Resource Allocation for Multi-user Video Transmission over Multi-carrier Networks

Yi Su and Mihaela van der Schaar

Electrical Engineering Department, UCLA

Email: {yisu,mihaela}@ee.ucla.edu

***Abstract*—This paper addresses 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 Physical (PHY) and Medium Access Control (MAC) layers, we exploit the unique property of state-of-the-art video coders that can provide bitstream prioritization in terms of distortion impact and solve the problem of allocating wireless resources, i.e., power and rate among multiple users such that the weighted sum of the overall video qualities is maximized. An optimality condition is derived to describe the achievable convex utility region. We start from the two-user case and develop an algorithm for the optimal resource allocation. Inspired by the intuition gained from the two-user case, we extend the algorithm 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.**

***Keywords*—wireless resource allocation, wireless multimedia, multi-carrier networks, utility driven resource management**

I. Introduction

Multi-carrier communication is becoming the leading physical layer technology for many existing and emerging wireless networks and standards [1]. An important application over this network is bandwidth-intense multimedia streaming. Hence, the development of advanced resource allocation strategies for wireless multimedia 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. The optimization for independent and identically-distributed (i.i.d.) fading channels is studied in [2] and [3], where the optimal power allocation which maximizes the weighted sum of the rates over time is characterized. Another important result in the area of wireless resource allocation is determined based on a combination of information theory and queueing theory. An optimum policy named "Longest Queue Highest Possible Rate" (LQHPR) is found [4], which allows the system to obtain the shortest average delay. Related cross-layer approaches on queueing stability, delay, and adaptive coding and modulation schemes can be found in [5]-[7]. Recent studies [8][9] 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 [9]. However, prior works [8][9] 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.

In this paper, we address optimal resource (power and rate) allocation in multi-user video transmission. Existing cross-layer research focuses on the Physical (PHY), Medium Access Control (MAC), and Network layers [2]-[8]. Alternatively, we study this problem using an integrated 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 prioritizes the encoded video streams based on overall distortion impact [10], which results in a concave increase of the video quality as a function of rate. We develop a unified PHY-MAC-APP framework and study the optimal resource allocation policy which maximizes the weighted sum of utilities. We take an information theoretic approach throughout this paper. Since the Shannon 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 the operational rate distortion theory. Typical applications of the proposed solutions include multi-user video transmission (e.g. uploading movies) over WLAN or cognitive radio [11], and video surveillance form wireless camera.

Our main contributions in this paper are as follows. We demonstrate the convexity of the utility region measured in the Peak Signal-to-Noise Ratio (PSNR) sense. Without requiring full knowledge of the entire capacity region, we propose a procedure to determine the utility region for the two-user case and extend it to the multi-user case using a heuristic approach. The proposed algorithms make it tractable to describe the entire utility region and greatly reduce the complexity of the weighted sum maximization of the utilities.

This paper is organized as follows. Section II describes the model of multi-carrier wireless networks for multi-user video streaming, explains the deployed utility function, and formulates the multi-user resource allocation into an optimization problem. In Section III, the optimality condition is derived and iterative approaches are developed to optimize the resource allocation. Section IV gives simulation results of the proposed algorithms. Conclusions are drawn in Section V.

II. Multi-user Multimedia Transmission

A. System Description

We focus on the Gaussian multiple-access channel in this paper. A Gaussian multiple-access channel is a multiple-access

This work was supported in part by NSF CAREER Award CCF-0541867 and in part by grants from UC Micro and ONR.

channel where the additive noise is Gaussian [12]. The system diagram is shown in Figure 1. The optimal transmission strategy for the multiple-access channels generally requires all the users to share the entire frequency band simultaneously. We focus on this strategy throughout this paper.

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

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

where $X_{ij}(n)$ is the transmitted symbol of user i at jth sub-carrier at time n and $N_j(n)$ is the additive white Gaussian noise (AWGN) with two-sided spectral density of $N_0/2$. The power that user i allocates at the jth sub-carrier is $E[\|X_{ij}(n)\|^2] = P_{ij}$.

