



Energy-Efficient Heterogeneous Cellular Networks with Spectrum Underlay and Overlay Access

DOI:
[10.1109/TVT.2017.2773506](https://doi.org/10.1109/TVT.2017.2773506)

Document Version
Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):
Tang, J., So, D. K. C., Alsusa, E., Hamdi, K., Shojaeifard, A., & Wong, K-K. (2017). Energy-Efficient Heterogeneous Cellular Networks with Spectrum Underlay and Overlay Access. *IEEE Transactions on Vehicular Technology*, 67(3), 2439-2453. <https://doi.org/10.1109/TVT.2017.2773506>

Published in:
IEEE Transactions on Vehicular Technology

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



Energy-Efficient Heterogeneous Cellular Networks with Spectrum Underlay and Overlay Access

Jie Tang, *Member, IEEE*, Daniel K. C. So, *Senior Member, IEEE*, Emad Alsusa, *Senior Member, IEEE*, Khairi Ashour Hamdi, *Senior Member, IEEE*, Arman Shojaeifard, *Member, IEEE*, and Kai-Kit Wong, *Fellow IEEE*

Abstract—In this paper, we provide joint subcarrier assignment and power allocation schemes for quality-of-service (QoS)-constrained energy-efficiency (EE) optimization in the downlink of an orthogonal frequency division multiple access (OFDMA)-based two-tier heterogeneous cellular network (HCN). Considering underlay transmission, where spectrum-efficiency (SE) is fully exploited, the EE solution involves tackling a complex mixed-combinatorial and non-convex optimization problem. With appropriate decomposition of the original problem and leveraging on the quasi-concavity of the EE function, we propose a dual-layer resource allocation approach and provide a complete solution using difference-of-two-concave-functions approximation, successive convex approximation and gradient-search method. On the other hand, the inherent inter-tier interference from spectrum underlay access may degrade EE particularly under dense small-cell deployment and large bandwidth utilization. We therefore develop a novel resource allocation approach based on the concepts of spectrum overlay access and resource efficiency (RE) (normalized EE-SE trade-off). Specifically, the optimization procedure is separated where the macro-cell optimal RE and the corresponding bandwidth is first determined, then the EE of small-cells utilizing the remaining spectrum is maximized. Simulation results confirm the theoretical findings and demonstrate that the proposed resource allocation schemes can approach the optimal EE with each strategy being superior under certain system settings.

Index Terms—Heterogeneous cellular network (HCN), orthogonal frequency division multiple access (OFDMA), energy-efficiency (EE), spectrum-efficiency (SE), resource efficiency (RE).

I. INTRODUCTION

The global mobile data traffic, thanks largely to the ever-growing use of applications on smart devices, increased by a tremendous 4k times in a decade from 2005 to 2015 and is expected to further grow going into 2020 and beyond [1]. It is well-understood that the conventional cellular network architecture using macro-cells only cannot possibly support demand going forward. This trend has driven the wireless

J. Tang is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou and with the State Key Laboratory of Integrated Services Networks, Xidian University, China. (e-mail: eejtang@scut.edu.cn).

D. K. C. So, E. Alsusa, and K. A. Hamdi are with the School of Electrical and Electronic Engineering, University of Manchester, Manchester M13 9PL, United Kingdom. (e-mail: d.so@manchester.ac.uk; e.alsusa@manchester.ac.uk; k.hamdi@manchester.ac.uk).

A. Shojaeifard and K.-K. Wong are with the Department of Electronic and Electrical Engineering, University College London, London, United Kingdom. (e-mail: a.shojaeifard@ucl.ac.uk; kai-kit.wong@ucl.ac.uk).

This work has been supported by the National Natural Science Foundation of China (Grant No. 61601186), Natural Science Foundation of Guangdong Province (Grant No. 1714050004216), Guangzhou Science Technology and Innovation Commission (Grant No. 201707010159), and the Engineering and Physical Sciences Research Council of UK (Grant No. EP/N008219/1).

industry to devise new technologies and standards for a new fifth-generation (5G) mobile network. A promising enabler for supporting user equipments (UEs) with increased density and quality-of-service (QoS) requirements is to deploy different types of base stations (BSs), thus forming what is referred to as heterogeneous cellular network (HCN) [2], [3]. The underlying air interface technology for HCN in the downlink is orthogonal frequency division multiple access (OFDMA) as specified in modern cellular standards [4].

Meanwhile, spectrum-efficiency (SE), a measure of the total amount of information transmitted per unit bandwidth, has been used as a key performance indicator in the design and analysis of cellular networks [5]–[7]. More recently, particular focus has been placed on the SE maximization problem in OFDMA-based HCNs. In particular, although intra-cell interference is suppressed via exclusive channel assignment in OFDMA, considering the dense and irregular deployment of nodes in HCNs, there remains inter-cell interference from both intra- and inter-tier sources. As a result, novel interference management strategies for HCNs has been an active area of research [8], [9]. In [10], a joint subcarrier and power allocation method is proposed to maximize SE considering densely deployed small-cells. In [11], a distributed resource allocation scheme using convex optimization is developed to enhance SE in a two-tier HCN. The authors in [12] propose a joint scheduling and power allocation scheme for maximizing the HCN sum rate. As a remedy to poor performance or/and high complexity, the authors in [13] propose a novel distributed interference management scheme. In addition, taking into account the presence of UEs with heterogeneous QoS requirements, the authors in [14] propose a joint subcarrier assignment and power allocation algorithm for the small-cells under an interference temperature limit to protect the macro-cell from harmful inter-tier interference. In [15], the problem of joint subcarrier assignment and power allocation is investigated in the context of downlink OFDMA-based HCNs where the sum rate of all small-cell UEs is maximized whilst protecting the minimum throughput requirements of the macro-cell UEs.

On the other hand, placing the focus solely on maximizing SE will lead to ever-rising network power consumption, which goes against global commitments for sustainable development jointly in terms of energy cost and environmental factors. Energy-efficiency (EE), defined as the total amount of information delivered per unit energy, is widely recognized as an important measure for joint spectrum- and energy-efficient cellular network design. The EE optimization problem has attracted great interest in the context of OFDMA-based systems

[16]–[19], and more recently OFDMA-based HCNs [20]–[24]. In [20], a resource allocation scheme for maximizing EE in spectrum underlay access OFDMA-based HCNs is proposed. In [21], the authors have jointly considered EE enhancement and interference control for OFDMA-based HCNs where the problem has been formulated as a Stackelberg game. In [22], the authors propose an energy-efficient spectrum sharing scheme among a macro-cell and multiple small-cells. In particular, the EE of the small-cells is maximized whilst preventing any potential severe interference leakage to the macro-cell UEs. The authors in [23] further investigate the joint power allocation and admission control problem in OFDMA-based HCNs. Specifically, considering spectrum underlay access, a novel resource allocation approach has been proposed with the goal of admitting the maximum possible number of UEs whilst keeping interference below a certain threshold.

A. Main Contributions

Previous works in the literature [20]–[23] investigated EE optimization in spectrum sharing OFDMA-based HCNs with spectrum underlay access, where interference constraints are imposed on the small-cell BSs in order to protect the QoS requirements of the macro-cell UEs. The inherent severity of inter-tier interference from spectrum underlay access may nevertheless degrade EE in certain cases. In addition, in contrast to the state-of-the-art studies on EE maximization [20]–[23] where the impact of spectrum utilization is not considered, in this paper we directly include bandwidth usage in the analysis by modeling the dynamic circuit power consumption as a linear function of the bandwidth [25], [26]. As a result, the global frequency reuse strategy may lead to higher circuit power consumption and degraded EE performance. On the other hand, utilizing overlay transmission, where the allocated bandwidth for the macro-cell and the small-cells are exclusive, is considered a promising strategy when it comes to densely deployed or bandwidth-abundant HCNs. In this paper, a fundamental study of EE in the context of an OFDMA-based two-tier HCN consisting of a macro-cell and multiple small-cells is provided. We consider both underlay and overlay transmission strategies and provide joint subcarrier assignment and power allocation schemes for maximizing EE subject to satisfying QoS constraints. It should be noted user association plays a pivotal role in enhancing the load balancing, the SE, and the EE in HCNs [27], [28]. However, similar to the works in [10]–[15], [20]–[23] where the user-BS association is considered fixed during runtime, this paper jointly investigates the power and bandwidth relationship under fixed user association strategy in order to optimize EE considering spectrum underlay and overlay access.

We first consider spectrum underlay access under constraints in terms of minimum user rate requirements as well as maximum transmit powers of the different BS types. The EE maximization problem under this setup, involving the apportion of limited radio resources to different UEs in different cells, is mixed-combinatorial and non-convex, and hence very challenging to solve. In order to tackle this problem, we decompose the original problem into a series of sub-problems with single inequality constraints. Based on the

quasi-concavity of the EE function, a dual-layer resource allocation approach is accordingly proposed for solving each sub-problem. We provide a complete solution where the inner-layer is solved using difference-of-two-concave-functions approximation and successive convex approximation while gradient-search is invoked for the outer-layer. The proposed joint subcarrier assignment and power allocation algorithm using underlay transmission may not be energy-efficient in certain cellular environments such as dense (due to severe inter-tier interference) or bandwidth-abundant (due to high circuit power) cases. In addition, the computational complexity for the proposed underlay-based approach can be high when the number of subcarriers or/and the number of UEs is relatively large. Consequently, motivated by our previous work on resource efficiency (RE) [29] and the idea of spectrum overlay access as an effective inter-tier interference mitigation strategy, a novel transmission scheme for the two-tier HCN is proposed where the optimization procedure for the macro-cell and the small-cells is separated. Specifically, we first optimize the RE at the macro-cell and determine the corresponding optimal bandwidth using joint subcarrier assignment and power allocation. We then assign the remaining spectrum to the small-cells and accordingly optimize their EE. Numerical results validate the effectiveness of the proposed algorithms and show that each algorithm has its strong point depending on scenario parameters such as the small-cells density and the bandwidth utilization.

B. Organization

The remainder of this paper is organized as follows. The system model and problem formulation is given in Section II. In Section III, the fundamentals for EE optimization in OFDMA-based HCNs is studied, where in particular the EE optimization problem with multiple inequality constraints is transformed into a multiple single inequality problems. In Section IV, joint subcarrier assignment and power allocation algorithm for the inner-layer is first introduced by exploiting some properties of the optimization problem. In particular, an efficient solution based on difference-of-two-concave-functions approximation for the inner-layer process is developed, followed by a complete solution to the dual-layer scheme. In Section V, a novel low-complexity algorithm is developed based on overlay transmission. Simulation results are provided in Section VI and conclusions are drawn in Section VII.

II. PRELIMINARIES

In this section, we describe the two-tier OFDMA-based HCN setup under consideration. The QoS-constrained EE optimization problem is then mathematically formulated.

A. System Model

We consider the downlink of an OFDMA-based two-tier HCN comprising of a macro-cell and L small-cells. The set of cells is denoted using $\mathcal{L} = \{0, 1, 2, \dots, L\}$, where indexes 0 and $\{1, 2, \dots, L\}$ correspond to the macro-cell and the small-cells, respectively. In addition, we assume that there are K_0

macro-cell UEs (MUEs) and K_l small-cell UEs (SUEs). For simplicity, the index of the UEs (MUEs and SUEs) associated with cell $l \in \mathcal{L}$ is denoted with \mathcal{K}_l . The HCN total available spectrum, W_{tot} , is divided into N_{tot} subcarriers with each having a bandwidth of $W_C = \frac{W_{tot}}{N_{tot}}$. Specifically, the set of all accessible frequencies is denoted with \mathcal{N} (where $|\mathcal{N}| = N_{tot}$). We consider exclusive channel assignment, where any OFDMA subcarrier can only be employed by at most one UE in a given cell at a given time, in order to prevent harmful collision. Note that the UE-BS association is considered fixed during runtime.

