IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE VEHICULAR TECHNOLOGY SOCIETY SECTION

Received July 25, 2021, accepted August 15, 2021, date of publication August 20, 2021, date of current version August 30, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3106495*

# Energy Optimization in Massive MIMO UAV-Aided MEC-Enabled Vehicular Networks

**EMMANOUEL T. MICHAILIDIS[1], NIKOLAOS I. MIRIDAKIS[2], (Senior Member, IEEE), ANGELOS MICHALAS[3], (Member, IEEE), EMMANOUIL SKONDRAS[4], DIMITRIOS J. VERGADOS[5], AND DIMITRIOS D. VERGADOS[4], (Senior Member, IEEE)**

[1]Department of Electrical and Electronics Engineering, University of West Attica, 12241 Egaleo, Greece
[2]Department of Informatics and Computer Engineering, University of West Attica, 12243 Egaleo, Greece
[3]Department of Electrical and Computer Engineering, University of Western Macedonia, 50131 Kozani, Greece
[4]Department of Informatics, University of Piraeus, 18534 Piraeus, Greece
[5]Department of Informatics, University of Western Macedonia, 52100 Kastoria, Greece

Corresponding author: Emmanouel T. Michailidis (emichail@uniwa.gr)

This work was supported by the Research Committee of the University of Western Macedonia, Kozani, Greece.

**ABSTRACT** This paper presents a novel unmanned aerial vehicle (UAV)-aided mobile edge computing (MEC) architecture for vehicular networks. It is considered that the vehicles should complete latency-critical computation-intensive tasks either locally with on-board computation units or by offloading part of their tasks to road side units (RSUs) with collocated MEC servers. In this direction, a hovering UAV can serve as an aerial RSU (ARSU) for task processing or act as an aerial relay and further offload the computation tasks to a ground RSU (GRSU). To significantly reduce the delay during data offloading and downloading, this architecture relies on the benefits of line-of-sight (LoS) massive multiple-input–multiple-output (MIMO). Therefore, it is considered that the vehicles, the ARSU, and the GRSU employ large-scale antennas. A three-dimensional (3-D) geometrical representation of the MEC-enabled network is introduced and an optimization method is proposed that minimizes the computation-based and communication-based weighted total energy consumption (WTEC) of vehicles and ARSU subject to transmit power allocation, task allocation, and time slot scheduling. The results verify the theoretical derivations, emphasize on the effectiveness of the LoS massive MIMO transmission, and provide useful engineering insights.

**INDEX TERMS** Computation offloading, energy efficiency, massive multiple-input multiple-output (MIMO), mobile edge computing (MEC), unmanned aerial vehicle (UAV), vehicular networks.

## I. INTRODUCTION

With the emergence of the big data era at vehicular networks, the Internet of Vehicles (IoV) paradigm, and the vehicular-to-everything (V2X) information interaction, a vast number of connected automobile terminals equipped with computation and multi-communication units will pave the path for novel services [1]. The next wave of applications, including augmented reality (AR), ultra-high-quality video streaming, and autonomous driving, is expected to reach the limits of current technologies and pose strict requirements in terms of computation, latency, and throughput. For locally intra-vehicle processed applications, a large amount of energy is consumed, which in turn reduces the driving range of energy-limited electric vehicles [2]. On the other hand, it is often infeasible for resource-constrained vehicles to timely handle computation-intensive tasks. To maintain the energy consumption at a low level, while meeting critical latency demands, partly or fully task offloading to mobile edge computing (MEC) servers has been suggested [3], [4]. In this regard, road side units (RSUs) along roads and in proximity to the vehicles can facilitate the provision of MEC services.

The associate editor coordinating the review of this manuscript and approving it for publication was Jie Gao.

### A. BACKGROUND

As the highly dynamic and random nature of MEC-enabled vehicular networks drastically affects the performance of data offloading, the computation overhead was minimized in [5], whereas a reliability-oriented stochastic optimization model was presented in [6] to maximize the lower bound of the expected reliability during offloading. The effective collaboration of cloud computing and MEC in vehicular networks was addressed in [7] and the computation offloading decisions were optimized through a

game-theoretic approach. Also, a vehicle edge computing (VEC) network architecture with the vehicles acting as MEC servers was proposed in [8] and the joint optimization of computing offloading and resource allocation was handled via deep reinforcement learning (DRL). A software-defined networking (SDN)-based and fiber-wireless (FiWi)-enhanced load-balancing task offloading policy was introduced in [9] to minimize the processing delay. Moreover, a computation offloading protocol that exploits geo-location information was developed in [10] enabling efficient and reliable data retrieval in VEC environments with hybrid vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) links. In order to minimize the total network delay, an edge intelligence empowered IoV framework was constructed in [11] and an online algorithm relying on Lyapunov optimization was designed to manage computation offloading and content caching.

Despite such promising computing capabilities, attaining ubiquitous connectivity and sufficient radio coverage between vehicles and MEC servers is challenging, since ground RSUs (GRSUs) often struggle in areas with obstacles and highly mobile and disperse nodes. In this respect, hovering aerial RSUs (ARSUs) based on unmanned aerial vehicles (UAVs) can fly over connected vehicles and effectively mitigate shadowing and blockage effects thus maintaining line-of-sight (LoS) propagation [12]–[14]. Most of current work on UAV-aided MEC-enabled networks has focused on energy-aware solutions both from ground users (GUs) and UAV perspective. In [15], a UAV was deployed to assist an access point (AP) to provide MEC services to GUs and an algorithm that minimizes the energy consumption was proposed. By adopting similar setups, the maximum delay [16], sum power [17], task completion time [18], average latency [19], and computation efficiency [20] were also optimized. A non-orthogonal multiple access (NOMA) scheme was studied in [21], whereas an edge-cloud system supporting virtualized network functions (VNFs) was proposed in [22]. Beyond the deterministic binary and partial task offloading, the concept of stochastic offloading was studied in [23]. In [24], the resource allocation and UAV's trajectory were optimized for a social IoV (SIoV) scenario, whereas an optimization framework for SDN-enabled computation offloading was proposed in [25] to minimize the execution time of computation tasks of vehicles, under energy and quality of service (QoS) constraints. The advantage of employing UAVs as MEC servers in Cyber-Physical Systems (CPSs) was outlined in [26] and the three-dimensional (3-D) UAV's trajectory was sub-optimally optimized to extend the UAV's endurance. In [27], the use of UAV-mounted edge nodes in Long Range Wide Area Networks (LoRaWANs) for disaster management was investigated. Wireless power transfer (WPT) was also introduced to prolong network's operation time and optimization problems were formulated to maximize the sum completed task-input bits [28] and the UAV's required energy [29]. An Internet of Things (IoT) scenario was examined in [30], while a multi-UAV-based MEC system was proposed in [31]. Moreover, a UAV-aided MEC-enabled IoV architecture that relies on reconfigurable intelligent surface (RIS) units was proposed in [32].

On another front, massive multiple-input–multiple-output (MIMO) technology has recently received unprecedented attention as a key enabler for increased spectral and energy efficiency, drastically reduced round-trip latency, and support of highly-intensive computation tasks for a vast number of connected users [33]. Beyond the conventional massive MIMO, where only the base stations employ multiple antennas [33], large-scale antenna configurations in vehicles [34]–[38] and UAVs [39], [40] were previously considered, without focusing on MEC applications. Besides, the double-sided massive MIMO concept [35], [37], [39], [41]–[44] was also envisioned to offer notably increased beamforming gain and counterbalance the signal propagation losses. Nevertheless, there exist some practical barriers (e.g., energy consumption and hardware complexity) towards the implementation of large-scale antennas, especially when fully digital beamforming is used. Recently, less complex hybrid analog-digital [42] and fully analog [42], [43] beamforming architectures were studied. Also, low-complexity transceivers with perfect [41] and partial [44] channel state information (CSI) requirements were proposed. In the near future, we believe that devices with reduced energy consumption and practical transceivers for challenging applications will be developed.

To the best of the authors' knowledge, the area of massive MIMO UAV-based MEC networks is unexplored and only single-antenna MEC solutions have been previously studied. However, these solutions cannot properly capture the characteristics of the massive MIMO channel. In the context of terrestrial MIMO MEC networks, a multi-antenna NOMA architecture that enables multi-user computation offloading over the same time/frequency resources was proposed in [45]. Also, the optimization of energy consumption and maximum delay, under perfect and imperfect CSI estimation, was studied for MIMO [46] and massive MIMO [47] systems. Moreover, single-cell [48] and multi-cell [49] MEC networks that enable the simultaneous offloading of multiple APs were previously presented. The benefits of combining massive MIMO and millimeter wave (mmWave) frequencies in wireless local area networks (WLANs) with MEC were underlined in [50], whereas a cell-free system consisting of multiple single/multi-antenna APs with MEC servers and a central cloud server was described in [51]. Notwithstanding, these works are improper for UAV-based MEC networks, since the UAVs fly in a 3-D space and above rooftops leading to peculiar link geometry and especial mobility characteristics. To reconcile these challenging issues, newer network architectures are indispensable.

### B. CONTRIBUTION

Motivated by the aforementioned observations, we investigate a massive MIMO UAV-aided MEC-enabled vehicular

network. The major contributions of this paper are summarized as follows:

- **A novel dual-MEC network architecture** is proposed, where a UAV operates as an ARSU equipped with a MEC server and also as an intermediate decode-and-forward (DF) aerial relay between vehicles and a GRSU. This architecture trades on the LoS massive MIMO transmission and the efficient use of all computing resources. Moreover, a MEC computation offloading and downloading protocol is presented with distinct operation phases. Based on this protocol, partial offloading is applied to obtain a trade-off between energy consumption and delay.
- In order to unlock the full potential of massive MIMO, we propose **the concept of triple-sided LoS massive MIMO**, as an extension of the single-sided and double-sided massive MIMO. Therefore, it is considered that the vehicles, the ARSU, and the GRSU employ two-dimensional (2-D) uniform rectangular planar arrays (URPAs) with a large number of antenna elements.
- **Realistic 3-D placement and mobility modeling** of the vehicles, the ARSU, the GRSU, and the URPAs is proposed. Also, position, distance, and velocity vectors are used to accommodate the geometrical representation of the proposed network architecture and construct the massive MIMO channel matrices.
- **A multi-variable optimization problem is formulated** that intends to minimize the weighted total energy consumption (WTEC) from both the vehicles and ARSU perspective and prolong their lifetime as major network segments. The Lagrange dual method is leveraged to derive closed-form solutions for the transmit power allocation, time slot scheduling, and computation bits allocation. A subgradient-based algorithm is also constructed to expedite the optimization process. The results depict the total computation-based and communication-based delay (TCCD) and WTEC, point out the advantages of massive MIMO transmission, and validate the effectiveness of the optimization procedure.

### C. STRUCTURE

The remainder of this paper is organized as follows. Section II introduces the system model and outlines the geometrical, mobility, and channel characteristics. Section III presents the computation offloading and downloading model. Section IV formulates the optimization problem and derives its solution, whereas Section V provides numerical results. Finally, conclusions and future directions are given in Section VI.

## II. SYSTEM MODEL

Consider a triple-sided massive MIMO MEC-enabled vehicular network that facilitates the computing offloading of $K$ vehicles moving along a unidirectional road segment. Each vehicle has a latency-sensitive and bit-wise-independent computation task that can be executed partly *locally* with the on-board computing processor and partly *remotely* by computation offloading to network MEC servers. In this direction, a fixed GRSU with sufficiently powerful computation capacity and grid power supply is situated along the road. It is assumed that the $k$-th vehicle cannot directly communicate with the GRSU owing to signal blockage or severe shadowing. Thus, an ARSU is employed to enable vehicle-to-ARSU (V2U) networking thus facilitating the MEC services. Contrary to GRSU, the ARSU has certain computing and energy limitations that depend on its type, weight, and battery size. To preserve its energy resources, the ARSU can determine the portion of tasks that can locally process and then act as an aerial relay forwarding the remaining part of the vehicles' offloaded tasks to GRSU. By employing a sufficiently large data buffer, the ARSU can separately store the offloaded data and computation results.

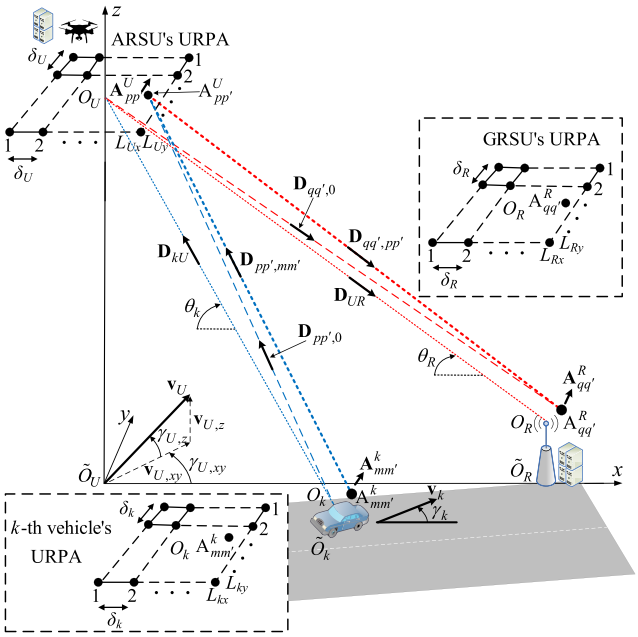


**FIGURE 1.** The geometrical characteristics of the proposed triple-sided massive MIMO UAV-aided MEC-enabled vehicular network, where an ARSU assists the k-th vehicle execute its offloaded computing task and also acts as an aerial relay to further transmit part of this task to a GRSU for computing.