We denote the CSI vector as $\boldsymbol{H} = (\boldsymbol{H}_1, \boldsymbol{H}_2, \cdots, \boldsymbol{H}_N)$ in which

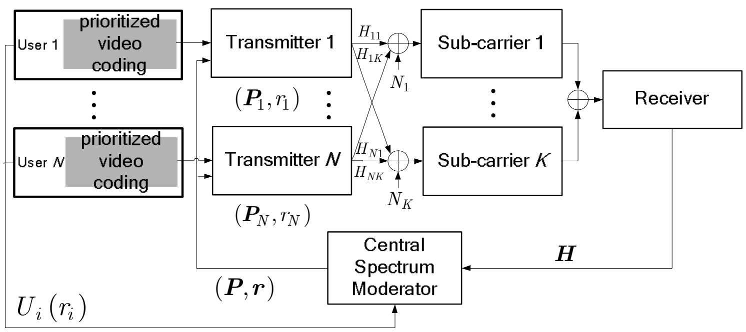

Figure 1. System Structure.

$\boldsymbol{H}_i = (H_{i1}, H_{i2}, \cdots, H_{iK})$, the power allocation vector as $\boldsymbol{P} = (\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ in which $\boldsymbol{P}_i = (P_{i1}, P_{i2}, \cdots, P_{iK})$, 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.

We assume that users are streaming pre-compressed video content over a multi-carrier wireless infrastructure. The Central Spectrum Moderator (CSM) collects the 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 optimally determines 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 $P_i^{\max}$:

$$\sum_{j=1}^{K} P_{ij} \le P_i^{\max}, \quad (2)$$

and we denote $\boldsymbol{P^{max}} = (P_1^{\max}, P_2^{\max}, \cdots, P_N^{\max})$.

B. Utility-Rate Functions

Here we define the utility to be the video quality in terms of the PSNR, which is the only widely accepted metric for assessing the video quality. It has been shown that partitioning the packets into different priority classes and adjusting the transmission strategies for each class can significantly improve the received quality and provide graceful degradation [13]. Several operational utility-rate models can accurately capture the performance of various coders for different video sequence characteristics [14][15]. 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 [15]. Based on this model, the utility (PSNR) for user i is given by

$$U_i(r_i) = 10 \log \frac{255^2 (r_i - R_{0i})}{D_{0i}(r_i - R_{0i}) + c_i}, \quad (1)$$

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. Throughout this paper, we assume that $r_i > R_{0i}$.

C. Problem Formulation

We denote the capacity region with the power constraint $\boldsymbol{P^{max}}$ as $C(\boldsymbol{P^{max}})$. Figure 2 shows the basic idea in formulating this problem, which can be summarized into two steps. First, for any power allocation $\boldsymbol{P} \preceq \boldsymbol{P^{max}}$, there is an achievable rate vector r within the capacity region $C(\boldsymbol{P^{max}})$, i.e., $\boldsymbol{r} \in C(\boldsymbol{P^{max}})$ [17]. Second, by mapping the rate vector $\boldsymbol{r}$ into the utility vector $\boldsymbol{u}$ using the utility-rate functions $U_i(r_i)$, we can obtain the utility region $U(\boldsymbol{P^{max}})$, which is defined to be:

$$U(\boldsymbol{P^{max}}) = \{(U_1(r_1), U_2(r_2), \cdots, U_n(r_n)) : (r_1, r_2, \cdots, r_n) \in C(\boldsymbol{P^{max}})\} \quad (2)$$

Note that $U_i(r_i)$ depends on the video characteristics and encoder-selected parameters. The objective function that we aim to maximize is the weighted sum of the overall utilities:

$$\max_{\boldsymbol{u}} \boldsymbol{\beta u}, \ s.t. \ \boldsymbol{u} \in U(\boldsymbol{P^{max}}), \quad (3)$$

where $\boldsymbol{\beta} \in R_+^N$, with $\|\boldsymbol{\beta}\|_1 = 1$, is a weighted vector whose components indicate the importance of the various users.

Figure 2 highlights that, in order to solve the optimization problem in (3), we need to describe the capacity region first and map it into the utility region. In the next section, we derive the optimality condition for achieving the utility boundary and develop efficient algorithms to solve the problem in (3).

III. Optimal Resource Allocation

A. Convexity of the Achievable Utility Region

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 $R_i' \le R_i, \forall i$ is also within the achievable capacity region $C(\boldsymbol{P^{max}})$.

Proof: It follows from the convex hull operation that forms the capacity region of a multiple access channel [12]. ■

Lemma 2: The utility-rate function in (1) is a monotonically increasing and concave function in r_i.