The channel power gain from the cell- m BS to the cell- l k -th UE over subcarrier n is denoted with $h_{[k,l,m]}^n$. The received signal-to-interference-plus-noise ratio (SINR) at the cell- l k -th UE over subcarrier n can be formulated as [30]

$$\gamma_{[k,l]}^n = \frac{h_{[k,l,l]}^n p_l^n}{\sum_{m \in \mathcal{L} \setminus \{l\}} h_{[k,l,m]}^n p_m^n + \sigma_{[k,l]}^n} \quad (1)$$

where p_l^n is the transmit power of the cell- l BS over subcarrier n and $\sigma_{[k,l]}^n$ is the noise power at the cell- l k -th UE over subcarrier n . Note that intra-cell interference is avoided under exclusive channel assignment. We thus can express the rate for the cell- l k -th UE over subcarrier n using

$$r_{[k,l]}^n = W_C \log_2 \left(1 + \gamma_{[k,l]}^n \right). \quad (2)$$

Hence, the rate achieved by UE $[k, l]$ is given by

$$R_{[k,l]} = \sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n \quad (3)$$

where $\rho_{[k,l]}^n \in \{1, 0\}$ indicates whether or not the n^{th} subcarrier is assigned to the UE $[k, l]$.

OFDM and OFDMA modulated signals exhibit high peak-to-average power ratio; thus suffering from severe nonlinearity effects. In practice, to circumvent the resulting performance degradation, input backoff is implemented by reducing the power of the input signal at the PA, so that the amplification stays within the linearity region as much as possible. This will lead to EE reduction because the PA efficiency is typically designed to peak near the saturation point and it usually drops rapidly as the input power decreases. On the other hand, it has been shown in [31] that the relations between relative RF output power and BS power consumption are nearly linear. Hence, considering that BSs are the dominant sources of energy consumption in cellular networks, a linear approximation of the BS power model appears justified,

$$P = \zeta P_T + P_C \quad (4)$$

where ζ , P_T and P_C denote the BS reciprocal of drain efficiency of the power amplifier, transmission power, and circuit power consumption, respectively. Furthermore, the circuit energy consumption consists of static (fixed) and dynamic parts, where the latter depends on the active links parameters. Motivated by the approach in [25] and [26], the circuit power consumption is considered to be proportional to the total utilized bandwidth for transmission. Consequently, the total

circuit power can be written as

$$P_c = P_s + \varrho W \quad (5)$$

where P_s is the static circuit power in transmission mode, ϱ is a constant corresponding to the dynamic power consumption per unit bandwidth, and W is the occupied bandwidth. As a result, the total power consumption in the two-tier OFDMA-based HCN under consideration is given by

$$P = \sum_{l \in \mathcal{L}} (\zeta P_T^{[l]} + P_C^{[l]}) \quad (6)$$

where $P_T^{[l]} = \sum_{n \in \mathcal{N}} p_l^n$ denotes the transmission power at the cell- l BS, and p_l^n is the power allocated on the n^{th} subcarrier at the cell- l ; $P_C^{[l]} = P_s^l + \varrho W_l$ represents the circuit power at the cell- l BS where P_s^l and W_l respectively denote the static circuit power and the occupied bandwidth at the cell- l .

B. Problem Formulation

Recall that EE is defined as the total number of successfully delivered bits per unit energy. The two-tier OFDMA-based HCN EE in the downlink can hence be described using the following equation

$$\lambda_{EE} \triangleq \frac{C}{P} = \frac{\sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}_l} R_{[k,l]}}{\sum_{l \in \mathcal{L}} (\zeta P_T^{[l]} + P_C^{[l]})} \quad (7)$$

where C and P are respectively used to denote the total data rate and the total power consumption of the HCN. Given that an OFDMA-based cellular network will occupy as much bandwidth as possible in order to maximize the EE [29], in this paper, we consider a two-tier OFDMA-based HCN where the entire spectrum is fully occupied in order to exploit the EE. Therefore, the whole bandwidth is occupied and shared by macro-cell and small-cells, i.e., $W_l = W_{tot}, l \in \mathcal{L}$. Moreover, we are concerned with the problem of achieving high EE whilst guaranteeing the required QoS constraints for each user under limited transmit power resources. In other words, the power allocated on the n^{th} at the cell- l , $p_l^n \geq 0$ and the subcarrier assignment $\rho_{[k,l]}^n \in \{1, 0\}$ are the optimization variables. Hence, we formulate an optimization problem for maximizing EE under a series of (minimum) rate requirements and maximum power budgets. Accordingly, the EE optimization problem for the two-tier OFDMA-based HCN is formulated as

$$\max_{\rho, \mathbf{p}} \lambda_{EE} \quad (8)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \quad \forall l \in \mathcal{L}, \quad (9)$$

$$\sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n \geq \delta_{[k,l]}, \quad \forall k \in \mathcal{K}_l, \forall l \in \mathcal{L}, \quad (10)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k,l]}^n = 1, \quad \forall n \in \mathcal{N}, \forall l \in \mathcal{L}, \quad (11)$$

where ρ denotes a feasible subcarrier assignment indicator vector and $\rho = \{[\rho_{[k,l]}^n]_{LKN \times 1} | \rho_{[k,l]}^n \in \{1, 0\}, \forall l \in \mathcal{L}, \forall k \in \mathcal{K}_l, \forall n \in \mathcal{N}\}$, \mathbf{p} denotes a feasible power allocation vector and $\mathbf{p} = \{[p_l^n]_{LN \times 1} | p_l^n \geq 0, \forall n \in \mathcal{N}, \forall l \in \mathcal{L}\}$, $P_{max}^{[l]}$ is the maximum transmit power of the cell- l BS and $\delta_{[k,l]}$ corresponds

to the minimum rate requirements for UE $[k, l]$. Therefore, constraints (9)-(10) are used to guarantee the maximum power budget and the minimum rate target in each cell. In addition, the constraint in (11) corresponds to the exclusive subcarrier assignment strategy in any cell.

The EE optimization problem here, which considers joint subcarrier assignment and power allocation in the presence of inter-cell interference, is mixed-combinatorial and non-convex. The solution is therefore nontrivial and cannot be obtained directly. As a result, in the following sections, we develop two different resource allocation approaches considering both spectrum underlay and overlay access.

III. FUNDAMENTALS OF EE OPTIMIZATION IN OFDMA-BASED HCNs WITH SPECTRUM UNDERLAY ACCESS

With spectrum underlay access, the small-cells share the available radio spectrum with the macro-cell and hence introduce inter-tier interference which renders the resource allocation problem significantly more challenging to tackle. In addition, the variables for subcarrier assignment and power allocation are coupled together and hence the non-convex optimization problem in (8)-(11) is extremely difficult to solve. In this section, we provide a fundamental study for energy-efficient design in OFDMA-based underlay HCNs. In particular, a relationship between optimal EE and UE's data rate is derived as in the following theorem.

Theorem 1. *For any rate vector for the UEs that satisfies the minimum rate constraint, $\mathbf{R} \geq \boldsymbol{\delta}$, achieved with power allocation $p_l^n, \forall (l, n) \in (\mathcal{L}, \mathcal{N})$ and subcarrier assignment $\rho_{[k,l]}^n$, the maximum achievable EE, namely,*

$$\lambda_{EE}^*(\mathbf{R}) \triangleq \max_{\rho_{[k,l]}^n, P_l^n} \lambda_{EE} \quad (12)$$

$$\text{s.t. } \mathbf{R} \geq \boldsymbol{\delta}, \forall k \in \mathcal{K}_l, \forall l \in \mathcal{L}, \quad (13)$$

$$\sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \forall l \in \mathcal{L}, \quad (14)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k,l]}^n = 1, \forall n \in \mathcal{N}, \forall l \in \mathcal{L}, \quad (15)$$

where $\mathbf{R} = [R_{[1,0]} \ R_{[2,0]} \ \dots \ R_{[K_L,L]}]$ and $\boldsymbol{\delta} = [\delta_{[1,0]} \ \delta_{[2,0]} \ \dots \ \delta_{[K_L,L]}]$, is strictly quasi-concave in \mathbf{R} .

Proof: See Appendix A.

By definition, a continuous and strictly quasi-concave function has a unique maximum value over a finite domain [32]. Therefore, *Theorem 1* indicates that there always exists a unique EE solution. On the other hand, the original EE optimization problem is very challenging to solve due to the multiple inequality constraints in (10). By extending the Lagrange dual decomposition method for single-cell multi-carrier systems [16] to our OFDMA-based underlay HCN setup, the gradient ascent approach can be invoked to generate \mathbf{R} , and

$$\mathbf{R}(n+1) = [\mathbf{R}(n) + \mu \nabla \lambda_{EE}(\mathbf{R}(n))]^+. \quad (16)$$

However, a closed-form expression for the gradient of this vector does not exist; hence, it is impossible to directly employ (16) in order to obtain the solution in (8)-(11). Nevertheless,

we can transform the gradient ascent method in (16) using the following approach [33]

$$\begin{aligned} R_0(n+1) &= [R_0(n) + \mu \nabla \lambda_{EE}(R_0(n))]^+, \\ &\vdots \\ R_L(n+1) &= [R_L(n) + \mu \nabla \lambda_{EE}(R_L(n))]^+. \end{aligned} \quad (17)$$

As a result, the vector gradient can be alternatively decomposed into multiple scalar gradient; thus making the optimization problem relatively easier to solve. Accordingly, we propose an iterative resource allocation scheme to tackle the EE optimization problem for two-tier OFDMA-based HCNs. Similar to the gradient decomposition approach in [33], by keeping one minimum rate constraint at a time and setting all others as equality constraints (fixed rate), the optimization problem with multiple inequality constraints can be decomposed into a series of optimization problems with single inequality constraint. Specifically, all other users will be under equality constraints whilst user $[k^*, l^*]$ is under inequality constraint (minimum rate requirement). Therefore, the EE maximization problem is transformed into

$$\max_{R_{[k^*, l^*]}} \lambda_{EE} \quad (18)$$

$$\text{s.t. } R_{[k,l]} = \bar{R}_{[k,l]}, \forall (k,l) \in (\mathcal{K}_l, \mathcal{L}) \setminus (k^*, l^*), \quad (19)$$

$$\bar{R}_{[k^*, l^*]} \geq \delta_{[k^*, l^*]}, \quad (20)$$

where $\bar{R}_{[k,l]}$ represents the optimal rate (for all other users apart from user $[k^*, l^*]$) obtained from the previous iteration. Under this setup, we can obtain the solution for user $[k^*, l^*]$, $\bar{R}_{[k^*, l^*]}$, by solving the above single inequality constrained optimization problem. Once this UE's rate is obtained, the next user is placed under inequality constraint (minimum rate requirement), i.e., $\forall (k^* + 1, l^*) \in (\mathcal{K}_l, \mathcal{L})$, while all other users have updated rate values from the previous iteration. In particular, we can rewrite the constraints in (19)-(20) as

$$R_{[k,l]} = \bar{R}_{[k,l]}, \forall (k,l) \in (\mathcal{K}_l, \mathcal{L}) \setminus \{k^* + 1, l^*\}, \quad (21)$$

$$\sum_{n \in \mathcal{N}} \rho_{[k^*+1, l^*]}^n r_{[k^*+1, l^*]}^n \geq \delta_{[k^*+1, l^*]}. \quad (22)$$

After solving the problem for the next user, the maximum EE value is stored in the buffer and the corresponding optimal rate $\bar{R}_{[k^*+1, l^*]}$ for user $[k^* + 1, l^*]$ is updated. This process is repeated cyclically for all cells until convergence, i.e., $\lambda_{EE}^{opt}(n+1) - \lambda_{EE}^{opt}(n) \leq \varepsilon$. We provide a pseudocode for the proposed iterative resource allocation scheme:

- (1) Initialize $[k^*, l^*]$ as the first user in $(\mathcal{K}_l, \mathcal{L})$ with inequality constraint;
- (2) Tackle the problem in (18)-(20) and store $\lambda_{EE}^{opt}(n)$ in the buffer;
- (3) Modify the constraints using (21)-(22) and update the corresponding rates;
- (4) Repeat steps (2) and (3) until convergence $\lambda_{EE}^{opt}(n+1) - \lambda_{EE}^{opt}(n) \leq \varepsilon$.

