Location-Dependent Augmented Reality Services in Wireless Edge-Enabled Metaverse Systems

HONGLIANG ZHANG [1] (Member, IEEE), SHIWEN MAO [2] (Fellow, IEEE),
DUSIT NIYATO [3] (Fellow, IEEE), AND ZHU HAN [4,5] (Fellow, IEEE)

[1]School of Electronics, Peking University, Beijing 100871, China

[2]Department of Electrical and Computer Engineering, Auburn University, AL 36849, USA

[3]School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798

[4]Department of Electrical and Computer Engineering, University of Houston, TX 77004, USA

[5]Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea

CORRESPONDING AUTHOR: H. ZHANG (e-mail: hongliang.zhang@pku.edu.cn)

ABSTRACT Metaverse is envisioned as the next-generation paradigm that links a virtual world to the physical world, providing immersive experience for human activities in the hypothetical environment with augmented reality (AR) and virtual reality (VR) technology. Recent development of wireless multi-access edge computing (MEC) has accelerated the realization of this vision. In this paper, we consider location-dependent AR services in MEC-enabled Metaverse systems. In such a system, each user performs simultaneous localization and communication first to estimate its position and request AR contents from the MEC server. Due to limited computation and communication resources, the MEC server adaptively adjusts the resolution of AR contents. However, two challenges arise in such a system. First, the localization error will degrade the user's quality-of-experience (QoE); and second, the coupling of communication and computation complicates the resource management. To address these issues, we first model the QoE taking the localization errors into consideration, and then propose a minimum QoE maximization scheme to achieve fairness among users by jointly optimizing the waveform at users, the transmit power at the base station (BS), and the resolution of transmitted videos. Simulation results verify the effectiveness of the proposed scheme in comparison with benchmark schemes.

INDEX TERMS Metaverse, location-dependent augmented reality, multi-access edge computing, resource management.

I. INTRODUCTION

METAVERSE is an emerging concept that describes how a hypothetical world is connected to the real world and is considered as a potential disruption technology for the next generation Internet [1], [2]. Facebook was rebranded as "Meta" in particular to demonstrate its dedication to developing the Metaverse technology. Due to recent changes in social interaction, users of the Metaverse expect to enjoy immersive experiences in the virtual world provided by augmented reality (AR) and virtual reality (VR) services [3]. First, since online meetings and instructions are becoming more common as a result of the pandemic, it will be crucial to offer immersive interactions in the digital world. Second, users may become "super heroes" by connecting all relevant information with the physical world using an AR glass to react to an event promptly [4]. For example, we can see the participant's LinkedIn profile when meeting a new people in a conference and you can know his background immediately.

Given the aforementioned use cases, it is conceivable that the Metaverse will be an *essential* component in our daily lives soon.

On the other hand, the recent development of wireless communications and multi-access edge computing (MEC) technology [5], which allows users to access the virtual environment whenever and wherever they are, has become a key enabler for the growth of the Metaverse [6]. Particularly, the Third Generation Partnership Project (3GPP) has begun standardizing the AR/VR support in New Radio (NR). A Release 17 study on NR performance for AR/VR services was undertaken [7], and the work will continue in Release 18 (the first 5G-Advanced release), with the goal of improving the performance of AR/VR services over the NR. Moreover, the wireless MEC-empowered Metaverse can offer seamless immersive experiences with lower latency, due to the faster response of the MEC than cloud computing [8].

In the literature, there are a few studies working on the wireless MEC-enabled Metaverse [9], [10] or AR applications [11], [12], [13], [14]. In [9], an advertising strategy was studied for VR content transmission in the wireless MEC-enabled Metaverse from the perspective of preserving privacy. In [10], a learning-based incentive mechanism was proposed for trading VR services between users and providers in the wireless MEC-enabled Metaverse. These two works considered VR services without physical location information, and thus cannot be applied to the AR services considered in this paper. In [11], the authors proposed to insert an edge layer between the user and cloud layers to deliver a better QoE for AR services. In [12], an energy-efficient resource allocation approach over both communication and computation resources was proposed for MEC-enabled AR applications. The authors in [13] investigated a server assignment and resolution selection problem to achieve the latency-accuracy trade-off. The work in [14] studied the popular content caching and performing computation at the MEC. However, these methods typically assume that the location estimations for users are perfect and neglect the impact of localization errors on the QoE, which motivates this paper.

In this paper, we consider location-dependent AR services in wireless MEC-assisted Metaverse systems as the virtual AR contents should be consistent wit the physical environment.[1] In the Metaverse system, two phases are performed to transmit AR contents: uplink and downlink. In the *uplink* phase, each user first performs simultaneous localization and communication to estimate its position and request AR contents from the MEC server according to its estimated position. However, the data volume of the entire virtual objects might be too large to be transmitted within one time slot. In the *downlink* phase, the MEC server

needs to select part of the AR content, and adjust its resolution to transmit based on the available computing and communication resources, aiming to maximize the user's Quality-of-Experience (QoE). However, two challenges arise to manage the resources in the Metaverse system. *First*, localization estimation errors are inevitable, which might degrade the QoE of the user. For example, one should place the virtual dolphin in the pool but the location errors might lead to the dolphin on the land. Therefore, it is important to consider the localization errors in the resolution adjustment. *Second*, compressing more data to save communication resources needs more computation resources. The coupling of communication and computation makes the resource management not trivial.

To address these challenges, we first model the QoE by taking location estimation errors into consideration, which are modeled with a known distribution within a given range. The distribution determines the probability of the field of view (FoV) of each user for the AR video. Next, we formulate a minimum QoE maximization problem for fairness among users by optimizing the waveform in the uplink phase, and the transmit power at the base station (BS) as well as the resolution of transmitted videos. To tackle this problem, we decouple the problem into two subproblems: the uplink and downlink subproblems, and propose a waveform design algorithm for the uplink subproblem and a joint resolution and transmit power optimization algorithm for the downlink subproblem.

Our contributions can be summarized as follows:

- We study a wireless MEC-enabled Metaverse system for location-dependent AR services with multiple users. Limited by the communication and computing resources owned by the system, we formulate a minimum QoE maximization problem by properly scheduling these resources for fairness among users, where the physical locations of users are considered in the QoE model.
- To address this problem, we decouple the problem into two subproblems: uplink and downlink subproblems. We propose a waveform design algorithm for the uplink subproblem and a joint resolution and transmit power optimization algorithm for the downlink subproblem.
- We perform extensive simulations for validation. Simulation results show that the proposed scheme outperforms the benchmark schemes in terms of QoE. The estimation error is shown to be necessary to be considered in the data transmission error, which can effectively improve the system QoE in the Metaverse.

The rest of this paper is organized as follows. In Section II, we introduce the system model for the wireless MEC-enabled Metaverse system. In Section III, we first introduce the QoE model to quantify the performance of the Metaverse system, and then formulate a minimum QoE maximization problem. Our proposed algorithm is presented in Section IV. Simulation results are shown in Section V, and finally conclusions are drawn in Section VI.

---

[1]. The location information should be considered in AR services because the provided AR contents need to be consistent with the physical world. For example, presenting a shark on a beach will confuse the Metaverse user, and thus degrading the QoE of Metaverse. Therefore, accurate localization will help maintain the consistency.
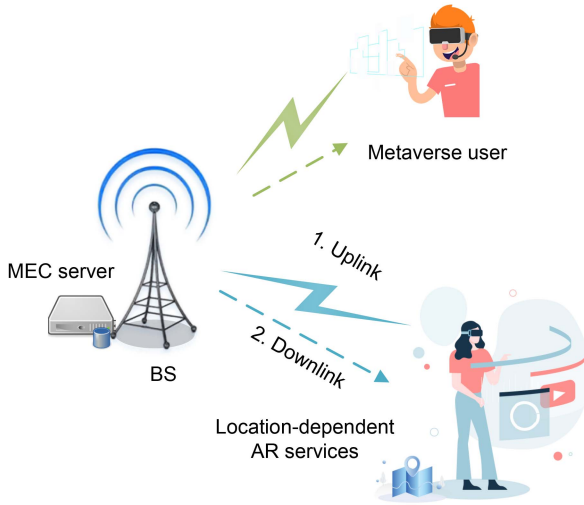
FIGURE 1. System model of wireless MEC-enabled Metaverse systems for location-dependent AR services.

