Yi Wang, Lin Xiang, and Anja Klein, “Completion Time Minimization for UAV-Based Communications with a Finite Buffer” in *IEEE International Conference on Communications (ICC), Rome, Italy, May 2023.*

Completion Time Minimization for UAV-Based Communications with a Finite Buffer

Yi Wang, Lin Xiang, and Anja Klein
Communications Engineering Lab, Technische Universität Darmstadt, Germany
Emails: {y.wang, l.xiang, a.klein}@nt.tu-darmstadt.de

***Abstract*—This paper considers a buffer-aided unmanned aerial vehicle (UAV) serving as an aerial relay for communication between a base station (BS) and multiple ground users (GUs). Thanks to its flexible mobility, the UAV can achieve high-rate communications with the GUs/BS by buffering the communication data and exploiting the favorable channel conditions on its flight trajectory for transmission and reception. However, the size of the buffer is limited in practice, which may severely restrict the throughput gains enabled by buffering. Whether it is beneficial to consider buffering at UAVs with a small buffer is an open research problem, which is tackled in this paper. Assuming a finite buffer mounted at the UAV, we consider joint optimization of the resource allocation, data buffering, and trajectory planning for minimizing the UAV's completion time required for delivering a given data volume from each GU to the BS, where the resource allocation contains power and bandwidth allocation. The formulated optimization problem is a mixed-integer nonconvex program, which is generally intractable. To solve this problem, we propose a novel low-complexity two-layer iterative suboptimal algorithm based on bisection search and penalty successive convex approximation (PSCA). Note that minimizing the completion time in turn maximizes the average throughput, i.e., the amount of data delivered from the GUs to the BS per unit of time. Simulation results show that the buffer with sufficiently large size can increase the UAV's average throughput by up to 123.8% compared to without buffering. Moreover, with our proposed scheme, 63.2% of the throughput gains can already be achieved using only a small buffer.**

I. Introduction

Aerial relays and base stations (BSs) enabled by unmanned aerial vehicles (UAVs) provide a promising technology candidate for future communication networks. In addition to extending the communication range and/or improving the transmission power efficiency, the UAVs can be deployed fast and reconfigured flexibly to satisfy both normal and emergency communication demands. However, mobile UAVs performing two-hop relaying may suffer from limited communication capacity. For example, consider UAV-aided communication between a ground user (GU) and a ground BS. When the UAV moves closer to the GU (or the BS), the path loss of the UAV-to-BS (or the UAV-to-GU) link may increase and limit the system capacity, which is given by the minimum of the two link capacities.

Buffer-aided UAV communication has recently been proposed to mitigate the capacity bottleneck of UAVs. The idea of buffer-aided communication was originally developed in terrestrial relaying networks in [1]–[3] and has been extensively investigated since then. It has been shown in [1]–[3] that buffer-aided terrestrial relay nodes can *opportunistically exploit* favorable realizations of a fading channel to increase the communication throughput. Unlike the terrestrial relays in [1]–[3], UAVs can flexibly move in the system to *proactively* adjust the channel conditions. Therefore, a buffer-aided UAV can *simultaneously exploit* the channel fading and the UAV's mobility to achieve an even higher throughput than terrestrial relays.

So far, buffer-aided UAV communication has been considered for delay-tolerant networks (DTN) [5], [6] such as wireless sensor networks and mobile ad hoc networks, where the network nodes are usually sparsely connected. The buffer-aided UAVs can store the delay-tolerant communication data within the buffer and carry them from data sources to destinations, which significantly extends the communication range and increases the delivery probability. Buffer-aided UAVs were deployed to assist data delivery from isolated DTN nodes to their destinations in [5], where the delivery probability was investigated via simulations. In [6], buffer-aided UAVs were exploited to simultaneously lower the delivery delay and improve the reliability of message forwarding for vehicular DTN nodes with intermittent network connections.

Meanwhile, buffer-aided UAVs were also exploited to increase the throughput of cellular networks [7]–[10]. In [7], joint power allocation and trajectory optimization for maximizing the throughput of buffer-aided UAV relaying was investigated. Considering buffer-aided full-duplex UAV relaying for user pairs sharing the same spectrum, joint transmit power allocation and trajectory optimization for maximizing the energy efficiency, i.e., throughput per Joule of energy consumption, was investigated in [8]. In [9] and [10], buffer-aided UAV communication was further investigated in a mixed free-space optical (FSO)/radio frequency (RF) relaying network with FSO and RF communications employed at the source-to-relay and the relay-to-destination links, respectively. The ergodic throughput of the considered mixed FSO/RF relaying network was analyzed in [9], while the optimal UAV trajectory for throughput maximization was investigated in [10].

Note that the existing literature [5]–[9] usually assumed an infinitely large buffer to achieve the maximum throughput gains of buffering and avoid data overflow. In practice, how large buffer size is sufficient, and for UAV communications

This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center and has been supported by the BMBF project Open6GHub under grant 16KISKO14.

whether buffering at UAVs with a small buffer is beneficial remain open research problems. To tackle this problem, in this paper, we consider a UAV with a finite buffer to assist uplink communication from multiple GUs to a BS. To investigate the maximum throughput gains of the finite buffer, we jointly optimize power and bandwidth allocation, data buffering and trajectory planning for minimizing the UAV's completion time required for communicating a given volume of data from each GU to the BS. To our knowledge, completion time minimization for buffer-aided UAV communication systems has not been reported in the literature yet, even for the case of infinitely large buffer. Here, we limit our discussion to only uplink communication and half-duplex UAVs that are widely adopted in practical systems; but the idea and proposed solution are equally applicable for full-duplex UAV communications. Our contributions are:

- We consider a UAV with only a finite buffer to assist uplink communication from multiple GUs to a BS. We jointly optimize the power and bandwidth allocation, dynamic buffering, and trajectory planning of the UAV for minimizing the completion time needed to deliver all the data from the GUs to the BS.
- The formulated optimization problem is a nonconvex mixed-integer nonlinear program, which is generally intractable. We propose a novel low-complexity two-layer iterative suboptimal algorithm based on bisection search and penalty successive convex approximation (PSCA), which requires no prior knowledge of a feasible resource allocation as the starting point of iteration.
- Simulation results show that the buffer-aided UAV communications can achieve a good average throughput gain even with a small buffer by jointly optimizing resource allocation, data buffering, and trajectory of the UAV. This reveals the great potential of buffer-aided UAV communications in practical systems with a finite or even limited buffer size.