The decomposed EE optimization problem in (18)-(20) has a single inequality constraint. With a fundamental study of the problem, we can arrive at the following theorem.

Theorem 2. *For any given rate for user $[k^*, l^*]$, $R_{[k^*, l^*]} \geq$*

$\delta_{[k^*, l^*]}$, achieved with power allocation $p_l^n, \forall (l, n) \in (\mathcal{L}, \mathcal{N})$ and subcarrier assignment $\rho_{[k, l]}^n$, the maximum EE, namely,

$$\lambda_{EE}^*(R_{[k^*, l^*]}) \triangleq \max_{\rho_{[k, l]}^n, p_l^n} \lambda_{EE} \quad (23)$$

$$\text{s.t. } R_{[k, l]} = \bar{R}_{[k, l]}, \forall (k, l) \in (\mathcal{K}_l, \mathcal{L}) \setminus (k^*, l^*), \quad (24)$$

$$\sum_{n \in \mathcal{N}} \rho_{[k^*, l^*]}^n r_{[k^*, l^*]}^n = R_{[k^*, l^*]} \geq \delta_{[k^*, l^*]}, \quad (25)$$

$$\sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \forall l \in \mathcal{L}, \quad (26)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k, l]}^n = 1, \forall n \in \mathcal{N}, \forall l \in \mathcal{L}, \quad (27)$$

is strictly quasi-concave in $R_{[k^*, l^*]}$.

Proof: Theorem II is a special case of Theorem I, thus a similar proof to that in Appendix A can be applied here.

Theorem III. In the rate region for user $[k^*, l^*]$ $[\delta_{[k^*, l^*]} \delta_{[k^*, l^*]}^{(max)}]$, the EE, $\lambda_{EE}^*(R_{[k^*, l^*]})$

(i) strictly decreases with $R_{[k^*, l^*]}$ and is maximized at $R_{[k^*, l^*]} = \delta_{[k^*, l^*]}$ if

$$\left. \frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}} \right|_{R_{[k^*, l^*]} = \delta_{[k^*, l^*]}} \leq 0,$$

(ii) strictly increases with $R_{[k^*, l^*]}$ and is maximized at $R_{[k^*, l^*]} = \delta_{[k^*, l^*]}^{(max)}$ if

$$\left. \frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}} \right|_{R_{[k^*, l^*]} = \delta_{[k^*, l^*]}^{(max)}} > 0$$

$$\text{and } \left. \frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}} \right|_{R_{[k^*, l^*]} = \delta_{[k^*, l^*]}^{(max)}} \geq 0,$$

(iii) first strictly increases and then strictly decreases with $R_{[k^*, l^*]}$ and is maximized at $R_{[k^*, l^*]}(\lambda_{EE}^{opt})$ if

$$\left. \frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}} \right|_{R_{[k^*, l^*]} = \delta_{[k^*, l^*]}^{opt}} > 0$$

$$\text{and } \left. \frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}} \right|_{R_{[k^*, l^*]} = \delta_{[k^*, l^*]}^{opt}} < 0,$$

(iv) infeasible if

$$\delta_{[k^*, l^*]} > \delta_{[k^*, l^*]}^{(max)},$$

where $\delta_{[k^*, l^*]}^{(max)}$ is the maximum feasible rate for user $[k^*, l^*]$ under the constraints in (24)-(27) and $R_{[k^*, l^*]}(\lambda_{EE}^{opt})$ is the rate for user $[k^*, l^*]$ that corresponds to the maximum EE under the constraints in (24)-(27).

Proof: see Appendix B.

The quasi-concavity property guarantees the existence of a unique maximum, hence Theorem II proves the existence of a unique EE solution. Moreover, the quasi-concavity of EE optimization problem (Theorem III) further indicates that $\lambda_{EE}(R_{[k^*, l^*]})$ either decreases or first increases and then decreases with $R_{[k^*, l^*]}$. Thus, problem (23)-(27) can be solved through a dual-layer decomposition method using the following processes:

(i) Inner-layer: Finds the maximum EE in cell- l^* , $\lambda_{EE}^*(R_{[k^*, l^*]})$, under a fixed rate, $R_{[k^*, l^*]}$.

(ii) Outer-layer: Obtains the optimal EE, λ_{EE}^{opt} , using heuristic search.

Note that the key challenge for adopting the proposed dual-layer decomposition method lies in the inner-layer mechanism, as discussed in the following section.

IV. JOINT SUBCARRIER ASSIGNMENT AND POWER ALLOCATION FOR EE MAXIMIZATION

In this section, we provide a joint subcarrier assignment and power allocation method for the inner-layer by exploiting the fundamental properties of the optimization problem. A complete solution to the proposed dual-layer resource allocation approach is then presented.

Given that a centralized control is capable of providing optimal resource allocation for the entire network and exhibits a fast convergence, we employ a centralized approach for underlay access where the network contains a single central entity that performs resource allocation. This central entity collects information, such as the channel quality and the resource demand from all users.

Given that the optimization problem in (23)-(27) involves fixed throughput requirements (equality constraints), it can be equivalently expressed in terms of the the following power minimization problem

$$\min_{\rho_{[k, l]}^n, p_l^n} \sum_{l \in \mathcal{L}} \sum_{n \in \mathcal{N}} p_l^n \quad (28)$$

$$\text{s.t. } \sum_{n \in \mathcal{N}} \rho_{[k, l]}^n r_{[k, l]}^n = R_{[k, l]}, \forall (k, l) \in (\mathcal{K}_l, \mathcal{L}), \quad (29)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k, l]}^n = 1, \forall n \in \mathcal{N}, \forall l \in \mathcal{L}, \quad (30)$$

$$\sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \forall l \in \mathcal{L}. \quad (31)$$

The above power minimization problem involves subcarrier assignment and power allocation, therefore, we can extend the iterative approach proposed in [12] to a HCN scenario.

The joint subcarrier assignment and power allocation process can be separated as

$$\underbrace{\boldsymbol{\rho}[0] \rightarrow \mathbf{p}[0]}_{\text{Initialization}} \rightarrow \cdots \rightarrow \underbrace{\boldsymbol{\rho}[t] \rightarrow \mathbf{p}[t]}_{\text{Iteration } t} \rightarrow \underbrace{\boldsymbol{\rho}^{opt} \rightarrow \mathbf{p}^{opt}}_{\text{Optimal Solution}}. \quad (32)$$

where $\mathbf{p}^n = [p_0^n, p_1^n, \dots, p_L^n]$, $\mathbf{p}_l = [p_l^1, p_l^2, \dots, p_l^N]$, $\mathbf{p} = \text{vec}[\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_L]$, $\boldsymbol{\rho}_{[k, l]} = [\rho_{[k, l]}^1, \rho_{[k, l]}^2, \dots, \rho_{[k, l]}^N]$, $\boldsymbol{\rho}_l = \text{vec}[\rho_{[1, l]}, \rho_{[2, l]}, \dots, \rho_{[K_l, l]}]$ and $\boldsymbol{\rho} = \text{vec}[\rho_{[0]}, \rho_{[1]}, \dots, \rho_{[L]}]$. Note that the number inside the square bracket denotes the iteration number.

At the initial moment of each iteration t , based on a given power allocation $\mathbf{p}[t-1]$ from the last iteration, we solve the subcarrier assignment problem and obtain the optimal $\boldsymbol{\rho}[t]$. We then find the optimal power allocation $\mathbf{p}[t]$ based on the fixed $\boldsymbol{\rho}[t]$ obtained from the previous step. This process is repeated until convergence, i.e., no further EE improvement is realized. Therefore, this iterative resource allocation approach simplifies the original EE problem by separating it into two sub-problems, namely the subcarrier assignment process and the power allocation process. More importantly, the number

of variables is decreased by nearly a half in each sub-problem hence allowing for more tractable algorithm designs.

A. Optimal Subcarrier Assignment for Power Minimization Problem

Having a fixed power allocation $\mathbf{p}[t-1]$ obtained from the last iteration, we attempt to obtain the optimal subcarrier assignment $\rho[t]$ at iteration t . The optimization problem in (28)-(31) is accordingly converted to

$$\max_{\rho_{[k,l]}^n} \sum_{l \in \mathcal{L}} \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}_l} r_{[k,l]}^n(\rho_{[k,l]}^n, \mathbf{p}^n[t-1]) \quad (33)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k,l]}^n = 1, \forall n \in \mathcal{N}, \forall l \in \mathcal{L} \quad (34)$$

where $r_{[k,l]}^n(\rho_{[k,l]}^n, \mathbf{p}^n[t-1])$ denotes the rate function with respect to the subcarrier assignment $\rho_{[k,l]}^n$ and the power allocation result obtained from the previous iteration. We can therefore arrive at the following theorem.

Theorem IV. *The solution of (33)-(34) involves assigning subcarriers to UEs that consume the lowest power (i.e., the highest SINR) on those subcarriers.*

Proof: see Appendix C.

The optimal subcarrier assignment for all UEs can be found using *Theorem IV*, thus avoiding the need for an exhaustive search approach which here would be exponentially computationally complex in the number of subcarriers. Furthermore, by extending the approach in [15] to a HCN scenario, the subcarrier assignment problem can be decomposed into L sub-problems such that

$$\max_{\rho_{[k,l]}^n} \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}_l} r_{[k,l]}^n(\rho_{[k,l]}^n, \mathbf{p}^n[t-1]) \quad (35)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}_l} \rho_{[k,l]}^n = 1, \forall n \in \mathcal{N}. \quad (36)$$

Recall from *Theorem IV* that the optimal solution of (35)-(36) is to assign each subcarrier to the UE with the highest SINR. We can therefore conclude the optimal subcarrier assignment strategy for cell $l \in \mathcal{L}$ at iteration t using

$$\rho_{[k,l]}^n[t] = \rho_{[k,l]}^{n*} = \begin{cases} 1, & \text{if } k = \operatorname{argmax}_{k \in \mathcal{K}_l} r_{[k,l]}^n(\mathbf{p}^n[t-1]) \\ 0, & \text{otherwise} \end{cases}. \quad (37)$$

B. Optimal Power Allocation for Power Minimization Problem

After determining the optimal subcarrier assignment $\rho[t]$ at iteration t , we aim to find the optimal power allocation. Therefore, problem (28)-(31) is now converted to

$$\min_{p_l^n} \sum_{l \in \mathcal{L}} \sum_{n \in \mathcal{N}} p_l^n \quad (38)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n = R_{[k,l]}, \forall (k, l) \in (\mathcal{K}_l, \mathcal{L}), \quad (39)$$

$$\sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \forall l \in \mathcal{L}. \quad (40)$$

It is easy to note that the power allocation problem (38)-(40) is non-convex as a result of the non-convexity of the

SINR and corresponding rate functions. Hence, we extend the successive convex approximation approach proposed in [15] to our work in order to solve the above non-convex problem. The methodology can be described as:

- 1 Initialize a power vector $\mathbf{p}[0]$ and $t_p = 1$.
- 2 Create the t_p -th convex sub-problem by estimating the non-concave rate function (involving SINR) with some concave function based on the previous result $\mathbf{p}[t_p-1]$.
- 3 Tackle the t_p -th sub-problem to achieve the solution $\mathbf{p}[t_p]$, and accordingly update the approximation parameters in Step 2.
- 4 Update $t_p = t_p + 1$ and iterate this process until $\mathbf{p}[t_p]$ converges.

