Sidelink-Aided Multiquality Tiled 360° Virtual Reality Video Multicast

Jianmei Dai, *Member, IEEE*, Guosen Yue, *Senior Member, IEEE*, Shiwen Mao, *Fellow, IEEE*, and Danpu Liu, *Senior Member, IEEE*

***Abstract*—Mobile/wireless virtual reality (VR) services, especially immersive 360° VR videos, have advanced unprecedentedly in recent years. However, the high bandwidth requirement of VR services has compounded the burden on wireless networks. Multicast is a high potential technique for alleviating the bandwidth requirement of 360° VR video streaming, but the multicast capacity is still constrained by the users with poor channel conditions, and it vanishes when the number of users increases while the number of the base station (BS) antennas is fixed. To overcome the drawbacks of multicast, sidelink, which is an adaptation of the core LTE standard that allows the device-to-device (D2D) communications without going through a BS, can be utilized. In this article, two sidelink-aided multicast scenarios (i.e., independent decoding and joint decoding) are studied for multiquality tiled 360° VR video transmission. We propose a utility model for each scenario, and quality level selection, sidelink sender/receiver selection, and transmission resource allocation are optimized to maximize the total utility of all users under the bandwidth constraints as well as the quality smoothness constraints for multiquality tiles. We then develop an iterative two-stage algorithm to obtain suboptimal solutions to the formulated mixed-integer nonlinear programming (MINLP) problems. Simulation results demonstrate the advantage of the proposed solutions over several baseline schemes.**

***Index Terms*—Multicast, resource allocation, sidelink, tiled 360° video, virtual reality (VR).**

I. Introduction

THE LAST few years have witnessed an unprecedented proliferation of interest in both academia and industry in mobile/wireless virtual reality (VR) services. It is forecasted that the VR ecosystem will become an $80 billion market and a trillion-dollar industry by the year 2025 and 2035, respectively, [1], [2]. The immersive panoramic (i.e., 360°) VR video is an important form of media created by the convergence of the advanced VR technology and fast video processing hardware, which can provide a 360° view angle to enable an omni-directional viewport of the scene [3], [4]. However, delivering an entire 360° VR video incurs four to five times higher bandwidth requirements than that of a traditional video, since the general 360° VR video data rate at the 4K resolution is 50 Mb/s and the data rate with 8K resolution increases to 200 Mb/s [5].

To alleviate the burden of wireless networks, an effective method is to transmit user's Field-of-View (FoV) data instead of the entire 360° VR video. The FoV is defined as the portion of the 360° video that is in the user's line of sight, and an FoV can be spatially divided into smaller parts referred to as "tiles," which can be encoded into multiple versions, each at a different quality level [6]–[9]. Bandwidth consumption can be reduced by only sending the tiles in the FoV in high resolution, while sending other tiles in lower resolutions or not at all [10].

Multicast is another effective technique for decreasing the bandwidth requirement of 360° VR video streaming. In such an application, users wear an individual head-mounted device (HMD) to watch a 360° video, and the number of concurrent transmissions for the same content will be much larger than that of conventional videos, even in a small geographical area [11]. By sharing spectrum with many users requesting the same video segments or tiles, multicast is highly promising to mitigate the high demand on network resources [12]–[17].

However, the multicast capacity is limited by the users in poor channel conditions and vanishes when the number of users increases due to the fixed number of base station (BS) antennas [18], [19]. To overcome this issue, device to device (D2D) communications can be used as an auxiliary means to improve the multicast performance. The D2D communication technology in long-term evolution (LTE) was introduced by the 3rd Generation Partnership Program (3GPP) to allow devices in proximity to communicate directly with each other. Therefore, D2D communications can effectively alleviate the burden of the BS and reduce the transmission delay. In order to support LTE D2D, 3GPP defined the PC5 interface, a new direct link between user equipments (UEs) termed "sidelink" in the access stratum layer. UEs can utilize the sidelink to exchange information when they are in close proximity. As a

Manuscript received April 11, 2021; revised July 4, 2021; accepted August 9, 2021. Date of publication August 16, 2021; date of current version March 7, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61971069 and Grant 61801051, and in part by the Beijing Natural Science Foundation under Grant L202003. The work of Shiwen Mao was supported in part by NSF under Grant ECCS-1923717. *(Corresponding author: Jianmei Dai.)*

Jianmei Dai is with the Space Engineering University and Beijing Laboratory of Advanced Information Network, Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: jammy_wirelesscomm@126.com).

Guosen Yue is with Wireless Research and Standards, Wireless Access Lab, Futurewei Technologies, Bridgewater, NJ 08807 USA (e-mail: guosenyue@gmail.com).

Shiwen Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

Danpu Liu is with the Beijing Laboratory of Advanced Information Network, Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: dpliu@bupt.edu.cn).

Digital Object Identifier 10.1109/JIOT.2021.3105100

2327-4662 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

new paradigm underlying cellular networks, sidelink communications have been standardized to enhance the performance of cellular networks and support vehicle-to-everything (V2X) communications, which enable a vehicle to collaborate with other vehicles, devices, and infrastructure.

Several enhanced multicast schemes for traditional services aided by sidelink have been proposed in [20]–[26]. The idea is to let the BS focus its transmission toward a suitably selected group of users with good channel quality, while leaving a few others in outage [27], [28]. The sidelinks are enabled to allow the users with unfavorable channel conditions to be served by their peers. The overall transmission capacity can be thus improved. However, these schemes are designed for traditional services, and may not be suitable for 360° VR videos since many unique features and challenges of VR video streaming are not considered. For example, unlike common videos, different 360° VR frame tiles are of different levels of importance to the viewer experience, depending on their spatial positions in the user's FoV. Apparently, the tiles which are in the center of the FoV are more important and should be delivered with a higher resolution, while the others may have a reduced resolution to save bandwidth and power. Furthermore, the tile smoothness should be considered in the design of VR transmission systems, in other words, the difference of quality between adjacent VR video tiles should be bounded to ensure a good viewer experience. Generally, the smaller the boundary, the higher the user's Quality of Experience (QoE).

In this work, we study the problem of sidelink-aided multicast optimization for 360° VR videos. Inspired by several prior works [12], [13], [16], we aim to design a sidelink-aided multicast solution for multiquality tiled 360° VR videos. Unlike these prior works, our solution is based on the joint consideration of multicast downlinks and sidelinks. In this system, not all the users are served by the BS multicast link; some users with poor channel conditions are selected to be served by sidelinks. Under such a circumstance, two resulting different decoding scenarios, i.e., an independent decoding scenario and a joint decoding scenario, are modeled and analyzed. For each scenario, we optimize the tile quality level selection, transmission resource allocation, and sidelink sender/receiver selection to maximize the total utility of all users under the transmission resource constraints as well as the quality smoothness constraints for mixed-quality tiles. Both problems are challenging mixed-integer nonlinear programming (MINLP) problems. We propose a novel iterative two-stage algorithm to obtain suboptimal, yet highly competitive solutions, using greedy search and continuous relaxation. To the best of our knowledge, such an effective design has not yet been analytically verified in the existing literature.

In summary, the novelty and technical contributions of this work are provided as follows.

1) We propose a sidelink-aided multicast system for multiquality tiled 360° VR video streaming from one BS to multiple users, where both BS-user multicast link and sidelink are jointly utilized.
2) We describe two sidelink-aided multicast scenarios, the independent decoding scenario and the joint decoding scenario, and formulate a utility maximization problem for each scenario under the transmission resource and tile quality smoothness constraints via optimized quality level selection, transmission resource allocation, and sidelink sender/receiver selection.
3) To tackle the NP-hardness of the formulated MINLP problems, we propose an iterative two-stage algorithm to obtain a suboptimal solution, using greedy search and continuous relaxation, which exhibits super fast convergence performance in our evaluations.
4) Our simulation results demonstrate that the proposed suboptimal solutions can provide significantly better utility performance than several baseline schemes, which effectively enhance the user QoE compared to conventional multicasting strategies and the baseline schemes.

The remainder of this article is organized as follows. Section II discusses the related work. Our system model and problem statement are presented in Section III and Section IV, respectively. In Section V and Section VI, the proposed two-stage algorithms for the two scenarios and their simulation performance validation are presented, respectively. Finally, Section VII concludes this article.

II. Related Work

Multicast is an up-and-coming solution to deal with the high network requirements of VR. Two optimal multicast transmission schemes for tiled 360° VR video were studied in [12] and [13], respectively. One was to maximize the received video quality in orthogonal frequency division multiple access (OFDMA) systems by optimizing subcarrier, transmission power, and transmission rate allocation [12]. The other was to minimize the average transmission energy by optimizing transmission time and power allocation [13]. Ahmadi *et al.* [14] proposed a multicast DASH-based tiled streaming solution, including a user's viewports-based tile weighting approach and a rate adaptation algorithm, to provide an immersive experience for VR users. View synthesis multicast was analyzed in [15]. Further, Long *et al.* [16] optimized the VR video quality level selection, transmission time allocation, and transmission power allocation to maximize the utility under transmission time and power allocation constraints as well as quality smoothness constraints for mixed-quality tiles. A multisession multicasting (MSM) system for multiquality tiled-video was proposed in [17], where user grouping, wireless-resource allocation, and tiled-video rate selection were jointly optimized.

Introducing D2D communications to the traditional multicast schemes can not only relieve the burden of the BS but greatly enhance resource utilization. Aiming to maximize the system throughput, a two-stage multicast scheme was investigated in [20], and the system throughput could be improved with the help of mobile relays. The idea of employing network controlled reliable multicasting for D2D communications was explored in [21], which was a scalable and efficient solution for file transfer and video streaming. The D2D assisted cooperative multicast was investigated in [22] and the optimal time allocation between the multicast and retransmission stages was derived. A mixed infrastructure-to-device (I2D) multicast and D2D-relaying offloading scheme

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

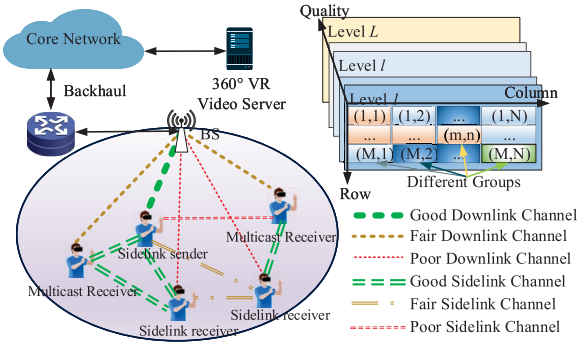


Fig. 1. Illustration of the sidelink-aided multicast system for multiquality tiled 360° VR video. Each 360° VR video segment is partitioned into a grid of $M \times N$ nonoverlapping rectangular tiles of the same size, and each tile has multiple quality representations. A 360° VR video server providing video streams connects to a BS through high-speed wired links. The BS equipped with one transmit antenna aims to announce streaming services and deliver VR video tiles of different qualities to multiple users within its range, wearing single-antenna, HMDs. Different tiles at various quality levels can be transmitted to users by the BS via the multicast downlink or by a selected user (sidelink sender) via sidelink.

based on reinforcement learning was proposed in [23], aiming to reduce the redundant traffic while guaranteeing timely delivery. Multihop D2D communication was studied in [24], with the objective of improving the multimedia multicast sessions and transmissions in terms of throughput and mean download time per client. A simple computationally efficient scheme was proposed for the D2D aided multicast scenario, in which only statistical channel knowledge at the transmitter was required [25]. An enhanced scheme for selecting a subset of UEs who cooperate to spread the common message across the rest of the network via D2D retransmissions was proposed in the new light of precoding capabilities at the BS [26].

