

Achieving Maximum Energy-Efficiency in Multi-Relay OFDMA Cellular Networks: A Fractional Programming Approach

Kent Tsz Kan Cheung, Shaoshi Yang, and Lajos Hanzo, *Fellow, IEEE*

Abstract—In this paper, the joint power and subcarrier allocation problem is solved in the context of maximizing the energy-efficiency (EE) of a multi-user, multi-relay orthogonal frequency division multiple access (OFDMA) cellular network, where the objective function is formulated as the ratio of the spectral-efficiency (SE) over the total power dissipation. It is proven that the fractional programming problem considered is quasi-concave so that Dinkelbach’s method may be employed for finding the optimal solution at a low complexity. This method solves the above-mentioned master problem by solving a series of parameterized concave secondary problems. These secondary problems are solved using a dual decomposition approach, where each secondary problem is further decomposed into a number of similar subproblems. The impact of various system parameters on the attainable EE and SE of the system employing both EE maximization (EEM) and SE maximization (SEM) algorithms is characterized. In particular, it is observed that increasing the number of relays for a range of cell sizes, although marginally increases the attainable SE, reduces the EE significantly. It is noted that the highest SE and EE are achieved, when the relays are placed closer to the BS to take advantage of the resultant line-of-sight link. Furthermore, increasing both the number of available subcarriers and the number of active user equipment (UE) increases both the EE and the total SE of the system as a benefit of the increased frequency and multi-user diversity, respectively. Finally, it is demonstrated that as expected, increasing the available power tends to improve the SE, when using the SEM algorithm. By contrast, given a sufficiently high available power, the EEM algorithm attains the maximum achievable EE and a suboptimal SE.

Index Terms—Subcarrier/power allocation, green communications, energy-efficiency, multiple relays, dual decomposition, fractional programming.

I. INTRODUCTION

ENERGY-efficiency (EE) is becoming of great concern in the telecommunications community owing to the rapidly increasing data rate requirements, increasing energy prices, and societal as well as political pressures on mobile phone

This research has been funded by the Industrial Companies who are Members of the Mobile VCE, with additional financial support from the UK Government’s Engineering & Physical Sciences Research Council (EPSRC). The financial support of the China Scholarship Council (CSC), of the Research Councils UK (RCUK) under the India-UK Advanced Technology Center (IU-ATC), of the EU under the auspices of the Concerto project, and of the European Research Council’s Senior Research Fellow Grant is also gratefully acknowledged.

K. T. K. Cheung is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (e-mail: ktkc106@ecs.soton.ac.uk).

S. Yang is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK. He is also with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China (e-mail: sy7g09@ecs.soton.ac.uk).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK (e-mail: lh@ecs.soton.ac.uk).

operators to reduce their ‘carbon footprint’ [1]. This has led to several joint academic and industrial research efforts dedicated to developing novel energy-saving techniques, such as the ‘green radio’ project [2], the GreenTouch alliance [3], and the energy aware radio and network technologies (EARTH) project [4]. Substantial research efforts have also been dedicated to the next-generation wireless networks, such as the third generation partnership project’s (3GPP) long term evolution-advanced (LTE-A) and IEEE 802.16 worldwide interoperability for microwave access (WiMAX) [5] standards, which may rely on relaying between the central base station (BS) and the user equipment (UE). As a benefit of reduced transmission distances, either the quality of the communication is maintained at reduced power requirements, or the transmission integrity is improved at the same power consumption. This allows the need for expensive deployment and maintenance of additional BSs to be circumvented. The two most popular relaying techniques are the amplify-and-forward (AF) and the decode-and-forward (DF) schemes [6]. The AF regime is less complex than DF, since the relay node (RN) needs only to receive and linearly amplify the source’s transmissions, before forwarding it to the destination. The effects of scheduling and frequency reuse in the context of the above-mentioned networks was studied in [7].

Both LTE-A and WiMAX employ the orthogonal frequency division multiple access (OFDMA) technique. In OFDMA, the whole channel’s bandwidth is divided into multiple subcarriers, where subsets of subcarriers may be allocated for transmission to different users [8]. In OFDMA, the system attains two types of diversity, which may be jointly exploited for improving the achievable sum-rate (SR) of the system. Firstly, multi-user diversity is attained with the aid of appropriate user mapping, because when the channel from the BS to a specific UE is undergoing severe fading on a particular subcarrier, then this subcarrier may be assigned for transmission to another user, whose channel might be more friendly. On the other hand, activating only those subcarriers that are suitable for high-quality transmission to a particular UE leads to frequency diversity. These philosophies underpin several contributions in the literature, where the goal is to assign the available resources, for example power and subcarriers, so that a system-wide metric is maximized. These methods belong to the family of resource allocation policies and typically aim for solving one of two problems: either the spectral-efficiency (SE)¹ [9], [10], [11] of the system is maximized while a maximum power constraint is enforced, or the power consumption is minimized under a minimum total system throughput or individual UE

¹For a given system bandwidth, the SR maximization and SE maximization (SEM) solutions are identical. To avoid any additional abbreviations, they are both henceforth referred to as SEM.

rate constraint² [12], [13], [14], [15], [16].

An example of the SE Maximization (SEM) problem was considered in [11], where the authors formulate the optimization problem for the downlink (DL) of an AF relaying-aided OFDMA cellular network, and their goal was to optimize the power and subcarrier allocation so that the SE of the system was maximized under maximum outage probability and total power constraints. In the class of power minimization problems, an example is the often-cited work by Wong *et. al* [12], where a heuristic bit allocation algorithm was conceived for a multi-user OFDMA system with the aim of minimizing the power consumption under a minimum individual user rate constraint. With a similar goal, Piazzo [13] developed a sub-optimal bit allocation algorithm for an orthogonal frequency division multiplexing (OFDM) system. This work was later extended to provide the optimal bit allocation in [15].

However, the SEM and the power minimization problems do not directly consider an EE objective function (OF), and in general they do not deliver the EE maximization (EEM) solution. In recent years, research into resource allocation using an EE OF has become increasingly popular. In reality, EEM may be viewed as an example of multi-objective optimization, since typically the goal is to maximize the SE achieved, whilst concurrently minimizing the power consumption required. From this perspective, [17] derives an aggregate OF, which consists of a weighted sum of the SR achieved and the power dissipated. However, selecting appropriate weights for the two OFs is not straightforward, and different combinations of weights can lead to very different results. Another example is given in [18], where the EEM problem is considered in a multi-relay network employing the AF protocol. However, both [17], [18] only optimize the user selection and power allocation without considering the subcarrier allocation in the network. Another formulation, demonstrated in [19], [20], considers power and subcarrier allocation in an OFDMA cellular network, but without a maximum total power constraint and without relaying. The authors of [21] formulate the EEM problem in a OFDMA cellular network under a maximum total power constraint, however relaying is not considered.

In light of the above discussions, this work focuses on a solution method for the EEM problem in a multi-relay, multi-user OFDMA cellular network, which jointly considers both power and subcarrier allocation as well as a maximum total power constraint. The contributions of this paper is summarized as follows:

- The EEM problem in the context of a multi-relay, multi-user OFDMA cellular network, in which both direct and relayed transmissions are employed, is formulated as a fractional programming problem, which jointly considers both the power and subcarrier allocation. In contrast to previous contributions such as [7], the aim is for finding the optimal power and subcarrier allocations within a network context. Furthermore, in contrast to [12], [13], [9], [10], [14], [11], [15], [18], [17], the focus is placed on an EE OF. It is demonstrated that in its original form the problem is a mixed-integer non-linear programming problem (MINLP) [22], which is challenging to solve. In order to make the problem more tractable, both a variable transformation and a relaxation of the integer variables is

²In the latter case, the minimum rate constraint may be viewed as ensuring fairness among the users, since each user achieves at least a minimum rate.

introduced.

- It is proven that the relaxed problem is quasi-concave and consequently Dinkelbach's method [23] may be employed for obtaining the optimal solution by solving a sequence of parameterized secondary problems. Each of these are solved using the dual decomposition approach of [24]. It is demonstrated that the EEM algorithm reaches the optimal solution within a low number of iterations and reaches the optimal solution obtained via an exhaustive search. Thus the original problem is solved at a low complexity.
- Comparisons are made between two multi-relay resource allocation problems, namely one that solves the EEM problem and another that considers SEM. As an example, it is shown that when the maximum affordable power is lower than a given threshold, the two problems have the same solutions. However, as the maximum affordable power is increased, the SEM algorithm attempts to achieve a higher SE at the cost of a lower EE, while given the total power, the EEM algorithm reaches the upper limit of the maximum achievable SE for the sake of maintaining the maximum EE.
- Since the system model is generalized, the EEM and SEM algorithms may be employed for gaining insights into network design, when the aim is for maximizing either the EE or SE. To that end, a comprehensive range of results are presented, which demonstrate both the effect of increasing the number of available subcarriers and UEs in the system, as well as quantifying the impact of increasing the number of RNs in the system and its relation not only to the cell radius, but also to the relays' positions. The algorithm may be used for characterizing the effects of many other system design choices on the maximum SE and EE.

