On Supporting Multiservices in UAV-Enabled Aerial Communication for Internet of Things

Hamed Hellaoui [iD], *Member, IEEE*, Miloud Bagaa [iD], *Senior Member, IEEE*, Ali Chelli [iD], *Member, IEEE*, Tarik Taleb, *Senior Member, IEEE*, and Bin Yang [iD], *Member, IEEE*

Abstract—Multiservices are of fundamental importance in unmanned aerial vehicle (UAV)-enabled aerial communications for the Internet of Things (IoT). However, the multiservices are challenging in terms of requirements and use of shared resources such that the traditional solutions for a single service are unsuitable for the multiservices. In this article, we consider a UAV-enabled aerial access network for ground IoT devices, each of which requires two types of services, namely, ultrareliable low-latency communication (uRLLC) and enhanced mobile broadband (eMBB), measured by transmission delay and effective rate, respectively. We first consider a communication model that accounts for most of the propagation phenomena experienced by wireless signals. Then, we derive the expressions of the effective rate and the transmission delay, and formulate each service type as an optimization problem with the constraints of resource allocation and UAV deployment to enable multiservice support for the IoT. These two optimization problems are nonlinear and nonconvex and are generally difficult to be solved. To this end, we transform them into linear optimization problems, and propose two iterative algorithms to solve them. Based on them, we further propose a linear program algorithm to jointly optimize the two service types, which achieves a tradeoff of the effective rate and the transmission delay. Extensive performance evaluations have been conducted to demonstrate the effectiveness of the proposed approach in reaching a tradeoff optimization that enhances the two services.

Index Terms—5G and beyond 5G, aerial communication, cellular networks, Internet of Things (IoT), multiservices support, unmanned aerial vehicles (UAVs).

Hamed Hellaoui is with the Communications and Networking Department, Aalto University, 02150 Espoo, Finland (e-mail: hamed.hellaoui@aalto.fi).

Miloud Bagaa is with the Communications and Networking Department, Aalto University, 02150 Espoo, Finland, and also with the Department of Electrical and Computer Engineering, Université du Québec à Trois-Rivières, Trois-Rivières, QC G8Z 4M3, Canada (e-mail: miloud.bagaa@aalto.fi).

Ali Chelli is with the Department of Business, Marketing and Law, University of South-Eastern Norway (Campus Ringerike), 3679 Notodden, Norway (e-mail: ali.chelli@usn.no).

Tarik Taleb is with the Information Technology and Electrical Engineering, University of Oulu, 90570 Oulu, Finland (e-mail: tarik.taleb@oulu.fi).

Bin Yang is with the School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China (e-mail: yangbinchi@gmail.com).

I. Introduction

CONSIDERED as one of the technologies that are reshaping our daily lives, unmanned aerial vehicles (UAVs) are getting more attention. According to *Valuates Reports* [2], the global UAV market size is projected to reach U.S. $133.5 billion by 2026, at a compound annual growth rate (CAGR) of 26.4% during 2021–2026. This growth will be associated with new applications and services never experienced before. More precisely, when equipped with adequate radio access technologies (RATs), UAVs can operate as flying base stations (BSs) to serve ground users, such as the Internet of Things (IoT). Their ability to move enables on-demand and easy to deploy aerial access solutions for ground devices that can be used to support terrestrial communication or to provide connectivity in less covered areas. Recently, UAVs have been recognized as a key enabler of various IoT services and applications in the era of 5G and beyond.

The use of IoT technology in various applications is diverse, one example being in mission-critical scenarios like wildfire monitoring. In such cases, firefighters utilize special IoT devices which are equipped with a camera, to provide video streaming on the uplink, and also with a temperature and wind sensors to provide near real-time measurements [3]. These two services will be used to build a fire map, allowing efficient coordination of the firefighters' efforts to surround and counter the fire. Given the criticality of this application, where fires can spread very fast, strict performance guarantees are required for real-time measurements (reduced delay) and the video streaming service (high throughput). This type of application requires multiservice support for ultraReliable Low Latency Communication (uRLLC) and enhanced Mobile Broadband (eMBB) services. Another typical application that requires multiservice support is enhancing the connectivity during a crowded event (e.g., in a concert event) [3], where the network is required to ensure a high throughput service on the uplink (for live streaming) while supporting a massive number of users at the same time. This type of application requires multiservice support for massive Machine-Type Communication (mMTC) and eMBB services. Traditional methods for handling a single service are no longer sufficient for handling these distinct service types in the same time. Such an issue has derived the development of the concept of network slicing, which is one of the pillars of the next generation of mobile networks (6G). As a result, many service types are distinguished, namely, uRLLC, eMBB, holographic, tactile communications, and mMTC. Many studies have focused on

multiservice support at the virtualized resources level by leveraging network function virtualization (NFV) and software-defined networking (SDN). However, fewer works tackle the part between the connected users and the access network. By doing so, we enable the support of different service types for the same device. This partly reflects an essential puzzle that allows unlocking the concept of transmission slicing.

The proposed work, in this article, tackles the support of multiservices in UAV-enabled aerial communication for the IoT, where each IoT device requires two different service types. A primary version of this work has been proposed in [1]. For the sake of simplicity and without losing the generality, we consider two services: 1) uRLLC and 2) eMBB, that have conflicting requirements. The presence of several services with different requirements and shared resources makes traditional solutions for a single service unsustainable to support multiservices. While an uRLLC service seeks to reduce the delay, an eMBB service seeks to enhance the rate. To this end, we consider the joint problem of resource allocation and UAV deployment to improve the two services. The optimization is conducted to minimize the transmission delay for uRLLC services, and the effective rate is maximized for eMBB services. The effective rate is defined as the achieved rate at the UAV, providing a better Quality-of-Service (QoS) evaluation. Furthermore, we also investigate the effect of packet retransmission, which is caused by failure of reception. We therefore consider an automatic repeat request (ARQ) scheme in the two service types. By jointly optimizing resource allocation and UAV deployment, enhanced QoS can be reached for each service type per IoT device.

The main contributions of this article are the following.

1) We consider a UAV-enabled aerial access network, where each ground IoT node requires two different service types measured by the effective rate and the transmission delay. In such a network, we adopt a communication model that accounts for most of the propagation phenomena experienced by wireless signals.
2) We derive the expressions of the effective rate and the transmission delay, and formulate each service type as an optimization problem with the constraints of resource allocation and UAV deployment. We also propose two iterative algorithms to solve them.
3) Based on them, we further propose a linear program algorithm to jointly optimize the two service types, which achieves a tradeoff of the effective rate and the transmission delay.
4) Extensive performance evaluations are presented to illustrate the effectiveness of the proposed approach in reaching a tradeoff optimization that enhances the two services.

The remainder of this article is organized in the following fashion. Section II presents some works on cellular UAVs. The system model are provided in Section III. Section IV gives the performance metrics and problem formulation. The proposed solution for resource allocation and UAV deployment to support uRLLC and eMBB for the IoT is introduced in Section V. Performance evaluations are provided in Section VI. This article concludes in Section VII.

II. RELATED WORKS

The potential of UAV applications has attracted attention from both industrial and academic communities. The 3rd Generation Partnership Project (3GPP) working activities in UAVs, translated into different technical/specification reports, including TR 36.777 [4], TR 22.825 [5], and TS 22.261 [6], demonstrate the interest of mobile network organizations in cellular UAVs. This interest is also materialized in real-field evaluations to investigate communication quality better. Compared to ground devices, the evaluations showed that flying UAVs could have poor link quality and even impact terrestrial communications [4], [7], [8].

Real-field evaluations have paved the way for different contributions in cellular-connected UAVs. Hellaoui et al. [9] considered the downlink scenario and studied the impact of UAV communications on cellular networks. The study highlighted two parameters that can impact cellular communication, namely, the transmission power of the UAVs and the employed subcarriers. Challita et al. [10] considered the problem of path planning for cellular UAVs. A dynamic game solution is proposed to achieve a tradeoff between maximizing the energy efficiency and minimizing the latency and the interference level caused on the ground. Hellaoui et al. [11] studied the uplink communication in cellular-connected UAVs. This scenario is particularly interesting to ensure command and control (C2) links to the UAVs. A joint subcarrier and power allocation approach is proposed to address this issue and enhance C2 links. In another work [12], Guan et al. considered the joint problem of subcarrier and power allocation for UAVs. A weighted mean square error (MSE) formulation is introduced along with a solution based on alternating optimization. Hellaoui et al. [13] studied the application of the principle of connection steering on cellular-connected UAVs for uplink communication. The aim is to take advantage of the availability of several mobile networks within the communication range of the flying UAVs, and to select the one ensuring the best QoS for the UAVs.

On the other hand, different works studied the use of UAVs as flying BSs to provide connectivity to ground users. Such applications are particularly interesting to provide communication support in a specific region or to extend the network coverage to rural areas. Mozaffari et al. [14] interested in the transmit power of UAVs, serving ground users, and the transmission rate requirements of these users. The authors considered transport theory and facility location to address the power minimization problem while satisfying users' requirements. Liu et al. [15] addressed the problem of power allocation for UAV-assisted wireless networks. A price-based power allocation scheme is proposed and a Stackelberg game is considered to model the interaction between the UAVs and the ground users. The problem of optimizing 3-D placement and the mobility of UAVs collecting data from ground IoT is investigated in [16]. The authors proposed a framework for jointly optimizing the 3-D placement and the mobility of the UAVs, device-UAV association, and uplink power control. Li et al. [17] studied the problem of subcarrier and transmission power selection for UAV-enabled wireless

TABLE I
SUMMARY OF NOTATIONS

Notation	Description
$\mathcal{U}$	Set of IoT devices.
$\mathcal{V}$	Set of UAVs.
$\mathcal{B}$	Set of sub-carriers.
$\mathcal{B}_u$	Set of sub-carriers assigned to the IoT node $u \in \mathcal{U}$.
$\mathcal{C}$	Set of service types. $\mathcal{C} = \{r, d\}$ where r refers to the service type uRLLC and d refers to the service type eMBB.
$\mathcal{A}_v$	IoT devices belonging to the cell partition associated to the UAV v.
$\mathcal{L}_v$	Set of locations for the UAV v.
uv	Link between the IoT device $u \in \mathcal{U}$ and its serving UAV $v \in \mathcal{V}$.
$t = 1, \ldots, N_v$	Refers to the N_v interfering nodes affecting the UAV v.
R_u^c	Transmission rate employed by the IoT $u \in \mathcal{U}$ for the service type $c \in \mathcal{C}$.
$R_{u,l,b}^{\text{eff}}$	Effective rate from the IoT device $u \in \mathcal{U}$ to its serving UAV $v \in \mathcal{V}$ deployed at the location $l \in \mathcal{L}_v$ over the sub-carrier b.
$D_{u,l,b,Q}$	Transmission delay between the IoT device u and its serving UAV $v \in \mathcal{V}$ deployed at the location $l \in \mathcal{L}_v$ over Q sub-carriers.
E^c	Maximum number of re-transmission for the service type c.
$\mathbb{E}(T_u)$	Average number of re-transmissions for the IoT device u.
$\mathcal{X}_{v,l}$	Boolean variable that indicates whether the UAV $v \in \mathcal{V}$ will be deployed at the location $l \in \mathcal{L}$.
$\mathcal{Z}_{u,b}^c$	Boolean variable that indicates whether the IoT device u uses the sub-carrier $b \in \mathcal{B}$ for transmitting data related to the service type $c \in \mathcal{C}$.
N-LP-R	Non-Linear Program for optimizing the Rate.
N-LP-D	Non-Linear Program for optimizing the Delay.
N-LP-F	Non-Linear Program for optimizing the rate and delay (Fair solution).
LP-R	Linear Program for optimizing the Rate.
LP-D	Linear Program for optimizing the Delay.
LP-F	Linear Program for optimizing the rate and delay (Fair solution).

communication. An iterative algorithm is proposed along with a Lagrangian dual decomposition method to solve it.

However, IoT applications can be associated with different service types having different requirements. Existing works do not consider the support of multiservices for each IoT node. This underpins this paper's focus in which we consider that each IoT node is requiring two service types, namely, uRLLC and eMBB. To this end, we address the joint problem of resource allocation and UAV deployment in order to maximize the efficiency of each service type per IoT node. The system model and problem formulation are provided in the next section.

