On the Spectral-Energy Efficiency and Rate Fairness Tradeoff in Relay-Aided Cooperative OFDMA Systems

Zhengyu Song, Qiang Ni, Keivan Navaie, Shujuan Hou, Siliang Wu, and Xin Sun

Abstract—In resource constraint wireless systems, achieving higher spectral efficiency (SE) and energy efficiency (EE), and greater rate fairness are conflicting objectives. Here a general framework is presented to analyze the tradeoff among these three performance metrics in cooperative OFDMA systems with decode-and-forward (DF) relaying, where subcarrier pairing and allocation, relay selection, choice of transmission strategy, and power allocation are jointly considered. In our analytical framework, rate fairness is represented utilizing α-fairness model and the resource allocation problem is formulated as a multi-objective optimization (MOO) problem. We then propose a cross-layer resource allocation algorithm across application and physical layers, and further devise a heuristic algorithm to tackle the computational complexity issue. The SE-EE tradeoff is characterized as a Pareto optimal set, and the efficiency and fairness tradeoff is investigated through the price of fairness (PoF). Simulations indicate that higher fairness results in a worse SE-EE tradeoff. It is also shown imposing fairness helps to reduce the outage probability. For a fixed number of relays, by increasing circuit power, the performance of SE-EE tradeoff is degraded. Interestingly, by increasing the number of relays, although the total circuit power is increased, the SE-EE tradeoff is not necessarily degraded. This is thanks to the extra degree of freedom provided in relay selection.

Index Terms—Cross-layer optimization, cooperative communications, energy efficiency, OFDMA, price of fairness, rate fairness, resource allocation, spectral efficiency.

I. INTRODUCTION

RELAY-aided cooperative communication is a promising technique for improving performance in cellular networks, including coverage area, transmission reliability, and system throughput [1]. Meanwhile, orthogonal frequency division multiple access (OFDMA) is the major access scheme in current cellular networks. Due to the inherent spectrum scarcity in wireless communications, great efforts have been made to improve the spectral efficiency (SE) in OFDMA cooperative systems [2]–[4]. Recently, because of the huge amount of energy consumption in wireless communication systems [5], increasing the energy efficiency (EE) has become an essential issue in the current and future fifth-generation (5G) cellular communication networks [6].

Z. Song and X. Sun are with the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China (email: z.song912@gmail.com, xsun@bjtu.edu.cn).

Q. Ni and K. Navaie are with the School of Computing and Communications, Lancaster University, LA1 4WA, UK (email: {q.ni, k.navaie}@lancaster.ac.uk).

S. Hou and S. Wu are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (email: {shujuanhou, siliangw}@bit.edu.cn).

Resource allocation schemes for maximizing SE in OFDMA cooperative systems have been extensively studied in the literature. See, e.g., [2]–[4], [7]. Authors in [2] investigate relay selection, subcarrier pairing and power allocation problems in multi-relay OFDMA systems with one single user to maximize SE, where amplify-and-forward (AF) protocol is used. Further in [3], a joint subcarrier and power allocation scheme is proposed to maximize the system throughput in LTE-Advanced cooperative networks, where subcarrier pairing and relay selection are fixed and thus excluded from the proposed resource allocation scheme. In [4], authors investigate QoS-aware relay selection and subcarrier assignment via branch-and-cut and dual method in multi-user OFDMA relay networks to maximize sum-rate.

Since EE will be a key issue in the future 5G cellular networks, energy-aware system design has become an immediate need in both industry and academia. Expanding bandwidth increases EE, however it also degrades the system SE [8]. Considering the spectrum scarcity in wireless communications, it is imperative to balance SE and EE as conflicting objectives. Existing studies on the relationship of SE and EE can be divided into two categories, and in both categories the rate fairness is excluded.

The first category is based on the approach in which EE is maximized [9]–[18], and the second category is focused on exploring the optimal achievable envelope of the SE-EE tradeoff [19]–[22]. In [9]–[11], [17], [18], energy-efficient designs are proposed where it is demonstrated that the power consumption can be reduced by performing EE optimization. Considering users' quality-of-service (QoS) requirements, [12] and [13] deal with the energy-efficient resource allocation problem in a multiuser OFDMA system. Considering a pre-assigned relay to each user, [14] formulates an EE maximization problem in AF relay cellular networks, where subcarrier pairing is excluded. In [15], EE is maximized while satisfying a SE requirement in a three-terminal relay network. Afterwards, the transmitter and receiver power consumption are jointly considered in [16] to maximize the EE thus increasing the battery life, but the SE requirement is not incorporated. All these schemes mainly focus on EE maximization. For those with given SE requirements, they are inflexible and often restrict the SE performance. For those without SE requirements in place, maximizing EE often leads to compromising SE performance.

The SE-EE tradeoff [19] has been investigated in the second category. For OFDMA systems in [20], a new metric, namely resource efficiency is introduced to enforce a balance between SE and EE, where bandwidth and transmission power are jointly optimized. The authors in [21], propose a new formulation to characterize the Pareto optimal set of SE-EE

tradeoff, but they do not provide specific resource allocation schemes. In [22], a relay cooperation scheme is proposed for MIMO cellular networks, where various relay decoding strategies are considered, and the SE and EE are then evaluated for this relay cooperation scheme. These works investigate the global relationship of SE and EE, and provide some flexible techniques to manage the SE-EE tradeoff.

Given the time varying nature of wireless channels, rate fairness is a critical performance indicator in cellular networks. Fairness-aware energy efficient radio resource allocation is considered in the literature in conventional OFDMA systems without relays. In [23], EE definition is generalized as the weighted number of delivered bits per unit energy to provide fairness to some extent. More recently, EE maximization resource allocation algorithms with proportional rate constraint are proposed in [24]–[26]. However, in these fairness-aware researches, the focus is on EE maximization without considering the negative impact on the SE performance.

To our best knowledge, the three-factor tradeoff among SE, EE and rate fairness has not been studied in the literature. In this paper, we present a general framework to analyze the SE, EE and rate fairness tradeoff in multi-user OFDMA cooperative systems. We devise a joint cross-layer radio resource allocation algorithm based on Lagrangian Dual Decomposition (LDD). Further, a heuristic resource allocation algorithm is developed to reduce the computational complexity.

The main contributions of this paper are:

1) A novel general framework based on multi-objective optimization is proposed to investigate the three-factor tradeoff among SE, EE and rate fairness, where rate fairness is represented utilizing α-fairness model. The SE-EE tradeoff is characterized as a Pareto optimal set and we introduce price of fairness (PoF) to quantitatively evaluate the efficiency and fairness tradeoff.

2) Unlike previous literature such as [3], [4] and [14], we exploit all degrees of freedom in resource allocation to jointly manage the three-factor tradeoff for achieving satisfactory performance. In our analytical framework, we first introduce the *virtual relay* concept which maps the problem of choosing the transmission strategy into a relay selection problem, and then propose a LDD-based cross-layer algorithm to jointly obtain the optimal decisions on relay selection, subcarrier pairing and allocation, and power allocation across application and physical layer optimization.

3) We propose a novel heuristic resource allocation algorithm to reduce the computational complexity of the LDD-based cross-layer algorithm. The heuristic algorithm first obtains subcarrier pairing and allocation, and relay selection, and then carries out the optimal power allocation. Simulation results show that the two algorithms closely follow the same pattern and achieve similar performance, while the computational complexity of the proposed heuristic algorithm is significantly lower than the LDD-based cross-layer algorithm.

Simulations indicate that higher fairness results in a worse SE-EE tradeoff. It is also shown imposing fairness helps to reduce the outage probability. Besides, we observe that for a fixed number of relays, by increasing circuit power, the performance of SE-EE tradeoff is degraded, i.e., for a given
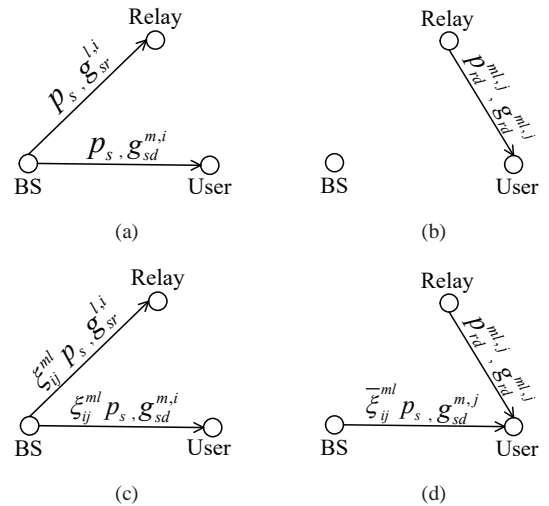

Fig. 1. (a) 1st slot of asynchronous mode. (b) 2nd slot of asynchronous mode. (c) 1st slot of synchronous mode. (d) 2nd slot of synchronous mode.

fairness level and SE (EE), a lower EE (SE) is achieved. Interestingly, by increasing the number of relays, although the total circuit power is increased, the SE-EE tradeoff is not necessarily degraded. This is because of the extra degree of freedom provided in relay selection.

The rest of this paper is organized as follows. In Section II, we present the system model and problem formulation. In Section III, the optimization problem is solved using LDD method and a cross-layer resource allocation algorithm is proposed. In Section IV, a novel heuristic algorithm is proposed to reduce the complexity in resource allocation. Simulation results are presented in Section V followed by the conclusions in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider the downlink in a multi-user relay-aided OFDMA cooperative cellular network with one base station (BS), M users, and L relays. The total bandwidth is divided into N orthogonal subcarriers, and the intra-cell interference is negligible. All the relays are assumed to be half-duplex, which means they cannot transmit and receive signals simultaneously. Perfect channel state information (CSI) is available at the receivers, and this information is fed back to the BS. Since relays in a cellular network are usually deployed to assist the transmission of users located at the cell-edge, in this paper, our focus is on the users located at the cell-edge and experiencing deteriorated wireless link from the BS due to the channel fading.

We consider three transmission modes, i.e., asynchronous mode, synchronous mode and direct mode. Each transmission frame is divided into two consequent slots, and the channel gains are assumed to be constant during the two slots. The normalized channel gains over noise on subcarrier $i \in \{1, 2, ..., N\}$ from BS to relay $l \in \{1, 2, ..., L\}$ and user $m \in \{1, 2, ..., M\}$ are denoted as $g_{sr}^{l,i}$ and $g_{sd}^{m,i}$, respectively. Similarly, for subcarrier $j \in \{1, 2, ..., N\}$, the normalized channel gains from relay l and BS to user m are denoted by $g_{rd}^{ml,j}$ and $g_{sd}^{m,j}$, respectively.

