Multi-user Multimedia Resource Allocation over Multi-carrier Wireless Networks

Yi Su* and Mihaela van der Schaar

ABSTRACT

We address the problem of multi-user video transmission over the uplink of multi-carrier networks from an information theoretic perspective. Under the constraints imposed by the underlying Physical (PHY) and Medium Access Control (MAC) layers, we exploit the unique property of state-of-the-art video coders that can provide inherent bitstream prioritization in terms of distortion impact and solve the problem of allocating wireless resources, i.e., power/rate and sub-carrier assignment, among multiple users such that the weighted sum of the overall video qualities is maximized. We focus on two different types of multiple access strategies and their corresponding achievable rate regions, i.e., Shannon capacity region and Frequency-Division Multiplexing Access (FDMA) capacity region, in the Gaussian multiple access channel. We propose two different approaches to optimize the multi-user multimedia transmission by considering the specific structures of both problems. First, for the general multiple access strategy, under the constraint of its Shannon capacity region, we propose an algorithm to describe the achievable convex utility region directly. Second, for the FDMA strategy, we study the problem by relaxing the original integer programming problem into a convex optimization problem, which makes it tractable to find near-optimal solution analytically. For both multiple access schemes, we start from the two-user case and develop algorithms for finding the (near) optimal resource allocation strategies. Inspired by the intuition gained from the two-user case, we extend the algorithms to the multiple-user case. Our numerical simulations show that the proposed resource allocation algorithms give significant performance improvements as compared to application-layer agnostic solutions that do not consider the quality impact.

This work was supported by NSF CAREER Award CCF-0541867, and grants from UC Micro and ONR. The authors are with the Dept. of Electrical Engineering (EE), University of California Los Angeles (UCLA), 66-147E Engineering IV Building, 420 Westwood Plaza, Los Angeles, CA, 90095-1594, USA, Phone: +1-310-825-5843, Fax: +1-310-206-4685, E-mail: {yisu, mihaela}@ee.ucla.edu

*Corresponding author. Address and contact information: see above.

Index Terms— wireless resource allocation, wireless multimedia, multi-carrier wireless networks, utility driven resource management

I. Introduction

Multi-carrier communication (in particular, Orthogonal Frequency Division Multiplexing (OFDM)) is becoming the leading physical layer technology for many existing and emerging wireless networks and standards [1]. An important application over these networks is bandwidth-intense multimedia streaming. Hence, the development of advanced resource allocation strategies for wireless multimedia applications has recently emerged as an important topic of research. In this paper, we study the problem of optimal resource allocation across multiple users transmitting video over the multi-carrier wireless network infrastructure from an information theoretic perspective.

There has been significant research dedicated to studying resource allocation strategies in wireless networks. Recent research has shown that significant performance gains can be achieved by using dynamic resource allocation. Because of the time-varying property of wireless channels, the knowledge of channel state information can help to allocate limited resources in order to achieve better performance. One example is opportunistic communication, which can effectively exploit the multi-user diversity, and thus increase the total system throughput. The optimization problem for independent and identically-distributed (i.i.d.) fading channels is studied in [2] and [3], where the optimal power allocation over time is characterized. Knopp and Humblet [2] first show that if the channel state information (CSI) is perfectly known by all transmitters, the sum-rate is maximized by a simple strategy: always allocate the power and rate to the user with the best instantaneous channel. Taking into account individual power constraints for the users, Tse and Hanly [3] indicate how to find the power control and rate allocation policies that maximize the weighted sum of the rates, by exploiting the polymatroid structure of the capacity region. Another important result in the area of wireless resource allocation is determined based on a combination of information theory and queueing theory. Yeh and Cohen [4][5] find an optimum policy named "Longest Queue Highest Possible Rate" (LQHPR), which allows the system to obtain the highest stable throughput, the shortest average queue length per user, and thus the shortest average delay. Related cross-layer approaches on queueing stability, delay, and adaptive coding and modulation schemes can be found in [6]-[8]. Recent studies [9][10] show that for multimedia transmission, the approaches above are not optimal from a video quality perspective because the characteristics of video streaming also need to be considered into the cross-layer framework. The optimal rate allocation policy, Largest Quality Improvement Highest Possible Rate (LQIHPR), is proposed to maximize the overall video quality in a single-carrier multiple access fading channel [10]. However, prior works [9][10] cannot be extended to multi-carrier systems directly because allocating power and rate across different sub-carriers in order to maximize the overall quality is non-trivial due to the underlying vector channels. For multi-carrier networks, OFDM systems in particular, most of the papers in the literature are dealing with the downlink rate-sum, downlink utility-sum, and uplink rate-sum maximization problems [11]-[16]. In downlink, the power constraint is imposed over the total transmission power rather than the power of an individual user. In such systems, it has been accepted as an optimal solution that each sub-carrier is allocated to the user with the best channel condition and power is allocated by the water-filling over sub-carriers [11][12]. However, the optimality in downlink does not hold for uplink because each user has its own power constraint. For uplink, joint sub-carrier and power allocation algorithms are proposed to maximize rate-sum capacity [15][16].

In this paper, we address the problem of optimal resource allocation (power, rate, and sub-carrier allocation) across multiple users transmitting video bitstreams. Most of the existing cross-layer research focuses on the interaction among Physical (PHY), Medium Access Control (MAC), and Network layers [2]-[8]. Alternatively, we study this problem using an integrated cross-layer approach that also considers the source coder employed at the Application (APP) layer and the resulting utility impact (i.e. the video quality). We exploit the unique property of state-of-the-art video coders that prioritize the encoded video streams based on overall distortion impact [17]. This prioritization results in a concave increase of the utility (in terms of video quality) as a function of the allocated rate. We develop a unified PHY-MAC-APP framework and study the optimal resource allocation policy which maximizes the weighted sum of video qualities across all users. More importantly, as opposed to conventional approaches, which usually perform optimization considering the capacity region, our approach focuses on explicitly describing the achievable utility region. The proposed solution has many applications in practical systems, where multiple wireless capturing devices are transmitting their content. Typical applications include multi-user video transmission (e.g. uploading movies) over wireless LAN or spectrum agile radio [18][19], video surveillance form wireless camera, and multimedia streaming in digital subscriber line (DSL) systems [20].

In particular, we focus on Gaussian multiple access channels using two different multiple access strategies. First, we consider the general multiple access strategy, which allows users to utilize the entire frequency band simultaneously.

Next, we discuss the Frequency-Division Multiplexing Access (FDMA) strategy in which users share the total bandwidth in an efficient manner [21]. Throughout this paper, we take an information theoretic approach in deriving the optimal resource allocation solutions. Since the capacity region is the fundamental characterization of the achievable rates, we can derive the limit of the achievable video quality of a specific video coder by using operational rate-distortion theory. Our main contributions in this paper are as follows.

First, for the general multiple access strategy, we demonstrate the convexity of the achievable utility region measured in the Peak Signal to Noise Ratio (PSNR) performance achieved by the various receivers. Without requiring full knowledge of the entire Shannon capacity region, we propose a procedure to determine the utility region for the two-user case and extend it to the multiple-user case using a heuristic approach. The proposed algorithms take advantage of the characteristics of the Shannon capacity region, make it tractable to describe the entire utility region, and greatly reduce the complexity of maximizing the weighted sum of the utilities compared to the exhaustive search.

Second, we also examine the case in which multiple users access sub-carriers in the FDMA fashion. In this case, the implementation is simplified, but the problem is converted into an integer programming problem, which makes it difficult to solve. Fortunately, convex optimization theory can provide near-optimal solution analytically. We develop an iterative search algorithm to find the optimal resource allocation strategy for the two-user case and extend it to the multiple-user case by a heuristic approach inspired by the intuition gained from the two-user case.

The rest of this paper is organized as follows. Section II describes the considered model of multi-carrier wireless networks for multi-user video streaming. Section III explains the deployed end-to-end utility objective function and formulates the multi-user resource allocation into an optimization problem. In Section IV, for the general multiple access strategy, we propose a procedure to determine the achievable utility region mapped from Shannon capacity region, and we develop an iterative approach to find the optimal solution for maximizing the weighted sum of the video qualities. Section V discusses optimal resource allocation for the FDMA strategy in detail. Section VI gives simulation results of the proposed algorithms to verify the effectiveness of our algorithms. Conclusions are drawn in Section VII.

II. Multi-carrier Networks for Multi-User Multimedia Transmission

A. *System Description*

In this paper, we focus on the Gaussian multiple access channel. A Gaussian multiple access channel refers to a multiple access channel where the additive noise is Gaussian [22]. The system diagram of the multi-carrier wireless networks for multi-user multimedia transmission is shown in Figure 1.

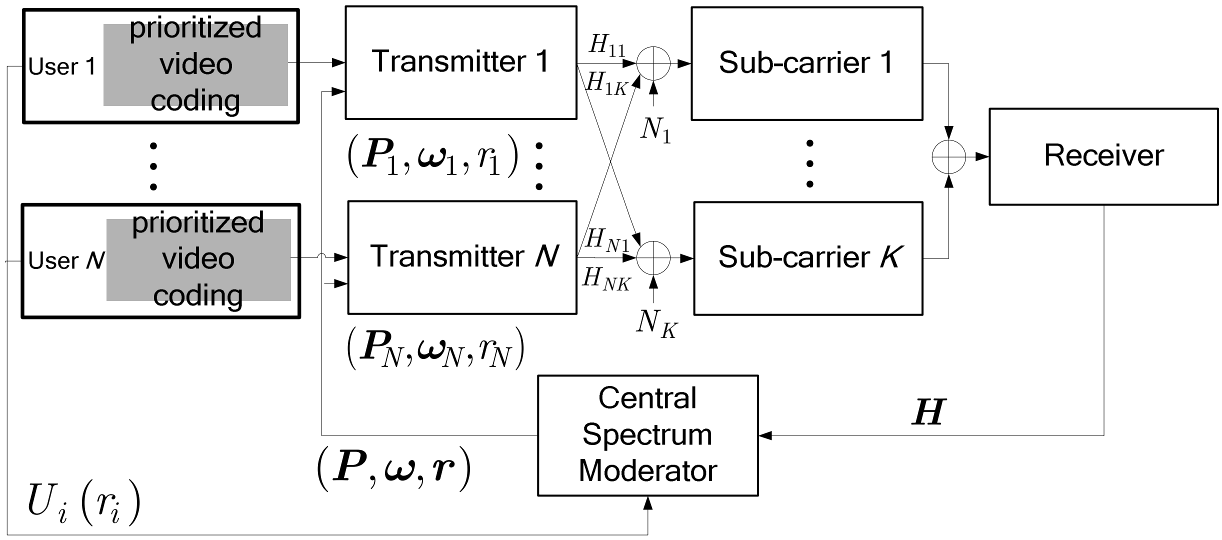

Figure 1. System Structure

Suppose there are N users in the system. The entire frequency band is divided into K sub-carriers and the available bandwidth of each sub-carrier is B. Each user experiences a flat fading channel within the bandwidth of each sub-carrier. We denote user i's channel gain at the jth sub-carrier as H_{ij}. In Figure 1, the received signal at the jth sub-carrier is given by

$$Y_j(n) = \sum_{i=1}^{N} \omega_{ij} H_{ij} X_{ij}(n) + N_j(n), \tag{1}$$

where $X_{ij}(n)$ is the transmitted symbol of user i at jth sub-carrier at time n, ω_{ij} is an indicator of whether user i occupies the jth sub-carrier, and $N_j(n)$ is the additive white Gaussian noise (AWGN) with two-sided spectral density of $N_0/2$. Each user i is subjected to a long-term average power constraint at the jth sub-carrier: $E\left[\left\|X_{ij}(n)\right\|^2\right] \le P_{ij}$.

Here, we denote the CSI vector as $\boldsymbol{H} = (\boldsymbol{H}_1, \boldsymbol{H}_2, \cdots, \boldsymbol{H}_N)$ in which $\boldsymbol{H}_i = (H_{i1}, H_{i2}, \cdots, H_{iK})$, the power allocation vector as $\boldsymbol{P} = (\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ in which $\boldsymbol{P}_i = (P_{i1}, P_{i2}, \cdots, P_{iK})$, the sub-carrier allocation vector as $\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \cdots, \boldsymbol{\omega}_N)$ in which $\boldsymbol{\omega}_i = (\omega_{i1}, \omega_{i2}, \cdots, \omega_{iK})$, and the achievable rate vector $\boldsymbol{r} = [r_1\ r_2 \cdots r_N]^T$ in which r_i is user i's achievable rate under current power and sub-carrier allocation.

Throughout this paper, we discuss two types of MAC strategies, i.e., the general multiple access strategy and the FDMA strategy. The optimal transmission strategy for the multiple access channel generally requires the entire frequency band to be shared by all the users simultaneously. We name this transmission strategy as "general multiple access strategy". Studying this strategy provides us with the upper bound of the utility. However, to achieve the Shannon capacity for a multiple access channel, joint decoding at the receiver is needed in general, which makes the implementation prohibitive. Although the FDMA strategy is not always optimal in the information theoretic sense, it is often desirable from a practical implementation point of view. FDMA transmission schemes allow different users to occupy orthogonal dimensions, so they can be separated at the receiver without joint decoding. This greatly simplifies the receiver design and it is especially suitable in OFDM systems. It allows exploring frequency multi-user diversity and also leads to good spectral efficiency. Hence, we also consider the FDMA strategy in our analysis.

