	Total time slots
$P_{\max}$	Total transmit power of UAV
p	Transmit power on each sub-channel ($p = P_{\max}/C$)
$\mathbf{r}_k^{\mathsf{u}}$	Fixed horizontal coordinate of GU k ($\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}})$)
r_0	Center location of the considered network area
R_0	Radius of the considered network area
r_{c}	Radius of the circular cluster area
σ^2	Power of the additive white Gaussian noise (AWGN)
T	Flight period
Δt	Element slot length ($\Delta t = T/N$)
$V_{\max}$	Maximum speed of UAV
S_{max}	Maximum horizontal distance that the UAV can travel in each time slot $\left(S_{\mathsf{max}} \triangleq V_{\max}\Delta t\right)$
W	Bandwidth of each sub-channel
Decision and Auxiliary Variables	
$\omega_{k,m}[n]$	Association between UAV m and GU k in time slot n
$\theta_{z,c}[n]$	The sub-channel c is assigned to GU k in time slot n
$\mathbf{q}_m[n]$	Time-variant horizontal coordinate of the UAV m in time slot n $\left(\mathbf{q}_m[n] = (x_m^{\mathsf{d}}[n], y_m^{\mathsf{d}}[n])\right)$
$\mathbf{\Omega}$	Vector of all UAVs-GUs association decision variables $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$
$\mathbf{\Theta}$	Vector of all sub-channel assignment decision variables $\mathbf{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$
$\mathbf{Q}$	Vector of all time-variant horizontal coordinate of the UAVs $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$
η	Minimum average rate of all GUs $\left(\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right) = \min_{k\in\mathcal{K}} \bar{R}_k\right)$
Functions	
$\mathcal{C}_m[n]$	Set of sub-channels used by UAV m in time slot n $\lvert\mathcal{C}_m[n]\rvert = \sum_{k=1}^{K}\sum_{c=1}^{C}\omega_{k,m}[n]\theta_{k,c}[n]$, where $\lvert\mathcal{C}_m[n]\rvert$ denotes the number of sub-channels used by UAV m in time slot n
$d_{k,m}[n]$	Distance between UAV m and GU k in time slot n
$g_{k,m}[n]$	Channel power gain from UAV m and GU k in time slot n
$\gamma_{k,m,c}[n]$	Signal-to-interference-plus-noise ratio (SINR) at GU k served by UAV m in time slot n
ρ_0	Channel power gain at the distance of 1 m
$R_{k,m,c}[n]$	Achievable rate of GU k served by UAV m in time slot n on the sub-channel c
$R_k[n]$	Total rate achieved by GU k in time slot n
$\bar{R}_k$	Average rate per all slots of GU k

Let C be the number of sub-channels available to support the wireless access links between UAVs and GUs. We denote the total transmit power of each UAV as $P_{\max} \geq 0$. We assume that the uniform power allocation is used by each UAV i.e., the transmit power on each sub-channel is equal the total transmit power $P_{\max}$ divided by the total sub-channels used for downlink communications and is given by

$$p = \frac{P_{\max}}{C}. \tag{5.1}$$

The list of key notations in this chapter is given in Table 5.2.

5.4.1 UAV-GU Association

Given the locations of the UAVs in each time slot, GUs need to be associated with the UAVs offering high-quality communications. For this, we define the binary UAV-GU association decision variable $\omega_{k,m}[n]$ which is equal to 1 if GU k is served by UAV m in time slot n and equal to 0, otherwise. Since each GU is associated with exactly one UAV in each time slot, they must meet the constraints

$$\sum_{m=1}^{M} \omega_{k,m}[n] = 1, \forall k, n. \tag{5.2}$$

5.4.2 Sub-channel Assignment

In addition to the UAV assignment, let W (MHz) denote the bandwidth of each sub-channel and $\mathcal{C} = \{1, ..., C\}$ denote the set of sub-channels. Besides, we have to decide the set of sub-channels to be assigned for each GU. The corresponding decision variables are defined as

$$\theta_{k,c}[n] = \begin{cases} 1 & \text{if sub-channel } c \text{ is assigned to GU } k \\ & \text{in time slot } n \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

The first requirement for the assignment is that each GU must be assigned at least one sub-channel at all times in order to maintain a continuous communication. This can be expressed as

$$\sum_{c=1}^{C} \theta_{k,c}[n] \geq 1, \forall k, n. \tag{5.4}$$

Another constraint on the assignment is that for a given UAV, a sub-channel can be used to support only one GU. This leads to a coupling between the ω and θ variables expressed as

$$\sum_{k=1}^{K} \omega_{k,m}[n]\theta_{k,c}[n] \leq 1,\ \forall m, n, c, \tag{5.5}$$

This constraint is the reason why we need to distinguish between the channels, even though they all have the same bandwidth.

5.4.3 UAV Trajectory Control

We optimize the UAVs' trajectories over the flight period T. This can be typically performed to achieve performance targets in data throughput and delay [41]. We assume that the UAV's energy is sufficiently large to cover its flight operation and wireless communications over the flight period T. This assumption is supported by [194] where a UAV equipped with a 3-cell, 3250mAh, and 11.1V LiPo battery can have a flight time of about 20 minutes.

The horizontal coordinate of UAV m in time slot n is denoted as $\mathbf{q}_m[n] = (x^{\mathsf{d}}_m[n], y^{\mathsf{d}}_m[n])$. We assume that each UAV m must come back to its initial position at the end of the flight period, i.e., its trajectory must satisfy the following constraint:

$$\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m \in \mathcal{M}. \tag{5.6}$$

The slot interval $\Delta t = T/N$ is set sufficiently small so that each UAV just flies a small distance during each time slot even at the maximum speed $V_{\max}$. Hence, the UAVs' trajectories must satisfy the following constraints:

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \leq S^2_{\mathsf{max}}, n=1, ..., N-1, \forall m, \tag{5.7}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d^2_{\mathsf{min}}, \forall n, m, j \neq m, \tag{5.8}$$

where $\|.\|$ denotes the Euclidean norm, $S_{\mathsf{max}} \triangleq V_{\max}\Delta t$ is the maximum horizontal distance that the UAV can travel in each time slot, d_{min} denotes the minimum inter-UAV distance and constraints (5.8) are imposed to ensure collision avoidance among UAVs.

5.4.4 Communication Model

Recall that we have assumed that the communication links from UAVs to GUs are dominated by the LoS propagation where the channel quality is mostly dependent on the UAV-GU distance. In time slot n, the distance between UAV m and GU k can be calculated as

$$d_{k,m}[n] = \sqrt{H^2 + \|\mathbf{q}_m[n] - \mathbf{r}_k^{\mathsf{u}}\|^2}. \tag{5.9}$$

The channel power gain from UAV m to GU k in time slot n on sub-channel c is assumed to follow the free-space path loss model and it can be expressed as

$$g_{k,m}[n] = \rho_0 d_{k,m}^{-2}[n] = \frac{\rho_0}{H^2 + \|\mathbf{q}_m[n] - \mathbf{r}_k^{\mathsf{u}}\|^2}, \tag{5.10}$$

where ρ_0 presents the channel power gain at the reference distance of 1 m. The received signal to interference plus noise ratio (SINR) at GU k on sub-channel c can be calculated as

$$\gamma_{k,m,c}[n]=\frac{p g_{k,m}[n]}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n]+\sigma^2}, \tag{5.11}$$

where σ^2 is the power of the additive white Gaussian noise (AWGN) at the receiver. The term $\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n]$ represents the interference at GU k on the sub-channel c due to the transmissions of other UAVs in time slot n on this sub-channel. The achievable rate of GU k served by UAV m in time slot n on the sub-channel c, denoted by $R_{k,m,c}[n]$ in bits/second (bps), can then be expressed as

$$R_{k,m,c}[n] = \omega_{k,m}[n]\theta_{k,c}[n] W \log_2\left(1 + \gamma_{k,m,c}[n]\right). \tag{5.12}$$

Therefore, the total rate achieved by GU k in time slot n, denoted by $R_k[n]$, can be written as

$$R_k[n] = \sum_{m=1}^{M} \sum_{c=1}^{C} R_{k,m,c}[n]. \tag{5.13}$$

As a result, the average rate per slot of GU k over N time slots can be expressed as

$$\begin{aligned} \bar{R}_k &= \frac{1}{N} \sum_{n=1}^{N} R_k[n] \\ &= \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{c=1}^{C} \omega_{k,m}[n] \theta_{k,c}[n] W \log_2 \left(1+\gamma_{k,m,c}[n]\right). \end{aligned} \tag{5.14}$$

5.4.5 Problem Formulation

For convenience, we gather different decision variables as $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$, $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$ and $\mathbf{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$. Our design goal is to maximize the minimum average rate achieved by all GUs by jointly optimizing the user association i.e., $\mathbf{\Omega}$, sub-channel assignment i.e., $\mathbf{\Theta}$, and UAV trajectory i.e., $\mathbf{Q}$ over all time slots of the flight period.

The average rate $\bar{R}_k$ in (5.14) is a non-linear function with respect to three decision variables $\mathbf{\Omega}, \mathbf{\Theta}$, and $\mathbf{Q}$. Instead of performing the max-min optimization of this non-linear function, we introduce the function $\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right) = \min_{k \in \mathcal{K}} \bar{R}_k$ as the minimum average rate of all GUs. Then, our optimization problem becomes equivalent to maximizing $\eta\left(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}\right)$, which is more tractable. Moreover, we assume that GU $k, \forall k \in \mathcal{K}$, has the minimum data transmission demand of D_k^{min}, which must be received in the downlink direction over the UAV flight period. Then, the joint UAV-GU association, sub-channel assignment, and UAV trajectory control optimization problem to maximize the minimum average rate over all GUs can be formulated as

$$
\begin{aligned}
\textbf{(P1):} \quad & \max_{\eta, \boldsymbol{\Omega}, \boldsymbol{\Theta}, \mathbf{Q}} \ \eta && (5.15)\\
& \text{s.t. } \bar{R}_k \geq \eta, \ \forall k, && (5.15a)\\
& \sum_{n=1}^{N} \Delta t R_k[n] \geq D_k^{\mathsf{min}}, \ \forall k, && (5.15b)\\
& \|r_0 - \mathbf{q}_m[n]\| \leq R_0, \ \forall m, n, && (5.15c)\\
& \sum_{m=1}^{M} \omega_{k,m}[n] = 1, \ \forall k, n, && (5.15d)\\
& \sum_{k=1}^{K} \omega_{k,m}[n]\theta_{k,c}[n] \leq 1, \forall m, n, c, && (5.15e)\\
& \sum_{c=1}^{C} \theta_{k,c}[n] \geq 1, \forall k, n, && (5.15f)\\
& \mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m, && (5.15g)\\
& \|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \leq S_{\mathsf{max}}^2, \ n{=}1, ..., N{-}1, && (5.15h)\\
& \|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d_{\mathsf{min}}^2, \ \forall n, m, j \neq m, && (5.15i)\\
& \omega_{k,m}[n] \in \{0, 1\}, \forall k, m, n, && (5.15j)\\
& \theta_{k,c}[n] \in \{0, 1\}, \forall k, c, n, && (5.15k)
\end{aligned}
$$

where R_0 represents the radius of the network area centered at r_0. Constraints (5.15b) capture the required data transmission demand for each GU over the flight period of T seconds. while constraints (5.15c) restrict the trajectories of all UAVs inside the desired network area. Moreover, (5.15d)-(5.15e) present the UAV-GU association constraints, (5.15e)-(5.15f) capture constraints on the sub-channel assignment, and (5.15g)-(5.15i) represent constraints on the UAVs' trajectories. It can be seen that the constraints (5.15a), (5.15b), and (5.15i) are non-linear and integer decision variables are involved in (5.15j) and (5.15k) for the UAV-GU association and sub-channel assignment, respectively. Hence, problem (5.15) is a mixed integer non-linear optimization problem (MINLP), which is difficult to solve optimally. In the following section, we describe how to compute good feasible solutions to this problem.

5.5 Proposed Algorithm

We adopt the alternating optimization approach to solve problem (5.15) where we iteratively optimize each set of variables given the values of other variables in the corresponding sub-problems until convergence. We describe how to solve these different sub-problems in the following.

5.5.1 UAV-GU Association Given Sub-channel Assignment and UAV Trajectory Control

For the given sub-channel assignment $\mathbf{\Theta}$ and UAV trajectory $\mathbf{Q}$, the problem of optimizing the UAV-GU association $\mathbf{\Omega} = \{\omega_{k,m}[n], \forall k, m, n\}$ to achieve the max-min average rate over all GUs is still a integer non-linear optimization problem. To make the problem more tractable, we relax the integer decision variables in $\mathbf{\Omega}$ into continuous decision variables, which yields the following problem

$$\textbf{(P1.1):} \max_{\eta, \mathbf{\Omega}} \ \eta \tag{5.16}$$

$$\text{s.t.} \quad 0 \le \omega_{k,m}[n] \le 1, \forall k, m, n, \tag{5.16a}$$

$$\text{constraints } (5.15\text{a}), (5.15\text{b}), (5.15\text{d}), (5.15\text{e}).$$

Even with this relaxation, problem (5.16) is still a non-convex optimization problem due to the non-convex constraints (5.15a) and (5.15b). To this end, $R_{k,m,c}[n]$, in constraints (5.15a) and (5.15b), can be re-written as

$$\begin{aligned} R_{k,m,c}[n] &= \omega_{k,m}[n]\theta_{k,c}[n] W \log_2\left(1 + \frac{p g_{k,m}[n]}{\sum_{j=1, j\neq m}^{M} \sum_{z=1, z\neq k}^{K} \omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2}\right) \\ &\ge \omega_{k,m}[n]\theta_{k,c}[n] W R^{\mathsf{A}}_{k,m,c}[n], \end{aligned} \tag{5.17}$$

where

$$R^{\mathsf{A}}_{k,m,c}[n] \le \log_2\left(1 + \frac{p g_{k,m}[n]}{\sum_{j=1, j\neq m}^{M} \sum_{z=1, z\neq k}^{K} \omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2}\right). \tag{5.18}$$

By introducing auxiliary variables $\mathbf{R}^{\mathsf{A}} = \{R^{\mathsf{A}}_{k,m,c}[n], \forall k, m, c, n\}$, problem (5.16) can be reformulated as

$$\textbf{(P1.1'):} \max_{\eta, \mathbf{\Omega}, \mathbf{R}^{\mathsf{A}}} \ \eta \tag{5.19}$$

$$\text{s.t.} \ \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C} \omega_{k,m}[n] R^{\mathsf{A}}_{k,m,c}[n]\theta_{k,c}[n] W \geq \eta, \tag{5.19a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C} \Delta t \omega_{k,m}[n] R^{\mathsf{A}}_{k,m,c}[n]\theta_{k,c}[n] W \geq D_k^{\mathsf{min}}, \tag{5.19b}$$

$$\text{constraints (5.15d), (5.15e), (5.16a), (5.18).}$$

It can be seen that the constraints (5.19a), (5.19b), and (5.18) are still non-linear. Thus, problem (5.19) is still a non-convex optimization problem. To tackle this challenge, the successive convex optimization technique can be applied. First, let us consider the left-hand side (LHS) of (5.19a) and (5.19b) with the variables of $\omega_{k,m}[n]$ and $R^{\mathsf{A}}_{k,m,c}[n]$, and based on the first-order Taylor expansion at the given points $\omega^r_{k,m}[n]$ and $R^{\mathsf{A},r}_{k,m,c}[n]$ in the r-th iteration of the approximation process, we can obtain the following inequality

$$\omega_{k,m}[n] R^{\mathsf{A}}_{k,m,c}[n] \geq \frac{1}{4}\Bigg[-\Big(\omega^r_{k,m}[n] + R^{\mathsf{A},r}_{k,m,c}[n]\Big)^2 + 2\Big(\omega^r_{k,m}[n] + R^{\mathsf{A},r}_{k,m,c}[n]\Big)\Big(\omega_{k,m}[n] + R^{\mathsf{A}}_{k,m,c}[n]\Big) - \Big(\omega_{k,m}[n] - R^{\mathsf{A}}_{k,m,c}[n]\Big)^2\Bigg] \triangleq R^{\mathsf{Alb},r}_{k,m,c}[n]. \tag{5.20}$$

Moreover, the right-hand side (RHS) of constraints (5.18) is convex with respect to $\omega_{z,j}[n]$. Thus, by applying the first-order Taylor expansion at the given points $\omega^r_{z,j}[n]$, we can obtain the lower bound $R^{\mathsf{AA},r}_{k,m,c}[n]$ as in (5.21).

$$\log_2\left(1 + \frac{p g_{k,m}[n]}{\sum_{j\neq m}\sum_{z\neq k} \omega_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2}\right) \geq \log_2\left(1 + \frac{p g_{k,m}[n]}{\sum_{j\neq m}\sum_{z\neq k} \omega^r_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2}\right) - \sum_{j\neq m}\sum_{z\neq k} A_{z,j,k,m,c}[n]\left(\omega_{z,j}[n] - \omega^r_{z,j}[n]\right) \triangleq R^{\mathsf{AA},r}_{k,m,c}[n], \tag{5.21}$$

where

$$A_{z,j,k,m,c}[n] = \frac{\omega^r_{z,j}[n]\theta_{z,c}[n] p^2 g_{k,j}[n] g_{k,m}[n] \log_2(e)}{\left(\sum_{j\neq m}\sum_{z\neq k} \omega^r_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2\right)\left(\sum_{j\neq m}\sum_{z\neq k} \omega^r_{z,j}[n]\theta_{z,c}[n] p g_{k,j}[n] + \sigma^2 + p g_{k,m}[n]\right)}.$$

Algorithm 5.1. SCA-based Algorithm to Solve (5.19)

1: **Initialization:** Set $r := 0$, generate an initial point $\left(\boldsymbol{\Omega}^0, \mathbf{R}^{\mathsf{A},0}\right)$ of (5.22);
2: **repeat**
3: $r := r + 1$;
4: Solve (5.22) to obtain optimal values $\left(\boldsymbol{\Omega}^*, \mathbf{R}^{\mathsf{A},*}\right)$;
5: Update $\left(\boldsymbol{\Omega}^r, \mathbf{R}^{\mathsf{A},r}\right) := \left(\boldsymbol{\Omega}^*, \mathbf{R}^{\mathsf{A},*}\right)$;
6: **until** Convergence
7: **Output** $\eta_{\mathsf{a}}^*, \boldsymbol{\Omega}^*, \mathbf{R}^{\mathsf{A},*}$.

Using the approximations above, problem (5.19) can be approximated by the following problem:

$$\textbf{(P1.1")}: \max_{\eta_{\mathsf{a}}^r, \boldsymbol{\Omega}, \mathbf{R}^{\mathsf{A}}} \quad \eta_{\mathsf{a}}^r \tag{5.22}$$

$$\text{s.t.} \quad \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C} R_{k,m,c}^{\mathsf{Alb},r}[n]\theta_{k,c}[n]W \geq \eta_{\mathsf{a}}^r, \forall k, \tag{5.22a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C} \Delta t R_{k,m,c}^{\mathsf{Alb},r}[n]\theta_{k,c}[n]W \geq D_k^{\mathsf{min}}, \forall k, \tag{5.22b}$$

$$R_{k,m,c}^{\mathsf{A}}[n] \leq R_{k,m,c}^{\mathsf{AA},r}[n], \forall k, m, c, n, \tag{5.22c}$$

$$\text{constraints (5.15d)}, (5.15e), (5.16a).$$

It can be seen that all constraints are linear. Hence, problem (5.22) is a standard convex optimization problem which can be solved efficiently by any convex optimization solvers such as CVX-Mosek [38]. Detailed description of our proposed algorithm to solve the UAV-GU association problem is given in Algorithm 5.1. In the solution obtained by Algorithm 5.1, if the UAV-GU association variables $\omega_{k,m}[n]$ are all binary, then the relaxation is tight and the obtained solution is also a feasible solution of problem **(P1)**. Otherwise, the UAV-GU association solution needs to be recovered by rounding it to the nearest integer of 0 or 1. Furthermore, since constraints (5.15d) and (5.15e) are met with equalities in the solution of (5.22), a binary solution can be recovered.

5.5.2 Sub-channel Assignment Given UAV-GU Association and UAV Trajectory

For the given UAV-GU association and UAV trajectory $\{\boldsymbol{\Omega}, \mathbf{Q}\}$, we optimize the sub-channel assignment $\boldsymbol{\Theta} = \{\theta_{k,c}[n], \forall k, c, n\}$ to achieve the max-min average rate among all GUs and this problem

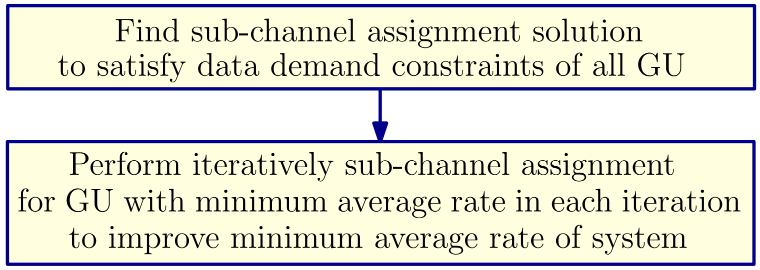

Figure 5.2 – Two main phases of proposed sub-channel assignment algorithm.

can be expressed as follows:

$$\textbf{(P1.2):} \max_{\eta, \boldsymbol{\Theta}} \ \eta \tag{5.23}$$
$$\text{s.t.} \quad \text{constraints (5.15a), (5.15b), (5.15e), (5.15f), (5.15k).}$$

Proposition 5.1. Problem **(P1.2)** is NP-hard.

Proof. The proof is given in Appendix 5.8.1. □

This integer non-convex optimization problem is difficult to solve because sub-channel assignments must be optimized over multiple UAVs, GUs, and time slots during the flight period. Hence, we propose a heuristic but efficient algorithm for sub-channel assignments. Key phases of the proposed algorithm are described in Fig. 5.2.

Recall that our design objective is to maximize the minimum average rate among all GUs and satisfy the data transmission demands of individual GUs, i.e., $D_k^{\mathsf{min}}, \forall k \in \mathcal{K}$. Hence, in the first phase, we perform sub-channel assignments for each GU to not only improve the design objective, but also ensure the constraints on data transmission demands of all GUs be satisfied. Specifically, we search a sub-channel assignment for each GU k associated with UAV m in a certain time slot n to achieve higher and maximum increase in the average rate of GU k and ensure the minimum average rate of the system is not decreasing in each assignment step.

After the required data transmission demands of all GUs are satisfied, the algorithm enters an iterative sub-channel assignment loop where in each iteration, it searches the GU with the minimum average rate and finds the best sub-channel assignment achieving the highest and better average rate for the underlying GU while improving the minimum average rate of the system. In fact, the

Algorithm 5.2. Iterative Sub-channel Assignment (ISA) Algorithm

Require: M UAVs, K GUs, C sub-channels;
1: ***Given*:** UAV-GU association, UAV trajectory control;
Ensure: Max-min average rate $(\bar{R}_k)$, η;
2: $k = 1$;
3: **while** $k \leq K$ **do**
4: **repeat**
5: Calculate the minimum average rate of the system: minrate $= \min_{k\in\mathcal{K}}\{\bar{R}_k\}$;
6: Given GU k, identify all UAV and time slot pairs $\{m, n\}$ with $\{\omega_{k,m}[n] = 1\}$;
7: Given GU k and each pair $\{m, n\}$ identified in step 6, find the sub-channel c for assignment to achieve the highest and better average rate for GU k;
8: Compare all potential sub-channel assignments for different pairs $\{m, n\}$ found in step 7, realize the best sub-channel assignment if it can improve the minimum average rate of the system, i.e., we calculate rate $= \min_{k\in\mathcal{K}}\{\bar{R}_k\}$ and the new sub-channel assignment must satisfy minrate $<$ rate;
9: **until** $\sum_n D_k[n] \geq D_k^{\mathsf{min}}$
10: $k \leftarrow k + 1$;
11: **end while**
12: **repeat**
13: Find GU $k = \mathrm{argmin}_{k\in\mathcal{K}}\{\bar{R}_k\}$;
14: Calculate the minimum average rate of the system: minrate$^* = \min_{k\in\mathcal{K}}\{\bar{R}_k\}$;
15: Given GU k, identify all UAV and time slot pairs $\{m, n\}$ with $\{\omega_{k,m}[n] = 1\}$;
16: Given GU k and each pair $\{m, n\}$ identified in step 15, find the sub-channel c for assignment to achieve the highest and better average rate for GU k;
17: Compare all potential sub-channel assignments for different pairs $\{m, n\}$ found in step 16, realize the best sub-channel assignment if it can improve the minimum average rate of the system, i.e., we calculate rate $= \min_{k\in\mathcal{K}}\{\bar{R}_k\}$ and the new sub-channel assignment must satisfy minrate$^* <$ rate;
18: Update minrate* = rate;
19: **until** Convergence
20: Update $\eta^* \leftarrow$ minrate*;
21: Return η^*, Θ^*.

method to determine the best sub-channel assignment solution in this loop is similar to that in the previous phase. The algorithm terminates when the minimum average rate of all GUs cannot be improved further.

Details of the proposed algorithm called *"Iterative Sub-channel Assignment (ISA) Algorithm"* are given in Algorithm 5.2. Let $D_k[n]$ denote the mount of data transmitted to GU k in time slot n. Then, the sum of $D_k[n]$ over different time slots of the flight period should be greater than the required data transmission demand for this GU, i.e., D_k^{min}. In the first phase of the proposed algorithm, we perform sub-channel assignments for each GU k until its required data transmission demand is satisfied. Details of this phase are described from step 4 to step 9.

Specifically, the UAV serving a GU of interest is identified based on the UAV-GU association decision variable, i.e., $\omega_{k,m}[n]$=1, which is the solution obtained from Section 5.5.1. In particular, step 6 identifies all potential pairs of UAV m and time slot n over the flight period for GU k. Then, we search the best sub-channel c for one UAV and time slot pair (m, n) among those found in step 6 and perform the corresponding sub-channel assignment, i.e., $\theta_{k,c}[n]$=1, to achieve higher average rate for GU k while ensuring the minimum average rate of the system not decreasing. These steps are presented from step 7 to step 8. The sub-channel assignment solution is identified by searching over all available sub-channels and time slots for GU k while maintaining constraints (5.15e), (5.15f) and (5.15k). After performing the best sub-channel assignment in a certain time slot for GU k, its data transmission demand constraint is verified. More sub-channel assignments can be performed until the data transmission demand constraint of each GU k is satisfied.

Then, we attempt to improve the minimum average rate of the system in the following steps. In each sub-channel assignment iteration, we find the GU k with the minimum average rate in step 13. For the identified GU k, we can find the best sub-channel assignment in a certain time slot for this user and ensure the minimum average rate of the system not decreasing similar to that in the first phase. These steps are described from step 14 to step 17. After each sub-channel assignment, the minimum average rate of the system is updated and the algorithm terminates when this value cannot be improved further.

5.5.3 UAV Trajectory Control Given UAV-GU Association and Sub-channel Assignment

Given the UAV-GU association and sub-channel assignment $\{\mathbf{\Omega}, \mathbf{\Theta}\}$, the problem optimizing the UAV trajectory control $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$ to achieve the max-min average rate over all GUs can be written as follows:

$$\begin{aligned} \textbf{(P1.3):} \max_{\eta, \mathbf{Q}} \quad & \eta \\ \text{s.t.} \quad & \text{constraints } (5.15\text{a}), (5.15\text{b}), (5.15\text{c}), (5.15\text{g}), (5.15\text{h}), (5.15\text{i}). \end{aligned} \tag{5.24}$$

This problem is a non-convex optimization problem due to the non-convex constraints (5.15a), (5.15b) and (5.15i). Therefore, it is difficult to solve this problem optimally. We design an algorithm with three main steps to solve this problem as follows. In Step 1, we introduce some auxiliary variables and transform problem (5.24) into an equivalent form. Then, we approximately convexify the corresponding problem in Step 2. Finally, we use a convex optimization solver to solve the obtained convex problem in Step 3.

5.5.3.1 Step 1 - Equivalent Transformation

We first re-write the GU's achievable rate in the difference of convex functions (DC) form (more details can be found in Appendix 5.8.2). First, we re-write $R_{k,m,c}[n]$ in constraints (5.15a) as

$$R_{k,m,c}[n] = \omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}_{k,m,c}[n]\right), \tag{5.25}$$

where

$$\hat{R}_{k,m,c}[n]=\log_2\left(\sum_{j=1}^{M}\sum_{z=1}^{K}R^{\mathsf{Ab}}_{z,j,c,k,m}[n]+\sigma^2\right), \tag{5.26}$$

$$\tilde{R}_{k,m,c}[n]=\log_2\left(\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}R^{\mathsf{Ab}}_{z,j,c,k,m}[n]+\sigma^2\right), \tag{5.27}$$

in which

$$R^{\mathsf{Ab}}_{z,j,c,k,m}[n] = \frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+\|\mathbf{q}_j[n]-\mathbf{r}^{\mathsf{u}}_k\|^2}. \tag{5.28}$$

We now introduce auxiliary variables $\mathbf{S} = \left\{S_{k,j}[n] \leq \|\mathbf{q}_j[n]-\mathbf{r}^{\mathsf{u}}_k\|^2, \forall j,k,n\right\}$. Applying this to (5.28), we have

$$R^{\mathsf{Ab}}_{z,j,c,k,m}[n] \leq \frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+S_{k,j}[n]}. \tag{5.29}$$

Then, the problem (5.24) can be reformulated as

$$\textbf{(P1.3')}: \max_{\eta,\mathbf{Q},\mathbf{S},\mathbf{R}} \quad \eta \tag{5.30}$$

$$\text{s.t. } \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\omega_{k,m}[n]\theta_{k,c}[n]W\Big(\hat{R}_{k,m,c}[n]-\tilde{R}_{k,m,c}[n]\Big) \geq \eta,\ \forall k, \tag{5.30a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\Delta t\omega_{k,m}[n]\theta_{k,c}[n]W\Big(\hat{R}_{k,m,c}[n]-\tilde{R}_{k,m,c}[n]\Big) \geq D_k^{\mathsf{min}},\ \forall k, \tag{5.30b}$$

$$R^{\mathsf{Ab}}_{z,j,c,k,m}[n] \leq \frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+S_{k,j}[n]}, \tag{5.30c}$$

$$S_{k,m}[n] \leq \|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\|^2, \forall k,m,n, \tag{5.30d}$$

$$\|r_0-\mathbf{q}_m[n]\| \leq R_0,\ \forall m,n, \tag{5.30e}$$

$$\mathbf{q}_m[1]=\mathbf{q}_m[N], \forall m, \tag{5.30f}$$

$$\|\mathbf{q}_m[n+1]-\mathbf{q}_m[n]\|^2 \leq S^2_{\mathsf{max}}, n=1,...,N-1, \tag{5.30g}$$

$$\|\mathbf{q}_m[n]-\mathbf{q}_j[n]\|^2 \geq d^2_{\mathsf{min}},\ \forall n,m,j\neq m, \tag{5.30h}$$

where $\mathbf{R}=\{R^{\mathsf{Ab}}_{z,j,c,k,m}[n],\forall k,m,z,j,c,n\}$.

It can be verified that $\hat{R}_{k,m,c}[n]$ and $\tilde{R}_{k,m,c}[n]$ are concave with respect to $R^{\mathsf{Ab}}_{z,j,c,k,m}[n]$. Moreover, constraints (5.30a) and (5.30b) are in the DC form. However, the constraints in (5.30c), (5.30d), and (5.30h) are still non-convex so that problem (5.30) is still a non-convex optimization problem.

5.5.3.2 Step 2 - Convex Approximation

We can handle the non-convex constraints (5.30a) and (5.30b) using the successive convex approximation (SCA) technique. The non-convex constraint functions are approximated by convex functions and the resulting optimization problem is solved iteratively. Specifically, we define $\mathbf{Q}^r=\{\mathbf{q}^r_m[n],\forall m,n\}$ to represent the trajectories of UAVs and $S^r_{k,j}[n]=\left\|\mathbf{q}^r_j[n]-\mathbf{r}^{\mathsf{u}}_k\right\|^2$ to denote the distance between UAVs and GUs in the r-th iteration of this approximation. In addition, let $\mathbf{S}^r,\mathbf{R}^r$ be the achieved feasible variables in the r-th iteration.

We now describe how to convexify this problem. We first consider $\tilde{R}_{k,m,c}[n]$ a concave function with respect to $R^{\mathsf{Ab}}_{z,j,c,k,m}[n]$. Recall that any concave function is upper-bounded by its first-order

Taylor expansion at any point. Hence, we have

$$\tilde{R}_{k,m,c}[n] \leq \tilde{R}^{\mathsf{ub}}_{k,m,c}[n], \tag{5.31}$$

where $\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]$ is described in more details in Appendix 5.8.2.

