Joint Mobility, Communication and Computation Optimization for UAVs in Air-Ground Cooperative Networks

Jianshan Zhou, Daxin Tian, *Senior Member, IEEE*, Zhengguo Sheng, *Senior Member, IEEE*, Xuting Duan, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—Unmanned aerial vehicles (UAVs) play a significant role in various 5G or Beyond-5G (B5G)-enabled Internet-of-Things (IoT) applications. However, the UAV performance in an air-ground cooperative network is significantly affected by its mobility and air-to-ground (A2G) communication and computation behaviors. In this paper, a UAV-oriented computation offloading system is investigated, where the UAV desires to complete its onboard computation demands with the assistance of a ground edge-computing infrastructure, i.e., a road-side unit (RSU). The objective is to maximize the energy efficiency of the UAV. Specifically, a non-convex constrained optimal control problem is formulated to optimize the overall energy efficiency of UAV by jointly considering the coupled effects of UAV's longitudinal mobility, A2G communication, and computation dynamics. To address the coupled complexity and non-convexity of the original problem, a primal decomposition approach is developed to transform the problem into a convex subproblem and a primary problem, and then a closed-form optimal transmission power control is derived by solving the subproblem, which is dependent on mobility information. By embedding the closed-form optimal power control into the primary problem, a gradient projection-based iterative algorithm is proposed to obtain a joint optimal solution for both the longitudinal acceleration control and the power control, the feasibility and convergence of which is also theoretically proven. Extensive simulations have been conducted to validate the effectiveness of the proposed method in terms of constraint satisfaction and convergence speed, and comparative results also demonstrate that it can outperform other benchmark methods in terms of global energy efficiency.**

***Index Terms*—Air-ground cooperative networks, computation offloading, convex optimization, trajectory optimization, unmanned air vehicles (UAVs).**

Manuscript received September 10, 2020; revised February 8, 2021; accepted February 14, 2021. Date of publication February 17, 2021; date of current version April 2, 2021. This work was supported in part by the "Zhuoyue" Program of Beihang University (Postdoctoral Fellowship), in part by the China Postdoctoral Science Foundation under Grant 2020M680299, in part by the National Natural Science Foundation of China under Grants 61822101 and U20A20155, in part by the Beijing Municipal Natural Science Foundation L191001 and 4181002, the Newton Advanced Fellowship under Grant 62061130221, in part by the H2020-MSCA-RISE (101006411-SEEDS), and in part by the Royal Society Kan Tong Po International Fellowship under Grant KTP-R1-201007. The review of this article was coordinated by Dr. A.-C. Pang. *(Corresponding author: Daxin Tian.)*

Jianshan Zhou, Daxin Tian, and Xuting Duan are with the Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems & Safety Control, School of Transportation Science and Engineering, Beihang University, Beijing 100191, China (e-mail: jianshanzhou@foxmail.com; tiandaxin@gmail.com; duanxuting@buaa.edu.cn).

Zhengguo Sheng is with the Department of Engineering and Design, University of Sussex, Richmond 3A09, U.K. (e-mail: z.sheng@sussex.ac.uk).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/TVT.2021.3059964

I. Introduction

THE integration of unmanned aerial vehicles (UAVs) and advanced information and communications technologies (ICTs) has spawned a wide variety of applications ranging from remote sensing to disaster rescue [1], [2]. Generally, UAVs can form an aerial subnetwork and help the ground network (e.g., vehicular networks) through air-to-ground (A2G) communications, which motivates a novel networking architecture termed "Air-Ground Cooperative Networks" or "Air-Ground Integrated Networks" [3]. In the envisioned beyond 5G (B5G) or sixth-generation mobile networks (6G), UAVs are also considered as a key role to coordinate with Low Earth Orbit (LEO) and geostationary (GEO) satellites-based systems, which can create a space-air-ground integrated network (SAGIN) [4], [5]. Such an architecture enables seamless and flexible communication coverage, information services, and many other emerging Internet-of-Things (IoT) applications such as urban computing.

In addition, mobile edge computing (MEC) is currently considered as an emerging paradigm to support latency-critical and computation-intensive applications in B5G/6G, which can provide the computing resources at network edges to resource-hungry mobile users in close proximity to the edges [6]–[9]. Hence, many researchers are engaged in designing novel cooperation architectures, protocols, and algorithms for MEC-enabled communication and networking systems, such as low-latency fog-radio access network architectures [10], cooperative fog computing methods [11], and user's cooperative communication and computation approaches [12]. These conventional works mainly consider the optimization design of communication and computing in ground mobility scenarios. In the visions of B5G/6G, UAVs can also be equipped with computing and storage resources such that they are treated as flying MEC nodes to provide on-demand communication and computing supports for IoT applications. Thus, it is critical to investigate the air-ground cooperation of flying UAVs and ground infrastructure in both communication and computation.

Despite of many potential appealing opportunities provided by UAVs, there are some significant challenges to be addressed for the practical implementation of the air-ground cooperative network and the space-air-ground integrated network such as the

dynamic nature of wireless links due to the high mobility, communication and computation resource constraints. More importantly, the UAV's mobility, computation offloading, and communication power control are inherently coupled in a MEC-enabled air-ground cooperative network, which makes the system design even more complex. Energy consumption is one of the most important factors related to the mobility, communication and computation of UAVs. Without proper joint optimization design, the limited energy resources carried by a UAV can neither efficiently support itself to collect and process its computation tasks nor prolong its service lifetime. Therefore, it is of great significance to enhance the overall energy efficiency of UAVs from the perspective of joint optimization design, meanwhile maintaining the system reliability to process various applications on the energy resource-constrained flying platforms.

A. Motivation

UAVs can be combined with MEC technology to optimize their onboard resources. In both academia and industry, many research efforts are currently devoted to developing energy-efficient systems for UAV-assisted MEC systems [13]–[33]. Specifically, in many studies such as [15]–[17], [19], [27], UAVs are employed as mobile cloudlets to serve ground end-users by sharing their onboard computing resources. In the above works, UAVs are treated as service providers in the MEC scenario. However, due to the size and load constraints, the energy resource carried by a flying UAV is quite limited and also its operation lifetime heavily depends on the efficiency of resource utilization. The UAV-mounted cloudlet needs to be properly designed so as to guarantee the system practicability. UAVs can also have their own intensive computation demands, especially when they are used as airspace sensors to collect and process a large volume of sensing data, such as aerial images and video streams. Hence, UAVs, as IoT nodes, face the challenge arising from onboard applications that require massive storage and computing capacities. In this case, UAVs are resource-hungry service requesters rather than resource-rich service providers.

Inspired by the flexible mobility and low-cost commercialization of UAVs, many researchers have also dedicated themselves to developing novel deployment and mobility solutions for various UAVs-assisted communication systems [34]–[36]. For example, Chou *et al.* have proposed an energy-aware 3D deployment algorithm for a small cell-mounted UAV swarm [34]. In [35], an intelligent reflecting surface (IRS) optimization model is developed to the mean signal-to-interference-plus-noise ratio (SINR) of an air-ground communication network while mitigating inter-cell interferences. In [36], Chen *et al.* propose a unified consensus model by combining distributed energy minimization, grey prediction, and a dynamic spanning tree algorithm to achieve the robust deployment of a large-scale UAV swarm. The above works mainly focus on UAVs-oriented deployment and mobility optimization, while the issue of joint optimization of air-ground communication, computation scheduling, and UAVs' mobility remains to be explored at full length.

Specifically, UAVs need not only appropriately schedule the offloading of its computation demands but also autonomously adapt its motion under a certain control constraint set. The overall energy-efficient performance of UAVs inherently depends on the complex and coupled effects of mobility, communication and computation, which poses an important challenge on the optimization design. On one side, the successful offloading rate between UAVs and Road-Side Units (RSUs) relies on the performance of the A2G communication, meanwhile the computation partition of the UAV should also take into account its local computation performance. On the other side, the motion of the UAV will affect the time-varying relative distance between the UAV and a RSU, which further influences the dynamics of A2G communication. The mobility, communication and computation of the UAV account for most of its energy consumption, and thus should be optimally controlled from a global perspective. Therefore, in this paper, we investigate how UAVs and ground infrastructure can process the computation demands in a cooperative manner, and propose a novel UAV-oriented energy-efficient system via joint optimization. In our considered situation, a flying UAV is enabled to adaptively and dynamically offload the whole or partial computation demands to a ground edge-computing provider, i.e., a RSU, through A2G communication to optimize the efficiency of its limited energy resources. Meanwhile, the UAV is also required to complete its flight mission such as target tracking and ground traffic detection, i.e., satisfying a set of mobility control constraints.

B. Main Contributions

Towards this end, we investigate the air-ground cooperative networking scenario and propose a joint mobility, communication and computation optimization method for a UAV to maximize its overall energy efficiency. To be specific, we formulate different models with respect to the UAV's longitudinal motion, the wireless A2G communication, and the local computation, respectively. Based on the perspective of optimal control theory, we further develop a joint optimal control model that takes into consideration the coupled dynamics of the mobility, communication and computation of the UAV. The transmission power and the longitudinal acceleration are treated as the decision variables of the UAV, which are time-varying and constrained. To deal with the challenge in the optimal control system, we propose a primal decomposition approach to transform the original model into a subproblem and a primary problem, which reduces the complexity and thus facilitates the algorithm design. A closed-form optimal power control that relies on the mobility information has been theoretically derived by solving the subproblem. Then, a gradient projection-based algorithm combined with the optimal power control has been proposed to solve the primary problem, which results in the joint optimal control solution over a finite horizon. To the best of our knowledge, our work presents the first joint optimal control on the mobility, communication and computation of a UAV in an air-ground cooperative network, which can provide meaningful guidelines for air-ground cooperative computing and cross-layer system optimization.

The main contributions of our work are twofold with respect to modeling and optimization design:

- We develop a joint optimal control model by considering the coupled dynamics of the UAV's longitudinal mobility, A2G communication and local computation, and then

decompose such a complicated constrained optimization model into a tractable one.
- We theoretically derive and validate a closed-form and mobility-dependent formulation for the optimal transmission power control of the UAV, which is used to dynamically adapt the computation offloading according to the UAV's motion and exogenous application demands.
- Based on the closed-form expression of the optimal power control, we further propose a gradient projection-based algorithm to solve the joint optimal control model. The descent feasibility and convergence of the proposed algorithm in the closed control constraint domain are guaranteed by theoretical analysis.
- We have carried out extensive simulation experiments to validate the effectiveness of the proposed method and demonstrate its advantage. We compare our method with other conventional methods based on the aggressive mobility, aggressive offloading, and local computation mechanisms, respectively, and show that our method can outperform the conventional methods in terms of the global energy efficiency.

The rest of this paper is organized as follows: we review the related work in Section II and formulate the system model of the joint optimal control in Section III. In Section IV, we present a primal decomposition for model transformation and derive the closed-form solution for a subproblem. In Section V, a gradient projection-based algorithm is proposed with the closed-form solution and its convergence is theoretically analyzed. Section VI conducts the performance evaluation, and Section VII concludes our paper and outlines the future work.

II. Related Work

UAVs are currently playing significant roles in various domains due to their highly flexible mobility. As a result, diverse UAV-based intelligent systems have been attracting much research attention from both academia and industry. In particular, UAVs equipped with emerging ICT and control technologies are envisioned to boost a technological shift from conventional ground infrastructure-based communication and networking paradigms to air-ground cooperative paradigms [1], [2], [5], [9]. For instance, [3] presents a novel architecture of aerial-ground cooperative vehicular networks in which multiple UAVs are used to help forward vehicular information and enhance the vehicular connectivity. [13] develops an anti-jamming vehicular ad-hoc networks by integrating the relaying functionality of a UAV.

Specifically, there are many researchers currently focusing on joint communication and trajectory optimization design for UAV-based networking systems. In [28], the researchers have formulated a system throughput maximization model to optimize the relaying UAV's transmission power along with its flying trajectory and proposed an iterative algorithm based on successive convex optimization technique. In [26], the researchers also aim at maximizing the sum data rate of a UAV-based vehicular communication system, for which the UAV works as an aerial base station and a trajectory control solution has been designed by resorting to the theory of reinforcement learning. In [24], the reinforcement learning theory is also adopted to design the UAV's trajectory with the goal of maximizing the expected uplink sum rate of a UAV-aided cellular network. The work [20] leverages deep reinforcement learning for energy-efficient control design of UAVs, which aims at improving the coverage and connectivity of a UAV-based cellular communication system. In both [29], [30], the problem of the coverage and deployment optimization of multiple UAVs has also been studied in terms of energy efficiency, for which the optimal transport and the block coordinate descent methods are applied, respectively. In [14], [23], the authors have investigated the issue on energy-efficient UAV communication and trajectory optimization. In their system designs, the communication throughput in the sense of energy efficiency is treated as the optimization objective. Differently, [14] exploits the sequential convex optimization technique to solve their joint optimization model, while an iterative algorithm has been proposed based on the alternating directional method of multipliers (ADMM) in [23]. Besides focusing on joint communication and trajectory design, many researchers have also been engaged in developing UAV-oriented secure communication methods by jointly considering UAV communication and mobility control such as [25], [31]. From these aforementioned studies [14], [20], [23]–[26], [28]–[31], it can be summarized that many significant and novel paradigms have been designed and well validated from the perspective of joint UAV communication and mobility optimization. Nevertheless, the decision-making modeling and optimization of mobile edge computing in air-ground cooperative scenarios has been ignored in these works, while the computation-related energy consumption can actually affect the overall energy efficiency of UAVs.