### A. GEOMETRICAL CHARACTERISTICS AND MOBILITY MODEL

Fig. 1 illustrates the 3-D geometrical characteristics of the proposed dual-MEC network architecture. To aid our analysis, the subscripts $k$, $U$, and $R$, where $k \in \{1, 2, \ldots K\}$, are associated with the $k$-th vehicle, ARSU, and GRSU, respectively. The $(x, y, z)$ axes designate the global coordinate system (GCS), which controls the position of each network segment, with the projection $\tilde{O}_U$ of ARSU's array center $O_U$ positioned at the origin $(x = 0, y = 0, z = 0)$, where $\tilde{O}$ denotes the projection of a point $O$ onto the $xy$ plane. Since the height of $k$-th vehicle (GRSU) is relatively low compared

to that of ARSU, $O_k$ ($O_R$) is almost identical to $\tilde{O}_k$ $\left(\tilde{O}_R\right)$. It is considered that the $k$-th vehicle, ARSU, and GRSU employ URPAs. These URPAs are defined by the local coordinate system (LCS), the origins of which are at URPAs' centers. More specifically, the $k$-th vehicle is equipped with an URPA with $L_k = L_{kx} \times L_{ky}$ antennas spanning $L_{kx}$ rows along the $x$-axis and $L_{ky}$ columns along the $y$-axis of the LCS with equal inter-element spacing $\delta_k$, where $L_k \gg 1$. The position of URPAs in the GCS is specified by the transformation from GCS to LCS. Based on [52], [53], we consider an arbitrary 3-D rotation of the LCS with respect to the GCS given by the angles $\beta_x \in [-\pi/2, \pi/2]$ (slant angle), $\beta_y \in [-\pi/2, \pi/2]$ (downtilt angle), and $\beta_z \in [0, 2\pi]$ (bearing angle). These angles designate a 3-D counterclockwise rotation around the reference coordinate axes of GCS and can also describe ARSU's roll, pitch, and yaw, respectively. Then, a sequence of rotations assigns the URPAs' orientation with respect to GCS. For the position vector $\mathbf{A}^k_{mm'}$ of antenna element $A^k_{mm'}$ with $m = 1, 2, \ldots, L_{kx}, m' = 1, 2, \ldots, L_{ky}$, we obtain $\mathbf{A}^k_{mm'} = \mathbf{R}\mathbf{A}'^k_{mm'}$, where

$$\mathbf{R} = \mathbf{R}_X(\beta_x)\,\mathbf{R}_Y\left(\beta_y\right)\mathbf{R}_Z(\beta_z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta_x & -\sin\beta_x \\ 0 & \sin\beta_x & \cos\beta_x \end{pmatrix} \times \begin{pmatrix} \cos\beta_y & 0 & \sin\beta_y \\ 0 & 1 & 0 \\ -\sin\beta_y & 0 & \cos\beta_y \end{pmatrix} \begin{pmatrix} \cos\beta_z & -\sin\beta_z & 0 \\ \sin\beta_z & \cos\beta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

is the rotation matrix, $\mathbf{A}'^k_{mm'} = \left[x'^k_{mm'}, y'^k_{mm'}, 0\right]^T$ is the position vector of antenna element $A^k_{mm'}$ in the LCS,

$$x'^k_{mm'} = \begin{cases} -\left(L_{kx} - 2m + 1\right)\delta_k/2, & m < \left(L_{kx}+1\right)/2 \\ \left(L_{kx} - 2m + 1\right)\delta_k/2, & m \geq \left(L_{kx}+1\right)/2 \end{cases}, \quad (2)$$

$$y'^k_{mm'} = \begin{cases} -\left(L_{ky} - 2m' + 1\right)\delta_k/2, & m' < \left(L_{ky}+1\right)/2 \\ \left(L_{ky} - 2m' + 1\right)\delta_k/2, & m' \geq \left(L_{ky}+1\right)/2 \end{cases}, \quad (3)$$

and $(\cdot)^T$ denotes the transpose operation. By replacing the indices, the position vectors $\mathbf{A}^U_{pp'}$ and $\mathbf{A}^R_{qq'}$ of antenna elements $A^U_{pp'}$ and $A^R_{qq'}$, respectively, are similarly defined.

Based on Fig. 1, $\mathbf{D}_{pp',mm'} = \mathbf{D}_{pp',0} - \mathbf{A}^k_{mm'}$ represents the distance vector between $\mathrm{A}^k_{mm'}$ and $\mathrm{A}^U_{pp'}$, $\mathbf{D}_{pp',0} = \mathbf{D}_{kU} + \mathbf{A}^U_{pp'}$ is the distance vector between the $k$-vehicle array center and $\mathrm{A}^U_{pp'}$, $\mathbf{D}_{kU} = [h_U/\tan\theta_k, 0, h_U]^T$ is the distance vector between the $k$-vehicle and ARSU array centers, $h_U$ is the altitude of the ARSU antenna array, and $\theta_k$ is the elevation angle of ARSU relative to $O_k$. Note that the other distance vectors can be similarly defined. Moreover, $\mathbf{v}_k = v_k[\cos\gamma_k, \sin\gamma_k, 0]^T$ and $\mathbf{v}_U = v_U\left[\cos\gamma_{U,xy}\cos\gamma_{U,z}, \sin\gamma_{U,xy}\cos\gamma_{U,z}, \sin\gamma_{U,z}\right]^T$ are the velocity vectors of the $k$-vehicle and the ARSU, respectively, $v_k$ ($v_U$) is the instantaneous velocity of the $k$-vehicle (ARSU), $\gamma_k$ $\left(\gamma_{U,xy}\right)$ is the azimuth angle that controls the moving direction of the $k$-vehicle (ARSU) in the azimuth domain, and $\gamma_{U,z}$ is the elevation angle that characterizes possible rising, diving, and hovering operations of the ARSU.

The total energy of ARSU is limited. Thus, the flying period is restrained by $T_U$. For convenience, adequately small constant $\tau$ is used to divide $T_U$ into $N$ time slots. In each time slot $n \in \{1, 2, \ldots N\}$, the $k$-th vehicle and the ARSU can be considered to be static, whereas their antenna position vectors are updated, respectively, as $\mathbf{A}^k_{mm'}[n+1] = \mathbf{A}^k_{mm'}[n] + \mathbf{v}_k[n]\tau$ and $\mathbf{A}^U_{pp'}[n+1] = \mathbf{A}^U_{pp'}[n] + \mathbf{v}_U[n]\tau$ with $\mathbf{A}^k_{mm'}(t) = \mathbf{A}^k_{mm'}(\tau n) = \mathbf{A}^k_{mm'}[n]$ and $\mathbf{A}^U_{pp'}(t) = \mathbf{A}^U_{pp'}(\tau n) = \mathbf{A}^U_{pp'}[n]$. Note that the distance vectors are also updated accordingly.

Based on [26], we model the energy consumption during flight in the $n$-th time slot for a fixed-wing ARSU as

$$E_{fl,fw}[n] = \tau\left(c_1\left\|\mathbf{v}_{U,xy}[n]\right\|^3 + \frac{c_2}{\left\|\mathbf{v}_{U,xy}[n]\right\|} + c_3\left\|\mathbf{v}_{U,z}[n]\right\|\right), \quad (4)$$

where $c_1$ and $c_2$ are constants depending on the ARSU's weight, wing area, and air density, $c_3$ is a constant associated with ARSU's descending/ascending, $\mathbf{v}_{U,xy}[n]$ and $\mathbf{v}_{U,z}[n]$ are the horizontal and vertical ARSU velocity vector, respectively, with $\mathbf{v}_U[n] = \mathbf{v}_{U,xy}[n] + \mathbf{v}_{U,z}[n]$, and $\|\cdot\|$ is the Euclidean norm. For a rotary-wing ARSU, the energy consumption during flight can be modelled as [26]

$$\begin{aligned} E_{fl,rw}[n] = \tau\Bigg(&P_0\left(1 + \frac{3\left\|\mathbf{v}_{U,xy}[n]\right\|^2}{v^2_{\text{tip}}}\right) \\ &+ \frac{1}{2}d_r s\rho G\left\|\mathbf{v}_{U,xy}[n]\right\|^3 \\ &+ P_1\sqrt{\sqrt{1 + \frac{\left\|\mathbf{v}_{U,xy}[n]\right\|^4}{4v_0^2}} - \frac{\left\|\mathbf{v}_{U,xy}[n]\right\|^2}{2v_0^2}} \\ &+ P_2\left\|\mathbf{v}_{U,z}[n]\right\|\Bigg), \end{aligned} \quad (5)$$

where $P_0$ and $P_1$ describe the blade profile power and induced power, respectively, $P_2$ controls the descending/ascending power, $v_{\text{tip}}$ is the tip speed of rotor blade, $v_0$ is the mean rotor induced velocity, $d_r$ is the fuselage drag ratio, $s$ is the rotor solidity, $\rho$ is the air density, and $G$ is the rotor disc area.

### B. WIRELESS TRANSMISSION MODEL

In each time slot the $k$-th vehicle and the ARSU move over a small distance. Thus, the channel coefficients are keeping unchanged and can be estimated using uplink and downlink orthogonal pilot sequences (prior to data transmission) at the start of each time slot [54]. Overall, the channel is described by a series of channel snapshots for different placement of the $k$-th vehicle and the ARSU in each time slot. In practice, the communication links between vehicles (ARSU) and ARSU (GRSU) may be vulnerable to fading and blockage effects due to large scattering objects (e.g., buildings, poles, trees, hills or human bodies) in the propagation environment. However, the flying altitude of ARSU is much higher

than that of the ground nodes, whereas the ARSU can also approach these nodes. Thus, the ARSU has a significant chance of achieving line-of-sight (LoS) communication links with respect to vehicles and GRSU, thus enhancing coverage and connectivity. In this paper, it is assumed that the V2U, ARSU to GRSU (U2R), GRSU to ARSU (R2U), and ARSU to $k$-vehicle (U2V) channels are dominated by LoS links within the short frame $T_U$, as indicated by recent measurements in several propagation environments [55]–[58] and also assumed by previous work (e.g., [15]–[18], [20], [24], [30]). As the CSI related to the LoS propagation changes much slower than the CSI related to the multipath propagation, perfect CSI is considered at all ends of the links. Based on [39], [52], the channel coefficient between antenna elements $\mathrm{A}^{k}_{mm'}$ and $\mathrm{A}^{U}_{pp'}$ can be written as follows

$$g_{pp',mm'}[n] = \sqrt{\beta_{kU}[n]}\, h_{pp',mm'}[n], \tag{6}$$

where

$$\beta_{kU}[n] = \beta_0 \left\|\mathbf{D}_{kU}[n]\right\|^{-\alpha}, \tag{7}$$

$$h_{pp',mm'}[n] = \exp\left[j\left(2\pi f_{pp',mm'}[n] + \varphi_{pp',mm'}[n]\right)\right], \tag{8}$$

$$f_{pp',mm'}[n] = \frac{\left\langle \mathbf{D}_{kU}[n] + \mathbf{A}^{U}_{pp'}[n] - \mathbf{A}^{k}_{mm'}[n],\, \mathbf{v}_k[n] - \mathbf{v}_U[n] \right\rangle}{\lambda \left\| \mathbf{D}_{kU}[n] + \mathbf{A}^{U}_{pp'}[n] - \mathbf{A}^{k}_{mm'}[n] \right\|}, \tag{9}$$

$$\varphi_{pp',mm'}[n] = \frac{2\pi \left\| \mathbf{D}_{kU}[n] + \mathbf{A}^{U}_{pp'}[n] - \mathbf{A}^{k}_{mm'}[n] \right\|}{\lambda}, \tag{10}$$

$\beta_0$ is the channel gain at a reference distance $d_0 = 1$ m, $\alpha$ is the path-loss exponent, $\langle \cdot, \cdot \rangle$ is the inner product operator, and $\lambda$ is the carrier wavelength. As mentioned in Section II.A, the position vectors $\mathbf{A}^{k}_{mm'}$ and $\mathbf{A}^{U}_{pp'}$ of antenna elements $A^{k}_{mm'}$ and $A^{U}_{pp'}$, respectively, can be defined by a sequence of rotations using (1)-(3). Since $\|\mathbf{D}_{kU}[n]\| \gg \{\delta_k, \delta_U\}$, the path-loss parameter $\beta_{kU}[n]$ is invariable over $L_k$ and $L_U$. The massive MIMO channel between the $k$-vehicle and ARSU in the $n$-th time slot can be described by the channel matrix $\mathbf{G}_{kU}[n] = \left[g_{pp',mm'}[n]\right]_{L_U \times L_k} \in \mathbb{C}^{L_U \times L_k}$, which can be expressed as

$$\mathbf{G}_{kU}[n] = \mathbf{H}_{kU}[n]\, \mathbf{F}^{1/2}_{kU}[n], \tag{11}$$

where $\mathbf{H}_{kU}[n] = \left[h_{pp',mm'}[n]\right]_{L_U \times L_k}$ is a $L_U \times L_k$ matrix and $\mathbf{F}_{kU}[n] = [\beta_{kU}[n]]_{L_k \times L_k}$ is a $L_k \times L_k$ diagonal matrix. The channel matrices for the other links can be similarly defined by using (11) and properly replacing the indices. However, in the U2R and R2U cases, the channel coefficients are only affected by the movement of ARSU, since GRSU is static.

The achievable rate of the V2U massive MIMO channel can be expressed as [50]

$$r_{kU}[n] = B \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p^{\mathrm{off}}_k[n]\, \lambda^2_{kU,l}[n]}{B N_0 L_k}\right), \tag{12}$$

where $B$ is the allocated bandwidth, $p^{\mathrm{off}}_k[n]$ is the transmit power of the $k$-th vehicle, $N_0$ is the variance of the additive white Gaussian noise (AWGN) at ARSU, and $\left\{\lambda_{kU,l}[n]\right\}_{l=1}^{\min(L_k,L_U)}$ are the singular values of $\mathbf{G}_{kU}[n]$. Without loss of generality, we assume that $N_0$ is the variance of the AWGN at any network node. Using (12) and appropriately replacing the indices, the achievable rates $r_{UR}$, $r_{RU}$, and $r_{Uk}$ of the U2R, R2U, and U2V massive MIMO channels, respectively, can be obtained.

