

Energy-Efficient Resource Allocation for OFDMA Heterogeneous Networks

Nam-Tran Le, Le-Nam Tran, *Senior Member, IEEE*, Quang-Doanh Vu, *Member, IEEE*,
Dhammika Jayalath, *Senior Member, IEEE*

Abstract—We proposed several energy-efficient resource allocation algorithms for the downlink of an orthogonal frequency-division-multiple-access (OFDMA) based femtocell heterogeneous networks (HetNets). Heterogeneous QoS and fairness in rate are investigated in the proposed resource allocation problem. A dense deployment of femtocells in the coverage area of a central macrocell is considered and energy usage of both femtocell and macrocell users are optimized simultaneously. We aim to maximize the weighted sum of the individual energy efficiencies (WSEEMax) and the network energy efficiency (NEEMax) while satisfying the following: (1) minimum throughput for delay-sensitive (DS) users, (2) fairness constraint for delay-tolerant (DT) users, (3) required constraints of OFDMA systems. The problem is formulated in three different forms: mixed 0-1 integer programming formulation, time-sharing formulation and sparsity-inducing formulation. The proposed resource block (RB) and power optimization problems are combinatorial and highly non-convex due to the fractional form of the objective function, the integer constraint of OFDMA RBs and non-affine fairness. We adopt the successive convex approximation (SCA) approach and transform the problems into a sequence of convex subproblems. With the proposed algorithms, we show that the overall joint RB and power allocation schemes converge to suboptimal solutions. Numerical examples confirm the merits of the proposed algorithms.

Index Terms—Femtocells, OFDMA, heterogeneous networks, resource allocation, successive convex approximation.

I. INTRODUCTION

Network densification is considered as one of the promising ways to jointly meet the exponentially rising demand for higher throughput and ubiquitous coverage. The main idea is to increase explosively number of end-users which are served by a large number of small cell base stations (BSs) [1]. Among

This work was funded by the Australian Government under the Endeavour Postgraduate Scholarship Program and by the Science and Engineering Faculty, Queensland University of Technology, Australia under the Faculty Write-up Scholarship. This publication has emanated from research supported in part by a Grant from Science Foundation Ireland under Grant number 17/CDA/4786. The work of Vu was supported by the projects "Flexible Uplink-Downlink Resource Management for Energy and Spectral Efficiency Enhancing in Future Wireless Networks (FURMESFuN)" funded by the Academy of Finland under Grant 31089, and "6Genesis Flagship" funded by the Academy of Finland under Grant 318927.

Nam-Tran Le and Dhammika Jayalath are with the Science and Engineering Faculty, Queensland University of Technology, Queensland, Australia. (email: namtran.le@hdr.qut.edu.au; dhammika.jayalath@qut.edu.au).

Le-Nam Tran is with the School of Electrical and Electronic Engineering, University College Dublin, Dublin D04 V1W8, Ireland (e-mail: nam.tran@ucd.ie).

Quang-Doanh Vu is with the Centre for Wireless Communications, University of Oulu, P.O.Box 4500, FI-90014, Oulu, Finland (email: doanh.vu@oulu.fi).

different kinds of small cells, femtocell is perhaps the most popular because of its unique features including small form-factor, IP based backhaul, and end-users supported ad-hoc deployment without any network planing and site survey [2]. Recent surveys show that 70% of mobile data and 50% of voice services originate from indoor femtocell BSs [3], [4].

Femtocell heterogeneous networks (Hetnets) are the deployment of femtocells in an existing area on top of traditional macrocells. Such an architecture can provide many benefits including higher coverage area and low-power transmission, improved network capacity, reduced macrocell loading and cost benefits, to name but a few [3], [5], [6]. Consequently, dense small cell HetNets are going to be the main network architecture for the fifth generation (5G) systems [7]. On the other hand, orthogonal frequency division multiple access (OFDMA) technology is expected to improve the system capacity without intra-cell interference. Collectively, dense OFDMA femtocell Hetnets are likely to be the core network in the near future [8].

Over the last decade, energy efficiency (EE) has become one of the main criteria for designing wireless systems due to the concern of sustainable development [9]. In dense small cell Hetnets, EE design has received significant attention since operating a massive number of BSs may consume enormous energy. There are two common metrics for EE [9]: network EE (NEE) defined as a ratio between sum of all individual throughputs and total consumed power, and weighted sum of the individual EE (WSEE). The former focuses on the performance of overall network while the latter balances the individual EEs of different BSs leading to the EE fairness. Beside the two mentioned EE metrics, the exponentially-weighted product of the energy efficiencies over all subcarriers and all BSs was considered in [10]. Herein, we will focus on maximizing NEE and WSEE.

Energy efficiency resource allocation should account for different practical requirements. Specifically, beside the transmit power budget at BSs, minimum data rate for delay-sensitive (DS) users and the fairness between delay-tolerant (DT) users also need to be considered. Here, the minimum data rate requirement is to satisfy some quality-of-service (QoS) or to stabilize queuing (e.g. keeping the queuing sizes small). On the other hand, the fairness between DT users should be maintained due to the fact that users with less favourable channel gains may not be able to receive any data at a given time, because most of resources are assigned to the users of better channel gains.

A. Related work

Dense small cell networks were investigated in many works with different scenarios. For examples, in [11], authors developed a model of an artificial immune system that automatically turns on/off depending on traffic demands. Assuming stochastic traffic arrivals, an optimization-based scheme for the EE resource management in pico-transceiver Hetnets was proposed in [12]. The work of [13] studied the problem of determining the number of deployed antennas, transmit power levels and optimal BS density for EE maximization in dense small cell networks with massive MIMO technique. An EE optimization framework that minimizes small cells' energy under the guarantee of minimum average user rate or instantaneous user rate was considered in [14]. [15] investigated WSEE and NEE maximization problems in multicarrier wireless networks. However, the approach in [15] cannot be applied to OFDMA systems since it does not consider the OFDMA principle that a specific resource block (RB) can only be allocated to one user.

There exist many works considering joint subcarrier and power allocation in OFDMA systems. In [16], an energy optimization algorithm was proposed for dense OFDM networks, considering the load factor of OFDM systems. For EE resource allocation in OFDMA systems, an algorithm of power minimization was proposed in [17]. In [18], a resource allocation algorithm for EE optimization was studied in conventional OFDMA systems. An EE optimization scheme for both RB and power allocation was investigated for MIMO-OFDMA systems in [19]. [20] studied resource allocation for EE in MIMO OFDMA wireless networks. However, DT and DS constraints were not considered. [21] studied thWSEE maximization problem under the DS constraint only. In [22], EE resource allocation was considered with both DS and DT services. Generally, these works considered conventional OFDM systems where there is one BS, i.e. the intercell interference does not exist. Thus the approaches in these works are not straightforwardly applicable to OFDM Hetnets, especially for dense deployment, where co-tier and cross-tier interference should be carefully controlled.

Resource allocation for OFDM Hetnets was studied in [8], [23]–[25]. In particular, [8], [23] proposed spectral efficiency (SE) frameworks which maximize throughput under constraints of users' power. Recently, a joint allocation of RB and power for EE maximization was proposed in [24], which only includes DS constraints. Therein, the intercell interference is assumed to be zero. The assumption is quite strong in practical implementation. Power minimization only with DS constraints was considered in [25]. An interference-aware EE scheme was developed for both tiers in OFDMA femtocell Hetnets in [26], which did not include DT constraints. [27] proposed a power minimization transmission method for the two-tier LTE macro-femtocell network. To the best of our knowledge, resource allocation for EE maximization in OFDMA-based femtocell Hetnets under DS and DT constraints has not been addressed previously.

B. Main contributions

Motivated by the above discussions, we consider the EE resource allocation problem in OFDMA-based femtocell Hetnets with dense deployment. The aim is to jointly design RB and power allocation at macrocell and femtocell BSs such that the EE performance is maximized under both DS and DT constraints. Both WSEE and NEE metrics are considered. Here, for OFDMA-based systems, the joint allocation of RB and power may achieve better EE performance compared to only power optimization, which can be understood from the definition of EE. With the property of OFDMA systems that a BS allocate a RB to only one user, both throughput and power consumption are the functions of RB and power.

The considered design problems are highly intractable which combines the difficulties from combinatoric nature and (continuous) nonconvex functions. Thus, we focus on developing efficient sub-optimal solutions. To this end, we tackle the combinatoric nature by three different approaches: (1) mixed-integer programming formulation, (2) time-sharing formulation and (3) sparsity-inducing formulation. These formulations lead to the relaxed problems which are continuous nonconvex programs. Inspired by the efficiency of successive convex approximation (SCA) framework [28] in wireless network design, we then customize this technique to deal with the nonconvexity. Consequently, the solutions are achieved by iterative procedures which are provably convergent. In particular, in each iteration, only one second-order cone program (SOCP) needs to be solved. The convergence and effectiveness of the proposed algorithms are then evaluated by extensive simulations.

We summarized the contributions of this paper as follows:

- Dense deployment of femtocells with both co-tier and cross-tier interference is considered.
- The proposed EE resource allocation algorithms support heterogeneous services including delay-sensitive (DS) users with minimum data rate requirement and delay-tolerant (DT) users with fairness in data rate.
- Both WSEE and NEE maximization problems are considered in the three formulations.

The rest of the paper is organized as follows. The system model, power consumption model and EE metrics are discussed in Section II. The three formulations of EE problems are presented in Section III. The proposed solutions for the formulations are provided in Sections IV, V, and VI. We discuss the convergence property and computational complexity of the proposed solutions in Section VII. The Numerical results are provided in Section VIII. Finally, Section IX summarizes the work.

Notation: The following notations are used throughout the paper. Bold lower and upper case letters represent vectors and matrices, respectively; calligraphic upper case letters represent sets; $|\mathcal{A}|$ denotes cardinality of set \mathcal{A} ; $\|\mathbf{a}\|_q$ is the ℓ_q -norm of the vector \mathbf{a} ; notation \odot denotes Schur-Hadamard (element-wise) multiplication of two matrices; \mathbf{A}^T is normal transpose of matrix \mathbf{A} ; $\mathcal{CN}(0, a)$ denotes complex Gaussian random variable with zero mean and variance a .

II. ENERGY EFFICIENCY FRAMEWORK OF OFDMA HETEROGENEOUS NETWORKS

A. System Model

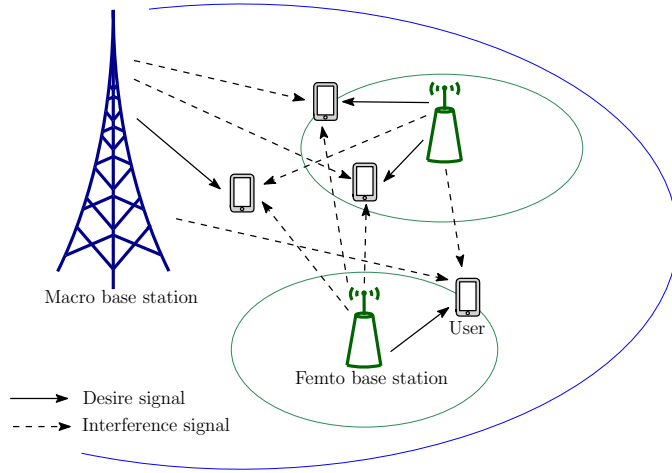


Fig. 1. A scenario of downlink transmission in OFDMA HetNets.

We consider a downlink transmission in OFDMA HetNets as shown in Fig. 1 where there are one macro base station (BS) and a set of K femto BSs. Let $\mathcal{K} \triangleq \{0, 1, 2, \dots, K\}$ denote the set of all BSs where macro BS is referred as BS 0. BS k , $k \in \mathcal{K}$, serves a set of U_k users denoted by $\mathcal{U}_k \triangleq \{1, 2, \dots, U_k\}$. We assume that each user is served by only one BS, i.e. $\mathcal{U}_k \cap \mathcal{U}_j = \emptyset$ where $k, j \in \mathcal{K}$ and $k \neq j$. Here, the heterogeneous services include two categories: delay-sensitive (DS) data with minimum data rate requirements and delay-tolerant (DT) data with proportional fairness [29], [30]. Let $\mathcal{U}_k^S \subseteq \mathcal{U}_k$ and $\mathcal{U}_k^T \subseteq \mathcal{U}_k$ be the set of DS users and DT users served by BS k , respectively. We assume that each user only receives one kind of service category at a given time, i.e. $\mathcal{U}_k^S \cap \mathcal{U}_k^T = \emptyset$ and $|\mathcal{U}_k^S| + |\mathcal{U}_k^T| = U_k$ for all k . Let N be the number of orthogonal resource blocks (RB) with a bandwidth W .¹ From now on, without loss of generality, let us suppose $W = 1$ Hz for notational convenience.² For OFDMA systems, BS k assigns a specific RB to only one user of \mathcal{U}_k so that there exists no intra-cell interference. Let k_u denote user u served by BS k , $h_{j,k_u,n}$ denote the channel gain between BS j and user k_u on RB n , and $p_{k_u,n}$ denote the power allocated to user k_u on RB n . Note that $p_{k_u,n}$ is nonzero only when RB n is assigned to user k_u . With these notations and the above assumptions, the signal-to-interference-plus-noise ratio (SINR) at user k_u on RB n can be written as [8]

$$\gamma_{k_u,n}(\mathbf{p}) = \frac{p_{k_u,n} h_{k,k_u,n}}{\sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} p_{jv,n} h_{j,k_u,n} + \sigma_{k_u,n}^2} \quad (1)$$

where \mathbf{p} denotes a vector stacking all $p_{k_u,n}$, and $\sigma_{k_u,n}^2$ is the variance of the zero mean additive white Gaussian noise

¹In Long-Term Evolution (LTE) system, one RB includes a group of 12 adjacent subcarriers and 6 or 7 consecutive OFDM symbols [31].