The rest of this paper is organized as follows. In Section II, the multi-user, multi-relay OFDMA cellular network model is described, which is followed by a formulation of the optimization problem in Section III. Upon invoking a transformation of variables and a relaxation of the integer variables, it is proven that the OF is quasi-concave. The combined solution algorithm of Dinkelbach's method [23] and dual decomposition [24] is outlined in Section IV. The performance of the EEM algorithm is demonstrated in Section V, which includes results obtained when the EEM and SEM algorithms are employed for characterizing the effect of different system design choices on the achievable SE and EE. Lastly, conclusions are given in Section VI, where future work ideas are also listed.

II. SYSTEM MODEL

Consider an OFDMA DL cellular system relying on a single BS, M fixed RNs and K uniformly-distributed UEs, as shown in Fig. 1³. The cell is divided into M sectors, where each sector is served by one of the fixed RNs. Naturally, the path-loss is a major factor in determining the receiver's signal-to-noise ratios (SNRs) at the UEs, and thus has a substantial effect on the EE. Therefore, in order to minimize the RN-to-UE pathloss, all the UEs in a specific sector are only

³Although it is more realistic to consider a multi-cell system, which would lead to inter-cell interference, our system model assumes that intelligent interference coordination or mitigation techniques are employed such that the level of inter-cell interference is negligible [25].

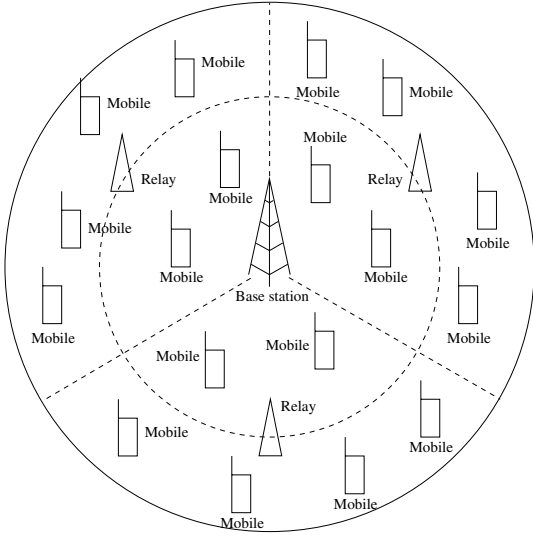


Figure 1: An example of a cellular network with $M = 3$ RNs and $K = 18$ UEs.

supported by that sector's RN, and therefore relay selection is implicitly accomplished. Although the model may be readily extended to include relay selection, for the sake of mathematical tractability, it is not included in this work. The model accounts for both the AF relayed link as well as for the direct link between the BS and UEs, while the variables related to these two communication protocols are distinguishable by the superscripts A and D , respectively. When defining links, the subscript 0 is used for indicating the BS, whilst $\mathcal{M}(k) \in \{1, \dots, M\}$ indicates the RN selected for assisting the DL-transmissions to user k . The proportion of the BS-to-RN distance to the cell radius is denoted by D_r , while the total available instantaneous transmission power of the network is P_{max} . Although it is more realistic to consider a system with separate power constraints for each transmitting entity, for simplicity, a certain total power constraint is considered⁴. The results obtained provide insights into holistic system design by granting a higher grade of freedom in terms of sharing the power among the transmitting entities, and thus attaining a higher performance.

Using the direct transmission protocol, the receiver's SNR at UE k on subcarrier n may be expressed as $\Gamma_k^{D,n}(\mathcal{P})$, whereas when using the AF relaying protocol, the receiver's SNR at UE k on subcarrier n may be expressed as [6]

$$\Gamma_k^{A,n}(\mathcal{P}) = \frac{\gamma_{0,\mathcal{M}(k)}^{A,n} \gamma_{\mathcal{M}(k),k}^{A,n}}{\left(\gamma_{0,\mathcal{M}(k)}^{A,n} + \gamma_{\mathcal{M}(k),k}^{A,n} + 1\right)}, \quad (1)$$

where $\gamma_{a,b}^{X,n} = P_{a,b}^{X,n} G_{a,b}^n / \Delta\gamma N_0 W$ is the SNR at receiver $b \in \{1, \dots, M, 1, \dots, K\}$ on subcarrier $n \in \{1, \dots, N\}$, and $P_{a,b}^{X,n}$ is allocated to transmitter $a \in \{0, \dots, M\}$ using protocol $X \in \{D, A\}$ for transmission to receiver b . Furthermore, $G_{a,b}^n$ represents the channel's attenuation between transmitter a and receiver b on subcarrier n , which is assumed to be known at the BS for all links. The channel's attenuation is modeled by the path-loss and the Rayleigh fading between the transmitter and receiver. Furthermore, N_0 is the additive white

Gaussian noise (AWGN) variance and W is the bandwidth of a single subcarrier. Still referring to (1), $\Delta\gamma$ is the SNR gap at the system's bit error ratio (BER) target between the SNR required at the discrete-input continuous-output memoryless channel (DCMC) capacity and the actual SNR required the modulation and coding schemes of the practical physical layer transceivers employed. For example, making the simplifying assumption that idealized transceivers operating exactly at the DCMC capacity are employed, then $\Delta\gamma = 0$ dB. Although it is not possible to operate exactly at the DCMC channel capacity, several physical layer transceiver designs exist that operate arbitrarily close to it [26]. Additionally, the power allocation policy of the system is denoted by \mathcal{P} , which determines the values of $P_{a,b}^{X,n}$.

Assuming sufficiently high receiver's SNR values, the following approximation can be made

$$\Gamma_k^{A,n}(\mathcal{P}) \approx \frac{P_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n P_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n}{\Delta\gamma N_0 W \left(P_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n + P_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n \right)}, \quad (2)$$

which is valid⁵ for $P_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n + P_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n \gg \Delta\gamma N_0 W$. The SE of an AF link to UE k on subcarrier n is then given by

$$R_k^{A,n}(\mathcal{P}) = \frac{1}{2} \log_2 \left(1 + \Gamma_k^{A,n} \right) \text{ [bits/s/Hz]}, \quad (3)$$

where the factor of $\frac{1}{2}$ accounts for the fact that two time slots are required for the two-hop AF transmission. The SE of a direct link to UE k on subcarrier n is similarly given by

$$R_k^{D,n}(\mathcal{P}) = \log_2 \left(1 + \Gamma_k^{D,n} \right) \text{ [bits/s/Hz]}. \quad (4)$$

The subcarrier indicator variable $s_k^{X,n} \in \{0, 1\}$ is now introduced, which denotes the allocation of subcarrier n for transmission to user k using protocol X for $s_k^{X,n} = 1$, and $s_k^{X,n} = 0$ otherwise. The weighted average SE of the system is calculated as

$$R_T(\mathcal{P}, \mathcal{S}) = \frac{1}{N} \sum_{k=1}^K \omega_k \sum_{n=1}^N s_k^{D,n} \log_2 \left(1 + \Gamma_k^{D,n} \right) + \frac{s_k^{A,n}}{2} \log_2 \left(1 + \Gamma_k^{A,n} \right) \text{ [bits/s/Hz]}, \quad (5)$$

where \mathcal{S} denotes the subcarrier allocation policy of the system, which determines the values of the subcarrier indicator variable $s_k^{X,n}$. The weighting factor ω_k may be varied for ensuring fairness amongst users. However, since ensuring fairness is not the focus of this work, $\omega_k = 1, \forall k$ is assumed then the effect of ω_k may be ignored.

In order to compute the energy used in these transmissions, a model similar to [27] is adopted and the total power consumption of the system is assumed to be governed by a constant term and a term that varies with the transmission powers, which may be written as (6).

Here, $P_C^{(B)}$ and $P_C^{(R)}$ represent the fixed power consumption of each BS and each RN, respectively, while $\xi^{(B)} > 1$ and $\xi^{(R)} > 1$ denote the reciprocal of the drain efficiencies of the power amplifiers employed at the BS and the RNs,

⁴Additionally, it was empirically shown the dual decomposition approach only obtains a local optimum when separate BS and RN power constraints are imposed.

⁵It is plausible that in next-generation systems, through the combination of multi-user and frequency diversity, this assumption holds true when an intelligent scheduler is employed [7].

$$P_T(\mathcal{P}, \mathcal{S}) = \left(P_C^{(B)} + M \cdot P_C^{(R)} \right) + \sum_{k=1}^K \sum_{n=1}^N s_k^{D,n} \xi^{(B)} P_{0,k}^{D,n} + \frac{1}{2} s_k^{A,n} \cdot \left(\xi^{(B)} P_{0,\mathcal{M}(k)}^{A,n} + \xi^{(R)} P_{\mathcal{M}(k),k}^{A,n} \right) \quad [\text{Watts}] \quad (6)$$

respectively. For example, an amplifier having a 25% drain efficiency would have $\xi = \frac{1}{0.25} = 4$.

