

Energy-Efficient Resource Assignment and Power Allocation in Heterogeneous Cloud Radio Access Networks

Mugen Peng, *Senior Member, IEEE*, Kecheng Zhang, Jiamo Jiang, Jiaheng Wang, *Member, IEEE*, and Wenbo Wang, *Member, IEEE*

Abstract—Taking full advantage of both heterogeneous networks and cloud access radio access networks, heterogeneous cloud radio access networks (H-CRANs) are presented to enhance both spectral and energy efficiencies, where remote radio heads (RRHs) are mainly used to provide high data rates for users with high quality of service (QoS) requirements, whereas the high-power node (HPN) is deployed to guarantee seamless coverage and serve users with low-QoS requirements. To mitigate the intertier interference and improve energy efficiency (EE) performances in H-CRANs, characterizing user association with RRH/HPN is considered in this paper, and the traditional soft fractional frequency reuse (S-FFR) is enhanced. Based on the RRH/HPN association constraint and the enhanced S-FFR, an energy-efficient optimization problem with the resource assignment and power allocation for the orthogonal-frequency-division-multiple-access-based H-CRANs is formulated as a nonconvex objective function. To deal with the nonconvexity, an equivalent convex feasibility problem is reformulated, and closed-form expressions for the energy-efficient resource allocation solution to jointly allocate the resource block and transmit power are derived by the Lagrange dual decomposition method. Simulation results confirm that the H-CRAN architecture and the corresponding resource allocation solution can enhance the EE significantly.

Index Terms—Fifth-generation (5G), fractional frequency reuse (FFR), green communication, heterogeneous cloud radio access network (H-CRAN), resource allocation.

Manuscript received March 11, 2014; revised July 12, 2014, September 10, 2014, and November 25, 2014; accepted December 5, 2014. Date of publication December 12, 2014; date of current version November 10, 2015. The work of M. Peng, K. Zhang, and W. Wang was supported in part by the National Natural Science Foundation of China under Grant 61222103, by the National Basic Research Program of China (973 Program) under Grant 2013CB336600 and Grant 2012CB316005, by the State Major Science and Technology Special Projects under Grant 2013ZX03001001, by the Beijing Natural Science Foundation under Grant 4131003, and by the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20120005140002. The work of J. Wang was supported by the National Natural Science Foundation of China under Grant 61201174, by the Natural Science Foundation of Jiangsu Province under Grant BK2012325, and by the Fundamental Research Funds for the Central Universities. The review of this paper was coordinated by Dr. F. Gunnarsson.

M. Peng, K. Zhang, and W. Wang are with the Key Laboratory of Universal Wireless Communications for Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 215233, China (e-mail: pmg@bupt.edu.cn; buptzkc@163.com; wbwang@bupt.edu.cn).

J. Jiang is with the Institute of Communication Standards Research, China Academy of Telecommunication Research of MIIT, Beijing 100191, China (e-mail: jiamo.jiang@gmail.com).

J. Wang is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: jhwang@seu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2014.2379922

I. INTRODUCTION

CLOUD radio access networks (C-RANs) are being recognized to curtail capital and operating expenditures, as well as to provide a high transmission bit rate with fantastic energy efficiency (EE) performances [1], [2]. Remote radio heads (RRHs) operate as soft relay by compressing and forwarding the received signals from the mobile user equipment (UE) to the centralized baseband unit (BBU) pool through wire/wireless fronthaul links. To highlight the advantages of C-RAN, the joint decompression and decoding schemes are executed in the BBU pool [3]. However, the nonideal fronthaul with limited capacity and long time delay degrades performances of C-RANs. Furthermore, it is critical to decouple the control and user planes in C-RANs, and RRHs are efficient to provide high capacity without considering functions of control planes. How to alleviate the negative influence of the constrained fronthaul on EE performances and how to broadcast control signalings to UEs without RRHs are still not straightforward in C-RANs [4].

Meanwhile, high-power nodes (HPNs; e.g., macro or micro base stations) existing in heterogeneous networks (HetNets) are still critical to guarantee the backward compatibility with the traditional cellular networks and support the seamless coverage since low-power nodes (LPNs) are mainly deployed to provide high bit rates in special zones [5]. Under the help of HPNs, the unnecessary handover can be avoided, and the synchronous constraints among LPNs can be alleviated. Accurately, although the HetNet is a good alternative to improve both coverage and capacity simultaneously, there are two remarkable challenges to block its commercial applications: 1) Coordinated multipoint transmission and reception needs a huge number of signalings in backhaul links to mitigate the intertier interferences between HPNs and LPNs, whereas backhaul capacity is often constrained; 2) the ultradense LPNs improve capacity with the cost of consuming too much energy, which results in low EE performances.

To overcome these aforementioned challenges in both C-RANs and HetNets, a new architecture and technology known as the heterogeneous cloud radio access network (H-CRAN) is presented as a promising paradigm for future heterogeneous converged networks [6]. The H-CRAN architecture shown in Fig. 1 takes full advantage of both C-RANs and HetNets, where RRHs with low energy consumption cooperate with each other in the BBU pool to achieve high cooperative gains. The BBU pool is interfaced to HPNs for coordinating the

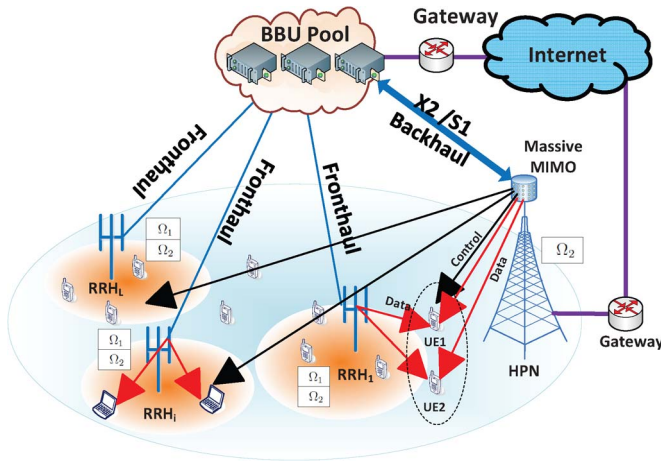


Fig. 1. System architecture of the proposed H-CRANs.