III. SYSTEM MODEL

We consider a set $\mathcal{U}$ of IoT devices deployed in a geographical area $\mathcal{A} \subset \mathbb{R}^2$. Each IoT device is sending two types of packets requiring different service types. Let $c \in \mathcal{C} = \{r, d\}$ denote the service type, where r refers to the service type uRLLC and d refers to the service type eMBB. More notations are provided in Table I. Without loss of generality, we reflect the service type uRLLC by the transmission delay, while the service type eMBB is reflected by the effective rate. Furthermore, we also investigate the effect caused by packet retransmission. Indeed, a successful reception requires a random number of packet retransmission. To this end, we consider an ARQ scheme until a successful reception or a maximum number of retransmission E^c for the service $c \in \mathcal{C}$

is reached (each node is equipped with a buffer to store the packets before their transmission). On the other hand, a set $\mathcal{V}$ of UAVs is considered as flying BSs to provide uplink wireless communication to the ground IoT devices. Each of IoT nodes and UAVs has a single antenna. As shown in Fig. 1, the area $\mathcal{A}$ is divided into cell partitions where $\mathcal{A}_v$ refers to the partition that gathers the IoT nodes served by the UAV $v \in \mathcal{V}$. Note that the partitions are disjoint (e.g., $\forall v_1, v_2 \in \mathcal{V}; \mathcal{A}_{v_1} \cap \mathcal{A}_{v_2} = \emptyset$ and $\cup_{v \in \mathcal{V}} \mathcal{A}_v = \mathcal{U}$). In order to effectively serve the ground IoT nodes, each UAV needs to head toward an adequate location and provide connection to its associated IoT nodes. The problem of UAV deployment and subcarrier resource allocation is crucial to enable multiservice support for IoT devices. The 3-D plane where the UAV v can be deployed is denoted $\mathcal{L}_v$. Therefore, we derive in this section the expressions of the effective rate and the transmission delay for the IoT nodes connected to their serving UAVs.

Let us denote by $u \in \mathcal{U}$ the source IoT node and by $v \in \mathcal{V}$ the serving UAV. The latter employs an orthogonal frequency division multiple access (OFDMA) technique to serve the connected devices. Intracell interference is thus neglected and the interference can only be caused by nonserved IoT devices, as shown in Fig. 1. For a node u, the interference originates from nodes in neighboring cell partitions that use the same subcarriers as u. Let $\mathcal{B}$ denote the set of subcarriers. The complex-value fading coefficient for the link uv is denoted by h_{uv} and follows a Nakagami-m distribution. Note that both
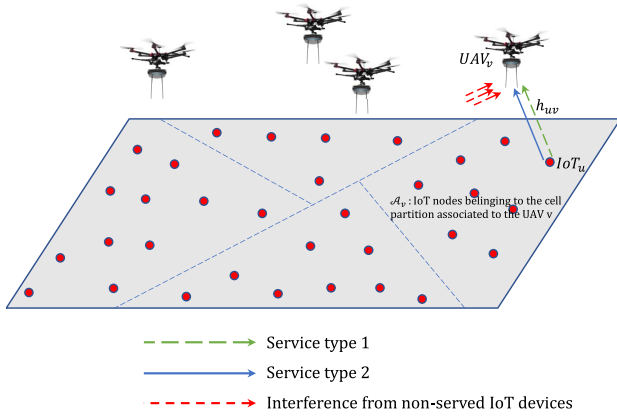
Fig. 1. System model (uplink scenario): each IoT node has two types of traffic, with different requirements, to be sent to the serving UAV.

Line-of-Sight (LoS) and non-LoS conditions can be modeled by adjusting the parameters of the Nakagami-m distribution. Then, the received signal r_v can be expressed as

$$r_v = h_{uv}\sqrt{p_u}x_u + \sum_{t=1}^{N_v} h_{tv}\sqrt{p_t}x_t + n_v. \quad (1)$$

The second term in the right-hand side of (1) represents the interference impact from nodes t ($t = 1, \ldots, N_v$, and N_v is the number of interfering nodes on the UAV v). The nodes u and t transmit the symbols x_u and x_t with the powers p_u and p_t, respectively. As for the third term in (1), n_v, it accounts for a zero-mean additive white Gaussian noise with variance N_0. The instantaneous received signal-to-interference-plus-noise ratio (SINR) for the link uv can be defined as

$$\text{SINR}_{uv} = \frac{\gamma_{uv}}{1 + \sum \gamma_{tv}} \approx \frac{\gamma_{uv}}{\sum \gamma_{tv}} \quad (2)$$

where γ_{uv} and γ_{tv}, respectively, stand for the instantaneous received signal-to-noise ratio (SNR) of the links uv and tv. The approximation in (2) is valid if the noise power can be neglected compared to the interference power. This is generally a well-accepted assumption in the literature and is known as an interference-limited regime. Then, we can express the SNR of uv, γ_{uv}, as

$$\gamma_{uv} = p_u h_{uv}^2 / N_0 \quad (3)$$

and the mean value of γ_{uv}, denoted by $\bar{\gamma}_{uv}$, can be determined as

$$\bar{\gamma}_{uv} = p_u \mathbb{E}\left[h_{uv}^2\right]/N_0 = p_u \times 10^{-\frac{\text{PL}_{uv}}{10}}/N_0 \quad (4)$$

where $\mathbb{E}[h_{uv}^2]$ reflects the channel variance and $\mathbb{E}[\cdot]$ stands for the expectation operator. The former can be computed as $\mathbb{E}[h_{uv}^2] = 10^{-(\text{PL}_{uv}/10)}$, where PL_{uv} is the path loss in dB scale. In this study, we consider the path-loss model adopted by 3GPP [4] as

$$\text{PL}_{uv} = 28.0 + 22\ \log_{10}\left(d_{uv}^{3D}\right) + 20 \log_{10}(f_c). \quad (5)$$

d_{uv}^{3D} reflects the Euclidean distance between the transmitter and the receiver, while f_c accounts for the carrier frequency.

To denote that the random variable (RV) X follows a Gamma distribution with parameters α and β, we use the shorthand notation $X \sim \mathcal{G}(\alpha, \beta)$. The SNR γ_{uv} is Gamma distributed with parameters α_{uv} and $\beta_{uv} = \bar{\gamma}_{uv}/\alpha_{uv}$ and can thus be expressed as $\gamma_{uv} \sim \mathcal{G}(\alpha_{uv}, \beta_{uv})$. The total interference at the UAV, $\gamma_{I_v} = \sum_{t=1}^{N_v} p_t h_{tv}^2 = \sum_{t=1}^{N_v} \gamma_{I_{tv}}$, is the sum of N_v independent nonidentically distributed Gamma RV, with $t = 1, \ldots, N_v$ refers to the N_v interfering nodes affecting the UAV v and $\gamma_{I_{tv}} \sim \mathcal{G}(\alpha_{tv}, \beta_{tv})$. In addition, the probability density function (PDF) of γ_{I_v} can be approximated by a Gamma distribution with parameters α_v and β_v [i.e., $\gamma_{I_v} \sim \mathcal{G}(\alpha_v, \beta_v)$], with

$$\alpha_v = \frac{\left(\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}\right)^2}{\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}^2} \quad (6)$$

$$\beta_v = \frac{\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}^2}{\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}}. \quad (7)$$

IV. PERFORMANCE METRICS AND PROBLEM FORMULATION

To evaluate each service type between an IoT node and the serving UAV, we first define two performance metrics and then derive their expressions. We further provide the problem formulation of performance optimization via subcarrier resource allocation and UAV deployment.

A. Performance Metrics

Effective Rate: It is defined as the achievable rate without occurring a packet transmission failure from an IoT node to its receiving UAV.

Transmission Delay: It is defined as the time duration from the time when a packet arrives at an IoT node to the time when the packet is received by the UAV.

The expressions of these two performance metrics are provided in the following theorems.

Theorem 1 (Effective Rate): For a node $u \in \mathcal{U}$ transmitting data on ARQ mode with a rate R_u^r over the subcarrier b to its serving UAV $v \in \mathcal{V}$ deployed at the location $l \in \mathcal{L}_v$, the average effective rate at the receiving UAV can be expressed as

$$R_{u,l,b}^{\text{eff}} = \frac{R_u^r \times (\mathcal{I}_u)^2}{1 - (1 - \mathcal{I}_u)^{E^r}}. \quad (8)$$

The function $\mathcal{I}_u = \mathcal{I}(2^{R_u^r} - 1, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v)$ is expressed as

$$\mathcal{I}(x, \alpha, \beta, \alpha_v, \beta_v) = \left(\frac{x\beta_v}{\beta}\right)^{-\alpha_v} \frac{\Gamma(\alpha + \alpha_v)}{\Gamma(\alpha)\Gamma(1 + \alpha_v)} {}_2F_1\left(\alpha_v, \alpha + \alpha_v, 1 + \alpha_v, \frac{-\beta}{x\beta_v}\right) \quad (9)$$

where ${}_2F_1(a, b, c, z)$ is the Gauss hypergeometric function.

Proof: See the Appendix. ∎

The above theorem provides the effective rate on the uplink scenario. The underlying equations consider path loss, fast fading, and interference. The theorem also considers the outage probability, which is expressed as $1 - \mathcal{I}_u$ as detailed in the Appendix. We also derive the expression of the transmission delay on the uplink. As mentioned earlier, each node is equipped

with a buffer to store the packets before their transmission. The use of buffers allows to control the packet flow and is considered as the main source for the delay [18]. The data generated by the node u for the delay-sensitive service (uRLLC) are assumed to follow a Poisson distribution with parameter λ_u. In order to model the delay over multi subcarriers, we consider a parallel $M/M/1$ queuing model where the traffic is equitably shared among the different queues. Consequently, the arrival rate λ_u of the IoT node u will be divided on the number of parallel queues Q_u. Therefore, the expected delay for the direct communication, between the node u and its serving UAV v over the subcarrier b, can be expressed as follows.

Theorem 2 (Transmission Delay): For a node $u \in \mathcal{U}$ transmitting data on ARQ mode with a rate R_u^d over the subcarrier b to its serving UAV $v \in \mathcal{V}$ deployed at the location $l \in \mathcal{L}_v$, the transmission delay can be expressed as

$$
D_{u,l,b,Q_u} = \frac{\lambda_u T_F^2}{Q_u 2(1-\rho_u)} \left(\frac{1 - (2E^d - 1)(1 - \mathcal{I}_u)^{E^d}}{\mathcal{I}_u} \right.
$$
$$
\left. + \frac{2(1 - \mathcal{I}_u)\left(1 - (1 - \mathcal{I}_u)^{E^d - 1}\right)}{\mathcal{I}_u^2} \right)
$$
$$
+ \frac{T_F}{2} + \frac{1 - (1 - \mathcal{I}_u)^{E^d}}{\mathcal{I}_u} T_F \qquad (10)
$$

where T_F refers to time required for a single transmission of a fixed size uRLLC packet, $\mathcal{I}_u = \mathcal{I}(2^{R_u^d} - 1, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v)$ and ρ_u is provided as

$$
\rho_u = \frac{\lambda_u \left(1 - (1 - \mathcal{I}_u)^{E^d}\right) T_F}{Q_u \mathcal{I}_u}. \qquad (11)
$$

Proof: See the Appendix. ∎

Note that the transmission delay expression formulated in Theorem 2 includes the queuing delay, the delivery delay and the effect of retransmission. It also considers the outage probability, which is expressed as $1 - \mathcal{I}_u$ as detailed in the Appendix. The expressions provided in Theorems 1 and 2 are original and can not be found in the literature. The average delay over all the used subcarriers by the IoT node u can therefore be expressed as

$$
\mathrm{Avr}(D_{u,l,Q_u}) = \sum_{b \in \mathcal{B}} \frac{1}{Q} D_{u,l,b,Q_u} = \sum_{b \in \mathcal{B}} \bar{D}_{u,l,b,Q_u}. \qquad (12)
$$

The expressions of the effective rate and the transmission delay in Theorems 1 and 2 consider unchanged fading coefficients between the IoT nodes and the UAVs. Indeed, if the duration of the time slot is much smaller than the coherence time of the channel, this channel is constant during the whole time slot duration [19], [20]. On the other hand, for systems operating at a carrier frequency of 2.5 GHz where the receiver is moving with a speed of 100 km/h, the coherence time is equal to 4 ms [21]. Note that most of the frequencies in 5G are smaller than 2.5 GHz. Furthermore, the time slot duration in 5G standards is smaller than 0.5 ms when considering subcarrier spacing higher than 30 kHz [22]. Therefore, given the fact that the deployed IoT nodes do not move, while the UAVs operate at a low speed (to position in their locations),

the assumption that the time slot is much smaller than the coherence time holds.