However, these existing schemes do not consider the mixed-quality and the smoothness of VR tiles, and thus may not achieve a good performance for VR streaming.

III. System Model

A. Network Model and Problem Statement

A sidelink-assisted live streaming of popular 360° videos over 4G/5G cellular networks that support multicast services, such as evolved multimedia broadcast and multicast services (eMBMS) in LTE networks [29], is considered in this article. The system architecture is illustrated in Fig. 1 and the notation used in this article is summarized in Table I.

In our system model, the 360° VR video segment is divided into a grid of $M \times N$ smaller nonoverlapping rectangular tiles of the same size, where M and N represent the numbers of tiles in each column and each row, respectively. The (m, n)th ($1 \le m \le M$ and $1 \le n \le N$) tile refers to the tile in the mth row and nth column. Similar to regular mobile multimedia streaming and the model in [30], a 360° VR video server providing video streams is connected to a BS through high-speed wired links. The BS, equipped with one transmit antenna, aims at announcing streaming services and delivering VR video tiles of different qualities to a set $\mathcal{U} \triangleq \{1, 2, \ldots, u, \ldots, |\mathcal{U}|\}$ of users wearing HMDs, that are within its coverage.

TABLE I
Notation

Symbol	*Description*
b	The maximum encoding rate of all the tile quality levels
$\mathcal{G}, \|\mathcal{G}\|, g$	User group set, the maximum group numbers, g-th group
$\mathcal{L}, L, l$	Quality level set, the best quality level, and the l-th quality level
$M, N, (m, n)$	Tile numbers in each row and column, the (m, n)-th tile, $1 \le m \le M, 1 \le n \le N$
$\mathcal{U}, \|\mathcal{U}\|, u$	Total user set, total user numbers, user index
$\mathcal{U}_g, \|\mathcal{U}_g\|$	User set of group g,user numbers of $\mathcal{U}_g$
$\mathcal{J}_g, \|\mathcal{J}_g\|, j_g$	Sidelink receiver set in $\mathcal{U}_g$, sidelink receiver numbers of $\mathcal{J}_g$, sidelink receiver index
$\mathcal{K}_g, \|\mathcal{K}_g\|, k_g$	Multicast user set in $\mathcal{U}_g$, multicast user numbers of $\mathcal{K}_g$, user index in $\mathcal{K}_g$
$B^{\mathrm{DL}}, B_g^{\mathrm{DL}}$	Overall bandwidth of the downlink of BS, the downlink bandwidth allocated for group g
$B^{\mathrm{SL}}, B_g^{\mathrm{SL}}$	Overall bandwidth of the sidelink, the allocated sidelink bandwidth for group g
$C_g^{\mathrm{DL}}, C_g^{\mathrm{SL}}$	Multicast and sidelink capacity of group g
$P_{\mathrm{BS}}, P_{\mathrm{U}}$	Power of BS and user
$W_{m,n}$	Overall weight of the (m, n)-th tile
$W_{m,n}^{\mathrm{global}}, W_{m,n}^{\mathrm{FoV}}$	Long term and FoV weight of the (m, n)-th tile
$x_{m,n}, y_{m,n}$	Tile's quality level selected by BS and transcoded by sidelink sender
Δ	Quality difference between two adjacent tiles
Φ, Φ_g	Tiles set transmitted to all the user groups, and the tiles set transmitted to user group g

The users who request the same tiles forms one multicast group g, and there are at most $|\mathcal{G}|$ groups for all the users, where $|\mathcal{G}| \le \min(|\mathcal{U}|, M \times N)$. We denote $\mathcal{G} = \{1, 2, \ldots, g, \ldots, |\mathcal{G}|\}$, $\mathcal{U}_g$, and $|\mathcal{U}_g|$ as the set of user groups, the user set in group g, and the number of users in group g, respectively. The special multicast groups where $|\mathcal{U}_g| = 1$ (i.e., the tiles are requested only by one user) are beyond the scope of this article, since the sidelink cannot be utilized in this case.

Tiles in each group are independently encoded, meaning that the user can either decode them or fail to decode them via the multicast downlink channel. We denote subset $\mathcal{K}_g$ (with $|\mathcal{K}_g|$ users) as the users who can receive and decode VR video tiles successfully via the multicast downlink, while the other subset $\mathcal{J}_g$ with $|\mathcal{J}_g|$ users can only decode the same video tiles correctly with the help of a sidelink sender. The sidelink sender is a selected user in $\mathcal{K}_g$ who has successfully decoded the video tiles and can help users in $\mathcal{J}_g$. We have $\mathcal{K}_g \cup \mathcal{J}_g = \mathcal{U}_g$, $\mathcal{K}_g \cap \mathcal{J}_g = \emptyset$, $|\mathcal{J}_g| = |\mathcal{U}_g| - |\mathcal{K}_g|$, and $\mathcal{U}_g \subseteq \mathcal{U}$. In this context, based on the average channel qualities, e.g., the wideband channel quality indicator (CQI) or the average signal-to-noise ratio (SNR), or the average rate of the sidelinks, which are reported by the users to the BS, the BS can allocate its radio resources to a selected subset of users of each group with optimized $\mathcal{K}_g$, and one of which in turn can cooperate to spread the video data across the rest of the network.

The following two-phase scheme [25], [26], [31], [32] has been adopted for data transmissions, taking two time slots.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

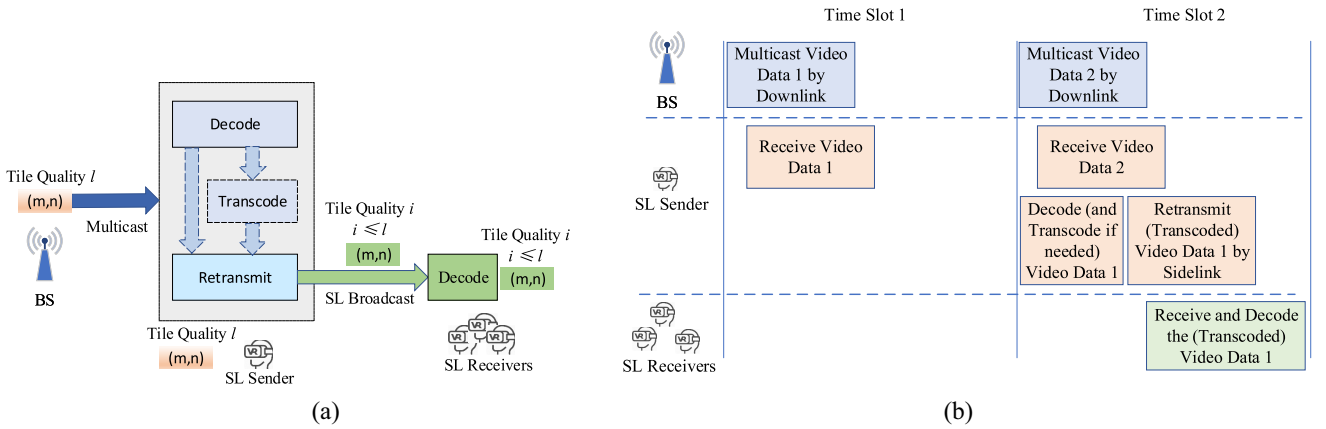


Fig. 2. Illustration of Scenario InD: The sidelink sender that decodes the tile transcodes it (if needed) and broadcasts it at the same quality or a degraded quality to the sidelink receivers. (a) Data process. (b) Time schedule for the BS and users.

Phase 1: The BS conveys the video tiles using orthogonal multicast transmissions to users.

Phase 2: A sidelink sender is selected for each group for sidelink assisted transmissions. Since the frequencies of multicast downlink and sidelink are different, the sidelink sender can receive new frames/packets from the BS while transmitting decoded frames/packets on the sidelink in parallel. Assuming channel status does not change drastically in a short period of time (i.e., coherence time), then mathematically, we can assume that the two transmissions are simultaneous without violating the causal status.

Specifically, we consider that the sidelink receivers can decode the tiles independently or jointly in Phase 2, which are described in the following.

1) Independent Decoding Scenario (Hereinafter Referred to as Scenario InD): Scenario InD is illustrated in Fig. 2. The sidelink sender in $\mathcal{K}_g$ decodes the tile, transcodes it (if needed), and broadcasts the tile to all the other users in $\mathcal{J}_g$, who receive the tile with the same quality or with a degraded quality. We use $x_{m,n}$ and $y_{m,n}$ to denote the original tile quality level selected by the BS and the transcoded tile quality level transmitted by the sidelink sender, respectively. We then have

$$x_{m,n} \geq y_{m,n}, \ \ x_{m,n} \in \mathcal{L}, \ \ y_{m,n} \in \mathcal{L}, \ \ (m,n) \in \Phi \tag{1}$$

where $\mathcal{L} \triangleq \{1, 2, \ldots, l, \ldots, L\}$ is the set of quality representation for each tile, considering the heterogeneous channel conditions of different users and the limited transmission resource.

Video transcoding is the process of converting one encoding format into another. This is necessary when the network conditions between a sidelink sender and sidelink receivers are worse than the network conditions between the BS and multicast receivers. Considering the latency-sensitive nature of VR services, conventional video transcoding methods, e.g., FFmpeg-based transcoding [33], are not feasible for this system since they do not operate in realtime. In [34], a distributed transcoder system for tiled streaming of ultrahigh resolution 360° VR video was proposed, which has the advantages of realtime transcoding of 16K videos. Therefore, with abundant computing capability, a real-time transcoder can be deployed at the sidelink sender. As shown in Fig. 2(b), the video transcoding process is divided into two phases: 1) decoding the encoded data of quality level $x_{m,n}$ into the original tiles and 2) re-encoding these tiles at a lower quality level $y_{m,n}$.

2) Joint Decoding Scenario (Hereinafter Referred to as Scenario JnD): In this scenario, we assume the theoretical optimal decoding at the physical layer based on the received signals in two phases. Hence, we apply the information-theoretic results, i.e., the achievable rate for the same information through parallel Gaussian channels at different SNR levels (due to different channel gains).

The data process and time schedule of Scenario JnD are illustrated in Fig. 3(a) and (b). As shown in Fig. 3(a), the sidelink sender k_g^*, which has decoded the tile, broadcasts the same information with additional coded bits (i.e., with incremental redundancy) via the sidelink. As shown in Fig. 3(b), the users in $\mathcal{J}_g$ receive and perform joint decoding on both received signals (i.e., from both the downlink and sidelink) to decode the tile with the same quality. We denote by $z_{m,n}$ the tile quality level both conveyed by the BS and the sidelink sender in this joint decoding scenario. Then, we have

$$z_{m,n} \in \mathcal{L}, \quad (m,n) \in \Phi. \tag{2}$$

The main objective of this article is to determine which qualities of the required tiles among the groups should be selected and transmitted by the BS and the sidelink senders, aiming to maximize the overall video quality of all the users.