²The practical value of bandwidth W will be considered in the simulation section.

(AWGN). And the total achievable data rate of user k_u is given by

$$R_{k_u}(\mathbf{p}) = \sum_{n=1}^N \log(1 + \gamma_{k_u,n}(\mathbf{p})). \quad (2)$$

The considered system model is inspired from the coordinated beamforming, which is one of the coordinated multi point (CoMP) transmission techniques in LTE networks. In coordinated beamforming, the base stations cooperate to design their beamforming vectors (i.e. the power levels herein), but data is just transmitted from a serving base station. On the other hand, joint transmission CoMP (i.e. a user can simultaneously receive data from multiple BSs) can improve the performance. But this transmission scheme is very hard to implement in reality since it requires very strict synchronization between BSs.³

B. Power Consumption Model

We develop the power consumption model for the considered network based on the one widely used in the works related to EE designs [32]–[34]. In particular, the consumed power at BS k operating in the transmit mode, can be divided into two parts called static power and dynamic power. The static part includes the power for operating circuits and preparing the transmitted signal including filters, coolers, baseband processing, frequency synthesizer, RF chains, etc. This part is commonly assumed to be fixed. The dynamic part includes the power dissipated in the power amplifier for sending the RF signal, which is linearly dependent with the radiated power. In case of the users, who operate in the reception mode, the main consumed power is for operating circuits which is also fixed. Based on the discussion, let us denote by \tilde{P}_k^{sta} and $\tilde{P}_{k_u}^{\text{sta}}$ the static power consumed at BS k and user k_u , respectively. Then the total power consumed by BS k and its served users \mathcal{U}_k can be written as

$$g_k(\mathbf{p}) = \underbrace{\frac{1}{\xi_k} \left(\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N p_{k_u,n} \right)}_{\text{dynamic power}} + \underbrace{\tilde{P}_k^{\text{sta}} + \sum_{u \in \mathcal{U}_k} \tilde{P}_{k_u}^{\text{sta}}}_{\text{static power}} \quad (3)$$

where $\xi_k \in (0, 1)$ is the efficiency of the power amplifier at BS k .

C. Energy Efficiency Metrics

In this paper we consider two widely used EE design criteria. The first one is called weighted sum EE (WSEE) defined as [33]–[35]

$$E_{\text{WS}}(\mathbf{p}) \triangleq \sum_{k \in \mathcal{K}} w_k \frac{\sum_{u \in \mathcal{U}_k} R_{k_u}(\mathbf{p})}{g_k(\mathbf{p})} \quad (4)$$

where the term $\frac{\sum_{u \in \mathcal{U}_k} R_{k_u}(\mathbf{p})}{g_k(\mathbf{p})}$ is the EE of cell k , and the weighting factor w_k is introduced to achieve some degree of EE fairness among cells, which is critical in HetNets.

³It is possible to modify the proposed solutions to deal with joint transmission CoMP. In this case, the SINR given in (1) is a linear combination of the power allocation coefficients. The same observation is also applied to the denominator in (1).

The second criterion is to focus on the EE of the overall system called network EE (NEE) which is defined as [36], [37]

$$E_{\text{Net}}(\mathbf{p}) \triangleq \frac{\sum_{k \in \mathcal{K}} \sum_{u \in \mathcal{U}_k} R_{k_u}(\mathbf{p})}{\sum_{k \in \mathcal{K}} g_k(\mathbf{p})} \quad (5)$$

where the numerator and denominator account for the total transmitted data and total consumed power over entire system, respectively.

By definitions, EE design criteria (4) and (5) are measured in bits (or nats) per Joule.

III. ENERGY EFFICIENCY PROBLEM FORMULATIONS

We focus on the problems of WSEE maximization (WSEEMax) and NEE maximization (NEEMax), subject to some specific constraints on BSs' transmit power, QoS, and users' fairness. As mentioned in the previous section, the requirement in OFDMA systems is that a RB can only be assigned to a single user by a BS. This naturally makes the considered problems a combinatorial program (commonly known as mixed-integer program). In general, a mixed-integer program is NP-hard [38]. Additionally, both $E_{\text{WS}}(\mathbf{p})$ and $E_{\text{Net}}(\mathbf{p})$ are intractable due to the SINR term. In the rest of the section, we introduce three different formulations that lead to efficient solutions. The motivation for applying different approaches is that one of these may produce a better performance than the others for a specific scenario.

A. Formulation 1: Mixed 0-1 Integer Programming Approach

We start with a mixed 0-1 integer programming formulation that arises naturally from the principle of OFDMA systems. Let us introduce a Boolean variable $a_{k_u, n}$ which represents the connection status between RB n and user k_u . Specifically,

$$a_{k_u, n} = \begin{cases} 1 & \text{if RB } n \text{ is allocated to user } k_u \\ 0 & \text{otherwise.} \end{cases}$$

Then the WSEEMax and NEEMax problems can be formulated as

$$\underset{\mathbf{a}, \mathbf{p}}{\text{maximize}} f(\mathbf{p}) \quad (6a)$$

$$\text{subject to } p_{k_u, n} \leq a_{k_u, n} P_k^{\max}, \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (6b)$$

$$\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N p_{k_u, n} \leq P_k^{\max}, \forall k \in \mathcal{K} \quad (6c)$$

$$R_{k_u}(\mathbf{p}) \geq R_{k_u}^{\min}, \forall u \in \mathcal{U}_k^S, k \in \mathcal{K} \quad (6d)$$

$$p_{k_u, n} \geq 0, \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (6e)$$

$$a_{k_u, n} \in \{0, 1\}, \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (6f)$$

$$\sum_{u \in \mathcal{U}_k} a_{k_u, n} = 1, \forall k \in \mathcal{K}, n = 1, \dots, N \quad (6g)$$

$$\frac{R_{k_u}(\mathbf{p})}{\sum_{v \in \mathcal{U}_k^T} R_{k_v}(\mathbf{p})} = c_{k_u}, \forall u \in \mathcal{U}_k^T, k \in \mathcal{K} \quad (6h)$$

where \mathbf{a} denotes a vector stacking all $a_{k_u, n}$'s; $f(\mathbf{p})$ stands for $E_{\text{WS}}(\mathbf{p})$ and $E_{\text{Net}}(\mathbf{p})$ in the WSEEMax and NEEMax

problems, respectively; P_k^{\max} denotes the maximum transmit power at BS k ; $R_{k_u}^{\min}$ denotes the minimum data rate for DS user k_u , $u \in \mathcal{U}_k^S$; and c_{k_u} represents the priority of user k_u among the DT users where $\sum_{u \in \mathcal{U}_k^T} c_{k_u} = 1$.

Next, let us discuss the physical interpretation of the constraints in (6). Constraint (6b) based on the big-M formulation theory is to ensure that if $a_{k_u, n} = 0$, then $p_{k_u, n} = 0$ [38]. This means no power is allocated to user k_u on RB n , if RB n is not assigned to user k_u . (6c) represents the power constraint at the BSs. Constraint (6d) is to guarantee the QoS for the DS users. (6g) is taken into account to make sure that a specific RB is assigned to only a single user in each cell. Constraint (6h) provides the proportional fairness among the DT users. In practice, (6h) is hardly achieved as the equality constraint is relatively strict, which makes the feasible region really small. Here, we overcome the issue by adopting a simple relaxed version of (6h), which is written as

$$(1 - \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} R_{k_v}(\mathbf{p}) \right) \leq R_{k_u}(\mathbf{p}) \\ \leq (1 + \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} R_{k_v}(\mathbf{p}) \right), \forall u \in \mathcal{U}_k^T, k \in \mathcal{K} \quad (7)$$

where $\alpha_f \in [0, 1]$ is the parameter controlling the degree of relaxation, i.e. the smaller α_f , the tighter relaxation. Clearly, when $\alpha_f \rightarrow 0$, (7) approaches (6h).

B. Formulation 2: Time-Sharing Approach

The second formulation is obtained based on the time-sharing approach, the technique widely used in designing OFDMA systems [20], [24], [39], [40]. By abuse of notation, we let $a_{k_u, n} \in [0, 1]$ stand for the time-sharing factor of user k_u over RB n . More explicitly, $a_{k_u, n}$ can be viewed as the fraction of time that RB n is assigned to user k_u during one transmission frame. Accordingly, the actual power allocated to user k_u is $a_{k_u, n} p_{k_u, n}$. Thus the SINR in (1) is modified into [40]

$$\gamma_{k_u, n}(\mathbf{p}, \mathbf{a}) = \frac{a_{k_u, n} p_{k_u, n} h_{k, k_u, n}}{\sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} a_{j, v, n} p_{j, v, n} h_{j, k_u, n} + \sigma_{k_u, n}^2}. \quad (8)$$

Similarly, the sum rate at user k_u in (2) and the total transmit power of cell k in (3) are rewritten as [40]

$$R_{k_u}(\mathbf{p}, \mathbf{a}) = \sum_{n=1}^N a_{k_u, n} \log(1 + \gamma_{k_u, n}(\mathbf{p}, \mathbf{a})) \quad (9)$$

$$g_k(\mathbf{p}, \mathbf{a}) = \frac{1}{\xi_k} \left(\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N a_{k_u, n} p_{k_u, n} \right) + P_k^{\text{sta}} \quad (10)$$

respectively, where $P_k^{\text{sta}} = \tilde{P}_k^{\text{sta}} + \sum_{u \in \mathcal{U}_k} \tilde{P}_{k_u}^{\text{sta}}$. Clearly, (8), (9), and (10) become (1), (2), and (3), respectively, when \mathbf{a} is a Boolean vector. With the introduced time-sharing factors, the WSEEMax and NEEMax problems are now formulated as

$$\underset{\mathbf{p}, \mathbf{a}}{\text{maximize}} \tilde{f}(\mathbf{p}, \mathbf{a}) \quad (11a)$$

$$\text{subject to } \sum_{u \in \mathcal{U}_k} \sum_{n=1}^N a_{k_u,n} p_{k_u,n} \leq P_k^{\max}, \forall k \in \mathcal{K} \quad (11b)$$

$$R_{k_u}(\mathbf{p}, \mathbf{a}) \geq R_{k_u}^{\min}, \forall u \in \mathcal{U}_k^S, \forall k \in \mathcal{K} \quad (11c)$$

$$p_{k_u,n} \geq 0, \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (11d)$$

$$a_{k_u,n} \in [0, 1], \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (11e)$$

$$\sum_{u \in \mathcal{U}_k} a_{k_u,n} = 1, \forall k \in \mathcal{K}, n = 1, \dots, N \quad (11f)$$

$$(1 - \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} R_{k_v}(\mathbf{p}) \right) \leq R_{k_u}(\mathbf{p}), \quad \forall u \in \mathcal{U}_k^T, k \in \mathcal{K} \quad (11g)$$

$$R_{k_u}(\mathbf{p}) \leq (1 + \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} R_{k_v}(\mathbf{p}) \right), \quad \forall u \in \mathcal{U}_k^T, k \in \mathcal{K}. \quad (11h)$$

where $\tilde{f}(\mathbf{p}, \mathbf{a})$ represents the two EE metrics constructed from $R_{k_u}(\mathbf{p}, \mathbf{a})$ and $g_k(\mathbf{p}, \mathbf{a})$.

We note that Formulation 2 is not a linear relaxation of Formulation 1. Formulation 2 uses the time sharing relaxation, which means that RB_n is assigned to user k_u in a fraction of one time slot. With this understanding, there is no intra-tier interference in the same cell. On the other hand, Formulation 1 with continuous relaxation suffer the intra-tier interference, which is different from Formulation 2.

C. Formulation 3: Sparsity Inducing Approach

The third formulation is inspired from the context of sparsity inducing regularization methods [32], [41]. Let $\mathbf{p}_{k,n} = [p_{k_1,n}, p_{k_2,n}, \dots, p_{k_{U_k},n}]^T$ be the vector composing of all allocated powers at BS k on RB n . We recall that the principle of OFDMA systems implies that only one element of $\mathbf{p}_{k,n}$ are nonzero, which is mathematically written as $\|\mathbf{p}_{k,n}\|_0 = 1$. Accordingly, the WSEEMax and NEEMax problems can be formulated as

$$\underset{\mathbf{p}}{\text{maximize}} f(\mathbf{p}) \quad (12a)$$

$$\text{subject to } \|\mathbf{p}_{k,n}\|_0 = 1, \forall k \in \mathcal{K}, n = 1, \dots, N. \quad (12b)$$

$$(6c)-(6e), (11g), (11h). \quad (12c)$$

An advantage of formulation (12) is that it does not require introducing variables related to RB assignment, and thus the number of variables is smaller compared to that of the other formulations.