intertier interferences between RRHs and HPNs. Only the front radio frequency and simple symbol processing functionalities are configured in RRHs, whereas the other important baseband physical processing and procedures of the upper layers are executed in the BBU pool. By contrast, the entire communication functionalities from the physical to network layers are implemented in HPNs. The data and control interfaces between the BBU pool and HPNs are S1 and X2, respectively, which are inherited from definitions of the Third-Generation Partnership Project (3GPP) standards. Compared with the traditional C-RAN architecture in [1], H-CRANs alleviate the fronthaul constraints between RRHs and the BBU pool by incorporating HPNs. The control and broadcast functionalities are shifted from RRHs to HPNs, which alleviates capacity and time delay constraints on the fronthaul and supports the burst traffic efficiently. The adaptive signaling/control mechanism between connection-oriented and connectionless is supported in H-CRANs, which can achieve significant overhead savings in the radio connection/release by moving away from a pure connection-oriented mechanism.

In H-CRANs, the RRH/HPN association strategy is critical for improving EE performances, and main differences from the traditional cell association techniques are twofold. First, the transmit power of RRHs and HPNs is significantly different. Second, the inter-RRH interferences can be jointly coordinated through the centralized cooperative processing in the BBU pool, whereas the intertier interferences between HPNs and RRHs are severe and difficult to mitigate. Consequently, it is not always efficient for UEs to be associated with neighbor RRHs/HPNs via the strongest receiving power mechanism [7]. As shown in Fig. 1, though obtaining the same receiving power from RRH₁ and HPN, both UE1 and UE2 prefer to associate with RRH₁ because lower transmit power is needed and more radio resources are allocated from RRH₁ than those from the HPN. Meanwhile, the association with RRHs can decrease energy consumption by saving massive use of air conditioning.

Based on the aforementioned RRH/HPN association characteristics, the joint optimization solution for resource block (RB) assignment and power allocation to maximize EE per-

formances in the orthogonal frequency-division multiple access (OFDMA)-based H-CRANs is researched in this paper.

A. Related Work

OFDMA is a promising multiaccess technique for exploiting channel variations in both frequency and time domains to provide high data rates in the fourth-generation (4G) and beyond cellular networks. To be backward compatible with 4G systems, the OFDMA is adopted in H-CRANs by assigning different RBs to different UEs. Recently, the radio resource allocation (RA) to maximize the spectral efficiency (SE) or meet diverse quality of service (QoS) requirements in OFDMA systems has attracted considerable attention [8]–[11]. The relay selection problem in a network with multiple UEs and multiple amplify-and-forward relays is investigated in [8], where an optimal relay selection scheme whose complexity is quadratic in the number of UEs and relays is presented. Li *et al.* in [9] proposed an asymptotical resource allocation algorithm via leveraging the cognitive radio (CR) technique in open-access OFDMA femtocell networks. The resource optimization for spectrum sharing with interference control in CR systems is researched in [10], where the achievable transmission rate of the secondary user over Rayleigh channels subject to a peak power constraint at the secondary transmitter and an average interference power constraint at the primary receiver is maximized. In [11], the optimal power allocation for minimizing the outage probability in point-to-point fading channels with the energy-harvesting constraints is investigated.

The EE performance metric has become a new design goal due to the sharp increase of the carbon emission and operating cost of wireless communication systems [12]. The EE-oriented radio resource allocation has been studied in various networks [13]–[16]. In [13], the distributed power allocation for multicell OFDMA networks taking both EE and intercell interference mitigation into account is investigated, where a biobjective problem is formulated based on the multiobjective optimization theory. To maximize the average EE performance of multiple UEs each with a transceiver of constant circuit power, the power allocation, RB allocation, and relay selection are jointly optimized in [14]. The active number of subcarriers and the number of bits allocated to each RB at the source nodes are optimized to maximize EE performances in [15], where the optimal solution turns out to be a bidirectional waterfilling bit allocation to minimize the overall transmit power. To maximize EE performances under constraints of total transmit power and interference in CR systems, an optimal power allocation algorithm using equivalent conversion is proposed in [16].

Intuitively, the intercell or intertier interference mitigation is key to improving both SE and EE performances. Some advanced algorithms in HetNets, such as cell association and fractional frequency reuse (FFR), have been proposed in [17] and [18], respectively. In [19], a network utility maximization formulation with a proportional fairness objective is presented, where prices are updated in the dual domain via coordinate descent. In [20], energy-efficient cellular networks through the employment of the base station with sleep-mode strategies,

as well as small cells, are researched, and the corresponding tradeoff issue is discussed as well.

To the best of our knowledge, there is lack of solutions for maximizing the EE performance in H-CRANS. Particularly, the RRH/HPN association strategy should be enhanced from the traditional strongest received power strategy. Furthermore, the radio resource allocation to achieve an optimal EE performance in H-CRANS is still not straightforward. To deal with these problems, the joint optimization solution with the RB assignment and power allocation subject to the RRH/HPN association and interference mitigation should be investigated.