In the remainder of this paper, Section II presents the system model of buffer-aided UAV communications. The formulation and the solution of the completion time minimization problem are provided in Sections III and IV, respectively. Section V evaluates the performance of the proposed scheme. Finally, Section VI concludes the paper.

Notation: Throughout this paper, $\mathbb{N}$ and $\mathbb{C}$ denote the sets of natural and complex numbers, respectively. $\|\cdot\|$ and $|\cdot|$ denote the Euclidean distance and the absolute value, respectively. $\mathcal{CN}\left(0, \sigma^2\right)$ indicates a complex-valued zero-mean Gaussian random variable with variance σ^2.

II. System Model

In this section, we present the system model including the channel model and buffer status of the considered buffer-aided UAV communication.

A. Buffer-aided UAV Communications

As depicted in Figure 1, we consider a single-antenna UAV with a half-duplex wireless transceiver and a finite buffer,

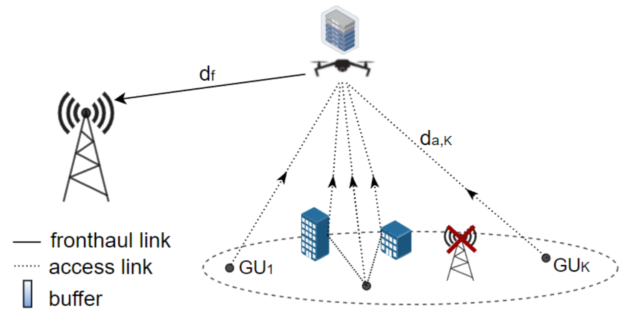


Fig. 1. System diagram of buffer-aided UAV communications.

which is deployed to provide uplink communication for K single-antenna GUs that are out of cellular coverage due to, e.g., malfunction of their local serving BS or roaming to remote areas. The GU k is located at position $\boldsymbol{u}_k^{\mathrm{GU}}, k \in \mathcal{K} \triangleq \{1, ..., K\}$. Each GU k has a certain amount of data D_k that needs to be communicated to the BS. The UAV aims to recover the uplink communication for the GUs by serving as an aerial relay between the GUs and a remote BS located at position $\boldsymbol{u}^{\mathrm{BS}}$. The UAV may hover at or fly toward any of the positions between the GUs and the BS. A limited buffer enables the UAV to store data received from the GUs before forwarding it to the remote BS. The transceiving strategy depends on the UAV's positions and channel conditions.

For the considered uplink communication, the terms fronthaul and access links refer to the UAV-to-BS and the GU-to-UAV links, respectively. To facilitate trajectory planning, a discrete time system is applied, and each time slot has equal duration $\Delta > 0$. Let i be the index of time slot. We assume that Δ is sufficiently small, i.e., shorter than the channel coherence time of fronthaul and access links [11], such that the position and channel conditions of the UAV are approximated as constants in each time slot. Moreover, the system has a total bandwidth of B, which is divided into small non-overlapping frequency sub-channels via frequency-division multiple access (FDMA). Sub-channels with bandwidths $b_{\mathrm{a},k}[i]$ and $b_{\mathrm{f},k}[i]$ are allocated to GU k for access and fronthaul communication, respectively. FDMA is considered as a continuous approximation of the orthogonal frequency division multiple access (OFDMA) that has a sufficiently large number of subcarriers.

B. Channel Models and Achievable Data Rates

Let $\boldsymbol{u}[i]$ be the position of the UAV in time slot i. We consider line-of-sight (LoS)-dominant channels for both fronthaul and access links. The channel gains of the fronthaul and the access links for communicating with GU k in time slot i, denoted as $H_{\mathrm{f},k}[i]$ and $H_{\mathrm{a},k}[i]$, are modeled as

$$H_{m,k}[i] = \sqrt{A_0} \cdot h_{m,k}[i] \cdot d_{m,k}^{-\alpha/2}[i],\ m \in \{\mathrm{f}, \mathrm{a}\}, \tag{1}$$

where $d_{\mathrm{f},k}[i] \equiv d_{\mathrm{f}}[i] \triangleq \left\|\boldsymbol{u}[i] - \boldsymbol{u}^{\mathrm{BS}}\right\|$ and $d_{\mathrm{a},k}[i] \triangleq \left\|\boldsymbol{u}[i] - \boldsymbol{u}_k^{\mathrm{GU}}\right\|$ are the distances of the fronthaul and access links for communicating with GU k, respectively. α is the path loss exponent with $\alpha \geq 2$. Moreover, $A_0 d_{m,k}^{-\alpha}$ and $h_{m,k}[i] \in \mathbb{C}$, $m \in \{\mathrm{f}, \mathrm{a}\}$ denote the path losses and the channel fading over the fronthaul/access links in time slot i, respectively. For the considered LoS-dominant channels, $h_{m,k}[i]$, $m \in \{\mathrm{f}, \mathrm{a}\}$ are usually modeled as [12]

$$h_{m,k}[i] = \sqrt{\frac{\kappa_m}{\kappa_m+1}}\sigma_{m,k} e^{j\theta_{m,k}[i]} + \sqrt{\frac{1}{\kappa_m+1}}\mathcal{CN}\left(0, \sigma_{m,k}^2\right), \quad (2)$$

where $h_{\mathrm{f},k} \equiv h_{\mathrm{f}}$ is assumed for any k. In (2), the first and the second terms capture the LoS and the non-LoS (NLoS) components of propagation paths, respectively, where the phases $\theta_{m,k}[i]$ are uniformly distributed in $[0, 2\pi]$ and $\theta_{m,k}[i]$ of different paths are independent. Thus, $|h_{m,k}[i]|^2$ follows the Rician distribution with factor $\kappa_m \geq 0$, where κ_m is the energy ratio between the LoS and the NLoS components. A channel with dominating LoS component has a large κ_m [12].