In addition, constraints (5.30c) can be equivalently written as

$$R^{\mathsf{Ab}}_{z,j,c,k,m}[n]S_{k,j}[n] \leq \omega_{z,j}[n]\theta_{z,c}[n]p\rho_0 - R^{\mathsf{Ab}}_{z,j,c,k,m}[n]H^2.$$

We can express the left-hand side (LHS) of this constraint in the DC form [175]. Hence, based on the first-order Taylor expansion at the given points $R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]$ and $S^r_{k,j}[n]$ in the r-th iteration of the approximation process, it can be approximated as

$$R^{\mathsf{Ab}}_{z,j,c,k,m}[n]S_{k,j}[n] \leq R^{\mathsf{App},r}[n], \tag{5.32}$$

where $R^{\mathsf{App},r}[n]$ is given in Appendix 5.8.2. Thus, the constraints (5.30c) can be approximated as

$$\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0 - R^{\mathsf{Ab}}_{z,j,c,k,m}[n]H^2 \geq R^{\mathsf{App},r}[n]. \tag{5.33}$$

Moreover, since $\|\mathbf{q}_m[n] - \mathbf{r}^{\mathsf{u}}_k\|^2$ in the constraints (5.30d) is a convex function with respect to $\mathbf{q}_m[n]$, we have the following inequality by applying the first-order Taylor expansion at the given point $\mathbf{q}^r_m[n]$:

$$\|\mathbf{q}_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2 \geq \|\mathbf{q}^r_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{r}^{\mathsf{u}}_k\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}^r_m[n]\right). \tag{5.34}$$

Furthermore, by applying the first-order Taylor expansion at the given point $\mathbf{q}^r_m[n]$ and $\mathbf{q}^r_j[n]$ to $\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2$, the LHS of constraint (5.30h) can be approximated by its lower bound as

$$\|\mathbf{q}_m[n]-\mathbf{q}_j[n]\|^2 \geq -\left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 + 2\left(\mathbf{q}^r_m[n] - \mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right), \forall j\neq m, n. \tag{5.35}$$

Algorithm 5.3. SCA-based Algorithm to Solve (5.24)

1: **Initialization:** Set $r := 0$, generate an initial point $\left(\mathbf{Q}^0, \mathbf{S}^0, \mathbf{R}^0\right)$ of (5.36);
2: **repeat**
3: $r := r + 1$;
4: Solve (5.36) to obtain optimal values $(\mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*)$;
5: Update $(\mathbf{Q}^r, \mathbf{S}^r, \mathbf{R}^r) := (\mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*)$;
6: **until** Convergence
7: **Output** $\eta^*_{\mathsf{trj}}, \mathbf{Q}^*, \mathbf{S}^*, \mathbf{R}^*$.

5.5.3.3 Step 3 - Solving Approximated Convex Problem

Using the approximations above, problem (5.30) can be approximated as

$$\textbf{(P1.3”):} \max_{\eta^r_{\mathsf{trj}}, \mathbf{Q}, \mathbf{S}, \mathbf{R}} \quad \eta^r_{\mathsf{trj}} \tag{5.36}$$

$$\text{s.t. } \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]\right) \geq \eta^r_{\mathsf{trj}},\ \forall k, \tag{5.36a}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{c=1}^{C}\Delta t\omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}^{\mathsf{ub}}_{k,m,c}[n]\right) \geq D^{\mathsf{min}}_k,\ \forall k, \tag{5.36b}$$

$$S_{k,m}[n] \leq \|\mathbf{q}^r_m[n] - \mathbf{r}^{\mathsf{u}}_k\|^2 + 2\left(\mathbf{q}^r_m[n] - \mathbf{r}^{\mathsf{u}}_k\right)^T\left(\mathbf{q}_m[n] - \mathbf{q}^r_m[n]\right), \forall k, m, n, \tag{5.36c}$$

$$d^2_{\mathsf{min}} \leq -\left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right), \forall j \neq m, n, \tag{5.36d}$$

$$(5.30e) - (5.30g), (5.33).$$

In this problem, constraints (5.30g) are convex while all remaining constraints are linear. Hence, problem (5.36) is a standard convex optimization problem which can be solved efficiently by any convex optimization solvers such as CVX-Mosek [38]. Detailed description of our proposed algorithm to solve the UAV trajectory control optimization problem is given in Algorithm 5.3.

5.5.4 Integrated UAV-GU Association, Sub-channel Assignment and UAV Trajectory Control

Using the results presented in Sections 5.5.1, 5.5.2 and 5.5.3, our proposed algorithm based on the alternating optimization method is described in Algorithm 5.4. The convergence of this algorithm is stated in the following proposition.

Algorithm 5.4. Integrated UAV-GU Association, Sub-channel Assignment and UAV Trajectory Control

Require: M UAVs, K GUs, C sub-channels and T;
Ensure: Max-min average rate $(\bar{R}_k)$, η; Let $r = 1$;
1: **repeat**
2: Optimize the UAV-GU association given the sub-channel assignment and UAVs' trajectories by solving sub-problem using Algorithm 5.1 to obtain $\mathbf{\Omega}^r$;
3: Optimize the sub-channel assignment given the UAV-GU association and UAVs' trajectories by solving sub-problem using Algorithm 5.2 to obtain $\mathbf{\Theta}^r$;
4: Optimize the UAVs' trajectories given the UAV-GU association and sub-channel assignment using Algorithm 5.3 to obtain $\mathbf{Q}^r$;
5: Update $r = r + 1$;
6: **until** Convergence
7: Return $\eta^*, \mathbf{\Omega}^*, \mathbf{\Theta}^*, \mathbf{Q}^*$.;

Proposition 5.2. The proposed Algorithm 5.4 creates a sequence of feasible solutions where the objective value monotonically increases over iterations. As a result, the algorithm converges to a feasible solution.

Proof. The proof is given in Appendix 5.8.3. □

5.5.5 Complexity Analysis

We now analyze the complexity of our proposed algorithm evaluated in the number of required arithmetic operations. For the UAV-GU association sub-problem, since CVX [38] invokes the interior-point method to solve the underlying problem (5.22), the involved complexity is $\mathcal{O}(m_1^{1/2}(m_1 + m_2)m_2^2)$, where m_1 is the number of inequality constraints, m_2 denotes the number of variables [195], and $\mathcal{O}$ denotes the big-O notation. Hence, the complexity of this step is $\mathcal{O}(L_1 N K^{\frac{7}{2}})$, where L_1 is the number of iterations required to achieve the convergence of Algorithm 5.1.

We now analyze the complexity of the proposed iterative sub-channel assignment (ISA) algorithm to solve the sub-channel assignment sub-problem i.e., Algorithm 5.2. In the first phase of this algorithm, we find the sub-channel assignment to satisfy the demand constraints for all GUs. For each GU, the required complexity is dominated by operations in steps 6 to 8, which investigate all sets of GUs, UAVs, sub-channels and time slots to identify the best assignment. The computational complexity is $\mathcal{O}(I_1(KMNC))$, where I_1 denotes the average number of iterations needed to ensure that the data transmission demand of each GU is satisfied. Hence, the complexity of this phase is $\mathcal{O}(I_1(K^2MNC))$. In the second phase, step 13 finds the UAV with the minimum average rate and its complexity is $\mathcal{O}(KMNC)$. In addition, step 14 to step 17 are similar to those in the first phase where the involved complexity is $\mathcal{O}(KMNC)$. Thus, the second loop has the computation complexity of $\mathcal{O}(2I_2(KMNC))$, where I_2 denotes the number of iterations needed to achieve convergence. Hence, the complexity of Algorithm 5.2 is $\mathcal{O}(I_1(K^2MNC) + 2I_2(KMNC))$.

We now analyze the complexity of the SCA-based algorithm 5.3 to solve the UAV trajectory control sub-problem. In step 4, we use the CVX solver to solve the convex optimization problem (5.36) by using the interior-point method with the complexity of $\mathcal{O}(L_2NK^{\frac{7}{2}})$, where L_2 is the number of iterations required to achieve the convergence of Algorithm 5.3.

Let L denote the number of iterations needed in the outer loop to reach convergence for Algorithm 5.4. Then, the overall computation complexity of the proposed algorithm is $\mathcal{O}\Big(L(I_1(K^2MNC) + 2I_2(KMNC))\Big)$.

5.5.6 Heuristic Sub-channel Assignment

For benchmarking purposes, a sub-channel assignment with interference management (SAIM) algorithm is presented whose key steps are illustrated in Fig. 5.3. The performance comparison between the SAIM and ISA algorithms are be discussed in Section 5.6.

In the SAIM algorithm, we sequentially assign sub-channels to each GU until its data transmission demand is satisfied. As a result, the constraint on the data transmission demand of each GU is tracked during the assignments to get a feasible solution. The reason for this approach is to minimize utilizing a sub-channel for different UAVs to mitigate the co-channel interference. In particular, the set of all sub-channels, sets of sub-channels used by UAVs m and m' in time slot n, which are denoted as $\mathcal{C}, \mathcal{C}_m[n], \mathcal{C}_{m'}[n]$, respectively, are investigated to perform each sub-channel assignment. If all sub-channels are assigned to GUs and certain data transmission demand constraints are still not satisfied, we reuse and allocate a sub-channel already assigned for certain GU k' served by UAV $m' \neq m$ with maximum distance $d_{k,k'}$ from GU k of interest to mitigate the interference.

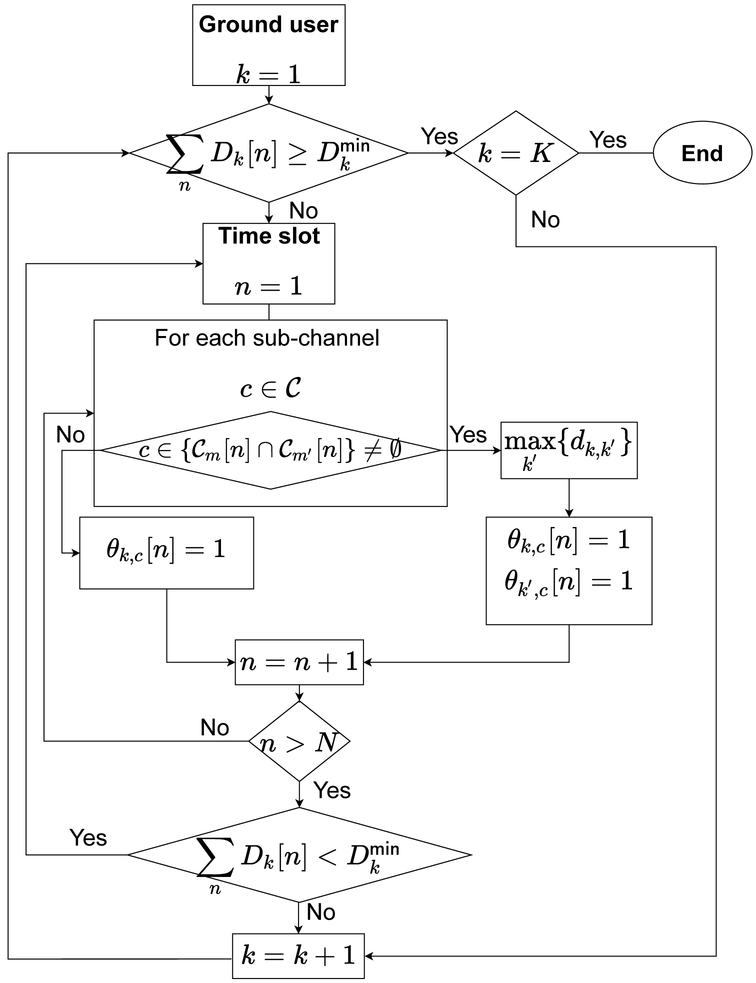

Figure 5.3 – Sub-channel assignment with interference management (SAIM).

5.6 Numerical Results

5.6.1 Simulation Setting

In this section, we evaluate the performance of the proposed algorithm. The parameter setting for the simulations is similar to that in [39–41] and are summarized in Table 5.3. We consider a circular network area with radius $R_0 = 500$m with two or more clusters i.e., hotspots, of GUs. The radius of each circular cluster area is $r_{\mathsf{c}} = 200$m and different clusters are placed far enough apart not to overlap. The distance between two neighboring clusters' centers is set to satisfy the constraint $D^0 \geq d_{\min} + 2 \times r_{\mathsf{c}}$(m). The altitude of all UAVs is assumed to be fixed at $H = 100$m. Moreover, the required transmission data demand for each GU k ($D_k^{\min}$) is set according to the size of short videos e.g., video files with the resolution of 30 frames per second (fps) [42]).

Table 5.3 – Simulation parameters

Parameter	Description	Value
M	Number of UAVs	$[2, 3]$
K	Number of GUs	$[8, 10, 12, 14, 16]$
C	Number of sub-channels	$[20, 30, 40, 50, 60]$
W	Bandwidth of each sub-channel	1 MHz
T	Flight period	$[20, 40]$ s
Δt	Length of each time slot	0.5 s
H	Altitude of UAVs	100 m
r_c	Radius of cluster	200 m
P_{max}	Transmit power of UAV	30 dBm
σ^2	Noise power	-110 dBm
f_c	Carrier frequency	2.5 GHz
d_{min}	Minimum inter-UAV distance	20 m
$V_{\max}$	Maximum speed of UAV	$[10-80]$ m/s

5.6.2 Algorithm Initialization

To execute our algorithm, we need to initialize the values of all variables. Here, we present a simple method to initialize the variables related to UAV-GU association and UAV trajectory control.

5.6.2.1 Initial UAV-GU Association

We assume that each GU k is initially associated with the UAV providing the highest average received signal strength (RSS) as follows:

$$\omega_{k,m}[n] = \begin{cases} 1 & m = \arg\max_{k} RSS_{k,m}[n], \\ 0 & \text{otherwise.} \end{cases} \tag{5.37}$$

where the average RSS for the GU k and UAV m can be expressed as

$$RSS_{k,m}[n](\text{dBm}) = p(\text{dBm}) - g_{k,m}[n](\text{dBm}), \tag{5.38}$$

where p is the transmit power of UAV m used in its communication with GU k in time slot n on each assigned sub-channel.

5.6.2.2 Initial Circular UAV Trajectory

To set the initial circular UAV trajectory, we assume that each UAV serves a circular network partition i.e., serving a cluster of GUs, with radius r_{c} (m). For each network partition, we need to determine its center $\mathbf{c}_m^{\mathsf{init}} = (x_m^{\mathsf{init}}, y_m^{\mathsf{init}})$ and radius d_m^{init} for UAV m. These points are the centers of the corresponding network partitions computed using the k-means clustering algorithm [196] for given locations of the K GUs. The initial circular radius for UAV m is given by $d_m^{\mathsf{init}} = \min\left(d_m^{\mathsf{max1}}, d_m^{\mathsf{max2}}, r_{\mathsf{c}}\right)$, where $d_m^{\mathsf{max1}} = R_0 - \|r_0 - \mathbf{c}_m^{\mathsf{init}}\|$ denotes the maximum radius of circular trajectory of UAV m. The UAV remains inside the desired area during the flight period and d_m^{max2} is the maximum radius of the circular UAV trajectory with the same starting and ending points.

To determine d_m^{max2}, we approximate the largest circumference of a circle as the maximum distance, denoted as $D = V_{\mathsf{max}}T$, that the UAV can travel during the flight period. Therefore, we have $d_m^{\mathsf{max2}} \approx D/2\pi$. Let $\phi_n \triangleq 2\pi\frac{n-1}{N-1}, \forall n$, we can initialize $\mathbf{Q}^0 = \{\mathbf{q}_m^0[n], \forall m, n\}$ as follows:

$$\mathbf{q}_m^0[n] = \left(x_m^{\mathsf{init}} + d_m^{\mathsf{init}}\cos\phi_n, y_m^{\mathsf{init}} + d_m^{\mathsf{init}}\sin\phi_n\right), \ \forall m, n. \tag{5.39}$$

For the multi-UAV system, d_{min} denotes the minimum inter-UAV distance to ensure collision avoidance, which must be maintained as we initialize the UAVs' trajectories. Fig. 5.4 illustrates the initial circular trajectories for the network with 3 UAVs.

5.6.3 Numerical Results

5.6.3.1 UAV-GU Association

We first compare the performance achieved by the RSS based method and optimized solution from (5.16) for UAV-GU association with 2 UAVs, 40 sub-channels, initial circular UAV trajectories, T = 20s, and the minimum required transmission data D_k^{min} = 5MB for each GU. The max-min average rates achieved by the RSS based method and optimized solution by using Algorithm 5.1 are presented in TABLE 5.4 for the UAV flight period T = 20s. In addition, the SAIM algorithm for sub-channel assignment and the UAVs' circular trajectories are used to evaluate the system performance for different numbers of GUs. It can be seen that the optimized algorithm outperforms the RSS based method for UAV-GU association.

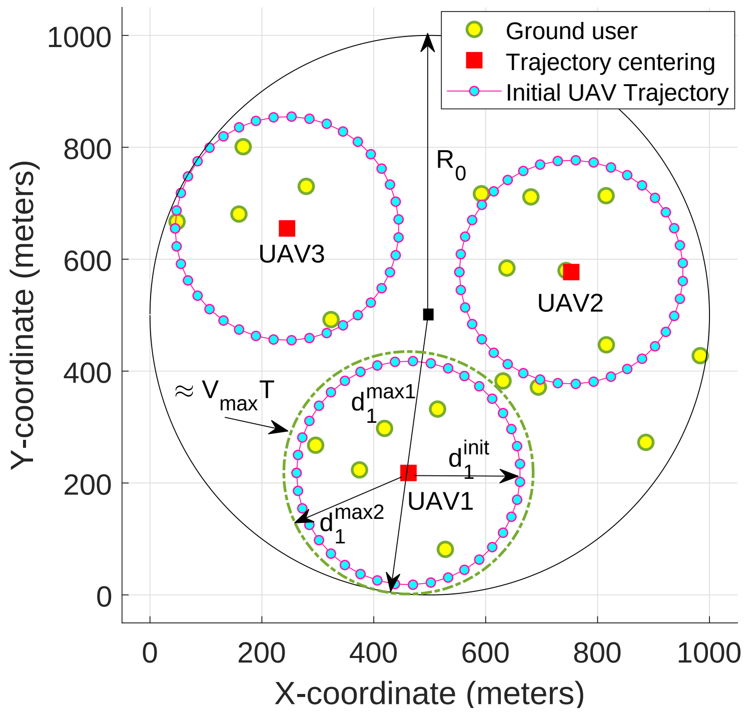

Figure 5.4 – Initial UAVs' trajectories.

Table 5.4 – Comparison of RSS method and optimized solution for UAV-GU association.

Max-min rate comparison (Mbps)	Number of GUs				
	8	**10**	**12**	**14**	**16**
RSS method	10.5821	9.1238	7.6425	6.2541	5.1228
Optimized solution using Algorithm 5.1	11.1149	9.8503	8.5216	7.4138	6.4206

5.6.3.2 Sub-channel Assignment

We now evaluate the performance of the proposed ISA algorithm described in Alg. 5.2 and the SAIM algorithm illustrated in Fig. 5.3. Specifically, the max-min average rates due to different schemes are shown in Fig. 5.5 for the network with 2 UAVs, 40 sub-channels, UAV's flight period $T = 20$s, and the maximum velocity of UAVs $V_{\mathsf{max}} = 40$ (m/s). We see that the proposed design with ISA algorithm, optimized UAV-GU association and trajectory control achieves the highest max-min average rate among the considered schemes. In addition, the rate gaps between the proposed ISA algorithm and other schemes increase when the number of GUs increases. For a given number of sub-channels, more sub-channels are likely to be reused by different UAVs to meet the GUs' data transmission demands when the number of GUs increases and this will likely lead to stronger co-channel interference. The results in Fig. 5.5 imply that the proposed ISA algorithm can effectively manage interference and resources.

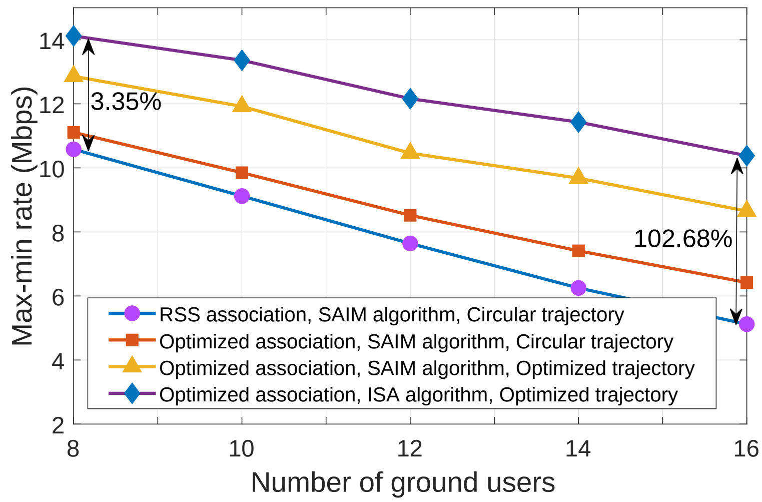

Figure 5.5 – Performance comparison of different schemes with 2 UAVs and $T = 20s$.

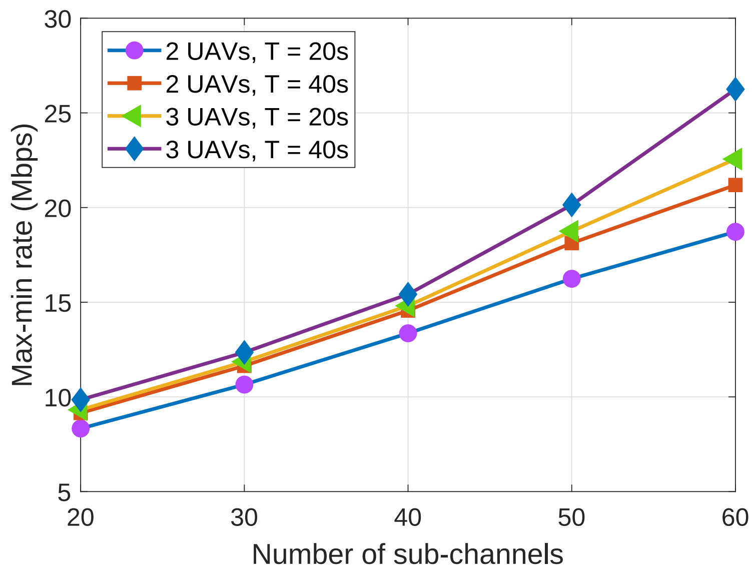

Figure 5.6 – Max-min rate under different number of sub-channels.

Fig. 5.6 illustrates the max-min average rates achieved by the proposed algorithm with different number of sub-channels, for the 2-UAV and 3-UAV scenarios, 10 GUs, flight period $T = [20, 40]$s, and the maximum velocity of UAVs $V_{\mathsf{max}} = 40$ (m/s). We see that the max-min average rate increases almost linearly as the number of sub-channels increases in the scenario with 2 UAVs and the UAV's flight period $T = 20$s. Furthermore, the max-min average rate for the scenario with 3 UAVs is higher than that with 2 UAVs where the rate difference becomes larger when the number of sub-channels increases.

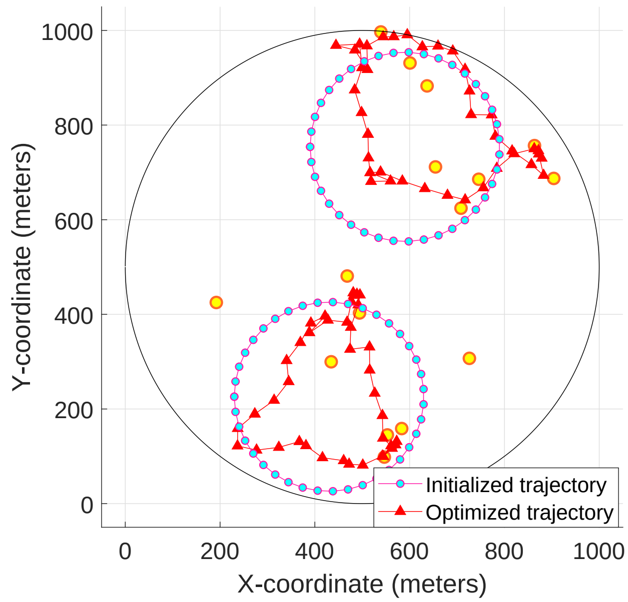

Figure 5.7 – Optimized trajectories for 2 UAVs, 16 GUs, and $T = 20$s.

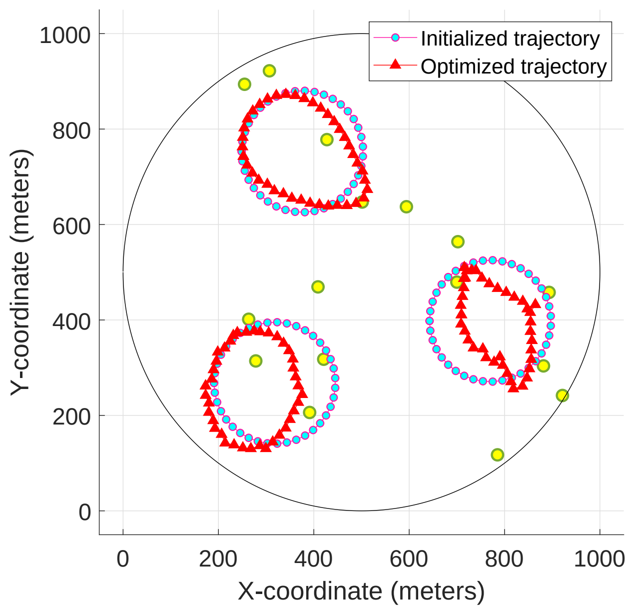

Figure 5.8 – Optimized trajectories for 3 UAVs, 16 GUs, and $T = 20$s.

5.6.3.3 UAV Trajectory

Figs. 5.7 and 5.8 show the UAV trajectories for the scenarios with 2 UAVs and 3 UAVs, 16 GUs, and UAV flight period $T = 20$s. For the setting with 2 UAVs and 16 GUs, it can be seen that the optimized trajectories are not smooth and the UAVs fly close to the corresponding clusters of

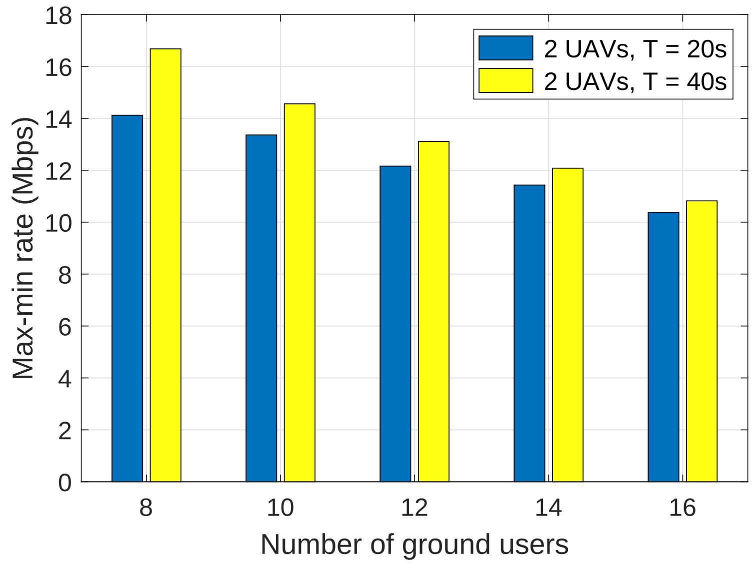

Figure 5.9 – Max-min rates with 2 UAVs.

GUs they are serving. Similar observations can be drawn for the scenario with 3 UAVs and 16 GUs where UAVs tend to serve the GUs closer to them.

Fig. 5.9 shows the max-min average rate for the 2-UAV scenario under different number of GUs and UAV flight period $T = [20, 40]$s, maximum velocity of UAVs $V_{\mathsf{max}} = 40$ (m/s), and 40 sub-channels. The results show that the max-min average rate increases when the UAV's flight period becomes longer. In addition, the rate difference between scenarios with the flight period $T = 20$s and $T = 40$s decreases when the number of GUs increases. This is because larger number of GUs improves the radio utilization efficiency.

We show the max-min average rates achieved for scenarios with 2 UAVs and 3 UAVs under different number of GUs in Fig. 5.10 with parameters $T = 20$s, $V_{\mathsf{max}} = 40$ (m/s) and 40 sub-channels. This figure shows that the max-min average rate achieved by 3 UAVs is larger than that with 2 UAVs. This is because the UAV-GU distance is typically smaller when a larger number of UAVs serves a particular number of GUs. However, the rate gap between the two scenarios decrease when the number of GUs increases due to stronger interference with the larger number of UAVs.

We study the impact of the maximum UAV's velocity V_{max} on the max-min average rate in Fig. 5.11 for scenarios with 2 UAVs and 3 UAVs, 10 GUs, and 40 sub-channels where V_{max} varies in the range of $10 - 80$ (m/s). It can be seen that the peaks of the max-min average rate are achieved at the maximum UAV's velocity of 40 (m/s) and 30 (m/s) for $T = [20, 40]$s, respectively. Moreover, the

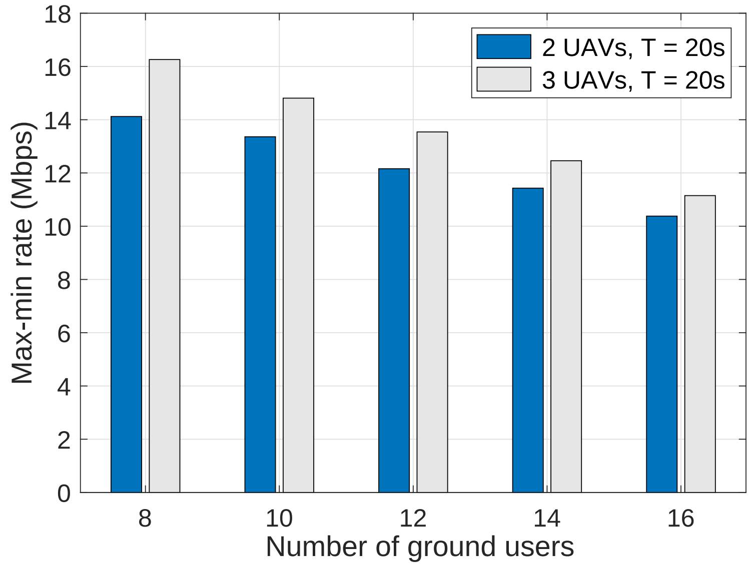

Figure 5.10 – Max-min rates with 2 UAVs and 3 UAVs.

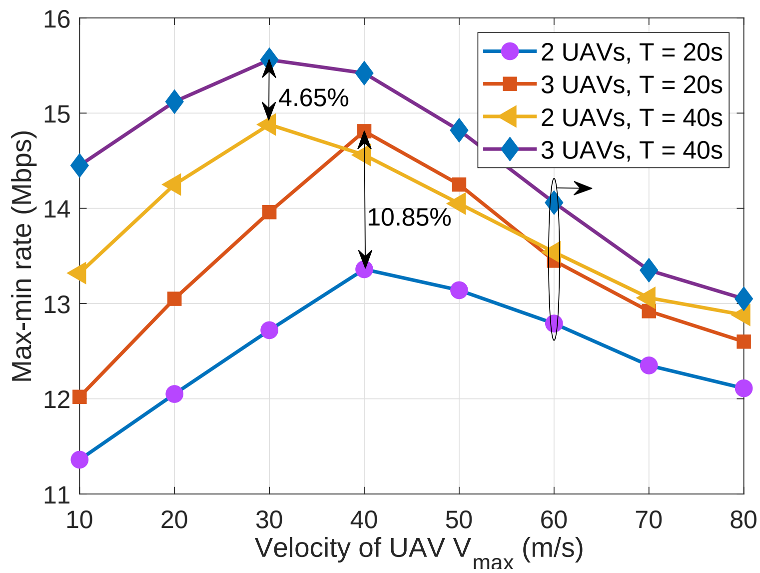

Figure 5.11 – Max-min rate under different velocity of UAV V_{max}.

rate gains at the peak rates for the 3-UAV setting versus the 2-UAV setting are 4.65% and 10.85% for $V_{\mathsf{max}} = [30, 40]$ (m/s) and $T = [20, 40]$s, respectively. However, this rate gain tends to decrease with the higher maximum velocity of UAVs. In fact, with the restricted network area of radius, $d_m^{\mathsf{init}}, (\forall m)$ given in Eq. (5.39), the velocity of UAVs strongly impacts the initial and the optimized trajectories of UAVs. This is because when the UAVs fly faster, the inter-UAV distances can become smaller in larger portions of the flight and the co-channel interference would be stronger, especially with a large number of UAVs. Specifically, the max-min average rate with $V_{\mathsf{max}} \geq 60$ (m/s) in the 3-UAV deployment and $T = 20$s is smaller than that in the 2-UAV scenario with $T = 40$s.

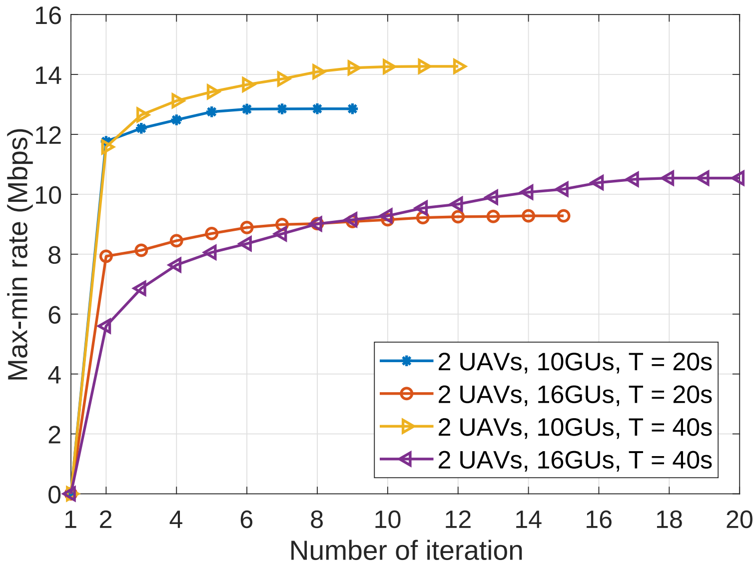

Figure 5.12 – Convergence of the proposed algorithm.