Notation: Boldface lower and upper case symbols denote vectors and matrices, respectively. $\mathbb{C}^{M \times N}$ denotes a complex matrix with dimension $M \times N$. Conjugate transpose is denoted by $(\cdot)^H$, transpose is denoted by $(\cdot)^T$, and the trace of matrix $\boldsymbol{A}$ is denoted by $\text{Tr}(\boldsymbol{A})$. Expectation operator is denoted by $\mathbb{E}[\cdot]$. Matrix $\boldsymbol{A} \succeq \boldsymbol{B}$ indicate that $\boldsymbol{A} - \boldsymbol{B}$ is positive semi-definite.

II. SYSTEM MODEL
In Section II-A, we present the scenario of a wireless MEC-enabled Metaverse system for location-dependent AR services, and introduce its uplink and downlink phases. In Section II-B, we describe the simultaneous localization and communication models in the uplink phase. In Section II-C, we formulate the communication and computation models in the downlink phase.

A. SCENARIO DESCRIPTION
We study a wireless MEC-enabled Metaverse system for location-dependent AR services, as depicted in Fig. 1. In this system, there exists a BS equipped with M antennas, denoted by $\mathcal{M} = \{1, \ldots, M\}$, and an MEC server is installed at the BS to provide AR services for Metaverse users. K Metaverse users, denoted by $\mathcal{K} = \{1, \ldots, K\}$, run AR applications and periodically request videos from the BS. Suppose that each user's AR device has N antennas. The system owns K channels, the bandwidth of which is W. To avoid the mutual interference, each channel is used to serve one user.

The system is assumed to be well synchronized. To be specific, the timeline is divided into time slots and we have the following two phases in each time slot:

- *Uplink phase:* In this phase, the user performs simultaneous localization and transmission. Specifically, the AR device will transmit wireless signals to update its own position and orientation. At the same time, it will

also transmit a video update request according to the estimated results.
- *Downlink phase:* After the uplink phase, the MEC server processes the video, and then the BS transmits the processed video back to the user. They need to take the computation and communication resources into consideration when processing these videos.

It is important to note that the virtual objects will remain unchanged within the considered time window. In other words, the total amount of data to be transmitted within the time window is fixed. However, for certain AR applications, the total data volume could be large or the number of users can be large [15]. Therefore, we consider an adaptive data transmission approach for AR data, where the data cannot be fully transmitted within a time slot and only a portion of the data can be transmitted. Moreover, the data transmitted in the previous time slots will be cached in the headset and there is no need to re-transmit it again in the following time slots. Details of these two phases will be presented in the following two subsections. Note that these two phases are for each time slot t, and we omit the index t in the following unless mentioned otherwise.

B. UPLINK PHASE
In the uplink phase, each user will send wireless signals for simultaneous communications and localization, as elaborated below.

1) COMMUNICATION
To guarantee that the BS can receive video request from each user, the received signal-to-noise ratio (SNR) should be larger than a predefined threshold. To be specific, denote the channel between the BS and user k as $\boldsymbol{H}_k \in \mathbb{C}^{N \times M}$ and the transmitted signals from user k as $\boldsymbol{x}_k \in \mathbb{C}^{(A+1) \times 1}$, where A is the number of single-antenna anchor nodes for localization. We also define the beamformer for user k as $\boldsymbol{W}_k \in \mathbb{C}^{N \times (A+1)}$. Therefore, the received signal at the BS from user k can be expressed as [16]

$$\boldsymbol{y}_k = \boldsymbol{H}_k^T \boldsymbol{W}_k \boldsymbol{x}_k + \boldsymbol{n}_k, \tag{1}$$

where $\boldsymbol{n}_k$ is the addictive Gaussian noise with zero mean and variance σ_c^2.

Let $|\boldsymbol{x}_k|^2 = 1$, and the SNR constraint can be expressed as

$$\gamma_k = \frac{(\boldsymbol{H}_k^T \boldsymbol{W}_k)(\boldsymbol{H}_k^T \boldsymbol{W}_k)^H}{\sigma_c^2} \geq \eta^k, \tag{2}$$

where η_k is the SNR threshold for AR user k, which is specific for the AR service provided.

Moreover, subject to the power budget for each user, the beamformer vector should satisfy

$$\text{Tr}\left(\boldsymbol{W}_k \boldsymbol{W}_k^H\right) \leq P_{max}, \tag{3}$$

where P_{max} is the power budget for each user.

2) LOCALIZATION

At the same time, each user also generate beams for localization estimation [17]. Let the channel between user k and anchor node i be $\boldsymbol{h}_{ik} \in \mathbb{C}^{1 \times N}$, and thus, the received signal can be expressed as

$$y_{ik} = \boldsymbol{h}_{ik} \boldsymbol{W}_k \boldsymbol{x}_k + n_{ik}, \tag{4}$$

where n_{ik} is also the addictive Gaussian noise with zero mean and power σ^2. The SNR for at each anchor node can be expressed as

$$\gamma_{ik} = \frac{(\boldsymbol{h}_{ik} \boldsymbol{W}_k)(\boldsymbol{h}_{ik} \boldsymbol{W}_k)^H}{\sigma^2}. \tag{5}$$

Each user can gather the signals received at these anchor nodes and estimate the location from these signals. However, the estimation will not be perfect and will definitely have estimation errors. To be specific, the location estimation error can be expressed as

$$e_k = \mathbb{E}\left[|\boldsymbol{l}_k - \tilde{\boldsymbol{l}}_k|^2 \right], \tag{6}$$

where $\tilde{\boldsymbol{l}}_k$ is the estimated location and $\boldsymbol{l}_k$ is the ground-truth. For any unbiased estimator of the user's position, the estimation error is bounded by the Cramer-Rao bound [18]. To be specific, the estimation error can be approximated as

$$e_k \approx \mathrm{tr}\left(\boldsymbol{J}_k^{-1} \right), \tag{7}$$

where $\boldsymbol{J}_k$ is the Fisher information matrix (FIM). According to the results in [19], the FIM can be defined as

$$\boldsymbol{J}_k = \frac{8\pi^2 W^2}{c^2} \sum_{i=1}^{A} \gamma_{ik} \boldsymbol{q}_{ik} \boldsymbol{q}_{ik}^T, \tag{8}$$

where c is the propagation speed, and $\boldsymbol{q}_{ik} = [\cos \phi_{ik}, \sin \phi_{ik}]^T$ with ϕ_{ik} being the angle between user k and anchor i, which is a constant obtained by estimating the angle of arrival (AoA).

For orientation, we assume that the headset is equipped with an accelerometer, gyroscope, and magnetometer [20], and the orientation can be obtained through these sensors, which is assumed to be highly accurate.

C. DOWNLINK PHASE

In the downlink phase, the BS needs to transmit the data of virtual objects to users. However, limited by the communication resources, the BS might not be able to transmit all the required data to users within one time slot. Instead, the BS only sends the data which might be within the FoV of the user relating to the current position, and thus the transmitted data will change according to the movement (including position and orientation) of the user, as shown in Fig. 2. On the other hand, even though the BS only transmits the data of the FOV, the bandwidth might not be sufficient for the highest resolution. Therefore, the BS will transmit a portion of the video according to the available communication
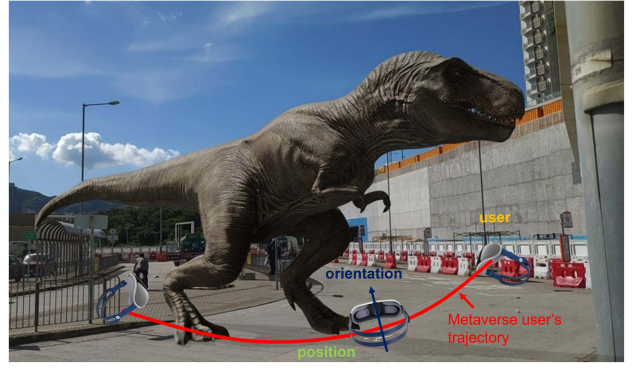

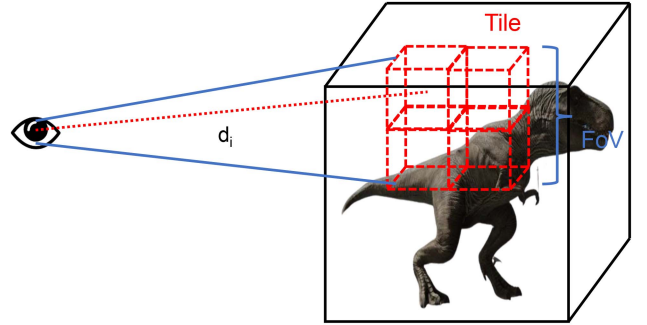**FIGURE 2.** Adaptive AR communications under Metaverse user's movement.

