This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

Power-Efficient Wireless Streaming of Multi-Quality Tiled 360 VR Video in MIMO-OFDMA Systems

Chengjun Guo, Lingzhi Zhao, Ying Cui, *Member, IEEE*, Zhi Liu, *Senior Member, IEEE*, and Derrick Wing Kwan Ng, *Fellow, IEEE*

***Abstract*—In this paper, we study the optimal wireless streaming of a multi-quality tiled 360 virtual reality (VR) video from a multi-antenna server to multiple single-antenna users in a multiple-input multiple-output (MIMO)-orthogonal frequency division multiple access (OFDMA) system. In the scenario without user transcoding, we jointly optimize beamforming and subcarrier, transmission power, and rate allocation to minimize the total transmission power. This problem is a challenging mixed discrete-continuous optimization problem. We obtain a globally optimal solution for small multicast groups, an asymptotically optimal solution for a large antenna array, and a suboptimal solution for the general case. In the scenario with user transcoding, we jointly optimize the quality level selection, beamforming, and subcarrier, transmission power, and rate allocation to minimize the weighted sum of the average total transmission power and the transcoding power. This problem is a two-timescale mixed discrete-continuous optimization problem, which is even more challenging than the problem for the scenario without user transcoding. We obtain a globally optimal solution for small multicast groups, an asymptotically optimal solution for a large antenna array, and a low-complexity suboptimal solution for the general case. Finally, numerical results demonstrate the significant gains of proposed solutions over the existing solutions.**

***Index Terms*—Wireless streaming, virtual reality, resource allocation, bitrate adaptation, MIMO-OFDMA, transcoding, beamforming, multicast, optimization.**

I. Introduction

A 360 virtual reality (VR) video can be generated by capturing a scene of interest in all directions simultaneously with an omnidirectional camera [2]. It is predicted that the VR market will reach 87.97 billion USD by 2025 [3]. Transmitting a 360 VR video over wireless networks enables users to experience immersive environments without geographical or behavioral restrictions. As a 360 VR video has a much larger file size than a traditional video, streaming an entire 360 VR video brings a heavy burden to wireless networks [4]–[6]. When watching a 360 VR video, a user perceives it from only one viewing direction at any time, which corresponds to one part of the 360 VR video, known as field-of-view (FoV). The tiling technique is widely adopted to improve the transmission efficiency for 360 VR videos [7]. Specifically, a 360 VR video is divided into smaller rectangular segments of the same size, known as tiles. Transmitting the set of tiles covering each predicted FoV can reduce the required communication resources without reducing the quality of experience [2], [7]. In practice, users may have heterogeneous conditions (e.g., channel conditions, display resolutions of devices, etc.). Pre-encoding each tile into multiple representations with different quality levels allows bitrate (quality) adaptation according to a user's condition. Therefore, wireless streaming of a multi-quality tiled 360 VR video has received growing interest.

Recently, [8]–[10] study optimal wireless streaming of a multi-quality tiled 360 VR video in single-user networks. Specifically, [8]–[10] optimize the quality level selection [8], [10] or transmission rate [9] to minimize the total distortion [8], [9], or total utility [10]. The proposed solutions for single-user networks in [8]–[10] are not applicable in multi-user networks, as optimal resource sharing among users with heterogeneous channel conditions is not considered. In several VR applications, such as VR gaming, VR military training, and VR sports [11], [12], a 360 VR video has to be transmitted to multiple users simultaneously. When a tile is required by multiple users concurrently, multicast opportunities can be utilized to improve transmission efficiency. In [13]–[19], the authors study the optimal wireless streaming of a multi-quality tiled 360 VR video to multiple users by exploiting multicast opportunities. In particular, in our previous works [13], [14], we optimize transmission resource allocation to minimize the average transmission power for given video quality requirements of all users and optimize the encoding rate of each tile to maximize the received video quality for a given transmission power budget. In [15]–[19], the authors focus on optimizing quality level selection for each tile. Specifically, in [15]–[17], the authors maximize the total utility of all users [15], [16] or minimize the distortion of video scenes [17], without considering any constraints on quality variation. Consequently, more multicast opportunities can be exploited to further improve transmission efficiency. Nevertheless, the obtained quality levels of adjacent tiles may vary significantly, leading to poor viewing experiences [15]–[17]. In [18], the authors impose some constraints on quality variation while maximizing the total utility of all users to address this issue. Although the restrictions on quality variation in

Manuscript received July 24, 2020; revised December 31, 2020; accepted February 14, 2021. The work of Y. Cui was supported in part by STCSM 18DZ2270700 and in part by the Natural Science Foundation of Shanghai under Grant 20ZR1425300. The work of Z. Liu was supported in part by JSPS KAKENHI under Grants 19H04092, 20H04174, in part by ROIS NII Open Collaborative Research 2020(20FA02). The work of D. W. K. Ng was supported by funding from the UNSW Digital Grid Futures Institute, UNSW, Sydney, under a cross-disciplinary fund scheme and by the Australian Research Council's Discovery Project (DP210102169). This paper will be presented in part at the IEEE ICC 2021 [1]. The associate editor coordinating the review of this paper and approving it for publication was Dongxiao Yu. *(Corresponding author: Ying Cui.)*

C. Guo, L. Zhao and Y. Cui are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: cuiying@sjtu.edu.cn).

Z. Liu is with Graduate School of Informatics and Engineering, the University of Electro-Communications, Tokyo 182-8585, Japan.

D. W. K. Ng is with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
The final version of record is available at http://dx.doi.org/10.1109/TWC.2021.3067803

[18] can alleviate quality variation in an FoV to a certain extent, they cannot guarantee quality smoothness and are less mathematically tractable. In [19], user transcoding is adopted to ensure that the quality levels of all received tiles in each FoV are identical when maximizing the total utility of all users.

Despite the fruitful research in the literature, the performance of wireless streaming of a tiled 360 VR video is still unsatisfactory. The results in [13]–[19] are all for single-antenna servers, which cannot exploit spatial degrees of freedom for effective resource utilization. The performance of wireless systems can be improved by deploying multiple antennas at a server and designing efficient beamformers. Among various multi-antenna technologies, MIMO-OFDMA is the dominant air interface for 5G broadband wireless communications, as it can provide more reliable communications at high speeds. For instance, in [20]–[22], the authors consider single-group multicast [20] and multi-group multicast [21], [22]. Specifically, in [20], the authors consider the optimization of beamforming and power allocation to maximize the minimum user data rate. In [21], the authors consider the subcarrier allocation and power allocation to maximize the sum rate. However, the solutions proposed in [20], [21] are heuristic and hence have no performance guarantee. [22] studies the optimization of beamforming to minimize the total transmission power. A stationary point of the beamforming design problem is obtained based on successive convex approximation. Note that in [22], messages on each subcarrier are associated with different beamforming vectors, resulting in a substantial increase in the number of variables and the computational complexity for solving the optimization problem.

This paper considers the optimal wireless streaming of a multi-quality tiled 360 VR video from one server to multiple users in a MIMO-OFDMA system in the scenarios without and with user transcoding. With more advanced physical layer techniques than those in [13]–[15], [17]–[19], we expect the stringent requirements for 360 VR video transmission to be better satisfied. Assume that the set of tiles to be transmitted to each user has been determined and each user's quality requirement is given. The main contributions of this paper are summarized below.

- In the scenario without user transcoding, we exploit natural multicast opportunities and formulate the minimization of the total transmission power with respect to (w.r.t.) beamforming, subcarrier allocation, transmission power, and rate allocation as a multi-group multicast problem in the MIMO-OFDMA system. This problem is a challenging mixed discrete-continuous optimization problem. We obtain its optimal solutions for two special cases, namely, the case of small multicast groups (for different sets of tiles) and the case of a large antenna array, exploiting decomposition, continuous relaxation, and Karush-Kuhn-Tucker (KKT) conditions. We also obtain a suboptimal solution for the general case using continuous relaxation and difference of convex (DC) programming. Note that previous works studying multi-group multicast in MIMO-OFDMA systems do not investigate special cases in which optimal solutions can be obtained [21], [22]. Besides, the proposed multi-group multicast formulation with one beamforming vector for each subcarrier can achieve the same performance as the multi-group multicast formulation in [22] which has one beamforming vector for each user and each subcarrier, but yields a substantially lower computational complexity for the general case.
- In the scenario with user transcoding, a flexible trade-off between computation and communications resource consumptions can be struck via exploiting transcoding-enabled multicast opportunities. We utilize natural multicast opportunities and transcoding-enabled multicast opportunities and minimize the weighted sum of the average total transmission power and the transcoding power by optimizing the quality level selection, beamforming, and subcarrier, transmission power, and rate allocation. This problem extends the multi-group multicast optimization in the scenario without user transcoding, and it is a more challenging two-timescale mixed optimization problem. For two special cases, we obtain the corresponding optimal solutions. For the general case, we obtain a low-complexity suboptimal solution. Note that the formulations in [21] and [22], which consider only natural multicast opportunities, cannot provide an effective design in the scenario with user transcoding.
- Finally, numerical results show substantial gains achieved by the proposed solutions over existing schemes and demonstrate the advantages of multicast and user transcoding in wireless streaming of a multi-quality tiled 360 VR video. Furthermore, numerical results illustrate that the proposed low-complexity optimal solution obtained for a large antenna array can achieve promising performance when the number of antennas is moderate, demonstrating its effectiveness.

Note that this work extends our previous results on wireless streaming of a multi-quality tiled 360 VR video in time division multiple access (TDMA) systems [13], [18], [19] and OFDMA systems [14]. The extensions are highly nontrivial due to the non-convexity w.r.t beamforming vectors. To the best of our knowledge, this is the first work providing an optimization-based design for wireless streaming of a multi-quality tiled 360 VR video in MIMO-OFDMA systems.

Notation: For a Hermitian matrix $\mathbf{A}$, $\mathbf{A} \succeq \mathbf{0}$ means that $\mathbf{A}$ is an Hermitian positive semidefinite matrix. The symbol $(\cdot)^H$ denotes complex conjugate transpose operator. $\mathrm{tr}(\cdot)$ and $\mathrm{rank}(\cdot)$ denote the trace and the rank, respectively. $\mathbb{E}[\cdot]$ denotes the statistical expectation.

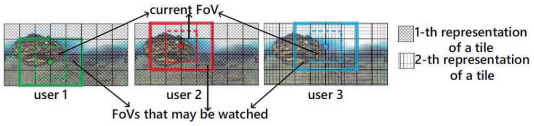
II. System Model

Wireless streaming of a multi-quality tiled 360 VR video from a server (e.g., access point or base station) to multiple users arises in several VR applications such as VR concert, VR military training, and VR sports. This paper aims to optimize the wireless streaming of a multi-quality tiled 360 VR video from a server to K users in a MIMO-OFDMA system as illustrated in Fig. 1.[1] The server is equipped with

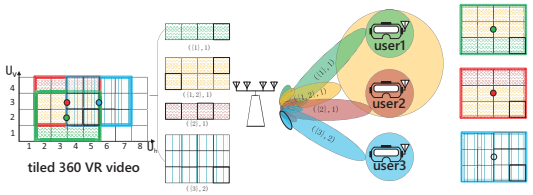
[1]We adopt a multi-quality tiled 360 VR video model which is similar to those in our previous works [13], [14], [18], [19], and the details are presented here for completeness.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

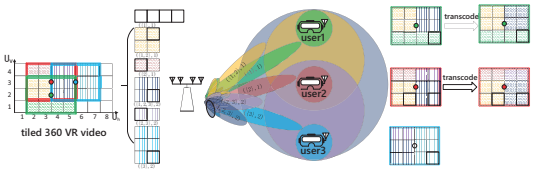
This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications



(a) Multi-quality tiled 360 VR video required by multiple users.



(b) Scenario without user transcoding.



(c) Scenario with user transcoding.

Fig. 1. System models of wireless streaming of a multi-quality tiled 360 VR video in the two scenarios. $K = 3$, $\mathbf{r} = (1, 1, 2)$, $U_h \times U_v = 8 \times 4$, $M = 4$, $\mathcal{I} = \{\{1\}, \{2\}, \{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$, $\mathcal{P}_{\{1\}} = \{(2, 1), (3, 1), (4, 1), (5, 1)\}$, $\mathcal{P}_{\{2\}} = \{(2, 4), (3, 4)\}$, $\mathcal{P}_{\{3\}} = \{(6, 2), (6, 3), (6, 4), (7, 2), (7, 3), (7, 4)\}$, $\mathcal{P}_{\{1,2\}} = \{(2, 2), (2, 3), (3, 2), (3, 3)\}$, $\mathcal{P}_{\{2,3\}} = \{(4, 4), (5, 4)\}$, $\mathcal{P}_{\{1,2,3\}} = \{(4, 2), (4, 3), (5, 2), (5, 3)\}$, $\mathcal{K}_{\{1\},1} = \{1\}$, $\mathcal{K}_{\{1,2\},1} = \{1, 2\}$, $\mathcal{K}_{\{2\},1} = \{2\}$, $\mathcal{K}_{\{3\},2} = \{3\}$.

M transmit antennas and each user wears a single-antenna VR headset. Denote $\mathcal{K} \triangleq \{1, \ldots, K\}$ as the set of user indices. When a VR user is interested in one viewing direction of a 360 VR video, the user watches a rectangular FoV of size $F_h \times F_v$ (in rad×rad), the center of which is referred to as the viewing direction. Besides, a user can freely switch views when watching a 360 VR video.

Tiling is adopted to improve the transmission efficiency of the 360 VR video [7]. In particular, the 360 VR video is divided into $U_h \times U_v$ rectangular segments of the same size, referred to as tiles, where U_h and U_v represent the numbers of segments in each row and column, respectively. Define $\mathcal{U}_h \triangleq \{1, \ldots, U_h\}$ and $\mathcal{U}_v \triangleq \{1, \ldots, U_v\}$. The (u_h, u_v)-th tile corresponds to the tile in the u_h-th row and the u_v-th column, for all $u_h \in \mathcal{U}_h$ and $u_v \in \mathcal{U}_v$. Considering user heterogeneity (e.g., channel conditions, display resolutions of devices, etc.), each tile is pre-encoded into L representations corresponding to L quality levels using High Efficiency Video Coding (HEVC), as in Dynamic Adaptive Streaming over HTTP (DASH). Denote $\mathcal{L} \triangleq \{1, \ldots, L\}$ as the set of quality levels. For all $l \in \mathcal{L}$, the l-th representation of each tile corresponds to the l-th lowest quality. For ease of exposition, we assume that the encoding rates of the tiles with the same quality level are identical. Let D_l (in bits/s) denote the encoding rate of the l-th representation of a tile. Note that $D_1 < D_2 < \ldots < D_L$. We study the system for the duration of the playback time of multiple groups of pictures (GOPs).[2] In this duration, the FoV of each user does not change. Denote $r_k \in \mathcal{L}$ as the quality level for the FoV of user $k \in \mathcal{K}$. Because of the video coding structure, $\mathbf{r} \triangleq (r_k)_{k \in \mathcal{K}}$ should not change during the considered time duration.