The convergence of the Alg. 5.4 is illustrated in Fig. 5.12 for the network setting with 2 UAVs, 10 and 16 GUs, $T = [20, 40]$s, $V_{\mathsf{max}} = 40$ (m/s) and 40 sub-channels. In particular, we optimize UAV-GU association, sub-channel assignment, and UAV trajectory control until convergence in each iteration of the proposed algorithm. It can be seen that the value of the objective function improves over iterations. Our algorithm converges more slowly to the feasible solution with the larger flight period T. This is because the optimization space becomes larger with larger flight period T. Moreover, the number of iterations increases when the number of GUs increases due to increasing complexity for algorithm.

5.7 Conclusion

In this chapter, we studied the joint UAV-GU association, sub-channel assignment, and UAV trajectory control problem to achieve fair resource sharing among GUs considering their data transmission demands and spectrum reuse. To solve the underlying MINLP problem, we used the alternating optimization approach and developed an efficient integrated algorithm. In particular, the iterative sub-channel assignment (ISA) algorithm was proposed to solve the sub-channel assignment sub-problem.

We used successive convex approximation to solve the UAV trajectory control sub-problem. Numerical results have demonstrated the effectiveness of the proposed algorithm. Specifically, the optimized UAV trajectory can result in a non-negligible rate gain compared to the case where UAVs' trajectories are set to be circular around the corresponding clusters of GUs. Moreover, we showed that the number of deployed UAVs, number of sub-channels, and UAV's maximum velocity have strong impacts on the achieved max-min average rate.

5.8 Appendices

5.8.1 Proof of Proposition 5.1

In this appendix, we prove that the considered optimization problem is NP-hard [197]. We construct an instance of problem **(P1.2)** where sub-channels assigned to GUs in different time slots. Let $\mathcal{K}, \mathcal{C}$, and $\mathcal{N}$ be three disjoint sets of GUs, sub-channels, and time slots, respectively. Sets $\mathcal{K}, \mathcal{C}$, and $\mathcal{N}$ satisfy $\mathcal{K} \cap \mathcal{C} = \varnothing$, $\mathcal{K} \cap \mathcal{N} = \varnothing$, and $\mathcal{C} \cap \mathcal{N} = \varnothing$. Let $\mathcal{Q}$ be a collection of ordered triples $\mathcal{Q} \subseteq \mathcal{K} \times \mathcal{C} \times \mathcal{N}$, where each element in $\mathcal{Q}$ corresponds to a sub-channel with the corresponding GU in a particular time slot, i.e., $\mathcal{Q}_i = (\mathcal{K}_i, \mathcal{C}_i, \mathcal{N}_i) \in \mathcal{Q}$. For convenience, we set $L = \min\{|\mathcal{K}|, |\mathcal{C}|, |\mathcal{N}|\}$. There exists $\mathcal{Q}' \subseteq \mathcal{Q}$ that the following holds: (1) $|\mathcal{Q}'| = |L| =$ L; (2) for any two distinct triples $(\mathcal{K}_i, \mathcal{C}_i, \mathcal{N}_i) \in \mathcal{Q}'$ and $(\mathcal{K}_j, \mathcal{C}_j, \mathcal{N}_j) \in \mathcal{Q}'$, we have $i \neq j$.

Hence, $\mathcal{Q}'$ is a three dimension matching (3-DM). Since the 3-DM problem has been proved to be NP-complete in [198], the constructed instance of problem is also NP-complete. Therefore, the problem **(P1.2)** is NP-hard.

5.8.2 Detailed Description of The Development in Section 5.5.3

In this appendix, we provide details of the developments presented in Section 5.5.3.

- First, $R_{k,m,c}[n]$, in constraints (5.15a), can be re-written as

$$\begin{aligned}
R_{k,m,c}[n] &= \omega_{k,m}[n]\theta_{k,c}[n]W\log_2\left(1+\frac{pg_{k,m}[n]}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\omega_{z,j}[n]\theta_{z,c}[n]pg_{k,j}[n]+\sigma^2}\right)\\
&= \omega_{k,m}[n]\theta_{k,c}[n]W\log_2\left(1+\frac{\frac{p\rho_0}{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+\left\|\mathbf{q}_j[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}+\sigma^2}\right)\\
&= \omega_{k,m}[n]\theta_{k,c}[n]W\log_2\left(\frac{\sum_{j=1}^{M}\sum_{z=1}^{K}\frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+\left\|\mathbf{q}_j[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}+\sigma^2}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+\left\|\mathbf{q}_j[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}+\sigma^2}\right)\\
&= \omega_{k,m}[n]\theta_{k,c}[n]W\Bigg[\log_2\left(\sum_{j=1}^{M}\sum_{z=1}^{K}\frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+\left\|\mathbf{q}_j[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}+\sigma^2\right)-\\
&\qquad \log_2\left(\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\frac{\omega_{z,j}[n]\theta_{z,c}[n]p\rho_0}{H^2+\left\|\mathbf{q}_j[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}+\sigma^2\right)\Bigg]\\
&= \omega_{k,m}[n]\theta_{k,c}[n]W\left(\hat{R}_{k,m,c}[n]-\tilde{R}_{k,m,c}[n]\right). \qquad (5.40)
\end{aligned}$$

- Second, the upper bound of $\tilde{R}_{k,m,c}[n]$ can be expressed as

$$\begin{aligned}
\tilde{R}^{\mathsf{ub}}_{k,m,c}[n] &= \log_2\left(\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]+\sigma^2\right)+\\
&\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}\frac{\log_2(e)}{\sum_{j=1,j\neq m}^{M}\sum_{z=1,z\neq k}^{K}R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]+\sigma^2}\left(R^{\mathsf{Ab}}_{z,j,c,k,m}[n]-R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]\right). \qquad (5.41)
\end{aligned}$$

- Third, based on the first-order Taylor expansion at the given points $R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]$ and $S^r_{k,j}[n]$ in the r-th iteration of the approximation process, it can be approximated as

$$\begin{aligned}
R^{\mathsf{Ab}}_{z,j,c,k,m}[n]S_{k,j}[n] \leq \frac{1}{4}\Bigg[&\left(R^{\mathsf{Ab}}_{z,j,c,k,m}[n]+S_{k,j}[n]\right)^2-2\left(R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]-S^r_{k,j}[n]\right)\left(R^{\mathsf{Ab}}_{z,j,c,k,m}[n]-S_{k,j}[n]\right)\\
&+\left(R^{\mathsf{Ab},r}_{z,j,c,k,m}[n]-S^r_{k,j}[n]\right)^2\Bigg] \triangleq R^{\mathsf{App},r}[n]. \qquad (5.42)
\end{aligned}$$

5.8.3 Proof of Proposition 5.2

In this appendix, we prove that the Algorithm 5.4 creates a non-decreasing sequence of objective values of problem **(P1)** and converges to a feasible solution. First, it can be verified after the initialization step or after each r-iteration of the approximation process, we achieve a feasible solution of $\mathbf{\Omega}^r$, $\mathbf{\Theta}^r$, and $\mathbf{Q}^r$. For step 2 of Algorithm 5.4, since the optimal solution of **(P1.1)** is obtained for given $\mathbf{\Theta}^r$ and $\mathbf{Q}^r$, we have

$$\eta(\mathbf{\Omega}^r, \mathbf{\Theta}^r, \mathbf{Q}^r) \leq \eta(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^r, \mathbf{Q}^r), \tag{5.43}$$

where $\eta(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q})$ is defined in the formulation for problem (5.15). Moreover, for given $\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^r$, and $\mathbf{Q}^r$ obtained in step 3 of Algorithm 5.4, it follows that

$$\eta(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^r, \mathbf{Q}^r) \leq \eta(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^{r+1}, \mathbf{Q}^r). \tag{5.44}$$

This is because problem **(P1.2)** is solved under the iterative sub-channel assignment (ISA) algorithm presented in Section 5.5.2. Finally, for given $\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^{r+1}$, and $\mathbf{Q}^r$ obtained in step 4 of Algorithm 5.4, we have

$$\begin{aligned} \eta(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^{r+1}, \mathbf{Q}^r) &\overset{(a)}{\leq} \eta_{\mathsf{trj}}^{\mathsf{lb,r}}(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^{r+1}, \mathbf{Q}^{r+1}) \\ &\overset{(b)}{\leq} \eta(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^{r+1}, \mathbf{Q}^{r+1}), \end{aligned} \tag{5.45}$$

where we define $\eta_{\mathsf{trj}}^{\mathsf{lb,r}}(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}) = \eta_{\mathsf{trj}}^r$ as the objective value of problem (5.36). And (a) holds since Algorithm 5.4 achieves the solution $\mathbf{Q}^{r+1}$ for given $\mathbf{\Omega}^{r+1}$ and $\mathbf{\Theta}^{r+1}$; (b) holds since the objective value of problem (5.36) is the lower bound of the objective of its original problem (5.24) at $\mathbf{Q}^{r+1}$ because the SCA method is applied. Using the results in (5.43)-(5.45), we obtain

$$\eta(\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{Q}) \leq \eta(\mathbf{\Omega}^{r+1}, \mathbf{\Theta}^{r+1}, \mathbf{Q}^{r+1}), \tag{5.46}$$

which indicates that the objective value of problem **(P1)** is non-decreasing after each iteration of Algorithm 5.4. Since the objective value of problem **(P1)** is upper bounded by a finite value, Algorithm 5.4 is guaranteed to converge to a feasible solution. This completes the proof.

Chapter 6

UAV Placement and Resource Allocation for Intelligent Reflecting Surface Assisted UAV-Based Wireless Networks

The content of this chapter was published in the following paper[1]:

Minh Dat Nguyen, Long Bao Le, and André Girard , "UAV Placement and Resource Allocation for Intelligent Reflecting Surface Assisted UAV-Based Wireless Networks," *IEEE Communications Letters*, vol. 26, no. 5, pp. 1106–1110, May 2022.

6.1 Abstract

We design an unmanned aerial vehicle (UAV) based wireless network with wireless access and backhaul links leveraging an intelligent reflecting surface (IRS). This design aims to maximize the sum rate achieved by ground users (GUs) through optimizing the UAV placement, IRS phase shifts, and sub-channel assignments considering the wireless backhaul capacity constraint. To tackle the

[1]The system model in this chapter is illustrated in Fig. 4.2.

underlying mixed integer non-linear optimization problem (MINLP), we first derive the closed-form IRS phase shift solution; we then optimize the sub-channel assignment and UAV placement by using the alternating optimization method. Specifically, we propose an iterative sub-channel assignment method to efficiently utilize the bandwidth and balance bandwidth allocation for wireless access and backhaul links while maintaining the backhaul capacity constraint. Moreover, we employ the successive convex approximation (SCA) method to solve the UAV placement optimization sub-problem. We show the effectiveness of our proposed design via extensive numerical studies.

6.2 Introduction

Unmanned aerial vehicle (UAV) communications enable network operators to significantly enhance the performance of next-generation wireless communications networks thanks to their attributes such as mobility, flexibility, and adaptive altitude [70, 199]. There are several recent work addressing the UAV placement problem considering capacity constrained backhaul where UAVs act as flying base stations (BSs) [75, 200]. In [200], the authors investigated the joint optimization of UAV placement, spectrum allocation, and power control to maximize the sum rate. Meanwhile, joint bandwidth allocation and UAV placement considering mixed line-of-sight (LoS) and non-LoS propagation were optimized in [75].

To enhance communications quality between UAVs and ground users (GUs), intelligent reflecting surface (IRS) assisted communication has been proposed for UAV-based wireless networks [26, 141]. In particular, the IRSs are installed on UAVs or deployed on building walls [27, 43, 146, 153]. The setting with a single UAV using an IRS was studied in [146] where the underlying design aims to maximize the energy efficiency. On the other hand, the optimization of UAV trajectory and transmit power control for a single IRS deployed on a building wall was investigated in [27, 43, 153] with the objective of maximizing the secrecy rate in [43] and the sum rate in [27, 153]. Besides, the sub-channel assignment design was addressed in [153]. In these work, the optimal IRS phase shifts that are aligned with the phases of the channel coefficients were derived to maximize the achievable rate.

To the best of our knowledge, none of the existing work has studied the multi-carrier IRS-assisted UAV-based wireless network taking into account the constrained capacity of wireless backhauls. To fill this research gap, we study the joint optimization of UAV placement, IRS phase shifts, and sub-channel assignments for wireless access and backhaul links where our design objective is to maximize the sum rate achieved by GUs. To solve the underlying mixed integer nonlinear program (MINLP), we first derive the closed-form IRS phase shift solution and then optimize the sub-channel assignment and UAV placement in an iterative manner by using the alternating optimization method. The sets of sub-channels assigned for the access and backhaul links are iteratively updated to efficiently utilize the available bandwidth while maintaining the backhaul capacity constraint. Moreover, we employ the successive convex approximation (SCA) technique to solve the UAV placement sub-problem. Numerical results are presented to study the impacts of different parameters on the achieved sum rate.

6.3 System Model

We consider downlink communications between a UAV and a set of GUs in an IRS-assisted wireless network with the backhaul link between the UAV and a base station (BS). We define $\mathcal{K}$ as the set of GUs, i.e., $\mathcal{K} = \{1, ..., K\}$, located on the ground at fixed horizontal coordinates $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. We assume that the UAV is placed at the altitude H with the horizontal coordinate $\mathbf{q} = (x^{\mathsf{d}}, y^{\mathsf{d}})$. The UAV acts as an airborne BS connected to the core network wirelessly through a cellular BS which is placed at the coordinate $\mathbf{r}^{\mathsf{b}} = (x^{\mathsf{b}}, y^{\mathsf{b}})$ and a fixed altitude H^{b}.

We assume that a single IRS is installed on the surface of a building wall at the altitude H^{i} and horizontal coordinate $\mathbf{w}^{\mathsf{i}} = (x^{\mathsf{i}}, y^{\mathsf{i}})$. The IRS is made up of $I_r \times I_c$ passive reflection elements units installed as a uniform planar array (UPA) with I_c and I_r elements on each column and each row, respectively. The distance between any two adjacent elements of the IRS is denoted by d. The phase shift matrix of the IRS is denoted by $\mathbf{\Phi} = \text{diag}\left\{e^{j\phi_{1,1}}, \ldots, e^{j\phi_{i_r,i_c}}, \ldots, e^{j\phi_{I_r,I_c}}\right\} \in \mathbb{C}^{I_r \times I_c}$, where $\phi_{i_r,i_c} \in [0, 2\pi)\,, \forall i_r = 1, \ldots, I_r$, and $i_c = 1, \ldots, I_c$. In the following, we describe the sub-channel assignment, the communication model and present the problem formulation.

6.3.1 Sub-channel Assignment

We assume that orthogonal frequency-division multiple access (OFDMA) is employed for both wireless access and backhaul links where $\mathcal{C} = \{1, ..., C\}$ denotes the set of available sub-channels and the bandwidth of each sub-channel is W (Hz). Let $\psi^{\mathsf{A}}_{k,c}$ denote sub-channel assignment variables for the access links between the UAV and K GUs, where $\psi^{\mathsf{A}}_{k,c} = 1$, if sub-channel c is assigned for GU k and $\psi^{\mathsf{A}}_{k,c} = 0$, otherwise. Similary, we define $\psi^{\mathsf{B}}_{0,c}$ as sub-channel assignment variables for the backhaul link, where $\psi^{\mathsf{B}}_{0,c} = 1$, if sub-channel c is assigned for the backhaul link and $\psi^{\mathsf{B}}_{0,c} = 0$, otherwise. The first requirement for the assignment is that each GU must be assigned at least one sub-channel in order to maintain a communication and this can be expressed as $\sum_{c\in\mathcal{C}} \psi^{\mathsf{A}}_{k,c} \geq 1, \forall k$. Another constraint is that each sub-channel can be used either to support only one GU or the backhaul link. This constraint can be expressed as $\sum_{k\in\mathcal{K}} \psi^{\mathsf{A}}_{k,c} + \psi^{\mathsf{B}}_{0,c} \leq 1, \forall c$.

6.3.2 Channel Model

We assume that all BS-UAV, UAV-IRS, and UAV-GU[2] communication links are dominated by the LoS propagation while communications channels between the IRS and GUs experience Rayleigh channel fading due to blockages. Hence, the distances among BS, UAV, IRS, and GUs can be calculated based on their coordinates as $d^{\mathsf{BU}} = \sqrt{\|\mathbf{r}^{\mathsf{b}} - \mathbf{q}\|^2 + (H^{\mathsf{b}} - H)^2}$, $d^{\mathsf{UI}} = \sqrt{\|\mathbf{q} - \mathbf{w}^{\mathsf{i}}\|^2 + (H - H^{\mathsf{i}})^2}$, $d^{\mathsf{UG}}_k = \sqrt{\|\mathbf{q} - \mathbf{r}^{\mathsf{u}}_k\|^2 + H^2}, \forall k$, $d^{\mathsf{IG}}_k = \sqrt{\|\mathbf{w}^{\mathsf{i}} - \mathbf{r}^{\mathsf{u}}_k\|^2 + (H^{\mathsf{i}})^2}, \forall k$, corresponding to the distances from BS to UAV, UAV to IRS, UAV to GU k, and IRS to GU k, respectively. Moreover, we assume that the power of the signals that are reflected by the IRS two or more times is negligible and thus ignored due to their large path loss.

6.3.2.1 Channel Model for IRS-assisted UAV Communication

As discussed in [43], the received signal at GU k due to the communications from the UAV is given by $y_k = \sqrt{p}\left((\mathbf{h}^{\mathsf{IG}}_k)^H \boldsymbol{\Phi} \mathbf{h}^{\mathsf{UI}} + h^{\mathsf{UG}}_k\right) x_k + n^{\mathsf{G}}$, where x_k represents the transmitted symbol from the UAV, which satisfies $\mathbb{E}(|x_k|^2) = 1$, and p denotes the transmit power of the UAV for GU k on each sub-channel, i.e., $p = P_{\max}/C$ assuming uniform power allocation where $P_{\max}$ is the total transmit power of UAV, and n^{G} denotes the additive white Gaussian noise (AWGN) at GU, with zero mean and variance σ^2. Also, let h^{UG}_k, $\mathbf{h}^{\mathsf{UI}}$, and $\mathbf{h}^{\mathsf{IG}}_k$ denote the channel coefficients of the links between UAV and GU k, UAV and IRS, IRS and GU k, respectively, which are expressed as $h^{\mathsf{UG}}_k = \sqrt{\frac{\beta_0}{(d^{\mathsf{UG}}_k)^2}}, \forall k$,

[2] In our setting, the probability of LoS propagation between the UAV and the GUs is more than 95% in suburban and 75% – 80% in urban areas [72].

and

$$\mathbf{h}^{\mathsf{UI}} = \sqrt{\frac{\beta_0}{(d^{\mathsf{UI}})^2}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}}}\right]^H \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}}\right]^H, \tag{6.1}$$

$$\mathbf{h}_k^{\mathsf{IG}} = \sqrt{\frac{\beta_0}{(d_k^{\mathsf{IG}})^\kappa}} \times \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_r-1)\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}}\right]^H \otimes \left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}, ..., e^{-j\frac{2\pi d}{\lambda}(I_c-1)\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}}\right]^H \times \alpha^{\mathsf{IG}}, \forall k, \tag{6.2}$$

where β_0 denotes the channel gain at the reference distance of 1 meter, κ is the path loss exponent, λ is the wavelength of the carrier wave, and α^{IG} is the random scattering components modeled by a circularly symmetric complex Gaussian random variable with zero mean and unit variance. In addition, $(\theta^{\mathsf{UI}}, \xi^{\mathsf{UI}})$ and $(\theta_k^{\mathsf{IG}}, \xi_k^{\mathsf{IG}})$ represent the vertical and horizontal angle-of-departures from the UAV to the IRS and from the IRS to GU k, respectively, which can be calculated from $\sin\theta^{\mathsf{UI}} = \frac{|H-H^{\mathsf{i}}|}{d^{\mathsf{UI}}}$, $\sin\xi^{\mathsf{UI}} = \frac{|x^{\mathsf{i}}-x^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\cos\xi^{\mathsf{UI}} = \frac{|y^{\mathsf{i}}-y^{\mathsf{d}}|}{\sqrt{\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\|^2}}$, $\sin\theta_k^{\mathsf{IG}} = \frac{H^{\mathsf{i}}}{d_k^{\mathsf{IG}}}$, $\sin\xi_k^{\mathsf{IG}} = \frac{|x^{\mathsf{i}}-x_k^{\mathsf{u}}|}{\sqrt{\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\|^2}}$, and $\cos\xi_k^{\mathsf{IG}} = \frac{|y^{\mathsf{i}}-y_k^{\mathsf{u}}|}{\sqrt{\|\mathbf{w}^{\mathsf{i}}-\mathbf{r}_k^{\mathsf{u}}\|^2}}$, $\forall k \in \mathcal{K}$.

6.3.2.2 Achievable Rate

As presented in [43], the achievable rate for GU k served by the UAV on sub-channel c can be expressed as

$$R_{k,c}^{\mathsf{A}} = \psi_{k,c}^{\mathsf{A}} W \log_2\left(1+\frac{p}{\sigma^2}\left|\frac{\sqrt{\beta_0}}{d_k^{\mathsf{UG}}}+\frac{\beta_0 f_k|\alpha^{\mathsf{IG}}|}{(d_k^{\mathsf{IG}})^{\kappa/2}d^{\mathsf{UI}}}\right|^2\right), \tag{6.3}$$

where $f_k = \sum_{i_c=1}^{I_c}\sum_{i_r=1}^{I_r} e^{j\left(F_k^{i_r,i_c}+\phi_{i_r,i_c}\right)}, \forall k$, and $F_k^{i_r,i_c} = -\frac{2\pi d}{\lambda}((i_r-1)(\sin\theta_k^{\mathsf{IG}}\cos\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\cos\xi^{\mathsf{UI}})+(i_c-1)(\sin\theta_k^{\mathsf{IG}}\sin\xi_k^{\mathsf{IG}}+\sin\theta^{\mathsf{UI}}\sin\xi^{\mathsf{UI}}))-\arg(\alpha^{\mathsf{IG}})$.

Also, the achievable rate of the backhaul link on sub-channel c can be expressed as

$$R_{0,c}^{\mathsf{B}} = \psi_{0,c}^{\mathsf{B}} W \log_2\left(1+\frac{p_0\beta_0}{(d^{\mathsf{BU}})^2\sigma^2}\right), \tag{6.4}$$

where p_0 denotes the transmit power of the cellular BS.

Moreover, to maintain the good end-to-end performance, the total data rate of all access links from the UAV to all the GUs should not exceed the backhaul rate. This constraint can be described as $\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}} R^{\mathsf{B}}_{0,c}$.

6.3.3 Problem Formulation

Let $\mathbf{\Psi} = \{\psi^{\mathsf{A}}_{k,c}, \psi^{\mathsf{B}}_{0,c}, \forall k, c\}$, $\mathbf{\Phi}$, and $\mathbf{Q} = \{\mathbf{q}\}$ denote vectors of all decision variables for sub-channel assignment, IRS phase shifts, and UAV placement, respectively. We want to maximize the sum rate of all GUs by optimizing all variables $\mathbf{\Psi}, \mathbf{\Phi}$, and $\mathbf{Q}$. This design problem can be formulated as

$$\textbf{(P2):} \quad \max_{\mathbf{\Psi},\mathbf{\Phi},\mathbf{Q}} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} R^{\mathsf{A}}_{k,c} \tag{6.5}$$

$$\text{s.t.} \sum_{c\in\mathcal{C}} \psi^{\mathsf{A}}_{k,c} \geq 1, \forall k, \tag{6.5a}$$

$$\sum_{k\in\mathcal{K}} \psi^{\mathsf{A}}_{k,c} + \psi^{\mathsf{B}}_{0,c} \leq 1, \forall c, \tag{6.5b}$$

$$\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}} R^{\mathsf{B}}_{0,c}, \tag{6.5c}$$

$$\psi^{\mathsf{A}}_{k,c}, \psi^{\mathsf{B}}_{0,c} \in \{0,1\}, \forall k, c, \tag{6.5d}$$

$$\phi_{i_r,i_c} \in [0, 2\pi)\,, \forall i_r = 1,\ldots, I_r; \forall i_c = 1, \ldots, I_c. \tag{6.5e}$$

Because of the non-convex constraint (6.5c) and integer variables in (6.5d), problem (6.5) is a non-convex mixed integer nonlinear optimization program (MINLP), which is difficult to solve. One might argue that adding constraint (6.5c) is a trivial modification of the previous models. While this is certainly true as far as writing the mathematical model, this constraint is not convex and thus makes the design of an efficient solution algorithm much more complicated. In the following, we describe the details of our proposed algorithm.

6.4 Proposed Algorithm

To solve problem **(P2)** we first derive the closed-form phase shift solution and then optimize the sub-channel assignment and UAV placement iteratively. Let $\mathcal{C}^{\mathsf{A}}$ and $\mathcal{C}^{\mathsf{B}}$ be the sets of sub-channels assigned for access and backhaul links, respectively where $\mathcal{C} = \mathcal{C}^{\mathsf{A}} \cup \mathcal{C}^{\mathsf{B}}$. Initially, the number of sub-channels allocated in $\mathcal{C}^{\mathsf{A}}$ is equal to the number of GUs K to ensure each GU is assigned at least one sub-channel and all remaining sub-channels are allocated to $\mathcal{C}^{\mathsf{B}}$. Then, the sets $\mathcal{C}^{\mathsf{A}}$ and $\mathcal{C}^{\mathsf{B}}$ are updated by taking sub-channels from $\mathcal{C}^{\mathsf{B}}$ and re-allocating to $\mathcal{C}^{\mathsf{A}}$ and the IRS phase shifts and UAV placement are optimized accordingly while maintaining the backhaul capacity constraint. Details of the proposed algorithm is described in the following.

To obtain the maximum access rate $R^{\mathsf{A}}_{k,c}$ given in (6.3) and hence the sum rate, i.e., the objective function, the IRS phase shift $\mathbf{\Phi}^*$ must be aligned with the phases of channel coefficients. Such optimal IRS phase shifts, which result in $f^*_k = I_c I_r$, can be expressed as

$$\begin{aligned}\phi^*_{i_r,i_c} \quad = \quad & \frac{2\pi d}{\lambda}\Big((i_r-1)(\sin\theta^{\mathsf{IG}}_k \cos\xi^{\mathsf{IG}}_k + \sin\theta^{\mathsf{UI}} \cos\xi^{\mathsf{UI}}) \\ & +(i_c-1)(\sin\theta^{\mathsf{IG}}_k \sin\xi^{\mathsf{IG}}_k + \sin\theta^{\mathsf{UI}} \sin\xi^{\mathsf{UI}})\Big) \\ & + \arg(\alpha^{\mathsf{IG}}). \end{aligned} \tag{6.6}$$

Substituting this IRS phase shifts into problem **(P2)** still results in a non-convex MINLP problem. Thus, we employ the alternating optimization approach to tackle this problem where we iteratively optimize each set of optimization variables given the values of other variables until convergence.

6.4.1 Optimization of Sub-channel Assignment

For given $\mathbf{\Phi}$ and $\mathbf{Q}$, the sub-problem to optimize the sub-channel assignment $\mathbf{\Psi}$ can be stated as

$$\textbf{(P2.1):} \max_{\mathbf{\Psi}} \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{A}}_{k,c} \tag{6.7}$$

$$\text{s.t.} \quad \sum_{c\in\mathcal{C}^{\mathsf{A}}} \psi^{\mathsf{A}}_{k,c} \geq 1, \forall k, \tag{6.7a}$$

$$\sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{A}}_{k,c} \leq \sum_{c\in\mathcal{C}^{\mathsf{B}}} R^{\mathsf{B}}_{0,c}, \tag{6.7b}$$

$$\text{constraints} \;\; (6.5\text{b}), (6.5\text{d}).$$

This is a standard mixed integer linear program (MILP), which can be solved efficiently by using the CVX-Mosek solver [38].

6.4.2 Optimization of UAV Placement

For given $\mathbf{\Psi}$ and $\mathbf{\Phi}$, the sub-problem to optimize the UAV placement $\mathbf{Q}$ is non-convex. To solve this problem, we first introduce some auxiliary variables and then solve the transformed problem by using the SCA method. Specifically, we introduce variables $\nu_k \geq \|\mathbf{q}-\mathbf{r}_k^{\mathsf{u}}\|^2 + H^2, \forall k$, $\mu \geq \left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2 + (H-H^{\mathsf{i}})^2$, and $\epsilon \geq \left\|\mathbf{r}^{\mathsf{b}}-\mathbf{q}\right\|^2 + (H^{\mathsf{b}}-H)^2$. From (6.3) and (6.4), we have

$$\begin{aligned} R_{k,c}^{\mathsf{Aq}} &= \psi_{k,c}^{\mathsf{A}} W \log_2\left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k}+\frac{Y_k^2}{\mu}+\frac{2X_kY_k}{\nu_k^{1/2}\mu^{1/2}}\right)\right) \\ &\leq R_{k,c}^{\mathsf{A}}, \end{aligned} \tag{6.8}$$

$$R_{0,c}^{\mathsf{Bq}} = \psi_{0,c}^{\mathsf{B}} W \log_2\left(1+\frac{Z}{\epsilon}\right) \leq R_{0,c}^{\mathsf{B}}, \tag{6.9}$$

where $X_k = \sqrt{\beta_0}$, $Y_k = \beta_0 f_k^* |\alpha^{\mathsf{IG}}|(d_k^{\mathsf{IG}})^{-\kappa/2}$, $Z = p_0\beta_0/\sigma^2$, in which $f_k^* = I_c I_r$ is a solution given by the IRS phase shifts expressed in (6.6). It can be verified that $R_{k,c}^{\mathsf{Aq}}$ is a convex function with respect to ν_k and μ and it can be lower-bounded by its first-order Taylor expansion at r-th iteration in the approximation process as follows:

$$\begin{aligned} R_{k,c}^{\mathsf{Aq}} &\geq \psi_{k,c}^{\mathsf{A}} W \log_2 D^r + \frac{L^r}{D^r}(\nu_k-\nu_k^r) + \frac{S^r}{D^r}(\mu-\mu^r) \\ &\triangleq R_{k,c}^{\mathsf{Aqlb}}, \end{aligned} \tag{6.10}$$

where

$$\begin{aligned} D^r &= \left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k^r}+\frac{Y_k^2}{\mu^r}+\frac{2X_kY_k}{\nu_k^{1/2,r}\mu^{1/2,r}}\right)\right), \\ L^r &= -\psi_{k,c}^{\mathsf{A}} W \log_2(e)\left(\frac{p}{\sigma^2}\left(\frac{X_k^2}{\nu_k^{2,r}}+\frac{X_kY_k}{\nu_k^{3/2,r}\mu^{1/2,r}}\right)\right), \\ S^r &= -\psi_{k,c}^{\mathsf{A}} W \log_2(e)\left(\frac{p}{\sigma^2}\left(\frac{Y_k^2}{\mu^{2,r}}+\frac{X_kY_k}{\nu_k^{1/2,r}\mu^{3/2,r}}\right)\right). \end{aligned}$$

Similarly, since $R^{\mathsf{Bq}}_{0,c}$ is convex with respect to ϵ, by applying the first-order Taylor expansion at the given point ϵ^r, it can be lower-bounded as

$$\begin{aligned} R^{\mathsf{Bq}}_{0,c} &\geq \psi^{\mathsf{B}}_{0,c} W \log_2\left(1+\frac{Z}{\epsilon^r}\right) - \psi^{\mathsf{B}}_{0,c} W \frac{\log_2(e) Z}{\epsilon^r(\epsilon^r + Z)}(\epsilon - \epsilon^r) \\ &\triangleq R^{\mathsf{Bqlb}}_{0,c}, \end{aligned} \tag{6.11}$$

Moreover, the upper-bound of the access rate given in (6.3) can be expressed by introducing auxiliary variables $\alpha_k \leq \|\mathbf{q}-\mathbf{r}^{\mathsf{u}}_k\|^2 + H^2, \forall k$, $\gamma \leq \left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2 + (H-H^{\mathsf{i}})^2$ and we have

$$\begin{aligned} R^{\mathsf{Aub}}_{k,c} &= \psi^{\mathsf{A}}_{k,c} W \log_2\left(1+\frac{p}{\sigma^2}\left(\frac{X_k^2}{\alpha_k}+\frac{Y_k^2}{\gamma}+\frac{2X_kY_k}{\alpha_k^{1/2}\gamma^{1/2}}\right)\right) \\ &\geq R^{\mathsf{A}}_{k,c}, \end{aligned} \tag{6.12}$$

It also can be verified that $R^{\mathsf{Aub}}_{k,c}$ is a convex function with respect to α_k and γ. Besides, since $\|\mathbf{q}-\mathbf{r}^{\mathsf{u}}_k\|^2$ and $\left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2$ are convex functions with respect to $\mathbf{q}$, by applying the first-order Taylor expansion at the given point $\mathbf{q}^r$, we have $\|\mathbf{q}-\mathbf{r}^{\mathsf{u}}_k\|^2 \geq \|\mathbf{q}^r-\mathbf{r}^{\mathsf{u}}_k\|^2 + 2\left(\mathbf{q}^r-\mathbf{r}^{\mathsf{u}}_k\right)^T(\mathbf{q}-\mathbf{q}^r)$ and $\left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2 \geq \left\|\mathbf{q}^r-\mathbf{w}^{\mathsf{i}}\right\|^2 + 2\left(\mathbf{q}^r-\mathbf{w}^{\mathsf{i}}\right)^T(\mathbf{q}-\mathbf{q}^r)$.