FIGURE 3. Model of a video and illustration of FoV.

resources by adjusting the resolution of the video. As different parts of a video contribute differently to the QoE, their corresponding resolutions might vary. In the following, we first explain how the system adjusts the resolution of the video.

As shown in Fig. 3, the video for each user is partitioned into I 3D tiles and only part of these tiles will fall inside the user's FoV. Here, we define $\mathcal{I}^k = \{1, \ldots, I_k\}$ as the set of tiles for user k, and I_k is the total number of the tile. However, at a certain position and orientation, as different tiles contribute differently to the user's QoE, the resolution for different tiles should vary. For user k, define the resolution for the i-th tile as $\beta_i^k \in (0, B]$, where B is the highest resolution that a tile can be [21]. Therefore, the resolution for each tile should be adjusted according to the available computation and communication resources. It is worth pointing out that the highest resolution level for each tile will vary in different time slots. To be specific, the transmitted video will be cached in the headset, and thus those data does not need to be retransmitted. Based on these facts, we elaborate on the computation and communication models in the following.

1) COMMUNICATION MODEL

The volume of transmitted data for each tile is proportional to the resolution of this tile [22]. Define κ as the coefficient, and thus, the data traffic to be transmitted to user k can be

expressed as

$$D_k = \sum_{i \in \mathcal{I}^k} \kappa \beta_i^k. \qquad (9)$$

Let p_m^k be the transmit power of antenna $m \in \mathcal{M}$ for user k, and $\boldsymbol{p}^k = [p_1^k, \ldots, p_M^k]^T$ be the transmit power vector at the BS for user k. The transmission rate between the BS and user k can be expressed as

$$R_k = W \log_2 \left(1 + \frac{\boldsymbol{H}^k (\boldsymbol{p}^k)^{1/2} \left(\boldsymbol{H}^k (\boldsymbol{p}^k)^{1/2} \right)^H}{\sigma^2} \right), \qquad (10)$$

where σ^2 is the power of the additive Gaussian white noise. To guarantee that all the data can be transmitted to the user, the data rate should meet the following constraint:

$$D_k \le R_k. \qquad (11)$$

Moreover, the transmit power of each antenna over all the channels cannot exceed its energy budget, i.e.,

$$\sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} p_m^k \le P, \qquad (12)$$

where P is the energy budget at the BS.

2) COMPUTATION MODEL

To adjust the video resolution, the edge server needs to utilize its central processing units (CPUs). Let the number of overall CPU cycles to process the videos be V within a time slot. According to the video model shown in Fig. 3, for user k, the original data volume is $\kappa I_k B$ while the transmitted volume is D_k given in (9). Therefore, the compressed data volume is

$$C_k = \kappa I_k B - D_k. \qquad (13)$$

It should be noted that the maximum resolution B varies across different slots. When more data is cached in the headset, the required maximum resolution will decrease.

The required number of CPU cycles is linear with respect to the compressed data volume [23]. Let ν be the data bits that can be processed in a CPU cycle, the required number of CPU cycles for user k can be expressed as

$$\mu_k = \frac{C_k}{\nu}. \qquad (14)$$

Therefore, to ensure that all the computation tasks can be completed on time, we have

$$\sum_{k \in \mathcal{K}} \mu_k \le V. \qquad (15)$$

III. PROBLEM FORMULATION

In Section III-A, we first elaborate on how the estimation error influences the QoE. In Section III-B, we introduce the QoE model to quantify the performance of the Metaverse system. Finally, we formulate the minimum QoE maximization problem in Section III-C.
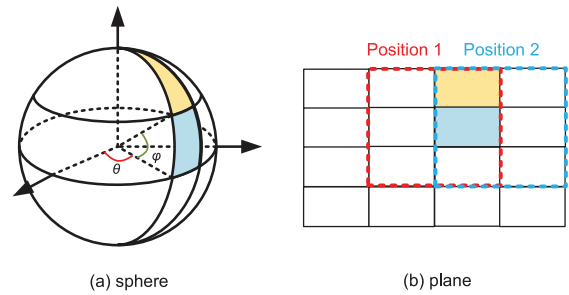

FIGURE 4. Spherical mapping of tiles.

A. IMPACT OF ESTIMATION ERROR

As we have introduced in Section II-B2, the localization estimation cannot be perfect and will introduce some errors here. As a result, the estimation error will affect the probability that each tile falls within the FOV. In the following, we show the details about how the estimation error influences the probability.

In this paper, we assume that the FoV consists of $b_w \times b_h$ tiles. However, as shown in Fig. 4, the estimation error of position will cause the shift of the FOV, where the yellow and blue ones indicate different tiles in spherical and planar representations. Therefore, to quantify the impact of the estimation error, we use the probability for each tile that falls in the FOV. Intuitively speaking, if the estimation error is smaller, the tiles within the FOV associated with the estimated position will have a higher probability to be shown in the FOV.

We assume that the ground-truth position is within the error range centered at the estimated position [24]. To be specific, we divide the range of location errors into several discrete levels so that the corresponding FOV moves left (right) or up (down), as illustrated by positions 1 and 2 in Fig. 4. Moreover, the orientation including azumith and evaluation angles will also influence the shift of the FOV as discussed below.

Without loss of generality, the position error range is divided into L levels with fixed gap ΔL, denoted by $\mathcal{L}$. We define the azumith and evaluation angles for user k as θ_k and φ_k, respectively. We also consider that the position follows a certain known distribution within the error range, denoted by $\boldsymbol{f}^p$. For tile i, we use x_l^i to indicate whether this tile is within FOV[2] l under azumith angle θ_k and evaluation angle φ_k, where

$$x_l^i = \begin{cases} 1, & \text{if tile } i \text{ is within FOV } l; \\ 0, & \text{otherwise.} \end{cases} \qquad (16)$$

Therefore, the probability that tile i falls in the FOV can be expressed as

$$f_i = \sum_{l \in \mathcal{L}} x_l^i f_l^p, \qquad (17)$$

2. This indicates that the FOV is associated with the l-th position in $\mathcal{L}$.

where f_l^p is the probability that the l-th location is the ground-truth.

Remark 1: The estimation range is a circle plane, and the FOV shift caused by the estimation error of the position is horizontal, as shown in Fig. 4.

Remark 2: It should be noted that the number of tiles I_k is highly related to the location estimation error. In general, the higher estimation error, the more tiles should be transmitted, leading to degradation of resolution.

B. QOE MODEL

According to the previous discussions, different tiles contribute to the total QoE differently since its probability to be presented in the FOV varies. Before presenting the optimization problem, we need to first model the QoE as contributed by each tile.

According to the results in [25], the QoE contributed by tile i for user k can be expressed as

$$Q_i^k = \omega_i^k \left(\beta_i^k + \bar{\beta}_i^k \right) f_i^k. \tag{18}$$

Here, ω_i^k is the quality weight for the tile, which can be obtained from historical data [33]. In other words, this parameter indicates the importance of tile i in the video for user k. $\bar{\beta}_i^k$ is the data of tile i that has been stored in the headset, and f_i^k is the probability that tile i falls within the FoV of user k, which is calculated by (17).