B. Main Contributions

The goal of this paper is to investigate the joint optimization problem with the RB assignment and power allocation subject to the RRH/HPN association and intertier interference mitigation to maximize EE performances in the OFDMA-based H-CRAN system. The EE performance optimization is highly challenging because the energy-efficient resource allocation in H-CRANS is a nonconvex objective problem. Different from the published radio resource optimization in HetNets and C-RANs, the characteristics of H-CRANS should be highlighted and modeled. To simplify the coordinated scheduling between RRHs and the HPN, an enhanced soft fractional frequency reuse (S-FFR) scheme is presented to improve performances of cell-center-zone UEs served by RRHs with individual spectrum frequency resources. The contributions of this paper can be summarized as follows.

- To overcome challenges in HetNets and C-RANs, H-CRANS are presented as cost-efficient potential solutions to improve spectral and energy efficiencies, in which RRHs are mainly used to provide high data rates for UEs with high-QoS requirements in the hot spots, whereas HPNs are deployed to guarantee seamless coverage for UEs with low-QoS requirements.
- To mitigate the intertier interference between RRHs and HPNs, an enhanced S-FFR scheme is proposed, where the total frequency band is divided into two parts. Only partial spectral resources are shared by RRHs and HPNs, whereas the other is solely occupied by RRHs. The exclusive RBs are allocated to UEs with high-rate-constrained QoS requirements, whereas the shared RBs are allocated to UEs with low-rate-constrained QoS requirements.
- An energy-efficient optimization problem with the RB assignment and power allocation under constraints of the intertier interference mitigation and RRH/HPN association is formulated as a nonconvex objective function. To deal with this nonconvexity, an equivalent convex feasibility problem is reformulated, based on which, an iterative algorithm consisting of both outer- and inner-loop optimizations is proposed to achieve the global optimal solution.
- We numerically evaluate EE performance gains of H-CRANS and the corresponding resource allocation optimization solution. Simulation results demonstrate that EE performance gain of the H-CRAN architecture over

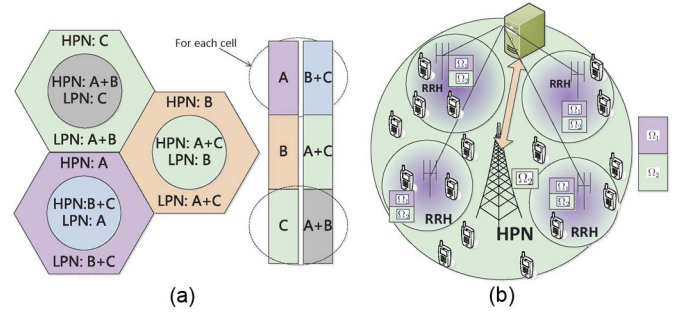


Fig. 2. Principle of the proposed enhanced S-FFR scheme. (a) Traditional S-FFR. (b) Enhanced S-FFR.

the traditional C-RAN/HetNet is significant. The proposed iterative solution is converged, and its EE performance gain over the baseline algorithms is impressive.

The remainder of this paper is organized as follows. In Section II, we describe the system model of H-CRANS, the proposed enhanced S-FFR, and the problem formulation. The optimization framework is introduced in Section III. Section IV provides simulations to verify the effectiveness of the proposed H-CRAN architecture and the corresponding solutions. Finally, we conclude this paper in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The traditional S-FFR is considered as an efficient intercell and intertier interference coordination technique, in which the service area is partitioned into spatial subregions, and each subregion is assigned to different frequency subbands. As shown in Fig. 2(a), the cell-edge-zone UEs do not interfere with cell-center-zone UEs, and the intercell interference can be suppressed with an efficient channel allocation method [21]. The HPN is mainly used to deliver the control signalings and guarantee the seamless coverage for UEs accessing the HPN (denoted by HUEs) with low-QoS requirements. By contrast, the QoS requirement for UEs accessing RRH (denoted by RUEs) is often with a higher priority. Consequently, as shown in Fig. 2(b), an enhanced S-FFR scheme is proposed to mitigate the intertier interference between HPNs and RRHs, in which only partial radio resources are allocated to both RUEs and HUEs with low-QoS requirements, and the remaining radio resources are allocated to RUEs with high-QoS requirements. In the proposed enhanced S-FFR, RUEs with low-QoS requirements share the same radio resources with HUEs, which is absolutely different from that in the traditional S-FFR. If the traditional S-FFR is utilized in H-CRANS, only the cell-center-zone RUEs share the same radio resources with HUEs, which decreases the SE performance significantly. Furthermore, it is challenging to judge whether UEs are located in the cell-edge or cell-center zone for the traditional S-FFR.

These aforementioned problems are avoided in the proposed enhanced S-FFR, where only the QoS requirement should be distinguished for RUEs. On the one hand, to avoid the intertier interference, the out-band frequency is preferred for use according to standards of HetNets in 3GPP [22], which suggests that RBs for HPNs should be orthogonal with those for RRHs. On

the other hand, to save the occupied frequency bands, the in-band strategy is defined as well in 3GPP [22], which indicates that both RUEs and HUEs share the same RBs although the intertier interference is severe. To be completely compatible with both in-band and out-band strategies in 3GPP, only two RB sets Ω_1 and Ω_2 are divided in this proposed enhanced S-FFR scheme. Obviously, if locations of RUEs could be known and the traffic volume in different zones is clearly anticipated, more RB sets in Ω_1 could be divided to achieve higher performance gains. The division of two RB sets in the proposed enhanced S-FFR is a good tradeoff between performance gains and implementing complexity/flexibility.