In the following part, an efficient power allocation approach based on difference-of-two-concave-functions approximation is proposed to update Step 2 and 3.

To solve Step 2 and 3 in the above successive convex approximation process, we formulate the data rate function (2) in a difference-of-two-concave-functions approximation form

$$\sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n(\mathbf{p}^n) = f_l(\mathbf{p}) - g_l(\mathbf{p}), \quad (41)$$

where $\rho_{[k,l]}^n$ denotes the subcarrier assignment obtained previously, $f_l(\mathbf{p})$ and $g_l(\mathbf{p})$ are respectively representing two concave functions which are defined as

$$f_l(\mathbf{p}) = \sum_{n \in \mathcal{N}} \rho_{[k,l]}^n \ln \left(\sum_{m \in \mathcal{L}} h_{[k,l,m]}^n p_m^n + \sigma_{[k,l]}^n \right) \quad (42)$$

and

$$g_l(\mathbf{p}) = \sum_{n \in \mathcal{N}} \rho_{[k,l]}^n \ln \left(\sum_{m \in \mathcal{L} \setminus \{l\}} h_{[k,l,m]}^n p_m^n + \sigma_{[k,l]}^n \right). \quad (43)$$

We then approximate $g_l(\mathbf{p})$ based on a fixed $\mathbf{p}[t_p-1]$ (obtain from iteration t_p-1) such that [34]

$$g_l(\mathbf{p}) \approx g_l(\mathbf{p}[t_p-1]) + \nabla g_l^T[\mathbf{p}[t_p-1]](\mathbf{p} - \mathbf{p}[t_p-1]) \quad (44)$$

where $\nabla g_l(\mathbf{p})$ is a vector with length $(L+1)N$, and its corresponding entry is defined as

$$\nabla g_l(\mathbf{p})^{(Nj+n)} = \begin{cases} 0, & \text{if } j = l \\ \frac{\rho_{[k,l]}^n h_{[k,l,j]}^n}{\sum_{s \in \mathcal{L} \setminus \{l\}} h_{[k,l,s]}^n p_s^n + \sigma_{[k,l]}^n}, & \text{if } j = \mathcal{L} \setminus \{l\} \end{cases}. \quad (45)$$

Therefore, combining (41) with (45), we can obtain

$$\sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n(\mathbf{p}^n) \approx f_l(\mathbf{p}) - g_l(\mathbf{p}[t_p-1]) - \nabla g_l^T[\mathbf{p}[t_p-1]](\mathbf{p} - \mathbf{p}[t_p-1]). \quad (46)$$

It should be noted that the right-hand side of the above equation is concave in \mathbf{p} .

Therefore, (46) enables us to reformulate the problem (38)-(40) into a series of convex optimization sub-problems. Particularly, the t_p -th iteration (t_p sub-problem) is established as

$$\min_{p_l^n} \|\mathbf{p}\| \quad (47)$$

$$\begin{aligned} \text{s.t. } f_l(\mathbf{p}) - g_l(\mathbf{p}[t_p - 1]) - \nabla g_l^T[t_p - 1](\mathbf{p} - \mathbf{p}[t_p - 1]) \\ = R_{[k,l]}, \forall (k, l) \in (\mathcal{K}_l, \mathcal{L}), \end{aligned} \quad (48)$$

$$\sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \forall l \in \mathcal{L}, \quad (49)$$

where $\mathbf{p}[t_p - 1]$ has been determined from the last iteration $t_p - 1$. Since the objective function and the constraints are all convex, this problem (t_p sub-problem) can be efficiently solved using Branch and Bound (BnB) method [35]. However, the BnB method has very high complexity and hence not practical to use. Introducing a new indication vector $\mathbf{s}_l = [\mathbf{0}_{(l-1)N} \ \mathbf{1}_N \ \mathbf{0}_{(L-l)N}]^T$, where $\mathbf{0}_m$ is a $1 \times m$ vector whose elements are all zeros, and $\mathbf{1}_m$ is a $1 \times m$ vector whose elements are all ones, constraint (49) can be reformulated as,

$$\mathbf{p}\mathbf{s}_l^T \leq P_{max}^{[l]}, \forall l \in \mathcal{L}. \quad (50)$$

Therefore, the optimization problem can be transformed into second order cone programming (SOCP) form, which can be solved efficiently using interior point methods [32]. Once the sub-problem (47)-(49) is found, $\mathbf{p}[t_p]$ is obtained and updated in (44) to solve the ($t_p + 1$) sub-problem in the next iteration.

C. A Complete Solution to the Dual-Layer Approach

The inner-layer, which under a fixed rate for user- $[k^*, l^*]$, $R_{[k^*, l^*]}$, finds the maximum EE, $\lambda_{EE}^*(R_{[k^*, l^*]})$, can be efficiently solved based on the proposed joint subcarrier assignment and power allocation algorithm. Next, we propose an approach for the outer-layer process using gradient-search.

With an initial setting $R_{[k^*, l^*]}(1)$, $\lambda_{EE}^*(R_{[k^*, l^*]}(1))$ can be obtained using the proposed joint subcarrier assignment and power allocation algorithm. On the basis of *Theorem II* and *Theorem III*, we can then update $R_{[k^*, l^*]}$ using the following approach

$$R_{[k^*, l^*]}(n+1) = \begin{cases} \frac{R_{[k^*, l^*]}(n)}{\varpi} & \left. \frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}} \right|_{R_{[k^*, l^*]}} > 0 \\ \varpi R_{[k^*, l^*]}(n) & \text{otherwise} \end{cases} \quad (51)$$

where $\varpi > 1$ denotes the search step size. Furthermore, we need to reduce the step size ϖ if the gradient $\frac{d\lambda_{EE}^*(R_{[k^*, l^*]})}{dR_{[k^*, l^*]}}$ changes its sign as

$$\varpi(n+1) = \frac{\varpi(n)}{2}, \quad (52)$$

and (51) is repeated until convergence. Since $\lambda_{EE}^*(R_{[k^*, l^*]})$ is strictly quasi-concave in $R_{[k^*, l^*]}$, $\lambda_{EE}^*(R_{[k^*, l^*]})$ either strictly decreases or first increases and then strictly decreases with $R_{[k^*, l^*]}$, and the proposed approach will terminate with either convergence, i.e., $|\lambda_{EE}^*[R_{[k^*, l^*]}(n+1)] - \lambda_{EE}^*[R_{[k^*, l^*]}(n)]| \leq \epsilon$, or $\delta_{[k^*, l^*]}$ if $\lambda_{EE}^*(R_{[k^*, l^*]})$ is monotonically decreasing and $\delta_{[k^*, l^*]}(max)$ if $\lambda_{EE}^*(R_{[k^*, l^*]})$ is monotonically increasing. As stated in *Theorem III*, the problem would be infeasible if $\delta_{[k^*, l^*]} > \delta_{[k^*, l^*]}(max)$ holds. Nonetheless this condition only occurs if $\delta_{[k^*, l^*]}(max)$ is very small, which rarely happens in general. In addition, we can obtain $\delta_{[k^*, l^*]}(max)$ by solving the

- 1) Initialize $R_{[k^*, l^*]}(1) = \delta_{[k^*, l^*]}$, and set $n = 1$;
- 2) **IF** $\delta_{[k^*, l^*]} > \delta_{[k^*, l^*]}(max)$
- 3) **RETURN** Infeasible;
- 4) **ELSE**
- 5) **REPEAT**
- 6) Obtain the maximum EE $\lambda_{EE}^*(R_{[k^*, l^*]})$ using the proposed joint subcarrier assignment and power allocation approach in Section IV.A and Section IV.B;
- 7) Update $R_{[k^*, l^*]}(n)$ using (51); $n = n + 1$;
- 8) **UNTIL** $|\lambda_{EE}^*[R_{[k^*, l^*]}(n+1)] - \lambda_{EE}^*[R_{[k^*, l^*]}(n)]| \leq \epsilon$;
- 9) **END**

TABLE I
COMPLETE SOLUTION TO THE EE OPTIMIZATION PROBLEM

following problem

$$\max_{\rho_{[k,l]}^n, p_l^n} R_{[k^*, l^*]} \quad (53)$$

$$\text{s.t. } \sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n = R_{[k,l]}, \forall (k, l) \in (\mathcal{K}_l, \mathcal{L}) \setminus (k^*, l^*), \quad (54)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k,l]}^n = 1, \forall n \in \mathcal{N}, \forall l \in \mathcal{L}, \quad (55)$$

$$\sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \forall l \in \mathcal{L}, \quad (56)$$

which can be efficiently solved using the joint subchannel and power allocation algorithm in [15]. It should be noted that the computational complexity of the outer-layer algorithm depends on the number of iterations and is linear with $\frac{1}{\varpi^2}$ [36]. Therefore, choosing an appropriate ϖ to balance the convergence speed, accuracy and complexity is very important. The complete solution to the EE optimization problem in (23)-(27) for the OFDMA-based two-tier HCN with spectrum underlay access is summarized in Table I.

V. SOLUTION BASED ON SPECTRUM OVERLAY ACCESS AND RESOURCE EFFICIENCY

The inter-tier interference from the proposed underlay-based approach may degrade EE especially under high throughput requirements in densely deployed scenarios. In addition, considering that the available bandwidth is fully exploited by the macro-cell and the small-cells at the same time, a higher circuit power consumption and hence reduced EE performance may be incurred. Furthermore, although the proposed iterative joint subcarrier and power allocation algorithm is numerically stable, its computational complexity depends on the number of optimizing variables, which can be large if the number of subcarriers or the number of UEs is large. Hence the complexity of this scheme is comparatively high. As a result, based on the idea of spectrum overlay access, we next develop a low-complexity resource allocation approach for the two-tier HCN under consideration, where the allocated bandwidth for the macro-cell and the small-cells are mutually exclusive. In other words, although the entire bandwidth is still fully occupied by the network, the proportion allocated for macro-cell and small-cells can be optimized. Therefore, under this

setup, a distributed control mechanism is used where it does not require a central entity and allows BSs to make resource allocation decisions by themselves through the interaction between BSs. Hence, this distributed control is attractive owing to its low implementational complexity and low signaling overhead. It is particularly suitable for large networks, especially for HCNs associated with many autonomous small-cells. In order to consider transmission-bandwidth requirements, a new paradigm named resource efficiency (RE) has been proposed in [29] that is capable of exploiting the tradeoff between EE and SE by balancing consumption power and occupied bandwidth (simultaneously optimizing both EE and SE). Therefore, we first optimize the RE for macro-cell in order to determine the optimal proportion of bandwidth where the total circuit power plays a very important role in this process. After determining the spectrum for macro-cell, the EE of small-cells utilizing the remaining spectrum is maximized in order to further improve the system.