According to the definition in (18), the QoE for videos with different resolutions will be different, leading to unfairness among users. To reduce the QoE variation caused by the resolution, we define the QoE for user k as

$$Q^k = \tilde{Q} + s \log_2 \left(\frac{\sum\limits_{i \in \mathcal{I}^k} Q_i^k}{\sum\limits_{i \in \mathcal{I}^k} \omega_i^k \left(B + \bar{\beta}_i^k \right) f_i^k} \right), \tag{19}$$

where $\tilde{Q}$ is the constant representing the maximum QoE that the system can achieve and s is a scaling factor.

C. PROBLEM STATEMENT

As communication and computation resources are limited, it is not possible to guarantee the QoE for all the users. For fairness among these users, we aim to maximize the minimum QoE among all the users, which aims to align the QoE's of different users under limited resource constraints [26]. Mathematically, the problem can be written as

$$\max_{\{\beta_i^k\}, \{p^k\}, \{W_k\}} \quad \min_{k \in \mathcal{K}} Q^k, \\ s.t. \quad (2), (3), (6), (11), (12), \text{ and } (15). \tag{20}$$

Problem (20) is a max-min problem, which is not easy to deal with. To address this issue, we first introduce an auxiliary variable η, and replace the objective as the maximization of ϵ. To guarantee that ϵ is the minimum of Q^k among $\mathcal{K}$, we need to introduce the following constraint:

$$\epsilon \le Q^k, \quad \forall k \in \mathcal{K}. \tag{21}$$

Moreover, according to Remark 2, the uplink and downlink phases are coupled, which makes the problem hard to solve. To effectively solve the original problem, we decouple it into two subproblems: uplink and downlink subproblems. In this way, the estimation errors are obtained in the uplink phase, and thus, the number of tiles to be transmitted will be fixed in the downlink phase, making the problem easier to tackle.

1) Uplink Subproblem: Before introducing the uplink subproblem, we would like to give a proposition to show how the QoE changes with a larger estimation error range.

Proposition 1: Under a scaled distribution, i.e., the mean and variance of the distribution are multiplied by the same scaling factor by which the range of the estimation error is increased, the QoE will decrease with a larger estimation error range.

Proof: Please see Appendix A. ∎

Based on this proposition, in the uplink phase, the subproblem can be transformed into the minimization of the estimation error range. Mathematically, as each Metaverse user optimizes its waveform independently, the optimization problem can be written as

$$\min_{W_k} \quad e_k, \\ s.t. \quad (2) \text{ and } (3). \tag{22}$$

2) Downlink Subproblem: In this subproblem, we can obtain the localization error range according to the waveform optimized in the uplink phase $\{W_k\}$. After that, the BS needs to determine the transmit power for each antenna and the resolution for each tile. Mathematically, the optimization problem can be written as

$$\min_{\{\beta_i^k\}, \{p^k\}, \epsilon} \quad -\epsilon, \\ s.t. \quad (11), (12), (15), \text{ and } (21). \tag{23}$$

It is worth pointing out that these two subproblems are not solved iteratively. Instead, these two subproblems are solved one by one in a time slot. This is because the joint optimization will require frequent signaling between Metaverse users and the BS as the uplink phase is performed by the Metaverse user while the downlink phase is executed by the MEC server and the BS.

IV. ALGORITHM DESIGN

In this section, we elaborate on how to solve the aforementioned two subproblems in the following two subsections, respectively.

A. WAVEFORM DESIGN FOR THE UPLINK SUBPROBLEM

In this subproblem, we aim to minimize the estimation error under the SNR constraint and power budget. It is easy to check that the constraints are convex. Therefore, we elaborate on how to tackle with objective e_k.

According to [27], such a problem can be converted into a semi-definite program (SDP). To be specific, we

first introduce an auxiliary matrix Z_k and introduce a new constraint as:

$$Z_k \succeq J_k^{-1}. \tag{24}$$

Since J_k is a positive semi-definite matrix, due to the property of Schur complement, the inequality is equivalent to

$$\begin{bmatrix} Z_k & I \\ I & J_k \end{bmatrix} \succeq 0. \tag{25}$$

Therefore, the problem for user k can be rewritten as

$$\min_{W_k, Z_k} \quad \text{Tr}(Z_k),$$
$$s.t. \quad (2), \ (3), \ \text{and} \ (25). \tag{26}$$

Such a problem is an SDP, which can be solved by existing convex optimization techniques [28].

Remark 3: The computational complexity to solve this subproblem is $\mathcal{O}(KN^6)$ according to [29]. This shows that the subproblem can be solved within a polynomial time.

B. JOINT RESOLUTION AND TRANSMIT POWER OPTIMIZATION FOR THE DOWNLINK SUBPROBLEM

In this subproblem, we use the waveform obtained in the uplink subproblem to calculate the estimation error range. In other words, the probability that each tile falls within the FOV f_i^k is constant in this subproblem as defined in Section III-A.

As the constraints define a convex set, the problem can be solved by existing convex optimization techniques. In the following, we introduce the algorithm with more details.

Let λ_1^k, λ_2^m, λ_3, and λ_4^k be the Lagrangian multipliers corresponding to (11), (12), (15), and (21), respectively. Therefore, the Lagrangian can be expressed as

$$U = -\epsilon + \sum_{k \in \mathcal{K}} \lambda_1^k (D_k - R_k) + \sum_{m \in \mathcal{M}} \lambda_2^m \left(\sum_{k \in \mathcal{K}} p_m^k - P \right)$$
$$+ \lambda_3 \left(\kappa IBK - \sum_{k \in \mathcal{K}} D_k - \nu V \right) + \sum_{k \in \mathcal{K}} \lambda_4^k (\epsilon - Q_k), \tag{27}$$

and the Lagrangian dual problem can be written as

$$\max_{\{\lambda_1^k\}, \{\lambda_2^m\}, \lambda_3, \{\lambda_4^k\}} \min_{\{\beta_i^k\}, \{p^k\}, \epsilon} U. \tag{28}$$

We can solve the dual problem iteratively by decomposing it into master and slave subproblems as follows:

a) Slave subproblem: According to the Karush-Kuhn-Tucker (KKT) conditions, we can obtain the optimal solution by equaling the first partial derivative of the Lagrangian function U over each variable to 0. Specifically, we have $\frac{\partial U}{\partial p_m^k} = 0$, $\frac{\partial U}{\partial \beta_i^k} = 0$, and $\frac{\partial U}{\partial \epsilon} = 0$. After some transformation, we can have

$$\text{Tr}\left(\left(H^k \right)^H P \left(H^k \right) \right) = \frac{\lambda_1^k W \left(h_m^k \right)^H h_m^k}{\lambda_2^m} - \sigma^2, \tag{29}$$

$$\sum_{i \in \mathcal{I}^k} \omega_i^k \left(\beta_i^k + \bar{\beta}_i^k \right) f_i^k = \frac{(\lambda_1^k - \lambda_3)\kappa}{\lambda_4^k \omega_i^k f_i^k \ln 2}, \tag{30}$$

$$\sum_{k \in \mathcal{K}} \lambda_4^k = 1. \tag{31}$$

We can see that p_m^k and β_i^k can be obtained through solving the linear programs defined in (29) and (30), respectively. In (31), it gives a condition on the Lagrangian multipliers. For ϵ, it can be set as the minimum Q_k among all $k \in \mathcal{K}$.

b) Master subproblem: Once the results of the slave subproblem are obtained, the solution of the dual problem can be obtained by a subgradient method [30]. The update rule is given as

$$\lambda_1^{k,(t+1)} = \left[\lambda_1^{k,(t)} - \delta_1 (D_k - R_k) \right]^+, \forall k \in \mathcal{K}, \tag{32}$$

$$\lambda_2^{m,(t+1)} = \left[\lambda_2^{m,(t)} - \delta_2 \left(\sum_{k \in \mathcal{K}} p_m^k - P \right) \right]^+, \forall m \in \mathcal{M}, \tag{33}$$

$$\lambda_3^{(t+1)} = \left[\lambda_3^{(t)} - \delta_3 \left(\kappa IBK - \nu V - \sum_{k \in \mathcal{K}} D_k \right) \right]^+, \tag{34}$$

$$\lambda_4^{k,(t+1)} = w_k^{t+1} \left[\lambda_4^{k,(t)} - \delta_4 (\epsilon - Q_k) \right]^+, \forall k \in \mathcal{K}, \tag{35}$$

where t is the iteration indicator, $[x]^+ = \max\{0, x\}$, w_k^{t+1} is weighting factor to satisfy (31), and δ_1, δ_2, δ_3, and δ_4 are the step sizes to guarantee convergence.