The achievable data rates over the fronthaul and access links, denoted by $r_{m,k}[i]$, $m \in \{\mathrm{f}, \mathrm{a}\}$, are calculated based on Shannon's channel capacity formula [12],

$$r_{m,k}[i] = b_{m,k}[i] \cdot \log_2\left(1 + \frac{|H_{m,k}[i]|^2 \cdot p_{m,k}[i]}{\sigma_{\mathrm{n}}^2 \cdot b_{m,k}[i]}\right), \quad (3)$$

where $p_{m,k}[i]$, $m \in \{\mathrm{f}, \mathrm{a}\}$ denotes the transmit power. $p_{\mathrm{a},k}[i]$ is the transmit power of GU k in time slot i, and $p_{\mathrm{f},k}[i]$ is the transmit power that the UAV allocates to GU k in the fronthaul link in time slot i. σ_{n}^2 is the power spectral density of the channel noise.

C. Buffer Status Evolution

Let $Q_{\max}$ be the size of the buffer at the UAV. We assume that the total buffer size $Q_{\max}$ is dynamically allocated to the GUs based on their channel conditions and the UAV's position in each time slot. Let $Q_k[i]$ be the amount of data buffered at the UAV, also referred to as the queue length, for GU k in time slot i. We have

$$Q_k[i] = \max\left\{Q_k[i-1] - r_{\mathrm{f},k}[i] \cdot \Delta, 0\right\} + r_{\mathrm{a},k}[i] \cdot \Delta, \quad (4)$$

where $\sum_{k=1}^{K} Q_k[i] \leq Q_{\max}$ should hold to avoid buffer overflow. Moreover, $\max\{\cdot, 0\}$ prevents the buffer from underflow. We assume that the data received in time slot i can only be forwarded to the BS in the next time slot. Once the data is transmitted to the BS in time slot i, the corresponding buffer space can be immediately freed up.

Note that the achievable data rates of the fronthaul and access links, $r_{m,k}[i]$, $m \in \{\mathrm{f}, \mathrm{a}\}$, may significantly differ from each other, due to different path losses and independent channel fadings. If $r_{\mathrm{f},k}[i] < r_{\mathrm{a},k}[i]$ and $\sum_{k=1}^{K} Q_k[i] \leq Q_{\max}$, the buffer at the UAV can be used to store the data received from the GUs; otherwise, pre-stored data from GUs will be forwarded to the BS over the fronthaul link. Consequently, buffering facilitates the exploitation of the links with strong channel conditions for increasing the communication throughput.

III. Problem Formulation

To maximize the average throughput of buffer-aided UAV communications in the small/finite buffer regime, joint resource (frequency and power) allocation, data buffering, and trajectory optimization is investigated in this section. The aim is to minimize the UAV's completion time, i.e., the time required for completing data delivery for each GU. This problem in turn maximizes the average throughput, namely the amount of data delivered from the GUs to the BS per time slot, within the optimal completion time [14, Theorem 1].

Assume that the channel conditions of the fronthaul and access links are known. Given the UAV's starting point $\boldsymbol{u}[0]$, the resulting optimization problem is formulated as

$$\begin{aligned}
&\mathcal{P}_1: \min_{\boldsymbol{u}[i], \boldsymbol{p}_k[i], \boldsymbol{b}_k[i]} T_{\mathrm{MT}} && (5)\\
&\text{s.t. } \mathrm{C}_1: \|\boldsymbol{u}[i] - \boldsymbol{u}[i-1]\|/\Delta \leq V_{\max}, \quad i \in \mathcal{I}\\
&\mathrm{C}_2: \sum\nolimits_{k=1}^{K} p_{\mathrm{f},k}[i] \leq P_{\mathrm{u},\max}, \quad i \in \mathcal{I}\\
&\mathrm{C}_3: p_{\mathrm{a},k}[i] \leq P_{k,\max}, \quad i \in \mathcal{I}, k \in \mathcal{K}\\
&\mathrm{C}_4: \sum\nolimits_{k=1}^{K} \left(b_{\mathrm{a},k}[i] + b_{\mathrm{f},k}[i]\right) \leq B, \quad i \in \mathcal{I}\\
&\mathrm{C}_5: r_{m,k}[i] \geq b_{m,k}[i] \cdot E_m^{\min},\ m \in \{\mathrm{f}, \mathrm{a}\}, i \in \mathcal{I}, k \in \mathcal{K}\\
&\mathrm{C}_6: \sum\nolimits_{i=1}^{j} \left(r_{\mathrm{f},k}[i] - r_{\mathrm{a},k}[i-1]\right)\Delta \leq 0,\ j \in \mathcal{I},\ k \in \mathcal{K}\\
&\mathrm{C}_7: \sum_{k=1}^{K}\sum_{i=1}^{j} \left(r_{\mathrm{a},k}[i-1] - r_{\mathrm{f},k}[i]\right)\Delta \leq Q_{\max},\ j \in \mathcal{I}\\
&\mathrm{C}_8: \sum\nolimits_{i=1}^{T_{\mathrm{MT}}} r_{\mathrm{f},k}[i]\Delta \geq D_k, \quad k \in \mathcal{K},
\end{aligned}$$