Proof: This can be verified by taking the first and second derivatives of the utility-rate function. Note that the monotonically increasing and concave property comes from the inherent prioritization of the video bitstream. ■

Proposition 1: The utility region $U(\boldsymbol{P^{max}})$ is convex.

Proof: The convexity can be proven using Lemma 1 and 2.

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 $\boldsymbol{U}(\boldsymbol{r}) = [U_1(r_1) \ U_2(r_2) \cdots U_N(r_N)]^T$.

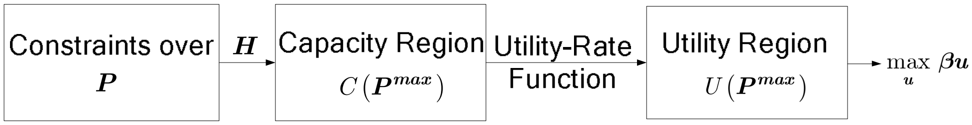

Figure 2. Problem Interpretation.

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

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:

$$\sum_{k=1}^{K} \hat{P}_{ik} = \left[\alpha\sum_{k=1}^{K} P_{ik} + (1-\alpha)\sum_{k=1}^{K} P'_{ik}\right] \le P_i^{\max}, \forall i = 1,2,\cdots,N. \quad (4)$$

The rates are concave functions in $\boldsymbol{P}$ [17], therefore

$$\alpha\mathcal{R}(\boldsymbol{P}) + (1-\alpha)\mathcal{R}(\boldsymbol{P}') \preceq \mathcal{R}(\alpha\boldsymbol{P} + (1-\alpha)\boldsymbol{P}') = \mathcal{R}(\hat{\boldsymbol{P}}). \quad (5)$$

By the property in Lemma 2, (5) can be converted into

$$\alpha\mathcal{U}(\boldsymbol{P}) + (1-\alpha)\mathcal{U}(\boldsymbol{P}') = \alpha\boldsymbol{U}(\mathcal{R}(\boldsymbol{P})) + (1-\alpha)\boldsymbol{U}(\mathcal{R}(\boldsymbol{P}')) \preceq \boldsymbol{U}(\alpha\mathcal{R}(\boldsymbol{P}) + (1-\alpha)\mathcal{R}(\boldsymbol{P}')) \preceq \boldsymbol{U}(\mathcal{R}(\alpha\boldsymbol{P} + (1-\alpha)\boldsymbol{P}')) = \mathcal{U}(\hat{\boldsymbol{P}}). \quad (6)$$

By the monotonically increasing property of $U_i(r_i)$, from (6), we know that there exists a rate vector $\bar{\boldsymbol{r}}$ satisfying

$$\boldsymbol{U}(\bar{\boldsymbol{r}}) = \alpha\mathcal{U}(\boldsymbol{P}) + (1-\alpha)\mathcal{U}(\boldsymbol{P}') \Rightarrow \bar{\boldsymbol{r}} \preceq \mathcal{R}(\alpha\boldsymbol{P} + (1-\alpha)\boldsymbol{P}').$$

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

$$\bar{\boldsymbol{r}} = \mathcal{R}(\bar{\boldsymbol{P}}). \quad (7)$$

Therefore, $\forall\, \mathcal{U}(\boldsymbol{P}), \mathcal{U}(\boldsymbol{P}') \in U(\boldsymbol{P}^{max}), 0 \le \alpha \le 1$

$$\Rightarrow \alpha\mathcal{U}(\boldsymbol{P}) + (1-\alpha)\mathcal{U}(\boldsymbol{P}') = \mathcal{U}(\bar{\boldsymbol{P}}) \in U(\boldsymbol{P}^{max}).$$

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

Thus, explicitly characterizing the entire convex utility region $U(\boldsymbol{P}^{max})$ is equivalent to solving the optimization problem in (3) for all possible $\boldsymbol{\beta} \in R_+^N$ and $\|\boldsymbol{\beta}\|_1 = 1$.