Therefore, the UAV placement optimization problem can be approximated by

$$\textbf{(P2.2):} \max_{\mathbf{Q},\nu_k,\mu,\epsilon,\alpha_k,\gamma} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{Aqlb}}_{k,c} \tag{6.13}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}^{\mathsf{A}}} R^{\mathsf{Aub}}_{k,c} - \sum_{c\in\mathcal{C}^{\mathsf{B}}} R^{\mathsf{Bqlb}}_{0,c} \leq 0, \tag{6.13a}$$

$$\nu_k \geq \|\mathbf{q}-\mathbf{r}^{\mathsf{u}}_k\|^2 + H^2, \forall k;\ \mu \geq \left\|\mathbf{q}-\mathbf{w}^{\mathsf{i}}\right\|^2 + (H-H^{\mathsf{i}})^2, \tag{6.13b}$$

$$\alpha_k \leq \|\mathbf{q}^r-\mathbf{r}^{\mathsf{u}}_k\|^2 + 2\left(\mathbf{q}^r-\mathbf{r}^{\mathsf{u}}_k\right)^T(\mathbf{q}-\mathbf{q}^r) + H^2, \forall k, \tag{6.13c}$$

$$\gamma \leq \left\|\mathbf{q}^r-\mathbf{w}^{\mathsf{i}}\right\|^2 + 2\left(\mathbf{q}^r-\mathbf{w}^{\mathsf{i}}\right)^T(\mathbf{q}-\mathbf{q}^r) + (H-H^{\mathsf{i}})^2, \tag{6.13d}$$

$$\epsilon \geq \left\|\mathbf{r}^{\mathsf{b}}-\mathbf{q}\right\|^2 + (H^{\mathsf{b}}-H)^2. \tag{6.13e}$$

This is a convex problem, which can be solved efficiently by using the CVX-Mosek solver [38].

Solutions of these sub-problems are used in our proposed algorithm which is described in Algorithm 6.1. In each iteration of this algorithm, we take a sub-channel c from $\mathcal{C}^{\mathsf{B}}$ and re-allocate to $\mathcal{C}^{\mathsf{A}}$ in step 4. Steps 5 to 9 describe the iterative method to solve problem **(P2)** for given sub-channel

Algorithm 6.1. Joint Algorithm for Sub-channel Assignment, IRS Phase Shifts, and UAV Placement

1: Initialization: $\mathcal{C}^{\mathsf{A}}$,$\mathcal{C}^{\mathsf{B}}$,$\mathbf{Q}^0$,$\mathbf{\Phi}^0$,$\mathbf{\Psi}^0$, $S^* = 10^2, S = S_1 = 0$, $t = 0$;
2: **repeat**
3: $\quad S^* = S$ and $t = t + 1$;
4: $\quad$ Take a sub-channel c from $\mathcal{C}^{\mathsf{B}}$; update $\mathcal{C}^{\mathsf{A}} = \mathcal{C}^{\mathsf{A}} \cup \{c\}$ and $\mathcal{C}^{\mathsf{B}} \setminus \{c\}$;
5: $\quad$ **repeat**
6: $\qquad$ Given $\mathbf{\Phi}^{r,*}$ in (6.6);
7: $\qquad$ Solve **(P2.1)** iteratively until convergence to obtain $\mathbf{\Psi}^r$;
8: $\qquad$ Solve **(P2.2)** iteratively until convergence to obtain $\mathbf{Q}^r$;
9: $\quad$ **until** Convergence
10: $\quad$ **if** Obtain a feasible solution with maximum sum rate S_1 and $S<S_1$ **then**
11: $\qquad$ Update $S = S_1$ and $\{\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*\} = \{\mathbf{\Psi}^t, \mathbf{Q}^t, \mathbf{\Phi}^t\}$;
12: $\quad$ **else**
13: $\qquad$ Update $\mathcal{C}^{\mathsf{B}} = \mathcal{C}^{\mathsf{B}} \cup \{c\}$ and $\mathcal{C}^{\mathsf{A}} \setminus \{c\}$;
14: $\qquad$ Update $\{\mathbf{\Psi}^*, \mathbf{Q}^*, \mathbf{\Phi}^*\} = \{\mathbf{\Psi}^{t-1}, \mathbf{Q}^{t-1}, \mathbf{\Phi}^{t-1}\}$;
15: $\quad$ **end if**
16: **until** $|S - S^*| < 10^{-6}$
17: Return $\mathbf{\Psi}^*$,$\mathbf{Q}^*$, $\mathbf{\Phi}^*$;

allocation sets $\mathcal{C}^{\mathsf{A}}$ and $\mathcal{C}^{\mathsf{B}}$. Specifically, we alternatively optimize sub-channel assignment and UAV placement sub-problems given the optimal IRS phase shift solution given in (6.6) to obtain a feasible solution. The CVX-Mosek solver employs the interior point method to solve these sub-problems; hence, the complexity of these steps is $\mathcal{O}(L_1 K C^{\frac{7}{2}})$, where L_1 denotes the number of iterations needed to reach the convergence. Steps 10 to 15 verify the feasibility of the obtained solution and update the optimization variables and the achieved sum rate accordingly. Specifically, if a feasible solution with a higher sum rate $S_1 > S$ is obtained, the objective value and the optimization variables are updated in step 11. Otherwise, additional allocation of sub-channel c in $\mathcal{C}^{\mathsf{B}}$ to $\mathcal{C}^{\mathsf{A}}$ could not maintain the backhaul capacity constraint, i.e., backhaul bandwidth is not sufficient to support the total access rate; hence, we reverse the sub-channel assignment to that in the previous iteration by updating the sets $\mathcal{C}^{\mathsf{A}}, \mathcal{C}^{\mathsf{B}}$ in step 13 and update the optimization variables accordingly in step 14. The whole algorithm terminates when the achieved sum rate cannot be improved further. The complexity of the whole algorithm is $\mathcal{O}(L(L_1 K C^{\frac{7}{2}}))$, where L denotes the number of iterations needed to achieve convergence.

6.5 Numerical Results

We consider a rectangular network area with size $1000 \times 1000(\mathrm{m}^2)$. The altitude of the UAV is fixed at $H = 120$m and the BS is located at $(0, 0, 20)$m. In addition, the IRS is fixed at $(500, 500, 50)$m

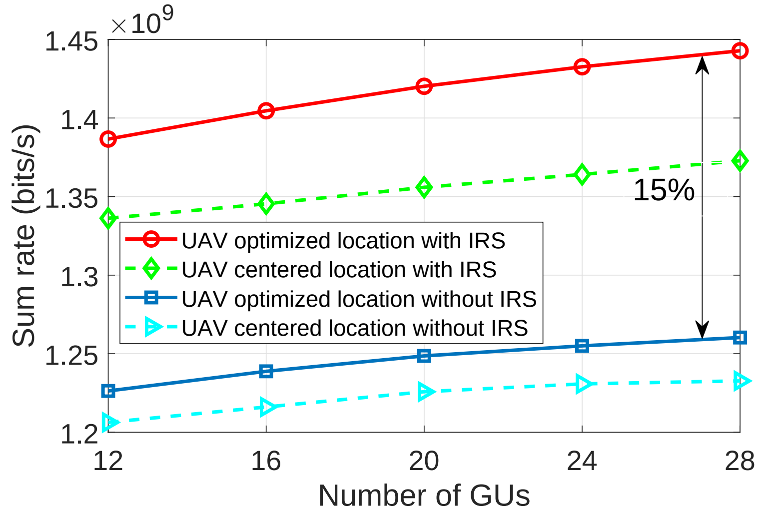

Figure 6.1 – Sum rate for different number of GUs.

and GUs are placed inside circular clusters with a radius of $r_c = 200$m. We initially locate the UAV at the center of the GUs' cluster. The remaining parameters are set as $p_0 = 33$dBm, $P_{\mathsf{max}} = 30$dBm, $W = 1$MHz, $\sigma^2 = -110$dBm, $f_c = 2.5$GHz, $d = \lambda/2$, and $\kappa = 2$. Square IRSs with $I_r = I_c$ will be considered where the number of IRS elements is denoted by $I = I_r I_c$.

Fig. 6.1 – Fig. 6.3 show the sum rate achieved by the proposed algorithm, i.e., Alg. 6.1, and compared with the case where the UAV is placed at the cluster's center, which are indicated as "UAV optimized location" and "UAV centered location", respectively. Fig. 6.1 shows the sum rate for different number of GUs with $C = 60$ and $I = 64$. It can be seen that the sum rate slightly increases with increasing number of GUs and the difference in achieved sum rate between the optimized and centered location of UAV becomes larger as the number of GUs increases. The rate gain due to the proposed algorithm with and without leveraging the IRS is about 15%.

Fig. 6.2 describes the sum rate for different number of sub-channels with 20 GUs and $I = 64$. The sum rate almost linearly increases with the number of sub-channels and the rate gain due to IRS becomes higher for larger number of sub-channels. Fig. 6.3 illustrates the sum rate for different number of IRS elements with 20 GUs and $C = 60$. In fact, larger numbers of sub-channels or IRS elements lead to higher system diversity, which improves the achieved sum rate. These results confirm the effectiveness of the proposed algorithm in optimizing the UAV placement and sub-channel assignment in the IRS-assisted UAV communications.

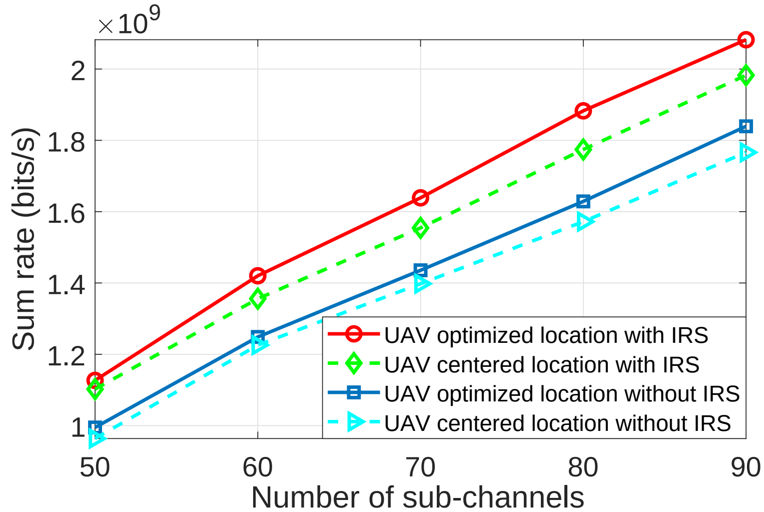

Figure 6.2 – Sum rate for different number of sub-channels.

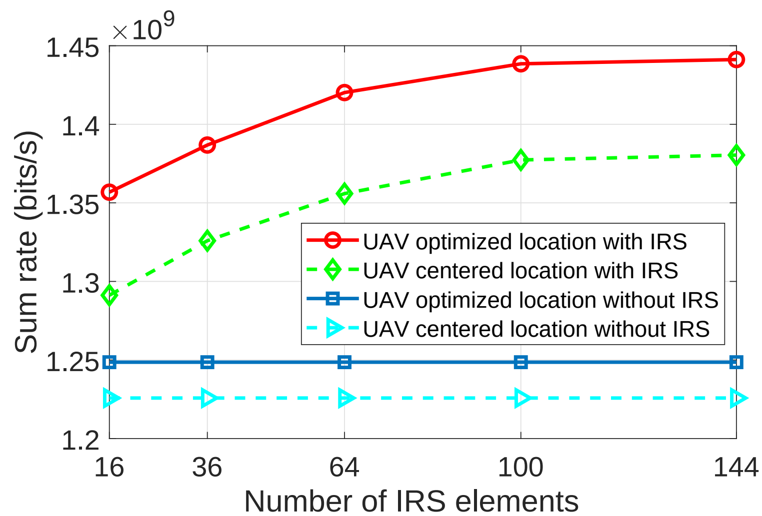

Figure 6.3 – Sum rate for different number of IRS elements.

6.6 Conclusion

In this chapter, we studied the joint optimization of UAV placement, IRS phase shifts, and sub-channel assignment. We proposed an algorithm to solve the underlying MINLP problem and the SCA method was used to tackle the non-convex UAV placement sub-problem. Numerical results have demonstrated the effectiveness of the proposed algorithms where the optimized algorithms leveraging the IRS achieves a significant rate gain compared to the case without IRS.

Chapter 7

Integrated Computation Offloading, UAV Trajectory Control, User Scheduling, Resource Allocation, and Admission Control in SAGIN

The content of this chapter was presented in the following paper:

Minh Dat Nguyen, Long Bao Le, and André Girard, "Integrated Computation Offloading, UAV Trajectory Control, User Scheduling, Resource Allocation, and Admission Control in SAGIN," submitted to *IEEE Transactions Vehicular Technology*, Oct. 2022.

7.1 Abstract

In this chapter, we study the computation offloading problem in space-air-ground integrated networks (SAGIN), where joint optimization of partial computation offloading, unmanned aerial vehicle (UAV) trajectory control, user scheduling, computation, resource allocation, and admission control is performed. Specifically, the considered SAGIN employs multiple UAV-mounted edge servers with controllable UAV trajectory and a cloud sever which can be reached by ground users (GUs)

via multi-hop low-earth-orbit (LEO) satellite communications. This design aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of underlying computation tasks. To tackle the underlying non-convex mixed integer non-linear optimization problem, we use the alternating optimization approach where we iteratively solve four sub-problems, namely user scheduling, partial offloading control and bit allocation over time slots, computation resource and bandwidth allocation, and multi-UAV trajectory control until convergence. Moreover, feasibility verification and admission control strategies are proposed to handle overloaded network scenarios. Furthermore, the successive convex approximation (SCA) method is employed to convexify and solve the non-convex computation resource and bandwidth allocation and UAV trajectory control sub-problems. Via extensive numerical studies, we illustrate the effectiveness of our proposed design compared to baselines.

7.2 Introduction

It is expected that future wireless networks provide higher capacity, lower latency, better communications reliability, and support emerging Internet of Things (IoTs) applications [7]. To this end, a number of promising technologies have been under consideration, including satellite communications, unmanned aerial vehicle (UAV) communications, and mobile edge computing (MEC) [12, 13]. Satellite communications enabled by low-earth-orbit (LEO) satellites possess unique advantages such as low propagation delay, high communication rates, and seamless communication services for wide geographical areas. Satellite communications are, therefore, vital for areas where ground base stations are not available or damaged by natural disasters [47, 91, 92]. By leveraging the complementary strengths of the space, air, and ground network segments, space-air-ground integrated networks (SAGIN) provide effective means for high quality and ubiquitous communications [31–33].

Computation offloading and resource allocation in the SAGIN have attracted great attention [126, 157–159] where computation tasks can be offloaded from ground users (GUs) to UAVs and satellites to save GUs' energy and/or improve computation latency. Several designs for uplink communications and computation offloading in the SAGIN have been investigated in [201–203]. In particular, the uplink communications with ultra-dense LEO satellite constellations and multiple UAVs were studied in [201] to optimize the data gathering efficiency. The joint optimization of task scheduling and computation resource allocation was considered in [202] where multiple satellites and fixed UAVs' trajectories were assumed where the design objective is to minimize the total system cost due to the task delay, users' energy consumption, and server usage. Moreover, the UAV's position and power allocation were optimized in [203] where the objective is to maximize the localization accuracy for the entire area of interest.

There have been some research efforts on developing efficient communication and edge computing management strategies considering inter-satellite links (ISLs) in dense LEO satellite constellations [204, 205]. Specifically, efficient algorithms for dynamic establishment of the inter-plane ISLs in LEO constellations were proposed in [204] where the underlying design aims to maximize the sum rates in the inter-plane ISLs of the LEO constellations and greedy algorithms were developed to solve the satellite matching and resource allocation problem. Meanwhile, Gost *et al.* investigated the energy minimization problem that jointly optimizes the computing and communication resource allocation in [205]. However, efficient utilization of edge computing resources in the SAGIN requires much further research to address various major challenges. First, computing delay and bandwidth constraints must be taken into consideration in engineering the SAGIN. Second, ground users and high altitude platforms such as UAVs typically have limited energy; therefore, energy-efficient design in SAGIN is an important research issue. Finally, many emerging IoT applications have complex design requirements and functionalities such as demanding data transmission, e.g., video downloads, data processing and analysis, e.g., video analysis and speech recognition, and content caching. Therefore, low-latency computation task processing and efficient resource management are needed to enable edge computing based applications in SAGIN.

To the best of our knowledge, none of the previous work has studied computation offloading for the hybrid edge-cloud SAGIN considering user scheduling design over the UAV flight period and multi-hop communications in the satellite network segment while satisfying the maximum delay requirements of underlying computation tasks. To fill this research gap, the current work investigates the integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control for SAGIN with multi-hop satellite communications. The main contributions can be summarized as follows:

- We study partial computation offloading in SAGIN where fractions of computation tasks from GUs are processed locally and/or offloaded and processed at the UAV-mounted edge servers and cloud server leveraging multi-hop LEO satellite communications. We formulate an optimization problem that aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of underlying computation tasks by jointly optimizing the user scheduling, partial offloading control and bit allocation over time, computation resource and bandwidth allocation, and UAV trajectory control.

- The alternating optimization approach is employed to solve the underlying non-convex mixed-integer non-linear optimization problem (MINLP). Moreover, the successive convex approximation (SCA) method is employed to solve the computation resource and bandwidth allocation and UAV trajectory control sub-problems.

- We propose efficient strategies for feasibility verification and admission control in the overloaded network scenarios. Specifically, an iterative algorithm is proposed to solve the feasibility verification problem and an efficient user removal strategy is developed for admission control while satisfying all GUs' and system constraints.

- Numerical results are presented to show the impacts of different parameters including the hop count in the multi-hop satellite communications, number of GUs, bandwidth, and computation task size on the achievable performance and the gains due to optimizing the UAV trajectory control, user scheduling, resource allocation, and computation offloading. Moreover, the admission ratio of GUs that are actually served in the different scenarios is presented.

The remainder of this chapter is organized as follows. In Section 7.3, we discuss the related work on computation offloading and resource allocation in the SAGIN. Section 7.4 presents the system model and problem formulation. In Section 7.5, we describe our proposed integrated algorithm to solve the considered problem when it is feasible. Section 7.6 presents our proposed admission control and network management design. Moreover, we provide the convergence and complexity analysis of the proposed algorithm. Section 7.7 discusses numerical studies for performance evaluations of the developed algorithms. Finally, Section 7.8 concludes the chapter.

7.3 Related Work

The related work on computation offloading and resource allocation in SAGIN with key design features are summarized in Table 7.1. In the computation offloading and resource allocation design

Table 7.1 – Related work on computation offloading and resource allocation in SAGIN

Ref.	Objective	Maximum Delay Constraint	Multi-hop Satellite	UAV Trajectory Optimization	Parallel Offloading Control	User Association	Admission Control
[160]	Maximize long-term time-averaged system computation rate			✓			
[161]	Minimize long-term time-averaged network operation cost					✓	
[162]	Minimize the expected total latency					✓	
[163]	Minimize difference between allocated and required data rate					✓	
[164]	Maximize system energy efficiency					✓	
[165]	Maximize the system capacity			✓		✓	
[48]	Minimize weighted energy consumption of GUs and UAVs			✓		✓	
[166]	Maximize average throughput among GUs			✓		✓	
[44]	Minimize the maximum computation delay among IoT devices	✓			✓	✓	
[45]	Minimize the average latency of all GUs				✓	✓	
[167]	Maximize the sum rate of the small cell	✓			✓	✓	
[206]	Minimize weighted energy consumption of GUs and UAVs	✓	✓	✓	✓	✓ ✓User scheduling	
This Work	Minimize weighted energy consumption of GUs and UAVs	✓	✓	✓	✓	✓ ✓User scheduling	✓

for SAGIN, consideration of different key aspects including maximum delay constraint, multi-hop satellite communications, UAV trajectory optimization, parallel offloading control, user association/scheduling, and admission control is crucial but making the design very challenging. As shown in Table 7.1, most related work only considered the association between GUs and UAVs, i.e., edge servers or satellites but not both. In addition, to guarantee the maximum delay constraints, admission control is necessary in overloaded network scenarios. Furthermore, multi-hop satellite communications must be explicitly modeled to appropriately capture the communications and computing delay in the SAGIN.

Table 7.1 shows that our current work considering all key design aspects provides the most comprehensive design compared to the existing literature as clarified in the following. Liu *et al.* [160] considered a simple SAGIN setting with a single LEO satellite providing cloud computing capability and a UAV-mounted MEC server providing computing resources near the GUs where the design aims to maximize the long-term time-averaged total system computation rate by optimizing the computing resource, power allocation, and UAV trajectory control. Optimization of user association between the GUs and multiple UAV-mounted edge servers was studied in [161–164]. However, these work only optimized the UAV placement or assumed that the UAVs' trajectories are predetermined. Particularly, the joint optimization of user association, task assignment, computing resource allocation at GUs and UAVs to minimize the long-term time averaged network operation cost was investigated in [161]. Meanwhile, Chen *et al.* [162] studied the joint optimization of access control, task scheduling, and computation resource allocation to minimize the expected latency considering multiple satellites and fixed UAVs' trajectories. Moreover, user association, sub-channel assignment, and power control to minimize the difference between the allocated and required data rates were addressed in [163]. Furthermore, the joint optimization of sub-channel selection, power control, and UAV deployment was conducted in [164] where the design objective is to maximize the system energy efficiency.

Optimization of UAV trajectory control and user association for SAGIN has been performed considering different design objectives as functions of throughput/capacity and energy consumption [48, 165, 166]. Specifically, the SAGIN setting with one LEO satellite and multiple UAVs was considered in [165] where the design aims to maximize the system capacity by jointly optimizing user association, power control, and UAV trajectory. Meanwhile, Liu *et al.* [48] studied the SAGIN with one satellite and multiple UAVs and the work aims to minimize the weighted energy consumption via joint optimization of device association, resource partitioning, bit allocation, and UAV trajectory control. Moreover, Pervez *et al.* [166] considered the downlink communications for SAGIN with a single satellite, multiple UAVs, and base station with user terminals and the design aims to maximize the average throughput among GUs by jointly optimizing user association, power control, and UAV trajectory. However, the above existing work have not considered the maximum delay constraints in the computation offloading control design and they only tackled the binary task assignment or task partitioning between the local device and the MEC or cloud servers.

The general computation offloading designs with the parallel task execution at local devices, MEC, and/or cloud servers were investigated in [44, 45, 167]. Specifically, minimization of the maximum delay experienced by different GUs by jointly optimizing UAV-device association, task assignment, power control, bandwidth allocation, computation resource, and UAV placement was studied in [44]. Feng *et al.* [45] considered a multi-user MEC system and optimized user association and task partitioning to achieve minimum average latency for all GUs where independent and dependent sub-tasks are explored. Moreover, a SAGIN setting with a single satellite, a single UAV, and multiple small cells was studied in [167] whose design aims to maximize the sum rate of the small cells by jointly optimizing the user association, sub-channel and power allocation considering the maximum delay constraint.

Several key design aspects were not addressed satisfactorily in the aforementioned existing work. First, partial offloading for efficient computation load balancing among GUs, edge and cloud servers while considering radio resource allocation, UAV trajectory control, and multi-hop satellite communications has not been studied in the literature. Second, the maximum delay constraints imposed by underlying computation tasks may not be achieved due to limited radio and computation resources. To this end, admission control is an important research problem, which has not been addressed in the SAGIN context. Our current work will address these open issues where we propose a comprehensive design framework for SAGIN where the joint optimization of UAV trajectory control, parallel offloading control, user association and scheduling, and admission control will be addressed to achieve minimum weighted energy consumption of GUs and UAVs considering maximum delay constraints of GUs and multi-hop satellite communications. In our preliminary work [206], we tackled the joint problem of computation offloading, user scheduling, resource allocation, and UAV trajectory control with multi-hop satellite communications when the underlying problem is feasible. Our current work makes several significant extensions compared to the conference publication as follows. First, feasibility verification and admission control designs are addressed for overloaded scenarios. Second, complexity analysis for the proposed algorithms is conducted. Finally, much more extensive numerical studies are performed in this journal version compared to those in the preliminary work [206].

7.4 System Model

We consider the computation offloading design in the SAGIN-based edge-cloud system as illustrated in Fig. 7.1 where the terrestrial network comprises K GUs located on the ground, the aerial network layer employs M UAVs, and the space network layer relies on LEO satellites for connections to a distant cloud server. We denote the sets of satellites, UAVs, and GUs as $\mathcal{S} = \{1, ..., S\}$, $\mathcal{M} = \{1, ..., M\}$, and $\mathcal{K} = \{1, ..., K\}$, respectively.

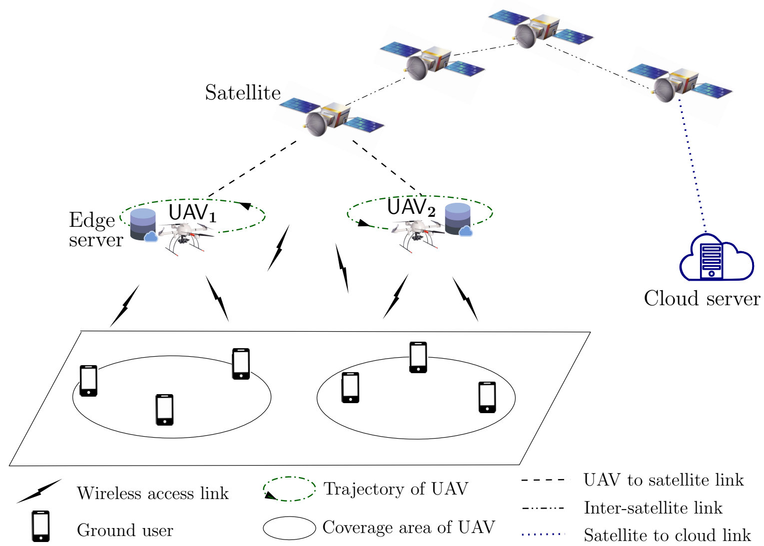

Figure 7.1 – SAGIN with multi-hop satellite communications.

We assume that partial computation offloading is employed for a computation task of each GU. Specifically, each GU partitions its computation task into three sub-tasks where the first sub-task is processed locally and the other two sub-tasks are offloaded and processed at the UAV-mounted edge server and the cloud server, respectively. Moreover, the data related to the second sub-task must be transmitted from the associated GU to the connected UAV while the data related to the third sub-task must be transmitted from the GU to the cloud server via a multi-hop satellite communication path. This network model reflects a practical scenario where a cloud server is deployed far away from the considered terrestrial network area, e.g., the network area in Montreal (45.50°N, 73.56°W) and the cloud server in Vancouver (49.28°N, 123.12°W) approximately 4500km away.

All GUs located on the ground at zero altitude are assumed to have fixed horizontal coordinates of $\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}}), \forall k \in \mathcal{K}$. Besides, we assume that the UAVs fly at a fixed altitude H over a flight period of $T > 0$ seconds. We divide the flight period into N time slots where the set of time slots is denoted as $\mathcal{N} = \{1, ..., N\}$. Moreover, we assume that uplink communications from multiple GUs to their associated UAVs employ the frequency division multiple access (FDMA). Specifically, let W denote the total bandwidth available to support uplink communications from GUs to UAVs. We assume that the available bandwidth is partitioned into orthogonal sub-bands each of which is allocated to one corresponding UAV to serve its associated GU. We denote the bandwidth allocated for UAV m as W_m^{u} then we have $\sum_{m\in\mathcal{M}} W_m^{\mathsf{u}} = W$. In practical LEO satellite deployments, each satellite covers a large area on the ground, e.g., Starlink satellites are deployed at the height of

$H^{\mathsf{s}} = 550$km and each Starlink satellite provides a coverage area with the radius of 580km [207]. As a result, we assume that the associations between GUs and UAVs and between GUs and satellites are fixed during the computation offloading process. Furthermore, we assume that the data size corresponding to the computation results is much smaller than that of the offloading data so that we can neglect the download time of the computation results in the offloading process. For ease of reference, the list of key notations in this chapter is given in Table 7.2.

7.4.1 Computation Task Model

We assume that each GU k has one delay-constrained computation task represented by $U_k = (f_k, s_k, c_k, T_k^{\mathsf{max}})$, where f_k denotes the computation demand expressed by the number of central process unit (CPU) cycles per second (CPU cycles/second), s_k (bits) represents the size of input raw data, c_k (CPU cycles/bit) denotes the computation resource required for 1-bit input data, and T_k^{max} (seconds) describes the maximum tolerable latency of computation task U_k.

We assume that each GU's computation task is partitioned into three sub-tasks that are processed in parallel at the GU, the UAV-mounted edge server, and the cloud server reached via the multi-hop LEO satellite communication as considered in [44, 45]. Then, the task processing time for GU k can be expressed as

$$T_k = \max\left\{T_k^{\mathsf{lo}}, T_k^{\mathsf{ed}}, T_k^{\mathsf{cl}}\right\}, \tag{7.1}$$

where T_k^{lo}, T_k^{ed}, and T_k^{cl} represent the total data transmission and task execution time at the GU, UAV-mounted edge server, and cloud server, respectively. Specifically, T_k^{ed} includes both the data transmission time from GU k to the associated UAV and the execution time of the sub-task from GU k at the associated UAV. We will describe in more detail how to calculate this execution time later. Hence, the delay constraint for GU k can be expressed as $T_k \leq T_k^{\mathsf{max}}$.

To model the task partitioning for GU k, we introduce variables λ_k^{lo} and λ_k^{ed}, $(0 \leq \lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}} \leq 1)$ that represent the fractions of input data to be processed locally at GU k and to be offloaded and processed at the UAV-mounted edge server, respectively. Hence, $(1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}})$ represents the fraction of input data from GU k to be offloaded and processed at the distant cloud server.

Table 7.2 – Key notations

	Description
c_k	Computation resource required for 1-bit input data (CPU cycles/bit)
d_{min}	Minimum inter-UAV distance
e^{ed}	Energy consumption per CPU cycle at UAV-mounted edge server (W/GHz)
f_k	Computation resource of GU (CPU cycles/s)
F_m^{max}	Maximum computation resource of UAV (CPU cycles/s)
H	Fixed altitude of UAVs
κ	Effective switched capacitance depending on the chip architecture
α_1, α_2	Weight factors of the energy consumption of UAVs and GUs, respectively
K	Number of GUs
$\mathcal{K}$	Set of GUs
M	Number of UAVs
$\mathcal{M}$	Set of UAVs
N	Total time slots
$\mathcal{N}$	Set of time slots
P_k^{u}, P_k^{s}	Transmit power of GU with UAV and satellite
P_m^{f}	Flying power consumption of UAV
$\mathbf{r}_k^{\mathsf{u}}$	Fixed horizontal coordinate of GU k $\left(\mathbf{r}_k^{\mathsf{u}} = (x_k^{\mathsf{u}}, y_k^{\mathsf{u}})\right)$
R_k^{s}, R_i^{ss}, R^{cl}	Transmission rate from GU to satellite, inter-satellite, and from satellite to cloud server, respectively
s_k	Size of input data at GU k (bits)
T	Flight period
T_k^{max}	Maximum delay of GU k
T_k^{prop}	Total propagation delay from GU k to cloud server
Δt	Element slot length ($\Delta t = T/N$)
N_k	Number of consecutive time slots, $N_k = \left\lceil T_k^{\mathsf{max}}/\Delta t \right\rceil$, where $\lceil . \rceil$ denotes the round-up operation.
$V_{\max}$	Maximum speed of UAV
$D_{\max}$	Maximum horizontal distance that the UAV can travel in each time slot $\left(D_{\max} \triangleq V_{\max}\Delta t\right)$
W	Total bandwidth
	Decision and Auxiliary Variables
δ	Feasibility checking variable
L	Number of hop-count satellites
$\phi_{k,m}^{\mathsf{u}}[n]$	Association between UAV m and GU k in time slot n
$\mathbf{\Theta}$	Vector of all user scheduling variables: $\mathbf{\Theta} = \{\theta_k[n], \forall k, n\}$
$\mathbf{\Lambda}$	Vector of partial offloading control variables: $\mathbf{\Lambda} = \{\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, \forall k\}$
$\mathbf{L}$	Vector of bit allocation variables: $\mathbf{L} = \{l_k^{\mathsf{u}}[n], \forall k, n\}$
β	Vector of bandwidth allocation variables: $\boldsymbol{\beta} = \{\beta_k^{\mathsf{u}}[n], \forall k, n\}$
$\mathbf{F}$	Vector of computation resource allocation variables: $\mathbf{F} = \{f_k^{\mathsf{u}}[n], \forall k, n\}$
$\mathbf{Q}$	Vector of all time-variant horizontal coordinate of UAVs: $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$
	Functions
E_k^{lo}, E_k^{ed}, E_k^{s}	Energy consumption of GU k for execution or offloading the sub-task at local, edge server, and cloud server
E_m^{edf}	Flying energy consumption of UAV
$E_{k,m}^{\mathsf{edt}}[n]$	Energy consumption for data transmission from GU k to the associated UAV m in time slot n
E^{sum}	Weighted energy consumption of UAVs and GUs $E^{\mathsf{sum}}(\mathbf{\Theta}, \mathbf{L}, \mathbf{\Lambda}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q})$
$g_{k,m}[n]$	Channel power gain from UAV m and GU k in time slot n
$\mathcal{K}^{\mathsf{ac}}$	Set of GUs admitted
$R_{k,m}^{\mathsf{u}}[n]$	Achievable rate of uplink transmission from GU k to the associated UAV m in time slot n
T_k, T_k^{lo}, T_k^{ed}, T_k^{cl}	Processing time for GU k, executed at local, edge server, and cloud server, respectively
$\mathcal{W}_m^{\mathsf{u}}$	Bandwidth allocated to UAV m, $\sum_{m\in\mathcal{M}} W_m^{\mathsf{u}} = W$

7.4.2 UAV Trajectory Control

In practice, the UAV flight period T should be chosen to achieve a good balance between the user throughput and access delay [41]. We assume that the UAV's energy is sufficiently large to cover its flight operation and wireless communications over the flight period T. This assumption holds in many practical scenarios, e.g., a UAV equipped with a 3-cell, 3250mAh, and 11.1V LiPo battery can have a flight time of about 20 minutes [194].