where the objective function of $\mathcal{P}_1$ is the UAV's completion time $T_{\mathrm{MT}} \in \mathbb{N}$. The optimization space includes the trajectory $\boldsymbol{u}[i]$ of the UAV, the power allocation $\boldsymbol{p}_k[i] \triangleq (p_{\mathrm{a},k}[i], p_{\mathrm{f},k}[i])$, and the bandwidth allocation $\boldsymbol{b}_k[i] \triangleq (b_{\mathrm{a},k}[i], b_{\mathrm{f},k}[i])$ for $i \in \mathcal{I} \triangleq \{1, ..., T_{\mathrm{MT}}\}$. Constraint C_1 limits the maximum flight speed of the UAV by $V_{\max}$. C_2 and C_3 guarantee that the total power consumed at the UAV and at GU k for communication does not exceed $P_{\mathrm{u},\max}$ and $P_{k,\max}$, respectively. C_4 constrains the maximum total bandwidth allocated for the access and fronthaul links by B. C_5 ensures a minimum signal-to-noise ratio (SNR) of $2^{E_m^{\min}} - 1$ for each active GU with $b_{m,k}[i] > 0$ in time slot i; otherwise, for inactive GUs C_5 can be ignored. C_6 guarantees causality for buffer-aided data transmission, i.e., for any $j \in \mathcal{I}$, the total throughput of the fronthaul link within time period $[1, j]$ should not exceed that of the access link to avoid buffer underflow. C_7 limits the maximum use of buffer space to avoid buffer overflow. Implicit in C_7, dynamic buffering is guaranteed since the fractions of the buffer size allocated to the different GUs are flexibly adjusted. Finally, C_8 requires that the data D_k of GU k should be completely delivered to the BS until time slot T_{MT}.

$\mathcal{P}_1$ is a mixed-integer nonconvex problem since the objective function and constraint C_8 involve an integer variable T_{MT}, and constraints C_5, C_6, C_7 and C_8 involve the nonconvex function $r_{m,k}[i]$, $m \in \{\mathrm{f}, \mathrm{a}\}$. Moreover, as $\mathcal{I}$ is a function of T_{MT}, minimization of mission completion time T_{MT} is coupled with optimization of the UAV's resource allocation and trajectory $\{\mathbf{u}[i], \boldsymbol{p}_k[i], \boldsymbol{b}_k[i]\}$ [14]. Such type of problem is generally difficult to be optimally solved within a polynomial computation time. In Section IV, we propose a low-complexity two-layer iterative suboptimal algorithm based on bisection search and PSCA to solve $\mathcal{P}_1$.

IV. Problem Solution

In this section, the nonconvex constraints of problem $\mathcal{P}_1$ are first transformed into equivalent differences of convex (DC) forms, and tackled using PSCA. Then the two-layer suboptimal iterative algorithm is proposed to solve the mixed-integer nonconvex problem by further decoupling the optimization of T_{MT} and $\{\mathbf{u}[i], \boldsymbol{p}_k[i], \boldsymbol{b}_k[i]\}$.

A. Problem Transformation

It is convenient to directly optimize $r_{m,k}[i]$, $m \in \{\mathrm{f}, \mathrm{a}\}$, since constraints C_5, C_6, C_7, and C_8 are convex with respect to $r_{m,k}[i]$. According to (3), we replace $p_{m,k}[i]$ with $r_{m,k}[i]$ in C_2 and C_3 using

$$p_{m,k}[i] = C_{m,k}[i] d_{m,k}[i] g\left(b_{m,k}[i], r_{m,k}[i]\right), m \in \{\mathrm{f}, \mathrm{a}\}, \quad (6)$$

where we define $C_{m,k}[i] \triangleq \frac{\sigma_{\mathrm{n}}^2}{A_0 |h_{m,k}[i]|^2}$ and $g\left(b_{m,k}[i], r_{m,k}[i]\right) \triangleq b_{m,k}[i]\left(\exp\left(\frac{r_{m,k}[i] \ln 2}{b_{m,k}[i]}\right) - 1\right)$. Note that $g\left(b_{m,k}[i], r_{m,k}[i]\right)$, $m \in \{\mathrm{f}, \mathrm{a}\}$ are perspective functions of the exponential function, for which they are jointly convex with respect to $b_{m,k}[i]$ and $r_{m,k}[i]$ [13].

Using (6) and introducing slack variables $\tau_i \geq 0$ and $\tau_{k,i} \geq 0$, constraints C_2 and C_3 can be rewritten as

$$\mathrm{C}_{2a} : \sum_{k=1}^{K} g\left(b_{\mathrm{f},k}[i], r_{\mathrm{f},k}[i]\right) \leq \tau_i, \quad \forall i \quad (7)$$

$$\mathrm{C}_{2b} : C_{\mathrm{f},k}[i] d_{\mathrm{f},k}[i] \leq P_{\mathrm{u,max}} / \tau_i, \quad \forall i \quad (8)$$

$$\mathrm{C}_{3a} : g\left(b_{\mathrm{a},k}[i], r_{\mathrm{a},k}[i]\right) \leq \tau_{k,i}, \quad \forall k, i \quad (9)$$

$$\mathrm{C}_{3b} : C_{\mathrm{a},k}[i] d_{\mathrm{a},k}[i] \leq P_{k,\max} / \tau_{k,i}, \quad \forall k, i. \quad (10)$$

Particularly, C_2 will be replaced by C_{2a} and C_{2b}, and C_3 will be replaced by C_{3a} and C_{3b}. Note that C_{2a} and C_{3a} are convex constraints. However, C_{2b} and C_{3b} are nonconvex DC constraints, which can be further tackled below using the PSCA method.

B. Proposed Two-Layer Iterative Solution

Problem $\mathcal{P}_1$ is still difficult to solve due to the nonconvex constraints C_{2b} and C_{3b}, and the integer variable T_{MT}, which is coupled with the other optimization variables. Here, we propose to solve $\mathcal{P}_1$ in two layers iteratively. In each iteration, an outer layer searches over time to find the minimum mission completion time T_{MT}^*. To this end, a bisection search is employed because the objective function $T_{\mathrm{MT}} \in \mathbb{N}$ can be regarded as a quasi-convex function of the absolute continuous time [14].