B. Transmission Model

Different from [35], in the considered system, the multicast downlink and broadcast sidelink are on different frequency channels; thus the interference between the downlink and sidelink is negligible. We denote by B^{DL} and B_g^{DL} the overall multicast downlink bandwidth available at the BS and the bandwidth allocated for user group g, respectively. We allocate

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

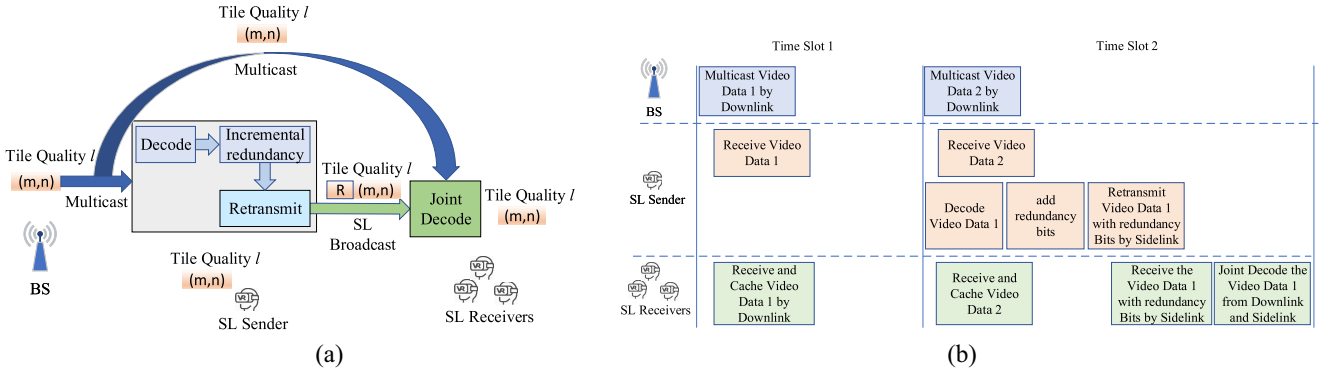


Fig. 3. Illustration of Scenario JnD. The sidelink sender broadcasts the same quality tile with additional coded bits to the sidelink receivers who perform joint decoding of the signals from both the downlink and sidelink. (a) Data process. (b) Time schedule for BS and users.

B^{DL} to all the groups, given by

$$\sum_{g\in|\mathcal{G}|} B_g^{\mathrm{DL}} = B^{\mathrm{DL}}, \quad B_g^{\mathrm{DL}} > 0. \tag{3}$$

Let B^{SL} and B_g^{SL} be the overall bandwidth for the sidelinks and the bandwidth allocated for user group g, respectively. Similar to (3), we also have

$$\sum_{g\in|\mathcal{G}|} B_g^{\mathrm{SL}} = B^{\mathrm{SL}}, \quad B_g^{\mathrm{SL}} > 0. \tag{4}$$

We consider shadow fading P_{LS} and the Rayleigh fading P_{LR} with free space path loss in this model, which follow the log-normal distribution and complex Gaussian distribution, respectively. For user $u \in \mathcal{U}$, the free space path loss is calculated as

$$P_L(x,u) = 32.4 + 20\lg\left(f^x\right)(\mathrm{GHz}) + 20\lg\left(d_{x,u}\right)(\mathrm{m}) \tag{5}$$

where x can be either the BS or a user u' ($u' \neq u, u' \in \mathcal{U}$), and f^x and $d_{x,u}$ denote the center frequency of x and distance between x and u, respectively. The received power of user u can be expressed as

$$P(x,u) = P_x + G_x + G_{\mathrm{U}} - P_L(x,u) - P_{LS} - P_{LR} \tag{6}$$

where G_x and P_x denote the transmit antenna gain and the transmit power of the BS, P_{BS} (or the transmit power of user device, P_{U}), respectively. For simplicity, we assume that all the user devices have the same antenna gain and transmit power.

1) Downlink and Multicast Capacity: We denote $\mathrm{SINR}^{\mathrm{DL}}(B_g^{\mathrm{DL}}, k_g)$ the average signal to interference plus noise ratio (SINR) for user k_g on the downlink, which is given by

$$\mathrm{SINR}^{\mathrm{DL}}\left(B_g^{\mathrm{DL}}, k_g\right) = \frac{P\left(\mathrm{BS}, k_g\right)}{N_0 B_g^{\mathrm{DL}} + \sum_{u'\in U/k_g} P\left(u', k_g\right)} \tag{7}$$

where N_0 is the one-sided power spectral density of the white Gaussian noise.

Subsequently, the downlink capacity between user k_g and the BS is given by

$$C^{\mathrm{DL}}\left(B_g^{\mathrm{DL}}, k_g\right) = B_g^{\mathrm{DL}}\log_2\left(1 + \mathrm{SINR}^{\mathrm{DL}}\left(B_g^{\mathrm{DL}}, k_g\right)\right). \tag{8}$$

The multicast downlink capacity C_g^{DL} of group g can be described as follows, since it is limited by the users with poor channel conditions:

$$C_g^{\mathrm{DL}} = \min_{k_g\in K_g} C^{\mathrm{DL}}\left(B_g^{\mathrm{DL}}, k_g\right). \tag{9}$$

2) Sidelink Capacity: In a similar way, the capacity between user k_g and user j_g on the sidelink is denoted by $C^{\mathrm{SL}}(B_g^{\mathrm{SL}}, k_g, j_g)$, which can be derived as

$$C^{\mathrm{SL}}\left(B_g^{\mathrm{SL}}, k_g, j_g\right) = B_g^{\mathrm{SL}}\log_2\left(1 + \mathrm{SINR}^{\mathrm{SL}}\left(B_g^{\mathrm{SL}}, k_g, j_g\right)\right) \tag{10}$$

where $\mathrm{SINR}^{\mathrm{SL}}$ is the average SINR for user j_g on the sidelink, and can be expressed as

$$\mathrm{SINR}^{\mathrm{SL}}\left(B_g^{\mathrm{SL}}, k_g, j_g\right) = \frac{P\left(k_g, j_g\right)}{N_0 B_g^{\mathrm{SL}} + \sum_{u'\in\mathcal{U}/k_g, j_g} P\left(u', j_g\right)}. \tag{11}$$

Correspondingly, the sidelink capacity C_g^{SL} in group g is limited by the best channel condition among $\min_{j_g\in\mathcal{J}_g} C^{\mathrm{SL}}(B_g^{\mathrm{SL}}, k_g, j_g)$, which can be written as

$$C_g^{\mathrm{SL}} = \max_{k_g\in\mathcal{K}_g}\min_{j_g\in\mathcal{J}_g} C^{\mathrm{SL}}\left(B_g^{\mathrm{SL}}, k_g, j_g\right). \tag{12}$$

C. 360° VR Video Model

Generally, a user watching a 360° VR video is interested in only one FoV at a time, and different users may be watching different FoVs at the same time. Following [36], we assume that the FoV always contains complete tiles exactly; if part of a tile is in the FoV, the entire tile will be included in the FoV. Therefore, there are always an integer number of tiles in the FoV. We denote by Φ_g the set of indices of tiles transmitted to user group g, and $\Phi \triangleq \bigcup_{g\in\mathcal{G}} \Phi_g$ indicates that the user group is formed following the tile groups.

To maintain a high user QoE, we assume that any two adjacent tiles have similar quality levels (i.e., no abrupt changes). Therefore, the tile spatial smoothness is introduced in this model. In addition, considering the different popularity of video tiles, weights are introduced to represent tile popularity.

1) Tile Spatial Smoothness: Following [16], the quality difference between two adjacent tiles is bounded by a parameter $\Delta \in \mathcal{L}\cup\{0\}$. In addition, considering that the first column of a tile is adjacent to the last column of the previous tile, and the

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

first row of a tile is adjacent to the last row of the tile above, we have

$$\left|x_{m,n} - x_{m,(n+1) \bmod N}\right| \leq \Delta, \text{ for } (m,n) \in \Phi, (m, (n+1) \bmod N) \in \Phi \tag{13}$$

$$\left|x_{m,n} - x_{(m+1) \bmod M,n}\right| \leq \Delta \text{for } (m,n) \in \Phi, ((m+1) \bmod M, n) \in \Phi \tag{14}$$

$$\left|y_{m,n} - y_{m,(n+1) \bmod N}\right| \leq \Delta \text{for } (m,n) \in \Phi, (m, (n+1) \bmod N) \in \Phi \tag{15}$$

$$\left|y_{m,n} - y_{(m+1) \bmod M,n}\right| \leq \Delta \text{for } (m,n) \in \Phi, ((m+1) \bmod M, n) \in \Phi \tag{16}$$

$$\left|z_{m,n} - z_{m,(n+1) \bmod N}\right| \leq \Delta \text{for } (m,n) \in \Phi, (m, (n+1) \bmod N) \in \Phi \tag{17}$$

$$\left|z_{m,n} - z_{(m+1) \bmod M,n}\right| \leq \Delta \text{for } (m,n) \in \Phi, ((m+1) \bmod M, n) \in \Phi. \tag{18}$$

2) Tile Weight: Tile weight represents the degree of a spatial region in video segments that attracts attention of users (i.e., popularity). Consider that tiles with higher weight scores represent regions with more attractive textures or objects, higher quality level should be used by the BS and the sidelink sender for such tiles.

In our model, the weight of a tile consists of two parts. From the long-term perspective, the more the requesting users, the higher the weight. We denote by $W_{m,n}^{\text{global}}$ the long-term weight of the (m,n)th tile, which follows the $\mathcal{N}(0,1)$ normal distribution [36]. Usually the tiles at the same (m,n) position in different segments have different weights. On the other hand, for a tile in a user's current FoV, its weight should take into account its position in the FoV [37], [38]: the tile should have a higher weight when it is closer to the center of the FoV. We denote by $W_{m,n}^{\text{FoV}}$ such short-term weight of the (m,n)th tile. Noted that users may have their own interest in viewing videos, and the same tile can be at either the center or margin in different users' FoVs. Thus, $W_{m,n}^{\text{FoV}}$ will be an arithmetic average weight of all the concurrent FoVs.

The overall weight of tile $W_{m,n}$ is thus defined as the product of $W_{m,n}^{\text{global}}$ and $W_{m,n}^{\text{FoV}}$, given by

$$W_{m,n} = W_{m,n}^{\text{global}} \cdot W_{m,n}^{\text{FoV}}. \tag{19}$$

D. Utility Model

The most popular metric to quantify the video QoE is the logarithmic law definition [39]–[41], which is specified as

$$Q = \alpha \log\left(\frac{\beta r_i}{r}\right) \tag{20}$$

where r_i and r denote the actual video rate and requested (maximum) video rate, respectively. The constant parameters α and β are both positive coefficients to normalize the utility Q to remain between 0 and 1, and they can be empirically determined for different applications and videos.