As in [13], [14], [18], [19], the server collects a user's information such as head orientation (tracked by a 3DoF or 6DoF VR headset) and location (tracked by a 6DoF VR headset) from the user's headset via the uplink transmission, predicts the user's FoV and determines the set of tiles to be transmitted to the user.[3] Denote $\mathcal{G}_k$ as the set of indices of the tiles that need to be transmitted to user k. Denote $\mathcal{G} \triangleq \bigcup_{k \in \mathcal{K}} \mathcal{G}_k$ as the set of indices of the tiles that need to be sent to all K users. For all $\mathcal{S} \subseteq \mathcal{K}$, $\mathcal{S} \neq \emptyset$, denote $\mathcal{P}_\mathcal{S} \triangleq \left(\bigcap_{k \in \mathcal{S}} \mathcal{G}_k\right) \bigcap \left(\mathcal{G} - \bigcup_{k \in \mathcal{K} \setminus \mathcal{S}} \mathcal{G}_k\right)$ as the set of indices of the tiles that are needed by all users in $\mathcal{S}$ and are not needed by the users in $\mathcal{K} \setminus \mathcal{S}$. Then $\mathcal{P} \triangleq \{\mathcal{P}_\mathcal{S} | \mathcal{P}_\mathcal{S} \neq \emptyset,\ \mathcal{S} \subseteq \mathcal{K},\ \mathcal{S} \neq \emptyset\}$ forms a partition of $\mathcal{G}$ and $\mathcal{I} \triangleq \{\mathcal{S} | \mathcal{P}_\mathcal{S} \neq \emptyset,\ \mathcal{S} \subseteq \mathcal{K},\ \mathcal{S} \neq \emptyset\}$ specifies the user sets corresponding to the partition. Denote $I \triangleq |\mathcal{I}|$. Let $\mathcal{I}_k \triangleq \{\mathcal{S} | \mathcal{S} \subseteq \mathcal{I}, k \in \mathcal{S}\}$, $k \in \mathcal{K}$. The tiles in $\mathcal{P}_\mathcal{S}$, $\mathcal{S} \in \mathcal{I}_k$ are required by user k. We consider the tiles in each set jointly rather than separately, to reduce the complexity for transmission and resource allocation. For all $l \in \mathcal{L}$ and $\mathcal{S} \subseteq \mathcal{K}$, the encoding (source coding) bits of the l-th representations of the tiles in $\mathcal{P}_\mathcal{S}$ are "aggregated" into one message indexed by $(\mathcal{S}, l)$, which is transmitted at most once to the users in $\mathcal{S}$ that will utilize it, to improve transmission efficiency. For all $\mathcal{S} \in \mathcal{I}$ and $l \in \mathcal{L}$, let $\mathcal{K}_{\mathcal{S},l} \triangleq \{k \in \mathcal{S} | r_k = l\}$. If there is only one user in $\mathcal{K}_{\mathcal{S},l}$, the transmission of message $(\mathcal{S}, l)$ corresponds to unicast; and if there are multiple users in $\mathcal{K}_{\mathcal{S},l}$, the transmission of message $(\mathcal{S}, l)$ corresponds to multicast. Thus, the multi-quality tiled 360 VR video transmission to the K users may involve both unicast and multicast. An illustration example can be seen in Fig. 1 (b).

Let $\mathcal{N} \triangleq \{1, \ldots, N\}$, where N is the number of subcarriers. The bandwidth of each subcarrier is B (in Hz). We assume block fading, i.e., the small-scale channel fading coefficients do not change within one frame. Let $\mathbf{H}_{n,k} \in \mathbb{C}^{M \times 1}$ denote the random small-scale fading coefficient between the server and user k on subcarrier n. Denote $\mathbf{H} \triangleq (\mathbf{H}_{n,k})_{n \in \mathcal{N}, k \in \mathcal{K}}$. Let $\mathbf{h} \triangleq (\mathbf{h}_{n,k})_{n \in \mathcal{N}, k \in \mathcal{K}}$ represent a realization of $\mathbf{H}$ (which can be obtained by the server via channel estimation), where $\mathbf{h}_{n,k} \in \mathbb{C}^{M \times 1}$ is a realization of $\mathbf{H}_{n,k}$. Let $\beta_k > 0$ denote the large-scale channel fading gain between the server and user k, which remains constant during the considered time duration and is known to the server.

Denote $\mu_{\mathcal{S},l,n}(\mathbf{h}) \in \{0, 1\}$ as the subcarrier assignment indicator for subcarrier n and message $(\mathcal{S}, l)$ under $\mathbf{h}$, where $\mu_{\mathcal{S},l,n}(\mathbf{h}) = 1$ indicates that subcarrier n is assigned to transmit the symbols for message $(\mathcal{S}, l)$, $\mu_{\mathcal{S},l,n}(\mathbf{h}) = 0$ otherwise. For ease of implementation, assume that each subcarrier is assigned to transmit symbols of only one message in each frame [23], [24]. Thus, subcarrier allocation constraints are

[2]The duration of the playback time of one GOP is usually 0.06-1 seconds.

[3]A widely adopted mechanism for dealing with possible prediction errors is to transmit the tiles in the predicted FoV plus a safe margin [18], [19]. A more significant prediction error yields a larger safe margin, leading to more transmission resource consumption. Note that the proposed framework does not rely on any particular prediction method or transmission mechanism and only focuses on the optimal design of transmitting the scheduled tiles.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

given by

$$\mu_{\mathcal{S},l,n}(\mathbf{h}) \in \{0,1\},\ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, n \in \mathcal{N}, \quad (1)$$

$$\sum\nolimits_{\mathcal{S}\in\mathcal{I}} \sum\nolimits_{l\in\mathcal{L}} \mu_{\mathcal{S},l,n}(\mathbf{h}) = 1,\ n \in \mathcal{N}. \quad (2)$$

To capture the scaling of the transmission power with M for studying the optimal power allocation at large M, let $\frac{\eta_{\mathcal{S},l,n}(\mathbf{h})}{M}$ denote the transmission power for the symbols for message $(\mathcal{S},l)$ on subcarrier n under $\mathbf{h}$, where

$$\eta_{\mathcal{S},l,n}(\mathbf{h}) \geq 0,\ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, n \in \mathcal{N}. \quad (3)$$

The total transmission power is $\sum\limits_{n\in\mathcal{N}} \sum\limits_{\mathcal{S}\in\mathcal{I}} \sum\limits_{l\in\mathcal{L}} \frac{\mu_{\mathcal{S},l,n}(\mathbf{h})\eta_{\mathcal{S},l,n}(\mathbf{h})}{M}$.

Suppose that subcarrier n is assigned to transmit the symbols for message $(\mathcal{S},l)$. Let $s_{\mathcal{S},l,n}$ represent the symbols for message $(\mathcal{S},l)$ transmitted on subcarrier n. Assume $\mathbb{E}[|s_{\mathcal{S},l,n}|^2] = 1$ for all $l \in \mathcal{L}, \mathcal{S} \in \mathcal{I}$ and $n \in \mathcal{N}$. Let $\mathbf{w}_n(\mathbf{h}) \in \mathbb{C}^{M\times 1}$ denote the beamforming vector for the message transmitted on subcarrier n under $\mathbf{h}$, where

$$\|\mathbf{w}_n(\mathbf{h})\|_2 = 1,\ n \in \mathcal{N}. \quad (4)$$

The received signal at user k on subcarrier n is given by

$$y_{\mathcal{S},l,k,n} = \sqrt{\beta_k \frac{\eta_{\mathcal{S},l,n}(\mathbf{h})}{M}} \mathbf{h}_{n,k}^H \mathbf{w}_n(\mathbf{h}) s_{\mathcal{S},l,n} + z_{n,k},\ k \in \mathcal{K}, n \in \mathcal{N},$$

where $z_{n,k} \sim \mathcal{CN}(0,\sigma^2)$ represents the noise at user k on subcarrier n. To obtain design insights, we consider the application of a capacity achieving code [23], [25], [26]. The maximum transmission rate for the symbols for message $(\mathcal{S},l)$ to user $k \in \mathcal{S}$ on subcarrier n under $\mathbf{h}$ is given by $B\log_2\left(1+\frac{\eta_{\mathcal{S},l,n}(\mathbf{h})\beta_k|\mathbf{h}_{n,k}^H\mathbf{w}_n(\mathbf{h})|^2}{M\sigma^2}\right)$ (in bit/s). Let $c_{\mathcal{S},l,n}(\mathbf{h})$ denote the transmission rate for the symbols for message $(\mathcal{S},l)$ on subcarrier n under $\mathbf{h}$, where

$$c_{\mathcal{S},l,n}(\mathbf{h}) \geq 0,\ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, n \in \mathcal{N}. \quad (5)$$

Then, $\sum_{n\in\mathcal{N}} c_{\mathcal{S},l,n}$ represents the transmission rate of message $(\mathcal{S},l)$.

In Section III, we will consider the scenario where users cannot perform transcoding but directly play the received messages. In Section IV, we will consider the scenario where users can first perform transcoding, i.e., convert a tile representation at a certain quality level to a lower quality level using transcoding tools such as Fast Forward Mpeg (FFmpeg), and then play the received video.

III. Total Transmission Power Minimization Without User Transcoding

In this section, we consider the scenario without user transcoding. For all $\mathcal{S} \in \mathcal{I}$, let $\mathcal{L}_\mathcal{S} \triangleq \{r_k | k \in \mathcal{S}\}$. When $|\mathcal{K}_{\mathcal{S},l}| > 1$, message $(\mathcal{S},l)$, where $\mathcal{S} \in \mathcal{I}$ and $l \in \mathcal{L}_\mathcal{S}$ can be transmitted to all users in $\mathcal{K}_{\mathcal{S},l}$ simultaneously via multicast. This type of multicast opportunities are referred to as natural multicast opportunities [13], [14], [18], [19]. An illustration example is shown in Fig. 1 (b). In this example, by using natural multicast opportunities, the server multicasts message $(\{1,2\},1)$ to user 1 and user 2.

Consider one frame with small-scale fading coefficients $\mathbf{h}$. To guarantee that message $(\mathcal{S},l)$ can be successfully sent to each user $k \in \mathcal{K}_{\mathcal{S},l}$ on subcarrier n under $\mathbf{h}$, we have

$$\mu_{\mathcal{S},l,n}(\mathbf{h}) B\log_2\left(1+\frac{\eta_{\mathcal{S},l,n}(\mathbf{h})\beta_k|\mathbf{h}_{n,k}^H\mathbf{w}_n(\mathbf{h})|^2}{M\sigma^2}\right) \geq c_{\mathcal{S},l,n}(\mathbf{h}), \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}, k \in \mathcal{K}_{\mathcal{S},l}, n \in \mathcal{N}. \quad (6)$$

To maximally avoid rebuffering and reduce startup delay, we require that the transmission rate of each message $(\mathcal{S},l)$ is no smaller than its encoding rate

$$\sum\nolimits_{n\in\mathcal{N}} c_{\mathcal{S},l,n}(\mathbf{h}) \geq |\mathcal{P}_\mathcal{S}| D_l,\ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}, \quad (7)$$

where $|\mathcal{P}_\mathcal{S}|$ denotes the number of tiles in $\mathcal{P}_\mathcal{S}$.

For convenience, denote $\boldsymbol{\mu}(\mathbf{h}) \triangleq (\mu_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S},n\in\mathcal{N}}$, $\boldsymbol{\eta}(\mathbf{h}) \triangleq (\eta_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S},n\in\mathcal{N}}$, $\mathbf{c}(\mathbf{h}) \triangleq (c_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S},n\in\mathcal{N}}$ and $\mathbf{w}(\mathbf{h}) \triangleq (\mathbf{w}_n(\mathbf{h}))_{n\in\mathcal{N}}$. We consider $\boldsymbol{\mu}(\mathbf{h}), \boldsymbol{\eta}(\mathbf{h}), \mathbf{c}(\mathbf{h}), \mathbf{w}(\mathbf{h})$ as functions of $\mathbf{h}$, respectively. Given the quality levels of all users $\mathbf{r}$, we would like to optimize $\boldsymbol{\mu}(\mathbf{h})$, $\boldsymbol{\eta}(\mathbf{h})$, $\mathbf{c}(\mathbf{h})$, and $\mathbf{w}(\mathbf{h})$ to minimize the average total transmission power

$$\mathbb{E}\left[\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_\mathcal{S}} \frac{\mu_{\mathcal{S},l,n}(\mathbf{H})\eta_{\mathcal{S},l,n}(\mathbf{H})}{M}\right],$$

where the expectation is taken over $\mathbf{H}$, subject to the constraints in (1)-(4) and (5)-(7). This is a variational problem due to the calculus of variation in the objective function. Note that for each $\mathbf{h}$, the number of optimization variables is the same. Also, the optimization variables and constraints are separated for all $\mathbf{h}$. Consequently, it is equivalent to optimize $\boldsymbol{\mu}(\mathbf{h})$, $\boldsymbol{\eta}(\mathbf{h})$, $\mathbf{c}(\mathbf{h})$, and $\mathbf{w}(\mathbf{h})$ to minimize $\frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_\mathcal{S}} \mu_{\mathcal{S},l,n}(\mathbf{h})\eta_{\mathcal{S},l,n}(\mathbf{h})$ subject to the constraints in (1)-(4), and (5)-(7), for all $\mathbf{h}$. Thus, we consider the following problem.

Problem 1 (Total Transmission Power Minimization for $\mathbf{h}$):

$$E^\star(\mathbf{h}) \triangleq \min_{\boldsymbol{\mu}(\mathbf{h}),\boldsymbol{\eta}(\mathbf{h}),\mathbf{c}(\mathbf{h}),\mathbf{w}(\mathbf{h})} \sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_\mathcal{S}} \frac{\mu_{\mathcal{S},l,n}(\mathbf{h})\eta_{\mathcal{S},l,n}(\mathbf{h})}{M}$$
$$\text{s.t.}\ \ (1),\ (2),\ (3),\ (4),\ (5),\ (6),\ (7).$$

It can be observed that Problem 1 is a challenging mixed discrete-continuous optimization problem. In Section III-A, we first obtain a globally optimal solution for small multicast groups and an asymptotically optimal solution for a large antenna array. Then, in Section III-B, we obtain a suboptimal solution for the general case.[4]

A. Solutions in Special Cases

In this subsection, we solve Problem 1 in two special cases, by solving the following equivalent problem of Problem 1.

Problem 2 (Equivalent Problem of Problem 1 for $\mathbf{h}$):

$$E^\dagger(\mathbf{h}) \triangleq \min_{\boldsymbol{\mu}(\mathbf{h}),\mathbf{P}(\mathbf{h})} \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_\mathcal{S}} P_{\mathcal{S},l,n}(\mathbf{h})$$
$$\text{s.t.}\ \ (1),\ (2),$$
$$P_{\mathcal{S},l,n}(\mathbf{h}) \geq 0,\ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}, n \in \mathcal{N}, \quad (8)$$
$$\sum_{n\in\mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}) B\log_2\left(1+\frac{P_{\mathcal{S},l,n}(\mathbf{h})}{\mu_{\mathcal{S},l,n}(\mathbf{h})Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}\right) \geq |\mathcal{P}_\mathcal{S}| D_l,\ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}, k \in \mathcal{K}_{\mathcal{S},l}, \quad (9)$$

[4]Note that the goal of solving a nonconvex problem is usually to design an iterative algorithm to obtain a stationary point or a KKT point. In general, it is impossible to analytically or numerically show the gap between a globally optimal solution and a suboptimal solution, as a globally optimal solution cannot be obtained analytically or numerically with effective and efficient methods [27].

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

where $\mathbf{P}(\mathbf{h}) \triangleq (P_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S},n\in\mathcal{N}}$ and $Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})$ is given by the following problem. Let $(\boldsymbol{\mu}^\dagger(\mathbf{h}), \mathbf{P}^\dagger(\mathbf{h}))$ denote an optimal solution of Problem 2.

Problem 3 (Subproblem of Problem 2 for $\mathbf{h}$):

$$Q^\dagger_{\mathcal{S},l,n}(\mathbf{h}) \triangleq \min_{\mathbf{V}_{\mathcal{S},l,n}\in\mathbb{C}^{M\times M}} \operatorname{tr}(\mathbf{V}_{\mathcal{S},l,n})$$
$$\text{s.t.} \quad \frac{\operatorname{tr}(\beta_k \mathbf{h}_{k,n}\mathbf{h}_{k,n}^H \mathbf{V}_{\mathcal{S},l,n})}{M\sigma^2} \geq 1, \ k \in \mathcal{K}_{\mathcal{S},l}, \tag{10}$$
$$\mathbf{V}_{\mathcal{S},l,n} \succeq \mathbf{0},$$
$$\operatorname{rank}(\mathbf{V}_{\mathcal{S},l,n}) = 1. \tag{11}$$

Let $\mathbf{V}^\dagger_{\mathcal{S},l,n}(\mathbf{h})$ denote an optimal solution of Problem 3, which can be written as $\mathbf{V}^\dagger_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{v}^\dagger_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^\dagger_{\mathcal{S},l,n}(\mathbf{h}))^H$ for some $\mathbf{v}^\dagger_{\mathcal{S},l,n}(\mathbf{h}) \in \mathbb{C}^{M\times 1}$.

By exploring structures of Problem 1, Problem 2, and Problem 3, we have the following result.