In the scenario of the general multiple access strategy, each user can transmit symbols using any sub-carrier. Therefore, $\omega_{ij} = 1, \forall i, j$. In the case of the FDMA strategy, we have that $\omega_{ij} = 0$ or 1, $\forall i, j$, and $\sum_{i=1}^{N} \omega_{ij} \le 1$, $\forall j$.

We assume that the N users are streaming pre-compressed video content over a shared multi-carrier wireless infrastructure. The Central Spectrum Moderator (CSM) collects the accurate CSI $\boldsymbol{H}$ and utility-rate functions from all users and performs resource allocation to maximize the overall video utility based on the collected information. To perform the resource allocation, the CSM needs to optimally determine the power allocation vector $\boldsymbol{P}$. Here we assume each user is subjected to its maximum power constraint and the maximum allowable power for user i is $P_i^{\max}$:

$$\sum_{j=1}^{K} P_{ij} \le P_i^{\max}. \tag{2}$$

We denote $\boldsymbol{P^{max}} = (P_1^{\max}, P_2^{\max}, \cdots, P_N^{\max})$. Note that if users share the whole frequency band in the FDMA fashion, the CSM also needs to determine $\boldsymbol{\omega}$ in order to allocate each sub-carrier to different users.

B. Utility-Rate Functions

In multimedia applications, the utility represents the video quality, which is determined by the bit rate, video sequences as well as video coder performance. Throughout this paper, we define the utility to be the video quality in terms of the PSNR, as this is the only widely accepted metric for assessing the video quality.

It has been shown that partitioning the packets into different priority classes and correspondingly adjusting the transmission strategies for each class can significantly improve the overall received quality and provide graceful degradation [23][24]. Several operational utility-rate models for video coders based have been proposed [25][26]. It has been shown that these utility-rate models can accurately capture the performance of various coders for different video sequence characteristics and practical video streaming considerations [25]-[27]. Our focus is to characterize the information-theoretic upper bounds for the performance for the achievable video quality region of operational video source schemes. In this paper, we use a popular utility rate model that is well-suited for the operational performance of state-of-the-art prioritized video coders [26]. Based on this model, the utility (PSNR) for user i is given by

$$U_i\left(r_i\right) = 10\log\frac{255^2\left(r_i - R_{0i}\right)}{D_{0i}\left(r_i - R_{0i}\right) + c_i}, \tag{3}$$

where R_{0i}, D_{0i}, c_i are the parameters for this model, which are dependent on the video sequence characteristics and operational encoder-selected parameters. It is easy to see that this is a continuous utility-rate function with a continuously decreasing positive slope, reflecting the inherent prioritization of the video bitstreams[1]. Throughout this paper, we assume that $r_i > R_{0i}$.

III. PROBLEM FORMULATION

In this section, we formulate the multi-user resource allocation into an optimization problem, briefly review the results on achievable capacity regions for the general multiple access strategy and the FDMA strategy from prior literature, and highlight the main challenges in solving the optimization problem in the utility domain for both multiple access strategies.

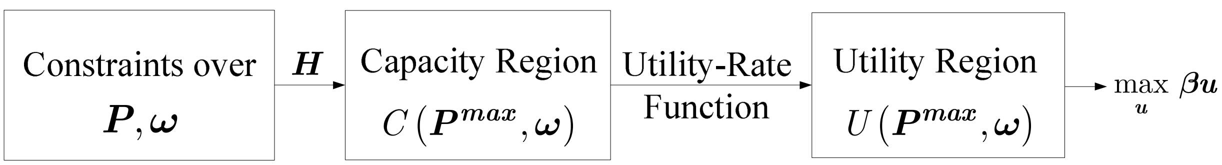

Figure 2. Problem Interpretation

Figure 2 shows the basic idea in formulating this problem, which can be summarized into two steps. First, for any given power allocation $\boldsymbol{P}$ and sub-carrier allocation $\boldsymbol{\omega}$ that satisfy the constraints, there will be a corresponding achievable rate vector $\boldsymbol{r}$ within the capacity region $C\left(\boldsymbol{P}^{max},\boldsymbol{\omega}\right)$, i.e., $\boldsymbol{r} \in C\left(\boldsymbol{P}^{max},\boldsymbol{\omega}\right)$ [22][28]. Second, by mapping the rate vector $\boldsymbol{r}$ into the utility vector $\boldsymbol{u}$ based on the utility-rate functions $U_i\left(r_i\right)$, we can obtain the corresponding utility region $U\left(\boldsymbol{P}^{max},\boldsymbol{\omega}\right)$, which is defined as follows:

$$U\left(\boldsymbol{P}^{max},\boldsymbol{\omega}\right) = \left\{\left(U_1\left(r_1\right),U_2\left(r_2\right),\cdots,U_n\left(r_n\right)\right) : \left(r_1,r_2,\cdots,r_n\right) \in C\left(\boldsymbol{P}^{max},\boldsymbol{\omega}\right)\right\} \tag{4}$$

[1]Please note that all the results in this paper can be extended to other applications with utility-rate functions that are monotonically increasing and strictly concave.

It should be pointed out that $U_i(r_i)$ depends on the video sequence characteristics and operational encoder-selected parameters. The objective function that we aim to maximize is the weighted sum of all the users' video qualities:

$$\max_{\boldsymbol{u}} \boldsymbol{\beta u}, \ s.t. \ \boldsymbol{u} \in U(\boldsymbol{P}^{max}, \boldsymbol{\omega}), \tag{5}$$

where $\boldsymbol{\beta} \in R_+^N$, with $\|\boldsymbol{\beta}\|_1 = 1$, is a given weighted vector that indicates the importance of the various users. Note that $\boldsymbol{\beta}$ can also be interpreted as the bargaining power of the Nash bargaining solution [30], where the possible strategies and effects of assigning bargaining powers in discussed in details.

Figure 2 highlights that, in order to determine an adequate solution for the optimization problem in (5), it is essential to determine whether we can describe the capacity region of each multiple access strategy.

Here we first briefly review the results on the capacity regions for both the general multiple access strategy and the FDMA strategy in Gaussian multiple access channels. For Gaussian multiple access channels with inter-symbol interference (ISI), the optimal multiple access strategy is the general multiple access strategy, which results in the Shannon capacity region. The Shannon capacity region for multiple access multi-carrier system was characterized in [29]. Because each user has different channels, finding the optimal allocation of power over the frequencies is not a trivial task. The optimal power allocation for different points in the capacity region can be found numerically and efficiently [28][29]. In the case of the FDMA strategy, especially in the discrete version, the resource allocation problem is essentially a sub-carrier assignment problem. Unfortunately, an exact solution for this sub-carrier assignment problem usually involves an exhaustive search, which is computationally prohibitive when the number of sub-carriers is large. Therefore, the exact corresponding FDMA capacity region is generally hard to describe [15].

Based on whether or not the capacity regions of both multiple access strategies can be accurately and efficiently described, we propose two approaches to solve the optimization problem in (5).

The first idea is to characterize the achievable capacity region $C(\boldsymbol{P}^{max}, \boldsymbol{\omega})$, map it into the utility region $U(\boldsymbol{P}^{max}, \boldsymbol{\omega})$ using the deployed utility-rate functions and find the corresponding optimum on the boundary of $U(\boldsymbol{P}^{max}, \boldsymbol{\omega})$. This approach is desirable in the general multiple access strategy, where the capacity region can be accurately and efficiently described. However, in the FDMA case, because the capacity boundary is hard to describe in general [15], this approach is not desirable. Moreover, in general, this approach requires the knowledge of the entire capacity boundary to find the optimal solution. Since the capacity boundary consists of infinite number of points and usually lacks closed form expression, the computation complexity of optimal solution is prohibitive and low-complexity method is needed. We will discuss this approach in Section IV in detail.

The second approach is quite straightforward. We can describe the corresponding utility vector $\boldsymbol{u}$ as a function of P_{ij} and ω_{ij} analytically, and then maximize the weighted sum of the utilities directly. For example, in the FDMA case, for any given sub-carrier allocation pattern $\boldsymbol{\omega}$, we can get the closed form expression of the utility vector $\boldsymbol{u}$, and subsequently formulate the optimization problem. In Section V, in the context of the FDMA strategy, we will discuss this approach and its continuous relaxation thoroughly. Although the problem could also be formulated in a similar manner for the general multiple access strategy, this approach is not suitable because in this case, when performing channel decoding, we have to consider decoding order with a total number of $N!$ possibilities to derive the achievable rate vector $\boldsymbol{r}$ and the objective function in (5) is non-convex in the power allocation vector $\boldsymbol{P}$.

IV. Optimal Resource Allocation For The General Multiple Access Strategy

In this section, for the general multiple access strategy, we explain the first approach in solving problem (5) in detail. Based on the fact that the Shannon capacity region under the general multiple access strategy is convex and the utility-rate functions are monotonically increasing and concave, we demonstrate the convexity of achievable utility region. Then we derive the optimality condition for achieving the utility boundary. For the two-user case, we develop a computationally efficient solution, which requires no full knowledge of the entire Shannon capacity region, to determine the boundary of achievable utility region. Subsequently, we extend it to the multiple-user case by heuristic approach. Based on these algorithms, we can solve the original maximization problem in the weighted sum of the utilities and these algorithms could be extended to some other utility fairness criteria (see e.g. [30]).

Note that for the general multiple access strategy, $\omega_{ij} = 1, \forall i, j$. In this section, we neglect $\boldsymbol{\omega}$ and simply denote the capacity region $C(\boldsymbol{P}^{max}, \boldsymbol{\omega})$ and the utility region $U(\boldsymbol{P}^{max}, \boldsymbol{\omega})$ as $C(\boldsymbol{P}^{max})$ and $U(\boldsymbol{P}^{max})$. We denote $\mathcal{R}(\boldsymbol{P})$ and $\mathcal{U}(\boldsymbol{P})$ as the rate vector and utility vector associated with a power allocation vector $\boldsymbol{P}$. We also denote the utility vector associated with a rate vector $\boldsymbol{r} = [r_1\ r_2 \cdots r_N]^T$ as $\boldsymbol{U}(\boldsymbol{r})$, where $\boldsymbol{U}(\boldsymbol{r}) = [U_1(r_1)\ U_2(r_2) \cdots U_N(r_N)]^T$.

A. *Convexity of the Achievable Utility Region*

In this subsection, we show that the achievable utility region $U(\boldsymbol{P}^{max})$ of the general multiple access strategy is convex.

Lemma 1: If a rate vector $\boldsymbol{R} = (R_1, R_2, \cdots, R_N)$ is achievable, any rate vector $\boldsymbol{R}' = (R_1', R_2', \cdots, R_N')$ that satisfies $R_i' \le R_i, \forall i = 1,2,\cdots,N$ is also within the achievable capacity region $C(\boldsymbol{P^{max}})$.

Proof: This property follows from the convex hull operation that forms the capacity region of a Gaussian multiple access channel [22]. ■

Lemma 2: The utility-rate function in (3) is a monotonically increasing and concave function in r_i.

Proof: The monotonically increasing property is straightforward. The concavity can be proved by taking the second derivative of $U_i(r_i)$:

$$\frac{d^2 U_i(r_i)}{dr_i^2} = -10 \frac{c_i \left[2D_{0i}(r_i - R_{0i}) + c_i\right]}{(r_i - R_{0i})^2 \left[D_{0i}(r_i - R_{0i}) + c_i\right]^2}. \tag{6}$$

Therefore, if $r_i > R_{0i}$, $\frac{d^2 U_i(r_i)}{dr_i^2} < 0$. The utility-rate function in (3) is concave in r_i. Note that the monotonically increasing and concave property comes from the inherent prioritization of the video bitstream. ■

This conclusion reflects the fact that efficient video coders prioritize the encoded video streams based on their impacts on the overall distortion, i.e., the more important bits would be sent before less important ones [27].

Proposition 1: The achievable utility region $U(\boldsymbol{P^{max}})$ of the general multiple access strategy is convex.

Proof: The convexity of $U(\boldsymbol{P^{max}})$ can be proven as a direct consequence of Lemma 1 and Lemma 2.

First, let us consider two power vectors $\boldsymbol{P}$ and $\boldsymbol{P}'$ that satisfy $\sum_{k=1}^{K} P_{ik} \le P_i^{\max}$ and $\sum_{k=1}^{K} P_{ik}' \le P_i^{\max}, \forall i = 1,2,\cdots,N$.