For the asynchronous mode, as shown in Fig. 1 (a) and (b), in the first slot, BS broadcasts the signals over subcarrier i with power p_s to all the relays and users. In the second slot, relays transmit their received signals over subcarrier j with power $p_{rd}^{ml,j}$ to users using decode-and-forward (DF) protocol. While for the synchronous mode, as shown in Fig. 1 (c) and (d), BS broadcasts over subcarrier i with power $\xi_{ij}^{ml} p_s$ in the first slot, where $\xi_{ij}^{ml} \in (0,1)$. In the second slot, in addition that relays forwards signals over subcarrier j with power $p_{rd}^{ml,j}$, BS also transmits over subcarrier j with power $\overline{\xi}_{ij}^{ml} p_s$ to cooperate with relays, where $\overline{\xi}_{ij}^{ml} = 1 - \xi_{ij}^{ml}$. Users combine the received signals within the two slots utilizing maximum ratio combining (MRC) followed by decoding.

Therefore, for a cooperative transmission link (ml, ij) with DF protocol, the normalized maximum achievable data rate over bandwidth is [3]

$$
\begin{aligned}
r_{\mathrm{DF}}^{ml,ij} = \max_{0 \le \xi_{ij}^{ml} \le 1} \ \min \frac{1}{2} \Big\{ & \log_2 \left(1 + g_{sr}^{l,i} \xi_{ij}^{ml} p_s \right), \\
& \log_2 \Big[1 + g_{sd}^{m,i} \xi_{ij}^{ml} p_s + \Big(\sqrt{g_{sd}^{m,j} \overline{\xi}_{ij}^{ml} p_s} + \sqrt{g_{rd}^{ml,j} p_{rd}^{ml,j}} \Big)^2 \Big] \Big\}
\end{aligned}
\tag{1}
$$

Here, we set $\xi_{ij}^{ml} \in [0,1]$ and $\overline{\xi}_{ij}^{ml} = 1 - \xi_{ij}^{ml}$. ξ_{ij}^{ml} is the portion of the BS transmit power used to cooperate with relays. The maximum data rate is achieved when the decoding rate at the relay is equal to the destination decoding rate. Therefore, one can jointly adjust p_s, $p_{rd}^{ml,j}$ and ξ_{ij}^{ml} such that the two rates become equal. If the relay decoding rate is lower than that of the destination, we set $\xi_{ij}^{ml} = 1$, which is known as the "asynchronous mode" [3]. Otherwise, if the destination decoding rate is lower, ξ_{ij}^{ml} is reduced until these two rates become equal. This case is referred to as the "synchronous mode" [3].

For the asynchronous mode, if setting $p_{s,\mathrm{asyn}} + p_{rd,\mathrm{asyn}}^{ml,j} = p_{\mathrm{asyn}}^{ml,ij}$, we then express the normalized maximum achievable data rate for the transmission link (ml, ij) as

$$
r_{\mathrm{asyn}}^{ml,ij} = \frac{1}{2} \log_2 \left(1 + g_{\mathrm{asyn}}^{ml,ij} p_{\mathrm{asyn}}^{ml,ij} \right),
\tag{2}
$$

where the power allocation and equivalent channel gain is given by

$$
p_{s,\mathrm{asyn}} = \frac{g_{rd}^{ml,j}}{g_{rd}^{ml,j} + g_{sr}^{l,i} - g_{sd}^{m,i}} p_{\mathrm{asyn}}^{ml,ij},
\tag{3}
$$

$$
p_{rd,\mathrm{asyn}}^{ml,j} = \frac{g_{sr}^{l,i} - g_{sd}^{m,i}}{g_{rd}^{ml,j} + g_{sr}^{l,i} - g_{sd}^{m,i}} p_{\mathrm{asyn}}^{ml,ij},
\tag{4}
$$

$$
g_{\mathrm{asyn}}^{ml,ij} = \frac{g_{rd}^{ml,j} g_{sr}^{l,i}}{g_{rd}^{ml,j} + g_{sr}^{l,i} - g_{sd}^{m,i}}.
\tag{5}
$$

For the synchronous mode, where $0 \le \xi_{ij}^{ml} < 1$, and $p_{s,\mathrm{syn}} + p_{rd,\mathrm{syn}}^{ml,j} = p_{\mathrm{syn}}^{ml,ij}$, the normalized maximum achievable data rate for the transmission link (ml, ij) is

$$
r_{\mathrm{syn}}^{ml,ij} = \frac{1}{2} \log_2 \left(1 + g_{\mathrm{syn}}^{ml,ij} p_{\mathrm{syn}}^{ml,ij} \right),
\tag{6}
$$

where the optimal power division and equivalent channel gain is given by

$$
\xi_{ij}^{ml} p_{s,\mathrm{syn}} = \frac{G_1^{ml,ij}}{G_2^{ml,ij}} p_{\mathrm{syn}}^{ml,ij},
\tag{7}
$$

$$
\overline{\xi}_{ij}^{ml} p_{s,\mathrm{syn}} = \frac{g_{sd}^{m,j} G_3^{ml,ij}}{G_1^{ml,ij} G_2^{ml,ij}} p_{\mathrm{syn}}^{ml,ij},
\tag{8}
$$

$$
p_{rd,\mathrm{syn}}^{ml,j} = \frac{g_{rd}^{ml,j} G_3^{ml,ij}}{G_1^{ml,ij} G_2^{ml,ij}} p_{\mathrm{syn}}^{ml,ij},
\tag{9}
$$

$$
g_{\mathrm{syn}}^{ml,ij} = \frac{g_{sr}^{l,i} G_1^{ml,ij}}{G_2^{ml,ij}}.
\tag{10}
$$

where $G_1^{ml,ij} = g_{sd}^{m,j} + g_{rd}^{ml,j}$, $G_2^{ml,ij} = g_{rd}^{ml,j} - g_{sd}^{m,i} + g_{sr}^{l,i} + g_{sd}^{m,j}$ and $G_3^{ml,ij} = g_{sr}^{l,i} - g_{sd}^{m,i}$.

For the direct mode, the BS transmits to user m on subcarrier i in the first slot and on subcarrier j in the second slot. Therefore, the normalized achievable data rate is

$$
r_{\mathrm{DT}}^{m,ij} = \frac{1}{2} \left[\log_2 \left(1 + g_{sd}^{m,i} p_{sd}^{m,i} \right) + \log_2 \left(1 + g_{sd}^{m,j} p_{sd}^{m,j} \right) \right],
\tag{11}
$$

where $p_{sd}^{m,i}$, and $p_{sd}^{m,j}$ are the BS transmit power in the first and second slots, respectively, and

$$
p_{\mathrm{DT}}^{m,ij} \triangleq p_{sd}^{m,i} + p_{sd}^{m,j}.
\tag{12}
$$

A. The Unified Framework

In order to analyze asynchronous, synchronous and direct modes of transmission in a unified framework, here in addition to L real relays, we introduce $L+1$ *virtual relays* denoted as $\{0\}$ and $\{L+1, L+2, ..., 2L\}$. Therefore, $l \in \{0, 1, 2, ..., 2L\}$. For a virtual relay $l \in \{L+1, L+2, ..., 2L\}$, the normalized channel gains are $g_{sr}^{l,i} = g_{sr}^{(l-L),i}$, and $g_{rd}^{ml,j} = g_{rd}^{m(l-L),j}$. In asynchronous mode, a real relay $l \in \{1, 2, ..., L\}$ is selected, whereas in synchronous mode, a virtual relay $l \in \{L+1, L+2, ..., 2L\}$ is chosen. For the direct mode, however, $l = 0$. By introducing *virtual relays*, the choice of transmission strategy can be incorporated into relay selection problem.

Therefore, combining (2), (6) and (11), the normalized maximum achievable data rate of the transmission link (ml, ij) is

$$
r_{ij}^{ml} = \begin{cases} \frac{1}{2} \log_2 \left(1 + g_{ij}^{ml} p_{ij}^{ml} \right), & l \in \{1, 2, ..., 2L\}, \\ \frac{1}{2} \Big[\log_2 \left(1 + g_{sd}^{m,i} p_{sd}^{m,i} \right) \\ \qquad + \log_2 \left(1 + g_{sd}^{m,j} p_{sd}^{m,j} \right) \Big], & l = 0, \end{cases}
\tag{13}
$$

where

$$
g_{ij}^{ml} = \begin{cases} \dfrac{g_{rd}^{ml,j} g_{sr}^{l,i}}{g_{rd}^{ml,j} + g_{sr}^{l,i} - g_{sd}^{m,i}}, & l \in \{1, 2, ..., L\}, \\ \dfrac{g_{sr}^{l,i} \left(g_{sd}^{m,j} + g_{rd}^{ml,j} \right)}{g_{rd}^{ml,j} - g_{sd}^{m,i} + g_{sr}^{l,i} + g_{sd}^{m,j}}, & l \in \{L+1, L+2, ..., 2L\}, \end{cases}
\tag{14}
$$

$$
p_{ij}^{ml} = \begin{cases} p_{s,\mathrm{asyn}} + p_{rd,\mathrm{asyn}}^{ml,j}, & l \in \{1, 2, ..., L\}, \\ p_{s,\mathrm{syn}} + p_{rd,\mathrm{syn}}^{ml,j}, & l \in \{L+1, L+2, ..., 2L\}, \\ p_{sd}^{m,i} + p_{sd}^{m,j}, & l = 0. \end{cases}
\tag{15}
$$

We further define $\rho_{ij} \in \{0, 1\}$ as the subcarrier pairing indicator which is equal to 1 if subcarriers i, and j are paired

in the two transmission slots, and 0, otherwise. We denote $\beta_{ij}^{ml} \in \{0,1\}$ as the subcarrier pair allocation and relay selection indicator, which is equal to 1, if user m is assisted by relay l using subcarrier pair (i, j), and 0, otherwise. Then the total normalized achievable data rate over bandwidth for user m can be expressed as

$$R_m = \sum_{l=0}^{2L} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \beta_{ij}^{ml} r_{ij}^{ml}. \tag{16}$$

B. Power Consumption Model

The total power consumption for transmission consists of the circuit power, and dynamic amplifier power of the BS and relay nodes. Circuit power consumption is assumed to be fixed. The total transmit power of the BS and all relay nodes is

$$P_{\mathrm{t}} = \frac{1}{2} \sum_{l=0}^{2L} \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \beta_{ij}^{ml} p_{ij}^{ml}. \tag{17}$$

The total power consumption is therefore the summation of the circuit power and the amplifiers' power:

$$P_{\mathrm{total}} = P_{\mathrm{C}} + \varepsilon P_{\mathrm{t}}, \tag{18}$$

where P_{C} is the total fixed circuit power of the BS and all relays, and $1/\varepsilon$ is the amplifiers' efficiency. Here, for brevity we assume the amplifiers' efficiency in the BS and relays is equal.