Theorem 1 (Equivalence between Problem 1 and Problem 2): The optimal values of Problem 1 and Problem 2 are identical. In addition, $(\boldsymbol{\mu}^\dagger(\mathbf{h}), \boldsymbol{\eta}^\dagger(\mathbf{h}), \mathbf{c}^\dagger(\mathbf{h}), \mathbf{w}^\dagger(\mathbf{h}))$ is an optimal solution of Problem 1, where

$$\boldsymbol{\eta}^\dagger(\mathbf{h}) = \mathbf{P}^\dagger(\mathbf{h}), \tag{12}$$

$$\mathbf{w}^\dagger_n(\mathbf{h}) = \sum_{\mathcal{S}\in\mathcal{I}} \sum_{l\in\mathcal{L}_\mathcal{S}} \mu^\dagger_{\mathcal{S},l,n}(\mathbf{h}) \frac{\mathbf{v}^\dagger_{\mathcal{S},l,n}(\mathbf{h})}{\sqrt{Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}}, \ n \in \mathcal{N}, \tag{13}$$

and $\mathbf{c}^\dagger(\mathbf{h}) \triangleq (c^\dagger_{\mathcal{S},l,n}(\mathbf{h}))_{\mathbf{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S},n\in\mathcal{N}}$ with

$$c^\dagger_{\mathcal{S},l,n}(\mathbf{h}) = \mu^\dagger_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{P^\dagger_{\mathcal{S},l,n}(\mathbf{h})}{Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}\right). \tag{14}$$

Proof 1: See Appendix A.

According to Theorem 1, to obtain an optimal solution of Problem 1, we can first get $\mathbf{w}^\dagger(\mathbf{h})$ by solving Problem 3, and then get $\boldsymbol{\mu}^\dagger(\mathbf{h})$, $\boldsymbol{\eta}^\dagger(\mathbf{h})$, and $\mathbf{c}^\dagger(\mathbf{h})$ by solving Problem 2. Notice that Problem 3 is nonconvex due to the rank-one constraint in (11), while Problem 2 is a nonconvex problem because of the binary constraints in (1). Both problems are pretty challenging. In the following, we solve Problem 3 and Problem 2 for two special cases.

1) Case of Small Multicast Groups: In this part, we consider the case where for all $\mathcal{S} \in \mathcal{I}$ and $l \in \mathcal{L}_\mathcal{S}$, there are at most three users who need message $(\mathcal{S}, l)$ (i.e., $|\mathcal{K}_{\mathcal{S},l}| \leq 3$). First, we obtain an optimal solution of Problem 3 by applying semidefinite relaxation and rank reduction proposed in [28]. Specifically, we relax the constraint in (11), and get an SDP, which is convex and can be solved effectively. Under the condition that $|\mathcal{K}_{\mathcal{S},l}| \leq 3$, a rank-one solution can be constructed based on an optimal solution of the SDP using rank reduction [28]. Then, substituting $Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})$, $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$ into Problem 2 and relaxing the constraints in (1) to

$$\mu_{\mathcal{S},l,n}(\mathbf{h}) \geq 0, \quad \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}, n \in \mathcal{N}, \tag{15}$$

we obtain a relaxed problem of Problem 2, which is convex. Using the KKT conditions, we know that under a mild condition, there exists an optimal solution of the relaxed problem of Problem 2 which provides binary subcarrier assignment [14]. For all $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$ and $n \in \mathcal{N}$, define:

$$f_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) \triangleq \left[\frac{B\lambda_{\mathcal{S},l}}{\ln 2} - Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})\right]^+, \tag{16}$$

$$\mathcal{W}_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) \triangleq \lambda_{\mathcal{S},l} B \left(\log_2\left(1 + \frac{f_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l})}{Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}\right) - \frac{f_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l})}{\left(Q^\dagger_{\mathcal{S},l,n}(\mathbf{h}) + f_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l})\right)\ln 2}\right), \tag{17}$$

$$\mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) \triangleq \begin{cases} 1, \ (\mathcal{S}, l) = \underset{\mathcal{S}'\in\mathcal{I},l'\in\mathcal{L}}{\operatorname{argmax}} \mathcal{W}_{\mathcal{S}',l',n}(\mathbf{h}, \lambda_{\mathcal{S}',l'}), \\ 0, \ \text{otherwise}, \end{cases} \tag{18}$$

$$P_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) \triangleq \mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) f_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}). \tag{19}$$

Let $\lambda^\dagger_{\mathcal{S},l}(\mathbf{h})$ denote a root of

$$\sum_{n\in\mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) B \log_2\left(1 + \frac{P_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l})}{\mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda_{\mathcal{S},l}) Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}\right) = |\mathcal{P}_\mathcal{S}| D_l, \ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}. \tag{20}$$

An optimal solution of Problem 2 is given below [14].

Claim 1 (Optimal Solution of Problem 2 for $\mathbf{h}$): Suppose that for all $n \in \mathcal{N}$, there exists a unique pair $(\mathcal{S}_n, l_n)$ such that $\mathcal{W}_{\mathcal{S}_n,l_n,n}(\mathbf{h}, \lambda^\dagger_{\mathcal{S}_n,l_n}(\mathbf{h})) = \max_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S}} \mathcal{W}_{\mathcal{S},l,n}(\mathbf{h}, \lambda^\dagger_{\mathcal{S},l}(\mathbf{h}))$. Then, an optimal solution of Problem 2 is given by $\mu^\dagger_{\mathcal{S},l,n}(\mathbf{h}) = \mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda^\dagger_{\mathcal{S},l}(\mathbf{h}))$, $P^\dagger_{\mathcal{S},l,n}(\mathbf{h}) = P_{\mathcal{S},l,n}(\mathbf{h}, \lambda^\dagger_{\mathcal{S},l}(\mathbf{h}))$, $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$ and $n \in \mathcal{N}$.

Note that $\mathcal{W}_{\mathcal{S},l,n}(\mathbf{h}, \lambda^\dagger_{\mathcal{S},l}(\mathbf{h}))$ monotonically increases with $Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})$ and $Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})$, $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$ are different for user groups $\mathcal{K}_{\mathcal{S},l}$, $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$ (as $Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})$ captures both large-scale fading and small-scale fading effects). Thus, $\mathcal{W}_{\mathcal{S},l,n}(\mathbf{h}, \lambda^\dagger_{\mathcal{S},l}(\mathbf{h}))$, $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$ are usually different, and the condition in Claim 1 can be easily satisfied [14]. Note that $\lambda^\dagger_{\mathcal{S},l}(\mathbf{h})$ can be obtained using a subgradient method as in [18].

The details for obtaining an optimal solution of Problem 1 via solving Problem 3 and Problem 2 are summarized in Algorithm 1. Specifically, in Steps 1-13, we solve Problem 3 for all $\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_\mathcal{S}$, and $n \in \mathcal{N}$ with computational complexity $O(M^6NKI)$; in Steps 14-20, we solve Problem 2 with computational complexity $O(NLKI^2)$; in Steps 21-25, we compute the optimal solution of Problem 1 based on the solutions of Problem 2 and Problem 3 with computational complexity $O(NKI)$. Therefore, the computational complexity of Algorithm 1 is $O(NLKI^2)$.

2) Case of a Large Antenna Array: In this part, we consider the case where the server is equipped with a large antenna array. For the sake of presentation, in this part, we explicitly write the optimal value of Problem 3 as a function of M, i.e., $Q^{\dagger(M)}_{\mathcal{S},l,n}(\mathbf{h})$. Following the proofs for Theorem 1 and Theorem 3 in [29], we can show the following result.

Theorem 2 (Asymptotically Optimal Solution of Problem 3): For all $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_\mathcal{S}$, $n \in \mathcal{N}$ and $\mathbf{h}$, $\mathbf{V}^*_{\mathcal{S},l,n}(\mathbf{h}) =$

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
The final version of record is available at http://dx.doi.org/10.1109/TWC.2021.3067803

Algorithm 1 Globally Optimal Solution of Problem 1 for Case of Small Multicast Groups

for $\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}$ and $n \in \mathcal{N}$ **do**
  Find an optimal solution $\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})$ (with arbitrary ranks) of Problem 3 without the rank-one constraint in (11);
  while $\text{rank}(\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})) > 1$ **do**
    Set $\psi = \text{rank}(\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))$;
    Decompose $\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{U}\mathbf{U}^H$;
    Find a nonzero solution $\mathbf{\Delta}$ of the system of linear equations: $\text{tr}(\mathbf{U}^H \beta_k \mathbf{h}_{k,n}\mathbf{h}^H_{k,n}\mathbf{U}\mathbf{\Delta}) = 0$, $k \in \mathcal{K}_{\mathcal{S},l}$, where $\mathbf{\Delta}$ is a $\psi \times \psi$ Hermitian matrix;
    Evaluate the eigenvalues $\delta_1, \ldots, \delta_\psi$ of $\mathbf{\Delta}$;
    Determine i_0 such that $|\delta_{i_0}| = \max\{|\delta_i| : 1 \le i \le \psi\}$;
    Compute $\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{U}(\mathbf{I}_\psi - (1/\delta_{i_0})\mathbf{\Delta})\mathbf{U}^H$;
  end while
  Compute $Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \text{tr}(\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))$;
  Decompose $\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))^H$;
end for
Initialize $\boldsymbol{\lambda}^{(0)}$. Set iteration index $t = 0$;
repeat
  For all $\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}$ and $n \in \mathcal{N}$, compute $W_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l})$ according to (17);
  For all $\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}$ and $n \in \mathcal{N}$, compute $\mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l})$ and $P_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l})$ according to (18) and (19), respectively;
  For all $\mathcal{S} \in \mathcal{I}$ and $l \in \mathcal{L}_{\mathcal{S}}$, compute $\lambda^{(t+1)}_{\mathcal{S},l}$ according to:

$$\lambda^{(t+1)}_{\mathcal{S},l} = \left[\lambda^{(t)}_{\mathcal{S},l} - \delta^{(t)}\left(\sum_{n\in\mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l}) B \log_2 (1+ \frac{P_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l})}{\mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l}) Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}\right) - |\mathcal{P}_{\mathcal{S}}| D_l\right)\right]^+,$$

  where $\delta^{(t)}$, $t = 1, 2, \ldots$ satisfy

$$\delta^{(t)} > 0, \ \sum_{t=0}^{\infty}(\delta^{(t)})^2 < \infty, \ \sum_{t=0}^{\infty}\delta^{(t)} = \infty, \ \lim_{t\to\infty}\delta^{(t)} = 0; \quad (21)$$

  Set $t = t + 1$;
until convergence criteria is met
for $n \in \mathcal{N}$ **do**
  Compute $\mathbf{w}^{\dagger}_n(\mathbf{h})$ according to (13);
end for
For all $\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}$ and $n \in \mathcal{N}$, set $\mu^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mu_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l})$ and $P^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = P_{\mathcal{S},l,n}(\mathbf{h}, \lambda^{(t)}_{\mathcal{S},l})$;
Set $\boldsymbol{\eta}^{\dagger}(\mathbf{h}) = \mathbf{P}^{\dagger}(\mathbf{h})$, and compute $\mathbf{c}^{\dagger}(\mathbf{h})$ according to (13).

$\mathbf{v}^*_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^*_{\mathcal{S},l,n}(\mathbf{h}))^H$ is an asymptotically optimal solution of Problem 3 at large M, where

$$\mathbf{v}^*_{\mathcal{S},l,n}(\mathbf{h}) = \frac{\sum\limits_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{\mathbf{h}_{n,k}}{\sqrt{\beta_k}}}{\left\|\sum\limits_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{\mathbf{h}_{n,k}}{\sqrt{\beta_k}}\right\|_2} \sqrt{\frac{M\sigma^2}{\min\limits_{k\in\mathcal{K}_{\mathcal{S},l}} \beta_k \frac{\left|\sum\limits_{j\in\mathcal{K}_{\mathcal{S},l}} \frac{\mathbf{h}^H_{n,k}\mathbf{h}_{n,j}}{\sqrt{\beta_j}}\right|^2}{\left\|\sum\limits_{j\in\mathcal{K}_{\mathcal{S},l}} \frac{\mathbf{h}_{n,j}}{\sqrt{\beta_j}}\right\|_2^2}}}. \quad (22)$$

Proof 2: See Appendix B.

Substituting $Q^{\dagger(M)}_{\mathcal{S},l,n}(\mathbf{h}) = \text{tr}(\mathbf{V}^*_{\mathcal{S},l,n}(\mathbf{h}))$ into Problem 2 and using the same method as in Section III-A1 for solving Problem 2, an asymptotically optimal solution of Problem 1 (which can achieve competitive performance at large M) can be obtained.

B. Suboptimal Solution in General Case

In the general case, i.e., there exists $(\mathcal{S}, l)$ such that $|\mathcal{K}_{\mathcal{S},l}| > 3$ and the number of antennas equipped at the server is not large, we cannot obtain a globally optimal solution of the nonconvex problem in Problem 3. Thus, we cannot solve Problem 1 by solving its equivalent form in Problem 2. In this subsection, we directly tackle Problem 1, and develop a low-complexity algorithm to obtain a suboptimal solution of Problem 1 using relaxation and DC programming.

First, by relaxing the constraints in (1) of Problem 1 to the constraints in (15), we can obtain the relaxed continuous problem of Problem 1. Next, we convert the relaxed continuous problem of Problem 1 to DC programming. Let $\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h}) \triangleq \sqrt{\eta_{\mathcal{S},l,n}(\mathbf{h})\mu_{\mathcal{S},l,n}(\mathbf{h})}\mathbf{w}_n(\mathbf{h})$. Thus, the constraints in (3), (4), and (6) can be equivalently transformed to the following ones.

$$\mu_{\mathcal{S},l,n}(\mathbf{h})\left(2^{\frac{c_{\mathcal{S},l,n}(\mathbf{h})}{B\mu_{\mathcal{S},l,n}(\mathbf{h})}} - 1\right) - \frac{\beta_k |\mathbf{h}^H_{n,k}\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h})|^2}{M\sigma^2} \le 0,$$
$$\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}, k \in \mathcal{K}_{\mathcal{S},l}, n \in \mathcal{N}. \quad (23)$$

Thus, the relaxed continuous problem of Problem 1 is given as follows.

Problem 4 (DC Problem of Relaxed Problem 1 for $\mathbf{h}$*):*

$$\min_{\mathbf{W}(\mathbf{h}),\boldsymbol{\mu}(\mathbf{h}),\mathbf{c}(\mathbf{h})} \ \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} \|\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h})\|_2^2$$
$$\text{s.t.} \ \ (2), \ (5), \ (7), \ (15), \ (23).$$

Note that the objective function of Problem 4 and the constraints in (2), (5), (7), and (15) are all convex. Besides, each of the constraints in (23) can be regarded as a difference of two convex functions, i.e., $\mu_{\mathcal{S},l,n}(\mathbf{h})\left(2^{\frac{c_{\mathcal{S},l,n}(\mathbf{h})}{B\mu_{\mathcal{S},l,n}(\mathbf{h})}} - 1\right)$ and $\frac{\beta_k |\mathbf{h}^H_{n,k}\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h})|^2}{M\sigma^2}$. Thus, Problem 4 is a standard DC programming and can be handled by using the DC algorithm [30]. The core idea is to solve a sequence of convex approximations of Problem 4 iteratively, each of which is obtained by linearizing the concave function, i.e., $-\frac{\beta_k |\mathbf{h}^H_{n,k}\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h})|^2}{M\sigma^2}$ in (23). Specifically, at the t-th iteration, the convex approximation of Problem 4 is given below.

Problem 5 (Convex Approximation of Problem 4 for $\mathbf{h}$ *at* t*-th Iteration):*

$$E^{(t)}(\mathbf{h}) \triangleq \min_{\mathbf{W}(\mathbf{h}),\boldsymbol{\mu}(\mathbf{h}),\mathbf{c}(\mathbf{h})} \ \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}} \|\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h})\|^2$$
$$\text{s.t.} \ \ (2), \ (5), \ (7), \ (15), \ (24),$$

where (24) is shown at the top of the next page. Let $(\mathbf{W}^{(t)}(\mathbf{h}), \boldsymbol{\mu}^{(t)}(\mathbf{h}), \mathbf{c}^{(t)}(\mathbf{h}))$ denote an optimal solution at the t-th iteration.