Now, define $\hat{\boldsymbol{P}} = \alpha \boldsymbol{P} + (1-\alpha)\boldsymbol{P}', 0 \le \alpha \le 1$. Obviously, for this convex combination to be in $U(\boldsymbol{P^{max}})$, it must satisfy the power constraints:

$$\sum_{k=1}^{K} \hat{P}_{ik} = \left[\alpha \sum_{k=1}^{K} P_{ik} + (1-\alpha) \sum_{k=1}^{K} P_{ik}'\right] \le P_i^{\max}, \forall i = 1,2,\cdots,N. \tag{7}$$

For Shannon capacity region of multi-carrier systems, the achievable rates are concave functions in $\boldsymbol{P}$ [28], therefore

$$\alpha \mathcal{R}(\boldsymbol{P}) + (1-\alpha)\mathcal{R}(\boldsymbol{P}') \preceq \mathcal{R}(\alpha \boldsymbol{P} + (1-\alpha)\boldsymbol{P}') = \mathcal{R}(\hat{\boldsymbol{P}}). \tag{8}$$

By the concave and monotonically increasing property in Lemma 2, (8) can be converted into

$$\begin{aligned}&\alpha\mathcal{U}(\boldsymbol{P})+(1-\alpha)\mathcal{U}(\boldsymbol{P}')=\alpha\boldsymbol{U}(\mathcal{R}(\boldsymbol{P}))+(1-\alpha)\boldsymbol{U}(\mathcal{R}(\boldsymbol{P}'))\\&\preceq\boldsymbol{U}(\alpha\mathcal{R}(\boldsymbol{P})+(1-\alpha)\mathcal{R}(\boldsymbol{P}'))\preceq\boldsymbol{U}(\mathcal{R}(\alpha\boldsymbol{P}+(1-\alpha)\boldsymbol{P}'))=\mathcal{U}(\hat{\boldsymbol{P}})\end{aligned}. \tag{9}$$

By the monotonically increasing property of $U_i(r_i)$, from (9), we know that there exists a rate vector $\bar{\boldsymbol{r}}$ satisfying $\boldsymbol{U}(\bar{\boldsymbol{r}})=\alpha\mathcal{U}(\boldsymbol{P})+(1-\alpha)\mathcal{U}(\boldsymbol{P}')$. Obviously, $\bar{\boldsymbol{r}}\preceq\mathcal{R}(\alpha\boldsymbol{P}+(1-\alpha)\boldsymbol{P}')$.

By Lemma 1, we can conclude that $\bar{\boldsymbol{r}}$ can be achieved directly by a certain power allocation vector $\bar{\boldsymbol{P}}$:

$$\bar{\boldsymbol{r}}=\mathcal{R}(\bar{\boldsymbol{P}}). \tag{10}$$

Therefore, $\forall\ \mathcal{U}(\boldsymbol{P}),\mathcal{U}(\boldsymbol{P}')\in U(\boldsymbol{P}^{max})$, $0\le\alpha\le1\ \Rightarrow\ \alpha\mathcal{U}(\boldsymbol{P})+(1-\alpha)\mathcal{U}(\boldsymbol{P}')=\mathcal{U}(\bar{\boldsymbol{P}})\in U(\boldsymbol{P}^{max})$.

Hence, we can conclude that the utility region is convex. ■

Thus, explicitly characterizing the entire achievable convex utility region $U(\boldsymbol{P}^{max})$ is equivalent to solving the following optimization problem

$$\max_{\boldsymbol{u}}\ \boldsymbol{\beta u}\quad s.t.\ \boldsymbol{u}\in U(\boldsymbol{P}^{max}), \tag{11}$$

for all possible $\boldsymbol{\beta}\in R_+^N$ and $\|\boldsymbol{\beta}\|_1=1$. The problem in (11) is exactly the same form as the weighted sum maximization.

B. *Optimality Condition*

Due to the monotonically increasing property of the utility rate function, for any given power allocation reaching the boundary of $C(\boldsymbol{P}^{max})$, there exists a corresponding point on the boundary of $U(\boldsymbol{P}^{max})$. In this subsection, we derive the mapping function that projects the normal vector to the tangent hyperplane at each boundary point from the capacity region $C(\boldsymbol{P}^{max})$ to the utility region $U(\boldsymbol{P}^{max})$. This mapping function provides the optimality condition under which the problem in (11) reaches the optimum.

Conventional approaches in solving (11), such as Largest Quality Improvement Highest Possible Rate (LQIHPR), searches the optimum along the utility rate function continuously until reaching the boundary of the capacity region [10]. In the case of scalar non-fading AWGN channel and fading channel with given power control, the capacity regions exhibit the polymatroid structure, which makes it possible to characterize the entire boundary with finite inequalities. Therefore, by checking these inequalities, we can examine whether or not a certain rate vector reaches the capacity boundary. However, for multi-carrier networks, this approach is computationally intensive because it is in general impossible to describe the capacity region in finite inequalities and thus, all the points on the boundary surface of the capacity region need to be calculated in advance. Besides, how to characterize the feasible utility region is also of great interest since it has never been described explicitly for the general case in previous approaches [10][30].

In order to solve the problem in (11), we could take advantage of existing numerical algorithms in describing the Shannon capacity region of multi-carrier multiple access channel. For the general multiple access strategy, we can trace out the entire capacity region by efficiently solving the maximization of the weighted sum of the rates for all possible $\boldsymbol{\mu}$:

$$\max_{\boldsymbol{r}} \boldsymbol{\mu r} \quad s.t.\ \boldsymbol{r} \in C\left(\boldsymbol{P}^{max}\right), \tag{12}$$

in which $\boldsymbol{\mu} \in R_+^N$, with $\|\boldsymbol{\mu}\|_1 = 1$, is a given weighted vector whose components represent the relative priority for each user. In this paper, we modify the iterative waterfilling algorithm for the Gaussian vector multiple access channel in [28] to solve (12). (See Appendix A for the details.) The problem now is reduced to finding the optimal solution of problem (11) efficiently, based on the assumption that we can already efficiently solve problem (12) for all possible $\boldsymbol{\mu}$.

We have already shown that the achievable utility region $U\left(\boldsymbol{P}^{max}\right)$ is convex. Therefore, every point on the utility boundary is Pareto optimal [32]. Recall that the utility-rate functions monotonically increase in r_i. We can conclude that each point on the utility boundary is mapped from one point that lies on the capacity boundary. Now we derive the mapping function, which projects the normal vector to the tangent hyperplane at each boundary point from the capacity region $C\left(\boldsymbol{P}^{max}\right)$ to the utility region $U\left(\boldsymbol{P}^{max}\right)$. Suppose a power allocation $\boldsymbol{P}$ reaches the boundaries of $C\left(\boldsymbol{P}^{max}\right)$ and $U\left(\boldsymbol{P}^{max}\right)$. We denote the normal vectors to the tangent hyperplanes at the capacity and utility boundary points as $\boldsymbol{\mu}(\boldsymbol{P}) = \arg \max_{\boldsymbol{\mu} \in R_+^N \ and\ \|\boldsymbol{\mu}\|_1 = 1} \boldsymbol{\mu} \mathcal{R}(\boldsymbol{P})$ and $\boldsymbol{\beta}(\boldsymbol{P}) = \arg \max_{\boldsymbol{\beta} \in R_+^N \ and\ \|\boldsymbol{\beta}\|_1 = 1} \boldsymbol{\beta} \mathcal{U}(\boldsymbol{P})$ respectively.

Proposition 2: For any power allocation $\boldsymbol{P} = (\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ that achieves the boundary of $U\left(\boldsymbol{P}^{max}\right)$ and satisfies (11), the relation between $\boldsymbol{\mu}(\boldsymbol{P})$ and $\boldsymbol{\beta}$ is given by

$$\frac{\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P})}{\left\|\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P})\right\|_1} = \boldsymbol{\mu}(\boldsymbol{P}), \tag{13}$$

in which $\boldsymbol{\lambda}(\boldsymbol{P}) = \left[\left.\frac{\partial U_1(x)}{\partial x}\right|_{x=r_1} \left.\frac{\partial U_2(x)}{\partial x}\right|_{x=r_2} \cdots \left.\frac{\partial U_N(x)}{\partial x}\right|_{x=r_N} \right]$, $\boldsymbol{r} = [r_1\ r_2 \cdots r_N]^T = \mathcal{R}(\boldsymbol{P})$, and $\odot$ represents the Hadamard product [33].

Proof : Since the capacity region is convex, it can be described by infinite inequality constraints

$$C\left(\boldsymbol{P}^{max}\right)=\bigcap_{\left\{\boldsymbol{\mu}^i\right\}}\left\{\boldsymbol{r}\mid \boldsymbol{\mu}^i\boldsymbol{r}\le\boldsymbol{\mu}^i\boldsymbol{r}_{\boldsymbol{\mu}^i}\right\},\ \ \forall\boldsymbol{\mu}^i\in R_+^N\ \ and\ \left\|\boldsymbol{\mu}^i\right\|_1=1, \tag{14}$$

where $\boldsymbol{r}_{\boldsymbol{\mu}^i}=\arg\max_{\boldsymbol{r}\in C\left(\boldsymbol{P}^{max}\right)}\boldsymbol{\mu}^i\boldsymbol{r},$. Form the Lagrangian of (5) as follows:

$$\mathcal{L}\left(\boldsymbol{r},v\right)=\boldsymbol{\beta}\left(\boldsymbol{P}\right)\boldsymbol{u}+\sum_{i=1}^{\infty}v_i\left(\boldsymbol{\mu}^i\boldsymbol{r}_{\boldsymbol{\mu}^i}-\boldsymbol{\mu}^i\boldsymbol{r}\right), \tag{15}$$

in which $\boldsymbol{u}=\mathcal{U}\left(\boldsymbol{P}\right)=\boldsymbol{U}\left(\mathcal{R}\left(\boldsymbol{P}\right)\right)=\boldsymbol{U}\left(\boldsymbol{r}\right)$ and $v_i\ge 0,\ i=1,2,\cdots$. Note that in general, for the Shannon capacity region of the Gaussian multiple access channel with ISI, $\boldsymbol{r}_{\boldsymbol{\mu}}\ne\boldsymbol{r}_{\boldsymbol{\mu}'}$, if $\boldsymbol{\mu}\ne\boldsymbol{\mu}'$. By using the Karush-Kuhn-Tucker (KKT) condition, we take the derivative of (15) with respect to $\boldsymbol{r}$. At the optimum, only one inequality constraint in (14) holds with equality. We denote that active constraint to be $\boldsymbol{\mu}^{opt}\boldsymbol{r}\le\boldsymbol{\mu}^{opt}\boldsymbol{r}_{\boldsymbol{\mu}^{opt}}$. According to complementary slackness [32], $v_{opt}>0$, and

$$\boldsymbol{\beta}\left(\boldsymbol{P}\right)\odot\boldsymbol{\lambda}\left(\boldsymbol{P}\right)=v_{opt}\boldsymbol{\mu}^{opt}, \tag{16}$$

in which $\boldsymbol{\mu}^{opt}=\boldsymbol{\mu}\left(\boldsymbol{P}\right)=\arg\max_{\boldsymbol{\mu}\in R_+^N\ and\ \left\|\boldsymbol{\mu}\right\|_1=1}\boldsymbol{\mu}\mathcal{R}\left(\boldsymbol{P}\right)$. Note that (16) is identical to (13), because

$$\frac{\boldsymbol{\beta}\left(\boldsymbol{P}\right)\odot\boldsymbol{\lambda}\left(\boldsymbol{P}\right)}{\left\|\boldsymbol{\beta}\left(\boldsymbol{P}\right)\odot\boldsymbol{\lambda}\left(\boldsymbol{P}\right)\right\|_1}=\frac{v_{opt}\boldsymbol{\mu}^{opt}}{\left\|v_{opt}\boldsymbol{\mu}^{opt}\right\|_1}=\frac{\boldsymbol{\mu}^{opt}}{\left\|\boldsymbol{\mu}^{opt}\right\|_1}=\boldsymbol{\mu}\left(\boldsymbol{P}\right), \tag{17}$$

where $\boldsymbol{\mu}\left(\boldsymbol{P}\right)$ and $\boldsymbol{\beta}\left(\boldsymbol{P}\right)$ are the normal vectors to the tangent hyperplanes at the boundary points of the capacity and utility region respectively, and $\boldsymbol{\lambda}\left(\boldsymbol{P}\right)$ consists of the first order derivatives of the utility-rate functions. Therefore, this equality gives the optimality condition for linking the boundary points of capacity region and utility region. ■

By the equality in (13), we can project the normal vector to the tangent hyperplane at each boundary point from the capacity region $C\left(\boldsymbol{P}^{max}\right)$ to the utility region $U\left(\boldsymbol{P}^{max}\right)$. This mapping process is illustrated by the right pointing arrow in Figure 3 for the two-user case. For $\forall\boldsymbol{\mu}\in R_+^N$ and $\left\|\boldsymbol{\mu}\right\|_1=1$, we solve the problem (12), get the boundary point $\boldsymbol{r}$, and subsequently use (13) to calculate the normal vector $\boldsymbol{\beta}$. The right pointing arrow provides a possible solution of the problem in (11), that is, enumerate all the possible $\boldsymbol{\mu}$ until the normal vector after mapping coincides with the original $\boldsymbol{\beta}$ in (11). This solution is impractical because there are infinite possible choices in $\boldsymbol{\mu}$. As illustrated by the left pointing arrow in Figure 3, if $\boldsymbol{\beta}$ is given, we are more interested in how to search $\boldsymbol{\mu}$ until the optimality condition in (13) holds.