The three formulations is highly intractable due to the mixed integer and/or nonconvexity, computing globally optimal solutions is challenging and is often not of practical interest. In the following sections, we propose efficient methods finding high-quality feasible solutions to these problems.

IV. PROPOSED SOLUTION TO FORMULATION 1

A. A Tighter Big-M Formulation for Continuous Relaxation

The common approach dealing with discrete variables is relaxing them into continuous domain. To this end, we will manipulate continuous relaxation (CR) of (6), and thus a tight

CR is desired. We remark that big-M formulation (6b) usually offers a relatively poor CR. In order to obtain a tighter CR, we propose to replace (6b) by

$$p_{k_u,n} \leq (a_{k_u,n})^q P_k^{\max}, \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (13)$$

for any integer $q \geq 1$. Note that $a_{k_u,n} = (a_{k_u,n})^q$ for $a_{k_u,n} \in \{0, 1\}$, and thus (6b) and (13) are indeed equivalent. However, the CR results in different tightness. Indeed, when $a_{k_u,n}$ is relaxed to be continuous over the interval $[0, 1]$, it holds that $(a_{k_u,n})^q \leq a_{k_u,n}$ for $q \geq 1$. Thus, if $a_{k_u,n}$ is feasible to (13), it is also feasible to (6b). Let f^* and f_q^* be the optimal objective of (6) and the CR of (6) with (13), respectively. Then we have the following inequality

$$f^* \leq \dots \leq f_q^* \leq \dots \leq f_1^*. \quad (14)$$

which comes from the fact that the wider the feasible set, the bigger the optimal value. Clearly, inequality (14) promotes the use of (13) in the sequel.

B. Solution to WSEEMax

Let us first focus on the WSEEMax problem. From the discussion in the previous subsection, the CR of (6) can be written as

$$\underset{\mathbf{a}, \mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k \eta_k \quad (15a)$$

$$\text{subject to } \frac{\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N x_{k_u,n}}{g_k(\mathbf{p})} \geq \eta_k, \forall k \quad (15b)$$

$$\frac{p_{k_u,n} h_{k,k_u,n}}{\beta_{k_u,n}(\mathbf{p})} \geq \theta_{k_u,n}, \forall k, u \in \mathcal{U}_k, n \quad (15c)$$

$$p_{k_u,n} \leq (a_{k_u,n})^q P_k^{\max}, \forall k, u \in \mathcal{U}_k, n \quad (15d)$$

$$\log(1 + \theta_{k_u,n}) \geq x_{k_u,n}, \forall k, u \in \mathcal{U}_k, n \quad (15e)$$

$$\sum_{n=1}^N x_{k_u,n} \geq R_{k_u}^{\min}, \forall u \in \mathcal{U}_k^S, \forall k \quad (15f)$$

$$(1 - \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} \sum_{n=1}^N x_{k_v,n} \right) \leq \sum_{n=1}^N x_{k_u,n}, \quad \forall u \in \mathcal{U}_k^T, \forall k \quad (15g)$$

$$\sum_{n=1}^N x_{k_u,n} \leq (1 + \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} \sum_{n=1}^N x_{k_v,n} \right), \quad \forall u \in \mathcal{U}_k^T, \forall k \quad (15h)$$

$$0 \leq a_{k_u,n} \leq 1, \forall k, u \in \mathcal{U}_k, n \quad (15i)$$

$$(6c), (6e), (6g). \quad (15j)$$

where $\beta_{k_u,n}(\mathbf{p}) \triangleq \sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} p_{j_v,n} h_{j,k_u,n} + \sum_{v \in \mathcal{U}_k \setminus \{u\}} p_{k_v,n} h_{k,k_u,n} + \sigma_{k_u,n}^2$; $\boldsymbol{\eta} \triangleq \{\eta_k\}_{k \in \mathcal{K}}$, $\boldsymbol{\theta} \triangleq \{\theta_{k_u,n}\}_{k,u,n}$, and $\mathbf{x} \triangleq \{x_{k_u,n}\}_{k,u,n}$ are the newly introduced slack variables. The purpose of the variable introduction is to reveal the hidden convexity in the problem. The second term in the denominator of the fraction in (15c), i.e. $\beta_{k_u,n}(\mathbf{p})$, is added due to the CR. It is easy to justify (by contradiction) that at the optimality, both (15b) and (15c) hold with equality. Therefore (15) is indeed the CR of (6). In

this regard, η_k represents the EE of cell k , and $\theta_{k_u,n}$ is the SINR of user k_u over RB n .

The nonconvexity of (15) is due to that of (15b), (15c) and (15d). To solve (15) we apply the notion of successive convex approximation (SCA) [42], [43], an efficient optimization techniques overcoming nonconvex problems. The core idea is to iteratively solve a sequence of convex approximate programs of (15) until convergence is achieved. To proceed, let us first consider (15b) and equivalently rewrite it as

$$\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N x_{k_u,n} \geq \eta_k g_k(\mathbf{p}), \forall k. \quad (16)$$

Following the principle of SCA, we need to find a convex upper bound of the right-hand-side (RHS) of (16). For this purpose, we recall the following inequality [43]

$$\eta_k g_k(\mathbf{p}) \leq \frac{y_k}{2} \eta_k^2 + \frac{1}{2y_k} g_k^2(\mathbf{p}) \quad (17)$$

for arbitrary constant $y_k > 0$. Inequality (17) holds with equality when $y_k = \frac{g_k(\mathbf{p})}{\eta_k}$.

We now consider (15c). Similar to (15b), we rewrite (15c) as $p_{k_u,n} h_{k,k_u,n} \geq \theta_{k_u,n} \beta_{k_u,n}(\mathbf{p})$, and have a convex upper bound of the RHS given as

$$\theta_{k_u,n} \beta_{k_u,n}(\mathbf{p}) \leq \frac{\varphi_{k_u,n}}{2} \theta_{k_u,n}^2 + \frac{1}{2\varphi_{k_u,n}} (\beta_{k_u,n}(\mathbf{p}))^2 \quad (18)$$

for all constant $\varphi_{k_u,n} > 0$ where the equality holds at $\varphi_{k_u,n} = \frac{\beta_{k_u,n}(\mathbf{p})}{\theta_{k_u,n}}$.

We turn our attention to (15d) and note that its RHS is convex for $q > 1, a_{k_u,n} \geq 0$. Again, following the SCA framework, the RHS of (15d) is to be replaced by a lower bound, which can be easily obtained by the first order Taylor series approximation due to its convexity. Specifically, we have

$$(a_{k_u,n})^q \geq (a_{k_u,n}^{(i)})^q + q(a_{k_u,n}^{(i)})^{q-1} (a_{k_u,n} - a_{k_u,n}^{(i)}), \forall k, u, n \quad (19)$$

where $a_{k_u,n}^{(i)} \in [0, 1]$. In fact, we have linearized $(a_{k_u,n})^q$ around the operating point $a_{k_u,n}^{(i)}$.

Remark 1. [Computationally Efficient Approximation]: With the approximations provided in (17), (18), and (19), we are ready to obtain a convex approximation of (15). However, the resulted approximate problem is a generic convex program due to the logarithm function in (15e), which leads to some computational disadvantages those are lack of efficient off-the-shelf solvers and, more importantly, the computational complexity scales fast with the size of the problem [44]. These would become the major drawbacks in the considered systems since the number of RBs (and thus the number of variables) is usually large. In order to overcome the issue, we use a concave lower bound of the logarithm function obtained based on the SCA principle, which admits a second-order cone (SOC) representation. In particular, we first have the lower bound of $\log(1 + \theta_{k_u,n})$ given as [45]

$$\log(1 + \theta_{k_u,n}) \geq \log\left(1 + \theta_{k_u,n}^{(i)}\right) + 1 - \frac{1 + \theta_{k_u,n}^{(i)}}{1 + \theta_{k_u,n}}. \quad (20)$$

Algorithm 1 The proposed algorithm solving (15)

- 1: **Initialization:** Set $i := 0$ and generate a feasible point of (15) denoted by $(\mathbf{a}^{(0)}, \mathbf{p}^{(0)}, \boldsymbol{\eta}^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{x}^{(0)})$.
- 2: **repeat**
- 3: Set $\varphi_{k_u,n} := \frac{\beta_{k_u,n}(\mathbf{p}^{(i)})}{\theta_{k_u,n}^{(i)}}$, $y_k := \frac{g_k(\mathbf{p}^{(i)})}{\eta_k^{(i)}}$.
- 4: Solve (23) and denote the optimal as $(\mathbf{a}^*, \mathbf{p}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$.
- 5: Update $i := i + 1$, and $(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)}) := (\mathbf{a}^*, \mathbf{p}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$.
- 6: **until** Convergence
- 7: **Output:** $(\mathbf{a}^*, \mathbf{p}^*)$

where $\theta_{k_u,n}^{(i)} \geq 0$. From (20), we achieve a safe approximation of (15e) given by

$$\left(\log(1 + \theta_{k_u,n}^{(i)}) + 1 - x_{k_u,n}\right) (1 + \theta_{k_u,n}) \geq 1 + \theta_{k_u,n}^{(i)}. \quad (21)$$

The above inequality is in fact a rotated SOC, i.e. it can be written as the following SOC [46]

$$\log(1 + \theta_{k_u,n}^{(i)}) + 2 - x_{k_u,n} + \theta_{k_u,n} \geq \sqrt{4 \left(1 + \theta_{k_u,n}^{(i)}\right) + \left(\log(1 + \theta_{k_u,n}^{(i)}) - x_{k_u,n} - \theta_{k_u,n}\right)^2}. \quad (22)$$

In summary, let $(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)})$ be a feasible point of (15), then a convex approximation program of (15) is given by

$$\text{maximize}_{\mathbf{a}, \mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}} \sum_{k \in \mathcal{K}} w_k \eta_k \quad (23a)$$

$$\text{subject to} \sum_{u \in \mathcal{U}_k} \sum_{n=1}^N x_{k_u,n} \geq \frac{y_k}{2} \eta_k^2 + \frac{1}{2y_k} g_k^2(\mathbf{p}), \forall k \quad (23b)$$

$$h_{k,k_u,n} p_{k_u,n} \geq \frac{\varphi_{k_u,n}}{2} (\theta_{k_u,n})^2 + \frac{1}{2\varphi_{k_u,n}} (\beta_{k_u,n}(\mathbf{p}))^2, \quad \forall k, u, n \quad (23c)$$

$$p_{k_u,n} \leq P_k^{\max} (a_{k_u,n}^{(i)})^q + P_k^{\max} q (a_{k_u,n}^{(i)})^{q-1} (a_{k_u,n} - a_{k_u,n}^{(i)}), \quad \forall k, u, n \quad (23d)$$

$$(15f)-(15j), (22). \quad (23e)$$

which is an SOCP. And the proposed SCA-based procedure solving (15) is outlined in **Algorithm 1**.

1) *Finding a Feasible Point of (15):* To start **Algorithm 1**, we need to find a feasible point of (15) which is, in general, as difficult as solving (15). To overcome this issue, we use an efficient heuristic method based on a regularized formulation of (23) which is given by [47]

$$\text{maximize}_{\mathbf{a}, \mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}} \sum_{k \in \mathcal{K}} w_k \eta_k - \lambda \left(\sum_{k \in \mathcal{K}} \sum_{u \in \mathcal{U}_k^S} \epsilon_{k_u} \right) \quad (24a)$$

$$\text{subject to} \sum_{n=1}^N x_{k_u,n} + \epsilon_{k_u} \geq R_k^{\min}, \forall u \in \mathcal{U}_k^S, k \quad (24b)$$

$$\epsilon_{k_u} \geq 0, \forall u \in \mathcal{U}_k^S, k \quad (24c)$$

$$(15g)-(15j), (22), (23c), (23d). \quad (24d)$$

Algorithm 2 The proposed algorithm to find a feasible point of (15).