Problem 5 is a convex problem. We can use standard convex optimization techniques to solve it. According to [30], for any initial point which is a feasible solution of Problem 4, as $t \to \infty$, $(\mathbf{W}^{(t)}(\mathbf{h}), \boldsymbol{\mu}^{(t)}(\mathbf{h}), \mathbf{c}^{(t)}(\mathbf{h})) \to (\mathbf{W}^{(\infty)}(\mathbf{h}), \boldsymbol{\mu}^{(\infty)}(\mathbf{h}), \mathbf{c}^{(\infty)}(\mathbf{h}))$, which is a stationary point of the relaxed Problem 1, and $E^{(t)}(\mathbf{h}) \to E^{(\infty)}(\mathbf{h})$. Note that $\boldsymbol{\mu}^{(\infty)}(\mathbf{h})$ may not be binary, and hence $(\mathbf{W}^{(\infty)}(\mathbf{h}), \boldsymbol{\mu}^{(\infty)}(\mathbf{h}), \mathbf{c}^{(\infty)}(\mathbf{h}))$ may not be a feasible solution of Problem 1. By using the KKT conditions, we can obtain an optimal solution of Problem 5 for the $t^\diamond$-the iteration where $t^\diamond$ satisfies some convergence criteria. It provides binary

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
The final version of record is available at http://dx.doi.org/10.1109/TWC.2021.3067803

$$\mu_{\mathcal{S},l,n}(\mathbf{h})\left(2^{\frac{c_{\mathcal{S},l,n}(\mathbf{h})}{B\mu_{\mathcal{S},l,n}(\mathbf{h})}}-1\right)-\frac{2\beta_k R\left\{(\mathbf{W}_{\mathcal{S},l,n}^{(t-1)}(\mathbf{h}))^H\mathbf{h}_{n,k}\mathbf{h}_{n,k}^H\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h})\right\}}{M\sigma^2}+\frac{|\mathbf{h}_{n,k}^H\mathbf{W}_{\mathcal{S},l,n}^{(t-1)}(\mathbf{h})|^2}{M\sigma^2}\le 0,\ \mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}},k\in\mathcal{K}_{\mathcal{S},l},n\in\mathcal{N}. \tag{24}$$

subcarrier assignment under a mild condition, and can be treated as a suboptimal solution of Problem 1.

Let $\boldsymbol{\lambda}_{\mathcal{S},l,n}\triangleq(\lambda_{\mathcal{S},l,n,k})_{k\in\mathcal{K}_{\mathcal{S},l}}$. For all $\mathcal{S}\in\mathcal{I}$, $l\in\mathcal{L}_{\mathcal{S}}$ and $n\in\mathcal{N}$, define:

$$G_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})\triangleq\gamma_{\mathcal{S},l}\log_2\frac{\gamma_{\mathcal{S},l}}{\ln 2\sum_{k\in\mathcal{S}}\lambda_{\mathcal{S},l,n,k}}-\frac{\gamma_{\mathcal{S},l}B}{\ln 2}+\sum_{k\in\mathcal{S}}\lambda_{\mathcal{S},l,n,k}, \tag{25}$$

$$\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})=\begin{cases}1,\ (\mathcal{S},l)=\underset{\mathcal{S}'\in\mathcal{I},l'\in\mathcal{L}_{\mathcal{S}}}{\operatorname{argmax}}\ G_{\mathcal{S}',l',n}(\mathbf{h},\gamma_{\mathcal{S}',l'},\boldsymbol{\lambda}_{\mathcal{S}',l',n}),\\0,\ \text{otherwise},\end{cases} \tag{26}$$

$$c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})=\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})B\left[\log_2\frac{\gamma_{\mathcal{S},l}}{\ln 2\sum_{k\in\mathcal{S}}\lambda_{\mathcal{S},l,n,k}}\right]^+, \tag{27}$$

$$\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})=\frac{\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})\mathbf{A}_{\mathcal{S},l,n}^{(t-1)}\sum\limits_{k\in\mathcal{S}}\lambda_{\mathcal{S},l,n,k}\beta_k|\mathbf{h}_{n,k}^H\mathbf{W}_{\mathcal{S},l,n}^{(t-1)}(\mathbf{h})|^2}{\|\mathbf{A}_{\mathcal{S},l,n}^{(t-1)}\|_2^2}. \tag{28}$$

where

$$\mathbf{A}_{\mathcal{S},l,n}^{(t-1)}=\sum\nolimits_{k\in\mathcal{S}}\lambda_{\mathcal{S},l,n,k}\beta_k(\mathbf{W}_{\mathcal{S},l,n}^{(t-1)}(\mathbf{h}))^H\mathbf{h}_{n,k}\mathbf{h}_{n,k}^H. \tag{29}$$

Let $\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h})$ and $\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})$ denote the roots of

$$\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})\left(2^{\frac{c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})}{B\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})}}-1\right)-\frac{2\beta_k R\left\{(\mathbf{W}_{\mathcal{S},l,n}^{(t^\diamond)}(\mathbf{h}))^H\mathbf{h}_{n,k}\mathbf{h}_{n,k}^H\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})\right\}}{M\sigma^2}$$
$$+\frac{|\mathbf{h}_{n,k}^H\mathbf{W}_{\mathcal{S},l,n}^{(t^\diamond)}(\mathbf{h})|^2}{M\sigma^2}=0,\ \mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}},k\in\mathcal{K}_{\mathcal{S},l},n\in\mathcal{N},$$
$$\sum\nolimits_{n\in\mathcal{N}}c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l},\boldsymbol{\lambda}_{\mathcal{S},l,n})=|\mathcal{P}_{\mathcal{S}}|D_l,\ \mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}}.$$

An optimal solution of Problem 5 for the $t^\diamond$-th iteration which provides binary subcarrier assignment is given below.

Claim 2 (Optimal Solution of Problem 5 for $t^\diamond$): Suppose that there exists a unique pair $(\mathcal{S}_n,l_n)$ such that

$$G_{\mathcal{S}_n,l_n,n}(\mathbf{h},\gamma_{\mathcal{S}_n,l_n}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S}_n,l_n,n}^{\diamond}(\mathbf{h}))$$
$$=\max\nolimits_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}}}G_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})),n\in\mathcal{N},$$

Then, an optimal solution of Problem 5 for $t^\diamond$ is given by $\mathbf{W}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})=\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))$, $\mu_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})=\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))$ and $c_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})=c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))$.

Similar to the condition stated in Claim 1, the condition in Claim 2 can be easily satisfied. Note that $\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h})$ and $\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})$ can be obtained using a subgradient method. The details for obtaining a suboptimal solution $(\boldsymbol{\mu}^\diamond(\mathbf{h}),\boldsymbol{\eta}^\diamond(\mathbf{h}),\mathbf{c}^\diamond(\mathbf{h}),\mathbf{w}^\diamond(\mathbf{h}))$ of Problem 1 are summarized in Algorithm 2. Specifically, in Steps 1-5, we solve Problem 4 with computational complexity $O(M^3K^4N^{3.5}I^{3.5})$; in Steps 6-13, we obtain an optimal solution of Problem 5 based on the optimal solution of Problem 4 with computational complexity $O(M^5NK^2I)$; in Steps 14-15, we compute a suboptimal solution of Problem 1 based on the optimal solution of Problem 5 with computational complexity $O(MNKI)$. Therefore, the computational complexity of Algorithm 2 is $O(M^3K^4N^{3.5}I^{3.5})$.

Algorithm 2 Suboptimal Solution of Problem 1 for the General Case

1: Find a random feasible point of Problem 4 as the initial point $(\mathbf{W}^{(0)}(\mathbf{h}),\boldsymbol{\mu}^{(0)}(\mathbf{h}),\mathbf{c}^{(0)}(\mathbf{h}))$, and set $t=0$;
2: **repeat**
3: Set $t=t+1$;
4: Obtain $(\mathbf{W}^{(t)}(\mathbf{h}),\boldsymbol{\mu}^{(t)}(\mathbf{h}),\mathbf{c}^{(t)}(\mathbf{h}))$ by solving Problem 5 using standard convex optimization techniques;
5: **until** convergence criteria are met
6: Set $t^\diamond=t$;
7: Initialize $\boldsymbol{\gamma}^{(1)}$ and $\boldsymbol{\lambda}^{(1)}$, and set $i=0$;
8: **repeat**
9: Set $i=i+1$;
10: For all $\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}}$ and $n\in\mathcal{N}$, compute $G_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{(i)},\boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})$, $\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{(i)},\boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})$, $c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{(i)},\boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})$ and $\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{(i)},\boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})$ according to (25), (26), (27) and (28), respectively;
11: For all $\mathcal{S}\in\mathcal{I}$, $l\in\mathcal{L}_{\mathcal{S}}$, $n\in\mathcal{N}$ and $k\in\mathcal{K}_{\mathcal{S},l}$, compute $\lambda_{\mathcal{S},l,n,k}^{(i+1)}$ according to (30), where (30) is shown at the top of the next page, and $\delta^{(i)}$, $i=1,2,\ldots$ satisfy (21);
12: For all $\mathcal{S}\in\mathcal{I}$ and $l\in\mathcal{L}_{\mathcal{S}}$, compute $\gamma_{\mathcal{S},l}^{(i+1)}$ according to:
$$\gamma_{\mathcal{S},l}^{(i+1)}=\left[\gamma_{\mathcal{S},l}^{(i)}-\delta^{(i)}\left(\sum_{n\in\mathcal{N}}c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{(i)},\boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})-|\mathcal{P}_{\mathcal{S}}|D_l\right)\right]^+,$$
where $\delta^{(i)}$, $i=1,2,\ldots$ satisfy (21);
13: **until** convergence criteria are met
14: Set $\boldsymbol{\gamma}^\diamond(\mathbf{h})=\boldsymbol{\gamma}^{(i)}$ and $\boldsymbol{\lambda}^\diamond(\mathbf{h})=\boldsymbol{\lambda}^{(i)}$;
15: For all $\mathcal{S}\in\mathcal{I}$, $l\in\mathcal{L}_{\mathcal{S}}$ and $n\in\mathcal{N}$, set $\mu_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})=\mu_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))$, $\eta_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})=\|\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))\|_2$, $c_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})=c_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))$ and $\mathbf{w}_n^{\diamond}(\mathbf{h})=\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}}\mu_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})\frac{\mathbf{W}_{\mathcal{S},l,n}(\mathbf{h},\gamma_{\mathcal{S},l}^{\diamond}(\mathbf{h}),\boldsymbol{\lambda}_{\mathcal{S},l,n}^{\diamond}(\mathbf{h}))}{\eta_{\mathcal{S},l,n}^{\diamond}(\mathbf{h})}$.

IV. Average Transmission Power Minimization With User Transcoding

In this section, we consider the case with user transcoding. Although message $(\mathcal{S},l)$, where $\mathcal{S}\in\mathcal{I}$ and $l\in\mathcal{L}$, is requested only by the users in $\mathcal{K}_{\mathcal{S},l}$, it may be transmitted to all users in $\mathcal{K}_{\mathcal{S},l}$ and $\mathcal{K}_{\mathcal{S},l'}$ for some $l'<l$ simultaneously via multicast. The users in $\mathcal{K}_{\mathcal{S},l}$ directly use message $(\mathcal{S},l)$. In contrast, the users in $\mathcal{K}_{\mathcal{S},l'}$, $l'<l$ first convert message $(\mathcal{S},l)$ to message $(\mathcal{S},l')$ using transcoding, before using it. This type of multicast opportunities are referred to as transcoding-enabled multicast opportunities [19]. An illustration example is shown in Fig. 1 (c). In this example, by making use of natural multicast opportunities, the server multicasts message $(\{1,2\},1)$ to user 1 and user 2; by making use of transcoding-enabled multicast opportunities, the server multicasts message $(\{2,3\},2)$ to user 2 and user 3 and multicasts message $(\{1,2,3\},2)$ to user 1, user 2 and user 3. By comparing Fig. 1 (b) and Fig. 1 (c), we can see that transcoding provides more multicast opportunities.

To model user transcoding [31], let $\mathbf{x}\triangleq(x_{\mathcal{S},l,k})_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L},k\in\mathcal{S}}$ denote the quality level selection variables, where

$$x_{\mathcal{S},l,k}\in\{0,1\},\ \mathcal{S}\in\mathcal{I},l\in\mathcal{L},k\in\mathcal{S}, \tag{31}$$

$$\sum\nolimits_{l\in\mathcal{L}}x_{\mathcal{S},l,k}=1,\ \mathcal{S}\in\mathcal{I},k\in\mathcal{S}. \tag{32}$$

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

The final version of record is available at http://dx.doi.org/10.1109/TWC.2021.3067803

$$\lambda_{\mathcal{S},l,n,k}^{(i+1)} = \left[\lambda_{\mathcal{S},l,n,k}^{(i)} - \delta^{(i)} \left(\mu_{\mathcal{S},l,n}(\mathbf{h}, \gamma_{\mathcal{S},l}^{(i)}, \boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)}) \left(2^{\frac{c_{\mathcal{S},l,n}(\mathbf{h}, \gamma_{\mathcal{S},l}^{(i)}, \boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})}{B \mu_{\mathcal{S},l,n}(\mathbf{h}, \gamma_{\mathcal{S},l}^{(i)}, \boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)})}} - 1 \right) - \frac{2\beta_k R \left\{ (\mathbf{W}_{\mathcal{S},l,n}^{(t^\diamond)}(\mathbf{h}))^H \mathbf{h}_{n,k} \mathbf{h}_{n,k}^H \mathbf{W}_{\mathcal{S},l,n}(\mathbf{h}, \gamma_{\mathcal{S},l}^{(i)}, \boldsymbol{\lambda}_{\mathcal{S},l,n}^{(i)}) \right\}}{M \sigma^2} + \frac{|\mathbf{h}_{n,k}^H \mathbf{W}_{\mathcal{S},l,n}^{(t^\diamond)}(\mathbf{h})|^2}{M \sigma^2} \right) \right]^+, \tag{30}$$

Here, $x_{\mathcal{S},l,k} = 1$ indicates that the server will transmit message $(\mathcal{S}, l)$ to user k, and $x_{\mathcal{S},l,k} = 0$ otherwise. Note that the constraints in (32) ensure that the server transmits only one of the messages $(\mathcal{S}, l)$, $l \in \mathcal{L}$ which has quality level $\sum_{l' \in \mathcal{L}} l' x_{\mathcal{S},l',k}$ to user k. Note that $\mathbf{x}$ should not change during the considered time duration because of the video coding structure. With transcoding, to ensure that user k can play his FoV at quality level r_k, it is sufficient to require:

$$\sum\nolimits_{l \in \mathcal{L}} l x_{\mathcal{S},l,k} \geq r_k, \ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, k \in \mathcal{S}. \tag{33}$$

Then, the successful transmission constraints in (6) become:

$$\mu_{\mathcal{S},l,n}(\mathbf{h}) B \log_2 \left(1 + \frac{\eta_{\mathcal{S},l,n}(\mathbf{h}) \beta_k |\mathbf{h}_{n,k}^H \mathbf{w}_n(\mathbf{h})|^2}{M \sigma^2} \right) \geq c_{\mathcal{S},l,n}(\mathbf{h}) x_{\mathcal{S},l,k}, \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, k \in \mathcal{S}, n \in \mathcal{N}, \tag{34}$$

and the transmission rate constraints in (7) become:

$$\sum_{n \in \mathcal{N}} c_{\mathcal{S},l,n}(\mathbf{h}) \geq |\mathcal{P}_{\mathcal{S}}| D_l x_{\mathcal{S},l,k}, \ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, k \in \mathcal{S}. \tag{35}$$

On the other hand, user transcoding also consumes power. For ease of exposition, we assume that the transcoding powers for reducing the quality levels of all tiles by one are identical at each user. Denote E_k as the transcoding power at user k for lowering the quality level of the representation of a tile by one. Since different users have heterogeneous hardware conditions, we allow E_k, $k \in \mathcal{K}$ to be different. Then, the total transcoding power at all users is $E_{\text{tc}}(\mathbf{x}) \triangleq \sum_{\mathcal{S} \in \mathcal{I}} \sum_{k \in \mathcal{S}} \left(\sum_{l \in \mathcal{L}} l x_{\mathcal{S},l,k} - r_k \right) |\mathcal{P}_{\mathcal{S}}| E_k$. The weighted sum of the average transmission power and the transcoding power is

$$\mathbb{E}\left[\frac{1}{M} \sum_{n \in \mathcal{N}} \sum_{\mathcal{S} \in \mathcal{I}} \sum_{l \in \mathcal{L}} \mu_{\mathcal{S},l,n}(\mathbf{H}) \eta_{\mathcal{S},l,n}(\mathbf{H}) \right] + \alpha E_{\text{tc}}(\mathbf{x}),$$

where $\alpha \geq 1$ is the corresponding weight factor, and the expectation is taken over $\mathbf{H}$. Note that $\alpha > 1$ means that a higher cost on the power consumption for user devices is incurred due to the limited battery powers of user devices.