Remark 4: Subproblem (23) has a non-empty feasible set only when

$$\sum_{k \in \mathcal{K}} R_k^* \geq \kappa IBK - \nu V, \tag{36}$$

where $\sum_{k \in \mathcal{K}} R_k^*$ is the maximum sum-rate under the power budget constraint in (12).

Proof: Please see Appendix B. ∎

Remark 4 provides the conditions when both computation and communication requirements are satisfied. This provides the guidelines on how to deploy communication and computation resources, which is important for practical system design.

Remark 5: The number of iterations to converge for the gradient decent algorithm is $\mathcal{O}(\log(1/\pi))$, where π is the error tolerance level [28]. The complexity in each iteration should be linear with the number of users, i.e., $\mathcal{O}(K)$.

V. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed scheme. We first present our simulation settings, and then discuss the evaluation results.

A. SIMULATION SETTING

The simulation setting is presented in Fig. 5, where a virtual object is placed in the center of the room, and we have $K = 5$ users who walk around the virtual object in a circular trajectory as suggested in [15] whose radius is 10 m.
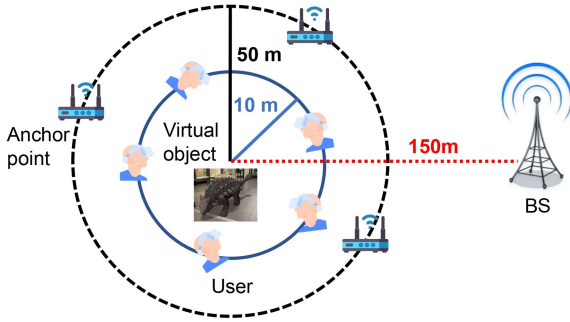
FIGURE 5. The AR scenario considered in our simulation study.

TABLE 1. Weight for tiles.

Column / Row	1	2	3	4	5	6
1	0.001	0.002	0.006	0.008	0.005	0.001
2	0.011	0.026	0.115	0.137	0.049	0.015
3	0.032	0.052	0.158	0.180	0.061	0.045
4	0.006	0.011	0.032	0.033	0.006	0.008

$A = 3$ single-antenna anchor points for localization are randomly located in a circle whose radius is 50 m. The distance between the virtual object and the BS is set to 150 m, and the height of the BS is 10 m. According to [31], we assume that the possible positions are uniformly distributed within the error range.

For video transmission, the videos are segmented into chunks of one second length, i.e., the duration for each time slot is one second [32], and we consider a total duration of $T = 10$ s. The video is encoded by FFMPEG with X.264 and its best quality is 4K,[3] corresponding to a bitrate of 35 Mbps. In other words, $B = 3.686 \times 10^5$ and $\kappa = 2.086$. The FOV is divided into 4 rows and 6 columns, and the tile weights are given in Table 1 [33]. The maximum QoE constant is $\tilde{Q} = 30$ and the scaling factor is $s = 10$.

We assume that the AR device is equipped with 4 antennas, and the BS is equipped with 16 antennas. The channel is Rayleigh faded and the channel gain is modeled as given in [34] where the decay factor is 3.76 and the power gain factor is -17.7 dB. The CPU capacity for edge server is $V = 6 \times 10^{10}$ cycles/second and the required CPU cycles for the compression task is $\mu = 2,000$ cycles/bit [35]. Other simulation settings are listed in Table 2 [36].

B. EVALUATION RESULTS

To evaluate the performance of the proposed scheme, we compare it with the following two benchmark schemes:

- *Fixed location:* the edge server only transmits the VR content according to the estimated location without considering the estimation error. The resolution adjustment, communication and computation resource management schemes are the same as that in the proposed scheme;

---

3. 4K video means 4096×2160 pixels.

TABLE 2. Parameters for our simulation study.

Parameters	Values
Power budget for each user P_{max}	20 dBm
Transmission bandwidth W	1 MHz
Carrier frequency	1.9 GHz
Noise figure	5 dB
SNR threshold η_k	0 dB
Error tolerance level π	0.1
Power budget for the BS P	23 dBm

- *Uniform resolution:* the edge server does not recognize the weight of each tile and uses the same resolution for all the possible tiles within the FOV. The resolution needs to be optimized according to the available computation and communication resources.

Fig. 6 (a) shows how QoE changes for different number of users K. In general, the system QoE, i.e., the minimum QoE among all the users, decreases as the number of users increases. The reasons are two-fold: 1) The more severe interference among more users will make the estimation less accurate. As a result, more data needs to be transmitted; and 2) The communication and computation resources for each user decrease with more users. On the other hand, compared with benchmark schemes, we can observe that the proposed scheme achieves the best performance, indicating the effectiveness of the proposed scheme. In comparison with the fixed location scheme, our scheme is similar to the idea of robust optimization [37], which allocates communication and computation resources to those FOVs with a lower probability to reduce the QoE degradation caused by the estimation error. Moreover, compared with the uniform resolution scheme, we find that proper resolution matching, i.e., transmitting tiles with a higher weight with a higher resolution[4] will result in a higher total QoE. Furthermore, we can also observe that the QoE obtained by the uniform resolution scheme is higher than that obtained by the fixed location scheme, indicating that the estimation error has a more significant impact on the QoE than the resolution matching.

In Fig. 6 (b), we plot the QoE vs. estimation error range with different video quality levels, i.e., 2K, 4K, and 8K. It is observed that the QoE decreases with a larger estimation error range, which is consistent with Proposition 1. We can also learn that the drop rate is increasing with a larger estimation error. This is because a larger range results in a larger potential FOV range, and thus more data is required to be transmitted. Moreover, we can observe that the QoE will decrease for higher-quality videos with limited communication resources. The reason under this observation is that more data is to be transmitted to maintain the same

---

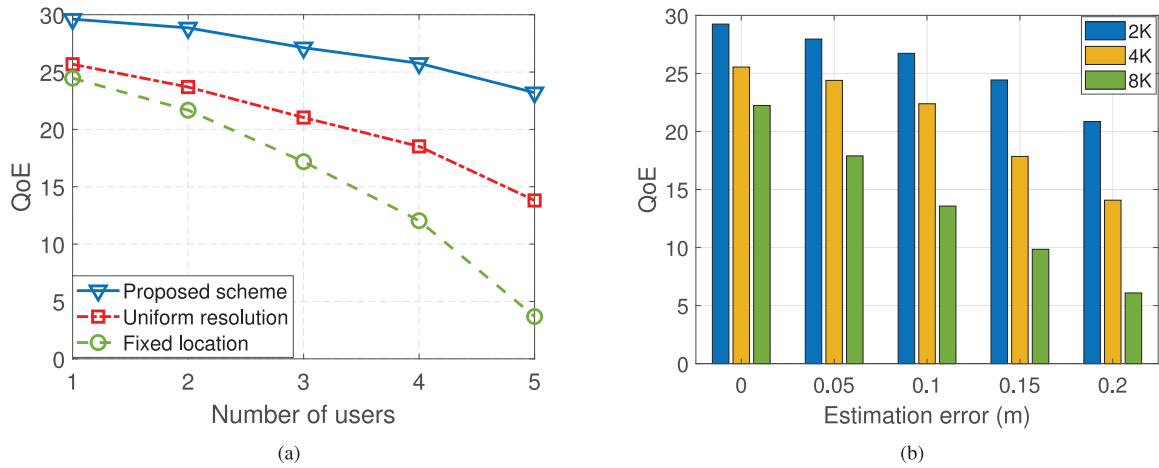4. The idea is similar to maximum ratio combining in wireless communications [38].