The problem is converted into how to find the power allocation $\boldsymbol{P}$ that satisfies the optimality condition for a given $\boldsymbol{\beta}$. As shown in Figure 4, suppose that we randomly choose a rate vector $\boldsymbol{r} = [r_1\ r_2 \cdots r_N]^T$ on the capacity boundary, there will be a corresponding vector $\boldsymbol{\lambda}'$, which satisfies $\boldsymbol{\lambda}' = \left[\left.\frac{\partial U_1(x)}{\partial x}\right|_{x=r_1} \left.\frac{\partial U_2(x)}{\partial x}\right|_{x=r_2} \cdots \left.\frac{\partial U_N(x)}{\partial x}\right|_{x=r_N} \right]$. We can get a new weighted vector $\boldsymbol{\mu}'$ by the equality in (13). Then, we solve the problem in (12), and denote its solution as $\boldsymbol{r}'$. Generally, $\boldsymbol{r} \neq \boldsymbol{r}'$, otherwise the optimality condition holds. In the following, we will discuss how to search and find the optimal solution of (13) efficiently starting from the two-user case. However, if $N > 2$, the problem will be significantly more complicated. Based on the insights gained from the two-user case, we develop a low-complexity heuristic search algorithm to seek the optimum for the multiple-user case.

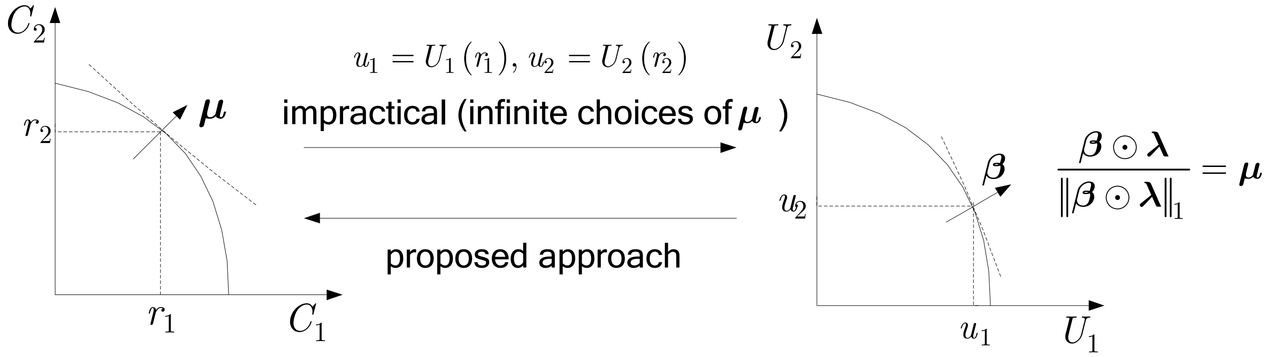

Figure 3. Optimality Condition for the Two-user Case

Below, we highlight several monotonic properties upon which the search algorithms of finding the power allocation $\boldsymbol{P}$, which satisfies (13), are based.

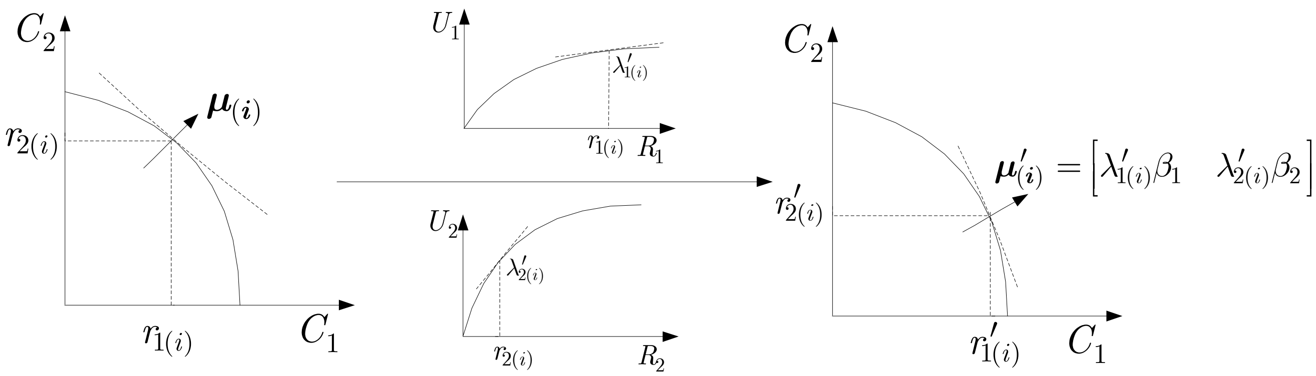

Figure 4. Bisection Search Algorithm for the Two-user case

Proposition 3: For any given $\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\beta} \succeq \boldsymbol{0}$ satisfying $\|\boldsymbol{\mu}\|_1 = \|\boldsymbol{\beta}\|_1 = 1$ and

$$\frac{\boldsymbol{\beta}\odot\boldsymbol{\lambda}}{\left\|\boldsymbol{\beta}\odot\boldsymbol{\lambda}\right\|_1}=\boldsymbol{\mu}\,,$$

where $\boldsymbol{\lambda}=\left[\left.\frac{\partial U_1(x)}{\partial x}\right|_{x=r_1}\ \left.\frac{\partial U_2(x)}{\partial x}\right|_{x=r_2}\ \cdots\left.\frac{\partial U_N(x)}{\partial x}\right|_{x=r_N}\right]$ and $\boldsymbol{r}=[r_1\ r_2\cdots r_N]^T$, if the mth component r_m of $\boldsymbol{r}$ is increased and the other components are held fixed, the mth component μ_m in $\boldsymbol{\mu}$ decreases.

Proof: The equality is identical to

$$\boldsymbol{\mu}=\frac{\left[\lambda_1\beta_1\quad\lambda_2\beta_2\quad\cdots\quad\lambda_N\beta_N\right]}{\left\|\left[\lambda_1\beta_1\quad\lambda_2\beta_2\quad\cdots\quad\lambda_N\beta_N\right]\right\|_1}\,.$$

Suppose we increase r_i for any i and fix all the other $r_j\ (j\neq i)$. Due to the monotonically increasing and concave property of the utility rate functions $U_i(r_i)$, λ_i will decrease and $\lambda_i>0$, while all the other $\lambda_j\ (j\neq i)$ remain fixed. Consequently, μ_i will monotonically decrease. ■

Proposition 4: Suppose $\boldsymbol{\mu}=[\mu_1\ \mu_2\cdots\mu_N]\in R_+^N$, $\|\boldsymbol{\mu}\|_1=1$, and $\boldsymbol{r}_{\boldsymbol{\mu}}=\max_{\boldsymbol{r}}\ \boldsymbol{\mu}\boldsymbol{r}$. For all m, if the mth component μ_m of the weighted vector $\boldsymbol{\mu}$ is increased and the other components are held fixed, the mth component of the rate vector $\boldsymbol{r}_{\boldsymbol{\mu}}$ remains the same or increases while all the other components of $\boldsymbol{r}_{\boldsymbol{\mu}}$ decreases.

Proof: See Lemma 6 in [34]. ■

C. *Algorithm for Describing the Two-User Utility Region*

Now we consider the two-user case. For the Shannon capacity region, we are able to solve (12) for arbitrary $\boldsymbol{\mu}=\left[\mu_1\ \mu_2\right]\in R_+^2$ and $\|\boldsymbol{\mu}\|_1=1$ and the utility functions of both users are available. Note that $\mu_1=1-\mu_2$, hence solving $\boldsymbol{\mu}$ is equivalent to finding μ_1. Proposition 3 and 4 enable us to use the bisection algorithm, which does not require full knowledge of the entire capacity boundary, to solve the problem in (11) efficiently.

As shown in Figure 4, suppose that the rate vector $\boldsymbol{r}=[r_1\ r_2\cdots r_N]^T$ is the maximizer of $\boldsymbol{\mu}\boldsymbol{r}$ in the capacity region. Then, we take the corresponding vector $\boldsymbol{\lambda}'=\left[\left.\frac{\partial U_1(x)}{\partial x}\right|_{x=r_1}\ \left.\frac{\partial U_2(x)}{\partial x}\right|_{x=r_2}\ \cdots\left.\frac{\partial U_N(x)}{\partial x}\right|_{x=r_N}\right]$ and get a new weighted vector $\boldsymbol{\mu}'$, by $\boldsymbol{\mu}'=\frac{\boldsymbol{\beta}\odot\boldsymbol{\lambda}'}{\left\|\boldsymbol{\beta}\odot\boldsymbol{\lambda}'\right\|}$. Denoting $\mu_1'=g(\mu_1)$ and $f(\mu_1)=\mu_1-g(\mu_1)$, the bisection search algorithm in describing the two-user utility region is summarized in Algorithm 1.

Proposition 5: Algorithm 1 converges to the unique solution of problem (11).

Proof: By the concavity of $U_i(r_i)$ and proposition 3 and 4, $g(\mu_1)$ is a monotonically decreasing function. Therefore, $f(\mu_1)$ is monotonic in μ_1. Moreover, for the convex utility region, there exist solutions for $f(\mu_1)=0$. The monotonicity of $f(\mu_1)$ guarantees that there exists a unique zero of the function $f \in \mathbb{R}[0,1]$. The convergence of the bisection search is guaranteed and we can use it to find the boundary point where the optimality condition in (13) holds [32]. ■

Algorithm 1: Two-user optimum search algorithm for the general multiple access strategy.

Input: $\boldsymbol{H}, \boldsymbol{P}^{max}, \boldsymbol{\beta}$; error tolerance ε; both users' utility rate models;

Initialization: Take $\boldsymbol{\mu}_{(0)}$ randomly which satisfies $\left\|\boldsymbol{\mu}_{(0)}\right\|_1 = 1$, $i=1$, flag = 0

Repeat:

1) Use Algorithm 5 in Appendix A to find the point on the capacity boundary which maximizes $\boldsymbol{\mu}_{(i)}\boldsymbol{r}$;

2) Denote the solution in 1) as $\boldsymbol{r}_{(i)} = \left[r_{1(i)}\ r_{2(i)}\right]^T$ and get the corresponding slopes $\lambda_{1(i)}, \lambda_{2(i)}$ on each utility rate curve;

3) Calculate $\boldsymbol{\mu}'_{(i)} = \left[\lambda_{1(i)}\beta_1 \quad \lambda_{2(i)}\beta_2\right] \Big/ \left\|\left[\lambda_{1(i)}\beta_1 \quad \lambda_{2(i)}\beta_2\right]\right\|_1$ and $f\left(\mu_{1(i)}\right) = \mu_{1(i)} - g\left(\mu_{1(i)}\right) = \mu_{1(i)} - \mu'_{1(i)}$;

4) If flag = 0

$a_{(0)} = \mu_{1(0)}, b_{(0)} = \mu'_{1(0)}$, flag = 1;

Else

If $f\left(a_{(i)}\right) \cdot f\left(\mu_{1(i)}\right) < 0$, $\mu_{1(i+1)} = \left(a_{(i)} + \mu_{1(i)}\right)\big/2, a_{(i+1)} = a_{(i)}, b_{(i+1)} = \mu_{1(i)}$;

If $f\left(a_{(i)}\right) \cdot f\left(\mu_{1(i)}\right) > 0$, $\mu_{1(i+1)} = \left(\mu_{1(i)} + b_{(i)}\right)\big/2, a_{(i+1)} = \mu_{1(i)}, b_{(i+1)} = b_{(i)}$;

5) $\mu_{2(i+1)} = 1 - \mu_{1(i+1)}$, $\boldsymbol{\mu}_{(i+1)} = \left[\mu_{1(i+1)}\ \mu_{2(i+1)}\right]$, $i = i+1$;

Until: $\left|f\left(\mu_{1(i)}\right)\right| < \varepsilon$

Return: $\boldsymbol{r}_{(i)}$ and its corresponding power allocation $\boldsymbol{P}$

Note that we choose bisection algorithm, because no closed form expression exists in general for $f(\mu_1)$. Within the i th iteration in Algorithm 1, $\mu_{1(i)}$ lies in the interval $\left[a_{(i)}, b_{(i)}\right]$ and $r_{1(i)}, r_{2(i)}, r'_{1(i)}, r'_{2(i)}$ are on the Shannon capacity boundary. The monotonic properties in proposition 3 and 4 ensure that both $\left|\mu_{1(i)} - \mu'_{1(i)}\right|$ and $\left|\mu_{2(i)} - \mu'_{2(i)}\right|$ decrease after each iteration until the optimality condition holds. The number of iterations for this bisection search is upper bounded by $\log_2\lceil 1/\varepsilon\rceil$.