With slight abuse of notation, denote $\boldsymbol{\mu}(\mathbf{h}) \triangleq (\mu_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, n \in \mathcal{N}}$, $\boldsymbol{\eta}(\mathbf{h}) \triangleq (\eta_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, n \in \mathcal{N}}$, $\mathbf{c}(\mathbf{h}) \triangleq (c_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S} \in \mathcal{I}, l \in \mathcal{L}, n \in \mathcal{N}}$ and $\mathbf{w}(\mathbf{h}) \triangleq (\mathbf{w}_n(\mathbf{h}))_{n \in \mathcal{N}}$. Similarly, we treat $\boldsymbol{\mu}(\mathbf{h}), \boldsymbol{\eta}(\mathbf{h}), \mathbf{c}(\mathbf{h}), \mathbf{w}(\mathbf{h})$ as functions of $\mathbf{h}$, respectively. For given quality requirements of all users $\mathbf{r}$, we would like to minimize the weighted sum of the average transmission power and the transcoding power under the constraints in (1)-(4), (5), and (31)-(35), by optimizing $\boldsymbol{\mu}(\mathbf{h})$, $\boldsymbol{\eta}(\mathbf{h})$, $\mathbf{c}(\mathbf{h})$, $\mathbf{w}(\mathbf{h})$, and $\mathbf{x}$. Specifically, for given $\mathbf{r}$, we have

Problem 6 (Average Total Transmission Power and Transcoding Power Minimization):

$$\min_{\boldsymbol{\mu}(\mathbf{h}), \boldsymbol{\eta}(\mathbf{h}), \mathbf{c}(\mathbf{h}), \mathbf{w}(\mathbf{h}), \mathbf{x}} \ \mathbb{E}\left[\frac{1}{M} \sum_{n \in \mathcal{N}} \sum_{\mathcal{S} \in \mathcal{I}} \sum_{l \in \mathcal{L}} \mu_{\mathcal{S},l,n}(\mathbf{H}) \eta_{\mathcal{S},l,n}(\mathbf{H}) \right] + \alpha E_{\text{tc}}(\mathbf{x})$$

$$\text{s.t.} \ \ (1), \ (2), \ (3), \ (4), \ (5), \ (31), \ (32), \ (33), \ (34), \ (35).$$

Problem 6 is a variational problem. Besides, it is a two-timescale mixed optimization problem, and is more challenging than Problem 1.[5] Specifically, quality level selection is in a larger timescale and adapts to the channel distribution; subcarrier, power, and rate allocation and beamforming design are in a shorter timescale and are adaptive to instantaneous channel states.[6] Problem 6 generalizes multi-group multicast problems in MIMO-OFDMA systems because it allows optimizing multicast groups to a certain extent via exploiting transcoding-enabled multicast opportunities. In the following, Problem 6 is solved using two methods. The first method provides optimal solutions for some special cases, and the second method offers a suboptimal solution for the general case.

A. Solutions for Special Cases

It can be easily veryfied that an optimal solution satisfies

$$\sum\nolimits_{l \in \mathcal{L}} l x_{\mathcal{S},l,k} \in \mathcal{L}_{\mathcal{S}}, \ \mathcal{S} \in \mathcal{I}, \ l \in \mathcal{L}, \ k \in \mathcal{S}. \tag{36}$$

Thus, we can impose the extra constraints in (36) without loss of optimality. In two special cases, we solve Problem 6 by solving the following equivalent problem of Problem 6.

Problem 7 (Equivalent Problem of Problem 6):

$$\min\nolimits_{\mathbf{x} \in \boldsymbol{\mathcal{X}}} \ \mathbb{E}\left[E^\star(\mathbf{x}, \mathbf{H}) \right] + \alpha E_{\text{tc}}(\mathbf{x})$$

where $\boldsymbol{\mathcal{X}} \triangleq \{\mathbf{x} | \mathbf{x} \text{ satisfies (31), (32), (33), (36)}\}$, and $E^\star(\mathbf{x}, \mathbf{h})$ is given by the following problem.

Problem 8 (Subproblem of Problem 7 for $\mathbf{x} \in \boldsymbol{\mathcal{X}}$ and $\mathbf{h}$):

$$E^\star(\mathbf{x}, \mathbf{h}) \triangleq \min_{\mathbf{w}(\mathbf{h}), \boldsymbol{\mu}(\mathbf{h}), \boldsymbol{\eta}(\mathbf{h}), \mathbf{c}(\mathbf{h})} \ \sum_{n \in \mathcal{N}} \sum_{\mathcal{S} \in \mathcal{I}} \sum_{l \in \mathcal{L}} \frac{\mu_{\mathcal{S},l,n}(\mathbf{h}) \eta_{\mathcal{S},l,n}(\mathbf{h})}{M}$$

$$\text{s.t.} \ \ (1), \ (2), \ (3), \ (4), \ (5), \ (34), \ (35).$$

Problem 8 has the same structure as Problem 1 and can be equivalently converted to the following problem.

Problem 9 (Equivalent Problem of Problem 8 for $\mathbf{x} \in \boldsymbol{\mathcal{X}}$ and $\mathbf{h}$):

$$\min_{\boldsymbol{\mu}(\mathbf{h}), \mathbf{P}(\mathbf{h})} \ \frac{1}{M} \sum\nolimits_{n \in \mathcal{N}} \sum\nolimits_{\mathcal{S} \in \mathcal{I}} \sum\nolimits_{l \in \mathcal{L}_{\mathcal{S}}} P_{\mathcal{S},l,n}(\mathbf{h})$$

$$\text{s.t.} \ \ (1), \ (2), \ (8),$$

$$\sum\nolimits_{n \in \mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}) B \log_2 \left(1 + \frac{P_{\mathcal{S},l,n}(\mathbf{h})}{\mu_{\mathcal{S},l,n}(\mathbf{h}) Q_{\mathcal{S},l,n}^\dagger(\mathbf{h})} \right) \geq |\mathcal{P}_{\mathcal{S}}| D_l x_{\mathcal{S},l,k}, \ \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}, k \in \mathcal{K}_{\mathcal{S},l}, \tag{37}$$

where $Q_{\mathcal{S},l,n}^\dagger(\mathbf{h})$ is given by Problem 3. Let $(\hat{\boldsymbol{\mu}}^\dagger(\mathbf{h}), \hat{\mathbf{P}}^\dagger(\mathbf{h}))$ denote an optimal solution of Problem 9.

Analogously, by exploring structures of Problem 8, Problem 3, and Problem 9, we have the following result.

[5]Similarly, it is in general impossible to analytically or numerically show the gap between a globally optimal solution and a suboptimal solution [27].

[6]The optimal quality level selection can be used until $\mathcal{G}_k, k \in \mathcal{K}$ change. For any given $\mathcal{G}_k, k \in \mathcal{K}$, we only need to solve Problem 6 once and then solve Problem 8 (which is similar to Problem 1) for each subsequent frame.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

Theorem 3 (Equivalence between Problem 8 and Problem 9): The optimal values of Problem 8 and Problem 9 are identical. In addition, $(\hat{\boldsymbol{\mu}}^\dagger(\mathbf{h}), \hat{\boldsymbol{\eta}}^\dagger(\mathbf{h}), \hat{\mathbf{c}}^\dagger(\mathbf{h}), \hat{\mathbf{w}}^\dagger(\mathbf{h}))$ is an optimal solution of Problem 8, where $\hat{\boldsymbol{\eta}}^\dagger(\mathbf{h}) = \hat{\mathbf{P}}^\dagger(\mathbf{h})$, $\hat{\mathbf{w}}_n^\dagger(\mathbf{h}) = \sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_\mathcal{S}} \hat{\mu}^\dagger_{\mathcal{S},l,n}(\mathbf{h})\frac{\mathbf{v}^\dagger_{\mathcal{S},l,n}(\mathbf{h})}{\sqrt{Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}}$, and $\hat{\mathbf{c}}^\dagger(\mathbf{h}) \triangleq (\hat{c}^\dagger_{\mathcal{S},l,n}(\mathbf{h}))_{\mathbf{S}\in\mathcal{I},l\in\mathcal{L}_\mathcal{S},n\in\mathcal{N}}$ with $\hat{c}^\dagger_{\mathcal{S},l,n}(\mathbf{h}) = \hat{\mu}^\dagger_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1+\frac{\hat{P}^\dagger_{\mathcal{S},l,n}(\mathbf{h})}{Q^\dagger_{\mathcal{S},l,n}(\mathbf{h})}\right)$.

According to Theorem 3, to obtain an optimal solution of Problem 8, we can first obtain $\hat{\mathbf{w}}^\dagger(\mathbf{h})$ by solving Problem 3, and then obtain $\hat{\boldsymbol{\mu}}^\dagger(\mathbf{h})$, $\hat{\boldsymbol{\eta}}^\dagger(\mathbf{h})$, and $\hat{\mathbf{c}}^\dagger(\mathbf{h})$ by solving Problem 9. In the case where each group $\mathcal{S}\in\mathcal{I}$ has at most 3 users, i.e., $|\mathcal{S}|\le 3$, $\mathcal{S}\in\mathcal{I}$, we can obtain an optimal solution of Problem 9 by using Algorithm 1 in Section III-A1. In the case of a large antenna array, we can get an asymptotically optimal solution of Problem 9 using the method in Section III-A2. After obtaining $E^\star(\mathbf{x},\mathbf{h})$ for all $\mathbf{x}\in\boldsymbol{\mathcal{X}}$ and $\mathbf{h}$, we can numerically compute $\mathbb{E}[E^\star(\mathbf{x},\mathbf{H})]$, for all $\mathbf{x}\in\boldsymbol{\mathcal{X}}$, and then solve Problem 7 using the exhaustive search. The exhaustive search is over $\prod_{\mathcal{S}\in\mathcal{I}} L_\mathcal{S}$ possible values of $\mathbf{x}$, i.e., the computational complexity scales with $\prod_{\mathcal{S}\in\mathcal{I}} L_\mathcal{S}$, where $L_\mathcal{S}\triangleq|\mathcal{L}_\mathcal{S}|$.[7]

B. Suboptimal Solution in General Case

In the general case, we directly tackle Problem 6 and develop a low-complexity algorithm to obtain a suboptimal solution. Specifically, we get a suboptimal quality level selection by solving an approximation of Problem 6 using DC programming and obtain a suboptimal solution of Problem 8 with the obtained $\mathbf{x}$ using Algorithm 2 in Section III-B.

First, we get an approximation of Problem 6, which has only one timescale and has a much smaller number of variables than Problem 6 [32]. Specifically, (35) and (34) are replaced by

$$\sum\nolimits_{n\in\mathcal{N}} \bar{c}_{\mathcal{S},l,n} \ge |\mathcal{P}_\mathcal{S}| D_l x_{\mathcal{S},l,k},\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, \tag{38}$$

$$\bar{\mu}_{\mathcal{S},l,n} B\log_2\left(1+\frac{\bar{\eta}_{\mathcal{S},l,n}}{\bar{Q}_k}\right) \ge \bar{c}_{\mathcal{S},l,n} x_{\mathcal{S},l,k},$$
$$\mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, n\in\mathcal{N}, \tag{39}$$

respectively. Here, $\bar{\mu}_{\mathcal{S},l,n}$, $\bar{\eta}_{\mathcal{S},l,n}$ and $\bar{c}_{\mathcal{S},l,n}$ approximately characterize $\mathbb{E}[\mu_{\mathcal{S},l,n}(\mathbf{H})]$, $\mathbb{E}[\eta_{\mathcal{S},l,n}(\mathbf{H})]$ and $\mathbb{E}[c_{\mathcal{S},l,n}(\mathbf{H})]$, and

$$\bar{Q}_k \triangleq \mathbb{E}\left[\min_{\mathbf{w}_n(\mathbf{H})\in\{\mathbf{w}|\mathbf{w}\in\mathbb{C}^{M\times 1},\|\mathbf{w}\|_2=1\}} \frac{M\sigma^2}{\beta_k|\mathbf{h}_{n,k}^H\mathbf{w}_n(\mathbf{H})|^2}\right]$$
$$= \frac{M\sigma^2}{\beta_k\mathbb{E}\left[\max_{\mathbf{w}_n(\mathbf{H})\in\{\mathbf{w}|\mathbf{w}\in\mathbb{C}^{M\times 1},\|\mathbf{w}\|_2=1\}}|\mathbf{h}_{n,k}^H\mathbf{w}_n(\mathbf{H})|^2\right]} = \frac{\sigma^2}{\beta_k}.$$

Then, by introducing $\bar{p}_{\mathcal{S},l,n}\triangleq\bar{\mu}_{\mathcal{S},l,n}\bar{\eta}_{\mathcal{S},l,n}$ and eliminating $\bar{\mathbf{c}}$ as well as $\bar{\boldsymbol{\eta}}$, we simplify the constraints in (38) and (39) to

$$\sum\nolimits_{n\in\mathcal{N}} \bar{\mu}_{\mathcal{S},l,n} B\log_2\left(1+\frac{\bar{p}_{\mathcal{S},l,n}}{\bar{Q}_k\bar{\mu}_{\mathcal{S},l,n}}\right) \ge |\mathcal{P}_\mathcal{S}| D_l x_{\mathcal{S},l,k},$$
$$\mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, n\in\mathcal{N}. \tag{40}$$

Therefore, we can obtain the following problem.

[7]When $\prod_{\mathcal{S}\in\mathcal{I}} L_\mathcal{S}$ is large, one can adopt Algorithm 2 to obtain a suboptimal solution with relatively low computational complexity.

Problem 10 (Approximation of Problem 6):

$$E_t^\dagger \triangleq \min_{\bar{\boldsymbol{\mu}},\bar{\mathbf{p}},\mathbf{x}} \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}} \bar{p}_{\mathcal{S},l,n} + \alpha E_{\mathrm{tc}}(\mathbf{x})$$
$$\text{s.t. } (31),\ (32),\ (33),\ (40),$$
$$\bar{\mu}_{\mathcal{S},l,n}\ge 0,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, n\in\mathcal{N}, \tag{41}$$
$$\bar{p}_{\mathcal{S},l,n}\ge 0,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, n\in\mathcal{N}, \tag{42}$$
$$\sum\nolimits_{\mathcal{S}\in\mathcal{I}}\sum\nolimits_{l\in\mathcal{L}}\bar{\mu}_{\mathcal{S},l,n}=1, n\in\mathcal{N}, \tag{43}$$

where $\bar{\boldsymbol{\mu}}\triangleq(\bar{\mu}_{\mathcal{S},l,n})_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L},n\in\mathcal{N}}$ and $\bar{\boldsymbol{p}}\triangleq(\bar{\eta}_{\mathcal{S},l,n})_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L},n\in\mathcal{N}}$. Let $(\bar{\boldsymbol{\mu}}^\dagger,\bar{\boldsymbol{p}}^\dagger,\mathbf{x}^\dagger)$ denote an optimal solution.

Problem 10 is a single timescale mixed discrete-continuous problem with $LI(2N+K)$ variables, much simpler than Problem 6. The dimensions of $\bar{\boldsymbol{\mu}}$ and $\bar{\boldsymbol{p}}$ are both LNI. To further reduce the computational complexity of Problem 10, we convert it to an equivalent problem.