B. Optimality Condition

Conventional approaches in solving (3), such as LQIHPR [9], continuously searches the optimum along the utility rate functions until reaching the boundary of the capacity region. In the case of scalar non-fading AWGN channel, the capacity regions exhibit the polymatroid structure, and the boundary can be characterized by finite inequalities. Therefore, by checking these inequalities, we know that whether a rate vector reaches the capacity boundary. However, for multi-carrier cases, this approach is computationally intensive because it is in general impossible to describe the capacity region in finite inequalities. The only way to trace out the entire capacity region is to maximize the weighted sum of the rates for all possible $\boldsymbol{\mu}$:

$$\max_{\boldsymbol{r}} \boldsymbol{\mu r} \quad s.t.\ \boldsymbol{r} \in C(\boldsymbol{P}^{max}), \quad (8)$$

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. The iterative water-filling algorithm provides an efficient solution of the problem in (8) for the Gaussian vector multiple-access channel [17]. We also use this algorithm in this paper. Now the problem is reduced to how to find the optimal solution of problem (3) efficiently, based on the assumption that we are already able to efficiently solve problem (8) for all possible $\boldsymbol{\mu}$.

Now we derive the mapping function, which projects the normal vector to the tangent hyperplane at each boundary point from the capacity region $C(\boldsymbol{P}^{max})$ to the utility region $U(\boldsymbol{P}^{max})$. Suppose a power allocation $\boldsymbol{P}$ reaches the boundaries of $C(\boldsymbol{P}^{max})$ and $U(\boldsymbol{P}^{max})$. We denote the normal vectors to the tangent hyperplanes at the capacity and utility boundary as $\boldsymbol{\mu}(\boldsymbol{P}) = \arg\max_{\boldsymbol{\mu}\in R_+^N\ and\ \|\boldsymbol{\mu}\|_1=1} \boldsymbol{\mu}\mathcal{R}(\boldsymbol{P})$ and $\boldsymbol{\beta}(\boldsymbol{P}) = \arg\max_{\boldsymbol{\beta}\in R_+^N\ and\ \|\boldsymbol{\beta}\|_1=1} \boldsymbol{\beta}\mathcal{U}(\boldsymbol{P})$.

Proposition 2: For any power allocation $\boldsymbol{P} = (\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)$ that achieves the boundary of $U(\boldsymbol{P}^{max})$ and satisfies (3), the relation between $\boldsymbol{\mu}(\boldsymbol{P})$ and $\boldsymbol{\beta}$ is given by

$$\frac{\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P})}{\|\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P})\|_1} = \boldsymbol{\mu}(\boldsymbol{P}), \quad (9)$$

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 [19].

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

$$C(\boldsymbol{P}^{max}) = \bigcap_{\{\boldsymbol{\mu}^i\}} \left\{\boldsymbol{r} \mid \boldsymbol{\mu}^i\boldsymbol{r} \le \boldsymbol{\mu}^i\boldsymbol{r}_{\boldsymbol{\mu}^i}\right\},\ \forall\boldsymbol{\mu}^i \in R_+^N\ and\ \|\boldsymbol{\mu}^i\|_1 = 1, \quad (10)$$

where $\boldsymbol{r}_{\boldsymbol{\mu}^i} = \arg\max_{\boldsymbol{r}\in C(\boldsymbol{P}^{max})} \boldsymbol{\mu}^i\boldsymbol{r}$. Form the Lagrangian of (3):

$$\mathcal{L}(\boldsymbol{r}, v) = \boldsymbol{\beta}(\boldsymbol{P})\boldsymbol{u} + \sum_{i=1}^{\infty} v_i\left(\boldsymbol{\mu}^i\boldsymbol{r}_{\boldsymbol{\mu}^i} - \boldsymbol{\mu}^i\boldsymbol{r}\right), \quad (11)$$

in which $\boldsymbol{u} = \mathcal{U}(\boldsymbol{P}) = \boldsymbol{U}(\mathcal{R}(\boldsymbol{P})) = \boldsymbol{U}(\boldsymbol{r})$ and $v_i \ge 0, i = 1,2,\cdots$. 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 (11) with respect to $\boldsymbol{r}$. At the optimum, only one inequality constraint in (10) 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 [18], $v_{opt} > 0$, and

$$\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P}) = v_{opt}\boldsymbol{\mu}^{opt}, \quad (12)$$

in which $\boldsymbol{\mu}^{opt} = \boldsymbol{\mu}(\boldsymbol{P}) = \arg\max_{\boldsymbol{\mu}\in R_+^N\ and\ \|\boldsymbol{\mu}\|_1=1} \boldsymbol{\mu}\mathcal{R}(\boldsymbol{P})$. Note that (12) is identical to (9), because

