Radio Map-Assisted Approach for Interference-Aware Predictive UAV Communications

Bowen Li, *Graduate Student Member, IEEE*, and Junting Chen, *Member, IEEE*

***Abstract*— Herein, an interference-aware predictive aerial-and-terrestrial communication problem is studied, where an unmanned aerial vehicle (UAV) delivers some data payload to a few nodes within a communication deadline. The first challenge is the possible interference to the ground base stations (BSs) and users possibly at unknown locations. This paper develops a radio-map-based approach to predict the channel to the receivers and the unintended nodes. Therefore, a predictive communication strategy can be optimized ahead of time to reduce the interference power and duration for the ground nodes. Such predictive optimization raises the second challenge of developing a low-complexity solution for a batch of transmission strategies over T time slots for N receivers before the flight. Mathematically, while the proposed interference-aware predictive communication problem is non-convex, it is converted into a relaxed convex problem, and solved by a novel dual-based algorithm, which is shown to achieve global optimality at asymptotically small slot duration. The proposed algorithm demonstrates orders of magnitude of the computational time saving compared to several existing solvers. Simulations show that the radio-map-assisted scheme can reduce the interference to the unintended receivers at known locations below a prescribed threshold and significantly reduce the interference to the users at unknown locations.**

***Index Terms*— UAV communication, air-to-ground interference, radio map, predictive communication.**

Manuscript received 2 February 2024; revised 1 July 2024; accepted 9 August 2024. Date of publication 27 August 2024; date of current version 13 November 2024. This work was supported in part by the National Science Foundation of China (NSFC) under Grant 62171398; in part by the Basic Research Project of Hetao Shenzhen-HK S&T Cooperation Zone under Grant HZQB-KCZYZ-2021067; in part by NSFC under Grant 62293482; in part by Shenzhen Science and Technology Program under Grant JCYJ20210324134612033, Grant KJZD20230923115104009, and Grant KQTD20200909114730003; in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515011206; in part by Guangdong Research under Grant 2019QN01X895, Grant 2017ZT07X152, and Grant 2019CX01X104; in part by Shenzhen Outstanding Talents Training Fund under Grant 202002; in part by Guangdong Provincial Key Laboratory of Future Networks of Intelligence under Grant 2022B1212010001; in part by the National Key Research and Development Program of China under Grant 2018YFB1800800; and in part by the Key Area Research and Development Program of Guangdong Province under Grant 2018B030338001. The associate editor coordinating the review of this article and approving it for publication was M. Kaneko. *(Corresponding author: Junting Chen.)*

The authors are with the School of Science and Engineering and Shenzhen Future Network of Intelligence Institute (FNii–Shenzhen), The Chinese University of Hong Kong, Shenzhen (CUHK–Shenzhen), Shenzhen, Guangdong 518172, China (e-mail: bowenli6@link.cuhk.edu.cn; juntingc@cuhk.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2024.3446240.

Digital Object Identifier 10.1109/TWC.2024.3446240

I. Introduction

LOW-ALTITUDE UAV activities have grown significantly over the last decade [1]. It is important to establish a reliable communication network for real-time navigation, control, and surveillance of the UAV network. For example, it may be crucial to acquire the field of view of the UAV to monitor the operation status during the mission of the UAV, and this requires a reliable communication network.

One potential solution for UAV network communication is to construct self-organized networks (SONs) that utilize the orthogonal spectrum of those used in terrestrial cellular networks for interference avoidance [2], [3], [4]. However, SONs may not be reliable as the topology of the UAV network can be significantly time-varying. Another viable solution is to extend the coverage of terrestrial cellular networks to assist with UAV communication. However, the transmission from the UAV to ground BSs may generate strong interference with other ground terminals due to the high probability of line-of-sight (LOS) conditions from the sky [5], [6], [7], [8], [9], [10], [11], [12], [13].

A majority of existing research that studied integrated UAV-and-terrestrial communications ignored the interference [14], [15], [16], [17], [18] or assumed orthogonal transmissions between aerial nodes and ground nodes [19], [20], [21]. Some recent works [5], [6], [7], [8], [9], [10], [11], [12] attempted to mitigate the interference from the UAVs to the ground nodes by optimizing the UAV trajectory, power control, sub-channel allocation, and MIMO beamforming. However, these approaches may not apply to some UAV networks where communication is not the primary mission of the UAVs. First, the UAV trajectory may not be altered for communication purposes during the mission, and therefore, the approaches based on trajectory optimization [5], [6], [7] are not suitable here. Secondly, as the topology of the UAV network is time-varying, it becomes challenging to meet transmission deadlines or age-of-information requirements for delivering large and time-sensitive content, and these factors are not considered in [8], [9], [10], and [11].

This paper studies an integrated aerial-and-terrestrial communication scenario, where a UAV node uses the cellular spectrum to transmit to N nodes, which can be other UAV nodes or ground BSs with known locations. The objective is to deliver a given amount of data within a communication deadline, resulting in a planning problem for the communication

timing and resource allocation. The main challenge is the air-to-ground interference to the unintended BSs during the uplink (UL) and users on the ground during the downlink (DL) due to the transmission of the UAV, since it is possible that some BSs are at UL while some others are at DL. A general interference mitigation may require the channel state information (CSI) of the unintended nodes for the entire transmission, which can be challenging for a resource planning problem where the future CSI cannot be obtained via online measurements. While the locations of the BSs can be known by the network for a rough CSI prediction, the locations of the ground users are usually unknown due to their mobility. In addition, joint UAV and BS interference-aware resource allocation is challenging because a low-altitude UAV may interfere with many BSs and ground users due to possible LOS conditions; as a consequence, conventional interference-aware resource allocation will require centralized processing with global CSI across a large number of cells, which is almost prohibitive in implementation.

To tackle the interference to the nodes with known locations, we employ a *radio map* approach, where the radio map captures the large-scale channel information between any two locations, including path loss, shadowing, and statistics of the small-scale fading; but the actual channel is not available on the radio map. The opportunity that drives the radio map approach is the fact that aerial nodes mostly have predetermined trajectories. For example, the trajectories of cargo delivery UAVs and many patrol surveillance UAVs are determined by the operators prior to the flight. As a consequence, at least the large-scale statistics of the future CSI of the nodes with known locations can be predicted by the UAV trajectory using the radio map, and therefore, one can optimize for the *predictive UAV communications* based on the predictive large-scale CSI statistics. To tackle the interference to the unknown ground nodes, we introduce a sleep strategy aiming at reducing the transmission duration and lowering the chance of interfering with the ground nodes with unknown channel status. To summarize, this paper attempts to address the following challenges:

- *How to exploit radio maps to optimize for interference-aware communication.* First, as radio maps may only capture large-scale CSI, the uncertainty of the future channel quality in predictive communication needs to be addressed. In addition, the impact of the air-to-ground interference for ground nodes at unknown locations needs to be considered.
- *How to develop low-complexity solutions over a large horizon for predictive communication.* The predictive UAV communication problem involves determining a sequence of transmission strategies over T time slots *before* the flight. Therefore, optimizing the solution based on standard off-the-shelf solvers can be computationally prohibitive for a large T, even for convex problems.

Mathematically, we formulate a radio-map-assisted predictive communication problem for N-receiver interference-aware UAV communications with a delay constraint and interference constraints. While previous research has considered the impact of small-scale fading on transmission strategy design, for example, prediction of the future small-scale fading [22], online adjustment based on real-time small-scale fading [23], and constant attenuation on the channel [24], these approaches cannot be applied to the transmission planning in predictive communications. To address the challenge of discontinuity due to the on-off control for the interference-aware sleep mode optimization, we propose a relax-then-round algorithm for an efficient solution with the asymptotic optimality guarantee. Moreover, we develop a dual-based algorithm, substantially reducing the computational complexity without sacrificing the optimality. In summary, we make the following contributions:

- We formulate an interference-constrained predictive communication problem exploiting radio maps. In this formulation, we address the uncertainty of the future CSI by developing a deterministic expected capacity lower bound, and mitigate the air-to-ground interference for ground nodes at unknown locations by penalizing the transmission duration.
- While the problem is non-convex, we develop a relax-then-round optimization strategy, which is proven to have the asymptotic optimality guarantee. Based on the Lagrangian dual technique, we develop a low-complexity algorithm, which achieves orders of magnitude of complexity savings, compared to state-of-the-art Lagrangian relaxation (LR)-based and successive convex approximation (SCA)-based algorithms for a large horizon T.
- Our simulations show that the proposed radio-map-assisted scheme can reduce the interference at the unintended receivers at known locations below a prescribed threshold and also significantly reduce the interference to the users at unknown locations. Moreover, the proposed relax-then-round can achieve global optimality at the asymptotically small slot duration, and the running time is 1000 times less than the LR & SCA scheme.

The rest of the paper is organized as follows. Section II presents the communication system model and the problem formulation. Section III reformulates the problem to handle the uncertainty from the small-scale fading, then relaxes the non-convex and non-smooth problem to a convex problem and proposes a dual-based algorithm. We develop a rounding strategy to modify the relaxed results to suit the original problem in Section IV. Numerical results are demonstrated in Section V, and conclusions are given in Section VI.

II. System Model

Consider a multiuser transmission system with a single transmitter, N receiver nodes, and M neighboring nodes that may be interfered with by the transmitter, as shown in Fig. 1. The positions of the transmitter, receivers, and neighboring nodes are known and denoted as $\boldsymbol{q}_j \in \mathbb{R}^3$, where $j = 0, 1, \ldots, N+M$. Here, $j = 0$ refers to the transmitter, while $1 \le j \le N$ refers to the receiver nodes, and $N+1 \le j \le N+M$ refers to the neighboring nodes.

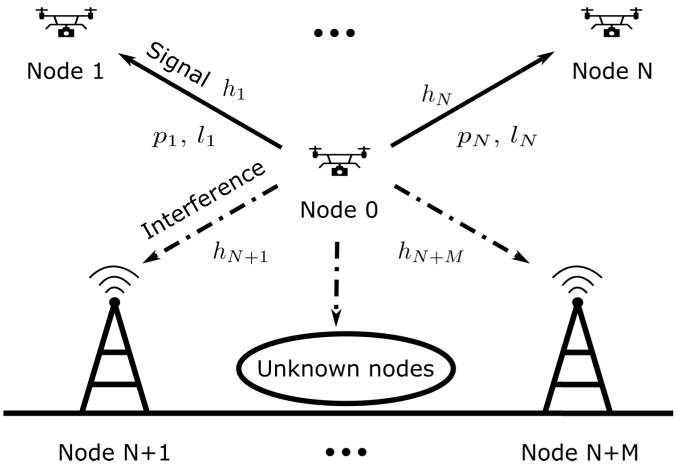

Fig. 1. Slotted communication system illustration.

A. Channel and Radio Map Models

We consider a flat fading channel model, in which the channel power gain from the transmitter to node $1 \leq j \leq N+M$ is given by

$$h_j = g_j \xi_j \tag{1}$$

where g_j and ξ_j represent the channel gain due to the large- and small-scale fading, respectively. The small-scale fading ξ_j is random and follows a Gamma distribution $\mathrm{G}(\kappa_j, 1/\kappa_j)$ with shape parameter κ_j, and the scale parameter is chosen as $1/\kappa_j$ to normalize the mean of ξ_j to 1. Note that Gamma distribution can be used to model various fading models, including Rayleigh and Nakagami. For example, the Gamma distribution with $\kappa_j = 1$ degenerates to an exponential distribution, which corresponds to the power gain of the Rayleigh fading channel.

Assume that the system has access to the large-scale channel parameters, including the large-scale channel gain g_j and the Gamma distribution parameter κ_j, via a radio map $\boldsymbol{\Theta}(\boldsymbol{q}_0, \boldsymbol{q}_j) = (g_j, \kappa_j)$ that records the large-scale channel parameters g_j and κ_j as a function of the transmitter and receiver locations $\boldsymbol{q}_0$ and $\boldsymbol{q}_j$ [25], [26], [27], [28]. These characteristics change only in a long timescale and can be tracked using off-line approaches, and thus, they are considered static in our problem formulation. As a consequence, the statistics of the channel h_j are available ahead of time when the trajectories of the nodes $\boldsymbol{q}_j$ are available.

B. Communication Model

Specifically, at each time slot $t \in \mathcal{T} \triangleq \{1, 2, \cdots, T\}$, the channel power gain from the transmitter to node $1 \leq j \leq N+M$ is represented as $h_j(t)$, the large-scale fading as $g_j(t)$, and the small-scale fading as $\xi_j(t)$. Consider slotted transmission, where in each time slot $t \in \mathcal{T}$, a fraction of non-overlapping frequency resource $l_n(t)$ is allocated to the nth receiver, $n \in \mathcal{N} \triangleq \{1, 2, \ldots, N\}$. Denote the frequency allocation on time t as $\boldsymbol{l}(t) \in \mathcal{L}$, where

$$\mathcal{L} \triangleq \left\{ [l_n(t)]_{n\in\mathcal{N}} : l_n(t) \in [0,1], \sum_{n\in\mathcal{N}} l_n(t) \in [0,1] \right\}$$

denotes the set of feasible frequency resources, in which the total resources are normalized to 1.[1] As a result, the total transmission duration among T time slots can be computed as $\sum_{t\in\mathcal{T}} \mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$.

Let $p_n(t)$ be the power allocated to receiver n at time slot t, where $p_n(t) \in \mathcal{P} \triangleq \{p : p \geq 0\}$. As orthogonal frequency resources are used for the N receivers, the total normalized throughput for the nth receiver over T time slots can be computed as

$$\Upsilon_n(\boldsymbol{p}_n, \boldsymbol{l}_n) = \sum_{t\in\mathcal{T}} \log\left(1 + p_n(t)\, h_n(t)\right) l_n(t) \tag{2}$$

where $\boldsymbol{p}_n = [p_n(t)]_{t\in\mathcal{T}} \in \mathbb{R}^{T\times 1}$ and $\boldsymbol{l}_n = [l_n(t)]_{t\in\mathcal{T}} \in \mathbb{R}^{T\times 1}$ and log represents logarithm with base 2 throughout the paper.

Assume that both aerial and terrestrial communications use the same frequency band. Thus, the aerial transmission creates interference to neighboring node m at time slot t as

$$I_{m,n}(t) = h_m(t)\, p_n(t).$$

Note that there could be other ground nodes at unknown locations, and their channels are not available at the transmitter node. To reduce the impact on those unknown ground nodes, we propose to minimize both the transmission power and the total transmission time, resulting in the following cost function

$$F(\boldsymbol{P}, \boldsymbol{L}) = \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} p_n(t)\, l_n(t) + \lambda \sum_{t\in\mathcal{T}} \mathbb{I}\left\{ \sum_{n\in\mathcal{N}} l_n(t) > 0 \right\} \tag{3}$$

where $\boldsymbol{P} = [\boldsymbol{p}_n]_{n\in\mathcal{N}} \in \mathbb{R}^{T\times N}$, $\boldsymbol{L} = [\boldsymbol{l}_n]_{n\in\mathcal{N}} \in \mathbb{R}^{T\times N}$, and $\mathbb{I}\{\mathcal{A}\}$ is an indicator function defined as $\mathbb{I}\{\mathcal{A}\} = 1$ if condition $\mathcal{A}$ is satisfied, and $\mathbb{I}\{\mathcal{A}\} = 0$ otherwise. Specifically, $\sum_{n\in\mathcal{N}} l_n(t) > 0$ indicates that slot t is active and $\sum_{t\in\mathcal{T}} \mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$ represents the number of active slots. As a result, the first term in (3) captures the transmission power, the second term captures the total transmission time, and λ is the weighting factor for the two terms.

This paper aims to control the interference to the ground BSs at known locations and generate less interference to the ground users at unknown locations. As a result, an interference-constrained predictive communication problem is formulated to plan the transmission strategy $\boldsymbol{P}$ and $\boldsymbol{L}$ for T time slots ahead of the transmission

$$\underset{\boldsymbol{P},\boldsymbol{L}}{\text{minimize}} \quad F(\boldsymbol{P}, \boldsymbol{L}) \tag{4}$$

$$\text{subject to} \quad \mathbb{E}\{\Upsilon_n(\boldsymbol{p}_n, \boldsymbol{l}_n)\} \geq S_n, \forall n \tag{5}$$

$$\mathbb{E}\{I_{m,n}(t)\} \leq I_{\text{bs}}, \forall m, n, t \tag{6}$$

$$p_n(t) \in \mathcal{P}, \forall n, t,\ \boldsymbol{l}(t) \in \mathcal{L}, \forall t. \tag{7}$$

The objective function (4) represents the cost defined in (3), incorporating the transmission energy and duration. Constraint (5) specifies the minimum expected throughput S_n, and constraint (6) ensures the expected air-to-ground interference is within thresholds for nodes with known locations. Constraint (7) requires the allocated power to be non-negative and the allocated frequency resources to be non-overlapping.