Problem 11 (Equivalent Problem of Problem 10):

$$\bar{E}_t^* \triangleq \min_{\bar{\mathbf{N}},\bar{\mathbf{P}},\mathbf{x}} \frac{1}{M}\sum\nolimits_{\mathcal{S}\in\mathcal{I}}\sum\nolimits_{l\in\mathcal{L}}\bar{P}_{\mathcal{S},l} + \alpha E_{\mathrm{tc}}(\mathbf{x})$$
$$\text{s.t. } (31),\ (32),\ (33),$$
$$\bar{N}_{\mathcal{S},l}\ge 0,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, \tag{44}$$
$$\bar{P}_{\mathcal{S},l}\ge 0,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, \tag{45}$$
$$\sum\nolimits_{\mathcal{S}\in\mathcal{I}}\sum\nolimits_{l\in\mathcal{L}}\bar{N}_{\mathcal{S},l}=N, \tag{46}$$
$$\bar{N}_{\mathcal{S},l}B\log_2\left(1+\frac{\bar{P}_{\mathcal{S},l}}{\bar{N}_{\mathcal{S},l}\bar{Q}_k}\right)\ge|\mathcal{P}_\mathcal{S}|D_l x_{\mathcal{S},l,k},$$
$$\mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, n\in\mathcal{N}, \tag{47}$$

where $\bar{\mathbf{N}}\triangleq(\bar{N}_{\mathcal{S},l})_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}}$ and $\bar{\mathbf{P}}\triangleq(\bar{P}_{\mathcal{S},l})_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}}$. Let $(\bar{\mathbf{N}}^*,\bar{\mathbf{P}}^*,\mathbf{x}^*)$ denote an optimal solution.

Theorem 4 (Equivalence Between Problem 10 and Problem 11): There exist an optimal solution of Problem 10 (i.e., $(\bar{\boldsymbol{\mu}}^\dagger,\bar{\boldsymbol{p}}^\dagger,\mathbf{x}^\dagger)$) and an optimal solution of Problem 11 (i.e., $(\bar{\mathbf{N}}^*,\bar{\mathbf{P}}^*,\mathbf{x}^*)$) such that $\mathbf{x}^*=\mathbf{x}^\dagger$, and $\bar{N}^*_{\mathcal{S},l}=\sum_{n\in\mathcal{N}}\bar{\mu}^\dagger_{\mathcal{S},l,n}$, $\bar{P}^*_{\mathcal{S},l}=\sum_{n\in\mathcal{N}}\bar{p}^\dagger_{\mathcal{S},l,n}$, $\mathcal{S}\in\mathcal{I}$, $l\in\mathcal{L}$.

Proof 3: See Appendix C.

By noting that the dimensions of $\bar{\mathbf{N}}$ and $\bar{\mathbf{P}}$ are both LI, i,e, $\frac{1}{N}$ of those of $\bar{\boldsymbol{\mu}}$ and $\bar{\boldsymbol{\eta}}$, Problem 11 is much simpler than Problem 10. In the following, a low-complexity algorithm is developed to obtain a suboptimal solution of Problem 11 using the DC algorithm [30].

First, we convert Problem 11 to a penalized DC problem. Specifically, we equivalently convert the discrete constraints in (31) to the following continuous constraints:

$$0\le x_{\mathcal{S},l,k}\le 1,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, \tag{48}$$
$$x_{\mathcal{S},l,k}(1-x_{\mathcal{S},l,k})\le 0,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}. \tag{49}$$

By augmenting the constraints in (49) to the objective function via the penalty method [18], Problem 11 can be equivalently converted to the following problem.

Problem 12 (Penalized DC Problem of Problem 11):

$$\min_{\bar{\mathbf{N}},\bar{\mathbf{P}},\mathbf{x}} \sum\nolimits_{\mathcal{S}\in\mathcal{I}}\sum\nolimits_{l\in\mathcal{L}}\bar{P}_{\mathcal{S},l} + \alpha E_{\mathrm{tc}}(\mathbf{x}) + \rho\chi(\mathbf{x})$$
$$\text{s.t. } (44),\ (45),\ (46),\ (47),\ (33),\ (32),$$
$$0\le x_{\mathcal{S},l,k}\le 1,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, \tag{50}$$

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

where the penalty parameter $\rho > 0$ and the penalty function $\chi(\mathbf{x}) \triangleq \sum_{\mathcal{S}\in\mathcal{I}} \sum_{k\in\mathcal{S}} \sum_{l\in\mathcal{L}} x_{\mathcal{S},l,k}(1 - x_{\mathcal{S},l,k})$.

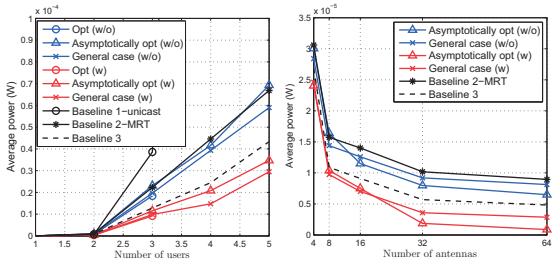
Note that we can regard the objective function of Problem 12 as a difference of two convex functions and the constraints of Problem 12 are all convex. Thus, we can view Problem 12 as a penalized DC problem of Problem 11. When the feasible set of Problem 11 is nonempty, there exists $\rho_0 > 0$ such that for all $\rho > \rho_0$, Problem 12 and Problem 11 are equivalent [30]. By solving Problem 12 using a DC algorithm [30], we obtain a stationary point $(\bar{\mathbf{P}}^\diamond, \bar{\mathbf{N}}^\diamond, \mathbf{x}^\diamond)$ of Problem 11.

Next, by substituting $\mathbf{x}^\diamond$ into Problem 8, we can get a suboptimal solution of Problem 8 for each $\mathbf{h}$, denoted by $(\boldsymbol{\mu}^\diamond(\mathbf{h}), \boldsymbol{\eta}^\diamond(\mathbf{h}), \mathbf{c}^\diamond(\mathbf{h}), \mathbf{w}^\diamond(\mathbf{h}))$, using Algorithm 2 in Section III-B.

V. Numerical Results

This section considers the two scenarios without (w/o) and with (w) user transcoding, and compares the proposed solutions in Section III and Section IV with baseline schemes. In the scenario without user transcoding, we consider the following two baseline schemes. Baseline 1 serves K users separately (i.e., adopts unicast) and adopts the normalized maximum ratio transmission (MRT) beamformer for each user on each subcarrier. Baseline 2 jointly considers the FoVs of all users (i.e., adopts multicast for a message, if there exists a multicast opportunity) as in this paper and adopts the normalized MRT beamformer for a massage on each subcarrier obtained based on the channel power gain matrix of all users requiring this message on each subcarrier [33]. Then, for each baseline scheme, the optimal subcarrier, power, and rate allocation is obtained by solving Problem 2 for the respective MRT using the method proposed in Section III-A. In the scenario with user transcoding, we consider one baseline scheme, i.e., Baseline 3, which transmits message $(\mathcal{S}, r_{\mathcal{S},\max})$ to all users in $\mathcal{S}$ using multicast, where $r_{\mathcal{S},\max} \triangleq \max_{k\in\mathcal{S}} r_k$, and uses the optimal subcarrier, power, and rate allocation and beamforming design as in Section III-B. We evaluate the average total transmission power in the scenario without user transcoding and the sum of the average total transmission power and transcoding power in the scenario with user transcoding. In the following, both measurement metrics are referred to as average power for short. We implement the proposed solutions and baseline schemes using Matlab and CVX (a Matlab software for disciplined convex programming).

In this simulation, we set $\beta_k = 1$ and $E_k = 10^{-6}$ W for all $k \in \mathcal{K}$, $F_h = F_v = 100°$, $U_h \times U_v = 30 \times 15$, $B = 39$ kHz, $N = 64$, $n_0 = 10^{-9}$ W, and $\alpha = 1$. The elements of $\mathbf{H}_{n,k}$, $n \in \mathcal{N}$, $k \in \mathcal{K}$, are independent and identically distributed according to $\mathcal{CN}(0, \mathbf{1}_{M\times M})$. We consider the 3DoF VR video sequence Venice [34] and use the viewing directions of 30 users for the 15-th frame of this video sequence obtained from real measurements in [34] as the predicted viewing directions. To deal with possible prediction errors, an extra 15° in the four directions of the predicted FoV is transmitted for each user [13], [14]. The 360 VR video encoder named Kvazaar is adopted. Set $L = 5$, and choose $D_l, l \in \mathcal{L}$ as in [18]. For any $\mathcal{G}_k, k \in \mathcal{K}$, we evaluate the average power over 100 random realizations of small-scale channel fading coefficients.



(a) Average power versus K. $M = 4$, $\mathbf{r} = (2, 2, 3, 3, 4)$. (b) Average power versus M. $K = 4$, $\mathbf{r} = (2, 3, 3, 4)$.

Fig. 2. Average power versus K and M.

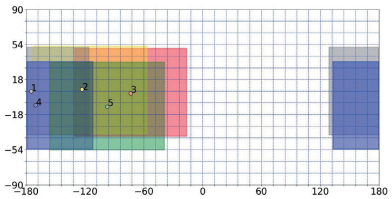


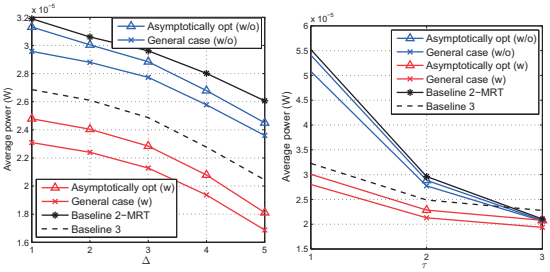
Fig. 3. Viewing directions and corresponding FoVs of 5 users [34].

First, we evaluate the average power over 1,000 random choices for the viewing directions of 1-5 users from 30 users from [34]. Fig. 2 (a) illustrates the average power versus the number of users K. Since the proposed optimal solutions for small multicast groups in the scenarios without and with user transcoding are valid only for $|\mathcal{K}_{\mathcal{S},l}| \leq 3, \mathcal{S} \in \mathcal{I}, l \in \mathcal{L}_{\mathcal{S}}$, and $|\mathcal{S}| \leq 3, \mathcal{S} \in \mathcal{I}$, respectively. Therefore, we do not show their average powers for $K > 3$ where the above-mentioned conditions are not satisfied. We can observe that the average powers of the proposed solutions and baseline schemes increase with K, as the transmission load increases with K. When $K \leq 3$, the proposed solutions for the general case achieve close-to-optimal powers. Given the unsatisfactory performance of Baseline 1, we no longer compare with it in the remaining figures. Fig. 2 (b) illustrates the average power versus the number of antennas M. We can observe that the powers achieved by the proposed solutions and baseline schemes decrease with M. Also, when M is sufficiently large, the proposed asymptotically optimal solutions reach close-to-optimal average powers, demonstrating the asymptotically optimalities of the proposed solutions for the case of a large antenna array.

Next, we show the impacts of the concentration of the viewing directions of all users and the similarity of the required quality levels of all users. We choose the viewing directions of 5 users out of 30 users from [34], i.e., $(\nu_k, \gamma_k)_{k\in\{1,\ldots,5\}}$, as shown in Fig. 3. To show the impact of the concentration of the viewing directions of all users, based on the chosen viewing directions, we consider five sets of viewing directions, i.e., $(\nu_1+\Delta, \gamma_1), (\nu_2+\Delta, \gamma_2), (\nu_3, \gamma_3), (\nu_4-\Delta, \gamma_4)$, and $(\nu_5-\Delta, \gamma_5)$, $\Delta = 1, \ldots, 5$, and evaluate the corresponding average powers. Note that Δ reflects the concentration of the viewing directions of the 5 users. In particular, the concentration increases with Δ. Fig. 4 (a) shows the average power versus the concentration parameter Δ. We can see that each multicast scheme's average power decreases with Δ, since multicast opportunities increase with Δ. To show the impact of the similarity of the required quality levels, we consider three sets of required quality levels, i.e., $\bar{\mathbf{r}}_\tau = (\min\{\tau, 3\}, \min\{\tau+1, 3\}, 3, \max\{3, 5-\tau\}, \max\{3, 6-\tau\})$, $\tau = 1, 2, 3$. Note that τ indicates the similarity of the

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
The final version of this article is available at http://dx.doi.org/10.1109/TWC.2021.3067803



(a) Average power versus Δ. (b) Average power versus τ. $K = 5$, $M = 4$, $\mathbf{r} = K = 5$, $M = 4$. $(2, 2, 3, 3, 4)$.

Fig. 4. Average power versus Δ and τ.

required quality levels of the users. Specifically, the required quality levels are closer when τ is larger. Fig. 4 (b) illustrates the average power versus the similarity parameter τ. We can see that when τ increases, the average power of each multicast scheme decreases, due to the rise of natural multicast opportunities. Furthermore, as τ increases, the gaps between the average powers of the multicast schemes in the scenario without user transcoding and those in the scenario with user transcoding decrease, as transcoding-enabled multicast opportunities decrease.

Fig. 2 and Fig. 4 show that the proposed solutions in the scenario with user transcoding outperform those in the scenario without user transcoding, which demonstrates the importance of using transcoding-enabled multicast opportunities in reducing power consumption. Fig. 2 and Fig. 4 also show that the proposed solutions outperform the baseline schemes. Specifically, the proposed solutions in the scenario without user transcoding outperform Baseline 1, as the proposed solutions utilizing multicast transmission offer higher spectral efficiency. The proposed solutions in the scenario without user transcoding outperform Baseline 2, as the proposed solutions carefully choose beamforming vectors. The proposed solutions in the scenario with user transcoding outperform Baseline 3, as the proposed solutions optimally exploit transcoding-enabled multicast opportunities to reduce power consumption.

VI. Conclusion

This paper studied the optimal wireless streaming of a multi-quality tiled 360 VR video to multiple users in a MIMO-OFDMA system. In the scenario without user transcoding, we minimized the total transmission power by optimizing the beamforming, and subcarrier, transmission power and rate allocation. This is a challenging mixed discrete-continuous optimization problem. We obtained a globally optimal solution for small multicast groups, an asymptotically optimal solution for a large antenna array, and a suboptimal solution for the general case. In the scenario with user transcoding, we minimized the weighted sum of the average total transmission power and the transcoding power by optimizing the quality level selection, beamforming, and subcarrier, transmission power, and rate allocation. This is a more challenging two-timescale mixed discrete-continuous optimization problem. We obtained a globally optimal solution for small multicast groups, an asymptotically optimal solution for a large antenna array, and a low-complexity suboptimal solution for the general case. Finally, numerical results showed that the proposed solutions have significant gains over existing schemes.

Appendix A: Proof of Theorem 1

First, we obtain a problem with the same optimal value as Problem 1. Let $(\boldsymbol{\mu}^\star(\mathbf{h}), \boldsymbol{\eta}^\star(\mathbf{h}), \mathbf{c}^\star(\mathbf{h}), \mathbf{w}^\star(\mathbf{h}))$ denote an optimal solution of Problem 1. By contradiction, we can easily show

$$c^\star_{\mathcal{S},l,n}(\mathbf{h}) = \mu^\star_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{\eta^\star_{\mathcal{S},l,n}(\mathbf{h}) \min\limits_{k\in\mathcal{K}_{\mathcal{S},l}} \beta_k |\mathbf{h}^H_{k,n}\mathbf{w}^\star_n(\mathbf{h})|^2}{M\sigma^2}\right).$$

Thus, equivalently, we can eliminate $\mathbf{c}(\mathbf{h})$ and replace the constraints in (5), (6), and (7) with

$$\sum_{n\in\mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{\eta_{\mathcal{S},l,n}(\mathbf{h}) \min\limits_{k\in\mathcal{K}_{\mathcal{S},l}} \beta_k |\mathbf{h}^H_{k,n}\mathbf{w}_n(\mathbf{h})|^2}{M\sigma^2}\right) \geq |\mathcal{P}_{\mathcal{S}}| D_l,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}. \tag{51}$$

Define $P_{\mathcal{S},l,n}(\mathbf{h}) = \mu_{\mathcal{S},l,n}(\mathbf{h})\eta_{\mathcal{S},l,n}(\mathbf{h})$, $\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{w}_n(\mathbf{h})$, and

$$\widetilde{Q}_{\mathcal{S},l,n}(\mathbf{h}) = \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k |\mathbf{h}^H_{k,n}\mathbf{w}_n(\mathbf{h})|^2}, \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}, n\in\mathcal{N}. \tag{52}$$

By the change of variables of $\mathbf{P}(\mathbf{h}) \triangleq (P_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}},n\in\mathcal{N}}$ and $\tilde{\mathbf{w}}(\mathbf{h}) \triangleq (\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}},n\in\mathcal{N}}$, and introducing the auxiliary variable $\widetilde{\mathbf{Q}}(\mathbf{h}) \triangleq (\widetilde{Q}_{\mathcal{S},l,n}(\mathbf{h}))_{\mathcal{S}\in\mathcal{I},l\in\mathcal{L}_{\mathcal{S}},n\in\mathcal{N}}$, the constraints in (4) and (51) can be transformed to (52) and the following constraints:

$$\|\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h})\|_2 = 1,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}, n\in\mathcal{N}, \tag{53}$$

$$\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h}) = \tilde{\mathbf{w}}_{\mathcal{S}',l',n}(\mathbf{h}),\ \mathcal{S},\mathcal{S}'\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}, l'\in\mathcal{L}_{\mathcal{S}'}, n\in\mathcal{N}, \tag{54}$$

$$\sum_{n\in\mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{P_{\mathcal{S},l,n}(\mathbf{h})}{\mu_{\mathcal{S},l,n}(\mathbf{h})\widetilde{Q}_{\mathcal{S},l,n}(\mathbf{h})}\right) \geq |\mathcal{P}_{\mathcal{S}}| D_l,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}. \tag{55}$$

In addition, by contradiction, we can easily show that (52) can be equivalently replaced by the following constraints:

$$\widetilde{Q}_{\mathcal{S},l,n}(\mathbf{h}) \geq \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k |\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h})|^2},\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}, n\in\mathcal{N}. \tag{56}$$

Thus, Problem 1 and the following problem have the same optimal value.