As MEC is widely considered as a promising technology for tackling the challenge arising from computation/data-intensive applications in smart but resource-hungry mobile devices, there are also many research efforts that have been made to combine MEC with UAV-based systems. For instance, [27] targets maximizing the computation rate of a UAV-enabled wireless-powered MEC system, which proposes a two-stage and a three-stage algorithms for computation offloading of ground end-users. In such a situation, a UAV with constant-velocity motion is used as computing resource providers to serve ground end-users. In [4], UAVs are expected to be integrated with both ground and space networking systems to form a space-air-ground integrated network, and a joint resource allocation and task scheduling method is proposed for UAV edge servers. In their scenarios, the mobility impacts of the UAVs have been neglected to reduce the modeling complexity. Additionally, some other works such as [15]–[19], [21], [22], [32], [33] differentiate their studies from the above ones [4], [27] by introducing the idea of joint computation offloading and trajectory optimization. Specifically, in [15], a UAV is working as a cloudlet, and its trajectory and bit allocation are jointly optimized by exploiting successive convex approximation so as to support computation offloading from ground mobile end-users. Nonetheless, the goal of their system design is to minimize the overall energy consumption of all the ground end-users instead of the energy cost of the UAV. Similarly, in [16], the authors have studied the problem of joint ground users' computation offloading and a UAV server's trajectory design, and they proposed a penalty algorithm based on dual decomposition to handle the non-convexity of the system model. The optimization objective is to minimize the

sum of the maximum latency of the overall ground end-users. In [18], the authors aim to optimize the average weighted energy consumption of both ground end-users and a UAV by jointly considering the ground end-users' computation offloading, the UAV's trajectory scheduling and resource allocation. In their work, a stochastic solution is presented by a Lyapunov-based queueing analysis and a decomposition-based algorithm design. In [19], a successive convex approximation-based algorithm has been employed to minimize the energy consumption of a UAV, in which the terminal CPU frequencies, the offloading data amount, the transmit power of ground end-users and the UAV's trajectory are jointly considered and optimized. Although the system model of [19] has taken into account the computation and mobility-related energy consumption, it does not aim to control or optimize the transmission power, and the initial or terminal position and velocity of the UAV are treated as free design variables to reduce the complexity in its trajectory optimization. In reality, the initial or terminal state of a UAV is usually restricted, i.e., the UAV is expected to start at a specified initial position with a given velocity and end at a targeted state while the flying trajectory is optimized, especially when it is assigned with a flight mission. In [21], a successive convex approximation approach combined with the Dinkelbach's fractional programming algorithm has also been adopted to deal with the problem of jointly optimizing the UAV trajectory, the ground user's transmit power, and computation load allocation. In [22], the UAV's trajectory is the only design factor and a mixed-integer non-convex model has been formulated to maximize the sum rate of edge users served by the UAV. In [33], the authors jointly optimize the bandwidth allocation, the computation resource, and the UAV's trajectory. The researchers proposed a three-step optimization scheme to address the modeling complexity [22], in which they first fix the UAV's trajectory and the allocated bandwidth to schedule the computation resource, then optimize the bandwidth allocation by fixing the obtained computation resource allocation in the second step, and finally design the UAV's trajectory. However, the overall solution obtained by the three-step design may not be optimal in terms of global optimality, since the inherent coupling effects in their original model have not been incorporated the design steps. Besides those works mentioned above, successive convex approximation-based methods have been widely adopted to tackle the joint flight trajectory and wireless caching optimization issues, which, for instance, can be found in [17], [32]. Nevertheless, even though the successive convex approximation technique is powerful to deal with the modeling non-convexity, it may fail to satisfy the solution feasibility during iterations due to the potential existence of a gap between the approximated and the actual objective or constraint equations.

In this paper, we differentiate our contributions from these aforementioned works by jointly modeling UAV's longitudinal mobility, A2G wireless communication and onboard computation, and by proposing a novel joint optimization design based on primal decomposition and gradient projection mechanisms. To be specific, unlike the aforementioned works, we treat a UAV as a resource-hungry requester instead of a resource-rich server, and both the longitudinal acceleration and the transmission power of the UAV are design variables. The trajectory of the UAV is not directly optimized in our work. Instead, the energy-efficient trajectory is obtained by the UAV's kinematics that is driven by the optimal acceleration control signal. The initial and terminal constraints are also considered in our model. Besides, the coupling effects in the UAV's mobility, communication and computation are explicitly handled, and we can derive a closed-form and mobility information-dependent optimal power control. Based on the closed-form power control, we further propose a gradient projection method to solve the primary problem, the iterative feasibility and convergence of which are theoretically proven and well guaranteed.

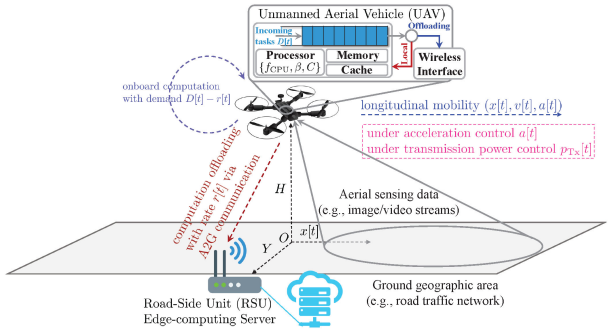

Fig. 1. An exemplary application scenario of an air-ground cooperative network where a cruising UAV cooperates with a ground RSU (an edge-computing server with a wireless interface) to process the real-time aerial sensing data via A2G computation offloading. The UAV is enabled to maximize its overall energy efficiency by joint optimization of mobility, communication and computation related control variables.

III. System Model

As shown in Fig. 1, we consider a general A2G computation offloading scenario where a UAV with limited capacities of communication, computing and mobility has to offload a part of its computation or its whole computation to a ground infrastructure (e.g., an RSU) for remote execution. In such a situation, the UAV needs to fly from a given initial position to a targeted position under a proper speed control strategy. From a systematic perspective, the mobility of the UAV will significantly affect the dynamics of the relative geographical distance between itself and the RSU over time, which will have a further impact on the A2G communication link performance. On the other side, the computation offloading, which depends on the A2G link transmission, will be also influenced by the UAV mobility. At this point, the mobility, communication and computing dynamics of the UAV are inherently coupled, hence we aim at developing a joint optimization framework. For practical implementation, we consider that the continuous time horizon $[t_0, t_f]$ is discretized into N discrete intervals each with the duration of $\Delta\tau = (t_f - t_0)/N$ seconds, i.e., $[t_0, t_f] = [t_0, t_1) \cup [t_1, t_2) \cup \cdots \cup [t_{N-1}, t_N]$, where $t_0 \geq 0$ and $t_f > t_0$ denote the initial and the terminal time instants, respectively. Let t be the time interval index such that $t = 0, 1, \ldots, N$. In the following subsections, we formulate different models for the UAV in terms of mobility, communication and computing.

A. Mobility Modeling

For simplicity but without loss of generality, we consider to optimize the longitudinal mobility of the targeted UAV, while

assuming that the flight altitude and the lateral position do not change within $[t_0, t_f]$. Nevertheless, it is remarked here that the mobility modeling approach in this paper can be easily extended to multiple dimensions, e.g., the 3-dimensional coordinate space. Let the flight altitude and the lateral position of the UAV be $H > 0$ and $Y > 0$, respectively, while the time-varying longitudinal position is denoted by $x[t]$. Besides, the longitudinal velocity is denoted by $v[t]$. With these notations, we can represent a time-varying mobility state of the UAV as a column vector $\mathbf{s}[t] = [x[t], v[t]]^{\mathrm{T}}$. We also let the time-varying velocity control be $a[t]$. Now, we further propose the dynamical system using a time-discrete double-integral model to describe the UAV mobility as follows

$$\mathbf{s}[t+1] = \mathbf{A}\mathbf{s}[t] + \mathbf{B}a[t], \ \ \forall t, \tag{1}$$

where $\mathbf{A}$ and $\mathbf{B}$ are the state matrix and the control matrix, respectively, which are given as

$$\mathbf{A} \triangleq \begin{bmatrix} 1 & \Delta\tau \\ 0 & 1 \end{bmatrix}, \ \mathbf{B} \triangleq \begin{bmatrix} \frac{1}{2}(\Delta\tau)^2 \\ \Delta\tau \end{bmatrix}. \tag{2}$$

To model the mobility-related energy consumption of the UAV, we assume that the UAV is of rotary-wing type as widely considered in most existing literature. Let $E_{\mathbf{mob}}[t]$ denote its propulsive energy consumption at t. According to [14], [15], [37], $E_{\mathrm{mob}}[t]$ can be approximately modeled by

$$E_{\mathbf{mob}}[t] = a_1 \left\|v[t]\right\|^3 + \frac{a_2}{\left\|v[t]\right\|}\left(1 + \frac{\left\|a[t]\right\|^2}{g^2}\right), \tag{3}$$

where $a_1 > 0$ and $a_2 > 0$ are two positive coefficients related to the fluid dynamics and aerodynamic layout of the rotary-wing UAV. g is the constant gravitational acceleration, whose typical value is 9.8 $\mathrm{m/s^2}$.

B. Communication Modeling

Let the position of the UAV be $\mathbf{p}_{\mathrm{UAV}}[t] = [x[t], Y, H]^{\mathrm{T}}$ and the specific RSU be $\mathbf{p}_{\mathrm{RSU}}$. The relative distance between the UAV and the RSU can be formulated as follows

$$d[t] = \left\|\mathbf{p}_{\mathrm{UAV}}[t] - \mathbf{p}_{\mathrm{RSU}}\right\|, \ \ \forall t. \tag{4}$$

To model the physical-layer transmission over the A2G link, we resort to the data rate formula in the Shannon's form that has been widely used in recent studies [13]–[19], [30]. Let the available bandwidth assigned for this A2G link be ω, the path loss exponent over this link be α, the average noise power be N_0, and the transmission power of the UAV at t be $p_{\mathrm{Tx}}[t]$. When adopting the power control $p_{\mathrm{Tx}}[t]$, the communication energy consumption of the UAV at t is

$$E_{\mathrm{Tx}}[t] = p_{\mathrm{Tx}}[t]\Delta\tau. \tag{5}$$

Additionally, the achievable data rate of the A2G link at t is then formulated as a logarithmic function of the transmission power $p_{\mathrm{Tx}}[t]$ and the relative distance $d[t]$ as follows

$$r[t] = \omega \log_2\left(1 + a_3 p_{\mathrm{Tx}}[t] d^{-\alpha}[t]\right), \tag{6}$$

where $a_3 = a_4/N_0$ and $a_4 > 0$ is an additional scalar factor used to characterize the attenuation of the A2G link. When the A2G link is a line-of-sight (LoS) link, a_4 can be fixed at $a_4 = 1$, while $a_4 \in (0, 1)$ when the link is a non-line-of-sight (NLoS). Besides, some current works [30], [38], [39] also suggest that a_4 can be specified as the average attenuation of the A2G link in a stochastic situation, i.e., $a_4 = p_{\mathrm{LoS}} + \eta p_{\mathrm{NLoS}}$, where p_{LoS} and p_{NLoS} denote the probabilities of the LoS and the NLoS situations, respectively, and η denotes the attenuation factor due to the NLoS effect.

C. Computation Modeling

Regarding the computation offloading of the UAV, we denote the arrival computation demand in bits at t by $D[t]$, which is the exogenous factor of the aerial-ground integrated edge-computing network. Given the offloading data rate, i.e., $r[t]$ in (6), the UAV needs to consume certain onboard energy to process the partial data, $q[t] = D[t] - r[t]\Delta\tau$. Specifically, letting the effective switched capacitance of the onboard CPU be β, the clock frequency of the onboard CPU be f_{CPU}, and the number of the CPU clock cycles required to process per-bit input data is C, we can model the computation energy consumption for processing $q[t]$-bit input data as follows

$$E_{\mathrm{com}}[t] = \beta C q[t] f_{\mathrm{CPU}}^2 = \beta C \left(D[t] - r[t]\Delta\tau\right) f_{\mathrm{CPU}}^2. \tag{7}$$

D. Joint Optimization Model

Using the mobility, communication and computing models developed above, we can further propose the joint optimization model under a set of state and control bound constraints. To be specific, let the initial state be $\mathbf{s}_0 = [x_0, v_0]^{\mathrm{T}}$, where x_0 is the initial longitudinal position, and v_0 is the initial longitudinal velocity. We further denote the upper bound of the longitudinal position of the UAV during motion as $x_f > x_0$. Thus, the longitudinal position bound is assumed to be $\mathcal{X} \triangleq [x_0, x_f]$, while the longitudinal velocity bound is $\mathcal{V} \triangleq [v_{\min}, v_{\max}]$ where $v_{\min}$ and $v_{\max}$ are the upper and the lower bounds on the longitudinal velocity of the UAV, respectively. Then, we can present the state bound constraint as $\mathbf{s}[t] \in \mathcal{S} \triangleq \mathcal{X} \times \mathcal{V}$ for $t = 0, 1, \ldots, N-1$. At the terminal time $t = N$, we denote the terminal state space as $\mathcal{S}_f$, i.e., letting $\mathbf{s}[N] \in \mathcal{S}_f$. In addition, we let the allowable minimum and maximum velocity control be $a_{\min}$ and $a_{\max}$, respectively, and the allowable minimum and maximum power control be $p_{\min}$ and $p_{\max}$, respectively. The joint system control at t can be represented by $\mathbf{u}[t] = [a[t], p_{\mathrm{Tx}}[t]]^{\mathrm{T}}$.