[1]The continuous frequency resource allocation, *i.e.*, $l_n(t) \in [0,1]$, can be extended to a discrete allocation in a straightforward manner.

III. Radio-Map-Assisted Optimization

The radio-map-assisted optimization faces a significant challenge in dealing with channel uncertainty, as only large-scale information is available. This section addresses this challenge by analyzing the expected throughput and interference using large-scale information from the radio map. A robust formulation is thus constructed, then transformed and relaxed into a convex problem that enables efficient and effective solution techniques to be applied. Furthermore, we propose an efficient dual-based algorithm to accelerate the search for optimal solutions.

A. Radio-Map-Assisted Reformulation

With the aid of the radio map $\boldsymbol{\Theta}$ and the known locations of the transmitter, receivers, and neighboring nodes at time t, $\boldsymbol{q}_j(t)$, the large-scale channel parameters at time t are $(g_j(t), \kappa_j(t)) = \boldsymbol{\Theta}(\boldsymbol{q}_0(t), \boldsymbol{q}_j(t))$. Thus, the power gain at time t, $h_j(t)$ defined in (1), follows Gamma distribution $\mathrm{G}(\kappa_j(t), g_j(t)/\kappa_j(t))$. Based on the statistical information of $h_n(t)$, $n \in \mathcal{N}$, the expected channel capacity can be explicitly lower bounded as follows.

Lemma 1 (A deterministic capacity lower bound): The expected capacity $\mathbb{E}\{\log(1+p_n h_n)\}$ *is lower bounded by* $\log(1+p_n g_n) - \epsilon_n$ *where* $\epsilon_n = \log(e)/\kappa_n - \log(1+(2\kappa_n)^{-1})$.

Proof: See Appendix A. □

The gap ϵ_n is always positive, and tends to 0 when κ_n goes to infinity where the channel is asymptotically deterministic.

As a result of Lemma 1, the expected throughput $\mathbb{E}\{\Upsilon_n(\boldsymbol{p}_n, \boldsymbol{l}_n)\}$ is lower bounded by

$$\mathbb{E}\{\Upsilon_n(\boldsymbol{p}_n, \boldsymbol{l}_n)\} \geq \sum_{t\in\mathcal{T}} \left(\log\left(1+p_n(t)\, g_n(t)\right) - \epsilon_n\right) l_n(t).$$

In other words, the expected throughput constraint (5) can be relaxed to

$$\sum_{t\in\mathcal{T}} c_n(t)\, l_n(t) \geq S_n, \forall n \tag{8}$$

where $c_n(t) = \log\left(1+p_n(t)\, g_n(t)\right) - \epsilon_n(t)$ is the approximated channel capacity.

In addition, using the statistical information of $h_m(t)$, $m \in \mathcal{M} \triangleq \{N+1, \cdots, N+M\}$, the expected interference $\mathbb{E}\{I_{m,n}\}$ at node m can be computed as $\mathbb{E}\{I_{m,n}\} = g_m(t) \cdot p_n(t)$. As a result, the interference constraint (6) can be equivalently transformed into a power constraint as

$$p_n(t) \leq \bar{p}(t) \triangleq I_{\mathrm{bs}} / \max_{m\in\mathcal{M}} g_m(t), \forall n, t. \tag{9}$$

Using the approximated throughput constraint (8) and interference-equivalent power constraint (9), the original problem can be relaxed into the following problem

$$\begin{aligned} \mathscr{P}1: \quad \underset{\boldsymbol{P},\boldsymbol{L}}{\text{minimize}} \quad & F(\boldsymbol{P}, \boldsymbol{L}) \\ \text{subject to} \quad & \sum_{t\in\mathcal{T}} c_n(t)\, l_n(t) \geq S_n, \forall n \quad (10) \\ & p_n(t) \in \mathcal{P}', \forall n, t,\, \boldsymbol{l}(t) \in \mathcal{L}, \forall t \quad (11) \end{aligned}$$

where $\mathcal{P}' \triangleq \{p(t) : 0 \leq p(t) \leq \bar{p}(t)\}$.

However, this problem is non-convex due to the indicator function in the objective function in (3). The state-of-the-art techniques to handle the indicator function include LR [29] and SCA [30], but the optimality is unclear, and the computational complexity can be high.

In the rest of the paper, we develop a cost-aware relax-then-round scheme to solve problem $\mathscr{P}1$. First, the indicator functions $\mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$ in the objective (3) are relaxed to continuous ones, then we transform the non-convex problem to a convex equivalent problem and propose an efficient algorithm. Subsequently, in Section IV, a cost-aware rounding strategy is developed to ensure asymptotic optimality.

B. Convex Relaxation

1) Relaxation from Indicator Function: Recall that the presence of the indicator function in the objective (3) implies the need for integer programming to optimize the on-off mode. We replace the indicator $\mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$ by a continuous variable $\sum_{n\in\mathcal{N}} l_n(t)$ which can take continuous value from 0 to 1. This relaxation results in the following objective

$$\tilde{F}(\boldsymbol{P}, \boldsymbol{L}) = \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} \left(p_n(t) + \lambda\right) l_n(t). \tag{12}$$

It is clear that the relaxed objective (12) gives a lower bound of the cost function (3), because $\sum_{n\in\mathcal{N}} l_n(t) \leq \mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$ with equality for $\sum_{n\in\mathcal{N}} l_n(t) = 0$ or 1. Such a property motivates our rounding strategy, which will be discussed in Section IV. As a result, problem $\mathscr{P}1$ becomes

$$\mathscr{P}2: \quad \underset{\boldsymbol{P},\boldsymbol{L}}{\text{minimize}} \quad \tilde{F}(\boldsymbol{P}, \boldsymbol{L}), \quad \text{subject to} \quad (10), (11).$$

Note that the optimal value of problem $\mathscr{P}2$ is a lower bound on the optimal value of $\mathscr{P}1$, because $\tilde{F}(\boldsymbol{P}, \boldsymbol{L}) \leq F(\boldsymbol{P}, \boldsymbol{L})$ for all $(\boldsymbol{P}, \boldsymbol{L})$ and the feasible sets of the two problems are the same.

2) Convex Transformation: To handle the non-convex constraint (10), we introduce a new variable

$$\phi_n(t) \triangleq \left(\log\left(1+p_n(t)\, g_n(t)\right) - \epsilon_n(t)\right) l_n(t) \tag{13}$$

to replace the variable $p_n(t)$. As a result, the throughput constraint (10) becomes linear

$$\sum_{t\in\mathcal{T}} \phi_n(t) \geq S_n, \forall n \tag{14}$$

and the power constant in (11) also becomes linear

$$\phi_n(t) \in \left[-\epsilon_n(t)\, l_n(t), \bar{c}_n(t)\, l_n(t)\right] \tag{15}$$

where $\bar{c}_n(t) \triangleq \log(1+\bar{p}_n(t) g_n(t)) - \epsilon_n(t)$, due to the monotonic increasing property of the function $\phi_n(t)$ over $p_n(t)$. In addition, the variable $p_n(t)$ can be derived to a function of $\phi_n(t)$ and $l_n(t)$, as

$$p_n(t) = \begin{cases} \left(2^{\phi_n(t)/l_n(t) + \epsilon_n(t)} - 1\right)/g_n(t) & l_n(t) > 0 \\ 0 & l_n(t) = 0. \end{cases} \tag{16}$$

As a result, the cost function (12) can be expressed as

$$\sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}}\left(\frac{2^{\frac{\phi_n(t)}{l_n(t)}+\epsilon_n(t)}-1}{g_n(t)}+\lambda\right)l_n(t) \tag{17}$$

which is convex, because for each t and n, the summand $((2^{\phi_n(t)/l_n(t)+\epsilon_n(t)}-1)/g_n(t)+\lambda)l_n(t)$ is a perspective function of a convex function $\varphi(x)\triangleq(2^{x+\epsilon_n(t)}-1)/g_n(t)+\lambda$, and thus is convex [31].

Using the transformed objective function (17), and the transformed constraints (14)–(15), the problem $\mathscr{P}2$ is transformed into the following problem

$$\begin{aligned}\mathscr{P}3:\quad \underset{\boldsymbol{\Phi},\boldsymbol{L}}{\text{minimize}}\quad & \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}}\left(\frac{2^{\frac{\phi_n(t)}{l_n(t)}+\epsilon_n(t)}-1}{g_n(t)}+\lambda\right)l_n(t)\\ \text{subject to}\quad & \sum_{t\in\mathcal{T}}\phi_n(t)\geq S_n,\forall n\\ & \phi_n(t)\in[-\epsilon_n(t)l_n(t),\bar{c}_n(t)l_n(t)],\forall n,t\\ & \boldsymbol{l}(t)\in\mathcal{L},\forall t\end{aligned}$$

where $\boldsymbol{\Phi}\triangleq[\phi_n(t)]_{t\in\mathcal{T},n\in\mathcal{N}}\in\mathbb{R}^{T\times N}$.

The transformed problem $\mathscr{P}3$ is convex because the objective function is convex, and all the constraints are linear. Therefore, it can be solved efficiently.

C. Optimal Solution for $N=1$

While problem $\mathscr{P}3$ is convex, solving problem $\mathscr{P}3$ by general solvers, such as the (iterative) subgradient or interior-point methods, has complexity higher than $\mathcal{O}(N^3T^3)$ per iteration [32], which will be overwhelming for a large number of time slots T when we want to optimize over a large time horizon or for smaller time slots. As a result, the key challenge in radio-map-assisted predictive communication is to develop efficient algorithms to solve the optimization problem for a large T. To tackle this challenge, we first develop a semi-closed form solution to $\mathscr{P}3$ under $N=1$.

For simplicity, we omit the subscript n in the variables such as ϕ_n and l_n since $N=1$. Then in problem $\mathscr{P}3$, the resource allocation variables become $\boldsymbol{\phi}=[\phi(t)]_{t\in\mathcal{T}}$ and $\boldsymbol{l}=[l(t)]_{t\in\mathcal{T}}$, and the cost is written as $\sum_{t\in\mathcal{T}}((2^{\phi(t)/l(t)+\epsilon(t)}-1)/g(t)+\lambda)l(t)$. Since $\mathscr{P}3$ is convex, the solution to $\mathscr{P}3$ satisfies the Karush-Kuhn-Tucker (KKT) conditions which can be derived as follows.

Lemma 2: let x_0 be the solution to $\vartheta(x)=0$, where

$$\vartheta(x)\triangleq\left(2^{x+\epsilon(t)}-1\right)/g(t)+\lambda-x\left(\ln 2\cdot 2^{x+\epsilon(t)}/g(t)\right) \tag{18}$$

where ln *represents logarithm with base e throughout the paper. The root x_0 is unique in the region $x\geq 0$. Moreover, $0<x_0\leq\max\{2/\ln 2,\log(\lambda g(t)-1)-\epsilon(t)\}$.*

Proof: See Appendix B □

Denote $\hat{c}(t)$ as the solution to $\vartheta(x;g(t),\epsilon(t))=0$ for $x\geq 0$. From Lemma 2, the mapping from $g(t)$ and $\epsilon(t)$ to $\hat{c}(t)$ is unique. In addition, denote $\hat{\mu}(t)\triangleq\ln 2\cdot 2^{\hat{c}(t)+\epsilon(t)}/g(t)$ representing the point with maximal energy efficiency, $\bar{\mu}(t)\triangleq\ln 2\cdot(1/g(t)+\bar{p}(t))$ representing the point with maximal transmission power, and $\tilde{\mu}(t)\triangleq(\bar{p}(t)+\lambda)/(\log(1+\bar{p}(t)g(t))-\epsilon(t))$ representing the point with maximal energy efficiency and maximal transmission power, which happens only when $\hat{c}(t)>\bar{c}(t)$. We have the following result to characterize the solution to $\mathscr{P}3$ under $N=1$.

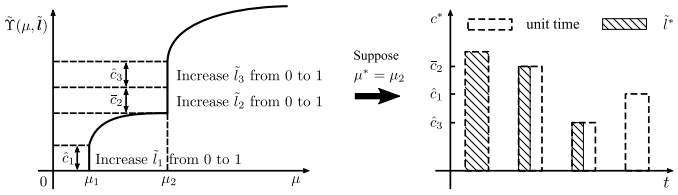

Fig. 2. Illustration of $\tilde{\Upsilon}(\mu,\tilde{\boldsymbol{l}})$ over μ and $\tilde{\boldsymbol{l}}$ (left), where one slot is activated when $\mu=\mu_1$, and two slots are activated when $\mu=\mu_2$, and a resource allocation example (right) based on Proposition 1.

Proposition 1 (Optimal Solution Under $N=1$): The optimal solution $(\boldsymbol{\phi},\boldsymbol{l})$ to $\mathscr{P}3$ under $N=1$ is given by

(i) for $t\in\mathcal{T}_1\triangleq\{t\in\mathcal{T}:\hat{c}(t)\leq\bar{c}(t)\}$

$$\phi(t)=\begin{cases}0 & \mu<\hat{\mu}(t)\\ \hat{c}(t)\tilde{l}(t) & \mu=\hat{\mu}(t)\\ \log(\mu g(t)/\ln 2)-\epsilon(t) & \hat{\mu}(t)<\mu<\bar{\mu}(t)\\ \bar{c}(t) & \mu\geq\bar{\mu}(t)\end{cases} \tag{19}$$

$$l(t)=\mathbb{I}\{\mu>\hat{\mu}(t)\}+\tilde{l}(t)\mathbb{I}\{\mu=\hat{\mu}(t)\} \tag{20}$$

(ii) for $t\in\mathcal{T}_2\triangleq\{t\in\mathcal{T}:\hat{c}(t)>\bar{c}(t)\}$

$$\phi(t)=\begin{cases}0 & \mu<\tilde{\mu}(t)\\ \bar{c}(t)\tilde{l}(t) & \mu=\tilde{\mu}(t)\\ \bar{c}(t) & \mu>\tilde{\mu}(t)\end{cases} \tag{21}$$

$$l(t)=\mathbb{I}\{\mu>\tilde{\mu}(t)\}+\tilde{l}(t)\mathbb{I}\{\mu=\tilde{\mu}(t)\} \tag{22}$$

where $\{\mu\geq 0,\tilde{\boldsymbol{l}}\triangleq[\tilde{l}(t)]_{t\in\mathcal{T}}\}$ are parameters chosen to satisfy

$$\tilde{\Upsilon}(\mu,\tilde{\boldsymbol{l}})\triangleq\sum_{t\in\mathcal{T}}\phi(t)=S \tag{23}$$

and $\tilde{l}(t)\in[0,1]$, $\forall t\in\mathcal{T}$.

Proof: See Appendix C. □

As a result, we can obtain the optimal solution to $\mathscr{P}3$ by solving equation (23). In addition, the throughput function $\tilde{\Upsilon}(\mu,\tilde{\boldsymbol{l}})$ is proven to be monotonically non-decreasing over μ and $\tilde{\boldsymbol{l}}$, as shown in the following proposition.

Proposition 2 (Monotonicity): For any $\mu_1>\mu_2\geq 0$ and $\mathbf{1}\succcurlyeq\tilde{\boldsymbol{l}}_1,\tilde{\boldsymbol{l}}_2\succcurlyeq\mathbf{0}$,[2] it holds that $\tilde{\Upsilon}(\mu_1,\tilde{\boldsymbol{l}}_1)\geq\tilde{\Upsilon}(\mu_2,\tilde{\boldsymbol{l}}_2)$. For any $\mathbf{1}\succcurlyeq\tilde{\boldsymbol{l}}_1\succ\tilde{\boldsymbol{l}}_2\succcurlyeq\mathbf{0}$ and $\mu\geq 0$, it holds that $\tilde{\Upsilon}(\mu,\tilde{\boldsymbol{l}}_1)\geq\tilde{\Upsilon}(\mu,\tilde{\boldsymbol{l}}_2)$.