## III. COMPUTATION OFFLOADING AND DONWLOADING MODEL

We define $l_k = \{c_k, b_k, \xi_k\}$ the computation task of the $k$-th vehicle, where $c_k > 0$ is the number of required central processing unit (CPU) cycles per bit, $b_k$ is the task-input data size (in bits), $c_k b_k$ is the total required number of CPU cycles, and $\xi_k$ is the proportionality ratio between offloaded data and computed results. The maximum CPU frequency at the $k$-th vehicle and the ARSU is denoted as $f_{k,\max}$ and $f_{U,\max}$, respectively, with $f_{k,\max} < f_{U,\max}$, whereas $c_U > 0$ denotes the number of required CPU cycles per bit at the ARSU. Since the computational resources of the $k$-th vehicle are limited, more computing power is required to accomplish its task within the maximum allowable latency (task deadline) $\eta_k \leq T_U$. In this regard, partial task offloading is exploited. Hence, the $k$-th vehicle offloads to ARSU and GRSU (via relaying) part of its task in each time slot. The computation task at the $k$-vehicle in a given time slot is partitioned as

$$b_k[n] = b_{k,l}[n] + b_{k,U}[n] + b_{k,R}[n] \geq b_{k,\min}[n], \tag{13}$$

where $b_{k,l}[n]$, $b_{k,U}[n]$, and $b_{k,R}[n]$ are the computation bits allocated for local computing, offloading to ARSU for computing, and offloading to GRSU for computing via ARSU, respectively. Besides, $b_{k,\min}$ are the minimum task bits that should be periodically completed in each time slot. Note that the case that $\eta_k = T_U$ is only considered in this paper $\forall k$.

### A. TRANSMISSION DELAY AND COMPUTATION DELAY

In order to implement the data offloading and downloading processes, while avoiding interference among the vehicles, each time slot is fairly divided into $K$ equal durations $\{\tau_k[n]\}_{k=1}^{K}$ with $\sum_{k=1}^{K} \tau_k[n] = \tau$. Contrary to previous work on conventional UAV-based MEC networks with single antennas (e.g., [15], [30]), the use of massive MIMO can meaningfully shorten the transmission time thus rendering the duration of data offloading, computing, and downloading comparable. Therefore, a protocol for data transmission and data computation with five distinct operation phases is proposed. As shown in Fig. 2, $\tau^{\mathrm{off}}_k[n] = b_{k,UR}[n]/r_{kU}$ is the transmission time for offloading the bits $b_{k,UR}[n] = b_{k,U}[n] + b_{k,R}[n]$ from the $k$-th vehicle to ARSU (Phase 1); $\tau^{\mathrm{off}}_{k,U}[n] = b_{k,R}[n]/r_{UR}$ is the transmission time for offloading the bits $b_{k,R}[n]$ from ARSU to GRSU via relaying[1] (Phase 2); $\tau_{k,cU}[n] = c_U b_{k,U}[n]/f_{U,\max}$ is the

[1] In this paper half-duplex relaying is considered. Although full-duplex relaying is generally more efficient and allows the simultaneous data transmission and reception at the same time instance, loop interference cancellation mechanisms and increased signal processing capabilities at the ARSU are required, which raise the complexity and cost.

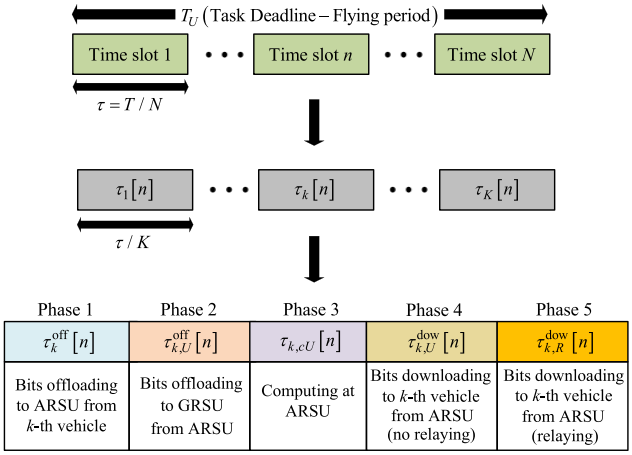


**FIGURE 2.** The data transmission and data computation protocol.

computation delay at the ARSU (Phase 3); $\tau_{k,U}^{\text{dow}}[n] = \tilde{b}_{k,U}[n]/r_{Uk}$ is the transmission time for downloading the bits $\tilde{b}_{k,U}[n] = \xi_k b_{k,U}[n]$ from the ARSU to $k$-th vehicle (Phase 4); and $\tau_{k,R}^{\text{dow}}[n] = \tilde{b}_{k,R}[n]/r_{Uk}$ is the transmission time for downloading the bits $\tilde{b}_{k,R}[n] = \xi_k b_{k,R}[n]$ from the ARSU to $k$-th vehicle (Phase 5). Note that the downloading transmission time for the R2U channel is omitted, since there are no transmit power constraints at the GRSU side. Also, the computing time at the GRSU is negligible, whereas the decision time for task partitioning is very short compared to the entire latency and can be neglected as well. It is assumed that the $k$-th vehicle can simultaneously carry out local computing and bits offloading, whereas the local computation delay $\tau_{k,cl}[n] = c_k b_{k,l}[n]/f_{k,\max}$ can span a time slot $\tau$. Overall, the following time allocation constraints should be satisfied:

$$0 \leq \left\{ \tau_k^{\text{off}}[n], \tau_{k,U}^{\text{off}}[n], \tau_{k,cU}[n], \tau_{k,U}^{\text{dow}}[n], \tau_{k,R}^{\text{dow}}[n], \frac{\tau_{k,cl}[n]}{K} \right\} \leq \frac{\tau}{K}, \tag{14}$$

$$\tau_k^{\text{off}}[n] + \tau_{k,U}^{\text{off}}[n] + \tau_{k,cU}[n] + \tau_{k,U}^{\text{dow}}[n] + \tau_{k,R}^{\text{dow}}[n] \leq \frac{\tau}{K}. \tag{15}$$

As the bits offloading (or bits downloading) cannot outreach the rate capabilities of the massive MIMO channels, we obtain the following constraints:

$$b_{k,UR}[n] \leq \tau_k^{\text{off}}[n] r_{kU}[n], \tag{16}$$

$$b_{k,R}[n] \leq \tau_{k,U}^{\text{off}}[n] r_{UR}[n], \tag{17}$$

$$\tilde{b}_{k,U}[n] \leq \tau_{k,U}^{\text{dow}}[n] r_{Uk}[n], \tag{18}$$

$$\tilde{b}_{k,R}[n] \leq \tau_{k,R}^{\text{dow}}[n] r_{Uk}[n]. \tag{19}$$

### B. ENERGY CONSUMPTION

The energy consumed for data offloading (or computed data downloading) at Phases 1, 2, 4, and 5, respectively, can be expressed as

$$E_k^{\text{off}}[n] = p_k^{\text{off}}[n] \tau_k^{\text{off}}[n] \leq P_{k,\max}^{\text{off}} \tau_k^{\text{off}}[n], \tag{20}$$

$$E_{k,U}^{\text{off}}[n] = p_{k,U}^{\text{off}}[n] \tau_{k,U}^{\text{off}}[n] \leq P_{k,A,\max}^{\text{off}} \tau_{k,U}^{\text{off}}[n], \tag{21}$$

$$E_{k,U}^{\text{dow}}[n] = p_{k,U}^{\text{dow}}[n] \tau_{k,U}^{\text{dow}}[n] \leq P_{k,U,\max}^{\text{dow}} \tau_{k,U}^{\text{dow}}[n], \tag{22}$$

$$E_{k,R}^{\text{dow}}[n] = p_{k,R}^{\text{dow}}[n] \tau_{k,R}^{\text{dow}}[n] \leq P_{k,R,\max}^{\text{dow}} \tau_{k,R}^{\text{dow}}[n], \tag{23}$$

where $p_{k,U}^{\text{off}}[n], p_{k,U}^{\text{dow}}[n]$, and $p_{k,R}^{\text{dow}}[n]$ are the transmit powers at Phases 2, 4, and 5, respectively, and $P_{k,\max}^{\text{off}}, P_{k,U,\max}^{\text{off}}, P_{k,U,\max}^{\text{dow}}$, and $P_{k,R,\max}^{\text{dow}}$ are the maximum transmit powers at Phases 1, 2, 4, and, 5, respectively. In each time slot, the energy consumption for local intra-vehicle and ARSU computing can be, respectively, expressed as [59]

$$E_{k,cl}[n] = P_{k,cl}\tau_{k,cl}[n] \equiv \kappa_k c_k^3 \left(b_{k,l}[n]\right)^3 \tau^{-2}, \tag{24}$$

$$E_{k,cU}[n] = P_{k,cU}\tau_{k,cU}[n] \equiv \kappa_U c_U^3 K^2 \left(b_{k,U}[n]\right)^3 \tau^{-2}, \tag{25}$$

where $P_{k,cl} = \kappa_k f_{k,\max}^3$ and $P_{k,cU} = \kappa_U f_{U,\max}^3$ is the power consumption of the CPU at the $k$-th vehicle and at the ARSU, respectively [59], and $\kappa_k > 0$ and $\kappa_U > 0$ are the effective capacitance coefficient that rely on the chip architecture at the $k$-th vehicle and at the ARSU, respectively.

## IV. PROBLEM FORMULATION AND OPTIMIZATION

Considering a dual-MEC UAV-aided massive MIMO vehicular network with energy-limited battery-powered vehicles and ARSU, a novel multi-variable optimization problem is formulated to minimize the WTEC, including the computation-based and communication-based energy consumption. This problem is explicitly subjected to physical layer parameters, such as the transmit power allocation, as well as the massive MIMO uplink and downlink data rates. Also, this problem accounts for time slot scheduling and task allocation. Although the energy consumption during flight is significant and is directly controlled by the ARSU's trajectory, it is excluded from the optimization process. More specifically, this paper considers a pre-determined ARSU's trajectory and the optimization of the trajectory, which can further decrease the WTEC, is left as a future work. Towards this end, the optimization problem is formulated as

$$(\text{P1}): \min_{\mathbf{B},\mathbf{P},\boldsymbol{\tau}} \sum_{n=1}^{N} \left( \left( \sum_{k=1}^{K} w_k E_k[n] \right) + w_U E_U[n] \right) \tag{26a}$$

$$\text{s.t. } b_{k,l}[n] + b_{k,U}[n] + b_{k,R}[n] \geq b_{k,\min}[n] \tag{26b}$$

$$b_{k,l}[n] \geq 0, b_{k,U}[n] \geq 0, b_{k,R}[n] \geq 0 \tag{26c}$$

$$0 \leq \tau_k^{\text{off}}[n] \leq \frac{\tau}{K}, \ 0 \leq \tau_{k,U}^{\text{off}}[n] \leq \frac{\tau}{K},$$

$$0 \leq \frac{c_U b_{k,U}[n]}{f_{U,\max}} \leq \frac{\tau}{K},$$

$$0 \leq \tau_{k,U}^{\text{dow}}[n] \leq \frac{\tau}{K}, \ 0 \leq \tau_{k,R}^{\text{dow}}[n] \leq \frac{\tau}{K},$$

$$0 \leq \frac{c_k b_{k,l}[n]}{f_{k,\max}} \leq \tau \tag{26d}$$

$$\tau_k^{\text{off}}[n] + \tau_{k,U}^{\text{off}}[n] + \frac{c_U b_{k,U}[n]}{f_{U,\max}}$$

$$+\tau_{k,U}^{\mathrm{dow}}[n]+\tau_{k,R}^{\mathrm{dow}}[n]\leq\frac{\tau}{K} \tag{26e}$$

$$b_{k,U}[n]+b_{k,R}[n]\leq\tau_{k}^{\mathrm{off}}[n]\,r_{kU}\left(\frac{E_{k}^{\mathrm{off}}[n]}{\tau_{k}^{\mathrm{off}}[n]}\right) \tag{26f}$$

$$b_{k,R}[n]\leq\tau_{k,U}^{\mathrm{off}}[n]\,r_{UR}\left(\frac{E_{k,U}^{\mathrm{off}}[n]}{\tau_{k,U}^{\mathrm{off}}[n]}\right) \tag{26g}$$

$$\xi_{k}b_{k,U}[n]\leq\tau_{k,U}^{\mathrm{dow}}[n]\,r_{Uk}\left(\frac{E_{k,U}^{\mathrm{dow}}[n]}{\tau_{k,U}^{\mathrm{dow}}[n]}\right) \tag{26h}$$

$$\xi_{k}b_{k,R}[n]\leq\tau_{k,R}^{\mathrm{dow}}[n]\,r_{Uk}\left(\frac{E_{k,R}^{\mathrm{dow}}[n]}{\tau_{k,R}^{\mathrm{dow}}[n]}\right) \tag{26i}$$

$$\begin{aligned}0\leq E_{k}^{\mathrm{off}}[n]&=p_{k}^{\mathrm{off}}[n]\,\tau_{k}^{\mathrm{off}}[n]\\&\leq P_{k,\max}^{\mathrm{off}}\tau_{k}^{\mathrm{off}}[n]\end{aligned} \tag{26j}$$

$$\begin{aligned}0\leq E_{k,U}^{\mathrm{off}}[n]&=p_{k,U}^{\mathrm{off}}[n]\,\tau_{k,U}^{\mathrm{off}}[n]\\&\leq P_{k,A,\max}^{\mathrm{off}}\tau_{k,U}^{\mathrm{off}}[n]\end{aligned} \tag{26k}$$

$$\begin{aligned}0\leq E_{k,U}^{\mathrm{dow}}[n]&=p_{k,U}^{\mathrm{dow}}[n]\,\tau_{k,U}^{\mathrm{dow}}[n]\\&\leq P_{k,U,\max}^{\mathrm{dow}}\tau_{k,U}^{\mathrm{dow}}[n]\end{aligned} \tag{26l}$$

$$\begin{aligned}0\leq E_{k,R}^{\mathrm{dow}}[n]&=p_{k,R}^{\mathrm{dow}}[n]\,\tau_{k,R}^{\mathrm{dow}}[n]\\&\leq P_{k,R,\max}^{\mathrm{dow}}\tau_{k,R}^{\mathrm{dow}}[n]\end{aligned} \tag{26m}$$