D. *Low-Complexity Heuristic Algorithm for Describing the Multiple-User Utility Region*

Since proposition 3 and 4 only guarantee component-wise monotonicity, Algorithm 1 cannot be extended to the multiple-user case directly. In the multiple-user case, the only way to find the optimum is to characterize the entire capacity region $C(\boldsymbol{P}^{max})$ first, solve the utility-maximization problem subject to the constraint of this capacity region, and find the optimal rate vector. This algorithm is impractical, because the boundary of $C(\boldsymbol{P}^{max})$ consists of infinite number of points and usually lacks closed form expression. In this subsection, we will discuss how to attempt to find the optimum by developing a heuristic search algorithm inspired by the intuition gained from the two-user search algorithm.

In the two-user case, both $\left|r_{1(i)} - r'_{1(i)}\right|$ and $\left|r_{2(i)} - r'_{2(i)}\right|$ decrease after each iteration. Intuitively, in the multiple-user case, we should update $\boldsymbol{\mu}$ so that $\left|r_{n(i)} - r'_{n(i)}\right|$ decreases for $n = 1,2,...,N$. Similar to Algorithm 1, we still calculate $\boldsymbol{\mu}_{(i)}$ and $\boldsymbol{\mu}'_{(i)}$ within each iteration. Here, we denote the maximizers of $\boldsymbol{\mu}_{(i)}\boldsymbol{r}$ and $\boldsymbol{\mu}'_{(i)}\boldsymbol{r}$ as $\boldsymbol{r}_{(i)}$ and $\boldsymbol{r}'_{(i)}$, and the m th components of $\boldsymbol{\mu}_{(i)}$ and $\boldsymbol{r}_{(i)}$ as $\mu_{m(i)}$ and $r_{m(i)}$. The basic idea of this N-user heuristic search algorithm is to partition the users into two groups according to the component-wise relationship between $\boldsymbol{r}_{(i)}$ and $\boldsymbol{r}'_{(i)}$, i.e., if $r_{m(i)} < r'_{m(i)}$, we put user i in group 1, otherwise group 2. For users in group 1, because $r_{m(i)} < r'_{m(i)}$, we should update $\boldsymbol{\mu}_{(i+1)}$ in order to cause $r_{m(i+1)}$ to increase. Similarly, we should decrease $r_{m(i+1)}$ for users in group 2. From proposition 4, intuitively, we can infer that if we increase $\mu_{m(i)}$, there is a high possibility that $r_{m(i)}$ also increases. Therefore, to update $\boldsymbol{\mu}_{(i+1)}$, we can simply partition the users into two groups mentioned above, keep the ratio of $\mu_{(i)}$ within each group and adjust the ratio between the two groups. The new $\boldsymbol{\mu}_{(i+1)}$ will cause $\left|r_{m(i)} - r'_{m(i)}\right|$ to decrease for all $m = 1,2,\cdots,N$ with large probability. We can repeat this procedure until the optimality condition in (13) holds. The low-complexity heuristic search for the general multiple access strategy in the multiple-user case is summarized in Algorithm 2.

It should be pointed out that the step size δ should be carefully chosen. The step size δ not only affects the rate of convergence, but also determines the accuracy that the algorithm can achieve. Large step sizes δ will have fast rates of convergence, but small step sizes δ will result in better achieved accuracy. Therefore, the step size δ could be chosen according to the specific requirement of convergence-rate and desired accuracy. The step size δ can also be updated adaptively at each iteration [35].

Algorithm 2: Multiple-user low-complexity heuristic search for the general multiple access strategy.

Input: $\boldsymbol{H}, \boldsymbol{P}^{max}, \boldsymbol{\beta}$; error tolerance ε ; step size δ ; all users' utility rate functions; maximum iteration number $I_{\max}$

Initialization: Take $\boldsymbol{\mu}_{(0)}$ randomly which satisfies $\left\|\boldsymbol{\mu}_{(0)}\right\|_1 = 1, i = 1$

Repeat:

1) Use Algorithm 5 in Appendix A to find the point on the capacity boundary which maximizes $\boldsymbol{\mu}_{(i)}\boldsymbol{r}$;

2) Denote the solution in 1) as $\boldsymbol{r}_{(i)} = [r_{1(i)}\ r_{2(i)} \cdots r_{N(i)}]^T$ and get the corresponding slopes $\boldsymbol{\lambda}_{(i)} = [\lambda_{1(i)}\ \lambda_{2(i)} \cdots \lambda_{N(i)}]$ on each utility rate curve;

3) Calculate $\boldsymbol{\mu}'_{(i)} = \boldsymbol{\beta} \odot \boldsymbol{\lambda}_{(i)} \Big/ \left\|\boldsymbol{\beta} \odot \boldsymbol{\lambda}_{(i)}\right\|_1$;

4) Use Algorithm 5 in Appendix A to find the point on the capacity boundary which maximizes $\boldsymbol{\mu}'_{(i)}\boldsymbol{r}$ and denote the solution as $\boldsymbol{r}'_{(i)} = [r'_{1(i)}\ r'_{2(i)} \cdots r'_{N(i)}]^T$;

5) Let $\boldsymbol{e}_{(i)} = \left[e_{1(i)}\ e_{2(i)} \cdots e_{N(i)}\right]^T$, where $e_{n(i)} = I\left(r_{n(i)} - r'_{n(i)} < 0\right), n = 1,2,\cdots,N$ and $I(\bullet)$ is the indicator function;

6) $\boldsymbol{\mu}_{(i+1)} = \dfrac{\boldsymbol{\mu}_{(i)} \odot (1+\delta)\boldsymbol{e}_{(i)} + \boldsymbol{\mu}_{(i)} \odot (1-\delta)\left(\mathbf{1} - \boldsymbol{e}_{(i)}\right)}{\left\|\boldsymbol{\mu}_{(i)} \odot (1+\delta)\boldsymbol{e}_{(i)} + \boldsymbol{\mu}_{(i)} \odot (1-\delta)\left(\mathbf{1} - \boldsymbol{e}_{(i)}\right)\right\|_1}$, where $\mathbf{1} = [1\,1 \cdots 1]$, $i = i + 1$;

Until: $\left|r'_{ni} - r_{ni}\right| < \varepsilon, n = 1,2,\cdots,N$ or $i = I_{\max}$

Return: $\boldsymbol{r}_{(i)}$ and its corresponding power allocation $\boldsymbol{P}$

The algorithms discussed in this section can also be applied to solve other utility-fair resource allocation problems. For example, in [30], the authors adopt the Kalai-Smorodinsky bargaining solution (KSBS) when performing the resource allocation in 802.11 Wireless LAN. The wireless resources are allocated so that in the application-specific utility domain the achieved utility of every participating station incurs the same quality penalty. By adjusting the weighted vector, our algorithm can provide an efficient solution to find an optimum in the context of KSBS bargaining without the need of calculating all the boundary points of utility region.

V. Optimal Resource Allocation For The FDMA Strategy

In the Gaussian multiple access channel, successive decoding is indispensable to achieve the boundary of capacity region. However, it will greatly increase the complexity of the receiver. Although the FDMA strategy is not optimal in the information theoretical sense, the frequency-division multiple access technique is often desirable from a practical implementation point of view. A FDMA transmission scheme allows different users to occupy orthogonal dimensions, so they can be separated at the receiver without joint decoding. This scheme can utilize frequency multi-user diversity and lead to good spectral efficiency. In this section, we examine how to optimize the resource allocation in order to maximize the weighted sum of the utilities.

If multiple users access the system in the FDMA fashion, the channel capacity at jth sub-carrier for user i is determined by

$$C_{ij} = \omega_{ij} B \log\left\{1 + \frac{P_{ij}\left|H_{ij}\right|^2}{N_0 \omega_{ij} B}\right\}. \tag{18}$$

The user i's achievable rate r_i cannot exceed its capacity, which leads to

$$r_i \le C_i = \sum_{j=1}^{K} C_{ij}\ . \tag{19}$$

From an information theoretic point of view, the channel capacity defines the maximum achievable rate and hence, it determines the corresponding maximum video quality. Therefore, the resource allocation problem can be formulated as

$$\begin{array}{ll} \max\limits_{\omega_{ij}, P_{ij}} & \sum\limits_{i=1}^{N} \beta_i \log \dfrac{(C_i - R_{0i})}{D_{0i}(C_i - R_{0i}) + c_i} \\ s.t. & \sum\limits_{j=1}^{K} P_{ij} \le P_i^{\max} \quad ,\forall i \\ & P_{ij} \ge 0 \qquad\quad ,\forall i,j \\ & \sum\limits_{i=1}^{N} \omega_{ij} \le 1 \qquad ,\forall j \\ & \omega_{ij} = 0 \text{ or } 1 \quad ,\forall i,j \end{array} , \tag{20}$$

in which C_i satisfies[2]

$$C_i = \sum_{j=1}^{K} \omega_{ij} B \log\left\{1 + \frac{P_{ij}\left|H_{ij}\right|^2}{N_0 \omega_{ij} B}\right\}. \tag{21}$$

Unfortunately, this power and sub-carrier assignment problem belongs to the class of integer programming problem, for which an exact solution usually requires an exhaustive search. However, this becomes unacceptable when the number of sub-carriers and users is large because it is generally computationally prohibitive. In the following subsection, we will discuss a feasible approach using continuous relaxation, which makes the problem tractable while

[2]The rate models in (21) can be modified to approximate the real systems by adding a SNR-gap term [31], which defines the gap between a practical coding and modulation scheme and the channel capacity, into the expression of the Shannon capacity.

achieving a near-optimal performance.

A. *Continuous Relaxation*

Fortunately, we can approximate the original integer programming problem in (20) by its continuous relaxation. Instead of forcing the original optimization variable to be either 0 or 1, the last constraint in (20) can be relaxed to

$$0 \le \omega_{ij} \le 1\,. \tag{22}$$

Mathematically, the continuous relaxation of the original problem in (20) can be posed as follows:

$$\begin{aligned} \max_{\omega_{ij},P_{ij}} & \sum_{i=1}^{N} \beta_i \log \frac{(C_i - R_{0i})}{D_{0i}(C_i - R_{0i}) + c_i} \\ s.t. \quad & \sum_{j=1}^{K} P_{ij} \le P_i^{\max} \quad , \forall i \\ & \sum_{i=1}^{N} \omega_{ij} \le 1 \qquad , \forall j \\ & P_{ij} \ge 0 \qquad\quad , \forall i,j \\ & \omega_{ij} \ge 0 \qquad\quad , \forall i,j \end{aligned} \quad , \tag{23}$$

in which $C_i = \sum_{j=1}^{K} \omega_{ij} B \log\left\{1 + \frac{P_{ij} H_{ij}^2}{N_0 \omega_{ij} B}\right\}$.

Lemma 3*:* The objective function in the optimization problem in (23) is a concave function in (P_{ij}, ω_{ij}).

Proof: In [15], it has been shown that C_i is a two-dimensional concave function in (P_{ij}, ω_{ij}). Note that the utility functions $U_i(r_i)$ are concave and monotonically increasing. By the property of operations which preserve convexity [32], the objective function here is still concave. ■

The constraint set in (23) is convex because the constraints in the optimization problem are linear. Hence, the essence of the problem is to maximize a concave function subject to a convex constraint. Several numerical search algorithms exist to obtain solutions efficiently [32]. However, these general numerical algorithms do not shed much insight on these resource allocation problems. Instead, we explore the specific problem structure in order to lead to intuitions on the structure of the optimal solution that cannot be gained from purely numerical methods. In the following, we again start with the two-user case, discuss its near-optimal solution, and gain intuition that can be used for a low-complexity heuristic solution for the multiple-user case.