- 1: **Initialization:** set $i := 0$ and randomly generate $(\mathbf{a}^{(0)}, \mathbf{p}^{(0)}, \boldsymbol{\eta}^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{x}^{(0)})$ that satisfies (15b)-(15e), (15g)-(15j).
- 2: **repeat**
- 3: Set $\varphi_{k_u, n} := \frac{\beta_{k_u, n}(\mathbf{p}^{(i)})}{\theta_{k_u, n}^{(i)}}$, $y_k := \frac{g_k(\mathbf{p}^{(i)})}{\eta_k^{(i)}}$.
- 4: Solve (24) and denote the optimal as $(\mathbf{a}^*, \mathbf{p}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*, \boldsymbol{\epsilon}^*)$.
- 5: Update $i := i + 1$, and $(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)}) := (\mathbf{a}^*, \mathbf{p}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$.
- 6: **until** $\boldsymbol{\epsilon}^* = 0$
- 7: **Output:** $(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)})$

where $\boldsymbol{\epsilon} \triangleq \{\epsilon_{k_u}\}_{k_u \in \mathcal{U}_k^S}$ and λ is a positive parameter. We can see that the set $\mathcal{S} \triangleq \{\mathbf{a}, \mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}, \boldsymbol{\epsilon} \mid (15b)-(15e), (15g)-(15j), (24b), (24c)\}$ is a relaxed version of the feasible set of (15), and a point of \mathcal{S} with $\boldsymbol{\epsilon} = 0$ is also a feasible point of (15). In addition, it is easy to find a point in \mathcal{S} . Particularly, it is simple to generate $(\mathbf{a}^{(0)}, \mathbf{p}^{(0)}, \boldsymbol{\eta}^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{x}^{(0)})$ such that (15b)-(15e), (15g)-(15j) are satisfied. Then there always exists $\boldsymbol{\epsilon} \geq 0$ satisfying (24b). With this point of \mathcal{S} , we can run an iterative procedure similar to **Algorithm 1**, which solves (24) in each iteration. And, by the maximization in (24a), it is expected that $\boldsymbol{\epsilon}$ will eventually decrease to zero after some iterations. As so, we obtain a feasible point of (15), which can be used to execute **Algorithm 1**. For physical interpretation, $\boldsymbol{\epsilon}$ represents the level of constraint violence in (15f). The second term in the objective in (24a) is the penalty of the violence. And constant λ determines a trade-off between optimization and feasibility. In summary, the procedure to find a feasible point of (15) is presented in **Algorithm 2**.

2) *Post Processing Phase:* We recall that (15) is the continuous relaxation version of (6). Thus the output of **Algorithm 1** needs to be processed in order to obtain a solution of (6). The post processing phase includes two steps that are mapping the RBs to the users and refining the allocated power. For the RB mapping, since a RB is exclusively assigned to only one user in an OFDMA system, it is reasonable to select user k_u with a value of $a_{k_u, n}^*$ closest to 1 for a particular RB n . Hence, we propose the following mapping rule

$$a_{k_u, n} = \begin{cases} 1 & \text{if } u = \arg \max_{v \in \mathcal{U}_k} a_{k_v, n}^* \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

Once \mathbf{a} is determined, power vector \mathbf{p} is recalculated. In particular, let $\mathcal{N}_{k_u} \in \{1, \dots, N\}$ denote the set of RBs allocated to user k_u , then the problem of power allocation, given vector \mathbf{a} , is given by

$$\underset{\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k \eta_k \quad (26a)$$

$$\text{subject to } \frac{\sum_{u \in \mathcal{U}_k} \sum_{n \in \mathcal{N}_{k_u}} x_{k_u, n}}{\frac{1}{\xi_k} \left(\sum_{u \in \mathcal{U}_k} \sum_{n \in \mathcal{N}_{k_u}} p_{k_u, n} \right) + P_k^{\text{sta}}} \geq \eta_k, \forall k \quad (26b)$$

$$\frac{p_{k_u, n} h_{k, k_u, n}}{\sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} p_{j_v, n} h_{j, k_u, n} + \sigma_{k_u, n}^2} \geq \theta_{k_u, n}, \quad \forall k, u \in \mathcal{U}_k, n \in \mathcal{N}_{k_u} \quad (26c)$$

$$\log(1 + \theta_{k_u, n}) \geq x_{k_u, n}, \quad \forall k, u \in \mathcal{U}_k, n \in \mathcal{N}_{k_u} \quad (26d)$$

$$\sum_{n \in \mathcal{N}_{k_u}} x_{k_u, n} \geq R_{k_u}^{\min}, \quad \forall u \in \mathcal{U}_k^S, \forall k \quad (26e)$$

$$(1 - \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} \sum_{n \in \mathcal{N}_{k_v}} x_{k_v, n} \right) \leq \sum_{n \in \mathcal{N}_{k_u}} x_{k_u, n}, \quad \forall u \in \mathcal{U}_k^T, \forall k \quad (26f)$$

$$\sum_{n \in \mathcal{N}_{k_u}} x_{k_u, n} \leq (1 + \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^T} \sum_{n \in \mathcal{N}_{k_v}} x_{k_u, n} \right), \quad \forall u \in \mathcal{U}_k^T, \forall k \quad (26g)$$

$$\sum_{u \in \mathcal{U}_k} \sum_{n \in \mathcal{N}_{k_u}} p_{k_u, n} \leq P_k^{\max}, \quad \forall k, p_{k_u, n} \geq 0, \quad \forall k, u \in \mathcal{U}_k, n \in \mathcal{N}_{k_u} \quad (26h)$$

in which the note that the number of entries in each of vectors \mathbf{p} , $\boldsymbol{\theta}$, and \mathbf{x} is only KN . Although problem (26) is nonconvex, we can use an SCA-based procedure similar to **Algorithm 1** in order to obtain a solution.

We recall that, in a specific cell, a RB is assigned to one user only (due to the OFDMA system principle). Thus, if the relaxed assignment variables are far from ‘0’ or ‘1’ when Algorithm 1 converges (i.e indecisive cases), we need to determine to whom the RB is allocated by means of rounding. Without the post processing phase, significant performance loss may occur since the power allocation is not optimized for the rounded assignment variables, or even worst, the resulting solution becomes infeasible. In our numerical experiments, we observe that the relaxed assignment variables are uncertain for most cases. Thus the post processing phase is normally required.

C. Solution to NEEMax

Now we turn our consideration to the NEEMax problem. It should be noted that the NEEMax is different from the WSEEMax problem only in objective function. So we can achieve a solution for the NEEMax problem by slightly modifying the procedure presented in Section IV-B. In particular, a convex approximation of the CR version of the NEEMax problem is given by

$$\underset{\mathbf{a}, \mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}}{\text{maximize}} \eta \quad (27a)$$

$$\text{subject to } \sum_{k \in \mathcal{K}} \sum_{u \in \mathcal{U}_k} \sum_{n=1}^N x_{k_u, n} \geq \frac{y}{2} \eta^2 + \frac{1}{2y} \left(\sum_{k \in \mathcal{K}} g_k(\mathbf{p}) \right)^2 \quad (27b)$$

$$(23c)-(23e). \quad (27c)$$

Also, after solving the CR problem, the post processing phase is conducted in order to obtain a solution for the NEEMax problem.

V. PROPOSED SOLUTION TO FORMULATION 2

In this section, we develop the method solving (11). For notational brevity, let us introduce variables $\mathbf{s} \triangleq \{s_{k_u,n}\}_{k,u,n}$ which represent the actual power allocated to the users, i.e.,

$$s_{k_u,n} = a_{k_u,n} p_{k_u,n}, \forall k \in \mathcal{K}, u \in \mathcal{U}_k, n = 1, \dots, N \quad (28)$$

Then, the SINR in (8) becomes

$$\gamma_{k_u,n}(\mathbf{s}) = \frac{s_{k_u,n} h_{k,k_u,n}}{\sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} s_{j_v,n} h_{j,k_u,n} + \sigma_{k_u,n}^2}. \quad (29)$$

Also, the data rate for user k_u and the total consumed power of cell $k \in \mathcal{K}$ (in (9) and (10)) are respectively rewritten as

$$R_{k_u}(\mathbf{s}, \mathbf{a}) = \sum_{n=1}^N a_{k_u,n} \log(1 + \gamma_{k_u,n}(\mathbf{s})) \quad (30)$$

$$g_k(\mathbf{s}) = \frac{1}{\xi_k} \left(\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N s_{k_u,n} \right) + P_k^{\text{sta}} \quad (31)$$

With these introduced notations, problem (11) can be reformulated as

$$\underset{\mathbf{s}, \mathbf{a}}{\text{maximize}} \quad \tilde{f}(\mathbf{s}, \mathbf{a}) \quad (32a)$$

$$\text{subject to} \quad \sum_{u \in \mathcal{U}_k} \sum_{n=1}^N s_{k_u,n} \leq P_k^{\text{max}}, \forall k \quad (32b)$$

$$R_{k_u}(\mathbf{s}, \mathbf{a}) \geq R_{k_u}^{\text{min}}, \forall k, u \in \mathcal{U}_k^{\text{S}} \quad (32c)$$

$$s_{k_u,n} \geq 0, \forall k, u, n \quad (32d)$$

$$a_{k_u,n} \in [0, 1], \forall k, u, n \quad (32e)$$

$$\sum_{u \in \mathcal{U}_k} a_{k_u,n} = 1, \forall k, n \quad (32f)$$

$$(1 - \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^{\text{T}}} R_{k_v}(\mathbf{s}, \mathbf{a}) \right) \leq R_{k_u}(\mathbf{s}, \mathbf{a}), \quad \forall u \in \mathcal{U}_k^{\text{T}}, k \in \mathcal{K} \quad (32g)$$

$$R_{k_u}(\mathbf{s}, \mathbf{a}) \leq (1 + \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^{\text{T}}} R_{k_v}(\mathbf{s}, \mathbf{a}) \right), \quad \forall u \in \mathcal{U}_k^{\text{T}}, k \in \mathcal{K}. \quad (32h)$$

The nonconvexity of problem (32) comes from the fractional form of the objective functions and (32c)-(32h). Again, we adopt the SCA framework to tackle (32). Let us focus on the WSEEMax problem, then the NEEmax problem is treated similarly. Similar to (15), we first use slack variables to reveal the hidden convexity in (32). Particularly, (32) is equivalent to

$$\underset{\mathbf{a}, \mathbf{s}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}}{\text{maximize}} \quad \sum_{k \in \mathcal{K}} w_k \eta_k \quad (33a)$$

$$\text{subject to} \quad \frac{\sum_{u \in \mathcal{U}_k} \sum_{n=1}^N x_{k_u,n}}{g_k(\mathbf{s})} \geq \eta_k, \forall k \quad (33b)$$

$$\frac{s_{k_u,n} h_{k,k_u,n}}{\sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} s_{j_v,n} h_{j,k_u,n} + \sigma_{k_u,n}^2} \geq \theta_{k_u,n}, \quad \forall k, u, n \quad (33c)$$

$$a_{k_u,n} \log(1 + \theta_{k_u,n}) \geq x_{k_u,n}, \forall k, u, n \quad (33d)$$

$$\sum_{n=1}^N x_{k_u,n} \geq R_{k_u}^{\text{min}}, \forall k, u \in \mathcal{U}_k^{\text{S}} \quad (33e)$$

$$(1 - \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^{\text{T}}} \sum_{n=1}^N x_{k_v,n} \right) \leq \sum_{n=1}^N x_{k_u,n}, \quad \forall u \in \mathcal{U}_k^{\text{T}}, \forall k \quad (33f)$$

$$\sum_{n=1}^N x_{k_u,n} \leq (1 + \alpha_f) c_{k_u} \left(\sum_{v \in \mathcal{U}_k^{\text{T}}} \sum_{n=1}^N x_{k_v,n} \right), \quad \forall u \in \mathcal{U}_k^{\text{T}}, \forall k \quad (33g)$$

$$(32b)-(32f) \quad (33h)$$

Clearly, the difficulty in solving (33) is due to nonconvex constraints (33b)-(33d), which is approximated to achieve the convex approximated problem. In particular, the (33b) and (33c) are approximated in the same way with (15b) and (15c), respectively. In order to approximate (33d), we first equivalently rewrite it as

$$\theta_{k_u,n} \log(1 + \theta_{k_u,n}) \geq \frac{x_{k_u,n} \theta_{k_u,n}}{a_{k_u,n}} \quad (34)$$

which is then rearranged as

$$\underbrace{\theta_{k_u,n} \log(1 + \theta_{k_u,n}) + \frac{(x_{k_u,n} - \theta_{k_u,n})^2}{4a_{k_u,n}}}_{\triangleq h(x_{k_u,n}, \theta_{k_u,n}, a_{k_u,n})} \geq \frac{(x_{k_u,n} - \theta_{k_u,n})^2}{4a_{k_u,n}}. \quad (35)$$

The first term in LHS of (35) is a negative entropy function and the second term is a quadratic-over-linear function. Thus $h(x_{k_u,n}, \theta_{k_u,n}, a_{k_u,n})$ is convex with the involved variables [48]. And the concave lower bound of $h(x_{k_u,n}, \theta_{k_u,n}, a_{k_u,n})$ can be obtained by its first order Taylor series approximation, i.e. the convex approximation of (35) is

$$\begin{aligned} & h(x_{k_u,n}^{(i)}, \theta_{k_u,n}^{(i)}, a_{k_u,n}^{(i)}) + (\theta_{k_u,n} - \theta_{k_u,n}^{(i)}) \times \\ & \left(-\frac{(x_{k_u,n}^{(i)} - \theta_{k_u,n}^{(i)})}{2a_{k_u,n}^{(i)}} + \frac{\theta_{k_u,n}^{(i)}}{\theta_{k_u,n}^{(i)} + 1} + \ln(\theta_{k_u,n}^{(i)} + 1) \right) \\ & + (x_{k_u,n} - x_{k_u,n}^{(i)}) \frac{(x_{k_u,n}^{(i)} - \theta_{k_u,n}^{(i)})}{2a_{k_u,n}^{(i)}} \\ & + (a_{k_u,n} - a_{k_u,n}^{(i)}) \left(-\frac{(x_{k_u,n}^{(i)} - \theta_{k_u,n}^{(i)})^2}{4(a_{k_u,n}^{(i)})^2} \right) \\ & \geq \frac{(x_{k_u,n} + \theta_{k_u,n})^2}{4a_{k_u,n}}, \forall k, u, n \quad (36) \end{aligned}$$

where $(\mathbf{a}^{(i)}, \mathbf{s}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)})$ is a feasible point of (33).

