

# Statistical Delay QoS Driven Energy Efficiency and Effective Capacity Tradeoff for Uplink Multi-User Multi-Carrier Systems

Wenjuan Yu, Leila Musavian and Qiang Ni

**Abstract**—In this paper, the total system effective capacity (EC) maximization problem for the uplink transmission, in a multi-user multi-carrier OFDMA system, is formulated as a combinatorial integer programming problem, subject to each user's link-layer energy efficiency (EE) requirement as well as the individual's average transmission power limit. To solve this challenging problem, we first decouple it into a frequency provisioning problem and an independent multi-carrier link-layer EE-EC tradeoff problem for each user. In order to obtain the subcarrier assignment solution, a low-complexity heuristic algorithm is proposed, which not only offers close-to-optimal solutions, while serving as many users as possible, but also has a complexity linearly relating to the size of the problem. After obtaining the subcarrier assignment matrix, the multi-carrier link-layer EE-EC tradeoff problem for each user is formulated and solved by using Karush-Kuhn-Tucker (KKT) conditions. The per-user optimal power allocation strategy, which is across both frequency and time domains, is then derived. Further, we theoretically investigate the impact of the circuit power and the EE requirement factor on each user's EE level and optimal average power value. The low-complexity heuristic algorithm is then simulated to compare with the traditional exhaustive algorithm and a fair-exhaustive algorithm. Simulation results confirm our proofs and design intentions, and further show the effects of delay quality-of-service (QoS) exponent, the total number of users and the number of subcarriers on the system tradeoff performance.

**Index Terms**—Link-layer energy-rate tradeoff, delay-outage probability, effective capacity, energy efficiency.

## I. INTRODUCTION

Green communication networks, which not only emphasize on spectrum efficiency (SE), but also promise high energy efficiency (EE), have become imperative needs of future communication systems. However, by nature, EE and SE could require conflicting design approaches. From the information-theoretic point of view, the EE-SE tradeoff problem in a downlink orthogonal frequency division multiple access (OFDMA) network was analyzed in [2], in which the impact of the channel power gain and the circuit power on the EE-SE relation was discussed. Considering the cognitive radio networks, a multi-objective optimization was formulated in [3], in which the ergodic capacity was maximized and the total

transmission power of femtocell base stations was minimized. A general power consumption model in multi-user OFDMA systems, including the transmission power, signal processing power, and circuit power from both the transmitter and the receiver sides, was first established in [4]. Then the authors in [4] proposed a joint optimization method to iteratively find the optimal solution for the EE-maximization problem, subject to a peak transmit power constraint and a minimum system data rate requirement. The EE and SE tradeoff problem has also been extensively studied for other kinds of wireless communication networks, such as energy-constrained wireless multi-hop networks with a single source-destination pair [5], general narrowband interference-limited systems [6] and OFDMA-based cooperative cognitive radio networks [7]. In the aforementioned studies, however, the system throughput was given by Shannon limit, without taking into account delay constraints. For systems with delay-sensitive applications, such as video conferencing and online gaming, the physical-layer based power and rate adaptation techniques may not be efficient. In fact, 5G, the next generation of mobile communication technology, has been anticipated to offer >1 Gbps downlink data rate, sub-1ms end-to-end latency and 90% reduction in network energy usage [8]. This infers that the future wireless communication networks are targeted at satisfying the end-user applications' delay quality-of-service (QoS) requirements, while at the same time increasing EE and SE for green communications.

In order to fulfill these requirements, extensive studies in the context of power control, scheduling, and admission control have been widely provided in [9]–[21]. A cross-layer optimization framework for delay-sensitive applications over a single wireless link was formulated in [9], in which some characteristics, e.g., delay deadlines, dependencies, distortion impacts, are considered and discussed. The authors in [10] provided energy-efficient transmission techniques for a group of  $M$  packets subject to individual packet transmission delay constraints. The above works all characterize the delay QoS requirement for a dynamic queuing system in a deterministic way, where the delay is bounded within a certain threshold [11]. Although this sounds reasonable for real-time services, satisfying fixed QoS guarantees is especially challenging in fading communication scenarios, due to the random variations experienced in channel conditions, user mobility and changing environment [12], which could lead to settling for non-necessarily low data rates. In contrast to the above deterministic delay QoS bounds, in this paper we concentrate on the delay QoS requirement in a statistical way, which

This work was supported in part by the China Scholarship Council, UK EPSRC under grant number EP/K011693/1, and grant number EP/N032268/1, and the EU FP7 under grant number PIRSES-GA-2013-610524. Part of this work was presented in [1]. (Corresponding author: Qiang Ni)

W. Yu and Q. Ni are with the School of Computing and Communications, InfoLab21, Lancaster University, LA1 4WA, UK (Emails: {w.yu1, q.ni}@lancaster.ac.uk). L. Musavian is with the School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ, UK (Email: leila.musavian@essex.ac.uk).

considers and confines the delay bound violation probability to a required value range. In this direction, the authors in [13] introduced a link-layer capacity notion supporting statistical delay QoS requirements, which is the concept of effective capacity (EC). Formulated as the dual of the effective bandwidth, EC specifies the maximum arrival rate that can be supported by a wireless channel given that a target delay-outage probability requirement is guaranteed [13] [19]. Therefore, EC can be regarded as the link-layer SE. The link-layer EE, henceforth, can be formulated as the ratio of the EC to the total power expenditure [14].

Due to the inconsistent property of the link-layer EE and EC, many researchers have elaborately studied how to balance the two metrics. Considering frequency flat-fading channels, an optimal power allocation strategy to maximize EC subject to a link-layer EE constraint, for delay-limited mobile multimedia applications was obtained in [15]. For a Rayleigh flat-fading channel under delay-outage probability constraints, a multi-objective optimization problem to jointly maximize EE and EC was formulated and solved in [16]. The above mentioned papers, however, focus on a point-to-point single-channel communication system.

We note that based on the theory of Shannon limit, the total average rate of a multi-carrier system is a linear summation of each subcarrier's achievable average rate. This, however, does not apply to systems with limited statistical delay requirements. Specifically, in delay-constrained systems, the concavity and monotonicity of the EC do not remain homogeneous for single-carrier and multi-carrier systems [17]. In addition, for systems with statistical delay QoS constraints, it has been proven that the optimal power allocation strategy for single-carrier communications cannot be simply extended to the multi-carrier communications [17]. Hence, considering a single-user multi-carrier link over a frequency-selective fading channel, the delay-constrained EC maximization and EE maximization problem were separately addressed in [17] and [18], respectively. However, the link-layer EE-EC tradeoff problem for the multi-carrier communications is not investigated and analyzed in the literature. Especially, when we consider a multi-user multi-carrier network, the link-layer EE-EC tradeoff problem becomes more challenging. The formulated problem will be a complex combinatorial integer programming problem, rather than a convex optimization problem in [17] which was solved using Lagrangian method. In [20] and [21], an EE optimization problem with statistical delay provisioning and per-user's EC requirement constraint was analyzed for a downlink multi-user OFDMA network. In these papers, the power allocation for each subcarrier is assumed to be only related to its subcarrier's channel power gain, and not related to the same user's other subcarriers' channel power gains. Therefore, based on this assumption and the independent and identically distributed (i.i.d.) property of all subcarriers, the EC value for a single-user multi-carrier system can be formulated as a linear summation of the EC values of all subcarriers. While this independent optimization approach is optimal in maximizing the Shannon capacity (e.g., water-filling power control for multi-carrier transmissions), it is not the optimal policy to maximize the EC-based problems for an arbitrary statistical delay provisioning [17]. In this paper, we will not

make this assumption, and aim to derive the optimal power allocation strategy for each user, which is not only across the time domain, but also across the frequency domain.

In this paper, we target to maximize the system total EC for the uplink transmission in a multi-user multi-carrier OFDMA network, subject to each user's required link-layer EE performance level and its individual resource limits. We decouple the problem into two parts and provide the subcarrier assignment solution and optimal power allocation strategy for each user. In more detail, we propose a low-complexity heuristic algorithm, which first allocates each served user the exact number of its required subcarriers, and then implements the optimal per-user power allocation strategy to calculate each user's current EC value. Finally, the remaining subcarriers will be allocated by adopting the strategy that the user with current minimum EC value has the allocation priority.

To sum up, this paper has the following contributions:

- A novel total EC maximization problem for the uplink transmission, in a multi-user multi-carrier OFDMA system, is formulated as a complex combinatorial integer programming problem, subject to each user's link-layer EE requirement and the individual's average input power limit. A new adjustable EE requirement factor is defined to further tune each user's EE constraint value, which transforms the formulated problem into a tradeoff problem between the system total EC and the users' individual EE achievements.
- The formulated challenging problem is first decoupled into a frequency provisioning problem and an independent link-layer multi-carrier EE-EC tradeoff problem for each user. The traditional exhaustive algorithm and a fair-exhaustive algorithm are introduced first, followed by a low-complexity heuristic algorithm, which cares about user fairness, offers a close-to-optimal performance, and also has a complexity linearly relating to the size of the problem.
- The independent multi-carrier power-constrained link-layer EE-EC tradeoff problem is then solved and analyzed for each user, given a subcarrier assignment matrix. The optimal power allocation strategy, which is across frequency and time domains, and the Pseudocode of the power allocation process are derived and proposed.
- We prove that each user's average optimal power level monotonically decreases with its EE requirement factor. Furthermore, we prove that each user's link-layer EE value monotonically decreases with its circuit power value, but increases with its EE requirement factor.
- Simulation results reveal that when there is a link-layer EE constraint, each user's operational tradeoff EC value<sup>1</sup> will not show a monotonic trend with its delay QoS exponent. Further, the tradeoff EC value achieved with a smaller number of available subcarriers may be higher than the one obtained with larger number of subcarriers.

<sup>1</sup>Here each user's operational tradeoff EC value is the calculated final EC value achieved at its EE requirement equality.

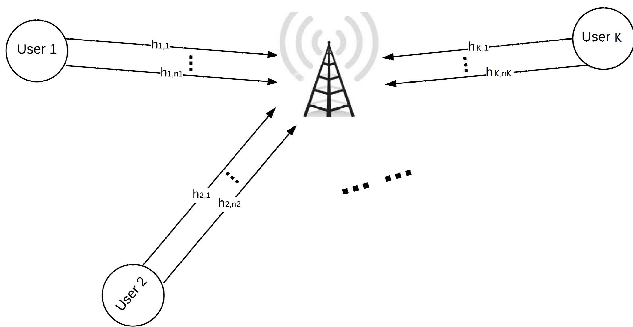


Fig. 1: Uplink transmission in a multi-user multi-carrier network.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. Multi-user Multi-carrier System Model

We consider the uplink transmission, where the  $K$  active users send their own information to the base station, in a multi-user multi-carrier OFDMA system depicted in Fig. 1. A total bandwidth of  $B$  is divided into  $N$  subcarriers, each with a bandwidth of  $\frac{B}{N}$ . Assume that each subcarrier is exclusively assigned to at most one user at each time to avoid interference among different users. The total number of allocated subcarriers for all users does not exceed the available frequency resources. Therefore, a feasible subcarrier assignment indicator matrix can be denoted as  $\phi$ , which satisfies

$$\phi \in \Phi \triangleq \left\{ [\phi_{k,n}]_{K \times N} \mid \phi_{k,n} \in \{0, 1\}, \sum_{k=1}^K \phi_{k,n} \leq 1, \sum_{k=1}^K \sum_{n=1}^N \phi_{k,n} \leq N, k \in \mathcal{K}_0, n \in \mathcal{N}_0 \right\}. \quad (1)$$

Here,  $\Phi$  denotes the set of all possible subcarrier allocation indicator matrices, and  $\mathcal{K}_0 = \{1, 2, \dots, K\}$ ,  $\mathcal{N}_0 = \{1, 2, \dots, N\}$  denote the set of all users and all subcarriers, respectively. The number of allocated subcarriers for the  $k^{\text{th}}$  user is denoted by  $N_k$ , namely,  $N_k = \sum_{n=1}^N \phi_{k,n}$ , and the bandwidth allocated to the  $k^{\text{th}}$  user is denoted by  $B_k$ , i.e.,  $B_k = N_k \frac{B}{N}$ .