B. Problem Formulation

In order to enable multiservice support for ground IoT devices, this article addresses the joint problem of UAV deployment and resource allocation. To this end, we first start by modeling the problem as a nonlinear integer program. We define the Boolean variable $\mathcal{X}_{v,l}$ that indicates whether the UAV $v \in \mathcal{V}$ will be deployed at the location $l \in \mathcal{L}_v$ as

$$
\mathcal{X}_{v,l} = \begin{cases} 1, & \text{if the UAV } v \text{ will be deployed at the} \\ & \text{location } l \in \mathcal{L}_v \\ 0, & \text{otherwise.} \end{cases} \qquad (13)
$$

As for the problem of resource allocation, we define the variable $\mathcal{Z}_{u,b}^c$ as

$$
\mathcal{Z}_{u,b}^c = \begin{cases} 1, & \text{if the IoT device } u \text{ uses the subcarrier } b \in \mathcal{B} \\ & \text{for transmitting data related to the service} \\ & \text{type } c \in \mathcal{C} \\ 0, & \text{otherwise.} \end{cases} \qquad (14)
$$

Considering the service type eMBB, the corresponding optimization problem for optimizing the effective rate, N-LP-R (which stands for nonlinear program for optimizing the rate), can be formulated as follows:

N-LP-R:

$$
\underset{\{\mathcal{X}_{v,l}\}, \{\mathcal{Z}_{u,b}^c\}}{\text{maximize}} \min_{u \in \mathcal{U}} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \mathcal{X}_{v,l} \mathcal{Z}_{u,b}^r R_{u,l,b}^{\text{eff}} \qquad (15)
$$

s.t.

$$
\forall v \in \mathcal{V} \quad \forall l \in \mathcal{L}_v; \mathcal{X}_{v,l} \in \{0, 1\} \qquad (16)
$$

$$
\forall u \in \mathcal{U} \quad \forall c \in \mathcal{C} \quad \forall b \in \mathcal{B}; \mathcal{Z}_{u,b}^c \in \{0, 1\} \qquad (17)
$$

$$
\forall v \in \mathcal{V}; \sum_{l \in \mathcal{L}_v} \mathcal{X}_{v,l} = 1 \qquad (18)
$$

$$
\forall u \in \mathcal{U} \quad \forall c \in \mathcal{C}; \sum_{b \in \mathcal{B}} \mathcal{Z}_{u,b}^c \geq 1 \qquad (19)
$$

$$
\forall v \in \mathcal{V} \quad \forall b \in \mathcal{B}; \sum_{c \in \mathcal{C}} \sum_{u \in \mathcal{A}_v} \mathcal{Z}_{u,b}^c \leq 1. \qquad (20)
$$

The objective function (15) of the above optimization problem aims to maximize the minimum effective rate for the set $\mathcal{U}$ of IoT devices, while ensuring constraints (16)–(20). Constraints (16) and (17) limit the value of the decision variables $\mathcal{X}_{v,l}$ and $\mathcal{Z}_{u,b}^c$ to $\{0, 1\}$. Constraint (18) forces each UAV to choose one and only one location. Constraint (19) states that a node u will use a certain number of subcarriers for transmitting data related to the service type c to its serving UAV. Constraint (20) ensures that a subcarrier b within one cell partition $\mathcal{A}_v$ will be used at most by one node.

As for the service type uRLLC, the aim is to reduce the transmission delay. Following the same logic considered for the service type eMBB, the optimization problem for optimizing the delay, N-LP-D (which stands for nonlinear program for optimizing the delay), can be expressed as

N-LP-D:

$$
\underset{\{\mathcal{X}_{v,l}\}, \{\mathcal{Z}_{u,b}^c\}}{\text{minimize}} \max_{u \in \mathcal{U}} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \mathcal{X}_{v,l} \mathcal{Z}_{u,b}^d \sum_{1 \leq i \leq |\mathcal{B}|} P_{u,i} \bar{D}_{u,l,b,i} \qquad (21)
$$

s.t.

$(16)-(20)$

$$\forall u \in \mathcal{U}; Q_u = \sum_{b \in \mathcal{B}} \mathcal{Z}_{u,b}^d \tag{22}$$

$$\forall u \in \mathcal{U} \quad \forall i \in [1, \dots, |\mathcal{B}|]; P_{u,i} \in \{0, 1\} \tag{23}$$

$$\forall u \in \mathcal{U}; \sum_{1 \le i \le |\mathcal{B}|} P_{u,i} = 1 \tag{24}$$

$$\forall u \in \mathcal{U} \quad \forall i \in [1, \dots, |\mathcal{B}|]; Q_u \le i + (1 - P_{u,i}) \times \infty \tag{25}$$

$$\forall u \in \mathcal{U} \quad \forall i \in [1, \dots, |\mathcal{B}|]; i \le Q_u + (1 - P_{u,i}) \times \infty. \tag{26}$$

The above optimization problem aims to minimize the delay for the set of IoT devices $\mathcal{U}$, by selecting the optimal allocation of subcarriers and locations of the serving UAVs. As captured in the objective function (21), the delay for the IoT node u is expressed as the average from the Q_u used subcarriers for the node u, as defined in (12); indeed, condition (22) enables computing the total number of selected subcarriers for each IoT node u. $P_{u,i}$ is a binary variable, as specified in condition (23), that indicates the number of selected subcarriers; i.e., $P_{u,i} = 1 \iff Q_u = i$. The delay function is expressed as defined in (12). Finally, condition (24) forces one $P_{u,i}$ to equal to 1 for each node u, while conditions (25) and (26) ensure that this corresponds to the case where $Q_u = i$.

However, the optimization problem N-LP-R is not linear, which is due to the expression of the effective rate in the objective function (15). Indeed, computing the effective rate for an IoT node u depends on the chosen subcarriers by this node but also on the selected subcarriers by nodes connected to the other UAVs, as the expression of the effective rate considers the interference impact. Moreover, the objective function (15) expresses a product of variables ($\mathcal{X}_{v,l}$ and $\mathcal{Z}_{u,b}^d$). On the other hand, the optimization problem N-LP-D is also not linear. This is due to the objective function (21), that expresses a product of variables ($\mathcal{X}_{v,l}$, $\mathcal{Z}_{u,b}^d$, and $P_{u,i}$), and also to the expression of the delay which is not linear.

Furthermore, the two optimization problems N-LP-R and N-LP-D will optimize each service type separately and will not reach a tradeoff solution optimizing the two service types. In the next section, we introduce our solution to jointly optimize subcarrier allocation and UAV deployment in a way to enhance multiservices in aerial communication for the IoT.

V. Multiservice Support Solution for in Aerial Communication for the IoT

This section introduces the proposed solution for joint resource allocation and UAV deployment to support multiservices in cellular communication for the IoT. The proposed solution relies on the optimization problems N-LP-R and N-LP-D defined in the previous section. To this end: 1) we introduce a set of transformations allowing to linearize the constraints in the previous optimization problems; 2) we also propose an iterative algorithm allowing to linearize the expressions of the effective rate and the transmission delay; and furthermore; and 3) a tradeoff solution is thereafter provided to jointly optimize the two service types.

Fig. 2 illustrates the relationship between the different optimization problems. The optimization problems N-LP-R
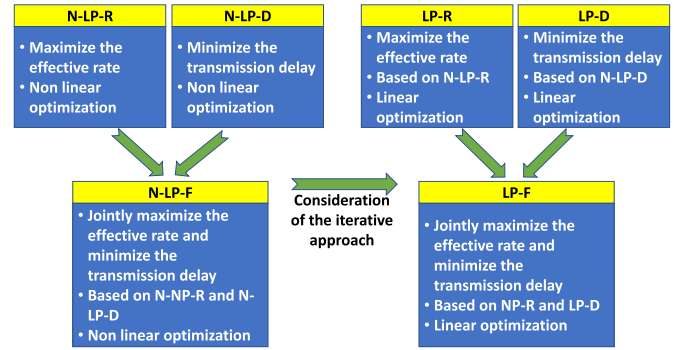

Fig. 2. Relationship between the different optimization problems.

and N-LP-D aim to enhance the effective rate and the transmission delay, respectively. Based on the N-LP-R and N-LP-D, we propose a Nonlinear Program N-LP-F, which can obtain a tradeoff of the effective rate and the transmission delay. Note that these three optimization problems are not linear. Thus, we further linearize these optimization problems. Specifically, we use a Linear Program LP-R to improve the effective rate performance, and use a Linear Program LP-D to improve the transmission delay performance. Based on the LP-R and LP-D, we further use a Linear Program LP-F to achieve the tradeoff of these two performance metrics.

A. LP-R Optimization Problem and Algorithm

Considering N-LP-R, the objective function expresses the product of Boolean variables ($\mathcal{X}_{v,l}$ and $\mathcal{Z}_{u,b}^d$). We therefore define a new Boolean variable $T_{u,l,b}^r = \mathcal{X}_{v,l} \mathcal{Z}_{u,b}^r$ which will be forced to equal 1 when $\mathcal{X}_{v,l} = \mathcal{Z}_{u,b}^r = 1$ by considering the following constraints:

$$T_{u,l,b}^r \le \mathcal{X}_{v,l} \tag{27}$$

$$T_{u,l,b}^r \le \mathcal{Z}_{u,b}^r \tag{28}$$

$$T_{u,l,b}^r \ge \mathcal{X}_{v,l} + \mathcal{Z}_{u,b}^r - 1. \tag{29}$$

The two conditions (27) and (28) together will force $T_{u,l,b}^r$ to 0 when $\mathcal{X}_{v,l}$ or $\mathcal{Z}_{u,b}^r$ is equal to 0. As for condition (29), it forces $T_{u,l,b}^r$ to 1 when both $\mathcal{X}_{v,l}$ and $\mathcal{Z}_{u,b}^r$ are equal to 1. However, the expression of the effective rate is not linear and is more complex. The underlying complexity is due to the fact the expression of the effective rate depends on the interfering nodes using the same subcarriers. Thus, computing the effective rate also depends on the values of the decision variables $\mathcal{Z}_{u,b}^c$ associated to the nodes connected to the other UAVs.