Similarly, we introduce the VR video utility model based on a function of video tile qualities, since the tile quality level is in direct proportion to video rate. Inspired by the models in [39]–[41], we denote Υ_{ind} and Υ_{jnd} as the VR video tile utility of Scenario InD and Scenario JnD, respectively, as

$$\Upsilon_{ind}(m,n) = W_{m,n}\alpha \log\left(\beta \frac{|\mathcal{K}_g| x_{m,n} + \left(|\mathcal{U}_g| - |K_g|\right) y_{m,n}}{|\mathcal{U}_g| L}\right) \tag{21}$$

$$\Upsilon_{jnd}(m,n) = W_{m,n}\alpha \log\left(\beta \frac{z_{m,n}}{L}\right). \tag{22}$$

IV. Problem Formulation

In this article, we aim to achieve the maximum sum quality for all the tiles requested by all user groups by optimizing the quality levels $x_{m,n}$, $y_{m,n}$, and $z_{m,n}$, the multicast receiver users $|\mathcal{K}_g|$ (i.e., the user with the worst wireless condition, $k_{g,\min}$), the sidelink sender k_g^*, and the transmission resource allocation B_g^{DL} and B_g^{SL}. We formulate the optimization problems in this section.

For Scenario InD, the mathematical formulation of the problem is given by

$$\max_{\substack{\{x_{m,n}\},\{y_{m,n}\},\{|\mathcal{K}_g|\},\\ \{k_g^*\},\{B_g^{\text{DL}}\},\{B_g^{\text{SL}}\}}} \quad \mathcal{P}_{ind} = \frac{1}{|\mathcal{G}|} \sum_{g \in |\mathcal{G}|} \sum_{(m,n) \in \Phi_g} \Upsilon_{ind}(m,n) \tag{23a}$$

$$\text{s.t. } b \sum_{(m,n) \in \Phi_g} x_{m,n} \leq C_g^{\text{DL}}, \ \forall g \in \mathcal{G} \tag{23b}$$

$$b \sum_{(m,n) \in \Phi_g} y_{m,n} \leq C_g^{\text{SL}}, \ \forall g \in \mathcal{G} \tag{23c}$$

$$(1),\ (3),\ (4),\ (13)-(16). \tag{23d}$$

The constraints can be explained as follows. Constraints (23b) and (23c) indicate that the overall data size of requested tiles for one user group should not exceed the capacity of the multicast downlink and the sidelink, respectively; constraint (1) ensures that the quality level of tiles sent by BS is no less than that transmitted by the sidelink sender; constraints (3) and (4) enforce that both the BS and sidelink sender allocate all its downlink and sidelink bandwidth for all the groups, respectively; and constraints (13)–(16) specify the smoothness of nearby tiles.

For Scenario JnD, the problem formulation is given by

$$\max_{\substack{\{z_{m,n}\},\{|\mathcal{K}_g|\},\\ \{k_g^*\},\{B_g^{\text{DL}}\},\{B_g^{\text{SL}}\}}} \quad \mathcal{P}_{jnd} = \frac{1}{|\mathcal{G}|} \sum_{g \in |\mathcal{G}|} \sum_{(m,n) \in \Phi_g} \Upsilon_{jnd}(m,n) \tag{24a}$$

$$\text{s.t. } b \sum_{(m,n) \in \Phi_g} z_{m,n} \leq C_g^{jnd}, \ \forall g \in |\mathcal{G}| \tag{24b}$$

$$(2),\ (3),\ (4),\ (17),\ (18) \tag{24c}$$

where C_g^{jnd} in Problem (24a) is given by

$$C_g^{jnd} = \min\left\{C_g^{\text{DL}},\ C^{\text{DL}}\left(B_g^{\text{DL}}, j_g\right) + C_g^{\text{SL}}\right\} \tag{25}$$

which means that the overall data size of requested tiles for one user group should not exceed the minimum value of group's multicast capacity and the sum of sidelink capacity and the downlink capacity of the user with the worst wireless channel.

Both Problems (23a) and (24a) are MINLP problems, which are hard to solve in polynomial time. A general method for a suboptimal solution is to make a continuous relaxation of the

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

discrete constraints. Specifically, constraints (1) and (2) can be relaxed as

$$x_{m,n} \geq y_{m,n}, x_{m,n} \in [1, L], y_{m,n} \in [1, L], (m, n) \in \Phi \quad (26)$$

$$z_{m,n} \in [1, L], \ (m, n) \in \Phi. \quad (27)$$

Furthermore, k_g^* and $|\mathcal{K}_g|$ are also discrete variables, which are related to bandwidth allocation, and the selected tile quality level $x_{m,n}$, $y_{m,n}$, and $z_{m,n}$. It is hard to make a continuous relaxation by common methods. Under this circumstance, an intuitive approach to solving the problems is exhaustive search, i.e., finding the maximum value corresponding to all the combinations of k_g^* and $\mathcal{K}_g$, where $g \in \mathcal{G}$. When k_g^* and $\mathcal{K}_g$ are fixed, the original problems are reduced to a standard convex problem and can be solved by many optimization tools.

Proposition 1: The computational complexity of the exhaustive search approach for Problem (23a) is $\mathcal{O}(|\mathcal{U}|^{|\mathcal{G}|}(Q^2R^{2.5} + R^{3.5}))$, where $Q = 2(MN + |\mathcal{G}|)$ and $R = 7MN + 2|\mathcal{G}|$.

Proof: When k_g^* and $|\mathcal{K}_g|$ are fixed, Problem (23a) can be solved by popular optimization tools. We choose the SeDuMi solver for its low asymptotic computational complexity, which is $\mathcal{O}(Q^2R^{2.5} + R^{3.5})$, where Q and R are the numbers of decision variables and linear matrix inequalities (LMIs), respectively [42].

Then we can see that there are MN $\{x_{m,n}\}$, MN $\{y_{m,n}\}$, $|\mathcal{G}|$ $\{B_g^{\mathrm{DL}}\}$, and $|\mathcal{G}|$ $\{B_g^{\mathrm{SL}}\}$ decision variables in Problem (23a). Thus, we have $Q = 2(MN + |\mathcal{G}|)$. In addition, it is noted that (1), (3), (4), (13)–(16), (23b), and (23c), are LMIs. Then we have $R = 7MN + 2|\mathcal{G}|$.

For all the groups, $\prod_{g\in\mathcal{G}}(|\mathcal{U}_g| - 1)$ optimizations are needed to obtain the optimal solution, since there are $|\mathcal{U}_g| - 1$ possible selections for k_g^* and $\mathcal{K}_g$ in one group $\mathcal{U}_g$. Considering that $|\mathcal{U}_g| \leq |\mathcal{U}|$, the computational complexity of the exhaustive search approach is $\mathcal{O}(|\mathcal{U}|^{|\mathcal{G}|}(Q^2R^{2.5} + R^{3.5}))$, where $Q = 2(MN + |\mathcal{G}|)$ and $R = 7MN + 2|\mathcal{G}|$. ■

Proposition 2: The computational complexity of the exhaustive search approach for Problem (24a) is $\mathcal{O}(|\mathcal{U}|^{|\mathcal{G}|}(Q^2R^{2.5} + R^{3.5}))$, where $Q = MN + 2|\mathcal{G}|$ and $R = 3MN + 2|\mathcal{G}|$.

Proof: Similar to Proposition 1, there are MN $\{z_{m,n}\}$, $|\mathcal{G}|$ $\{B_g^{\mathrm{DL}}\}$, and $|\mathcal{G}|$ $\{B_g^{\mathrm{SL}}\}$ decision variables in Problem (24a). Thus, we have $Q = MN + 2|\mathcal{G}|$. In addition, it is noted that (3), (4), (17), (18), and (24b) are LMIs, and we have $R = 3MN + 2|\mathcal{G}|$.

For all the groups, there are also $|\mathcal{U}|^{|\mathcal{G}|}$ iterations for obtaining the optimal solution, and the computational complexity is $\mathcal{O}(|\mathcal{U}|^{|\mathcal{G}|}(Q^2R^{2.5} + R^{3.5}))$, where $Q = MN + 2|\mathcal{G}|$ and $R = 3MN + 2|\mathcal{G}|$. ■

V. Two-Stage Optimization Algorithm

According to Propositions 1 and 2, the exhaustive search approach could be extremely time consuming and the optimal solution cannot be found in polynomial time as the number of users and the number of groups $|\mathcal{G}|$ increase. Therefore, in this section, a two-stage optimization algorithm with much lower computational complexity is proposed to compute a suboptimal solution. Since Problems (23a) and (24a) have the same structure, we will focus on the design of the two-stage optimization algorithm for Problem (23a) for brevity. The algorithm for solving Problem (24a) has a similar design.

Specifically in Stage I, we fix the bandwidth allocation to find a quasi-optimal solution k_g^* and $|\mathcal{K}_g|$ with a sidelink receivers and sender selecting algorithm. At Stage II, we execute a bandwidth allocation and tile quality level selection scheme based on the k_g^* and $|\mathcal{K}_g|$ solved in Stage I. Considering that the change of allocated bandwidth in each group may cause some changes on the multicast downlink rate and sidelink rate, and consequently, lead to changes on the value of k_g^* and $|\mathcal{K}_g|$, the solution procedure will iterate over these two stages until the result is convergent. The details of the two-stage approach are given in the following.

A. Stage I: Sidelink Receivers and Sender Selection

The goal of Stage I is to determine the specific sidelink sender k_g^* and the number of users $|\mathcal{K}_g|$ for each user group g. Accordingly, the user with the worst channel condition in $\mathcal{K}_g$ and $\mathcal{J}_g$, which is denoted as $k_{g,\min}$ and $j_{g,\min}$, respectively, can be identified. By virtually allocating fixed amounts of bandwidth $B_{\mathrm{fixed}}^{\mathrm{DL}}$ and $B_{\mathrm{fixed}}^{\mathrm{SL}}$ to each group, the original problem can be decoupled into $|\mathcal{G}|$ independent problems. The decoupled problem of group g can be formulated as

$$\max_{\{x_{m,n}\},\{y_{m,n}\},\{|\mathcal{K}_g|\},\left\{k_g^*\right\}} \quad \Upsilon_{ind}^{\dagger}(g) = \sum_{(m,n)\in\Phi_g} \Upsilon_{ind}(m, n) \quad (28a)$$

$$\text{s.t.} \ \ b \sum_{(m,n)\in\Phi_g} x_{m,n} \leq C_g^{\mathrm{DL}\dagger} \quad (28b)$$

$$b \sum_{(m,n)\in\Phi_g} y_{m,n} \leq C_g^{\mathrm{SL}\dagger} \quad (28c)$$

$$(13)-(16), \ (26) \quad (28d)$$

where $C_g^{\mathrm{DL}\dagger} = \min_{k_g\in K_g} C^{\mathrm{DL}}(B_{\mathrm{fixed}}^{\mathrm{DL}}, k_g)$ and $C_g^{\mathrm{SL}\dagger} = \max_{k_g\in\mathcal{K}_g} \min_{j_g\in\mathcal{J}_g} C^{\mathrm{SL}}(B_{\mathrm{fixed}}^{\mathrm{SL}}, k_g, j_g)$.

The sidelink receivers and sender selection algorithm (SRSSA) is summarized in Algorithm 1. First, the users in group g are sorted according to their wireless channel conditions in the descending order, and the information of users are restored in $\mathcal{U}_g$. Second, we choose the top $|\mathcal{K}_g|$ users in $\mathcal{U}_g$ to compute $C_g^{\mathrm{DL}\dagger}$ and $C_g^{\mathrm{SL}\dagger}$, and the objective function (28a) is solved by a standard optimization solver afterwards. To find the optimal object value for (28a), $|\mathcal{K}_g|$ is changed from $|\mathcal{U}_g| - 1$ to 1 by detaching one user with the worst channel condition at a time. Third, find k_g^* and $|\mathcal{K}_g|$ according to the maximum value among the $|\mathcal{U}_g| - 1$ results.