$$\frac{\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P})}{\|\boldsymbol{\beta}(\boldsymbol{P}) \odot \boldsymbol{\lambda}(\boldsymbol{P})\|_1} = \frac{v_{opt}\boldsymbol{\mu}^{opt}}{\|v_{opt}\boldsymbol{\mu}^{opt}\|_1} = \frac{\boldsymbol{\mu}^{opt}}{\|\boldsymbol{\mu}^{opt}\|_1} = \boldsymbol{\mu}(\boldsymbol{P}), \quad (13)$$

where $\boldsymbol{\mu}(\boldsymbol{P})$ and $\boldsymbol{\beta}(\boldsymbol{P})$ are the normal vectors to the tangent hyperplanes at the boundary of the capacity and utility region respectively, and $\boldsymbol{\lambda}(\boldsymbol{P})$ consists of the first order derivatives of the utility-rate functions. This equality links the boundary points of capacity region and utility region. ■

By the equality in (9), we can project the normal vector to the tangent hyperplane at each boundary point from $C(\boldsymbol{P}^{max})$ to $U(\boldsymbol{P}^{max})$. This process is illustrated by the right pointing arrow in Figure 3. For $\forall\boldsymbol{\mu} \in R_+^N$ and $\|\boldsymbol{\mu}\|_1 = 1$, we solve the problem (8), get the boundary point $\boldsymbol{r}$, and subsequently use (9) to calculate

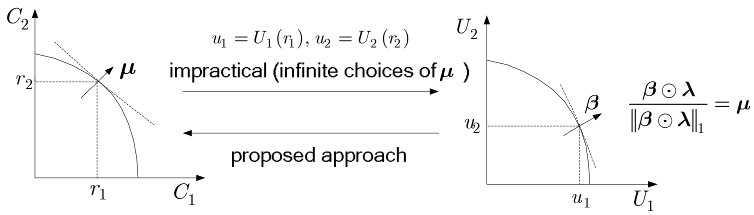

Figure 3. Optimality Condition for the Two-user

the vector β. As mentioned before, the right pointing arrow provides a possible solution of the problem in (3), that is, enumerate all the possible μ until the normal vector after mapping coincides with the original β in (3). This solution is impractical because there are infinite possible choices in μ. As illustrated by the left pointing arrow in Figure 3, if β is given, we are interested in how to search μ until the optimality condition in (9) holds.

We discuss how to search and find the optimal solution of (9) efficiently starting from the two-user case. Based on the insights gained from the two-user case, we develop a practical heuristic search algorithm for the multiple-user case. We first highlight several monotonic properties upon which the search algorithms of optimal power allocation are based.

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

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

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

Proof: The equality is identical to

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

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

Proposition 4: Suppose $\mu = [\mu_1 \; \mu_2 \cdots \mu_N] \in R_+^N$, $\|\mu\|_1 = 1$, and $r_\mu = \max_r \mu r$. For all m, if the m th component μ_m of the weighted vector μ increases and the other components are fixed, the m th component of the rate vector r_μ remains the same or increases while all the other components of r_μ decreases.

Proof: See Lemma 6 in [20]. ■

C. Algorithm for Describing the Two-User Utility Region

Now we consider the two-user case. In the Shannon capacity region, we are able to solve (8) for arbitrary $\mu = [\mu_1 \; \mu_2] \in R_+^2$ and $\|\mu\|_1 = 1$. Note that $\mu_1 = 1 - \mu_2$, hence finding μ that satisfies the equality in (9) 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 (3) efficiently.

Suppose that the rate vector $r = [r_1 \; r_2 \cdots r_N]^T$ maximizes μr. we can take the vector $\lambda' = \left[\frac{\partial U_1(x)}{\partial x}\Big|_{x=r_1} \;\; \frac{\partial U_2(x)}{\partial x}\Big|_{x=r_2} \cdots \frac{\partial U_N(x)}{\partial x}\Big|_{x=r_N} \right]$ and get a new vector μ' by $\mu' = \beta \odot \lambda' / \|\beta \odot \lambda'\|$. Denoting $\mu_1' = g(\mu_1)$ and $f(\mu_1) = \mu_1 - g(\mu_1)$, the bisection search algorithm for the two-user case is summarized in Algorithm 1.