Each transmitter implements a first-in-first-out (FIFO) buffer, which prevents loss of packets that could occur when the source rate is higher than the service rate, at the expense of increasing the delay [13]. The upper-layer packets are divided into frames at the data-link layer and are stored at the transmit buffer. The frames are then split into bit streams at the physical layer. By utilizing perfect channel state information (CSI) knowledge fed back from the receiver and the predetermined statistical QoS constraint, adaptive modulation and coding (AMC) and adaptive power control policy are applied at the transmitter side [17]. Then, the bit streams are read out of the buffer and are transmitted through the wireless fading subcarriers. At the receiver side, the reverse operations are performed and the frames are recovered for further processing. We assume that each subcarrier experiences block fading, i.e., the channel gains of  $N$  subcarriers are invariant within a

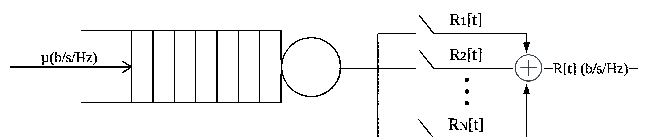


Fig. 2: Queuing system model for each transmitter.

fading-block's time duration  $T_f$ , but independently varies from one fading block to another. In addition, the length of each fading-block,  $T_f$ , is considered to be an integer multiple of the symbol duration  $T_s$ , and is assumed to be less than the fading coherence time [17].

For the  $k^{\text{th}}$  user on the  $n^{\text{th}}$  subcarrier at the fading-block index  $t$ , the subcarrier power gain is denoted by  $\gamma_{k,n}[t]$ ,  $k \in \mathcal{K}_0, n \in \mathcal{N}_0$ . Also, each subcarrier is assumed to experience i.i.d. additive white Gaussian noise (AWGN) with power spectral density  $\frac{\eta_0}{2}$ . Therefore, the instantaneous maximum achievable rate of the  $k^{\text{th}}$  user on the  $n^{\text{th}}$  subcarrier at the  $t^{\text{th}}$  fading-block is given by

$$R_{k,n}[t] = \frac{B}{N} T_f \log_2 \left( 1 + P_{k,n}[t] \frac{\gamma_{k,n}[t]}{P_{\mathcal{L}}^k \eta_0 \left( \frac{B}{N} \right)} \right) \text{ (bits)}, \quad (2)$$

where  $P_{\mathcal{L}}^k$  denotes the distance-based path-loss power and  $P_{k,n}[t]$  is the nonnegative transmission power for the  $k^{\text{th}}$  user on the  $n^{\text{th}}$  subcarrier, at the  $t^{\text{th}}$  fading-block, i.e.,  $P_{k,n}[t] \geq 0$ . Specifically, for the  $k^{\text{th}}$  user, the subcarrier power allocation vector is denoted as  $\mathbf{P}_k[t] = [P_{k,1}[t] \ P_{k,2}[t] \ \dots \ P_{k,N}[t]]^T$ . The total achievable rate over all allocated subcarriers for the  $k^{\text{th}}$  user, which depends on the subcarrier allocation indicator matrix  $\phi$  and the subcarrier power allocation vector  $\mathbf{P}_k$ , can be denoted as  $\mathbf{R}_k(\phi, \mathbf{P}_k) = \sum_{n \in \mathcal{N}_k} \phi_{k,n} R_{k,n}$ , where  $\mathcal{N}_k$  is the set of subcarriers allocated to the  $k^{\text{th}}$  user.

### B. Multi-user Multi-carrier Effective Capacity and Link-layer Energy Efficiency

For each transmitter, the FIFO buffer is assumed to be a dynamic queueing system with stationary ergodic arrival and service processes, depicted in Fig. 2 [22]. By using the large deviation theory, the queue length process  $Q(t)$  converges in distribution to a steady-state queue length  $Q(\infty)$  such that [22]

$$-\lim_{x \rightarrow \infty} \frac{\ln(\Pr\{Q(\infty) > x\})}{x} = \theta, \quad (3)$$

where  $\Pr\{a > b\}$  shows the probability that  $a > b$  holds. This definition implies that the probability of the queue length exceeding a certain threshold  $x$  decays exponentially fast as  $x$  increases [23]. Note that in (3), the parameter  $\theta$  ( $\theta > 0$ ) indicates the exponential decay rate of the QoS violation probability. A smaller value of  $\theta$  denotes a looser QoS requirement, while larger  $\theta$  implies a lower probability of violating the queue length and a more stringent delay constraint. Particularly, when

<sup>2</sup>Since the service rate process of the  $k^{\text{th}}$  user on the  $n^{\text{th}}$  subcarrier is considered to be stationary and ergodic [17], hereafter, the block index  $t$  could be omitted for simplicity.

$\theta \rightarrow 0$ , which refers to a system with no delay constraint, the optimum power allocation strategy is the traditional water-filling approach and the maximum achievable rate is ergodic capacity. For a transmitter with  $\theta \rightarrow \infty$ , the optimum power allocation is the channel inversion with fixed rate transmission technique, under which the delay-limited capacity can be achieved. In other words, the ergodic capacity and the delay-limited capacity can be considered as two extreme cases of the effective capacity.

Taking the delay experienced by a source packet arriving at time  $t$ , defined by  $D(t)$ , into consideration, the probability that the delay exceeds a maximum delay bound  $D_{\max}$ , can be estimated as [13]

$$P_{\text{delay}}^{\text{out}} = \Pr\{D(t) > D_{\max}\} \approx \Pr\{Q(t) > 0\}e^{-\theta\mu D_{\max}}, \quad (4)$$

where  $P_{\text{delay}}^{\text{out}}$  presents the delay-outage probability,  $D_{\max}$  is in the unit of a symbol period,  $\Pr\{Q(t) > 0\}$  denotes the probability of a non-empty buffer at time  $t$ , and can be approximated by the ratio of the constant arrival rate to the average service rate [17], [22], i.e.,  $\Pr\{Q(t) > 0\} \approx \frac{\mu}{\mathbb{E}[R(t)]}$ . Hence, in order to meet a target delay-outage probability limit  $P_{\text{delay}}^{\text{out}}$ , a source needs to limit its data rate to the maximum of  $\mu$ , where  $\mu$  is the solution to (4).

Assume that the Gartner-Ellis theorem [24, Pages 34-36] is satisfied. For the  $k^{\text{th}}$  user, the EC value, in b/s/Hz, over a multi-carrier transmission with a total bandwidth  $B_k$  can be expressed as [13]

$$E_c^k(\theta_k, \phi, \mathbf{P}_k) = -\frac{1}{\theta_k T_f B_k} \ln \left( \mathbb{E} \left[ e^{-\theta_k \mathbf{R}_k(\phi, \mathbf{P}_k)} \right] \right), \quad (5)$$

where  $\theta_k$  stands for the delay QoS exponent of the  $k^{\text{th}}$  user which is associated with the statistical delay QoS requirement and  $\mathbb{E}[\cdot]$  indicates the expectation operator. Henceforth, EC of the  $k^{\text{th}}$  user becomes a function of  $\theta_k$ ,  $\phi$ , and  $\mathbf{P}_k$ .

By expanding  $\mathbf{R}_k(\phi, \mathbf{P}_k)$  and inserting it into (5), EC of the  $k^{\text{th}}$  user can be further expressed as

$$E_c^k(\theta_k, \phi, \mathbf{P}_k) = -\frac{1}{\theta_k T_f B_k} \ln \left( \mathbb{E} \left[ e^{-\theta_k \sum_{n \in \mathcal{N}_k} \phi_{k,n} R_{k,n}} \right] \right). \quad (6)$$

For the multi-user OFDMA network, the overall EC value can be expressed as

$$E_c(\theta, \phi, \mathbb{P}) = \frac{\sum_{k=1}^K N_k E_c^k(\theta_k, \phi, \mathbf{P}_k)}{\sum_{k=1}^K N_k} \quad (\text{b/s/Hz}), \quad (7)$$

where  $\theta = [\theta_1 \ \theta_2 \ \dots \ \theta_K]$  is the  $K \times 1$  vector of delay exponents for all  $K$  users.  $\mathbb{P}$  denotes the transmission power allocation matrix, for all users over all subcarriers, i.e.,  $\mathbb{P} \in \mathcal{P} \triangleq \left\{ [P_{k,n}]_{K \times N} \in \mathbb{R}_+ \mid \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^N \phi_{k,n} P_{k,n} \right] \leq P_{\max}^k, k \in \mathcal{K}_0 \right\}$ . Here,  $\mathcal{P}$  is all the possible power allocation matrices,  $\mathbb{E}_{\gamma_k}[\cdot]$  indicates the expectation over the PDF of  $\gamma_k$ , where  $\gamma_k$  is the  $k^{\text{th}}$  user's subcarrier power gains, i.e.,  $\gamma_k = [\gamma_{k,1} \ \gamma_{k,2} \ \dots \ \gamma_{k,N_k}]$ .  $P_{\max}^k$  represents the maximum average power limit of the  $k^{\text{th}}$  user.

Moreover, for the  $k^{\text{th}}$  user, we define the link-layer EE as the ratio of EC to the sum of its circuit power  $P_c^k$ , and the average

transmission power scaled by the power amplifier efficiency  $\epsilon$ , yielding

$$\text{EE}^k(\theta_k, \phi, \mathbf{P}_k) = \frac{E_c^k(\theta_k, \phi, \mathbf{P}_k)}{P_c^k + \frac{1}{\epsilon} \mathbb{E}_{\gamma_k} \left[ \sum_{n \in \mathcal{N}_k} \phi_{k,n} P_{k,n} \right]}. \quad (8)$$

### C. Problem Formulation

From a system point of view, the overall EC value needs to be maximized to achieve the best system performance. On the other hand, from the individual user point of view, each user has its own link-layer EE requirement, average transmission power limit and delay QoS constraint. Therefore, considering a multi-user multi-carrier network, the overall system throughput maximization problem, subject to each user's resource constraints, can be formulated as

$$Q1 : \max_{\phi \in \Phi, \mathbb{P} \in \mathcal{P}} E_c(\theta, \phi, \mathbb{P}) \quad (9a)$$

$$\text{subject to: } \text{EE}^k(\theta_k, \phi, \mathbf{P}_k) \geq \eta_{\text{req}}^k, \forall k, \quad (9b)$$

$$\mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^N \phi_{k,n} P_{k,n} \right] \leq P_{\max}^k, \forall k, \quad (9c)$$

$$\sum_{k=1}^K \phi_{k,n} \leq 1, \forall n, \quad (9d)$$

$$\sum_{k=1}^K \sum_{n=1}^N \phi_{k,n} \leq N, \quad (9e)$$

$$\phi_{k,n} \in \{0, 1\}, \forall k, \forall n, \quad (9f)$$

$$P_{k,n} \geq 0, \forall k, \forall n, \quad (9g)$$

where  $\eta_{\text{req}}^k$  is the  $k^{\text{th}}$  user's required link-layer EE level, defined by a certain ratio of its maximum achievable link-layer EE value, i.e.,  $\eta_{\text{req}}^k = \chi_{\text{EE}}^k \times \eta_{\max}^{k,N}$ . Here,  $\eta_{\max}^{k,N} = \text{EE}^k \Big|_{\substack{N_k=N \\ P_k=P_{\text{EE}}^{k,*}}}$

denotes the  $k^{\text{th}}$  user's maximum achievable EE value, when all  $N$  subcarriers in the system are allocated to it.  $P_{\text{EE}}^{k,*}$  is the operational average input power which achieves  $\eta_{\max}^{k,N}$ . Further,  $\chi_{\text{EE}}^k \in [0, 1]$  is an adjustable EE requirement factor, which reveals the strictness of the  $k^{\text{th}}$  user's required EE level and directly influences the system performance. In particular,  $\chi_{\text{EE}}^k = 0$  indicates that the  $k^{\text{th}}$  user has no EE requirement, while  $\chi_{\text{EE}}^k = 1$  means that user  $k$  requires an operational EE value at  $\eta_{\max}^{k,N}$ . Since  $\eta_{\max}^{k,N}$  depends on the individual user's delay QoS exponent and its maximum average power limit, its value is different for each user. Therefore, the  $k^{\text{th}}$  user's required EE level  $\eta_{\text{req}}^k$  is different from the other users, even when they have the same EE requirement factors.