where

$$\begin{aligned}E_{k}[n]&=E_{k,cl}[n]+E_{k}^{\mathrm{off}}[n]\\&=\kappa_{k}c_{k}^{3}\left(b_{k,l}[n]\right)^{3}\tau^{-2}+p_{k}^{\mathrm{off}}[n]\,\tau_{k}^{\mathrm{off}}[n],\end{aligned} \tag{27}$$

$$\begin{aligned}E_{U}[n]&=\sum_{k=1}^{K}\left(E_{k,U}^{\mathrm{off}}[n]+E_{k,cU}[n]+E_{k,U}^{\mathrm{dow}}[n]+E_{k,R}^{\mathrm{dow}}[n]\right)\\&=\sum_{k=1}^{K}\Big(p_{k,U}^{\mathrm{off}}[n]\,\tau_{k,U}^{\mathrm{off}}[n]+\kappa_{U}c_{U}^{3}K^{2}\left(b_{k,U}[n]\right)^{3}\tau^{-2}\\&\quad+p_{k,U}^{\mathrm{dow}}[n]\,\tau_{k,U}^{\mathrm{dow}}[n]+p_{k,R}^{\mathrm{dow}}[n]\,\tau_{k,R}^{\mathrm{dow}}[n]\Big)\end{aligned} \tag{28}$$

stand for the energy consumption of the $k$-vehicle and ARSU, respectively, in each time slot, including the energy consumed for bits offloading, bits downloading, and bits computation. In (P1), (26a) is the objective function for minimizing the WTEC, $w_k \geq 0$ and $w_U \geq 0$ are the weight factors of energy consumption of $k$-th vehicle and ARSU, respectively, and $\mathbf{B} \triangleq \left\{b_{k,l}[n], b_{k,U}[n], b_{k,R}[n]\right\}$, $\mathbf{P} \triangleq \left\{p_{k}^{\mathrm{off}}[n], p_{k,U}^{\mathrm{off}}[n], p_{k,U}^{\mathrm{dow}}[n], p_{k,R}^{\mathrm{dow}}[n]\right\}$, and $\boldsymbol{\tau} \triangleq \left\{\tau_{k}^{\mathrm{off}}[n], \tau_{k,U}^{\mathrm{off}}[n], \tau_{k,U}^{\mathrm{dow}}[n], \tau_{k,R}^{\mathrm{dow}}[n]\right\}$ are the optimizing variables. Depending on the requirements of an envisioned application, the weight factors can be easily modified to satisfy the energy demands and trade-offs. Also, these weight factors provide priority and fairness among the vehicles. As $w_k$ ($w_U$) increases, the $k$-th vehicle (ARSU) has higher priority and thus the minimization of the $k$-th vehicle's (ARSU's) energy consumption becomes more important. It is noted that the constraint (26b) designates the task allocation, the constraint (26c) ensures the non-negativeness of the computation bits, the constraints (26d) and (26e) indicate the transmission and computation delay limitations, the constraints (26f)-(26i) describe the data offloading and downloading limitations, and the constraints (26j), (26k), (26l), and (26m) specify the range of transmit power at Phases 1, 2, 4, and, 5, respectively. One observes that $E_{total}$ in problem (P1) is an increasing function of the offloaded data. Hence, bits offloading should be realized, only if local computation violates latency constraints.

*Lemma 1:* Problem (P1) is a convex problem.

*Proof:* From (26a) (27), and (28), we can conclude that the objective function of problem (P1) is convex with respect to (w.r.t.) $\mathbf{P}$, $b_{k,l}[n]$, and $b_{k,U}[n]$, since its Hessian matrix is positive semidefinite. Also, the expressions in constraints (26b)-(26e) and (26j)-(26m) are linear. Moreover, the right-hand-side of (26f)-(26i) are concave, since $f(x,t) = t\log(1+x/t)$ with $t > 0$ is concave [60]. Since the sum of convex functions preserves convexity, the objective function of problem (P1) is jointly convex w.r.t. all the optimizing variables. Therefore, Problem (P1) is a convex problem.

To solve Problem (P1), the Lagrangian dual method is leveraged. The optimal solutions of problem (P1) are given in the following proposition.

*Proposition 1: The optimal computation bits allocated for local computing, offloading to ARSU for computing, and offloading to GRSU for computing via ARSU can be obtained, respectively, as*

$$b_{k,l}^{*}[n]=\tau\left[\sqrt{\frac{\chi_{1,k,n}}{3w_{k}\kappa_{k}c_{k}^{3}}}\right]_{0}^{\frac{f_{k,\max}}{c_{k}}}, \tag{29}$$

$$b_{k,U}^{*}[n]=\begin{cases}\dfrac{\tau}{K}\left[\sqrt{\dfrac{\zeta_{k,n}}{3w_{U}\kappa_{U}c_{U}^{3}f_{U,\max}}}\right]_{0}^{\frac{f_{U,\max}}{c_{U}}}, & \zeta_{k,n}\geq 0\\ 0, & \zeta_{k,n}<0\end{cases}, \tag{30}$$

$$b_{k,R}^{*}[n]=\begin{cases}0, & \chi_{3,k,n}+\chi_{4,k,n}+\chi_{6,k,n}\xi_{k}-\chi_{1,k,n}>0\\ \omega, & \chi_{3,k,n}+\chi_{4,k,n}+\chi_{6,k,n}\xi_{k}-\chi_{1,k,n}=0\end{cases}, \tag{31}$$

*where* $\zeta_{k,n} = f_{U,\max}\left(\chi_{1,k,n}-\chi_{3,k,n}-\chi_{5,k,n}\xi_{k}\right) - \chi_{2,k,n}c_{U}$, $\omega$ *is an arbitrary non-negative constant, and* $\chi_{1,k,n}$, $\chi_{2,k,n}$, $\chi_{3,k,n}$, $\chi_{4,k,n}$, $\chi_{5,k,n}$, *and* $\chi_{6,k,n}$ *are the non-negative Lagrange multipliers (dual variables) related with the constraints in (26b), (26e), (26f), (26g), (26h), and (26i), respectively. Also, the optimal transmit power* $p_{k}^{\mathrm{off}*}[n]$ *at Phase 1 can be obtained by solving the following equation using numerical solving techniques:*

$$w_{k}-\chi_{3,k,n}B\sum_{l=1}^{\min(L_{k},L_{U})}\frac{\lambda_{kU,l}^{2}[n]}{BN_{0}L_{k}\log_{2}\left(1+\frac{p_{k}^{\mathrm{off}}[n]\lambda_{kU,l}^{2}[n]}{BN_{0}L_{k}}\right)}=0. \tag{32}$$

*Moreover, the optimal transmission delay* $\tau_k^{\text{off}*}[n]$ *can be expressed as in (33), as shown at the bottom of the next page.*

*Proof:* See Appendix A.

Using (29) and (30), the optimal computation delays $\tau_{k,cl}^*[n] = c_k b_{k,l}^*[n]/f_{k,\max}$ and $\tau_{k,cU}^*[n] = c_U b_{k,U}^*[n]/f_{U,\max}$ can be obtained. Also, the optimal energy consumption of the $k$-th vehicle for data offloading at Phase 1 can be expressed as

$$E_k^{\text{off}*}[n] = p_k^{\text{off}*}[n]\,\tau_k^{\text{off}*}[n]. \tag{34}$$

In addition, using (32)-(34) and properly replacing the indices, the optimal transmit powers, transmission delays, and energy consumption at Phases 2, 4, and 5 can be obtained.

*Remark 1: The expressions in (29) and (30) indicate that the weight factor directly controls the division of the task-input bits and the corresponding computation delay. As* $w_k$ *and* $w_U$ *increase, less task-input data is processed locally and at ARSU, respectively. Also,* $b_{k,U}^*[n]$ *decreases, as* $\xi_k$ *increases. Moreover, the k-th vehicle (ARSU) would choose to offload data to ARSU (GRSU) for computing, as far as* $b_k[n] > b_{k,l}^*[n]\left(b_{k,UR}[n] > b_{k,U}^*[n]\right)$.

To derive closed-form solutions for $p_k^{\text{off}*}[n]$ and provide network design recommendations, the special, but common, case of highly correlated rank-1 LoS massive MIMO channels is initially studied. This case leads to the lower bound of the achievable rate [50], [54]. Then, $\mathbf{H}_{kU}[n]$ has one nonzero singular value. Next, the upper bound of the achievable rate is investigated, where this matrix is full-rank and all of its singular values are nonzero and equal. Unlike in Rayleigh channels, orthogonality among the spatially multiplexed signals can be attained in this special case, under strict geometrical constraints and specific orientation of the arrays [52]. Although it is infeasible to adjust the placement of URPAs in highly mobile vehicular scenarios, we investigate the special case of full-rank channels, since it corresponds to the theoretical upper bound of the achievable rate.

*Proposition 2: The optimal transmit power and offloading time at Phase 1 for the lower bound of the achievable rate can be, respectively, obtained as*

$$p_{k,lb}^{\text{off}*}[n] = \left[B\left(\frac{\chi_{3,k,n}}{w_k \ln(2)} - \frac{N_0 L_k}{\Phi_{kU}[n]}\right)\right]_0^{P_{k,\max}^{\text{off}}} \tag{35a}$$

*and as in (35b), as shown at the bottom of the next page, where* $\Phi_{kU}[n] = tr\left(\mathbf{G}_{kU}[n]\,\mathbf{G}_{kU}^H[n]\right) = \sum_{l=1}^{\min(L_k,L_U)} \lambda_{kU,l}^2[n]$ *and* $(\cdot)^H$ *and* $tr(\cdot)$ *represent the conjugate transpose and trace of matrix, respectively. Moreover, the optimal transmit power and offloading time at Phase 1 for the upper bound of the achievable rate can be, respectively, obtained as*

$$p_{k,ub}^{\text{off}*}[n] = \left[\min(L_k, L_U)\, p_{k,lb}^{\text{off}*}[n]\right]_0^{P_{k,\max}^{\text{off}}} \tag{36a}$$

*and as in (36b), as shown at the bottom of the next page.*

*Proof:* See Appendix B.

*Remark 2: From (35a) and (36a), one concludes that there exists a linear relationship between the optimized transmitted power of a rank-1 and that of a full-rank channel, whereas* $p_{k,lb}^{\text{off}*}[n] \le p_{k,ub}^{\text{off}*}[n]$. *As* $\min(L_k, L_U)$ *increases,* $p_{k,ub}^{\text{off}*}[n]$ *linearly increases and leads to increased WTEC. Clearly, the massive MIMO channels and system geometry directly affect the energy optimization. In addition, (35a) and (36a) indicate that larger values of* $w_k$ *correspond to decreased optimized transmit power and thus decreased* $E_k^{\text{off}*}[n]$.

Henceforth, predetermined dual variables are considered. The next step of the optimization procedure is to obtain the optimal dual variables by following the procedure described in Appendix C. Since $\boldsymbol{\tau}^*$ and $b_{k,R}^*[n]$ are not unique, we formulate the following linear programming problem:

$$(\text{P2}): \min_{b_{k,R}[n],\boldsymbol{\tau}} \sum_{n=1}^{N}\left[\left(\sum_{k=1}^{K} w_k E_k^{\text{off}}[n]\right) + w_U E_U'[n]\right] \tag{37a}$$

$$\text{s.t. (26c), (26d), (26e), (26j)} - \text{(26m)} \tag{37b}$$

$$b_{k,l}^*[n] + b_{k,U}^*[n] + b_{k,R}[n] \ge b_{k,\min}[n] \tag{37c}$$

$$b_{k,U}^*[n] + b_{k,R}[n] \le \tau_k^{\text{off}}[n]\, r_{kU}\left(p_k^{\text{off}*}[n]\right) \tag{37d}$$

$$b_{k,R}[n] \le \tau_{k,U}^{\text{off}}[n]\, r_{UR}\left(p_{k,U}^{\text{off}*}[n]\right) \tag{37e}$$

$$b_{k,U}^*[n]\,\xi_k \le \tau_{k,U}^{\text{dow}}[n]\, r_{Uk}\left(p_{k,U}^{\text{dow}*}[n]\right) \tag{37f}$$

$$b_{k,R}[n]\,\xi_k \le \tau_{k,R}^{\text{dow}}[n]\, r_{Uk}\left(p_{k,R}^{\text{dow}*}[n]\right) \tag{37g}$$

where $E_U'[n] = \sum_{k=1}^{K}\left(E_{k,U}^{\text{off}}[n] + E_{k,U}^{\text{dow}}[n] + E_{k,R}^{\text{dow}}[n]\right)$.

In order to obtain the optimal solution to primal problem (P1), problem (P2) should be solved. Based on the previous results and observations and adopting the ellipsoid method [60], the subgradient-based Algorithm 1 is proposed to optimally solve this problem. Based on [60], the convergence of this algorithm can be guaranteed. Also, the complexity and running time of Algorithm 1 depends on the number of time slots and the number of vehicles. More importantly, the main complexity of Algorithm 1 lies in steps 4, 5, and 6, where the complexity is $O(KN)$, $O(KN)$, and $O\left(K^2N^2\right)$, respectively [60]. Hence, Algorithm 1 has a total complexity of $O\left(K^4N^4\right)$. Finally, in Step 9, the complexity mainly depends on solving problem (P2) by CVX [61].

## V. NUMERICAL RESULTS AND DISCUSSION

In this section, numerical results are presented to illustrate the TCCD $\tau_{TCCD} = \sum_{n=1}^{N}\sum_{k=1}^{K}\tau_k[n]$ and the WTEC for different values of the key system parameters and under latency constraints. The effectiveness of the LoS massive MIMO and the optimization method is also studied, whereas the convergence performance of the proposed algorithm is also evaluated. The results take into account the number of antennas, the number of vehicles, the computation task size, the relative location of the ARSU w.r.t. the vehicles (GRSU), the time horizon $T_U$, the velocity and weight factor of energy consumption of ARSU, and the proportionality ratio between offloaded data and computed results. The vector $\mathbf{B}_V \in \mathbb{R}^{1\times k}$ is used to represent the set of required computation data

(in Mbits), in which the $k$-th entry stands for the required computation task for the $k$-th vehicle per time slot. Mobile applications that typically compound multiple procedures and support partial offloading are considered, such as the computation components in AR applications [62]. Without loss of generality, it is assumed that $L_k = L_U = L_R$, whereas the vehicles have identical task requirement. Unless otherwise stated, the values of key parameters are listed in Table 1.