Finally, the average EE metric of the system is expressed as

$$\eta_E(\mathcal{P}, \mathcal{S}) = \frac{R_T(\mathcal{P}, \mathcal{S})}{P_T(\mathcal{P}, \mathcal{S})} \quad [\text{bits/Joule/Hz}]. \quad (7)$$

III. PROBLEM FORMULATION

The aim of this work is to maximize the energy efficiency metric of (7) subject to a maximum total instantaneous transmit power constraint. In its current form, (7) is dependent on $3KN$ continuous power variables $P_{0,k}^{D,n}$, $P_{0,\mathcal{M}(k)}^{A,n}$ and $P_{\mathcal{M}(k),k}^{A,n}$, $\forall k, n$, and $2KN$ binary subcarrier indicator variables $s_k^{D,n}$ and $s_k^{A,n}$, $\forall k, n$. Thus, it may be regarded as a MINLP problem [22], and can be solved using the branch-and-bound method of [28]. However, the computational effort required for branch-and-bound techniques typically increases exponentially with the problem size. Therefore, a simpler solution is derived by relaxing the binary constraint imposed on the subcarrier indicator variables, $s_k^{D,n}$ and $s_k^{A,n}$, so that they may assume continuous values from the interval $[0, 1]$, as demonstrated in [12], [29]. Furthermore, the variables $\tilde{P}_{0,k}^{D,n} = P_{0,k}^{D,n} s_k^{D,n}$, $\tilde{P}_{0,\mathcal{M}(k)}^{A,n} = P_{0,\mathcal{M}(k)}^{A,n} s_k^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n} = P_{\mathcal{M}(k),k}^{A,n} s_k^{A,n}$ are introduced.

The relaxation of the binary constraints imposed on the variables $s_k^{D,n}$ and $s_k^{A,n}$ allows them to assume continuous values, which leads to a time-sharing subcarrier allocation between the UEs. Naturally, the original problem is not actually solved. However, it has been shown that solving the dual of the relaxed problem provides solutions that are arbitrarily close to the original, non-relaxed problem, provided that the number of available subcarriers tends to infinity [29]. It has empirically been shown that in some cases only 8 subcarriers are required for obtaining close-to-optimal results [30]. It shall be demonstrated in Section V that even for as few as two subcarriers, the solution algorithm employed in this work approaches the optimal EE achieved by an exhaustive search.

The optimization problem is formulated as shown as follows:

Relaxed Problem (P):

$$\text{maximize} \quad \frac{\tilde{R}_T}{\tilde{P}_T} \quad (8)$$

$$\text{subject to} \quad \sum_{k=1}^K \sum_{n=1}^N \tilde{P}_{0,k}^{D,n} + \tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n} \leq P_{max}, \quad (9)$$

$$s_k^{D,n} + s_k^{A,n} \leq 1, \quad \forall k, n, \quad (10)$$

$$\sum_{k=1}^K s_k^{D,n} + s_k^{A,n} \leq 1, \quad \forall n, \quad (11)$$

$$\tilde{P}_{0,k}^{D,n}, \tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \in \mathbb{R}_+, \quad \forall k, n, \quad (12)$$

$$0 \leq s_k^{D,n}, s_k^{A,n} \leq 1, \quad \forall k, n, \quad (13)$$

where the objective function is the ratio between (14) and (15).

In this formulation, the variables to be optimized are $s_k^{D,n}$, $s_k^{A,n}$, $\tilde{P}_{0,k}^{D,n}$, $\tilde{P}_{0,\mathcal{M}(k)}^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$, $\forall k, n$. Physically, the constraint (9) ensures that the sum of the power allocated to variables $\tilde{P}_{0,k}^{D,n}$, $\tilde{P}_{0,\mathcal{M}(k)}^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$ does not exceed the maximum power budget of the system. Constraint (10) ensures that a single transmission protocol, either direct or AF, is chosen for each user-subcarrier pair. The constraint (11) guarantees that each subcarrier is only allocated to at most one user, thus intra-cell interference is avoided. The constraints (12) and (13) describe the feasible region of the optimization variables. The following is a proof that the OF of problem (P) is quasi-concave [31].

A. Proving that the OF in problem (P) is quasi-concave

A function, $f : \mathbf{R}^n \rightarrow \mathbf{R}$, is quasi-concave if its domain is convex, and all its superlevel sets are convex, i.e. if the domain $S_\alpha = \{x \in \text{dom } f \mid f(x) \geq \alpha\}$ is convex for $\alpha \in \mathbf{R}$ [32]. For a fractional function, $g(x)/h(x)$, the inequality $g(x)/h(x) \geq \alpha$ is equivalent to $[g(x) - \alpha h(x)] \geq 0$, assuming $h(x) > 0$, $\forall x$. Therefore, in order to prove that (8) is quasi-concave, it is sufficient to show that the numerator is concave and the denominator is both affine and positive, whilst the domain is convex. It is plausible that the denominator is both affine and positive, since it is the linear combination of multiple nonnegative variables and a positive constant. The proof that the numerator is concave is as follows.

Firstly, the concavity of $f_1 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right) = \frac{\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^{n} \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^{n}}{\Delta \gamma N_0 W \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^{n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^{n} \right)}$ is proven. This may be accomplished by examining the Hessian matrix of $f_1 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right)$ with respect to (w.r.t.) the variables $\tilde{P}_{0,\mathcal{M}(k)}^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$ [32]. The Hessian has the eigenvalues $e_1 = 0$ and

$$e_2 = - \frac{2 \left(G_{0,\mathcal{M}(k)}^{n} G_{\mathcal{M}(k),k}^{n} \right)^2 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right)}{\Delta \gamma N_0 W \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^{n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^{n} \right)^3}, \quad (16)$$

which are non-positive, indicating that the Hessian is negative-semidefinite. This indicates that $f_1 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right)$ is concave w.r.t. the variables $\tilde{P}_{0,\mathcal{M}(k)}^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$.

Examination of the composition $f_2 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right) = \log_2 \left[1 + f_1 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right) \right]$ reveals that $f_2 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right)$ is concave, since $\log_2(\cdot)$ is concave as well as non-decreasing and $1 + f_1 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right)$ is concave [32].

The second term in the summation of (14) may be denoted by (17). This may be obtained using the perspective transfor-

$$\begin{aligned} \tilde{R}_T &= \sum_{k=1}^K \sum_{n=1}^N s_k^{D,n} \log_2 \left(1 + \frac{\tilde{P}_{0,k}^{D,n} G_{0,k}^n}{s_k^{D,n} \Delta \gamma N_0 W} \right) \\ &\quad + \frac{s_k^{A,n}}{2} \log_2 \left(1 + \frac{\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n}{s_k^{A,n} \Delta \gamma N_0 W \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n + \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n \right)} \right) \end{aligned} \quad (14)$$

$$\tilde{P}_T = \left(P_C^{(B)} + M \cdot P_C^{(R)} \right) + \sum_{k=1}^K \sum_{n=1}^N \xi^{(B)} \tilde{P}_{0,k}^{D,n} + \frac{1}{2} \left(\xi^{(B)} \tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \xi^{(R)} \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right) \quad (15)$$

$$f_3 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n}, s_k^{A,n} \right) = s_{k,n}^{A,n} \cdot \log_2 \left(1 + \frac{\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n}{s_k^{A,n} \Delta \gamma N_0 W \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n + \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n \right)} \right). \quad (17)$$

mation⁶ of [32] yielding

$$f_3 \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n}, \tilde{P}_{\mathcal{M}(k),k}^{A,n}, s_k^{A,n} \right) = s_k^{A,n} \cdot f_2 \left(\frac{\tilde{P}_{0,\mathcal{M}(k)}^{A,n}}{s_k^{A,n}}, \frac{\tilde{P}_{\mathcal{M}(k),k}^{A,n}}{s_k^{A,n}} \right), \quad (18)$$

which preserves concavity. Using similar arguments, $s_k^{D,n} \log_2 \left(1 + \frac{\tilde{P}_{0,k}^{D,n} G_{0,k}^n}{s_k^{D,n} \Delta \gamma N_0 W} \right)$ is proven to be concave w.r.t. the variables $s_k^{D,n}$ and $\tilde{P}_{0,k}^{D,n}$.

Finally, the numerator is shown to be concave w.r.t the variables $s_k^{A,n}$, $s_k^{D,n}$, $\tilde{P}_{0,k}^{D,n}$, $\tilde{P}_{0,\mathcal{M}(k)}^{A,n}$, $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$, $\forall k, n$, since it is the non-negative sum of multiple concave functions. Thus, the OF in problem (P) has a numerator that is concave, while its denominator is affine. Hence, the OF of problem (P) is quasi-concave.

B. Problem solution methods

Quasi-concavity may be viewed as a type of generalized concavity, since it can describe discontinuous functions as well as functions that have multiple stationary points. This means that a local maximum is not guaranteed to be a global maximum and thus standard convex optimization techniques, such as interior-point or ellipsoid methods, cannot be readily applied for finding the optimal solution [32]. However, a quasi-concave function has convex superlevel sets, hence the bisection method [18] may be used for iteratively closing the gap between an upper and lower bound solution, until the difference between the two becomes lower than a predefined tolerance. The drawback of this method is that there is no exact method of finding the initial upper as well as lower bounds. Additionally, a convex feasibility problem [32] must be solved in each iteration, which may become computationally undesirable. In light of these discussions, the method detailed in [23] is employed, which allows the quasi-concave problem to be solved as a sequence of parameterized concave programming problems. For clarity, the algorithm is summarized in Fig. 2, which is discussed in the following section.

IV. DINKELBACH'S METHOD FOR SOLVING PROBLEM (P)

A. Introduction to Dinkelbach's method

Dinkelbach's method [23], [31] is an iterative algorithm that can be used for solving a quasi-concave problem in a parameterized concave form. The algorithm is summarized in Table I.