Due to the conflicting property of the total system EC and each user's personal EE achievement, after introducing  $\chi_{\text{EE}}^k$ , the formulated problem Q1 becomes an adjustable tradeoff problem. To be more specific, if the total system EC value has a high priority, each user's EE requirement factor value can be required to be very low, which results in a low link-layer EE level for each user. Correspondingly, if the total system EC value has a low priority, each user's EE requirement factor can be relatively high, so that each user will have a satisfied high level of link-layer EE.

### III. OPTIMAL AND SUB-OPTIMAL SOLUTIONS

Since we assume that one subcarrier can be assigned to only one user at a time, therefore there could be  $K^N$  possible subcarrier assignments [25]. Hence, the complexity of the above combinatorial integer programming problem in finding the jointly optimal subcarrier and power allocation grows exponentially with the number of subcarriers. Furthermore, we note that it is very difficult to jointly obtain the optimal subcarrier allocation sets and all power allocation values in every frame, due to the reasons below. Firstly, from (6), we can notice that the EC formulation of the  $k^{\text{th}}$  user not only requires the multiplication of two unknown parameters, i.e.,  $\phi_{k,n}$ , and  $R_{k,n}$ , but also involves the expectation over the joint PDF of all subcarriers' channel power gains, i.e.,  $\gamma_k$ . Secondly, the expectation and the multiplication operations cannot be interchanged, even if all subcarriers are assumed to be i.i.d., and that is because the power allocation value on each subcarrier is related to the other subcarriers.

Henceforth, in order to make the formulated problem  $Q1$  tractable, we divide the solving process into two steps: frequency provisioning which decides the number of subcarriers to be allocated to each user; and then optimal power allocation for each user over all its allocated subcarriers. Specifically, the proposed frequency provisioning algorithms, which are independent of the instantaneous CSI knowledge in each frame, will be implemented only once within a period of time. On the other hand, for each user, the proposed optimal power allocation strategy on each subcarrier, not only relies on the instantaneous CSI of this subcarrier, but also depends on the other subcarriers' CSI knowledge in each frame.

We start from introducing three frequency provisioning algorithms: traditional exhaustive algorithm, fair-exhaustive algorithm and our proposed low-complexity heuristic frequency allocation algorithm. After obtaining the subcarrier assignments, the optimal power allocation strategy for each single-user multi-carrier system will then be derived and obtained in Section III-B.

#### A. Frequency Provisioning Algorithms

By applying frequency provisioning, we assume that all subcarriers follows the same distribution. It is the number of designated subcarriers which matters, regardless where those subcarriers are located in the frequency band [25]. To reduce the problem complexity and the solving time, we first build a pre-calculated offline database  $\mathcal{D}$  which stores all users' maximum achievable link-layer EE values, i.e.,  $\eta_{\max} = [\eta_{\max}^1 \ \eta_{\max}^2 \ \dots \ \eta_{\max}^K]^T$ , in terms of certain settings of  $P_c$ ,  $\theta$  and  $N$ . Here  $\eta_{\max}^k$  is a  $1 \times N$  vector of the  $k^{\text{th}}$  user's maximum achievable EE values with different number of allocated subcarriers, i.e.,  $\eta_{\max}^k = [\eta_{\max}^{k,1} \ \eta_{\max}^{k,2} \ \dots \ \eta_{\max}^{k,N}]$ . Define  $\eta_{\text{req}} = [\eta_{\text{req}}^1 \ \eta_{\text{req}}^2 \ \dots \ \eta_{\text{req}}^K]^T$  as the  $K \times 1$  vector of the EE requirement values for all  $K$  users, then we will transform  $\eta_{\text{req}}$  to a  $K \times 1$  vector which specifies all users' required number of subcarriers, i.e.,  $S_{\text{req}} = [S_{\text{req}}^1 \ S_{\text{req}}^2 \ \dots \ S_{\text{req}}^K]^T$ . Let us consider the  $k^{\text{th}}$  user as an example. Its required link-layer EE value is denoted by  $\eta_{\text{req}}^k$ , and correspondingly, its subcarrier requirement value will be stored as  $S_{\text{req}}^k$ .

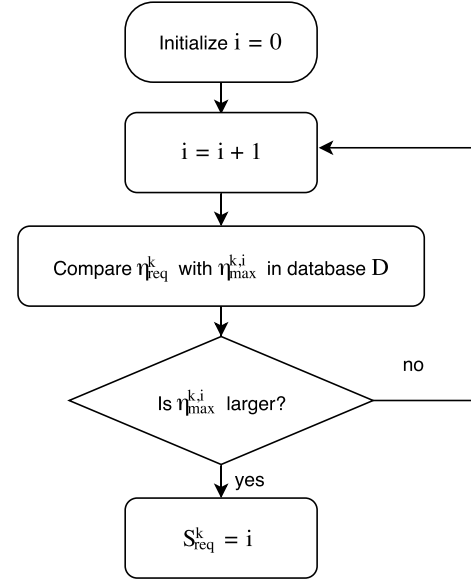


Fig. 3: Transform  $\eta_{\text{req}}^k$  to  $S_{\text{req}}^k$ .

To obtain  $S_{\text{req}}^k$ , we provide a flowchart in Fig. 3 to compare  $\eta_{\text{req}}^k$  with  $\eta_{\max}^k$ . If the maximum achievable EE value obtained with  $i$  subcarriers is larger than the required EE value, i.e.,  $\eta_{\max}^{k,i} \geq \eta_{\text{req}}^k$ , then we can conclude that the minimum number of subcarriers required to satisfy the  $k^{\text{th}}$  user's EE requirement  $\eta_{\text{req}}^k$ , is  $i$ , i.e.,  $S_{\text{req}}^k = i$ . Henceforth, all  $K$  users' EE requirements in  $\eta_{\text{req}}$  can be transformed to the subcarrier requirement vector  $S_{\text{req}}$ , by utilizing the flowchart in Fig. 3. In this way, the feasibility of each user's EE constraint can be easily checked by comparing the number of allocated subcarriers with the number of required subcarriers.

#### 1) Traditional Exhaustive Algorithm

The traditional way to solve an NP-hard problem, like the one we formulated in (9a)-(9g), is to carry out an exhaustive search, which systematically enumerates all possible combinations and finally locates the solution which optimizes the objective function and satisfies all the problem constraints [25]. Specifically, for problem  $Q1$ , the set of feasible combinations is found first. Then, the optimal power allocation strategy proposed in the next section, will be applied to all the feasible combinations. Finally, the feasible combination which offers the maximum system throughput will be chosen as the optimal solution. Although exhaustive search is able to find the optimal frequency provisioning solution, it also lacks user fairness and has a high computational complexity which exponentially grows with the size of the problem.

#### 2) Fair-Exhaustive Algorithm

To further find the optimal frequency provisioning solution which not only maximizes the total system EC value, while satisfying each user's link-layer EE requirement, but also serves the maximum number of users that can be allowed, we propose a fair exhaustive algorithm. Firstly, the sum of all users' required subcarriers is compared with the total number of subcarriers  $N$  to find the maximum number of users that can be served. For example, let us assume  $N = 8$ , and the subcarrier requirement vector for all users is  $[1, 2, 2, 4]$ . Hence, the total available subcarriers can serve 3 users at

most. Secondly, the set of feasible subcarrier allocation vectors is found, in which each allocation vector not only satisfies all served users' subcarrier requirements, but also serves the maximum allowed number of users. Then, the optimal power allocation strategy proposed in Section III-B will be applied to all feasible allocation vectors to locate the fair and optimal solution which outperforms the others.

Clearly, by enumerating all possible subcarrier allocation vectors which can serve the allowed maximum number of users, the above proposed algorithm exhaustively find the optimal solution in a fair way. Although the fair-exhaustive algorithm is less complex compared to the traditional exhaustive algorithm, but its computational complexity is still very high, especially when the number of available subcarriers  $N$  is large. To further reduce algorithm complexity, we provide the following heuristic algorithm, which is simple, fair and close-to-optimal.

### 3) Heuristic Algorithm

There are three steps included in the proposed heuristic frequency provisioning algorithm, which are allocation process, calculation process and check process. Firstly, in order to serve as many users as possible, in the allocation process, we start from the user which requires the minimum number of subcarriers. Each served user will be allocated the exact number of its required subcarriers, so that all the allocated users can satisfy their EE requirements. The allocation will be repeated until the remaining subcarriers run out, or there are not enough subcarriers to satisfy the next user's EE requirement, or all users' subcarrier requirements have already been satisfied. Then, the calculation process starts, in which each served user operates the optimal power allocation strategy described in Table II to obtain its corresponding EC value. In the check process, we aim to maximize the system throughput, based on the strategy that the user with current minimum EC value has the allocation priority. Therefore, the remaining subcarriers will be assigned one-by-one to the user who has the current minimum EC value, until all subcarriers run out.

Assume the final subcarrier allocation vector is denoted by  $N = [N_1 \ N_2 \ \dots \ N_K]$ . The Pseudocode of the proposed heuristic algorithm is illustrated in Table I. We note that, the proposed algorithm only needs at most  $K - 1$  comparisons per iteration, given that each user's EC values with various number of subcarriers, is pre-calculated off-line and is stored in a database. Therefore, the heuristic algorithm offers a relatively low computational complexity comparing to the two exhaustive algorithms whose complexity exponentially increase with the number of subcarriers. On the other hand, later, in simulation results, we will demonstrate that the proposed low-complexity algorithm offers a close performance with the fair-exhaustive algorithm.

Now, let us analyze and explain the strategies utilized in the proposed heuristic algorithm. Firstly, in the allocation process, the heuristic algorithm starts the allocation from the user which has the minimum subcarrier requirement. Assume that user  $i$  has the relatively small number of required subcarriers,  $S_{\text{req}}^i$ . By regarding  $\theta_i$  and  $\chi_{\text{EE}}^i$  as the two influencing parameters on  $S_{\text{req}}^i$ , a small value of  $S_{\text{req}}^i$  may result from the following two possibilities: 1) user  $i$  has a small delay QoS exponent

TABLE I: Heuristic Algorithm

#### Initialization:

Calculate  $S_{\text{req}}$ , using  $\eta_{\text{req}}$  and the pre-calculated database  $\mathcal{D}$ .

Define  $S_{\text{tol}} = N$ ,  $H = S_{\text{req}}$ .

#### Allocation Process:

**While**  $S_{\text{tol}} > 0$

**If**  $H = 0$

Break;

**End**

Find  $H_i = \min(H)$ , and  $H_i > 0$ ;

**If**  $S_{\text{tol}} > S_{\text{req}}^i$

$N_i = S_{\text{req}}^i$ ;

$S_{\text{tol}} = S_{\text{tol}} - S_{\text{req}}^i$ ;

$H_i = 0$ ;

**Else**

Break;

**End**

**End**

#### Calculation Process:

For each user  $i$  with  $H_i = 0$ , apply the optimal power allocation process in Table II.

Calculate the  $i^{\text{th}}$  user's EC value  $J_i$  and define

$J = [J_1 \ J_2 \ \dots \ J_K]$ .

#### Check Process:

**While**  $S_{\text{tol}} > 0$

Find  $J_i = \min(J)$ , in which user  $i$  satisfies  $H_i = 0$ ;

$N_i = N_i + 1$ ;

Apply the optimal power allocation process to user  $i$  and update  $J_i$ .

**End**

**Output:**  $N$ ;  $E_c$  given in (7).

$\theta_i$  and the same  $\chi_{\text{EE}}^i$  value, comparing to the other users; 2) user  $i$  has a small EE requirement factor  $\chi_{\text{EE}}^i$ , and the same  $\theta_i$  value, comparing to the others. For the first situation, a small delay QoS exponent  $\theta_i$  means a loose requirement on delay QoS, which will offer a bigger EC value, when the allocated number of subcarriers and  $\chi_{\text{EE}}^i$  are fixed. Meanwhile, for the second situation, a small value of  $\chi_{\text{EE}}^i$  also provides a larger EC value, because now the EE requirement constraint is easy to be satisfied and the multi-carrier system will have more resource and flexibility to maximize the EC performance. Consequently, the design idea of the allocation process not only makes sure that as many users as possible can be served, but also intends to serve the user which can contribute a larger EC value.