The horizontal coordinates of UAV m in time slot n are denoted as $\mathbf{q}_m[n] = (x^{\mathsf{d}}_m[n], y^{\mathsf{d}}_m[n])$. We assume that each UAV must come back to its initial position at the end of the flight period, i.e., $\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m \in \mathcal{M}$. In addition, the slot interval $\Delta t = \frac{T}{N}$ is chosen to be sufficiently small so that the UAVs' locations are within a bounded small neighborhood in each time slot even at the maximum flight speed $V_{\max}$ in meter/second (m/s). Hence, the trajectories of the UAVs must meet the following constraints:

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \leq D^2_{\max}, \forall m, n = 1, ..., N-1, \tag{7.2}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d^2_{\min}, \forall n, m, j \neq m, \tag{7.3}$$

where $\|.\|$ denotes the Euclidean norm, $D_{\max} \triangleq V_{\max}\Delta t$ is the maximum horizontal distance that an UAV can travel in each time slot, $d_{\min}$ denotes the minimum inter-UAV distance between UAVs to ensure collision avoidance.

7.4.3 User Scheduling

Let $\phi^{\mathsf{u}}_{k,m}[n]$ denote binary decision variables for the association between the GUs and UAVs over flight period T, where $\phi^{\mathsf{u}}_{k,m}[n] = 1$ if GU k is served by UAV m in time slot n and $\phi^{\mathsf{u}}_{k,m}[n] = 0$, otherwise. The first requirement for the association is that each GU can offload its computation sub-task to at most one UAV in each time slot, i.e., $\sum_{m\in\mathcal{M}} \phi^{\mathsf{u}}_{k,m}[n] \leq 1$. We assume that each GU k is initially associated with the UAV providing the highest average received signal strength (RSS), i.e., $\phi^{\mathsf{u}}_{k,m}[n] = 1$ with $m = \arg\max_k(RSS_{k,m}[n])$, where $RSS_{k,m}[n](\text{dBm}) = P^{\mathsf{u}}_k(\text{dBm}) - g_{k,m}[n](\text{dBm})$ with P^{u}_k denotes the transmit power of GU k to its associated UAV and $g_{k,m}[n]$ stands for channel power gain from GU k to UAV m. To satisfy the delay constraint of each GU, the number of *consecutive time slots* required to completely process the computation task of GU k can be denoted as $N_k = \lceil T_k^{\mathsf{max}}/\Delta t \rceil$, where $\lceil . \rceil$ denotes the round-up operation. We now introduce binary user scheduling variables $\theta_k[n]$, where $\theta_k[n] = 1$ if GU k is scheduled to transmit to its associated UAV in time slot n and $\theta_k[n] = 0$, otherwise. We need to impose the following constraints on the user scheduling decisions:

$$\sum_{m\in\mathcal{M}} \sum_{t=0}^{N_k-1} \theta_k[n+t]\phi^{\mathsf{u}}_{k,m}[n+t] = N_k, \forall k, n\in\{1, ..., N-N_k\}. \tag{7.4}$$

7.4.4 Computing Models

The computation task execution time and energy consumption are discussed in the following.

7.4.4.1 Local Computing Model

The local task execution time at GU k can be expressed as

$$T_k^{\mathsf{lo}} = \frac{\lambda_k^{\mathsf{lo}} s_k c_k}{f_k}. \tag{7.5}$$

The delay constraint imposed to the local processing can be expressed as $T_k^{\mathsf{lo}} \leq T_k^{\mathsf{max}}$. The energy consumption due to local task execution can be calculated as

$$E_k^{\mathsf{lo}} = \kappa \lambda_k^{\mathsf{lo}} s_k c_k (f_k)^2, \tag{7.6}$$

where κ is the effective switched capacitance depending on the chip architecture [46].

7.4.4.2 UAV-Mounted Edge Computing Model

For the partitioned sub-tasks offloaded to the UAVs, let $l_k^{\mathsf{u}}[n]$ denote the number of offloading bits from GU k to the associated UAV over time slot n. Besides, let us denote the computing resource of UAV m allocated to handle the sub-task offloaded from GU k in time slot n by $f_k^{\mathsf{u}}[n]$ (CPU cycles/second). Then, the task execution time at the associated UAV in time slot n can be computed as

$$T_k^{\mathsf{edc}}[n] = \frac{l_k^{\mathsf{u}}[n] c_k}{f_k^{\mathsf{u}}[n]}. \tag{7.7}$$

Moreover, the energy consumption for executing the offloaded sub-task from GU k at the associated UAV in time slot n can be calculated as

$$E_k^{\mathsf{edc}}[n] = l_k^{\mathsf{u}}[n] c_k e^{\mathsf{ed}}, \tag{7.8}$$

where e^{ed} denotes the energy consumption per CPU cycle of the UAV-mounted edge server [46]. Hence, the total energy consumption at the associated UAVs to process the offloading sub-task from GU k can be calculated as

$$E_k^{\mathsf{ed}} = \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] (l_k^{\mathsf{u}}[n] c_k e^{\mathsf{ed}}). \tag{7.9}$$

In addition, we assume that the communication links from the GUs to UAVs are dominated by the line-of-sight (LoS) propagation where the channel quality is mostly dependent on the UAV-GU distance. The distance between GU k and UAV m in time slot n can be calculated as $d_{k,m}[n]{=}\sqrt{H^2{+}\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}$. Moreover, the channel power gain from GU k to UAV m in time slot n is assumed to follow the free-space path loss model, which can be expressed as

$$g_{k,m}[n]{=}\rho_0 (d_{k,m}[n])^{-2}{=}\frac{\rho_0}{H^2+\left\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\right\|^2}, \tag{7.10}$$

where ρ_0 presents the channel power gain at the reference distance of 1 m. Hence, the achievable rate of the uplink transmission from GU k to the associated UAV m in time slot n, denoted by $R_{k,m}^{\mathsf{u}}[n]$ in bits/second (bps), can be expressed as

$$R_{k,m}^{\mathsf{u}}[n] = \beta_k^{\mathsf{u}}[n] \log_2\left(1+\frac{P_k^{\mathsf{u}} g_{k,m}[n]}{\beta_k^{\mathsf{u}}[n]\sigma^2}\right), \tag{7.11}$$

where $\beta_k^{\mathsf{u}}[n]$ and P_k^{u} represent the bandwidth allocated to GU k in time slot n and the transmit power of GU k for its uplink transmission, respectively, and σ^2 denotes the power density of the additive white Gaussian noise (AWGN) at the receiver. Then, the transmission time from GU k to UAV m in time slot n for the data related to the underlying offloaded sub-task can be expressed as

$$T_{k,m}^{\mathsf{edt}}[n] = \frac{l_k^{\mathsf{u}}[n]}{R_{k,m}^{\mathsf{u}}[n]}. \tag{7.12}$$

The energy consumption for the data transmission from GU k to UAV m in time slot n can be calculated as

$$E^{\mathsf{edt}}_{k,m}[n] = \frac{l^{\mathsf{u}}_k[n] P^{\mathsf{u}}_k}{R^{\mathsf{u}}_{k,m}[n]}. \tag{7.13}$$

Moreover, we assume that the partial task from GU k is offloaded and processed completely at each associated UAV in each time slot. Then we have following constraints

$$T^{\mathsf{ed}}_{k,m}[n]{=}\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n] c_k}{f^{\mathsf{u}}_k[n]}{+}\frac{l^{\mathsf{u}}_k[n]}{R^{\mathsf{u}}_{k,m}[n]}\right) \leq \Delta t, \forall k, m, n. \tag{7.14}$$

Then, the total processing time at the UAVs to serve GU k can be written as

$$T^{\mathsf{ed}}_k = \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} \theta_k[n] T^{\mathsf{ed}}_{k,m}[n]. \tag{7.15}$$

Furthermore, to remain in the air, each UAV consumes some energy during its hovering time. Specifically, the flying energy consumption of UAV m can be expressed as $E^{\mathsf{edf}}_m = P^{\mathsf{f}}_m T$, where P^{f}_m denotes the flying power of UAV m.

7.4.4.3 Satellite Cloud Computing Model

A cloud server deployed on each satellite typically has abundant computation resource to process offloaded computation sub-tasks. Therefore, we omit the processing time at the cloud server and we also ignore the cloud energy consumption involved in computation task execution and transmission of the computation results from the cloud server to GUs. Moreover, by using advanced communication such as multiple input multiple output (MIMO) communications, we assume GUs on the ground can communicate directly with the satellites [187]. For the considered scenario where the terrestrial network area is far away from the cloud server, we rely on multi-hop satellite communication to transmit data related to the offloaded sub-task from each GU to the cloud server.

In a recent work [2], an algorithm to determine the number of hops, i.e., the number of inter-satellite links (ISLs), and the corresponding satellites to establish the multi-hop communication path between two locations on the ground was proposed, i.e., see Algorithm 1 of [2]. By using this algorithm, we can determine the set of satellites involved in the transmission of the offloaded sub-task and we denote this satellite set by $\mathcal{S}' \subset \mathcal{S}$. The number of hops between the first and the last satellites connecting the considered terrestrial network area and the cloud server can be calculated as $L = |\mathcal{S}'| - 1$, where $|\mathcal{S}'|$ is the number of satellites in the set $\mathcal{S}'$. Because of the large coverage radius of each satellite, e.g., each Starlink satellite has the coverage radius of 580 km [207], we assume that all GUs are connected and offload their computation sub-tasks to the first satellite in $\mathcal{S}'$ while the last satellite in $\mathcal{S}'$ can be directly connected to the ground cloud server. Hence, the total data processing time and the propagation time from GU k to the cloud server can be calculated as

$$T_k^{\mathsf{cl}} = (1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}}) s_k \bigg(\frac{1}{R_k^{\mathsf{s}}} + \sum_{i=1}^{L} (\frac{1}{R_i^{\mathsf{ss}}}) + \frac{1}{R^{\mathsf{cl}}} \bigg) + T_k^{\mathsf{prop}}, \tag{7.16}$$

where $R_k^{\mathsf{s}}, R_i^{\mathsf{ss}}, R^{\mathsf{cl}}$ stand for the transmission rates between the GU k and the first satellite, between the satellites in the i-th hop, and between the last satellite and the cloud server, respectively. Here, T_k^{prop} represents the total propagation delay from GU k to the first satellite, between satellites over the L ISLs, and from the last satellite to the cloud server. Moreover, the energy consumption of GU k for transmitting the data related to the offloaded sub-task to the first satellite can be calculated as

$$E_k^{\mathsf{s}} = \frac{(1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}}) s_k P_k^{\mathsf{s}}}{R_k^{\mathsf{s}}}, \tag{7.17}$$

where P_k^{s} represents the transmission power of GU k to the satellite.

7.4.5 Problem Formulation

In this work, we are interested in minimizing the weighted energy consumption of all GUs and UAVs for all involved computation tasks, which can be expressed as

$$\begin{aligned} E^{\mathsf{sum}} \;\; = \;\; & \alpha_1 \bigg(\sum_{k \in \mathcal{K}} E_k^{\mathsf{ed}} + \sum_{m \in \mathcal{M}} P_m^{\mathsf{f}} T \bigg) \\ & + \alpha_2 \sum_{k \in \mathcal{K}} \bigg(E_k^{\mathsf{lo}} + \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] E_{k,m}^{\mathsf{edt}}[n] + E_k^{\mathsf{s}} \bigg), \end{aligned} \tag{7.18}$$

where $\alpha_1, \alpha_2 \in [0, 1]$ represent the weight factors of the energy consumption of the UAVs and GUs, respectively, which strike to balance the energy consumption between the UAVs and GUs.

For convenience, we gather different decision variables and define the corresponding groups of variables as follows: user scheduling $\mathbf{\Theta} = \{\theta_k[n], \forall k, n\}$, partial offloading control $\mathbf{\Lambda} = \{\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, \forall k\}$, bit allocation $\mathbf{L} = \{l_k^{\mathsf{u}}[n], \forall k, n\}$, bandwidth allocation $\boldsymbol{\beta} = \{\beta_k^{\mathsf{u}}[n], \forall k, n\}$, computation resource allocation $\mathbf{F} = \{f_k^{\mathsf{u}}[n], \forall k, n\}$, and UAV trajectory control $\mathbf{Q} = \{\mathbf{q}_m[n], \forall m, n\}$. Our design aims to minimize the weighted energy consumption of the GUs and UAVs while satisfying the maximum delay constraints of individual computation tasks. The optimization problem can be formulated as

$$\textbf{(P3):} \quad \min_{\mathbf{\Theta},\mathbf{\Lambda},\mathbf{L},\boldsymbol{\beta},\mathbf{F},\mathbf{Q}} \quad E^{\mathsf{sum}} \tag{7.19}$$

$$\text{s.t. } T_k^{\mathsf{lo}} \leq T_k^{\mathsf{max}}, \forall k, \tag{7.19a}$$

$$T_k^{\mathsf{cl}} \leq T_k^{\mathsf{max}}, \forall k, \tag{7.19b}$$

$$\sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] \leq T_k^{\mathsf{max}}, \forall k, \tag{7.19c}$$

$$\theta_k[n] T_{k,m}^{\mathsf{ed}}[n] \leq \Delta t, \forall k, m, n, \tag{7.19d}$$

$$\sum_{m \in \mathcal{M}} \sum_{t=0}^{N_k-1} \theta_k[n+t] \phi_{k,m}^{\mathsf{u}}[n+t] = N_k, \forall k, n \in \{1, ..., N-N_k\}, \tag{7.19e}$$

$$\sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] l_k^{\mathsf{u}}[n] = \lambda_k^{\mathsf{ed}} s_k, \forall k, \tag{7.19f}$$

$$\sum_{k \in \mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] \beta_k^{\mathsf{u}}[n] \leq W_m^{\mathsf{u}}, \forall m, n, \tag{7.19g}$$

$$\sum_{k \in \mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] f_k^{\mathsf{u}}[n] \leq F_m^{\mathsf{max}}, \forall m, n, \tag{7.19h}$$

$$\mathbf{q}_m[1] = \mathbf{q}_m[N], \forall m, \tag{7.19i}$$

$$\|\mathbf{q}_m[n+1] - \mathbf{q}_m[n]\|^2 \leq D_{\max}^2, \forall m, n = 1, ..., N-1, \tag{7.19j}$$

$$\|\mathbf{q}_m[n] - \mathbf{q}_j[n]\|^2 \geq d_{\min}^2, \ \forall n, m, j \neq m, \tag{7.19k}$$

$$\theta_k[n] \in \{0, 1\}, \forall k, n, \tag{7.19l}$$

$$0 \leq \lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}}, 1 - \lambda_k^{\mathsf{lo}} - \lambda_k^{\mathsf{ed}} \leq 1, \forall k, \tag{7.19m}$$

$$\beta_k^{\mathsf{u}}[n], f_k^{\mathsf{u}}[n], l_k^{\mathsf{u}}[n] \geq 0, \forall k, n, \tag{7.19n}$$

where constraints (7.19a)-(7.19d) capture the delay requirements for the GUs. Constraints (7.19e) and (7.19l) describe the binary user scheduling constraints for the GUs served by the associated UAVs. Constraints (7.19g) capture the bandwidth allocation for transmission between the GUs and UAVs while constraints (7.19h) present the UAVs' computation constraints where F_m^{max} denotes the maximum computation resource of UAV m. It can be seen that the objective and constraint functions (7.19a)-(7.19d) are non-linear and integer decision variables are involved in (7.19l) for the user scheduling. Hence, problem (7.19) is a non-convex mixed integer non-linear optimization problem (MINLP), which is difficult to solve optimally.

7.5 Proposed Algorithm

In the section, we develop an algorithm to solve the formulated problem when it is feasible. Specifically, we adopt the alternating optimization approach to solve problem (7.19) where we iteratively optimize each set of variables given the values of other variables in the corresponding sub-problems until convergence. We describe how to solve different sub-problems in the following.

7.5.1 Optimization of User Scheduling

Given $\{\mathbf{L}, \mathbf{\Lambda}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, the user scheduling sub-problem to optimize $\mathbf{\Theta}$ can be formulated as

$$\begin{aligned} \textbf{(P3.1):} \min_{\mathbf{\Theta}} \ & E^{\mathsf{sum}} \\ \text{s.t. } & \text{constraints } (7.19c) - (7.19h), (7.19l). \end{aligned} \tag{7.20}$$

It can be verified that problem (7.20) is a standard mixed integer linear program (MILP), which can be solved efficiently by using the CVX-Gurobi solver [38].

7.5.2 Optimization of Partial Offloading Control and Bit Allocation Over Time Slots

Given $\{\mathbf{\Theta}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, the sub-problem optimizing the partial offloading control and bit allocation $\{\mathbf{\Lambda}, \mathbf{L}\}$ can be formulated as

$$\begin{aligned} \textbf{(P3.2):} \ \min_{\mathbf{\Lambda}, \mathbf{L}} \ & E^{\mathsf{sum}} \\ \text{s.t. } \ & (7.19a) - (7.19d), (7.19f), (7.19m), (7.19n). \end{aligned} \tag{7.21}$$

It can be verified that problem (7.21) is a linear problem (LP), it can be solved by using the CVX-Gurobi solver [38].

7.5.3 Optimization of Computation Resource and Bandwidth Allocation

Given $\{\mathbf{\Theta}, \mathbf{\Lambda}, \mathbf{L}, \mathbf{Q}\}$, the sub-problem optimizing the computation resource and bandwidth allocation $\{\mathbf{F}, \boldsymbol{\beta}\}$ can be stated as

$$\begin{aligned}\textbf{(P3.3):}\ \min_{\mathbf{F},\boldsymbol{\beta}}\ &\alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{R^{\mathsf{u}}_{k,m}[n]}\Big)\\ &+E^{\mathsf{sum1}} \qquad (7.22)\\ \text{s.t. constraints } &(7.19\text{c}), (7.19\text{d}), (7.19\text{g}), (7.19\text{h}), (7.19\text{n}),\end{aligned}$$

where

$$\begin{aligned}E^{\mathsf{sum1}}=&\alpha_2\Big(\sum_{k\in\mathcal{K}}(\kappa\lambda^{\mathsf{lo}}_k s_k c_k(f_k)^2+\frac{(1-\lambda^{\mathsf{lo}}_k-\lambda^{\mathsf{ed}}_k)s_kP^{\mathsf{s}}_k}{R^{\mathsf{s}}_k})\Big)+\\ &\alpha_1\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n](l^{\mathsf{u}}_k[n]c_ke^{\mathsf{ed}})+\sum_m P^{\mathsf{f}}_mT\Big). \qquad (7.23)\end{aligned}$$

We first introduce auxiliary variables

$$\xi_{k,m}[n]=R^{\mathsf{u}}_{k,m}[n]=\beta^{\mathsf{u}}_k[n]\log_2\Big(1+\frac{B_{k,m}[n]}{\beta^{\mathsf{u}}_k[n]}\Big), \qquad (7.24)$$

where $B_{k,m}[n]=\frac{P^{\mathsf{u}}_k g_{k,m}[n]}{\sigma^2}$. Then, problem **(P3.3)** can be transformed to the following equivalent problem:

$$\begin{aligned}\textbf{(P3.3.1):}\ \min_{\mathbf{F},\boldsymbol{\beta},\mathbf{\Xi}}\ &\alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{\xi_{k,m}[n]}\Big)\\ &+E^{\mathsf{sum1}} \qquad (7.25)\\ \text{s.t. constraints } &(7.19\text{c}), (7.19\text{d}), (7.19\text{g}), (7.19\text{h}), (7.19\text{n}), (7.24),\end{aligned}$$

where $\mathbf{\Xi}=\{\xi_{k,m}[n], \forall k, m, n\}$.

It can be verified that $\beta_k^{\mathsf{u}}[n]\log_2(1+\frac{B_{k,m}[n]}{\beta_k^{\mathsf{u}}[n]})$ is a concave function with respect to $\beta_k^{\mathsf{u}}[n]$[1]. Using the successive convex approximation (SCA) method, the upper-bound for this concave function by using the first-order Taylor expansion at the given point $\beta_k^{\mathsf{u},r}[n]$ in the r-th iteration of the approximation process can be derived as

$$\begin{aligned}\beta_k^{\mathsf{u}}[n]\log_2\left(1+\frac{B_{k,m}[n]}{\beta_k^{\mathsf{u}}[n]}\right) \leq\ & \beta_k^{\mathsf{u},r}[n]\log_2\left(1+\frac{B_{k,m}[n]}{\beta_k^{\mathsf{u},r}[n]}\right)+ \\ & \left(\log_2(1+\frac{B_{k,m}[n]}{\beta_k^{\mathsf{u},r}[n]})-\frac{\log_2(e)B_{k,m}[n]}{B_{k,m}[n]+\beta_k^{\mathsf{u},r}[n]}\right)(\beta_k^{\mathsf{u}}[n]-\beta_k^{\mathsf{u},r}[n]) \\ & \triangleq R_{k,m}^{\mathsf{ub}}[n]. \end{aligned} \tag{7.26}$$

Then, problem (7.25) can be approximated by the following problem:

$$\textbf{(P3.3.2):} \min_{\mathbf{F},\beta,\mathbf{\Xi}} \quad \alpha_2\left(\sum_{k,m,n}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]l_k^{\mathsf{u}}[n]P_k^{\mathsf{u}}\frac{1}{\xi_{k,m}[n]}\right) + E^{\mathsf{sum1}} \tag{7.27}$$

$$\text{s.t.} \quad \sum_{m,n}\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}+\frac{l_k^{\mathsf{u}}[n]}{\xi_{k,m}[n]}\right)\leq T_k^{\mathsf{max}}, \forall k, \tag{7.27a}$$

$$\theta_k[n]\phi_{k,m}^{\mathsf{u}}[n]\left(\frac{l_k^{\mathsf{u}}[n]c_k}{f_k^{\mathsf{u}}[n]}+\frac{l_k^{\mathsf{u}}[n]}{\xi_{k,m}[n]}\right) \leq \Delta t, \forall k,m,n, \tag{7.27b}$$

$$\xi_{k,m}[n] \leq R_{k,m}^{\mathsf{ub}}[n], \forall k,m,n, \tag{7.27c}$$

$$\text{constraints (7.19g), (7.19h), (7.19n).}$$

Since $\frac{1}{f_k^{\mathsf{u}}[n]}$ and $\frac{1}{\xi_{k,m}[n]}$ are convex functions with respect to $f_k^{\mathsf{u}}[n]$ and $\xi_{k,m}[n]$, respectively, it can be seen that the objective function is convex and all constraints are linear. Hence, problem (7.27) is a convex problem, which can be solved effectively by using the CVX-Gurobi solver [38].

---

[1]The Hessian of function $f(x) = x\log_2(1+a/x)$ can be derived as $\triangledown^2 f(x) = -\frac{\log_2(e)a^2}{x(a+x)^2} < 0, \forall x, a > 0$. Therefore, $f(x)$ is a concave function with respect to x.

7.5.4 Optimization of Multi-UAV Trajectory

Given $\{\boldsymbol{\Theta}, \boldsymbol{\Lambda}, \mathbf{L}, \mathbf{F}, \boldsymbol{\beta}\}$, the sub-problem optimizing multi-UAV trajectory control variables $\mathbf{Q}$ can be formulated as

$$\begin{aligned}\textbf{(P3.4):}\ \min_{\mathbf{Q}}\ & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{R^{\mathsf{u}}_{k,m}[n]}\Big)\\ & + E^{\mathsf{sum1}} \qquad (7.28)\\ \text{s.t.}\ & \text{constraints (7.19c), (7.19d), (7.19i), (7.19j), (7.19k).}\end{aligned}$$

To approximate this problem, we introduce auxiliary variables $\gamma_{k,m}[n] = R^{\mathsf{u}}_{k,m}[n]$ and $S_{k,m}[n] \leq H^2 + \|\mathbf{q}_m[n] - \mathbf{r}^{\mathsf{u}}_k\|^2$ and we have

$$\begin{aligned}\gamma_{k,m}[n] &= \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2(H^2+\|\mathbf{q}_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2)}\right)\\ &\leq \beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2 S_{k,m}[n]}\right). \qquad (7.29)\end{aligned}$$

Then, problem **(P3.4)** can be reformulated as

$$\begin{aligned}\textbf{(P3.4.1):}\ \min_{\mathbf{Q},\boldsymbol{\Gamma},\mathbf{S}}\ & \alpha_2\Big(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{\gamma_{k,m}[n]}\Big)\\ & + E^{\mathsf{sum1}} \qquad (7.30)\\ \text{s.t.}\ & S_{k,m}[n] \leq H^2+\|\mathbf{q}_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2, \forall k,m,n, \qquad (7.30\text{a})\\ & \text{constraints (7.19c), (7.19d), (7.19i), (7.19j), (7.19k), (7.29),}\end{aligned}$$

where $\boldsymbol{\Gamma}=\{\gamma_{k,m}[n], \forall k,m,n\}$, $\mathbf{S} = \{S_{k,m}[n], \forall k,m,n\}$.

It can be verified that $\beta^{\mathsf{u}}_k[n]\log_2\left(1+\frac{R_k[n]}{S_{k,m}[n]}\right)$ is a convex function with respect to $S_{k,m}[n]$, where $R_k[n]=\frac{P^{\mathsf{u}}_k\rho_0}{\beta^{\mathsf{u}}_k[n]\sigma^2}$. By applying the SCA method, the lower-bound for the right hand side (RHS) of (7.29) derived by using the first-order Taylor expansion at the given point $S^r_{k,m}[n]$ in the r-th iteration of the approximation process can be expressed as

$$\beta_k^{\mathsf{u}}[n] \log_2 \left(1+\frac{R_k[n]}{S_{k,m}[n]}\right) \geq \beta_k^{\mathsf{u}}[n]\Bigg(\log_2 \left(S_{k,m}[n]+R_k[n]\right) - \log_2(S^r_{k,m}[n]) - \frac{\log_2(e)}{S^r_{k,m}[n]}\left(S_{k,m}[n]-S^r_{k,m}[n]\right)\Bigg) \triangleq R^{\mathsf{lb}}_{k,m}[n]. \tag{7.31}$$

Since $\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\|^2$ is a convex function with respect to $\mathbf{q}_m[n]$, we have the following inequality by applying the first-order Taylor expansion at the given point $\mathbf{q}^r_m[n]$:

$$\|\mathbf{q}_m[n]-\mathbf{r}_k^{\mathsf{u}}\|^2 \geq \|\mathbf{q}^r_m[n]-\mathbf{r}_k^{\mathsf{u}}\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{r}_k^{\mathsf{u}}\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}^r_m[n]\right), \forall k,m,n. \tag{7.32}$$

Furthermore, by applying the first-order Taylor expansion at the given point $\mathbf{q}^r_m[n]$ and $\mathbf{q}^r_j[n]$ to $\|\mathbf{q}_m[n]-\mathbf{q}_j[n]\|^2$, the left hand side (LHS) of constraints (7.19k) can be approximated as

$$\|\mathbf{q}_m[n]-\mathbf{q}_j[n]\|^2 \geq -\left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right), \forall j\neq m, n. \tag{7.33}$$

Therefore, the optimization problem (7.30) can be approximated by the following problem:

$$\textbf{(P3.4.2):} \min_{\mathbf{Q},\mathbf{\Gamma},\mathbf{S}} \ \alpha_2\Bigg(\sum_{k,m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]l^{\mathsf{u}}_k[n]P^{\mathsf{u}}_k\frac{1}{\gamma_{k,m}[n]}\Bigg) + E^{\mathsf{sum1}} \tag{7.34}$$

$$\text{s.t.} \ \sum_{m,n}\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\gamma_{k,m}[n]}\right) \leq T^{\mathsf{max}}_k, \forall k, \tag{7.34a}$$

$$\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\gamma_{k,m}[n]}\right) \leq \Delta t, \forall k,m,n, \tag{7.34b}$$

$$\gamma_{k,m}[n] \leq R^{\mathsf{lb}}_{k,m}[n], \forall k,m,n, \tag{7.34c}$$

$$S_{k,m}[n] \leq \|\mathbf{q}^r_m[n]-\mathbf{r}_k^{\mathsf{u}}\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{r}_k^{\mathsf{u}}\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}^r_m[n]\right) + H^2, \forall k,m,n, \tag{7.34d}$$

$$d^2_{\min} \leq -\left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 + 2\left(\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right), \forall j\neq m, n, \tag{7.34e}$$

$$\text{constraints (7.19i), (7.19j).}$$

Algorithm 7.1. Integrated User Scheduling, Partial Offloading, Computation, Bandwidth Allocation, and Multi-UAV Trajectory Control Algorithm

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, and locations of GUs, satellites and cloud server;
1: **Initialization:** $\mathbf{L}^0, \mathbf{\Lambda}^0, \mathbf{F}^0, \beta^0, \mathbf{Q}^0$;
Ensure: Min weighted energy consumption (E^{sum}); Let $r = 1$;
2: **repeat**
3: Solve sub-problem (7.20) to obtain $\mathbf{\Theta}^r$;
4: Solve sub-problem (7.21) to obtain $\mathbf{L}^r$ and $\mathbf{\Lambda}^r$;
5: Solve sub-problem (7.27) to obtain β^r and $\mathbf{F}^r$;
6: Solve sub-problem (7.34) to obtain $\mathbf{Q}^r$;
7: Update $r = r + 1$;
8: **until** Convergence
9: Return $E^{\mathsf{sum},*}, \mathbf{\Theta}^*, \mathbf{L}^*, \mathbf{\Lambda}^*, \mathbf{F}^*, \beta^*, \mathbf{Q}^*$.

Since $\frac{1}{\gamma_{k,m}[n]}$ is a convex function with respect to $\gamma_{k,m}[n]$, the objective function is convex. In addition, all constraints are linear. Therefore, problem (7.34) is a convex problem, which can be solved effectively by using the CVX-Gurobi solver [38].

7.5.5 Integrated User Scheduling, Partial Offloading Control, Computation Resource, Bandwidth Allocation, and Multi-UAV Trajectory Control Algorithm

Using the results above, we can develop an integrated algorithm based on the alternating optimization method as described in Algorithm 7.1. The convergence of this algorithm is stated in the following proposition.

Proposition 7.1. The proposed Algorithm 7.1 creates a sequence of feasible solutions where the objective value monotonically decreases over iterations. As a result, the algorithm converges to a feasible solution.

Proof. The proof is given in Appendix 7.9.1. □

7.6 Joint Admission Control and Network Management Design

If problem **(P3)** is feasible then Algorithm 7.1 converges to a feasible solution. However, problem **(P3)** can be infeasible in certain overloaded scenarios. To this end, we develop an algorithm to verify the feasibility of problem **(P3)** and propose a joint admission control and network management algorithm to tackle problem **(P3)** in a generic scenario where this problem can be feasible or infeasible.