To tackle this issue, we propose an iterative process where each iteration consists of linear optimization problems. Indeed, when optimizing the effective rate only for the IoT nodes connected to a given UAV v, the objective function becomes linear. Here, the decision variables, which are related to the nodes connected to the other UAVs, are not being changed. We can enhance the effective rate of the served IoT devices $u \in \mathcal{A}_v$ by optimizing the UAV deployment and subcarrier allocation for a given UAV $v \in \mathcal{V}$. The following optimization problem LP-R is proposed for an iteration, where v refers to the UAV

in question and C_{total} is a constant

$$\textbf{LP-R}(v, C_{\text{total}}):$$

$$\underset{\{\mathcal{X}_{v,l}\},\left\{\mathcal{Z}_{u,b}^c\right\}}{\text{maximize}} \min_{u \in \mathcal{A}_v} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} T_{u,l,b}^r R_{u,l,b}^{\text{eff}}, \tag{30}$$

$$\textbf{s.t.}$$

$$\forall l \in \mathcal{L}_v; \mathcal{X}_{v,l} \in \{0, 1\} \tag{31}$$

$$\forall u \in \mathcal{A}_v \quad \forall c \in \mathcal{C} \quad \forall b \in \mathcal{B}; \mathcal{Z}_{u,b}^c \in \{0, 1\} \tag{32}$$

$$\sum_{l \in \mathcal{L}_v} \mathcal{X}_{v,l} = 1 \tag{33}$$

$$\forall u \in \mathcal{A}_v \quad \forall c \in \mathcal{C}; \sum_{b \in \mathcal{B}} \mathcal{Z}_{u,b}^c \geq 1 \tag{34}$$

$$\forall b \in \mathcal{B}; \sum_{c \in \mathcal{C}} \sum_{u \in \mathcal{A}_v} \mathcal{Z}_{u,b}^c \leq 1 \tag{35}$$

$$\forall u \in \mathcal{A}_v \quad \forall l \in \mathcal{L}_v \quad \forall b \in \mathcal{B}; T_{u,l,b}^r \leq \mathcal{X}_{v,l} \tag{36}$$

$$\forall u \in \mathcal{A}_v \quad \forall l \in \mathcal{L}_v \quad \forall b \in \mathcal{B}; T_{u,l,b}^r \leq \mathcal{Z}_{u,b}^r \tag{37}$$

$$\forall u \in \mathcal{A}_v \quad \forall l \in \mathcal{L}_v \quad \forall b \in \mathcal{B}; T_{u,l,b}^r \geq \mathcal{X}_{v,l} + \mathcal{Z}_{u,b}^r - 1 \tag{38}$$

$$\text{SUM\_RATE}(\mathcal{V}) \geq C_{\text{total}}. \tag{39}$$

The objective function (30) in the above optimization problem is derived from the objective function (15) of N-LP-R with focus on the nodes $u \in \mathcal{A}_v$; it aims to maximize the effective rate for these nodes. Moreover, the constraints (31)–(35) are also derived from those of N-LP-R considering the IoT nodes $u \in \mathcal{A}_v$. Note that the linear transformations of the constraints have been considered in LP-R. As for constraint (39), it aims to express a global condition imposing to increase the sum of the effective rate for all the nodes above a constant C_{total}. This would also ensure that optimizing the effective rate for the nodes $u \in \mathcal{A}_v$ will not come at the expense of other nodes. More precisely, the function SUM_RATE($\mathcal{V}$) is defined as

$$\text{SUM\_RATE}(\mathcal{V}) = \sum_{u \in \mathcal{A}_v} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} T_{u,l,b}^r R_{u,l,b}^{\text{eff}}$$
$$+ \sum_{u \in \mathcal{V} \setminus \mathcal{A}_v} \sum_{b \in \mathcal{B}_u} \left(\sum_{\bar{u} \in \mathcal{A}_v} \sum_{c \in \mathcal{C}} \mathcal{Z}_{\bar{u},b}^c R_{u,l_u,b}^{\text{eff}} + \xi_{v,b} R_{u,l_u,b}^{\text{eff}} \right) \tag{40}$$

where the first term of the right-hand side of (40) refers to the sum of the effective rate for the nodes $u \in \mathcal{A}_v$, while the second term refers to the sum of the effective rate for the rest of the nodes, i.e., $u \in \mathcal{V} \setminus \mathcal{A}_v$. The latter is defined over the set of subcarriers, $\mathcal{B}_u$, already assigned to the IoT $u \in \mathcal{V} \setminus \mathcal{A}_v$. The term $\sum_{\bar{u} \in \mathcal{A}_v} \sum_{c \in \mathcal{C}} \mathcal{Z}_{\bar{u},b}^c R_{u,l_u,b}^{\text{eff}}$ corresponds to the case that the subcarriers of u will also be used by nodes connected to v, while the term $\xi_{v,b} R_{u,l_u,b}^{\text{eff}}$ corresponds to the opposite case. Here, l_u is the effective location of node $u \in \mathcal{V} \setminus \mathcal{A}_v$. The variable $\xi_{v,b}$ is a Boolean variable. If no node connected to the UAV v is using the subcarrier b, it is equal to 1. Otherwise, it is equal to 0. It is therefore defined by the two following conditions:

$$\forall b \in \mathcal{B}; \xi_{v,b} \geq 1 - \sum_{u \in \mathcal{A}_v} \sum_{c \in \mathcal{C}} \mathcal{Z}_{u,b}^c \tag{41}$$

$$\forall b \in \mathcal{B} \quad \forall u \in \mathcal{A}_v \quad \forall c \in \mathcal{C}; \xi_{v,b} \leq 1 - \mathcal{Z}_{u,b}^c. \tag{42}$$

Algorithm 1 Iterative Algorithm for Optimizing the Effective Rate

Input:
1: $C_{total} = 0$
2: **while** True **do**
3:    Stable = True
4:    **for** $v \in \mathcal{V}$ **do**
5:       LP-R(v, C_{total})
6:       **if** $C_{total} > $ SUM_RATE($\mathcal{V}$) **then**
7:          Stable = False
8:       **end if**
9:       $C_{total} = $ SUM_RATE($\mathcal{V}$)
10:    **end for**
11:    **if** Stable **then**
12:       break
13:    **end if**
14: **end while**

Note that the two previous conditions together will force $\xi_{v,b}$ to 1 only if all the variables $\mathcal{Z}_{u,b}^c$ are equal to 0 for nodes connected to the UAV v using the subcarrier b.

As we can see, an iteration considers the optimization problem LP-R to successively optimize the effective rate for the nodes connected to each UAV $v \in \mathcal{V}$. The iterative process is expressed in Algorithm 1.

Algorithm 1 shows the proposed iterative process for optimizing the effective rate. The linear optimization problem denoted byLP-R(v, C_{total}) will be considered for each UAV $v \in \mathcal{V}$ (lines 4–10). This allows to optimize the effective rate for the served IoT nodes connected to v, while maintaining the total effective sum rate of all the nodes above a constant C_{total}. The latter is updated with the new effective sum rate obtained after each optimization (line 9). This process is repeated until reaching a stability, which reflects a situation that the total effective sum rate can no longer be increased.

B. LP-D Optimization Problem and Algorithm

As for the optimization problem N-LP-D, its objective function includes a product of Boolean variables. To this end, we introduce the Boolean variable $T_{u,l,b,i}^d = \mathcal{X}_{v,l} \mathcal{Z}_{u,b}^d P_{u,i}$. In addition, we define the following conditions allowing to force $T_{u,l,b,i}^d$ to 1 only when $\mathcal{X}_{v,l}$, $\mathcal{Z}_{u,b}^d$ and $P_{u,i}$ are equal to 1 as

$$T_{u,l,b,i}^d \leq \mathcal{X}_{v,l} \tag{43}$$

$$T_{u,l,b,i}^d \leq \mathcal{Z}_{u,b}^d \tag{44}$$

$$T_{u,l,b,i}^d \leq P_{u,i} \tag{45}$$

$$T_{u,l,b,i}^d \geq \mathcal{X}_{v,l} + \mathcal{Z}_{u,b}^d + P_{u,i} - 2. \tag{46}$$

Note that the expression of the delay is not linear since the delay over a subcarrier, as per (12), depends on the interfering nodes using the same subcarrier. In order to tackle this issue, we propose a similar iterative approach as the one introduced in LP-R. Indeed, the objective function becomes linear when optimizing the delay only for the IoT nodes connected to a given UAV. The following optimization problem LP-D is therefore proposed for an iteration, where v refers to the UAV in

question and C_{total} is a constant

LP-D(v, C_{total}) :

$$\underset{\{\mathcal{X}_{v,l}\}, \{\mathcal{Z}_{u,b}^c\}}{\text{minimize}} \max_{u \in \mathcal{A}_v} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \sum_{1 \leq i \leq |\mathcal{B}|} T_{u,l,b,i}^d \bar{D}_{u,l,b,i} \qquad (47)$$

s.t.

$$(31)-(35)$$

$$\forall u \in \mathcal{A}_v; \; Q_u = \sum_{b \in \mathcal{B}} \mathcal{Z}_{u,b}^d \qquad (48)$$

$$\forall u \in \mathcal{A}_v \quad \forall i \in [1, \ldots, |\mathcal{B}|]; \; P_{u,i} \in \{0, 1\} \qquad (49)$$

$$\forall u \in \mathcal{A}_v; \; \sum_{1 \leq i \leq |\mathcal{B}|} P_{u,i} = 1 \qquad (50)$$

$$\forall u \in \mathcal{A}_v \quad \forall i \in [1, \ldots, |\mathcal{B}|]; \; Q_u \leq i + (1 - P_{u,i}) \times \infty \quad (51)$$

$$\forall u \in \mathcal{A}_v \quad \forall i \in [1, \ldots, |\mathcal{B}|]; \; i \leq Q_u + (1 - P_{u,i}) \times \infty \quad (52)$$

$$\forall u \in \mathcal{A}_v \quad \forall l \in \mathcal{L}_v \quad \forall b \in \mathcal{B} \quad \forall i \in [1, \ldots, |\mathcal{B}|]$$

$$T_{u,l,b,i}^d \leq \mathcal{X}_{v,l} \qquad (53)$$

$$T_{u,l,b,i}^d \leq \mathcal{Z}_{u,b}^d \qquad (54)$$

$$T_{u,l,b,i}^d \leq P_{u,i} \qquad (55)$$

$$T_{u,l,b,i}^d \geq \mathcal{X}_{v,l} + \mathcal{Z}_{u,b}^d + P_{u,i} - 2 \qquad (56)$$

$$\text{SUM\_DELAY}(\mathcal{V}) \leq ct. \qquad (57)$$

The optimization problem LP-D is derived from N-LP-D, in the same way as LP-R. This optimization problem focuses on the IoT nodes $u \in \mathcal{A}_v$ and aims to minimize the delay of these nodes. Constraint (57) aims to express a global condition imposing to decrease the sum of the delays of all the nodes below a constant C_{total}. The function SUM_DELAY($\mathcal{V}$) is defined in the same way as the function SUM_RATE($\mathcal{V}$). More precisely, it is defined as

$$\text{SUM\_DELAY}(\mathcal{V}) = \sum_{u \in \mathcal{A}_v} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \sum_{1 \leq i \leq |\mathcal{B}|} T_{u,l,b,i}^d \bar{D}_{u,l,b,i}$$

$$+ \sum_{u \in \mathcal{V} \setminus \mathcal{A}_v} \sum_{b \in \mathcal{B}_u} \left(\sum_{\bar{u} \in \mathcal{A}_v} \sum_{c \in \mathcal{C}} \mathcal{Z}_{\bar{u},b}^c \bar{D}_{u,l_u,b,Q_u} + \xi_{v,b} \bar{D}_{u,l_u,b,Q_u} \right).$$

$$(58)$$

Algorithm 2 shows the iterative process for optimizing the delay of the IoT devices in the network. It is defined in the same way as Algorithm 1; The linear optimization problem (LP-D(v, C_{total})) will be considered for each UAV $v \in \mathcal{V}$ (lines 4–10). This allows to optimize the transmission delay for the served IoT nodes connected to v, while maintaining the total transmission sum delay of all the nodes below a constant C_{total}. The latter is updated with the new transmission sum delay obtained after each optimization (line 9). This process is repeated until reaching a stability, which reflects a situation that the total transmission sum delay can no longer be decreased.