On the other hand, the design strategy of the check process, i.e., the user with current minimum EC value has the allocation priority, comes from Fig. 4, which describes the results of maximum EC versus delay QoS exponent  $\theta$ , for various values of  $N$ , in a single-user multi-carrier system. Specifically, Fig. 4 reveals that the user with current minimum EC value has a high possibility to offer the largest EC-increase, if given one more subcarrier. In more detail, from Fig. 4, we notice that for two users with the same values of  $\theta$ , if we allocate more subcarriers to them, the user with current smaller EC value, i.e., the one which has smaller number of subcarriers, will get a larger EC-increase. Furthermore, for two users with the same

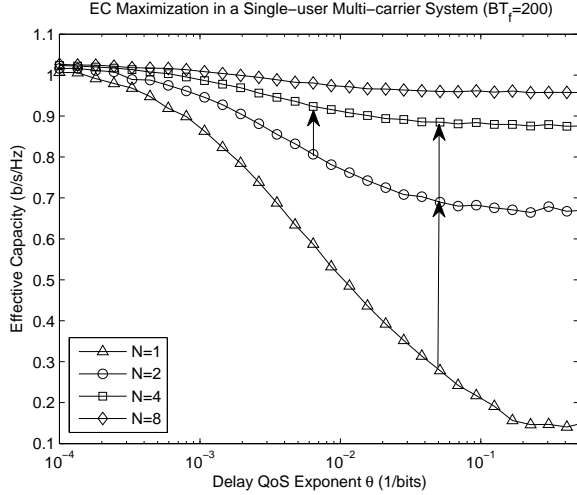


Fig. 4: Effective capacity versus delay QoS exponent  $\theta$ , for various values of  $N$ .

number of subcarriers, when we allocate each user two more subcarriers, the user with relatively smaller EC value, namely, the one which has larger delay QoS exponent, will provide a larger EC-increase. Simulation results in Section IV confirm the effectiveness of our design method, and inform that the proposed heuristic algorithm offers very close performance with the fair-exhaustive algorithm.

### B. Optimal Power Allocation For A Single-user Multi-carrier System

Given a subcarrier assignment matrix  $\phi$ , the multi-user OFDMA system can be viewed as a frequency-division multiple access (FDMA) system, where each user transmits data through a number of assigned subcarriers independently [26]. Therefore, the original total EC maximization problem, subject to each user's link-layer EE requirement and maximum average power limit, can be transformed into a link-layer EE-EC tradeoff problem for each single-user multi-carrier system.

Specifically, for the  $k^{\text{th}}$  user, the problem can be expressed as

$$Q2: \max_{\substack{P_{k,n} \geq 0 \\ n \in \mathcal{N}_k}} E_c^k(\theta_k, \mathbf{P}_k) \quad (10a)$$

$$\text{s.t. } EE^k(\theta_k, \mathbf{P}_k) \geq \eta_{\text{req}}^k, \quad (10b)$$

$$\mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n} \right] \leq P_{\max}^k. \quad (10c)$$

By recalling that the total bandwidth allocated to the  $k^{\text{th}}$  user is  $B_k$ , the total instantaneous service rate of the  $k^{\text{th}}$  user is given by

$$R_k = \frac{B_k}{N_k} T_f \sum_{n=1}^{N_k} \log_2 \left( 1 + P_{k,n} \frac{\gamma_{k,n}}{P_{\ell}^k \eta_0 \left( \frac{B_k}{N_k} \right)} \right) \text{ (bits)}. \quad (11)$$

By inserting (11) into (5), we get the mathematical expression of EC for the  $k^{\text{th}}$  user. Correspondingly, the link-layer EE for the  $k^{\text{th}}$  user, as the ratio of EC to the total power expenditure,

can be obtained. Therefore, problem Q2 can be expanded as

$$Q3: \max_{\substack{P_{k,n} \geq 0 \\ n \in \mathcal{N}_k}} -\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} \left( 1 + N_k P_{k,n}^r \gamma_{k,n} \right)^{-\frac{\alpha_k}{N_k}} \right] \right) \quad (12a)$$

$$\text{s.t. } \frac{-\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} \left( 1 + N_k P_{k,n}^r \gamma_{k,n} \right)^{-\frac{\alpha_k}{N_k}} \right] \right)}{K_{\ell}^k \left( P_{c_r}^k + \frac{1}{\epsilon} \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right] \right)} \geq \eta_{\text{req}}^k, \quad (12b)$$

$$K_{\ell}^k \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right] \leq P_{\max}^k, \quad (12c)$$

where  $\alpha_k \equiv \frac{\theta_k T_f B_k}{\ln(2)}$ ,  $P_{k,n}^r = \frac{P_{k,n}}{K_{\ell}^k}$ , and  $P_{c_r}^k = \frac{P_c^k}{K_{\ell}^k}$ . Here  $K_{\ell}^k = P_{\ell}^k \eta_0 B_k$ , which denotes the path loss factor, including both AWGN power and path loss power. Set  $\hat{\eta}_{\text{req}}^k = K_{\ell}^k \eta_{\text{req}}^k$ , and  $\hat{P}_{\max}^k = P_{\max}^k / K_{\ell}^k$ . Then,  $K_{\ell}^k$  in (12a)-(12c) can be canceled to scale the system performance with respect to the path loss factor.

From (12a)-(12c), one can notice that the EC expression in a single-user multi-carrier system is not a linear summation of each subcarrier's achievable EC value. Hence, the concavity and monotonicity of the EC function in a single-subcarrier system cannot be simply extended to the multi-carrier system. In order to find the joint energy and spectral efficient power allocation strategy in a single-user multi-carrier system, we start from analyzing the proposed problem Q3.

By referring to the scaled multi-carrier transmit power vector as  $\mathbf{P}_k^r = [P_{k,1}^r, P_{k,2}^r, \dots, P_{k,N}^r]$ , we note that the objective function (12a) is concave in  $\mathbf{P}_k^r$  [18]. Then, the link-layer EE, as the ratio of a concave function over a non-negative affine function in  $\mathbf{P}_k^r$ , is a quasi-concave function in subcarrier power allocations [18]. Therefore, its upper contour set defined by (12b) is convex [27]. Hence, (12a)-(12c) is a concave optimization problem and the Karush-Kuhn-Tucker (KKT) conditions are both sufficient and necessary for the global optimum value. Specifically, the proposed optimal power allocation strategy for the  $k^{\text{th}}$  user is related to the joint probability density function (PDF) of the subcarrier power gains  $\gamma_k$ , given by  $\rho(\gamma_k)$ .

To solve the concave optimization problem (12a)-(12c), we start from analyzing the power-unconstrained problem (12a)-(12b), which paves the way for the power-constrained optimization problem. By transforming (12b) to

$$-\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} \left( 1 + N_k P_{k,n}^r \gamma_{k,n} \right)^{-\frac{\alpha_k}{N_k}} \right] \right) - \hat{\eta}_{\text{req}}^k \left( P_{c_r}^k + \frac{1}{\epsilon} \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right] \right) \geq 0, \quad (13)$$

we get the Lagrangian function as follows

$$\begin{aligned} \mathcal{L}(\mathbf{P}_k^r, \lambda) = & -\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right] \right) \\ & + \lambda \left( -\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right] \right) \right. \\ & \left. - \hat{\eta}_{\text{req}}^k \left( P_{c_r}^k + \frac{1}{\epsilon} \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right] \right) \right) - \sum_{n=1}^N \mu_n P_{k,n}^r, \end{aligned} \quad (14)$$

where  $\lambda \in R$  is the Lagrange multiplier associated to (13) and  $\mu_n$  is the Lagrange multiplier associated to the constraint  $P_{k,n}^r \geq 0, \forall n \in \mathcal{N}_k$ .

At the optimal power allocation, we have

$$\frac{\partial \mathcal{L}(\mathbf{P}_k^r, \lambda)}{\partial \mathbf{P}_k^r} = 0. \quad (15)$$

Because of the complementary slackness condition [27], if  $P_{k,n}^r > 0$ , then  $\mu_n = 0, \forall n \in \mathcal{N}_k$ . On the other hand, if  $P_{k,n}^r = 0, \exists n \in \mathcal{N}_k$ , then  $\mu_n \neq 0$ . Thus, the following two cases need to be considered to find the optimal power allocation strategy.

1) *Case 1:*  $P_{k,n}^r > 0, \forall n \in \mathcal{N}_k$

In this case, all  $N_k$  subcarriers are allocated non-zero transmission power. Therefore, based on the complementary slackness,  $\{\mu_n\}_{n=1}^{N_k} = 0$ . Then, the KKT condition (15) can be simplified as

$$\prod_{i=1}^{N_k} (1 + N_k P_{k,i}^r \gamma_{k,i})^{-\frac{\alpha_k}{N_k}} = \frac{\beta}{\gamma_{k,n}} (1 + N_k P_{k,n}^r \gamma_{k,n}), \quad \forall n \in \mathcal{N}_k, \quad (16)$$

$$\text{where } \beta = \frac{\lambda \hat{\eta}_{\text{req}}^k}{\epsilon(\lambda + 1) \log_2 e} \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right].$$

By multiplying the right and left-hand sides of the  $N_k$  equations in (16), the optimal power allocation strategy can be obtained as

$$P_{k,n}^r = \frac{1}{N_k} \left[ \frac{1}{\beta^{\frac{1}{\alpha_k+1}} \prod_{i=1}^{N_k} \gamma_{k,i}^{\frac{\alpha_k}{(\alpha_k+1)N_k}}} - \frac{1}{\gamma_{k,n}} \right], \forall n \in \mathcal{N}_k. \quad (17)$$

The derived power allocation strategy (17) is optimal only when all subcarriers are assigned with positive powers. If there are one or more subcarriers which are allocated non-positive powers, then the second case needs to be taken into consideration.

2) *Case 2:*  $P_{k,j}^r = 0, \exists j \in \mathcal{N}_k$

If there exists  $P_{k,j}^r \leq 0$ , then the set of subcarriers, which only positive powers should be assigned, needs to be found.

Firstly, we define  $\hat{\mathcal{N}}_k$  as

$$\hat{\mathcal{N}}_k = \left\{ n \in \mathcal{N}_k \mid \frac{1}{N_k} \left[ \frac{1}{\beta^{\frac{1}{\alpha_k+1}} \prod_{i=1}^{N_k} \gamma_{k,i}^{\frac{\alpha_k}{(\alpha_k+1)N_k}}} - \frac{1}{\gamma_{k,n}} \right] \geq 0 \right\}.$$

According to Lemma 1 in [17], the total power must be assigned to the subcarriers which belong to  $\hat{\mathcal{N}}_k$ , while the

subcarriers  $n \notin \hat{\mathcal{N}}_k$  should not be allocated any power. Therefore, a new power-unconstrained optimization problem could be expressed as

$$Q4 : \max_{\substack{P_{k,n}^r \geq 0 \\ n \in \hat{\mathcal{N}}_k}} -\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{\hat{N}_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right] \right) \quad (18a)$$

$$\text{s.t. } \frac{-\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{\hat{N}_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right] \right)}{K_\ell^k \left( P_{c_r}^k + \frac{1}{\epsilon} \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{\hat{N}_k} P_{k,n}^r \right] \right)} \geq \eta_{\text{req}}^k, \quad (18b)$$

where  $\hat{N}_k = |\hat{\mathcal{N}}_k|$  represents the cardinality of  $\hat{\mathcal{N}}_k$ .

Therefore, if  $P_{k,n}^r > 0, \forall n \in \hat{\mathcal{N}}_k$ , then, the optimization problem can be solved exactly like Case 1. Otherwise, if there are subcarriers  $n \in \hat{\mathcal{N}}_k$  having  $P_{k,n}^r = 0$ , then  $\hat{\mathcal{N}}_k$  must be further partitioned by recursively repeating the above process until a set  $\mathcal{N}_k^*$  can be found, in which all subcarriers are allocated positive powers [18].

After obtaining  $\mathcal{N}_k^*$ , the optimal power allocations are computed as

$$P_{k,n}^r = \begin{cases} \frac{1}{N_k} \left[ \frac{1}{\beta^{\frac{1}{N_k+\alpha_k N_k^*}} \prod_{i \in \mathcal{N}_k^*} \gamma_{k,i}^{\frac{\alpha_k}{N_k+\alpha_k N_k^*}}} - \frac{1}{\gamma_{k,n}} \right], & n \in \mathcal{N}_k^* \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

where  $N_k^* = |\mathcal{N}_k^*|$ .