⁶The perspective transformation of the function $f(x)$ is given by $tf(x/t)$.

Table I: Dinkelbach's method for energy efficiency maximization.

Algorithm 1 Dinkelbach's method for energy efficiency maximization	
Input:	I_{outer}^D (maximum number of iterations) $\epsilon_{outer} > 0$ (convergence tolerance)
1:	$q_0 \leftarrow 0$
2:	$i \leftarrow 0$
3:	do while $q_i - q_{i-1} > \epsilon_{outer}^D$ and $i < I_{outer}^D$
4:	$i \leftarrow i + 1$
5:	Solve $\max_{\mathcal{P}, \mathcal{S}} R_T(\mathcal{P}, \mathcal{S}) - q_{i-1} P_T(\mathcal{P}, \mathcal{S})$ to obtain the optimal solution \mathcal{P}_i^* and \mathcal{S}_i^* (inner loop)
6:	$q_i \leftarrow R_T(\mathcal{P}_i^*, \mathcal{S}_i^*) / P_T(\mathcal{P}_i^*, \mathcal{S}_i^*)$
7:	end do
8:	return

The concave form of the fractional program (P) is formed by denoting the OF value as q so that a subtractive form of the OF may be written as $F(q) = R_T(\mathcal{P}, \mathcal{S}) - q P_T(\mathcal{P}, \mathcal{S})$, which is concave. Since the parameter q now acts as a negative weight on the total power consumption of system, it may be intuitively viewed as the 'price' of the system's power consumption. At the optimal OF value of q^* , the following holds true

$$\max_{\mathcal{P}, \mathcal{S}} \{F(q^*)\} = \max_{\mathcal{P}, \mathcal{S}} \{R_T(\mathcal{P}, \mathcal{S}) - q^* P_T(\mathcal{P}, \mathcal{S})\} = 0. \quad (19)$$

Explicitly, the solution of $F(q^*)$ is equivalent to the solution of the fractional problem (P). Dinkelbach [23] proposed an iterative method to find increasing q values, which are feasible, by solving the parameterized problem of $\max_{\mathcal{P}, \mathcal{S}} \{F(q_{i-1})\}$ at each iteration. Hence, it can be shown that the method produces an increasing sequence of q values, which converges to the optimal value at a superlinear convergence rate. As shown in Table I, each outer iteration corresponds to solving $\max_{\mathcal{P}, \mathcal{S}} \{F(q_{i-1})\}$, where q_{i-1} is a given value of the parameter q , to obtain \mathcal{P}_i^* and \mathcal{S}_i^* , which at the optimal power and subcarrier values at the i th iteration of Dinkelbach's method. For further details and a proof of convergence, please refer to [23].

B. Solving the inner loop maximization problem

Dinkelbach's method relies on solving $\max_{\mathcal{P}, \mathcal{S}} F(q_{i-1})$, in each iteration, which will henceforth be referred to as $(P_{q_{i-1}})$. Since it has been shown that $R_T(\mathcal{P}, \mathcal{S})$ is concave whilst $P_T(\mathcal{P}, \mathcal{S})$ is affine, then $(P_{q_{i-1}})$ is concave w.r.t. the variables \mathcal{P} and \mathcal{S} . Assuming the existence of an interior point (Slater's condition), there is a zero duality gap between the dual problem of $(P_{q_{i-1}})$ and the primal problem of (P) [32]. Thus

solving the dual problem of $(P_{q_{i-1}})$ is equivalent to solving the primal problem of (P).

The Lagrangian of $(P_{q_{i-1}})$ is given by (20), where $\lambda \geq 0$ is the Lagrangian multiplier associated with the constraint (9). The feasible region constraints (12) and (13), and constraints (10) and (11) will be considered when deriving the optimal solution, which is detailed later.

The dual problem of $(P_{q_{i-1}})$ may be written as [24]

$$\min_{\lambda \geq 0} g(\lambda) = \min_{\lambda \geq 0} \max_{\mathcal{P}, \mathcal{S}} \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda), \quad (21)$$

which is solved using the dual decomposition approach [24]. Using dual decomposition, (21) may be readily solved via solving NK similar subproblems to obtain both the power as well as subcarrier allocations, and by solving a master problem to update λ . The dual decomposition approach is outlined in the following.

1) *Solving the subproblem of power and subcarrier allocation:* For a fixed λ and q_{i-1} , $\max_{\mathcal{P}, \mathcal{S}} \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)$ is solved to obtain the corresponding power and subcarrier allocations. Since the problem is now in a standard concave form, the Karush–Kuhn–Tucker (KKT) conditions [32], which are first-order necessary and sufficient conditions for optimality, may be used in order to find the optimal solution. All optimal variables are denoted by a superscript asterisk. The total transmit power assigned for AF transmission to user k over subcarrier n is now denoted by $\tilde{P}_k^{A,n} = \tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n}$. Then, by substituting $\tilde{P}_{\mathcal{M}(k),k}^{A,n} = \tilde{P}_k^{A,n} - \tilde{P}_{0,\mathcal{M}(k)}^{A,n}$ into (20), the following first-order derivatives may be obtained

$$\left. \frac{\partial \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)}{\partial \tilde{P}_{0,k}^{D,n}} \right|_{\tilde{P}_{0,k}^{D,n} = \tilde{P}_{0,k}^{D,n*}} = 0, \quad (22)$$

$$\left. \frac{\partial \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)}{\partial \tilde{P}_k^{A,n}} \right|_{\tilde{P}_k^{A,n} = \tilde{P}_k^{A,n*}} = 0 \quad (23)$$

and

$$\left. \frac{\partial \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)}{\partial \tilde{P}_{0,\mathcal{M}(k)}^{A,n}} \right|_{\tilde{P}_{0,\mathcal{M}(k)}^{A,n} = \tilde{P}_{0,\mathcal{M}(k)}^{A,n*}} = 0. \quad (24)$$

The optimal values of $\tilde{P}_{0,k}^{D,n}$ may be readily obtained from (22) as

$$P_{0,k}^{D,n*} = \left[\frac{1}{\ln 2 (q_{i-1} \xi^{(B)} + \lambda)} - \frac{1}{\alpha_k^{D,n}} \right]^+, \quad (25)$$

where the effective channel gain of the direct transmission is given by

$$\alpha_k^{D,n} = \frac{G_{0,k}^n}{\Delta \gamma N_0 W} \quad (26)$$

and $[\cdot]^+$ denotes $\max(0, \cdot)$ since the powers allocated have to be nonnegative due to the constraint (12). Similarly the optimal values of $\tilde{P}_{0,\mathcal{M}(k)}^{A,n}$ and $\tilde{P}_{\mathcal{M}(k),k}^{A,n}$ may be obtained by equating (23) and (24) to give

$$P_{0,\mathcal{M}(k)}^{A,n*} = \beta_k^{A,n} P_k^{A,n*} \quad (27)$$

and

$$P_{\mathcal{M}(k),k}^{A,n*} = \left(1 - \beta_k^{A,n}\right) P_k^{A,n*}, \quad (28)$$

where the total transmit power assigned for the AF transmission to user k over subcarrier n is given by (29), (30) and (31). Observe that (31) is the fraction of the total AF

transmit power that is allocated for the BS-to-RN link while obeying $0 \leq \beta_k^{A,n} \leq 1$.

Having calculated the optimal power allocations, the optimal subcarrier allocations may be derived using the first-order derivatives as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)}{\partial s_k^{D,n}} &= \log_2 \left(1 + \alpha_k^{D,n} P_{0,k}^{D,n*}\right) \\ &\quad - \frac{\alpha_k^{D,n} P_{0,k}^{D,n*}}{\ln 2 \left(1 + \alpha_k^{D,n} P_{0,k}^{D,n*}\right)} \\ &= D_k^n \begin{cases} < 0 & \text{if } s_k^{D,n*} = 0, \\ = 0 & \text{if } s_k^{D,n*} \in (0, 1), \\ > 0 & \text{if } s_k^{D,n*} = 1 \end{cases} \end{aligned} \quad (32)$$

and

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)}{\partial s_k^{A,n}} &= \frac{1}{2} \log_2 \left(1 + \alpha_k^{A,n} \tilde{P}_k^{A,n*}\right) \\ &\quad - \frac{\alpha_k^{A,n} \tilde{P}_k^{A,n*}}{2 \ln 2 \left(1 + \alpha_k^{A,n} \tilde{P}_k^{A,n*}\right)} \\ &= A_k^n \begin{cases} < 0 & \text{if } s_k^{A,n*} = 0, \\ = 0 & \text{if } s_k^{A,n*} \in (0, 1), \\ > 0 & \text{if } s_k^{A,n*} = 1. \end{cases} \end{aligned} \quad (34)$$

(32) and (34) stem from the fact that if the optimal value of $s_k^{X,n}$ occurs at the boundary of the feasible region, then $\mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)$ must be decreasing with the values of $s_k^{X,n}$ that approach the interior of the feasible region. By contrast, for example, the derivative $D_k^n = 0$ if the optimal $s_k^{D,n}$ is obtained in the interior of the feasible region [12]. However, since each subcarrier may only be used for transmission to a single user, each subcarrier n is allocated to the specific user k having the highest value of $\max(A_k^n, D_k^n)$ in order to achieve the highest increase in $\mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)$. The optimal allocation for subcarrier n is as follows⁷

$$s_k^{D,n*} = \begin{cases} 1, & \text{if } D_k^n = \max_j [\max(A_j^n, D_j^n)] \text{ and } D_k^n \geq 0, \\ 0, & \text{otherwise,} \end{cases} \quad (35)$$

and

$$s_k^{A,n*} = \begin{cases} 1, & \text{if } A_k^n = \max_j [\max(A_j^n, D_j^n)] \text{ and } A_k^n \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (36)$$

Thus constraints (10)- (13) are satisfied and the optimal primal variables are obtained for a given λ . Observe that the optimal power allocations given by (25) and (29) are indeed customized water-filling solutions [33], where the effective channel gains are given by $\alpha_k^{D,n}$ and $\alpha_k^{A,n}$, respectively, and where the water levels are determined both by the cost of allocating power, λ , as well as the current cost of power to the EE given by q_{i-1} .