B. *Near-Optimal Solution for the Two-User Case*

If $N=2$, the Lagrangian function of (23) can be written as a function of ω_{ij}, P_{ij}

$$\begin{aligned}\mathcal{L}\left(\omega_{ij}, P_{ij}\right) &= \sum_{i=1}^{2}\beta_i \log\frac{\left(C_i - R_{0i}\right)}{D_{0i}\left(C_i - R_{0i}\right)+c_i} + \sum_{j=1}^{K}\kappa_j^1\left(\sum_{i=1}^{2}\omega_{ij}-1\right) \\ &+\sum_{i=1}^{2}\kappa_i^2\left(\sum_{j=1}^{K}P_{ij}-P_i^{\max}\right)-\sum_{i=1}^{2}\sum_{j=1}^{K}\nu_{ij}^1 P_{ij}-\sum_{i=1}^{2}\sum_{j=1}^{K}\nu_{ij}^2\omega_{ij}\end{aligned}, \tag{24}$$

where $\kappa_j^1, \kappa_i^2, \nu_{ij}^1$, and ν_{ij}^2 are Lagrangian multipliers. Taking the derivative with respect to P_{ij} gives the KKT condition corresponding to the usual waterfilling solution, which means there exist positive constants K_i, such that for all $i=1,2$ and for all $j=1,2,\cdots,K$, if $P_{ij}>0$, then:

$$\frac{P_{ij}}{\omega_{ij}}+\frac{1}{g_{ij}} = K_i = -\frac{\beta_i B}{\kappa_i^2}\cdot\frac{\partial}{\partial C_i}\left[\log\frac{\left(C_i - R_{0i}\right)}{D_{0i}\left(C_i - R_{0i}\right)+c_i}\right], \tag{25}$$

and if $P_{ij}=0$, then

$$\frac{1}{g_{ij}} \geq K_i, \tag{26}$$

in which $g_{ij}=\dfrac{H_{ij}^2}{N_0 B}$ and the water-level K_i is actually a function of β_i and C_i. By using the KKT condition, we also take the derivative of (24) with respect to ω_{ij}, and have for all $j=1,2,\cdots,K$, if $\omega_{1j}>0$ and $\omega_{2j}>0$, i.e., the *j*-th sub-carrier is by both users, then

$$\beta_1\frac{\log\left\{1+\dfrac{P_{1j}g_{1j}}{\omega_{1j}}\right\}-\dfrac{\dfrac{P_{1j}g_{1j}}{\omega_{1j}}}{1+\dfrac{P_{1j}g_{1j}}{\omega_{1j}}}}{\left(C_1-R_{01}\right)\left[D_{01}\left(C_1-R_{01}\right)+c_1\right]} = \beta_2\frac{\log\left\{1+\dfrac{P_{2j}g_{2j}}{\omega_{2j}}\right\}-\dfrac{\dfrac{P_{2j}g_{2j}}{\omega_{2j}}}{1+\dfrac{P_{2j}g_{2j}}{\omega_{2j}}}}{\left(C_2-R_{02}\right)\left[D_{02}\left(C_2-R_{02}\right)+c_2\right]}. \tag{27}$$

Denoting $\rho_i = \dfrac{\beta_i}{\left(C_i-R_{0i}\right)\left[D_{0i}\left(C_i-R_{0i}\right)+c_i\right]}$, we have

$$\rho_1\left\{\log\left(g_{1j}K_1\right)+\frac{1}{g_{1j}K_1}-1\right\} = \rho_2\left\{\log\left(g_{2j}K_2\right)+\frac{1}{g_{2j}K_2}-1\right\}. \tag{28}$$

The equality is satisfied only if that sub-carrier is shared. Multimedia applications generally require high data rates to achieve reasonable quality. Therefore, the system usually operates at high SNR, and $\dfrac{1}{g_{ij}K_i}$ on either side of (28) approaches 0. We take the difference between the left-hand side and right-hand side as a function f of $g_{1j}, g_{2j}, \rho_1, \rho_2$,

$$f\left(\frac{g_{1j}^{\rho_1}}{g_{2j}^{\rho_2}}\right) \approx \log_2\left(\frac{g_{1j}^{\rho_1}}{g_{2j}^{\rho_2}}\right) + \log_2\left(\frac{K_1^{\rho_1}}{K_2^{\rho_2}}\right) + \rho_2 - \rho_1. \tag{29}$$

Similarly with the Theorem 2 in [15], the optimal frequency partition that maximizes $\beta_1 U_1 + \beta_2 U_2$ consists of two contiguous frequency bands with user 1 using the lower frequency sub-carriers and user 2 using the higher frequency sub-carriers. Assume we pre-arrange the index of sub-carriers to make $g_{1j}^{\rho_1} / g_{2j}^{\rho_2}$ decreasing in *j*. Since for any fixed ρ_1 and ρ_2, $g_{1j}^{\rho_1} / g_{2j}^{\rho_2}$ decreases in j, we are able to decide the optimal sub-carrier allocation by checking whether or not $f\left(g_{1j}^{\rho_1} / g_{2j}^{\rho_2}\right) > 0$. If $f\left(g_{1j}^{\rho_1} / g_{2j}^{\rho_2}\right) > 0$, that sub-carrier will be assigned to user 1, otherwise it should be assigned to user 2.

We develop an iterative search algorithm to find the near-optimal resource allocation strategy for the two-user case. The basic idea is to update sub-carrier allocations, ρ_1 and ρ_2 repeatedly. Within each iteration, we fix ρ_i and use the optimality condition to achieve the desired point in the current achievable utility region. Then, we update ρ_i according to the new sub-carrier allocation. The iterative algorithm converges when (27) is satisfied. The proposed algorithm is summarized in Algorithm 3.

Algorithm 3 Two-user resource allocation algorithm for the FDMA strategy.

Input: $\boldsymbol{H}, \boldsymbol{P}^{max}, \boldsymbol{\beta}$

Initialization: Make an initial sub-carrier allocation so that $C_1 > R_{01}$ and $C_2 > R_{02}$, calculate ρ_1 and ρ_2

Repeat:

(1) Sort the sub-carriers according to $g_{1j}^{\rho_1} / g_{2j}^{\rho_2}$ from the largest to the smallest.

(2) For *j*=0,…,*K*

water-fill for user 1 using sub-carrier 1 to *j*

water-fill for user 2 using sub-carrier *j*+1 to *K*

(3) Choose the frequency partition boundary to be the one that maximizes $\sum_{i=1}^{2} \beta_i \log \frac{(C_i - R_{0i})}{D_{0i}(C_i - R_{0i}) + c_i}$

(4) Update ρ_i according to $\rho_i = \frac{\beta_i}{(C_i - R_{0i})\left[D_{0i}(C_i - R_{0i}) + c_i\right]}$

Until: No improvement can be achieved in step (3)

Return: Sub-carrier assignment $\boldsymbol{\omega}$ and power allocation $\boldsymbol{P}$

The difference between the proposed sub-carrier allocation and the continuous-relaxed optimal sub-carrier allocation is that only one sub-carrier is allocated differently. More specifically, the optimum of continuous-relaxed problem indicates that the sub-carrier satisfying (27) should be divided into two smaller bins and allocated to each user separately, and our proposed algorithm assign it to only one user. Since the continuous-relaxation provides an upper bound of the optimum of original integer programming problem, the proposed algorithm can achieve near optimal performance.

C. Low-Complexity Heuristic Solution for the Multiple-User Case

If the number of users is larger than two, the continuous relaxation optimization problem in (23) can only be solved numerically. However, as stated before, pure numerical algorithms generally do not offer much insight into the original problem. Besides, most existing research works focus on solving the resource allocation problem in OFDM system in a centralized fashion [13]-[14], in which the computational complexity are generally high with an increasing number of sub-carriers and users. Note that in the case of multiple users, the original problem could be viewed as a composition of many two-user sub-problems. Without loss of generality, assuming N is even, the problem in (5) is identical to

$$\begin{aligned}
&\max_{\omega_{ij},P_{ij}} \sum_{i=1}^{N} \beta_i U_i\left(r_i\right) = \max_{K_i} \sum_{i=0}^{N/2-1} \max_{\omega_{jk},P_{jk}} \sum_{j=1}^{2} \beta_{2i+j} U_{2i+j}\left(r_{2i+j}\right) \\
&s.t. \quad \sum_{k=1}^{K} P_{jk} \le P_j^{\max} \quad ,\forall j \\
&\qquad\ P_{jk} \ge 0 \quad ,\forall j,k \\
&\qquad \sum_{j=1}^{2} \omega_{jk} \le 1 \quad ,\forall k \in K_i \\
&\qquad\ \omega_{jk} = 0 \text{ or } 1 \quad ,\forall k \in K_i \\
&\qquad\ \omega_{jk} = 0 \quad ,\forall j,\ if\ k \notin K_i
\end{aligned} \quad , \tag{30}$$

where K_i is a set of the indices of the sub-carriers that are allocated to user $2i$ and $2i+1$, $\bigcup_{i=0}^{\frac{N}{2}-1} K_i = \{1,2,\cdots,K\}$ and $\bigcap_{i=0}^{\frac{N}{2}-1} K_i = \varnothing$.

Because the sub-carriers are allocated orthogonally among different users in FDMA systems, the optimal solution of the original problem in (5) also achieves the optimal solution of each two-user sub-problem. Therefore, we can decompose the multi-user sub-carrier allocation problem into multiple two-user resource allocation problems and apply Algorithm 3 to solve each sub-problem. Here, we propose a method based on the criteria of "serving the highest demand first". We define the "demand value" D_i for each user as

$$D_i = \beta_i \left.\frac{\partial U_i(x)}{\partial x}\right|_{x=r_i} \tag{31}$$

Algorithm 4 N -user "serving the highest demand first" allocation algorithm for the FDMA strategy

Input: $\boldsymbol{H}, \boldsymbol{P}^{max}, \boldsymbol{\beta}$

Initialization: Make an initial sub-carrier allocation, sort $\beta_i \frac{\partial U_i(x)}{\partial x}\Big|_{x=r_i}$ in the descending order, and construct sub-carrier exchange matrix A with $A_{ij} = 0, \forall i, j$

Repeat:

Loop: For i =1,…, N

For j =1,…, N $(j \neq i)$

If $A_{ij} = 0$

Apply Algorithm 3 to re-allocate user i and j 's sub-carriers and calculate $\Delta\left\{\beta_i U_i(r_i) + \beta_j U_j(r_j)\right\}$

If $\Delta\left\{\beta_i U_i(r_i) + \beta_j U_j(r_j)\right\} > 0$

(1) Set $A_{ik} = A_{jk} = A_{ki} = A_{kj} = 0, \forall k$

(2) Re-sort the "demand values" $\beta_i \frac{\partial U_i(x)}{\partial x}\Big|_{x=r_i}$ in the descending order

(3) Adjust the column and row of A correspondingly

(4) Go to Loop;

Else

Set $A_{ij} = A_{ji} = 1$

Until: $i = j = N \text{ and } A_{mn} = 1, \forall m, n$

Return: Sub-carrier assignment $\boldsymbol{\omega}$ and power allocation $\boldsymbol{P}$

The basic idea of "serving the highest demand first" is to allow the user with the largest value of D_i to negotiate its sub-carriers allocation with other users first. If the user with the highest demand cannot improve the weighted sum of the utilities by negotiating with other users, we consider the user with the second highest demand and let it negotiate with the other users, and so forth. If any exchange of sub-carrier between two users is made, we re-calculate the new value of D_i for those two users who exchanged their sub-carriers, re-sort their demand values, and restart the process of "serving the highest demand first". We repeat this procedure iteratively until no further improvement can be made. This algorithm will results in a "local optimum" in the sense that the performance cannot be further improved by exchanging resources between any two users. Compared with the optimal algorithm of which the computational complexity is N^K, the overall complexity for each iteration of our proposed scheme is at most $(N-1)^2 K \log_2 K$ [12].

To further reduce the complexity of the algorithm, we introduce the sub-carrier exchange matrix A, whose entry

A_{ij} indicates the history whether user i and j have negotiated with each other under the current sub-carrier allocation pattern. If these two users have negotiated before under the same sub-carrier allocation pattern, it means that no improvement can be made, thus we do not need to re-allocate sub-carriers between them. As soon as any exchange of sub-carrier is made between two users, the corresponding entries of those two users will be set to zero. The algorithm of "serving the highest demand first" is summarized in Algorithm 4.

VI. SIMULATION RESULTS

The performances of the proposed algorithms are examined in this section. For the purpose of illustration, we consider a multi-carrier system with only 8 sub-carriers. We assume the bandwidth of each sub-carrier is $B = 50$ kHz.

Now we consider the two-user case. The channel conditions of the sub-carriers for the two users are given in Table I. We choose the weighted vector $\boldsymbol{\beta} = \left[0.9\ 0.1\right]$ for illustration and $P_1^{\max} = 10^{1.2}, P_2^{\max} = 10^2$. The parameter values for the utility-rate function deployed for these experiments are determined based on a state-of-the-art wavelet video coder [36]. In this case, we assume that user 1 wants to transmit the *Mobile* video (CIF, 15Hz) with $D_{01} = 1$, $R_{01} = 44.04$ kbps, and $c_1 = 38230$ kbps, while user 2 has the the *Foreman* video (CIF, 15Hz) for transmission with $D_{02} = 1$, $R_{02} = 20.72$ kbps, and $c_2 = 2760$ kbps. R_{0i}, D_{0i}, c_i are the parameters of the utility-rate model in (3).

TABLE I

CHANNEL CONDITION OF A TWO-USER SYSTEM ($N_0 B = 1$)

User	$\lvert H_{i1}\rvert^2$	$\lvert H_{i2}\rvert^2$	$\lvert H_{i3}\rvert^2$	$\lvert H_{i4}\rvert^2$	$\lvert H_{i5}\rvert^2$	$\lvert H_{i6}\rvert^2$	$\lvert H_{i7}\rvert^2$	$\lvert H_{i8}\rvert^2$
User 1 $(i = 1)$	0.5718	1.4196	0.0466	1.3392	1.3138	2.3280	0.4179	2.2805
User 2 $(i = 2)$	1.4406	1.3182	0.5150	0.6160	0.1048	0.0625	0.4122	1.0255

First, we simulate the general multiple access strategy. As shown in Figure 5, by varying $\boldsymbol{\mu}$ and solving the problem in (12) via the iterative water-filling algorithm, we can obtain a series of rate vectors and trace out the corresponding Shannon capacity region.