The QoS requirement is treated as the minimum transmission rate in this paper, which is also called the rate-constrained QoS requirement. In this paper, the high- and low-rate-constrained QoS requirements are denoted as η_R and η_{ER} , respectively. For simplicity, it is assumed that there are N and M RUEs per RRH occupying the RB sets Ω_1 and Ω_2 , respectively. In the OFDMA-based downlink H-CRANs, there are total K RBs (denoted as Ω_T) with bandwidth B_0 . These K RBs in Ω_T are categorized into two types: Ω_1 is only allocated to RUEs with high-rate-constrained QoS requirements, and Ω_2 is allocated to RUEs and HUEs with low-rate-constrained QoS requirements. Since all signal processing for different RRHs is executed on the BBU pool centrally, the inter-RRH interferences can be coordinated, and the same radio resources can be shared among RRHs. Hence, the channel-to-interference-plus-noise ratio (CINR) for the n th RUE occupying the k th RB can be divided into two parts, i.e.,

$$\sigma_{n,k} = \begin{cases} \frac{d_n^R h_{n,k}^R}{B_0 N_0}, & k \in \Omega_1 \\ \frac{d_n^R h_{n,k}^R}{(P^M d_n^M h_{n,k}^M + B_0 N_0)}, & k \in \Omega_2 \end{cases} \quad (1)$$

where d_n^R and d_n^M denote the path loss from the served RRH and the reference HPN to RUE n , respectively. $h_{n,k}^R$ and $h_{n,k}^M$ represent the channel gain from the RRH and HPN to RUE n on the k th RB, respectively. $P^M = P_{\max}^M/M$ is the allowed transmit power allocated on each RB in HPN, and P_{\max}^M denotes the maximum allowable transmit power of HPN. N_0 denotes the estimated power spectrum density of both the sum of noise and weak inter-RRH interference (in dBm/Hz).

The sum data rate for each RRH can be expressed as

$$C(\mathbf{a}, \mathbf{p}) = \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) \quad (2)$$

where $n \in \{1, \dots, N\}$ denotes the RUE allocated to the RB set Ω_1 , and $n \in \{N+1, \dots, N+M\}$ denotes the RUE allocated to the RB set Ω_2 . The $(N+M) \times K$ matrices $\mathbf{a} = [a_{n,k}]_{(N+M) \times K}$ and $\mathbf{p} = [p_{n,k}]_{(N+M) \times K}$ represent the feasible RB and power allocation policies, respectively. $a_{n,k}$ is defined as the RB allocation indicator, which can only be 1 or 0, indicating whether the k th RB is allocated to RUE n . $p_{n,k}$ denotes the transmit power allocated to RUE n on the k th RB.

According to [12], the total power consumption $P(\mathbf{a}, \mathbf{p})$ for H-CRANs mainly depends on the transmit power and circuit

power. When the power consumption for the fronthaul is considered, the total power consumption per RRH is written as

$$P(\mathbf{a}, \mathbf{p}) = \varphi_{\text{eff}} \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} p_{n,k} + P_c^R + P_{\text{bh}} \quad (3)$$

where φ_{eff} , P_c^R , and P_{bh} denote the efficiency of the power amplifier, circuit power, and power consumption of the fronthaul link, respectively.

Similarly, the sum data rate for the HPN can be calculated as

$$C_M(\mathbf{a}^M, \mathbf{p}^M) = \sum_{t=1}^T \sum_{m=1}^M a_{t,m} B_0 \log_2(1 + \sigma_{t,m} p_{t,m}) \quad (4)$$

where $t \in \{1, \dots, T\}$ denotes the HUE allocated to the RB set Ω_2 . The $T \times M$ matrices $\mathbf{a}^M = [a_{t,m}]_{T \times M}$ and $\mathbf{p}^M = [p_{t,m}^m]_{T \times M}$ represent the feasible RB and power allocation policies for the HPN, respectively. $a_{t,m}$ is defined as the RB allocation indicator, which can only be 1 or 0, indicating whether the m th RB is allocated to the HUE t . $p_{t,m}^m$ denotes the transmit power allocated to the HUE t on the m th RB. $\sigma_{t,m}$ represents the CINR of the t th HUE on the m th RB. The total power consumption of the HPN can be given by

$$P_M(\mathbf{a}^M, \mathbf{p}^M) = \varphi_{\text{eff}}^M \sum_{t=1}^T \sum_{m=1}^M a_{t,m} p_{t,m}^M + P_c^M + P_{\text{Mbh}} \quad (5)$$

where $p_{t,m}^M$ denotes the transmission baseband power for the HPN when the m th RB is allocated to the t th HUE, which forms the transmit power vector \mathbf{p}^M for all HUEs. φ_{eff}^M , P_c^M , and P_{Mbh} denote the efficiency of power amplifier, the circuit power, and the power consumption of the backhaul link between the HPN and BBU pool, respectively. Considering that the HPN is mainly utilized to extend the coverage and provide basic services for HUEs, the same downlink transmit power for different HUEs over different RBs is assumed as $p_{t,m}^M = P^M$, where P^M has been defined in (1).

The overall EE performance for the H-CRAN with L RRHs and 1 HPN can be written as

$$\gamma = \frac{L * C(\mathbf{a}, \mathbf{p}) + C_M(\mathbf{a}^M, \mathbf{p}^M)}{L * P(\mathbf{a}, \mathbf{p}) + P_M(\mathbf{a}^M, \mathbf{p}^M)}. \quad (6)$$

For the dense RRH deployed H-CRAN, L is much larger than 1. The intertier interference from RRHs to HUEs remains constant when the density of RRHs is sufficiently high, and hence, the downlink SE and EE performances for the HPN can be assumed to be stable. Therefore, $C(\mathbf{a}, \mathbf{p})$ is much larger than $C_M(\mathbf{a}^M, \mathbf{p}^M)$, and $P_M(\mathbf{a}^M, \mathbf{p}^M)$ can be ignored if $L * P(\mathbf{a}, \mathbf{p})$ is sufficiently large. When L is sufficiently large, the overall EE performance in (6) for the H-CRAN can be approximated as

$$\gamma \approx \frac{L * C(\mathbf{a}, \mathbf{p})}{L * P(\mathbf{a}, \mathbf{p})} = \frac{C(\mathbf{a}, \mathbf{p})}{P(\mathbf{a}, \mathbf{p})}. \quad (7)$$

According to (7), the overall EE performance mainly depends on the EE optimization of each RRH. To make sure that HUEs meet the low-rate-constrained QoS requirement, the interference from RRHs should be constrained and not larger

than the predefined threshold δ_0 . Consequently, to optimize downlink EE performances, the core problem is converted to optimize EE performance of each RRH with constraints on the intertier interference to the HPN from RRHs when the density of RRHs is sufficiently high.