A. Resource Efficiency Optimization for Macro-cell

In [29], RE is defined as a weighted EE-SE trade-off using a normalizing factor β

$$\lambda_{RE} \triangleq \frac{R}{P} (1 + \beta \frac{\eta_P}{\eta_W}) \quad (57)$$

where η_P and η_W respectively denote the power utilization and bandwidth utilization such that

$$\eta_P \triangleq \frac{P}{P_{tot}}, \quad \eta_W \triangleq \frac{W}{W_{tot}}. \quad (58)$$

The notion behind RE maximization requires EE and SE as input vectors in a multi-objective optimization problem. Hence, there does not exist a-priori correspondence between a weight vector and a solution vector. This implies that the weights that control EE and SE has to be decided by the decision maker. In addition, it has been shown in [29] that the corresponding EE is decreasing with increasing β while the corresponding SE is increasing with increasing β . As a result, we modify the RE formulation to a more generalized expression

$$\begin{aligned} \lambda_{RE} &\triangleq \alpha \frac{R}{P} + (1 - \alpha) \tau \frac{R}{W} \\ &= \frac{R}{P} \left(\alpha + (1 - \alpha) \frac{\eta_P}{\eta_W} \right) \end{aligned} \quad (59)$$

where $0 \leq \alpha \leq 1$ and $\tau = \frac{W_{tot}}{P_{tot}}$.

The generalized RE optimization problem in the downlink of the macro-cell can be mathematically formulated as

$$\max_{\rho, \mathbf{p}, \alpha} \frac{\sum_{k \in \mathcal{K}_0} \sum_{n \in \mathcal{N}} \rho_{[k,0]}^n r_{[k,0]}^n}{\zeta P_T^{[0]} + P_C^{[0]}} (\alpha + (1 - \alpha) \frac{\eta_P}{\eta_W}) \quad (60)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} \rho_{[k,0]}^n r_{[k,0]}^n \geq \delta_{[k,0]}, \quad \forall k \in \mathcal{K}_0, \quad (61)$$

$$\sum_{k \in \mathcal{K}_0} \rho_{[k,0]}^n = 1, \quad \forall n \in \mathcal{N}, \quad (62)$$

$$\sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}_0} \rho_{[k,0]}^n \leq N_{tot}, \quad (63)$$

$$\sum_{n \in \mathcal{N}} p_0^n \leq P_{max}^{[0]} \quad (64)$$

where $P_T^{[0]} = \sum_{n \in \mathcal{N}} p_0^n$ and $P_C^{[0]} = P_s^l + \varrho W_C \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \rho_{[k,0]}^n$. Problem (60)-(64) is mixed-combinatorial and non-convex. In order to tackle this, the subcarrier assignment and power allocation procedures are separated. Specifically, we first analyze the fundamental properties of the case with a given weight α and a given subcarrier assignment set, which are summarized in the following paragraphs. For simplicity, here, we remove the index from the macro-cell parameters, e.g., transmission power is changed from $P_T^{[0]}$ to P_T , MUEs set is changed from \mathcal{K}_0 to \mathcal{K} , the rate and the corresponding power for user $[k, 0]$ on the subcarrier n is changed to $r_{k,n}$ and $p_{k,n}$, respectively.

Considering a given weight α , subcarrier allocation vector ρ and its corresponding UEs set $\mathcal{S}_k (\forall k \in \mathcal{K})$, the maximum RE at a certain transmit power, P_T , namely,

$$\lambda_{RE}(P_T) \triangleq \max_{p_{k,n} \geq 0} \frac{\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{S}_k} r_{k,n}}{\zeta P_T + P_C} (\alpha + (1 - \alpha) \frac{\eta_P}{\eta_W}) \quad (65)$$

subject to

$$\sum_{n \in \mathcal{S}_k} r_{k,n} \geq \delta_k, \quad \forall k \in \mathcal{K}, \quad (66)$$

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{S}_k} p_{k,n} = P_T \leq P_{max}, \quad (67)$$

has the following properties:

- (i) $\lambda_{RE}(P_T)$ is a continuously differentiable quasi-concave function with respect to P_T ,
- (ii) the derivative of $\lambda_{RE}(P_T)$ meets the following condition

$$\frac{d\lambda_{RE}(P_T)}{dP_T} = \frac{(\alpha + (1 - \alpha) \frac{\eta_P}{\eta_W}) \frac{d\bar{R}(P_T)}{dP_T} - \alpha \zeta \lambda_{EE}(P_T)}{\zeta P_T + P_C} \quad (68)$$

where

$$\lambda_{EE}(P_T) = \frac{\bar{R}(P_T)}{\zeta P_T + P_C}, \quad (69)$$

$$\bar{R}(P_T) \triangleq \max_{p_{k,n} \geq 0} R(P_T) = \max_{p_{k,n} \geq 0} \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{S}_k} r_{k,n}, \quad (70)$$

$$r_{k,n} = W_C \log_2 (1 + p_{k,n} g_{k,n}), \quad (71)$$

and $\bar{R}(P_T)$ represents the maximum sum rate under the maximum power constraint and minimum rate constraint (66)-(67) meeting the condition

$$\frac{d\bar{R}(P_T)}{dP_T} = \max_{k \in \mathcal{K}, n \in \mathcal{S}_k} \frac{W_C g_{k,n} \log_2 e}{1 + p_{k,n}^* g_{k,n}} \quad (72)$$

where $p_{k,n}$ denotes the power allocated for the k^{th} user on the n^{th} subcarrier, $g_{k,n} \triangleq \frac{|h_k^n|^2}{\sigma_k^2}$ and $p_{k,n}^*$ ($n \in \mathcal{S}_k$) are respectively representing the channel-power-to-noise ratio (CNR) of the k -th UE on the n -th OFDMA subcarrier and the optimal allocated power on the n -th subcarrier to obtain $\bar{R}(P_T)$.

Proof: See Appendix D.

For the case with fixed transmission power P_T and subcarrier assignment vector ρ , we can rewrite the RE of the two-tier HCN as

$$\lambda_{RE}(P_T) = \max_{\rho_{k,n} \geq 0} R\left(\frac{\alpha}{P} + (1-\alpha)\frac{W_{tot}}{WP_{tot}}\right) = \omega \bar{R}(P_T), \quad (73)$$

where $\omega \triangleq \frac{\alpha}{P} + (1-\alpha)\frac{W_{tot}}{WP_{tot}}$. To derive the optimal power allocation, we can extend the multi-level water-filling scheme [29] to a HCN scenario as follows

$$\tilde{p}_{k,n} = \left(\mu_k - \frac{1}{g_{k,n}}\right)^+, \quad \forall n \in \mathcal{S}_k, \quad (74)$$

$$\sum_{n \in \mathcal{S}_k, \tilde{p}_{k,n} > 0} W_C \log_2(\mu_k g_{k,n}) = \delta_k, \quad (75)$$

$$p_{k,n}^* = \tilde{p}_{k,n} + \left(\mu - \frac{1}{g_{k,n}} - \tilde{p}_{k,n}\right)^+, \quad (76)$$

$$\sum_{k \in \mathcal{K}} \sum_{n \in \{\mathcal{S}_k | \tilde{p}_{k,n} > \tilde{p}_{k,n}\}} \left(\mu - \frac{1}{g_{k,n}} - \tilde{p}_{k,n}\right) = P_T - \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{S}_k} \tilde{p}_{k,n}, \quad (77)$$

where μ_k and μ are used to denote the intermediate variables. The multi-level water-filling approach consists of two steps. Firstly, the power is allocated in order to satisfy the minimum rate requirement of the macro-cell UEs, where the allocated power in this step is $P_S = \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{S}_k} \tilde{p}_{k,n}$. Next, the remaining power is allocated in order to further improve the sum rate. Since quasi-concave function guarantees the existence of a unique maximum, we thus apply the gradient method to search for the optimal power. In particular, for a fixed subcarrier assignment set, gradient-based power adaptation can be used with single-UE water-filling in (74)-(75), and the multi-level water-filling in (76)-(77), where the power is updated using the gradient of RE

$$P_T(n) = P_T(n-1) + t \times \frac{d\lambda_{RE}(P_T)}{dP_T} \quad (78)$$

with t being the step size. Since the quasi-concavity property implies that $\lambda_{RE}(P_T)$ either strictly decreases or first increases and then strictly decreases with P_T , the proposed algorithm will terminate with either convergence or $P_0 = C_0^{-1}(\mathcal{S}_k, \gamma_k)$ if $\lambda_{RE}(P_T)$ is monotonically decreasing in $[P_0, P_{max}]$ and P_{max} if $\lambda_{RE}(P_T)$ is monotonically increasing in $[P_0, P_{max}]$.

We are now ready to investigate the subcarrier assignment strategy. *Theorem IV* in Section IV can be applied here. Specifically, we assign each subcarrier to the MUE that would achieve the highest SNR on that subcarrier.

$$\rho_k^{n*} = \begin{cases} 1, & \text{if } k = \arg \max_{k \in \mathcal{K}} \gamma_k^n \\ 0, & \text{otherwise} \end{cases} \quad (79)$$

Note that the proposed power allocation approach only aims to solve the RE optimization problem (60)-(64) with fixed normalizing factor α . Let $\lambda_{RE}(\alpha)$ denote the objective value of problem (60)-(64) with a given α . It is easy to see that $\lambda_{RE}(1)$ denotes the maximum RE value of problem (60)-(64) that only aims to maximize EE without taking SE into account. On the other hand, $\lambda_{RE}(0)$ corresponds to the case in which SE is maximized without taking EE into account. Therefore, based on this result, we further develop a bi-section approach to numerically search for the optimal value of α per described in Table II.

By employing the proposed bi-section based subcarrier

- 1) Initialize $\alpha_{min} = 0$ and $\alpha_{max} = 1$;
- 2) **REPEAT**;
- 3) Let $\alpha_{mid} = \frac{\alpha_{min} + \alpha_{max}}{2}$, perform the subcarrier assignment approach using (79);
- 4) **FOR** each UE $k \in \mathcal{K}$
- 5) Perform single-UE water-filling using (74)-(75) to obtain $\tilde{p}_{k,n}$ and μ_k , calculate the power consumption P_S ;
- 6) **END FOR**
- 7) **IF** $P_S > P_{max}$
- 8) infeasible;
- 9) **ELSE**
- 10) Initial Power $P_T(1) \in [P_S, P_{max}]$;
- 11) **REPEAT**
- 12) For the remaining power, perform the multi-level water-filling in (76)-(77);
- 13) The transmission power is updated using the gradient of RE in (78);
- 14) **UNTIL** when $|P_T(n) - P_T(n-1)| \leq \varepsilon$;
- 15) **END IF**
- 16) **IF** $\lambda_{RE}(\alpha_{mid}) \geq \lambda_{RE}(\alpha_{max})$, let $\alpha_{max} = \alpha_{mid}$, **OTHERWISE** $\alpha_{min} = \alpha_{mid}$;
- 17) **UNTIL** $|\alpha_{max} - \alpha_{min}| \leq \varepsilon$.

TABLE II
PROPOSED BI-SECTION-BASED SUBCARRIER ASSIGNMENT AND POWER ALLOCATION APPROACH

- 1) **FOR** $W_c = W_{min} : W_{tot}$
- 2) Perform the proposed bi-section based subcarrier assignment and power allocation approach;
- 3) Save the current RE value in the buffer;
- 4) **END FOR**
- 5) The maximum RE value and its corresponding bandwidth is determined.