FIGURE 6. a) QoE vs. number of users *K*; b) QoE vs. estimation error with different video quality levels.
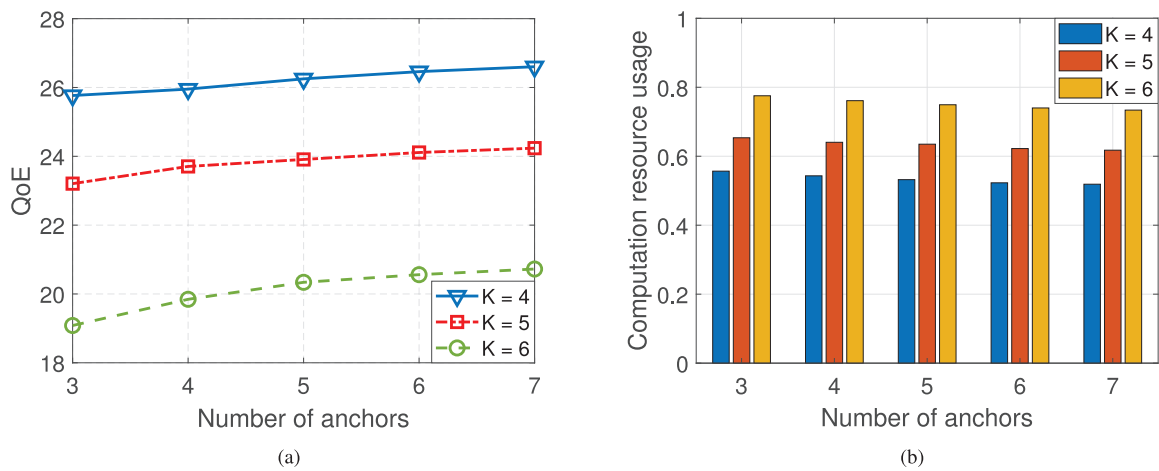

FIGURE 7. a) QoE vs. number of anchors *A*; b) Computation resource usage.

QoE level while the communication resource might not be sufficient to support more data transmission.

In Fig. 7, we show how the number of anchors influences the system performance. From Fig. 7(a), we can observe that the QoE slightly increases with the number of anchors. This is because the estimation error can be reduced with more anchor nodes, leading to fewer data to be transmitted and higher resolution for each tile under limited communication resources, which is also verified by the results shown in Fig. 7(b). In this subfigure, we can observe that less data is compressed with more anchors, indicating that 1) less tiles are involved in the possible FOVs (caused by the shrinking estimation error range); or 2) higher resolution video for each tile is transmitted. On the other hand, we find that the performance improvement brought by increasing the number of anchors *A* is much less than the impact of the number of users *K*, because the sensing accuracy improvement is limited when the number of anchors is larger than a fixed threshold. Moreover, from this figure, we can conclude that the system tends to maximize the usage of communication resources as the computation resources are not fully utilized. The reason under this observation is that the QoE

metric motivates the system to transmit video with as higher resolution as possible.

In Fig. 8, we show how the QoE changes in different time slots. Fig. 8(a) shows the QoE of a static user. From this figure, we can observe that the QoE increases as time progresses because the video resolution will be higher as time goes. This is because the data transmitted in the previous time slots will be stored in the device. For the proposed scheme and the uniform resolution scheme, the QoE will eventually reach its maximum value as all the data at the edge server will be eventually transmitted to the users. In contrast, the QoE obtained by the fixed location scheme will achieve a relatively lower saturated QoE because the scheme will not transmit the tiles in the shifted FoV caused by estimation error. Moreover, we can also see that the proposed scheme has a higher convergence rate, indicating that the adaptive resolution allocation can achieve a better QoE.

In Fig. 8(b), we present the case where the Metaverse user moves randomly. As the speed of the user is typically slow, the FOVs of two successive time slots overlap with each other. We assume that 30% FOVs are the same for two successive time slots in this case. From this figure, we can
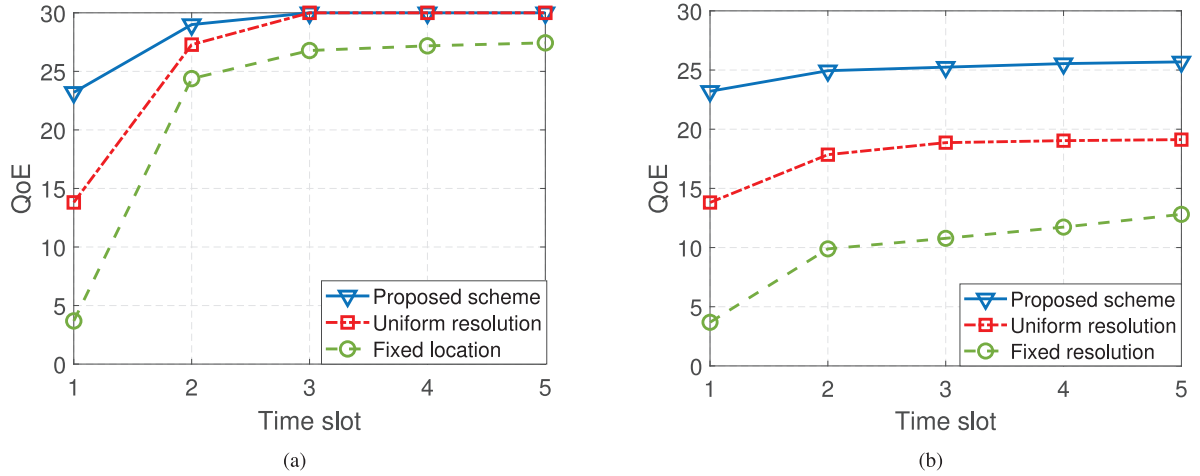
FIGURE 8. a) QoE in different time slots of a static user; b) QoE in different time slots of a mobile user.
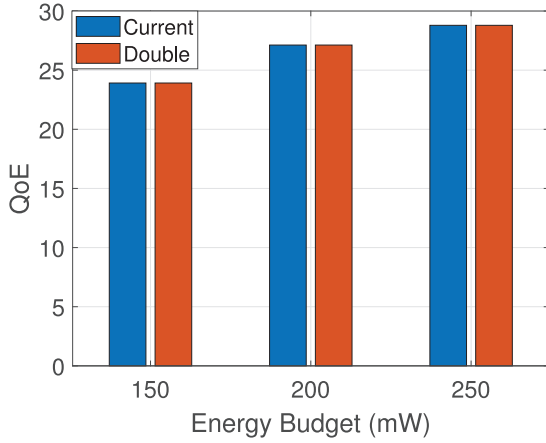

FIGURE 9. Sensitivity of QoE with respect to communication and computation resources.

see that the QoE also increases as time goes because the data that has been cached in the headset can help save communication resources in the following time slots. Moreover, we can also observe that the proposed scheme outperforms the benchmark schemes in the mobile case.

In Fig. 9, we show the sensitivity of QoE with respect to communication and computation resources when the number of Metaverse users is $K = 3$. From this figure, we can observe that the QoE increases with higher energy budget for data transmission when the computation resource is fixed. However, when we double the CPU cycles, the QoE almost remains the same. This is because the QoE is positively related to the communication resource but independent to the computation resource. If more communication resource is deployed in the system, the BS will tend to transmit higher-resolution videos since the QoE will be improved, while increasing the computation resources will not change the QoE as the computation requirement has been satisfied.

VI. CONCLUSION AND FUTURE WORK

In this paper, we studied providing a location-dependent AR services in wireless MEC-enabled Metaverse systems. To maximize the minimum QoE among users, we first optimized the waveform in the uplink phase to achieve a low localization error, and then, in the downlink phase, we proposed a joint resolution and transmit power optimization algorithm for a high QoE. Simulation results verified the effectiveness of the proposed scheme, with the following findings: 1) Considering the localization estimation error can effectively improve the system QoE in Metaverse; 2) The QoE metric encourages the system to fully utilize the communication resources to transmit videos with a higher resolution; and 3) The impact of the number of anchor nodes on the system QoE is much weaker than that of the number of Metaverse users.