7.6.1 Feasibility Verification

We address the feasibility verification for problem **(P3)** in this section. We introduce a new variable δ and use it to all inequality constraints of problem **(P3)** and consider a related optimization problem aiming to minimize δ. This feasibility verification problem can be formulated as

$$\textbf{(P3')}: \min_{\mathbf{\Theta},\mathbf{\Lambda},\mathbf{L},\beta,\mathbf{F},\mathbf{Q},\delta} \quad \delta \tag{7.35}$$

$$\text{s.t.} \quad T_k^{\mathsf{lo}} - T_k^{\mathsf{max}} - \delta \leq 0, \forall k, \tag{7.35a}$$

$$T_k^{\mathsf{cl}} - T_k^{\mathsf{max}} - \delta \leq 0, \forall k, \tag{7.35b}$$

$$\sum_{m\in\mathcal{M}}\sum_{n\in\mathcal{N}} \theta_k[n] T_{k,m}^{\mathsf{ed}}[n] - T_k^{\mathsf{max}} - \delta \leq 0, \forall k, \tag{7.35c}$$

$$\theta_k[n] T_{k,m}^{\mathsf{ed}}[n] - \Delta t - \delta \leq 0, \forall k, m, n, \tag{7.35d}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] \beta_k^{\mathsf{u}}[n] - W_m^{\mathsf{u}} - \delta \leq 0, \forall m, n, \tag{7.35e}$$

$$\sum_{k\in\mathcal{K}} \theta_k[n] \phi_{k,m}^{\mathsf{u}}[n] f_k^{\mathsf{u}}[n] - F_m^{\mathsf{max}} - \delta \leq 0, \forall m, n, \tag{7.35f}$$

$$\|\mathbf{q}_m[n+1]-\mathbf{q}_m[n]\|^2 - D_{\max}^2 - \delta \leq 0, \forall m, n = 1, ..., N-1, \tag{7.35g}$$

$$d_{\min}^2 - \|\mathbf{q}_m[n]-\mathbf{q}_j[n]\|^2 - \delta \leq 0, \ \forall n, m, j \neq m, \tag{7.35h}$$

$$\text{constraints (7.19e), (7.19f), (7.19i), (7.19l), (7.19m), (7.19n).}$$

Note that this problem is feasible and there exists an optimal value of δ that can be used to determine the feasibility of problem **(P3)** as follows. Specifically, problem **(P3)** is feasible, i.e., all constraints are satisfied, if $\delta \leq 0$ and it is infeasible, otherwise. However, problem **(P3')** is also a mixed integer non-linear optimization problem (MINLP), which is difficult to solve optimally. Using the alternating optimization method again, we can solve problem **(P3')** efficiently as described in the following.

7.6.1.1 Optimization of User Scheduling

Given $\{\mathbf{L}, \mathbf{\Lambda}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, the sub-problem optimizing user scheduling $\mathbf{\Theta}$ can be formulated as

$$\textbf{(P3.1'):} \min_{\delta, \mathbf{\Theta}} \ \delta \tag{7.36}$$
$$\text{s.t. constraints } (7.35a) - (7.35h), (7.19l).$$

Problem (7.36) is a standard mixed integer linear program (MILP), which can be solved efficiently by using the CVX-Gurobi solver [38].

7.6.1.2 Optimization of Partial Offloading Control and Bit Allocation Over Time Slots

Given $\{\mathbf{\Theta}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}\}$, the sub-problem optimizing partial offloading control and bit allocation $\{\mathbf{\Lambda}, \mathbf{L}\}$ can be formulated as

$$\textbf{(P3.2'):} \ \min_{\delta, \mathbf{\Lambda}, \mathbf{L}} \ \delta \tag{7.37}$$
$$\text{s.t. } \ (7.35a) - (7.35h), (7.19f), (7.19m), (7.19n).$$

Problem (7.37) is a linear problem (LP), it can be solved by using the CVX-Gurobi solver [38].

7.6.1.3 Optimization of Computation Resource and Bandwidth Allocation

Given $\{\mathbf{\Theta}, \mathbf{\Lambda}, \mathbf{L}, \mathbf{Q}\}$, the sub-problem optimizing computation resource and bandwidth allocation $\{\mathbf{F}, \boldsymbol{\beta}\}$ can be approximated as

$$\textbf{(P3.3')}: \min_{\delta,\mathbf{F},\boldsymbol{\beta},\mathbf{\Xi}} \ \delta \tag{7.38}$$

$$\text{s.t.} \ \sum_{m,n} \theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\xi_{k,m}[n]}\right)-T^{\mathsf{max}}_k-\delta \le 0, \forall k, \tag{7.38a}$$

$$\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\xi_{k,m}[n]}\right) - \Delta t - \delta \le 0, \forall k, m, n, \tag{7.38b}$$

$$\xi_{k,m}[n] - R^{\mathsf{ub}}_{k,m}[n] \le 0, \forall k, m, n, \tag{7.38c}$$

$$\text{constraints (7.35a), (7.35b), (7.35e)} - \text{(7.35h), (7.19n).}$$

Problem (7.38) is a convex problem, which can be solved effectively by using the CVX-Gurobi solver [38].

7.6.1.4 Optimization of Multi-UAV Trajectory Control

Given $\{\mathbf{\Theta}, \mathbf{\Lambda}, \mathbf{L}, \mathbf{F}, \boldsymbol{\beta}\}$, the sub-problem optimizing multi-UAV trajectory control $\mathbf{Q}$ can be approximated as

$$\textbf{(P3.4')}: \min_{\delta,\mathbf{Q},\mathbf{\Gamma},\mathbf{S}} \ \delta \tag{7.39}$$

$$\text{s.t.} \ \sum_{m,n} \theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\gamma_{k,m}[n]}\right)-T^{\mathsf{max}}_k-\delta\le 0, \forall k, \tag{7.39a}$$

$$\theta_k[n]\phi^{\mathsf{u}}_{k,m}[n]\left(\frac{l^{\mathsf{u}}_k[n]c_k}{f^{\mathsf{u}}_k[n]}+\frac{l^{\mathsf{u}}_k[n]}{\gamma_{k,m}[n]}\right) - \Delta t - \delta \le 0, \forall k, m, n, \tag{7.39b}$$

$$\gamma_{k,m}[n] - R^{\mathsf{lb}}_{k,m}[n] \le 0, \forall k, m, n, \tag{7.39c}$$

$$S_{k,m}[n] - \|\mathbf{q}^r_m[n]-\mathbf{r}^{\mathsf{u}}_k\|^2 - 2\left(\mathbf{q}^r_m[n]-\mathbf{r}^{\mathsf{u}}_k\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}^r_m[n]\right) - H^2 \le 0, \forall k, m, n, \tag{7.39d}$$

$$d^2_{\min} + \left\|\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right\|^2 - 2\left(\mathbf{q}^r_m[n]-\mathbf{q}^r_j[n]\right)^T\left(\mathbf{q}_m[n]-\mathbf{q}_j[n]\right)-\delta \le 0, \forall j\neq m, n, \tag{7.39e}$$

$$\text{constraints (7.19i), (7.35a), (7.35b), (7.35e)} - \text{(7.35g).}$$

Problem (7.39) is a convex problem, which can be solved effectively by using the CVX-Gurobi solver [38].

7.6.1.5 Feasibility Verification Algorithm

Summary of the feasibility verification algorithm is given in Algorithm 7.2. Initially, we set *feasibility* = *true*, and initialize all variables $\boldsymbol{\omega}^0 = \{\boldsymbol{\Theta}^0, \mathbf{L}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0\}$. Then, we apply the alternating optimization method and iteratively solve each set of variables given the values of other variables until convergence to a stable value of δ^* as described from step 1 to step 19. Specifically, after solving each sub-problem in the r-th iteration, we check the obtained objective value δ^r as follows: if $\delta^r > 0$, we return this value and $\boldsymbol{\omega}^r$ and then break the *"repeat-until"* loop. Otherwise, if $\delta^r \leq 0$, we continue solving next sub-problem. Steps 20 to 24 check the obtained value of δ^* and output the *feasibility result* in step 25. If *feasibility* = *true*, all constraints of problem **(P3)** are satisfied and we can solve the considered optimization problem to obtain a feasible solution. Otherwise, if *feasibility* = *false*, problem **(P3)** is infeasible, i.e., certain constraints of problem **(P3)** cannot be satisfied. The outputs of Algorithm 7.2 are *feasibility* results and $\boldsymbol{\omega}^*$.

7.6.2 Admission Control and Network Management Algorithm

We have described an integrated algorithm that can be applied as the considered problem is feasible and feasibility verification algorithm to check the feasibility of problem **(P3)**. When the problem is infeasible, i.e., the output of the proposed feasibility verification algorithm is *feasibility* = *false*, it becomes necessary to perform admission control to ensure all constraints are satisfied. To this end, we design a user removal strategy that attempts to remove the smallest number of GUs while ensuring all constraints be satisfied for the system with the remaining GUs, i.e., the removed GUs are temporarily denied services so they do not utilize any network resources.

For problem **(P3)**, the maximum delay constraints and the constraints requiring partial tasks from GUs be offloaded and processed completely at the associated UAVs in each time slot are challenging ones to satisfy. We propose a user removal strategy that iteratively removes in each removal step one "worst" GU that requires the largest amount of resource to satisfy its stringent delay constraint. Specifically, given the output of the Algorithm 7.2, the total data transmission,

Algorithm 7.2. Feasibility Verification Algorithm

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, and locations of GUs, satellites and cloud server;
Ensure: Min δ; Let $r = 1$; *feasibility* $=$ *true*, and $\boldsymbol{\omega}^0 = \{\boldsymbol{\Theta}^0, \mathbf{L}^0, \boldsymbol{\Lambda}^0, \mathbf{F}^0, \boldsymbol{\beta}^0, \mathbf{Q}^0\}$;

repeat
  Solve sub-problem (7.36) to obtain δ^r and $\boldsymbol{\Theta}^r$;
  if $\delta^r > 0$ **then**
    Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^{r-1}, \boldsymbol{\Lambda}^{r-1}, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}$,
    Break **repeat** loop;
  end if
  Solve sub-problem (7.37) to obtain δ^r, $\mathbf{L}^r$, and $\boldsymbol{\Lambda}^r$;
  if $\delta^r > 0$ **then**
    Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}$,
    Break **repeat** loop;
  end if
  Solve sub-problem (7.38) to obtain δ^r, $\boldsymbol{\beta}^r$, and $\mathbf{F}^r$;
  if $\delta^r > 0$ **then**
    Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^{r-1}\}$,
    Break **repeat** loop;
  end if
  Solve sub-problem (7.39) to obtain δ^r and $\mathbf{Q}^r$;
  if $\delta^r > 0$ **then**
    Return δ^r and $\boldsymbol{\omega}^r = \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r\}$,
    Break **repeat** loop;
  end if
  Update $r = r + 1$;
until Convergence δ^*
if $\delta^* \leq 0$ or $\delta^r \leq 0$ **then**
  feasibility $=$ *true*;
else
  feasibility $=$ *false*;
end if
Output *feasibility result* and

$$\boldsymbol{\omega}^* \in \Big\{ \{\boldsymbol{\Theta}^r, \mathbf{L}^{r-1}, \boldsymbol{\Lambda}^{r-1}, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^{r-1}, \boldsymbol{\beta}^{r-1}, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^{r-1}\}, \{\boldsymbol{\Theta}^r, \mathbf{L}^r, \boldsymbol{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r\} \Big\};$$

propagation, and task processing time for each GU could be calculated as in (7.1). Then, we find the "worst" GU k that achieves the maximum value of T_k/T_k^{max}, remove it and update the set of remaining GUs $\mathcal{K}^{\mathsf{ac}}$ accordingly, i.e., removing the identified GU k from the set $\mathcal{K}$.

We propose a joint admission control and network management algorithm to solve problem **(P3)** to achieve the minimum weighted energy consumption of the GUs and UAVs as in Algorithm 7.3. First, we initialize all variables except the user scheduling variables in step 1. In step 2, using the algorithm in [2], we can determine the satellite hop-count L for the considered network setting, while the bandwidth allocated to each UAV W_m^{u} is determined in step 3. Moreover, we initialize the variable *feasibility* $=$ *true* as in step 4. The main steps of the algorithm are described from step 5 to step 17.

Algorithm 7.3. Joint Admission Control and Network Management Algorithm

Require: $\mathcal{M}$, $\mathcal{K}$, W, T, and locations of GUs, satellites and cloud server;
1: **Initialization:** $\mathbf{L}^0, \mathbf{\Lambda}^0, \mathbf{F}^0, \beta^0, \mathbf{Q}^0, \mathcal{K}^{\mathsf{ac}} = \mathcal{K}$;
Ensure: Min weighted energy consumption (E^{sum});
2: Find number of hop-count L by running the Alg. 1 in [2];
3: Determine the bandwidth W_m^{u}, i.e., $\sum_{m\in\mathcal{M}} W_m^{\mathsf{u}} \leq W$;
4: Let *feasibility* $=$ *true*;
5: **repeat**
6: Run feasibility verification Algorithm 7.2;
7: **if** *feasibility* $=$ *true* **then**
8: Run Algorithm 7.1
9: Break **repeat** loop;
10: **else**
11: Given $\boldsymbol{\omega}^*$ obtained from Algorithm 7.2, calculate T_k based on (7.1);
12: Find the worst GU $k = \mathrm{argmax}_{k\in\mathcal{K}} \left(T_k/T_k^{\mathsf{max}}\right)$;
13: Assign $\mathcal{K}^{\mathsf{ac}} = \mathcal{K} \setminus \{k\}$;
14: Update $\mathcal{K} \leftarrow \mathcal{K}^{\mathsf{ac}}$;
15: **end if**
16: **until** $\mathcal{K} = \emptyset$
17: Return $E^{\mathsf{sum},*}, \mathcal{K}^{\mathsf{ac}}, \mathbf{\Theta}^*, \mathbf{L}^*, \mathbf{\Lambda}^*, \mathbf{F}^*, \beta^*, \mathbf{Q}^*$.

Specifically, we run feasibility verification Algorithm 7.2 in step 6 in each iteration of the *"repeat-until"* loop. If Algorithm 7.2 outputs *feasibility* $=$ *true*, which means all constraints of the problem are satisfied, we run the integrated Algorithm 7.1 to obtain a feasible solution in step 8 then we break the loop in step 9 to end the whole algorithm. In contrast, if Algorithm 7.2 outputs *feasibility* $=$ *false* meaning certain constraints are not satisfied, we perform user removal to remove one GU and update the new set of GUs $\mathcal{K}$ as described in steps 11-14. The whole algorithm terminates when a feasible solution is achieved or when the set of GUs is empty meaning the network is not capable of serving even a single GU.

7.6.3 Complexity Analysis

We now analyze the complexity of our proposed algorithms in the number of required arithmetic operations. In Algorithm 7.3, the most complex operations are related to optimization steps required to solving different sub-problems in the feasibility verification and integrated algorithms in step 6 and step 8. First, in executing the feasibility verification Algorithm 7.2 in step 6, the CVX [38] is used to solve different underlying sub-problems. Since this solver employs the interior-point method, the complexity involved in each step is $\mathcal{O}(m_1^{1/2}(m_1+m_2)m_2^2)$, where m_1 is the number of inequality constraints, m_2 denotes the number of variables [195], and $\mathcal{O}$ denotes the big-O notation. Hence, the complexity of this algorithm is $\mathcal{O}(L_1 MNK^{\frac{7}{2}})$, where L_1 is the number of iterations required to achieve the convergence for Algorithm 7.2. In Step 8 of Algorithm 7.3, the integrated Algorithm. 7.1

is executed and the CVX solver is used again to solve different underlying sub-problems. The complexity involved in this step is $\mathcal{O}(L_2 MNK^{\frac{7}{2}})$, where L_2 is the number of iterations required to achieve the convergence for Algorithm. 7.1 . In step 12, we search the "worst" GU for removal and this search has the complexity of $\mathcal{O}(K)$. Let L denote the number of iterations required before Algorithm 7.3 terminates. Therefore, the overall computation complexity of the proposed algorithm is $\mathcal{O}(L(L_1 + L_2)MNK^{\frac{7}{2}})$.

7.6.4 Algorithm Initialization

We describe the initialization of UAV trajectory and other variables in the following.

7.6.4.1 Initial Circular UAV Trajectory

We describe how to set the initial circular UAV trajectory assuming that each UAV serves a circular network partition, i.e., each UAV serves a cluster of GUs located inside a circle with radius r_{c} (m). For each network partition served by UAV m, we need to determine its center $\mathbf{c}_m^{\mathsf{init}} = (x_m^{\mathsf{init}}, y_m^{\mathsf{init}})$ and radius d_m^{init}. Given the locations of K GUs, we employ the k-means clustering algorithm [196] to determine the centers of the corresponding network partitions. The radius of the initial circular network partition served by UAV m is given by $d_m^{\mathsf{init}} = \min\left(d_m^{\mathsf{max1}}, r_{\mathsf{c}}\right)$. It is assumed that each UAV must remain inside the area of its network partition during the flight period and d_m^{max1} is the maximum radius of the circular UAV trajectory with the same starting and ending point.

To determine d_m^{max1}, we approximate the largest circumference of a circle as the maximum distance, denoted as $D = V_{\mathsf{max}}T$, that the UAV can travel during the flight period. Therefore, we have $d_m^{\mathsf{max1}} \approx D/2\pi$. Let $\phi_n \triangleq 2\pi\frac{n-1}{N-1}, \forall n$, we can initialize $\mathbf{Q}^0 = \{\mathbf{q}_m^0[n], \forall m, n\}$ as follows:

$$\mathbf{q}_m^0[n] = \left(x_m^{\mathsf{init}} + d_m^{\mathsf{init}} \cos\phi_n, y_m^{\mathsf{init}} + d_m^{\mathsf{init}} \sin\phi_n\right), \ \forall m, n. \tag{7.40}$$

For the multi-UAV system, d_{min} denotes the minimum inter-UAV distance to ensure collision avoidance, which must be maintained as we initialize the UAVs' trajectories. Fig. 7.2 illustrates the initial circular trajectories for the network with 2 UAVs.

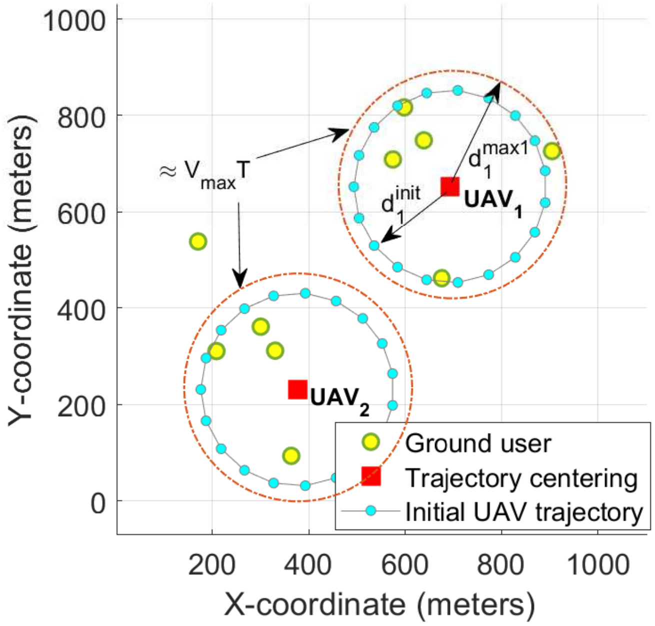

Figure 7.2 – Initial UAVs' trajectories.

7.6.4.2 Initial Partial Offloading Control, Computation Resource, and Bandwidth Allocation Variables

The task size values s_k are set randomly in range of $[1, 10]$Mbits and the values of the maximum tolerable delay T_k^{max} are also set randomly in range of $[1, 3]$(seconds). Moreover, the initial values of partial offloading control variables are randomly generated in $\lambda_k^{\mathsf{lo}}, \lambda_k^{\mathsf{ed}} \in [0, 0.5]$, and a uniform allocation of bit, computation resource, and bandwidth is applied, i.e., $l_k^{\mathsf{u}}[n] = \lambda_k^{\mathsf{ed}} s_k/N_k$, $f_k^{\mathsf{u}}[n] = MF_m^{\mathsf{max}}/K$, and $\beta_k^{\mathsf{u}}[n] = W/K$, respectively.

To investigate the effectiveness of the proposed algorithms, we consider the following baselines. In an "early scheduling" baseline, all GUs are scheduled continuously from the first time slot of the UAV flight period. In the second baseline, called "baseline edge", we initially set circular UAVs' trajectories to serve the corresponding groups of GUs, the values of partial offloading control variables are randomly set and a uniform allocation of bit, computation resource, and bandwidth is applied as described above. For comparison, the "optimized edge" strategy represents our proposed design where all variables are optimized.

Table 7.3 – Simulation parameters

Parameter	Description	Value
M	Number of UAVs	$[2, 3]$
K	Number of GUs	$[10, 12, 14, 16, 18]$
T	Flight period	$[10, 15]$ s
Δt	Length of each time slot	0.5 s
H	Altitude of UAVs	100 m
r_c	Radius of cluster	200 m
P_k^{u}	Transmit power of GU with UAV	24 dBm
P_k^{s}	Transmit power of GU with satellite	30 dBm
P_m^{f}	Flying power consumption of UAV	33 dBm
c_k	Computation resource required for 1-bit input data (CPU cycles/bit)	300
f_k	Computation resource of GU (CPU cycles/s)	2×10^8
F_m^{max}	Maximum computation resource of UAV (CPU cycles/s)	3×10^9
σ^2	Noise power	-174 dBm
κ	Effective switched capacitance depending on the chip architecture	10^{-28}
e^{ed}	Energy consumption per CPU cycle at the UAV-mounted edge server	1 W/GHz
R_k^{s}	Transmission rate from GU to satellite	10 Mbps
R_i^{ss}	Transmission rate inter-satellites	100 Mbps
R^{cl}	Transmission rate from satellite to cloud server	10 Mbps
T_k^{prop}	Total propagation delay from GU to cloud server	100 ms
d_{min}	Minimum inter-UAV distance	20 m
$[\alpha_1, \alpha_2]$	Weighted factors of energy consumption of UAVs and GUs	$[0.2, 0.8]$
V_{max}	Maximum speed of UAV	50 m/s
W	Total available bandwidth	$[10 - 50]$ MHz

7.7 Numerical Results

We consider different scenarios in which a cloud server is far away from a considered network area. For particular, the group of GUs is located in Montreal (45.50°N, 73.56°W) while the cloud server is located in Vancouver (49.28°N, 123.12°W). By running the Alg. 1 in [2], we can determine the number of satellite hops $L = 4$. The parameters for our simulations are set similarly to those in [44, 46–48] and the chosen values of key parameters are summarized in Table 7.3.

We first study a specific setting with 10 GUs located on the ground and their task size values are set as $s_k = [6, 10, 5, 4, 3, 5, 6, 8, 10, 10]$Mbits with the corresponding values of maximum tolerable delay of $T_k^{\mathsf{max}} = [2, 3, 2, 2, 1, 3, 2, 3, 3, 3]$(seconds). Fig. 7.3 illustrates the optimized trajectories of UAVs in the scenarios with 10 GUs, $T = [10, 15]$s, and $L = 4$. This figure shows that the UAVs must follow longer trajectories to account for the longer flight time as T increases, and the optimized trajectories are smoother for larger T. In Fig. 7.4, we show the scheduled time slots for individual GUs during the UAV flight periods of $T = [10, 15]$s and $L = 4$. It can be seen that GUs are mostly

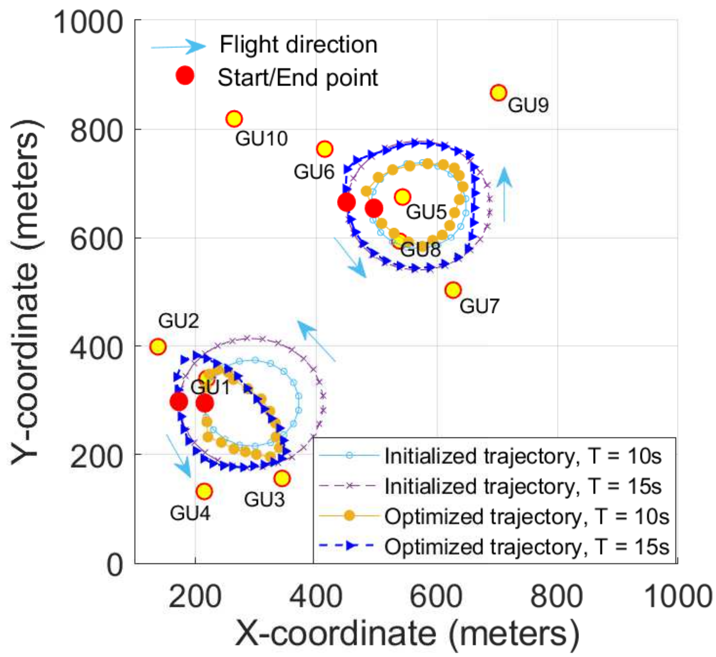

Figure 7.3 – Optimized UAV trajectories with different flight time T.

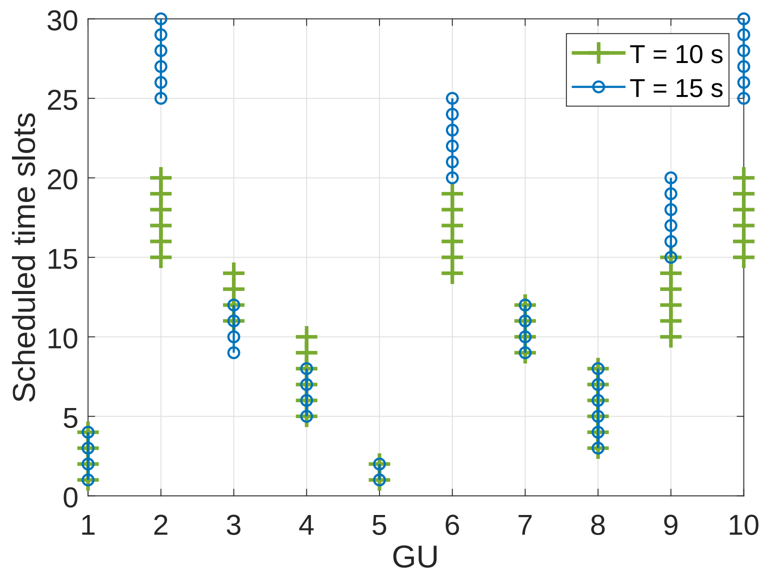

Figure 7.4 – Scheduled time slots for different GUs

scheduled when the GUs are physically closer to their associated UAVs. This is reasonable because smaller GU-UAV distance leads to a higher channel power gain and transmission rate that reduce the communication latency for the offloaded data of underlying computation tasks.

We now evaluate the effectiveness of the proposed algorithm in comparison with the baselines. The following results are obtained for random task size s_k and maximum delay requirements T_k^{max} as described in Section. 7.6.4.2. Moreover, the results are obtained by averaging over 100 simulation

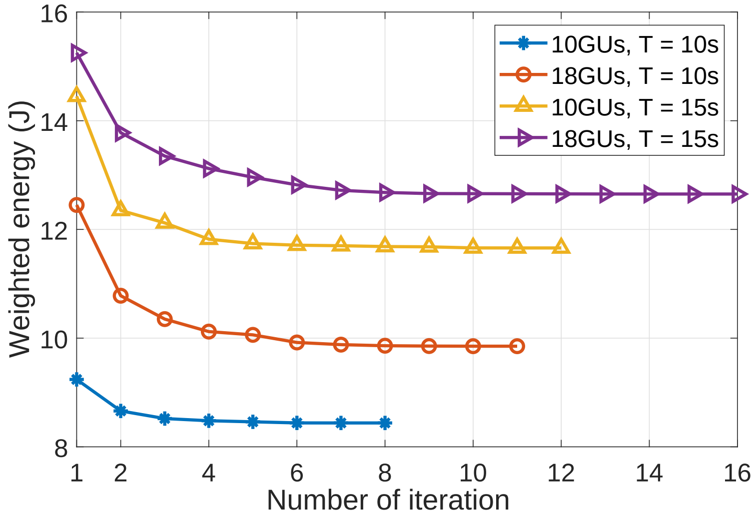

Figure 7.5 – Convergence of the proposed integrated algorithm.

runs, each with different random locations of the GUs and initial parameter settings. The convergence of the integrated Algorithm 7.1 is illustrated in Fig. 7.5 for the cases with 2 UAVs, $L = 4$, $W = 10$ MHz, 10 and 18 GUs, and $T = [10, 15]$s. In particular, we optimize user scheduling, partial offloading control and bit allocation, computation and bandwidth allocation, and UAV trajectory control until convergence in each iteration of the proposed algorithm. It can be seen that the value of the objective function decreases over iterations. Our algorithm converges more slowly to a feasible solution with the larger flight period T. This is because the optimization space becomes larger with larger flight period T. Moreover, the number of iterations increases as the number of GUs increases due to increasing computation complexity.

Fig. 7.6 illustrates the weighted sum of energy for different number of GUs, 2 UAVs, $W = 10$ MHz, $L = 4$, and $T = [10, 15]$s. It can be seen that the weighted sum of energy becomes higher with larger number of GUs and the proposed algorithm achieves the smallest weighted sum of energy compared to those due to other baselines in both scenarios with $T = [10, 15]$s. For 18 GUs, the weighted sum of energy can be reduced by 18.05% and 9.64% compared to the corresponding values due to the "early scheduling" and "baseline edge" baselines with $T = 10$s and $T = 15$s, respectively.

Fig. 7.7 shows the weighted sum of energy for different values of bandwidth allocated for the uplink communications of the GUs. These results are obtained by running the proposed algorithm for the scenarios with 2 and 3 UAVs, 10 GUs, $W = 10$ MHz, $L = 4$, and $T = [10, 15]$s. This figure shows that the weighted sum of energy decreases as the total bandwidth increases. This is because larger bandwidth allocated for GUs can help increase the transmission rates that allow

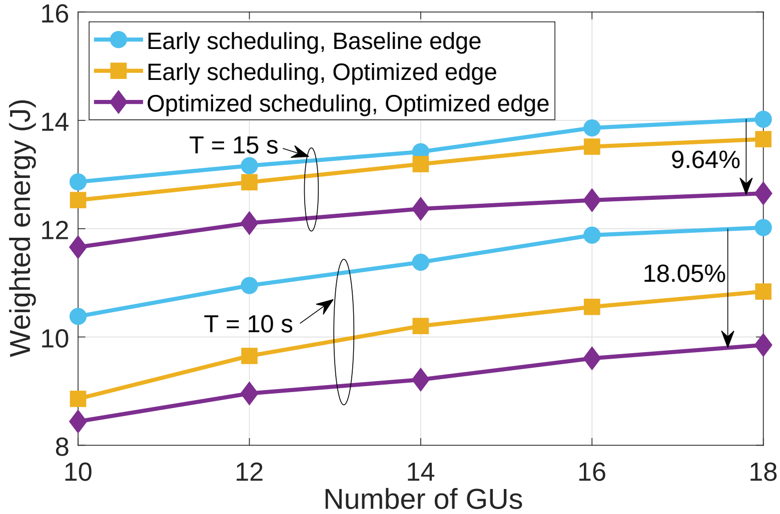

Figure 7.6 – Weighted sum of energy for different number of GUs.

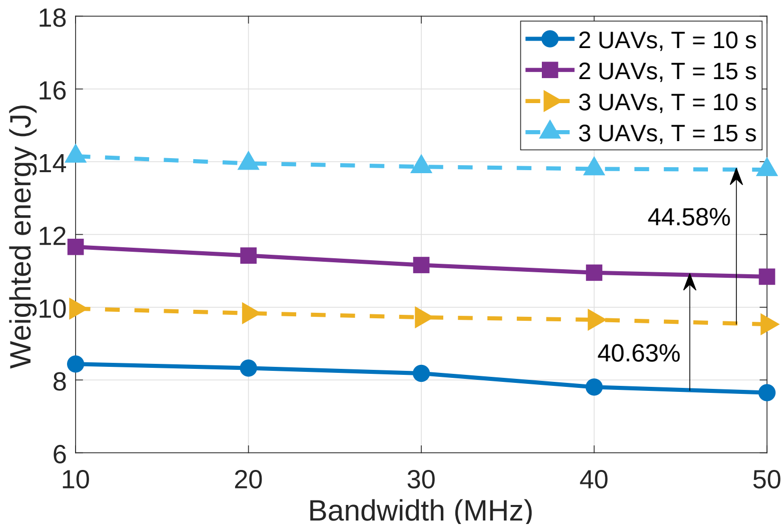

Figure 7.7 – Weighted sum of energy for different bandwidth values.

larger fractions of computation load to be offloaded and processed at the more resourceful edge servers achieving energy saving for both the edge servers and GUs. It can also be seen that for given UAV flight time, the network with 3 UAVs achieves higher weighted sum of energy compared to that due to the network with 2 UAVs. Besides, the difference in the weighted sum of energy for $T = 15$s and $T = 10$s is larger for the network setting with 3 UAVs compared to that with 2 UAVs, which are 44.58% and 40.63%, respectively.

Fig. 7.8 and Fig. 7.9 illustrate the weighted sum of energy for the different satellite hop counts, UAV flight periods, and number of UAVs. In particular, in Fig. 7.8, we show the weighted sum of energy for scenarios with 2 UAVs, 10 GUs, $T = [10, 15]$s, $W = 10$ MHz, and different satellite

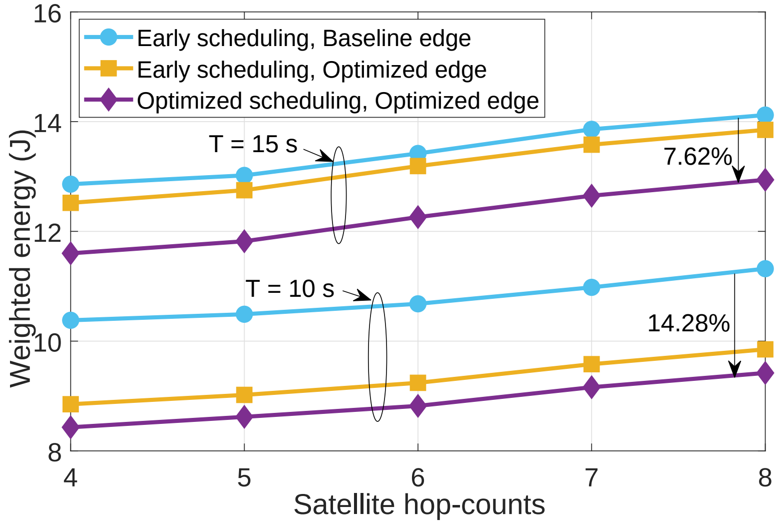

Figure 7.8 – Weighted sum of energy for different number of satellite hops and flight time.