Algorithm 1: Two-user optimum search algorithm

Input: H, P^{max}, β; error tolerance ε; both users' utility rate models;

Initialization: Take $\mu_{(0)}$ randomly which satisfies $\|\mu_{(0)}\|_1 = 1$, $i = 1$, flag = 0

Repeat:

1) Use the iterative water-filling Algorithm in [17] to find the point on the capacity boundary which maximizes $\mu_{(i)} r$;

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

3) Calculate $\mu'_{(i)} = \left[\lambda_{1(i)}\beta_1 \;\; \lambda_{2(i)}\beta_2\right] / \left\|\left[\lambda_{1(i)}\beta_1 \;\; \lambda_{2(i)}\beta_2\right]\right\|_1$ and

$f(\mu_{1(i)}) = \mu_{1(i)} - g(\mu_{1(i)}) = \mu_{1(i)} - \mu'_{1(i)}$;

4) If flag = 0

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

Else

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

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

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

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

Return: $r_{(i)}$ and its corresponding power allocation P

Proposition 5: Algorithm 1 converges to the solution of (3).

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. For the convex utility region, there exist solutions for $f(\mu_1) = 0$. The monotonicity of $f(\mu_1)$ ensures that there is a unique zero of $f \in \mathbb{R}[0,1]$. The convergence of the bisection search is guaranteed and we can use it to find the boundary point where the optimality condition in (9) holds. ■

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 $[a_{(i)}, b_{(i)}]$ 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 $\log_2[1/\varepsilon]$.

D. Algorithm for Describing the Multi-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, map the capacity region into the utility region, and use (13) to find the optimal rate vector. This algorithm is impractical, because the boundary of $C(\boldsymbol{P}^{max})$ consists of infinite points and lacks closed form expression. In this subsection, we will discuss how to search the optimum by developing a heuristic search algorithm inspired by the intuition gained from the two-user search algorithm.

Algorithm 2: Multiple-user low-complexity heuristic search.

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 $\|\boldsymbol{\mu}_{(0)}\|_1 = 1$, $i = 1$

Repeat:

1) Use the iterative water-filling Algorithm in [17] 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 $\boldsymbol{\mu}'_{(i)} = \boldsymbol{\beta} \odot \boldsymbol{\lambda}_{(i)} / \|\boldsymbol{\beta} \odot \boldsymbol{\lambda}_{(i)}\|_1$;

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

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

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

In the two-user case, both $|r_{1(i)} - r'_{1(i)}|$ and $|r_{2(i)} - r'_{2(i)}|$ decrease after each iteration. Intuitively, in the multiple-user case, we should update $\boldsymbol{\mu}$ so that $|r_{n(i)} - r'_{n(i)}|$ 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 $|r_{m(i)} - r'_{m(i)}|$ to decrease for all $m = 1,2,\cdots,N$ with large probability. We can repeat this procedure until the optimality condition in (9) holds. The low-complexity heuristic search in the multiple-user case is summarized in Algorithm 2.

It should be pointed out that the step size δ should be carefully chosen. 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 [21].

IV. 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 $\boldsymbol{\beta} = [0.9\ 0.1]$ for illustration and $P_1^{\max} = 10^{1.2}, P_2^{\max} = 10^2$. The parameters values for the utility-rate function deployed for these experiments are determined based on a state-of-the-art wavelet video coder. We assume that user 1 wants to transmit the *Mobile* video (CIF, 15Hz) with $D_{01} = 1$, $R_{01} =$ 44.04kbps, and $c_1 =$ 38230kbps, while user 2 has the the *Foreman* video (CIF, 15Hz) for transmission with $D_{02} = 1$, $R_{02} =$ 20.72kbps, and $c_2 =$ 2760kbps. R_{0i}, D_{0i}, and c_i are the parameters of the utility-rate model in (1).

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 that Algorithm 1 returns is $[685.23\ 786.87]^T$ (kbps) and 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 will maximize $r_1 + r_2$, which result in 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.