For each T_{MT} specified by the outer layer, an inner layer then checks whether a feasible resource allocation, buffering, and trajectory planning exist via solving the following problem

$$\mathcal{P}_2 : \text{find} \left(\mathbf{u}[i], \boldsymbol{r}_k[i], \boldsymbol{b}_k[i]\right), i \in \mathcal{I}, k \in \mathcal{K} \quad (11)$$
$$\text{s.t. } \mathrm{C}_1, \mathrm{C}_{2a-b}, \mathrm{C}_{3a-b}, \mathrm{C}_4 - \mathrm{C}_8,$$

where $\boldsymbol{r}_k[i] \triangleq \left(r_{\mathrm{a},k}[i], r_{\mathrm{f},k}[i]\right)$ is the rate allocation vector. The iteration continues until the outer-layer search converges.

Algorithm 1: Bisection Search in the Outer Layer

Input: T_{ini}, $P_{\mathrm{u,max}}$, B, $E_m^{\min}$, $\{P_{k,\max}, D_k\}_{k=1}^K$, $Q_{\max}$, $V_{\max}$
Output: T_{MT}^*

initialize $T_{\mathrm{fea}} = T_{\mathrm{ini}}$, $\quad T_{\mathrm{infea}} = 0$
while $T_{\mathrm{fea}} - T_{\mathrm{infea}} > 1$ **do**
  $T_{\mathrm{MT}} = \lceil \frac{T_{\mathrm{fea}} - T_{\mathrm{infea}}}{2} \rceil$
  if $\mathcal{P}_2$ *is feasible* **then**
    $T_{\mathrm{fea}} = T_{\mathrm{MT}}$
  else
    $T_{\mathrm{infea}} = T_{\mathrm{MT}}$
  end
end
$T_{\mathrm{MT}}^* = T_{\mathrm{fea}}$

A similar two-layer algorithm has been proposed in our prior work [14]; but unlike [14], the inner-layer problem $\mathcal{P}_2$ involves nonconvex constraints C_{2b} and C_{3b}, and is nontrivial to solve.

The procedure of outer-layer bisection search is summarized in Algorithm 1. In Algorithm 1, T_{ini} is the starting point of T_{MT}, and it is a large value which makes $\mathcal{P}_2$ feasible. After each search, the feasible value T_{fea} and infeasible value T_{infea} of T_{MT} will be updated according to the feasibility of $\mathcal{P}_2$. First, for solving $\mathcal{P}_2$, successive convex approximation (SCA) is adopted to tackle DC constraints C_{2b} and C_{3b} [15]. Unfortunately, SCA cannot be applied for solving a nonconvex feasibility-check problem such as $\mathcal{P}_2$. This is because SCA requires a feasible initial point $(\tau_i^{(0)}, \tau_{k,i}^{(0)})$, which is unavailable without solving $\mathcal{P}_2$. We tackle this issue by adopting the PSCA method [16], which eliminates the requirement of feasible initial points.

With SCA, C_{2b} and C_{3b} are approximated iteratively, with iteration indexed by q, as

$$C_{\mathrm{f},k}[i] d_{\mathrm{f},k}[i] - P_{\mathrm{u,max}} f(\tau_i; \tau_i^{(q)}) \leq 0, \forall i \quad (12)$$

$$C_{\mathrm{a},k}[i] d_{\mathrm{a},k}[i] - P_{k,\max} f(\tau_{k,i}; \tau_{k,i}^{(q)}) \leq 0, \forall k, i, \quad (13)$$

where $f(\tau, \tau^{(q)})$ is an affine (convex) underestimation of the convex function $1/\tau$ obtained via, e.g., Taylor expansion. An example of $f(\tau, \tau^{(q)})$ is given as

$$f(\tau, \tau^{(q)}) = \frac{2}{\tau^{(q)}} - \frac{\tau}{(\tau^{(q)})^2}, \quad (14)$$

where $f(\tau, \tau^{(q)}) \leq 1/\tau, \forall \tau^{(q)}$.

We start with a random point $(\tau_i^{(0)}, \tau_{k,i}^{(0)})$, which may be infeasible for $\mathcal{P}_2$. Instead of directly solving $\mathcal{P}_2$, the PSCA first relaxes (12) and (13) by adding nonnegative slack variables μ_i, $\mu_{k,i} \geq 0$,

$$\mathrm{C}_{2b}^{(q)} : C_{\mathrm{f},k}[i] d_{\mathrm{f},k}[i] - P_{\mathrm{u,max}} f(\tau_i; \tau_i^{(q)}) \leq \mu_i, \ \forall i \quad (15)$$

$$\mathrm{C}_{3b}^{(q)} : C_{\mathrm{a},k}[i] d_{\mathrm{a},k}[i] - P_{k,\max} f(\tau_{k,i}; \tau_{k,i}^{(q)}) \leq \mu_{k,i}, \forall k, i, \quad (16)$$

where the values of μ_i and $\mu_{k,i}$ represent the violations of constraints (12) and (13). By minimizing the sum of constraint violations μ_i, $\mu_{k,i}$ iteratively, the resulting optimal solution in

Algorithm 2: PSCA based Feasibility Check in the Inner Layer

Input: T_{MT}, $P_{\text{u,max}}$, B, $E_m^{\min}$, $V_{\max}$, $\gamma_{\max}$, $\eta > 1$, ξ, $\xi_{\text{violation}}$, $Q_{\max}$, $\{P_{k,\max}, D_k\}_{k=1}^{K}$

Output: the feasibility of $\mathcal{P}_2$ for given T_{MT}

1 initialize $q = 0$, $\{\tau_i^{(0)}, \tau_{k,i}^{(0)}, \gamma^{(0)}\} > 0$, $\forall i, k$,

2 **repeat**

3 Solve $\mathcal{P}_3^{(q)}$.

4 Update: $\tau_k^{(q+1)} = \tau_k^{(q)}$, $\tau_{k,i}^{(q+1)} = \tau_{k,i}^{(q)}$, and $\gamma^{(q+1)} = \min(\eta\gamma^{(q)}, \gamma_{\max})$.