Fig. 3 shows the non-optimized and optimized TCCD of Phases 1-5 as a function of the number of vehicles for a rotary-wing ARSU, URPAs with different number of antenna elements, and task requirement $b_k[n] = 0.5$ Mbit per time slot. Clearly, the delay substantially decreases with the number of antennas and grows with the number of vehicles. Besides, the number of supported vehicles changes with the number of antenna elements. To provide computing services to three vehicles, while satisfying the stringent latency constraints, URPAs with at least 16-elements are required. Meanwhile, the optimized scheme supports a larger number of vehicles, when compared with the non-optimized one, thus revealing the effectiveness of our optimization method. However, the optimized scheme can support up to a certain number of vehicles for a specific number of antennas, under task deadline constraints.

Fig. 4 depicts the non-optimized and optimized TCCD as a function of the number of antenna elements for a rotary-wing ARSU and varying task requirement $\mathbf{B}_V$. One observes that the TCCD significantly decreases, as the number of antennas increases, owing to the higher data rates and the lower transmission delay. As the number of antennas increases from 16 to 64, up to 1 Mbits and 1.55 Mbits can be supported for the non-optimized and optimized scheme, respectively. Also, using URPAS with relatively small dimensions, e.g., 36-element URPAs of 0.45 m x 0.45 m size for the commonly used 2 GHz carrier frequency, up to 0.82 Mbits can be computed per time slot. Besides, using a similar setup and single-antennas less than 0.2 Mbits can be timely executed. Thus, the benefits and feasibility from integrating large-scale antennas on size-constrained conventional vehicles and currently available commercial off-the-shelf UAVs is affirmed. Overall, a reasonable number of antennas should be employed, according to the amount of offloaded data, in order to attain acceptable TCCD, while satisfying practical antenna size constraints. By using mmWave frequency bands, which are potentially available for air-to-ground communications [63], the antenna arrays can be even more compact and more demanding tasks can be handled.

Fig. 5 shows the WTEC per time slot as a function of the delay requirement per time slot for a fixed-wing ARSU and task requirement $\mathbf{B}_V = [b_1, b_2, b_3] = [0.8, 0.8, 0.8]$ Mbits. More specifically, the maximum allowable time (i.e., latency constraint) needed for data offloading, data computation, and data downloading during operation phases 1 to 5 (see Fig. 2) is considered. It can be seen that the optimization scheme surpasses the non-optimized one for any delay requirement value. For instance, 5.8 Joules and 5.4 Joules per time slot should be consumed for the non-optimized and optimized

$$\tau_k^{\text{off}*}[n] = \begin{cases} = \frac{\tau}{K}, w_k p_k^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p_k^{\text{off}*}[n]\,\lambda_{kU,l}^2[n]}{BN_0 L_k}\right) < 0 \\ \in \left[0, \frac{\tau}{K}\right], w_k p_k^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p_k^{\text{off}*}[n]\,\lambda_{kU,l}^2[n]}{BN_0 L_k}\right) = 0 \\ = 0, w_k p_k^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p_k^{\text{off}*}[n]\,\lambda_{kU,l}^2[n]}{BN_0 L_k}\right) > 0 \end{cases} \tag{33}$$

$$\tau_{k,lb}^{\text{off}*}[n] \begin{cases} = \frac{\tau}{K}, w_k p_{k,lb}^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \log_2\left(1 + \frac{p_{k,lb}^{\text{off}*}[n]\,\Phi_{kU}[n]}{BN_0 L_k}\right) < 0 \\ \in \left[0, \frac{\tau}{K}\right], w_k p_{k,lb}^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \log_2\left(1 + \frac{p_{k,lb}^{\text{off}*}[n]\,\Phi_{kU}[n]}{BN_0 L_k}\right) = 0 \\ = 0, w_k p_{k,lb}^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \log_2\left(1 + \frac{p_{k,lb}^{\text{off}*}[n]\,\Phi_{kU}[n]}{BN_0 L_k}\right) > 0 \end{cases} \tag{35b}$$

$$\tau_{k,ub}^{\text{off}*}[n] = \begin{cases} = \frac{\tau}{K}, w_k p_{k,ub}^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \min(L_k, L_U) \log_2\left(1 + \frac{p_{k,ub}^{\text{off}*}[n]\,\Phi_{kU}[n]}{BN_0 L_k \min(L_k, L_U)}\right) < 0 \\ \in \left[0, \frac{\tau}{K}\right], w_k p_{k,ub}^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \min(L_k, L_U) \log_2\left(1 + \frac{p_{k,ub}^{\text{off}*}[n]\,\Phi_{kU}[n]}{BN_0 L_k \min(L_k, L_U)}\right) = 0 \\ = 0, w_k p_{k,ub}^{\text{off}*}[n] + \chi_{2,k,n} - \chi_{3,k,n} B \min(L_k, L_U) \log_2\left(1 + \frac{p_{k,ub}^{\text{off}*}[n]\,\Phi_{kU}[n]}{BN_0 L_k \min(L_k, L_U)}\right) > 0 \end{cases} \tag{36b}$$

**Algorithm 1** Optimal Solution to Problem (P1)

1: **Set** $K$, $\{w_k\}$, $w_U$, $T_U$, $\tau$, $f_{k,\max}$, $f_{U,\max}$, $\{c_k\}$, $c_U$ $\{\kappa_k\}$, $\kappa_U$, $\{\xi_k\}$, $h_U$, $\{\theta_k\}$, $\theta_R$, $\{v_k\}$, $\{\gamma_k\}$, $v_U$, $\gamma_{U,xy}$, $\gamma_{U,z}$, $c_1$, $c_2$, $c_3$, $v_{tip}$, $v_0$, $d_r$, $s$, $\rho$, $G$, $P_0$, $P_1$, $P_2$, $\lambda$, $\alpha$, $B$, $\beta_0$, $N_0$, $\left\{P^{\text{off}}_{k,\max}\right\}$, $\left\{P^{\text{off}}_{k,U,\max}\right\}$, $\left\{P^{\text{dow}}_{k,U,\max}\right\}$, $\left\{P^{\text{dow}}_{k,R,\max}\right\}$, $\{L_k\}$, $L_U$, $L_R$, $\{\delta_k\}$, $\delta_U$, $\delta_R$, $\beta_x$, $\beta_y$, $\beta_z$, and the tolerant threshold $\varepsilon$.
2: **Initialize** the iteration index, the non-optimized dual variables $\{\mathbf{x}_\delta\}_{\delta=1}^{6}$ and the ellipsoid (as described in Appendix C). Then, obtain the channel matrices using (11) and decompose these matrices via singular value decomposition (SVD) to obtain the singular values.
3: **Repeat**
4: Use (29) and (30) and obtain $b^*_{k,l}[n]$ and $b^*_{k,U}[n]$, respectively. Then, numerically solve the equation in (32) and obtain $\mathbf{P}^*$. Also, use (38) and obtain $\boldsymbol{\tau}^*$. Calculate the WTEC.
5: Solve problem P1-dual defined in (A.3a)-(A.3c) in Appendix A by calculating the subgradients defined in (C.1a)-(C1.f) in Appendix C.
6: Update $\{\mathbf{x}_\delta\}_{\delta=1}^{6}$ according to the ellipsoid method.
7: **End Repeat** until convergence.
8: Let $\left\{\mathbf{x}^*_\delta\right\}_{\delta=1}^{6} \leftarrow \{\mathbf{x}_\delta\}_{\delta=1}^{6}$
9: Use (29) and (30) and obtain $b^*_{k,l}[n]$ and $b^*_{k,U}[n]$, respectively. Update $\mathbf{P}^*$ by re-solving the equation in (32). Then, obtain $b^*_{k,R}[n]$ and $\boldsymbol{\tau}^*$ by solving problem (P2) by CVX. Finally, obtain the minimum WTEC.

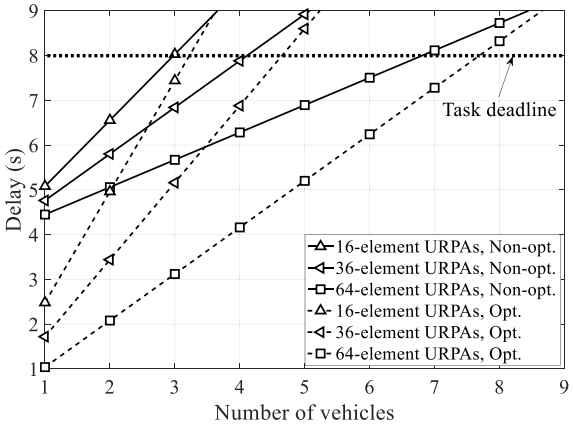


**FIGURE 3.** The non-optimized and optimized TCCD as a function of the number of vehicles for varying number of antennas.

scheme, respectively, in order to satisfy the delay requirement of 20 ms of AR and virtual reality (VR) applications [64]. Moreover, the WTEC decreases, as the delay requirement increases for both schemes. This is because a lower data rate is needed, under less tight latency constraints, which further leads to lower energy consumption. However, when the delay requirement is large enough, the decrement in the WTEC becomes insignificant. Also, Fig. 5 indicates that there exists a trade-off between energy consumption and task execution latency. Thus, considering both the limited battery lifetime of network nodes and the latency-sensitive applications, efficient offloading decisions should be made.

**TABLE 1.** Definition, notation, and values of key network parameters.

| **System Parameters** | **Value** |
|---|---|
| Number of vehicles: $K$ | 3 |
| Weight factor for energy consumption for $k$-th vehicle (ARSU): $w_k\,(w_U)$ | 1 (0.1) |
| Parameters of fixed-wing ARSU: $c_1$, $c_2$, $c_3$ | $9.26\cdot10^{-4}$, 2250, 3.33 [26] |
| Parameters of rotary-wing ARSU: $v_{tip}$, $v_0$, $d_r$, $s$, $\rho$, $G$, $P_0$, $P_1$, $P_2$ | 120, 4.3, 0.6, 0.05, 1.225, 0.503, $12\cdot30^3\cdot0.4^3\rho sG/8$, $1.1\cdot20^{3/2}/\sqrt{2\rho G}$, 11.46 [26] |
| **Computation Parameters** | **Value** |
| Task deadline (flight duration of ARSU): $T_U$ | 8 s |
| Time slot length: $\tau$ | 0.2 s [30] |
| Maximum CPU frequency at $k$-th vehicle (ARSU): $f_{k,\max}\,(f_{U,\max})$ | 1 GHz (3 GHz) |
| Required CPU cycles per bit at $k$-th vehicle (ARSU): $c_k\,(c_U)$ | $10^3$ ($10^3$) cycles/bit [30] |
| CPU capacitance coefficient at $k$-th vehicle (ARSU): $\kappa_k\,(\kappa_U)$ | $10^{-27}$ ($10^{-27}$) [30] |
| Task size ratio of output data to input data: $\xi_k$ | 0.8 |
| **Geometrical and Mobility Parameters** | **Value** |
| Initial height of ARSU: $h_U$ | 10 m [15] |
| Initial elevation angle of ARSU relative to $O_1$, $O_2$, $O_3$, $O_R$, respectively: $\theta_1$, $\theta_2$, $\theta_3$, $\theta_R$ | $\pi/3$, $\pi/4$, $\pi/6$, $\pi/3$ |
| Slant, downtilt, and bearing angle, respectively: $\beta_x$, $\beta_y$, $\beta_z$ | $\pi/3$, $\pi/4$, $\pi/3$ |
| Velocity and moving direction of $k$-th vehicle in the azimuth domain, respectively: $v_k$, $\gamma_k$ | 60 km/h, $\pi/3$ |
| Velocity and moving direction of ARSU in the azimuth (elevation) domain, respectively: $v_U$, $\gamma_{U,xy}$ $(\gamma_{U,z})$ | 10 m/s, $\pi/3$ ($\pi/9$) |
| **Wireless Transmission Parameters** | **Value** |
| Carrier wavelength: $\lambda$ | 0.15 m |
| Path-loss exponent: $\alpha$ | 2 [58] |
| Bandwidth for uplink (or downlink): $B$ | 5 MHz |
| Channel gain at reference distance $d_0 = 1$ m: $\beta_0$ | -50 dB [30] |
| Variance of AWGN at $k$-th vehicle, ARSU, and GRSU: $N_0$ | -130 dBm/Hz [30] |
| Max. transmit power in Phase 1, 2, 4, and, 5, respectively: $P^{\text{off}}_{k,\max}$, $P^{\text{off}}_{k,U,\max}$, $P^{\text{dow}}_{k,U,\max}$, $P^{\text{dow}}_{k,R,\max}$ | 35 dBm [30] |
| Number of antennas at $k$-th vehicle, ARSU, and GRSU array, respectively: $L_k$, $L_U$, $L_R$ | 36, 36, 36 |
| Inter-element spacing at $k$-th vehicle, ARSU, and GRSU antenna array, respectively: $\delta_k$, $\delta_U$, $\delta_R$ | $\lambda/2$ |

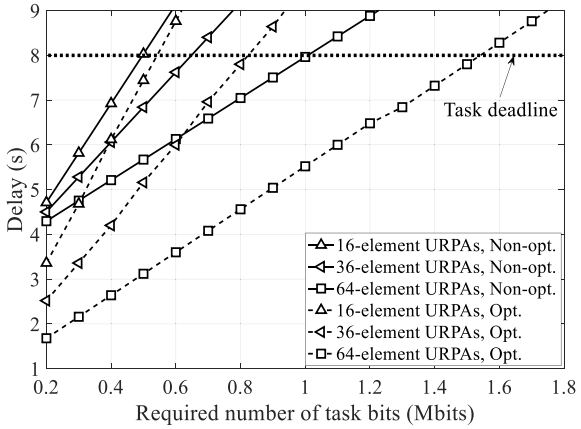


**FIGURE 4.** The non-optimized and optimized TCCD as a function of the task requirement per time slot for varying number of antennas.