Problem 13 (Equivalent Problem of Problem 1):

$$E^\star(\mathbf{h}) = \min_{\boldsymbol{\mu}(\mathbf{h}),\mathbf{P}(\mathbf{h}),\tilde{\mathbf{w}}(\mathbf{h}),\widetilde{\mathbf{Q}}(\mathbf{h})} \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} P_{\mathcal{S},l,n}(\mathbf{h})$$
$$\text{s.t.}\quad (1),\ (2),\ (8),\ (53),\ (54),\ (55),\ (56).$$

Denote $(\boldsymbol{\mu}^\star(\mathbf{h}), \boldsymbol{P}^\star(\mathbf{h}), \tilde{\mathbf{w}}^\star(\mathbf{h}), \widetilde{\mathbf{Q}}^\star(\mathbf{h}))$ an optimal solution.

Then, we obtain a problem with the same optimal value as Problem 13. Consider the following problem.

Problem 14 (Equivalent Problem of Problem 13):

$$E^\ddagger(\mathbf{h}) = \min_{\boldsymbol{\mu}(\mathbf{h}),\mathbf{P}(\mathbf{h}),\tilde{\mathbf{w}}(\mathbf{h}),\widetilde{\mathbf{Q}}(\mathbf{h})} \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} P_{\mathcal{S},l,n}(\mathbf{h})$$
$$\text{s.t.}\quad (1),\ (2),\ (8),\ (53),\ (55),\ (56).$$

Let $(\boldsymbol{\mu}^\ddagger(\mathbf{h}), \boldsymbol{P}^\ddagger(\mathbf{h}), \tilde{\mathbf{w}}^\ddagger(\mathbf{h}), \widetilde{\mathbf{Q}}^\ddagger(\mathbf{h}))$ denote an optimal solution. As Problem 13 has extra constraints, i.e., (54), compared to Problem 14, $E^\star(\mathbf{h}) \geq E^\ddagger(\mathbf{h})$. Thus, it remains to show that $E^\star(\mathbf{h}) \leq E^\ddagger(\mathbf{h})$. Based on an optimal solution of Problem 14,

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication. The final version of record is available at http://dx.doi.org/10.1109/TWC.2021.3067803

i.e., $(\boldsymbol{\mu}^{\ddagger}(\mathbf{h}), \boldsymbol{P}^{\ddagger}(\mathbf{h}), \tilde{\mathbf{w}}^{\ddagger}(\mathbf{h}), \widetilde{\mathbf{Q}}^{\ddagger}(\mathbf{h}))$, we construct a feasible solution of Problem 13, whose objective value is $E^{\ddagger}(\mathbf{h})$. Specifically, for all $n \in \mathcal{N}$, we construct $\tilde{\mathbf{w}}^{\diamond}_{\mathcal{S},l,n}(\mathbf{h}) = \tilde{\mathbf{w}}^{\ddagger}_{\mathcal{S}_n,l_n,n}(\mathbf{h})$, $\mathcal{S} \in \mathcal{I}$, $l \in \mathcal{L}_{\mathcal{S}}$, where $(\mathcal{S}_n, l_n)$ satisfies $\mu^{\ddagger}_{\mathcal{S},l,n}(\mathbf{h}) = 1$.[8] It is obvious that $(\boldsymbol{\mu}^{\ddagger}(\mathbf{h}), \boldsymbol{P}^{\ddagger}(\mathbf{h}), \tilde{\mathbf{w}}^{\diamond}(\mathbf{h}), \widetilde{\mathbf{Q}}^{\ddagger}(\mathbf{h}))$ is a feasible solution of Problem 13. Thus, we have $E^{\star}(\mathbf{h}) \leq \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} P^{\ddagger}_{\mathcal{S},l,n}(\mathbf{h}) = E^{\ddagger}(\mathbf{h})$. By $E^{\star}(\mathbf{h}) \geq E^{\ddagger}(\mathbf{h})$ and $E^{\star}(\mathbf{h}) \leq E^{\ddagger}(\mathbf{h})$, we have $E^{\star}(\mathbf{h}) = E^{\ddagger}(\mathbf{h})$.

Next, we show that Problem 14 and Problem 2 have the same optimal value. It is obvious that Problem 14 is equivalent to the following problem.

Problem 15 (Equivalent Problem of Problem 14 for $\mathbf{h}$*):*

$$E^{\ddagger}(\mathbf{h}) = \min_{\boldsymbol{\mu}(\mathbf{h}),\mathbf{P}(\mathbf{h})} \frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} P_{\mathcal{S},l,n}(\mathbf{h})$$

$$\text{s.t.}\quad (1),\ (2),\ (8),$$

$$\sum\nolimits_{n\in\mathcal{N}} \mu_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{P_{\mathcal{S},l,n}(\mathbf{h})}{\mu_{\mathcal{S},l,n}(\mathbf{h})\widetilde{Q}^{\star}_{\mathcal{S},l,n}(\mathbf{h})}\right) \geq |\mathcal{P}_{\mathcal{S}}|D_l,\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}_{\mathcal{S}}, k\in\mathcal{K}_{\mathcal{S},l}, \tag{57}$$

where $\widetilde{Q}^{\star}_{\mathcal{S},l,n}(\mathbf{h})$ is given by the following problem.

$$\widetilde{Q}^{\star}_{\mathcal{S},l,n}(\mathbf{h}) \triangleq \min_{\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h})} \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k|\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}_{\mathcal{S},l,n}(\mathbf{h})|^2} \quad \text{s.t.}\quad (53). \tag{58}$$

As

$$\begin{aligned} Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) &\overset{(a)}{=} \frac{\mathrm{tr}(\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))}{\|\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})\|_2^2} \\ &\overset{(b)}{=} \left(\min_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{\mathrm{tr}(\beta_k \mathbf{h}_{k,n}\mathbf{h}^H_{k,n}\mathbf{V}^{\dagger}_{\mathcal{S},l,n})}{M\sigma^2}\right)\left(\max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k|\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})|^2}\right) \\ &\overset{(c)}{=} \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k|\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})|^2} = \tilde{Q}^{\star}_{\mathcal{S},l,n}(\mathbf{h}), \end{aligned} \tag{59}$$

where (a) is due to (53), (b) is due to Claim 2 in [23], and (c) is due to that $\frac{\min_{k\in\mathcal{K}_{\mathcal{S},l}} \mathrm{tr}(\beta_k \mathbf{h}_{k,n}\mathbf{h}^H_{k,n}\mathbf{V}^{\dagger}_{\mathcal{S},l,n})}{M\sigma^2} = 1$ (which can be easily shown by contradiction), $E^{\ddagger}(\mathbf{h}) = E^{\dagger}(\mathbf{h})$.

Finally, we show that Problem 1 and Problem 2 have the same optimal value and characterize the relation between their optimal solutions. As $E^{\star}(\mathbf{h}) = E^{\ddagger}(\mathbf{h})$ and $E^{\ddagger}(\mathbf{h}) = E^{\dagger}(\mathbf{h})$, we know that the optimal values of Problem 1 and Problem 2 are identical, i.e.,

$$E^{\star}(\mathbf{h}) = E^{\dagger}(\mathbf{h}). \tag{60}$$

In the sequel, we show that $(\boldsymbol{\mu}^{\dagger}(\mathbf{h}), \boldsymbol{\eta}^{\dagger}(\mathbf{h}), \mathbf{c}^{\dagger}(\mathbf{h}), \mathbf{w}^{\dagger}(\mathbf{h}))$ is an optimal solution of Problem 1. It is obvious that $(\boldsymbol{\mu}^{\dagger}(\mathbf{h}), \boldsymbol{\eta}^{\dagger}(\mathbf{h}), \mathbf{c}^{\dagger}(\mathbf{h}), \mathbf{w}^{\dagger}(\mathbf{h}))$ satisfies the constraints in (1), (2), (3), (5), and (7). It remains to show that $(\boldsymbol{\mu}^{\dagger}(\mathbf{h}), \boldsymbol{\eta}^{\dagger}(\mathbf{h}), \mathbf{c}^{\dagger}(\mathbf{h}), \mathbf{w}^{\dagger}(\mathbf{h}))$ satisfies (4) and (6). We have

$$\|\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\|_2 \overset{(d)}{=} \sqrt{\mathrm{tr}\left(\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\right)} = \sqrt{Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}, \tag{61}$$

where (d) is due to $\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))^H$. Thus, we have

$$\|\mathbf{w}^{\dagger}_n(\mathbf{h})\|_2 \overset{(e)}{=} \|\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} \mu^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\frac{\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{\|\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\|_2}\|_2 \overset{(f)}{=} 1, \tag{62}$$

where (e) is due to (13) and (61), and (f) is due to (1) and (2). Thus, $\mathbf{w}^{\dagger}(\mathbf{h})$ satisfies (4). Besides, we have

$$\begin{aligned} &\mu^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{\eta^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\beta_k|\mathbf{h}^H_{n,k}\mathbf{w}_n(\mathbf{h})|^2}{M\sigma^2}\right) \\ &\overset{(g)}{\geq} \mu^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{P^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\beta_k|\mathbf{h}^H_{n,k}\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})|^2}{M\sigma^2 Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}\right) \\ &\overset{(h)}{\geq} \mu^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) B \log_2\left(1 + \frac{P^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}\right) = c^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}), \end{aligned} \tag{63}$$

where (g) is due to (12), (13), (1) and (2), and (h) is due to (10) and $\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))^H$. Thus, $(\boldsymbol{\mu}^{\dagger}(\mathbf{h}), \boldsymbol{\eta}^{\dagger}(\mathbf{h}), \mathbf{c}^{\dagger}(\mathbf{h}), \mathbf{w}^{\dagger}(\mathbf{h}))$ satisfies (6). Therefore, $(\boldsymbol{\mu}^{\dagger}(\mathbf{h}), \boldsymbol{\eta}^{\dagger}(\mathbf{h}), \mathbf{c}^{\dagger}(\mathbf{h}), \mathbf{w}^{\dagger}(\mathbf{h}))$ is a feasible solution of Problem 1, implying

$$\begin{aligned} E^{\star}(\mathbf{h}) &\leq \sum\nolimits_{n\in\mathcal{N}}\sum\nolimits_{\mathcal{S}\in\mathcal{I}}\sum\nolimits_{l\in\mathcal{L}_{\mathcal{S}}} \frac{\mu^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\eta^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{M} \\ &\overset{(i)}{\leq} \sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}_{\mathcal{S}}} \frac{P^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{M} = E^{\dagger}(\mathbf{h}) \overset{(j)}{=} E^{\star}(\mathbf{h}), \end{aligned} \tag{64}$$

where (i) is due to (1), (2) and (12), and (j) is due to (60). Thus, we can conclude that $(\boldsymbol{\mu}^{\dagger}(\mathbf{h}), \boldsymbol{\eta}^{\dagger}(\mathbf{h}), \mathbf{c}^{\dagger}(\mathbf{h}), \mathbf{w}^{\dagger}(\mathbf{h}))$ achieves the optimal value of Problem 1 and is an optimal solution of Problem 1.

Appendix B: Proof of Theorem 2

First, we show that the problem in (58) and Problem 3 are equivalent. We have

$$\begin{aligned} \tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h}) &\overset{(a)}{=} \frac{\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})}{\|\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})\|} \overset{(b)}{=} \frac{\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{\|\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})\|} \overset{(c)}{=} \frac{\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{\sqrt{Q^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}} \\ &\overset{(d)}{=} \frac{\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})}{\sqrt{\max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k|\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})|^2}}}, \end{aligned} \tag{65}$$

where (a) is due to (53), (b) is due to Claim 2 in [23], (c) is due to (61), and (d) is due to (59). Thus, we have

$$\begin{aligned} &\mathbf{V}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^{\dagger}_{\mathcal{S},l,n}(\mathbf{h}))^H \\ &= \tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})(\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h}))^H \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k|\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}^{\star}_{\mathcal{S},l,n}(\mathbf{h})|^2}. \end{aligned} \tag{66}$$

By (59) and (66), we can conclude that the problem in (58) and Problem 3 are equivalent.

Next, we obtain an asymptotically optimal solution of the problem in (58). Following the proof of Theorem 1 in [29], we can show that $\tilde{\mathbf{w}}^{*}_{\mathcal{S},l,n}(\mathbf{h}) = \frac{\sum_{k\in\mathcal{K}_{\mathcal{S},l}} \xi^{*}_{n,k}\mathbf{h}_{n,k}}{\left\|\sum_{k\in\mathcal{K}_{\mathcal{S},l}} \xi^{*}_{n,k}\mathbf{h}_{n,k}\right\|_2}$ is asymptotically optimal for the problem in (58), where $\xi^{*}_{n,k}$ is an optimal solution of the following problem:

$$\min_{\xi_{n,k}} \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{\sigma^2\sum_{j\in\mathcal{K}}\xi^2_{n,j}}{\beta_k\xi^2_{n,k}}. \tag{67}$$

[8] Due to the constraints in (1) and (2), there exists only one message $(\mathcal{S}_n, l_n)$ with $\mu^{\ddagger}_{\mathcal{S}_n,l_n,n}(\mathbf{h}) = 1$, for all $n \in \mathcal{N}$.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications

This problem is similar to Problem Q in [29]. Using the method proposed in [29], we have $\xi^*_{n,k} = \frac{1}{\sqrt{\beta_k}}$. Thus, the asymptotically optimal solution of the problem in (58) can be written as $\tilde{\mathbf{w}}^*_{\mathcal{S},l,n}(\mathbf{h}) = \frac{\sum_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{1}{\sqrt{\beta_k}}\mathbf{h}_{n,k}}{\left\|\sum_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{1}{\sqrt{\beta_k}}\mathbf{h}_{n,k}\right\|_2}$.

Finally, we show that $\mathbf{V}^*_{\mathcal{S},l,n}(\mathbf{h})$ is an asymptotically optimal solution of Problem 3. By $\mathbf{V}^*_{\mathcal{S},l,n}(\mathbf{h}) = \mathbf{v}^*_{\mathcal{S},l,n}(\mathbf{h})(\mathbf{v}^*_{\mathcal{S},l,n}(\mathbf{h}))^H$ and (22), we have

$$\mathbf{V}^*_{\mathcal{S},l,n}(\mathbf{h}) = \tilde{\mathbf{w}}^*_{\mathcal{S},l,n}(\mathbf{h})(\tilde{\mathbf{w}}^*_{\mathcal{S},l,n}(\mathbf{h}))^H \max_{k\in\mathcal{K}_{\mathcal{S},l}} \frac{M\sigma^2}{\beta_k|\mathbf{h}^H_{k,n}\tilde{\mathbf{w}}^*_{\mathcal{S},l,n}(\mathbf{h})|^2}. \tag{68}$$

Since $\tilde{\mathbf{w}}^*_{\mathcal{S},l,n}(\mathbf{h})$ is an asymptotically optimal solution of the problem in (58) , by (66) and (68), we can conclude that $\mathbf{V}^*_{\mathcal{S},l,n}(\mathbf{h})$ is an asymptotically optimal solution of Problem 3.