B. Stage II: Bandwidth Allocation and Tile Quality Level Selection

The goal of Stage II is to properly allocate bandwidth and select the tile quality level. Based on the results of Algorithm 1, we rewrite constraints (23b) and (23c) as

$$b \sum_{(m,n)\in\Phi_g} x_{m,n} \leq C^{\mathrm{DL}}\left(B_g^{\mathrm{DL}}, k_{g,\min}\right) \quad (29)$$

$$b \sum_{(m,n)\in\Phi_g} y_{m,n} \leq C^{\mathrm{SL}}\left(B_g^{\mathrm{SL}}, k_g^*, j_{g,\min}\right) \quad (30)$$

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

Algorithm 1 SRSSA

Input: $\mathcal{U}_g$, $|\mathcal{U}_g|$, $B_{\text{fixed}}^{\text{DL}}$, and $B_{\text{fixed}}^{\text{SL}}$;
Output: k_g^*, $|\mathcal{K}_g|$ (namely, $k_{g,\min}$), and $j_{g,\min}$;
1: Sort $\mathcal{U}_g$ in descending order according to channel condition;
2: **for** $1 \le |\mathcal{J}_g| \le |\mathcal{U}_g| - 1$ **do**
3: j_g is the index of the user with the $|\mathcal{J}_g|$ worst channel condition;
4: **for** $k_g \in \mathcal{K}_g$ **do**
5: Compute $C^{\text{DL}}\left(B_{\text{fixed}}^{\text{DL}}, k_g\right)$;
6: Compute $C^{\text{SL}}\left(B_{fixed}^{\text{SL}}, k_g, j_g\right)$;
7: **end for**
8: Find $\min\limits_{k_g \in \mathcal{K}_g} C^{\text{DL}}\left(B_{\text{fixed}}^{\text{DL}}, k_g\right)$;
9: Find $\max\limits_{k_g \in \mathcal{K}_g} C^{\text{SL}}\left(B_{fixed}^{\text{SL}}, k_g, j_g\right)$;
10: Compute temporary $\widetilde{\Upsilon}_{ind}^{\dagger}(g)$ by a standard optimization solver;
11: **end for**
12: Obtain k_g^*, $|\mathcal{K}_g|$, $k_{g,\min}$, and $j_{g,\min}$ according to the maximum value among all the $\widetilde{\Upsilon}_{ind}^{\dagger}(g)$s;

Algorithm 2 Bandwidth Allocating Algorithm (BAA)

Input: k_g^*, $|\mathcal{K}_g|$, $k_{g,\min}$ and $j_{g,\min}$;
Output: $\{x_{m,n}^*\}$,$\{y_{m,n}^*\}$, $\{B_g^{\text{DL}*}\}$, and $\{B_g^{\text{SL}*}\}$;
1: Solve Problem (31a) with a convex optimization solver;

Algorithm 3 Two-Stage Iteration Optimization Algorithm

Input: ε;
Output: $B_g^{\text{DL}*}$, $B_g^{\text{SL}*}$, $x_{m,n}^*$, $y_{m,n}^*$, k_g^*, $|\mathcal{K}_g|$, and Υ_{ind}^*;
1: Set $\ell = 1$, $\ell_{max} = |\mathcal{G}| \times \max_{g \in |\mathcal{G}|} |\mathcal{U}_g|$, and $\Upsilon_{ind}^* = 0$;
2: **repeat**
3: **if** $\ell == 1$ **then**
4: Obtain $k_g^*(\ell)$ and $|\mathcal{K}_g(\ell)|$ by Algorithm 1 with B_{fixed}^{DL} and B_{fixed}^{SL};
5: **else**
6: Obtain $k_g^*(\ell)$ and $|\mathcal{K}_g(\ell)|$ by Algorithm 1 with $B_g^{\text{DL}*}(\ell)$, $B_g^{\text{SL}*}(\ell)$;
7: **end if**
8: Compute $B_g^{\text{DL}*}(\ell)$, $B_g^{\text{SL}*}(\ell)$, $x_{m,n}^*(\ell)$, $y_{m,n}^*(\ell)$, and $\Upsilon_{ind}^*(\ell)$ by Algorithm 2 with $k_g^*(\ell)$ and $|\mathcal{K}_g(\ell)|$;
9: **if** $\Upsilon_{ind}^* < \Upsilon_{ind}^*(\ell)$ **then**
10: $\Upsilon_{ind}^* = \Upsilon_{ind}^*(\ell)$;
11: $B_g^{\text{DL}*} = B_g^{\text{DL}*}(\ell)$ and $B_g^{\text{SL}*} = B_g^{\text{SL}*}(\ell)$;
12: $k_g^* = k_g^*(\ell)$ and $|\mathcal{K}_g| = |\mathcal{K}_g(\ell)|$;
13: $x_{m,n}^* = x_{m,n}^*(\ell)$;
14: $y_{m,n}^* = y_{m,n}^*(\ell)$;
15: **end if**
16: $\ell = \ell + 1$;
17: **until** $\ell \ge \ell_{max}$
18: $x_{m,n}^* = \lfloor x_{m,n}^* \rfloor$;
19: $y_{m,n}^* = \lfloor y_{m,n}^* \rfloor$;

where k_g^*, $k_{g,\min}$, and $j_{g,\min}$ are imported from Algorithm 1.

Problem (23a) is then reformulated as follows:

$$\max_{\substack{\{x_{m,n}^*\},\{y_{m,n}^*\},\\ \{B_g^{\text{DL}*}\},\{B_g^{\text{SL}*}\}}} \quad \Upsilon_{ind}^{\ddagger} = \frac{1}{|\mathcal{G}|} \sum_{g \in |\mathcal{G}|} \sum_{(m,n) \in \Phi_g} \Upsilon_{ind}(m,n) \tag{31a}$$

$$\text{s.t.} \quad (3),\ (4),\ (13)-(16),\ (29),\ (30),\ (26). \tag{31b}$$

All the constraints in (31b) are convex, and the objective function (31a) is concave. Thus, Problem (31a) is a concave optimization, and can be solved effectively using a standard convex optimization solver. The bandwidth allocation algorithm is presented in Algorithm 2.

C. Two-Stage Iteration

Considering that a different bandwidth allocation for each group usually leads to change of k_g^*, $|\mathcal{K}_g|$ (namely, $k_{g,\min}$), and $j_{g,\min}$, we take multiround iterations over the above two stages to obtain the optimal solution. The optimal solution can be achieved after, at most, $|\mathcal{G}| \cdot \max_{g \in |\mathcal{G}|} |\mathcal{U}_g|$ iterations. The iterative procedure is shown in Algorithm 3.

Note that, we have relaxed the discrete constraint (1) to a continuous constraint (26). Thus, the solution after iterations may not be feasible for Problem (23a). Since $\Upsilon_{ind}^* > \Upsilon_{ind}$, we round down $x_{m,n}^*$ and $y_{m,n}^*$ to $\lfloor x_{m,n}^* \rfloor$ and $\lfloor y_{m,n}^* \rfloor$, where $\lfloor x_{m,n}^* \rfloor$ and $\lfloor y_{m,n}^* \rfloor$ denote the greatest integer less than or equal to $x_{m,n}^*$ and $y_{m,n}^*$. Finally, the suboptimal solution to Problem (23a) is denoted by $\{\lfloor x_{m,n}^* \rfloor, \lfloor y_{m,n}^* \rfloor, k_g^*, |\mathcal{K}_g|, B_g^{\text{DL}*}, B_g^{\text{SL}*}\}$. We have the following proposition on computational complexity of the procedure.

Proposition 3: The computational complexity of the two-stage algorithm for Scenario InD is $\mathcal{O}(|\mathcal{U}|^3 - |\mathcal{U}| + |\mathcal{G}|^2 U^2 (Q^2 R^{2.5} + R^{3.5}))$, where $Q = 2G(MN + 2)$ and $R = MN(2G + 5)$.

Proof: There are two parts of computational complexity in the two-stage algorithm.

1) *The Computational Complexity of $C_g^{DL\dagger}$ and $C_g^{SL\dagger}$:* For all the groups, $C_g^{\text{DL}\dagger}$ requires $\sum_{g \in \mathcal{G}} |\mathcal{U}_g|(|\mathcal{U}_g| - 1)/2$ iterations, since $|\mathcal{K}_g|$ starts from $(|\mathcal{U}_g| - 1)$ to 1 in $(|\mathcal{U}_g| - 1)$ loops. Considering that $|\mathcal{U}_g| \le |\mathcal{U}|$, the computational complexity of $C_g^{\text{DL}\dagger}$ is upper bounded by $\mathcal{O}(|\mathcal{U}|(|\mathcal{U}| - 1))$. Computing $C_g^{\text{SL}\dagger}$ takes $\sum_{|\mathcal{K}_g|=1}^{|\mathcal{U}_g|-1} |\mathcal{K}_g|(|\mathcal{U}_g| - |\mathcal{K}_g|)$ iterations, which is $(|\mathcal{U}_g|^3 - |\mathcal{U}_g|)/6$. Considering that $|\mathcal{U}_g| \le |\mathcal{U}|$, the computational complexity of $C_g^{\text{SL}\dagger}$ is in the order of $\mathcal{O}((|\mathcal{U}|^3 - |\mathcal{U}|))$. The $C_g^{\text{DL}\dagger}$ and $C_g^{\text{SL}\dagger}$ are computed jointly in Algorithm 1, and the computational complexity should be $\mathcal{O}((|\mathcal{U}|^3 - |\mathcal{U}|))$.
2) *The Computational Complexity of Solving the Convex Problem:* This is the computational complexity for finding the optimal values for $x_{m,n}$, $y_{m,n}$, B_g^{DL}, and B_g^{SL}. An optimization solver is to be called $\sum_{g \in |\mathcal{G}|} (|\mathcal{U}_g| - 1) + 1$ times, and there are at most $|\mathcal{G}| \cdot \max_{g \in |\mathcal{G}|} |\mathcal{U}_g|$ iterations. In Problem (23a), we have $Q = 2G(MN + 2)$ decision variables and $R = MN(2G + 5)$ LMIs.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

Therefore, the computational complexity of the two-stage algorithm is $\mathcal{O}(|\mathcal{U}|^3 - |\mathcal{U}| + |\mathcal{G}|^2 U^2 (Q^2 R^{2.5} + R^{3.5}))$. ■

Proposition 4: The space complexity of the two-stage algorithm for Scenario InD is $\mathcal{O}(|\mathcal{U}|)$.

Proof: It requires at most $|\mathcal{U}|$ units of storage for computing $C_g^{\mathrm{DL}\dagger}$ and $C_g^{\mathrm{SL}\dagger}$, and the variables of the two-stage algorithm are linearly proportional to the number of users and the number of tiles in each segment. Therefore, the space complexity of the two-stage algorithm is $\mathcal{O}(|\mathcal{U}|)$. ■

D. Two-Stage Algorithm for Scenario JnD

Similar to Scenario InD, the two-stage algorithm can be utilized to solve Problem (24a) as well with the following slight adaptations.