Accordingly, we formulate the global objective of jointly minimizing the whole energy consumption required by the communication, computing and mobility of the UAV as

$$\begin{aligned} &J\left(\mathbf{s}[0]; \mathbf{u}[0], \ldots, \mathbf{u}[N-1]\right) \\ &\quad = g(\mathbf{s}[N], N) + \sum_{t=0}^{N-1} g(\mathbf{s}[t], \mathbf{u}[t], t), \end{aligned} \tag{8}$$

where $g(\mathbf{s}[t], \mathbf{u}[t], t)$ is the cost function of the system in state $\mathbf{s}[t]$ under control $\mathbf{u}[t]$ at t, which combines the energy consumption related to the UAV mobility, communication and computing, i.e.,

$$g(\mathbf{s}[t], \mathbf{u}[t], t) = E_{\mathrm{mob}}[t] + E_{\mathrm{Tx}}[t] + E_{\mathrm{com}}[t], \tag{9}$$

while $g(\mathbf{s}[N], N)$ is the terminal cost depending on the terminal state the UAV actually reaches, $\mathbf{s}[N]$. To model the terminal cost function, we can set $\mathbf{s}_f = [x_f, v_f]^{\mathrm{T}} \in \mathcal{S}_f$ to be an expected

terminal state for the UAV at $t = N$, where v_f denotes the expected terminal velocity. Based on this, we propose an utility function for $g(\mathbf{s}[N], N)$ as follows

$$g(\mathbf{s}[N], N) = \begin{cases} a_1 \left\| \left[\frac{x_f - x[N]}{\Delta\tau}, v_f - v[N] \right] \right\|^3, \mathbf{s}[N] \in \mathcal{S}_f; \\ C_{\text{penalty}}, \text{ otherwise}; \end{cases} \tag{10}$$

in which a_1 is given as in (3), and C_{penalty} is a sufficiently large penalty factor when the UAV cannot reach the expected terminal state $\mathbf{s}_f$ at the end. From (9), it can be found that $0 \leq g(\mathbf{s}[N], N) \leq C_{\text{penalty}}$. When $\mathbf{s}[N] = \mathbf{s}_f$, the terminal cost $g(\mathbf{s}[N], N)$ is zero, while the terminal cost $g(\mathbf{s}[N], N)$ is positive when $\mathbf{s}[N]$ deviates $\mathbf{s}_f$. The closer the terminal state $\mathbf{s}[N]$ is to $\mathbf{s}_f$, the smaller the terminal cost $g(\mathbf{s}[N], N)$ is.

To optimize the overall energy efficiency, we further propose the following joint optimization model $\mathcal{M}_1$ by treating $a[t]$ and $p_{\text{Tx}}[t]$ as the control variables

$$\min_{\{\mathbf{u}[t], t=0,\ldots,N-1\}} : J(\mathbf{s}[0]; \mathbf{u}[0], \ldots, \mathbf{u}[N-1])$$
$$\text{s.t.} \begin{cases} \mathbf{s}[t+1] = \mathbf{A}\mathbf{s}[t] + \mathbf{B}a[t], t = 0, \ldots, N-1; \\ \mathbf{s}[0] = \mathbf{s}_0; \\ \mathbf{s}[t] \in \mathcal{S}, t = 0, \ldots, N-1; \\ a[t] \in [v_{\min}, v_{\max}], t = 0, \ldots, N-1; \\ p_{\text{Tx}}[t] \in [p_{\min}, p_{\max}], t = 0, \ldots, N-1. \end{cases} \tag{11}$$

From (11), it can be seen that $\mathcal{M}_1$ is a complex dynamical optimal control problem under a set of bound constraints. Indeed, it is difficult or even impossible to solve $\mathcal{M}_1$ directly by using traditional optimization algorithms such as the Newton's methods. In the following sections, we would like to develop a novel approach for solving $\mathcal{M}_1$ by combining a primal decomposition method and a gradient projection-based method.

IV. Primal Decomposition Method With Convex Optimization

In the primal model $\mathcal{M}_1$, we can find that the velocity control $a[t]$ is directly coupled with the communication energy consumption, $E_{\text{com}}[t]$, with respect to the transmit power control $p_{\text{Tx}}[t]$ via the distance-dependent transmission data rate, i.e., $r[t]$ as in (6). This motivates us to propose a primal decomposition transformation to convert $\mathcal{M}_1$ into another equivalent form that is mathematically tractable.

A. Primal Decomposition

Specifically, at each t we fix the distance between the UAV and the RSU, $d[t]$, which depends on the current system state $\mathbf{s}[t]$, and thus we are allowed to separately optimize the objective function with respect $p_{\text{Tx}}[t]$. At this point, we formulate the following model $\mathcal{M}_2$ given the fixed $d[t]$ at t based on (5), (6) and (7)

$$h(d[t]) = \min_{p_{\text{Tx}}[t]} : p_{\text{Tx}}[t]\Delta\tau + \beta C(D[t] - r[t]\Delta\tau) f_{\text{CPU}}^2$$
$$\text{s.t.} \begin{cases} r[t] = \omega \log_2 \left(1 + a_3 p_{\text{Tx}}[t] d^{-\alpha}[t]\right) \\ p_{\text{Tx}}[t] \in [p_{\min}, p_{\max}], \end{cases} \tag{12}$$

in which $h(d[t])$ denotes the optimal value of $\mathcal{M}_2$.

Following $\mathcal{M}_2$, the primal model $\mathcal{M}_1$ is then reformulated as follows, which is denoted by $\mathcal{M}_3$

$$\min_{\{a[t], t=0,\ldots,N-1\}} : g(\mathbf{s}[N], N) + \sum_{t=0}^{N-1} \left[E_{\text{mob}}[t] + h(d[t])\right]$$
$$\text{s.t.} \begin{cases} \mathbf{s}[t+1] = \mathbf{A}\mathbf{s}[t] + \mathbf{B}a[t], t = 0, \ldots, N-1; \\ \mathbf{s}[0] = \mathbf{s}_0; \\ \mathbf{s}[t] \in \mathcal{S}, t = 0, \ldots, N-1; \\ a[t] \in [v_{\min}, v_{\max}], t = 0, \ldots, N-1. \end{cases} \tag{13}$$

As demonstrated above, using the decomposition structure, we can solve the subproblem $\mathcal{M}_2$ with respect to only a decision variable $p_{\text{Tx}}[t]$, and then further address the primal problem with respect to $a[t]$. The subproblem $\mathcal{M}_2$ and the primal problem $\mathcal{M}_3$ actually share the common information incorporated in $d[t]$.

B. Convex Optimization for Subproblem $\mathcal{M}_2$

To deal with $\mathcal{M}_2$, we first derive the following results by proving its convexity.

Corollary 1: The subproblem $\mathcal{M}_2$ guarantees that there only exists an unique global optimum point in $[p_{\min}, p_{\max}]$.

Proof: For simplicity, we denote the objective function of $\mathcal{M}_2$ by $F(p_{\text{Tx}}[t]) = p_{\text{Tx}}[t]\Delta\tau + \beta C(D[t] - r[t]\Delta\tau) f_{\text{CPU}}^2$. The second-order derivative of $F(p_{\text{Tx}}[t])$ with respect to $F(p_{\text{Tx}}[t])$ is expressed as

$$\frac{d^2 F(p_{\text{Tx}}[t])}{d(p_{\text{Tx}}[t])^2} = \frac{\beta C f_{\text{CPU}}^2 \Delta\tau \omega a_3^2 \, d^{-2\alpha}[t]}{\ln 2 \left(1 + a_3 p_{\text{Tx}}[t] d^{-\alpha}[t]\right)^2}. \tag{14}$$

Recalling $d[t] > 0$, we can have $\frac{d^2 F(p_{\text{Tx}}[t])}{d(p_{\text{Tx}}[t])^2} > 0$. The result indicates that $F(p_{\text{Tx}}[t])$ is a strictly convex function with respect to $p_{\text{Tx}}[t]$. In addition, $\mathcal{M}_2$ is bounded by the closed domain $[p_{\min}, p_{\max}]$. Therefore, it well guarantees the strict convexity, which proves the corollary. ■

Following Corollary 1, we define the Lagrangian function associated with $\mathcal{M}_2$ as follows

$$L(p_{\text{Tx}}[t], \lambda_1, \lambda_2) = F(p_{\text{Tx}}[t]) - \lambda_1 \left(p_{\text{Tx}}[t] - p_{\min}\right) - \lambda_2 \left(p_{\max} - p_{\text{Tx}}[t]\right), \tag{15}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are two nonnegative Lagrangian multipliers. Based on the first-order necessary conditions (i.e., the Karush-Kuhn-Tucker (KKT) conditions), we obtain the following result:

Corollary 2: For $\mathcal{M}_2$, suppose that $p_{\text{Tx}}^{\text{opt}}[t]$ is its unique global optimal solution. The Lagrangian multipliers λ_1 and λ_2 can be identical to zero simultaneously if $p_{\min} < p_{\text{Tx}}^{\text{opt}}[t] < p_{\max}$; Otherwise, we only have $\lambda_2 = 0$ if $p_{\text{Tx}}^{\text{opt}}[t] = p_{\min}$, and $\lambda_1 = 0$ if $p_{\text{Tx}}^{\text{opt}}[t] = p_{\max}$.

Proof: According to Corollary 1, the subproblem $\mathcal{M}_2$ guarantees an unique global point $p_{\text{Tx}}^{\text{opt}}[t]$. It satisfies the KKT conditions as follows

$$\begin{cases} \nabla L(p_{\text{Tx}}^{\text{opt}}[t], \lambda_1, \lambda_2) = \nabla F(p_{\text{Tx}}^{\text{opt}}[t]) - \lambda_1 + \lambda_2 = 0; \\ \lambda_1 \left(p_{\text{Tx}}^{\text{opt}}[t] - p_{\min}\right) = 0; \\ \lambda_2 \left(p_{\max} - p_{\text{Tx}}^{\text{opt}}[t]\right) = 0; \\ \lambda_1 \geq 0, \ \lambda_2 \geq 0, \ p_{\text{Tx}}^{\text{opt}}[t] \in [p_{\min}, p_{\max}]. \end{cases} \tag{16}$$

In (16), when the optimal point is an interior point, i.e., $p_{\text{Tx}}^{\text{opt}}[t] \in (p_{\min}, p_{\max})$, the complementary slackness (i.e., the

second and the third equations) and the non-negativity of the Lagrangian multipliers result in $\lambda_1 = \lambda_2 = 0$. Otherwise, when $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\min}$, the third complementary slackness leads to $\lambda_2 = 0$, while, when $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\max}$, the second complementary slackness leads to $\lambda_1 = 0$. ■

From Corollary 2, it can be seen that when the optimal point of $\mathcal{M}_2$ is a boundary point of $[p_{\min}, p_{\max}]$, these two Lagrangian multipliers, λ_1 and λ_2, cannot be positive simultaneously. That is, it always holds that $\lambda_1 \times \lambda_2 = 0$. Based on Corollary 2, we derive the closed-form optimal solution for $\mathcal{M}_2$ as given in the following theorem.