C. LP-F Optimization Problem and Algorithm

As mentioned before, the two service types have different requirements and each of the underlying optimization problems aims to maximize the corresponding service type individually. In order to reach a tradeoff solution, we adopt the

---

Algorithm 2 Iterative Algorithm for Optimizing the Delay

Input:

1: $C_{total} = \max$
2: **while** True **do**
3:     Stable = True
4:     **for** $v \in \mathcal{V}$ **do**
5:         **LP-D**(v, C_{total})
6:         **if** $C_{total} < \text{SUM\_DELAY}(\mathcal{V})$ **then**
7:             Stable = False
8:         **end if**
9:         $C_{total} = \text{SUM\_DELAY}(\mathcal{V})$
10:     **end for**
11:     **if** Stable **then**
12:         break
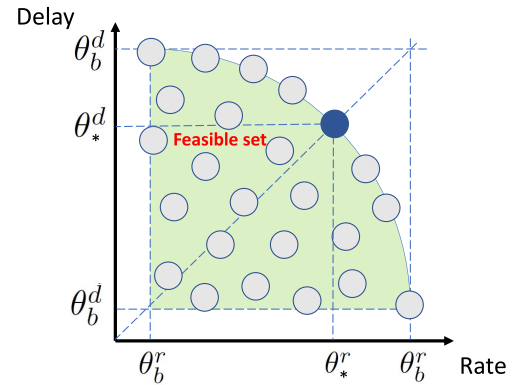13:   **end if**
14: **end while**

Fig. 3. Tradeoff solution.

approach considered in [23]. This allows achieving a tradeoff between the different service types by sharing the same utility function. We introduce a shared utility function, F, and two points, $\theta_b = (\theta_b^r, \theta_b^d)$ and $\theta_d = (\theta_d^r, \theta_d^d)$. $\theta_b = (\theta_b^r, \theta_b^d)$ reflects best utility that can be achieved for each service type, while $\theta_d = (\theta_d^r, \theta_d^d)$ represents the worst one (disagreement). As illustrated in Fig. 3, the tradeoff solution aims to jointly maximizing the effective rate and minimizing the transmission delay. This is performed by sharing the same utility $F = ([\theta_*^r - \theta_d^r]/[\theta_b^r - \theta_d^r]) = ([\theta_d^d - \theta_*^d]/[\theta_d^d - \theta_b^d])$ allowing to maximize the distance between the optimal solution and worst one, while considering the scale of the values of the two objective functions. Let $\dot{\mathcal{X}}$ and $\ddot{\mathcal{X}}$ be the two matrices of $\mathcal{X}_{v,l}$ obtained from resolving LP-R and LP-D, respectively. In addition, let $\dot{\mathcal{Z}}$ and $\ddot{\mathcal{Z}}$ be the two matrices of $\mathcal{Z}_{u,b}^c$ obtained by, respectively, resolving the same optimization problems. Then, $\theta_b = (\theta_b^r, \theta_b^d)$ and $\theta_d = (\theta_d^r, \theta_d^d)$ can be computed as

$$\begin{cases} \theta_b^r = \min_{u \in \mathcal{U}} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \dot{\mathcal{X}}_{v,l} \dot{\mathcal{Z}}_{u,b}^r R_{u,l,b}^{\text{eff}} \\ \theta_d^r = \min_{u \in \mathcal{U}} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \ddot{\mathcal{X}}_{v,l} \ddot{\mathcal{Z}}_{u,b}^r R_{u,l,b}^{\text{eff}} \\ \theta_b^d = \max_{u \in \mathcal{U}} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \ddot{\mathcal{X}}_{v,l} \ddot{\mathcal{Z}}_{u,b}^d \bar{D}_{u,l,b,Q_u} \\ \theta_d^d = \max_{u \in \mathcal{U}} \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \dot{\mathcal{X}}_{v,l} \dot{\mathcal{Z}}_{u,b}^d \bar{D}_{u,l,b,Q_u}. \end{cases} \qquad (59)$$

The tradeoff solution, N-LP-F, can therefore be expressed as follows:

N-LP-F :

$$\underset{\{\mathcal{X}_{v,l}\},\left\{\mathcal{Z}_{u,b}^c\right\}}{\text{maximize}} \; F \tag{60}$$

  s.t.

(16)−(20)

(22)−(26)

$$F = \frac{\vartheta^r - \theta_d^r}{\theta_b^r - \theta_d^r} \tag{61}$$

$$F = \frac{\theta_d^d - \vartheta^d}{\theta_d^d - \theta_b^d} \tag{62}$$

$$\forall u \in \mathcal{U}; \; \vartheta^r \leq \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \mathcal{X}_{v,l} \mathcal{Z}_{u,b}^r R_{u,l,b}^{\text{eff}} \tag{63}$$

$$\forall u \in \mathcal{U}; \; \vartheta^d \geq \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \mathcal{X}_{v,l} \mathcal{Z}_{u,b}^d \sum_{1 \leq i \leq |\mathcal{B}|} P_{u,i} \bar{D}_{u,l,b,i}. \tag{64}$$

To jointly optimize the effective rate and the transmission delay, the optimization problem N-LP-F defines a new objective function F which is shared between the two service types, as reflected in conditions (61) and (62). The variables ϑ^r and ϑ^d in these two conditions, respectively, reflect the minimum effective rate and the maximum transmission delay, as expressed in (63) and (64). Therefore, the objective function of the above optimization problem will jointly maximize the minimum rate and minimize the maximum delay.

However, the optimization problem N-LP-F is not linear, which is due to constraints (63) and (64). We therefore consider the proposed iterative algorithm adopted in LP-R and LP-D. The linear optimization problem LP-F for $v \in \mathcal{V}$ can be expressed as

LP-F $\left(v, C_{\text{total}}^r, C_{\text{total}}^d\right)$:

$$\underset{\{\mathcal{X}_{v,l}\},\left\{\mathcal{Z}_{u,b}^c\right\}}{\text{maximize}} \; F \tag{65}$$

  s.t.

$$F = \frac{\vartheta^r - \theta_d^r}{\theta_b^r - \theta_d^r} \tag{66}$$

$$F = \frac{\theta_d^d - \vartheta^d}{\theta_d^d - \theta_b^d} \tag{67}$$

(31)−(38)

(48)−(56)

$$\forall u \in \mathcal{A}_v; \; \vartheta^r \leq \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} T_{u,l,b}^r R_{u,l,b}^{\text{eff}} \tag{68}$$

$$\forall u \in \mathcal{A}_v; \; \vartheta^d \geq \sum_{l \in \mathcal{L}_v} \sum_{b \in \mathcal{B}} \sum_{1 \leq i \leq |\mathcal{B}|} T_{u,l,b,i}^d \bar{D}_{u,l,b,i} \tag{69}$$

$$\text{SUM\_RATE}(\mathcal{V}) \geq C_{\text{total}}^r \tag{70}$$

$$\text{SUM\_DELAY}(\mathcal{V}) \leq C_{\text{total}}^d. \tag{71}$$

The optimization problem LP-F focuses on the IoT devices $u \in \mathcal{A}_v$ and jointly maximizes the effective rate of the transmission delay of these nodes. The constraints (70) and (71) aim to express global conditions imposing to increase the sum rate and decrease the sum delay for all the IoT nodes.

Algorithm 3 Iterative Algorithm for Joint Rate and Delay Optimization

Input:
1: $C_{total}^r = 0$
2: $C_{total}^d = \max$
3: **while** True **do**
4:     Stable = True
5:     **for** $v \in \mathcal{V}$ **do**
6:       LP-F($v, C_{total}^r, C_{total}^d$)
7:       **if** $C_{total}^r >$ SUM_RATE($\mathcal{V}$) or $C_{total}^d <$ SUM_DELAY($\mathcal{V}$) **then**
8:         Stable = False
9:       **end if**
10:      $C_{total}^r = $ SUM_RATE($\mathcal{V}$)
11:      $C_{total}^d = $ SUM_DELAY($\mathcal{V}$)
12:     **end for**
13:     **if** Stable **then**
14:      break
15:     **end if**
16: **end while**

The iterative process for the tradeoff optimization is shown in Algorithm 3. This algorithm is similar to Algorithms 1 and 2; the linear optimization problem (LP-F($v, C_{\text{total}}^r, C_{\text{total}}^d$)) will be considered for each UAV $v \in \mathcal{V}$ (lines 5–12). This allows to optimize the effective rate and the transmission delay for the served IoT nodes connected to v, while maintaining the total effective sum rate of all the nodes above a constant C_{total}^r and the total transmission sum delay of all the nodes below a constant C_{total}^d. The latter are, respectively, updated with the new effective sum rate and transmission sum delay obtained after each optimization (lines 10 and 11). This process is repeated until reaching a stability, which reflects a situation that the total effective sum rate and the total transmission sum delay can no longer be decreased.

To evaluate the proposed tradeoff approach, we consider a baseline solution. We therefore define the following LP which reflects a multiobjective optimization characterized by the parameter α as:

LP-α $\left(v, C_{\text{total}}^r, C_{\text{total}}^d\right)$:

$$\underset{\{\mathcal{X}_{v,l}\},\left\{\mathcal{Z}_{u,b}^c\right\}}{\text{maximize}} \; F \tag{72}$$

  s.t.

$$F = \alpha \vartheta^r + (1-\alpha)\left(D_{\max} - \vartheta^d\right) \tag{73}$$

$$\begin{cases} (31)-(38) \\ (48)-(56) \\ (68), \; (69) \\ (70), \; (71). \end{cases} \tag{74}$$

The above optimization problem characterizing the baseline solutions is defined in the same way as the proposed LP-F; It introduces a weighting parameter α between the two objective functions [as reflected in the condition (73)] and aims to jointly optimize them using the iterative approach [which is subject

TABLE II
NUMBER OF VARIABLES AND CONDITIONS FOR EACH SOLUTION

Solution	Variables	Number of variables	Number of conditions
LP-R	$\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^r_{u,l,b}$, $\xi_{v,b}$	$(\mid \cup_{v\in\mathcal{V}} \mathcal{L}_v\mid) + (\mid\mathcal{C}\mid \times \mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid) + (\mid\mathcal{V}\mid \times \mid\mathcal{B}\mid)$	$\sum_{v\in\mathcal{V}}(1 + (\mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid) + \mid\mathcal{B}\mid + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid \times 3) + 1 + \mid\mathcal{B}\mid + (\mid\mathcal{B}\mid \times \mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid))$
LP-D	$\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^d_{u,l,b,i}$, $P_{u,i}$, Q_u, $\xi_{v,b}$	$(\mid \cup_{v\in\mathcal{V}} \mathcal{L}_v\mid) + (\mid\mathcal{C}\mid \times \mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid\times\mid\mathcal{L}_v\mid)\times\mid\mathcal{B}\mid^2) + (\mid\mathcal{U}\mid\times\mid\mathcal{B}\mid) + (\mid\mathcal{U}\mid) + (\mid\mathcal{V}\mid \times \mid\mathcal{B}\mid)$	$\sum_{v\in\mathcal{V}}(1 + (\mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid) + \mid\mathcal{B}\mid + (\mid\mathcal{A}_v\mid \times 2) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid^2 \times 4) + 1 + \mid\mathcal{B}\mid + (\mid\mathcal{B}\mid \times \mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid))$
LP-F	$\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^r_{u,l,b}$, $T^d_{u,l,b,i}$, $P_{u,i}$, Q_u, F, ϑ^r, ϑ^d, $\xi_{v,b}$	$(\mid \cup_{v\in\mathcal{V}} \mathcal{L}_v\mid) + (\mid\mathcal{C}\mid \times \mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid) + (\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid^2) + (\mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\mid\mathcal{U}\mid) + 3 + (\mid\mathcal{V}\mid \times \mid\mathcal{B}\mid)$	$\sum_{v\in\mathcal{V}}(1 + (\mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid) + \mid\mathcal{B}\mid + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times 2) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid^2 \times 4) + (\mid\mathcal{A}_v\mid \times 2) + 4 + \mid\mathcal{B}\mid + (\mid\mathcal{B}\mid \times \mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid))$
LP-α	$\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^r_{u,l,b}$, $T^d_{u,l,b,i}$, $P_{u,i}$, Q_u, F, ϑ^r, ϑ^d, $\xi_{v,b}$	$(\mid \cup_{v\in\mathcal{V}} \mathcal{L}_v\mid) + (\mid\mathcal{C}\mid \times \mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid) + (\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid^2) + (\mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\mid\mathcal{U}\mid) + 3 + (\mid\mathcal{V}\mid \times \mid\mathcal{B}\mid)$	$\sum_{v\in\mathcal{V}}(1 + (\mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid) + \mid\mathcal{B}\mid + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times 2) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid^2 \times 4) + (\mid\mathcal{A}_v\mid \times 2) + 3 + \mid\mathcal{B}\mid + (\mid\mathcal{B}\mid \times \mid\mathcal{A}_v\mid \times \mid\mathcal{C}\mid))$

to the set of constraints defined in (74)]. Note that ϑ^r and ϑ^d, respectively, reflect the minimum effective rate and the maximum transmission delay, as defined in (68) and (69). We have considered three values of α which are 0.2, 0.5, and 0.8.