C. SE, EE and Rate Fairness

In this paper, SE is defined as the normalized system throughput over bandwidth, i.e., $SE = \sum_{m=1}^{M} R_m$, while EE is defined as the delivered bits per unit energy, i.e., $EE = \left(\sum_{m=1}^{M} R_m \right) \Big/ P_{\mathrm{total}}$.

Furthermore, due to the random nature of wireless channels, in cellular radio communications, the users with better channels may achieve much higher data rate compared with those with worse channels, leading to rate unfairness among users. To incorporate rate fairness in resource allocation, here we adopt α-fair utility function as defined in [27]:

$$u_\alpha(R_m) = \begin{cases} \ln(R_m), & \text{if } \alpha = 1, \\ R_m^{1-\alpha} / (1-\alpha), & \text{if } \alpha \neq 1, \alpha \geq 0. \end{cases} \tag{19}$$

As it is seen, α-fair utility function represents a family of utility functions, where the values of α indicate different levels of rate fairness. Maximizing the sum-utility of all users in the coverage area results in an α-fair resource allocation. Adjusting α, one can examine the tradeoff between the system efficiency and different levels of fairness. For instance, if no fairness is required, i.e., $\alpha = 0$, then $U_\alpha(R_m) = R_m$. Therefore, maximizing the sum-utility of all users is equivalent to maximizing the total network throughput. In this case, the highest total throughput is achieved while rate fairness among users is completely ignored. We also note that for $\alpha > 0$, since α-fair utility function is strictly increasing and concave, its marginal utility diminishes when the data rate increases. Therefore, it can balance efficiency and fairness. To be specific, by increasing α, the rate fairness among users

rises while the efficiency declines. In particular, $\alpha = 1$, and $\alpha \to \infty$, are corresponding to the proportional, and max-min fairness among users, respectively.

In reality, achieving higher SE, EE and rate fairness are three conflicting performance objectives. In the following, we will investigate the three-factor tradeoff in resource allocation.

D. Problem Formulation

As maximizing the sum-utility of all users results in an α-fair resource allocation and achieves efficiency and fairness tradeoff, studying the three-factor tradeoff among SE, EE and rate fairness is actually equivalent to maximizing the sum-utility and minimizing the total power consumption simultaneously. Therefore, we formulate it as a multi-objective optimization (MOO) problem as following:

$$\max_{\mathbf{p}, \boldsymbol{\rho}, \boldsymbol{\beta}} \quad \sum_{m=1}^{M} u_\alpha(R_m), \tag{20a}$$

$$\max_{\mathbf{p}, \boldsymbol{\rho}, \boldsymbol{\beta}} \quad -P_{\mathrm{total}}, \tag{20b}$$

$$\text{s.t.} \quad \mathrm{C1}: P_t \leq P_{\mathrm{T}}, \tag{20c}$$

$$\mathrm{C2}: p_{ij}^{ml} \geq 0, \ \forall m, l, i, j, \tag{20d}$$

$$\mathrm{C3}: \sum_{i=1}^{N} \rho_{ij} = 1, \ \forall j, \tag{20e}$$

$$\mathrm{C4}: \sum_{j=1}^{N} \rho_{ij} = 1, \ \forall i, \tag{20f}$$

$$\mathrm{C5}: \sum_{m=1}^{M} \sum_{l=0}^{2L} \beta_{ij}^{ml} = 1, \ \forall i, j, \tag{20g}$$

$$\mathrm{C6}: \rho_{ij}, \beta_{ij}^{ml} \in \{0,1\}, \ \forall m, l, i, j, \tag{20h}$$

where $\mathbf{p} = \{p_{ij}^{ml}\}$, $\boldsymbol{\rho} = \{\rho_{ij}\}$, $\boldsymbol{\beta} = \{\beta_{ij}^{ml}\}$, and P_{T} is the maximum summation transmit power of BS and all relay nodes. Constraint C6 ensures that the subcarrier and relay assignment indicators are binary variables. Along with C6, constraints C3 and C4 ensure that each subcarrier is only paired with one subcarrier in each frame. C5 further enforces exclusive assignment of subcarrier pair (i, j) to only one relay (inclusive of virtual relays) and user pair (m, l).

III. CROSS-LAYER OPTIMIZATION FOR SE, EE AND RATE FAIRNESS TRADEOFF

To understand the tradeoff among SE, EE, and rate fairness, we need to investigate the solution set of the optimization problem in (20). In this section, we propose a cross-layer algorithm based on Lagrangian dual decomposition (LDD) method, to obtain the solution set of (20).

A. Transformation to Single Objective Optimization

To obtain the optimal solution set of (20), we employ weighted sum method [28] to transfer the MOO problem into a single-objective optimization (SOO) problem. In order to

ensure a consistent comparison, we normalize the objective functions in (20) as the following:

$$\max_{\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} \left[\sum_{m=1}^{M} u_\alpha\left(R_m\right) - U_{\min}^\alpha\right] \bigg/ \left(U_{\max}^\alpha - U_{\min}^\alpha\right), \quad (21a)$$

$$\max_{\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} \; -P_{\text{total}}/P_{\max}, \quad (21b)$$

where $P_{\max}$ is the maximum total power consumption, and $U_{\max}^\alpha$ and $U_{\min}^\alpha$ are the maximum and minimum achievable sum-utility with fairness parameter α under the constraints C1-C6:

$$P_{\max} = P_{\text{C}} + \varepsilon P_{\text{T}}, \quad (22a)$$

$$U_{\max}^\alpha = \max_{\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} \sum_{m=1}^{M} u_\alpha\left(R_m\right)\big|_{P_{\text{t}}=P_{\text{T}}}, \quad (22b)$$

$$U_{\min}^\alpha = \begin{cases} \displaystyle\sum_{m=1}^{M} \ln\left(\delta\right), & \text{if } \alpha = 1, \\ \displaystyle\sum_{m=1}^{M} \delta^{1-\alpha}/(1-\alpha), & \text{if } \alpha \neq 1, \alpha \geq 0, \end{cases} \quad (22c)$$

where δ is a predefined and sufficiently small value. We exclude the case where the data rate of each user is zero, thus we assume $R_m \geq \delta$ for all users.

Applying the weighted sum method, the MOO problem can be then converted into an equivalent SOO problem as the following:

$$\max_{\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} \; w \frac{\displaystyle\sum_{m=1}^{M} u_\alpha\left(R_m\right) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - (1-w) \frac{P_{\text{total}}}{P_{\max}}, \quad (23a)$$

$$\text{s.t. } \text{C1} - \text{C6}, \quad (23b)$$

where $w \in [0, 1]$ is a weighting parameter, which can be used to reflect the importance level of the two objectives.

B. The Cross-Layer Optimal Solution Set

For a given α, finding the optimal solutions to the SOO problem in (23) for different values of w, forms a Pareto optimal solution set for the original MOO problem in (20) [28]. Pareto optimal solution set provides the best achievable values of the conflicting objective functions for any given value of fairness parameter α. In the following, we adopt cross-layer optimization based on Lagrangian dual decomposition (LDD) method [29] to find the optimal solutions to the SOO problem in (23). We consider two cases: $\alpha > 0$, and $\alpha = 0$.

1) Case I: $\alpha > 0$: To enable the cross-layer optimization, similar to [30], we introduce an auxiliary vector $\mathbf{t} = [t_1, t_2, ..., t_m]^T$ and rewrite the SOO problem (23) as

$$\max_{\mathbf{t},\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} \; w \frac{\displaystyle\sum_{m=1}^{M} u_\alpha\left(t_m\right) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - (1-w) \frac{P_{\text{total}}}{P_{\max}}, \quad (24a)$$

$$\text{s.t. } \text{C1} - \text{C6}, \quad (24b)$$

$$\text{C7}: \; t_m \leq R_m, \; \forall m. \quad (24c)$$

The auxiliary variable t_m is defined in the application layer, which represents the demand of data rate in the application

layer for user m, while R_m represents the supply in the physical layer. Hence, constraint C7 means the application-layer demand of data rate must be less than or equal to the physical-layer supply. In fact, since α-fair utility function is a strictly increasing function, at the optimal point, t_m must be equal to R_m. Therefore, (24) must have the same optimal solution as (23).

The key step in adopting LDD-based cross-layer optimization is to relax C7. To characterize the duality gap between the primal and dual solutions, time-sharing condition is defined in [31] where it is shown that holding this condition, the duality gap is zero even if the original optimization problem is not convex. In practical multicarrier systems with a large number of subcarriers, channel conditions in the adjacent subcarriers are often similar. In such case, the time-sharing condition is readily satisfied, and accordingly the duality gap is nearly zero [31]. Thus, by relaxing C7, Lagrangian function associated to (24) is

$$\begin{aligned} L\left(\mathbf{t},\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\boldsymbol{\lambda}\right) &= w \frac{\sum_{m=1}^{M} u_\alpha\left(t_m\right) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - (1-w) \frac{P_{\text{total}}}{P_{\max}} \\ &\quad + \sum_{m=1}^{M} \lambda_m \left(R_m - t_m\right) \\ &= \left[w \frac{\sum_{m=1}^{M} u_\alpha\left(t_m\right) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - \sum_{m=1}^{M} \lambda_m t_m\right] \\ &\quad + \left[\sum_{m=1}^{M} \lambda_m R_m - (1-w) \frac{P_{\text{total}}}{P_{\max}}\right], \end{aligned}$$
$$\quad (25)$$

where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_m]^T$ is the dual vector for constraint C7 corresponding to each user. Therefore, the dual function is

$$h\left(\boldsymbol{\lambda}\right) = \max_{\mathbf{t},\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} L\left(\mathbf{t},\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\boldsymbol{\lambda}\right), \quad (26a)$$

$$\text{s.t. } \text{C1} - \text{C6}. \quad (26b)$$

The corresponding dual problem is then

$$\min_{\boldsymbol{\lambda} \geq 0} h\left(\boldsymbol{\lambda}\right). \quad (27)$$

To obtain the optimal solution to (27), we use the sub-gradient method, where the dual variables are updated as

$$\lambda_m^{k+1} = \left[\lambda_m^k + \chi^k \left(t_m - R_m\right)\right]^+, \quad (28)$$

and χ^k is the diminishing step size at the kth iteration. The above sub-gradient update is guaranteed to converge to the optimal solution as long as the step size is chosen to be sufficiently small [31], [32].