1) Compute $C^{\mathrm{DL}}(B_{\mathrm{fixed}}^{\mathrm{DL}}, j_{g,\min})$.
2) Compute C_g^{jnd}.
3) The tile quality received by multicast receivers and sidelink receivers in the same group is consistent.

Proposition 5: The computational complexity of the two-stage algorithm for Scenario JnD is $\mathcal{O}(|\mathcal{U}|^3 - |\mathcal{U}| + |\mathcal{G}|^2 U^2 (Q^2 R^{2.5} + R^{3.5}))$, where $Q = |\mathcal{G}|(MN + 4)$ and $R = MN(|\mathcal{G}| + 2)$.

Proof: Similar to the two-stage algorithm for Scenario InD, there are two parts of computational complexity in the two-stage algorithm for Scenario JnD.

The computational complexity of C_g^{jnd} in Stage I is $\mathcal{O}((|\mathcal{U}|^3 - |\mathcal{U}|))$, since C_g^{DL} and C_g^{SL} should be computed as well in Scenario JnD.

There are $Q = |\mathcal{G}|(MN + 4)$ decision variables and $R = MN(|\mathcal{G}| + 2)$ LMIs for solving Problem (24a).

Therefore, the computational complexity of the two-stage algorithm for Scenario InD is $\mathcal{O}(|\mathcal{U}|^3 - |\mathcal{U}| + |\mathcal{G}|^2 U^2 (Q^2 R^{2.5} + R^{3.5}))$. ■

VI. Simulation Evaluation and Discussion

In this section, the performance of the proposed algorithms for the two scenarios are evaluated with simulations, respectively. The parameters used in our simulations are presented in Table II. Specifically, we consider one cell with a coverage area of 250×250 m^2. 50 users are randomly distributed in the cell area, at distances larger than 100 m from the BS and no more than 6 m from each other, respectively. Similar to the standard LTE parapemers [43], the carrier frequency of BS-user downlink and sidelink are 2.6 and 2.4 GHz, respectively. We consider communications with 20-MHz bandwidth for both the downlink and the sidelink. The log-normal shadowing has an 8-dB standard deviation and the Rayleigh fading follows a complex Gaussian process with a 1-dB standard deviation. The one-sided power spectral density of the white Gaussian noise is $N_0 = -134$ dBm/MHz. The video library hosts 100 360° videos, each with 1000 frames. Each 360° video frame contains $M \times N = 8 \times 10$ tiles, and each user requires 40 tiles uniformly. There are $L = 10$ levels of quality for each tile and $b = 8.408 \times 10^5$ [16].

The proposed algorithms for solving the problems of Scenario InD and Scenario JnD are termed by the independent decoding algorithm (InDec) and joint decoding algorithm (JnDec), respectively. The following baseline schemes are also simulated for comparison purpose.

1) The optimal solutions (Exhaustive-InD and Exhaustive-JnD), which are the upper bound of the proposed algorithms and are obtained by exhaustive search.
2) The equal-bandwidth allocation algorithm with sidelink transmissions (Eq-SL).
3) The equal-bandwidth allocation without sidelink transmissions (Eq-NoSL).
4) The bandwidth allocation optimization algorithm without sidelink transmissions (Ba-NoSL), which considers multicast transmissions only. The Ba-NoSL algorithm is a simplified version of Algorithm 1 proposed in [12], but the power is not optimized for a consistent performance comparison.
5) Random downlink (and sidelink) bandwidth allocation, which is termed "Random."

The equal-resource allocation schemes (i.e., Eq-SL and Eq-NoSL) allocate the same amount of bandwidth resource to each multicast downlink (and sidelink) transmission. In each simulation, 100 random independent channel realizations are generated using MATLAB and the Sedumi toolbox.

TABLE II
Default Simulated Parameters

Parameters	*Values*
Cell area	250 m × 250 m
Distance between BS and Users	$\geq$ 100 m
Distance among users	$[2, 65]$ m
Number of users	50
Bandwidth of BS downlink	20 MHz
Bandwidth of sidelink	20 MHz
N_0	-134 dBm/MHz
Shadow fading	Log-normal, mean = 0, sigma=8
Rayleigh fast fading	complex Gaussian, mean = 0, sigma=1
Carrier Frequency of BS downlink	2.6 GHz
Carrier Frequency of sidelink	2.4 GHz
BS transmit power	46 dBm
User equipment transmit power	23 dBm
Antenna gain	User: -1 dBi; BS: 16 dBi
Antenna configure for BS	SISO (1Tx 1Rx)
Antennas for User	1 for downlink; 1 for sidelink
Tiles per 360° video frame	80
Frames per video	1000
Number of videos	100
Tile quality smoothness Δ	1
Number of user groups	$[10, 50]$

A. Complexity of the Algorithms

According to Propositions 1–3 and 5, the proposed algorithms have a much lower complexity than that of the exhaustive search, and the Algorithm JnDec has a lower complexity than that of the Algorithm InDec. In this simulation, we set a fixed number of groups (i.e., $|\mathcal{G}| = 10$) and fixed number of group users (i.e., $|\mathcal{U}_g| = 5, g \in |\mathcal{G}|$) to verify the complexity of the algorithms. We find the simulation results are consistent with the propositions. As shown in Table III, the proposed

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

TABLE III
EXECUTION TIMES AND TOTAL UTILITIES OF THE FOUR PROPOSED ALGORITHMS WITH $B^{\mathrm{DL}} = B^{\mathrm{SL}} = 20$ MHz

Algorithms	Computation time (s)		Total utility
	Stage I	Stage II	
Exhaustive-InD	2,002,780.16		130
Exhaustive-JnD	1,405,091.84		149
Algorithm InDec	17.01	1.91	122
Algorithm JnDec	15.08	1.34	144

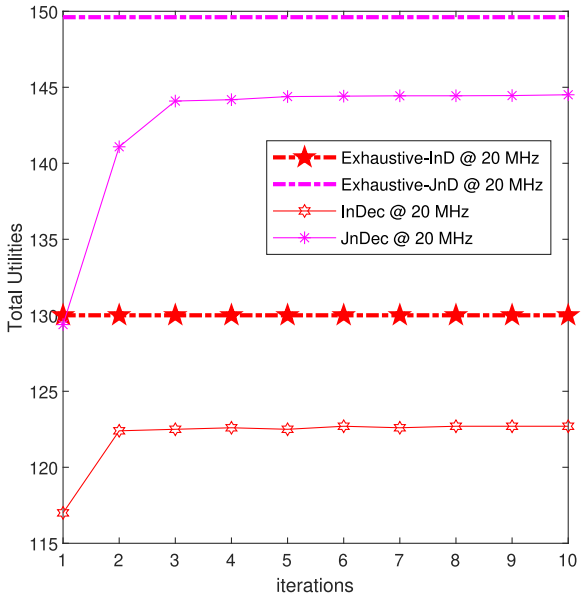


Fig. 4. Convergence of the proposed algorithms with 20-MHz downlink and 20-MHz sidelink bandwidth.

algorithms save more than 20 000 times of execution time than the exhaustive search-based algorithms. We also find that the Algorithm InDec is more time consuming while it has a similar utility performance as the Algorithm JnDec. However, there are also disadvantages for JnDec, since it requires users buffer the received signals, which increases the receiver complexity.

B. Algorithm Convergence Performance

Fig. 4 shows the convergence of the two proposed algorithms with 20-MHz multicast and 20-MHz sidelink bandwidth. We can see that the suboptimal solutions of both Algorithm InDec and Algorithm JnDec are obtained after at most four iterations. Both proposed iterative algorithms converge very fast, and the total utilities they achieve are also close to the optimal solutions produced by the time-consuming exhaustive search-based algorithms, respectively.

C. Total Utility Versus Bandwidth

Fig. 5 demonstrates the total utilities achieved by the proposed algorithms (InDec and JnDec), Exhaustive-InD algorithm, Exhaustive-JnD algorithm, Eq-SL algorithm, Eq-NoSL algorithm, Ba-NoSL algorithm, and Random algorithm with respect to BS downlink bandwidth B^{DL} and sidelink bandwidth B^{SL}, respectively. From Fig. 5(a), we can see that the

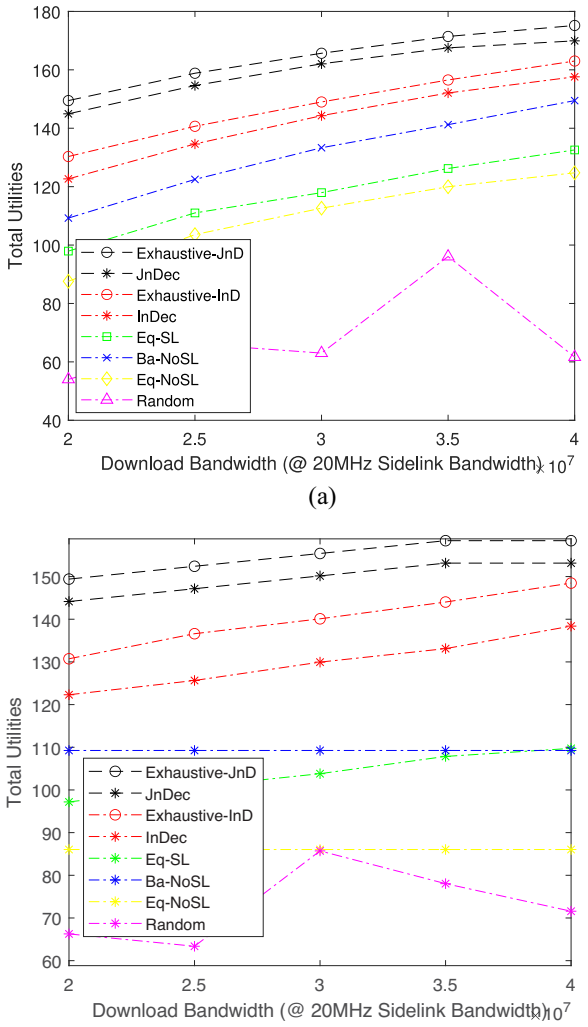


Fig. 5. Total utilities achieved by the algorithms with different bandwidths. (a) Total utilities with different B^{DL}. (b) Total utilities with different B^{SL}.