Hence, an approximate problem of (33) is

$$\underset{\mathbf{a}, \mathbf{s}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}}{\text{maximize}} \quad \sum_{k \in \mathcal{K}} w_k \eta_k \quad (37a)$$

$$\text{subject to} \quad \sum_{u \in \mathcal{U}_k} \sum_{n=1}^N x_{k_u,n} \geq \frac{y_k}{2} \eta_k^2 + \frac{1}{2y_k} g_k^2(\mathbf{s}), \forall k \quad (37b)$$

$$s_{k_u,n} h_{k,k_u,n} \geq \frac{\varphi_{k_u,n}}{2} (\theta_{k_u,n})^2 + \frac{1}{2\varphi_{k_u,n}} (\beta_{k_u,n}(\mathbf{s}))^2,$$

Algorithm 3 The proposed procedure solving (32).

- 1: **Initialization:** Set $i := 0$ and generate a feasible point of (33) denoted by $(\mathbf{a}^{(0)}, \mathbf{s}^{(0)}, \boldsymbol{\eta}^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{x}^{(0)})$.
- 2: **repeat**
- 3: Set $y_k^{(i)} := \frac{g_k(\mathbf{s}^{(i)})}{\eta_k^{(i)}}$ and $\varphi_{k_u, n}^{(i)} := \frac{\beta_{k_u, n}(\mathbf{s}^{(i)})}{\theta_{k_u, n}^{(i)}}$.
- 4: Solve (37) and denote the optimal point by $(\mathbf{a}^*, \mathbf{s}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$
- 5: Update $i := i + 1$, and $(\mathbf{a}^{(i)}, \mathbf{s}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)}) := (\mathbf{a}^*, \mathbf{s}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$
- 6: **until** Convergence
- 7: **Output:** $(\mathbf{a}^*, \mathbf{s}^*)$

$$\forall k, u, n \quad (37c)$$

$$(33e)-(33h), (36) \quad (37d)$$

where $\beta_{k_u, n}(\mathbf{s}) \triangleq \sum_{j \in \mathcal{K} \setminus \{k\}} \sum_{v \in \mathcal{U}_j} s_{j_v, n} h_{j, k_u, n} + \sigma_{k_u, n}^2$; $\{y_k\}_k$ and $\{\varphi_{k_u, n}\}_{k, u, n}$ are the approximated parameters.

Finally, the proposed procedure solving (32) is outlined in **Algorithm 3**. A feasible point for initializing the algorithm can be obtained by using the regularized form of (37) and the procedure similar to **Algorithm 2**. Also, output of **Algorithm 3** \mathbf{a}^* may be not Boolean. Therefore, the post processing phase presented in Section IV-B2 is performed to reconstruct the allocated power.

VI. PROPOSED SOLUTION TO FORMULATION 3

In this section, we present the solution for the third formulation. We note that the SCA framework cannot be directly applied to (12) due to the presence of ℓ_0 -norm in (12b), which causes (12) to be a combinatorial optimization program [49]. Here, following [50]–[52], we overcome the issue by regularizing the objective function with a penalty term of the reweighted ℓ_1 -norm, i.e.,

$$\underset{\mathbf{p}}{\text{maximize}} \quad f(\mathbf{p}) - \phi \left(\sum_{k \in \mathcal{K}} \sum_{n=1}^N \|\boldsymbol{\psi}_{k, n} \odot \mathbf{p}_{k, n}\|_1 \right) \quad (38a)$$

$$\text{subject to} \quad (6c)-(6e), (11g), (11h) \quad (38b)$$

where $\phi > 0$ is the regularization parameter determining a tradeoff between the desired objective and the degree of sparsity in the solution of \mathbf{p} ; and $\boldsymbol{\psi}_{k, n} \triangleq \{\psi_{k_u, n}\}_{u \in \mathcal{U}_k}$ is the nonnegative weight vector, which could improve the power reconstruction, if it is set suitably. One possible approach for setting the weights is using the power magnitude on the ℓ_1 -norm penalty function as [49, Eq. (5)]

$$\psi_{k_u, n} = \begin{cases} \frac{1}{p_{k_u, n}}, & p_{k_u, n} \neq 0 \\ \infty, & p_{k_u, n} = 0. \end{cases} \quad (39)$$

(39) means that the larger entry in $\mathbf{p}_{k, n}$ has the smaller weight. And, with the maximization in (38a), the small entries in $\mathbf{p}_{k, n}$ is expected to be forced to zero.

The penalty term in (38a) is convex, and the remaining term in (38a) as well as the constraints in (38) are the same as those in (6). Hence, we now can use the SCA technique to solve (38). In particular, for the WSEEMax problem, the convex

Algorithm 4 The proposed algorithm to solve (38).

- 1: **Initialization:** Set $i := 0$, $\psi_{k_u, n} := 1$ for all k, u, n , and generate a feasible point $(\mathbf{p}^{(0)}, \boldsymbol{\eta}^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{x}^{(0)})$.
- 2: **repeat**
- 3: Set $\varphi_{k_u, n}^{(i)} := \frac{\beta_{k_u, n}(\mathbf{p}^{(i)})}{\theta_{k_u, n}^{(i)}}$, $y_k := \frac{g_k(\mathbf{p}^{(i)})}{\eta_k^{(i)}}$
- 4: Solve (40), and denote the optimal solution by $(\mathbf{p}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$
- 5: Update $i := i + 1$, $(\mathbf{p}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)}) := (\mathbf{p}^*, \boldsymbol{\eta}^*, \boldsymbol{\theta}^*, \mathbf{x}^*)$, and $\psi_{k_u, n} := \begin{cases} \frac{1}{p_{k_u, n}^{(i)}}, & p_{k_u, n}^{(i)} \neq 0 \\ \infty, & p_{k_u, n}^{(i)} = 0. \end{cases}$ for all k, u, n .
- 6: **until** Convergence
- 7: **Output:** \mathbf{p}^*

approximation program solved in each SCA iteration is given by

$$\underset{\mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{x}}{\text{maximize}} \quad \sum_{k \in \mathcal{K}} w_k \eta_k - \phi \left(\sum_{k \in \mathcal{K}} \sum_{n=1}^N \|\boldsymbol{\psi}_{k, n} \odot \mathbf{p}_{k, n}\|_1 \right) \quad (40a)$$

$$\text{subject to} \quad (6c), (6e), (15f)-(15h), (22), (23b), (23c) \quad (40b)$$

The iterative procedure is presented in **Algorithm 4** where the weights $\{\psi_{k_u, n}\}_{u \in \mathcal{U}_k}$ are also updated in each iteration (i.e. reweighting). It is worth noting that, for stability in practical implementation, the weights can alternatively be updated as $\psi_{k_u, n} := \frac{1}{p_{k_u, n}^{(i)} + \tau}$ for all k, u, n where $\tau > 0$ is a tolerant parameter which should be sufficiently small [49]. On the other hand, in case the output of **Algorithm 4** does not satisfy (12b), the post processing phase is conducted to determine the allocated power.⁴

The similar procedure is used to tackle the NEEmax problem in formulation (12) which is skipped for the sake of brevity.

VII. CONVERGENCE AND COMPLEXITY ANALYSIS

In this section, we discuss the convergence property as well as the computational complexity of the proposed methods developed in the previous sections.

A. Convergence Discussion

We first focus on **Algorithm 1**. Following the same arguments as those in [53], it can be shown that the obtained objective sequence $\{\sum_{k \in \mathcal{K}} w_k \eta_k^{(i)}\}_{i=0}^{\infty}$ is nondecreasing and guaranteed to converge. More explicitly, consider problem (23) in the $(i + 1)$ th iteration, it can be easily examined that all the constraints in the problem are satisfied by the point $(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\eta}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{x}^{(i)})$, the solution of (23) in the i th iteration. Thus, we have $\sum_{k \in \mathcal{K}} w_k \eta_k^{(i)} \leq \sum_{k \in \mathcal{K}} w_k \eta_k^{(i+1)}$. Moreover, the problem is bounded above due to the power constraints in (6c). Hence, $\{\sum_{k \in \mathcal{K}} w_k \eta_k^{(i)}\}_{i=0}^{\infty}$ converges. The same justification can be applied to **Algorithm 3** pointing out that sequence $\{\sum_{k \in \mathcal{K}} w_k \eta_k^{(i)}\}_{i=0}^{\infty}$ obtained by this algorithm is nondecreasing and converges.

⁴The mapping step is based on \mathbf{p} instead of \mathbf{a} as in **Algorithm 1**.

For **Algorithm 4**, let us recall that the function $\pi(\mathbf{p}) \triangleq \sum_{k \in \mathcal{K}} \sum_{u \in \mathcal{U}_k} \sum_{n=1}^N \log(p_{k_u, n} + \tau)$ is concave and monotonic increasing. Thus, in the $(i+1)$ th iteration, the term $\pi(\mathbf{p}^{(i)}) + \sum_{k \in \mathcal{K}} \sum_{n=1}^N \left\| \psi_{k,n} \odot (\mathbf{p}_{k,n} - \mathbf{p}_{k,n}^{(i)}) \right\|_1$ is indeed a convex upper bound of $\pi(\mathbf{p})$ at $\mathbf{p}^{(i)}$. Thus, following the convergence results of the SCA technique, we have the sequence $\{\sum_{k \in \mathcal{K}} w_k \eta_k^{(i)} - \phi \pi(\mathbf{p}^{(i)})\}_{i=0}^{\infty}$ is nondecreasing and converges since $(\sum_{k \in \mathcal{K}} w_k \eta_k - \phi \pi(\mathbf{p}))$ is upper bounded. This property is used to stop **Algorithm 4**.

We note that the proposed algorithm is guaranteed to converge and they can only provide a locally optimal solution, which may not be a globally optimal due to the nonconvexity of the considered problem.

B. Computational Complexity Discussion

We now discuss the computational cost of the proposed methods. In particular, we provide the worst-case complexity of a general interior point method solving SOCPs in each of iterations based on the complexity estimates in [44, Chap. 6].

Let us denote by $U = \sum_{k \in \mathcal{K}} U_k$ the total number of users in the network. In problem (23), there are $4NU + K + 1$ real variables, $NU + K + 1$ SOC constraints of dimension 4, NU SOC constraints of dimension 2, and $(3N+1)U + (K+1)(N+1) + \sum_{k \in \mathcal{K}} |\mathcal{U}_k^T|$ linear constraints. Hence, the worst-case computational cost for solving (23) is $\mathcal{O}(((5N+1)U + (K+1)(N+2) + \sum_{k \in \mathcal{K}} |\mathcal{U}_k^T|)^{0.5} (4NU + K + 1)^2 ((9N+1)U + (K+1)(N+5) + \sum_{k \in \mathcal{K}} |\mathcal{U}_k^T|))$, which can be rewritten in a simple form as $\mathcal{O}(N^{1.5} (5U + K)^{0.5} (4NU + K)^2 (9U + K))$. This is also the worst-case computational cost for solving (37) since the two SOCPs have the same dimensions.

Problem (40) includes $3NU + K + 1$ real variables, $NU + K + 1$ SOC constraints of dimension 4, NU SOC constraints of dimension 2, and $(N+1)U + K + 1 + \sum_{k \in \mathcal{K}} |\mathcal{U}_k^T|$ linear constraints. So, the worst-case computational cost for solving (40) is $\mathcal{O}(((3N+1)U + 2(K+1) + \sum_{k \in \mathcal{K}} |\mathcal{U}_k^T|)^{0.5} (3NU + K + 1)^2 ((7N+1)U + 5(K+1) + \sum_{k \in \mathcal{K}} |\mathcal{U}_k^T|))$, which can be rewritten in a simple form as $\mathcal{O}((3NU + 2K)^{0.5} (3NU + K)^2 (7NU + 5K))$. With smaller numbers of variables and constraints, the worst-case computational complexity of solving (40) is lower compared to (23) and (37). The proposed algorithms show much lower complexity when comparing with the exhaustive search for RB allocation, which has the worst-case computational cost of $\mathcal{O}(\sum_{k \in \mathcal{K}} U_k^N)$ [40].

VIII. NUMERICAL RESULTS

In this section, we numerically investigate the EE performances of the OFDMA-based HetNets. In the simulation model, the macro BS covers a circle area with radius of 500m centered at the origin. The femto BSs are randomly placed inside the coverage area of macro BS, and the minimum distance between femto BSs is 5m. Each femtocell covers a circle area with radius of 20m centered at its BS. The minimum distance between the users and the macro BS is 40m. In each femtocell, the minimum distance between a user and the femto BS is 3m.