Theorem 1: For $\mathcal{M}_2$, define $p_{\mathrm{Tx}}^*[t]$, λ_1^*, and λ_2^* as follows

$$\begin{cases} p_{\mathrm{Tx}}^*[t] = \frac{\beta C f_{\mathrm{CPU}}^2 \omega}{\ln 2} - \frac{1}{a_3\, d^{-\alpha}[t]}; \\ \lambda_1^* = \Delta\tau - \frac{\beta C f_{\mathrm{CPU}}^2 \Delta\tau\omega a_3\, d^{-\alpha}[t]}{\ln 2(1+a_3 p_{\min} d^{-\alpha}[t])}; \\ \lambda_2^* = \frac{\beta C f_{\mathrm{CPU}}^2 \Delta\tau\omega a_3\, d^{-\alpha}[t]}{\ln 2(1+a_3 p_{\max} d^{-\alpha}[t])} - \Delta\tau. \end{cases} \tag{17}$$

The optimal solution of $\mathcal{M}_2$, $p_{\mathrm{Tx}}^{\mathrm{opt}}[t]$, and the Lagrangian multipliers, λ_1 and λ_2, must satisfy one of the following three situations:

i) If $p_{\min} \le p_{\mathrm{Tx}}^*[t] \le p_{\max}$, the optimal solution is $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\mathrm{Tx}}^*[t]$. The Lagrangian multipliers satisfy $\lambda_1 = \lambda_2 = 0$.

ii) If $\lambda_1^* > 0$, the optimal solution is $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\min}$ and the Lagrangian multipliers are $\lambda_1 = \lambda_1^*$ and $\lambda_2 = 0$, respectively.

iii) If $\lambda_2^* > 0$, the optimal solution is $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\max}$ and the Lagrangian multipliers are $\lambda_1 = 0$ and $\lambda_2 = \lambda_2^*$, respectively.

Proof: According to Corollary 2, we can only have three situations for the values of λ_1 and λ_2, i.e., (i) $\lambda_1 = \lambda_2 = 0$, (ii) $\lambda_1 > \lambda_2 = 0$, and (iii) $0 = \lambda_1 < \lambda_2$. In the first situation, we can solve the gradient condition in (16), i.e.,

$$\Delta\tau - \frac{\beta C f_{\mathrm{CPU}}^2 \Delta\tau\omega a_3\, d^{-\alpha}[t]}{\ln 2\,(1 + a_3 p_{\mathrm{Tx}}[t] d^{-\alpha}[t])} - \lambda_1 + \lambda_2 = 0, \tag{18}$$

with $\lambda_1 = \lambda_2 = 0$, which results in $p_{\mathrm{Tx}}^*[t]$ as in (17). Thus, the optimal solution is $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\mathrm{Tx}}^*[t]$ if $p_{\mathrm{Tx}}^*[t] \in [p_{\min}, p_{\max}]$.

In the second situation, $\lambda_1 > 0$ indicates that the optimal solution must be $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\min}$. Combining this result with $\lambda_2 = 0$, we can solve λ_1 from the gradient condition, i.e.,

$$\Delta\tau - \frac{\beta C f_{\mathrm{CPU}}^2 \Delta\tau\omega a_3\, d^{-\alpha}[t]}{\ln 2\,(1 + a_3 p_{\min} d^{-\alpha}[t])} - \lambda_1 = 0, \tag{19}$$

which results in λ_1^* as in (17). Thus, we can set $\lambda_1 = \lambda_1^*$ if $\lambda_1^* > 0$.

Following the same logic, in the third situation, $\lambda_2 > 0$ indicates that the optimal solution must be $p_{\mathrm{Tx}}^{\mathrm{opt}}[t] = p_{\max}$. Using $\lambda_1 = 0$, we can also derive λ_2 from the gradient condition, i.e.,

$$\Delta\tau - \frac{\beta C f_{\mathrm{CPU}}^2 \Delta\tau\omega a_3\, d^{-\alpha}[t]}{\ln 2\,(1 + a_3 p_{\max} d^{-\alpha}[t])} + \lambda_2 = 0, \tag{20}$$

which results in λ_2^* as in (17). Hence, we can set $\lambda_2 = \lambda_2^*$ if $\lambda_2^* > 0$.

Recall that $\mathcal{M}_2$ guarantees the strict convexity. It has only an unique optimal point that must satisfy one of the three situations above. At this point, the theorem is proven. ■

Based on Theorem 1, we can further derive the closed-form expression of the optimal objective function $h(d[t])$ in three cases as follows:

TABLE I
LIST OF SYMBOLS

Symbol	Definition
t	time slot index
N	number of time slots
t_0	initial time instant
t_f	terminal time instant
$\Delta\tau$	duration of a time slot
$x[t]$	UAV's longitudinal position at t
H	UAV's flight height
Y	UAV's lateral position
$v[t]$	UAV's longitudinal velocity at t
$v_{\min}$, $v_{\max}$	minimum and maximum longitudinal velocities
$a[t]$	UAV's motion control variable at t
$a_{\min}$, $a_{\max}$	minimum and maximum accelerations
$\mathbf{s}[t]$	UAV's motion state at t
$\mathcal{X}$, $\mathcal{V}$, $\mathcal{S}$	position, velocity and motion state constraint sets
$E_{\mathrm{mob}}[t]$	mobility-related energy consumption at t
a_1, a_2	UAV's propulsion energy consumption parameters
$\mathbf{p}_{\mathrm{UAV}}[t]$	UAV's position vector at t
$\mathbf{p}_{\mathrm{RSU}}$	RSU's position vector
$d[t]$	relative distance between UAV and RSU at t
$p_{\mathrm{Tx}}[t]$	UAV's power control variable at t
$p_{\min}$, $p_{\max}$	minimum and maximum power levels
$E_{\mathrm{Tx}}[t]$	communication-related energy consumption at t
ω	available bandwidth
α	path loss exponent
N_0	environmental noise power
a_3, a_4	A2G channel attenuation factors
$r[t]$	computation offloading rate at t
β	CPU effective switched capacitance
C	number of CPU cycles for processing 1 bit
f_{CPU}	CPU frequency
$D[t]$	arrival rate of computation tasks at t
$E_{\mathrm{com}}[t]$	computation-related energy consumption at t

i) When $p_{\min} \le p_{\mathrm{Tx}}^*[t] \le p_{\max}$ holds true, the optimal energy consumption of communication and computing is

$$h(d[t]) = \left(\frac{\beta C f_{\mathrm{CPU}}^2 \omega}{\ln 2} - \frac{1}{a_3\, d^{-\alpha}[t]}\right)\Delta\tau + \beta C f_{\mathrm{CPU}}^2 D[t] - \beta C f_{\mathrm{CPU}}^2 \omega \log_2\left(\frac{a_3\, d^{-\alpha}[t]\beta C f_{\mathrm{CPU}}^2 \omega}{\ln 2}\right)\Delta\tau. \tag{21}$$

ii) When $\lambda_1^* > 0$ holds true, the optimal energy consumption of communication and computing is

$$h(d[t]) = p_{\min}\Delta\tau + \beta C f_{\mathrm{CPU}}^2 D[t] - \beta C f_{\mathrm{CPU}}^2 \omega \log_2\left(1 + a_3 p_{\min} d^{-\alpha}[t]\right)\Delta\tau. \tag{22}$$

ii) When $\lambda_2^* > 0$ holds true, the optimal energy consumption of communication and computing is

$$h(d[t]) = p_{\max}\Delta\tau + \beta C f_{\mathrm{CPU}}^2 D[t] - \beta C f_{\mathrm{CPU}}^2 \omega \log_2\left(1 + a_3 p_{\max} d^{-\alpha}[t]\right)\Delta\tau. \tag{23}$$

To verify the proposed theorem above, we conduct a number of experiments and then compare the theoretical results obtained by the closed-form model with the numerical solution-based results. To be specific, a UAV is simulated to be cruising over a road traffic network and needs to process its collected computation tasks on demand (e.g., real-time traffic sensing data) via the air-ground cooperation between itself and a RSU as illustrated in Figure. 1. According to the existing literature [14], [15], [37], [40], the simulation parameters related to the mobility,

TABLE II
SIMULATION PARAMETERS

Parameter	Value
$\Delta\tau$, g	50 ms, 9.8 m/s^2
a_1, a_2	0.0037, 5.0206
$v_{\min}$, $v_{\max}$	1 m/s, 30 m/s
$a_{\min}$, $a_{\max}$	-5 m/s^2, 5 m/s^2
α, ω	3, 10 MHz
a_4, N_0	1.0, -50 dBm
$p_{\min}$, $p_{\max}$	-50 dBm, 30 dBm
β	10^{-28}
C	1550.7
f_{CPU}	800 MHz

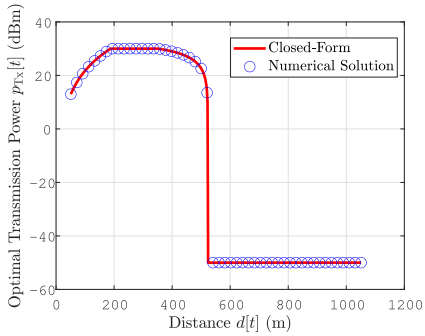

Fig. 2. Optimal transmission power against different relative distances between the UAV and the RSU.

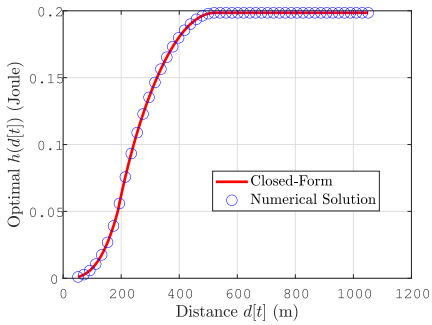

Fig. 3. Optimal communication and computation energy consumption against different relative distances between the UAV and the RSU.

communication and computation models of the UAV are given in Table II. In addition, we fix the flight height and the latitude of the UAV at 50 m and 10 m, respectively, while the altitude and latitude positions of the RSU are set to zero. The longitudinal position of the RSU is set to $\frac{x_0+x_f}{2}$, i.e., it is located at the middle position of the UAV's trajectory. The initial and the terminal velocities of the UAV are set to $v_0 = v_f = 1\,\text{m/s}$, the flight duration is set to $t_f - t_0 = 30$ seconds, and the computation demand is fixed at $D[t] = 2 \times 10^6$ bits for all t. It is noted that the UAV is set to move with the constant velocity since we do not jointly optimize its mobility here. The joint optimization method will be validated in the Performance Evaluation section.

We calculate the optimal transmission power against the varying relative distance between the UAV and the RSU, $d[t]$. The results are illustrated in Fig. 2. In Fig. 3, the optimal energy consumption in offloading and local computing is also shown

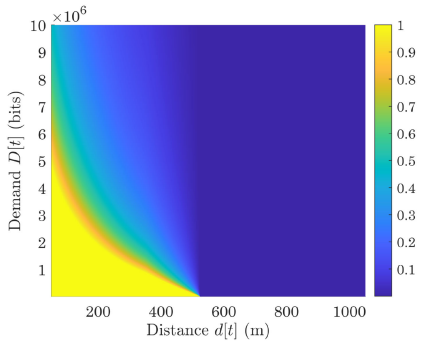

Fig. 4. Optimal offloading ratio against different computation demands and relative distances.

along with varying $d[t]$. From Fig. 2, it is seen that when the relative distance $d[t]$ is lower than about 200 m, the UAV can increase its transmission power up to the maximum level in order to improve the offload rate and get better benefit from remote computing. When $d[t]$ is larger than about 500 m and the UAV is flying away from the RSU, i.e., $d[t]$ arising, the UAV reduces its transmission power down to the minimum level and processes the computation demand mainly by local computing. This means that when the relative distance becomes larger, the UAV cannot get much benefit from computation offloading, and thus it is expected to adopt the local computation rather than the remote computation. Fig. 3 shows that increasing the relative distance will increase the overall energy cost of onboard communication and computation. Besides, from both the figures, it can be seen that the closed-form results are the same as those of the numerical solution, which confirms the proposed theorem. In Fig. 4, we evaluate the optimal ratio of the data volume to be offloaded to the RSU over the total computation demand under different situations. Fig. 4 also shows that the UAV reduces the offloaded data volume by reducing the transmission power when it is far away from the RSU and the optimal offloading ratio decreases with increasing the total computation demand. The result implies that the UAV cannot always save its energy by computation offloading. The optimal offloading strategy depends on the computation demand and the relative distance.