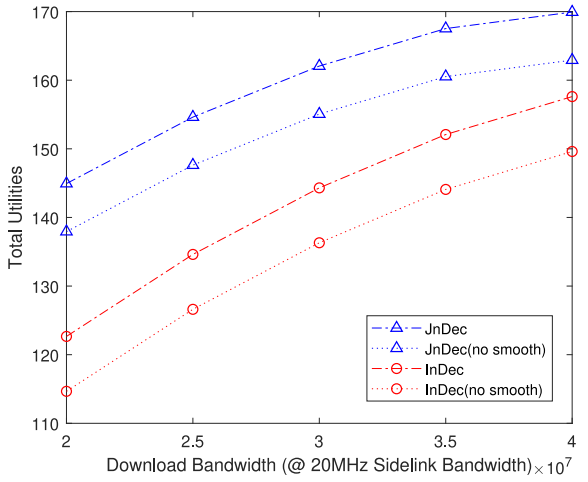


Fig. 6. Total utilities performance of the proposed algorithms with respect to tile quality smoothness.

total utilities of all the algorithms (except the Random algorithm) increase with B^{DL}, while the two proposed algorithms always outperform the other baseline algorithms and achieve

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

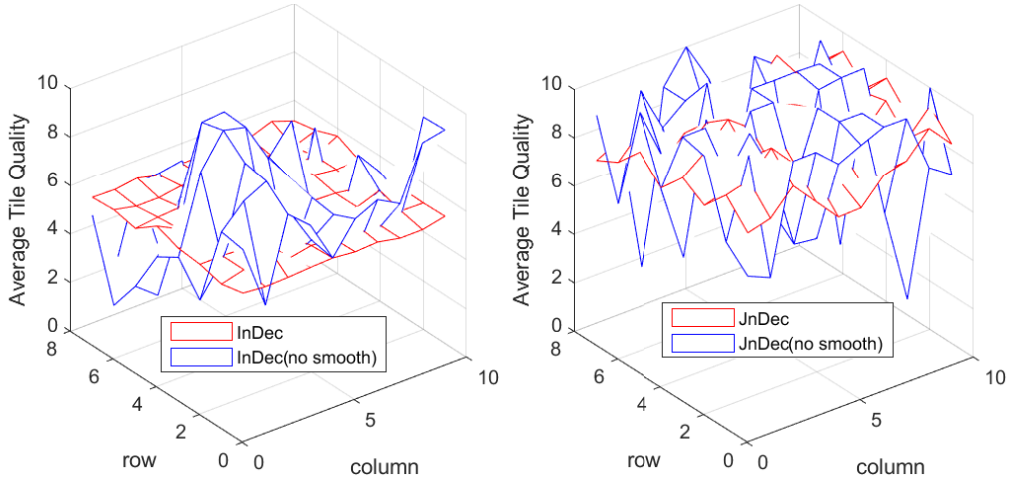


Fig. 7. Average tile quality level performance of the proposed algorithms with respect to tile quality smoothness.

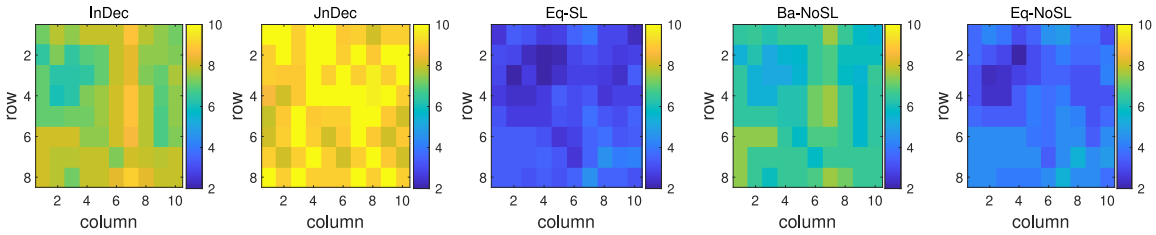


Fig. 8. Average tile quality performance of different algorithms when $B^{\mathrm{DL}} = B^{\mathrm{SL}} = 20$ MHz.

near-optimal performance (as their total utilities are close to the corresponding upper bounds, i.e., the utilities achieved by the Exhaustive-InD and Exhaustive-JnD, respectively). These results demonstrate the effectiveness of the proposed suboptimal solutions. It is shown that the proposed algorithms achieve 12%, 32%, and 46% performance gain over Ba-NoSL, Eq-SL, and Eq-NoSL, respectively. For the average bandwidth allocation algorithms, the utility performance of the algorithm Eq-SL is still 11% higher than that of Algorithm Eq-NoSL, which demonstrate the efficacy of sidelink transmissions.

As shown in Fig. 5(b), our proposed algorithms also outperform the other baseline algorithms with increased sidelink bandwidth. It is worth noting that the performance of JnDec no longer grows when the sidelink bandwidth exceeds 35 MHz. The reason is that the video tile quality is limited by C_g^{jnd} defined in (25), which is the minimum value of C_g^{DL} and $C^{\mathrm{DL}}(B_g^{\mathrm{DL}}, j_g) + C_g^{\mathrm{SL}}$. As C_g^{SL} increases as more bandwidth for the sidelink, C_g^{jnd} will be equal to C_g^{DL} (a constant value, fixed 20-MHz downlink bandwidth), and the total utility will become constant. Furthermore, we can see that the total utilities of all the algorithms rise slowly with the increase of B^{SL}, since the tile quality sent by the sidelink should not be higher than that of the BS multicast downlink.

D. Total Utility Versus Smoothness

Obviously, the results of problems (23a) and (24a) will be different if the tile quality smoothness constraints, i.e., (13)–(18), are not considered. However, the total utilities can still be obtained with the proposed two-stage optimization algorithms. The new algorithms without considering tile quality smoothness in Scenario JnD and InD are termed JnDec (no smooth) and InDec (no smooth), respectively. Figs. 6 and 7 illustrate the impact of tile quality smoothness on the total utility and the average tile quality performance. We can see in Fig. 6 that the total utilities of the JnDec (no smooth) and InDec (no smooth) algorithms are lower than that of the JnDec and InDec algorithms, respectively. In Fig. 7, the tile quality levels of the standard JnDec and InDec algorithms are much smoother than that of JnDec (no smooth) and InDec (no smooth). The latter consists of many high peaks and deep valleys, which greatly degrade the user QoE.

E. Tile Quality Level Performance

As shown in Fig. 8, we present the average tile quality level performance among different algorithms at 20-MHz downlink and 20-MHz sidelink bandwidth in the color block charts. The chart with 8×10 blocks represents the video frame, and each block with different colors stands for one tile. The average tile quality level increases as the block color changes from blue to yellow. For example, the average tile quality level of a yellow block is higher than those of blue blocks. We can see that there are more block in yellow colors when the InDec and JnDec algorithms are used, meaning that among all the compared algorithms, the proposed algorithms have the ability to select the highest quality representations for users, which effectively increases the user QoE. Note that we do not show the average tile quality level performance of the Random algorithm, which is poor due to the stochastic behavior in different simulations.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

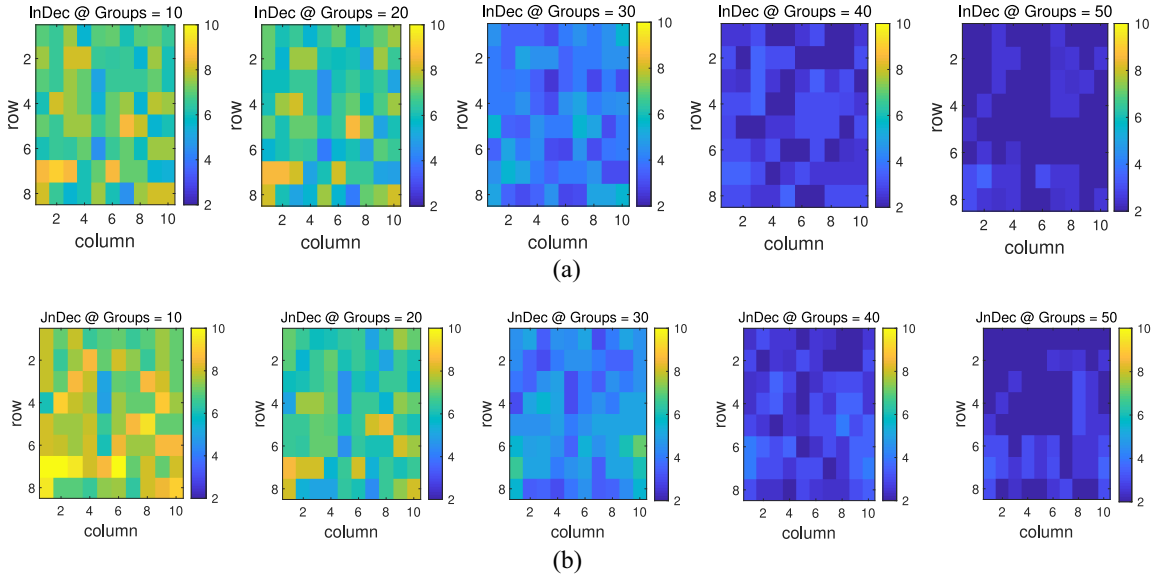


Fig. 9. Average tile quality performance of the proposed algorithm with respect to different numbers of user groups when $B^{\mathrm{DL}} = B^{\mathrm{SL}} = 20$MHz. (a) Average tile quality performance of InDec. (b) Average tile quality performance of JnDec.

Fig. 9 verifies the utility performance with different numbers of user groups, while the multicast downlink and the sidelink bandwidths are both fixed at 20 MHz. We can see that the tile quality level decreases when the number of groups is increased from 10 to 50, since less bandwidth can be allocated to each group as there are more groups. However, the proposed algorithms still outperform the baseline algorithms, which means that their tile quality levels are much higher than those of the baseline algorithms.

VII. Conclusion

In this article, a sidelink-aided multicast system for multi-quality tiled 360° VR video was proposed to achieve high user experience under the constraints of bandwidth resource and tile quality smoothness in overloaded situations. We formulated optimization problems that took into account the relationship among tile quality level selection, sidelink sender selection, and bandwidth allocation to maximize the total utility of all users for the independent decode scenario and the joint decode scenario, respectively. Two-stage iterative algorithms were proposed to obtain suboptimal solutions with low computational complexity. The simulation results demonstrated the many advantages of the proposed algorithm by exploiting sidelink transmissions for tiled 360° VR video multicast.

References

[1] H. Bellini. (Feb. 2016). *The Real Deal With Virtual and Augmented Reality*. [Online]. Available: http://www.goldmansachs.com/our-thinking/pages/virtual-and-augmented-reality.html

[2] L. Graham. (Oct. 2016). *Citi Eyes a Trillion-Dollar Industry in Virtual Reality Technology*. [Online]. Available: http://www.cnbc.com/2016/10/14/citi-eyesa-trillion-dollar-industry-in-virtual-reality-technology.html

[3] J. Dai, Z. Zhang, S. Mao, and D. Liu, "A view synthesis-based 360° VR caching system over MEC-enabled C-RAN," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 10, pp. 3843–3855, Oct. 2020.

[4] T. Zhang and S. Mao, "Joint power and channel resource optimization in soft multi-view video delivery," *IEEE Access*, vol. 7, pp. 148084–148097, 2019.

[5] Huawei, Inc. (2016). *White Paper: Hosted Network Requirements Oriented on VR Service*. [Online]. Available: http://www.imxdata.com/archives/17346

[6] M. Hosseini and V. Swaminathan, "Adaptive 360 VR video streaming: Divide and conquer," in *Proc. IEEE ISM*, San Jose, CA, USA, Dec. 2016, pp. 107–110.

[7] M. Graf, C. Timmerer, and C. Mueller, "Towards bandwidth efficient adaptive streaming of omnidirectional video over HTTP: Design, implementation, and evaluation," in *Proc. ACM MMSys*, Taipei, Taiwan, Jun. 2017, pp. 261–271.

[8] A. Zare, A. Aminlou, M. M. Hannuksela, and M. Gabbouj, "HEVC-compliant tile-based streaming of panoramic video for virtual reality applications," in *Proc. ACM Int. Conf. Multimedia*, Amsterdam, The Netherlands, Oct. 2016, pp. 601–605.