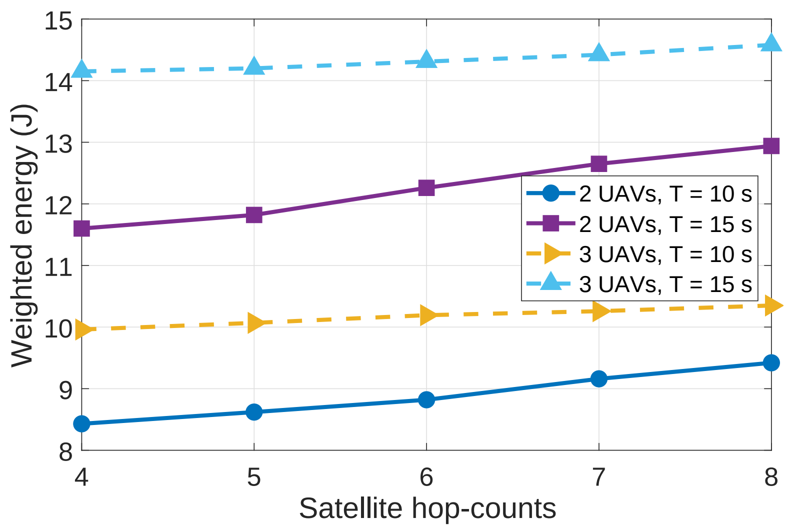

Figure 7.9 – Weighted sum of energy for different numbers of satellite hops and UAVs.

hop counts L in the space network layer. It can be seen that the weighted sum of energy increases with the number of satellite hop counts. This is because larger satellite hop counts imply longer transmission and propagation delay in the space network layer, which could force higher computation load to be processed at the GUs and edge servers. In addition, the weighted sum of energy also increases and the proposed algorithm can save less energy compared to those required by the "early scheduling" and "baseline edge" baselines for the longer flight time T. Specifically, the weighted sum of energy can be reduced by 14.28% and 7.62% compared to those due to the "early scheduling" and "baseline edge" strategies with $T = 10$s and $T = 15$s, respectively.

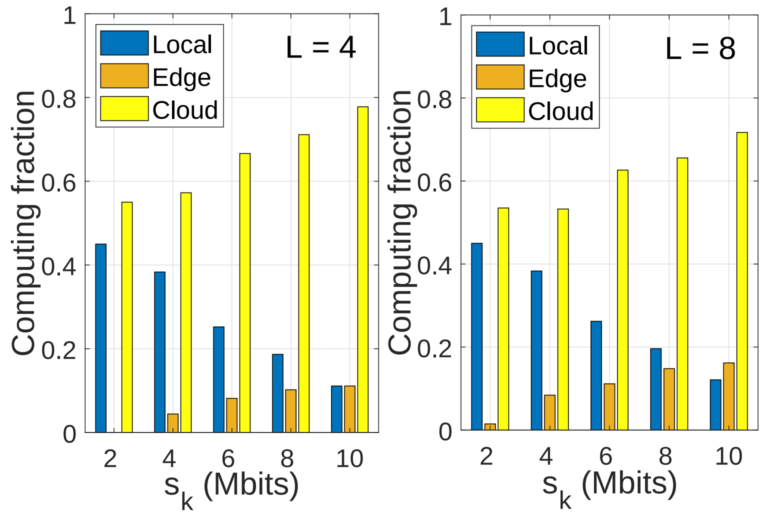

Figure 7.10 – Computation distribution for different task size values.

In Fig. 7.9, we illustrate the weighted sum of energy for different number of satellite hops L and UAVs. These results are obtained by running the proposed algorithm for the scenarios with 2 and 3 UAVs, 10 GUs, $W = 10$ MHz, and $T = [10, 15]$s. It can be seen that the weighted sum of energy increases with the number of satellite hops. However, the weighted sum of energy increases more sharply with the number of satellite hops for the network with 2 UAVs compared to that with 3 UAVs. This result implies that the network with 3 UAVs can save more energy compared to that with 2 UAVs for large satellite hop counts. As discussed above, with the larger satellite hop counts, higher computation load would be processed at the edge servers; therefore, more energy efficiency can be achieved with larger number of UAVs, i.e., edge servers. Fig. 7.10 illustrates the computation load distribution over network layers. This figure shows that larger task size values lead to less computation load distributed at the GUs while larger satellite hop counts result in higher computation load to be processed at the edge servers.

To evaluate the performance achieved by the proposed admission control design, we define an admission ratio as the ratio between the number of actual GUs served to the total number of GUs, i.e., $\frac{|\mathcal{K}^{\mathsf{ac}}|}{|\mathcal{K}_0|}$, where $\mathcal{K}_0$ denotes the set of original GUs and $\mathcal{K}^{\mathsf{ac}}$ represents the set of GUs admitted for which a feasible solution can be found by the proposed algorithm. Fig. 7.11 shows the admission ratio for different number of GUs for the networks with 2 and 3 UAVs, $W = 10$ MHz, $L = 4$, and $T = [10, 15]$s. It can be seen that the admission ratio decreases as the number of GUs increases. This is because given the fixed radio and computation resources, the number of GUs that the network can support is limited. Hence, a larger number of GUs would be removed from the system as the

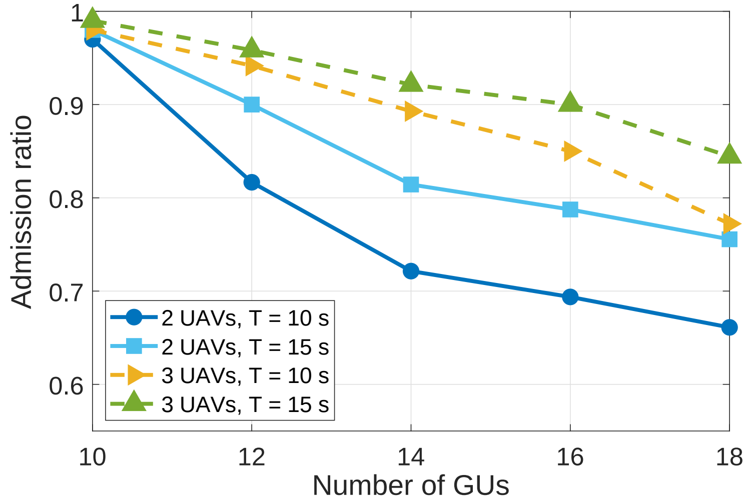

Figure 7.11 – Admission ratio for different number of GUs.

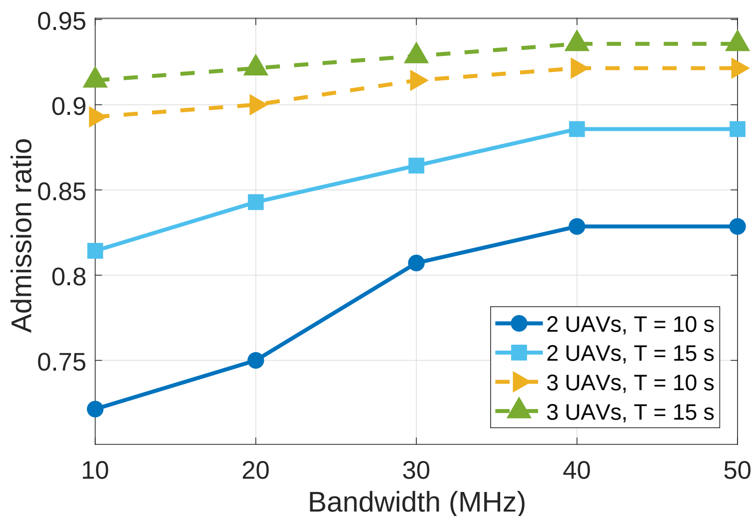

Figure 7.12 – Admission ratio for different bandwidth values.

number of GUs increases resulting in a decreasing admission ratio. It can also be seen that the difference in the admission ratios for the two scenarios with $T = 10$s and $T = 15$s is larger for 2 UAVs compared to that for 3 UAVs. In fact, for the network setting with a larger number of UAVs, i.e., edge servers, and larger UAV flight period T, the network can be covered better; therefore, a larger number of GUs can be served.

We present the variation of the admission ratio with system bandwidth for the networks with 2 and 3 UAVs, 14 GUs, $L = 4$, and $T = [10, 15]$s in Fig. 7.12. This figure shows that the admission ratio increases with larger system bandwidth. This is reasonable because larger system bandwidth allows to enhance transmission rates of GUs or equivalently more GUs can be supported while still

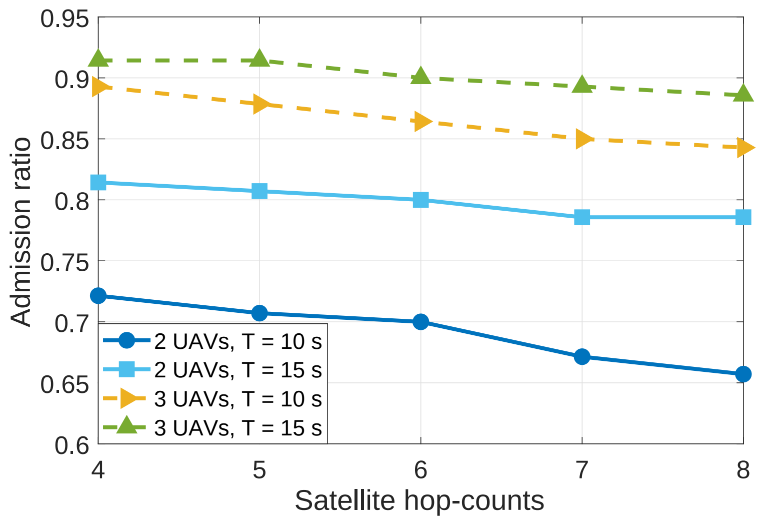

Figure 7.13 – Admission ratio for different number of satellite hops.

satisfying all GUs' and system constraints. Meanwhile, the admission ratio slightly decreases as the number of satellite hops increases as illustrated in Fig. 7.13 that presents the results for the networks with 2 and 3 UAVs, 14 GUs, $W = 10$ MHz, and $T = [10, 15]$s. In fact, larger satellite hop-counts would force higher computation load to be processed at the edge servers to satisfy the delay constraints of underlying computation tasks. However, limited computing resources at the edge servers also constrain the amount of partial tasks to be offloaded and processed completely in each time slot. As a result, the edge servers could be overloaded for larger satellite hop-counts and more GUs can be removed from the network. Moreover, it can be seen that the network with 3 UAVs can serve more GUs compared to that with 2 UAVs thanks to the better network coverage with more deployed UAVs.

7.8 Conclusion

In this chapter, we have studied the joint optimization of user scheduling, partial offloading control and bit allocation, computation resource, bandwidth allocation, admission control, and UAV trajectory control for the SAGIN. Specifically, we have used the alternating optimization approach to solve the underlying problem and leveraged the SCA method to tackle the non-convex bandwidth allocation and UAV trajectory control sub-problems. Moreover, we have proposed efficient strategies for feasibility verification and admission control, which can be employed in the overloaded network scenarios. Numerical results have demonstrated the effectiveness of the proposed algorithm compared to various baselines. We have also studied the impacts of various system parameters such as the number of GUs, bandwidth, the satellite hop-counts as well as the data size to computation load distribution and the energy consumption. Moreover, the admission ratio has been evaluated for different number of GUs, varying system bandwidth, and the satellite hop-counts.

7.9 Appendices

7.9.1 Proof of Proposition 7.1

In this appendix, we prove that the Algorithm 7.1 creates a non-increasing sequence of objective values of problem **(P3)** and converges to a feasible solution. First, it can be verified after the initialization step or after each r-iteration of the approximation process, we achieve a feasible solution of $\mathbf{\Theta}^r, \mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r$, and $\mathbf{Q}^r$. For step 3 of Algorithm 7.1, since the optimal solution of **(P3.1)** is obtained for given $\mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r$, and $\mathbf{Q}^r$, we have

$$\begin{aligned} &E^{\mathsf{sum}}(\mathbf{\Theta}^r, \mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r) \geq \\ &\qquad E^{\mathsf{sum}}(\mathbf{\Theta}^{r+1}, \mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r), \end{aligned} \tag{7.41}$$

where $E^{\mathsf{sum}}(\mathbf{\Theta}^r, \mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r)$ is defined as the objective function in the formulation for problem (7.19). Moreover, for given $\mathbf{\Theta}^{r+1}, \mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r,$ and $\mathbf{Q}^r$ obtained in step 3 of Algorithm 7.1, it follows that

$$\begin{aligned} &E^{\mathsf{sum}}(\mathbf{\Theta}^{r+1}, \mathbf{L}^r, \mathbf{\Lambda}^r, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r) \geq \\ &\qquad E^{\mathsf{sum}}(\mathbf{\Theta}^{r+1}, \mathbf{L}^{r+1}, \mathbf{\Lambda}^{r+1}, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r). \end{aligned} \tag{7.42}$$

This is because problem **(P3.2)** is optimally solved to obtain $\mathbf{L}^{r+1}$ and $\mathbf{\Lambda}^{r+1}$. Next, for given $\mathbf{\Theta}^{r+1}, \mathbf{L}^{r+1}, \mathbf{\Lambda}^{r+1}, \mathbf{F}^r, \boldsymbol{\beta}^r,$ and $\mathbf{Q}^r$ achieved in step 4 of Algorithm 7.1, we also have

$$\begin{aligned} &E^{\mathsf{sum}}(\mathbf{\Theta}^{r+1}, \mathbf{L}^{r+1}, \mathbf{\Lambda}^{r+1}, \mathbf{F}^r, \boldsymbol{\beta}^r, \mathbf{Q}^r) \geq \\ &\qquad E^{\mathsf{sum}}(\mathbf{\Theta}^{r+1}, \mathbf{L}^{r+1}, \mathbf{\Lambda}^{r+1}, \mathbf{F}^{r+1}, \boldsymbol{\beta}^{r+1}, \mathbf{Q}^r). \end{aligned} \tag{7.43}$$

This result is obtained by solving problem **(P3.3.2)** based ob the SCA method. Finally, for given $\boldsymbol{\Theta}^{r+1}, \mathbf{L}^{r+1}, \boldsymbol{\Lambda}^{r+1}, \mathbf{F}^{r+1}, \boldsymbol{\beta}^{r+1}$, and $\mathbf{Q}^{r}$ obtained in step 5 of Algorithm 7.1, we have

$$\begin{aligned} & E^{\mathsf{sum}}(\boldsymbol{\Theta}^{r+1}, \mathbf{L}^{r+1}, \boldsymbol{\Lambda}^{r+1}, \mathbf{F}^{r+1}, \boldsymbol{\beta}^{r+1}, \mathbf{Q}^{r}) \geq \\ & \quad E^{\mathsf{sum}}(\boldsymbol{\Theta}^{r+1}, \mathbf{L}^{r+1}, \boldsymbol{\Lambda}^{r+1}, \mathbf{F}^{r+1}, \boldsymbol{\beta}^{r+1}, \mathbf{Q}^{r+1}). \end{aligned} \tag{7.44}$$

This result holds since the problem **(P3.4.2)** achieves the solution $\mathbf{Q}^{r+1}$ by applying the SCA method. Using the results in (7.41)-(7.44), we obtain

$$\begin{aligned} & E^{\mathsf{sum}}(\boldsymbol{\Theta}, \mathbf{L}, \boldsymbol{\Lambda}, \mathbf{F}, \boldsymbol{\beta}, \mathbf{Q}) \geq \\ & \quad E^{\mathsf{sum}}(\boldsymbol{\Theta}^{r+1}, \mathbf{L}^{r+1}, \boldsymbol{\Lambda}^{r+1}, \mathbf{F}^{r+1}, \boldsymbol{\beta}^{r+1}, \mathbf{Q}^{r+1}), \end{aligned} \tag{7.45}$$

which indicates that the objective value of problem **(P3)** is non-increasing after each iteration of Algorithm 7.1. Since the objective value of problem **(P3)** is lower bounded by a finite value, Algorithm 7.1 is guaranteed to converge to a feasible solution. This completes the proof.

Chapter 8

Conclusions and Further Research Directions

In this chapter, we summarize our research contributions and discuss some potential directions for further research directions.

8.1 Major Research Contributions

The research performed in this dissertation results in three set of contributions. The first set of contributions are related to publications [75, 139, 140, 208] in which we investigate the UAV deployment, i.e., UAV placement and trajectory control, and resource allocation for UAV-based wireless networks. In the preliminary work [75], we study the problem of UAV placement and bandwidth allocation for wireless networks with wireless backhaul links. Specifically, we consider different configurations of LoS and NLoS propagation conditions between UAVs and GUs based on which we derive the average data rate of each access link. In the second work [139], we propose a novel design of single UAV trajectory and sub-channel assignment for both wireless access and backhaul links. The data transmission demands of individual GUs are considered in this work. The third work [140] extends the design in previous work where we study the multi-UAV trajectory control and non-orthogonal sub-channel assignment with co-channel interference management. The data transmission demands of individual GUs and wireless backhual links are still considered in this work. Finally, the work [208] formulates the joint UAV-GU association, UAV trajectory control, and non-orthogonal sub-channel assignment problem for UAV-based wireless networks. In the previous work [140], we develop a heuristic algorithm for the UAV trajectory control and sub-channel assignment problem. The work [208] makes several significant extensions of this conference work. Specifically, we solve three subproblems, namely, the UAV-GU association, sub-channel assignment, and UAV trajectory control, and develop an integrated algorithm to solve the joint optimization problem of UAV-GU association, sub-channel assignment, and UAV trajectory control. Moreover, we perform complexity analysis and prove the convergence of the integrated algorithm. Consequently, much more extensive numerical results are presented in this article compared to those in the conference paper to demonstrate the efficiency and desirable performance of the proposed algorithm. The alternating optimization approach and the successive convex approximation (SCA) method are employed in these work.

The second set of contributions correspond to publication [209] in which we tackle the joint optimization of UAV placement, IRS phase shifts, and sub-channel assignments for wireless access and backhaul links where our design objective is to maximize the sum rate achieved by GUs. The models underlying the UAV placement and resource allocation for both access and backhaul links are similar to those in our preliminary work [75]. The underlying optimization problem is in form of non-convex MINLP problem. To tackle this problem, we first derive the closed-form IRS phase shift solution and then optimize the sub-channel assignment and UAV placement in an iterative manner by using the alternating optimization method. Specifically, the sets of sub-channels assigned for the access and backhaul links are iteratively updated to efficiently utilize the available bandwidth while maintaining the backhaul capacity constraint. Moreover, we employ the SCA technique to solve the UAV placement sub-problem.

In the final set of contributions corresponding to publications [206, 210, 211], we study the joint computation offloading, UAV deployment, i.e., UAV placement or trajectory control, and resource allocation in SAGIN with multi-hop LEO satellite communications. In addition, we study partial computation offloading where fractions of computation tasks from GUs are processed locally and/or offloaded and processed at the UAV-aided edger servers and cloud server leveraging multi-hop LEO satellite communications. The aim of these designs is to minimize the weighted energy consumption while satisfying the maximum delay constraints of underlying tasks. Particularly, the work [210] investigates joint computation offloading, UAV placement, and resource allocation.

The joint computation offloading, UAV trajectory control, user scheduling, and resource allocation is studied in [206]. Finally, the work [211] makes several significant extensions compared to the conference publications. The feasibility verification and admission control designs are addressed for overloaded scenarios. In addition, complexity analysis for the proposed algorithms is conducted. Moreover, much more extensive numerical studies are performed in this journal version compared to those in the preliminary work [206]. The alternating optimization approach and the SCA method are employed in these work.

8.2 Concluding Remarks

In this doctoral dissertation, we studied network planning and resource management for UAV-based wireless networks with three main research contributions. We can put this work in context as follows:

- The results of this dissertation can be useful to address some long-range planning problems over a horizon of a year or more. Real-time implementation issues are outside the scope of this dissertation.
- Besides, the models are deterministic optimization problems where all input data are known, e.g., horizontal coordinates of UAVs and GUs. The GUs can be viewed as aggregates of traffic sources over a small region. The traffic demands are also averages of the demand over the planning horizon.
- Moreover, the proposed models and designs can provide answer some questions that are relevant in this context. Examples are how many UAVs should the network provider buy, whether or not the UAVs should be fixed or moving, how many IRS to buy and where they should be installed, whether a cloud architecture is worth it and where it should be located, etc.
- Furthermore, the more real-time issues such as GU-UAV association, UAV trajectory control, or computation splitting and offloading can also be used as guidelines for the real-time algorithms.

8.3 Further Research Directions

We now discuss potential research directions for further research following the studies in this dissertation.

8.3.1 Machine Learning Approaches for Resource Allocation in UAV-based Wireless Networks

While various design optimizations in UAV-based wireless networks can be performed for each service period, development and deployment of efficient algorithms for these optimization tasks can be quite challenging and may not be efficient in practice. This is because such an optimization algorithm cannot account for factors occurring during the considered service period. To this end, online optimization algorithms are more desirable because they can better adapt to system dynamics.

Several techniques can be employed to achieve this design target. In particular, decentralized optimization methods can be applied to engineer the UAV-based wireless networks in which individual UAVs can make their own decisions by using local network information and in collaboration with other UAVs in the network. Moreover, reinforcement learning techniques could be employed to optimize network operations so as to optimize long-term performance.

8.3.2 Enabling Technology for IRS-assisted UAV-based Wireless Networks

First, we would like to extend our work [209] to consider the multi-UAV trajectory control and sub-channel assignment for wireless access and backhaul links. This design is more challenging because we have to deal with the multiple associations between the UAVs and GUs and it is not trivial to develop an efficient algorithm for sub-channel assignment. The formulation optimization problem captures the design in the 3D space and it could lead to an NP-hard problem that is difficult to solve.

Exploiting the IRS technology for satellite-UAV-terrestrial networks was considered in [212–215]. Furthermore, the massive MIMO and beamforming techniques [216–219] can enable the IRSs to enhance the communication quality between the BS, UAVs and GUs. Therefore, applications of these promising technologies to the IRS-based SAGIN will be explored in our future work.

8.3.3 Machine Learning Approaches for Space–Air–Ground Integrated Networks Assisted Vehicular Networks

In general, the maximum delay constraints of the underlying computation tasks are difficult to maintain in many practical scenarios such as vehicular networks exploiting the SAGIN architecture. Therefore, it is essential to develop novel and efficient algorithms to engineer the resource allocation, UAV planning, and user scheduling algorithms. Besides, efficient offloading strategies for partial computation tasks should be developed. To this end, machine learning approaches could be considered to tackle these challenges.

Some recent surveys on the opportunities and challenges of edge computing and AI convergence for UAVs in SAGIN are given in [220,221]. Besides, a novel network architecture considering machine learning for SAGIN-assisted vehicular networks was discussed in [222]. We would like to further explore these research directions in our future work.

8.4 List of Publications

8.4.1 Journals

[J3]. Minh Dat Nguyen, Long Bao Le, and André Girard, "Integrated Computation Offloading, UAV Trajectory Control, User Scheduling, Resource Allocation, and Admission Control in SAGIN," submitted to *IEEE Transactions on Vehicular Technology*, Oct. 2022.

[J2]. Minh Dat Nguyen, Long Bao Le, and André Girard, "UAV Placement and Resource Allocation for Intelligent Reflecting Surface Assisted UAV-Based Wireless Networks," *IEEE Communications Letters*, vol. 26, no. 5, pp. 1106–1110, May 2022.

[J1]. Minh Dat Nguyen, Long Bao Le, and André Girard, "Integrated UAV Trajectory Control and Resource Allocation for UAV-Based Wireless Networks With Co-Channel Interference Management," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12754–12769, Jul. 2022.

8.4.2 Conferences

[C5]. Minh Dat Nguyen, Long Bao Le, and André Girard, "Joint computation offloading, UAV trajectory, user scheduling, and resource allocation in SAGIN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022, to appear.

[C4]. Minh Dat Nguyen, Long Bao Le, and André Girard, "Computation offloading, UAV placement, and resource allocation in SAGIN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM) Workshops*, Dec. 2022, to appear.

[C3]. Minh Dat Nguyen, Long Bao Le, and André Girard, "Trajectory control and resource allocation for UAVbased networks with wireless backhauls," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021.

[C2]. Minh Dat Nguyen, Tai Manh Ho, Long Bao Le, and André Girard, "UAV trajectory and sub-channel assignment for UAV based wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020.

[C1]. Minh Dat Nguyen, Tai Manh Ho, Long Bao Le, and André Girard, "UAV placement and bandwidth allocation for UAV based wireless networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019.

Références

[1] B. P. S. Sahoo, D. Puthal, and P. K. Sharma, “Toward advanced UAV communications: Properties, research challenges, and future potential,” *IEEE Internet Things Mag.*, vol. 5, no. 1, pp. 154–159, Mar. 2022.

[2] Q. Chen, G. Giambene, L. Yang, C. Fan, and X. Chen, “Analysis of inter-satellite link paths for LEO mega-constellation networks,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2743–2755, Mar. 2021.

[3] M. M. Azari, S. Solanki, S. Chatzinotas, O. Kodheli, H. Sallouha, A. Colpaert, J. F. M. Montoya, S. Pollin, A. Haqiqatnejad, A. Mostaani, E. Lagunas, and B. Ottersten, “Evolution of non-terrestrial networks from 5G to 6G: A survey,” *IEEE Commun. Surveys Tuts.*, 2022, early access.

[4] G. Geraci, D. Lopez-Perez, M. Benzaghta, and S. Chatzinotas, “Integrating terrestrial and non-terrestrial networks: 3D opportunities and challenges,” Jul. 2022. [Online]. Available: https://arxiv.org/abs/2207.10385

[5] IEEE Future Networks, “IEEE 5G and beyond technology roadmap white paper,” Oct. 2017. [Online]. Available: https://futurenetworks.ieee.org/images/files/pdf/ieee-5g-roadmap-white-paper.pdf

[6] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, “Cellular architecture and key technologies for 5G wireless communication networks,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[7] Cisco, “Cisco annual internet report (2018–2023) white paper,” Mar. 2020, (accessed Sep. 2022). [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[8] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, Jun. 2020.

[9] B. Li, Z. Fei, and Y. Zhang, “UAV communications for 5G and beyond: Recent advances and future trends,” *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, Apr. 2019.

[10] G. Geraci, A. Garcia-Rodriguez, M. M. Azari, A. Lozano, M. Mezzavilla, S. Chatzinotas, Y. Chen, S. Rangan, and M. D. Renzo, “What will the future of UAV cellular communications be? A flight from 5G to 6G,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1304–1335, 3rd Quart. 2022.

[11] B. Tezergil and E. Onur, “Wireless backhaul in 5G and beyond: Issues, challenges and opportunities,” *IEEE Commun. Surveys Tuts.*, 2022, early access.

[12] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H. V. Poor, “6G internet of things: A comprehensive survey,” *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, Jan. 2022.

[13] 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.

[14] M. Marchese, A. Moheddine, and F. Patrone, “IoT and UAV integration in 5G hybrid terrestrial-satellite networks,” *Sensors*, vol. 19, no. 17, Aug. 2019.

[15] Z. Ullah, F. Al-Turjman, and L. Mostarda, “Cognition in UAV-aided 5G and beyond communications: A survey,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 872–891, Sep. 2020.

[16] L. Yang, P. Li, F. Meng, and S. Yu, “Performance analysis of RIS-assisted UAV communication systems,” *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 9078–9082, Aug. 2022.

[17] H. Ren, Z. Zhang, Z. Peng, L. Li, and C. Pan, “Energy minimization in RIS-assisted UAV-enabled wireless power transfer systems,” *IEEE Internet Things J.*, 2022, early access.

[18] L. Yang, F. Meng, J. Zhang, M. O. Hasna, and M. D. Renzo, “On the performance of RIS-assisted dual-hop UAV communication systems,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 10 385–10 390, Sep. 2020.

[19] S. Li, Q. Chen, W. Meng, and C. Li, “Civil aircraft assisted space-air-ground integrated networks: An innovative NTN of 5G and beyond,” *IEEE Wireless Commun.*, pp. 1–8, 2022, early access.

[20] S. Li, Q. Chen, Z. Li, W. Meng, and C. Li, “Civil aircraft assisted space-air-ground integrated networks: Architecture design and coverage analysis,” *China Commun.*, vol. 19, no. 1, pp. 29–39, Jan. 2022.

[21] H. Cui, J. Zhang, Y. Geng, Z. Xiao, T. Sun, N. Zhang, J. Liu, Q. Wu, and X. Cao, “Space-air-ground integrated network (SAGIN) for 6G: Requirements, architecture and challenges,” *China Commun.*, vol. 19, no. 2, pp. 90–108, Feb. 2022.

[22] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, “Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1117–1174, 2nd Quart. 2022.

[23] G. K. Kurt, M. G. Khoshkholgh, S. Alfattani, A. Ibrahim, T. S. Darwish, M. S. Alam, H. Yanikomeroglu, and A. Yongacoglu, “A vision and framework for the high altitude platform station (HAPS) networks of the future,” *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 729–779, 2nd Quart. 2021.

[24] R. Chen, M. Liu, Y. Hui, N. Cheng, and J. Li, “Reconfigurable intelligent surfaces for 6G IoT wireless positioning: A contemporary survey,” *IEEE Internet Things J.*, 2022, early access.

[25] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, “Reconfigurable intelligent surfaces: Principles and opportunities,” *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart. 2021.

[26] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, “Intelligent reflecting surface aided wireless communications: A tutorial,” *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[27] S. Li, B. Duo, X. Yuan, Y. Liang, and M. Di Renzo, “Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming,” *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 716–720, May 2020.

[28] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, “Wireless communications through reconfigurable intelligent surfaces,” *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.

[29] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[30] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, “Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications,” *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Sep. 2019.

[31] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, “Space-air-ground integrated network: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, 4th Quart. 2018.

[32] T. Hong, W. Zhao, R. Liu, and M. Kadoch, “Space-air-ground IoT network and related key technologies,” *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 96–104, Apr. 2020.

[33] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, “A survey on space-air-ground-sea integrated network security in 6G,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 53–87, 1st Quart. 2022.

[34] Z. Jia, M. Sheng, J. Li, and Z. Han, “Toward data collection and transmission in 6G space–air–ground integrated networks: Cooperative HAP and LEO satellite schemes,” *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10 516–10 528, Jul. 2022.

[35] Z. Zhang, Y. Li, C. Huang, Q. Guo, L. Liu, C. Yuen, and Y. L. Guan, “User activity detection and channel estimation for grant-free random access in LEO satellite-enabled internet of things,” *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8811–8825, Sep. 2020.

[36] M. Sheng, D. Zhou, R. Liu, Y. Wang, and J. Li, “Resource mobility in space information networks: Opportunities, challenges, and approaches,” *IEEE Netw.*, vol. 33, no. 1, pp. 128–135, Feb. 2019.

[37] S. Chandrasekharan, K. Gomez, A. Al-Hourani, S. Kandeepan, T. Rasheed, L. Goratti, L. Reynaud, D. Grace, I. Bucaille, T. Wirth, and S. Allsopp, “Designing and implementing future aerial communication networks,” *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 26–34, May 2016.

[38] M. Grant and S. Boyd, “CVX: MATLAB software for disciplined convex programming, version 2.1,” Dec. 2018. [Online]. Available: http://cvxr.com/cvx/

[39] Q. Wu and R. Zhang, “Common throughput maximization in UAV-enabled OFDMA systems with delay consideration,” *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6614–6627, Dec. 2018.

[40] Z. Li, M. Chen, C. Pan, N. Huang, Z. Yang, and A. Nallanathan, “Joint trajectory and communication design for secure UAV networks,” *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 636–639, Apr. 2019.

[41] Q. Wu, Y. Zeng, and R. Zhang, “Joint trajectory and communication design for multi-UAV enabled wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2109–2121, Mar. 2018.

[42] *Online Video Test Media*, (accessed on 10 May 2021). [Online]. Available: https://media.xiph.org/video/derf/

[43] S. Fang, G. Chen, and Y. Li, “Joint optimization for secure intelligent reflecting surface assisted UAV networks,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 276–280, Feb. 2021.

[44] S. Mao, S. He, and J. Wu, “Joint UAV position optimization and resource scheduling in space-air-ground integrated networks with mixed cloud-edge computing,” *IEEE Syst. J.*, vol. 15, no. 3, pp. 3992–4002, Sep. 2021.

[45] M. Feng, M. Krunz, and W. Zhang, “Joint task partitioning and user association for latency minimization in mobile edge computing networks,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8108–8121, Aug. 2021.