The channel coefficient between BS j and user k_u on RB n is modeled as $\sqrt{\Gamma_{j,k_u}} \chi_{j,k_u,n}$ where $\chi_{j,k_u,n}$ is a complex Gaussian distribution variable with zero mean and unit variance, i.e. $\chi_{j,k_u,n} \sim \mathcal{CN}(0, 1)$, and Γ_{j,k_u} composes of the shadowing fading and path loss. In particular, the path loss between a femto BS and a femto user follows the indoor model and that between the macro BS and a macro user follows the outdoor model, which are given by [54]

$$\begin{cases} 127 + 30 \log_{10}(d) \text{ [dB]} & \text{for indoor path loss} \\ 128.1 + 37.6 \log_{10}(d) \text{ [dB]} & \text{for outdoor path loss} \end{cases} \quad (41)$$

where d is the distance in meters. The outdoor wall penetration loss is 10 dB. The log normal shadowing standard deviation is 8.

The number of RBs is taken as $N = 50$, each of RBs has a bandwidth of 180 KHz. So, the bandwidth of system is 10 MHz. These parameters correspond to LTE systems [55].⁵ The noise power density is -174 dBm/Hz [56]. The maximum transmit power of macro BS is taken as $P_0^{\max} = 46$ dBm [54]. QoS of all DS users are the same. Without loss of generality, we take the amplifier efficiency as $\xi_k = 1, \forall k \in \mathcal{K}$, for simplicity. Also, for WSEEmax problem, we take $w_k = 1, \forall k \in \mathcal{K}$, i.e. all BSs have the same EE priority. The fairness coefficient $\alpha_f = 0.01$.

For solving convex programs in this section, we use the MATLAB modeling toolbox YALMIP [57] with internal solver MOSEK [58].

In Fig. 2, we numerically examine the convergence behavior of the proposed algorithms for the two EE problems. The results of the WSEEmax are shown in Fig. VIII and that of the NEEmax are shown in Fig. VIII. We note that the results include the procedure of finding feasible points. Therefore, in some schemes, there are negative values at some first iterations due to the large value of the penalty terms. We can clearly observe that the proposed procedures require less than 15 iterations to find feasible points and converge. The results indicate that our solutions achieve fast converge.

In Fig. 3, we study the variation of the EE performances of WSEEmax and NEEmax problems with the maximum transmit power at femto BSs. The major observation is that, in all cases of EE problems and formulations, the EE performances first increase then are unchanged when P_k^{\max} increases. The result can be mathematically explained as follows. When P_k^{\max} is small, the corresponding constraints has strong impact on performances, and thus increasing P_k^{\max} leads to the performance increase. However, when P_k^{\max} is large enough, the other constraints dominate the performances. So further increasing P_k^{\max} does not change the performances. For the system design, the results indicate that providing more transmit power budget for femtocell BSs does not always bring additional EE benefits, and thus suggest that the BSs with reasonable maximum transmit power should be used for saving the implementation cost. Another observation is that

⁵If the number of RBs is varied, the same trends in the simulation results are expected. However, we note that if the number of RBs is large, then more frequency diversity gains can be obtained. On the other hand, the complexity of the proposed algorithms increases accordingly.

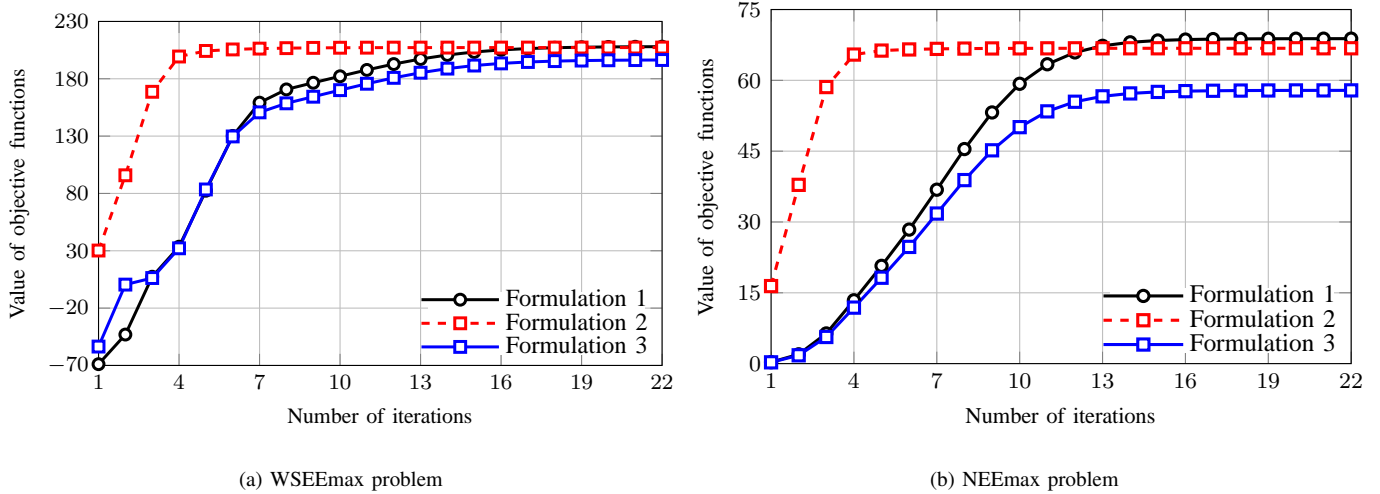


Fig. 2. Convergence performances of the proposed algorithms over a random channel realization. The network setting is taken as $K = 2$, $R_{k_u}^{\min} = 18$ Kbps $\forall k, u \in \mathcal{U}_k^S$, $P_k^{\max} = 24$ dBm $\forall k \in \mathcal{K} \setminus \{0\}$, $P_k^{\text{sta}} = 30$ dBm and $U_k = 4$, $\forall k \in \mathcal{K}$.

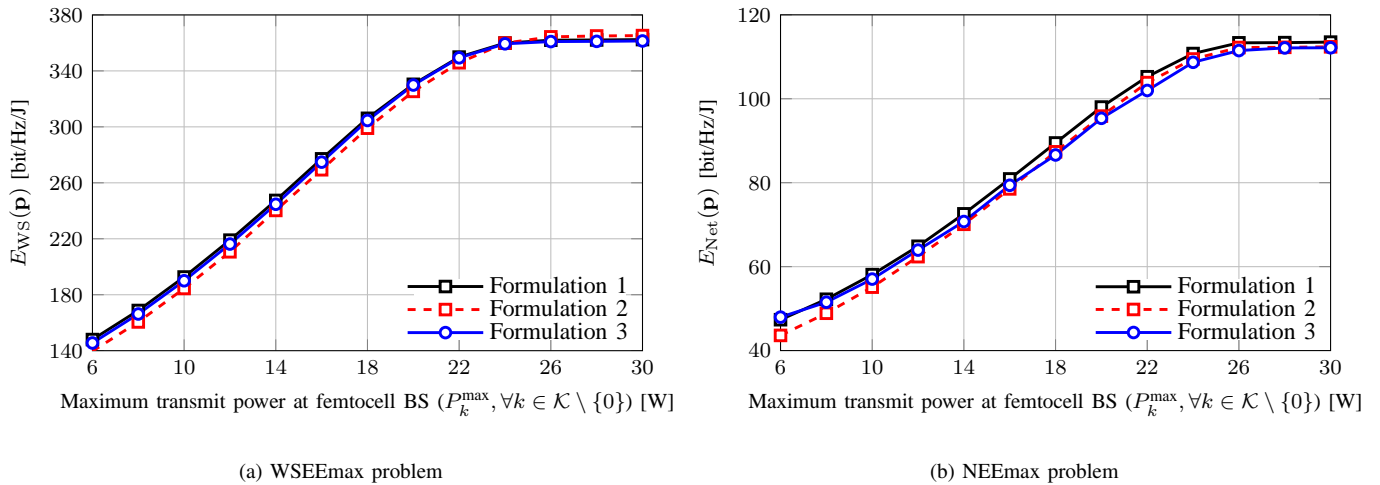


Fig. 3. The average EE performance of the two EE problems with the proposed algorithms versus the maximum transmit power at femtocell BSs, i.e. P_k^{\max} , $\forall k \in \mathcal{K} \setminus \{0\}$. The network setting is taken as $K = 2$, $R_{k_u}^{\min} = 18$ Kbps $\forall k, u \in \mathcal{U}_k^S$, $P_k^{\text{sta}} = 30$ dBm and $U_k = 4$ $\forall k \in \mathcal{K}$.

the performances of the three formulations are close together. Compared to **Formulations 1** and **3**, **Formulation 2** is slightly lower in the low regime of P_k^{\max} , but achieves good EE performances in the high regime of P_k^{\max} . **Formulations 1** and **3** are almost same together in WSEEmax problem while in WSEEmax problem, **Formulations 3** is slightly lower than **Formulations 1**.

In Fig. 4, we study the EE performances as functions of the static consumed power P_k^{sta} . We can observe from the figure is that, for the two EE problems, when P_k^{sta} increases, the EE performances decrease. The observation is reasonable since P_k^{sta} is a part of the denominators in $E_{\text{WS}}(\mathbf{p})$ and $E_{\text{Net}}(\mathbf{p})$. The result promotes a direction of designing hardware elements that is reducing the circuit operation power as much as possible since doing so always achieves better EE performances.

Fig. 5 plots the EE performances of the proposed EE schemes versus QoS of DS users $R_{k_u}^{\min}$. Again, all solutions behave in the same manner that is the performances degrade when $R_{k_u}^{\min}$ increases. This is reasonable since more power should be allocated to DS users, which could be under poor channel conditions, in order to satisfy the QoS constraints. Another observation is that, when $R_{k_u}^{\min}$ is large, **Formulation 1** outperforms the other formulations and the gap between them increases with $R_{k_u}^{\min}$. The result suggests using **Formulation 1** for the scenario with large QoS.

In Fig. 6, we investigate the relation between EE performances of the proposed algorithms with the number of femtocell BSs K . In case of WSEEmax, we can observe from Fig. 6(a) that the performance of all considered schemes increase when K increases since the number of elements in the summation of $E_{\text{WS}}(\mathbf{p})$ increases. However, the average

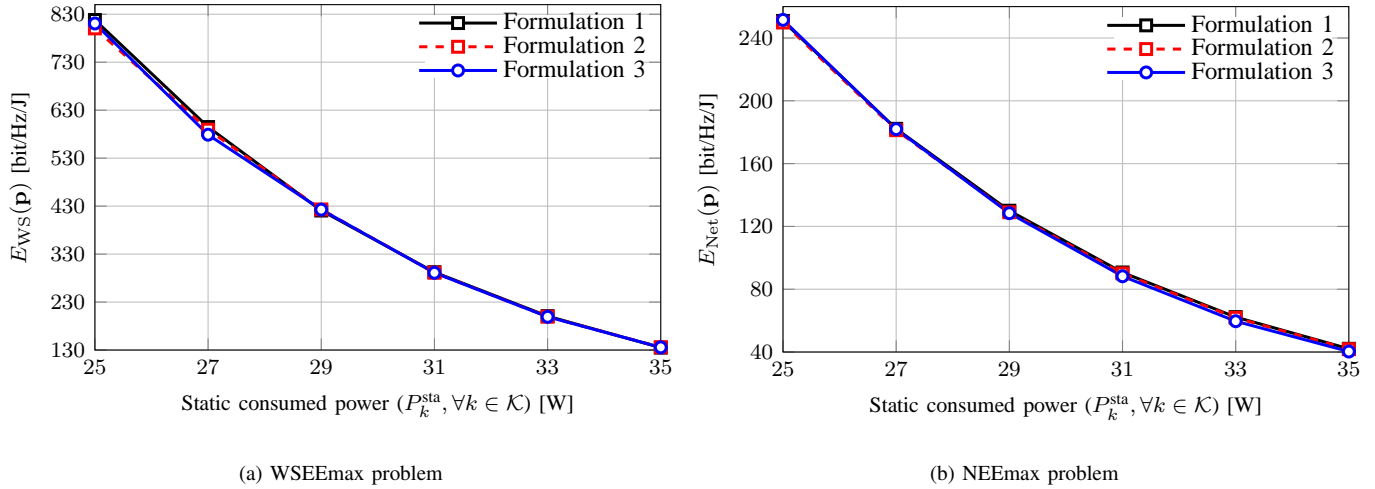


Fig. 4. The average EE performance of the two EE problems with the proposed algorithms versus the static consumed power $P_k^{sta}, \forall k \in \mathcal{K}$. The network setting is taken as $K = 2$, $R_{k_u}^{min} = 18$ Kbps $\forall k, u \in \mathcal{U}_k^S$, $P_k^{max} = 24$ dBm $\forall k \in \mathcal{K} \setminus \{0\}$, and $U_k = 4 \forall k \in \mathcal{K}$.