As it is seen in (25) and (26), the dual function $h\left(\boldsymbol{\lambda}\right)$ considers both application-layer and physical-layer variables, i.e., $\mathbf{t}$, which is defined in the application layer, and $\mathbf{p}$, $\boldsymbol{\rho}$, and $\boldsymbol{\beta}$, which are physical-layer variables. The dual vector $\boldsymbol{\lambda}$ interrelates the two layers, reflecting the cross-layer interaction between application layer and physical layer. Hence, this is a cross-layer optimization problem. Using LDD method, the dual function $h\left(\boldsymbol{\lambda}\right)$ can be decomposed into two maximization subproblems, namely *application layer* and *physical layer* subproblems.

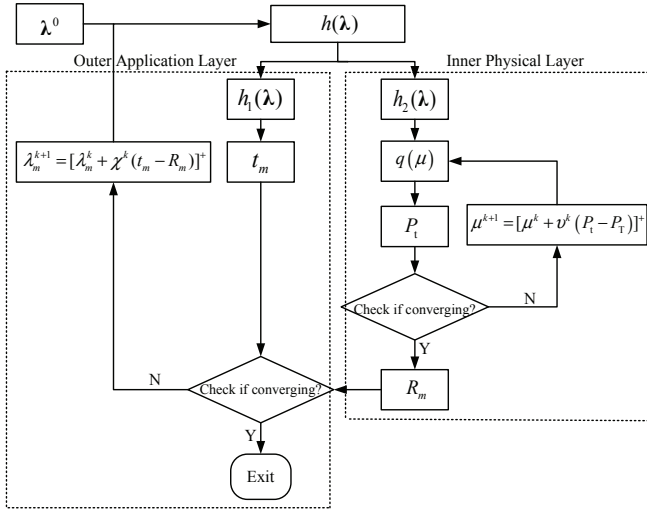The *application layer* subproblem is a utility maximization

Fig. 2. LDD-based cross-layer algorithm diagram.

problem as follows:

$$
h_1\left(\boldsymbol{\lambda}\right) = \max_{\mathbf{t}} f\left(\mathbf{t}\right) = w \frac{\left[\sum_{m=1}^{M} u_\alpha\left(t_m\right) - U_{\min}^{\alpha}\right]}{U_{\max}^{\alpha} - U_{\min}^{\alpha}} - \sum_{m=1}^{M} \lambda_m t_m. \tag{29}
$$

The *physical layer* subproblem is a joint subcarrier pairing and allocation, relay selection, and power allocation problem as the following:

$$
h_2\left(\boldsymbol{\lambda}\right) = \max_{\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} \sum_{m=1}^{M} \lambda_m R_m - (1-w)\frac{P_{total}}{P_{\max}}, \tag{30a}
$$

$$
\text{s.t. } C1 - C6. \tag{30b}
$$

Note that the two subproblems are not independent. They are interrelated by the dual vector $\boldsymbol{\lambda}$ across application layer and physical layer. Therefore, in the following, we will solve the two subproblems in (29) and (30) by cross-layer optimization. See Fig. 2 for the LDD-based cross-layer optimization diagram.

In (29), we note that both $u_\alpha\left(t_m\right)$ and $f\left(\mathbf{t}\right)$ are concave functions of t_m. Therefore, the optimal solutions of (29) are obtained by setting the derivation of $f\left(\mathbf{t}\right)$ with respect to t_m to zero. The optimal value of t_m is then obtained as

$$
t_m^* = \sqrt[\alpha]{\frac{w}{\lambda_m \left(U_{\max}^{\alpha} - U_{\min}^{\alpha}\right)}}. \tag{31}
$$

To solve the *physical layer* subproblem in (30), we adopt Lagrangian dual method where we further introduce another dual variable related to the price of transmit power to enable the second layer LDD. By relaxing C1, Lagrangian function associated to (30) is

$$
\begin{aligned}
Q\left(\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\mu\right) &= \sum_{m=1}^{M} \lambda_m R_m - (1-w)\frac{P_{total}}{P_{\max}} + \mu\left(P_{\mathrm{T}} - P_{\mathrm{t}}\right) \\
&= \sum_{m=1}^{M} \sum_{l=0}^{2L} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ \lambda_m r_{ij}^{ml} - \left[\frac{(1-w)\varepsilon}{2P_{\max}} + \frac{\mu}{2}\right] p_{ij}^{ml}\right\} \rho_{ij}\beta_{ij}^{ml} \\
&\quad - \frac{(1-w)P_{\mathrm{C}}}{P_{\max}} + \mu P_{\mathrm{T}},
\end{aligned} \tag{32}
$$

where μ is the dual variable defined corresponding to C1. The corresponding dual objective function is

$$
q\left(\mu\right) = \max_{\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta}} Q\left(\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\mu\right), \tag{33}
$$

and the dual problem is

$$
\min_{\mu \geq 0} q\left(\mu\right). \tag{34}
$$

Similar to the above, the derivative of $Q\left(\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\mu\right)$ with respect to $\mathbf{p}$ is set to zero. This provides us with the optimal power allocation for each transmission link (ml, ij) as the following:

$$
p_{ij}^{ml*} = \left[\Delta\left(m\right) - \frac{1}{g_{ij}^{ml}}\right]^{+}, \tag{35a}
$$

$$
p_{sd}^{m,i*} = \left[\Delta\left(m\right) - \frac{1}{g_{sd}^{m,i}}\right]^{+}, \tag{35b}
$$

$$
p_{sd}^{m,j*} = \left[\Delta\left(m\right) - \frac{1}{g_{sd}^{m,j}}\right]^{+}, \tag{35c}
$$

where $l \in \{1, 2, ..., 2L\}$, $(x)^{+} \triangleq \max\{0, x\}$, and

$$
\Delta\left(m\right) = \frac{\lambda_m}{\left[\frac{(1-w)\varepsilon}{P_{\max}} + \mu\right]\ln 2} \tag{36}
$$

is the water-filling level of user m. For $l = 0$, we further obtain

$$
p_{ij}^{ml*} = p_{sd}^{m,i*} + p_{sd}^{m,j*}, \quad l = 0. \tag{37}
$$

Here we define Ψ_{ij}^{ml} as the contribution of transmission link (ml, ij) to $Q\left(\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\mu\right)$. Therefore,

$$
\Psi_{ij}^{ml} = \lambda_m r_{ij}^{ml*} - \left[\frac{(1-w)\varepsilon}{2P_{\max}} + \frac{\mu}{2}\right]p_{ij}^{ml*}, \tag{38}
$$

where

$$
r_{ij}^{ml*} = \begin{cases} \frac{1}{2}\log_2\left(1 + g_{ij}^{ml} p_{ij}^{ml*}\right), & l \in \{1, 2, ..., 2L\}, \\ \frac{1}{2}\left[\log_2\left(1 + g_{sd}^{m,i} p_{sd}^{m,i*}\right) \right. \\ \quad \left. + \log_2\left(1 + g_{sd}^{m,j} p_{sd}^{m,j*}\right)\right], & l = 0. \end{cases} \tag{39}
$$

We further notice that the last two terms in (32) are both constants. Thus, the original Lagrangian function $Q\left(\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\mu\right)$ can be decomposed into $(2L+1)MN^2$ independent subproblems as the following:

$$
Q\left(p_{ij}^{ml}, \rho_{ij}, \beta_{ij}^{ml}, \mu\right) = \rho_{ij}\beta_{ij}^{ml}\Psi_{ij}^{ml}. \tag{40}
$$

According to C5 and C6, each subcarrier pair is exclusively assigned to only one relay-user pair. Therefore, to maximize $Q\left(\mathbf{p},\boldsymbol{\rho},\boldsymbol{\beta},\mu\right)$, the subcarrier pair (i, j) should be allocated to the relay-user pair (m, l) with maximum value of Ψ_{ij}^{ml}. This yields the optimal subcarrier pair allocation and relay selection indicator as

$$
\beta_{ij}^{ml*} = \begin{cases} 1, & \text{if } (m,l) = \arg\max_{m,l} \Psi_{ij}^{ml}, \\ 0, & \text{otherwise.} \end{cases} \tag{41}
$$

We then denote $\Omega_{ij} = \max_{m,l} \Psi_{ij}^{ml}$ and further simplify the

dual objective function (33) as the following:

$$q(\mu) = \max_{\boldsymbol{\rho}} \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \Omega_{ij}, \qquad (42)$$

which is in fact a two-dimensional assignment problem. Hungarian Algorithm (HA) is an efficient algorithm to obtain the solution of such assignment problem with the complexity of $O(N^3)$ [33]. Without loss of generality, we can express the subcarrier pairing result as

$$\rho_{ij}^{*} = \begin{cases} 1, & \text{if } (i,j) = \text{HA}(\Omega_{ij}), \\ 0, & \text{otherwise.} \end{cases} \qquad (43)$$

Finally, we use the sub-gradient method to minimize the dual objective function, where the dual variable is updated by the following:

$$\mu^{k+1} = \left[\mu^k + \upsilon^k (P_t - P_T)\right]^+, \qquad (44)$$

and υ^k is the diminishing step size at the kth iteration.