D. Complexity Analysis

We evaluate the complexity of the proposed approach. Indeed, the different optimization solutions are associated with different number of variables, which are detailed as follows: 1) LP-R is associated with the Boolean variables $\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^r_{u,l,b}$, and $\xi_{v,b}$; 2) LP-D is associated with the Boolean variables $\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^d_{u,l,b,i}$, $\xi_{v,b}$, and $P_{u,i}$, in addition to the integer variable Q_u; and 3) as for LP-F and LP-α, they both have the same number of variables. The two solutions are associated with the Boolean variables $\mathcal{X}_{v,l}$, $\mathcal{Z}^c_{u,b}$, $T^r_{u,l,b}$, $T^d_{u,l,b,i}$, $\xi_{v,b}$, and $P_{u,i}$, in addition to the integer variable Q_u, as well as the real variables F, ϑ^r, and ϑ^d. We can see that the two solutions LP-F and LP-α are associated with more variables compared to LP-R and LP-D. More precisely, LP-F has $(\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid^2) + (\mid\mathcal{U}\mid \times \mid\mathcal{B}\mid) + (\mid\mathcal{U}\mid) + 3$ more variables than LP-R and $(\sum_{v\in\mathcal{V}}(\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid) \times \mid\mathcal{B}\mid) + 3$ more variables than LP-D. This is due to the fact that the two latter optimize individual service types, whereas LP-F and LP-α jointly optimize the two services and incorporate variables from LP-R and LP-D. The detailed number of variables in an iteration for each solution is provided in Table II. This table also shows the number of constraints associated to an iteration of each solution. For LP-F and LP-α, they are associated with almost the same number of constraints (LP-F has one additional constraint compared to LP-α). Furthermore, these two solutions have more constraints compared to LP-R and LP-D. More precisely, LP-F has $\sum_{v\in\mathcal{V}}((\mid\mathcal{A}_v\mid \times 2) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid^2 \times 4) + (\mid\mathcal{A}_v\mid \times 2) + 3)$ more constraints than LP-R and $\sum_{v\in\mathcal{V}}((\mid\mathcal{A}_v\mid \times \mid\mathcal{L}_v\mid \times \mid\mathcal{B}\mid \times 3) + (\mid\mathcal{A}_v\mid \times 2) + 3)$ more constraints than LP-R. As mentioned earlier, both LP-F and LP-α aim to jointly optimize the two services and incorporate constraints from LP-R and LP-D. Note that $\mid\mathcal{V}\mid$ has a high impact on the number of constraints. This is due to the

fact that the iterative algorithms operate by performing the optimization for the IoT nodes connected to each UAV $v \in \mathcal{V}$.

VI. PERFORMANCE EVALUATION

This section conducts the performance evaluations of the proposed approach. We first give the parameter settings, and then provide the performance analysis.

A. Parameter Settings

We focus on the network in a 1000 m $\times$ 1000 m square area with varying numbers of UAVs and IoT nodes. We consider a carrier frequency f_c of 2 GHz. The altitude of the UAVs is set between 22.5 and 300 m, which is the feasibility range associated with the path-loss model [4]. The noise variance N_0 is -130 dBm [24]. We consider a single packet transmission time for an uRLLC packet, T_F, of 0.5 ms [25]. The Gurobi [26] is adopted as an optimizer to solve the linear integer programming models. Note that such optimizer can operate in a cloud environment to achieve faster optimization.

B. Performance Analysis

We first evaluate the effective rate and the transmission delay under our proposed approach. To this end, we explore the effect of the transmission rate and the maximum number of retransmission on these two performance metrics. In a network with 30 IoT nodes and 12 UAVs, the simulation results are summarized in Figs. 4 and 5. We can see from Fig. 4 that the average effective rate first increases with the transmission rate, and then decreases after reaching a certain threshold. We can also see from this figure that the average effective rate decreases with the maximum number of retransmission. Indeed, as a consequence of the application of ARQ scheme, the packets are retransmitted until reception or when reaching a maximum number of retransmission; the effective rate will decrease with average number of retransmission as shown in (75) of the Appendix. On the other hand, the averagetransmission delay increases with the transmission rate, as shown in Fig. 5. This is due to the fact that increasing the transmission rate results in higher outage probability and thus
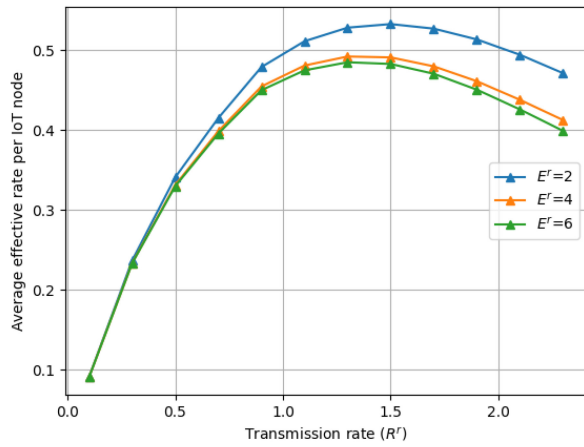
Fig. 4. Effect of the transmission rate and the max number of retransmission on the effective rate.
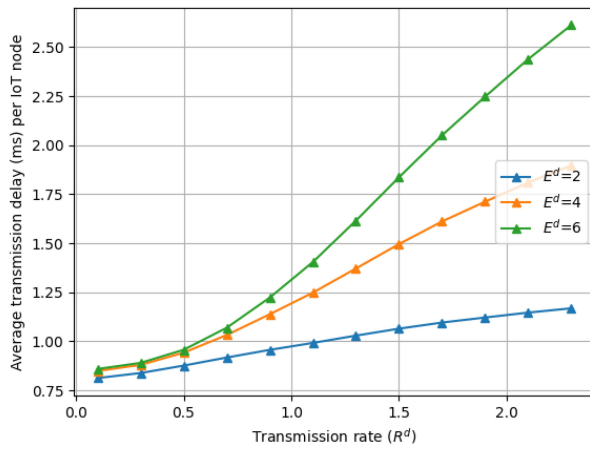

Fig. 6. Evaluation of the effective rate considering LP-R, LP-D, and LP-F.

Fig. 5. Effect of the transmission rate and the max number of retransmission on the transmission delay.
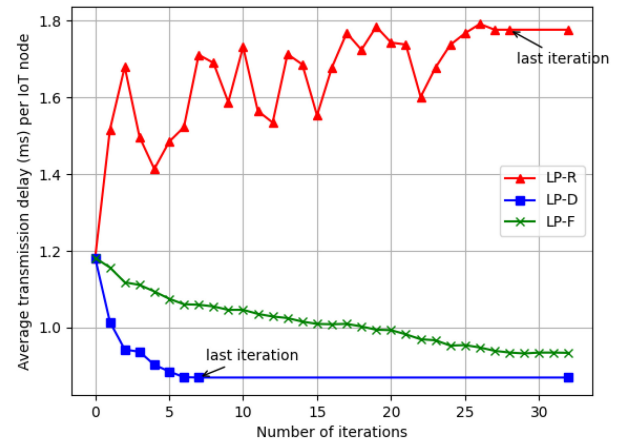
Fig. 7. Evaluation of the transmission delay considering LP-R, LP-D, and LP-F.

affects the waiting the retransmission delay. The averagetransmission delay does also increase with the maximum number of retransmission, which is a direct consequence of considering ARQ scheme. In the next evaluations, we consider a transmission rates of 0.5 bit/s/Hz for uRLLC service and 2.0 bit/s/Hz for eMBB service. We also set the maximum number of retransmission, E^c, to 4 for the service eMBB, and to 2 for the service uRLLC.

We evaluate these three optimization solutions, namely, LP-R, LP-D, and LP-F, in terms of the achieved effective rate at the serving UAVs and the transmission delay. The results are, respectively, depicted in Figs. 6 and 7. These evaluations consider 30 IoT nodes and 12 UAVs. In terms of optimizing the effective rate (Fig. 6), the two optimization solutions LP-R and LP-F achieve better results compared to LP-D. Indeed, starting from the initial allocation of resources and deployment of UAVs, the optimization solutions LP-R and LP-F increase the effective rate for the IoT nodes at each iteration whereas the optimization solution LP-D does not take it into account. The LP-R and LP-F algorithms achieve 47.82 and 27.29 times larger average effective rate than the LP-D algorithm, respectively. As for optimizing the transmission delay (Fig. 7), the optimization solutions LP-D and LP-F reached better results
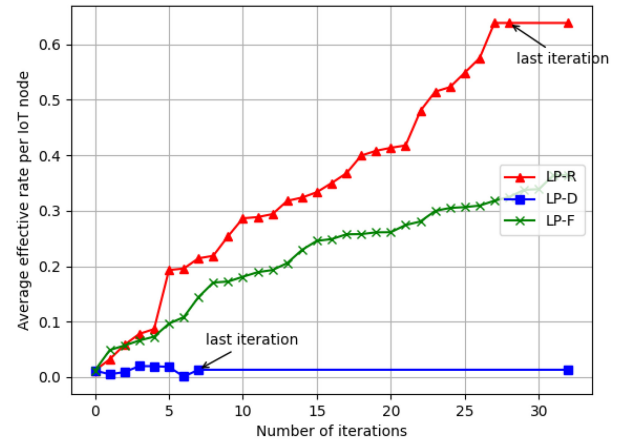
compared to LP-R. While the two solutions LP-D and LP-F managed to decrease the average transmission delay in each iteration of the algorithm, the optimization LP-R do not take it into account. The LP-D solution has 2.04 times shorter delay compared to the LP-R optimization while LP-F has reached 1.90 times shorter delay compared to the same optimization.

We evaluate the effective rate and the transmission delay of the proposed tradeoff approach against the initial allocation/deployment, while varying the number of IoT nodes with a fixed setting of 12 UAVs. The simulation results are summarized in Figs. 8 and 9, respectively. We can see that the average effective rate decreases with the number of considered IoT nodes, while the average transmission delay increases with the same number. This is because the more IoT nodes are considered in the network, the fewer subcarriers can be assigned to each node, which directly affects the average effective rate and transmission delay for those nodes. However, the proposed tradeoff approach achieves better results compared to the initial allocation. Indeed, it starts from the initial allocation and performs a number of iterations to successively reach optimal resource allocation and UAV deployment, leading to enhanced effective rate and transmission delay.

We now evaluate the proposed approach against the baseline solutions considering different numbers of IoT nodes and
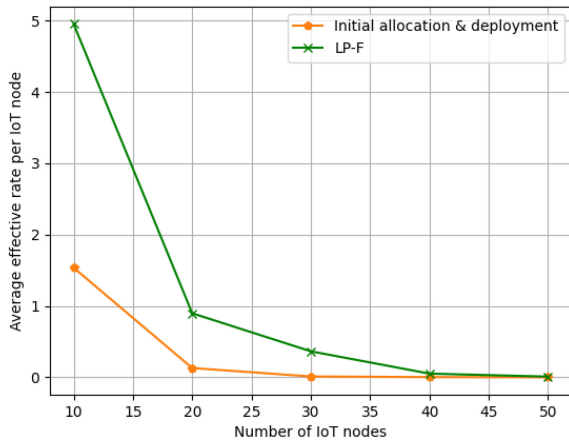
Fig. 8. Evaluation of the effective rate considering initial allocation/deployment and LP-F.
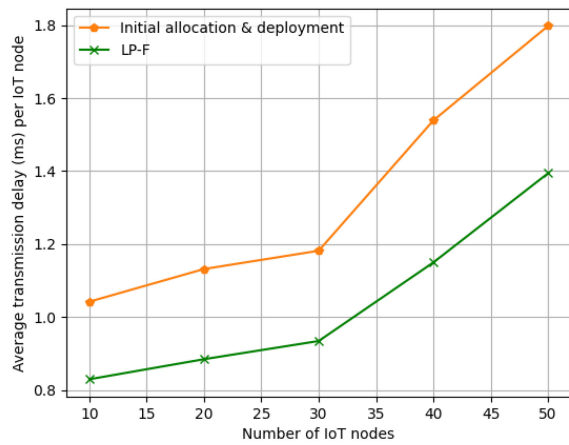

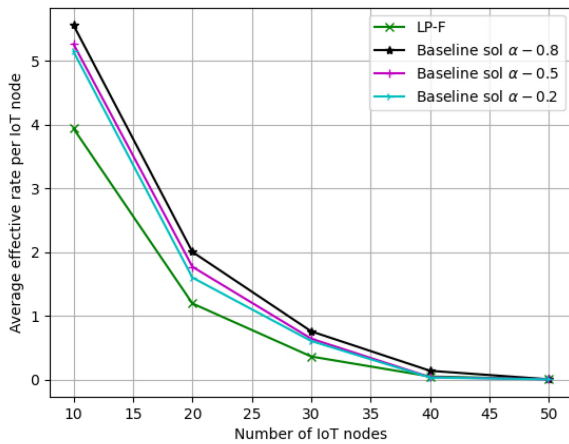Fig. 9. Evaluation of the transmission delay considering initial allocation/deployment and LP-F.