We apply Algorithm 1 to maximize $\boldsymbol{\beta u}$ and examine its convergence. As shown in Table II, Algorithm 1 converges after around 7 iterations. The rate vector which maximizes $\boldsymbol{\beta u}$ is around [685.23 kbps 786.87 kbps]^T and it can be verified that it approximately satisfies the optimality condition in proposition 2. The optimal power allocation is given in Table III. Under this power allocation, user 1's average PSNR is 30.5929dB, user 2's average PSNR is 41.5009dB and the weighted sum of PSNRs is 31.6837dB. As opposed to our algorithm, conventional sum-rate-maximizing approach that does not consider the video characteristics always chooses to maximize $r_1 + r_2$ in the two-user system. The allocation outcome is that user 1 experiences an average PSNR of 28.1790dB, user 2 experiences an average PSNR of 42.7704dB, and the weighted sum of PSNRs is 29.6382dB. For user 1, the sum-rate-maximizing approach will result in an unacceptable video quality below 30dB.

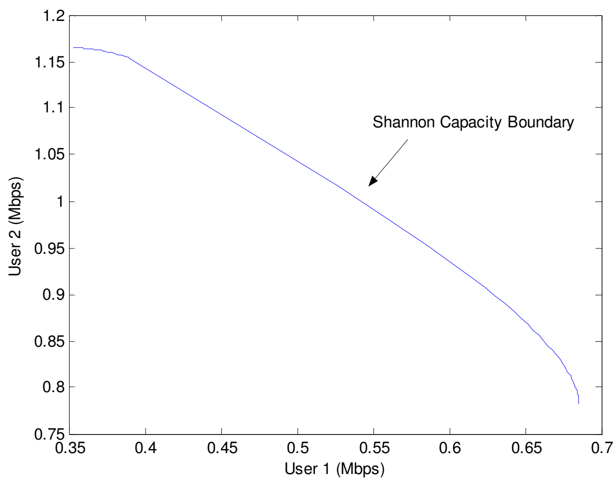

Figure 5 The Shannon Capacity Region of a Two-User System

Next, we consider the scenario of FDMA and still assume the same channel condition in Table I. We apply Algorithm 3 to derive the near-optimal power allocation and the solution is given in Table III. In the optimal power allocation scheme, users are allocated the sub-carriers at which they experience good channel condition and both of them water-fill their power across the sub-carriers assigned to them. Under this power allocation, user 1's average PSNR is 30.5941dB, user 2's average PSNR is 37.5560dB and the weighted sum of PSNRs is 31.2903dB.

As expected, under the same channel conditions, the weighted sum of the utilities in general multiple access strategy is larger than the FDMA strategy because the former one is optimal from the information theoretic view. Specifically, in this example, the weighted sum of PSNRs for the general multiple access strategy is only 0.3934dB larger than the FDMA strategy. Since the general multiple access strategy provides an upper bound of the optimal performance of FDMA strategy, the resource allocation scheme provided by Algorithm 3 can achieve near-optimal performance.

In the following, we consider the multiple-user case with $N = 3$ users. The channel conditions of the sub-carriers for the three users are given in Table IV. We choose the weighted vector $\boldsymbol{\beta} = \left[0.3\ 0.3\ 0.4\right]$ for illustration and $P_1^{\max} = 10, P_2^{\max} = 10^{1.5}, P_3^{\max} = 10^2$. We assume that user 1 wants to transmit the *Foreman* video (CIF, 15Hz) with

$D_{01} = 1$, $R_{01} = 20.72$ kbps, and $c_1 = 2760$ kbps, user 2 has the *Coastguard* video (CIF, 30Hz) for transmission with $D_{02} = 4.3$, $R_{02} = 0$ kbps, and $c_2 = 6329.7$ kbps, while user 3 wants to transmit the *Foreman* video (CIF, 30Hz) with $D_{03} = 3$, $R_{03} = 55.08$ kbps, and $c_3 = 4610$ kbps.

TABLE II

AN EXAMPLE OF ALGORITHM 1

Iteration	$\boldsymbol{\mu}_{(i)}$	$\boldsymbol{r}_{(i)}^{T}$ (kbps)	$\boldsymbol{\mu}'_{(i)}$	$\boldsymbol{r}'^{T}_{(i)}$ (kbps)	$[a_{(i)}\ b_{(i)}]$
$i = 1$	[0.5391 0.4609]	[601.39 934.16]	[0.9509 0.0491]	[685.33 785.25]	[0.5391 0.9509]
$i = 2$	[0.7450 0.2550]	[679.49 813.06]	[0.9342 0.0658]	[685.24 786.64]	[0.7450 0.9509]
$i = 3$	[0.8479 0.1521]	[684.10 795.66]	[0.9321 0.0679]	[685.23 786.82]	[0.8479 0.9509]
$i = 4$	[0.8994 0.1006]	[684.95 789.87]	[0.9314 0.0686]	[685.22 786.88]	[0.8994 0.9509]
$i = 5$	[0.9251 0.0749]	[685.18 787.44]	[0.9312 0.0688]	[685.22 786.91]	[0.9251 0.9509]
$i = 6$	[0.9380 0.0620]	[685.26 786.32]	[0.9310 0.0690]	[685.22 786.92]	[0.9251 0.9380]
$i = 7$	[0.9316 0.0684]	[685.23 786.87]	[0.9311 0.0689]	[685.22 786.91]	[0.9251 0.9316]

TABLE III

POWER ALLOCATION OF THE TWO-USER SYSTEM

MAC strategy	User	P_{i1}	P_{i2}	P_{i3}	P_{i4}	P_{i5}	P_{i6}	P_{i7}	P_{i8}
General MAC strategy	User 1 $(i = 1)$	1.4505	2.5221	0	2.5245	2.6773	3.0088	0.8308	2.8348
	User 2 $(i = 2)$	19.454	17.249	18.783	13.612	0	0	17.456	13.445
FDMA strategy	User 1 $(i = 1)$	0	3.0814	0	3.0392	3.0247	3.3563	0	3.3474
	User 2 $(i = 2)$	34.326	0	33.079	0	0	0	32.595	0

In the case of the general multiple access strategy, we apply Algorithm 2 to search the optimum. We set step size $\delta = 0.1$. As shown in Table V, in this case, Algorithm 2 converges after around 20 iterations. Based on the result after the 20$^{\text{th}}$ iteration, the rate vector which maximizes $\boldsymbol{\beta u}$ is approximately [586kbps 698.4kbps 937.1kbps]$^{\text{T}}$. By choosing smaller step size δ, we can reach a more accurate solution at the expense of decreasing the speed of convergence. Under this power allocation, user 1's average PSNR is 40.4379dB, user 2's average PSNR is 36.8719dB, user 3's average PSNR is 39.1417dB, and the weighted sum of PSNRs is 38.8496dB. Conventional sum-rate-maximizing approach with $\boldsymbol{\mu} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$ will cause the weighted sum of PSNRs to be 36.2988dB, which is of nearly 3dB degradation.

TABLE IV

CHANNEL CONDITION OF A THREE-USER SYSTEM ($N_0B = 1$)

User	$\lvert H_{i1}\rvert^2$	$\lvert H_{i2}\rvert^2$	$\lvert H_{i3}\rvert^2$	$\lvert H_{i4}\rvert^2$	$\lvert H_{i5}\rvert^2$	$\lvert H_{i6}\rvert^2$	$\lvert H_{i7}\rvert^2$	$\lvert H_{i8}\rvert^2$
User 1 $(i=1)$	0.3758	4.1200	2.0213	1.7739	0.8033	1.7400	0.9460	0.2058
User 2 $(i=2)$	4.9634	7.3880	3.1333	0.9054	0.0961	4.4658	0.9182	0.4643
User 3 $(i=3)$	2.2758	0.3953	0.8065	1.6528	1.1393	0.4865	2.1461	3.4382

TABLE V

AN EXAMPLE OF ALGORITHM 2

iteration	$\boldsymbol{\mu}_{(i)}$	$\boldsymbol{r}_{(i)}^T$ (Mbps)	$\boldsymbol{\mu}'_{(i)}$	$\boldsymbol{r}_{(i)}'^T$ (Mbps)
$i=1$	[0.6000 0.2000 0.2000]	[0.5865 0.4967 1.1388]	[0.4024 0.4130 0.1845]	[0.2974 1.1308 0.7544]
$i=2$	[0.5976 0.2032 0.1992]	[0.5865 0.6845 0.9509]	[0.4404 0.2995 0.2601]	[0.5853 0.7081 0.9276]
$i=3$	[0.5951 0.2065 0.1984]	[0.5865 0.6886 0.9467]	[0.4408 0.2973 0.2619]	[0.5855 0.7055 0.9301]
$i=4$	[0.5927 0.2098 0.1976]	[0.5865 0.6914 0.9437]	[0.4410 0.2959 0.2632]	[0.5857 0.7036 0.9319]
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$i=20$	[0.5507 0.2381 0.2112]	[0.5864 0.6984 0.9363]	[0.4415 0.2922 0.2663]	[0.5860 0.6990 0.9366]

Now consider the FDMA strategy under the same channel condition. We apply Algorithm 4 of "serving the highest demand first" to allocate sub-carriers and power among the three users. Our proposed scheme has the complexity of $(N-1)^2 K \log_2 K = 96$, which is only 1.46% of the complexity of optimal algorithm in which $N^K = 6561$. With the increasing in the number of sub-carriers, e.g., in IEEE 802.11a, $K = 48$ [37], the reduction in complexity would be more substantial. The result is shown in Table VI. Similar to the two-user case, the users are allocated the sub-carriers at which they experience good channel conditions and all of them water-fill across the sub-carriers assigned to them. In this case, Algorithm 4 converges to a power allocation scheme with the achievable rate vector $\boldsymbol{r} =$ [480.08kbps 856.34kbps 1125kbps]$^{\mathrm{T}}$ and it achieves almost the same performance as the general multiple access strategy. Under this power allocation, user 1's average PSNR is 39.6746dB, user 2's average PSNR is 37.4521dB, user 3's average PSNR is 39.6195dB, and the weighted sum of PSNRs is 38.9858dB. Note that the weighted sum of PSNRs here is slightly larger than in the general multiple access strategy case above, because we can only provide the optimal solution approximately by Algorithm 2.

TABLE VI

POWER ALLOCATION OF THE THREE-USER SYSTEM

MAC strategy	User	P_{i1}	P_{i2}	P_{i3}	P_{i4}	P_{i5}	P_{i6}	P_{i7}	P_{i8}
General MAC strategy	User 1 $(i=1)$	0	2.077	1.921	1.995	1.167	1.581	1.259	0
	User 2 $(i=2)$	5.307	8.709	8.445	0	0	9.162	0	0
	User 3 $(i=3)$	11.683	0	0	20.772	21.818	0	22.498	23.229
FDMA strategy	User 1 $(i=1)$	0	0	0	0	0	10	0	0
	User 2 $(i=2)$	5.969	5.7852	5.9029	4.9998	5.0153	0	3.9507	0
	User 3 $(i=3)$	0	0	0	0	0	0	0	100

VII. CONCLUSIONS

In this paper, we address the problem of multi-user video transmission in multi-carrier wireless networks. Focusing on two types of MAC strategy, i.e., the general multiple access strategy and the FDMA strategy, we propose two approaches to maximize the weighted sum of video qualities of all the users'. For the general multiple access strategy, we propose a general procedure to determine the achievable utility region under the constraints of a given capacity region. In the FDMA scenario, we resort to continuous relaxation to seek near-optimal solutions analytically. For both MAC schemes, we first develop iterative search algorithms to find the optimal resource allocation strategies for the two-user case. Subsequently, inspired by the intuition gained from the two-user case, we extend them to the multiple-user case using low-complexity heuristic approaches. Numerical experiments show that all these algorithms achieve significant performance improvements by explicitly considering the video utility impact and the specific rate-distortion performance of the operational video coder deployed.

REFERENCES

[1] A. R. S. Bahai, B. R. Saltzberg, M. Ergen, *Multi-carrier digital communications: Theory and applications of OFDM*. New York: Springer, 1999.

[2] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE ICC*, vol. 1, pp. 331–335, June 1995.

[3] D. Tse and S. Hanly, "Multiaccess fading channels-part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Info. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[4] E. Yeh and A. Cohen, “Delay optimal rate allocation in multiaccess fading communications,” in *Proc. of the Allerton Conf.*, Monticello, IL, pp. 140–149, Sep.-Oct. 2004.

[5] ——, “Throughput optimal power and rate control for multiaccess and broadcast communications,” in *Proc. of the ISIT*, Chicago, IL, p.112, June 2004.

[6] E. Telatar and R. Gallager, “Combining queueing theory with information theory for multiaccess,” *IEEE J. Select. Areas Commun.*, vol. 13, pp. 963–969, Aug. 1995.

[7] S. Raj, E. Telatar, and D. Tse, “Job scheduling and multiple access,” *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, March 2003.

[8] Q. Liu, S. Zhou, and G. B. Giannakis, “Queuing with adaptive modulation and coding over wireless links: Cross-layer analysis and design,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1142–1153, May 2005.