Proof: See Appendix D. □

The left figure in Fig. 2 illustrates monotonicity of $\tilde{\Upsilon}(\mu,\tilde{\boldsymbol{l}})$. It is shown that (i) When $\mu<\mu_1$, $\tilde{\Upsilon}=0$; (ii) When $\mu>\mu_1$, $\tilde{\Upsilon}$ is increasing over μ; (iii) When $\mu=\mu_i$, $\tilde{\Upsilon}$ is increasing over $\tilde{l}_i$. Based on the monotonicity, one solution μ^* and $\tilde{\boldsymbol{l}}^*$ can be found using bisection search as described in Algorithm 1. The right figure in Fig. 2 illustrates a resource allocation example, supposing $\mu^*=\mu_2$. It is shown that (i) for the slots with $\hat{\mu}(t)<\mu^*<\bar{\mu}(t)$, *e.g.*, $t=1$, $\phi(t)$ is related to μ^*; (ii) for

[2]Here, the notation $\boldsymbol{a}\succcurlyeq\boldsymbol{b}$ means that $\forall i$, $a_i\geq b_i$. The notation $\boldsymbol{a}\succ\boldsymbol{b}$ means that $\forall i$, $a_i\geq b_i$, and $\exists i, a_i>b_i$.

Algorithm 1 Fast Dual-Based Algorithm for $N = 1$

Bisection search to determine μ^*.
1) Set $\mu_{\max} = \max_{t\in\mathcal{T}}\{\bar{\mu}(t), \tilde{\mu}(t)\}$, $\mu_{\min} = 0$;
2) Set $\mu = (\mu_{\max} + \mu_{\min})/2$, if $\tilde{\Upsilon}(\mu, \mathbf{0}) > S$, set $\mu_{\max} = \mu$; otherwise, set $\mu_{\min} = \mu$;
3) Repeat step 2) until $|\mu_{\max} - \mu_{\min}|$ is small enough;
4) Set $\mu^* = \mu_{\min}$.

Bisection search to determine $\tilde{\boldsymbol{l}}^*$.
5) Set $l_{\max} = 1$, $l_{\min} = 0$;
6) Set $l = (l_{\max} + l_{\min})/2$ and $\tilde{\boldsymbol{l}} = l \cdot \mathbf{1}$, if $\tilde{\Upsilon}(\mu^*, \tilde{\boldsymbol{l}}) > S$, set $l_{\max} = l$; otherwise, set $l_{\min} = l$;
7) Repeat step 2) until $|l_{\max} - l_{\min}|$ is small enough;
8) Set $l^* = l_{\max}$ and $\tilde{\boldsymbol{l}} = l^* \cdot \mathbf{1}$.

Set $\boldsymbol{\phi}^*$ and $\boldsymbol{l}^*$ according to (19) – (22).

the slots with $\mu = \hat{\mu}(t)$ or $\mu = \tilde{\mu}(t)$, *e.g.*, $t = \{2, 3\}$, $\phi(t) = \hat{c}(t)\tilde{l}(t)$ or $\phi(t) = \bar{c}(t)\tilde{l}(t)$; (iii) for the slots with $\mu < \hat{\mu}(t)$ or $\mu < \tilde{\mu}(t)$, *e.g.*, $t = 4$, $\phi(t) = 0$.

The computational complexity of Algorithm 1 is $\mathcal{O}(T)$ consisted of $\mathcal{O}(1)$ steps for the bisection search of μ^* and $\tilde{\boldsymbol{l}}^*$, where each step requires $\mathcal{O}(T)$ for calculating $\tilde{\Upsilon}(\mu, \tilde{\boldsymbol{l}})$.

D. Efficient Algorithm for $N > 1$

It is observed from $\mathscr{P}3$ that for $N > 1$, the variables are coupled over n only by the constraint $\boldsymbol{l} \in \mathcal{L}$. This observation suggests that $\mathscr{P}3$ can be decomposed into N subproblems, provided the coupled constraint involving $l_n(t)$ can be handled. To this end, a partial-dual-based algorithm is proposed, where the partial dual problem is given by

$$\underset{\boldsymbol{v}\succeq 0}{\text{maximize}} \quad q(\boldsymbol{v})$$

where $q(\boldsymbol{v})$ is the partial dual function over the Lagrangian parameter $\boldsymbol{v} = [v(t)]_{t\in\mathcal{T}}$, as given by the value of the following problem

$$\begin{aligned}\mathscr{P}4: \quad \underset{\boldsymbol{\Phi},\boldsymbol{L}}{\text{minimize}} \quad & Q(\boldsymbol{\Phi}, \boldsymbol{L}; \boldsymbol{v}) \\ \text{subject to} \quad & \sum_{t\in\mathcal{T}} \phi_n(t) \geq S_n, \forall n \\ & \phi_n(t) \in [-\epsilon_n(t) l_n(t), \bar{c}_n(t) l_n(t)], \forall n, t \\ & l_n(t) \in [0, 1], \forall n, t\end{aligned}$$

and

$$Q(\boldsymbol{\Phi}, \boldsymbol{L}; \boldsymbol{v}) \triangleq \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} \left(\frac{2^{\frac{\phi_n(t)}{l_n(t)}+\epsilon_n(t)} - 1}{g_n(t)} + \lambda\right) l_n(t) + \sum_{t\in\mathcal{T}} v(t)\left(\sum_{n\in\mathcal{N}} l_n(t) - 1\right).$$

Denote $\lambda(t) \triangleq \lambda + v(t)$, the objective function can be expressed as

$$Q(\boldsymbol{\Phi}, \boldsymbol{L}; \boldsymbol{v}) = \sum_{n\in\mathcal{N}}\sum_{t\in\mathcal{T}} \left(\frac{2^{\frac{\phi_n(t)}{l_n(t)}+\epsilon_n(t)} - 1}{g_n(t)} + \lambda(t)\right) l_n(t) - \sum_{t\in\mathcal{T}} v(t).$$

Algorithm 2 Alternative Optimization Algorithm for $N > 1$

Initialization process: set $\boldsymbol{v}_0$.
1) Alternatively optimize $(\boldsymbol{\phi}_n, \boldsymbol{l}_n)$ from $n = 1$ to N by Algorithm 1;
2) Update $\boldsymbol{v}_{k+1}$ using gradient projection method and go to step 1 until $||\boldsymbol{v}_{k+1} - \boldsymbol{v}_k||_F$ is small enough.

Output $(\boldsymbol{\Phi}^{(k)}, \boldsymbol{L}^{(k)})$.

Then the problem $\mathscr{P}4$ can be decoupled to N following sub-problem

$$\begin{aligned}\underset{\boldsymbol{\phi}_n,\boldsymbol{l}_n}{\text{minimize}} \quad & \sum_{t\in\mathcal{T}} \left(\frac{2^{\frac{\phi_n(t)}{l_n(t)}+\epsilon_n(t)} - 1}{g_n(t)} + \lambda(t)\right) l_n(t) - \sum_{t\in\mathcal{T}} \frac{v(t)}{N} \\ \text{subject to} \quad & \sum_{t\in\mathcal{T}} \phi_n(t) \geq S_n \\ & \phi_n(t) \in [-\epsilon_n(t) l_n(t), \bar{c}_n(t) l_n(t)], \forall t \\ & l_n(t) \in [0, 1], \forall t.\end{aligned}$$

The above problem is equivalent to $\mathscr{P}3$ for $N = 1$ case discussed in Section III-C, where the solution is given in Proposition 1 with the constant penalty λ to be replaced by $\lambda(t)$, and a constant $-\sum_{t\in\mathcal{T}} v(t)/N$ to be added into the objective.

For maximizing $q(\boldsymbol{v})$, the gradient projection method is used, and the algorithm stops when the $||\boldsymbol{v}_{k+1} - \boldsymbol{v}_k||_F$ is small enough. The detailed algorithm is shown in Algorithm 2.

Algorithm 2 converges because the dual problem is convex [31]. Moreover, the converged values $(\boldsymbol{\Phi}^{(k)}, \boldsymbol{L}^{(k)})$ is the optimal solution to $\mathscr{P}3$ because Slater's condition holds and strong duality holds since the objective function in $\mathscr{P}3$ is convex and all the constraints in $\mathscr{P}3$ are affine [31].

The computational complexity of Algorithm 2 is $\mathcal{O}(NT)$ per iteration, constructed by N times of $\mathcal{O}(T)$ for calculating $(\boldsymbol{\phi}_n, \boldsymbol{l}_n)$.

IV. Rounding Strategy

Although we can efficiently solve for $\mathscr{P}3$ via Algorithm 2, which equivalently solves for $\mathscr{P}2$, the solution may not be optimal to $\mathscr{P}1$, since $\mathscr{P}2$ is only a relaxed version of $\mathscr{P}1$. In some extreme cases, the solution to $\mathscr{P}2$ can lead to an unexpectedly large gap from the optimal value of $\mathscr{P}1$, which will be demonstrated in subsection IV-A. In this section, we develop a cost-aware rounding strategy to adjust the solution obtained from $\mathscr{P}2$, and show that the solution with rounding strategy is asymptotically optimal in small slot duration.

A. Challenge Due to Non-Unique Solutions to $\mathscr{P}2$

One particular scenario where the solution from problem $\mathscr{P}2$ yields a significant performance gap from $\mathscr{P}1$ is when the channel $g_n(t)$ is constant over t. In this case, the optimal solutions to $\mathscr{P}2$ are not unique, and a possible solution is to allocate equal power to a number of time slots which can be fractionally used, i.e., $\sum_{n\in\mathcal{N}} l_n(t)$ is strictly less than 1. This phenomenon can be understood by the water-filling solution for capacity-achieving power allocation, where the time slots with the same channel gain will be allocated

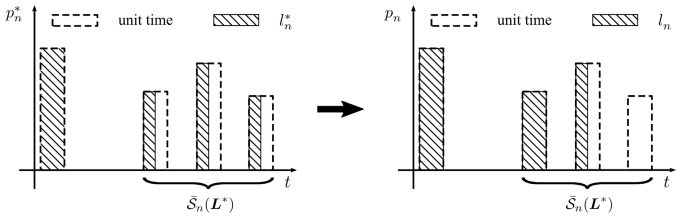

Fig. 3. The equivalence of multiple solutions under the optimality preservation conditions for receiver n.

Algorithm 3 Cost-Ware Rounding Algorithm

Initialization process: set $n \leftarrow 0$.
1) Calculate $\hat{\mathcal{S}}_n$ and $\hat{F}_n$ based on $(\boldsymbol{P}^*, \boldsymbol{L}^*)$;
2) Arbitrarily fixing the order of elements in the set $\hat{\mathcal{S}}_n$, and bisection search k and $\tilde{l}_n(t)$ according to (28);
3) Update $l_n^*(t)$ for $t \in \hat{\mathcal{S}}_n$ according to (27), set $n \leftarrow n+1$, and go to step 1 until $n > N$.

Output $\hat{\boldsymbol{L}} = \boldsymbol{L}^*$.

with the same power and the same frequency resource $l_n(t)$. While this solution is optimal in minimizing the relaxed cost function $\tilde{F}$ in (12), it is not optimal for the original cost function in (3), due to the discontinuous indicator function, $\mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$.

We construct a numerical example in Appendix E and show that directly applying the solution obtained in $\mathscr{P}2$ may yield an additional cost of $(T-1)\lambda$ in (3) above the minimum. Note that such a performance gap can be arbitrarily large for an arbitrarily large duration T.

Here, we investigate the conditions that characterize the set of equivalent solutions to $\mathscr{P}2$. Given a transmission strategy $(\boldsymbol{P}, \boldsymbol{L})$, let $\bar{\mathcal{S}}(\boldsymbol{L})$ denote the set of partially used slots defined as $\{t \in \mathcal{T} : \sum_{n\in\mathcal{N}} l_n(t) \notin \{0,1\}\}$. Let $\mathcal{T}_n(\boldsymbol{L})$ denote the set of slots allocated to node n as $\{t \in \mathcal{T} : l_n(t) > 0\}$. Then, the set of slots that are partially used and specifically used by node n is given by $\bar{\mathcal{S}}_n(\boldsymbol{L}) \triangleq \bar{\mathcal{S}}(\boldsymbol{L}) \cap \mathcal{T}_n(\boldsymbol{L})$.

Proposition 3 (Equivalent Solutions to $\mathscr{P}2$): For any optimal solution to $\mathscr{P}2$, $(\boldsymbol{P}^*, \boldsymbol{L}^*)$, a feasible solution $(\boldsymbol{P}, \boldsymbol{L})$ is also optimal if the following conditions are satisfied: (i) $p_n(t) = p_n^*(t)$, $\forall n \in \mathcal{N}, t \in \mathcal{T}$; (ii) $l_n(t) = l_n^*(t)$, $\forall t \notin \bar{\mathcal{S}}_n(\boldsymbol{L}^*)$, $\forall n \in \mathcal{N}$; (iii) $\sum_{t\in\bar{\mathcal{S}}_n(\boldsymbol{L}^*)}(p_n(t)+\lambda)l_n(t) = \sum_{t\in\bar{\mathcal{S}}_n(\boldsymbol{L}^*)}(p_n(t)+\lambda)l_n^*(t)$, $\forall n \in \mathcal{N}$.

Proof: See Appendix F. □

As a result, there are multiple optimal solutions to $\mathscr{P}2$ if the cardinality of $\bar{\mathcal{S}}_n(\boldsymbol{L}^*)$ is greater than 1. Fig. 3 illustrates the equivalence of multiple solutions under the conditions in Proposition 3. The solutions in the two subfigures of Fig. 3 have identical costs in (12), because the power allocation in the right subfigure is to concentrate the two half slots with the same power allocation in the left subfigure into one full slot. In addition, both solutions have identical throughput according to Proposition 3, because the cost in (12) consumed by the slots in set $\bar{\mathcal{S}}_n(\boldsymbol{L}^*)$ of the two solutions are identical. However, although the two solutions yield the same cost in (12), the solution on the right is more desired because it uses only 3 time slots as opposed to 4 time slots used by the solution on the left.

B. Cost-Aware Rounding Algorithm

To circumvent the issue from the non-unique optimal solutions, we develop an easy-to-implement strategy to find a better solution. At first, we prove that the cost gap is due to the number of partially used slots.

Proposition 4: (Cost upper bound) Denote F^* as the minimum cost of $\mathscr{P}1$. For any optimal solution $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ to $\mathscr{P}2$, the cost in (3) satisfies $F(\boldsymbol{P}^*, \boldsymbol{L}^*) - F^* \le |\bar{\mathcal{S}}(\boldsymbol{L}^*)|\lambda$.

Proof: See Appendix G. □

As a result, the performance gap to the optimum of $\mathscr{P}1$ can be decreased by reducing the number of the partially-used slot sets, $|\bar{\mathcal{S}}(\boldsymbol{L})|$.

Based on Proposition 3, one can derive a set of optimal solutions $\{(\boldsymbol{P}, \boldsymbol{L})\}$ based on a particular solution $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ to $\mathscr{P}2$. In addition, we are interested in the solution that minimizes $|\bar{\mathcal{S}}(\boldsymbol{L})|$ because of Proposition 4. Given a particular solution $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ to $\mathscr{P}2$, compute the parameter $\hat{\mathcal{S}}_n = \bar{\mathcal{S}}_n(\boldsymbol{L}^*)$ and $\hat{F}_n = \sum_{t\in\hat{\mathcal{S}}_n}(p_n^*(t)+\lambda)l_n^*(t)$. An equivalent solution to $\mathscr{P}2$ with a minimum $|\bar{\mathcal{S}}(\boldsymbol{L})|$ can be found by solving the following problem:

$$\begin{aligned}
\mathscr{P}5: \quad & \underset{\boldsymbol{L}}{\text{minimize}} \quad \left|\bar{\mathcal{S}}\left(\boldsymbol{L}\right)\right| \\
& \text{subject to} \quad \boldsymbol{l}\left(t\right) \in \mathcal{L}, \forall t \in \hat{\mathcal{S}}_n && (24)\\
& \qquad\qquad l_n\left(t\right) = l_n^*\left(t\right), \forall t \notin \hat{\mathcal{S}}_n, \forall n && (25)\\
& \qquad\qquad \sum_{t\in\hat{\mathcal{S}}_n}\left(p_n^*\left(t\right)+\lambda\right) l_n\left(t\right) = \hat{F}_n, \forall n && (26)
\end{aligned}$$

and $\boldsymbol{P} = \boldsymbol{P}^*$. It follows that $p_n(t) = p_n^*(t)$ meets the first condition in Proposition 3, and constraint (25) and constraint (26) correspond to the last two conditions in Proposition 3, respectively. Condition (24) ensures the solution is feasible to $\mathscr{P}2$. Thus, the optimal solution to $\mathscr{P}5$ is also an optimal solution to $\mathscr{P}2$ according to Proposition 3.