Problem 1 (EE Optimization): With the constraints on the required QoS, intertier interference and maximum transmit power allowance, the EE maximization problem in the downlink H-CRAN can be formulated as

$$\max_{\{\mathbf{a}, \mathbf{p}\}} \frac{C(\mathbf{a}, \mathbf{p})}{P(\mathbf{a}, \mathbf{p})} = \frac{\sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k})}{\varphi_{\text{eff}} \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} p_{n,k} + P_c^R + P_{\text{bh}}} \quad (8)$$

$$\text{s.t.} \quad \sum_{n=1}^{N+M} a_{n,k} = 1, a_{n,k} \in \{0, 1\} \quad \forall k \quad (9)$$

$$\sum_{k=1}^K C_{n,k} \geq \eta_R, 1 \leq n \leq N \quad (10)$$

$$\sum_{k=1}^K C_{n,k} \geq \eta_{\text{ER}}, N+1 \leq n \leq N+M \quad (11)$$

$$\sum_{n=N}^{N+M} a_{n,k} p_{n,k} d_k^{\text{R2M}} h_k^{\text{R2M}} \leq \delta_0, k \in \Omega_{\text{II}} \quad (12)$$

$$\sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} p_{n,k} \leq P_{\text{max}}^R, p_{n,k} \geq 0 \quad \forall k, \forall n \quad (13)$$

where $C_{n,k} = a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k})$, and constraint (9) denotes the RB allocation limitation that each RB cannot be allocated to more than one RUE at the same time. The constraints of (10) and (11) corresponding to the high- and low-rate-constrained QoS requirements specify the minimum data rate of η_R and η_{ER} , respectively. Equation (12) puts a limitation on $p_{n,k}$ to suppress the intertier interference from RRHs to HUEs that reuse the RB $k \in \Omega_{\text{II}}$. d_k^{R2M} and h_k^{R2M} represent the corresponding path loss and channel gain on the k th RB from the reference RRH to the interfering HUE, respectively. In (13), P_{max}^R denotes the maximum transmit power of the RRH.

Based on the enhanced S-FFR scheme, the interference to HUE is constrained, and the SINR threshold η_{HUE} , which is the minimum SINR requirement for decoding the signal of HUE successfully, is defined to represent the constraint of (12). When allocating the m th RB to the t th HUE, the SINR ($\eta_{t,m}$) larger than η_{HUE} can be given by

$$\eta_{t,m} = P^M d_m^M h_{t,m}^M / (L * \delta_0 + B_0 N_0) \geq \eta_{\text{HUE}}. \quad (14)$$

Obviously, the optimal RB allocation policy \mathbf{a}^* and power allocation policy \mathbf{p}^* with constraints of diverse QoS requirements and variable η_{HUE} to maximize the EE performance in *Problem 1* is a nonconvex optimization problem due to forms

of the objective function and the RB allocation constraint in (9), whose computing complexity exponentially increases with the number of binary variables [23]. Intuitively, *Problem 1* is a mixed-integer programming problem, and the fractional objective makes it complicated and difficult to be solved directly with the classical convex optimization methods.

III. ENERGY-EFFICIENT RESOURCE ALLOCATION OPTIMIZATION

Here, we propose an effective method to solve *Problem 1* in (8), where we first exploit the nonlinear fractional programming for converting the objective function in *Problem 1*, upon which we then develop an efficient iterative algorithm to solve this EE performance maximization problem.

A. Optimization Problem Reformulation

Since the objective function in *Problem 1* is classified as a nonlinear fractional programming [24], the EE performance of the reference RRH can be defined as a nonnegative variable γ in (7) with the optimal value $\gamma^* = C(\mathbf{a}^*, \mathbf{p}^*) / P(\mathbf{a}^*, \mathbf{p}^*)$.

Theorem 1 (Problem Equivalence): γ^* is achieved if and only if

$$\max_{\{\mathbf{a}, \mathbf{p}\}} C(\mathbf{a}, \mathbf{p}) - \gamma^* P(\mathbf{a}, \mathbf{p}) = C(\mathbf{a}^*, \mathbf{p}^*) - \gamma^* P(\mathbf{a}^*, \mathbf{p}^*) = 0 \quad (15)$$

where $\{\mathbf{a}, \mathbf{p}\}$ is any feasible solution of Problem 1 to satisfy constraints (9)–(13). ■

Proof: See Appendix A. ■

Based on the optimal condition stated in **Theorem 1**, *Problem 1* is equivalent to *Problem 2* as follows if we can find the optimal value γ^* . Although γ^* cannot be directly obtained, an iterative algorithm (see **Algorithm 1** below) is proposed to update γ while ensuring that the corresponding solution $\{\mathbf{a}, \mathbf{p}\}$ remains feasible in each iteration. The convergence can be proved, and the optimal RA for solving *Problem 2* can be derived.