Fig. 10. Evaluation of the effective rate for the proposed approach LP-F and the baseline solutions LP-α ($\alpha \in \{0.2, 0.5, 0.8\}$) as a function of the number of IoT nodes.
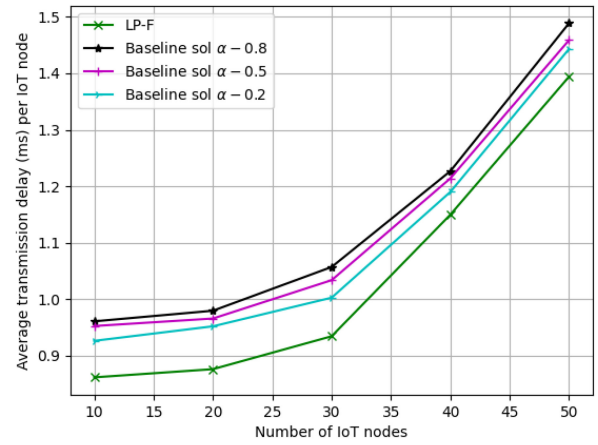

Fig. 11. Evaluation of the transmission delay for the proposed approach LP-F and the baseline solutions LP-α ($\alpha \in \{0.2, 0.5, 0.8\}$) as a function of the number of IoT nodes.

a fixed setting of 12 serving UAVs. The simulation results are summrized in Figs. 10 and 11, respectively. As stated earlier, the effective rate decreases with the number of considered IoT nodes while the transmission delay increase with this number. As we also can see, the three solutions LP-α ($\alpha \in \{0.2, 0.5, 0.8\}$) have comparable performance in terms of

both the effective rate and transmission delay. Furthermore, the three solutions have better optimization results for the effective rate (Fig. 10) than for the transmission delay (Fig. 11), which is similar to the behavior of the LP-R optimization solution. This is due to the fact that the above optimization problem for jointly optimizing the two objective functions does not take into consideration the scales of the values of the two functions, meaning that the value 0.5 for α can not reach the tradeoff between the two objective functions. Indeed, considering the condition (73) that defines the variable F to be maximized by the objective function, the values of the variable ϑ^r are in the scale of 10^0 (bit/s/Hz) while the values of the part ($D_{\max} - \vartheta^d$) are in the scale of 10^{-3} (ms). This favors the optimization of the effective rate when averaging between the two objective functions using the values 0.2, 0.5, and 0.8 for the parameter α. Note that the variables ϑ^r and ϑ^d are, respectively, defined in the two conditions (68) and (69). On the other hand, the proposed approach LP-F can achieve a tradeoff between the two objective functions while considering the scale of their values. This is thanks to the objective function between the two optimizations [as defined in (65), (66), and (67)].

We also evaluate the proposed approach against the baseline solutions with different numbers of UAVs serving a 30 IoT nodes. The simulation results are given in Figs. 12 and 13. We can see from the two figures that the average transmission rate increases with the number of serving UAVs, while the average transmission delay decreases with this number. As mentioned earlier, this is because that increasing the number of serving UAVs means that more subcarriers can be allocated to the IoT nodes. Furthermore, we can also see that the three baseline solutions provide similar results, which tend to favor the optimization of the effective rate at the expense of the transmission delay. As mentioned earlier, the baseline solutions do not consider the scale of the values related to the two objective functions, which is translated in favor of optimizing the effective rate for the values 0.2, 0.5, and 0.8 of α. On the other hand, the proposed LP-F approach achieves a better tradeoff by sharing the same objective function and considering the scale of the related values.
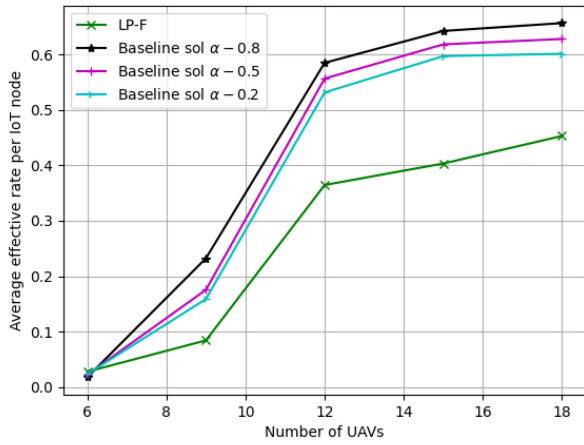
Fig. 12. Evaluation of the effective rate for the proposed approach LP-F and the baseline solutions LP-α ($\alpha \in \{0.2, 0.5, 0.8\}$) as a function of the number of UAVs.
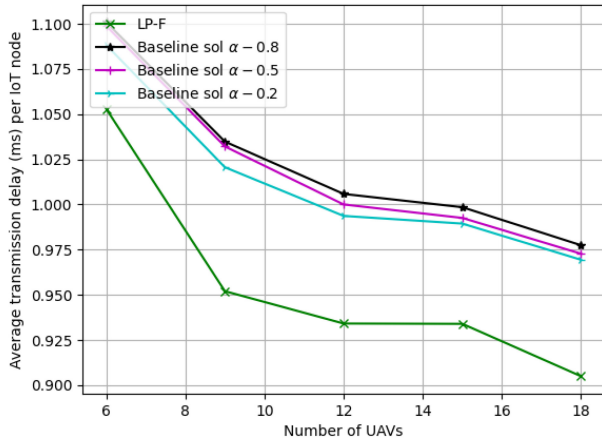

Fig. 13. Evaluation of the transmission delay for the proposed approach LP-F and the baseline solutions LP-α ($\alpha \in \{0.2, 0.5, 0.8\}$) as a function of the number of UAVs.

Furthermore, we evaluate the proposed iterative approach with the brute-force search solution. Indeed, the iterative approach, used in LP-R, LP-D, and LP-F, is based on considering linear optimizations at each iteration until reaching a stability. The latter is characterized by the situation that the objective function can not be enhanced further. Note that the value of the objective function at each iteration can only be enhanced or equal to the previous one. The iterative approach will therefore reach an optimal solution. We therefore implement the brute-fore search for N-LP-R and compare it with the iterative approach LP-R. We consider the case that each UAV is serving four IoT devices with an equitable number of subcarriers. In this case, the complexity of the brute-force search is $\mathcal{O}(\prod_{v \in \mathcal{V}} |\mathcal{L}_v| \times \prod_{v \in \mathcal{V}} |\mathcal{B}|!)$. As shown in Table III, the iterative approach achieves inferior results compared to the brute-force search. However, this difference is less than 25% in the considered scenarios. On the other hand, the brute-force search is associated with a huge complexity in comparison with the iterative approach based on the linear optimization.

TABLE III
COMPARISON OF THE ITERATIVE APPROACH WITH BRUTE-FORCE
SEARCH (SUM EFFECTIVE RATE—BIT/S/HZ)

Algorithm	3 UAVs	4 UAVs	5 UAVs
Brute force (N-LP-R)	3.01	4.17	5.37
iterative approach (LP-R)	2.67	3.41	4.03

VII. CONCLUSION

In this article, we investigated the multiservice performances in terms of the effective rate and transmission delay via subcarrier allocation and UAV deployment in UAV-enabled wireless communications for the IoT. We derived the expressions of these two performance metrics, and formulated each service type as an optimization problem with the constraints of resource allocation and UAV deployment. Then, we transformed them into linear optimization problems, which were solved by two iterative algorithms. We further proposed a linear program algorithm to jointly optimize the two service types. The results of performance evaluations validated the effectiveness of the proposed approach. We evaluated these three optimization solutions, namely, LP-R, LP-D, and LP-F, in terms of the achieved effective rate and the transmission delay. The LP-F algorithm achieves a larger average effective rate than the LP-D algorithm. Moreover, the LP-F algorithm reduces the transmission delay compared to the LP-R algorithm.

APPENDIX

PROOF OF THEOREMS 1 AND 2

This section provides the proof of Theorems 1 and 2, where we derive the expressions of the effective rate and the delay on the uplink communication. We start by providing the general expression of the effective rate which is given as

$$R_{u,l,b}^{\text{eff}} = \frac{R_u^r \times \left(1 - P_{\text{out}}\left(R_u^r\right)\right)}{\mathbb{E}(T_u)} \tag{75}$$

where R_u^r is the transmission rate of the source IoT node u. $P_{\text{out}}(R_u^r)$ is the probability of a packet transmission failure if the source IoT node uses a transmission rate R_u^r and $\mathbb{E}(T_u)$ denotes the average number of retramsission from the node u. In the ARQ mode, packers are retransmitted until a successful reception or when reaching a maximum number E^c of retransmission. A similar expression of the average effective rate has been provided in [27]. The expression of $P_{\text{out}}(R_u^r)$ can be provided as

$$P_{\text{out}}\left(R_u^r\right) = P\left[\log_2(1 + \text{SINR}) < R_u^r\right] \tag{76}$$

$$= P\left[\text{SINR} < 2^{R_u^r} - 1\right] \tag{77}$$

where the SINR stands for the SINR which is computed as

$$\text{SINR}_{uv} = \frac{\gamma_{uv}}{1 + \sum \gamma_{tv}} \approx \frac{\gamma_{uv}}{\sum \gamma_{tv}}. \tag{78}$$

The approximation in (78) is valid if the noise power can be neglected compared to the interference power. This is generally a well-accepted assumption in the literature and is known as an interference-limited regime.

In the uplink scenario, the source IoT node u transmits its packets to its serving UAV v. We can therefore define the outage probability for the link uv as

$$P_{\text{out}}(x) = P(\text{SINR}_{uv} < x) \tag{79}$$

$$= P\left(\frac{\gamma_{uv}}{\gamma_{I_v}} < x\right) = E_{\gamma_{I_v}}\left(P[\gamma_{uv} < x\gamma_{I_v}]\right) \tag{80}$$

$$= \int_0^\infty F_{\gamma_{uv}}(xy) f_{\gamma_{I_v}}(y) \, dy \tag{81}$$

where $F_{\gamma_{uv}}(\cdot)$ is the cumulative distribution function (CDF) of γ_{uv} and $f_{\gamma_{I_v}}(\cdot)$ is the PDF of γ_{I_v}. The channel coefficient h_{uv} for link uv is assumed to be Nakagami distribution, and thus γ_{uv} is Gamma distributed, i.e., $\gamma_{uv} \sim \mathcal{G}(\alpha_{uv}, \beta_{uv})$, with the corresponding PDF given as

$$f_{\gamma_{uv}}(x) = \frac{x^{\alpha_{uv}-1}}{\beta_{uv}^{\alpha_{uv}} \Gamma(\alpha_{uv})} \exp\left(-\frac{x}{\beta_{uv}}\right) \tag{82}$$

where α_{uv} is the Nakgami fading parameter for the link uv, and $\beta_{uv} = (\bar{\gamma}_{uv}/\alpha_{uv})$. The CDF of γ_{uv} can be computed as

$$F_{\gamma_{uv}}(x) = 1 - \frac{\Gamma(\alpha_{uv}, x/\beta_{uv})}{\Gamma(\alpha_{uv})} \tag{83}$$

where $\Gamma(\alpha_{uv}, x/\beta_{uv})$ is the upper incomplete gamma function defined as $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$. As for the PDF of $f_{\gamma_{I_v}}(y)$, it represents the PDF of the interference which is the sum of independent and nonidentical Gamma distributions, where $\gamma_{tv} \sim \mathcal{G}(\alpha_{tv}, \beta_{tv})$ and $\beta_{tv} = (\bar{\gamma}_{tv}/\alpha_{tv})$. The PDF of the total interference γ_{I_v} can be approximated by a Gamma distribution with parameters α_v and β_v which are given as

$$\alpha_v = \frac{(\mathbb{E}[\gamma_{I_v}])^2}{\text{Var}(\gamma_{I_v})} \tag{84}$$

$$\beta_v = \frac{\text{Var}(\gamma_{I_v})}{\mathbb{E}[\gamma_{I_v}]}. \tag{85}$$