Problem$\mathscr{P}5$ can be solved by a sequential approach by decomposing it into N subproblems by partially optimizing variable $l_n(t)$. Specially, given $l_m^*(t)$, $m \neq n$, the constraint $l_n(t) \in \mathcal{L}$ in $\mathscr{P}5$ becomes $0 \le l_n(t) \le 1 - \sum_{m\neq n} l_m^*(t)$, and the objective becomes $|\bar{\mathcal{S}}_n(\boldsymbol{L})|$. Then the nth subproblem of $\mathscr{P}5$ degenerates to minimizing the cardinality of the set $\bar{\mathcal{S}}_n\left(\boldsymbol{L}\right)$ with modified constraint $0 \le l_n(t) \le 1 - \sum_{m\neq n} l_m^*(t)$.

For simplification, we use a bisection searching method to approximately solve the subproblem, that is, updating $l_n\left(t\right)$ for $t \in \hat{\mathcal{S}}_n$ while leaving $l_n\left(t\right)$ for $t \notin \hat{\mathcal{S}}_n$ unchanged, as shown in Algorithm 3. By arbitrarily fixing the order of elements in the set $\bar{\mathcal{S}}_n(\boldsymbol{L})$, as $\{q_1, \cdots q_K\}$, where $K = |\hat{\mathcal{S}}_n|$, we opt for the first k slots in $\hat{\mathcal{S}}_n$ for full usage, and allocate the $(k+1)$th slot according to the constraint (26), while leaving the remaining $K-k-1$ slots unused. In other words,

$$l_n\left(t\right) = \begin{cases} 1 - \sum_{m\neq n} l_m^*(t) & t \in \{q_1, \cdots q_k\} \\ \tilde{l}_n\left(t\right) & t = q_{k+1} \\ 0 & t \in \{q_{k+1}, \cdots q_K\} \end{cases} \quad (27)$$

where $k \in \mathbb{Z}^+$ and $\tilde{l}_n(t) \in [0, 1 - \sum_{m\neq n} l_m^*(t))$ is chosen to satisfy condition (26), that is the solution of the following

equation

$$\sum_{n\in\{q_1,\cdots,q_k\}} \left(p_n^*(t)+\lambda\right)\left(1-\sum_{m\neq n} l_m^*(t)\right) + \left[\left(p_n^*(t)+\lambda\right)\tilde{l}_n(t)\right]_{n=q_{k+1}} = \hat{F}_n \quad (28)$$

which can be solved by bisection searching method because the summation function is monotonically increasing over k.

1) Complexity Analysis: The Algorithm 3 consists of N rounds of bisection searching for k and $\tilde{l}_n(t)$ with $\mathcal{O}(T\log(T))$ computational complexity, consisted of $\mathcal{O}(\log(T))$ steps for the bisection search of k^* and $\tilde{l}_n^*(t)$, where each step requires $\mathcal{O}(T)$ for calculating $\sum_{n\in\{q_1,\cdots,q_k\}}(p_n^*(t)+\lambda)(1-\sum_{m\neq n} l_m^*(t))$. As a result, the computational complexity of the Algorithm 3 is $\mathcal{O}(NT\log(T))$.

2) Performance Analysis: Denote the slot duration as δ, and the transmission duration as Γ, and consequently, we have $T=\Gamma/\delta$. We prove that the cost gap reduces to $N\delta\lambda$ by the cost-aware rounding algorithm.

Proposition 5 (Asymptotic optimality): Let F^* be the minimum cost of $\mathscr{P}1$ and $\hat{\boldsymbol{L}}$ be the output of Algorithm 3. Then, $F(\boldsymbol{P}^*,\hat{\boldsymbol{L}})-F^*\leq N\delta\lambda$, i.e.*, $F(\boldsymbol{P}^*,\hat{\boldsymbol{L}})\to F^*$ as $\delta\to 0$.*

Proof: See Appendix H. □

It is shown from Proposition 5 that the proposed rounding algorithm yields a solution that is asymptotically optimal for small slot duration δ.

The overall computational complexity is $\mathcal{O}(NT\omega + NT\log(T))$, where ω is the iteration number of Algorithm 2. This is contributed by the two sequential processes: optimization of the relaxed problem (Algorithm 2) with complexity $\mathcal{O}(NT)$ per iteration, and optimization of the rounding problem (Algorithm 3) with complexity $\mathcal{O}(NT\log(T))$. Moreover, when there is only one receiver, the overall computational complexity is $\mathcal{O}(T\log(T))$, where the optimization of the relaxed problem can be solved by Algorithm 1, with complexity $\mathcal{O}(T)$.

V. Simulation

Consider a 200×200 m^2 area, where 5 BSs, acting as the nodes with known locations, and $N_{\text{ue}}=100$ ground users, acting as the nodes with unknown locations, are randomly distributed, one UAV, serving as the transmitting node, navigates along the $y=100$ route, and four UAVs, two flying horizontally and two vertically, acting as the receiving nodes, traverse this area, as shown in Fig. 4. The height of users, BSs, vertical-route UAVs, and horizontal-route UAVs are set 0 m, 10 m, 95 m, and 100 m. The speeds of transmitting and receiving UAVs are 5 m/s and 3 m/s, respectively.

The channels are realized by $h_j=g_j\xi_j$ according to (1), where g_j includes path loss and shadowing $[-g_j]_{dB}=\text{PL}_j+\chi$, and ξ_j follows a Gamma distribution $\text{G}(\kappa_j,1/\kappa_j)$. The shape parameter κ_j is chosen randomly from 1 to 30. The shadowing is modeled by a log-normal distribution, where χ follows zero mean and 8 variance with the correlated distance 5 m. The propagation for LOS and non-line-of-sight (NLOS) links is taken from the 3GPP Urban Micro (UMi) model in [33] as

$$\text{PL}_j=\begin{cases}22.0+28.0\log_{10}(d_j)+20\log_{10}(f_c), & \text{LOS link}\\ 22.7+36.7\log_{10}(d_j)+26\log_{10}(f_c) & \text{NLOS link}\end{cases}$$

where d_j is the distance between the transmitting node and the receiving or interfered nodes, and $f_c=3$ Ghz represents the carrier frequency.

Since there is the low blockage probability between UAVs, we assume the channel between the transmitting node and the receiving nodes is always in LOS, *i.e.*, $\text{PL}_j=\text{PL}_{\text{LOS}}$, and there is no shadowing between them, *i.e.*, $\chi=0$.

The blockage status between UAV and ground nodes is simulated using the following steps. Initially, we employ the LOS probability model [34], [35]

$$\mathbb{P}(\text{LOS},\theta)=\frac{1}{1+a\exp\left(-b\left[\theta-a\right]\right)} \quad (29)$$

to generate LOS probability according to the elevation angle $\theta=\sin^{-1}(u_j/d_j)\times 180/\pi$, where u_j represents their relative height, and the parameters are set $a=11.95$ and $b=0.14$ [35]. Subsequently, the blockage status is determined based on a threshold $p_{\text{th}}=0.5$. The channel is considered blocked if $\mathbb{P}(\text{LOS},\theta)<p_{\text{th}}$, otherwise, the channel is in LOS.

We evaluate the following baseline schemes

1) Best Effort With Perfect CSI – Best Effort (Non-Predictive): This scheme determines the transmission strategy by the best-effort transmission approach based on perfect instantaneous CSI. Specifically, given the current instantaneous CSI, the transmission strategy is determined by maximizing the sum capacity for nodes that have not yet

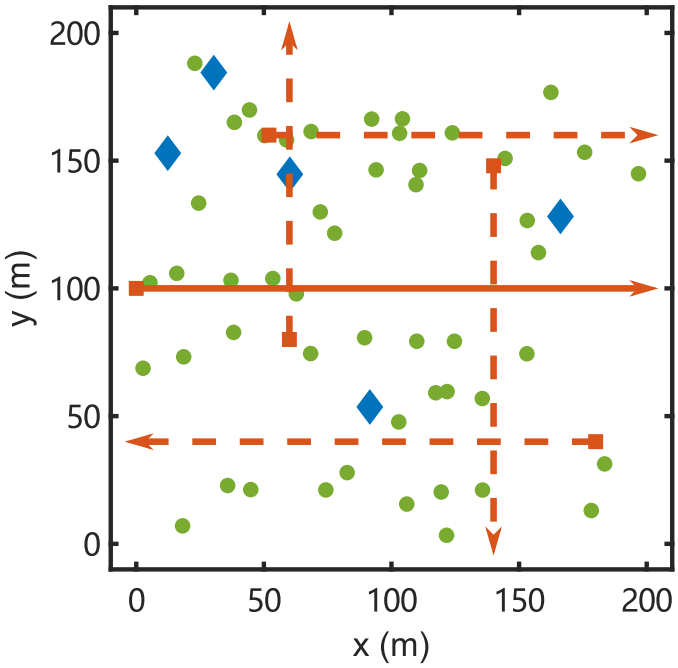

Fig. 4. Illustration for BSs positions (blue diamonds), user positions (green circles), and UAV routes (red lines with directions), where the solid line represents the route of transmitting UAV and the dashed lines represent the route of receiving UAV.

completed full data transmission in each time slot, subject to the power and interference constraints.

2) Predictive Communication Without a Radio Map – Predictive (Model-Based): When a radio map is not available, the probabilistic LOS model (29) is used for constructing an average channel gain model as $\bar{g}_j = \mathbb{P}(\text{LOS}, \theta) \cdot \text{PL}_{\text{LOS}} + (1 - \mathbb{P}(\text{LOS}, \theta)) \cdot \text{PL}_{\text{NLOS}}$. The predicted channel capacity is approximated by $\bar{c}_n \triangleq \log(1 + \beta p_n g_n)$, where $\beta = 0.5$ to back-off for the small-scale fading. The scheme then solves $\mathscr{P}1$ based on the predicted average channel gain $\bar{g}_j$ and the predicted channel capacity $\bar{c}_n$.

3) Predictive Communication With Non-Orthogonal Transmission Scheme – Predictive (Non-Orthogonal): This scheme uses the radio map information and adopts a non-orthogonal transmission approach, which uses superposition coding with multi-user detection and successive interference cancellation (SIC) to achieve capacity. Similar to work [36], a SCA-based algorithm is performed to make the transmission strategy. See Appendix I for details.

4) Proposed Predictive Optimization With Radio Maps – Predictive (Map-Based): The proposed scheme first calculates the attenuation $\epsilon_n(t)$ according to Lemma 1 and the interference-equivalent power upper bound $\bar{p}(t)$ according to (10), based on radio maps $\boldsymbol{\Theta}$. It then obtains an optimal solution to $\mathscr{P}2$ using Algorithm 2. Finally, the transmission strategy plan is obtained through rounding the solution to $\mathscr{P}2$ based on Algorithm 3.

A. Interference Leakage

We first evaluate the impact of interference leakage on BSs with the required data size $S_n = 500$ Mbits and interference threshold $I_{\text{bs}} = -70$ dBm. Fig. 5 shows the strongest interference power received among the 5 BSs at each time slot over a duration of $T = 40$ seconds. The results demonstrate that the proposed predictive optimization, which explores radio maps, can prevent all unintended receivers with known positions from experiencing interference (beyond the interference threshold). In contrast, when a radio map is not available, the baseline scheme cannot guarantee that the interference constraints are satisfied.

Subsequently, we evaluate the impact of interference leakage on users at unknown locations, as depicted in Fig. 6. For each time slot, a user is deemed to be interfered if the received interference power exceeds the interference threshold of users I_{ue}. Fig. 6 shows the cumulative ratio of interfered users at each time slot over a period of $T = 40$ seconds, where the ratio of interfered users at time t is computed as $r(t) = \sum_{q=1}^{N_{\text{ue}}} \mathbb{E}\{\mathbb{I}\{\max_{n\in\mathcal{N}} p_n(t) h_q(t) > I_{\text{ue}}\}\} / (N_{\text{ue}} T)$, where $h_q(t)$ is the channel gain between transmitting node to user q. It is observed that the cumulative ratio of interfered users for the best effort scheme over the 40 seconds, statistically roughly 50% of users experience at least -100 dBm interference and 20% of users experience at least -80 dBm interference for the entire period. For the proposed scheme with radio maps, these ratios reduce to 40% (at $\lambda = 100$) and 5% (at $\lambda = 0$), respectively, achieving a gain of 10%–15%. This is because the proposed scheme can reduce interference power by employing radio maps to concentrate the power on the good channels, and reduce the interference time by introducing the weighting factor to concentrate the transmission in fewer slots. The property of short transmission duration is also observed in Fig. 5 for the map-assisted scheme under $\lambda = 100$. In addition, compared with the scheme without an indicator function to quantify the interference-related cost in (3), that is, $\lambda = 0$, the proposed scheme reduces the number of interfered ground users by 20% under $\lambda = 100$. This result justifies the advantage of introducing the indicator function in the cost function in (3), a simple but effective strategy for interference reduction.

Fig. 7 evaluates the impact of interference leakage over the delivery of different data volume S_n, where the average ratio of interfered users is computed as $\sum_{t\in\mathcal{T}} r(t)$. Compared with the non-predictive transmission scheme, on average, the number of users affected by interference is reduced by over 10%, even in the case without a radio map. With the employment

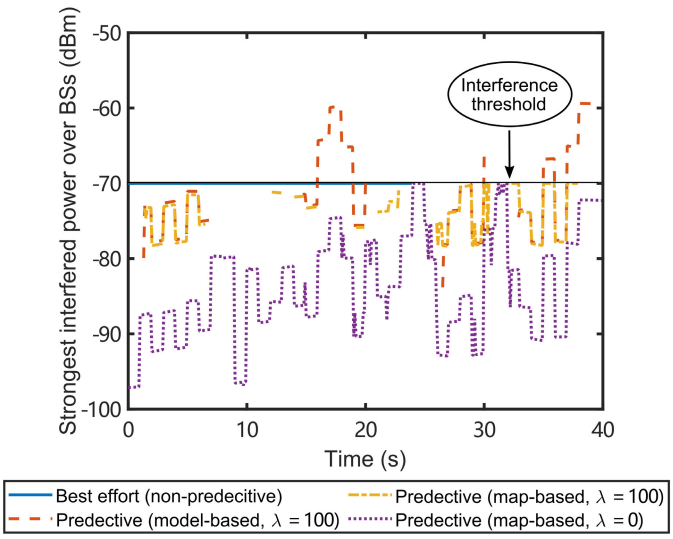

Fig. 5. Strongest interference power received among the 5 BSs at each time slot over a duration of 40 seconds.

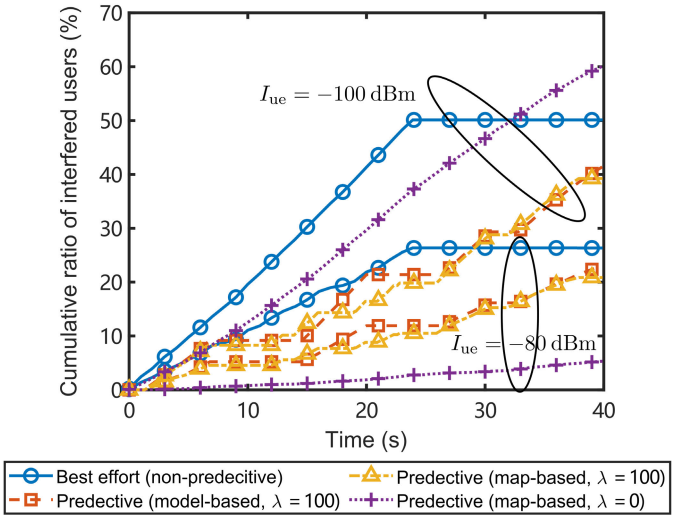

Fig. 6. Cumulative radio of interfered users at each time slot determined based on interference thresholds I_{ue}.