Algorithm 1 Energy-Efficient Resource Assignment and Power Allocation

- 1: Set the maximum number of iterations I_{max} , convergence condition ε_γ and the initial value $\gamma^{(1)} = 0$.
 - 2: Set the iteration index $i = 1$ and begin the iteration (Outer Loop).
 - 3: **for** $1 \leq i \leq I_{\text{max}}$
 - 4: Solve the resource allocation problem with $\gamma^{(i)}$ (Inner Loop);
 - 5: Obtain $\mathbf{a}^{(i)}, \mathbf{p}^{(i)}, C(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}), P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})$;
 - 6: **if** $C(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) - \gamma^{(i)} P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) < \varepsilon_\gamma$ **then**
 - 7: Set $\{\mathbf{a}^*, \mathbf{p}^*\} = \{\mathbf{a}^{(i)}, \mathbf{p}^{(i)}\}$ and $\gamma^* = \gamma^{(i)}$;
 - 8: **break**;
 - 9: **else**
 - 10: Set $\gamma^{(i+1)} = C(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) / P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})$ and $i = i + 1$;
 - 11: **end if**
 - 12: **end for**
-

Problem 2 (Transformed EE Optimization):

$$\begin{aligned} \max_{\{\mathbf{a}, \mathbf{p}\}} \quad & C(\mathbf{a}, \mathbf{p}) - \gamma^* P(\mathbf{a}, \mathbf{p}) \\ \text{s.t.} \quad & (9) - (13). \end{aligned} \quad (16)$$

Note that *Problem 2* is a tractable feasibility problem. ■

Hence, the objective function of the fractional form in (8) is transformed into the subtractive form. To design the efficient algorithm for solving *Problem 2*, we can define an equivalent function $F(\gamma) = \max_{\{\mathbf{a}, \mathbf{p}\}} C(\mathbf{a}, \mathbf{p}) - \gamma P(\mathbf{a}, \mathbf{p})$ with the following lemma.

Lemma 1: For all feasible \mathbf{a} , \mathbf{p} , and γ , $F(\gamma)$ is a strictly monotonic decreasing function in γ , and $F(\gamma) \geq 0$. ■

Proof: See Appendix B. ■

Due to the constraint of (9), the feasible domain of \mathbf{a} is a discrete and finite set consisting of all possible RB allocations, and thus, $F(\gamma)$ is generally a continuous but nondifferentiable function with respect to γ .

B. Proposed Iterative Algorithm

Based on **Lemma 1**, an iterative algorithm is proposed to solve the transformed *Problem 2* by updating γ in each iteration as **Algorithm 1**.

The proposed iterative **Algorithm 1** ensures that γ increases in each iteration. It can be observed that two nested loops executed in **Algorithm 1** can achieve the optimal solution to maximize EE performances. The outer loop updates $\gamma^{(i+1)}$ in each iteration with the $C(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})$ and $P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})$ obtained in the last iteration. In the inner loop, the optimal RB allocation policy $\mathbf{a}^{(i)}$ and power allocation policy $\mathbf{p}^{(i)}$ with a given value of $\gamma^{(i)}$ are derived by solving the following inner-loop *Problem 3*.

Problem 3 (Resource Allocation Optimization in the Inner Loop):

$$\begin{aligned} \max_{\{\mathbf{a}, \mathbf{p}\}} \quad & C(\mathbf{a}, \mathbf{p}) - \gamma^{(i)} P(\mathbf{a}, \mathbf{p}) \\ \text{s.t.} \quad & (9) - (13) \end{aligned} \quad (17)$$

where $\gamma^{(i)}$ is updated by the last iteration in the outer loop. ■

In fact, with the help of the proposed **Algorithm 1**, the solution to *Problem 3* converges, and the optimal solution is presented. The global convergence has the following theorem.

Theorem 2 (Global Convergence): Algorithm 1 always converges to the global optimal solution of Problem 3. ■

Proof: See Appendix C. ■

The optimization problem in (17) is nonconvex and hard to be solved directly. Generally speaking, if (17) can be solved by the dual method, there exists a nonzero duality gap [25]. The duality gap is defined as the difference between the optimal value of (17) (denoted by EE^*) and the optimal value of the dual problem for (17) (denoted by D^*). Fortunately, it can be demonstrated that the duality gap between (17) and the dual problem is nearly zero when the number of RBs is sufficiently large [26], which is illustrated as the following theorem.

Theorem 3 (Duality Gap): When the number of RBs is sufficiently large, the duality gap between (17) and its dual problem is nearly zero, i.e., $D^* - EE^* \approx 0$ holds. ■

Proof: See Appendix D. ■

C. Lagrange Dual Decomposition Method

Hence, according to **Theorem 3**, *Problem 3* in the i th outer loop can be solved by the dual decomposition method. With rearranging constraints (10)–(13), the Lagrangian function of the primal objective function is given by

$$\begin{aligned} L(\mathbf{a}, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) &= \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) \\ &\quad - \gamma^{(i)} \left(\varphi_{\text{eff}} \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} p_{n,k} + P_c^R + P_{\text{bh}} \right) \\ &\quad + \sum_{n=1}^N \beta_n \left[\sum_{k=1}^K a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) - \eta_R \right] \\ &\quad + \sum_{n=N+1}^{N+M} \beta_n \left[\sum_{k=1}^K a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) - \eta_{\text{ER}} \right] \\ &\quad + \sum_{k=1}^K \lambda_k \left(\delta_0 - \sum_{n=1}^{N+M} a_{n,k} p_{n,k} d_k^{\text{R2M}} h_k^{\text{R2M}} \right) \\ &\quad + \nu \left(P_{\text{max}}^R - \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} p_{n,k} \right) \end{aligned} \quad (18)$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_{N+M}) \succeq 0$ is the Lagrange multiplier vector associated with the required minimum data rate constraints (10) and (11). $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K) \succeq 0$ is the Lagrange multiplier vector corresponding to the intertier interference constraint (12), and $\lambda_k = 0$ for $k \in \Omega_I$. $\nu \geq 0$ is the Lagrange multiplier for the total transmit power constraint (13). The operator $\succeq 0$ indicates that the elements of the vector are all nonnegative.