2) Case II: $\alpha = 0$: In this case, $U_\alpha(R_m) = R_m$, and the problem in (23) is reduced to

$$\max_{\mathbf{p}, \boldsymbol{\rho}, \boldsymbol{\beta}} w \frac{\sum_{m=1}^{M} R_m - U_{\min}^0}{U_{\max}^0 - U_{\min}^0} - (1-w) \frac{P_{\text{total}}}{P_{\max}}, \qquad (45\text{a})$$

$$\text{s.t. } C1 - C6. \qquad (45\text{b})$$

This optimization problem can be directly solved adopting LDD method, and only the physical layer problem is involved. Here for brevity we skip the details as it follows the same line of argument as in case I without introducing the auxiliary vector $\mathbf{t}$. The optimal power allocation to transmission link (ml, ij) is then obtained as

$$p_{ij}^{ml*} = \left[\Gamma - \frac{1}{g_{ij}^{ml}}\right]^+, \qquad (46\text{a})$$

$$p_{sd}^{m,i*} = \left[\Gamma - \frac{1}{g_{sd}^{m,i}}\right]^+, \qquad (46\text{b})$$

$$p_{sd}^{m,j*} = \left[\Gamma - \frac{1}{g_{sd}^{m,j}}\right]^+, \qquad (46\text{c})$$

where $l \in \{1, 2, ..., 2L\}$, and

$$\Gamma = \frac{w}{(U_{\max}^0 - U_{\min}^0) \left[\frac{(1-w)\varepsilon}{P_{\max}} + \mu\right] \ln 2} \qquad (47)$$

is the water-filling level for all users. Similarly for $l = 0$, we also have

$$p_{ij}^{ml*} = p_{sd}^{m,i*} + p_{sd}^{m,j*}, \quad l = 0. \qquad (48)$$

The resource allocation indicators are the same as in (41) and (43). In this case, however, the contribution of transmission link (ml, ij) is different from (38), which is

$$\Psi_{ij}^{ml} = \frac{w}{(U_{\max}^0 - U_{\min}^0)} r_{ij}^{ml*} - \left[\frac{(1-w)\varepsilon}{2P_{\max}} + \frac{\mu}{2}\right] p_{ij}^{ml*}. \qquad (49)$$

From (35), we observe that in the case of $\alpha > 0$, the optimal power allocation is in fact a multi-level water-filling

---

Algorithm 1 LDD-based cross-layer joint resource allocation algorithm (LDDA)

Step 1. For a given weighting parameter w, initialize the dual variables $\boldsymbol{\lambda}^0$ (if $\alpha > 0$) and μ^0;

Step 2. For each transmission link (ml, ij), obtain the optimal power allocation via (35) (if $\alpha > 0$) or (46) (if $\alpha = 0$) at given w, $\boldsymbol{\lambda}$ and μ, and then obtain the resource allocation indicators by (41) and (43), respectively;

Step 3. Update dual variable μ by the sub-gradient method in (44);

Step 4. Repeat Step 2 and Step 3 until the inner physical layer subproblem converges. If $\alpha = 0$, the algorithm terminates. If $\alpha > 0$, go to Step 5;

Step 5. Update dual vector $\boldsymbol{\lambda}$ by the sub-gradient method in (28);

Step 6. Repeat Step 2 to Step 5 until the outer application layer subproblem converges.

---

problem. In other words, by imposing fairness, users would have different water-filling levels according to their subcarrier allocation and relay selection. On the contrary, if $\alpha = 0$, all of the users will have the same water-filling level as in (46), which is the common case of optimal power allocation.

C. LDD-Based Cross-Layer Algorithm (LDDA) for SE, EE and Rate Fairness Tradeoff

The LDD-based cross-layer joint resource allocation algorithm (LDDA) for SE, EE and rate fairness tradeoff is outlined in **Algorithm 1**. To derive the joint resource allocation algorithm, we consider two cases: $\alpha > 0$, and $\alpha = 0$. For $\alpha > 0$, the solution is obtained through an LDD-based cross-layer algorithm, and the algorithm diagram is shown in Fig. 2. By introducing an auxiliary vector $\mathbf{t}$, this problem consists of two-layer subproblems, where the application layer subproblem is the outer layer and the physical layer subproblem is the inner layer. The outer and inner layer subproblems are interrelated via the dual vector $\boldsymbol{\lambda}$, reflecting the cross-layer interaction between the two layers. Hence, this is a cross-layer optimization algorithm. In each iteration, the inner physical layer subproblem is first solved through LDD for a given $\boldsymbol{\lambda}$. Then, $\boldsymbol{\lambda}$ is updated according to t_m^* and the solution to the inner layer subproblem, R_m. The iterations continue until the outer application layer subproblem converges, and the algorithm terminates by converging the outer layer subproblem. This process is outlined in Steps 2-6 in **Algorithm 1**. While for the case of $\alpha = 0$, since the auxiliary vector $\mathbf{t}$ is unnecessary, only the physical layer problem is involved. In this case, it can be directly solved using LDD.

IV. A HEURISTIC ALGORITHM FOR SE, EE AND RATE FAIRNESS TRADEOFF

Since the LDD-based cross-layer resource allocation algorithm in Section III needs to iteratively converge to the optimal solution, this results in high computational complexity for a large number of users and subcarriers. To address this issue,

in this section, we develop a novel low-complexity heuristic resource allocation algorithm to obtain a sub-optimal solution to the SE, EE and rate fairness tradeoff. The proposed low-complexity heuristic algorithm includes two key steps. We first obtain subcarrier pairing and allocation and relay selection, followed by optimal power allocation as a multi-level water-filling.

To develop the low-complexity resource allocation algorithm, here we define a new resource allocation indicator $\eta_{ij}^{ml} \in \{0,1\}$ and set $\eta_{ij}^{ml} = \rho_{ij}\beta_{ij}^{ml}$. Then the normalized end-to-end data rate over bandwidth for user m is obtained from (16) as

$$R_m = \sum_{l=0}^{2L} \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_{ij}^{ml} r_{ij}^{ml}. \tag{50}$$

The total transmit power consumption is

$$P_t = \frac{1}{2} \sum_{l=0}^{2L} \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{j=1}^{N} p_{ij}^{ml}, \tag{51}$$

where for $\eta_{ij}^{ml} = 1$, $p_{ij}^{ml} > 0$, and otherwise, $p_{ij}^{ml} = 0$.

Accordingly, the SOO problem with total transmit power constraint in (23) is reduced as

$$\max_{\mathbf{P},\boldsymbol{\eta}} w \frac{\sum_{m=1}^{M} u_\alpha (R_m) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - (1-w)\frac{P_{\text{total}}}{P_{\max}}, \tag{52a}$$

$$\text{s.t. } P_t \leq P_T, \tag{52b}$$

$$p_{ij}^{ml} \geq 0, \ \forall m,l,i,j, \tag{52c}$$

$$\sum_{m=1}^{M} \sum_{l=0}^{2L} \eta_{ij}^{ml} = 1, \ \forall i,j, \tag{52d}$$

$$\eta_{ij}^{ml} \in \{0,1\}, \ \forall m,l,i,j. \tag{52e}$$

Considering the SOO problem in (52), we present the following proposition.

Proposition 1: To obtain the optimal solution of (52) for any given fairness parameter α and weighting parameter w, the subcarrier pair (i,j) and relay $l \in \{0,1,2,...,2L\}$ should be allocated to user m^*, where

$$m^* = \arg \max_m \frac{r_{ij}^{ml}}{(R_m)^\alpha}, \tag{53}$$

and the optimal power allocated to this transmission link (m^*l, ij) is

$$p_{ij}^{m^*l} = \left[\Theta(m^*) - \frac{1}{g_{ij}^{m^*l}}\right]^+, l \in \{1,2,...,2L\}, \tag{54a}$$

$$p_{sd}^{m^*,i} = \left[\Theta(m^*) - \frac{1}{g_{sd}^{m^*,i}}\right]^+, \tag{54b}$$

$$p_{sd}^{m^*,j} = \left[\Theta(m^*) - \frac{1}{g_{sd}^{m^*,j}}\right]^+, \tag{54c}$$

where

$$\Theta(m^*) = \frac{w}{(U_{\max}^\alpha - U_{\min}^\alpha)(R_{m^*})^\alpha \left[\frac{(1-w)\varepsilon}{P_{\max}} + \mu\right] \ln 2} \tag{55}$$

is the water-filling level of user m^*.

Proof: See Appendix A. ∎

Similar to the Section III.B, if $\alpha > 0$, the optimal power allocation in (54) is also a multi-level water-filling, while for the case of $\alpha = 0$, the water-filling levels of all users are identical.

A. A Low-Complexity Heuristic Algorithm (LCA) for SE, EE and Rate Fairness Tradeoff

As it is seen in (53) and (54), the transmission link assignment and power allocation are interrelated. This is the main reason of the high computational complexity of obtaining the optimal solutions. To address this issue, in the first step, we assume equal power distribution among all transmission links, i.e., $p_{ij}^{ml} = P_T/N, \forall (ml, ij)$. Also, for the direct mode, equal power allocation is assumed between the two slots within a frame, i.e., $p_{sd}^{m,i} = p_{sd}^{m,j} = p_{ij}^{ml}/2$. Then in the second step we perform optimal power allocation.

In the first step, subcarrier pair (i,j) and relay l are allocated to user m^* based on (53), and then R_{m^*} is updated as, $R_{m^*} = R_{m^*} + r_{ij}^{m^*l}$. Note that the subcarrier pairing and allocation is conducted per subcarrier basis, and both allocated subcarriers are then removed from the set of available subcarriers. This process repeats until all subcarriers are paired and allocated.

According to (53), user m^* which has better channel gains is given a higher priority for allocating a transmission link. R_{m^*} is then updated and thus becomes larger. This reduces the chance of allocating another transmission link to user m^* and helps to impose rate fairness among users. As expected, by increasing α, the chance of allocating new transmission link to user m^* is also decreased, which means a stricter rate fairness among users in resource allocation.

The solution obtained through the above algorithm is in fact sub-optimal. This is because we assume equal power distribution among all transmission links and subcarriers pairing is conducted per subcarrier basis, without applying Hungarian algorithm which is optimal for two-dimensional assignment problem. In the simulation results, however, it is observed that only a slight performance gap exists between the proposed low-complexity heuristic algorithm and the LDD-based cross-layer algorithm.