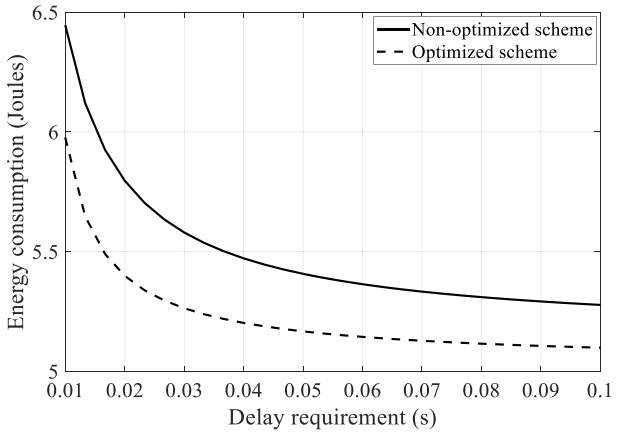


**FIGURE 5.** The non-optimized and optimized WTEC per time slot as a function of the delay requirement per time slot.

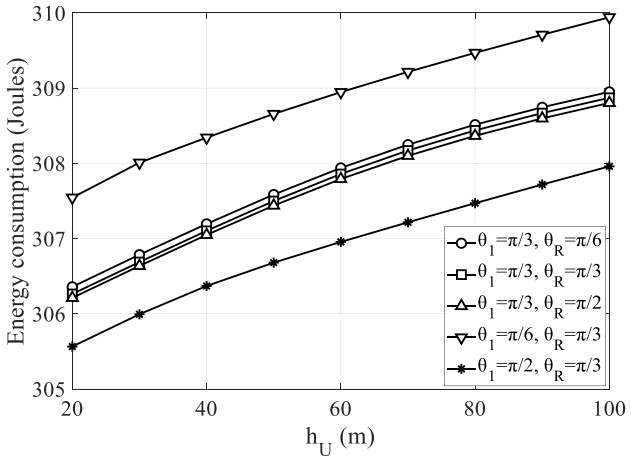


**FIGURE 6.** The non-optimized WTEC as a function of the altitude of the ARSU for varying elevation angle of the ARSU.

Fig. 6 investigates the impact of the altitude of a hovering rotary-wing ARSU on the non-optimized WTEC for $b_1[n] = 0.6$ Mbits, and varying initial elevation angle of the ARSU w.r.t. the vehicle, i.e., $\theta_1$, and w.r.t. the GRSU, i.e., $\theta_R$. As $h_U$ increases, the ARSU draws away from both the vehicle and GRSU and more energy is consumed. It is also evident that the WTEC fairly increases as $\theta_1$ decreases, since the quality of the V2U and U2V channels in terms of the path-loss and correlation is somehow degraded. However, changing $\theta_R$ is even less influential and negligibly affects the WTEC. Previous results on single-antenna configurations [15], stated that the UAV should be closed to ground nodes to ensure low offloading/downloading energy consumption and support large task sizes. Nevertheless, in this paper, the use of massive MIMO promises meaningfully lower transmission delays and enhanced rates that obliterate such indications and compensate the increased path-loss observed in larger distances. Therefore, this paper suggests that the ARSU should not necessarily approach the moving vehicles and/or GRSU to attain satisfactory WTEC and/or meliorate possible side effects of unstable $h_U$ due to obstacles and wind/pressure variableness. By avoiding aimless movements, a significant amount of propulsion energy can be saved thus extending the endurance of the ARSU.

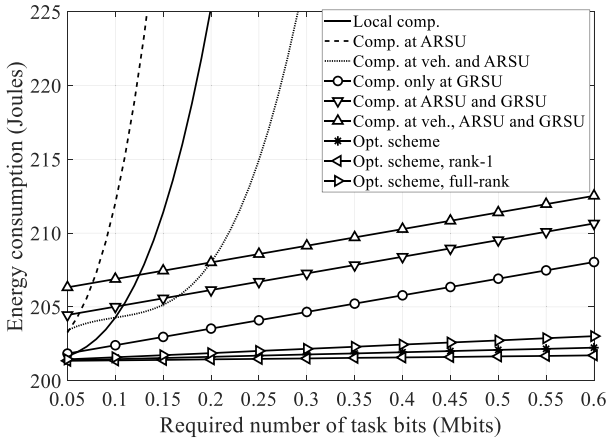


**FIGURE 7.** The non-optimized and optimized WTEC as a function of the task requirement per time slot for different computing scenarios.

Fig. 7 illustrates the non-optimized and optimized WTEC as a function of the task requirement per time slot for a fixed-wing ARSU and six computing scenarios, including local computing, full offloading and partial offloading. Apparently, the GRSU is not necessary to assist on computation for small values of task bits, e.g., 0.1 Mbits per time slot. However, local computing is subject to a maximum computing capability $\tau f_{k,\max}/c_k$, according to (14). Exploiting only the ARSU for computing leads to even greater WTEC, since wireless transmission consumes additional energy. Besides, as indicated in (25), the energy consumption for computing at ARSU exponentially increases with $K$. In order to extend the supported number of task bits, the vehicles and the ARSU may cooperatively handle the computation process with an acceptable growth of the WTEC. On the other hand, demanding computation tasks presuppose the participation of the GRSU for efficient edge computing and slight energy

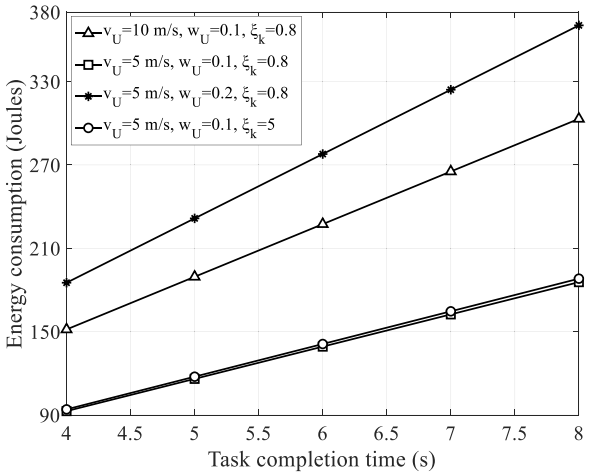


**FIGURE 8.** The optimized WTEC as a function of the task completion time (flying time of ARSU) for varying velocity of the ARSU, weight factor of the ARSU, and task size ratio of output data to input data.

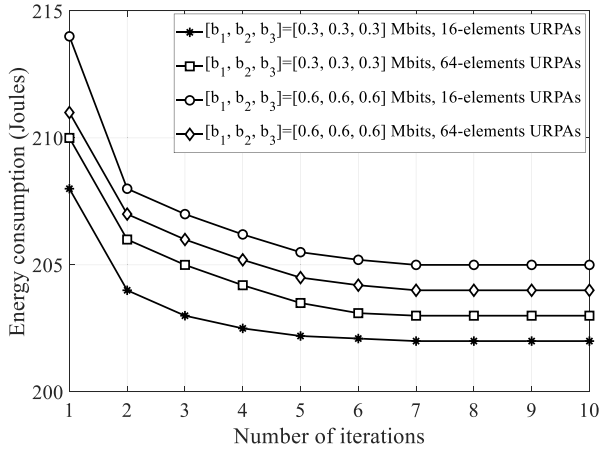


**FIGURE 9.** The optimized WTEC as a function of the number of iterations for varying task requirement and number of antenna elements at URPAs.

cost. As the task bits increase the vehicles should tend to transmit ideally the entire amount of task bits to the GRSU via the ARSU. Clearly, the curves of the optimized schemes outperform the non-optimized ones and depict the advantages of partial offloading by capitalizing on the local computation resources, as well as the MEC resources at ARSU and GRSU. Also, the difference between the non-optimized and optimized schemes enlarges with the task requirement, while the common scenario of rank-1 channels constitutes the most energy-efficient solution.

Fig. 8 shows the curves of the optimized WSEC as a function of the task completion time (flying time) for a fixed-wing ARSU for varying $v_U$, $w_U$, and $\xi_U$, when the task requirement is $\mathbf{B}_V = [b_1, b_2, b_3] = [0.6, 0.6, 0.6]$ Mbits and $\xi_1 = \xi_2 = \xi_3$. One observes that the consumed energy drastically and linearly increases, as the stringent deadline increases. It is also obvious that WTEC increases as $v_U$ and $w_U$ step up, since the propulsion energy contributes more to the WTEC. Highly-intensive computation tasks, e.g., video-rendering applications and delivery of 360° videos, may lead to $\xi_k \gg 1$ [65]. However, owing to the enhanced spectral efficiency offered by the massive MIMO channels, changing $\xi_k$ negligibly affects the WTEC. This is not the case for conventional single-antenna scenarios [15].

Finally, Fig. 9 investigates the convergence efficiency of the proposed Algorithm 1 and demonstrates the optimized WTEC for a fixed-wing ARSU and tolerant threshold $\varepsilon = 10^{-4}$ as a function of the iteration index. It can be seen that the proposed optimized WTEC scheme nearly converges after about 7 iterations, regardless of the task sizes and number of antennas, thus achieving computational effectiveness.

## VI. CONCLUSION AND FUTURE DIRECTIONS

In this paper, a novel WTEC optimization problem for a delay-constrained massive MIMO UAV-aided MEC-enabled vehicular network has been formulated. This problem can be decomposed into multiple convex subproblems that can be solved by the Lagrangian dual method along with an efficient subgradient-based algorithm. Capitalizing on the convenient form of the closed-form solutions, numerical calculations have been carried out to illustrate the mathematical derivations. We have shown that the number of antennas determines the number of supported vehicles and the size of offloaded data, under latency constraints. It has been also demonstrated that the vehicles may perform local computation for low task requirements. As task bits increase, partial task offloading is necessary. Since the velocity and weight factor of ARSU control the propulsion energy, the proposed approach has underlined that massive MIMO can counterbalance the distance-dependent path-loss and reduce purposeless mobility of ARSU.

This work can be expanded into various fertile research areas. As a pre-determined ARSU's trajectory is considered, the 3-D trajectory optimization is envisioned as a future work. Apart from LoS propagation, multipath fading may be included to more accurately describe the time-varying vehicular environment. In addition, the investigation of channel estimation errors and pilot contamination effects constitutes an interesting future direction. Moreover, multiple ARSUs and GRSUs can be utilized along with learning-based methods for intelligent control, in order to extend the network range. Also, massive connectivity can be ensured by adopting NOMA, while mm-wave frequencies can further increase the array gain.

## APPENDIX A
## PROOF OF PROPOSITION 1

The Lagrange function of problem (P1) can be expressed as in (A.1), as shown at the bottom of the next page, where $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5$, and $\mathbf{x}_6$ denote the sets of the dual variables $\chi_{1,k,n}$, $\chi_{2,k,n}$, $\chi_{3,k,n}$, $\chi_{4,k,n}$, $\chi_{5,k,n}$, and $\chi_{6,k,n}$, respectively. For arbitrary dual variables, the dual function of problem (P1) can be written as

$$\xi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6) = \min_{\mathbf{B},\mathbf{P},\boldsymbol{\tau}} \mathcal{L}(\mathbf{B}, \mathbf{P}, \boldsymbol{\tau}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6) \quad \text{(A.2a)}$$

$$\text{s.t. (26c), (26d), (26j)} - \text{(26m)} \tag{A.2b}$$

Based on the results in [30], $\xi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6)$ is bounded, provided that the inequality $\chi_{3,k,n} + \chi_{4,k,n} + \chi_{6,k,n}\xi_k - \chi_{1,k,n} \geq 0$ holds. Thus, the dual problem of problem (P1) can be expressed as

$$\text{P1-dual:} \max_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4,\mathbf{x}_5,\mathbf{x}_6} \xi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6) \tag{A.3a}$$

$$\text{s.t. } \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6\} \pm 0 \tag{A.3b}$$

$$\chi_{3,k,n} + \chi_{4,k,n} + \chi_{6,k,n}\xi_k - \chi_{1,k,n} \geq 0 \tag{A.3c}$$

Since problem (P1) is convex, the Slater's condition is satisfied [60]. Owing to the strong duality between (P1) and (P1-dual), we obtain the optimal solution of problem (P1) by solving its dual problem, i.e., problem (P1-dual). For arbitrary values of $\{\mathbf{x}_\delta\}_{\delta=1}^{6}$, we obtain the dual function by solving the problem defined in (A.3). This problem can be rewritten into a set of $KN$ independent subproblems. Thus, we can further decompose these subproblems into several subproblems w.r.t. each vehicle. These subproblems are convex and their solutions satisfy the Karush–Kuhn–Tucker (KKT) conditions [60]. We define the subproblems related with $b_{k,l}[n]$, $b_{k,U}[n]$, and $b_{k,R}[n]$, respectively as

$$\text{(L1)}: \min_{b_{k,l}[n]} w_k \kappa_k c_k^3 \left(b_{k,l}[n]\right)^3 \tau^{-2} - \chi_{1,k,n} b_{k,l}[n]$$
$$\text{s.t. (26c), (26d)}$$

$$\text{(L2)}: \min_{b_{k,U}[n]} w_U \kappa_U c_U^3 K^2 \left(b_{k,U}[n]\right)^3 \tau^{-2} + \left(\frac{\chi_{2,k,n} c_U}{f_{U,\max}} + \chi_{3,k,n} + \chi_{5,k,n}\xi_k - \chi_{1,k,n}\right) b_{k,U}[n]$$
$$\text{s.t. (26c), (26d)}$$

$$\text{(L3)}: \min_{b_{k,R}[n]} \left(\chi_{3,k,n} + \chi_{4,k,n} + \chi_{6,k,n}\xi_k - \chi_{1,k,n}\right) b_{k,R}[n]$$
$$\text{s.t. (26c)}$$