[46] Y. Dai, D. Xu, S. Maharjan, and Y. Zhang, “Joint computation offloading and user association in multi-task mobile edge computing,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12 313–12 325, Dec. 2018.

[47] N. Saeed, H. Almorad, H. Dahrouj, T. Y. Al-Naffouri, J. S. Shamma, and M.-S. Alouini, “Point-to-point communication in integrated satellite-aerial 6G networks: State-of-the-art and future challenges,” *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1505–1525, Jun. 2021.

[48] M. Liu, Y. Wang, Z. Li, X. Lyu, and Y. Chen, “Joint optimization of resource allocation and multi-UAV trajectory in space-air-ground IoRT networks,” in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2020, pp. 1–6.

[49] Ericsson., “Ericsson mobility report,” Jun. 2019, (accessed Sep. 2022). [Online]. Available: https://www.ericsson.com/assets/local/reports-papers/mobility-report/documents/2019/ericsson-mobility-report-june-2019.pdf

[50] Statista, “Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2021, with forecasts from 2022 to 2030,” Aug. 2022, (accessed Sep. 2022). [Online]. Available: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

[51] Cisco, “Annual internet report,” Mar. 2020, (accessed Sep. 2022). [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf

[52] Qualcomm Technologies, “LTE unmanned aircraft systems,” May 2017, (accessed Sep. 2022). [Online]. Available: https://www.qualcomm.com/media/documents/files/lte-unmanned-aircraft-systems-trial-report.pdf

[53] N. H. Motlagh, T. Taleb, and O. Arouk, "Low-altitude unmanned aerial vehicles-based internet of things services: Comprehensive survey and future perspectives," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 899–922, Dec. 2016.

[54] S. Sekander, H. Tabassum, and E. Hossain, "Multi-tier drone architecture for 5G/B5G cellular networks: Challenges, trends, and prospects," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 96–103, Mar. 2018.

[55] 3GPP Standard TR 36.777, "Study on enhanced LTE support for aerial vehicles (release 15)," Dec. 2017, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/36_series/36.777/

[56] 3GPP Standard TR 22.825, "Study on remote identification of unmanned aerial systems (UAS) (release 16)," Sep. 2018, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/22_series/22.825/

[57] 3GPP Standard TR 23.754, "Study on supporting unmanned aerial systems (UAS) connectivity, identification and tracking (release 17)," Mar. 2021, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/23_series/23.754/

[58] 3GPP Standard TR 23.755, "Study on application layer support for unmanned aerial systems (uas) (release 17)," Apr. 2021, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/23_series/23.755/

[59] 3GPP Standard TR 22.829, "Enhancement for unmaned aerial vehicles (UAVs) (release 17)," Sep. 2019, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/22_series/22.829/

[60] X. Lin, "An overview of 5G advanced evolution in 3GPP release 18," *arXiv*, 2022. [Online]. Available: https://arxiv.org/abs/2201.01358

[61] S. Zhang, Y. Zeng, and R. Zhang, "Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2580–2604, Mar. 2019.

[62] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 328–331, Jun. 2018.

[63] M. B. Ghorbel, D. Rodríguez-Duarte, H. Ghazzai, M. J. Hossain, and H. Menouar, "Joint position and travel path optimization for energy efficient wireless data gathering using unmanned aerial vehicles," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2165–2175, Mar. 2019.

[64] V. Sharma, R. Kumar, and R. Kaur, "UAV-assisted content-based sensor search in IoTs," *Electron. Lett.*, vol. 53, no. 11, pp. 724–726, May 2017.

[65] Y. Zeng, R. Zhang, and T. J. Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 4983–4996, Dec. 2016.

[66] J. Fan, M. Cui, G. Zhang, and Y. Chen, "Throughput improvement for multi-hop UAV relaying," *IEEE Access*, vol. 7, pp. 147 732–147 742, Dec. 2019.

[67] W. Wang, X. Li, M. Zhang, K. Cumanan, D. W. Kwan Ng, G. Zhang, J. Tang, and O. A. Dobre, "Energy-constrained UAV-assisted secure communications with position optimization and cooperative jamming," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4476–4489, Jul. 2020.

[68] H. Wang, J. Wang, G. Ding, J. Chen, Y. Li, and Z. Han, "Spectrum sharing planning for full-duplex UAV relaying systems with underlaid D2D communications," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1986–1999, Sep. 2018.

[69] R. Fan, J. Cui, S. Jin, K. Yang, and J. An, "Optimal node placement and resource allocation for UAV relaying network," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 808–811, Apr. 2018.

[70] M. Mozaffari, W. Saad, M. Bennis, Y. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 3rd Quart. 2019.

[71] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3417–3442, 4th Quart. 2019.

[72] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[73] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 604–607, Mar. 2017.

[74] P. Li and J. Xu, "Placement optimization for UAV-enabled wireless networks with multi-hop backhaul," *J. Commun. Inform. Netw.*, vol. 3, no. 4, pp. 312–316, Dec. 2018.

[75] M. D. Nguyen, T. M. Ho, L. B. Le, and A. Girard, "UAV placement and bandwidth allocation for UAV based wireless networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[76] H. Tang, Q. Wu, and B. Li, "An efficient solution for joint power and trajectory optimization in UAV-enabled wireless network," *IEEE Access*, vol. 7, pp. 59 640–59 652, May 2019.

[77] M. T. Nguyen and L. B. Le, "Flight scheduling and trajectory control in UAV-based wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[78] M. Alzenad, A. El-Keyi, and H. Yanikomeroglu, "3-D placement of an unmanned aerial vehicle base station for maximum coverage of users with different QoS requirements," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 38–41, Feb. 2018.

[79] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient Internet of Things communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7574–7589, Nov. 2017.

[80] H. Huang, C. Huang, and D. Ma, "A method for deploying the minimal number of UAV base stations in cellular networks," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 2, pp. 559–567, Mar. 2020.

[81] N. C. Luong, D. T. Hoang, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Data collection and wireless communication in internet of things (IoT) using economic analysis and pricing models: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2546–2590, 4th Quart. 2016.

[82] C. Zhan, Y. Zeng, and R. Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 328–331, Jun. 2018.

[83] J. Gong, T.-H. Chang, C. Shen, and X. Chen, "Flight time minimization of UAV for data collection over wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1942–1954, Sep. 2018.

[84] W. Wang, N. Zhao, L. Chen, X. Liu, Y. Chen, and D. Niyato, "UAV-assisted time-efficient data collection via uplink NOMA," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7851–7863, Nov. 2021.

[85] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[86] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, May 2019.

[87] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Aug. 2019.

[88] O. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, Dec. 2020.

[89] 3GPP, "Discussion about reconfigurable intelligent surfaces (RIS) (release 18)," Sep. 2021, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_AHs/2021_06_RAN_Rel18_WS/Docs

[90] A. U. Chaudhry and H. Yanikomeroglu, "Free space optics for next-generation satellite networks," *IEEE Consum. Electron. Mag.*, vol. 10, no. 6, pp. 21–31, Nov. 2021.

[91] Z. Jia, M. Sheng, J. Li, D. Niyato, and Z. Han, "LEO-satellite-assisted UAV: Joint trajectory and data collection for internet of remote things in 6G aerial access networks," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9814–9826, Jun. 2021.

[92] T. Darwish, G. K. Kurt, H. Yanikomeroglu, M. Bellemare, and G. Lamontagne, "LEO satellites in 5G and beyond networks: A review from a standardization perspective," *IEEE Access*, vol. 10, pp. 35 040–35 060, Mar. 2022.

[93] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 450–465, Feb. 2018.

[94] J. Zhang, X. Hu, Z. Ning, E. C. Ngai, L. Zhou, J. Wei, J. Cheng, B. Hu, and V. C. M. Leung, "Joint resource allocation for latency-sensitive services over mobile edge computing networks with caching," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4283–4294, Jun. 2019.

[95] X. Lin, S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, "5G from space: An overview of 3GPP non-terrestrial networks," *IEEE Commun. Standards Mag.*, vol. 5, no. 4, pp. 147–153, Dec. 2021.

[96] 3GPP Standard TR 38.811, "Study on new radio (NR) to support non-terrestrial networks (release 15)," Oct. 2020, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.811/

[97] 3GPP Standard TR 22.822, "Study on using satellite access in 5G (release 16)," Jul. 2018, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/22_series/22.822/

[98] 3GPP Standard TR 23.737, "Study on architecture aspects for using satellite access in 5G (release 17)," Mar. 2021, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/23_series/23.737/

[99] S. Cioni, R. De Gaudenzi, O. Del Rio Herrero, and N. Girault, "On the satellite role in the era of 5G massive machine type communications," *IEEE Netw.*, vol. 32, no. 5, pp. 54–61, Oct. 2018.

[100] 3GPP, "NTN & satellite in Rel-17 & 18," Oct. 2021, (accessed Sep. 2022). [Online]. Available: https://www.3gpp.org/news-events/partners-news/2254-ntn_rel17

[101] S. Alfattani, W. Jaafar, Y. Hmamouche, H. Yanikomeroglu, A. Yongaçoglu, N. D. Đào, and P. Zhu, "Aerial platforms with reconfigurable smart surfaces for 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 96–102, Jan. 2021.

[102] V. P. Hubenko, R. A. Raines, R. F. Mills, R. O. Baldwin, B. E. Mullins, and M. R. Grimaila, "Improving the global information grid's performance through satellite communications layer enhancements," *IEEE Commun. Mag.*, vol. 44, no. 11, pp. 66–72, Nov. 2006.

[103] M. Albuquerque, A. Ayyagari, M. A. Dorsett, and M. S. Foster, "Global information grid (GIG) edge network interface architecture," in *Proc. IEEE MILCOM*, Oct. 2007, pp. 1–7.

[104] C. K. J. Radtke and E. Stol, "Interactions of the space debris environment with mega constellations - Using the example of the OneWeb constellation," *Acta Astronautica*, vol. 131, pp. 55–68, Feb. 2017.

[105] T. Pultarova, "Telecommunications - Space tycoons go head to head over mega satellite network [news briefing]," *Eng. Technol.*, vol. 10, no. 2, pp. 20–20, 2015.

[106] H. Yao, L. Wang, X. Wang, Z. Lu, and Y. Liu, "The space-terrestrial integrated network: An overview," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 178–185, Sep. 2018.

[107] L. Zhang, W. Abderrahim, and B. Shihada, "Heterogeneous traffic offloading in space-air-ground integrated networks," *IEEE Access*, vol. 9, pp. 165 462–165 475, Dec. 2021.

[108] B. Cao, J. Zhang, X. Liu, Z. Sun, W. Cao, R. M. Nowak, and Z. Lv, "Edge–cloud resource scheduling in space–air–ground-integrated networks for internet of vehicles," *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5765–5772, Apr. 2022.

[109] X. Hou, J. Wang, Z. Fang, Y. Ren, K.-C. Chen, and L. Hanzo, "Edge intelligence for mission-critical 6G services in space-air-ground integrated networks," *IEEE Netw.*, vol. 36, no. 2, pp. 181–189, Apr. 2022.

[110] Y. Guo, Q. Li, Y. Li, N. Zhang, and S. Wang, "Service coordination in the space-air-ground integrated network," *IEEE Netw.*, vol. 35, no. 5, pp. 168–173, Oct. 2021.

[111] G. Wang, S. Zhou, S. Zhang, Z. Niu, and X. Shen, "SFC-based service provisioning for reconfigurable space-air-ground integrated networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 7, pp. 1478–1489, Jul. 2020.

[112] P. Zhang, P. Yang, N. Kumar, and M. Guizani, “Space-air-ground integrated network resource allocation based on service function chain,” *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7730–7738, Jul. 2022.

[113] J. Li, W. Shi, H. Wu, S. Zhang, and X. Shen, “Cost-aware dynamic SFC mapping and scheduling in SDN/NFV-enabled space–air–ground-integrated networks for internet of vehicles,” *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5824–5838, Apr. 2022.

[114] L. Qu, G. Xu, Z. Zeng, N. Zhang, and Q. Zhang, “UAV-assisted RF/FSO relay system for space-air-ground integrated network: A performance analysis,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6211–6225, Aug. 2022.

[115] J. Ye, S. Dang, B. Shihada, and M.-S. Alouini, “Space-air-ground integrated networks: Outage performance analysis,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7897–7912, Dec. 2020.

[116] H. Zhang, G. Pan, S. Ke, S. Wang, and J. An, “Outage analysis of cooperative satellite-aerial-terrestrial networks with spatially random terminals,” *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4972–4987, Jul. 2022.

[117] W. Huang, T. Song, and J. An, “QA2: QoS-guaranteed access assistance for space–air–ground internet of vehicle networks,” *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5684–5695, Apr. 2022.

[118] H. Wang, K. Fan, K. Zhang, Z. Wang, H. Li, and Y. Yang, “Encrypted data retrieval and sharing scheme in space–air–ground-integrated vehicular networks,” *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5957–5970, Apr. 2022.

[119] W. Sun, L. Wang, P. Wang, and Y. Zhang, “Collaborative blockchain for space-air-ground integrated networks,” *IEEE Wireless Commun.*, vol. 27, no. 6, pp. 82–89, Dec. 2020.

[120] Z. Wang, F. Zhang, Q. Yu, and T. Qin, “Blockchain-envisioned unmanned aerial vehicle communications in space-air-ground integrated network: A review,” *Intell. Converged Netw.*, vol. 2, no. 4, pp. 277–294, Dec. 2021.

[121] Y. Wang, Z. Su, J. Ni, N. Zhang, and X. Shen, “Blockchain-empowered space-air-ground integrated networks: Opportunities, challenges, and solutions,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 160–209, 1st Quart. 2022.

[122] B. Li, R. Liang, W. Zhou, H. Yin, H. Gao, and K. Cai, “LBS meets blockchain: An efficient method with security preserving trust in SAGIN,” *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5932–5942, Apr. 2022.

[123] N. Kato, Z. M. Fadlullah, F. Tang, B. Mao, S. Tani, A. Okamura, and J. Liu, “Optimizing space-air-ground integrated networks by artificial intelligence,” *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 140–147, Aug. 2019.

[124] F. Tang, H. Hofner, N. Kato, K. Kaneko, Y. Yamashita, and M. Hangai, “A deep reinforcement learning-based dynamic traffic offloading in space-air-ground integrated networks (SAGIN),” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 276–289, Jan. 2022.

[125] P. Zhang, C. Wang, N. Kumar, and L. Liu, “Space-air-ground integrated multi-domain network resource orchestration based on virtual network architecture: A DRL method,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2798–2808, Mar. 2022.

[126] B. Mao, F. Tang, Y. Kawamoto, and N. Kato, “Optimizing computation offloading in satellite-UAV-served 6G IoT: A deep learning approach,” *IEEE Netw.*, vol. 35, no. 4, pp. 102–108, Aug. 2021.

[127] C. Zhou, W. Wu, H. He, P. Yang, F. Lyu, N. Cheng, and X. Shen, “Deep reinforcement learning for delay-oriented IoT task scheduling in SAGIN,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 911–925, Feb. 2021.

[128] F. Tang, C. Wen, X. Chen, and N. Kato, “Federated learning for intelligent transmission with space-air-ground integrated network (SAGIN) toward 6G,” *IEEE Netw.*, pp. 1–7, 2022, early access.

[129] N. Cheng, W. Xu, W. Shi, Y. Zhou, N. Lu, H. Zhou, and X. Shen, “Air-ground integrated mobile edge networks: Architecture, challenges, and opportunities,” *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 26–32, Aug. 2018.

[130] F. Zhou, R. Q. Hu, Z. Li, and Y. Wang, “Mobile edge computing in unmanned aerial vehicle networks,” *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 140–146, Feb. 2020.

[131] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, “Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges,” *IEEE Access*, vol. 7, pp. 48 572–48 634, 2019.

[132] H. Kim, L. Mokdad, and J. Ben-Othman, “Designing UAV surveillance frameworks for smart city and extensive ocean with differential perspectives,” *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 98–104, Apr. 2018.

[133] R. Shakeri, M. A. Al-Garadi, A. Badawy, A. Mohamed, T. Khattab, A. K. Al-Ali, K. A. Harras, and M. Guizani, “Design challenges of multi-UAV systems in cyber-physical applications: A comprehensive survey and future directions,” *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3340–3385, 4th Quart. 2019.

[134] H. Mei, K. Yang, Q. Liu, and K. Wang, “Joint trajectory-resource optimization in UAV-enabled edge-cloud system with virtualized mobile clone,” *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5906–5921, Jul. 2020.

[135] F. Zeng, Z. Hu, Z. Xiao, H. Jiang, S. Zhou, W. Liu, and D. Liu, “Resource allocation and trajectory optimization for QoE provisioning in energy-efficient UAV-enabled wireless networks,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 7634–7647, Jul. 2020.

[136] Y. Huang, M. Cui, G. Zhang, and W. Chen, “Bandwidth, power and trajectory optimization for UAV base station networks with backhaul and user QoS constraints,” *IEEE Access*, vol. 8, pp. 67 625–67 634, Apr. 2020.

[137] S. Zhang, H. Zhang, B. Di, and L. Song, “Cellular UAV-to-X communications: Design and optimization for multi-UAV networks,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1346–1359, Feb. 2019.

[138] Y. Cai, Z. Wei, R. Li, D. W. K. Ng, and J. Yuan, “Joint trajectory and resource allocation design for energy-efficient secure UAV communication systems,” *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4536–4553, Jul. 2020.

[139] M. D. Nguyen, T. M. Ho, L. B. Le, and A. Girard, "UAV trajectory and sub-channel assignment for UAV based wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[140] M. D. Nguyen, L. B. Le, and A. Girard, "Trajectory control and resource allocation for UAV-based networks with wireless backhauls," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[141] J.-M. Kang, "Intelligent reflecting surface: Joint optimal training sequence and refection pattern," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1784–1788, Aug. 2020.

[142] H. Lu, Y. Zeng, S. Jin, and R. Zhang, "Aerial intelligent reflecting surface: Joint placement and passive beamforming design with 3D beam flattening," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4128–4143, Jul. 2021.

[143] S. Solanki, S. Gautam, S. K. Sharma, and S. Chatzinotas, "Ambient backscatter assisted co-existence in aerial-IRS wireless networks," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 608–621, Mar. 2022.

[144] S. Solanki, J. Park, and I. Lee, "On the performance of IRS-aided UAV networks with NOMA," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 9038–9043, Aug. 2022.

[145] Z. Mohamed and S. Aïssa, "Leveraging UAVs with intelligent reflecting surfaces for energy-efficient communications with cell-edge users," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Wkshps)*, Jun. 2020, pp. 1–6.

[146] T. Shafique, H. Tabassum, and E. Hossain, "Optimization of wireless relaying with flexible UAV-borne reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 309–325, Jan. 2021.

[147] S. Jiao, F. Fang, X. Zhou, and H. Zhang, "Joint beamforming and phase shift design in downlink UAV networks with IRS-assisted NOMA," *J. Commun. Inform. Netw.*, vol. 5, no. 2, pp. 138–149, Jun. 2020.

[148] D. Ma, M. Ding, and M. Hassan, "Enhancing cellular communications for UAVs via intelligent reflective surface," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[149] W. Wang, H. Tian, W. Ni, and M. Hua, "Intelligent reflecting surface aided secure UAV communications," Nov. 2020, [Online] Available: https://arxiv.org/pdf/2011.04339.pdf.

[150] X. Mu, Y. Liu, L. Guo, J. Lin, and H. V. Poor, "Intelligent reflecting surface enhanced multi-UAV NOMA networks," Jan. 2021, [Online] Available: https://arxiv.org/pdf/2101.09145.pdf.

[151] Y. Cai, Z. Wei, S. Hu, D. W. K. Ng, and J. Yuan, "Resource allocation for power-efficient IRS-assisted UAV communications," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Wkshps)*, Jun. 2020, pp. 1–7.

[152] L. Sixian, D. Bin, M. D. Renzo, T. Meixia, and Y. Xiaojun, "Robust secure UAV communications with the aid of reconfigurable intelligent surfaces," Dec. 2020, [Online] Available: https://arxiv.org/pdf/2008.09404.pdf.

[153] Z. Wei, Y. Cai, Z. Sun, D. W. K. Ng, J. Yuan, M. Zhou, and L. Sun, "Sum-rate maximization for IRS-assisted UAV OFDM communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2530–2550, Apr. 2021.

[154] Y. Pan, K. Wang, C. Pan, H. Zhu, and J. Wang, "UAV-assisted and intelligent reflecting surfaces-supported terahertz communications," Oct. 2020, [Online] Available: https://arxiv.org/pdf/2010.14223.pdf.

[155] L. Ge, P. Dong, H. Zhang, J. Wang, and X. You, "Joint beamforming and trajectory optimization for intelligent reflecting surfaces-assisted UAV communications," *IEEE Access*, vol. 8, pp. 78 702–78 712, Apr. 2020.

[156] M. Hua, L. Yang, Q. Wu, C. Pan, C. Li, and A. L. Swindlehurst, "UAV-assisted intelligent reflecting surface symbiotic radio system," Jan. 2021, [Online] Available: https://arxiv.org/pdf/2007.14029.pdf.

[157] B. Shang, Y. Yi, and L. Liu, "Computing over space-air-ground integrated networks: Challenges and opportunities," *IEEE Netw.*, vol. 35, no. 4, pp. 302–309, Aug. 2021.

[158] S. Yu, X. Gong, Q. Shi, X. Wang, and X. Chen, "EC-SAGINs: Edge-computing-enhanced space-air-ground-integrated networks for internet of vehicles," *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5742–5754, Apr. 2022.

[159] B. Cao, J. Zhang, X. Liu, Z. Sun, W. Cao, R. M. Nowak, and Z. Lv, "Edge–cloud resource scheduling in space–air–ground-integrated networks for internet of vehicles," *IEEE Internet Things J.*, vol. 9, no. 8, pp. 5765–5772, Apr. 2022.

[160] M. Liu, Y. Wang, Y. Chen, and H. Jia, "Joint stochastic computational resource and UAV trajectory for wireless-powered space-air-ground IoRT networks," *IEEE Access*, vol. 8, pp. 193 728–193 743, Oct. 2020.

[161] J. Liu, X. Zhao, P. Qin, S. Geng, and S. Meng, "Joint dynamic task offloading and resource scheduling for WPT enabled space-air-ground power internet of things," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 2, pp. 660–677, Mar. 2022.

[162] Y. Chen, B. Ai, Y. Niu, H. Zhang, and Z. Han, "Energy-constrained computation offloading in space-air-ground integrated networks using distributionally robust optimization," *IEEE Trans. Veh. Technol.*, vol. 70, no. 11, pp. 12 113–12 125, Nov. 2021.

[163] Q. Chen, W. Meng, S. Han, and C. Li, "Service-oriented fair resource allocation and auction for civil aircrafts augmented space-air-ground integrated networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 658–13 672, Nov. 2020.

[164] Z. Li, Y. Wang, M. Liu, R. Sun, Y. Chen, J. Yuan, and J. Li, "Energy efficient resource allocation for UAV-assisted space-air-ground internet of remote things networks," *IEEE Access*, vol. 7, pp. 145 348–145 362, Oct. 2019.

[165] Y. Wang, Z. Li, Y. Chen, M. Liu, X. Lyu, X. Hou, and J. Wang, "Joint resource allocation and UAV trajectory optimization for space–air–ground internet of remote things networks," *IEEE Syst. J.*, pp. 1–11, Oct. 2020.

[166] F. Pervez, L. Zhao, and C. Yang, "Joint user association, power optimization and trajectory control in an integrated satellite-aerial-terrestrial network," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3279–3290, May 2022.

[167] Y. Liu, H. Zhang, H. Zhou, K. Long, and V. C. M. Leung, "User association, subchannel and power allocation in space-air-ground integrated vehicular network with delay constraints," *IEEE Trans. Netw. Sci. Eng.*, Apr. 2022, early access.

[168] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[169] A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming.* Lecture Notes, Atlanta, GA, USA: Georgia Inst. Technol., 2004.

[170] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications.* Society for Industrial and Applied Mathematics, 2001.

[171] J. Nocedal and S. Wright, *Numerical optimization.* Springer Science & Business Media, 2006.

[172] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual, version 9.0.*, 2019. [Online]. Available: http://docs.mosek.com/9.0/toolbox/index.html

[173] M. Razaviyayn, "Successive convex approximation: Analysis and applications," Ph.D. dissertation, University of Minnesota, 2014.

[174] P. Dinh Tao and H. A. Le Thi, "Convex analysis approach to D.C. programming: Theory, Algorithm and Applications," *Acta Mathematica Vietnamica*, vol. 22, Jan. 1997.

[175] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.

[176] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 2898–2904.

[177] A. u. R. Khan, M. Othman, S. A. Madani, and S. U. Khan, "A survey of mobile cloud computing application models," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 393–413, 1st Quart. 2014.

[178] P. M. M. N. B. Chun, S. Ihm and A. Patti, "Clonecloud: Elastic execution between mobile device and cloud," in *Proc. ACM EuroSys*, 2011, p. 301–314.

[179] S. Kosta, A. Aucinas, P. Hui, R. Mortier, and X. Zhang, "Thinkair: Dynamic resource allocation and parallel execution in the cloud for mobile code offloading," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 945–953.

[180] C. Wang, F. R. Yu, C. Liang, Q. Chen, and L. Tang, "Joint computation offloading and interference management in wireless cellular networks with mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7432–7445, Aug. 2017.

[181] W. Zhang, Y. Wen, K. Guan, D. Kilper, H. Luo, and D. O. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4569–4581, Sep. 2013.

[182] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart. 2017.

[183] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, pp. 2795–2808, Oct. 2016.

[184] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in *Proc. NIPS*, 2014, p. 568–576.

[185] L. You, Q. Liao, N. Pappas, and D. Yuan, "Resource optimization with flexible numerology and frame structure for heterogeneous services," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2579–2582, Dec. 2018.

[186] S. E. Mahmoodi, R. N. Uma, and K. P. Subbalakshmi, "Optimal joint scheduling and cloud offloading for mobile applications," *IEEE Trans. Cloud Comp.*, vol. 7, no. 2, pp. 301–313, Jun. 2019.

[187] M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad, "Satellite communications supporting internet of remote things," *IEEE Internet Things J.*, vol. 3, no. 1, pp. 113–123, Feb. 2016.

[188] Y. Zhang, Q. Wu, Z. Lai, and H. Li, "Enabling low-latency-capable satellite-ground topology for emerging LEO satellite networks," in *Proc. IEEE INFOCOM - IEEE Conf. Comp. Commun.*, Jun. 2022, pp. 1329–1338.

[189] F. Ma, X. Zhang, X. Li, J. Cheng, F. Guo, J. Hu, and L. Pan, "Hybrid constellation design using a genetic algorithm for a LEO-based navigation augmentation system," *GPS Solut.*, vol. 24, no. 2, pp. 1–14, Apr. 2020.

[190] Y. Su, Y. Liu, Y. Zhou, J. Yuan, H. Cao, and J. Shi, "Broadband LEO satellite communications: Architectures and key technologies," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 55–61, Apr. 2019.

[191] V. W. Wong, R. Schober, D. W. K. Ng, and L. C. Wang, "Key technologies for 5G wireless systems." Cambridge, U.K.: Cambridge Univ. Press, 2017.

[192] Y. Zeng, J. Lyu, and R. Zhang, "Cellular-connected UAV: Potential, challenges and promising technologies," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 120–127, Feb. 2019.

[193] P. Yu, W. Li, F. Zhou, L. Feng, M. Yin, S. Guo, Z. Gao, and X. Qiu, "Capacity enhancement for 5G networks using MmWave aerial base stations: Self-organizing architecture and approach," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 58–64, Aug. 2018.

[194] A. Trotta, M. D. Felice, F. Montori, K. R. Chowdhury, and L. Bononi, "Joint coverage, connectivity, and charging strategies for distributed UAV networks," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 883–900, Aug. 2018.

[195] T. D. Hoang, L. B. Le, and T. Le-Ngoc, "Energy-efficient resource allocation for D2D communications in cellular networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, p. 6972–6986, Sep. 2016.

[196] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.

[197] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization.* John Wiley and Sons, Inc., New York, 1997.

[198] M. R. Garey and D. S. Johnson, "Computers and intractability: A guide to the theory of NP-completeness." 1st ed. New York, USA: Freeman, 1979, pp. 50–53.

[199] J.-M. Kang and C.-J. Chun, "Joint trajectory design, Tx power allocation, and Rx power splitting for UAV-enabled multicasting SWIPT systems," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3740–3743, Sep. 2020.

[200] Q.-V. Pham, N. Iradukunda, N. H. Tran, W.-J. Hwang, and S.-H. Chung, "Joint placement, power control, and spectrum allocation for UAV wireless backhaul networks," *IEEE Netw. Lett.*, vol. 3, no. 2, pp. 56–60, Jun. 2021.

[201] T. Ma, H. Zhou, B. Qian, N. Cheng, X. Shen, X. Chen, and B. Bai, "UAV-LEO integrated backbone: A ubiquitous data collection approach for B5G internet of remote things networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3491–3505, Nov. 2021.

[202] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, and X. Shen, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.

[203] Y. Zhao, Z. Li, N. Cheng, B. Hao, and X. Shen, "Joint UAV position and power optimization for accurate regional localization in space-air integrated localization network," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4841–4854, Mar. 2021.

[204] I. Leyva-Mayorga, B. Soret, and P. Popovski, "Inter-plane inter-satellite connectivity in dense LEO constellations," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3430–3443, Jun. 2021.

[205] M. M. Gost, I. Leyva-Mayorga, A. Perez-Neira, M. Vazquez-Angel, B. Soret, and M. Moretti, "Edge computing and communication for energy-efficient earth surveillance with LEO satellites," Nov. 2021. [Online]. Available: https://arxiv.org/abs/2111.09045

[206] M. D. Nguyen, L. B. Le, and A. Girard, "Joint computation offloading, UAV trajectory, user scheduling, and resource allocation in SAGIN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022, to appear.

[207] S. Cakaj, "The parameters comparison of the "Starlink" LEO satellites constellation for different orbital shells," *Frontiers Commun. Netw.*, vol. 2, p. 7, 2021.

[208] M. D. Nguyen, L. B. Le, and A. Girard, "Integrated UAV trajectory control and resource allocation for UAV-based wireless networks with co-channel interference management," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12 754–12 769, Jul. 2022.

[209] M. D. Nguyen, L. B. Le, and A. Girard, "UAV placement and resource allocation for intelligent reflecting surface assisted UAV-based wireless networks," *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1106–1110, May 2022.

[210] M. D. Nguyen, L. B. Le, and A. Girard, "Computation offloading, UAV placement, and resource allocation in SAGIN," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, to appear.

[211] M. D. Nguyen, L. B. Le, and A. Girard, "Integrated computation offloading, UAV trajectory control, user scheduling, resource allocation, and admission control in SAGIN," submitted to *IEEE Trans. Veh. Technol.*, Oct. 2022.

[212] K. Guo and K. An, "On the performance of RIS-assisted integrated satellite-UAV-terrestrial networks with hardware impairments and interference," *IEEE Wireless Commun. Lett.*, vol. 11, no. 1, pp. 131–135, Jan. 2022.

[213] B. Zheng, S. Lin, and R. Zhang, "Intelligent reflecting surface-aided LEO satellite communication: Cooperative passive beamforming and distributed channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 3057–3070, Oct. 2022.

[214] S. Xu, J. Liu, T. K. Rodrigues, and N. Kato, "Robust multiuser beamforming for IRS-enhanced near-space downlink communications coexisting with satellite system," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14 900–14 912, Aug. 2022.

[215] T. V. Nguyen, H. D. Le, and A. T. Pham, "On the design of RIS-UAV relay-assisted hybrid FSO/RF satellite-aerial-ground integrated network," *IEEE Trans. Aerospace Electron. Syst.*, 2022, early access.

[216] S.-N. Jin, D.-W. Yue, and H. H. Nguyen, "RIS-aided cell-free massive MIMO system: Joint design of transmit beamforming and phase shifts," *IEEE Syst. J.*, 2022, early access.

[217] Z. Peng, X. Chen, C. Pan, M. Elkashlan, and J. Wang, "Performance analysis and optimization for RIS-assisted multi-user massive MIMO systems with imperfect hardware," *IEEE Trans. Veh. Technol.*, 2022, early access.

[218] L. Hu, Z. Wang, H. Zhu, and Y. Zhou, "RIS-assisted over-the-air federated learning in millimeter wave MIMO networks," *J. Commun. Inform. Netw.*, vol. 7, no. 2, pp. 145–156, Jun. 2022.

[219] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, Jan. 2020.

[220] P. McEnroe, S. Wang, and M. Liyanage, "A survey on the convergence of edge computing and AI for UAVs: Opportunities and challenges," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15 435–15 459, 2022.

[221] S. Rezwan and W. Choi, "Artificial intelligence approaches for UAV navigation: Recent advances and future challenges," *IEEE Access*, vol. 10, pp. 26 320–26 339, Mar. 2022.

[222] F. Tang, C. Wen, M. Zhao, and N. Kato, "Machine learning for space–air–ground integrated network assisted vehicular network: A novel network architecture for vehicles," *IEEE Veh. Technol. Mag.*, vol. 17, no. 3, pp. 34–44, Sep. 2022.