Since each subcarrier pair (i,j) and relay l have been assigned to users, the optimal power allocation is performed in the second step. For notation brevity here, the resource allocation indicators η_{ij}^{ml} is substituted by its solution obtained in the first step, which is denoted as η_{ij}^{ml*}. Therefore, the normalized end-to-end data rate of user m is

$$R_m = \sum_{l=0}^{2L} \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_{ij}^{ml*} r_{ij}^{ml}. \tag{56}$$

Algorithm 2 Low-complexity heuristic resource allocation algorithm (LCA)

Step 1. Initialize $p_{ij}^{ml} = P_{\mathrm{T}}/N$ and $p_{sd}^{m,i} = p_{sd}^{m,j} = p_{ij}^{ml}/2, \forall m,l,i,j$;

Step 2. Initialize the data rate of each user as $R(m) = 1$. The available subcarrier sets in the first and second slot are Λ and Π, respectively, i.e., $i \in \Lambda$ and $j \in \Pi$;

Step 3. For each available subcarrier $i \in \Lambda$, allocate a subcarrier $j \in \Pi$, a relay $l \in \{0,1,2,...,2L\}$ to user m^* according to (53);

Step 4. Update available subcarrier sets as $\Lambda=\Lambda$-i and $\Pi=\Pi$-j; Update $R(m^*) = R(m^*) + r_{ij}^{m^*l}$;

Step 5. Repeat Step 3 and Step 4 until all available subcarrier pairs are allocated to users;

Step 6. Conduct optimal power allocation by standard convex optimization method based on the convex optimization problem in (57).

The SOO problem in (52) then becomes only a function of p_{ij}^{ml}, and it can therefore be simplified as the following:

$$\max_{\mathbf{P}} \quad w\frac{\sum_{m=1}^{M} u_\alpha(R_m) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - (1-w)\frac{P_{\mathrm{total}}}{P_{\max}}, \quad (57a)$$

$$\text{s.t.} \quad P_t \le P_{\mathrm{T}}, \quad (57b)$$

$$p_{ij}^{ml} \ge 0, \quad \forall m,l,i,j. \quad (57c)$$

Proposition 2: For any given fairness parameter α and weighting parameter w, (57) is a convex optimization problem.

Proof: See Appendix B. ∎

According to Proposition 2, (57) has a unique global optimal solution. There exist many efficient numerical algorithms such as the interior-point method to obtain the optimal solution.

The proposed low-complexity heuristic resource allocation algorithm (LCA) for SE, EE and rate fairness tradeoff is outlined in **Algorithm 2**.

B. Complexity Analysis

Here we compare computational complexity of the exhaustive search method and the two proposed algorithms. For the exhaustive search method, the complexity of subcarrier pair allocation is $O\left\{[(2L+1)M]^N\right\}$. For subcarrier pairing within the two slots, its complexity is $O(N!)$. Hence, the total complexity of exhaustive search method is $O\left\{N![(2L+1)M]^N\right\}$.

For LDDA, $(2L+1)MN^2$ times of optimal power allocation are calculated by (35) or (46). Further, the complexity of Hungarian Algorithm for subcarrier pairing is $O(N^3)$. For Lagrangian dual variable update based on sub-gradient method, its complexity is a polynomial function of the dual problem dimension, i.e., M for $h(\boldsymbol{\lambda})$, and 1 for $q(\mu)$ [30]. Therefore, the complexity of updating all dual variables is in the order of M^ϕ, where ϕ is a positive constant [2]. Hence, the total complexity of LDDA is $O\left\{\left[(2L+1)MN^2 + N^3\right]M^\phi\right\}$. For
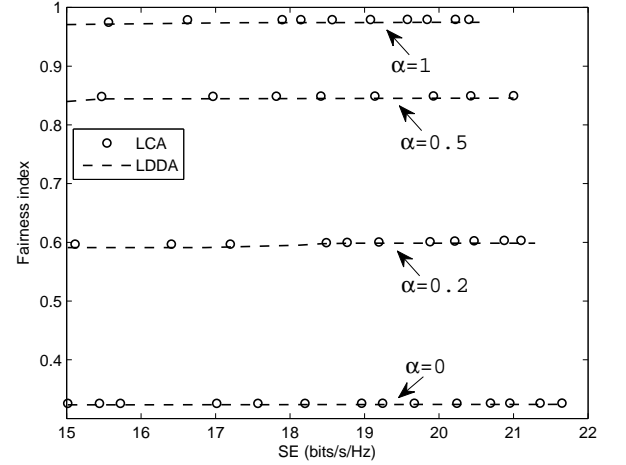

Fig. 3. Jain's fairness index *vs.* spectral efficiency.

LDDA without subcarrier pairing (LDDAwoSP), its complexity is $O\left\{\left[(2L+1)MN^2\right]M^\phi\right\}$.

For LCA, the complexity of subcarrier pairing is $N + (N-1) + ... + 2 + 1 = \frac{1}{2}N(N+1)$ for each relay and user pair. There are $(2L+1)M$ possible relay and user pairs, so the total complexity of LCA is $O\left[\frac{1}{2}N(N+1)(2L+1)M\right]$. Note that in wireless systems, the number of users, M, is usually much smaller than the number of subcarriers, N, i.e., $M \ll N$. Therefore, compared with LDDA and LDDAwoSP, the complexity of LCA is significantly reduced.

V. SIMULATION RESULTS

In this section, we investigate the performance of the two proposed algorithms through simulations. In the simulations, we consider a cellular network with $M = 8$ users randomly and uniformly distributed at the cell-edge region. BS is located at the center of the cell and the number of relays L, is equal to 3. Each relay is located on the axis of the corresponding sector with equal angle interval of $2\pi/L$, and the distance between each relay and BS is half of the cell radius. The number of subcarriers is $N = 128$ in both slots and the noise power spectral density is -174dBm/Hz. Without loss of generality, the circuit power of the BS and each relay is normalized to 1W while the drain efficiencies of the power amplifiers are assumed to be 38% as in [23]. The links between the BS and relays are in line-of-sight (LOS) and each user experiences independent frequency-selective Rayleigh fading. The modified Hata urban propagation model is adopted for the large-scale propagation loss, and the shadowing follows log-normal distribution with zero-mean and standard deviation of 8dB.

A. SE, EE and Rate Fairness Tradeoff

Fig. 3 shows the rate fairness performance of the two proposed algorithms with different values of α, where the fairness performance is measured by Jain's fairness index [7]. The value of Jain's fairness index is bounded between 0 and 1. If the index is 1, it means all users get the same data rate
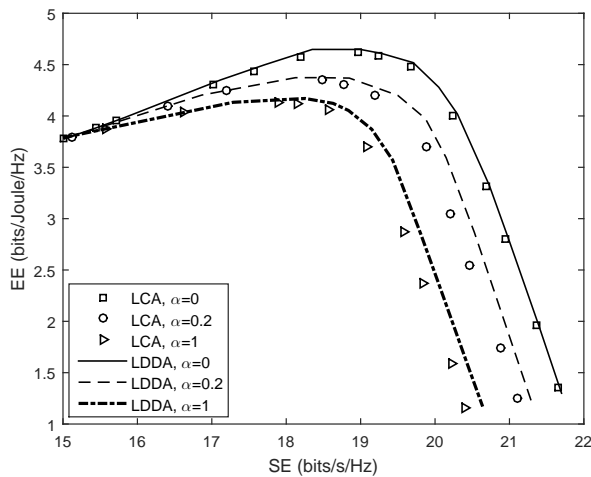
Fig. 4. Energy efficiency *vs.* spectral efficiency with different levels of fairness.

Fig. 5. Energy efficiency *vs.* spectral efficiency with different levels of fairness.

and the system is 100% fair. As the disparity of data rate increases, the index gradually decreases to 0. For simplicity, in Fig. 3, we only show four different values of α. As it is seen, by increasing α, the fairness among users is also increased. This verifies that the two proposed algorithms can both achieve different levels of rate fairness by adjusting α. We also observe that the two proposed algorithms achieve almost identical fairness performance for the same value of α. In fact, by continuously adjusting α, the algorithms can achieve any certain level of fairness, from no fairness ($\alpha = 0$) to absolute fairness ($\alpha \to \infty$).

Fig. 4 shows the Pareto optimal sets for SE-EE tradeoff with different levels of fairness. LDDA is based on dual method and for a large number of subcarriers, its duality gap becomes negligible. Therefore, we consider LDDA as the benchmark for performance comparisons. As shown in this figure, LCA follows the same pattern with LDDA. With the increase of SE, EE first increases and then decreases, and the performance gap between LCA and LDDA is slight. When SE is lower, since the radiated power is negligible compared with the circuit power, the growing of SE is much faster than that of the total power consumption. Thus, EE grows as SE increases. However, after the maximum point for EE, the circuit power does not dominate any longer and the increase of radiated power greatly affects the total power consumption. Under this circumstance, the growing of SE becomes slower than that of the total power consumption, and accordingly, EE gradually declines to a very low level. Note that as SE goes to infinity, EE will asymptotically approach zero. Therefore, the EE-SE relationship is actually quasiconcave.

In Fig. 4, it is also indicated that with larger α, for the same level of SE (EE), a lower EE (SE) is achieved compared with the case with smaller α. This means higher fairness results in a worse SE-EE tradeoff. As mentioned previously, since efficiency and fairness are conflicting objectives, fairness is usually enhanced at the cost of degrading the performance of SE-EE tradeoff. Besides, it is noteworthy that in this figure, after the maximum points of EE, a moderate reduction of SE
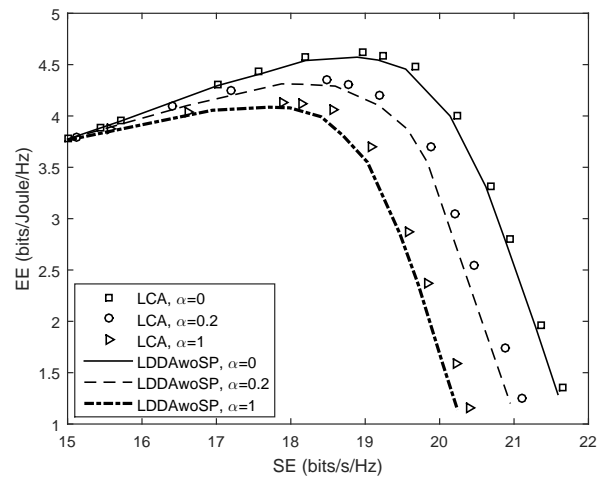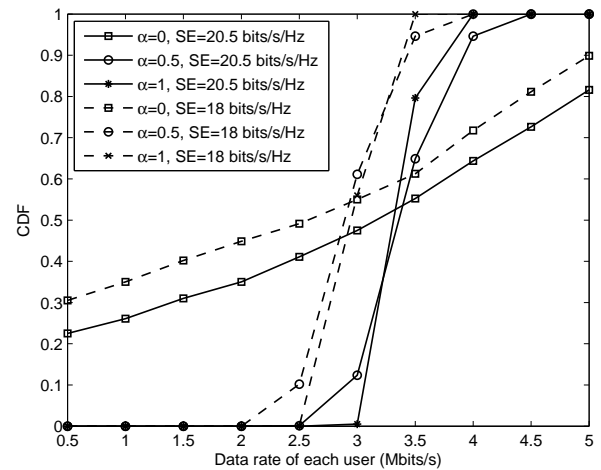

Fig. 6. CDF *vs.* data rate of each user.

may allow a significant improvement of EE, i.e., a noticeable energy saving. Fig. 4 provides the optimal envelop of the entire SE-EE region for different levels of fairness, so it is flexible to make SE-EE tradeoff for different preferences. But in practice, only the tradeoff after the maximum points of EE make sense. That is because before these maximum points, by increasing SE, EE also increases. In that case, SE and EE are not conflicting objectives.