The optimal value for  $\beta$ , referred to as  $\beta^*$ , is found when the  $k^{\text{th}}$  user's EE constraint is satisfied with equality, yielding

$$\begin{aligned} & -\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right] \right) \\ & - \hat{\eta}_{\text{req}}^k \left( P_{c_r}^k + \frac{1}{\epsilon} \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right] \right) = 0. \end{aligned} \quad (20)$$

Note that since EE versus EC is a bell shape curve, the required EE level, if possible, can be achieved at two different EC values, which means that there will be two solutions for  $\beta$ , i.e.,  $\beta_1$  and  $\beta_2$ , to satisfy (20). Assume that  $\bar{P}_{k1} = \bar{P}_k |_{\beta=\beta_1}$ , and  $\bar{P}_{k2} = \bar{P}_k |_{\beta=\beta_2}$ , where  $\bar{P}_k$  stands for  $K_\ell^k \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right]$ . Therefore, the feasible set of the average input power level satisfying the EE constraint (12b) can be written as  $[\bar{P}_{k1}, \bar{P}_{k2}]^3$ . Considering our intention to maximize EC and the fact that EC is a monotonically increasing function in  $\bar{P}_k$  [18], the optimal average input power value  $\bar{P}_k^*$ , which solves the power-unconstrained problem (12a)-(12b), is chosen as the larger one which satisfies (20), i.e.,  $\bar{P}_k^* = \max[\bar{P}_{k1}, \bar{P}_{k2}]$ . Based on the assumption that  $\bar{P}_{k2}$  is larger than  $\bar{P}_{k1}$ , therefore  $\bar{P}_k^* = \bar{P}_{k2}$ , and correspondingly,  $\beta^* = \beta_2$ . Here we complete the solving process of the optimal power allocation for the

<sup>3</sup>Without losing any generality, we assume that  $\bar{P}_{k2}$  is larger than  $\bar{P}_{k1}$ .



TABLE II: Optimal Power Allocation Process

<b>Input:</b> $[\phi, \theta_k, T_f, B, N, N_k, P_c^k, \epsilon, K_\ell^k, \gamma_k, P_{\max}^k, \eta_{\text{req}}^k]$
<b>Step 1:</b> Have a initial guess of $\beta$ . <b>Repeat</b> Create (20), using (17) or (19), which applies Monte Carlo method. Update $\beta$ using bisection method. <b>Until</b> find $\beta^*$ which solves (20). Calculate $P_{k,n}^r, n \in \mathcal{N}_k$ . Calculate $\overline{P}_k^* = K_\ell^k \mathbb{E}_{\gamma_k} \left[ \sum_{n=1}^{N_k} P_{k,n}^r \right] \Big _{\beta=\beta^*}$ .
<b>Step 2:</b> If $P_{\max}^k < \overline{P}_k^*$ . Create $\overline{P}_k^* = P_{\max}^k$ and update $\beta^*$ , correspondingly. Calculate $P_{k,n}^r, n \in \mathcal{N}_k$ , in (17) or (19).
<b>Step 3:</b> Calculate the value of $E_c^k$ given in (6) and the link-layer EE <sup>k</sup> value in (8).
<b>Output:</b> $[P_{k,n}^r, \overline{P}_k^*, E_c^k, \text{EE}^k]$

power-unconstrained problem (12a)-(12b).

By utilizing the above proposed optimal power allocation strategy, we start to analyze the optimization problem (12a)-(12c) with the average input power constraint. After the feasible set of the average power value for the EE constraint (12b) is found, the power-constrained EC maximization problem for the  $k^{\text{th}}$  user, subject to a link-layer EE constraint, can be simplified to

$$Q5: \max_{\substack{P_{k,n}^r \geq 0 \\ n \in \mathcal{N}_k}} -\frac{1}{\alpha_k} \log_2 \left( \mathbb{E}_{\gamma_k} \left[ \prod_{n=1}^{N_k} (1 + N_k P_{k,n}^r \gamma_{k,n})^{-\frac{\alpha_k}{N_k}} \right] \right) \quad (21a)$$

$$\text{s.t. } \overline{P}_k \in [\overline{P}_{k1}, \overline{P}_{k2}], \quad (21b)$$

$$\overline{P}_k \leq P_{\max}^k. \quad (21c)$$

We note that EC is a monotonically increasing function in  $\overline{P}_k$  [18], therefore, the optimal average power value which solves the problem in (21a)-(21c) will be achieved at one of the three endpoint values, i.e.,  $\overline{P}_{k1}$ ,  $\overline{P}_{k2}$ , or  $P_{\max}^k$ . In more detail, if  $\overline{P}_{k2} \leq P_{\max}^k$ , the optimal power level  $\overline{P}_k^*$  equals to  $\overline{P}_{k2}$ , and the optimal power allocation strategy (17) will be achieved and operated. On the other hand, if  $\overline{P}_{k1} < P_{\max}^k < \overline{P}_{k2}$ , the system has to operate at  $P_{\max}^k$  and the optimal power allocation to solve (12a)-(12b) is according to (17), wherein, optimal  $\beta^*$  is found such that  $\overline{P}_k^*|_{\beta=\beta^*} = P_{\max}^k$ . Moreover, if  $P_{\max}^k < \overline{P}_{k1}$ , the power-constrained problem Q5 has no feasible solution. For simplicity, we assume that each user's maximum available power is always sufficient to support the feasibility of its required EE value, i.e.,  $P_{\max}^k \geq \overline{P}_{k1}$ . Otherwise, the proposed problem will be infeasible.

To summarize, the Pseudocode of the optimal power allocation process to solve the power-constrained link-layer EE-EC tradeoff problem for the  $k^{\text{th}}$  user, through multiple subcarriers, is illustrated in Table II. After we obtain the optimal power allocation strategy and optimal operational average power

for problem Q3, further analysis is needed to thoroughly understand and investigate the impact of the  $k^{\text{th}}$  user's circuit power value and the EE requirement factor on its link-layer EE-EC tradeoff performance. Hence, we provide the following lemmas<sup>4</sup>.

### C. The effects of $P_c^k$ and $\chi_{EE}^k$ on the $k^{\text{th}}$ user's EE-EC tradeoff performance

**Lemma 1:** The  $k^{\text{th}}$  user's tradeoff link-layer EE value  $\text{EE}(\overline{P}_k^*)$  decreases with  $P_c^k$ .

*Proof:* The proof is provided in Appendix A. ■

Furthermore, the tradeoff optimal power value and the system performance can also be influenced by the introduced EE requirement factor. Specifically, when  $\chi_{EE}^k$  increases, the required link-layer EE level increases. Therefore, the final operational link-layer EE value which satisfies the EE requirement equality increases. Since the proposed tradeoff average power operates at the EE-EC conflicting region, therefore the corresponding EC value will decrease due to the increase in EE level. Hence, we can obtain the following lemma 2.

**Lemma 2:** The optimal average power value  $\overline{P}_k^*$  monotonically decreases with  $\chi_{EE}^k$ , but the corresponding link-layer EE value  $\text{EE}(\overline{P}_k^*)$  increases with  $\chi_{EE}^k$ .

*Proof:* The proof follows the above explanations and is omitted here due to page limit. ■

## IV. SIMULATION RESULTS

In this section, we simulate the uplink transmission in a multi-user multi-subcarrier system, in which the fading statistics of the different subcarriers are considered to be i.i.d. Rayleigh distributed such that the subcarrier power gains are realized as exponential random variables with unit mean. We numerically evaluate and compare the performance of the exhaustive algorithm, the fair-exhaustive algorithm, the heuristic algorithm and the algorithm proposed in [25], on the total system EC-maximization problem, under the constraints of each user's link-layer EE requirement and average transmission power limit. To further analyze the problem and confirm the lemmas proved in Section III-B, the impact of the delay QoS exponent  $\theta$ , the EE requirement factor  $\chi_{EE}$  and the circuit-to-noise power ratio  $P_{cr}$  on each user's operational EC value and the total system EC performance is simulated and analyzed. In the following simulations, we assume that  $B \cdot T_f = 200$ , the power amplifier efficiency  $\epsilon = 1$ , each user's individual average transmission power limit  $P_{\max} = 10\text{dB}$ , unless otherwise indicated.

In order to show the performance of the proposed heuristic algorithm, Fig. 5 shows the results of the total system EC versus the number of subcarriers  $N$ , for the heuristic algorithm, the exhaustive algorithm and the fair-exhaustive algorithm. To get Fig. 5, the number of users  $K$  is fixed, i.e.,  $K = 4$ , in which all users have the same settings of EE requirement factor, i.e.,  $\chi_{EE} = 0.7$ , and the circuit-to-noise power ratio, i.e.,  $P_{cr} = -10\text{dB}$ . The delay QoS exponent vector  $\theta$  is given by

<sup>4</sup>In these lemmas, we omit the influence of  $P_{\max}^k$ , by assuming that it is large enough to support the optimal power allocation strategy, i.e.,  $P_{\max}^k \geq \overline{P}_k^*$ .

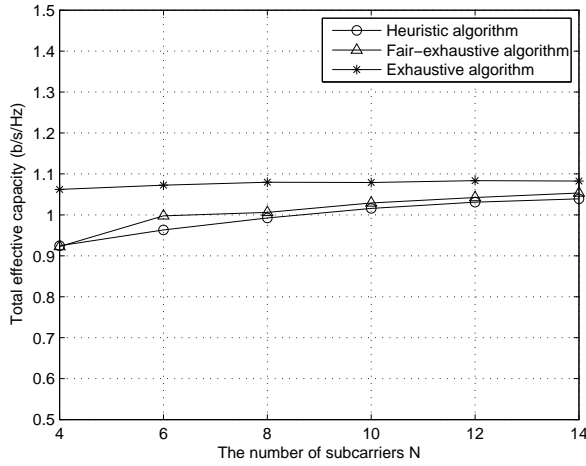


Fig. 5: The total effective capacity versus the number of subcarriers  $N$ , for heuristic algorithm, exhaustive algorithm and fair-exhaustive algorithm.

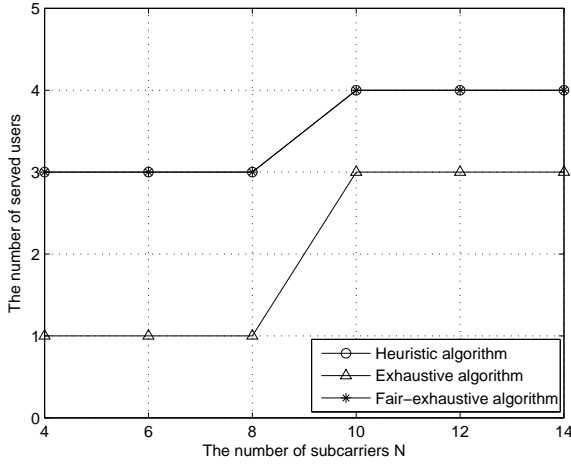


Fig. 6: The number of served users versus the number of subcarriers  $N$ , for heuristic algorithm, exhaustive algorithm and fair-exhaustive algorithm.

[0.001, 0.001, 0.01, 0.01]. For the exhaustive algorithm, when the number of subcarriers  $N$  increases, the total system EC value does not change very much, due to the loose delay QoS requirements for all the users. For fair-exhaustive algorithm and heuristic algorithm, the total EC performance curves are very close. This indicates that the proposed heuristic algorithm not only has a low complexity and guarantees user fairness, but also offers close-to-optimal performance.