5 Update iteration: $q = q + 1$.

6 **until** *the stopping criteria are satisfied*;

each iteration leads to "smaller" violations of constraints than the previous iteration. In order to search for a feasible point, the PSCA minimizes the ℓ_1-norm of these constraint violations by solving the following optimization problem

$$\mathcal{P}_3^{(q)}: \min_{\boldsymbol{u}[i], \boldsymbol{r}_k[i], \boldsymbol{b}_k[i], \boldsymbol{s}_i^{(q)}} \gamma^{(q)} \sum_{i=1}^{T_{\text{MT}}} \left(\mu_i^{(q)} + \sum_{k=1}^{K} \mu_{k,i}^{(q)} \right) \quad (17)$$

$$\text{s.t. } \mathrm{C}_1, \mathrm{C}_{2a}, \mathrm{C}_{2b}^{(q)}, \mathrm{C}_{3a}, \mathrm{C}_{3b}^{(q)}, \mathrm{C}_4 - \mathrm{C}_8, \mu_i \geq 0, \mu_{k,i} \geq 0,$$

where $\gamma^{(q)} > 0$ is a penalty factor to limit violations of constraints (12) and (13), and $\boldsymbol{s}_i^{(q)} \triangleq ([\tau_i^{(q)}, \tau_{k,i}^{(q)}, \mu_i^{(q)}, \mu_{k,i}^{(q)}])$ is the vector of slack variables. Note that $\mathcal{P}_3^{(q)}$ is a convex problem and can be solved using off-the-shelf solvers such as CVX.

Algorithm 2 summarizes the inner-layer feasibility check procedure. We start with a small $\gamma^{(q)}$ and possibly infeasible initial point. The constraint violations are further reduced by increasing $\gamma^{(q)}$. However, whenever the sum of penalty terms of $\mathcal{P}_3^{(q)}$ becomes zero, i.e., no constraint violation in (12) and (13), the solution of $\mathcal{P}_3^{(q)}$ is a feasible solution of $\mathcal{P}_2$. We assume that the relaxed constraints well approximates the original constraints without violations, when the sum of the violations is small enough, i.e., $\sum_{i=1}^{T_{\text{MT}}} \left(\mu_i^{(q)} + \sum_{k=1}^{K} \mu_{k,i}^{(q)} \right) \leq \xi_{\text{violation}}$, where $\xi_{\text{violation}} \geq 0$ is a predefined threshold for the maximum sum of allowable violations. In this case, the optimal solution in iteration q is approximately feasible for problem $\mathcal{P}_2$. Algorithm 2 stops when one of the following criteria is satisfied:

1. The sum of the violations is below the threshold $\xi_{\text{violation}}$, which implies that the given T_{MT} is feasible for the inner layer problem.
2. $q \geq q_{\max}$ and the sum of the violations is greater than $\xi_{\text{violation}}$, implying that the given T_{MT} is infeasible for the inner layer problem.
3. $q < q_{\max}$, but the improvement of the sum of the violations between two iterations is smaller than ξ, implying that the given T_{MT} is infeasible for the inner layer problem.

With Algorithm 2, the sum of constraint violations μ_i and $\mu_{k,i}$ will successively decrease until the algorithm converges. Here, an upper limit $\gamma_{\max}$ on $\gamma^{(q)}$ is set to avoid potential numerical issues caused by extremely large $\gamma^{(q)}$ [16].

Recall that the outer-layer problem is quasi-convex. This implies that Algorithm 1 is close-to-optimal if Algorithm 2 finds a feasible point for most or all feasible inner-layer problems. However, the latter may become difficult when $T_{\text{fea}} - T_{\text{infea}}$ is small. To overcome this issue, we can, e.g., employ multiple (possibly infeasible) initial points $(\tau_i^{(0)}, \tau_{k,i}^{(0)})$ and/or sample these initial points through some intelligent heuristics. As a feasible point of problem $\mathcal{P}_2$ may not be obtained within the scope of search even if $\mathcal{P}_2$ is feasible, the overall algorithm is generally suboptimal.

TABLE I
PARAMETER SETTINGS [14], [15]

Parameter	Value	Parameter	Value
B	1 MHz	Δ	40 ms
D	0.5 Mbits	$V_{\max}$	70 m/s
A_0	-30 dB	α	2
$P_{\{k=1,2\},\max}$	23 dBm	$E_m^{\min}$	0 dB
$P_{\text{u,max}}$	30 dBm	σ^2	-90 dBm
$\xi_{\text{violation}}$	1	ξ	0.8%
$[\gamma^{(0)}, \gamma_{\max}]$	[5, 1000]	η	2

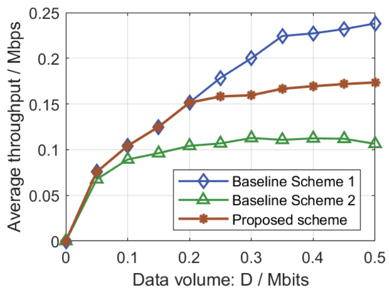
V. Simulation Results

In this section, the performance of the proposed scheme for optimizing the finite buffer-aided UAV communication is evaluated through simulations. We consider UAV-aided uplink communication from $K = 2$ GUs located at $\boldsymbol{u}_1^{\text{GU}} = (90, 90, 0)$ m and $\boldsymbol{u}_2^{\text{GU}} = (90, 50, 0)$ m to the BS located at $\boldsymbol{u}^{\text{BS}} = (-10, 0, 0)$ m. Each GU is assumed to have $D_1 = D_2 = D$ Mbits of data for delivery to the BS and the total available buffer space of the UAV is $Q_{\max}$ Mbits. Unless otherwise specified, the simulation parameters are set according to Table I. For performance comparison, we consider two baseline schemes:

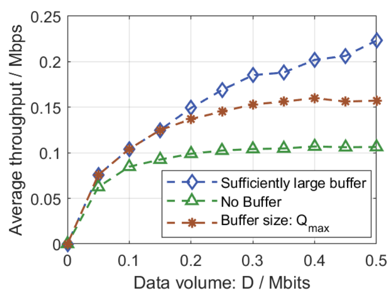
- *Baseline Scheme 1:* Unlike the proposed scheme, the UAV employs a sufficiently large buffer with $Q_{\max} \geq KD$.
- *Baseline Scheme 2:* The UAV has no buffer such that the data rate of each GU is given by the minimum capacity of the access and fronthaul links in each time slot.