2) *Updating the dual variable λ :*

Since (25), (27), (28), (35) and (36) give a unique solution for $\max_{\mathcal{P}, \mathcal{S}} \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)$, it follows that $g(\lambda)$ is differentiable and hence the gradient method [32], [24] may be readily used for

⁷If there are multiple users that tie for the value of $\max(A_k^n, D_k^n)$, then a random user from the maximal set is chosen.

$$\begin{aligned}
\mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda) &= \sum_{k=1}^K \sum_{n=1}^N s_k^{D,n} \log_2 \left(1 + \frac{\tilde{P}_{0,k}^{D,n} G_{0,k}^n}{s_k^{D,n} \Delta \gamma N_0 W} \right) \\
&+ \frac{s_k^{A,n}}{2} \log_2 \left(1 + \frac{\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n}{s_k^{A,n} \Delta \gamma N_0 W \left(\tilde{P}_{0,\mathcal{M}(k)}^{A,n} G_{0,\mathcal{M}(k)}^n + \tilde{P}_{\mathcal{M}(k),k}^{A,n} G_{\mathcal{M}(k),k}^n \right)} \right) \\
&- q_{i-1} \left[\left(P_C^{(B)} + M \cdot P_C^{(R)} \right) + \sum_{k=1}^K \sum_{n=1}^N \xi^{(B)} \tilde{P}_{0,k}^{D,n} + \frac{1}{2} \left(\xi^{(B)} \tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \xi^{(R)} \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right) \right] \\
&+ \lambda \left(P_{max} - \sum_{k=1}^K \sum_{n=1}^N \tilde{P}_{0,k}^{D,n} + \tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right). \tag{20}
\end{aligned}$$

$$P_k^{A,n*} = \left[\frac{1}{\ln 2 \left(\beta_k^{A,n} (q_{i-1} \xi^{(B)} + 2\lambda) + (1 - \beta_k^{A,n}) (q_{i-1} \xi^{(R)} + 2\lambda) \right)} - \frac{1}{\alpha_k^{A,n}} \right]^+ \tag{29}$$

$$\alpha_k^{A,n} = \frac{\beta_k^{A,n} (1 - \beta_k^{A,n}) G_{0,\mathcal{M}(k)}^n G_{\mathcal{M}(k),k}^n}{\left(\beta_k^{A,n} G_{0,\mathcal{M}(k)}^n + (1 - \beta_k^{A,n}) G_{\mathcal{M}(k),k}^n \right) \Delta \gamma N_0 W} \tag{30}$$

$$\beta_k^{A,n} = \frac{-G_{\mathcal{M}(k),k}^n (q_{i-1} \xi^{(R)} + 2\lambda) + \sqrt{G_{0,\mathcal{M}(k)}^n G_{\mathcal{M}(k),k}^n (q_{i-1} \xi^{(B)} + 2\lambda) (q_{i-1} \xi^{(R)} + 2\lambda)}}{G_{0,\mathcal{M}(k)}^n (q_{i-1} \xi^{(B)} + 2\lambda) - G_{\mathcal{M}(k),k}^n (q_{i-1} \xi^{(R)} + 2\lambda)} \tag{31}$$

updating the dual variables λ . The gradient of λ is given by

$$\frac{\partial \mathcal{L}(\mathcal{P}, \mathcal{S}, \lambda)}{\partial \lambda} = P_{max} - \sum_{k=1}^K \sum_{n=1}^N \left(\tilde{P}_{0,k}^{D,n} + \tilde{P}_{0,\mathcal{M}(k)}^{A,n} + \tilde{P}_{\mathcal{M}(k),k}^{A,n} \right). \tag{37}$$

Therefore, λ may be updated using the optimal variables to give (38), where $\alpha_\lambda(i)$ is the size of the step taken in the direction of the negative gradient for the dual variable λ at iteration i . For the performance investigations of Section V, a constant step size is used, since it is comparatively easier to find a value that strikes a balance between optimality and convergence speed. The process of computing the optimal power as well as subcarrier allocations and subsequently updating λ is repeated until convergence is attained, indicating that the dual optimal point has been reached. Since the primal problem ($P_{q_{i-1}}$) is concave, there is zero duality gap between the dual and primal solutions. Hence, solving the dual problem is equivalent to solving the primal problem. The inner loop solution method is summarized in Table II.

C. Summary of solution methodology

Again, for additional clarity, the solution methodology is summarized in Fig. 2. Firstly, the relaxed problem (P) expressed in a fractional form is rewritten as a subtractive, parameterized concave form, where q_i is the parameter. Solving this problem for a given q_i is termed an outer iteration, which is described in Section IV-A and is illustrated in the upper block of Fig. 2. The lower block of Fig. 2 illustrates the dual decomposition approach that is employed for solving the subtractive, concave problem. Each iteration of the dual decomposition approach is termed an inner iteration, which is further detailed in Section IV-B. In each inner iteration, $2NK$ subproblems are solved to obtain the optimal power

Table II: Inner loop solution method for obtaining the optimal power and subcarrier allocations for a given q_{i-1} .

Algorithm 2 Inner loop solution method for obtaining the optimal power and subcarrier allocations for a given q_{i-1} .	
Input:	I_{inner}^D (maximum number of iterations)
	$\epsilon_{inner}^D > 0$ (convergence tolerance)
	end do
1:	$i \leftarrow 0$
2:	do while $ \lambda(i) - \lambda(i-1) > \epsilon_{inner}^D$ and $i < I_{inner}^D$
3:	$i \leftarrow i + 1$
4:	for n from 1 to N
5:	for k from 1 to K
6:	Obtain the optimal power allocation using (25), (27) and (28)
7:	end for
8:	Obtain the optimal subcarrier allocation using (35) and (36)
9:	end for
10:	Update the dual variables λ using (38)
11:	end do
12:	return

and subcarrier variables for a given q_i and for the dual variable λ . The dual variable λ is then updated, depending on the power and subcarrier variables obtained. Multiple inner iterations are completed until convergence of the optimal dual and primal solutions is reached. The optimal \mathcal{P}^* and \mathcal{S}^* are then fed back into the upper block 2 to evaluate the updated value of q_i , which is used in the next outer iteration. Several outer iterations are completed until convergence to the optimal q_i is obtained. The corresponding optimal \mathcal{P}^* and \mathcal{S}^* values are the power and subcarrier allocation variables that solve the problem (P). The algorithmic complexity of this method is dominated by the comparison operations given by (35) and (36), which leads to a total complexity of $\mathcal{O}(I_{dual} \times 2NK)$ when NK is large, where I_{dual} is the total number of inner iterations required for reaching convergence in Dinkelbach's method.

$$\lambda(i+1) = \left[\lambda(i) - \alpha_{\lambda}(i) \left(P_{max} - \sum_{k=1}^K \sum_{n=1}^N \tilde{P}_{0,k}^{D,n*} + \tilde{P}_{0,\mathcal{M}(k)}^{A,n*} + \tilde{P}_{\mathcal{M}(k),k}^{A,n*} \right) \right]^+ \quad (38)$$

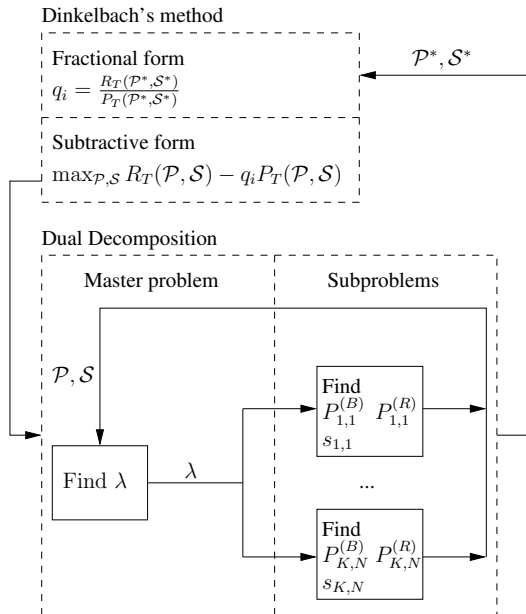


Figure 2: Summary of the solution methodology for the relaxed problem (P).

Table III: Simulation parameters used to obtain all results in this section unless otherwise specified.