In the current system, the users use orthogonal channels and interactions among users are not considered. In the future work, we will consider the interactions, e.g., cooperation or interference. To be specific, if multiple users cooperate to perform the same task provided by AR services, the QoE per user should not only consider the individual video display on each user, but also the synchronization among these users. Therefore, a protocol should be designed to facilitate the synchronization and the QoE model should be modified. Moreover, since the videos transmitted to these users are highly correlated, the BS can use some network coding techniques [39] to save the bandwidth where the users can share the same content using direct communications. Under such a setting, the interference among users will be inevitable, and proper wireless management will be important to alleviate the interference.

APPENDIX A
PROOF OF PROPOSITION 1

Let $\alpha > 1$ be the increasing scale for the extended range in which the tiles can possibly fall within the FOV. Therefore, the distribution among these tiles also follows the scaled distribution.

However, as the data that can be transmitted is limited, the average resolutions for each tile will decrease, i.e., $\hat{\beta}_k \leq \tilde{\beta}_k$,

where $\hat{\beta}_k$ and $\tilde{\beta}_k$ are the average resolutions for the extended and the original ranges for user k, respectively.

Therefore, we have

$$\sum_{i \in \hat{\mathcal{I}}_k} \hat{Q}_i^k = \hat{\beta}_k \sum_{i \in \hat{\mathcal{I}}_k} \hat{\omega}_i^k \hat{f}_i^k$$
$$\leq \tilde{\beta}_k \sum_{i \in \hat{\mathcal{I}}_k} \hat{\omega}_i^k \hat{f}_i^k \overset{(a)}{=} \tilde{\beta}_k \alpha \sum_{i \in \tilde{\mathcal{I}}_k} \tilde{\omega}_i^k \tilde{f}_i^k$$
$$= \alpha \sum_{i \in \tilde{\mathcal{I}}_k} \tilde{Q}_i^k, \tag{37}$$

where $\hat{\omega}_i^k$ and $\hat{f}_i^k$ are the weight and the probability for the i-th tile in the extended range, respectively, while $\tilde{\omega}_i^k$ and $\tilde{f}_i^k$ are the weight and the probability for the i-th tile in the original range, respectively. The equality (a) in equation (37) is achieved as the distribution is scaled.

Similarly, we can show that

$$\sum_{i \in \hat{\mathcal{I}}_k} \hat{\omega}_i^k \hat{B}_i^k \hat{f}_i^k = \alpha \sum_{i \in \tilde{\mathcal{I}}_k} \tilde{\omega}_i^k \tilde{B}_i^k \tilde{f}_i^k, \tag{38}$$

where $\hat{B}_i^k$ and $\tilde{B}_i^k$ are the maximum resolution and extended one for tile i of user k, respectively.

Therefore, the QoE satisfies the following inequality

$$\hat{Q}^k = \log_2 \left(\frac{\sum_{i \in \hat{\mathcal{I}}_k} \hat{Q}_i^k}{\sum_{i \in \hat{\mathcal{I}}_k} \hat{\omega}_i^k \hat{B}_i^k \hat{f}_i^k} \right)$$
$$\leq \log_2 \left(\frac{\sum_{i \in \tilde{\mathcal{I}}_k} \alpha \tilde{Q}_i^k}{\sum_{i \in \hat{\mathcal{I}}_k} \alpha \tilde{\omega}_i^k \tilde{B}_i^k \tilde{f}_i^k} \right) = \tilde{Q}^k. \tag{39}$$

This shows that the QoE decreases as the FOV range is extended.

APPENDIX B
PROOF OF REMARK 4
According to (13)-(15), we have

$$\sum_{k \in \mathcal{K}} D_k \geq \kappa IBK - \nu V. \tag{40}$$

On the other hand, we can derive from (11) that

$$\sum_{k \in \mathcal{K}} R_k \geq \sum_{k \in \mathcal{K}} D_k. \tag{41}$$

This indicates that we can find a feasible solution $\{D_k\}$ as long as the maximum value of $\sum_{k \in \mathcal{K}} R_k$ is guaranteed to be larger than $\kappa IBK - \nu V$. This completes the proof.

REFERENCES

[1] M. Xu et al., "A full dive into realizing the edge-enabled metaverse: Visions, enabling technologies, and challenges," 2022, *arXiv: 2203.05471*.

[2] L.-H. Lee et al., "All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda," 2021, *arXiv: 2110.05352*.

[3] W. Y. B. Lim et al., "Realizing the metaverse with edge intelligence: A match made in heaven," *IEEE Wireless Commun.*, early access, Jul. 4, 2022, doi: 10.1109/MWC.018.2100716.

[4] L. U. Khan, Z. Han, D. Niyato, E. Hossain, and C. S. Hong, "Metaverse for wireless systems: Vision, enablers, architecture, and future directions," 2022, *arXiv: 2207.00413*.

[5] Z. Wei, B. Li, R. Zhang, X. Cheng, and L. Yang, "OCVC: An overlapping-enabled cooperative vehicular fog computing protocol," *IEEE Trans. Mobile Comput.*, early access, Oct. 4, 2022, doi: 10.1109/TMC.2022.3211882.

[6] S. Dhelim, T. Kechadi, L. Chen, N. Aung, H. Ning, and L. Atzori, "Edge-enabled metaverse: The convergence of metaverse and mobile edge computing," 2022, *arXiv: 2205.02764*.

[7] "Study on XR (extended reality) evaluations for NR, v.17.0.0," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 38.838, Jan. 2022.

[8] P. Wang, Z. Zheng, B. Di, and L. Song, "HetMEC: Latency-optimal task assignment and resource allocation for heterogeneous multi-layer mobile edge computing," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4942–4956, Oct. 2019.

[9] H. Du, D. Niyato, J. Kang, D. I. Kim, and C. Miao, "Optimal targeted advertising strategy for secure wireless edge metaverse," 2021, *arXiv: 2111.00511*.

[10] M. Xu, D. Niyato, J. Kang, Z. Xiong, C. Miao, and D. I. Kim, "Wireless edge-empowered metaverse: A learning-based incentive mechanism for virtual reality," in *Proc. IEEE ICC*, Seoul, South Korea, May 2022, pp. 5220–5225.

[11] J. Ren, Y. He, G. Huang, G. Yu, Y. Cai, and Z. Zhang, "An edge-computing based architecture for mobile augmented reality," *IEEE Netw.*, vol. 33, no. 4, pp. 162–169, Jul./Aug. 2019.

[12] A. Al-Shuwaili and O. Simeone, "Energy-efficient resource allocation for mobile edge computing-based augmented reality applications," *IEEE Wireless Commun. Lett.*, vol. 6, no. 3, pp. 398–401, Jun. 2017.

[13] Q. Liu, S. Huang, J. Opadere, and T. Han, "An edge network orchestrator for mobile augmented reality," in *Proc. IEEE INFOCOM*, Honolulu, HI, USA, Apr. 2018, pp. 756–764.

[14] M. Erol-Kantarci and S. Sukhmani, "Caching and computing at the edge for mobile augmented reality and virtual reality (AR/VR) in 5G," *Ad Hoc Netw.*, vol. 223, pp. 169–177, Jan. 2018.

[15] X. Ran, C. Slocum, Y.-Z. Tsai, K. Apicharttrisorn, M. Gorlatova, and J. Chen, "Multi-user augmented reality with communication efficient and spatially consistent virtual objects," in *Proc. CoNEXT*, Barcelona, Spain, Dec. 2020, pp. 386–398.

[16] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[17] H. Zhang, B. Di, K. Bian, Z. Han, H. V. Poor, and L. Song, "Toward ubiquitous sensing and localization with reconfigurable intelligent surfaces," *Proc. IEEE*, vol. 110, no. 9, pp. 1401–1422, Sep. 2022.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing–Estimation Theory*. Englandwood Cliffs, NJ, USA: Prentice-Hall, 1993.