The Lagrangian dual function can be expressed as

$$\begin{aligned} g(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) &= \max_{\{\mathbf{a}, \mathbf{p}\}} L(\mathbf{a}, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) \\ &= \max_{\{\mathbf{a}, \mathbf{p}\}} \left\{ \sum_{k=1}^K \sum_{n=1}^{N+M} [(\beta_n + 1) a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) \right. \\ &\quad \left. - \gamma^{(i)} \varphi_{\text{eff}} a_{n,k} p_{n,k} \right. \\ &\quad \left. - \lambda_k a_{n,k} p_{n,k} d_k^{\text{R2M}} h_k^{\text{R2M}} - \nu a_{n,k} p_{n,k}] \right. \\ &\quad \left. - \gamma^{(i)} (P_c^R + P_{\text{bh}}) - \sum_{n=1}^N \beta_n \eta_R \right. \\ &\quad \left. - \sum_{n=N+1}^{N+M} \beta_n \eta_{\text{ER}} + \sum_{k=1}^K \lambda_k \delta_0 + \nu P_{\text{max}}^R \right\} \end{aligned} \quad (19)$$

and the dual optimization problem is reformulated as

$$\begin{aligned} \min_{\{\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu\}} \quad & g(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) \\ \text{s.t.} \quad & \boldsymbol{\beta} \succeq 0, \boldsymbol{\lambda} \succeq 0, \nu \geq 0. \end{aligned} \quad (20)$$

It is obvious that the dual optimization problem is always convex. In particular, the Lagrangian function $L(\mathbf{a}, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \nu)$ is linear with β_n , λ_k , and ν for any fixed $a_{n,k}$ and $p_{n,k}$, whereas the dual function $g(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu)$ is the maximum of these linear functions. We use the dual decomposition method to solve this dual problem, which is first decomposed into K independent problems as

$$g(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) = \sum_{k=1}^K g_k(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) - \gamma^{(i)} (P_c^R + P_{bh}) - \sum_{n=1}^N \beta_n \eta_R - \sum_{n=N+1}^{N+M} \beta_n \eta_{ER} + \sum_{k=1}^K \lambda_k \delta_0 + \nu P_{\max}^R \quad (21)$$

where

$$g_k(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) = \max_{\{\mathbf{a}, \mathbf{p}\}} \left\{ \sum_{n=1}^{N+M} [(1 + \beta_n) a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) - \gamma^{(i)} \varphi_{\text{eff}} a_{n,k} p_{n,k} - \lambda_k a_{n,k} p_{n,k} d_k^{\text{R2M}} h_k^{\text{R2M}} - \nu a_{n,k} p_{n,k}] \right\}. \quad (22)$$

Supposed that the k th RB is allocated to the n th UE, i.e., $a_{n,k} = 1$, it is obvious that (22) is concave in terms of $p_{n,k}$. With the Karush–Kuhn–Tucker condition, the optimal power allocation is derived by

$$p_{n,k}^* = \left[\omega_{n,k}^* - \frac{1}{\sigma_{n,k}} \right]^+ \quad (23)$$

where $[x]^+ = \max\{x, 0\}$, and the optimal waterfilling level $\omega_{n,k}^*$ is derived as

$$\omega_{n,k}^* = \frac{B_0(1 + \beta_n)}{\ln 2 (\gamma^{(i)} \varphi_{\text{eff}} + \lambda_k d_k^{\text{R2M}} h_k^{\text{R2M}} + \nu)}. \quad (24)$$

Then, substituting the optimal power allocation obtained by (23) into the decomposed optimization problem (22), we can have

$$g_k(\boldsymbol{\beta}, \boldsymbol{\lambda}, \nu) = \max_{1 \leq n \leq N+M} \left\{ (1 + \beta_n) B_0 [\log_2(\omega_{n,k}^* \sigma_{n,k})]^+ - (\gamma^{(i)} \varphi_{\text{eff}} + \lambda_k d_k^{\text{R2M}} h_k^{\text{R2M}} + \nu) \times \left[\omega_{n,k}^* - \frac{1}{\sigma_{n,k}} \right]^+ \right\}. \quad (25)$$

With (24) and (25), the optimal RB allocation indicator for the given dual variables can be expressed as

$$a_{n,k}^* = \begin{cases} 1, & n = \arg \max_{1 \leq n \leq N+M} H_{n,k} \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

where

$$H_{n,k} = [(1 + \beta_n) \log_2(\omega_{n,k}^* \sigma_{n,k})]^+ - \frac{(1 + \beta_n)}{\ln 2} \left[1 - \frac{1}{\omega_{n,k}^* \sigma_{n,k}} \right]^+. \quad (27)$$