TABLE III
PROPOSED RESOURCE EFFICIENCY MAXIMIZATION SCHEME

assignment and power allocation approach, the optimal RE can be obtained for a given bandwidth (given number of subcarriers). As a result, starting from $W_{min} = KW_C$ (i.e., each UE should be guaranteed at least one subcarrier), we apply the proposed gradient-based power adaptation and the subcarrier allocation policy to the current bandwidth setting, and store the optimal RE value in the buffer $\lambda_{RE}(W)$. Then, we increment the bandwidth using $W = W + W_C$. The proposed subcarrier assignment and power allocation approach is performed again to obtain the maximum RE of the macro-cell under the updated bandwidth. This procedure is repeated for all possible bandwidth options, i.e., from W_{min} to W_{tot} . Hence, the optimal RE and the corresponding bandwidth of the macro-cell is determined. The remaining bandwidth is then dedicated to the small-cell operation. A complete description of the algorithm can be found in Table III.

B. Energy-Efficiency Optimization for Small-Cells

Once the RE of the macro-cell is maximized, the remaining bandwidth is allocated exclusively to the small-cells. Since the

macro-cell and small-cells are not sharing the same spectrum, there is no inter-tier interference. Moreover, considering the maximum transmit power of the small-cells is usually low (small coverage), and small-cells are geographically separated, the intra-tier interference between small-cells is considerably small. Therefore, considering the effect of intra-tier interference as a raised thermal noise, we propose a low-complexity suboptimal resource allocation approach to maximize the EE of the small-cells. The SINR expression can therefore be rewritten as

$$\gamma_{[k,l]}^n = \frac{h_{[k,l]}^n p_l^n}{\bar{\sigma}_{[k,l]}^n}. \quad (80)$$

where $\bar{\sigma}_{[k,l]}^n$ is the combined noise and intra-tier interference. As a result, the optimization problem in (8)-(11) can be decomposed to a series of relatively isolated and simple optimization problems. In other words, one only needs to solve the EE maximization problem for each small-cell ($l \in \mathcal{L}$). This can be formulated as

$$\max_{\rho_{[k,l]}^n, p_l^n} \frac{\sum_{k \in \mathcal{K}_l} R_{[k,l]}}{\zeta P_T^{[l]} + P_C^{[l]}} \quad (81)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} p_l^n \leq P_{max}^{[l]}, \quad (82)$$

$$\sum_{n \in \mathcal{N}} \rho_{[k,l]}^n r_{[k,l]}^n \geq \delta_{[k,l]}, \quad \forall k \in \mathcal{K}_l, \quad (83)$$

$$\sum_{k \in \mathcal{K}_l} \rho_{[k,l]}^n = 1, \quad \forall n \in \mathcal{N}. \quad (84)$$

With a given total transmit power, P_T , the maximum EE can be written as $\lambda_{EE} \triangleq \frac{1}{(\alpha + (1-\alpha)\frac{\eta_P}{\eta_W})} \lambda_{RE}$. Since both P_T and W are constant, the maximum EE can be rewritten as $\lambda_{EE} = \nu \lambda_{RE}$. Therefore, the EE maximization problem in (81)-(84) is a special case of the RE maximization problem in (60)-(64). Therefore, the proposed gradient-based power adaptation and the subcarrier allocation policy can be applied here to solve the EE optimization problem.

As a result, we apply the proposed subcarrier allocation policy in (79) for the remaining bandwidth, and then perform the multi-level water-filling algorithm in (74)-(77) to obtain the optimal value. The power is then updated using the gradient of the EE as

$$P_T(n) = P_T(n-1) + t \times \frac{d\lambda_{EE}(P_T)}{dP_T} \quad (85)$$

where t is the step size. Since $\lambda_{EE}(P_T)$ either strictly decreases or first increases and then strictly decreases with P_T , the proposed approach will terminate with either convergence or $P_0 = \delta_k^{-1}(\mathcal{S}_k, \gamma_k)$ if $\lambda_{EE}(P_T)$ is monotonically decreasing in $[P_0, P_{max}]$ and P_{max} if $\lambda_{EE}(P_T)$ is monotonically increasing in $[P_0, P_{max}]$. This procedure is repeated for all the small-cells.

VI. SIMULATION RESULTS

In this section, we present numerical results in order to verify our theoretical findings and analyze the performance of the proposed underlay and overlay approaches in terms of EE. It is assumed that ten uniformly-distributed small-cells are

Central frequency of HCN	2 GHz
Subcarrier bandwidth, W_C	15 kHz
Number of small-cells	10
Number of UEs at the macro-cell, K_0	10
Number of UEs at the small-cells, K_l	3
Maximum transmit power of macro-cell, $P_{max}^{[0]}$	46 dBm
Maximum transmit power of small-cell, $P_{max}^{[l]}$	30 dBm
Path loss from macro-BS to UEs	$128.1 + 37.6 \log_{10} d_M$ (dB) [37]
Path loss from small-BS to UEs	$140.7 + 36.7 \log_{10} d_P$ (dB) [37]

TABLE IV
LIST OF SIMULATION PARAMETERS

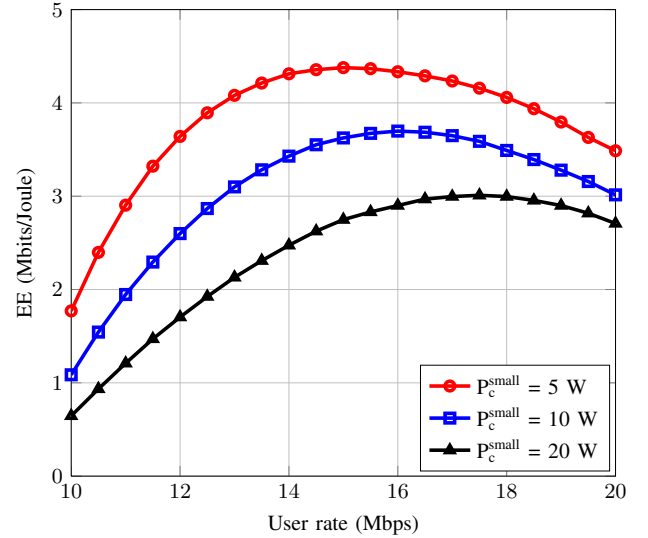


Fig. 1. EE versus SE curves with different small-cell circuit power parameters.

in the coverage area of a existing macro-cell, where ten and three uniformly-distributed UEs are serviced in the macro-cell and each small-cell, respectively. The radius of the macro-cell is set to 250 m, and that of the small-cells is set to 50 m. It should be noted that all results are obtained from various random locations of the UEs with identical and independent Rayleigh fading channels and Log-Normal shadowing with standard deviation of 8 dB. The minimum rate requirements per user for both macro-cell and small-cells are set to 10 Mbps. The drain efficiency of the power amplifier is 30% for our

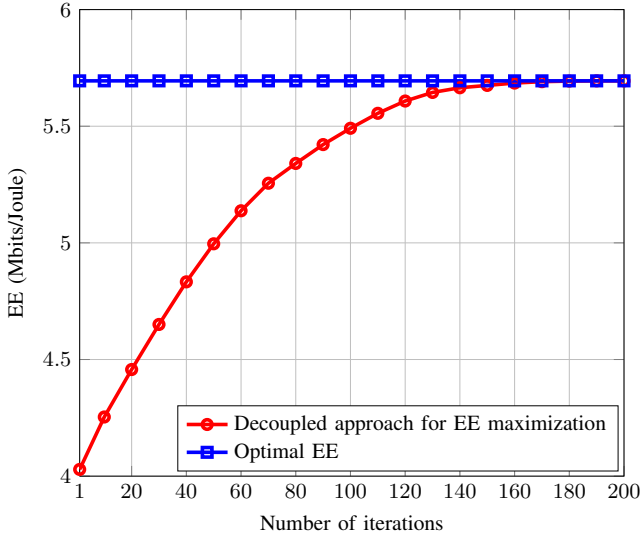


Fig. 2. The performance of the proposed dual-layer approach with respect to optimal EE.

simulation. Moreover, to balance the convergence speed and complexity, the stopping criteria ε and ϵ are set to 10^{-3} . Other simulation parameters are detailed in Table IV. In addition, these system parameters are merely chosen to demonstrate the EE optimization in an example and can easily be modified to any other values to address different scenarios.

In the first simulation, the performance of the proposed two solutions are studied. The EE-rate relationship (Theorem II) is first evaluated where the achievable rate of user under inequality constraint is varied. It has been shown in [36] that the EE in a MIMO-BC scenario converges after approximately 40 iterations when $\varpi = 2$, but is reduced to 17 iterations when a larger step size is chosen, e.g., $\varpi = 3$. Therefore, ϖ is set to 3 in our simulation in order to achieve a balance point between convergence speed, accuracy and complexity. It can be seen from Fig. 1 that the EE-rate relationship is quasi-concave and forms as a bell shape curve, where this quasi-concavity property is the basis of the proposed underlay-based approach. Furthermore, Fig. 1 also investigate the impact of circuit power on the EE-rate relationship. As anticipated, with increased circuit power, the corresponding optimal EE decreases due to higher power consumption. The performance of the proposed decomposition approach is then compared to the optimal EE. As it can be seen from Fig. 2, the proposed scheme successfully reaches the optimal EE after approximately 200 iterations. The validity of the proposed methodology is hence confirmed. On the other hand, Fig. 3 depicts the convergence behavior of the proposed overlay-based approach. It is observed that the optimal α in this case is very close to 0.5, this is inline with our original work on RE in [29] where the amount of reduction is only 4% on EE and 2% on SE for an equal weight.

Next, we evaluate the EE performance of the different schemes with spectrum underlay and overlay access. For the proposed RE-based overlay solution, we use a worst-case-method to model the interference. In particular, we assume

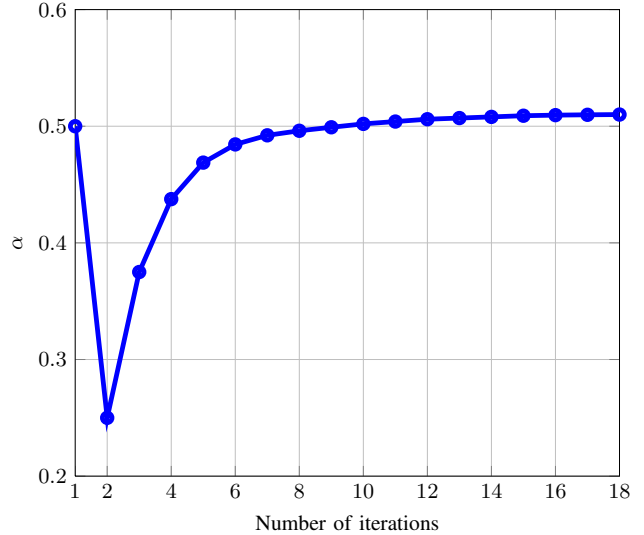


Fig. 3. Convergence analysis of the proposed bi-section based approach.

all the interfering small-cell BSs are operating using the maximum power budget (30 dBm) with equal power allocation over the remaining subcarriers. For comparison purposes, we compare the proposed schemes against the joint subcarrier assignment and power allocation scheme in [15]. The optimal EE is evaluated across the 0-20 dB CNR range. It should be noted that we vary the CNR by changing the channel gain where the results are averaged over 10000 randomly generated instances of channel. Therefore, the noise power is fixed in our simulation. In addition, the number of subcarrier is 1024 where the subcarrier bandwidth is 15 KHz, and the circuit power is set to 5 W for all cells. As it can be seen from Fig. 4, the EE achieved by the proposed bisection-based approach is very close to that of the proposed decomposition resource allocation approach whilst being much more efficient in terms of computational complexity. It is important to highlight, however, that the performance gap increases in high CNR region. This is because the overlay-based approach treats the intra-tier interference as noise, and will become dominant when the noise power diminishes at high CNR regime; hence resulting in reduced EE performance. Furthermore, both algorithms achieve higher EE compared to the scheme proposed in [15] which aims to maximize the sum rate.