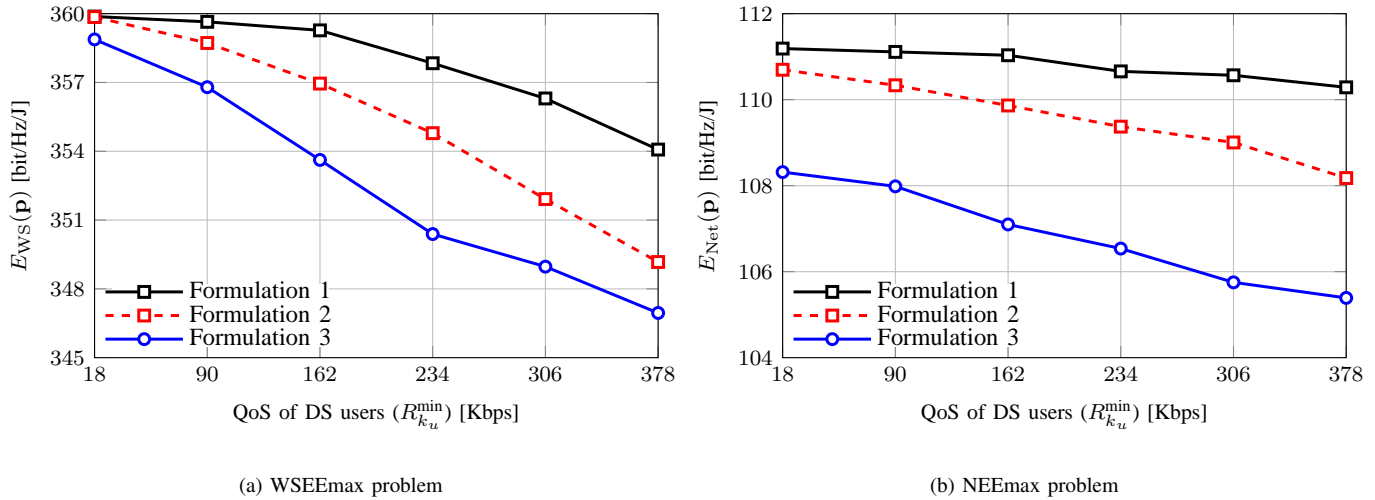


Fig. 5. The average EE performance of the two EE problems with the proposed algorithms versus the QoS of DS users $R_{k_u}^{min}, \forall k, u \in \mathcal{U}_k^S$. The network setting is taken as $K = 2$, $P_k^{max} = 24$ dBm $\forall k \in \mathcal{K} \setminus \{0\}$, $P_k^{sta} = 30$ dBm and $U_k = 4 \forall k \in \mathcal{K}$.

EE of individual cell, i.e. $E_{WS}(\mathbf{p})/K$, of the EE schemes decrease with K . This is because more femtocells results in more interference signals. For NEEMax problem, we can see that the performance first increases until a maximum then decreases with K . This result can be explained as follows. When K is small, there are few sources of inter-cell interference. Thus the performance increases with K due to the diversity gain. However, when K is large, the cells are close together making the interference signal strong. In addition, the total static consumed power also increases with K . These lead to the reduction of the performance. For the system design, the results in Fig. 6 imply that the number of active femtocells should not be too large in order to achieve good EE performances. Another important observation is that the gap between Formulation 1,2 and Formulation 3 becomes bigger when the size of the network increases. This result

indicates the best option of selecting Formulation 1 and 2, instead of Formulation 3 in cases of large number of femtocell BSs.

In Fig. 7, we report the average total run time of Algorithms 1,3,4 to converge as functions of number of RB. The results for WSEEmax and NEEmax are shown in Figs. 7(a) and 7(b), respectively. For WSEEmax, the total run time of the three formulations are almost same. On the other hand, for NEEmax, they are different where the Formulation 2 needs highest total run time and the Formulation 3 need lowest total run time. Another observation is that in all cases the run time increases with the number of resource block. This is because the size of the problems increase.

In the final set of experiments shown in Fig. 8, we demonstrate the benefit in terms of EE of our approaches by comparing to the performance of the sum rate maximization (SRmax)

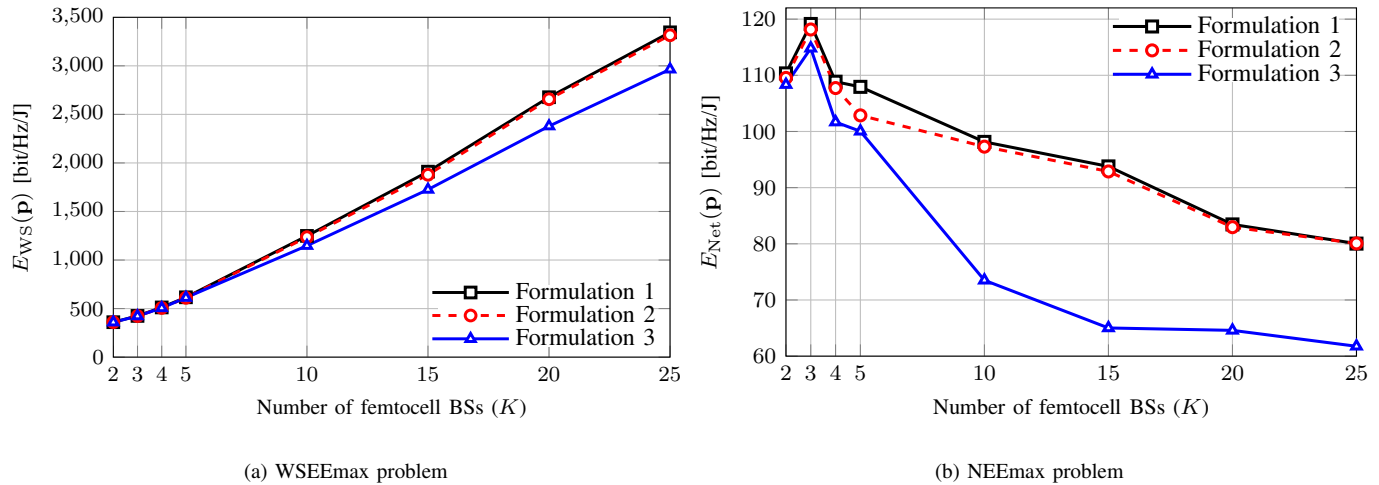


Fig. 6. The average EE performance of the two EE problems with the proposed algorithms versus the number of femtocell BSs K . The network setting is taken as $R_{k_u}^{\min} = 18$ Kbps $\forall k, u \in \mathcal{U}_k^S$, $P_k^{\max} = 24$ dBm $\forall k \in \mathcal{K} \setminus \{0\}$, $P_k^{\text{sta}} = 30$ dBm and $U_k = 4$ $\forall k \in \mathcal{K}$.

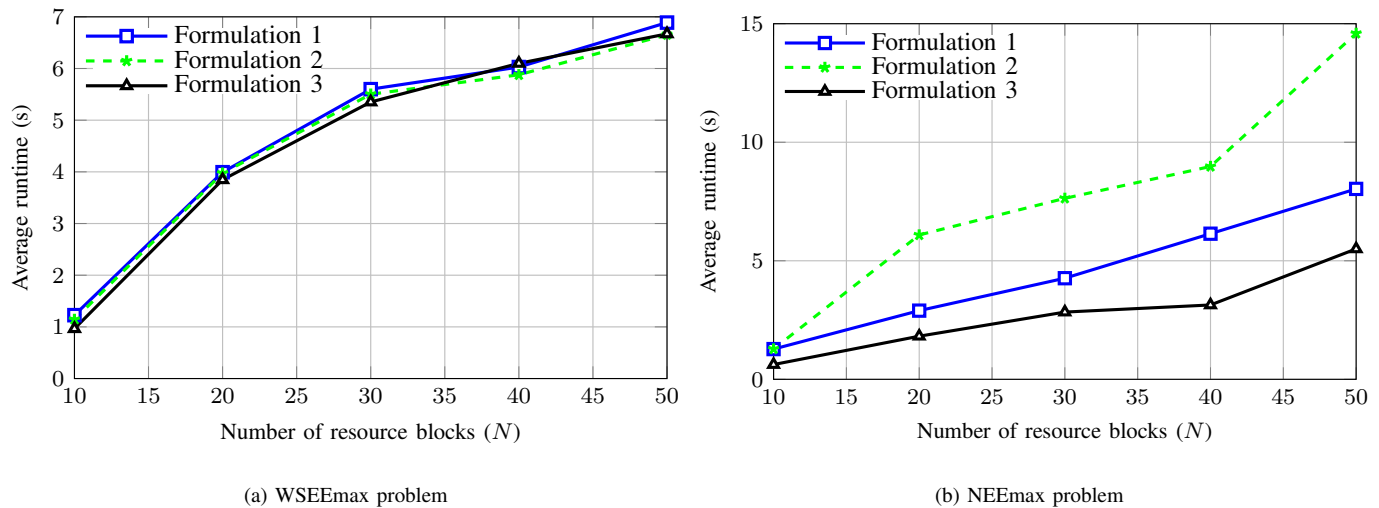


Fig. 7. The average run time to converge of Algorithms 1, 3, and 4 of the two EE problems versus the number of RB N . The network setting is taken as $R_{k_u}^{\min} = 18$ Kbps $\forall k, u \in \mathcal{U}_k^S$, $P_k^{\max} = 24$ dBm $\forall k \in \mathcal{K} \setminus \{0\}$, $P_k^{\text{sta}} = 30$ dBm and $U_k = 4$, $K = 2$ $\forall k \in \mathcal{K}$.

scheme in [8] which considered the scenario similar to that in this paper⁶. We can see that the gap between our EE approaches and SRmax is remarkably large in all cases of considered K , which confirms the validity of our approach in terms of EE performances. This gap can be explained as follows. The SRMax scheme is to maximize the spectral-efficiency and thus transmits at full power. Recall that operating at the high power regime is not energy-efficient. Meanwhile, the proposed algorithms aim to strike a balance between SE and the total power consumption to maximize the energy efficiency of the whole network. Therefore, the energy efficiency of the proposed solutions outperform than the SRmax scheme by a large margin.

⁶We consider SRmax scheme because there is no work proposing EE solution for the considered system. In this situation, the comparison between EE and SE schemes are commonly used [32], [56]

Throughout extensive simulation results presented above, it can be seen that there is always a trade-off between the achieved energy efficiency performance and the required complexity among the proposed formulations and their associated solutions. For example, solutions based on Formulation 1 generally have the best energy efficiency performance but they incur the most complexity. In contrast, Formulation 3 achieves the worst energy efficiency performance but is the most computationally efficient. Formulation 2 more or less stands in between. Therefore, for some network setting where the performance of the three formulations are almost similar (e.g. Fig. 3), we can choose Formulation 3 for computational efficiency. However, when the performance of these formulations are different to each other (e.g. Fig. 5), the trade-off should be considered. Thus, generally, it is hard to judge which of them is the best solution. The choice of any of the

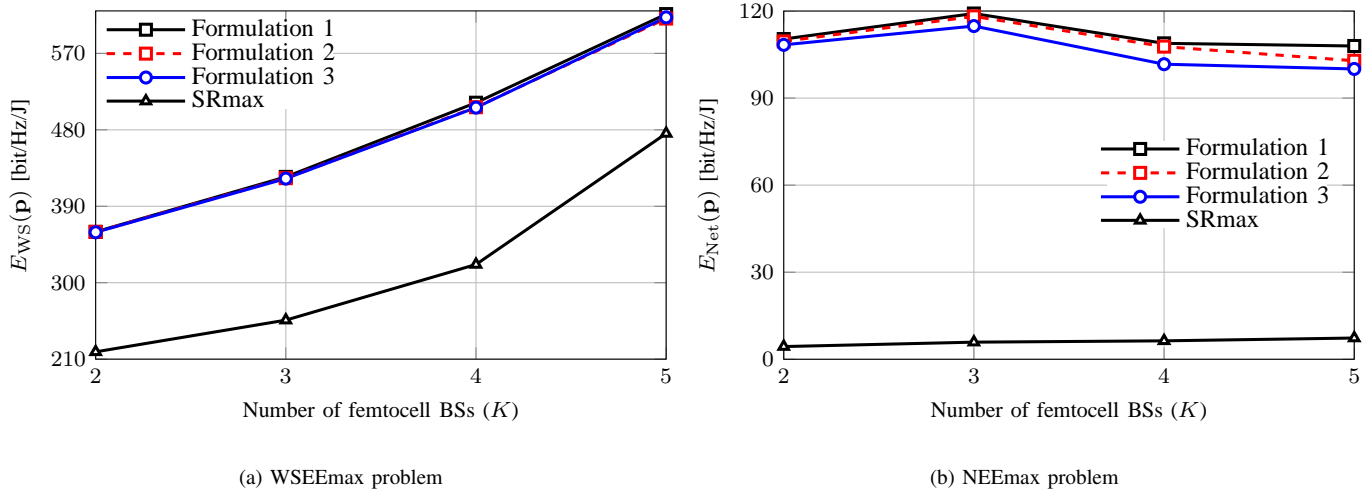


Fig. 8. The average EE performance of the two EE problems with the proposed algorithms and the SRmax scheme in [8] versus the number of femtocell BSs K . The network setting is taken as $R_{k_u}^{\min} = 18$ Kbps $\forall k, u \in \mathcal{U}_k^S$, $P_k^{\max} = 24$ dBm $\forall k \in \mathcal{K} \setminus \{0\}$, $P_k^{\text{sta}} = 30$ dBm and $U_k = 4$ $\forall k \in \mathcal{K}$.

proposed methods depends on the network situations and or the network operators. In terms of energy efficiency metrics, the metric NEEMax is the choice when performance of the overall network is concerned. However, in the scenario where the priority and/or fairness between BSs's performance should be guaranteed, the metric WSEEMax is more favorable which controls the priority and fairness by adjusting the weighting factor.