Simulation parameter	Value
Subcarrier bandwidth, W Hertz	12k
Number of RNs, M	{0, 1, 2, 3, 5, 6}
Number of subcarriers, N	{128, 512, 1024}
Number of UEs, K	{30, 60, 120}
Cell radius, km	{0.75, 1, 1.25, 1.5, 1.75, 2}
Ratio of BS-to-RN distance to the cell radius, D_r	{0.1, 0.3, 0.5, 0.7, 0.9}
SNR gap of wireless transceivers, $\Delta\gamma$ dB	0
Maximum total transmission power, P_{max} dBm	{0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60}
Fixed power consumption of the BS, $P_C^{(B)}$ Watts [27]	60
Fixed power consumption of RNs, $P_C^{(R)}$ Watts [27]	20
Reciprocal of the BS power amplifier's drain efficiency, $\xi^{(B)}$ [27]	2.6
Reciprocal of the RNs' power amplifier's drain efficiency, $\xi^{(R)}$ [27]	5
Noise power spectral density, N_0 dBm/Hz	-174
Maximum number of outer iterations in Dinkelbach's algorithm, I_{outer}^D	10
Maximum number of inner iterations in Dinkelbach's algorithm, I_{inner}^D	100
Convergence tolerance of iterative algorithms, $\epsilon_{outer}^D = \epsilon_{inner}^D$	10^{-8}
Number of channel samples	10^4

V. RESULTS AND DISCUSSIONS

This section presents the results of applying the EEM algorithm described in Section IV to the relay-aided cellular system shown in Fig. 1. Again, the channel is modeled by the path-loss [34] and uncorrelated Rayleigh fading obeying the complex normal distribution, $\mathcal{CN}(0, 1)$. It is assumed that the BS-to-RN link has line-of-sight (LOS) propagation, implying

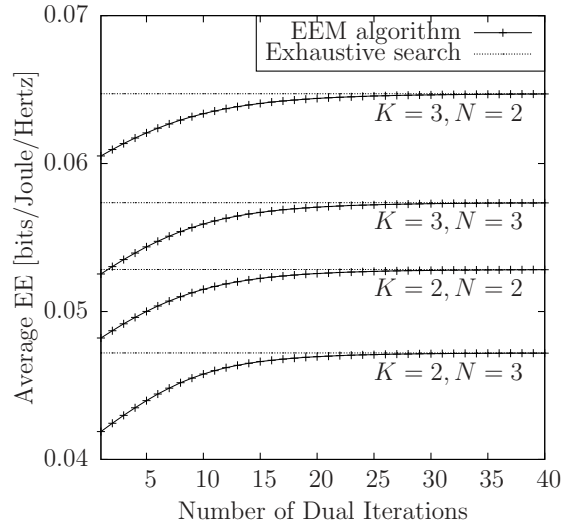


Figure 3: Average EE versus the total number of inner iterations of Dinkelbach's method required for reaching convergence when using the simulation parameters from Table III with $P_{max} = 0$ dBm, $D_r = 0.5$, $M = 0$ and with a cell radius of 1km.

that a RN was placed on a tall building. However, the BS-to-UE and RN-to-UE links typically have no LOS, since these links are likely to be blocked by buildings and other large obstructing objects. The RNs are evenly distributed at a fixed distance around the central BS and the UEs are uniformly distributed within the cell. An independently-random set of UE locations as well as fading channel realizations are generated for each channel sample. For fair comparisons, the metrics used are the average SE per subcarrier and the average EE per subcarrier. On the other hand, the sum-rate may be calculated by multiplying the average SE by NW . Additionally, ρ is introduced to denote the average fraction of the total number of subcarriers that are used for AF transmission. Thus, ρ quantifies the benefit attained from introducing RNs into the system. The main simulation parameters are given in Table III^{8,9}.

A. Convergence of iterative algorithms to optimal value

Fig. 3 illustrates the convergence behavior of Dinkelbach's method invoked for maximizing the EE for a selection of small-scale systems, averaged over 10^4 different channel realizations. Since the problem size is small, it is possible to generate also the exhaustive-search based solution within a reasonable computation time. As seen in Fig. 3, Dinkelbach's method converges to the optimal value within forty inner iterations. This result demonstrates that the EEM algorithm based on Dinkelbach's method indeed obtains the optimal

⁸For simplification, it is assumed that near-capacity transceivers are employed in the network.

⁹In all cases, the step size and the initial value of λ was empirically optimized to give the optimal objective value in as few iterations as possible, although the exact analytical method for achieving this still remains an open issue.

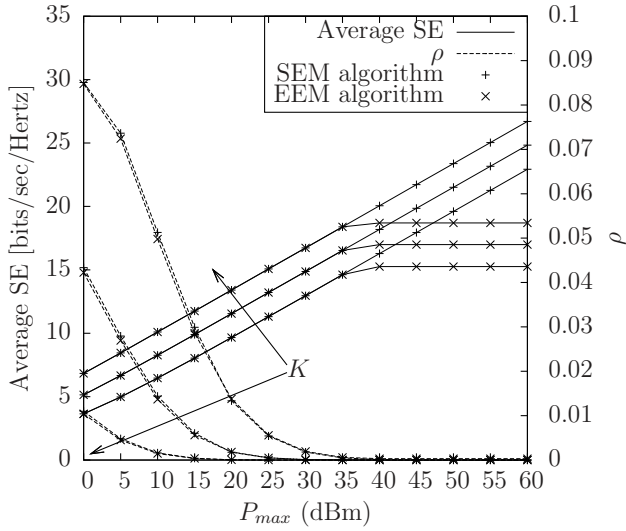
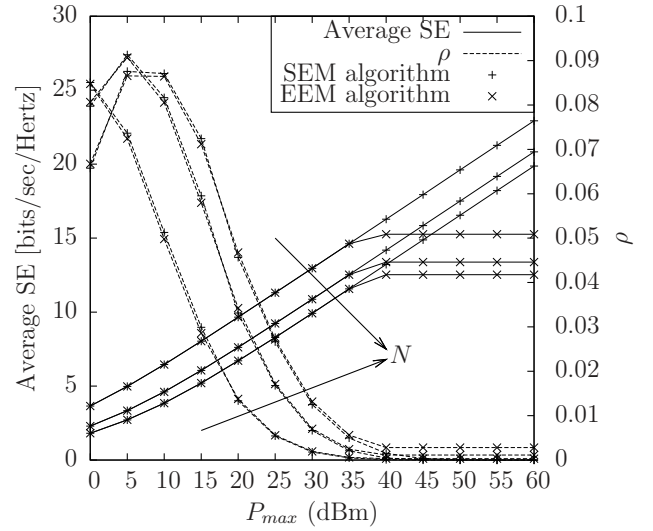
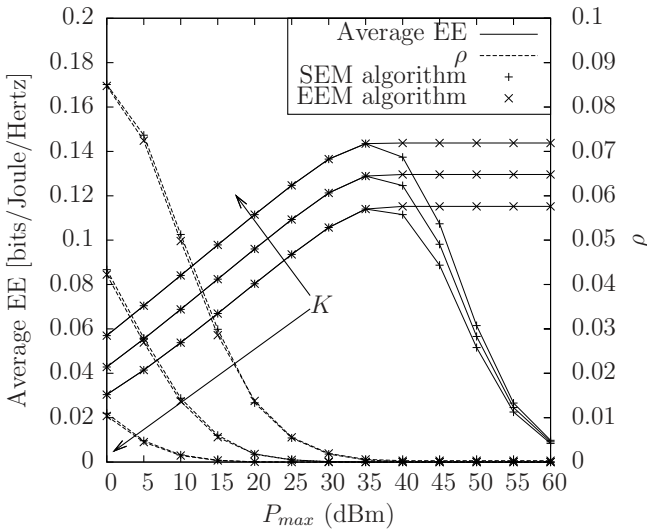
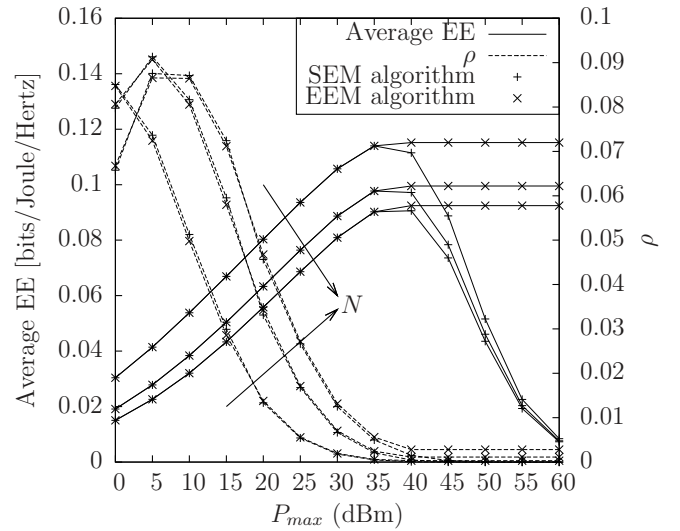
(a) Average SE and ρ versus P_{max} for $K = 30, 60$ and 120 .(a) Average SE and ρ versus P_{max} for $N = 128, 512$ and 1024 .(b) Average EE and ρ versus P_{max} for $K = 30, 60$ and 120 .(b) Average EE and ρ versus P_{max} for $N = 128, 512$ and 1024 .