By solving subproblems (L1), (L2), and (L3) with the aid of KKT conditions, we obtain the optimal solutions in (29), (30), and (31), respectively. Also, we define the subproblem, which is related with $p_k^{\text{off}}[n]$ and $\tau_k^{\text{off}}[n]$, as

$$\text{(L4)}: \min_{\tau_k^{\text{off}}[n], p_k^{\text{off}}[n]} \left(w_k p_k^{\text{off}}[n] + \chi_{2,k,n}\right) \tau_k^{\text{off}}[n] - \chi_{3,k,n} \tau_k^{\text{off}}[n] B_0 \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p_k^{\text{off}}[n] \lambda_{kU,l}^2[n]}{B_0 N_0 L_k}\right)$$
$$\text{s.t. (26d), (26j)}$$

The Lagrangian of subproblem (L4) can be written as

$$\begin{aligned}
&\mathcal{L}_4\left(\psi_{1,k,n}, \psi_{2,k,n}, \psi_{3,k,n}, \psi_{4,k,n}\right) \\
&= \left(w_k p_k^{\text{off}}[n] + \chi_{2,k,n}\right) \tau_k^{\text{off}}[n] - \chi_{3,k,n} \tau_k^{\text{off}}[n] B_0 \\
&\quad \times \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p_k^{\text{off}}[n] \lambda_{kU,l}^2[n]}{B_0 N_0 L_k}\right) \\
&\quad - \psi_{1,k,n} \tau_k^{\text{off}}[n] - \psi_{2,k,n}\left(\tau - \tau_k^{\text{off}}[n]\right) \\
&\quad - \psi_{3,k,n} p_k^{\text{off}}[n] \tau_k^{\text{off}}[n] \\
&\quad - \psi_{4,k,n}\left(P_{k,\max}^{\text{off}} \tau_k^{\text{off}}[n] - p_k^{\text{off}}[n] \tau_k^{\text{off}}[n]\right),
\end{aligned} \tag{A.4}$$

where $\psi_{1,k,n}, \psi_{2,k,n}, \psi_{3,k,n}$, and $\psi_{4,k,n}$ are non-negative Lagrange multipliers associated with the constraints $\tau_k^{\text{off}}[n] \geq 0, \tau_k^{\text{off}}[n] \leq \tau/K, p_k^{\text{off}}[n] \tau_k^{\text{off}}[n] \geq 0$, and

$$\begin{aligned}
&\mathcal{L}(\mathbf{B}, \mathbf{P}, \boldsymbol{\tau}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6) = \sum_{n=1}^{N}\sum_{k=1}^{K}\left[w_k \kappa_k c_k^3 \left(b_{k,l}[n]\right)^3 \tau^{-2} + w_k p_k^{\text{off}}[n] \tau_k^{\text{off}}[n]\right] \\
&+ w_U \sum_{n=1}^{N}\sum_{k=1}^{K}\left[p_{k,U}^{\text{off}}[n] \tau_{k,U}^{\text{off}}[n] + \kappa_U c_U^3 K^2 \left(b_{k,U}[n]\right)^3 \tau^{-2} + p_{k,U}^{\text{dow}}[n] \tau_{k,U}^{\text{dow}}[n] + p_{k,R}^{\text{dow}}[n] \tau_{k,R}^{\text{dow}}[n]\right)\right] \\
&+ \sum_{n=1}^{N}\sum_{k=1}^{K} \chi_{1,k,n} b_{k,\min}[n] - \sum_{n=1}^{N}\sum_{k=1}^{K} \chi_{1,k,n} b_{k,l}[n] + \sum_{n=1}^{N}\sum_{k=1}^{K}\left(\frac{\chi_{2,k,n} c_U}{f_{U,\max}} + \chi_{3,k,n} + \chi_{5,k,n}\xi_k - \chi_{1,k,n}\right) b_{k,U}[n] \\
&- \sum_{n=1}^{N}\sum_{k=1}^{K}\frac{\chi_{2,k,n}\tau}{K} + \sum_{n=1}^{N}\sum_{k=1}^{K}\left(\chi_{3,k,n} + \chi_{4,k,n} + \chi_{6,k,n}\xi_k - \chi_{1,k,n}\right) b_{k,R}[n] + \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{2,k,n}\tau_{k,U}^{\text{dow}}[n] + \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{2,k,n}\tau_k^{\text{off}}[n] \\
&+ \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{2,k,n}\tau_{k,U}^{\text{off}}[n] + \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{2,k,n}\tau_{k,R}^{\text{dow}}[n] - \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{3,k,n}\tau_k^{\text{off}}[n] B \sum_{l=1}^{\min(L_k, L_U)} \log_2\left(1 + \frac{p_k^{\text{off}}[n] \lambda_{kU,l}^2[n]}{BN_0 L_k}\right) \\
&- \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{4,k,n}\tau_{k,U}^{\text{off}}[n] B \sum_{l=1}^{\min(L_U, L_R)} \log_2\left(1 + \frac{p_{k,U}^{\text{off}}[n] \lambda_{UR,l}^2[n]}{BN_0 L_U}\right) - \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{5,k,n}\tau_{k,U}^{\text{dow}}[n] B \\
&\times \sum_{l=1}^{\min(L_U, L_k)} \log_2\left(1 + \frac{p_{k,U}^{\text{dow}}[n] \lambda_{Uk,l}^2[n]}{BN_0 L_U}\right) - \sum_{n=1}^{N}\sum_{k=1}^{K}\chi_{6,k,n}\tau_{k,R}^{\text{dow}}[n] B \sum_{l=1}^{\min(L_U, L_k)} \log_2\left(1 + \frac{p_{k,R}^{\text{dow}}[n] \lambda_{Uk,l}^2[n]}{BN_0 L_U}\right)
\end{aligned} \tag{A.1}$$

$p_k^{\text{off}}[n]\,\tau_k^{\text{off}}[n] \leq P_{k,\max}^{\text{off}}\tau_k^{\text{off}}[n]$, respectively, specified in (26d) and (26j). Based on KKT, the optimal transmit power at Phase 1 can be obtained by solving the equation $\partial\mathcal{L}_4\left(\psi_{1,k,n},\psi_{2,k,n},\psi_{3,k,n},\psi_{4,k,n}\right)/\partial p_k^{\text{off}}[n]=0$, which is defined in (32), using numerical solving techniques. Then, $\tau_k^{\text{off}*}[n]$ can be obtained by substituting $p_k^{\text{off}*}[n]$ into subproblem (L4) and is expressed as in (33).

## APPENDIX B
## PROOF OF PROPOSITION 2

The lower bound of the achievable rate can be expressed as [50], [54]

$$r_{kU,lb}[n]=B\log_2\left(1+\frac{p_k^{\text{off}}[n]\,\Phi_{kU}[n]}{BN_0L_k}\right). \tag{B.1}$$

Using (B.1) instead of (12), and solving the equation $\partial\mathcal{L}_4\left(\psi_{1,k,n},\psi_{2,k,n},\psi_{3,k,n},\psi_{4,k,n}\right)/\partial p_k^{\text{off}}[n]=0$, we obtain the optimal solution in (35a). Then, the optimal solution in (35b) can be obtained by substituting $p_{k,lb}^{\text{off}*}[n]$ into subproblem (L4). Also, the upper bound of the achievable rate is expressed as [50], [54]

$$r_{kU,ub}[n]=B\min\left(L_k,L_U\right)\log_2\left(1+\frac{p_k^{\text{off}}[n]\,\Phi_{kU}[n]}{BN_0L_k\min\left(L_k,L_U\right)}\right). \tag{B.2}$$

Similarly, using (B.2), one can obtain (36a) and then (36b).

## APPENDIX C
## DERIVATION OF OPTIMAL DUAL VARIABLES

In order to obtain the optimal dual variables, the problem (P1-dual), which is defined in (A3), should be solved. This problem is convex but non-differentiable and can be iteratively solved using the subgradient ellipsoid method [60]. The convergence of the ellipsoid method is guaranteed by the convexity of problem (P1-dual) [60]. Let $\Delta\mathbf{x}$ be the subgradient of the objective function in (A.3) w.r.t. $\mathbf{x}$. Then, we obtain

$$\Delta\mathbf{x}_1=b_{k,U}[n]+b_{k,R}[n]-\tau_k^{\text{off}}[n]\,r_{kU}\left(\frac{E_k^{\text{off}}[n]}{\tau_k^{\text{off}}[n]}\right), \tag{C.1a}$$

$$\Delta\mathbf{x}_2=b_{k,R}[n]-\tau_{k,U}^{\text{off}}[n]\,r_{UR}\left(\frac{E_{k,U}^{\text{off}}[n]}{\tau_{k,U}^{\text{off}}[n]}\right), \tag{C.1b}$$

$$\Delta\mathbf{x}_3=\xi_k b_{k,U}[n]-\tau_{k,U}^{\text{dow}}[n]\,r_{Uk}\left(\frac{E_{k,U}^{\text{dow}}[n]}{\tau_{k,U}^{\text{dow}}[n]}\right), \tag{C.1c}$$

$$\Delta\mathbf{x}_4=\xi_k b_{k,R}[n]-\tau_{k,R}^{\text{dow}}[n]\,r_{Uk}\left(\frac{E_{k,R}^{\text{dow}}[n]}{\tau_{k,R}^{\text{dow}}[n]}\right), \tag{C.1d}$$

$$\Delta\mathbf{x}_5=b_{k,\min}[n]-b_{k,l}[n]-b_{k,U}[n]-b_{k,R}[n], \tag{C.1e}$$

$$\Delta\mathbf{x}_6=\tau_k^{\text{off}}[n]+\tau_{k,U}^{\text{off}}[n]+\frac{c_U b_{k,U}[n]}{f_{U,\max}}+\tau_{k,U}^{\text{dow}}[n]+\tau_{k,R}^{\text{dow}}[n]-\tau/K. \tag{C.1f}$$

## REFERENCES


[1] W. Duan, J. Gu, M. Wen, G. Zhang, Y. Ji, and S. Mumtaz, "Emerging technologies for 5G-IoV networks: Applications, trends and opportunities," *IEEE Netw.*, vol. 34, no. 5, pp. 283–289, Sep. 2020.

[2] L. Thibault, G. De Nunzio, and A. Sciarretta, "A unified approach for electric vehicles range maximization via eco-routing, eco-driving, and energy consumption prediction," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 4, pp. 463–475, Dec. 2018.

[3] Z. Ning, X. Wang, and J. Huang, "Mobile edge computing-enabled 5G vehicular networks: Toward the integration of communication and computing," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 54–61, Mar. 2019.

[4] H. Guo, J. Liu, J. Ren, and Y. Zhang, "Intelligent task offloading in vehicular edge computing networks," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 126–132, Aug. 2020.

[5] J. Wang, D. Feng, S. Zhang, J. Tang, and T. Q. S. Quek, "Computation offloading for mobile edge computing enabled vehicular networks," *IEEE Access*, vol. 7, pp. 62624–62632, 2019.

[6] J. Zhou, D. Tian, Y. Wang, Z. Sheng, X. Duan, and V. Leung, "Reliability-oriented optimization of computation offloading for cooperative vehicle-infrastructure systems," *IEEE Signal Process. Lett.*, vol. 26, no. 1, pp. 104–108, Jan. 2019.

[7] J. Zhao, Q. Li, Y. Gong, and K. Zhang, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7944–7956, Aug. 2019.

[8] Y. Liu, H. Yu, S. Xie, and Y. Zhang, "Deep reinforcement learning for offloading and resource allocation in vehicle edge computing and networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11158–11168, Nov. 2019.

[9] J. Zhang, H. Guo, J. Liu, and Y. Zhang, "Task offloading in vehicular edge computing networks: A load-balancing solution," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2092–2104, Feb. 2020.

[10] A. Boukerche and V. Soto, "An efficient mobility-oriented retrieval protocol for computation offloading in vehicular edge multi-access network," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 6, pp. 2675–2688, Jun. 2020.

[11] Z. Ning, K. Zhang, X. Wang, L. Guo, X. Hu, J. Huang, B. Hu, and R. Y. K. Kwok, "Intelligent edge computing in internet of vehicles: A joint computation offloading and caching solution," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2212–2225, Apr. 2021, doi: 10.1109/TITS.2020.2997832.

[12] F. Zhou, R. Q. Hu, Z. Li, and Y. Wang, "Mobile edge computing in unmanned aerial vehicle networks," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 140–146, Feb. 2020.

[13] Y. Li, N. I. Miridakis, T. A. Tsiftsis, G. Yang, and M. Xia, "Air-to-air communications beyond 5G: A novel 3D CoMP transmission scheme," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7324–7338, Nov. 2020.

[14] N. Nomikos, E. T. Michailidis, P. Trakadas, D. Vouyioukas, H. Karl, J. Martrat, T. Zahariadis, K. Papadopoulos, and S. Voliotis, "A UAV-based moving 5G RAN for massive connectivity of mobile users and IoT devices," *Veh. Commun.*, vol. 25, Oct. 2020, Art. no. 100250.

[15] X. Hu, K. 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.

[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] Z. Yang, C. Pan, K. Wang, and M. Shikh-Bahaei, "Energy efficient resource allocation in UAV-enabled mobile edge computing networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4576–4589, Sep. 2019.

[18] C. Zhan, H. Hu, X. Sui, Z. Liu, and D. Niyato, "Completion time and energy optimization in the UAV-enabled mobile-edge computing system," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7808–7822, Aug. 2020.

[19] L. Zhang and N. Ansari, "Latency-aware IoT service provisioning in UAV-aided mobile-edge computing networks," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10573–10580, Oct. 2020.

[20] X. Zhang, Y. Zhong, P. Liu, F. Zhou, and Y. Wang, "Resource allocation for a UAV-enabled mobile-edge computing system: Computation efficiency maximization," *IEEE Access*, vol. 7, pp. 113345–113354, 2019.

[21] 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.

[22] H. Mei, K. Yang, Q. Liu, and K. Wang, "Joint trajectory-resource optimization in UAV-enabled edge-cloud system with virtualized mobile clone," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5906–5921, Jul. 2020.