As for $\mathbb{E}[\gamma_{I_v}]$ and $\text{Var}(\gamma_{I_v})$, they are computed as

$$\mathbb{E}[\gamma_{I_v}] = \sum_{t=1}^{N_v} \mathbb{E}[\gamma_{tv}] = \sum_{t=1}^{N_v} \alpha_t \beta_t \tag{86}$$

$$\text{Var}(\gamma_{I_v}) = \sum_{t=1}^{N_v} \text{Var}(\gamma_{tv}) = \sum_{t=1}^{N_v}\left(\mathbb{E}\left[(\gamma_{tv})^2\right] - (\mathbb{E}[\gamma_{tv}])^2\right). \tag{87}$$

Consequently, α_v and β_v will be computed as

$$\alpha_v = \frac{\left(\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}\right)^2}{\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}^2} \tag{88}$$

$$\beta_v = \frac{\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}^2}{\sum_{t=1}^{N_v} \alpha_{tv}\beta_{tv}}. \tag{89}$$

Thereafter, the outage probability can be expressed as

$$P_{\text{out}}(x) = 1 - \int_0^\infty \frac{\Gamma(\alpha_{uv}, xy/\beta_{uv})}{\Gamma(\alpha_{uv})} f_{\gamma_{I_v}}(y) \, dy \tag{90}$$

$$= 1 - \underbrace{\int_0^\infty \frac{\Gamma(\alpha_{uv}, xy/\beta_{uv})}{\Gamma(\alpha_{uv})} \frac{y^{\alpha_v-1}}{\beta_v^{\alpha_v}\Gamma(\alpha_v)} \exp\left(-\frac{y}{\beta_v}\right) dy}_{\mathcal{I}(x,\alpha_{uv},\beta_{uv},\alpha_v,\beta_v)}$$

$$= 1 - \mathcal{I}(x, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v). \tag{91}$$

The integral $\mathcal{I}(x, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v)$ can be computed as

$$\mathcal{I}(x, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v) = \left(\frac{x\beta_v}{\beta_{uv}}\right)^{-\alpha_v} \frac{\Gamma(\alpha_{uv}+\alpha_v)}{\Gamma(\alpha_{uv})\Gamma(1+\alpha_v)}$$
$$\times {}_2F_1\left(\alpha_v, \alpha_{uv}+\alpha_v, 1+\alpha_v, \frac{-\beta_{uv}}{x\beta_v}\right) \tag{92}$$

where ${}_2F_1(.)$ is the Gauss hypergeometric. Thus, expression of $P_{\text{out}}(x)$ can be written as

$$P_{\text{out}}(x) = 1 - \mathcal{I}(x, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v) \tag{93}$$

and the expression of the outage probability if the source IoT node uses a transmission rate R_u^r can thus be expressed as

$$P_{\text{out}}(R_u^r) = 1 - \mathcal{I}(R_u^r, \alpha_{uv}, \beta_{uv}, \alpha_v, \beta_v) \tag{94}$$

$$= 1 - \mathcal{I}_u. \tag{95}$$

As for the average number of retransmissions $\mathbb{E}(T_u)$, it can be computed as [28]

$$\mathbb{E}(T_u) = 1 + \sum_{e=1}^{E^c-1} P\left(F_u^1, \ldots, F_u^e\right) \tag{96}$$

$$= 1 + \sum_{e=1}^{E^c-1} (1 - \mathcal{I}_u) \tag{97}$$

$$= \sum_{e=0}^{E^c-1} (1 - \mathcal{I}_u) \tag{98}$$

$$= \frac{1 - (1 - \mathcal{I}_u)^{E^c}}{\mathcal{I}_u} \tag{99}$$

where $P(F_u^1, \ldots, F_u^e)$ refers to the probability of the reception failure at the $1^{st}, \ldots, e^{th}$ retransmissions for the IoT node u.

With the help of (75), (95), and (99), the effective rate expression for the link uv can be expressed as

$$R_{u,l,b}^{\text{eff}} = \frac{R_u^r \times (\mathcal{I}_u)^2}{1 - (1 - \mathcal{I}_u)^{E^r}}. \tag{100}$$

The result in (100) is the same as the effective rate provided in Theorem 1.

As for the expression of the delay, we consider a parallel $M/M/1$ queuing model where the traffic is equitably shared among the different queues. The arrival rate λ_u of the node u is therefore divided on the number of parallel queues Q_u. In this case, the delay can be evaluated using the Pollaczek–Khinchin equation as [29]

$$D_{u,l,b}^{[\lambda_u/Q_u]} = \mathcal{W}_{u,b}^{[\lambda_u/Q_u]} + \mathbb{E}(T_u)T_F \tag{101}$$

where $\mathcal{W}_{u,b}^{[\lambda_u/Q_u]}$ is the average waiting time for a data packet in the buffer of the IoT node u over the subcarrier b. $\mathcal{W}_{u,b}^{[\lambda_u/Q_u]}$ can be obtained as [29]

$$\mathcal{W}_{u,b}^{[\lambda_u/Q_u]} = \frac{\lambda_u \mathbb{E}(T_u^2)T_F^2}{Q_u 2(1-\rho_u)} + \frac{T_F}{2} \tag{102}$$

where ρ_u is represents a parameter which satisfies the stability condition

$$\rho_u = \frac{\lambda_u \mathbb{E}(T_u) T_F}{Q_u} < 1. \tag{103}$$

As for the term $\mathbb{E}(T_u^2)$, it represents the second-order moment of the number of retransmission T_u. This term can be derived as [28]

$$\mathbb{E}\left(T_u^2\right) = 1 + \sum_{e=1}^{E^c-1} (2e+1) P\left(F_u^1, \ldots, F_u^e\right) \tag{104}$$

$$= 1 + \sum_{e=1}^{E^c-1} (2e+1)(1 - \mathcal{I}_u) \tag{105}$$

$$= \sum_{e=0}^{E^c-1} (2e+1)(1 - \mathcal{I}_u) \tag{106}$$

$$= \frac{1 - (2E^c - 1)(1 - \mathcal{I}_u)^{Ec}}{\mathcal{I}_u}$$
$$+ \frac{2(1 - \mathcal{I}_u)\left(1 - (1 - \mathcal{I}_u)^{E^c-1}\right)}{\mathcal{I}_u^2}. \tag{107}$$

With the help of (101), (102), and (107), the delay can be expressed as

$$D_{u,l,b}^{[\lambda_u/Q_u]} = \frac{\lambda_u T_F^2}{Q_u 2(1 - \rho_u)}\left(\frac{1 - \left(2E^d - 1\right)(1 - \mathcal{I}_u)^{E^d}}{\mathcal{I}_u}\right.$$
$$+ \left.\frac{2(1 - \mathcal{I}_u)\left(1 - (1 - \mathcal{I}_u)^{E^d-1}\right)}{\mathcal{I}_u^2}\right)$$
$$+ \frac{T_F}{2} + \frac{1 - (1 - \mathcal{I}_u)^{E^d}}{\mathcal{I}_u} T_F. \tag{108}$$

The result in (108) is the same as the average delay provided in Theorem 2.

REFERENCES

[1] H. Hellaoui, A. Chelli, M. Bagaa, and T. Taleb, "UAV communication strategies in the next generation of mobile networks," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, 2020, pp. 1642–1647.

[2] "Global unmanned aerial vehicle market research report 2020." Valuates Reports. 2020. [Online]. Available: https://reports.valuates.com/

[3] "Use case specifications and requirements, H2020 5G!Drones D1.2." 2009. Accessed: Jan. 23, 2022. [Online]. Available: https://5gdrones.eu/wp-content/uploads/2020/05/D1.1-Use-case-specifications-and-requirements-v1.0.pdf

[4] "Study on enhanced LTE support for aerial vehicles," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 36.777, 2017.

[5] "Study on remote identification of unmanned aerial systems (UAS)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 22.825, 2018.

[6] *Service Requirements for the 5G System*, 3GPP Standard TS 22.261, 2017.

[7] B. V. Der Bergh, A. Chiumento, and S. Pollin, "LTE in the sky: Trading off propagation benefits with interference costs for aerial nodes," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 44–50, May 2016.

[8] X. Lin et al., "The sky is not the limit: LTE for unmanned aerial vehicles," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 204–210, Apr. 2018.

[9] H. Hellaoui, A. Chelli, M. Bagaa, and T. Taleb, "Towards mitigating the impact of UAVs on cellular communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–7.

[10] U. Challita, W. Saad, and C. Bettstetter, "Cellular-connected UAVs over 5G: Deep reinforcement learning for interference management," 2018, arXiv:1801.05500.

[11] H. Hellaoui, M. Bagaa, A. Chelli, and T. Taleb, "Joint sub-carrier and power allocation for efficient communication of cellular UAVs," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8287–8302, Dec. 2020.

[12] X. Guan, Y. Huang, and Q. Shi, "Joint subcarrier and power allocation for multi-UAV systems," *China Commun.*, vol. 16, no. 1, pp. 47–56, Jan. 2019.

[13] H. Hellaoui, A. Chelli, M. Bagaa, and T. Taleb, "Efficient steering mechanism for mobile network-enabled UAVs," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2019, pp. 1–6.

[14] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for power-efficient deployment of unmanned aerial vehicles," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[15] X. Liu, L. Li, F. Yang, X. Li, W. Chen, and W. Xu, "Price-based power allocation for multi-UAV enabled wireless networks," in *Proc. 28th Wireless Opt. Commun. Conf. (WOCC)*, 2019, pp. 1–5.

[16] 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.

[17] Z. Li et al., "Energy efficient resource allocation for UAV-assisted space-air-ground Internet of remote things networks," *IEEE Access*, vol. 7, pp. 145348–145362, 2019.

[18] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[19] M. Pätzold, *Mobile Radio Channels*. Chichester, U.K.: Wiley, 2011.

[20] A. F. Molisch, *Wireless Communications*, vol. 34. Chichester, U.K.: Wiley, 2012.

[21] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. Hoboken, NJ, USA: Pearson Educ., 2007.

[22] "Study on new radio access technology physical layer aspects," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 38.802, 2017.

[23] M. Bagaa, T. Taleb, A. Chelli, and H. Hellaoui, "Constraint hubs deployment for efficient machine-type communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 7936–7951, Dec. 2018.

[24] A. F. Molisch, *Wireless Communications*. Chichester, U.K.: Wiley, 2005.

[25] A. Chagdali, S. E. Elayoubi, and A. M. Masucci, "Impact of slice function placement on the performance of URLLC with redundant coverage," in *Proc. 16th Int. Conf. Wireless Mobile Comput. Netw. Commun. (WiMob)*, 2020, pp. 1–6.

[26] "Gurobi optimization." Gurobi. 2023. [Online]. Available: http://www.gurobi.com/

[27] W. Rui and V. K. N. Lau, "Combined cross-layer design and HARQ for multiuser systems with outdated channel state information at transmitter (CSIT) in slow fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2771–2777, Jul. 2008.

[28] A. Chelli, E. Zedini, M.-S. Alouini, J. R. Barry, and M. Pätzold, "Performance and delay analysis of hybrid ARQ with incremental redundancy over double Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6245–6258, Nov. 2014.

[29] W. C. Chan, T.-C. Lu, and R.-J. Chen, "Pollaczek–Khinchin formula for the M/G/1 queue in discrete time with vacations," *IEE Proc. Comput. Digit. Techn.*, vol. 144, no. 4, pp. 222–226, 1997.