Figure 4: Average SE, EE and ρ , and the effect of an increasing number of users, K , for a system with simulation parameters from Table III with $N = 128$, $M = 3$, $D_r = 0.5$ and with a cell radius of 1.5km.

Figure 5: Average SE, EE and ρ , and the effect of an increasing number of subcarriers, N , for a system with simulation parameters from Table III with $K = 30$, $M = 3$, $D_r = 0.5$ and with a cell radius of 1.5km.

power and subcarrier allocation, even though the relaxed problem is solved and a high receiver's SNR was assumed.

B. Effect of the number of UEs on the attainable SE and EE

Additionally, the EEM algorithm may be employed for evaluating the effects of system-level design choices on the network's SE and EE. The effect of K on the average EE and SE¹⁰ is depicted in Fig. 4. As expected, upon increasing K , the multi-user diversity of the system is increased, since the scheduler is allowed to choose its subcarrier allocations from a larger pool of channel gains. This results in an increase of both the maximum EE as well as of the SE attained. Furthermore, Fig. 4 shows that as P_{max} is increased, the SEM algorithm continues to allocate more power in order to achieve a higher average SE at the cost of EE, while the EEM algorithm attains the maximum EE and does not continue to increase

¹⁰N.B. The maximum SE is obtained in the first outer iteration of Dinkelbach's method with $q_0 = 0$, since this equates to zero penalty for any power consumption.

its attainable SE by sacrificing the achieved EE. On the other hand, ρ is inversely proportional to K . This indicates that as the multi-user diversity increases, the subcarriers are less likely to be allocated for AF transmissions, simply because there are more favorable BS-to-UE channels owing to having more UEs nearer to the cell-center. Moreover, the value of ρ decreases as P_{max} increases, because there is more power to allocate to the BS-to-UE links for UEs near the cell-center, which benefit from a reduced pathloss as well as from a more efficient power amplifier at the BS.

C. Effect of the number of subcarriers on the attainable SE and EE

Fig. 5 illustrates the effect of increasing N on the attainable SE and EE. Fig. 5 shows much of the same trends as Fig. 4. For example, the SE continues to rise at the cost of EE, as P_{max} increases when using the SEM algorithm, while the EEM algorithm attains the maximum EE and the corresponding SE. However, in Fig. 5 both the SE and EE,

averaged over N , decreases upon increasing N , which implies that the subcarriers are not used effectively, when more are available. On the other hand, it may be observed that the sum-rate achieved using the SEM algorithm increases with N , as expected due to frequency diversity.

In Fig. 5, it is noticeable that ρ increases upon increasing N , which is in contrast to the trend observed in Fig. 4. This may be understood by considering the UEs within the network. Since the UEs positions are fixed, as N increases the scheduler has access to a larger pool of channel gains for each individual UE, thus more cell-edge users may be supported for maximizing the SE or EE. However, increasing K does not have this effect, since a larger K value indicates that there are more UEs near the cell-center, and since both the SEM and EEM algorithms are greedy, these cell-center UEs are served before the cell-edge UEs, hence ρ decreases. The reduction of ρ when P_{max} is very small suggests that the total available power in the system is not high enough to take advantage of the AF transmissions.

D. Effect of the cell radius on the attainable SE and EE

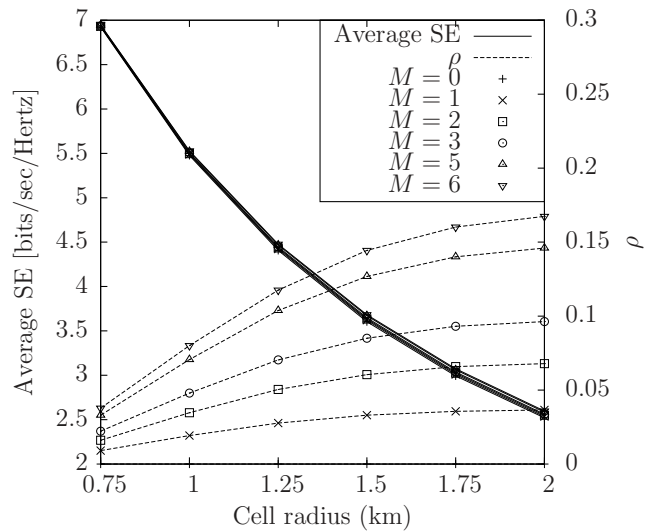
The effect of increasing the cell radius on the attainable SE and EE is characterized in Fig. 6. As expected, increasing the cell radius has a detrimental effect on both the SE and EE of the system regardless of the number of RNs employed owing to the increased pathlosses experienced. Additionally, it is noteworthy that ρ increases as the cell radius increases, indicating that relaying is more beneficial for larger cells. On the other hand, the increase in the SE obtained from employing RNs in a large cell is small. For example, the SE is improved by a factor of 1.03 when $M = 6$ RNs are used instead of $M = 0$ at a cell radius of 2km. This improvement is modest when compared to the reduction in EE of a factor of 0.34 due to having to support additional transmitting entities. This phenomenon stems from the fact that, since the UEs are uniformly distributed across the cell, the UEs nearer the cell-center are more likely to be allocated resources for maximizing the SE or EE as they may benefit from the more-favorable direct transmission. Thus, increasing the number of RNs in the system brings a marginal benefit in terms of SE or EE.

E. Effect of the relay's position on the attainable SE and EE

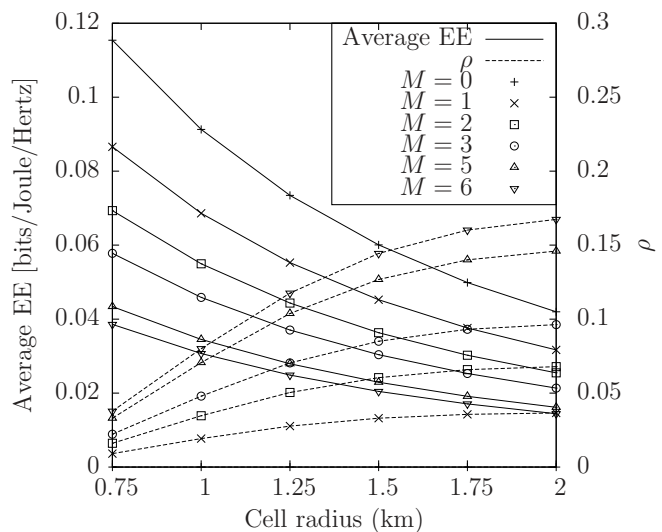
The effect of the RNs' position relative to the BS and the cell-edge is illustrated in Fig. 7, which clearly shows that the optimal SE and EE is obtained, when the RN is closer to the BS than to the UEs. This stems from the fact that the RN benefits from having a stronger LOS link to the BS, when it is placed closer to the BS, thus strengthening the AF links. However, it cannot be placed too close to the BS, since the benefits gleaned from having a stronger BS-to-RN link are then outweighed by having a more hostile RN-to-UE link.

VI. CONCLUSIONS

In this paper, the joint power and subcarrier allocation problem was formulated for maximizing the EE in a multi-relay aided multi-user OFDMA cellular network. The OF was proven to be quasi-concave and an iterative method, namely Dinkelbach's method, was employed for solving the associated optimization. Dinkelbach's method solves the optimization problem by solving a sequence of subtractive concave



(a) Average SE and ρ versus P_{max} for $M = 0, 1, 2, 3, 5$ and 6 .



(b) Average EE and ρ versus P_{max} for $M = 0, 1, 2, 3, 5$ and 6 .

Figure 6: Average SE, EE and ρ , and the effect of an increasing cell radius for a system with simulation parameters from Table III with $K = 30$, $N = 128$, $M \in \{0, 1, 2, 3, 5, 6\}$, $D_r = 0.5$ and with a $P_{max} = 0$ dBm.

problems, which were solved using the dual decomposition approach in this paper. The optimal power and subcarrier allocation were presented for solving each iteration of the dual decomposition algorithm, and simulations were performed to validate the algorithm.