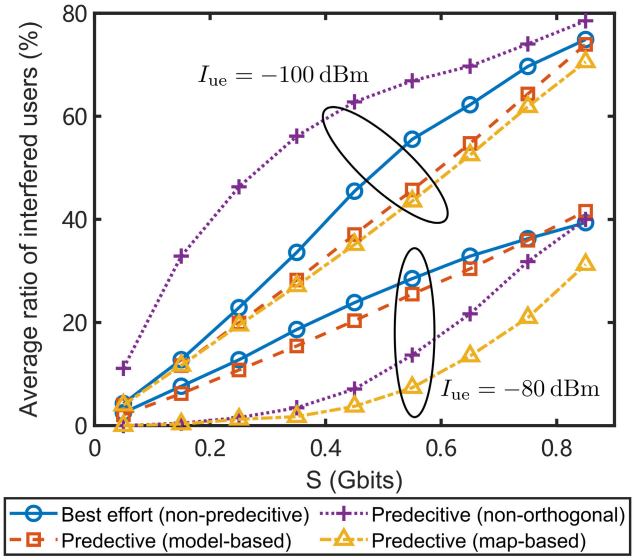

Fig. 7. Average ratio of interfered users over different required data size determined by different interference threshold I_{ue}.

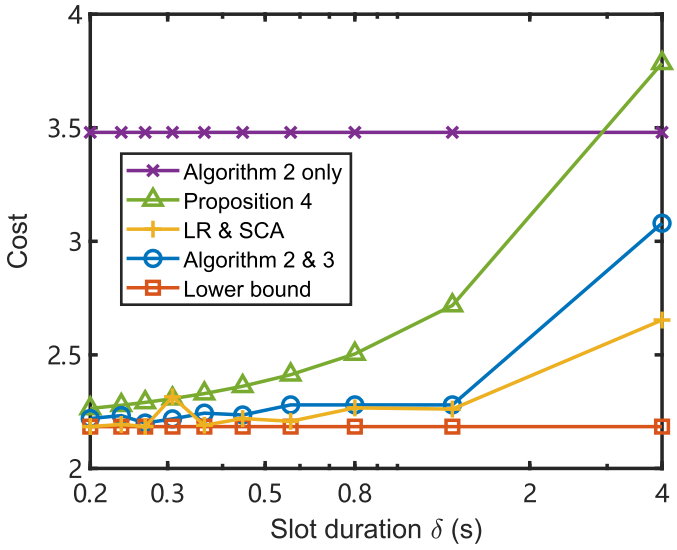

Fig. 8. Costs over different slot durations δ.

of radio maps, this figure is further reduced by more than 14% (at $I_{\text{ue}} = -100$ dBm) and 70% (at $I_{\text{ue}} = -80$ dBm), leading to fewer users being interfered with even without knowing their locations or channels. The superiority is significant when handling medium-sized data volumes, for example, when $S_n = 400$ Mbits, the proposed predictive scheme with radio maps outperforms the best effort scheme by 23% (at $I_{\text{ue}} = -100$ dBm) and 90% (at $I_{\text{ue}} = -80$ dBm). The non-orthogonal scheme performs worse than the proposed scheme, especially when the users are sensitive to the interference ($I_{\text{ue}} = -100$ dBm). This is because (1) the non-orthogonal scheme must use a looser capacity lower bound for handling the small-scale fading, such that higher power levels are planned, leading to a larger area of interference, and (2) the non-orthogonal scheme does not account for the duration of interference, resulting in prolonged interference periods.

B. Convergence and Complexity

We evaluate the following algorithms that also consider the universal interference and exploit the radio map, but use different strategies to plan resources.

1) LR & SCA: This algorithm solves the non-continuous non-convex problem $\mathscr{P}1$ through a four-step process. (i) Continuity Transformation: A new parameter $a(t) \triangleq \sum_{n\in\mathcal{N}} l_n(t)$ is introduced, replacing the indicator function $\mathbb{I}\{\sum_{n\in\mathcal{N}} l_n(t) > 0\}$, with constraints $a(t) \in [0,1]$ and $a(t)(1-a(t)) \leq 0$; (ii) LR [29]: The non-convex constraint $a(t)(1-a(t)) \leq 0$ is incorporated into the objective function with a penalty factor; (iii) SCA [30]: The local problem is approximated to a convex problem using Taylor's approximation; (iv) CVX solver: The local convex problem is solved using the CVX tool.

In addition, we evaluate the cost of the scheme without rounding, *i.e.*, applying the solution to Algorithm 2 for transmission planning, the cost lower bound, *i.e.*, the optimal value $\tilde{F}^*$ of the relaxed problem $\mathscr{P}2$, and the theoretical largest performance gap of the proposed relax-then-round strategy according to Proposition 4, *i.e.*, $\tilde{F}^* + N\delta\lambda$.

Fig. 8 depicts the cost with the slot duration δ. It is observed that the cost of the proposed scheme with cost-aware rounding (Algorithm 2 & 3) consistently remains below the theoretical upper bound (Proposition 4), and both converge towards the lower bound as the slot duration δ diminishes. This validates the performance gap solution in Proposition 4 and aligns with the asymptotic optimality of the proposed scheme. In comparison, a significant gap to the optimum is noted in the absence of a rounding process (Algorithm 2 only).

While LR & SCA and proposed schemes have similar performance, the computational complexity is significantly different. Table. I shows the running time over different slot lengths. The comparison includes the LR & SCA and CVX & Algorithm 3 schemes, with the latter addressing the proposed relax-then-rounding problem not with the proposed dual-based algorithm solving problem $\mathscr{P}3$, but with the CVX tool. It is demonstrated that the proposed dual-based algorithm achieves a time-saving factor of 8 times compared to the non-orthogonal algorithm, 50 times compared to the CVX scheme, and 1000 times compared to the LR & SCA scheme for the case of $N = 4$. These results suggest that the orthogonal scheme is easier to implement than the non-orthogonal scheme, the proposed dual-based algorithm is significantly more efficient than the general solver, and the proposed relax-then-round scheme outperforms the general LR-based and SCA-based algorithm in terms of efficiency. In addition, for the case of $N = 2$, the proposed dual-based algorithm achieves a time-saving factor of 100 times compared to the non-orthogonal algorithm, 1000 times compared to the CVX scheme, and 30,000 times compared to the LR & SCA scheme. This is because the parameter v in Algorithm 2 converges faster for smaller N, such that the running time is dominated by step 1 with computational complexity $\mathcal{O}(NT)$.

TABLE I
COMPUTATIONAL TIME (S) OVER DIFFERENTSLOT NUMBER T AND RECEIVING NODE N

T ($N = 4$)	10	50	90	130	170
LR & SCA	106.4	357.9	587.5	1165	1222
CVX & Algorithm 3	3.854	13.510	23.996	29.212	38.603
Non-orthogonal [36]	0.456	2.102	3.685	5.449	7.180
Algorithm 2 & 3	0.128	0.273	0.413	0.513	0.652
T ($N = 2$)	10	50	90	130	170
LR & SCA	57.78	114.7	121.3	232.9	306.5
CVX & Algorithm 3	2.306	6.099	10.26	14.18	18.32
Non-orthogonal [36]	0.144	0.717	1.341	1.934	2.40
Algorithm 2 & 3	0.002	0.003	0.004	0.008	0.009

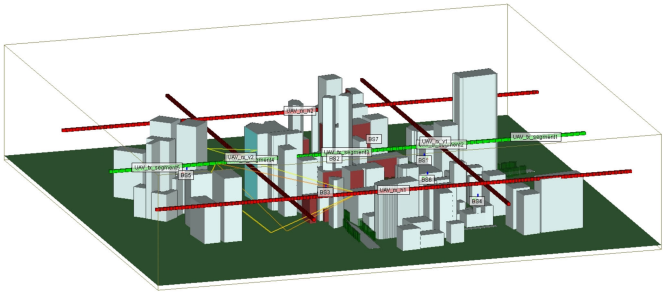

Fig. 9. The layout of the environment for low-altitude aerial communications and an example of propagation paths between two nodes.

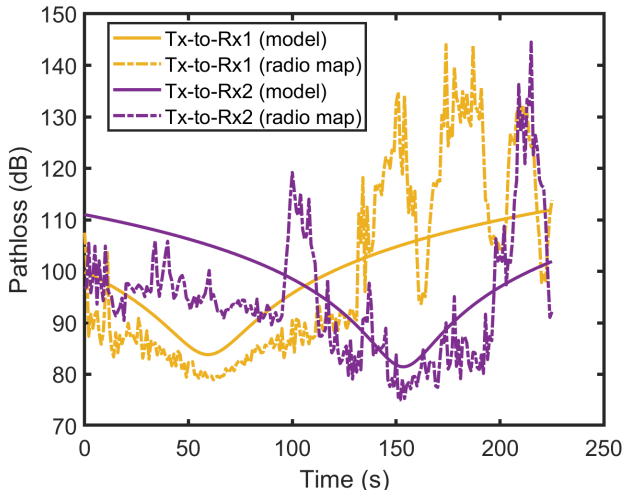

Fig. 10. Path loss comparison between model-based and radio-map-based predictions.

C. Experiment on a Synthetic Dataset Over a Realistic Environment

We conduct the experiments in a more realistic setting to verify the advantage of the radio map in predictive communication. Specifically, we use Wireless InSite to simulate aerial communication in San Francisco, as shown in Fig. 9, where the green vertical line represents the trajectory of node 0, and the other four red lines indicate the trajectories of the receiving UAVs. Figure 10 illustrates the path loss between the transmitter and two receivers, showing that channels cannot be easily predicted without a radio map.

We perform the proposed predictive transmission strategy on both the radio-map-predicted CSI and the model-predicted CSI, respectively. The results, shown in Fig. 11, indicate that

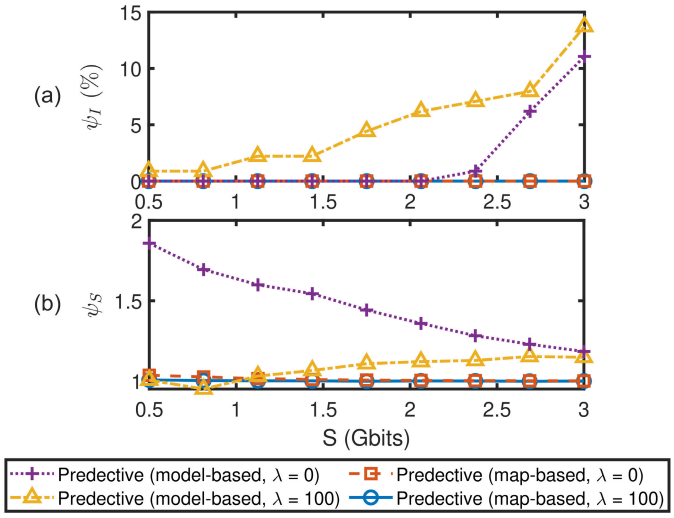

Fig. 11. (a) The time ratio of violation of the interference constraint, defined as $\psi_I \triangleq \sum_{t\in\mathcal{T}} \mathbb{I}\{\max_{m\in\mathcal{M},n\in\mathcal{N}} I_{m,n}(t) > I_{\text{bs}}\}/T$, over the data requirements. (b) Actual capacities versus throughput constraints, defined as $\psi_S \triangleq \min_{n\in\mathcal{N}} \hat{S}_n/S_n$, over the data requirements, where $\hat{S}_n$ is the actual capacity.

the transmission strategy using the radio-map-predicted CSI consistently ensures that the constraints are met. Specifically, the interference to BSs is always below the threshold, and the average capacity for the target user exceeds the throughput requirement. In contrast, the transmission strategy using the model-predicted CSI generates stronger interference and either cannot guarantee transmission requirements (when $\lambda = 100$) or occupies too many resources (when $\lambda = 0$). These results justify the advantage of using radio maps in predictive communication for better resource allocation and interference management.

VI. CONCLUSION

This paper exploits radio maps and develops a relax-then-round scheme for interference-aware predictive UAV communications. First, a radio-map-assisted optimization problem is formulated, addressing the uncertainty of future CSI and converting the interference limit to an equivalent power limit by exploiting radio maps. Then, the non-convex problem is relaxed and transformed into a convex one, and a dual-based algorithm with complexity $\mathcal{O}(NT)$ per iteration is proposed. Subsequently, a cost-aware rounding algorithm with complexity $\mathcal{O}(NT\log(T))$ is developed, which is theoretically proven to guarantee asymptotic optimality. Simulations validate that the radio-map-assisted scheme reduces the interference at the unintended receivers at known locations below a prescribed threshold, and substantially reduces the interference to the users at unknown locations. Furthermore, the numerical results confirm that the proposed relax-then-round scheme converges to the optimum as the slot duration diminishes, and the running time of the proposed scheme achieves orders of magnitude of reduction.

APPENDIX A
PROOF OF LEMMA 1

Since h_n follows Gamma distribution $\mathrm{G}(\kappa_n, g_n/\kappa_n)$, then $p_n h_n$ follows Gamma distribution $\mathrm{G}(\kappa_n, p_n g_n/\kappa_n)$. Based on

logarithmic Jensen's gap theory [37, Proposition 3], we have

$$\log\left(1+\mathbb{E}\left\{p_n h_n\right\}\right)-\mathbb{E}\left\{\log\left(1+p_n h_n\right)\right\}\leq\epsilon_n$$

where $\epsilon_n=\log(e)/\kappa_n-\log(1+1/(2\kappa_n))$. Thus we have

$$\begin{aligned}\mathbb{E}\left\{\log\left(1+p_n h_n\right)\right\}&\geq\log\left(1+\mathbb{E}\left\{p_n h_n\right\}\right)-\epsilon_n\\&=\log\left(1+p_n g_n\right)-\epsilon_n.\end{aligned}$$

Appendix B
Proof of Lemma 2

At first, the function $\vartheta(x)$ is monotonically decreasing in $x>0$, because its first derivative $\frac{\partial\vartheta}{\partial x}=-x(\ln 2)^2 2^{x+\epsilon(t)}/g\left(t\right)$ is less than 0 for $x>0$. Then, we have $\vartheta(0)=(2^{\epsilon(t)}-1)/g(t)+\lambda>0$ and

$$\begin{aligned}&\vartheta\left(\max\left\{2/\ln 2,\log\left(\lambda g\left(t\right)-1\right)-\epsilon\left(t\right)\right\}\right)\\&\leq\left(2^{x+\epsilon(t)}-1\right)/g\left(t\right)+\lambda-2\cdot\left(2^{x+\epsilon(t)}/g\left(t\right)\right)\\&=-2^{x+\epsilon(t)}/g\left(t\right)+\lambda-1/g\left(t\right)\\&\leq-\left(\lambda g\left(t\right)-1\right)/g\left(t\right)+\lambda-1/g\left(t\right)=0\end{aligned}$$

As a result, the function $\vartheta(x)$ has a unique root in $x\geq 0$, and the root lies in $(0,\max\{2/\ln 2,\log(\lambda g(t)-1)-\epsilon(t)\}]$.

Appendix C
Proof of Proposition 1

For any $t\in\mathcal{T}$, any feasible negative $\phi(t)$ will decrease the throughput and increase the cost, and thus there must be optimal $\phi^*(t)\geq 0$. In other words, the optimal solution to problem $\mathscr{P}3$ under $N=1$ is equivalent to the solution to the following problem

$$\begin{aligned}\mathscr{P}6:\quad&\underset{\boldsymbol{\phi},\boldsymbol{l}}{\text{minimize}}\quad\sum_{t\in\mathcal{T}}\left(\frac{2^{\frac{\phi(t)}{l(t)}+\epsilon(t)}-1}{g\left(t\right)}+\lambda\right)l\left(t\right)\\&\text{subject to}\quad\sum_{t\in\mathcal{T}}\phi\left(t\right)\geq S\\&\qquad\phi\left(t\right)\in\left[0,\bar{c}\left(t\right)l\left(t\right)\right],l\left(t\right)\in\left[0,1\right],\forall t.\end{aligned}$$

Since the objective function in $\mathscr{P}6$ is convex and all the constraints in $\mathscr{P}6$ are affine, then Slater's condition holds and strong duality holds [31]. In other words, the KKT conditions provide necessary and sufficient conditions for optimality [31].