For Baseline Schemes 1 and 2, the resource allocation and trajectory are further optimized using Algorithms 1 and 2. Therefore, Baseline Schemes 1 and 2 define a performance upper and lower bound on the proposed scheme, respectively.

Figure 2(a) shows the average throughput per GU for the considered schemes as a function of the data volume D, where $Q_{\max} = 0.2$ Mbits is assumed for the proposed scheme. From Figure 2(a), we observe that the average throughput of all considered schemes increase with D. This is because at the starting point $\boldsymbol{u}_0$ the UAV is far away from the GUs, which leads to a low data rate in the access link. As D increases, so does the completion time required for data delivery. Consequently, the UAV will fly closer to the GUs



(a)



(b)

Fig. 2. Average throughput per GU versus data volume D with (a) dynamic buffering and optimized bandwidth allocation, and (b) fixed buffer and bandwidth allocation

to achieve a higher data rate over the access links and hence, a higher average throughput. When a buffer is available, as D increases, the UAV can buffer more data and also moves closer to the BS for delivering the data at higher data rates over the fronthaul links. Thus, it achieves an even higher average throughput. For example, with no buffer at the UAV, the average throughput of Baseline Scheme 2 increases only slightly with D in the small data regime and saturates in the large data regime. The latter is because without a buffer, the UAV tends to hover at an optimal position when D is large. However, compared with Baseline Scheme 2, the average throughput of Baseline Scheme 1 increases by up to 123.8% when $D = 0.5$ Mbits. Moreover, despite a small buffer at the UAV, the average throughput of the proposed scheme increases by 63.2% compared to Baseline Scheme 2, because the UAV can move between the GUs and the BS to fetch and deliver data. Hence, the buffer can be reused in several rounds of flight and delivery. Furthermore, Figure 2(b) shows the average throughput per GU with *fixed* buffer allocation $Q_{\max}/2$ for each GU and *fixed* bandwidth $B/2$ for the fronthaul and access links. When comparing Figure 2(a) and Figure 2(b), we observe that optimizing buffering and the bandwidth allocation of the fronthaul and access links can improve the performance compared to fixing it, even for a small number of GUs. Note that, while we have only shown the results for two users, similar findings also hold for multi-user scenarios. In the later case, the proposed scheme can achieve additional throughput gains due to the resulting multi-user diversity.

Figure 3 shows the completion time T_{MT} of the proposed scheme as a function of the buffer size $Q_{\max}$. We observe that T_{MT} monotonically decreases with $Q_{\max}$ due to increased

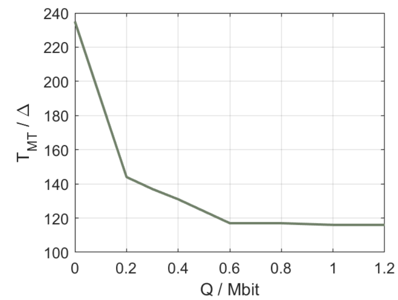
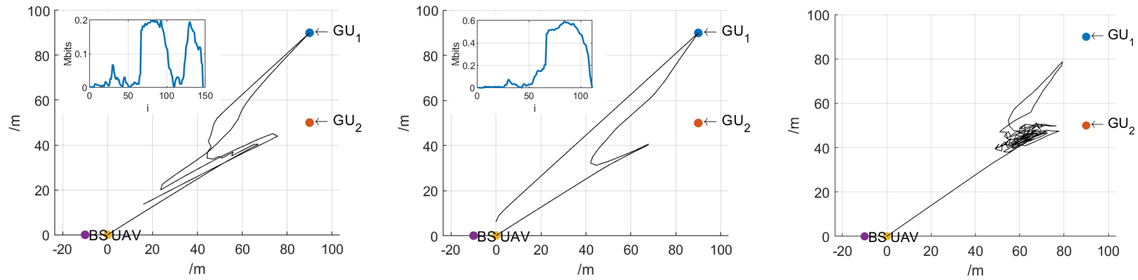


Fig. 3. Mission completion time versus buffer size

degrees of freedom for resource allocation. Moreover, a diminishing rate of decrease in the completion time is observed as $Q_{\max}$ increases. For example, the completion time reduces by 38.7% when $Q_{\max}$ increases from 0 Mbits (no buffer) to 0.2 Mbits. In contrast, T_{MT} only reduces by 18.8% when $Q_{\max}$ increases from 0.2 Mbits to 1 Mbits, and saturates when $Q_{\max}$ is sufficiently large. These results show that a small buffer can significantly reduce the completion time of the UAV.

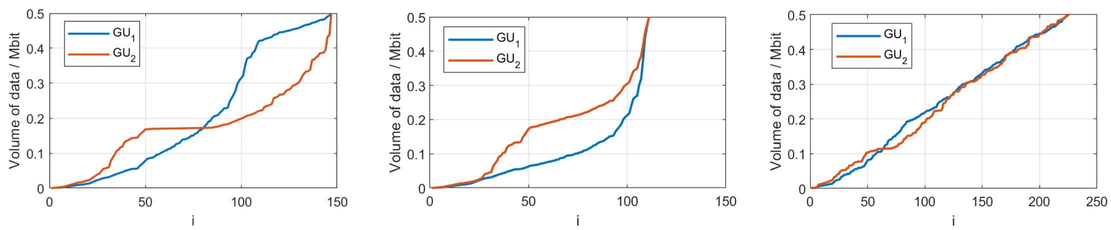
For more insights into the performance of our proposed scheme, Figure 4 shows the UAV's trajectory (projected onto the horizon) and buffer status for the proposed scheme, Baseline Scheme 1, and Baseline Scheme 2, where the UAV starts flight from the origin. From Figure 4(a), we observe that when the buffer size is limited, the UAV frequently flies back and forth between the BS and GUs to transmit the data to the BS without buffer overflow. Consequently, several peaks can be observed in the buffer status, where the UAV attempts to send almost all buffered data to the BS such that the buffer can be promptly reused for the next round of delivery. From Figure 4(b), we observe that when the buffer size is sufficiently large, the UAV collects and stores more data from the GUs than the proposed scheme before transmitting them to the BS. In this case, the UAV only has to fly a shorter total distance in order to lower the completion time. We also observe that the buffer size $Q_{\max} = 0.6$ Mbits is sufficiently large, although it is smaller than the total amount of data at the GUs. This is because in order to minimize the completion time, the buffered data will be delivered to the BS, whenever possible. Unlike Figure 4(a) and 4(b), we observe from Figure 4(c) that, when there is no buffer, the UAV stays in a small region most of the time and hence, cannot fully exploit its mobility for improving the average throughput.