To further compare the three algorithms, the plots for the number of served users versus the number of subcarriers  $N$  are included in Fig. 6. Although the exhaustive algorithm offers the best system performance in Fig. 5, Fig. 6 indicates that it serves the least number of users among all three algorithms. Especially, for the exhaustive algorithm, when  $N \in [4, 8]$ , it allocates all subcarriers to only one user, which shows a lack of fairness. On the contrary, for the heuristic algorithm and the fair-exhaustive algorithm, the number of served users shows an increasing trend until it equals to the total number of users  $K$ , i.e., 4. This happens because the increase of  $N$

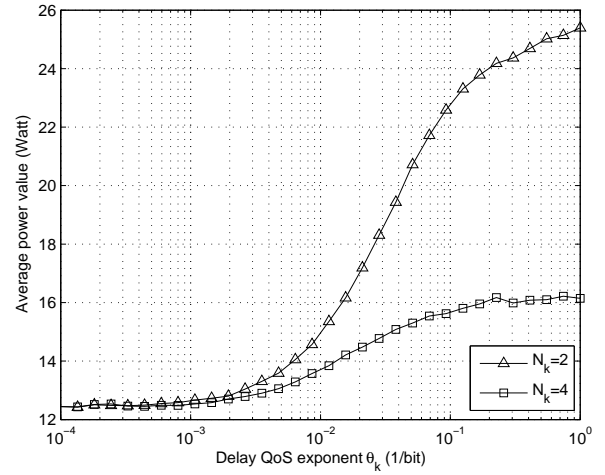


Fig. 7: Average tradeoff optimal power value versus delay QoS exponent  $\theta_k$ , for various values of  $N_k$ .

means more available frequency resources and the ability of supporting more users increases. We further note that normally, the traditional exhaustive algorithm prefers to choose the best user and allocates all subcarriers to it. However, Fig. 6 indicates that when  $N$  increases from 8 to 14, the number of served users for the exhaustive algorithm increases from 1 to 3 and then stays stable. This means that we may not have one single best user, when all users' delay QoS exponent values are different.

To understand the above phenomenon thoroughly, we consider the  $k^{\text{th}}$  user's multi-carrier system, and plot Fig. 7 and Fig. 8 which respectively include the curves of the average power versus  $\theta_k$ , and the tradeoff EC value versus  $\theta_k$ , for two different values of  $N_k$ , with  $\chi_{EE}^k = 0.2$ , and  $P_{cr}^k = -10\text{dB}$ . From Fig. 7, we note that with a fixed  $N_k$ , when  $\theta_k$  increases, the average power value increases. To explain this, we first recall that, the  $k^{\text{th}}$  user's EE requirement value is defined as a multiplication of  $\chi_{EE}^k$  and  $\eta_{\max}^{k, N_k}$ , in which  $\eta_{\max}^{k, N_k}$  is a function of  $\theta_k$  and  $N_k$ . With the fixed values of  $N_k$  and  $\chi_{EE}^k$ ,  $\eta_{\max}^{k, N_k}$  decreases with  $\theta_k$  [18], and in turn, the EE requirement value decreases. Furthermore, the curve of link-layer EE versus average power becomes wider when the user's delay QoS exponent becomes more stringent [15]. Therefore, when  $\theta_k$  increases, the optimal tradeoff average power obtained at a reduced EE requirement equality will become larger. Furthermore, Fig. 7 indicates that with a fixed value of  $\theta_k$ , when  $N_k$  becomes larger, the average power value reduces. This is due to the fact that when the values of  $\theta_k$  and  $\chi_{EE}^k$  are fixed,  $\eta_{\max}^{k, N_k}$  increases with  $N_k$  [18], as well as the required EE level. From Fig. 1 in [15], we note that a larger EE requirement will be satisfied at a smaller average power value. Hence, more available number of subcarriers could lead to less average power consumption.

Fig. 8 shows the relationship between the  $k^{\text{th}}$  user's tradeoff EC value and  $\theta_k$  for a single-user multi-carrier transmission system. This figure reveals two important conflicting situations and some insightful conclusions. Firstly, this figure indicates that one user's operational EC value will not show a monotonic trend with its delay QoS exponent, when there is a link-layer EE constraint. This phenomenon violates the monotonic trend

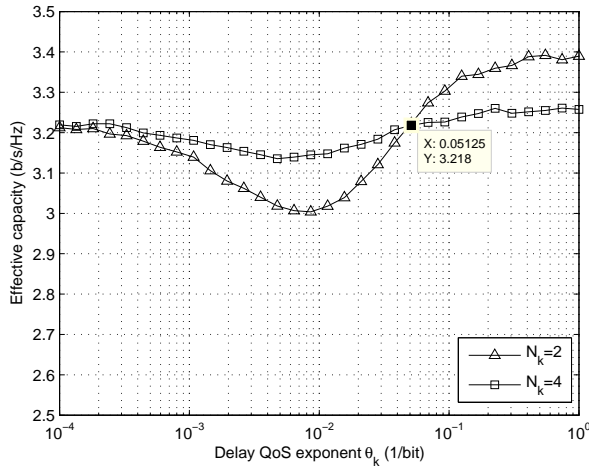


Fig. 8: Effective capacity versus delay QoS exponent  $\theta_k$ , for various values of  $N_k$ .

of EC versus delay QoS exponent, in the EC-maximization situation provided in [17]. From [17], we note that for a fixed delay QoS exponent, the EC value increases monotonically with the transmission power. Also, for a fixed transmission power, the EC value monotonically decreases with the delay QoS exponent. However, in our case, when  $\theta_k$  is small, the  $k^{\text{th}}$  user's link-layer EE requirement can be easily satisfied with a small value of transmission power. In contrast, when  $\theta_k$  becomes stringent, the required EE value has to be satisfied with a very large power value, like the trend indicated in Fig. 7. In other words, the operational average power value will increase with  $\theta_k$ . But, the increase of  $\theta_k$  and the increase of the average power will have a conflicting influence on the user's operational EC value. Therefore, with the inconsistent influence of these two parameters, EC will not show a monotonic trend, which can be confirmed from Fig. 8. Clearly, when  $\theta_k$  is loose, the tradeoff EC value will be more influenced by  $\theta_k$ . On the contrary, when  $\theta_k$  becomes stringent, the average power dominates the situation, therefore the operational EC value shows an increasing trend.

Secondly, Fig. 8 further reveals that one user's tradeoff EC value achieved at a smaller number of subcarriers may be higher than the one obtained with relatively larger number of subcarriers, when there is a link-layer EE constraint. Specifically, when  $\theta_k$  is loose, e.g.,  $\theta_k \in [10^{-4}, 0.05125]$ , the tradeoff EC value with 4 subcarriers is higher than the one obtained with 2 subcarriers. When  $\theta_k$  becomes stringent, e.g.,  $\theta_k \in [0.05125, 10^0]$ , the tradeoff EC value achieved with 4 subcarriers is lower than the one obtained with 2 subcarriers. This phenomenon also violates the monotonic trend of EC versus the number of subcarriers in EC-maximization situation analyzed in [17]. This is due to the fact that with a link-layer EE requirement, when  $N_k$  increase, the average power value required to satisfy the EE constraint decreases. Since the increase of  $N_k$  and the corresponding decrease of the average power will have a conflicting influence on the user's operational EC value, the EC will not show a monotonic trend. Apparently, Fig. 8 indicates that when  $\theta_k$  is loose, the tradeoff EC value will be more influenced by  $N_k$ . When  $\theta_k$  becomes stringent, the average power dominates the situation,

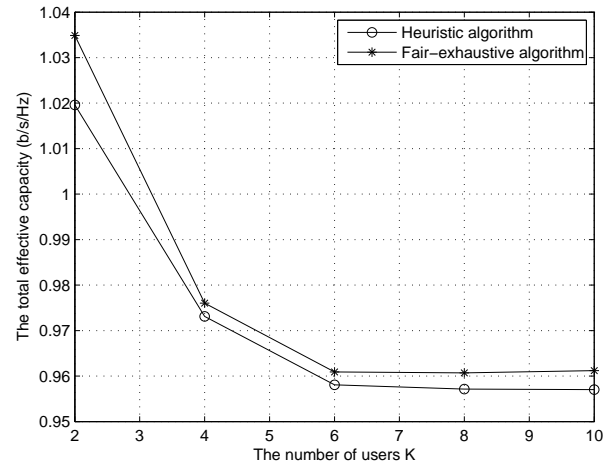


Fig. 9: The total effective capacity versus the number of users  $K$ , for heuristic algorithm and exhaustive algorithm.

therefore the operational EC value follows the same trend with the average power. In conclusion, Fig. 7 and Fig. 8 indicate that when there is a link-layer EE requirement, each user's operational tradeoff EC value may not show a monotonic trend with its delay QoS exponent value or its available number of subcarriers. This reveals that we may have multiple best users, when all factors vary. Hence, in Fig. 6, the exhaustive algorithm starts to serve more than one user when  $N > 8$ .

To examine the effect of the number of users on a multi-user multi-carrier system with limited resources, Fig. 9 includes the plots for the total EC value versus the number of users  $K$ , for the heuristic algorithm and the fair-exhaustive algorithm. Specifically, the total number of available subcarriers is fixed at  $N = 10$ . All users are assumed to have the same circuit-to-noise power ratio  $P_{cr} = -10\text{dB}$ , the same delay QoS exponent  $\theta = 0.01$ , and the same EE requirement factor  $\chi_{EE} = 0.7$ . When the number of users  $K$  increases, the total EC values calculated from the two algorithms decrease and then stabilize when  $K \geq 6$ . This happens because, when  $K$  increases from 2 to 6, the number of served users increases and correspondingly, the number of subcarriers allocated to each served user decreases. Henceforth, the achievable EC for each served user reduces and the total EC value calculated from (7) decreases. When  $K \geq 6$ , the number of served users remains the same, due to the limited number of available subcarriers. Hence, the total EC value stays stable when  $K$  becomes greater than 6.

To indicate that the proposed heuristic algorithm performs better than the other state-of-the-art algorithms, Fig. 10 plots the total EC versus the number of subcarriers for the proposed heuristic algorithm and the algorithm proposed in [25]. For the algorithm proposed in [25], an initial subcarrier allocation proportionate to each user's EC requirement was set at the beginning, and then the user with the highest slope of the EC versus signal-to-noise ratio (SNR) curve will keep releasing one subcarrier to the user with the lowest slope, until the total power consumption cannot be reduced any more. It is noted that [25] does not consider the tradeoff between EC and the link-layer EE, and only the total power consumption

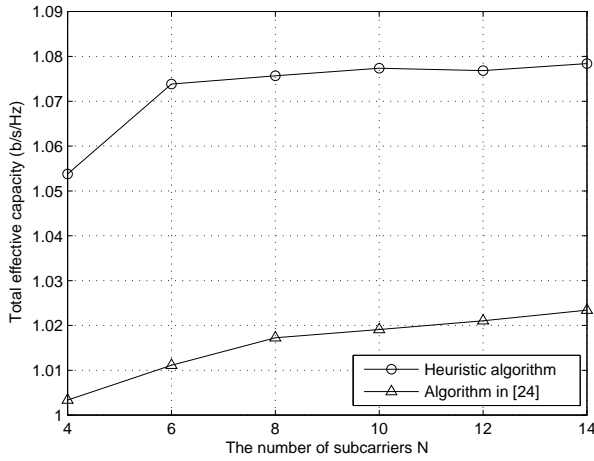


Fig. 10: The total effective capacity versus the number of subcarriers  $N$ , for heuristic algorithm and algorithm in [25].

needs to be minimized. Since in our paper, we have link-layer EE requirement for each user, for the initial subcarrier allocation vector, we will make the elements proportionate to the corresponding users' EE requirements<sup>5</sup>. The number of users is  $K = 4$ , and all the users are assumed to have the same delay QoS exponent, i.e.,  $\theta = 0.001$ , and the same EE requirement factor, i.e.,  $\chi_{EE} = 0.7$ . When the number of subcarriers increases, the total EC values calculated from the two algorithms first increase and then gradually stabilize. This indicates that when all users have loose delay QoS requirements, increasing the number of available subcarriers will not greatly improve the total system EC value. Further, from Fig. 10, it shows that our proposed heuristic algorithm performs better than the algorithm in [25]. Apart from this, the allocation algorithm in [25] also has no guarantee of each user's link-layer EE requirement, which can be confirmed from Fig. 14. Hence, we can conclude that, to solve the formulated total EC-maximization problem, subject to all users' EE requirement constraints, our proposed heuristic algorithm is more suitable, because it outperforms the algorithm in [25], and is also simple, fair and close-to-optimal.