We also simulate the multiple-user case with $N = 3$ users. The channel conditions for the three users are given in Table IV. We choose the weighted vector $\boldsymbol{\beta} = [0.3\ 0.3\ 0.4]$ for illustration and $P_1^{\max} = 10, P_2^{\max} = 10^{1.5}, P_3^{\max} = 10^2$. We assume that user 1 transmits the *Foreman* video (CIF, 15Hz) with $D_{01} = 1$, $R_{01} =$

20.72kbps, 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 transmits the *Foreman* video (CIF, 30Hz) with $D_{03} = 3$, $R_{03} = 55.08$kbps, and $c_3 = 4610$kbps.

We set step size $\delta = 0.1$ and apply Algorithm 2 to search the optimum. In this case, Algorithm 2 converges after around 20 iterations. The rate vector which approximately maximizes βu is $[586 \;\; 698.4 \;\; 937.1]^T$ (kbps) and the power allocation is given in Table V. Under this power allocation, user 1's average PSNR is 40.4379dB, user 2 is 36.8719dB, user 3 is 39.1417dB, and the weighted sum of PSNRs is 38.8496dB. The existing sum-rate-maximizing approach with $\mu = [1/3 \;\; 1/3 \;\; 1/3]$ leads the weighted sum of PSNRs to be 36.2988dB, which is of nearly 3dB degradation.

We can see that, by explicitly considering the operational rate-distortion performance of video coders, both Algorithm 1 and 2 can achieve significant performance improvement than the existing rate-maximizing approach.

V. CONCLUSION

In this paper, we address the problem of multi-user video transmission in multi-carrier wireless networks. A general procedure is proposed to determine the achievable utility region under the constraints of the capacity regions. We first develop iterative search algorithm to optimize the resource allocation 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 heuristic approaches. Experiments show that these algorithms achieve significant performance improvements by considering the video utility impact and the rate-distortion performance of the video coder.

REFERENCES

[1] A. Bahai, B. 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," *IEEE ICC*, pp. 331–335, 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," *Proc. of the Allerton Conf.*, pp.140–149, Sep. 2004.

[5] E. Telatar and R. Gallager, "Combining queueing theory with information theory for multiaccess," *IEEE JSAC*., vol. 13, pp. 963–969, Aug. 1995.

[6] S. Raj, E. Telatar, and D. Tse, "Job scheduling and multiple access," *DIMACS*, March 2003.

[7] 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.

[8] 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.

[9] C. Shen and M. van der Schaar, "Optimal resource allocation in wireless multiaccess video transmissions", *Proc.of ICC*, pp 4581-4586, June 2007

[10] J.R. Ohm, M. van der Schaar, J. Woods, "Interframe wavelet coding-Motion Picture Representation for Universal Scalability", *EURASIP Signal Processing: Image Communication*, 2004.

[11] J. Mitola, "Cognitive radio for flexible mobile multimedia communications," *Mobile Networks and Application*s, vol. 6, Sep. 2001.

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

[13] A. Albanese, M. Luby, "PET-priority encoding transmission," *High-Speed Networking for Multimedia Application*, Kluwer, 1996

[14] 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

[15] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *JSAC*, vol.18, pp.1012–1032, 2000.

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

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

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

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

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

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

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

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

TABLE II
AN EXAMPLE OF ALGORITHM 1

Iteration	$\mu_{(i)}$	$r_{(i)}^T$	$\mu'_{(i)}$	$r'^T_{(i)}$	$[a_{(i)} \; b_{(i)}]$
$i=1$	[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]
⋮	⋮	⋮	⋮	⋮	⋮
$i=7$	[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

User	P_{i1}	P_{i2}	P_{i3}	P_{i4}	P_{i5}	P_{i6}	P_{i7}	P_{i8}
User 1	1.4505	2.5221	0	2.5245	2.6773	3.0088	0.8308	2.8348
User 2	19.454	17.249	18.783	13.612	0	0	17.456	13.445

TABLE IV
CHANNEL CONDITION OF A THREE-USER SYSTEM ($N_0 B = 1$)

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

TABLE V
POWER ALLOCATION OF THE THREE-USER SYSTEM

User	P_{i1}	P_{i2}	P_{i3}	P_{i4}	P_{i5}	P_{i6}	P_{i7}	P_{i8}
User 1	0	2.077	1.921	1.995	1.167	1.581	1.259	0
User 2	5.307	8.709	8.445	0	0	9.162	0	0
User 3	11.683	0	0	20.772	21.818	0	22.498	23.229