The impact of the bandwidth on the optimal EE is illustrated in Fig. 5, where the circuit power is fixed to 5 W. Since the subcarrier bandwidth is fixed to 15 KHz, we vary the number of subcarriers in order to investigate the effect of the bandwidth. In addition, the CNR is set to 10 dB in this simulation. It can be seen that with a moderate number of subcarriers (small bandwidth), e.g., $N \leq 1000$, the EE achieved by the proposed overlay-based approach is lower than that of the proposed approach using underlay transmission. However this trend is reversed when the system has a larger bandwidth (e.g., $N \geq 1000$). The reason for this observation is that the proposed underlay-based approach allocates all available bandwidth to all cells. The excessive transmission-associated circuit power, which is modeled as a linear function

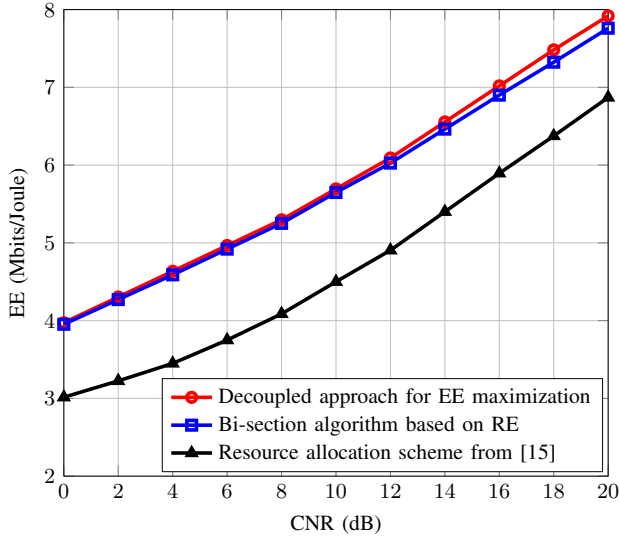


Fig. 4. Comparison of different subcarrier assignment and power allocation schemes in terms of EE.

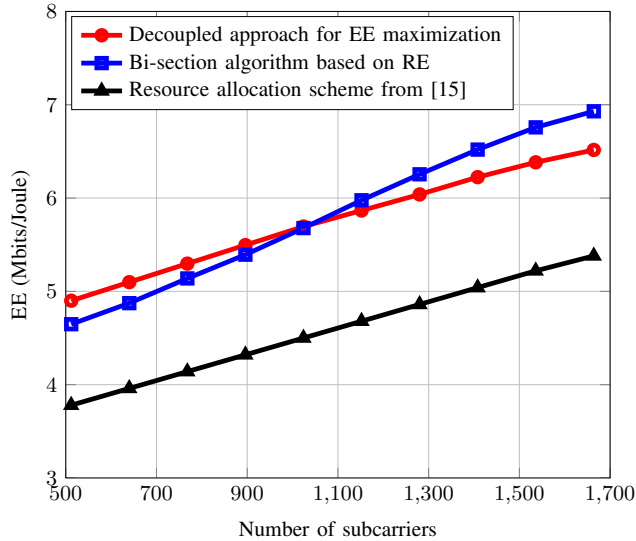


Fig. 5. Impact of the number of subcarriers on the EE performance of different schemes.

of the bandwidth, will reduce the EE performance in a system with large bandwidth. On the other hand, the proposed overlay-based approach allocates the exclusive spectrum parts to the macro-cell and small-cells and hence is more suitable for in the context of bandwidth-abundant HCNs.

Finally, the impact of small-cells density on the optimal EE is investigated in Fig. 6. In this simulation, the number of subcarrier is fixed to 1024 where the subcarrier bandwidth is 15 KHz, the circuit power is set to 5 W for all cells, and the CNR is set to 10 dB. As shown in the figure, with a lower small-cells density, e.g., $L \leq 10$, the EE achieved by the proposed bisection-based approach is lower than that of the proposed decomposed resource allocation approach. However, for the case of dense small-cells, e.g., $L \geq 10$, the EE achieved by the strategy using RE and overlay transmission is superior. This is because when the small-cells density is low,

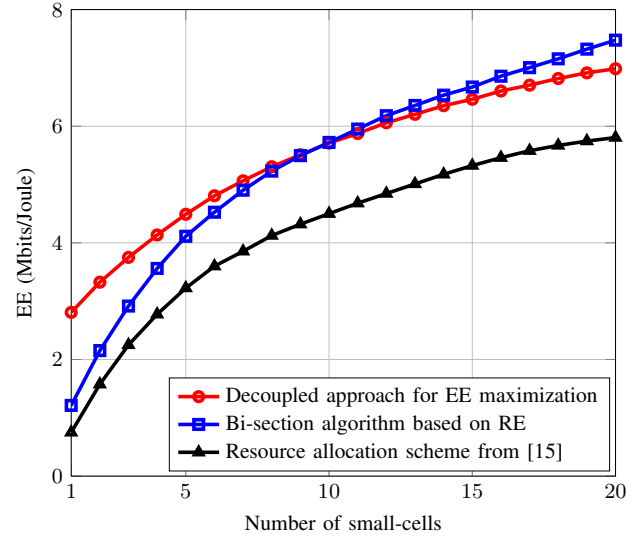


Fig. 6. Impact of small-cells density on the EE performance of different schemes.

the inter-tier interference will have less impact resulting in lower transmit power levels needed to satisfy the QoS targets. Therefore, the underlay transmission strategy is suitable under this setup. On the other hand, for the case of a dense deployed HCN, e.g., with $L \geq 10$, the inter-tier interference will become significant when spectrum is shared by the different tiers. Hence, extra power is required to maintain the throughput requirements of the UEs. Moreover, since macro-cell and small-cells occupy the whole bandwidth at the same time, the excessive transmission associated circuit power will further degrade EE performance. Consequently, the proposed overlay-based approach where different portions of the spectrum is allocated exclusively to the macro-cell and the small-cells is more suitable for dense multi-tier cellular environments.

VII. CONCLUSIONS

In this paper, we have addressed the EE optimization problem for OFDMA-based two-tier HCNs consisting of a macro-cell and multiple small-cells. Subcarrier assignment and power allocation policies were jointly investigated to optimize EE considering spectrum underlay and overlay access. Considering underlay transmission, where macro-cell and small-cells are sharing the available spectrum, we proved the relationship between EE and achievable rate is a quasi-concave function. On the basis of this property, we decomposed the original problem with multiple inequality constraints into multiple optimization problems with single inequality constraints. For each sub-problem, we separated the subcarrier assignment and power allocation process and developed an optimal solution based on difference-of-two-concave-functions approximation, successive convex approximation and gradient-search method. On the other hand, the underlay approach may not be energy-efficient due to severe inter-tier interference in a dense HCN scenario. In addition, it will lead to a higher power consumption in a bandwidth-abundant system and hence reduce the EE performance. Therefore, we developed a novel low-complexity

resource allocation scheme based on the idea of overlay transmission and RE. In this approach, we first optimized the RE of macro-cell and determined the optimal corresponding bandwidth, we then allocated the remaining bandwidth to small-cells and optimized the EE. Simulation results confirmed the theoretical findings and demonstrated that the proposed algorithms can efficiently approach the optimal EE.

APPENDIX A

PROOF OF THEOREM I

To prove $\lambda_{EE}^*(\mathbf{R})$ is a quasi-concave function, we denote the superlevel sets of $\lambda_{EE}^*(\mathbf{R})$ as

$$\mathcal{S}_\kappa = \{\mathbf{R} \geq \delta | \lambda_{EE}^*(\mathbf{R}) \geq \kappa\}. \quad (86)$$

In accordance with [32], for any real number κ , if the convexity for \mathcal{S}_κ is satisfied, $\lambda_{EE}^*(\mathbf{R})$ is strictly quasi-concave in \mathbf{R} . Therefore, we here divide the proof into two cases. For the case of $\kappa < 0$, since EE is always positive and hence there are no points on the counter, $\lambda_{EE}^*(\mathbf{R}) = \kappa$. For the case of $\kappa \geq 0$, λ_{EE} can be rewritten as

$$\lambda_{EE} = \frac{\sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}_l} R_{[k,l]}}{\zeta P_T(\mathbf{R}) + P_C}, \quad (87)$$

and hence \mathcal{S}_κ can be rewritten as $\kappa \zeta P_T(\mathbf{R}) + \kappa P_C - \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K}_l} R_{[k,l]} \leq 0$. In [33], it's been proved that $P_T(\mathbf{R})$ is convex in \mathbf{R} , therefore the convexity property of \mathcal{S}_κ holds and $\lambda_{EE}^*(\mathbf{R})$ is strictly quasi-concave in \mathbf{R} . This completes the proof of *Theorem I*. ■

APPENDIX B

PROOF OF THEOREM III

Proof: In order to prove *Theorem III*, we analyze the limit of $\lambda_{EE}^*(R_{[k^*,l^*]})$ as follows

$$\begin{aligned} & \lim_{R_{[k^*,l^*]} \rightarrow \infty} \lambda_{EE}^*(R_{[k^*,l^*]}) \\ &= \lim_{R_{[k^*,l^*]} \rightarrow \infty} \max_{\rho_{[k^*,l^*]}^n, \rho_{[k^*,l^*]}^l} \frac{\sum \sum_{l \in \mathcal{L}, k \in \mathcal{K}_l \setminus \{k^*,l^*\}} \bar{R}_{[k,l]} + R_{[k^*,l^*]}}{\zeta P_T^*(R_{[k^*,l^*]}) + P_C} \\ &= \lim_{R_{[k^*,l^*]} \rightarrow \infty} \frac{o(P_T^*(R_{[k^*,l^*]}))}{P_T^*(R_{[k^*,l^*]})} \\ &= 0. \end{aligned} \quad (88)$$

Hence, with strict concavity of $\lambda_{EE}^*(R_{[k^*,l^*]})$ which is proved in Appendix A, starting from $R_{[k^*,l^*]} = \delta_{[k^*,l^*]}$, $\lambda_{EE}^*(R_{[k^*,l^*]})$ either strictly decreases with $R_{[k^*,l^*]}$ if $\left. \frac{d\lambda_{EE}^*(R_{[k^*,l^*]})}{dR_{[k^*,l^*]}} \right|_{R_{[k^*,l^*]} = \delta_{[k^*,l^*]}} \leq 0$, or first strictly increases and then strictly decreases with $R_{[k^*,l^*]}$ if $\left. \frac{d\lambda_{EE}^*(R_{[k^*,l^*]})}{dR_{[k^*,l^*]}} \right|_{R_{[k^*,l^*]} = \delta_{[k^*,l^*]}} > 0$. The maximum EE in the rate region $[\delta_{[k^*,l^*]} \delta_{[k^*,l^*]}(max)]$ is straightforward as indicated in *Theorem III*. This completes the proof. ■