V. NONLINEAR PROGRAMMING WITH GRADIENT PROJECTION METHOD

A. Constrained Nonlinear Optimization Formulation

Following Section IV, we further propose a nonlinear programming method based on the gradient projection mechanism to deal with the complexity of the primal problem $\mathcal{M}_3$. For this goal, we re-arrange the system state equations as follows according to (1)

$$\mathbf{s}[t] = \mathbf{A}^t \mathbf{s}[0] + \sum_{l=0}^{t-1} \mathbf{A}^{t-1-l} \mathbf{B} a[l], t = 1, \ldots, N. \tag{24}$$

Based on (24), we can have for $t = 1, \ldots, N$

$$\begin{cases} x[t] = x[0] + t\Delta\tau v[0] + \sum_{l=0}^{t-1} \left[(t-1-l)\Delta^2\tau + \frac{1}{2}\Delta^2\tau\right] a[l]; \\ v[t] = v[0] + \Delta\tau \sum_{l=0}^{t-1} a[l]. \end{cases} \tag{25}$$

Let the sequence of the velocity controls over time be $\mathbf{a} = [a[0], a[1], \ldots, a[N-1]]^{\mathrm{T}}$, the sequence of the longitudinal positions over time be $\mathbf{x} = [x[1], x[2], \ldots, x[N]]^{\mathrm{T}}$, and the sequence of the longitudinal velocities over time be $\mathbf{v} = [v[1], v[2], \ldots, v[N]]^{\mathrm{T}}$. We can express (25) into a more compact form

$$\begin{cases} \mathbf{x} = \mathbf{b}_x + \mathbf{A}_x \mathbf{a}; \\ \mathbf{v} = \mathbf{b}_v + \mathbf{A}_v \mathbf{a}, \end{cases} \tag{26}$$

where $\mathbf{b}_x$ and $\mathbf{b}_v$ are two $N \times 1$ column vectors given as $\mathbf{b}_x = \mathrm{col}\{x[0] + \Delta\tau v[0], x[0] + 2\Delta\tau v[0], \ldots, x[0] + N\Delta\tau v[0]\}$ and $\mathbf{b}_v = \mathrm{col}\{v[0], \ldots, v[0]\}$, respectively. $\mathbf{A}_x$ and $\mathbf{A}_v$ are $N \times N$ lower triangular matrices, the m-th row and n-th column of which are $[\mathbf{A}_x]_{m,n} = [(m-n)\Delta^2\tau + \frac{1}{2}\Delta^2\tau] \times I_{m \geq n}$ and $[\mathbf{A}_v]_{m,n} = \Delta\tau \times I_{m \geq n}$, respectively. Here, $I_{m \geq n}$ is defined as an indicator function which is equal to 1 if and only if $m \geq n$, and otherwise 0.

Besides, we represent the lower and the upper bounds of the longitudinal position as $\mathbf{x}_{\min} = \mathrm{col}\{x_{\min}, \ldots, x_{\min}\}$ and $\mathbf{x}_{\max} = \mathrm{col}\{x_{\max}, \ldots, x_{\max}\}$, the lower and the upper velocity bounds as $\mathbf{v}_{\min} = \mathrm{col}\{v_{\min}, \ldots, v_{\min}\}$ and $\mathbf{v}_{\max} = \mathrm{col}\{v_{\max}, \ldots, v_{\max}\}$. The lower and the upper bounds of the velocity control sequence are given as $\mathbf{a}_{\min} \triangleq \mathrm{col}\{a_{\min}, \ldots, a_{\min}\}$ and $\mathbf{a}_{\max} \triangleq \mathrm{col}\{a_{\max}, \ldots, a_{\max}\}$, respectively. Thus, the bound constraints of the system states and the controls, i.e., $\mathbf{s}[t] \in \mathcal{S}$ and $a[t] \in [a_{\min}, a_{\max}]$ for $t = 1, \ldots, N$, can be represented as

$$\mathbf{C}\mathbf{a} \geq \mathbf{b}, \tag{27}$$

where $\mathbf{C} \triangleq \mathrm{col}\{\mathbf{A}_x, -\mathbf{A}_x, \mathbf{A}_v, -\mathbf{A}_v, \mathbf{I}_{N\times N}, -\mathbf{I}_{N\times N}\}$ and the constant column vector is $\mathbf{b} \triangleq \mathrm{col}\{\mathbf{x}_{\min} - \mathbf{b}_x, \mathbf{b}_x - \mathbf{x}_{\max}, \mathbf{v}_{\min} - \mathbf{b}_v, \mathbf{b}_v - \mathbf{v}_{\max}, \mathbf{a}_{\min}, -\mathbf{a}_{\max}\}$.

To re-arrange the terminal constraints of the system state, we define a $2 \times N$ matrix $\mathbf{E}$ as

$$\mathbf{E} \triangleq \begin{bmatrix} (N-1)\Delta^2\tau + \frac{1}{2}\Delta^2\tau, & \cdots, & \frac{1}{2}\Delta^2\tau \\ \Delta\tau, & \cdots, & \Delta\tau \end{bmatrix} \tag{28}$$

and a column vector $\mathbf{e}$ as

$$\mathbf{e} \triangleq \begin{bmatrix} x_f - x[0] - N\Delta\tau v[0] \\ v_f - v[0] \end{bmatrix}. \tag{29}$$

Since the UAV is expected to reach the specified terminal state $\mathbf{s}_f$, i.e., $\mathbf{s}[t] = \mathbf{s}_f$ for $t = N$, we can formulate the terminal equality constraint as follows by using (28) and (29)

$$\mathbf{E}\mathbf{a} = \mathbf{e}. \tag{30}$$

To proceed, the primal problem $\mathcal{M}_3$ boils down to the following constrained optimization model, denoted by $\mathcal{M}_4$,

$$\min_{\mathbf{a}} : W(\mathbf{a}) = \sum_{t=0}^{N-1} \left[E_{\mathrm{mob}}[t] + h(d[t])\right]$$
$$\text{s.t.} \begin{cases} \mathbf{C}\mathbf{a} \geq \mathbf{b}; \\ \mathbf{E}\mathbf{a} = \mathbf{e}. \end{cases} \tag{31}$$

B. Gradient Projection-Based Method

Motivated by the linear inequality and equality constraints in the model $\mathcal{M}_4$, we propose a gradient projection-based iterative framework to solve $\mathcal{M}_4$. The key idea is that we are allowed to search the optimal point in the direction of the negative gradient of the objective function $W(\mathbf{a})$ starting at a feasible point $\mathbf{a}$. When the starting point for iterations is located at the bounds of the feasible region, we can project the negative gradient of the objective function into the tangent subspace of those active constraints with respect to the starting point. Thus, we can always guarantee the feasibility of the search direction.

Let $\mathbf{a}_k$ be a feasible point at iteration k, i.e., satisfying $\mathbf{A}\mathbf{a}_k \geq \mathbf{b}$ and $\mathbf{E}\mathbf{a}_k = \mathbf{e}$ in $\mathcal{M}_4$. To construct the projection operator, we first construct a row full rank matrix $\mathbf{M}$, which is composed of those row vectors that are associated with the active constraints. Suppose for $\mathbf{a}_k$ $\mathbf{C}_1\mathbf{a}_k = \mathbf{b}_1$ and $\mathbf{C}_2\mathbf{a}_k > \mathbf{b}_2$, where $\mathbf{C}_1$ and $\mathbf{C}_2$ are two sub-blocks of $\mathbf{C}$, and $\mathbf{b}_1$ and $\mathbf{b}_2$ are two sub-blocks of $\mathbf{b}$, i.e., $\mathbf{C} = \mathrm{col}\{\mathbf{C}_1, \mathbf{C}_2\}$ and $\mathbf{b} = \mathrm{col}\{\mathbf{b}_1, \mathbf{b}_2\}$. We can let $\mathbf{M} = \mathrm{col}\{\mathbf{C}_1, \mathbf{E}\}$. The projection matrix that can be used to project any $N \times 1$ column vector into the null space of $\mathbf{M}$ is

$$\mathbf{P} = \mathbf{I}_{N\times N} - \mathbf{M}^{\mathrm{T}} \left(\mathbf{M}\mathbf{M}^{\mathrm{T}}\right)^{-1} \mathbf{M}, \tag{32}$$

where $\mathbf{I}_{N\times N}$ is an $N \times N$ identity matrix. For $\mathbf{P}$, we have the following results:

Lemma 1: It always holds true that $\mathbf{P} = \mathbf{P}^{\mathrm{T}}$, $\mathbf{P} = \mathbf{P}^2$ and $\mathbf{P}$ is a semidefinite matrix.

Proof: According to the formulation of $\mathbf{P}$ (32), it is easy to verify $\mathbf{P} = \mathbf{P}^{\mathrm{T}}$ and $\mathbf{P} = \mathbf{P}^2$. For any column vector $\mathbf{y} \in \mathbb{R}^{N\times 1}$, we can also observe $\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{P}\mathbf{y} = (\mathbf{P}\mathbf{y})^{\mathrm{T}}(\mathbf{P}\mathbf{y}) = \|\mathbf{P}\mathbf{y}\|^2 \geq 0$. This indicates that $\mathbf{P}$ is a semidefinite matrix. ■

Let the gradient of the objective function $W(\mathbf{a})$ with respect to $\mathbf{a}$ be $\nabla_{\mathbf{a}} W(\mathbf{a})$. When $\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \neq \mathbf{0}$ holds true at the feasible point $\mathbf{a}_k$, we show the projection feasibility of the negative gradient into the null space of $\mathbf{M}$, $\mathbf{Z}_k = -\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k)$, in the following theorem.

Theorem 2: Suppose that $\mathbf{a}_k$ is a feasible point of the model $\mathcal{M}_4$ and $\mathbf{P}$ is given by (32). If $\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \neq \mathbf{0}$ holds true, $\mathbf{Z}_k = -\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k)$ must be a descent feasible direction.

Proof: Based on Lemma 1, we can see

$$\begin{aligned} \left(\nabla_{\mathbf{a}} W(\mathbf{a}_k)\right)^{\mathrm{T}} \mathbf{Z}_k &= -\left(\nabla_{\mathbf{a}} W(\mathbf{a}_k)\right)^{\mathrm{T}} \mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= -\left\|\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k)\right\|^2 \leq 0, \end{aligned} \tag{33}$$

which indicates that $\mathbf{Z}_k$ is a descent direction for iterations.

Besides, it can also be seen that

$$\begin{aligned} \mathbf{M}\mathbf{Z}_k &= -\mathbf{M}\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= -\mathbf{M}\left(\mathbf{I}_{N\times N} - \mathbf{M}^{\mathrm{T}}\left(\mathbf{M}\mathbf{M}^{\mathrm{T}}\right)^{-1}\mathbf{M}\right)\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= (-\mathbf{M} + \mathbf{M})\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}, \end{aligned} \tag{34}$$

which implies that $\mathbf{C}_1\mathbf{Z}_k = \mathbf{0}$ and $\mathbf{E}\mathbf{Z}_k = \mathbf{0}$. According to the condition, $\mathbf{C}_2\mathbf{a}_k > \mathbf{b}_2$, there must exist a positive real number $\lambda_{\max} > 0$, such that for all $\lambda \in [0, \lambda_{\max}]$,

$$\mathbf{C}_2\left(\mathbf{a}_k + \lambda\mathbf{Z}_k\right) \geq \mathbf{b}_2. \tag{35}$$

Since $\mathbf{C}_1\mathbf{Z}_k = \mathbf{0}$ and $\mathbf{C}_1\mathbf{a}_k = \mathbf{b}_1$, we also get

$$\mathbf{C}_1\left(\mathbf{a}_k + \lambda\mathbf{Z}_k\right) = \mathbf{b}_1. \tag{36}$$

Thus, combining (35) and (36) can result in

$$\mathbf{C}\left(\mathbf{a}_k + \lambda\mathbf{Z}_k\right) \geq \mathbf{b}. \tag{37}$$

According to $\mathbf{E}\mathbf{Z}_k = \mathbf{0}$ and $\mathbf{E}\mathbf{a}_k = \mathbf{e}$, the following equation also holds true

$$\mathbf{E}\left(\mathbf{a}_k + \lambda \mathbf{Z}_k\right) = \mathbf{e}. \tag{38}$$

Combining (37) and (38) can prove the feasibility of the direction $\mathbf{Z}_k$. ■

In addition, based on (35) and recalling $\mathbf{b}_2 - \mathbf{C}_2\mathbf{a}_k < \mathbf{0}$, we further derive the closed-form expression for the upper bound of the step size λ, $\lambda_{\max}$, as follows

$$\lambda_{\max} = \begin{cases} \min\left\{ \frac{[\mathbf{b}_2 - \mathbf{C}_2\mathbf{a}_k]_l}{[\mathbf{C}_2\mathbf{Z}_k]_l} : [\mathbf{C}_2\mathbf{Z}_k]_l < 0 \right\}, \mathbf{C}_2\mathbf{Z}_k \ngeq \mathbf{0}; \\ +\infty, \mathbf{C}_2\mathbf{Z}_k \geq \mathbf{0}. \end{cases} \tag{39}$$

In (39), $\mathbf{C}_2\mathbf{Z}_k \ngeq \mathbf{0}$ implies that there exists at least one negative component in the vector $\mathbf{C}_2\mathbf{Z}_k$. In this case, the upper bound of λ is finite, while λ can be any positive real number when all the components in $\mathbf{C}_2\mathbf{Z}_k$ are positive. Accordingly, we can construct a new iterative point, $\mathbf{a}_{k+1}$, as follows

$$\mathbf{a}_{k+1} = \mathbf{a}_k + \lambda \mathbf{Z}_k \tag{40}$$

for $\lambda \in [0, \lambda_{\max}]$ and $\mathbf{Z}_k \neq \mathbf{0}$.

On the other hand, in such a situation where $\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}$, Theorem 2 may not be valid. At this point, we need to construct another projection matrix rather than using $\mathbf{P}$ directly. To be specific, we let $\mathbf{Q}$ be

$$\mathbf{Q} = \left(\mathbf{M}\mathbf{M}^{\mathrm{T}}\right)^{-1}\mathbf{M}\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}, \tag{41}$$

where $\mathbf{w}_1$ and $\mathbf{w}_2$ are two sub-blocks of $\mathbf{W}$, whose row indexes correspond to those of $\mathbf{C}_1$ and $\mathbf{E}$, respectively. Then, we obtain the following result.