A. Optimality Conditions

Let $\boldsymbol{\Lambda}=[\mu,\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4]$ be the Lagrangian parameter set, then the Lagrangian function of $\mathscr{P}6$ is given by

$$\begin{aligned}&L\left(\boldsymbol{\phi},\boldsymbol{l},\boldsymbol{\Lambda}\right)\\&=\sum_{t\in\mathcal{T}}\left(\frac{2^{\frac{\phi(t)}{l(t)}+\epsilon(t)}-1}{g\left(t\right)}+\lambda\right)l\left(t\right)+\mu\left(S-\sum_{t\in\mathcal{T}}\phi\left(t\right)\right)\\&\quad+\sum_{t\in\mathcal{T}}\left(v_1\left(t\right)\left(-\phi\left(t\right)\right)+v_2\left(t\right)\left(\phi\left(t\right)-\bar{c}\left(t\right)l\left(t\right)\right)\right)\\&\quad+\sum_{t\in\mathcal{T}}\left(v_3\left(t\right)\left(-l\left(t\right)\right)+v_4\left(t\right)\left(l\left(t\right)-1\right)\right).\end{aligned}$$

Let $(\boldsymbol{\phi}^*,\boldsymbol{l}^*)$ be the optimal solution to $\mathscr{P}6$ and $\boldsymbol{\Lambda}^*$ be the optimal Lagrange multiplier to its dual problem. Then we can derive from the KKT optimality conditions that $\forall t$, $(\boldsymbol{\phi}^*,\boldsymbol{l}^*,\boldsymbol{\Lambda}^*)$ should satisfy

$$\frac{\partial L}{\partial\phi\left(t\right)}=0\tag{30}$$

$$\frac{\partial L}{\partial l\left(t\right)}=0\tag{31}$$

$$-\phi\left(t\right)v_1\left(t\right)\geq 0,\ v_1\left(t\right)\geq 0\tag{32}$$

$$\left(\phi\left(t\right)-\bar{c}\left(t\right)l\left(t\right)\right)v_2\left(t\right)\geq 0,\ v_2\left(t\right)\geq 0\tag{33}$$

$$-l\left(t\right)v_3\left(t\right)\geq 0,\ v_3\left(t\right)\geq 0\tag{34}$$

$$\left(l\left(t\right)-1\right)v_4\left(t\right)\geq 0,\ v_4\left(t\right)\geq 0\tag{35}$$

$$\left(S-\sum_{t\in\mathcal{T}}\phi\left(t\right)\right)\mu\geq 0,\ \mu\geq 0.\tag{36}$$

Re-converting $\phi\left(t\right)$ by $c\left(t\right)$ and $l\left(t\right)$ according to $c\left(t\right)=\phi\left(t\right)/l\left(t\right)$, based on (30)-(36), we have that the optimal capacity and frequency allocation $(\boldsymbol{c}^*,\boldsymbol{l}^*)$ should satisfy $\forall t$, (34, 35) and

$$\ln 2\cdot 2^{c(t)+\epsilon(t)}/g\left(t\right)-\mu-v_1\left(t\right)+v_2\left(t\right)=0\tag{37}$$

$$\vartheta\left(c(t)\right)-v_2\left(t\right)\bar{c}\left(t\right)-v_3\left(t\right)+v_4\left(t\right)=0\tag{38}$$

$$-c\left(t\right)l\left(t\right)v_1\left(t\right)\geq 0,\ v_1\left(t\right)\geq 0\tag{39}$$

$$\left(c\left(t\right)l\left(t\right)-\bar{c}\left(t\right)l\left(t\right)\right)v_2\left(t\right)\geq 0,\ v_2\left(t\right)\geq 0\tag{40}$$

$$\left(S-\sum_{t\in\mathcal{T}}c\left(t\right)l\left(t\right)\right)\mu\geq 0,\ \mu\geq 0.\tag{41}$$

B. Optimal Allocation Strategy

Based on conditions (37, 39, 40), we can get the relationship between the optimal capacity $c^*(t)$ and the optimal Lagrangian multiplier μ^* during the non-zero allocated time $l^*(t)$. Note that when $l^*(t)$ is 0, the capacity $c^*(t)$ is 0, according to the definition of $p(t)$ in (16).

If $c^*(t)=0$, then $v_1\left(t\right)\geq 0$, $v_2\left(t\right)=0$ according to (39) and (40). Then, based on condition (37), we have $\mu^*\leq\ln 2\cdot 2^{c^*(t)+\epsilon(t)}/g(t)=\ln 2\cdot 2^{\epsilon(t)}/g\left(t\right)$. If $c^*(t)\in(0,\bar{c}(t))$, then $v_1\left(t\right)=0$, $v_2\left(t\right)=0$ according to (39) and (40). Then, based on condition (37), we have $\mu^*=\ln 2\cdot 2^{c^*(t)+\epsilon(t)}/g\left(t\right)$. If $c^*(t)=\bar{c}(t)$, then $v_1\left(t\right)=0$, $v_2\left(t\right)\geq 0$ according to (39) and (40). Then, based on condition (37), we have $\mu^*\geq\ln 2\cdot 2^{c^*(t)+\epsilon(t)}/g\left(t\right)=\ln 2\cdot 2^{\bar{c}(t)+\epsilon(t)}/g\left(t\right)$. In summary, the optimal allocated capacity can be expressed as a function of optimal Lagrangian multiplier μ^*, as

$$c^*\left(t\right)=\begin{cases}0 & \mu^*\leq\underline{\mu}\left(t\right)\\ \log\left(\mu^* g\left(t\right)/\ln 2\right)-\epsilon\left(t\right) & \underline{\mu}\left(t\right)<\mu^*<\overline{\mu}\left(t\right)\\ \bar{c}\left(t\right) & \mu^*\geq\overline{\mu}\left(t\right)\end{cases}\tag{42}$$

where $\underline{\mu}(t)=\ln 2\cdot 2^{\epsilon(t)}/g(t)$ and $\overline{\mu}(t)=\ln 2\cdot 2^{\bar{c}(t)+\epsilon(t)}/g(t)$.

Based on conditions (34, 35, 38, 40), we can get the relationship between the optimal time $l^*(t)$ and the optimal capacity $c^*(t)$.

When $c^*\left(t\right)=0$, there must be $l^*\left(t\right)=0$. Otherwise, $v_2\left(t\right)=0$ according to (40) and $v_3\left(t\right)=0$ according to (34), then condition (38) becomes

$$(2^{\epsilon(t)}-1)/g\left(t\right)+\lambda+v_4\left(t\right)=0$$

thus $(2^{\epsilon(t)}-1)/g(t)+\lambda \leq 0$, then $\lambda < 0$, which contradicts the setting of λ greater than 0.

When $c^*(t) > 0$, there are two cases based on whether the maximum generalized efficiency is achievable or not, where the generalized efficiency function is defined as $\eta(t) \triangleq c(t)/\varphi(c(t))$, where $\varphi(c(t)) = (2^{c(t)+\epsilon(t)}-1)/g(t)+\lambda$ represents the overall cost consumption. Since $\varphi(c(t))$ is convex over $c(t)$, then $\eta(t)$ is a quasi-concave function with respect to (w.r.t.) $c(t)$ [31], and the maximal generalized efficiency occurs where its first derivative $\frac{\partial \eta(t)}{\partial c(t)}$ equals to 0,

$$\frac{\left(2^{c(t)+\epsilon(t)}-1\right)/g(t)+\lambda - c(t)\left(\ln 2 \cdot 2^{c(t)+\epsilon(t)}/g(t)\right)}{\left(\left(2^{c(t)+\epsilon(t)}-1\right)/g(t)+\lambda\right)^2} = 0$$

that is, $\vartheta(c(t))/((2^{c(t)+\epsilon(t)}-1)/g(t)+\lambda)^2 = 0$ according to the definition of $\vartheta(c(t))$ in (18). Since $((2^{c(t)+\epsilon(t)}-1)/g(t)+\lambda)^2 > 0$ due to $c(t) \geq 0$, $\epsilon(t) \geq 0$, and $\lambda > 0$, the solution to the equation $\vartheta(c(t)) = 0$ with $c(t) \geq 0$ is the capacity with maximal generalized efficiency, denoted as $\hat{c}(t)$. Note that $\hat{c}(t)$ always exists and $\hat{c}(t) > 0$ according to Lemma 2. In other words,

$$\vartheta(c(t)) \begin{cases} > 0 & c(t) \in [0, \hat{c}(t)) \\ = 0 & c(t) = \hat{c}(t) \\ < 0 & c(t) > \hat{c}(t). \end{cases} \tag{43}$$

For the case (i) when the maximum generalized efficiency is achievable, that is, $t \in \mathcal{T}_1 \triangleq \{t \in \mathcal{T} : \hat{c}(t) \leq \bar{c}(t)\}$. If $c(t) \in [0, \hat{c}(t))$, we have $\vartheta(c(t)) > 0$ according to (43) and $v_2(t) = 0$ according to (40). Then based on (38), we have

$$-v_3(t) + v_4(t) < 0$$

which indicates $v_3(t) \neq 0$, thus $l^*(t) = 0$ according to (34).

If $c(t) = \hat{c}(t)$, we have $\vartheta(c(t)) = 0$ according to (43) and $v_2(t) = 0$ according to (40). Then based on (38), we have

$$-v_3(t) + v_4(t) = 0$$

which indicates $v_3(t) = v_4(t) \geq 0$, thus $l^*(t) \in [0,1]$ according to (34) and (35).

If $c(t) = (\hat{c}(t), \bar{c}(t)]$, we have $\vartheta(c(t)) < 0$ according to (43). Then based on (38), we have

$$-v_2(t)\bar{c}(t) - v_3(t) + v_4(t) > 0$$

which indicates $v_4(t) \neq 0$, thus $l^*(t) = 1$ according to (35).

In summary, the relationship between the optimal time $l^*(t)$ and the optimal capacity $c^*(t)$ is

$$l^*(t) \begin{cases} = 0 & c^*(t) \in [0, \hat{c}(t)) \\ \in [0,1] & c^*(t) = \hat{c}(t) \\ = 1 & c^*(t) \in (\hat{c}(t), \bar{c}(t)] \end{cases} \tag{44}$$

for $t \in \mathcal{T}_1 \triangleq \{t : \hat{c}(t) \leq \bar{c}(t)\}$.

Similarly, for $t \in \mathcal{T}_2 \triangleq \{t : \hat{c}(t) > \bar{c}(t)\}$, we have

$$l^*(t) \begin{cases} = 0 & c^*(t) \in [0, \bar{c}(t)) \\ \in [0,1] & c^*(t) = \bar{c}(t). \end{cases} \tag{45}$$

Finally, according to condition (41), the optimal solution $(c^*(t), l^*(t))$ is the solution to the following problem

$$\sum_{t \in \mathcal{T}} c(t)\, l(t) = S \tag{46}$$

otherwise, $\mu = 0$, then $c^*(t) = 0$ according to (42) and the throughput constraint cannot satisfy.

C. Dual Problem

Denote $\tilde{\boldsymbol{l}} \triangleq [\tilde{l}(t)]_{t \in \mathcal{T}}$ as the frequency allocation when $\mu = \hat{\mu}(t)$ for $t \in \mathcal{T}_1$ and when $\mu = \tilde{\mu}(t)$ for $t \in \mathcal{T}_2$. Combine the optimal capacity allocation strategy (35) and the optimal frequency allocation strategy, (44) and (45), the optimal $\phi(t)$ and $l(t)$ can be expressed as

(i) for $t \in \mathcal{T}_1$

$$\phi(t) = \begin{cases} 0 & \mu < \hat{\mu}(t) \\ \hat{c}(t)\,\tilde{l}(t) & \mu = \hat{\mu}(t) \\ \log(\mu g(t)/\ln 2) - \epsilon(t) & \hat{\mu}(t) < \mu < \bar{\mu}(t) \\ \bar{c}(t) & \mu \geq \bar{\mu}(t) \end{cases} \tag{47}$$

$$l(t) = \mathbb{I}\{\mu > \hat{\mu}(t)\} + \tilde{l}(t)\,\mathbb{I}\{\mu = \hat{\mu}(t)\} \tag{48}$$

(ii) for $t \in \mathcal{T}_2$

$$\phi(t) = \begin{cases} 0 & \mu < \tilde{\mu}(t) \\ \bar{c}(t)\,\tilde{l}(t) & \mu = \tilde{\mu}(t) \\ \bar{c}(t) & \mu > \tilde{\mu}(t) \end{cases} \tag{49}$$

$$l(t) = \mathbb{I}\{\mu > \tilde{\mu}(t)\} + \tilde{l}(t)\,\mathbb{I}\{\mu = \tilde{\mu}(t)\} \tag{50}$$

where (47) holds because $\underline{\mu}(t) < \hat{\mu}(t) \triangleq \ln 2 \cdot 2^{\hat{c}(t)+\epsilon(t)}/g(t)$ due to $\hat{c}(t) > 0$, and (49) holds because $\tilde{\mu}(t) = (\bar{p}(t)+\lambda)/(\log(1+\bar{p}(t)g(t)) - \epsilon(t))$ is the marginal rate of throughput cost when the slot is active at time t for $t \in \mathcal{T}_2$. Note that for $t \in \mathcal{T}_1$, the marginal rate of throughput cost, denoted as $\Delta\tilde{F}(t)/\Delta\phi(t)$, is always μ, according to Lemma 3.

Finally, combining the condition (46), the parameter $(\mu, \tilde{\boldsymbol{l}})$ are chosen from

$$\tilde{\Upsilon}\left(\mu, \tilde{\boldsymbol{l}}\right) \triangleq \sum_{t \in \mathcal{T}} \phi(t) = S.$$

Lemma 3: For $t \in \mathcal{T}_1$, the marginal rate of throughput cost is always μ.

Proof: There are two kinds of throughput increasing for $t \in \mathcal{T}_1$, i) increasing the allocated time Δt when $\mu = \hat{\mu}(t)$; ii) increasing the allocated power Δp when $\hat{\mu}(t) > \mu > \bar{\mu}(t)$. For case i), the marginal rate of throughput cost is

$$\frac{\Delta\tilde{F}(t)}{\Delta\phi(t)} = \frac{(\hat{p}(t)+\lambda)\cdot\Delta t}{\hat{c}(t)\cdot\Delta t} = \frac{(2^{\hat{c}(t)+\epsilon(t)}-1)/g(t)+\lambda}{\hat{c}(t)}. \tag{51}$$

Since $\vartheta(\hat{c}(t)) = 0$ according to (43), we have $(2^{\hat{c}(t)+\epsilon(t)} - 1)/g(t)+\lambda = \hat{c}(t)\cdot(\ln 2 \cdot 2^{\hat{c}(t)+\epsilon(t)}/g(t))$. Thus (51) becomes

$$\frac{\Delta\tilde{F}(t)}{\Delta\phi(t)} = \frac{\ln 2 \cdot 2^{\hat{c}(t)+\epsilon(t)}}{g(t)} \overset{(a)}{=} \hat{\mu}(t) = \mu$$

where (a) holds because $\hat{c}(t) = \log(\hat{\mu}(t)g(t)/\ln 2) - \epsilon(t)$ according to (42).

For case ii), the marginal rate of throughput cost is

$$\frac{\Delta\tilde{F}(t)}{\Delta\phi(t)} = \frac{\nabla_\mu p\,(t)\cdot\Delta\mu}{\nabla_\mu c\,(t)\cdot\Delta\mu} \tag{52}$$

where $c(t) = \log(\mu\cdot g(t)/\ln 2) - \epsilon(t)$ and $p(t) = (2^{c(t)+\epsilon(t)} - 1)/g(t) = (\mu g(t)/\ln 2 - 1)/g(t)$ according to (42). Thus (52) becomes

$$\frac{\Delta\tilde{F}(t)}{\Delta\phi(t)} = \frac{1\cdot\Delta\mu}{\frac{1}{\mu}\cdot\Delta\mu} = \mu.$$

□

Appendix D
Proof of Proposition 2

Function $\tilde{\Upsilon}(\mu, \tilde{\boldsymbol{l}})$ is a summation of non-negative function $\phi(t)$ over t according to its definition in (23). Thus, function $\tilde{\Upsilon}(\mu, \tilde{\boldsymbol{l}})$ has monotonicity property if function $\phi(t;\mu,l)$ over (μ, l) for all $t \in \mathcal{T}$ has monotonicity property. Accordingly, we will prove that for all $t \in \mathcal{T}$, any $\mu_1 > \mu_2 \geq 0$, and $1 \geq \tilde{l}_1, \tilde{l}_2 \geq 0$, it holds that $\phi(t;\mu_1,\tilde{l}_1) \geq \phi(t;\mu_2,\tilde{l}_2)$; then, prove that for all $t \in \mathcal{T}$, any $1 \geq \tilde{l}_1 > \tilde{l}_2 \geq 0$ and $\mu \geq 0$, it holds that $\phi(t;\mu,\tilde{l}_1) \geq \phi(t;\mu,\tilde{l}_2)$.