[9] A. Scaglione and M. van der Schaar, “Cross-layer resource allocation for delay constrained wireless video transmission”, in *Proc. of ICASSP*, vol. 5, pp.909-912, Mar. 2005.

[10] C. Shen and M. van der Schaar, “Optimal Resource Allocation in Wireless Multiaccess Video Transmissions”, *ICC* 2007, to appear.

[11] J. Jang and K. B. Lee, “Transmit power adaptation for multiuser OFDM systems,” *IEEE J. Select. Areas Commun.*, vol. 21, pp.171-178, Feb. 2003.

[12] G.-C. Song and Y. (G.) Li, “Cross-layer optimization for OFDM wireless networks – Part I: theoretical framework,” *IEEE Trans. on Wireless Commun.*, vol. 4, no. 2, pp. 614 – 624, March 2005.

[13] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, “Multiuser OFDM with adaptive subcarrier, bit, and power allocation,” *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999.

[14] W. Rhee and J. M. Cioffi, “Increase in capacity of multiuser OFDM system using dynamic subchannel allocation,” in *Proc. IEEE VTC*, pp. 1085-1089, 2000.

[15] W. Yu and J. M. Cioffi, "FDMA Capacity of Gaussian Multiple access Channels with ISI", *IEEE Trans. on Commu.*, vol. 50, no.1, pp.102-111, Jan. 2002.

[16] K. Kim, Y. Han, and S.-L. Kim, "Joint subcarrier and power allocation in uplink OFDMA systems," *IEEE Commun. Letter*, vol.9, no.6, pp.526–528, 2005.

[17] J.R. Ohm, M. van der Schaar, J. Woods, “Interframe wavelet coding-Motion Picture Representation for Universal Scalability”, *EURASIP Signal Processing: Image Communication, Special issue on Digital Cinema*, 2004.

[18] FEDERAL COMMUNICATIONS COMMISSION (2003a), “Notice for Proposed Rulemaking (NPRM 03 322): Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies”, ET Docket No. 03 108. Dec 2003.

[19] J. Mitola, “Cognitive radio for flexible mobile multimedia communications,” *Mobile Networks and Applications*, vol. 6, Sep. 2001.

[20] T. Bostoen, R. Oehen, J. Verlinden, "Optimizing DSL for Multimedia Services", *Alcatel Telecommunications Review*, pp. 155-159, 2nd Quarter 2005

[21] L.J. Cimini and N.R. Sollenberger, “OFDM with Diversity and Coding for Advanced Cellular Internet Services,” in *Proc. of IEEE GLOBECOM*, pp. 305-309, Nov. 1997.

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[23] A. Albanese, M. Luby, “PET-priority encoding transmission,” *High-Speed Networking for Multimedia Application*, Kluwer Academic Publishers, 1996

[24] M. van der Schaar, D. S. Turaga, “Cross-layer packetization and retransmission strategies for delay-senstive wireless multimedia transmission,” *IEEE Trans. On Multimedia*, vol. 9, pp. 185-197, Jan 2007.

[25] M. Wang and M. van der Schaar, “Operational Rate-Distortion Modeling for Wavelet Video Coders, ” *IEEE Trans. on Signal Processing*, vol. 54, pp. 3505-3517, Sep 2006

[26] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, “Analysis of video transmission over lossy channels,” *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1012–1032, June 2000.

[27] A. Ortega and K. Ramchandran, “Rate-Distortion Methods for Image and Video,” *IEEE Signal Processing Mag.*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[28] W. Yu, “Competition and Cooperation in Multi-user Communication Environments”, Ph.D. Dissertation, Stanford University, Jun. 2002.

[29] R. S. Cheng and S. Verdu, “Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling,” *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 773–785, May 1993.

[30] H. Park and M. van der Schaar, "Bargaining Strategies for Networked Multimedia Resource Management," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3496-3511, July 2007.

[31] X. Qiu and K. Chawla, “On the performance of adaptive modulation in cellular systems,” *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 884–895, Jun. 1999.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[33] R. A. Horn, C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1994.

[34] G. Wunder and T. Michel, “Optimal Resource Allocation for OFDM Multiuser Channels”, http://arxiv.org/abs/cs/0608023

[35] D. P. Bertsekas, *Nonlinear programming*, 2nd Edition. Belmont, MA: Athena Scientific, 1999.

[36] Y. Andreopoulos, A. Munteanu, J. Barbarien, M. van der Schaar, J. Cornelis, and P. Schelkens, “In-band motion compensated temporal filtering,” *Signal Processing: Image Communication* (*special issue on ”Subband/Wavelet Interframe Video Coding”*), vol. 19, no. 7, pp. 653–673, Aug. 2004.

[37] IEEE 802.11a-1999: High-speed physical layer in the 5 GHz band, 1999.

[38] H. Boche, E. A. Jorswieck, “Optimization of matrix monotone functions: saddle-point, worst case noise analysis, and applications”, *Proc. of ISIT*, pp.62, Jun 2004

APPENDIX A

In this appendix, we discuss how to describe the Shannon capacity region of the Gaussian multiple access channels with ISI. We modify the iterative waterfilling solution for the Gaussian vector multiple access channel to describe the Shannon capacity region of the multi-carrier Gaussian multiple access channels efficiently.

Multi-carrier multiple access systems could be viewed as simplified Gaussian vector multiple access channels. The input optimization problem for the Gaussian vector multiple access channel has been studied in the literature for several special cases. The capacity region of a Gaussian multiple access channel with ISI was characterized in [29]. For the multiple access channel with ISI, the input optimization problem can be formulated as a problem of optimal power allocation over frequencies, which leads to the multi-user waterfilling solution. In [28], it was shown that a reduction in computational complexity can be realized for the multiple access channel rate sum maximization problem by extending the single-user water-filling to the multi-user case. However, no similar efficient algorithm has been reported in characterizing the capacity region of the multi-carrier Gaussian multiple access channel with ISI.

As discussed in Section IV, for the convex Shannon capacity region, characterizing the entire capacity boundary is equivalent to solving (12). For a N-user Gaussian vector multiple access channel, in which each user is subjected to its power constraint $P_1^{\max},\cdots,P_N^{\max}$, the input distributions that maximize $\sum_{i=1}^{N}\mu_i R_i$, with $0 \le \mu_1 \le \cdots \le \mu_N$ and $\sum_{i=1}^{N}\mu_i = 1$, are Gaussian distributions whose covariance matrices $S_1,\cdots,S_N$ can be found by solving the following optimization problem [28], which is the equivalent expression for (12):

$$\begin{array}{llll} \max \; \mu_1 \cdot \frac{1}{2}\log\left|\sum_{i=1}^{N}\boldsymbol{H_i S_i H_i^T} + \boldsymbol{S_{zz}}\right| - \mu_N \cdot \frac{1}{2}\log\left|\boldsymbol{S_{zz}}\right| + \sum_{j=2}^{N}\left(\mu_j - \mu_{j-1}\right)\cdot\frac{1}{2}\log\left|\sum_{i=j}^{N}\boldsymbol{H_i S_i H_i^T} + \boldsymbol{S_{zz}}\right| & & & \\ s.t. & tr\left(\boldsymbol{S_i}\right) \le P_i^{\max} & i = 1,\cdots,N & . \\ & \boldsymbol{S_i} \succeq 0 & i = 1,\cdots,N & \end{array} \tag{32}$$

where $tr(A)$ represents the trace of matrix A and $|A|$ represents the determinant of matrix A.

The iterative water-filling algorithm is proposed to solve the sum rate maximization problem [28], in which $\boldsymbol{\mu} = \left[1\,1\cdots 1\right]$. However, it cannot be extended to general $\boldsymbol{\mu}$. In the multi-carrier multiple access networks, H_i are diagonal matrices because users occupy channels orthogonally in the frequency domain, which means $\boldsymbol{H_i} = diag\left(H_{i1}, H_{i2},\cdots,H_{iK}\right)$. For the simplicity of illustration, we assume that the noise at different sub-carriers are independent and have the same noise variance σ_n^2, $\boldsymbol{S_{zz}} = \sigma_n^2\boldsymbol{I}$. According to [38], the problem in (32) can be further converted into

$$\begin{array}{llll} \max \; \mu_1 \cdot \frac{1}{2}\log\left\{\prod_{k=1}^{K}\left[\left(\sum_{i=1}^{N}P_{ik}\left|H_{ik}\right|^2\right) + \sigma_n^2\right]\right\} - \mu_N \cdot \frac{1}{2}\log\sigma_n^{2K} + \sum_{j=2}^{N}\left(\mu_j - \mu_{j-1}\right)\cdot\frac{1}{2}\log\left\{\prod_{k=1}^{K}\left[\left(\sum_{i=j}^{N}P_{ik}\left|H_{ik}\right|^2\right) + \sigma_n^2\right]\right\} & & & \\ s.t. & \sum_{k=1}^{K}P_{ik} \le P_i^{\max} & i = 1,\cdots,N & , \\ & P_{ik} \ge 0 & i = 1,\cdots,N & \end{array} \tag{33}$$

In the case of multi-carrier multiple access networks, iterative waterfilling in [28] can be modified to describe all the boundary points. At each step of updating a specific user's power allocation, we can simply view the sum of all the other users' signals as noise. For user i, using Lagrange multipliers, the objective function in (33) can be rewritten as

$$\begin{aligned} \mathcal{L}\left(P_{i1}, P_{i2},\cdots,P_{iK}\right) &= \mu_1 \cdot \sum_{k=1}^{K}\log\left[P_{ik}\left|H_{ik}\right|^2 + \left(\sum_{m=1,m\ne i}^{N}P_{mk}\left|H_{mk}\right|^2 + \sigma_n^2\right)\right] \\ &+ \sum_{j=2}^{i}\left(\mu_j - \mu_{j-1}\right)\cdot\sum_{k=1}^{K}\log\left\{P_{ik}\left|H_{ik}\right|^2 + \left(\sum_{m=j,m\ne i}^{N}P_{mk}\left|H_{mk}\right|^2 + \sigma_n^2\right)\right\} - \lambda\left(\sum_{k=1}^{K}P_{ik}\right) \end{aligned} . \tag{34}$$

Differentiating with respect to P_{ik}, we have

$$\frac{\mu_1}{P_{ik}+\sigma'^2_{ik,1}}+\sum_{j=2}^{i}\frac{\mu_j-\mu_{j-1}}{P_{ik}+\sigma'^2_{ik,j}}-\lambda=0, \tag{35}$$

in which

$$\sigma'^2_{ik,j}=\frac{\sum_{m=j,m\neq i}^{N}P_{mk}\left|H_{mk}\right|^2+\sigma_n^2}{\left|H_{ik}\right|^2}. \tag{36}$$

Defining $f(P_{ik})=\frac{\mu_1}{P_{ik}+\sigma'^2_{ik,1}}+\sum_{j=2}^{i}\frac{\mu_j-\mu_{j-1}}{P_{ik}+\sigma'^2_{ik,j}}$ which monotonically decreases in P_{ik}, $f^{-1}(\bullet)$ is an injection function.

Noting that $P_{mk}\geq 0$, the optimal power is given by the KKT conditions

$$P_{ik}=\left(f^{-1}(\lambda)\right)^{+}. \tag{37}$$

where $(x)^{+}$ denotes the positive part of x, i.e., $(x)^{+}=\max\{x,0\}$.

The procedure of iterative waterfilling algorithm is summarized in Algorithm 5.

Algorithm 5 Iterative water-filling algorithm in describing the capacity region of the multi-carrier multiple access channel

Input: $\boldsymbol{H},\boldsymbol{P^{max}},\boldsymbol{\mu},\sigma_n^2$

Initialization : $P_{ik},\ i=1,2,\cdots,N,\ k=1,2,\cdots,K$

Repeat:

for $i=1$ to N

$$f(P_{i1},P_{i2},\cdots,P_{iK})=\mu_1\cdot\sum_{k=1}^{K}\log\left[P_{ik}\left|H_{ik}\right|^2+\left(\sum_{m=1,m\neq i}^{N}P_{mk}\left|H_{mk}\right|^2+\sigma_n^2\right)\right]$$
$$+\sum_{j=2}^{i}\left(\mu_j-\mu_{j-1}\right)\cdot\sum_{k=1}^{K}\log\left\{P_{ik}\left|H_{ik}\right|^2+\left(\sum_{m=j,m\neq i}^{N}P_{mk}\left|H_{mk}\right|^2+\sigma_n^2\right)\right\}$$

Use optimality condition in (35)-(37) to solve

$$\{P_{i1},P_{i2},\cdots,P_{iK}\}=\arg\max_{\{P_{i1},P_{i2},\cdots,P_{iK}\}}f(P_{i1},P_{i2},\cdots,P_{iK}),\ s.t.\ \sum_{k=1}^{K}P_{ik}\leq P_i^{\max},\ P_{ik}\geq 0, i=1,\cdots,N$$

end

Until: the desired accuracy is reached

Return: Power allocation $\boldsymbol{P}$