Theorem 3: Suppose that $\mathbf{a}_k$ is a feasible point of the model $\mathcal{M}_4$ and $\mathbf{P}$ is given by (32). If $\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}$ holds true, it follows that

i) if $\mathbf{w}_1 \leq \mathbf{0}$, $\mathbf{a}_k$ must be a local optimal point (a KKT point) for the model $\mathcal{M}_4$;

ii) if there exist negative components in $\mathbf{w}_1$, a new coefficient matrix $\widehat{\mathbf{C}}_1$ can be constructed by removing the rows whose indexes correspond to those with the negative components in $\mathbf{w}_1$, and a new projection matrix $\widehat{\mathbf{P}}$ can be defined as

$$\widehat{\mathbf{P}} = \mathbf{I}_{N\times N} - \widehat{\mathbf{M}}^{\mathrm{T}}\left(\widehat{\mathbf{M}}\widehat{\mathbf{M}}^{\mathrm{T}}\right)^{-1}\widehat{\mathbf{M}}, \tag{42}$$

where $\widehat{\mathbf{M}} \triangleq \mathrm{col}\{\widehat{\mathbf{C}}_1, \mathbf{E}\}$. Let $\widehat{\mathbf{Z}}_k = -\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k)$. $\widehat{\mathbf{Z}}_k$ must be a descent feasible direction.

Proof: In the first situation (i), since $\mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}$, we can have

$$\begin{aligned} \mathbf{0} = \mathbf{P}\nabla_{\mathbf{a}} W(\mathbf{a}_k) &= \left[\mathbf{I}_{N\times N} - \mathbf{M}^{\mathrm{T}}\left(\mathbf{M}\mathbf{M}^{\mathrm{T}}\right)^{-1}\mathbf{M}\right]\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= \nabla_{\mathbf{a}} W(\mathbf{a}_k) - \mathbf{M}^{\mathrm{T}}\left(\mathbf{M}\mathbf{M}^{\mathrm{T}}\right)^{-1}\mathbf{M}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= \nabla_{\mathbf{a}} W(\mathbf{a}_k) - \mathbf{C}_1^{\mathrm{T}}\mathbf{w}_1 - \mathbf{E}^{\mathrm{T}}\mathbf{w}_2. \end{aligned} \tag{43}$$

According to $\mathbf{w}_1 \geq \mathbf{0}$, (43) is in coincidence with the KKT conditions in which $\mathbf{w}_1$ can be treated as the nonnegative Lagrangian multipliers corresponding to the inequality constraints. Thus, $\mathbf{a}_k$ is a KKT point.

In the second situation (ii), let $[\mathbf{w}_1]_l < 0$ be one of the negative component in $\mathbf{w}_1$. We first prove by contradiction that $\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \neq \mathbf{0}$ must hold true. Suppose $\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}$. We can see

$$\begin{aligned} \mathbf{0} = \widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) &= \left[\mathbf{I}_{N\times N} - \widehat{\mathbf{M}}^{\mathrm{T}}\left(\widehat{\mathbf{M}}\widehat{\mathbf{M}}^{\mathrm{T}}\right)^{-1}\widehat{\mathbf{M}}\right]\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= \nabla_{\mathbf{a}} W(\mathbf{a}_k) - \widehat{\mathbf{M}}^{\mathrm{T}}\widehat{\mathbf{Q}}, \end{aligned} \tag{44}$$

where $\widehat{\mathbf{Q}}$ is defined as $\widehat{\mathbf{Q}} = (\widehat{\mathbf{M}}\widehat{\mathbf{M}}^{\mathrm{T}})^{-1}\widehat{\mathbf{M}}\nabla_{\mathbf{a}} W(\mathbf{a}_k)$. Let the l-th row of $\mathbf{C}_1$ be $[\mathbf{C}_1]_l$. Since

$$\begin{aligned} \mathbf{C}_1^{\mathrm{T}}\mathbf{w}_1 + \mathbf{E}^{\mathrm{T}}\mathbf{w}_2 &= \widehat{\mathbf{C}}_1^{\mathrm{T}}\widehat{\mathbf{w}}_1 + [\mathbf{w}_1]_l [\mathbf{C}_1]_l^{\mathrm{T}} + \mathbf{E}^{\mathrm{T}}\mathbf{w}_2 \\ &= \widehat{\mathbf{M}}^{\mathrm{T}}\overline{\mathbf{Q}} + [\mathbf{w}_1]_l [\mathbf{C}_1]_l^{\mathrm{T}}, \end{aligned} \tag{45}$$

where $\widehat{\mathbf{w}}_1$ is the vector composed of the components in $\mathbf{w}_1$ except $[\mathbf{w}_1]_l$. $\overline{\mathbf{Q}}$ is defined as $\overline{\mathbf{Q}} \triangleq \mathrm{col}\{\widehat{\mathbf{w}}_1, \mathbf{w}_2\}$. Now, substituting (45) into (43) can yield

$$\mathbf{0} = \nabla_{\mathbf{a}} W(\mathbf{a}_k) - \widehat{\mathbf{M}}^{\mathrm{T}}\overline{\mathbf{Q}} - [\mathbf{w}_1]_l [\mathbf{C}_1]_l^{\mathrm{T}}. \tag{46}$$

Furthermore, subtracting (46) from (44) can obtain

$$\mathbf{0} = \widehat{\mathbf{M}}^{\mathrm{T}}\left(\overline{\mathbf{Q}} - \widehat{\mathbf{Q}}\right) + [\mathbf{w}_1]_l [\mathbf{C}_1]_l^{\mathrm{T}}. \tag{47}$$

Notice that the right-side term of (47) is indeed a linear combination of the rows in the matrix $\mathbf{M}$. Since the combination coefficient $[\mathbf{w}_1]_l \neq 0$, (47) indicates that the row vectors of $\mathbf{M}$ are linearly dependent. This conclusion is contradictory to the fact that $\mathbf{M}$ is full row rank. At this point, a contradiction arises under the hypothesis of $\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}$. It is confirmed that $\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \neq \mathbf{0}$.

Since $\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \neq \mathbf{0}$, it can be seen that

$$\begin{aligned} \nabla_{\mathbf{a}} W(\mathbf{a}_k)^{\mathrm{T}}\widehat{\mathbf{Z}}_k &= -\nabla_{\mathbf{a}} W(\mathbf{a}_k)^{\mathrm{T}}\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= -\left\|\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k)\right\|^2 \leq 0, \end{aligned} \tag{48}$$

which means that $\widehat{\mathbf{Z}}_k$ is a descent direction. Besides, we can also observe

$$\begin{aligned} \widehat{\mathbf{M}}\widehat{\mathbf{Z}}_k &= -\widehat{\mathbf{M}}\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= -\widehat{\mathbf{M}}\left[\mathbf{I}_{N\times N} - \widehat{\mathbf{M}}^{\mathrm{T}}\left(\widehat{\mathbf{M}}\widehat{\mathbf{M}}^{\mathrm{T}}\right)^{-1}\widehat{\mathbf{M}}\right]\nabla_{\mathbf{a}} W(\mathbf{a}_k) \\ &= -\left[\widehat{\mathbf{M}} - \widehat{\mathbf{M}}\right]\nabla_{\mathbf{a}} W(\mathbf{a}_k) = \mathbf{0}. \end{aligned} \tag{49}$$

As $\widehat{\mathbf{M}} = \mathrm{col}\{\widehat{\mathbf{C}}_1, \mathbf{E}\}$, (49) indicates

$$\widehat{\mathbf{C}}_1\widehat{\mathbf{Z}}_k = \mathbf{0},\ \mathbf{E}\widehat{\mathbf{Z}}_k = \mathbf{0}. \tag{50}$$

Multiplying both sides of (46) with $[\mathbf{C}_1]_l\widehat{\mathbf{P}}$ can yield

$$0 = [\mathbf{C}_1]_l\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k) - [\mathbf{C}_1]_l\widehat{\mathbf{P}}\widehat{\mathbf{M}}^{\mathrm{T}}\overline{\mathbf{Q}} - [\mathbf{w}_1]_l [\mathbf{C}_1]_l\widehat{\mathbf{P}}[\mathbf{C}_1]_l^{\mathrm{T}}. \tag{51}$$

Recalling $\widehat{\mathbf{P}}\widehat{\mathbf{M}}^{\mathrm{T}} = \mathbf{0}$ and $\widehat{\mathbf{Z}}_k = -\widehat{\mathbf{P}}\nabla_{\mathbf{a}} W(\mathbf{a}_k)$, we further get

$$0 = [\mathbf{C}_1]_l\widehat{\mathbf{Z}}_k + [\mathbf{w}_1]_l [\mathbf{C}_1]_l\widehat{\mathbf{P}}[\mathbf{C}_1]_l^{\mathrm{T}}. \tag{52}$$

Since $\widehat{\mathbf{P}}$ is semidefinite and $[\mathbf{w}_1]_l < 0$, it can be found from (52) that

$$[\mathbf{C}_1]_l\widehat{\mathbf{Z}}_k = -[\mathbf{w}_1]_l [\mathbf{C}_1]_l\widehat{\mathbf{P}}[\mathbf{C}_1]_l^{\mathrm{T}} \geq 0. \tag{53}$$

Algorithm 1: Gradient Projection-based Algorithm.

/* Initialization */
Select a feasible initial point $\mathbf{a}_1$ and set $k = 1$.
/* Do iterations. */
while $k \leq K_{\max}$ **do**
  /* Decompose the coefficient matrix. */
  Decompose $\mathbf{C}$ into $\mathbf{C} = \mathrm{col}\{\mathbf{C}_1, \mathbf{C}_2\}$ and $\mathbf{b}$ into $\mathbf{b} = \mathrm{col}\{\mathbf{b}_1, \mathbf{b}_2\}$ such that $\mathbf{C}_1\mathbf{a}_k = \mathbf{b}_1$ and $\mathbf{C}_2\mathbf{a}_k > \mathbf{b}_2$.
  while TRUE **do**
    /* Construct the active matrix. */
    Construct $\mathbf{M} = \mathrm{col}\{\mathbf{C}_1, \mathbf{E}\}$.
    /* Construct the projection matrix. */
    Construct $\mathbf{P} = \mathbf{I}_{N\times N} - \mathbf{M}^{\mathrm{T}}(\mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1}\mathbf{M}$.
    /* Construct the search direction. */
    Set $\mathbf{Z}_k = -\mathbf{P}\nabla_{\mathbf{a}}W(\mathbf{a}_k)$.
    /* Check and do modification. */
    if $\|\mathbf{Z}_k\|_\infty \leq \epsilon$ **then**
      Construct $\mathbf{Q} = (\mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1}\mathbf{M}\nabla_{\mathbf{a}}W(\mathbf{a}_k)$.
      Decompose $\mathbf{Q}$ into $\mathbf{Q} = \mathrm{col}\{\mathbf{w}_1, \mathbf{w}_2\}$.
      if $\mathbf{w}_1 \geq \mathbf{0}$ **then**
        /* Obtain the KKT point. */
        Break the inner and outer While-loops.
        Return $\mathbf{a}_k$.
      else
        /* Modify the sub-block $\mathbf{C}_1$. */
        Select $[\mathbf{w}_1]_l < 0$.
        Remove $[\mathbf{C}_1]_l$ from $\mathbf{C}_1$.
    else
      Break the inner While-loop.
  Determine the upper bound of the step size $\lambda_{\max}$.
  Determine an optimal step size λ_{opt} by

$$\lambda_{\mathrm{opt}} = \underset{\lambda\in[0,\lambda_{\max}]}{\mathrm{argmin}}\ W(\mathbf{a}_k + \lambda\mathbf{Z}_k). \tag{55}$$

  Set $\mathbf{a}_{k+1} = \mathbf{a}_k + \lambda_{\mathrm{opt}}\mathbf{Z}_k$.
  Update $k = k + 1$.
Return $\mathbf{a}_k$.

Combining (53) and (50) results in

$$\mathbf{C}_1\widehat{\mathbf{Z}}_k \geq \mathbf{0},\ \mathbf{E}\widehat{\mathbf{Z}}_k = \mathbf{0}. \tag{54}$$

Therefore, following the same logic in Theorem 2 and (55), we can also prove that $\widehat{\mathbf{Z}}_k$ guarantees the feasibility of the descent direction. ■

Algorithm 1 summaries the proposed gradient projection based iterative algorithm based on Theorems 2 and 3. In Algorithm 1, we denote the maximum number of the iterations by $K_{\max}$ and the tolerant numerical error by $\epsilon \in (0, 1)$.