[19] S. Jeong, O. Simeone, A. Haimovich, and J. Kang, "Beamforming design for joint localization and data transmission in distributed antenna system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 62–76, Jan. 2015.

[20] "Sensors for AR/VR—Augmented and virtual reality: The next big thing in marketing?" Accessed: May 31, 2020. [Online]. Available: https://pressbooks.pub/augmentedrealitymarketing/chapter/sensors-for-arvr/

[21] J. Li, C. Zhang, Z. Liu, W. Sun, and Q. Li, "Joint communication and computational resource allocation for QoE-driven point cloud video streaming," in *Proc. IEEE ICC*, Dublin, Ireland, Jun. 2020, pp. 1–6.

[22] W. Yang, X. Chi, L. Zhao, and Z. Xiong, "QoE-based MEC-assisted predictive adaptive video streaming for on-road driving scenarios," *IEEE Wireless Commun. Lett.*, vol. 10, no. 11, pp. 2552–2556, Nov. 2021.

[23] D. Chen et al., "Matching-theory-based low-latency scheme for multitask federated learning in MEC networks," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11415–11426, Jul. 2021.

[24] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, Dec. 2014.

[25] W. Huang et al., "Utility-oriented resource allocation for 360-degree video transmission over heterogeneous networks," *Digit. Signal Process.*, vol. 84, pp. 1–14, Jan. 2019.

[26] J. M. Danskin, "The theory of max–min, with applications," *SIAM J. Appl. Math.*, vol. 14, no. 4, pp. 641–664, Jul. 1966.

[27] W. W.-L. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, "Robust power allocation for energy-efficient location-aware networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 6, pp. 1918–1930, Dec. 2013.

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

[29] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max–min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, Mar. 2008.

[30] S. Boyd, "Subgradient methods." May 2014. [Online]. Available: https://web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf

[31] H. A. B. F. Oliveira, E. F. Nakamura, A. A. F. Loureiro, and A. Boukerche, "Error analysis of localization systems for sensor networks," in *Proc. ACM GIS*, Bremen, Germany, Nov. 2005, pp. 71–78.

[32] Y. Zhang, P. Zhao, K. Bian, Y. Liu, L. Song, and X. Li, "DRL360: 360-degree video streaming with deep reinforcement learning," in *Proc. IEEE INFOCOM*, Paris, France, Apr. 2019, pp. 1252–1260.

[33] J. Chakareski, R. Aksu, X. Corbillon, G. Simon, and V. Swaminathan, "Viewport-driven rate-distortion optimized 360° video streaming," in *Proc. IEEE ICC*, Kansas City, MO, USA, May 2018, pp. 1–7.

[34] "Study on provision of low-cost machine-type communications (MTC) user equipments (UEs) based on LTE (release 12)," 3GPP, Sophia Antipolis, France, document 3GPP TR 36.888, Jun. 2013.

[35] D. Xu, Q. Li, and H. Zhu, "Energy-saving computation offloading by joint data compression and resource allocation for mobile-edge computing," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 704–707, Apr. 2019.

[36] H. Zhang, Y. Liao, and L. Song, "D2D-U: Device-to-device communications in unlicensed bands for 5G system," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3507–3519, Jun. 2017.

[37] Y. Cao et al., "Optimal energy management for multi-microgrid under a transactive energy framework with distributionally robust optimization," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 599–612, Jan. 2022.

[38] K. S. Ahn and R. W. Heath, "Performance analysis of maximum ratio combining with imperfect channel estimation in the presence of cochannel interferences," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1080–1085, Mar. 2009.

[39] R. Bassoli, H. Marques, J. Rodriguez, K. W. Shum, and R. Tafazolli, "Network coding theory: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1950–1978, 4th Quart., 2013.

SHIWEN MAO (Fellow, IEEE) is a Professor, an Earle C. Williams Eminent Scholar, and the Director of the Wireless Engineering Research and Education Center, Auburn University, Auburn, AL, USA. His research interest includes wireless networks, multimedia communications, and smart grid. He received the IEEE ComSoc TC-CSR Distinguished Technical Achievement Award in 2019 and the NSF CAREER Award in 2010. He is a co-recipient of the 2021 Best Paper Award of Elsevier/KeAi *Digital Communications and Networks* Journal, the 2021 IEEE INTERNET OF THINGS JOURNAL Best Paper Award, the 2021 IEEE Communications Society Outstanding Paper Award, the IEEE Vehicular Technology Society 2020 Jack Neubauer Memorial Award, the 2004 IEEE Communications Society Leonard G. Abraham Prize in the field of communications systems, and several conference best paper/demo awards. He is the Editor-in-Chief of IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING and is on the Editorial Board of several other IEEE and ACM journals. He is a Distinguished Lecturer of IEEE Communications Society and the IEEE Council of RFID.

HONGLIANG ZHANG (Member, IEEE) received the B.S. and Ph.D. degrees from the School of Electrical Engineering and Computer Science, Peking University, in 2014 and 2019, respectively, where he is currently an Assistant Professor with the School of Electronics. Prior to that, he was a Postdoctoral Associate with Princeton University and the University of Houston. His current research interests include reconfigurable intelligent surfaces, aerial access networks, optimization theory, and game theory. He received the Best Doctoral Thesis Award from Chinese Institute of Electronics in 2019. He is also the recipient of the 2021 IEEE Comsoc Heinrich Hertz Award for Best Communications Letters and the 2021 IEEE ComSoc Asia–Pacific Outstanding Paper Award. He has served as a TPC member and a workshop co-chair for many IEEE conferences. He is the winner of the Outstanding Leadership Award as the Publicity Chair for IEEE EUC in 2022. He is currently an Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE COMMUNICATIONS LETTERS, *IET Communications*, and *Frontiers in Signal Processing*. He has also served as a Guest Editor for several journals, such as IEEE INTERNET OF THINGS JOURNAL and *Journal of Communications and Networks*. He was an Exemplary Reviewer for IEEE TRANSACTIONS ON COMMUNICATIONS in 2020.

DUSIT NIYATO (Fellow, IEEE) received the B.E. degree from King Mongkuk's Institute of Technology Ladkrabang, Thailand, in 1999, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada, in 2008. He is currently the President's Chair Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests are in the areas of distributed collaborative machine learning, Internet of Things, edge intelligent metaverse, mobile and distributed computing, and wireless networks. He won the Best Young Researcher Award of IEEE Communications Society (ComSoc) Asia–Pacific, the 2011 IEEE Communications Society Fred W. Ellersick Prize Paper Award, the IEEE Computer Society Middle Career Researcher Award for Excellence in Scalable Computing in 2021, and the Distinguished Technical Achievement Recognition Award of IEEE ComSoc Technical Committee on Green Communications and Computing in 2022. He also won a number of best paper awards, including IEEE Wireless Communications and Networking Conference, IEEE International Conference on Communications, IEEE ComSoc Communication Systems Integration and Modelling Technical Committee, and IEEE ComSoc Signal Processing and Computing for Communications Technical Committee 2021. He was named the 2017–2021 Highly Cited Researcher in computer science. Currently, he is serving as the Editor-in-Chief for IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, an Area Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Associate Editor for IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE WIRELESS COMMUNICATIONS, IEEE NETWORK, and *ACM Computing Surveys*. He was a Guest Editor of IEEE JOURNAL ON SELECTED AREAS ON COMMUNICATIONS. He was a Distinguished Lecturer of the IEEE Communications Society for 2016–2017.

ZHU HAN (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. Currently, he is a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department as well as with the Computer Science Department, University of Houston, Houston, TX, USA. His main research targets on the novel game-theory-related concepts critical to enabling efficient and distributive use of wireless networks with limited resources. His other research interests include wireless resource allocation and management, wireless communications and networking, quantum computing, data science, smart grid, security, and privacy. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, the IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He is also the Winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks." He was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018, has been an AAAS Fellow since 2019, and has been an ACM Distinguished Member since 2019. He is a 1% highly cited researcher since 2017 according to Web of Science.