Appendix C: Proof of Theorem 3

First, we show that the optimal value of Problem 10 is no greater than that of Problem 11, i.e., $E^\dagger_t \le \bar{E}^*_t$. Based on an optimal solution of Problem 11, i.e., $(\bar{\mathbf{N}}^*, \bar{\mathbf{P}}^*, \mathbf{x}^*)$, we construct a feasible solution of Problem 10, whose objective value is $\bar{E}^*$. Specifically, we construct $\bar{\mu}^*_{\mathcal{S},l,n} = \frac{\bar{N}^*_{\mathcal{S},l}}{N}$, $\bar{p}^*_{\mathcal{S},l,n} = \frac{\bar{P}^*_{\mathcal{S},l}}{N}$, $\mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, n\in\mathcal{N}$. Then, it is obvious that $(\bar{\boldsymbol{\mu}}^*, \bar{\boldsymbol{p}}^*, \mathbf{x}^*)$ satisfies the constraints of Problem 10, implying that it is a feasible solution of Problem 10. Besides, we have

$$E^\dagger_t \le \sum_{n\in\mathcal{N}}\sum_{\mathcal{S}\in\mathcal{I}}\sum_{l\in\mathcal{L}} \frac{\bar{p}^*_{\mathcal{S},l,n}}{M} + \alpha E_{\mathrm{tc}}(\mathbf{x}^*) \overset{(a)}{=} \bar{E}^*_t, \tag{69}$$

where (a) is due to $\bar{\mu}^*_{\mathcal{S},l,n} = \frac{\bar{N}^*_{\mathcal{S},l}}{N}$, $\bar{p}^*_{\mathcal{S},l,n} = \frac{\bar{P}^*_{\mathcal{S},l}}{N}$.

Next, we show that the optimal value of Problem 10 is no smaller than that of Problem 11, i.e., $E^\dagger_t \ge \bar{E}^*_t$. Base on an optimal solution of Problem 10, i.e., $(\bar{\boldsymbol{\mu}}^\dagger, \bar{\boldsymbol{p}}^\dagger, \mathbf{x}^\dagger)$, we construct a feasible solution of Problem 11, whose objective value is $E^\dagger_t$. Specifically, we construct $\bar{N}^\dagger_{\mathcal{S},l} = \sum_{n\in\mathcal{N}} \bar{\mu}^\dagger_{\mathcal{S},l,n}$, $\bar{P}^\dagger_{\mathcal{S},l} = \sum_{n\in\mathcal{N}} \bar{p}^\dagger_{\mathcal{S},l,n}$, $\mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, n\in\mathcal{N}$. It is obvious that $(\bar{\mathbf{N}}^\dagger, \bar{\mathbf{P}}^\dagger, \mathbf{x}^\dagger)$ satisfies the constraints in (31), (32), (33), (44), (45), and (46). We also have

$$\begin{aligned}
&\frac{1}{N}\bar{N}^\dagger_{\mathcal{S},l} B\log_2\left(1+\frac{\bar{P}^\dagger_{\mathcal{S},l}}{\bar{Q}_k\bar{N}^\dagger_{\mathcal{S},l}}\right)\\
&= \left(\sum_{n\in\mathcal{N}}\frac{1}{N}\bar{\mu}^\dagger_{\mathcal{S},l,n}\right) B\log_2\left(1+\frac{\sum_{n\in\mathcal{N}}\frac{1}{N}\bar{p}^\dagger_{\mathcal{S},l,n}}{\bar{Q}_k\sum_{n\in\mathcal{N}}\frac{1}{N}\bar{\mu}_{\mathcal{S},l,n}}\right)\\
&\overset{(b)}{\ge} \sum_{n\in\mathcal{N}}\bar{\mu}^\dagger_{\mathcal{S},l,n} B\log_2\left(1+\frac{\bar{p}^\dagger_{\mathcal{S},l,n}}{\bar{Q}_k\bar{\mu}^\dagger_{\mathcal{S},l,n}}\right)\\
&\overset{(c)}{\ge} |\mathcal{P}_{\mathcal{S}}| D_l x_{\mathcal{S},l,k},\ \mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, k\in\mathcal{S}, n\in\mathcal{N},
\end{aligned} \tag{70}$$

where (b) is due to the concavity of $\bar{\mu}_{\mathcal{S},l,n} B\log_2\left(1+\frac{\bar{p}_{\mathcal{S},l,n}}{\bar{Q}_k\bar{\mu}_{\mathcal{S},l,n}}\right)$ in $(\bar{\mu}_{\mathcal{S},l,n}, \bar{p}_{\mathcal{S},l,n})$, and (c) is due to (40). By (70), we know that $(\bar{\mathbf{N}}^\dagger, \bar{\mathbf{P}}^\dagger, \mathbf{x}^\dagger)$ satisfies the constraints in (47). Thus, $(\bar{\mathbf{N}}^\dagger, \bar{\mathbf{P}}^\dagger, \mathbf{x}^\dagger)$ is a feasible solution of Problem 11. In addition, we have

$$E^\dagger_t \overset{(d)}{=} \frac{1}{M}\sum\nolimits_{\mathcal{S}\in\mathcal{I}}\sum\nolimits_{l\in\mathcal{L}} \bar{P}^\dagger_{\mathcal{S},l} + \alpha E_{\mathrm{tc}}(\mathbf{x}^\dagger) \ge \bar{E}^*_t, \tag{71}$$

where (d) is due to $\bar{N}^\dagger_{\mathcal{S},l} = \sum_{n\in\mathcal{N}}\bar{\mu}^\dagger_{\mathcal{S},l,n}$, $\bar{P}^\dagger_{\mathcal{S},l} = \sum_{n\in\mathcal{N}}\bar{p}^\dagger_{\mathcal{S},l,n}$, $\mathcal{S}\in\mathcal{I}, l\in\mathcal{L}, n\in\mathcal{N}$.

Finally, we show that the optimal values of Problem 10 and Problem 11 are identical and characterize the relatioship between their optimal solutions. By (69) and (71), we have $E^\dagger_t = \bar{E}^*_t$. By $E^\dagger_t = \bar{E}^*_t$ and (69), we know that $(\bar{\boldsymbol{\mu}}^*, \bar{\boldsymbol{p}}^*, \mathbf{x}^*)$ is an optimal solution of Problem 10. By $E^\dagger_t = \bar{E}^*_t$ and (71), we know that $(\bar{\mathbf{N}}^\dagger, \bar{\mathbf{P}}^\dagger, \mathbf{x}^\dagger)$ is an optimal solution of Problem 11. Therefore, the proof of Theorem 3 is completed.

References

[1] C. Guo, Y. Cui, Z. Liu, and D. Ng, "Optimal transmission of multi-quality tiled 360 VR video in MIMO-OFDMA systems," in *Proc. of IEEE ICC*, Jun. 2021.

[2] R. Ju, J. He, F. Sun, J. Li, F. Li, J. Zhu, and L. Han, "Ultra wide view based panoramic VR streaming," in *Proc. of the Workshop on VR/AR Network*, Aug. 2017, pp. 19–23.

[3] "Virtual reality (VR) market - growth, trends, and forecast (2020 - 2025)," https://www.mordorintelligence.com/industry-reports/virtual-reality-market/, Jan. 2020.

[4] T. Dang, and M. Peng, "Joint radio communication, caching, and computing design for mobile virtual reality delivery in fog radio access networks," *IEEE J. Select. Areas Commun.*, vol. 37, no. 7, pp. 1594-1607, Jul. 2019.

[5] M. Chen, W. Saad, and C. Yin, "Virtual reality over wireless networks: quality-of-service model and learning-based resource management," *IEEE Trans. Commun.,* vol. 66, no. 11, pp. 5621-5635, Nov. 2018.

[6] Y. Sun, Z. Chen, M. Tao, and H. Liu, "Communications, caching, and computing for mobile virtual reality: modeling and tradeoff," *IEEE Trans. Commun.,* vol. 67, no. 11, pp. 7573-7586, Nov. 2019.

[7] V. R. Gaddam, M. Riegler, R. Eg, C. Griwodz, and P. Halvorsen, "Tiling in interactive panoramic video: approaches and evaluation," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1819–1831, Sep. 2016.

[8] X. Lan, X. Zhang, and Z. Guo, "CLS: A cross-user learning based system for improving QoE in 360-degree video adaptive streaming," *in Proc. of ACM Multimedia*, Oct. 2018, pp. 564-572.

[9] C. Jacob, R. Aksu, X. Corbillon, G. Simon, and V. Swaminathan, "Viewport-driven rate-distortion optimized 360 video streaming," *in Proc. of IEEE ICC*, May 2018, pp. 1-7.

[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. Emerging and Select. Top. Circuits Syst.*, vol. 9, no. 1, pp. 29-42, Mar. 2019.

[11] C. Perfecto, M. S. Elbamby, J. D. Ser, and M. Bennis, "Taming the latency in multi-user VR 360: a QoE-aware deep learning-aided multicast framework," *IEEE Trans. on Commun*, vol. 68, no. 4, pp. 2491-2508, Apr. 2020.

[12] F. Hu, Y. Deng, and A. H. Aghvami, "Correlation-aware cooperative multigroup broadcast 360 video delivery network: a hierarchical deep reinforcement learning approach," *arXiv preprint arXiv:2010.11347*, Oct. 2020.

[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] ——, "Optimal multicast of tiled 360 VR video in OFDMA systems," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2563–2566, Oct. 2018.

[15] H. Ahmadi, O. Eltobgy, and M. Hefeeda, "Adaptive multicast streaming of virtual reality content to mobile users," in *Proc. of the on Thematic Workshops of ACM Multimedia*, Oct. 2017, pp. 170–178.

[16] Z. Zhilong, Z. Ma, Y. Sun, and D. Liu, "Wireless multicast of virtual reality videos with MPEG-I format," *IEEE Access*, vol. 7, pp. 176 693–176 705, Dec. 2019.

[17] N. Kan, C. Liu, J. Zou, C. Li, and H. Xiong, "A server-side optimized hybrid multicast-unicast strategy for multi-user adaptive 360-degree video streaming," in *Proc. of IEEE ICIP*, Sep. 2019, pp. 141–145.

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

[19] K. Long, Y. Cui, C. Ye, and Z. Liu, "Optimal wireless streaming of multi-quality 360 VR video by exploiting natural, relative smoothness-enabled and transcoding-enabled multicast opportunities," *IEEE Trans. Multimedia*, 2020.

[20] J. Joung, H. D. Nguyen, P. H. Tan, and S. Sun, "Multicast linear precoding for MIMO-OFDM systems," *IEEE Commun. Lett.*, vol. 19, no. 6, pp. 993–996, Jun. 2015.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3067803, IEEE Transactions on Wireless Communications
The final version of record is available at http://dx.doi.org/10.1109/TWC.2021.3067803

[21] J. Xu, S. Lee, W. Kang, and J. Seo, "Adaptive resource allocation for MIMO-OFDM based wireless multicast systems," *IEEE Trans. Broadcast.*, vol. 56, no. 1, pp. 98–102, Mar. 2010.

[22] G. Venkatraman, A. Tolli, M. Juntti, and L. Tran, "Multigroup multicast beamformer design for MISO-OFDM with antenna selection," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5832–5847, Nov. 2017.

[23] 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.

[24] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, Mar. 2015.

[25] J. G. Andrews and T. H. Meng, "Optimum power control for successive interference cancellation with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 375–383, Mar. 2003.

[26] J. Lindblom, E. Karipidis, and E. G. Larsson, "Efficient computation of pareto optimal beamforming vectors for the MISO interference channel with successive interference cancellation," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4782–4795, Oct. 2013.

[27] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.

[28] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[29] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE J. Sel. Areas. Commun.*, vol. 32, no. 6, pp. 1180–1193, Jun. 2014.

[30] T. Lipp and S. Boyd, "Variations and extension of the convex–concave procedure," *Optimization and Engineering*, vol. 17, no. 2, pp. 263–287, Jun. 2016.

[31] K. Long, C. Ye, Y. Cui, and Z. Liu, "Optimal transmission of multi-quality tiled 360 VR video by exploiting multicast opportunities," in *Proc. of IEEE GLOBECOM*, Dec. 2019, pp. 1–6.

[32] W. Xu, Y. Cui, and Z. Liu, "Optimal multi-view video transmission in multiuser wireless networks by exploiting natural and view synthesisenabled multicast opportunities," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1494–1507, Mar. 2020.

[33] C. Guo, Y. Cui, D. W. K. Ng, and Z. Liu, "Multi-quality multicast beamforming with scalable video coding," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5662–5677, Nov. 2018.

[34] "360-degree videos head movements dataset," http://dash.ipv6.enstb.fr/headMovements/.

Chengjun Guo received the B.E. degree in electronic and information engineering from Xi'an Jiao Tong University, China, in 2015. He is currently pursuing the Ph.D. degree with the Department of Electric Engineering, Shanghai Jiao Tong University, China. His research interests include multimedia transmission in wireless networks and optimization.

Lingzhi Zhao received the B.E. degree in information and communication engineering from Shanghai University, China, in 2019. He is currently pursuing the master degree with the Department of Electric Engineering, Shanghai Jiao Tong University, China. His research interests include video streaming and optimization.

Ying Cui (S'08-M'12) received the B.E. degree in electronic and information engineering from Xian Jiao Tong University, China, in 2007, and the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2011. From 2012 to 2013, she was a Post-Doctoral Research Associate with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. From 2013 to 2014, she was a Post-Doctoral Research Associate with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. Since 2015, she has been an Associate Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China. Her current research interests include optimization, cache-enabled wireless networks, mobile edge computing, and delay-sensitive cross-layer control. She was selected to the Thousand Talents Plan for Young Professionals of China in 2013. She was a recipient of the Best Paper Award at the IEEE ICC, London, U.K., in June 2015. She serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Zhi Liu (S'11-M'14-SM'19) received the B.E., from the University of Science and Technology of China, China and Ph.D. degree in informatics in National Institute of Informatics. He is currently an Associate Professor at The University of Electro-Communications.

His research interest includes video network transmission, vehicular networks and mobile edge computing. He is now an editorial board member of Springer wireless networks and has been a Guest Editor of ACM/Springer Mobile Networks and Applications, Springer Wireless networks and IEICE Transactions on Information and Systems. He is a senior member of IEEE.

Derrick Wing Kwan Ng (S'06-M'12-SM'17-F'21) received the bachelor degree with first-class honors and the Master of Philosophy (M.Phil.) degree in electronic engineering from the Hong Kong University of Science and Technology (HKUST) in 2006 and 2008, respectively. He received his Ph.D. degree from the University of British Columbia (UBC) in 2012. He was a senior postdoctoral fellow at the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany. He is now working as a Senior Lecturer and a Scientia Fellow at the University of New South Wales, Sydney, Australia. His research interests include convex and non-convex optimization, physical layer security, IRS-assisted communication, UAV-assisted communication, wireless information and power transfer, and green (energy-efficient) wireless communications.

Dr. Ng received the Australian Research Council (ARC) Discovery Early Career Researcher Award 2017, the Best Paper Awards at the WCSP 2020, IEEE TCGCC Best Journal Paper Award 2018, INISCOM 2018, IEEE International Conference on Communications (ICC) 2018, IEEE International Conference on Computing, Networking and Communications (ICNC) 2016, IEEE Wireless Communications and Networking Conference (WCNC) 2012, the IEEE Global Telecommunication Conference (Globecom) 2011, and the IEEE Third International Conference on Communications and Networking in China 2008. He has been serving as an editorial assistant to the Editor-in-Chief of the IEEE Transactions on Communications from Jan. 2012 to Dec. 2019. He is now serving as an editor for the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, and an area editor for the IEEE Open Journal of the Communications Society. Also, he has been listed as a Highly Cited Researcher by Clarivate Analytics since 2018.

1536-1276 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: Shanghai Jiaotong University. Downloaded on May 22,2021 at 07:38:03 UTC from IEEE Xplore. Restrictions apply.
Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.