Further simulation results show that when there is insufficient power for attaining the maximum achievable EE, both the EEM and the SEM algorithms have the same solution. As the system's total power is increased, the SEM algorithm continues to allocate more power in order to achieve ever higher values of SE, whereas the EEM algorithm reaches an upper bound and does not make use of the additional available power. Additionally, a comprehensive study of the effect of various system parameters on the achievable SE and EE is performed. To summarize, the achievable SE and EE is increased, when there is a larger number of UEs in the system owing to achieving a higher multi-user diversity. Increasing the number of available subcarriers, although increases the SR owing to frequency diversity, reduces the average SE

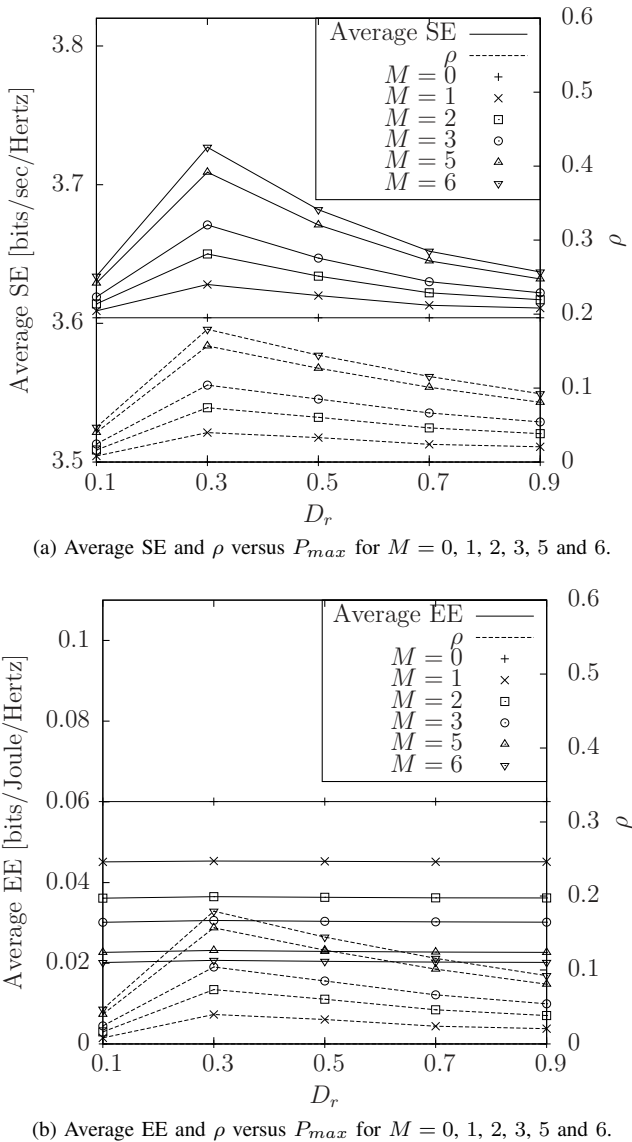


Figure 7: Average SE, EE and ρ , and the effect of an increasing D_r , with simulation parameters from Table III with $K = 30$, $N = 128$, $M \in \{0, 1, 2, 3, 5, 6\}$, $P_{max} = 0$ dBm and with a cell radius of 1.5 km.

since not all subcarriers are effectively utilized. The benefit of introducing RNs into the network for improving the achievable SE becomes more significant as the cell-size increases or the number of relays increases. However, the EE is then degraded due to the increased overhead power consumption. Furthermore, relaying is more beneficial, when the RNs are placed closer to the BS, if there exists a LOS link between the RNs and BS.

As a next step, EEM and SEM algorithms will be invoked for multi-cell systems, which are interference-limited, rather than noise-limited. This is in contrast to this work, which stipulated the idealized simplifying assumption that the inter-cell interference is sufficiently low, and thus may be ignored. Furthermore, on-line near-real-time optimization for mobile RNs may be considered.

REFERENCES

- [1] L. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. Van Der Perre, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66–72, Nov. 2010.
- [2] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. Grant, H. Haas, J. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. Nakhai, J. Zhang, and L. Hanzo, "Green radio: radio techniques to enable energy-efficient wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, Jun. 2011.
- [3] L. Collins, "Greening the global network," *Engineering Technology*, vol. 5, no. 4, pp. 64–65, Mar. 2010.
- [4] "EARTH (Energy Aware Radio and neTwork tecHnologies), EU Funded Research Project FP7-ICT-2009-4- 247733-EARTH," 2010–2012. [Online]. Available: <https://www.ict-earth.eu>
- [5] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMAX and LTE-advanced mobile systems," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 100–105, Oct. 2009.
- [6] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [7] O. Oyman, "Opportunistic scheduling and spectrum reuse in relay-based cellular networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 1074–1085, Mar. 2010.
- [8] L. Hanzo, Y. Akhtman, L. Wang, and M. Jiang, *MIMO-OFDM for LTE, WIFI and WIMAX: Coherent Versus Non-Coherent and Cooperative Turbo-Transceivers*. Wiley-IEEE Press, 2010.
- [9] G. Li and H. Liu, "Resource allocation for OFDMA relay networks with fairness constraints," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 11, pp. 2061–2069, Nov. 2006.
- [10] T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 328–339, Feb. 2007.
- [11] D. Ng and R. Schober, "Cross-layer scheduling for OFDMA amplify-and-forward relay networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1443–1458, Mar. 2010.
- [12] C. Y. Wong, R. Cheng, K. Letaief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [13] L. Piazzo, "Fast algorithm for power and bit allocation in OFDM systems," *Electronics Letters*, vol. 35, no. 25, pp. 2173–2174, Dec. 1999.
- [14] L. Xiao and L. Cuthbert, "Adaptive power allocation scheme for energy efficient OFDMA relay networks," in *Proceedings of the IEEE International Conference on Communication Systems (ICCS'08)*, Guangzhou, China, Nov. 2008, pp. 637–641.
- [15] L. Piazzo, "Optimal fast algorithm for power and bit allocation in OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1263–1265, Mar. 2011.
- [16] J. Joung and S. Sun, "Power efficient resource allocation for downlink OFDMA relay cellular networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2447–2459, May 2012.
- [17] R. Devarajan, S. Jha, U. Phuyal, and V. Bhargava, "Energy-aware resource allocation for cooperative cellular network using multi-objective optimization approach," *IEEE Transactions on Wireless Communications*, vol. 11, no. 5, pp. 1797–1807, May 2012.
- [18] H. Yu, R. Xiao, Y. Li, and J. Wang, "Energy-efficient multi-user relay networks," in *Proceedings of the International Conference on Wireless Communications and Signal Processing (WCSP'11)*, Nanjing, China, Nov. 2011.
- [19] G. Miao, N. Himayat, and G. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 545–554, Feb. 2010.
- [20] G. Miao, N. Himayat, G. Li, and S. Talwar, "Low-complexity energy-efficient scheduling for uplink OFDMA," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 112–120, Jan. 2012.
- [21] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation for secure OFDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 6, pp. 2572–2585, Jul. 2012.
- [22] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, Belmont, MA, USA, 1999.
- [23] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, pp. 492–498, Mar. 1967.
- [24] D. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.
- [25] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
- [26] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation*, 1st ed. Wiley-IEEE Press, 2009.
- [27] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular

networks,” in *Proceedings of the Future Network and Mobile Summit*, Florence, Italy, Jun. 2010.

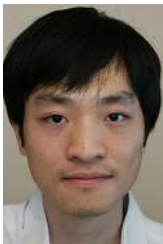
- [28] S. Boyd and J. Mattingley, “Branch and bound methods,” EE364b course notes, Stanford University, Mar. 2007.
- [29] W. Yu and R. Lui, “Dual methods for nonconvex spectrum optimization of multicarrier systems,” *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
- [30] K. Seong, M. Mohseni, and J. Cioffi, “Optimal resource allocation for OFDMA downlink systems,” in *Proceedings of the IEEE International Symposium on Information Theory (ISIT'06)*, Seattle, Washington, USA, Jul. 2006, pp. 1394–1398.
- [31] M. Avriel, W. E. Diewert, S. Schaible, and I. Zang, *Generalized Convexity*. Plenum Press, New York, NY, USA, 1988.
- [32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.
- [33] A. Goldsmith, *Wireless Communications*. Cambridge University Press, New York, NY, USA, 2005.
- [34] 3GPP, “TR 36.814 V9.0.0: further advancements for E-UTRA, physical layer aspects (release 9),” Mar. 2010.



Lajos Hanzo (<http://www.cspc.ecs.soton.ac.uk>) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded the honorary doctorate *Doctor Honoris Causa* by the Technical University of Budapest.

During his 35-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 80 PhD students, co-authored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10000 pages, published 1300+ research entries at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European IST Programme and the Mobile Virtual Centre of Excellence (VCE), UK.

He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008-2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. His research is funded by the European Research Council's Senior Research Fellow Grant. For further information on research in progress and associated publications please refer to <http://www.cspc.ecs.soton.ac.uk> Lajos has 17000+ citations.



Kent Tsz Kan Cheung (S'09) received his B.Eng. degree (first-class honors) in electronic engineering from the University of Southampton, Southampton, U.K., in 2009. Since then he has been working towards a Ph.D. degree in wireless communications at the same institution. He was a recipient of the EPSRC Industrial CASE award in 2009, and was involved with the Core 5 Green Radio project of the Virtual Centre of Excellence in Mobile and Personal Communications (Mobile VCE).

His research interests include energy-efficiency, multi-carrier MIMO communications, cooperative communications, resource allocation and optimization.



Shaoshi Yang (S'09) received the B.Eng. Degree in information engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2006. He is currently working toward the Ph.D. degree in wireless communications with the School of Electronics and Computer Science, University of Southampton, Southampton, U.K., through scholarships from both the University of Southampton and the China Scholarship Council.

From November 2008 to February 2009, he was an Intern Research Fellow with the Communications Technology Laboratory, Intel Labs China, Beijing, where he focused on Channel Quality Indicator Channel design for mobile WiMAX (802.16m). His research interests include multiuser detection/multiple-input multiple-output detection, multicell joint/distributed processing, cooperative communications, green radio, and interference management. He has published in excess of 20 research papers on IEEE journals and conferences.

Shaoshi is a recipient of the PMC-Sierra Telecommunications Technology Scholarship, and a Junior Member of the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK. He is also a TPC member of both the 23rd Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE PIMRC 2012), and of the 48th Annual IEEE International Conference on Communications (IEEE ICC 2013).