Assume that all  $K$  users, having the same  $\theta$  value at  $\theta = 0.01$ , and the same circuit-to-noise power ratio value at  $P_{cr} = -10\text{dB}$ , are split into two groups. In group 1, all  $K_1$  users are required to have the same  $\chi_{EE}$ , i.e.,  $\chi_{EE} = 0.1$ . Meanwhile, the EE requirement factor value  $\chi_{EE}$  for all  $K - K_1$  users in group 2 is 0.8. This indicates that the users in group 1 have looser EE requirements compared to the users in group 2. Set the total number of users  $K = 6$ , and the total number of subcarriers  $N = 12$ . Fig. 11 includes the plots for the results of the system total EC versus the number of users  $K_1$  in group 1, for various values of circuit-to-noise power ratio  $P_{cr}$ . With fixed  $P_{cr}$ , when  $K_1$  increases from 0 to 6, the total system EC value, in b/s/Hz, gradually increases. This is because when  $K_1$  increases, the number of users with  $\chi_{EE} = 0.1$  increases and correspondingly, the number of users with  $\chi_{EE} = 0.8$  reduces. We note that for

<sup>5</sup>Hence, we had to tweak the algorithm slightly to be able to provide these results.

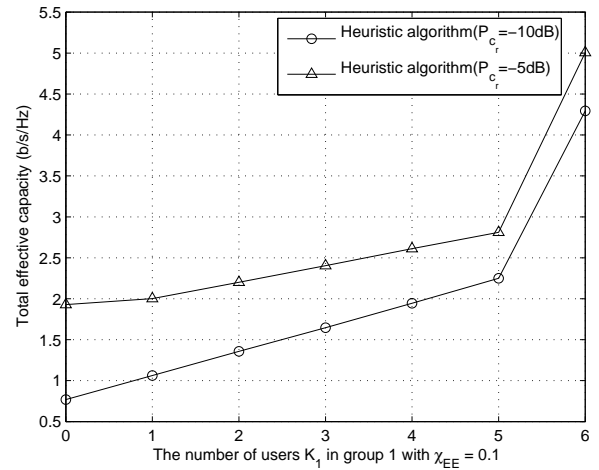


Fig. 11: The total effective capacity versus the number of users  $K_1$  in group 1, for various values of  $P_{cr}$ .

each user, a large  $\chi_{EE}$  value means that the user has a strict requirement on its link-layer EE value and will end up with a relatively small EC value. Therefore, when the number of users with  $\chi_{EE} = 0.8$  reduces, the system can save more resource to benefit the total system EC value, rather than sacrifice the system performance to support the strict EE requirements. When  $K_1$  increases from 5 to 6, the number of users with  $\chi_{EE} = 0.8$  reduces from 1 to 0 and the total EC value grows dramatically. This is due to the fact that in the check process of heuristic algorithm, the user having current minimum EC value will get the priority, which, in this case, corresponds to the one with  $\chi_{EE} = 0.8$ . Therefore, when  $K_1 = 5$ , the heuristic algorithm spends many resources on the user with  $\chi_{EE} = 0.8$ . When  $K_1 = 6$ , all users have the same loose EE requirements, i.e.,  $\chi_{EE} = 0.1$ , therefore, the system resources can be arranged evenly, which results in a great growth in the total EC value. Furthermore, from Fig. 11, we note that when  $P_{cr}$  becomes larger, the system total EC value increases. Since a bigger value of  $P_{cr}$  for all users will not change their relative difference, and correspondingly, will not change the subcarrier assignment solution, this phenomenon indicates that given a fixed subcarrier assignment, when one user's circuit power value increases, the system total EC value will increase, as well as its own EC value.

To analyze the impact of the circuit-to-noise power ratio  $P_{cr}^k$  and the EE requirement factor  $\chi_{EE}^k$  on the  $k^{\text{th}}$  user's multi-carrier system, Fig. 12 plots the results of the link-layer EE (on the left hand side (LHS) y-Axis, in solid lines) and the optimal tradeoff average power (on the right hand side (RHS) y-Axis, in dash lines) versus  $P_{cr}^k$ , for two different values of  $\chi_{EE}^k$ , considering  $N_k = 4$  and  $\theta_k = 0.01$ . When  $\chi_{EE}^k$  is fixed, the link-layer EE value decreases and the optimal average power increases with  $P_{cr}^k$ , which confirms the proved Lemma 1. Furthermore, with a fixed value of  $P_{cr}^k$ , when  $\chi_{EE}^k$  becomes larger, the  $k^{\text{th}}$  user's link-layer EE value increases, but the optimal average power decreases, which confirms the proposed Lemma 2 in Section III-B.

The plots of EE (on the RHS y-Axis, in dash lines) and EC (on the LHS y-Axis, in solid lines) versus  $\chi_{EE}^k$ , for two

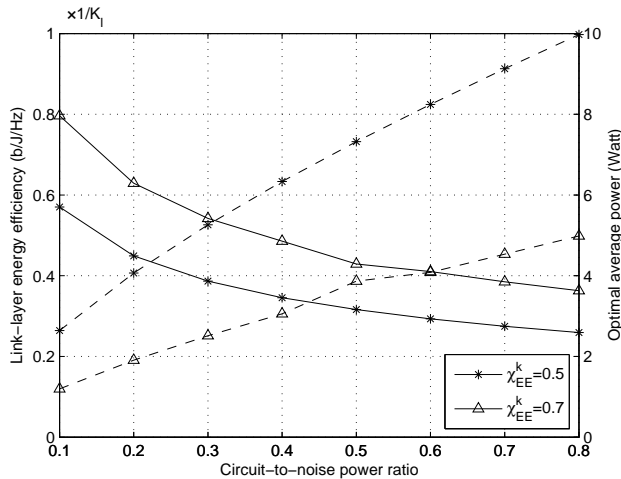


Fig. 12: Link-layer energy efficiency and the optimal average power versus circuit-to-noise power ratio  $P_{cr}^k$ , for various values of  $\chi_{EE}^k$ .

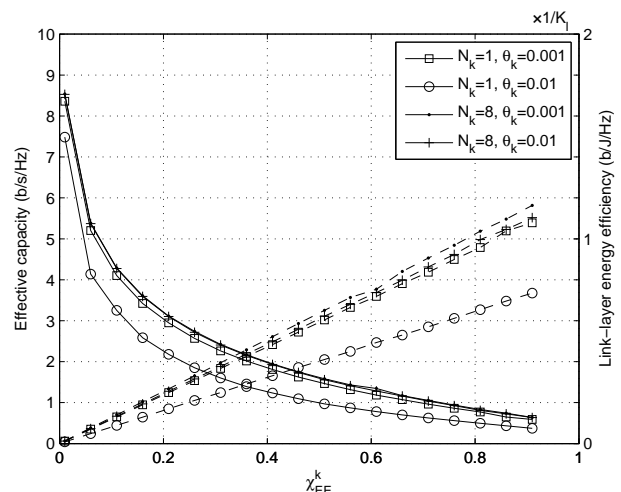


Fig. 13: Effective capacity and link-layer energy efficiency versus  $\chi_{EE}^k$ , for various values of  $\theta_k$  and  $N_k$ .

different values of delay QoS exponent  $\theta_k$  and various values of  $N_k$ , are included in Fig. 13. From this figure, we note that with fixed number of subcarriers  $N_k$ , when  $\chi_{EE}^k$  increases, EE increases. This confirms the proposed Lemma 2 in Section III-B. Furthermore, with a fixed  $N_k$ , EC decreases with  $\chi_{EE}^k$ . This is due to the fact that the tradeoff system operates in the conflicting region of EE and EC, therefore the EE-increases result from EC-reductions. Moreover, with fixed  $\chi_{EE}^k$  and  $\theta_k$ , as the number of subcarriers increases, both EC and EE increase. With fixed  $N_k$ , when the delay QoS exponent  $\theta_k$  increases from 0.001 to 0.01, both EE and EC decrease. Especially, when  $N_k = 1$ , the decreases of EE and EC, as a result of the increase in  $\theta_k$ , are significant. However, when  $N_k$  is larger, e.g.,  $N_k = 8$ , the decreases of EE and EC are minor. This indicates that the multi-carrier communication system is more robust against delay requirements, in comparison with single-carrier communication systems. In other words, when the system QoS requirement becomes more stringent, the multi-carrier system would sacrifice less EE and EC to guarantee the required  $\theta_k$ .

To further analyze and investigate the effect of  $\chi_{EE}$  on the

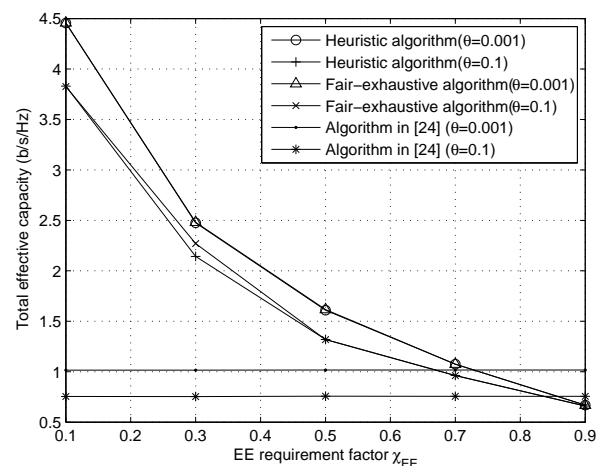


Fig. 14: The total effective capacity versus EE requirement factor  $\chi_{EE}$  for different values of  $\theta$  in heuristic algorithm, fair-exhaustive algorithm and algorithm in [25].

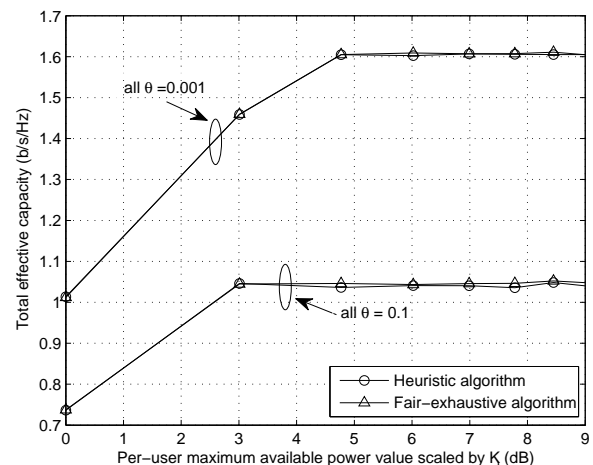


Fig. 15: The total effective capacity versus maximum power value, for different values of  $\theta$ .

multi-user multi-carrier system, Fig. 14 includes the plots for the total EC value versus EE requirement factor  $\chi_{EE}$ , with two different values of  $\theta$ , for the heuristic algorithm, the fair-exhaustive algorithm and the algorithm in [25]. Assume that the total number of users  $K = 4$  and the total number of available subcarriers  $N = 8$ . Specifically, all  $K$  users have the same settings of delay QoS exponent  $\theta$ ,  $\chi_{EE}$ , with  $P_{cr} = -10$  dB. When  $\chi_{EE}$  increases, the total EC value of the multi-user multi-carrier system decreases in the heuristic algorithm and the fair-exhaustive algorithm. Furthermore, when  $\theta = 0.001$ , the EC curves of the two algorithms are exactly the same. This indicates that for a system with  $K$  users having loose delay requirements, the difference between the total EC values calculated from the two algorithms is very little, even under different subcarrier assignment solutions. When  $\theta$  becomes larger, the total EC values become smaller. Since a larger  $\theta$  represents a more stringent delay QoS requirement, therefore each user's maximum achievable arrival rate that it can support to maintain the target delay requirement, namely, its EC value, becomes small. Henceforth, the total system EC value reduces, correspondingly. For the algorithm in [25],

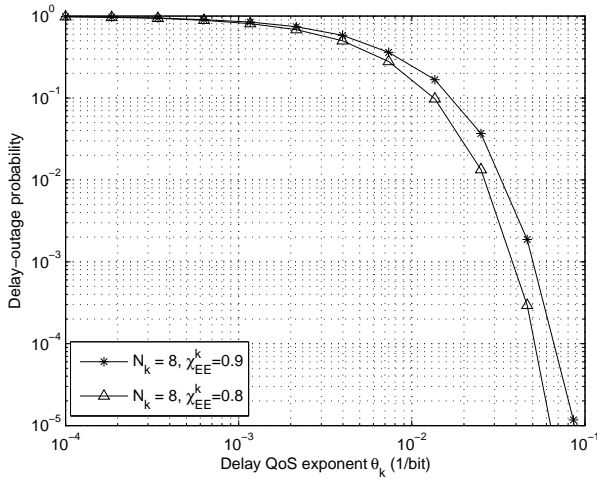


Fig. 16: Delay-outage probability versus delay QoS exponent  $\theta_k$ , for different values of  $\chi_{EE}$ .

when  $\chi_{EE}$  increases, the total EC value stays stable. This is because, this algorithm proposed in [25] does not guarantee all users' EE requirements. Therefore, in this case,  $\chi_{EE}$  has no influence on the total EC value.