IX. CONCLUSION

We have studied the energy-efficient resource allocation in OFDMA HetNets. In particular, we have considered both metrics of WSEE and NEE under the constraints of minimum data rate for DS users, rate fairness among DT users, and BSs' power budget. First, the interest problems have been formulated based on three different approaches those are big-M method, time-sharing technique, and sparsity inducing norm. For each formulation, we have developed an efficient optimization algorithm tackling the nonconvexity based on the framework of SCA. Specifically, the proposed methods only requires solving SOCP, whose solving cost is less sensitive to problem size compared to other convex structures, and thus are suitable for the considered system where the number of variables is usually large. Finally, the numerical results have been provided and analyzed. With the consideration of three formulations of two definitions of energy efficiency, the paper has provided a comprehensive insight and suggested some useful guidelines for OFDM HetNets design in terms of EE.

REFERENCES

- [1] J. G. Andrews, "Seven ways that hetnets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, 2013.
- [2] S. Mishra and C. S. R. Murthy, "Increasing energy efficiency via transmit power spreading in dense femto cell networks," *IEEE Systems Journal*, 2017.
- [3] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sept. 2008.

- [4] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, 2011.
- [5] G. Mansfield, "Femtocells in the US market-business drivers and consumer propositions," *FemtoCells Europe*, pp. 1927–1948, 2008.
- [6] J. Zhang and G. De la Roche, *Femtocells: technologies and deployment*. Wiley Online Library, 2010.
- [7] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE J. Select. Areas. Commun.*, vol. 30, no. 3, pp. 497–508, 2012.
- [8] D. T. Ngo, S. Khakurel, and T. Le-Ngoc, "Joint subchannel assignment and power allocation for OFDMA femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 342–355, Jan. 2014.
- [9] S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5g networks and challenges ahead," *IEEE J. Select. Areas. Commun.*, vol. 34, no. 4, pp. 697–709, 2016.
- [10] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink ofdma networks with base station coordination," *IEEE transactions on wireless communications*, vol. 14, no. 1, pp. 1–14, 2015.
- [11] H. Klessig, D. Öhmann, A. I. Reppas, H. Hatzikirou, M. Abedi, M. Simsek, and G. P. Fettweis, "From immune cells to self-organizing ultra-dense small cell networks," *IEEE J. Select. Areas. Commun.*, vol. 34, no. 4, pp. 800–811, 2016.
- [12] B. Zhuang, D. Guo, and M. L. Honig, "Energy-efficient cell activation, user association, and spectrum allocation in heterogeneous networks," *IEEE J. Select. Areas. Commun.*, vol. 34, no. 4, pp. 823–831, 2016.
- [13] E. Björnson, L. Sanguinetti, and M. Kountouris, "Deploying dense networks for maximal energy efficiency: Small cells meet massive MIMO," *IEEE J. Select. Areas. Commun.*, vol. 34, no. 4, pp. 832–847, 2016.
- [14] Z. Pan and Q. Zhu, "Energy efficiency optimization in 3-D small cell networks-based sleep strategy," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1131–1134, 2017.
- [15] S. D'Oro, A. Zappone, S. Palazzo, and M. Lops, "A learning approach for low-complexity optimization of energy efficiency in multicarrier wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3226–3241, 2018.
- [16] J. Chen, H. Zhuang, and Z. Luo, "Energy optimization in dense OFDM networks," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 189–192, 2016.
- [17] X. C. S. Z. N. P. W. Lu Liu, Yu Cheng, "Joint optimization of scheduling and power control in wireless networks: Multi-dimensional modeling and decomposition," *IEEE Transactions on Mobile Computing*, 2018.
- [18] Y. Li, M. Sheng, C. W. Tan, Y. Zhang, Y. Sun, X. Wang, Y. Shi, and J. Li, "Energy-efficient subcarrier assignment and power allocation in OFDMA systems with max-min fairness guarantees," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3183–3195, Sept. 2015.

- [19] X. Xiao, X. Tao, and J. Lu, "Energy-efficient resource allocation in LTE-based MIMO-OFDMA systems with user rate constraints," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 185–197, Jan. 2015.
- [20] A. Zappone, E. Jorswieck, and A. Leshem, "Distributed resource allocation for energy efficiency in MIMO OFDMA wireless networks," *IEEE J. Select. Areas. Commun.*, vol. 34, no. 12, pp. 3451–3465, 2016.
- [21] C. C. Zarakovitis, Q. Ni, and J. Spiliotis, "Energy-efficient green wireless communication systems with imperfect CSI and data outage," *IEEE J. Select. Areas. Commun.*, vol. 34, no. 12, pp. 3108–3126, 2016.
- [22] Q. Xu, X. Li, H. Ji, and X. Du, "Energy-efficient resource allocation for heterogeneous services in OFDMA downlink networks: systematic perspective," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2071–2082, June 2014.
- [23] R. Estrada, A. Jarray, H. Otrok, Z. Dziong, and H. Barada, "Energy-efficient resource-allocation model for ofdma macrocell/femtocell networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3429–3437, 2013.
- [24] K. Yang, S. Martin, D. Quadri, J. Wu, and G. Feng, "Energy-efficient downlink resource allocation in heterogeneous OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5086–5098, 2017.
- [25] A. Saeed, E. Katranaras, A. Zoha, A. Imran, M. A. Imran, and M. Dianati, "Energy efficient resource allocation for 5G heterogeneous networks," in *2015 IEEE 20th Int. Works. Comput. Aided Modelling and Design of Commun. Links and Networks (CAMAD)*.
- [26] W. Li, W. Zheng, H. Zhang, T. Su, and X. Wen, "Energy-efficient resource allocation with interference mitigation for two-tier OFDMA femtocell networks," in *2012 IEEE 23rd Int. Symp. Personal Indoor and Mobile Radio Commun. (PIMRC)*.
- [27] Y.-L. Chung, "An efficient power-saving transmission mechanism in LTE macrocell-femtocell hybrid networks," in *2014 Int. Conf. Inf. Networking (ICOIN)*.
- [28] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Trans. Signal Processing*, vol. 22, no. 7, pp. 804–808, 2015.
- [29] M. Tao, Y.-C. Liang, and F. Zhang, "Resource allocation for delay differentiated traffic in multiuser OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2190–2201, June 2008.
- [30] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.
- [31] S. Sesia, M. Baker, and I. Toufik, *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.
- [32] Q.-D. Vu, L.-N. Tran, M. J. Juntti, and E.-K. Hong, "Energy-efficient bandwidth and power allocation for multi-homing networks," *IEEE Trans. Signal Processing*, vol. 63, no. 7, pp. 1684–1699, 2015.
- [33] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Found. Trends Commun. Inf. Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.
- [34] S. He, Y. Huang, L. Yang, and B. Ottersten, "Coordinated multicell multiuser precoding for maximizing weighted sum energy efficiency," *IEEE Trans. Signal Processing*, vol. 62, no. 3, pp. 741–751, 2014.
- [35] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 1–14, 2015.
- [36] 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, 2013.
- [37] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618–3631, 2012.
- [38] J. Lee and S. Leyffer, *Mixed Integer Nonlinear Programming*. New York, USA: Springer, 2011.
- [39] 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, 2012.
- [40] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2366–2377, July 2014.
- [41] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Foundations and Trends® in Machine Learning*, vol. 4, no. 1, pp. 1–106, 2012.
- [42] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations research*, vol. 26, no. 4, pp. 681–683, 1978.
- [43] A. Beck, A. Ben-Tal, , and L. "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, pp. 29–51, 2010.
- [44] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*. Soc. Ind. Appl. Math. (SIAM), 2001.
- [45] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Network.*, 2017.
- [46] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Mathematical programming*, vol. 95, no. 1, pp. 3–51, 2003.
- [47] T. Lipp and S. Boyd, "Variations and extension of convex-concave procedure," *Optimization and Engineering*, no. 3, pp. 1–25, 2015.
- [48] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2004.
- [49] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5, pp. 877–905, 2008.
- [50] O. Mehanna, N. D. Sidiropoulos, and G. B. Giannakis, "Joint multicast beamforming and antenna selection," *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2660–2674, 2013.
- [51] F. Rinaldi, "New results on the equivalence between zero-one programming and continuous concave programming," *Optimization Letters*, vol. 3, no. 3, pp. 377–386, 2009.
- [52] F. Bach, R. Jenatton, J. Mairal, G. Obozinski *et al.*, "Structured sparsity through convex optimization," *Statistical Science*, vol. 27, no. 4, pp. 450–468, 2012.
- [53] L.-N. Tran, M. F. Hanif, A. Tolli, and M. Juntti, "Fast converging algorithm for weighted sum rate maximization in multicell MISO downlink," *IEEE Signal Process Lett.*, vol. 19, no. 12, pp. 872–875, Dec. 2012.
- [54] *Further advancements for E-UTRA physical layer aspects*, 3GPP TR 36.814 V9.2.0 March 2017.
- [55] T. Innovations, "LTE in a Nutshell: The Physical Layer [PDF online]," [https://home.zhaw.ch/kunr/NTM1/literatur/LTEin a Nutshell: The Physical Layer.pdf](https://home.zhaw.ch/kunr/NTM1/literatur/LTEin%20a%20Nutshell%20The%20Physical%20Layer.pdf), 2010.
- [56] Q.-D. Vu, L.-N. Tran, R. Farrell, and E.-K. Hong, "Energy-efficient zero-forcing precoding design for small-cell networks," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 790–804, 2016.
- [57] J. Lofberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *2004 IEEE Int. Symp. Comput. Aided Control Syst. Design*.
- [58] A. Mosek, "The mosek optimization toolbox for matlab manual," 2015.



Nam-Tran Le received the B.S. degree in Electronic and Telecommunication Engineering from Danang University of Technology, Danang City, Vietnam, in 2006, the M.S. degree in Communications and Computer Engineering from Tokyo Institute of Technology, Tokyo, Japan in 2011 and PhD degree in Electrical Engineering and Computer Science from Queensland University of Technology, Australia in 2019. Her current research interests include resource allocation, spectral and energy efficiency communications, OFDMA femtocell heterogeneous networks and applications of optimization techniques.

PLACE
PHOTO
HERE

Le-Nam Tran (M'10–SM'17) received the B.S. degree in electrical engineering from Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam, in 2003, and the M.S. and Ph.D. degrees in radio engineering from Kyung Hee University, Seoul, Korea, in 2006 and 2009, respectively. He is currently a Lecturer/Assistant Professor at the School of Electrical and Electronic Engineering, University College Dublin, Ireland. Prior to this, he was a Lecturer at the Department of Electronic Engineering, Maynooth University, Ireland. From 2010

to 2014, he had held postdoc positions at the Signal Processing Laboratory, ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden (2010-2011), and at Centre for Wireless Communications, University of Oulu, Finland (2011-2014). His research interests are mainly on applications of optimization techniques on wireless communications design. Some recent topics include energy-efficient communications, physical layer security, cloud radio access networks, massive MIMO, and full-duplex transmission. He has authored or co-authored in more than 90 papers published in international journals and conference proceedings.

Dr. Tran is an Associate Editor of EURASIP Journal on Wireless Communications and Networking. He was Symposium Co-Chair of Cognitive Computing and Networking Symposium of International Conference on Computing, Networking and Communication (ICNC 2016). He was awarded the Career Development Award from Science Foundation Ireland (SFI).

PLACE
PHOTO
HERE

Quang-Doanh Vu received the B.S. degree in Electrical Engineering from Ho Chi Minh National University of Technology, Ho Chi Minh City, Vietnam, in 2010, and the M.S. and PhD degrees both in Radio Engineering from Kyung Hee University, Republic of Korea, in 2012 and 2015, respectively. Since October 2015, he has been with the Centre for Wireless Communications, University of Oulu, Finland. His current research interests include resource allocation, energy-efficient communications, multiuser MIMO systems, wireless power transfer.

PLACE
PHOTO
HERE

Dhammika Jayalath received the B.Sc. degree in electronics and telecommunications engineering from the University of Moratuwa, Sri Lanka, the M.Eng. degree in telecommunications from the Asian Institute of Technology, Thailand, and the Ph.D. degree in wireless communications from Monash University, Australia, in 2002. He was a Fellow at the Australian National University and a Senior Researcher at the National ICT Australia. He has been an Academic with the Science and Engineering Faculty, Queensland University of Technol-

ogy since 2007. His research interests include the general areas of communications and signal processing. He has published significantly in these areas. His current research interests include cooperative communications, cognitive radios, statistical signal processing, and multiuser communications. He is a Senior Member of the IEEE.