APPENDIX C

PROOF OF THEOREM IV

Suppose that when we obtain the optimal solution of problem in (33)-(34), where for cell $l \in \mathcal{L}$, subcarrier $n \in \mathcal{N}$ is allocated to UE $k \in \mathcal{K}_l \setminus \{k^*(n,l)\}$, where $k^*(n,l)$ represents the UE that consumes the lowest power on subcarrier n . However, to maintain the minimum rate demand, if n is instead allocated to $k^*(n,l)$ and $p_{k^*(n,l)}^n < p_{k,l}^n$, the interference received by the UEs that use n will be decreased, hence reducing the power consumption for all UEs. This statement contradicts the initial assumption that the optimal assignment for $n \in \mathcal{N}$ is allocated to UE $K \in \mathcal{K}_l \setminus \{k^*(n,l)\}$. Therefore, subcarrier n should be assigned to $k^*(n,l)$. This completes the proof of *Theorem IV*. ■

APPENDIX D

With a given subcarrier assignment set ρ , $\bar{\lambda}_{RE}(P_T)$ could be rewritten as $\alpha \frac{\bar{R}(P_T)}{P} + \nu \bar{R}(P_T)$, where $\nu = (1 - \alpha) \frac{W_{tot}}{W_{P_{tot}}}$. Under the water-filling approach, the transmit power on each subcarrier is non-decreasing. With the assumption that $\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{S}_k} \Delta p_{k,n} = \Delta P_T$, the existence of the limit reveals that $\bar{R}(P_T)$ is continuously differentiable and satisfies the following equation

$$\frac{d\bar{R}(P_T)}{dP_T} = \frac{d\bar{R}(P_T)}{dp_{k,n}} = \max_{k \in \mathcal{K}, n \in \mathcal{S}_k} \frac{W_C g_{k,n} \log_2 e}{1 + g_{k,n} \bar{p}_{k,n}}. \quad (89)$$

Furthermore, for the case of $k \in \mathcal{K}$ and $n \in \mathcal{S}_k$, $\frac{W_C g_{k,n} \log_2 e}{1 + g_{k,n} \bar{p}_{k,n}}$ is non-increasing with P_T whilst $\max_{k \in \mathcal{K}, n \in \mathcal{S}_k} \frac{W_C g_{k,n} \log_2 e}{1 + g_{k,n} \bar{p}_{k,n}}$ is decreasing with respect to P_T . Hence, we can conclude that $\frac{d^2 \bar{R}(P_T)}{dP_T^2} < 0$ and $\bar{R}(P_T)$ is a strictly concave function with P_T .

Similar to the proof in Appendix A, if the convexity for \mathcal{S}_θ holds, $\frac{\bar{R}(P_T)}{P_T}$ is strictly quasi-concave in P_T . Therefore, we here divide the proof into two cases. For the case of $\theta < 0$, since rate is always positive and hence there are no points on the counter, $\frac{\bar{R}(P_T)}{P_T} = \theta$. For the case of $\theta \geq 0$, \mathcal{S}_θ can be rewritten as $\mathcal{S}_\theta = \{P_T \geq \sum_{k \in \mathcal{K}} R^{-1}(\mathcal{S}_k, \delta_k) | \theta \zeta P_T + \theta P_C - \bar{R}(P_T)\} \leq 0$, where $R^{-1}(\mathcal{S}_k, \delta_k)$ denotes the minimum transmit power required to satisfy the rate demand δ_k . Therefore, given that the concavity of $\bar{R}(P_T)$ holds and \mathcal{S}_θ is strictly convex in P_T , $\frac{\bar{R}(P_T)}{P_T}$ is continuously differentiable and quasi-concave with respect to P_T , this concludes the proof of property (i).

Moreover, considering $\bar{\lambda}_{RE}(P_T)$ as $\alpha \frac{\bar{R}(P_T)}{P} + \nu \bar{R}(P_T)$, the derivative of RE $\frac{d\bar{\lambda}_{RE}(P_T)}{dP_T}$ should satisfy the following equation

$$\frac{d\bar{\lambda}_{RE}(P_T)}{dP_T} = \alpha \frac{d\frac{\bar{R}(P_T)}{P}}{dP_T} + \nu \frac{d\bar{R}(P_T)}{dP_T}. \quad (90)$$

Thus, based on (90), $\frac{d\bar{\lambda}_{RE}(P_T)}{dP_T}$ can be further constructed as follows

$$\begin{aligned} \frac{d\frac{\bar{R}(P_T)}{P}}{dP_T} &= \lim_{\Delta P_T \rightarrow 0} \frac{\frac{\bar{R}(P_T + \Delta P_T)}{\zeta(P_T + \Delta P_T) + P_C} - \frac{\bar{R}P_T}{\zeta P_T + P_C}}{\Delta P_T} \\ &= \frac{\frac{d\bar{R}(P_T)}{dP_T} - \zeta \bar{\lambda}_{EE}(P_T)}{\zeta P_T + P_C}. \end{aligned} \quad (91)$$

Hence, we observe

$$\frac{d\bar{\lambda}_{RE}(P_T)}{dP_T} = \frac{(\alpha + (1 - \alpha)\frac{\eta_P}{\eta_W})\frac{d\bar{R}(P_T)}{P_T} - \alpha\zeta\bar{\lambda}_{EE}(P_T)}{\zeta P_T + P_C} \quad (92)$$

where $\frac{d\bar{R}(P_T)}{dP_T} = \max_{k \in \mathcal{K}, n \in \mathcal{S}_k} \frac{W_C g_{k,n} \log_2 e}{1 + g_{k,n} \bar{p}_{k,n}}$, $\bar{\lambda}_{EE}(P_T) = \frac{\bar{R}(P_T)}{\zeta P_T + P_C}$. This concludes the proof of property (ii). ■

REFERENCES

- [1] C. W. Paper, "Cisco visual networking index: Global mobile data traffic forecast, 2015-2020," June 2016.
- [2] J. G. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.
- [3] A. Shojaefard, K. A. Hamdi, E. Alsusa, D. K. C. So, and J. Tang, "A unified model for the design and analysis of spatially-correlated load-aware HetNets," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 1–16, Nov. 2014.
- [4] D. Lopez-Perez, A. Valcarce, G. de la Roche, and J. Zhang, "OFDMA femtocells: A roadmap on interference avoidance," *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 41–48, Sep. 2013.
- [5] M. Kang, M. S. Alouini, and L. Yang, "Outage probability and spectrum efficiency of cellular mobile radio systems with smart antennas," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1871–1877, Dec. 2002.
- [6] H. Dai and H. Poor, "Asymptotic spectral efficiency of multi-cell MIMO systems with frequency-flat fading," *IEEE Trans. Signal Process.*, vol. 51, pp. 2976–2988, Nov. 2003.
- [7] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, 2002.
- [8] V. Chandrasekhar, M. Kountouris, and J. G. Andrews, "Coverage in multi-antenna two-tier networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5314–5327, Oct. 2009.
- [9] M. Peng, C. Wang, J. Li, H. Xiang, and V. Lau, "Recent advances in underlay heterogeneous networks: Interference control, resource allocation, and self-organization," *Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 700–729, Mar. 2015.
- [10] Y. Kim, T. Kwon, and D. Hong, "Area spectral efficiency of shared spectrum hierarchical cell structure networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4145–4151, Oct. 2010.
- [11] L. Giupponi and C. Ibars, "Distributed interference control in OFDMA based femtocells," in *Proc. IEEE PIMRC*, Sep. 2010, pp. 1201–1206.
- [12] L. Venturino, N. Prasad, and X. Wang, "Coordinated scheduling and power allocation in downlink multicell OFDMA networks," *IEEE Trans. Veh. Tech.*, vol. 58, no. 6, pp. 2835–2848, July 2009.
- [13] K. Son, S. Lee, Y. Yi, and S. Chong, "REFIM: a practical interference management in heterogeneous wireless access networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1260–1272, June 2011.
- [14] 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. 2366–2376, July 2014.
- [15] 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.
- [16] C. Xiong, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.
- [17] S. Buzzi, G. Colavolpe, D. Saturnino, and A. Zappone, "Potential games for energy-efficient power control and subcarrier allocation in uplink multicell OFDMA systems," *IEEE Journal of Topics in Signal Processing*, vol. 6, no. 2, pp. 89–103, Apr. 2012.
- [18] X. Xiao, X. Tao, and J. Lu, "QoS-aware energy-efficient radio resource scheduling in multi-user OFDMA systems," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 75–78, Jan. 2013.
- [19] C. Xiong, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3767–3778, Dec. 2012.
- [20] G. Lim, C. Xiong, L. J. C. Jr., and G. Y. Li, "Energy-efficient resource allocation for OFDMA-based multi-RAT networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2696–2705, May 2014.
- [21] S. Bu, F. Yu, and H. Yanikomeroglu, "Interference-aware energy-efficient resource allocation for OFDMA-based heterogeneous networks with incomplete channel state information," *IEEE Trans. Veh. Tech.*, vol. 64, no. 3, pp. 1036–1050, Mar. 2015.
- [22] R. Ramamonjison and V. K. Bhargava, "Energy efficiency maximization framework in cognitive downlink two-tier networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1468–1479, Mar. 2015.
- [23] W. S. Lai, T. H. Chang, and T. S. Lee, "Joint power and admission control for spectral and energy efficiency maximization in heterogeneous OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3531–3547, May 2016.
- [24] J. Tang, D. K. C. So, A. Shojaefard, K. K. Wong, and J. Wen, "Joint antenna selection and spatial switching for energy efficient MIMO SWIPT system," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 7, pp. 4754–4769, July 2017.
- [25] S. Zhang, Y. Chen, and S. Xu, "Improving energy efficiency through bandwidth, power, and adaptive modulation," in *Proc. Veh. Technol. Conf. (VTC)*, Ottawa, ON, Sep. 2010.
- [26] —, "Joint bandwidth-power allocation for energy efficient transmission in multi-user systems," in *Proc. IEEE Globecom Workshops*, Miami, FL, Dec. 2010.
- [27] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," *IEEE Commun. Surv. Tut.*, vol. 18, no. 2, pp. 1018–1044, SECOND QUARTER 2016.
- [28] G. K. Tran, H. Shimodaira, R. E. Rezagah, K. Sakaguchi, and K. Araki, "Practical evaluation of on-demand smallcell ON/OFF based on traffic model for 5G cellular networks," in *Proc. IEEE WCNC*, Apr. 2016.
- [29] J. Tang, D. K. C. So, E. Alsusa, and K. A. Hamdi, "Resource efficiency: A new paradigm on energy efficiency and spectral efficiency tradeoff," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4656–4669, Aug. 2014.
- [30] A. Shojaefard, K. A. Hamdi, E. Alsusa, D. K. C. So, and J. Tang, "Exact SINR statistics in the presence of heterogeneous interferers," *IEEE Trans. on Inf. Theory*, vol. 61, no. 12, pp. 6759–6773, Dec 2015.
- [31] G. AUER, V. Giannini, C. Desset, I. Gódor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, pp. 40–49, Oct. 2011.
- [32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.
- [33] J. Tang, D. So, E. Alsusa, K. A. Hamdi, and A. Shojaefard, "Resource allocation for energy efficiency optimization in heterogeneous networks," *IEEE Journal on Sel. Areas in Commun.*, vol. 33, no. 10, pp. 2104–2117, Oct. 2015.
- [34] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.
- [35] S. Boyd, *Branch and Bound Methods*. Stanford University.
- [36] J. Tang, D. So, E. Alsusa, K. A. Hamdi, and A. Shojaefard, "Energy efficiency optimization with interference alignment in multi-cell MIMO interfering broadcast channels," *IEEE Trans. Commun.*, vol. 63, no. 7, pp. 2486–2499, July 2015.
- [37] 3GPP, "Coordinated multi-point operation for LTE physical layer aspects (Rel. 11)," Feb. 2011.