[9] R. Skupin, Y. Sanchez, D. Podborski, C. Hellge, and T. Schierl, "HEVC tile based streaming to head mounted displays," in *Proc. IEEE CCNC*, Las Vegas, NV, USA, Jan. 2017, pp. 613–615.

[10] D. V. Nguyen, H. T. T. Tran, A. T. Pham, and T. C. Thang, "An optimal tile-based approach for viewport-adaptive 360-degree video streaming," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 29–42, Mar. 2019.

[11] Y. Bao, T. Zhang, A. Pande, H. Wu, and X. Liu, "Motion-prediction-based multicast for 360-degree video transmissions," in *Proc. IEEE SECON*, San Diego, CA, USA, Jun. 2017, pp. 1–9.

[12] C. Guo, Y. Cui, and Z. Liu, "Optimal multicast of tiled 360 VR video in OFDMA systems," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2563–2566, Dec. 2018.

[13] C. Guo, Y. Cui, and Z. Liu, "Optimal multicast of tiled 360 VR video," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 145–148, Feb. 2019.

[14] H. Ahmadi, O. Eltobgy, and M. Hefeeda, "Adaptive multicast streaming of virtual reality content to mobile users," in *Proc. Thematic Workshops ACM MM*, Mountain View, CA, USA, Oct. 2017, pp. 170–178.

[15] W. Xu, Y. Wei, Y. Cui, and Z. Liu, "Energy-efficient multi-view video transmission with view synthesis-enabled multicast," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, 2018, pp. 1–7.

[16] K. Long, C. Ye, Y. Cui, and Z. Liu, "Optimal multi-quality multicast for 360 virtual reality video," in *Proc. IEEE GLOBECOM*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[17] J. Park, J.-N. Hwang, and H.-Y. Wei, "Cross-layer optimization for VR video multicast systems," in *Proc. IEEE GLOBECOM*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[18] N. Jindal and Z. Luo, "Capacity limits of multiple antenna multicast," in *Proc. IEEE ISIT*, Seattle, WA, USA, Jul. 2006, pp. 1841–1845.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

[19] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[20] F. Hou, L. X. Cai, P.-H. Ho, X. Shen, and J. Zhang, "A cooperative multicast scheduling scheme for multimedia services in IEEE 802.16 networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1508–1519, Mar. 2009.

[21] J. Seppälä, T. Kosksela, T. Chen, and S. Hakola, "Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks," in *Proc. IEEE WCNC*, Cancun, Mexico, Mar. 2011, pp. 986–991.

[22] C. Yin, Y. Wang, W. Lin, and J. Xu, "Device-to-device assisted two-stage cooperative multicast with optimal resource utilization," in *Proc. IEEE Globecom Workshops*, Austin, TX, USA, Dec. 2014, pp. 839–844.

[23] F. Rebecchi, L. Valerio, R. Bruno, V. Conan, M. D. de Amorim, and A. Passarella, "A joint multicast/D2D learning-based approach to LTE traffic offloading," *Elsevier Comput. Commun.*, vol. 72, pp. 26–37, Dec. 2015.

[24] G. Araniti *et al.*, "Novel D2D-based relaying method for multicast services over 3GPP LTE-A systems," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, Cagliari, Italy, Mar. 2017, pp. 1–5.

[25] T. V. Santana, R. Combes, and M. Kobayashi, "Device-to-device aided multicasting," in *Proc. IEEE ISIT*, Vail, CO, USA, Jun. 2018, pp. 771–775.

[26] P. Mursia, I. Atzeni, D. Gesbert, and M. Kobayashi, "D2D-aided multi-antenna multicasting," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, 2019, pp. 1–6.

[27] V. Ntranos, N. D. Sidiropoulos, and L. Tassiulas, "On multicast beamforming for minimum outage," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3172–3181, Jun. 2009.

[28] O. Mehanna, N. D. Sidiropoulos, and G. B. Giannakis, "Joint multicast beamforming and antenna selection," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2660–2674, May 2013.

[29] D. Lecompte and F. Gabin, "Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: Overview and Rel-11 enhancements," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 68–74, Nov. 2012.

[30] O. Eltobgy, "VRCast: Mobile streaming of live 360-degree videos," Ph.D. dissertation, Dept. School Comput. Sci., Simon Fraser Univ., Burnaby, BC, Canada, Dec. 2018.

[31] A. Khisti, U. Erez, and G. W. Wornell, "Fundamental limits and scaling behavior of cooperative multicasting in wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2762–2770, Jun. 2006.

[32] B. Sirkeci-Mergen and M. C. Gastpar, "On the broadcast capacity of wireless networks with cooperative relays," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3847–3861, Aug. 2010.

[33] Y. Chen, J. Zhu, T. A. Khan, and B. Kasikci, "CPU microarchitectural performance characterization of cloud video transcoding," in *Proc. IEEE Int. Symp. Workload Characterization (IISWC)*, Beijing, China, 2020, pp. 72–82.

[34] Y.-H. Kim, J. Huh, and J. Jeong, "Distributed video transcoding system for 8K 360° VR tiled streaming service," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Jeju, South Korea, 2018, pp. 592–595.

[35] Y. Ren, F. Liu, Z. Liu, C. Wang, and Y. Ji, "Power control in D2D-based vehicular communication networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5547–5562, Dec. 2015.

[36] W. Huang *et al.*, "Utility-oriented resource allocation for 360-degree video transmission over heterogeneous networks," *Elsevier Digit. Signal Process.*, vol. 84, no. 1, pp. 1–14, Jan. 2019.

[37] V. Sitzmann *et al.*, "Saliency in VR: How do people explore virtual environments?" *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 4, pp. 1633–1642, Apr. 2018.

[38] R. Monroy, S. Lutz, T. Chalasani, and A. Smolic, "SalNet360: Saliency maps for omni-directional images with CNN," *Elsevier Signal Process. Image Commun.*, vol. 69, pp. 26–34, Nov. 2018.

[39] W. Zhang, Y. Wen, Z. Chen, and A. Khisti, "QoE-driven cache management for HTTP adaptive bit rate streaming over wireless networks," *IEEE Trans. Multimedia*, vol. 15, no. 6, pp. 1431–1445, Oct. 2013.

[40] J. Park, J.-N. Hwang, Q. Li, Y. Xu, and W. Huang, "Optimal DASH-multicasting over LTE," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4487–4500, May 2018.

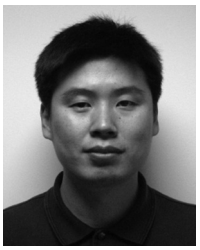
[41] J. Yang, J. Luo, D. Meng, and J.-N. Hwang, "QoE-driven resource allocation optimized for uplink delivery of delay-sensitive VR video over cellular network," *IEEE Access*, vol. 7, pp. 60672–60683, 2019.

[42] D. Peaucelle, D. Henrion, Y. Labit, and K. Taitz. (Sep. 2002). *User's Guide for SeDuMi Interface 1.04*. [Online]. Available: http://homepages.laas.fr/peaucell/software/sdmguide.pdf

[43] A. Awada *et al.*, "Field trial of LTE eMBMS network for TV distribution: Experimental results and analysis," *IEEE Trans. Broadcast.*, vol. 63, no. 2, pp. 321–337, Jun. 2017.

Jianmei Dai (Member, IEEE) received the B.S. degree in communication engineering, the M.S. degree in communication and information systems, and the Ph.D degree in information and communication engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2004, 2007, and 2021, respectively.

He was a Visiting Scholar with Auburn University, Auburn, AL, USA, from April 2019 to April 2020. He is currently a Lecturer with Space Engineering University, Beijing. His research interests include optimization theory and its applications in wireless video transmission and wireless networks.



Guosen Yue (Senior Member, IEEE) received the B.S. degree in physics and the M.S. degree in electrical engineering from Nanjing University, Nanjing, China, in 1994 and 1997, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2004.

He was a Senior Research Staff with NEC Laboratories America, Princeton, NJ, USA, where he conducted research on broadband wireless systems and mobile networks. From 2013 to 2015, he was with Broadcom Corporation, Matawan, NJ, USA, as a System Design Scientist. He is currently a Principal Research Staff with Futurewei Technologies, Bridgewater, NJ, USA. His research interests are in the general areas of wireless communications and signal processing.

Dr. Yue has served as an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and EURASIP Research Letters in Communications, the Guest Editor of *Journal of Wireless Communication and Networking* (EURASIP) special issue on interference management, *Physical Communication* (Elsevier) special issue on signal processing and coding. He served as the Symposium Co-Chair for IEEE GlobeCom 2019 and IEEE ICC 2010, the Track Co-Chair for IEEE ICCCN 2008, and the Steering Committee Member for IEEE RWS 2009. He also served as the Director of IEEE MMTC Communication Frontier.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.

Shiwen Mao (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Polytechnic University, Brooklyn, NY, USA, in 2004.

After joining Auburn University, Auburn, AL, USA, in 2006, he held the McWane Endowed Professorship from 2012 to 2015 and the Samuel Ginn Endowed Professorship from 2015 to 2020 with the Department of Electrical and Computer Engineering. He is currently a Professor, the Earle C. Williams Eminent Scholar Chair, and the Director of the Wireless Engineering Research and Education Center, Auburn University. His research interest includes wireless networks, multimedia communications, and smart grid.

Prof. Mao received the IEEE ComSoc TC-CSR Distinguished Technical Achievement Award in 2019 and the NSF CAREER Award in 2010. He was a co-recipient of the 2021 IEEE Communications Society Outstanding Paper Award, the IEEE Vehicular Technology Society 2020 Jack Neubauer Memorial Award, the IEEE ComSoc MMTC 2018 Best Journal Award and 2017 Best Conference Paper Award, the Best Demo Award of IEEE SECON 2017, the Best Paper Awards from IEEE GLOBECOM 2019, 2016, and 2015, IEEE WCNC 2015, IEEE ICC 2013, and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems. He is an Associate Editor-in-Chief of IEEE/CIC CHINA COMMUNICATIONS, an Area Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, and ACM GetMobile, and an Associate Editor of IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, IEEE TRANSACTIONS ON MOBILE COMPUTING, *IEEE Network Magazine*, IEEE MULTIMEDIA, and IEEE NETWORKING LETTERS. He is a Distinguished Lecturer of IEEE Communications Society and IEEE Council of RFID.

Danpu Liu (Senior Member, IEEE) received the Ph.D. degree in communication and electrical systems from Beijing University of Posts and Telecommunications, Beijing, China, in 1998.

She was a Visiting Scholar with the City University of Hong Kong, Hong Kong, in 2002, University of Manchester, Manchester, U.K., in 2005, and Georgia Institute of Technology, Atlanta, GA, USA, in 2014. She is currently working with Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications. She has published over 100 papers and three teaching books, and submitted 26 patent applications. Her research involved MIMO, OFDM as well as broadband wireless access systems. Her recent research interests include 60-GHz mmWave communication, wireless high definition video transmission, and wireless sensor network.

Authorized licensed use limited to: Auburn University. Downloaded on July 27,2022 at 06:08:14 UTC from IEEE Xplore. Restrictions apply.