Besides, we also compare the performance of LCA and LDDAwoSP (LDDA without subcarrier pairing). As shown in Fig. 5, LCA slightly outperforms LDDAwoSP. Furthermore, as mentioned in Section IV.B, the complexity of LCA is much lower than that of LDDAwoSP, especially when the numbers of users and subcarriers are very large.

To further illustrate the fairness performance, Fig. 6 shows the cumulative distribution function (CDF) of the data rate of each user. Here, we only take LCA as an example, and similar results can also be obtained for LDDA. Two cases of SE=20.5bits/s/Hz and SE=18bits/s/Hz are investigated. From

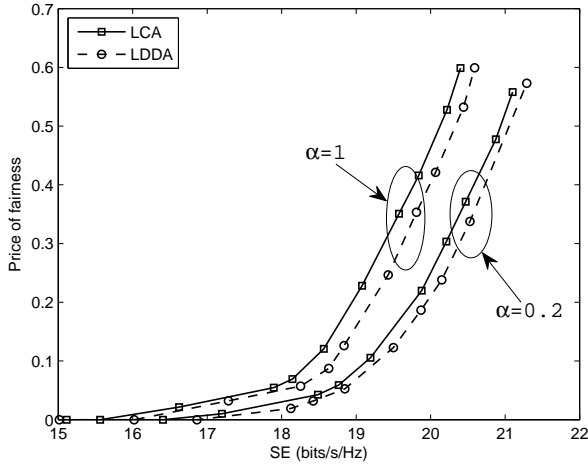Fig. 7. Price of fairness (PoF) *vs.* spectral efficiency.

Fig. 8. Outage probability for different levels of fairness.

this figure, we can see that for both cases of SE, when α is lower, the ratio of users with smaller data rate is relatively higher, and the CDF curves experience a slower increase. This means the data rate variance is larger, and accordingly, some users suffer from unfairness. While by increasing α, for example, when $\alpha = 1$, the CDF jumps to 100%, implying the users' data rate distribution is better balanced, and this is consistent with the high fairness index in Fig. 3. Besides, when the value of α is higher, the ratio of users with higher data rate is lower. It stems from the fact that due to the fairness requirement, users with better channel conditions sacrifice their data rates to compensate those users with worse channel conditions. Such compensations will degrade the performance of SE-EE tradeoff, which is consistent with the previous results in Fig. 4 and Fig. 5.

B. Price of Fairness

To quantify the tradeoff between efficiency and fairness, we adopt the metric of price of fairness (PoF) as in [34], [35]. The PoF is defined as $PoF\left(EE\left(\alpha\right)\right) = \frac{EE(0)-EE(\alpha)}{EE(0)}$, where $EE\left(\alpha\right)$, and $EE\left(0\right)$ are the system EE with α-fairness, and without fairness (i.e., $\alpha = 0$), respectively. For various values of α, PoF quantifies the reduction of system EE caused by imposing rate fairness, in comparison with the system without fairness.

The PoF versus SE is shown in Fig. 7. As illustrated, for a fixed SE, the PoF with a larger α is higher than the case with a smaller α. Therefore, the better the rate fairness, the higher is the EE loss. Interestingly, as SE increases, the PoF goes up very significantly. For instance, for LDDA, when $\alpha = 1$, by increasing SE from 19bits/s/Hz to 20.5bits/s/Hz, the PoF rises from 0.15 to 0.59 by up to three times. Therefore, if SE is increased in the SE-EE tradeoff, a higher EE sacrifice level is required to guarantee the fairness among users. Besides, with the same value of α, the PoF of LCA is always larger than LDDA for any certain SE. The metric of PoF enables the network operators to quantitatively evaluate the cost of fairness
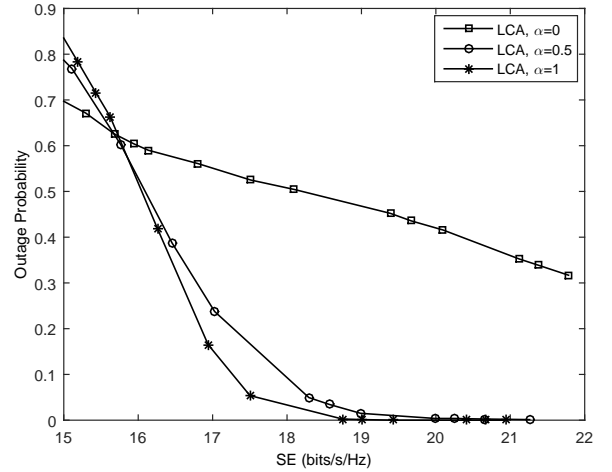
and balance the tradeoff between efficiency and fairness for different preferences.

C. Impact of Fairness on the Outage Probability

Fig. 8 demonstrates the outage probability for different levels of fairness. Outage probability is defined as the probability that user's data rate drops below the minimum rate requirement (MinR). Here the MinR is set as 2.5 Mbits/s for each user. As displayed, fairness requirement helps to reduce the outage probability dramatically, providing enhanced quality of experience (QoE) to more users than the scheme without fairness (i.e., $\alpha = 0$). This is because imposing fairness helps to balance the users' data rate distribution, as shown in Fig. 6. With the increase of system SE, the better balanced data rates among users exceed MinR and hence the outage probability plummets to near 0. Compared with them, if there is no fairness requirement, the data rate distribution would be highly concentrated to those users with better channel conditions. Thus, the outage probability declines slowly.

D. Impact of Number of Relays and Circuit Power

Fig. 9 illustrates the SE-EE tradeoff performance with different numbers of relays, where the maximum transmit power constraints are identical. We only show the case of $\alpha = 0.5$ as an example. As can be seen from this figure, since having more relays introduces higher circuit power, for most cases of SE, EE declines with the increase of the number of relays. However, when SE is larger, the EE may even become larger by increasing L. This is because with more relays, the system will have a higher degree of freedom in resource allocation, and accordingly, the achievable SE is increased. By exploring the diversity gains brought by more relays, the negative impact of increasing circuit power on EE might be less than the positive contribution of increased SE. Hence, introducing more relays does not necessarily always degrade the performance of SE-EE tradeoff. It is predicted that if the maximum transmit power (i.e., maximum achievable
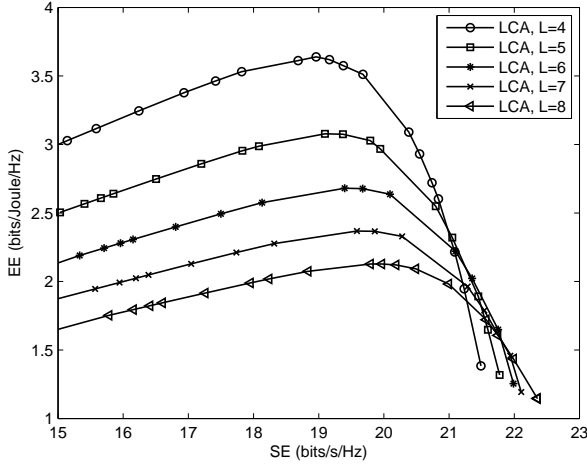
Fig. 9. Energy efficiency *vs.* spectral efficiency with different numbers of relays.
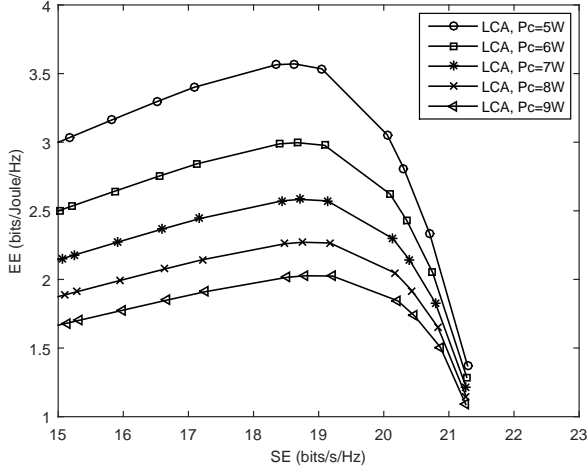

Fig. 10. Energy efficiency *vs.* spectral efficiency with different circuit power.

SE) is further increased, this trend will be more significant. Consequently, how many relays should be deployed in the cell is also a tradeoff for different preferences.

The SE-EE tradeoff for different circuit power with $\alpha = 0.5$ is displayed in Fig. 10, where the number of relays is fixed (i.e., $L = 3$). Different from Fig. 9, as circuit power has no impact on the achievable SE, EE is always reduced by increasing circuit power. Therefore, given a fixed number of relays, increasing circuit power always degrades the performance of SE-EE tradeoff.

VI. CONCLUSION

In this paper, a general framework is presented to analyze the three-factor tradeoff among SE, EE and rate fairness in relay-aided cooperative OFDMA systems. We formulate the problem as a MOO problem where rate fairness is represented using α-fairness model. A LDD-based cross-layer joint resource allocation algorithm (LDDA) and a low-complexity heuristic resource allocation algorithm (LCA) are proposed

to efficiently manage the three-factor tradeoff. The Pareto optimal solution is obtained to show the global relationship of SE and EE, while the PoF is applied to quantify the tradeoff of efficiency and fairness. Simulation results show that imposing a higher level of fairness may significantly reduce the outage probability. Besides, by increasing the number of relays, although the total circuit power is increased, the SE-EE tradeoff is not necessarily degraded due to the extra degree of freedom provided in relay selection.