[23] J. Zhang, L. Zhou, Q. Tang, E. C.-H. Ngai, X. Hu, H. Zhao, and J. Wei, "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.
[24] L. Zhang, Z. Zhao, Q. Wu, H. Zhao, H. Xu, and X. Wu, "Energy-aware dynamic resource allocation in UAV assisted mobile edge computing over social internet of vehicles," *IEEE Access*, vol. 6, pp. 56700–56715, 2018.
[25] L. Zhao, K. Yang, Z. Tan, X. Li, S. Sharma, and Z. Liu, "A novel cost optimization strategy for SDN-enabled UAV-assisted vehicular computation offloading," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3664–3674, Jun. 2021, doi: 10.1109/TITS.2020.3024186.
[26] 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.
[27] J. Xu, K. Ota, and M. Dong, "Big data on the fly: UAV-mounted mobile edge computing for disaster management," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 4, pp. 2620–2630, Oct. 2020.
[28] X. Hu, K.-K. Wong, and Y. Zhang, "Wireless-powered edge computing with cooperative UAV: Task, time scheduling and trajectory design," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8083–8098, Dec. 2020.
[29] 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.
[30] T. Zhang, Y. Xu, J. Loo, D. Yang, and L. Xiao, "Joint computation and communication design for UAV-assisted mobile edge computing in IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5505–5516, Aug. 2020.
[31] L. Yang, H. Yao, J. Wang, C. Jiang, A. Benslimane, and Y. Liu, "Multi-UAV-enabled load-balance mobile-edge computing for IoT networks," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6898–6908, Aug. 2020.
[32] E. T. Michailidis, N. I. Miridakis, A. Michalas, E. Skondras, and D. J. Vergados, "Energy optimization in dual-RIS UAV-aided MEC-enabled internet of vehicles," *Sensors*, vol. 21, no. 13, p. 4392, Jun. 2021.
[33] M. Zeng, W. Hao, O. A. Dobre, Z. Ding, and H. V. Poor, "Massive MIMO-assisted mobile edge computing: Exciting possibilities for computation offloading," *IEEE Veh. Technol. Mag.*, vol. 15, no. 2, pp. 31–38, Jun. 2020.
[34] T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, Mar. 2016, doi: 10.1109/MVT.2015.2496240.
[35] Y. Chen, L. Wang, Y. Ai, B. Jiao, and L. Hanzo, "Performance analysis of NOMA-SM in vehicle-to-vehicle massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2653–2666, Dec. 2017, doi: 10.1109/JSAC.2017.2726006.
[36] H. Jiang, Z. Zhang, J. Dang, and L. Wu, "A novel 3-D massive MIMO channel model for vehicle-to-vehicle communication environments," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 79–90, Jan. 2018, doi: 10.1109/TCOMM.2017.2751555.
[37] Z. Huang and X. Cheng, "A 3-D non-stationary model for beyond 5G and 6G vehicle-to-vehicle mmWave massive MIMO channels," *IEEE Trans. Intell. Transp. Syst.*, early access, Jun. 24, 2021, doi: 10.1109/TITS.2021.3077076.
[38] L. You, M. Xiao, X. Song, Y. Liu, W. Wang, X. Gao, and G. Fettweis, "Pilot reuse for vehicle-to-vehicle underlay massive MIMO transmission," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5693–5697, May 2020.
[39] E. T. Michailidis, N. Nomikos, P. Trakadas, and A. G. Kanatas, "Three-dimensional modeling of mmWave doubly massive MIMO aerial fading channels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1190–1202, Feb. 2020, doi: 10.1109/TVT.2019.2956460.
[40] Z. Guan, N. Cen, T. Melodia, and S. M. Pudlewski, "Distributed joint power, association and flight control for massive-MIMO self-organizing flying drones," *IEEE/ACM Trans. Netw.*, vol. 28, no. 4, pp. 1491–1505, Aug. 2020, doi: 10.1109/TNET.2020.2985972.
[41] S. Schwarz and M. Rupp, "Performance evaluation of low complexity double-sided massive MIMO transceivers," in *Proc. 13th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2016, pp. 582–588, doi: 10.1109/CCNC.2016.7444844.
[42] S. Buzzi and C. D'Andrea, "Energy efficiency and asymptotic performance evaluation of beamforming structures in doubly massive MIMO mmWave systems," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 2, pp. 385–396, Jun. 2018, doi: 10.1109/TGCN.2018.2800537.
[43] D.-W. Yue and H. H. Nguyen, "On the multiplexing capability of mmWave doubly-massive MIMO systems in LOS environments," *IEEE Access*, vol. 7, pp. 126973–126984, 2019, doi: 10.1109/ACCESS.2019.2940078.
[44] L. N. Ribeiro, S. Schwarz, and A. L. F. De Almeida, "Double-sided massive MIMO transceivers for mmWave communications," *IEEE Access*, vol. 7, pp. 157667–157679, 2019.
[45] W. Feng, J. Xu, and Z. Ding, "Multi-antenna NOMA for computation offloading in multiuser mobile edge computing systems," *IEEE Trans. Commun*, vol. 67, no. 3, pp. 2450–2463, Mar. 2019.
[46] T. T. Nguyen, L. Le, and Q. Le-Trung, "Computation offloading in MIMO based mobile edge computing systems under perfect and imperfect CSI estimation," *IEEE Trans. Services Comput.*, early access, Jan. 14, 2019, doi: 10.1109/TSC.2019.2892428.
[47] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, "Delay minimization for massive MIMO assisted mobile edge computing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6788–6792, Jun. 2020.
[48] W. Feng, J. Zheng, and W. Jiang, "Joint pilot and data transmission power control and computing resource allocation algorithm for massive MIMO-MEC networks," *IEEE Access*, vol. 8, pp. 80801–80811, 2020.
[49] R. Malik and M. Vu, "Energy-efficient computation offloading in delay-constrained massive MIMO enabled edge network using data partitioning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6977–6991, Oct. 2020.
[50] Y. Zhao, X. Xu, Y. Su, L. Huang, X. Du, and N. Guizani, "Multi-user MAC protocol for WLANs in mmWave massive MIMO systems with mobile edge computing," *IEEE Access*, vol. 7, pp. 181242–181256, 2019.
[51] S. Mukherjee and J. Lee, "Edge computing-enabled cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2884–2899, Apr. 2020.
[52] P. Chandhar, D. Danev, and E. Larsson, "Massive MIMO for communications with drone swarms," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1604–1629, Mar. 2018.
[53] *Study on Channel Model for Frequencies From 0.5 to 100 GHz*, Standard 38.901, Version 16.1.0, 3GPP, Sophia Antipolis, France, Tech. Rep., Dec. 2019.
[54] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
[55] W. Khawaja, I. Guvenc, D. W. Matolak, U.-C. Fiebig, and N. Schneckenburger, "A survey of air-to-ground propagation channel modeling for unmanned aerial vehicles," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2361–2391, 3rd Quart., 2019.
[56] D. W. Matolak and R. Sun, "Air–ground channel characterization for unmanned aircraft systems—Part I: Methods, measurements, and models for over-water settings," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 26–44, Jan. 2017.
[57] R. Sun and D. W. Matolak, "Air–ground channel characterization for unmanned aircraft systems—Part II: Hilly and mountainous settings," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 1913–1925, Mar. 2017.
[58] D. W. Matolak and R. Sun, "Air–ground channel characterization for unmanned aircraft systems—Part III: The suburban and near-urban environments," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6607–6618, Aug. 2017.
[59] W. Zhang, Y. Wen, K. Guan, D. Kilper, H. Luo, and D. O. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4569–4581, Sep. 2013.
[60] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, Mar. 2004.
[61] M. Grant, S. Boyd, and Y. Ye. (2013). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.0 Beta. 2013*. Accessed: Jul. 19, 2021. [Online]. Available: http://cvxr.com/cvx
[62] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart., 2017.
[63] W. Khawaja, O. Ozdemir, and I. Guvenc, "UAV air-to-ground channel characterization for mmWave systems," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Toronto, ON, Canada, Sep. 2017, pp. 1–5.
[64] D. Robbins, C. Cholas, M. Brennan, and K. Critchley, "Augmented and virtual reality for service providers," Revision 1.0, Immersive Media Bus. Brief, Intel Corp., Santa Clara, CA, USA, Tech. Rep. 337010001US, 2017. [Online]. Available: https://builders.intel.com/docs/networkbuilders/augmented-and-virtual-reality-for-service-providers.pdf
[65] S. Mangiante, G. Klas, A. Navon, Z. GuanHua, J. Ran, and M. D. Silva, "VR is on the edge: How to deliver 360° videos in mobile networks," in *Proc. Workshop Virtual Reality Augmented Reality Netw.*, Aug. 2017, pp. 30–35.

**EMMANOUEL T. MICHAILIDIS** was born in Athens, Greece. He received the M.Sc. degree in digital communications and networks from the University of Piraeus, Piraeus, Greece, in 2006, and the Ph.D. degree in broadband aerospace communications from the University of Piraeus, in 2011. Since 2018, he has been an Adjunct Lecturer with the Department of Electrical and Electronics Engineering, School of Engineering, University of West Attica, Egaleo, Greece. He has published more than 50 scientific articles and received several best paper awards in his areas of research. His current research interests include the channel modeling and performance analysis of next-generation terrestrial wireless, aerial, and satellite communication systems.

**NIKOLAOS I. MIRIDAKIS** (Senior Member, IEEE) was born in Athens, Greece, in 1982. He received the M.Sc. degree in networking and data communications from the Department of Information Systems, Kingston University, U.K., in 2008, and the Ph.D. degree in networking and data communications from the Department of Informatics, University of Piraeus, Greece, in 2012. Since 2012, he has been with the Department of Informatics, University of Piraeus, Greece, where he is a Senior Researcher. Since 2018, he has been with the School of Electrical and Information Engineering and the Institute of Physical Internet, Jinan University, Zhuhai, China, as a Distinguished Research Associate. He is currently an Assistant Professor with the Department of Informatics and Computer Engineering, University of West Attica, Egaleo, Greece. His main research interests include wireless communications, and more specifically interference analysis and management in wireless communications, multicarrier communications, MIMO systems, statistical signal processing, diversity reception, fading channels, and cooperative communications. He serves as a reviewer and a TPC member for several prestigious international journals and conferences. He was also recognized as an Exemplary Reviewer by IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and *Physical Communication* (Elsevier), in 2017. Since 2019, he serves as an Associate Editor for the IEEE COMMUNICATIONS LETTERS.

**ANGELOS MICHALAS** (Member, IEEE) is currently a Professor with the Electrical and Computer Engineering Department, University of Western Macedonia. Since 2015, he has been the Director of the M.Sc. Program ''Modern Information Technologies and Services'' co-organized from the Department of Informatics, UOWM, and the Department of Informatics, University of Piraeus. His research interests include design and performance evaluation of wired and wireless networks, advanced distributed network services, and quality of service/experience. He has participated in a large number of European and national research programs in these areas. He is an author of several peer-reviewed publications in these areas and serves as a technical program committee member and a reviewer in international conferences and journals. He is a member of scientific unions, including the ACM and the Technical Chamber of Greece.

**EMMANOUIL SKONDRAS** received the B.Sc. degree in informatics and computer technology from the Technological Educational Institution of Western Macedonia, the M.Sc. degree in computer science from the Department of Informatics, Athens University of Economics Business, and the Ph.D. degree in next generation wireless networks from the Department of Informatics, University of Piraeus. Additionally, he is an Adjunct Lecturer in M.Sc. program with the University of Piraeus and the University of Western Macedonia, while at the same time, he is working as a ICT Specialist with the Network Operations Center (NOC), University of Piraeus. He has also worked in research and development projects of information technology industry. He is currently a Postdoctoral Researcher in next generation wireless networks with the Department of Informatics, University of Piraeus. His research interests include sensor and communication networks, mainly in wireless networks, such as flying *ad hoc* networks (FANETs), vehicular *ad hoc* networks (VANETs), cellular, and mesh networks. He received two scholarships and one award in respect of his high degrees from the Technological Educational Institution of Western Macedonia.

**DIMITRIOS J. VERGADOS** was born in 1980. He received the B.Eng. and Ph.D. degrees from the School of Electrical and Computer Engineering, National Technical University of Athens (NTUA), Greece, in 2003 and 2009, respectively. He is currently an Assistant Professor with the Department of Informatics, University of Western Macedonia, Greece, and a Postdoctoral Researcher with the School of Electrical and Computer Engineering, NTUA. He has been employed in research projects and a visiting lecturer in several universities in Greece and Norway. His research interests include distributed systems, wireless networks, simulation modeling, scheduling algorithms, multihop networks, and smart-grids. He has authored several publications in these areas.

**DIMITRIOS D. VERGADOS** (Senior Member, IEEE) is currently a Professor and the Deputy Head with the Department of Informatics, University of Piraeus, and the Director of the M.Sc. Program Advanced Informatics and Computing Systems. He has significant research experience related to the project proposal. In particular, he has participated in several research projects and has published more than 220 publications (journals with high impact factor, books, conferences proceedings, and symposia) in research topics related with the project like wireless communications, wireless sensor networks, the Internet of Things and smart grid cloud computing, and security in information systems and telecommunications, and has more than 3200 citations in his research work (Source Scholar Google: H-index 23, i10-index 54). He has participated in many conferences (more than 60) (in the majority presenting his research work) and moreover he has also given more than 20 invited talks. He has also visited several European Universities through Staff Mobility of Erasmus/Erasmus+ Network Program. He has participated in several research projects funded by EU and National Agencies. He has served as a committee member and an evaluator and a reviewer in national and international organizations and agencies. He was also served as a Member of the Board of the Hellenic Telecommunications and Post Commission (EETT) (2013–2017) and an Invited Distinguished Trainer/Speaker of Europol at the 2nd Training Course on Payment Card Fraud Forensics and Investigations (July 2016) (See also Commissions of Trust and PI Achievements Document). He has been a member of the technical program committee in several international conferences (more than 100), since 2007. He is a guest editor and a member of the editorial board and a reviewer in several journals and conferences.