Considering that the value of the maximum available power for each user can influence the optimal results, we include Fig. 15 to show the performance of the heuristic algorithm and the fair-exhaustive algorithm, when the maximum available power value varies. The total number of subcarriers is fixed at  $N = 6$ , and the number of users is  $K = 4$ . All users are assumed to have the same settings for  $\chi_{EE}$ , i.e.,  $\chi_{EE} = 0.5$ , and the same circuit-to-noise power ratio, i.e.,  $P_{cr} = -10$  dB. In addition, two different scenarios of delay QoS exponent vector  $\theta$  are included in Fig. 15, i.e., all elements in  $\theta$  are either 0.001 or 0.1. Firstly, we note that in both scenarios, the calculated final EC values from the two algorithms are close, only with very little difference which makes the two curves difficult to distinguish. This confirms that the proposed heuristic algorithm indeed guarantees a close-to-optimal performance. Furthermore, when all users have more stringent delay QoS requirements, the total EC value reduces, which means that the value of EC needs to be sacrificed in this situation. More importantly, from Fig. 15, we can notice that for a fixed  $\theta$ , the curves first increase, and then stabilize. This is because when the maximum available power value is too small to support the proposed optimal power value, the system has to operate at  $P_{max}$ . Therefore, the final calculated EC value is smaller in this case, since it is obtained at  $P_{max}$ , rather than at the optimal power value. On the other hand, when the value of  $P_{max}$  becomes larger than the proposed optimal power value, then the system will operate at the optimal average power, which gives the maximum EC value, under each user's EE constraint. To find detailed analysis, please refer to Section III-B.

Fig. 16 plots the delay-outage probability for the  $k^{\text{th}}$  user,  $P_{\text{delay}}^{\text{out}}$ , versus delay QoS exponent  $\theta_k$ , for various values of  $\chi_{EE}^k$  with a maximum tolerable delay threshold  $D_{\text{max}} = 200$  and the circuit-to-noise power ratio  $P_{cr}^k = -10$  dB. This figure reveals that for loose delay-constrained situations, e.g.,  $\theta_k = 10^{-4}$ ,

the achievable  $P_{\text{delay}}^{\text{out}}$  values stay the same with different  $\chi_{EE}^k$  values. When the delay requirement is more stringent, e.g.,  $\theta_k = 10^{-2}$ , smaller  $\chi_{EE}^k$  ends up with less delay-outage probability. This happens because smaller  $\chi_{EE}^k$  value means more sacrifices of EE from its maximum value, and in turn, results in more increases in its EC value. Therefore, the probability that the buffer length exceeds  $D_{\text{max}}$  decreases, henceforth, the delay-outage probability reduces.

## V. CONCLUSIONS

A close-to-optimal subcarrier assignment solution jointed with an optimal power allocation strategy for end users to maximize the system total EC value, subject to all users' average transmission power limits and link-layer EE constraints in a multi-user multi-carrier uplink network, were proposed and developed in this paper. The traditional exhaustive algorithm was introduced, followed by a fair-exhaustive algorithm, which offers the optimal frequency provisioning solution in a fair manner. To reduce the computational complexity, we proposed a fair heuristic algorithm, which offers close-to-optimal solutions, and has a low complexity linearly relating to the size of the problem. Given the subcarrier allocation matrix, the power-constrained EC maximization problem in the single-user multi-carrier system, under each user's individual link-layer requirement level, was solved. To thoroughly analyze the tradeoff problem, the effects of the circuit power and the EE requirement factor were proved and investigated. Simulation results confirmed our design intentions, and further revealed that when there is a link-layer EE constraint, each user's tradeoff EC value may not monotonically decrease with its delay QoS provisioning, and the tradeoff EC value obtained with less subcarriers may be higher than the one achieved with more subcarriers.

## APPENDIX A: PROOF OF LEMMA 1

Assume that the calculated tradeoff average power value for the  $k^{\text{th}}$  user, obtained with a circuit power value  $P_{c,1}^k$ , can be denoted as  $\overline{P}_{k,1}^*$ . Meanwhile, when the allocated number of subcarriers is  $N$ , the maximum achievable EE value, i.e.,  $\eta_{\text{max},1}^{k,N} = EE^k \Big|_{\substack{N_k=N \\ P_c^k=P_{c,1}^k \\ \overline{P}_k=P_{EE,1}^*}}$ , achieves at  $\overline{P}_{EE,1}^*$ . If the  $k^{\text{th}}$  user has a higher circuit power, i.e.,  $P_{c,2}^k = P_{c,1}^k + \Delta P_c^k$ ,  $\Delta P_c^k > 0$ , the corresponding average power values which satisfy the EE requirement equality (20) and  $\eta_{\text{max},2}^{k,N}$  can be written as  $\overline{P}_{k,2}^*$  and  $\overline{P}_{EE,2}^*$ , respectively. From (20), we have the following equations:

$$EE^k \Big|_{\substack{P_c^k=P_{c,1}^k \\ \overline{P}_k=P_{EE,1}^*}} = \chi_{EE}^k \times EE^k \Big|_{\substack{N_k=N \\ P_c^k=P_{c,1}^k \\ \overline{P}_k=P_{EE,1}^*}}, \quad (22a)$$

$$EE^k \Big|_{\substack{P_c^k=P_{c,2}^k \\ \overline{P}_k=P_{EE,2}^*}} = \chi_{EE}^k \times EE^k \Big|_{\substack{N_k=N \\ P_c^k=P_{c,2}^k \\ \overline{P}_k=P_{EE,2}^*}}. \quad (22b)$$

In order to investigate the influence of circuit power  $P_c^k$  on the tradeoff EE value, we start from analyzing the effect of  $P_c^k$  on the maximum EE value  $\eta_{\text{max}}^{k,N}$ . For the system with  $P_{c,1}^k$ , if

we assume the operational average input power is  $\overline{P_{EE,2}^{k*}}$ , the corresponding link-layer EE value can be calculated as

$$EE^k \Big|_{\substack{N_k=N \\ P_c^k=P_{c,1}^k \\ \overline{P_k}=\overline{P_{EE,2}^{k*}}}} = \frac{E_c^k \Big|_{N_k=N \\ \overline{P_k}=\overline{P_{EE,2}^{k*}}}}{P_{c,1}^k + \frac{1}{\epsilon} \overline{P_{EE,2}^{k*}}}. \quad (23)$$

Meanwhile, for the system with  $P_{c,2}^k$ , the  $\eta_{\max,2}^{k,N}$  value is defined as

$$EE^k \Big|_{\substack{N_k=N \\ P_c^k=P_{c,2}^k \\ \overline{P_k}=\overline{P_{EE,2}^{k*}}}} = \frac{E_c^k \Big|_{N_k=N \\ \overline{P_k}=\overline{P_{EE,2}^{k*}}}}{P_{c,2}^k + \frac{1}{\epsilon} \overline{P_{EE,2}^{k*}}}. \quad (24)$$

Apparently, we can notice that the link-layer EE value in (23) is larger than the one in (24), because  $P_{c,1}^k < P_{c,2}^k$ . Henceforth, we can derive that the maximum achievable link-layer EE value  $\eta_{\max,1}^{k,N}$  for the system with circuit power  $P_{c,1}^k$  is larger than the one obtaining at a larger circuit power  $P_{c,2}^k$ . This means that when one user's circuit power becomes larger, its maximum achievable EE value reduces. From (22a)-(22b), we finally conclude that one user's tradeoff EE level also decreases with its circuit power.

## REFERENCES

- [1] W. Yu, L. Musavian, and Q. Ni, "Multi-carrier link-layer energy efficiency and effective capacity tradeoff," in *IEEE Int. Conf. Commun. (ICC) Workshop*, London, UK, Jun. 2015, pp. 2763 – 2768.
- [2] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.
- [3] M. R. Mili, L. Musavian, K. A. Hamdi, and F. Marvasti, "How to increase energy efficiency in cognitive radio networks," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1829 – 1843, May. 2016.
- [4] Q. Wu, W. Chen, M. Tao, J. Li, H. Tang, and J. Wu, "Resource allocation for joint transmitter and receiver energy efficiency maximization in downlink OFDMA systems," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 416 – 430, Feb. 2015.
- [5] C. Bae and W. E. Stark, "End-to-end energy-bandwidth tradeoff in multihop wireless networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4051–4066, Sept. 2009.
- [6] Y. Li *et al.*, "Energy efficiency and spectral efficiency tradeoff in interference-limited wireless networks," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1924–1927, Oct. 2013.
- [7] X. Chen and S. Ouyang, "Energy- and spectral-efficiency trade-off in OFDMA-based cooperative cognitive radio networks," *Int. J. Distrib. Sens. Netw.*, vol. 2014, Feb. 2014.
- [8] GSMA Intelligence. (2014, Dec.) Understanding 5G: Perspectives on future technological advancements in mobile. [Online]. Available: <https://gsmaintelligence.com/research/2014/12/understanding-5g/451/>
- [9] F. Fu and M. van der Schaar, "Decomposition principles and online learning in cross-layer optimization for delay-sensitive applications," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1401 – 1415, Mar. 2010.
- [10] W. Chen, M. J. Neely, and U. Mitra, "Energy-efficient transmissions with individual packet delay constraints," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2090 – 2109, May 2008.
- [11] X. Zhang and J. Tang, "Power-delay tradeoff over wireless networks," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3673 – 3684, Sep. 2013.
- [12] M. Ozmen and M. C. Gursoy, "Wireless throughput and energy efficiency with random arrivals and statistical queuing constraints," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1375 – 1395, Mar. 2016.
- [13] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality-of-service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, Jul. 2003.
- [14] M. Sinaie, A. Zappone, E. A. Jorswieck, and P. Azmi, "A novel power consumption model for effective energy efficiency in wireless networks," *IEEE Wireless Commun. Lett.*, vol. 5, no. 2, pp. 2162 – 2337, Apr. 2016.
- [15] L. Musavian and Q. Ni, "Effective capacity maximization with statistical delay and effective energy efficiency requirements," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3824–3835, Jul. 2015.
- [16] W. Yu, L. Musavian, and Q. Ni, "Tradeoff analysis and joint optimization of link-layer energy efficiency and effective capacity toward green communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3339–3353, Jan. 2016.
- [17] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation for multichannel communications over wireless links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 1536–1276, Dec. 2007.
- [18] A. Helmy, L. Musavian, and T. Le-Ngoc, "Energy-efficient power adaptation over a frequency-selective fading channel with delay and power constraints," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4529–4541, Sep. 2013.
- [19] W. Cheng, X. Zhang, and H. Zhang, "Joint spectrum and power efficiencies optimization for statistical QoS provisionings over SISO/MIMO wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 903–915, May 2013.
- [20] C. Xiong, G. Y. Li, Y. Liu, Y. Chen, and S. Xu, "Energy-efficient design for downlink OFDMA with delay-sensitive traffic," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 3085 – 3094, Jun. 2013.
- [21] T. Abrao, L. D. H. Sampaio, S. Yang, K. T. K. Cheung, P. J. E. Jeszensky, and L. Hanzo, "Energy efficient OFDMA networks maintaining statistical QoS guarantees for delay-sensitive traffic," *IEEE Access*, vol. 4, pp. 774 – 791, Feb. 2016.
- [22] C. S. Chang, "Stability, queue length, and delay of deterministic and stochastic queueing networks," *IEEE Trans. Autom. Control*, vol. 39, no. 5, pp. 913–931, May 1994.
- [23] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer-Verlag New York Inc., 2nd Edition, 1998.
- [24] J. A. Bucklew, *Introduction to Rare Event Simulation*. Springer-Verlag New York Inc., 2004.
- [25] A. Helmy and T. Le-Ngoc, "Low-complexity QoS-aware frequency provisioning in downlink multi-user multicarrier systems," in *IEEE Wireless Commun. Netw. Conf. (WCNC)*, Istanbul, Turkey, Apr. 2014, pp. 1785 – 1790.
- [26] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171 – 178, Feb. 2003.
- [27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.