APPENDIX A
PROOF OF PROPOSITION 1

Proof: We adopt Lagrangian dual method. Relaxing constraints C1 and C3 in (52), Lagrangian function associated to (52) is

$$G\left(\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mu\right) = w \frac{\sum_{m=1}^{M} u_\alpha\left(R_m\right) - U_{\min}^\alpha}{U_{\max}^\alpha - U_{\min}^\alpha} - (1-w)\frac{P_{\text{total}}}{P_{\max}} \\ + \sum_{i=1}^{N}\sum_{j=1}^{N}\theta_{ij}\left(1 - \sum_{m=1}^{M}\sum_{l=0}^{2L}\eta_{ij}^{ml}\right) + \mu\left(P_{\text{T}} - P_t\right), \quad (58)$$

where $\boldsymbol{\theta}$ and μ are Lagrangian dual variables.

By taking derivation of $G\left(\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mu\right)$ with respect to η_{ij}^{ml}, we obtain the necessary conditions for optimal resource allocation according to Karush-Kuhn-Tucker (KKT) conditions [36] as

$$\frac{\partial G\left(\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mu\right)}{\partial \eta_{ij}^{ml}} = \frac{w r_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha} - \theta_{ij} \le 0, \quad (59)$$

$$\eta_{ij}^{ml}\left(\frac{w r_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha} - \theta_{ij}\right) = 0. \quad (60)$$

Based on (59) and (60), if subcarrier pair (i, j) and relay l are allocated to user m, i.e., $\eta_{ij}^{ml} = 1$, then $\left(\frac{w r_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha} - \theta_{ij}\right) = 0$; otherwise, $\left(\frac{w r_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha} - \theta_{ij}\right) \le 0$. Therefore, subcarrier pair (i, j) and relay l should be allocated to user m with the highest value of $\left(\frac{w r_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha} - \theta_{ij}\right)$. Given w, $U_{\max}^\alpha$ and $U_{\min}^\alpha$ are all constants, the subcarrier pair (i, j) and relay l should be allocated by (53).

Similarly, by taking derivation of $G\left(\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mu\right)$ with respect to p_{ij}^{ml} ($l \neq 0$), we have

$$\frac{\partial G\left(\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mu\right)}{\partial p_{ij}^{ml}} \\ = \frac{w g_{ij}^{ml}\eta_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha\left(1 + g_{ij}^{ml}p_{ij}^{ml}\right)\ln 2} - \frac{(1-w)\varepsilon}{P_{\max}} - \mu \le 0, \quad (61)$$

$$p_{ij}^{ml}\left(\frac{w g_{ij}^{ml}\eta_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha\left(1 + g_{ij}^{ml}p_{ij}^{ml}\right)\ln 2} - \frac{(1-w)\varepsilon}{P_{\max}} - \mu\right) = 0. \quad (62)$$

If subcarrier pair (i, j) and relay l are allocated to user m, i.e., $\eta_{ij}^{ml} = 1$, then $p_{ij}^{ml} > 0$. Thus, from (62), we have

$$\frac{w g_{ij}^{ml}\eta_{ij}^{ml}}{\left(U_{\max}^\alpha - U_{\min}^\alpha\right)\left(R_m\right)^\alpha\left(1 + g_{ij}^{ml}p_{ij}^{ml}\right)\ln 2} - \frac{(1-w)\varepsilon}{P_{\max}} - \mu = 0. \quad (63)$$

Then, by substituting $\eta_{ij}^{ml} = 1$ into (63), the optimal power allocation p_{ij}^{ml} ($l \neq 0$) can be obtained as (54a).

For $l = 0$, following the same line of argument, the optimal power allocation $p_{sd}^{m,i}$ and $p_{sd}^{m,j}$ are obtained as (54b) and (54c). This completes the proof. ∎

APPENDIX B
PROOF OF PROPOSITION 2

Proof: It is easy to show that R_m is a concave function of p_{ij}^{ml}. Further, since α-fair utility function is strictly increasing and concave for any given α, according to (3.10) in [36], their composition, $u_\alpha(R_m)$, is also a concave function of p_{ij}^{ml}. Besides, $-P_{\text{total}}$ is a linear function of p_{ij}^{ml}, which is concave.

Therefore, the objective function in (57) can be viewed as a nonnegative weighted summation of concave functions. Hence, according to [36], the objective function is still a concave function of p_{ij}^{ml} for any given fairness parameter α and weighting parameter w. ∎

REFERENCES

[1] Y. A. Sambo *et al.*, "Expanding cellular coverage via cell-edge deployment in heterogeneous networks: spectral efficiency and backhaul power consumption perspectives," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 140–149, Jun. 2014.

[2] W. Dang, M. Tao, H. Mu, and J. Huang, "Subcarrier-pair based resource allocation cooperative multi-relay OFDM systems," *IEEE Trans. Wirel. Commun.*, vol. 9, no. 5, pp. 1640–1649, May 2010.

[3] X. Zhang, X. Shen, and L.-L. Xie, "Joint subcarrier and power allocation for cooperative communications in LTE-advanced networks," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 2, pp. 658–668, Feb. 2014.

[4] X. Zhang, X. Tao, Y. Li, N. Ge, and J. Lu, "On relay selection and subcarrier assignment for multi-user cooperative OFDMA networks with QoS guarantees," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4704–4717, Nov. 2014.

[5] G. Y. Li, Z. Xu, C. Xiong, C. Yang *et al.*, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.

[6] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[7] M. Salem, A. Adinoyi, M. Rahman, H. Yanikomeroglu *et al.*, "An overview of radio resource management in relay-enhanced OFDMA-based networks," *IEEE Commun. Surveys & Tutorials*, vol. 12, no. 3, pp. 422–437, Jul. 2010.

[8] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental tradeoffs on green wireless netowrks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, Jun. 2011.

[9] G. W. Miao, N. Himayat, G. Y. Li, and D. Bormann, "Energy efficient design in wireless OFDMA," in *Proc. IEEE Int. Conf. Commun.*, May 2008, pp. 3307–3312.

[10] G. W. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 545–554, Jan. 2010.

[11] G. W. Miao and A. Vastberg, "Energy efficiency in the wideband regime," in *IEEE Wireless Commun. and Networking Conf.*, Bangalore, India, Jan. 2013, pp. 261–267.

[12] X. Xiao, X. Tao, and J. Lu, "QoS-aware energy-efficient radio resource scheduling in multi-user OFDMA systems," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 75–78, Jan. 2013.

[13] C. C. Zarakovitis and Q. Ni, "Maximising energy efficiency in multi-user multi-carrier broadband wireless systems: convex relaxation and global optimisation techniques," *IEEE Trans. Veh. Technol.*, 2015.

[14] K. T. K. Cheung, S. Yang, and L. Hanzo, "Achieving maximum energy-efficiency in multi-relay OFDMA cellular networks: a fractional programming approach," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2746–2757, Jul. 2013.

[15] S. Huang, H. Chen, J. Cai, and F. Zhao, "Energy efficiency and spectral efficiency tradeoff in amplify-and-froward relay networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4366–4378, Nov. 2013.

[16] 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.

[17] F. Haider, C.-X. Wang, H. Haas, E. Hepsaydir, and X. Ge, "Energy-efficient subcarrier-and-bit allocation in multi-user OFDMA systems," in *Proc. IEEE VTC-Spring*, May 2012, pp. 1–5.

[18] X. Ge, X. Huang, Y. Wang, M. Chen *et al.*, "Energy efficiency optimization for MIMO-OFDM mobile multimedia communication systems with QoS constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2127–2138, Jun. 2014.

[19] X. Ge, T. Han, Y. Zhang, G. Mao *et al.*, "Spectrum and energy efficiency evaluation of two-tier femtocell networks with partially open channels," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1306–1319, Mar. 2014.

[20] J. Tang, D. K. C. So, E. Alsusa, and K. A. Hamdi, "Resource efficiency: a new paradigm on energy efficiency and spectral efficiency tradeoff," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 8, pp. 4656–4669, Aug. 2014.

[21] L. Deng, Y. Rui, P. Cheng, J. Zhang, Q. T. Zhang, and M. Li, "A unified energy efficiency and spectral efficiency tradeoff metric in wireless networks," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 55–58, Jan. 2013.

[22] I. Ku, C.-X. Wang, and J. Thompson, "Spectral-energy efficiency tradeoff in relay-aided cellular networks," *IEEE Trans. Wirel. Commun.*, vol. 12, no. 10, pp. 4970–4982, Oct. 2013.

[23] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3767–3778, Dec. 2012.

[24] K. Illanko, M. Naeem, A. Anpalagan, and D. Androutsos, "Frequency and power allocation for energy efficient OFDMA systems with proportional rate constraints," *IEEE Wirel. Commun. Lett.*, vol. 3, no. 3, pp. 313–316, Jun. 2014.

[25] C. He, G. Y. Li, F.-C. Zheng, and X. You, "Energy-efficient resource allocation in OFDM systems with distributed antennas," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1223–1231, Mar. 2014.

[26] Z. Ren, S. Chen, B. Hu, and W. Ma, "Energy-efficient resource allocation in downlink OFDM wireless systems with proportional rate constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2139–2150, Jun. 2014.

[27] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, Oct. 2000.

[28] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidisc. Optim.*, vol. 26, no. 6, pp. 369–395, Apr. 2004.

[29] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[30] T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 328–339, Feb. 2007.

[31] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[32] D. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.

[33] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics*, vol. 52, pp. 7–21, 2005.

[34] D. Bertsimas, V. F. Farias *et al.*, "The price of fairness," *Operations Research*, vol. 59, no. 1, pp. 17–31, Jan.-Feb. 2011.

[35] K. Davaslioglu and E. Ayanoglu, "Efficiency and fairness trade-offs in SC-FDMA schedulers," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 6, pp. 2991–3002, Jun. 2014.

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

Zhengyu Song received the B.E. and M.E. degrees from Beijing Jiaotong University and the Ph.D. degree from Beijing Institute of Technology, Beijing, China, all in Information and Communication Engineering. His research interests lies in the field of radio resource management in wireless communication systems and cooperative communications.