Finally, Figure 5 shows the volume of data received at the BS over time for the proposed scheme, Baseline Scheme 1, and Baseline Scheme 2. From Figure 5, we observe that the UAV completes the data delivery for both GUs at the same time, in order to minimize the mission completion time. As shown in Figure 5(b), when the buffer size is sufficiently large, the UAV gathers data simultaneously from both GUs and delivers it to the BS; then, the UAV flies back towards and transmits data to the BS before the end of the data delivery. Therefore, an exponential increase in the volume of delivered data appears in time period $[80, 110]$. This result is consistent with the trajectory in Figure 4(b). Unlike Figure 5(b), we can see



(a) (b) (c)

Fig. 4. Trajectory and buffer status of the UAV for (a) the proposed scheme with $Q_{\max} = 0.2$ Mbits, (b) Baseline Scheme 1, and (c) Baseline Scheme 2.



(a) (b) (c)

Fig. 5. Volume of data received at the BS from GU_1 and GU_2 over time with $D = 0.5$ Mbits for (a) the proposed scheme with $Q_{\max} = 0.2$ Mbits, (b) Baseline Scheme 1, and (c) Baseline Scheme 2.

that the UAV transmits data from different GUs to the BS at different times in Figure 5(a). This is because, in the case of a small buffer size, the UAV will exploit its mobility by flexibly adjusting its trajectory and resource allocation according to the geographical distribution of the GUs for minimizing the completion time. The trajectory of the UAV in Figure 4(a) also proves this. On the other hand, Figure 4(c) shows that, where there is no buffer, the data transmission rate stays constant for both GUs, where the UAV only moves in a small area.

VI. Conclusion

In this paper, we investigated the maximum average throughput gains of a finite buffer-aided UAV communication system which assists uplink communication from multiple GUs to a BS. To this end, joint optimization of power and bandwidth allocation, data buffering, and trajectory planning was considered for minimizing the completion time required for delivering a given volume of data from each GU. This in turn maximizes the average throughput achievable in the shortest completion time. To solve the formulated mixed-integer nonconvex problem, we proposed a low-complexity two-layer iterative suboptimal algorithm based on bisection search and PSCA, which is close-to-optimal under certain conditions. Simulation results showed that buffering with sufficiently large size can increase the UAV's average throughput by up to 123.8% compared to the case without buffering. Moreover, our proposed scheme using only a small buffer can intelligently exploit the UAV's mobility to achieve a throughput gain of 63.2% compared to without buffering.

References

[1] N. Zlatanov, R. Schober and P. Popovski, "Buffer-Aided Relaying with Adaptive Link Selection," *IEEE J. Sel. Area. Comm.*, vol. 31, no. 8, pp. 1530-1542, Aug. 2013.

[2] N. Nomikos et al., "A Survey on Buffer-Aided Relay Selection," *IEEE Commun Surv Tut*, vol. 18, no. 2, pp. 1073-1097, 2016.

[3] Z. Tian, Y. Gong, et al., "Buffer-Aided Relay Selection with Reduced Packet Delay in Cooperative Networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2567-2575, Mar. 2017.

[4] B. Li, S. Zhao, et al., "Joint Transmit Power and Trajectory Optimization for Two-Way Multi-Hop UAV Relaying Networks," in *Proc. IEEE ICC Workshops*, pp. 1-5, 2020.

[5] C. Giannini, A. A. Shaaban, et al., "Delay Tolerant Networking for Smart City through Drones," in *Proc. IEEE ISWCS*, pp. 603-607, 2016.

[6] Z. Du, C.Wu, et al., "A Routing Protocol for UAV-Assisted Vehicular Delay Tolerant Networks," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 85-98, 2021.

[7] Y. Zeng, R. Zhang and T. J. Lim, "Throughput Maximization for UAV-Enabled Mobile Relaying Systems," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 4983-4996, Dec. 2016.

[8] W. Wang, N. Qi, et al., "Energy-Efficient UAV-Relaying 5G/6G Spectrum Sharing Networks: Interference Coordination with Power Management and Trajectory Design," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1672-1687, 2022.

[9] H. Ajam, M. Najafi, et al., "Ergodic Sum Rate Analysis of UAV-Based Relay Networks with Mixed RF-FSO Channels," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 164-178, 2020.

[10] J.-H. Lee, K.-H. Park, et al., "Throughput Maximization of Mixed FSO/RF UAV-Aided Mobile Relaying with a Buffer," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 683-694, Jan. 2021.

[11] Y. Sun, D. Xu, et al., "Optimal 3D-Trajectory Design and Resource Allocation for Solar-Powered UAV Communication Systems," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4281-4298, Jun. 2019.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[13] S. P. Boyd and L. Vandenberg, *Convex Optimization*. Cambridge University Press, 2004.

[14] L. Xiang, D. W. K. Ng, et al., "Cross-Layer Optimization of Fast Video Delivery in Cache- and Buffer-Enabled Relaying Networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11366-11382, Dec. 2017.

[15] L. Xiang, L. Lei, et al., "Towards Power-Efficient Aerial Communications via Dynamic Multi-UAV Cooperation," in Proc. *IEEE WCNC*, pp. 1-7, 2020.

[16] T. Lipp and S. Boyd. "Variations and Extension of the Convex–Concave Procedure," *Optim. Eng.*, vol. 17, no. 2, pp. 263-287, 2016.