When $\mu_1 > \mu_2 \geq \bar{\mu}(t)$, $\phi(t;\mu_1,\tilde{l}_1) = \phi(t;\mu_2,\tilde{l}_2)$, $\forall \tilde{l}_1, \tilde{l}_2 \in [0,1]$. When $\bar{\mu}(t) > \mu_1 > \mu_2 > \hat{\mu}(t)$, the throughput $\log(\mu g(t)/\ln 2) - \epsilon(t)$ is positive and monotonically increasing over μ due to its first derivative $(\ln 2\cdot\mu)^{-1} > 0$, thus $\phi(t;\mu_1,\tilde{l}_1) > \phi(t;\mu_2,\tilde{l}_2)$, $\forall \tilde{l}_1, \tilde{l}_2 \in [0,1]$. When $\hat{\mu}(t) \geq \mu_1 > \mu_2 \geq 0$, $\lim_{\mu\to\hat{\mu}(t)^-}\phi(t) = 0 \leq \phi(t;\hat{\mu}(t),\tilde{l}) \leq \log(\hat{\mu}(t)g(t)/\ln 2) - \epsilon(t) = \lim_{\mu\to\hat{\mu}(t)^+}\phi(t)$ due to $\tilde{l} \in [0,1]$, thus $\phi(t;\mu_1,\tilde{l}_1) \geq \phi(t;\mu_2,\tilde{l}_2)$, $\forall \tilde{l}_1, \tilde{l}_2 \in [0,1]$.[3] In summary, for all $t \in \mathcal{T}$, any $\mu_1 > \mu_2 \geq 0$, and $1 \geq \tilde{l}_1, \tilde{l}_2 \geq 0$, it holds that $\phi(t;\mu_1,\tilde{l}_1) \geq \phi(t;\mu_2,\tilde{l}_2)$.

If $\mu \neq \hat{\mu}(t)$, $\tilde{l}$ has no effect on ϕ, then we have $\phi(t;\mu,\tilde{l}_1) = \phi(t;\mu,\tilde{l}_2)$ for any $1 \geq \tilde{l}_1 > \tilde{l}_2 \geq 0$. If $\mu \neq \hat{\mu}(t)$, ϕ is linear function over $\tilde{l}$, thus we have $\phi(t;\mu,\tilde{l}_1) > \phi(t;\mu,\tilde{l}_2)$ for any $1 \geq \tilde{l}_1 > \tilde{l}_2 \geq 0$. In summary, for all $t \in \mathcal{T}$, any $1 \geq \tilde{l}_1 > \tilde{l}_2 \geq 0$ and $\mu \geq 0$, it holds that $\phi(t;\mu,\tilde{l}_1) \geq \phi(t;\mu,\tilde{l}_2)$.

Appendix E
Large Gap Case

Consider a one-receiver zero-neighbor scenario with constant and deterministic channel, that is, $N = 1$, $M = 0$, $g(t) = g_\mathrm{c}$, and $\kappa(t) = \infty$ for all $t \in \mathcal{T}$. One of the optimal solutions to $\mathscr{P}2$ is $l'(t) = l_\mathrm{c}$ and $p'(t) = p_\mathrm{c}$ for all $t \in \mathcal{T}$, where $Tl_c\log(1 + p_c g_c) = S$. Suppose there is a throughput requirement S such that $\sum_{t\in\mathcal{T}} l'(t) \leq 1$, that is, $l_\mathrm{c}T \leq 1$, one can construct another transmission strategy $(\boldsymbol{P}'', \boldsymbol{L}'')$, where $l''(1) = l_\mathrm{c}T$, $p''(1) = p_\mathrm{c}$, and $l''(t) = 0$, $p''(t) = 0$ for $t \in \mathcal{T}\backslash\{1\}$.

The following can be observed for strategy $(\boldsymbol{P}'', \boldsymbol{L}'')$: (i) the strategy $(\boldsymbol{P}'', \boldsymbol{L}'')$ is also an optimal solution to problem $\mathscr{P}2$ because the relaxed cost (12) and throughput (2) under $(\boldsymbol{P}'', \boldsymbol{L}'')$ are the same as under $(\boldsymbol{P}', \boldsymbol{L}')$. (ii) The difference of the actual cost (3) of two strategies is $\lambda(T-1)$ because

$$\begin{aligned} &F\left(\boldsymbol{P}', \boldsymbol{L}'\right) - F\left(\boldsymbol{P}'', \boldsymbol{L}''\right) \\ &= \sum_{t\in\mathcal{T}} p'\,(t)\,l'\,(t) + \lambda\sum_{t\in\mathcal{T}}\mathbb{I}\left\{l'\,(t) > 0\right\} \\ &\quad - \sum_{t\in\mathcal{T}} p''\,(t)\,l''\,(t) + \lambda\sum_{t\in\mathcal{T}}\mathbb{I}\left\{l''\,(t) > 0\right\} \\ &= \lambda T - \lambda. \end{aligned}$$

In addition, the optimal value of $\mathscr{P}1$, denoted as F^*, is less than or equal to the cost in (3) by the solution to $\mathscr{P}2$, *i.e.*, $F^* \leq F(\boldsymbol{P}^*, \boldsymbol{L}^*)$, where $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ is an optimal solution to $\mathscr{P}2$, because $\mathscr{P}1$ and $\mathscr{P}2$ are two optimization problems with the same feasible set. As a result, the additional cost in (3) by the solution of $\mathscr{P}2$ can be given by

$$\begin{aligned} F\left(\boldsymbol{P}^*, \boldsymbol{L}^*\right) - F^* &\geq F\left(\boldsymbol{P}^*, \boldsymbol{L}^*\right) - F\left(\boldsymbol{P}'', \boldsymbol{L}''\right) \\ &\overset{(a)}{=} F\left(\boldsymbol{P}', \boldsymbol{L}'\right) - F\left(\boldsymbol{P}'', \boldsymbol{L}''\right) \\ &= \lambda\left(T-1\right). \end{aligned}$$

where (a) holds because $(\boldsymbol{P}', \boldsymbol{L}')$ is one of the optimal solution of $\mathscr{P}2$.

Appendix F
Proof of Proposition 3

Denote the generalized efficiency on the optimal solution as

$$\eta_n\left(t\right) = c_n^*\left(t\right)/\left(p_n^*\left(t\right) + \lambda\right) \tag{53}$$

where $c_n^*(t) = \log(1 + p_n^*(t)g_n(t)) - \epsilon_n(t)$.

Lemma 4: The generalized efficiencies of all partially-used slots by any receiver n are the same, that is, $\eta_n(t_1) = \eta_n(t_2)$, $\forall t_1, t_2 \in \bar{\mathcal{S}}_n(\boldsymbol{L}^*)$.

Proof: Suppose that there is an optimal solution to $\mathscr{P}2$, $(\boldsymbol{P}^*, \boldsymbol{L}^*)$, where there are two partially-used slots $t_1, t_2 \in \bar{\mathcal{S}}_n$ with different generalized efficiencies, without loss of generality (w.l.o.g.), $\eta_n(t_1) > \eta_n(t_2)$. One can construct another frequency allocation strategy $\boldsymbol{L}'$, where $l_n'(t_1) = l_n^*(t_1) + l_\mathrm{a}$, $l_n'(t_2) = l_n^*(t_2) - l_\mathrm{s}$, and $l_n'(t) = l_n^*(t)$ for all $t \in \mathcal{T}\backslash\{t_1, t_2\}$, such that $\Upsilon_n(\boldsymbol{P}^*, \boldsymbol{L}^*) = \Upsilon_n(\boldsymbol{P}^*, \boldsymbol{L}')$, that is

$$\begin{aligned} &\eta_n\left(t_1\right)\left(p_n^*\left(t_1\right) + \lambda\right) l_n^*\left(t_1\right) + \eta_n\left(t_2\right)\left(p_n^*\left(t_2\right) + \lambda\right) l_n^*\left(t_2\right) \\ &= \eta_n\left(t_1\right)\left(p_n^*\left(t_1\right) + \lambda\right) l_n'\left(t_1\right) + \eta_n\left(t_2\right)\left(p_n^*\left(t_2\right) + \lambda\right) l_n'\left(t_2\right) \end{aligned}$$

based on the definition of generalized efficiency in (53). Thus we have $\eta_n\left(t_1\right)\left(p_n^*\left(t_1\right) + \lambda\right) l_\mathrm{a} = \eta_n\left(t_2\right)\left(p_n^*\left(t_2\right) + \lambda\right) l_\mathrm{s}$.

However, the cost consumed by strategy $(\boldsymbol{P}^*, \boldsymbol{L}')$ is less than by $(\boldsymbol{P}^*, \boldsymbol{L}^*)$. This is due to the less cost for transmitting to receiver n

$$\begin{aligned} &F_n(\boldsymbol{P}^*, \boldsymbol{L}') \\ &= \left(p_n^*\left(t_1\right) + \lambda\right) l_n'\left(t_1\right) + \left(p_n^*\left(t_2\right) + \lambda\right) l_n'\left(t_2\right) + E_\mathrm{r} \\ &= E_n(\boldsymbol{P}^*, \boldsymbol{L}^*) + \left(p_n^*\left(t_2\right) + \lambda\right) l_\mathrm{s}\left(\eta_n\left(t_2\right)/\eta_n\left(t_1\right) - 1\right) \\ &\overset{(a)}{<} F_n(\boldsymbol{P}^*, \boldsymbol{L}^*) \end{aligned} \tag{54}$$

where $E_\mathrm{r} \triangleq \sum_{t\in\mathcal{T}\backslash\{t_1,t_2\}}(p_n^*(t) + \lambda)l_n^*(t)$, and (a) holds because $\eta_n(t_1) > \eta_n(t_2)$. Thus, $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ is a strictly sub-optimal, which contradicts the hypothesis. As a result, for all

[3] Here, a^- represents the left-sided limit and a^+ represents the right-sided limit.

$n \in \mathcal{N}$, the generalized efficiencies $\eta_n(t)$ of all partially-used slots $t \in \bar{\mathcal{S}}_n$ are the same. □

Then, we prove that the strategy $(\boldsymbol{P}, \boldsymbol{L})$ is also optimal by showing the cost and throughput are the same for both strategies $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ and $(\boldsymbol{P}, \boldsymbol{L})$.

The throughput on any receiver n under $(\boldsymbol{P}, \boldsymbol{L})$ is the same as under $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ because

$$\begin{aligned}\Upsilon_n(\boldsymbol{P}^*, \boldsymbol{L}^*) &= \sum_{t\in\bar{\mathcal{S}}_n^*} \phi_n^*(t) + \sum_{t\in\mathcal{T}\backslash\bar{\mathcal{S}}_n^*} \phi_n^*(t) \\ &\overset{(a)}{=} \sum_{t\in\bar{\mathcal{S}}_n^*} \phi_n^*(t) + \sum_{t\in\mathcal{T}\backslash\bar{\mathcal{S}}_n^*} \phi_n(t) \\ &\overset{(b)}{=} \sum_{t\in\bar{\mathcal{S}}_n^*} \phi_n(t) + \sum_{t\in\mathcal{T}\backslash\bar{\mathcal{S}}_n^*} \phi_n(t) \\ &= \Upsilon_n(\boldsymbol{P}, \boldsymbol{L})\end{aligned}$$

where $\bar{\mathcal{S}}_n^* \triangleq \bar{\mathcal{S}}_n(\boldsymbol{L}^*)$, (a) holds because $p_n(t) = p_n^*(t)$ and $l_n(t) = l_n^*(t)$ for all $t \in \mathcal{T}\backslash\bar{\mathcal{S}}_n^*$ according to condition (i) and (ii), and (b) holds because

$$\begin{aligned}&\sum_{t\in\bar{\mathcal{S}}_n^*} l_n^*(t)\, c_n^*(t) \overset{(c)}{=} \sum_{t\in\bar{\mathcal{S}}_n^*} l_n^*(t)\, \eta_n(t)\, (p_n^*(t) + \lambda) \\ &\overset{(d)}{=} \bar{\eta}_n \sum_{t\in\bar{\mathcal{S}}_n^*} l_n^*(t)\, (p_n^*(t) + \lambda) \overset{(e)}{=} \bar{\eta}_n \sum_{t\in\bar{\mathcal{S}}_n^*} l_n(t)\, (p_n(t) + \lambda) \\ &\overset{(f)}{=} \sum_{t\in\bar{\mathcal{S}}_n^*} \eta_n(t)\, l_n(t)\, (p_n(t) + \lambda) \overset{(g)}{=} \sum_{t\in\bar{\mathcal{S}}_n^*} l_n(t)\, c_n(t)\end{aligned}$$

where (c) and (g) hold according to (53), (d) and (f) holds according to Lemma 4, and (e) holds due to condition (iii).

The relaxed cost in (12) under $(\boldsymbol{P}, \boldsymbol{L})$ is the same as under $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ because

$$\begin{aligned}&\tilde{F}(\boldsymbol{P}^*, \boldsymbol{L}^*) \\ &= \sum_{n=1}^{N} \left(\sum_{t\in\bar{\mathcal{S}}_n^*} (p_n^*(t) + \lambda)\, l_n^*(t) + \sum_{t=\mathcal{T}\backslash\bar{\mathcal{S}}_n^*} (p_n^*(t) + \lambda)\, l_n^*(t) \right) \\ &\overset{(a)}{=} \sum_{n=1}^{N} \left(\sum_{t\in\bar{\mathcal{S}}_n^*} (p_n^*(t) + \lambda)\, l_n^*(t) + \sum_{t=\mathcal{T}\backslash\bar{\mathcal{S}}_n^*} (p_n(t) + \lambda)\, l_n(t) \right) \\ &\overset{(b)}{=} \sum_{n=1}^{N} \left(\sum_{t\in\bar{\mathcal{S}}_n^*} (p_n(t) + \lambda)\, l_n(t) + \sum_{t=\mathcal{T}\backslash\bar{\mathcal{S}}_n^*} (p_n(t) + \lambda)\, l_n(t) \right) \\ &= \tilde{F}(\boldsymbol{P}', \boldsymbol{L}')\end{aligned}$$

where (a) holds because $p_n(t) = p_n^*(t)$ and $l_n(t) = l_n^*(t)$ for all $t \in \mathcal{T}\backslash\bar{\mathcal{S}}_n^*$ according to condition (i) and (ii), and (b) holds according to condition (iii).

As a result, the strategy $(\boldsymbol{P}, \boldsymbol{L})$ is also optimal if the three conditions in Proposition 3 hold.

Appendix G
Proof of Proposition 4

The performance gap is bounded by

$$\begin{aligned}&F(\boldsymbol{P}^*, \boldsymbol{L}^*) - F^* \\ &\overset{(a)}{\le} F(\boldsymbol{P}^*, \boldsymbol{L}^*) - \tilde{F}(\boldsymbol{P}^*, \boldsymbol{L}^*) \\ &= \lambda \sum_{t\in\bar{\mathcal{S}}(\boldsymbol{L}^*)} \left(\mathbb{I}\left\{ \sum_{n\in\mathcal{N}} l_n^*(t) > 0 \right\} - \sum_{n\in\mathcal{N}} l_n^*(t) \right) \\ &\quad + \lambda \left(\sum_{t\in\mathcal{T}\backslash\bar{\mathcal{S}}(\boldsymbol{L}^*)} \mathbb{I}\left\{ \sum_{n\in\mathcal{N}} l_n^*(t) > 0 \right\} - \sum_{n\in\mathcal{N}} l_n^*(t) \right) \\ &\overset{(b)}{=} \lambda \sum_{t\in\bar{\mathcal{S}}(\boldsymbol{L}^*)} \left(1 - \sum_{n\in\mathcal{N}} l_n^*(t) \right) \\ &\le \lambda \sum_{t\in\bar{\mathcal{S}}(\boldsymbol{L}^*)} (1) = \left|\bar{\mathcal{S}}(\boldsymbol{L}^*)\right| \lambda\end{aligned}$$

where (a) holds because the optimal value of problem $\mathscr{P}2$ is a lower bound on the optimal value of $\mathscr{P}1$, and (b) holds because when $t \in \bar{\mathcal{S}}(\boldsymbol{L}^*)$, then $\sum_{n\in\mathcal{N}} l_n^*(t) \in (0, 1)$, thus $\mathbb{I}\{\sum_{n\in\mathcal{N}} l_n^*(t) > 0\} = 1$, while when $t \in \mathcal{T}\backslash\bar{\mathcal{S}}(\boldsymbol{L}^*)$, then $\mathbb{I}\{\sum_{n\in\mathcal{N}} l_n^*(t) > 0\} = \sum_{n\in\mathcal{N}} l_n^*(t)$ no matter of $\sum_{n\in\mathcal{N}} l_n^*(t) = 0$ or 1.