C. Complexity Analysis

From Algorithm 1, it can be observed that the computation complexity of the proposed gradient projection-based method is dominated by two parts, the construction of a feasible projection matrix $\mathbf{P}$ and the line search for obtaining an optimal descent step size λ_{opt}. For the program of constructing the projection matrix, the program might need to do the matrix modification several times as in the IF-THEN block in the worst-case situation. The major computational complexity of the projection matrix modification lies in calculating the matrix $\mathbf{Q}$. Let M be the row number of $\mathbf{M}$. We can have $M \leq (6N + 2)$ since $\mathbf{M} = \mathrm{col}\{\mathbf{C}_1, \mathbf{E}\}$ where $\mathbf{C}_1$ is a sub-block of $\mathbf{C} \in \mathbb{R}^{6N\times N}$ and $\mathbf{E} \in \mathbb{E}^{2\times N}$. Note that the computational complexity of the matrix transpose $\mathbf{M}^{\mathrm{T}}$, the matrix inversion $(\mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1}$, the matrix-by-matrix multiplications $(\mathbf{M}\mathbf{M}^{\mathrm{T}})$ and $(\mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1}\mathbf{M}$, the matrix-by-vector multiplication $(\mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1}\mathbf{M}\nabla_{\mathbf{a}}W(\mathbf{a}_k)$ is $\mathcal{O}(MN)$, $\mathcal{O}(M^3)$, $\mathcal{O}(2\,M^2\,N)$, and $\mathcal{O}(MN)$, respectively. Thus, the projection matrix modification has the complexity of $\mathcal{O}(MN + M^3 + 2\,M^2N + MN)$ in the worst-case situation. Similarly, the computational complexity in calculating $\mathbf{P}$ is in the order of $\mathcal{O}(2MN + M^3 + 2\,M^2N + N^2M + N^2)$. Hence, in the worst-case situation where it needs to do the matrix modification up to M times in the inner WHILE-DO loop of Algorithm V-B, the construction of a feasible projection matrix has the complexity of $\mathcal{O}(M(2MN + M^3 + 2\,M^2N + N^2M + N^2))$. Besides, for the line search in solving the optimal step size λ_{opt}, the complexity to achieve ϵ-optimality is in the order of $\mathcal{O}(\epsilon^{-2})$ by using Newton's first-order method [41]. With $K_{\max}$ iterations in the outer WHILE-DO loop of Algorithm V-B, the overall worst-case computational complexity of the proposed method is approximately in the order of $\mathcal{O}(K_{\max}M(2MN + M^3 + 2\,M^2N + N^2M + N^2) + K_{\max}\epsilon^{-2})$. At this point, the polynomial complexity can be efficiently realized by the proposed method. It is noticed that the worst-case computational complexity above is an upper bound of the actual complexity of the proposed method, and the sparsity of the active matrix $\mathbf{M}$ and some others can considerably reduce the computational complexity in actual application situations.

Additionally, it is remarked that the proposed method simply needs to allocate the memory for storing the decision vector $\mathbf{a}_k \in \mathbb{R}^{N\times 1}$, the constraint coefficient matrices $\mathbf{C} \in \mathbb{R}^{6N\times N}$ and $\mathbf{E} \in \mathbb{R}^{2\times N}$, the column vector $\mathbf{b} \in \mathbb{R}^{6N\times 1}$, the gradient vector $\nabla_{\mathbf{a}}W(\mathbf{a}_k) \in \mathbb{R}^{N\times 1}$, the active and projection matrices $\mathbf{M} \in \mathbb{R}^{M\times N}$ and $\mathbf{P} \in \mathbb{R}^{N\times N}$, and some others like $\mathbf{Z}_k \in \mathbb{R}^{N\times 1}$, $\mathbf{Q} \in \mathbb{R}^{M\times 1}$, $\mathbf{I}_{N\times N}$. The preallocated space for all the decisions and parameters involved in our algorithm will not change during iterations. According to the sizes of the above decisions and parameters depending on M and N, the space complexity is in the order of $\mathcal{O}(11N + 8\,N^2 + MN + M)$. Besides, all the constant-size parameters, including $\mathbf{e}, \lambda_{\mathrm{opt}}, \lambda_{\max}, k, K_{\max}$, have the space complexity of $\mathcal{O}(1)$. In actual computer programming, the preallocated space of the matrix-type or vector-type parameters can be locally reused or shared by using memory management techniques, e.g., the well-known Copy-on-Write (COW) technique, such that the space complexity can be practically reduced.

VI. Performance Evaluation

In this section, we evaluate the performance of the proposed joint optimization method for a UAV in an A2G-based computation offloading scenario, where the UAV is assumed to be

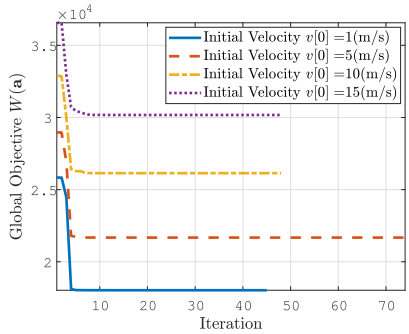

Fig. 5. The convergence of the proposed method.

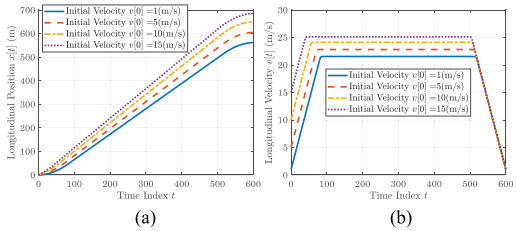

Fig. 6. Optimal longitudinal trajectory and velocity. (a) The optimal longitudinal trajectory. (b) The optimal longitudinal velocity.

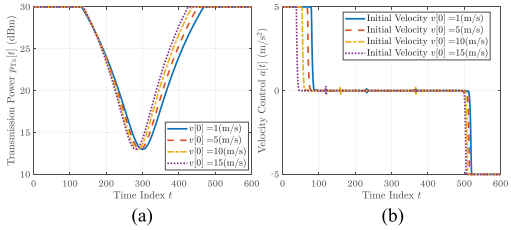

Fig. 7. Optimal power and velocity controls. (a) The optimal power control. (b) The optimal velocity control.

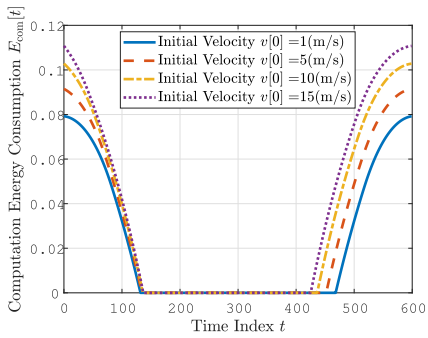

Fig. 8. The variation of the computation energy consumption.

cruising over a road traffic network and cooperate with a RSU (i.e., a network edge) to process computation-intensive tasks like real-time aerial sensing data. The UAV needs to jointly optimize its communication power, computation offloading rate, and longitudinal acceleration to maximize its energy efficiency meanwhile satisfying cruising motion constraints. In the simulation scenario, the convergence performance of the proposed algorithm is first demonstrated and then its global performance is compared with several other methods. The basic simulation parameters on the UAV's kinematics, physical-layer communication, and local computation are given in Table II and also adopted here to conduct the experiments.

A. Convergence Analysis

To validate the convergence of the proposed method, we consider that a rotary-wing UAV flies along a longitudinal trajectory with specified initial and terminal velocities and positions. The flight duration is set to 30 seconds, and the computation demand is set to $D[t] = 2 \times 10^6$ bits for all t. Besides, we restrict the terminal velocity of the UAV to $v[t_f] = 1$ (m/s), while varying its initial velocity $v[0]$ to simulate different situations, i.e., setting $v[0] \in \{1, 5, 10, 15\}$ (m/s). The convergence performance of the proposed joint optimization method with different initial velocities is illustrated in Fig. 5. As can be seen, the steady value of the global objective function that the algorithm converges to is different from each other since the initial velocity is varied, and more mobility-related energy is consumed with a larger initial velocity. Additionally, the proposed method is shown to converge effectively after a few iterations, even when the initial velocity is set to different situations. That is, the first-order optimality for convergence can be satisfied within only 10 iterations in all the situations.

In Fig. 6, the optimal longitudinal trajectory and velocity of the UAV during flying are illustrated. From both Fig. 6.(a) and Fig. 6.(b), it can be seen that the UAV first accelerates up to an optimal cruising velocity and then flies with the constant velocity. When it approaches the terminal position, the UAV decelerates to the specified terminal velocity. In addition, the optimal cruising velocity of the UAV with a different initial velocity is also different. In Fig. 7, we show the optimal transmission power for computation offloading over the A2G wireless link and the optimal velocity control during flying. It can be observed that when the relative distance between the UAV and the RSU is relatively large, i.e., the UAV is approaching the space above the RSU in the beginning phase (the acceleration phase) or flying far away from the RSU in the last phase (the deceleration phase), the UAV has to adopt the maximum transmission power to offload its computation tasks as many as possible. When the relative distance is reduced to a certain range, i.e., t ranging within (150 420), the power consumption also decreases. The main reason is that the quality of the A2G communication link becomes better when the UAV is closer to the RSU. The optimal velocity control curves in Fig. 7.(b) confirm that the UAV can always satisfy the control constraint and show that the UAV adopts the maximum acceleration and the maximum deceleration for achieving the optimal cruising velocity and the specified terminal velocity, respectively.

To illustrate the offloading behavior of the UAV, we further evaluate the computation energy cost related to the local computation in Fig. 8. Interestingly, we can see that the computation energy consumption $E_{\mathrm{com}}[t]$ is reduced to zero when $t \in (150, 420)$. This result indicates that the UAV is able to offload its whole computation tasks to the RSU and fully employs the remote computing mode when it is flying within a certain range sufficiently closed to the RSU. But when the UAV is flying away

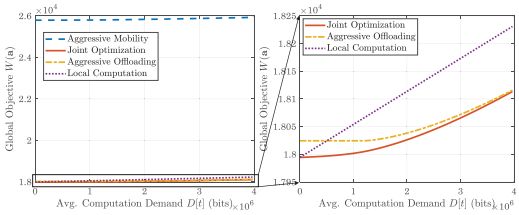

Fig. 9. Global performance comparison under different computation demands.

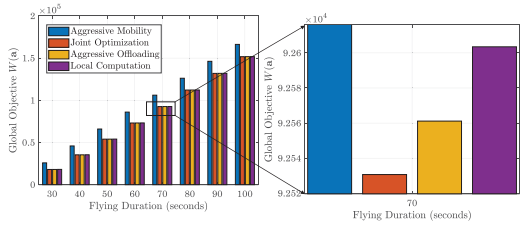

Fig. 10. Global performance comparison under different flying durations.

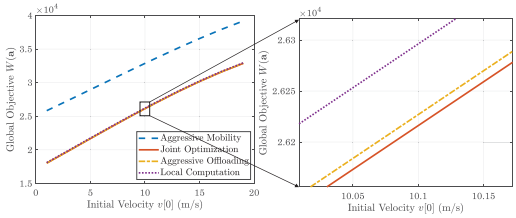

Fig. 11. Global performance comparison under different initial velocities.

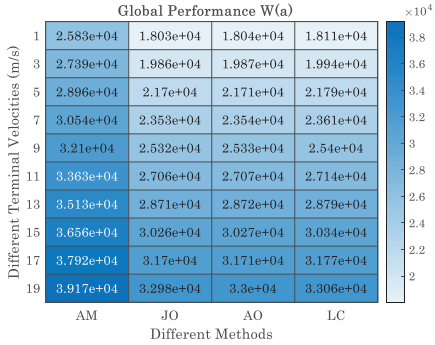

Fig. 12. Global performance comparison under different terminal velocities.

from the RSU, for instance, at $t \in (500\,600)$, the transmission power is increased and only a part of its computation tasks can be offloaded to the RSU. To sum up, the results from the figures above confirm the convergence and feasibility of the proposed method and show that it enables the UAV to adapt its mobility, communication and computation during flying.

B. Performance Comparison

To demonstrate the advantage of the proposed method, we further carry out a series of experiments for performance comparison under different situations. To be specific, we compare our method with three other conventional methods that are based on the aggressive mobility mechanism, the aggressive offloading mechanism, and the local computation mechanism, respectively. For the sake of simplicity, the numerical results corresponding to our proposed method are marked with "Joint Optimization" or "JO," while these other methods are marked with "Aggressive Mobility" or "AM," "Aggressive Offloading" or "AO," and "Local Computation" or "LC," respectively. With the aggressive mobility mechanism, the UAV will accelerate up to the maximum allowed velocity with the maximum acceleration and adopt the maximum velocity as its cruising velocity. In this situation, the trajectory of the UAV is not optimized. With the aggressive offloading mechanism, the UAV will adopt its maximum allowed transmission power for computation offloading all the time and it considers to jointly optimize the mobility and computation energy consumption. With the local computation mechanism, the UAV will only optimize its mobility and fully adopt the local computation mode to process its whole computation demand.