Appendix H
Proof of Proposition 5

Denote the nth round of solution of Algorithm 3 as $\boldsymbol{L}^{(n)}$. We will prove by induction that the transmission strategy $(\boldsymbol{P}^*, \boldsymbol{L}^{(n)}), \forall n \in \mathcal{N}$ is optimal to $\mathscr{P}2$. Firstly, $(\boldsymbol{P}^*, \boldsymbol{L}^{(0)})$ is a solution to $\mathscr{P}2$ because $(\boldsymbol{P}^*, \boldsymbol{L}^*)$ is a solution to $\mathscr{P}2$ and $\boldsymbol{L}^{(0)} = \boldsymbol{L}^*$. Secondly, suppose $(\boldsymbol{P}^*, \boldsymbol{L}^{(n-1)})$ is an optimum to $\mathscr{P}2$ for all $n \in \{1, \cdots, N-1\}$, then $(\boldsymbol{P}^*, \boldsymbol{L}^{(n)})$ is also an optimum to $\mathscr{P}2$ according to Proposition 3, because the constraints in problem $\mathscr{P}5$ ensure the conditions in Proposition 3 hold. As a result, the strategy $(\boldsymbol{P}^*, \hat{\boldsymbol{L}}) = (\boldsymbol{P}^*, \boldsymbol{L}^{(N)})$ is an optimal solution to $\mathscr{P}2$.

Next, we prove the cardinality of partly used slots in $\hat{\boldsymbol{L}}$ are less than N

$$\begin{aligned}\left|\bar{\mathcal{S}}\left(\hat{\boldsymbol{L}}\right)\right| &= \left|\bar{\mathcal{S}}\left(\boldsymbol{L}^{(N)}\right)\right| \overset{(a)}{\le} \sum_{n=1}^{N} \left|\bar{\mathcal{S}}_n\left(\boldsymbol{L}^{(N)}\right)\right| \\ &\overset{(b)}{\le} \bar{\mathcal{S}}_N\left(\boldsymbol{L}^{(N)}\right) + \sum_{n=1}^{N-1} \left|\bar{\mathcal{S}}_n\left(\boldsymbol{L}^{(N-1)}\right)\right| \\ &\cdots \\ &\overset{(c)}{\le} \sum_{n=1}^{N} \left|\bar{\mathcal{S}}_n\left(\boldsymbol{L}^{(n)}\right)\right| \overset{(d)}{\le} N\end{aligned}$$

where (a) holds because

$$\begin{aligned}\bar{\mathcal{S}}\left(\boldsymbol{L}^{(N)}\right) &= \bar{\mathcal{S}}\left(\boldsymbol{L}^{(N)}\right) \cap \bigcup_{n=1}^{N} \mathcal{T}_n\left(\boldsymbol{L}^{(N)}\right) \\ &= \bigcup_{n=1}^{N} \left(\bar{\mathcal{S}}\left(\boldsymbol{L}^{(N)}\right) \cap \mathcal{T}_n\left(\boldsymbol{L}^{(N)}\right)\right) \\ &= \bigcup_{n=1}^{N} \bar{\mathcal{S}}_n\left(\boldsymbol{L}^{(N)}\right)\end{aligned}$$

(d) holds because $\bar{\mathcal{S}}_n(\boldsymbol{L}^{(n)}) \subseteq \{q_{k+1}\}$ according to (27), and (b) to (c) hold because for any $n \in \{2, \cdots, N\}$

$$\begin{aligned}\sum_{i=1}^{n-1}\left|\bar{\mathcal{S}}_i\left(\boldsymbol{L}^{(n)}\right)\right| &= \sum_{i=1}^{n-1}\left|\bar{\mathcal{S}}\left(\boldsymbol{L}^{(n)}\right)\cap\mathcal{T}_i\left(\boldsymbol{L}^{(n)}\right)\right| \\ &\overset{(e)}{=} \sum_{i=1}^{n-1}\left|\bar{\mathcal{S}}\left(\boldsymbol{L}^{(n)}\right)\cap\mathcal{T}_i\left(\boldsymbol{L}^{(n-1)}\right)\right| \\ &\overset{(f)}{\leq} \sum_{i=1}^{n-1}\left|\bar{\mathcal{S}}\left(\boldsymbol{L}^{(n-1)}\right)\cap\mathcal{T}_i\left(\boldsymbol{L}^{(n-1)}\right)\right| \\ &= \sum_{i=1}^{n-1}\left|\bar{\mathcal{S}}_i\left(\boldsymbol{L}^{(n-1)}\right)\right|\end{aligned}$$

where (e) holds because $\mathcal{T}_i(\boldsymbol{L}^{(n)}) \subseteq \mathcal{T}_i(\boldsymbol{L}^{(n-1)})$ due to $\boldsymbol{l}_i^{(n)} = \boldsymbol{l}_i^{(n-1)}, \forall i \neq n$ in the nth round, and (f) holds because $\bar{\mathcal{S}}(\boldsymbol{L}^{(n)}) \subseteq \bar{\mathcal{S}}(\boldsymbol{L}^{(n-1)})$ since the step 3 in Algorithm 3 only modifies $\boldsymbol{l}_n(t)$ for $t \in \bar{\mathcal{S}}_n(\boldsymbol{L}^{(n-1)}) \subseteq \bar{\mathcal{S}}(\boldsymbol{L}^{(n-1)})$.

Finally, combining the results that $(\boldsymbol{P}^*, \hat{\boldsymbol{L}})$ is an optimal solution to $\mathscr{P}2$ and $|\bar{\mathcal{S}}(\hat{\boldsymbol{L}})| \leq N$, we can deduce that $F(\boldsymbol{P}^*, \hat{\boldsymbol{L}}) - F^* \leq N\delta\lambda$ according to Proposition 4.

Appendix I
Implementation Details for the Non-Orthogonal Transmission

In non-orthogonal transmission, the multiplexing is done in power domain, then the interference-constrained problem becomes

$$\begin{aligned}&\min_{\boldsymbol{P},\boldsymbol{L}} \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} p_n(t) + \lambda\sum_{t\in\mathcal{T}}\mathbb{I}\left\{\sum_{n\in\mathcal{N}} p_n(t) > 0\right\} \\ &\text{s. t. } \mathbb{E}\left\{\sum_{t\in\mathcal{T}}\log\left(1+\frac{p_n(t)\,h_n(t)}{\sum_{k\in\mathcal{N}_n^*} p_k(t)\,h_n(t)+1}\right)\right\} \geq S_n, \forall n \\ &\quad \mathbb{E}\left\{h_m(t)\sum_{n\in\mathcal{N}} p_n(t)\right\} \leq I_{\text{bs}}, \forall m, t \\ &\quad p_n(t) \geq 0, \forall n, t\end{aligned}$$

where $\mathcal{N}_n^*$ is the nodes that decode later than n.

The first challenge is estimating the lower bound of throughput due to the complex signal-to-interference-and-noise ratio (SINR) distribution, then, a fixed attenuation of 2.5 dB ($\varsigma = 0.56$) is added to ensure the expected throughput constraint. The second challenge is the non-convex and non-smooth nature of the objective function due to the indicator function, which is resolved by setting $\lambda = 0$. The third challenge involves the non-convex set constrained by throughput, handled by employing a local approximation convex method as in [38]. As a result, given the last round of transmission power $p_n^{(i)}(t)$, using the radio map, the problem becomes

$$\begin{aligned}&\min_{\boldsymbol{P},\boldsymbol{L}} \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} p_n(t) \\ &\text{s. t. } \sum_{t\in\mathcal{T}}\alpha_n(t)\log\left(1+\frac{\varsigma p_n(t)}{\sum_{k\in\mathcal{N}_n^*} p_k(t) + 1/g_n(t)}\right) \\ &\qquad + \beta_n(t) \geq S_n, \forall n \\ &\quad \sum_{n\in\mathcal{N}} p_n(t) \leq \bar{p}(t), \forall t \\ &\quad p_n(t) \geq 0, \forall n, t\end{aligned}$$

where $\alpha_n(t) = z_0(t)/(1+z_0(t))$, $\beta_n(t) = \log(1+z_0(t)) - \alpha_n(t)\log(z_0(t))$, and $z_0(t) = \varsigma p_n^{(i)}(t)/(\sum_{k\in\mathcal{N}_n^*} p_k^{(i)}(t) + 1/g_n(t))$. In order to solving this problem efficiently, a Lagrangian-based algorithm is employed, as proposed in [36], optimizing transmission power and Lagrangian parameter iteratively.

References

[1] Q. Wu et al., "A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 2912–2945, Oct. 2021.

[2] K. Yao et al., "Self-organizing slot access for neighboring cooperation in UAV swarms," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2800–2812, Apr. 2020.

[3] D. S. Lakew, U. Sa'ad, N. Dao, W. Na, and S. Cho, "Routing in flying ad hoc networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1071–1120, 2nd Quart., 2020.

[4] Z. Mou, F. Gao, J. Liu, and Q. Wu, "Resilient UAV swarm communications with graph convolutional neural network," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 393–411, Jan. 2022.

[5] 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.

[6] U. Challita, W. Saad, and C. Bettstetter, "Interference management for cellular-connected UAVs: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2125–2140, Apr. 2019.

[7] X. Zhong, Y. Guo, N. Li, Y. Chen, and S. Li, "Deployment optimization of UAV relay for malfunctioning base station: Model-free approaches," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 11971–11984, Dec. 2019.

[8] M. M. Azari, G. Geraci, A. Garcia-Rodriguez, and S. Pollin, "UAV-to-UAV communications in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6130–6144, Sep. 2020.

[9] M. Z. Hassan, G. Kaddoum, and O. Akhrif, "Interference management in cellular-connected Internet of Drones networks with drone-pairing and uplink rate-splitting multiple access," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 16060–16079, Sep. 2022.

[10] A. Rahmati et al., "Dynamic interference management for UAV-assisted wireless networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2637–2653, Apr. 2022.

[11] W. Tang, H. Zhang, Y. He, and M. Zhou, "Performance analysis of multi-antenna UAV networks with 3D interference coordination," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5145–5161, Jul. 2022.

[12] 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.

[13] B. Li and J. Chen, "Large timescale optimization for communications over aerial ad hoc networks with predetermined trajectories," *IEEE Trans. Commun.*, early access, Apr. 26, 2024, doi: 10.1109/TCOMM.2024.3393980.

[14] J. Zhang, Y. Zeng, and R. Zhang, "UAV-enabled radio access network: Multi-mode communication and trajectory design," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5269–5284, Oct. 2018.

[15] Y. Ji, Z. Yang, H. Shen, W. Xu, K. Wang, and X. Dong, "Multicell edge coverage enhancement using mobile UAV-relay," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7482–7494, Aug. 2020.

[16] L. Li, T.-H. Chang, and S. Cai, "UAV positioning and power control for two-way wireless relaying," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1008–1024, Feb. 2020.

[17] T. Ma et al., "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.

[18] B. Li and J. Chen, "Handover game for data transportation over dynamic UAV networks with predictable channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022, pp. 3724–3729.

[19] Q. Hu, Y. Cai, G. Yu, Z. Qin, M. Zhao, and G. Y. Li, "Joint offloading and trajectory design for UAV-enabled mobile edge computing systems," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1879–1892, Apr. 2019.

[20] Q. Hu, Y. Cai, A. Liu, G. Yu, and G. Y. Li, "Low-complexity joint resource allocation and trajectory design for UAV-aided relay networks with the segmented ray-tracing channel model," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6179–6195, Sep. 2020.

[21] A. Alsharoa and M. Yuksel, "Energy efficient D2D communications using multiple UAV relays," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 5337–5351, Aug. 2021.

[22] C. You and R. Zhang, "3D trajectory optimization in Rician fading for UAV-enabled data harvesting," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3192–3207, Jun. 2019.

[23] M. Samir, S. Sharafeddine, C. M. Assi, T. M. Nguyen, and A. Ghrayeb, "UAV trajectory planning for data collection from time-constrained IoT devices," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 34–46, Jan. 2020.

[24] J. Chen, U. Mitra, and D. Gesbert, "3D urban UAV relay placement: Linear complexity algorithm and analysis," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5243–5257, Aug. 2021.

[25] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, "RadioUNet: Fast radio map estimation with convolutional neural networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 4001–4015, Jun. 2021.

[26] K. Sato, K. Suto, K. Inage, K. Adachi, and T. Fujii, "Space-frequency-interpolated radio map," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 714–725, Jan. 2021.

[27] W. Liu and J. Chen, "UAV-aided radio map construction exploiting environment semantics," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6341–6355, Feb. 2023.

[28] Y. Zeng et al., "A tutorial on environment-aware communications via channel knowledge map for 6G," *IEEE Commun. Surveys Tuts.*, early access, Feb. 9, 2024, doi: 10.1109/COMST.2024.3364508.

[29] L. A. Wolsey, *Integer Programming*. Hoboken, NJ, USA: Wiley, 2020.

[30] M. Razaviyayn, "Successive convex approximation: Analysis and applications," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Minnesota, Minneapolis, MN, USA, 2014.

[31] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[32] D. Bertsekas, *Convex Optimization Algorithms*. Belmont, MA, USA: Athena Scientific, 2015.

[33] *Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects*, document TR 36.814, Release 9, 3GPP, Mar. 2017.

[34] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[35] 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.

[36] X. Zhang, H. Zhang, W. Du, K. Long, and G. K. Karagiannidis, "Joint resource allocation and reflecting design in IRS-UAV communication networks with SWIPT," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 2533–2546, Apr. 2024.

[37] J. Sebastian, C. Karakus, and S. Diggavi, "Approximate capacity of fast fading interference channels with no instantaneous CSIT," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6015–6027, Dec. 2018.

[38] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.

Bowen Li (Graduate Student Member, IEEE) received the B.Eng. degree from Harbin Engineering University, Harbin, China, in 2017, and the M.Eng. degree from the University of Chinese Academy of Sciences, China, in 2020. He is currently pursuing the Ph.D. degree with the School of Science and Engineering and the Shenzhen Future Network of Intelligence Institute (FNii–Shenzhen), The Chinese University of Hong Kong, Shenzhen (CUHK–Shenzhen), Guangdong, China. His research interests include resource allocation, interference management, and routing in aerial wireless communications.

Junting Chen (Member, IEEE) received the B.Sc. degree in electronic engineering from Nanjing University, Nanjing, China, in 2009, and the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology (HKUST), Hong Kong, SAR, China, in 2015.

From 2014 to 2015, he was a Visiting Student with the Wireless Information and Network Sciences Laboratory, MIT, Cambridge, MA, USA. He is currently an Assistant Professor with the School of Science and Engineering and Shenzhen Future Network of Intelligence Institute (FNii–Shenzhen), The Chinese University of Hong Kong, Shenzhen (CUHK–Shenzhen), Guangdong, China. Prior to joining CUHK–Shenzhen, he was a Post-Doctoral Research Associate with the Ming Hsieh Department of Electrical Engineering, University of Southern California (USC), Los Angeles, CA, USA, from 2016 to 2018, and with the Communication Systems Department of EURECOM, Sophia–Antipolis, France, from 2015 to 2016. His research interests include channel estimation, MIMO beamforming, machine learning, and optimization for wireless communications and localization. His current research interests include radio map sensing, construction, and application for wireless communications. He was a recipient of the HKTIIT Post-Graduate Excellence Scholarships in 2012. His paper received the Charles Kao Best Paper Award from WOCC 2022. He was nominated as the exemplary reviewer of *IEEE Wireless Communications Letters* in 2018.