Fig. 9 shows the global performance, $W(\mathbf{a})$, of the compared methods under different computation demands, $D[t]$. Specifically, $D[t]$ is varied from 10^3 bits to 4×10^6 bits. The flying duration is set to 30 seconds, while the initial and the terminal velocities are specified as $v[0] = v[t_f] = 1$ (m/s). From Fig. 9, we can observe that the aggressive mobility mechanism without optimization has the highest energy consumption in all the demand situations, while the global energy consumption levels of the other methods are of the same order of magnitude. However, when comparing our proposed method with the aggressive offloading and the local computation methods, we can find that the proposed joint optimization method can achieve the lowest energy consumption than the other two optimization methods under different computation demands. The global energy consumption of our method is about 14.3502 (Joule) and about 77.0035 (Joule) less than those of the aggressive offloading and the local computation methods on average, respectively. Besides, Fig. 9 demonstrates that when the computation demand is relatively large, for instance, $D[t] \geq 3 \times 10^6$ bits, the performance of the aggressive offloading mechanism tends to be closed to that of our proposed method. This is because the UAV needs to increase its transmission power for computation offloading in order to benefit from remote computing when the computation demand is increased substantially.

We also compare our method with the others under different flying durations. The numerical results are shown in Fig. 10. We see that increasing the flying duration can improve the whole energy cost as expected. The aggressive mobility performs worst in all the situations, since it does not exploit any optimization mechanism. Nonetheless, our method always achieves the lowest global energy cost. For instance, the right sub-figure in Fig. 10 shows the details on the numerical results of the other methods except the aggressive mobility method that are obtained under the flying duration of 70 seconds. In all the duration settings, our proposed joint optimization can reduce the energy consumption of about 29.7433 (Joule) and 75.8256 (Joule) on average when compared to the aggressive offloading and the local computation methods.

Finally, to demonstrate the effects of the initial and the terminal velocities on the global performance, we vary the initial velocity $v[0]$ and the terminal velocity $v[t_f]$ from 1 (m/s) to 19 (m/s) with an incremental step-size of 2 (m/s), respectively, and compare the different methods. Fig. 11 and Fig. 12

demonstrate their numerical results. It is shown that increasing the initial velocity or the terminal velocity can increase the mobility-related energy cost. Both the figures also show that our proposed method achieves the largest reduction in the global energy consumption among these compared methods. In particular, in Fig. 11, our method can reduce the energy consumption by about 10.9126 (Joule) and 80.7331 (Joule) on average when compared to the aggressive offloading and the local computation methods, respectively. Similarly, in Fig. 12, our method achieves the energy saving of about 10.9027 (Joule) and 80.6835 (Joule) on average, compared to these two optimization methods, respectively. These results confirm that jointly optimizing the mobility, communication and computation of the UAV is more advantageous than only planning the trajectory of the UAV and only optimizing either the computation offloading or the transmission power.

VII. Conclusion and Future Work

In this paper, we have investigated the problem of the UAV-oriented trajectory optimization, power control, and computation offloading in an air-ground-integrated network. We have proposed a joint mobility, communication and computation optimization method to maximize the energy efficiency of a UAV. To tackle the optimization complexity, we have proposed a primal decomposition approach and then theoretically derived a closed-form formulation for the optimal power control. Furthermore, the original joint optimization model is then converted to another tractable form with the closed-form power control. Based on the transformed model, we have proposed a novel gradient projection-based method and theoretically proved its feasibility and convergence. Comparative simulation results have validated the effectiveness and advantage of the proposed method. Our work can offer a valuable approach of joint control and optimization modeling for the UAV in air-ground-integrated scenarios. In the future, we will focus on the design of a joint optimization framework for vehicle-infrastructure-UAV-integrated systems.

References

[1] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[2] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5 G and beyond," *Proc. IEEE IRE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.

[3] Y. Zhou, N. Cheng, N. Lu, and X. S. Shen, "Multi-uav-aided networks: Aerial-ground cooperative vehicular networking architecture," *IEEE Veh. Technol. Mag.*, vol. 10, no. 4, pp. 36–44, Dec. 2015.

[4] N. Cheng *et al.*, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.

[5] J. Zhao, F. Gao, G. Ding, T. Zhang, W. Jia, and A. Nallanathan, "Integrating communications and control for UAV systems: Opportunities and challenges," *IEEE Access*, vol. 6, pp. 67519–67527, 2018.

[6] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[7] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5 G network edge cloud architecture and orchestration," *IEEE Commun. Surv. Tut.*, vol. 19, no. 3, pp. 1657–1681, Jul.-Sep. 2017.

[8] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surv. Tut.*, vol. 19, no. 4, pp. 2322–2358, Oct.-Dec. 2017.

[9] Z. Ullah, F. Al-Turjman, and L. Mostarda, "Cognition in UAV-aided 5 G and beyond communications: A survey," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 872–891, Sep. 2020.

[10] Y. Shih, W. Chung, A. Pang, T. Chiu, and H. Wei, "Enabling low-latency applications in fog-radio access networks," *IEEE Netw.*, vol. 31, no. 1, pp. 52–58, Jan. 2017.

[11] T. Chiu, A. Pang, W. Chung, and J. Zhang, "Latency-driven fog cooperation approach in fog radio access networks," *IEEE Trans. Serv. Comput.*, vol. 12, no. 5, pp. 698–711, Sep. 2019.

[12] X. Cao, F. Wang, J. Xu, R. Zhang, and S. Cui, "Joint computation and communication cooperation for energy-efficient mobile edge computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4188–4200, Jun. 2019.

[13] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "UAV relay in vanets against smart jamming with reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4087–4097, May 2018.

[14] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3747–3760, Jun. 2017.

[15] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a UAV-mounted cloudlet: Optimization of bit allocation and path planning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2049–2063, Mar. 2018.

[16] 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.

[17] H. Mei, K. Wang, D. Zhou, and K. Yang, "Joint trajectory-task-cache optimization in UAV-enabled mobile edge networks for cyber-physical system," *IEEE Access*, vol. 7, pp. 156476–156488, 2019.

[18] J. Zhang *et al.*, "Stochastic computation offloading and trajectory scheduling for UAV-assisted mobile edge computing," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3688–3699, Apr. 2019.

[19] Y. Liu, K. Xiong, Q. Ni, P. Fan, and K. B. Letaief, "UAV-assisted wireless powered cooperative mobile edge computing: Joint offloading, CPU control and trajectory optimization," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 2777–2790, Apr. 2020.

[20] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, Sep. 2018.

[21] M. Li, N. Cheng, J. Gao, Y. Wang, L. Zhao, and X. Shen, "Energy-efficient UAV-assisted mobile edge computing: Resource allocation and trajectory optimization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3424–3438, Mar. 2020.

[22] F. Cheng *et al.*, "UAV trajectory optimization for data offloading at the edge of multiple cells," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6732–6736, Jul. 2018.

[23] H. Tang, Q. Wu, J. Xu, W. Chen, and B. Li, "A novel alternative optimization method for joint power and trajectory design in UAV-enabled wireless network," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11358–11362, Nov. 2019.

[24] S. Yin, S. Zhao, Y. Zhao, and F. R. Yu, "Intelligent trajectory design in UAV-aided communications with reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8227–8231, Aug. 2019.

[25] X. Zhou, Q. Wu, S. Yan, F. Shu, and J. Li, "UAV-enabled secure communications: Joint trajectory and transmit power optimization," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4069–4073, Apr. 2019.

[26] L. Deng, G. Wu, J. Fu, Y. Zhang, and Y. Yang, "Joint resource allocation and trajectory control for UAV-enabled vehicular communications," *IEEE Access*, vol. 7, pp. 132806–132815, 2019.

[27] F. Zhou, Y. Wu, R. Q. Hu, and Y. Qian, "Computation rate maximization in UAV-enabled wireless-powered mobile-edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1927–1941, Sep. 2018.

[28] 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.

[29] X. Li, H. Yao, J. Wang, S. Wu, C. Jiang, and Y. Qian, "Rechargeable multi-UAV aided seamless coverage for QoS-guaranteed IoT networks," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10902–10914, Dec. 2019.

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

[31] Y. Cai, F. Cui, Q. Shi, M. Zhao, and G. Y. Li, "Dual-UAV-enabled secure communications: Joint trajectory design and user scheduling," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1972–1985, Sep. 2018.

[32] J. Ji, K. Zhu, D. Niyato, and R. Wang, "Joint cache placement, flight trajectory and transmission power optimization for multi-UAV assisted wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5389–5403, Aug. 2020.
[33] X. Hu, K. Wong, K. Yang, and Z. Zheng, "UAV-assisted relaying and edge computing: Scheduling and trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4738–4752, Oct. 2019.
[34] S. Chou, A. Pang, and Y. Yu, "Energy-aware 3D unmanned aerial vehicle deployment for network throughput optimization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 563–578, Jan. 2020.
[35] H. Hashida, Y. Kawamoto, and N. Kato, "Intelligent reflecting surface placement optimization in air-ground communication networks toward 6 G," *IEEE Wireless Commun.*, vol. 27, no. 6, pp. 146–151, Dec. 2020.
[36] W. Chen, J. Liu, and H. Guo, "Achieving robust and efficient consensus for large-scale drone swarm," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15867–15879, Dec. 2020.
[37] A. Filippone, "Flight performance of fixed and rotary wing aircraft," Reston, VA 20191-5807, USA: Amer. Inst. Aeronautics and Astronautics (AIAA), 2006. [Online]. Available: https://doi.org/10.2514/4.478390
[38] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. IEEE Glob. Commun. Conf.*, Dec. 2014, pp. 2898–2904.
[39] 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.
[40] W. Yuan and K. Nahrstedt, "Energy-efficient CPU scheduling for multimedia applications," *ACM Trans. Comput. Syst.*, vol. 24, no. 3, pp. 292–331, Aug. 2006.
[41] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.

Jianshan Zhou received the B.Sc., M.Sc., and Ph.D. degrees in traffic information engineering and control from Beihang University, Beijing, China, in 2013, 2016 and 2020, respectively. From 2017 to 2018, he was a Visiting Research Fellow with the School of Informatics and Engineering, University of Sussex, Brighton, U.K. He is currently a Postdoctoral Research Fellow supported by the Zhuoyue Program of Beihang University, and is or was the Technical Program Session Chair with the IEEE EDGE 2020 and the Youth Editorial Board Member of the Unmanned Systems Technology. He is the author or coauthor of more than 20 international scientific publications. His research interests include the modeling and optimization of vehicular communication networks and air–ground cooperative networks, the analysis and control of connected autonomous vehicles, and intelligent transportation systems. He was the recipient of the First Prize in the Science and Technology Award from the China Intelligent Transportation Systems Association in 2017, the First Prize in the Innovation and Development Award from the China Association of Productivity Promotion Centers in 2020, the National Scholarships in 2017 and 2019, the Outstanding Top-Ten Ph.D. Candidate Prize from Beihang University in 2018, and the Outstanding China-SAE Doctoral Dissertation Award in 2020.

Daxin Tian (Senior Member, IEEE) is currently a Professor with the School of Transportation Science and Engineering, Beihang University, Beijing, China. His current research interests include mobile computing, intelligent transportation systems, vehicular ad hoc networks, and swarm intelligent.

Zhengguo Sheng (Senior Member, IEEE) is currently a Senior Lecturer with the Department of Engineering and Design, University of Sussex, Brighton, U.K. He has authored more than 50 international conference and journal papers. His current research interests include IoT/M2M, vehicular communications, and edge or cloud computing.

Xuting Duan is currently an Assistant Professor with the School of Transportation Science and Engineering, Beihang University, Beijing, China. His current research focuses on vehicular ad hoc networks.

Xuemin Shen (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is currently a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests include network resource management, wireless network security, Internet of Things, 5G and beyond, and vehicular ad hoc and sensor networks. He is a registered Professional Engineer of Ontario, Canada, aFellow of the Engineering Institute of Canada, the Canadian Academy of Engineering, and the Royal Society of Canada, a Chinese Academy of Engineering Foreign Member, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and the IEEE Communications Society.

He was the Technical Program Committee Chair or Co-Chair of the IEEE Globecom'16, the IEEE Infocom'14, the IEEE VTC'10 Fall, and the IEEE Globecom'07, and the Chair of the IEEE Communications Society Technical Committee on Wireless Communications. He is the elected IEEE Communications Society Vice President for Technical and Educational Activities, the Vice President for Publications, a Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a Member of the IEEE ComSoc Fellow Selection Committee. He was or is the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL, the IEEE NETWORK, the IET *Communications*, and the *Peer-to-Peer Networking and Applications*. He was the recipient of the R.A. Fessenden Award in 2019 from IEEE Canada, the Award of Merit from the Federation of Chinese Canadian Professionals, Ontario, in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, the Joseph LoCicero Award in 2015, the Education Award in 2017 from the IEEE Communications Society, the Technical Recognition Award from the Wireless Communications Technical Committee in 2019 and AHSN Technical Committee in 2013, the Excellent Graduate Supervision Award in 2006 from the University of Waterloo, and the Premier's Research Excellence Award in 2003 from the Province of Ontario, Canada.