Then, the subgradient-based method [27] can be utilized to solve the given dual problem, and the subgradient of the dual function can be written as

$$\nabla \beta_n^{(l+1)} = \sum_{k=1}^K C_{n,k}^{(l)} - \eta_R, 1 \leq n \leq N \quad (28)$$

$$\nabla \beta_n^{(l+1)} = \sum_{k=1}^K C_{n,k}^{(l)} - \eta_{ER}, N+1 \leq n \leq N+M \quad (29)$$

$$\nabla \lambda_k^{(l+1)} = \begin{cases} 0, & \forall k \in \Omega_I \\ \delta_0 - \sum_{n=1}^N a_{n,k}^{(l)} p_{n,k}^{(l)} d_k^{\text{R2M}} h_k^{\text{R2M}}, & \forall k \in \Omega_{II} \end{cases} \quad (30)$$

$$\nabla \nu^{(l+1)} = P_{\max}^R - \sum_{n=1}^N \sum_{k=1}^K a_{n,k}^{(l)} p_{n,k}^{(l)} \quad (31)$$

where $a_{n,k}^{(l)}$ and $p_{n,k}^{(l)}$ represent the RB allocation and power allocation, which are derived by the dual variables of the l th iteration, respectively, and $C_{n,k}^{(l)} = a_{n,k}^{(l)} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}^{(l)})$. $\nabla \beta_n^{(l+1)}$, $\nabla \lambda_k^{(l+1)}$, and $\nabla \nu^{(l+1)}$ denote the subgradient utilized in the $(l+1)$ th inner-loop iteration. Hence, the update equations for the dual variables in the $(l+1)$ th iteration are given by

$$\beta_n^{(l+1)} = [\beta_n^{(l)} - \xi_\beta^{(l+1)} \times \nabla \beta_n^{(l+1)}]^+ \quad \forall n \quad (32)$$

$$\lambda_k^{(l+1)} = [\lambda_k^{(l)} - \xi_\lambda^{(l+1)} \times \nabla \lambda_k^{(l+1)}]^+ \quad \forall k \quad (33)$$

$$\nu^{(l+1)} = [\nu^{(l)} - \xi_\nu^{(l+1)} \times \nabla \nu^{(l+1)}]^+ \quad (34)$$

where $\xi_\beta^{(l+1)}$, $\xi_\lambda^{(l+1)}$, and $\xi_\nu^{(l+1)}$ are the positive step sizes.

IV. RESULTS AND DISCUSSIONS

Here, the EE performance of the proposed H-CRAN and corresponding optimization solutions are evaluated with simulations. There are $N = 10$ RUEs located in each RRH with the high-rate-constrained QoS and allocated by the orthogonal RB set Ω_1 . M is varied to denote the number of RUEs with low-rate-constrained QoS requirements. In the case of $M = 0$, there are no RUEs to share the same radio resources with HUEs, and thus, there are no intertier interferences. Otherwise, there are $M (> 0)$ RUEs with the low-rate-constrained QoS interfered with by the HPN. We assume that $d_n^P = 50$ m, $d_n^M = 450$ m in the case of $1 \leq n \leq N$, and $d_n^P = 75$ m, $d_n^M = 375$ m in the case of $N+1 \leq n \leq N+M$. The distance between the reference RRH and HUEs who reuse the k th RB is $d_k^{\text{R2M}} = 125$ m.

The number of total RBs is $K = 25$, and the system bandwidth is 5 MHz. The total transmit power of HPN is 43 dBm and equally allocated on all RBs. It is assumed that the path-loss model is expressed as $31.5 + 40.0 * \log_{10}(d)$ for the RRH-to-RUE link and $31.5 + 35.0 * \log_{10}(d)$ for the HPN-to-RUE and RRH-to-HUE links, where d denotes the distance between the transmitter and the receiver in meters. The number of simulation snapshots is set at 1000. In all snapshots, the fast-fading coefficients are all generated as independent and identically distributed Rayleigh random variables with unit variances. The low- and high-rate-constrained QoS requirements are assumed to be $\eta_{ER} = 64$ kbit/s and $\eta_R = 128$ kbit/s, respectively.

We assume the static circuit power consumption to be $P_c^R = 0.1$ W and the power efficiency to be $\varphi_{eff} = 2$ for the power amplifiers of RRHs. For HPNs, they are assumed to be $P_c^M = 10.0$ W and $\varphi_{eff}^M = 4$, respectively [28]. The power consumption for both RRHs and HPNs to receive the channel state indication and coordination signaling are taken into account, i.e., the power consumption of the fronthaul link P_{bh} , and the backhaul link between the HPN and the BBU pool P_{bh} are assumed to be 0.2 W.

A. H-CRAN Performance Comparisons

To evaluate EE performance gains of H-CRANs, the traditional C-RAN, two-tier HetNet, and one-tier HPN scenarios are presented as baselines. Only one HPN and one RRH are considered in the H-CRAN scenario. Similarly, only two HPNs are considered in the one-tier HPN scenario, and one HPN and one pico base station (PBS) are considered in the two-tier HetNet scenario. For the one-tier C-RAN scenario, two RRHs are considered, in which one RRH is used to replace the HPN in H-CRANs.

- One-tier HPN: All UEs access the HPN, and the RB set Ω_1 and Ω_2 are allocated to the cell-center-zone and cell-edge-zone UEs, respectively. The optimal power and RB allocations are based on the classical waterfilling and the maximum signal-to-interference-plus-noise ratio (SINR) scheduling algorithms, respectively. The number of cell-center-zone UEs is 10, and the number of cell-edge-zone UEs with the low-rate-constrained QoS requirement is varied from 1 to 10.
- Two-tier overlaid HetNet: The orthogonal RB set Ω_1 and Ω_2 are allocated to the PBS and HPN, respectively. The static circuit power consumption for the PBS is $P_c^P = 6.8$ W, and the power efficiency is $\varphi_{eff}^P = 4$ for the power amplifiers in PBS. Note that RBs allocated to the PBS and HPN are orthogonal with no intertier interferences. The number of RUEs served by the PBS keeps constant and is set at 10, and the number of HUEs served by HPN is varied. Note that UEs served by the HPN in this scenario are denoted by HUEs with low-rate-constrained QoS requirements to be consistent with other baselines.
- Two-tier underlaid HetNet: All RBs in the set Ω_1 and Ω_2 are fully shared by UEs within the covering areas of the PBS and HPN. The optimal power and RB allocations are adopted to enhance EE performances [9]. The locations

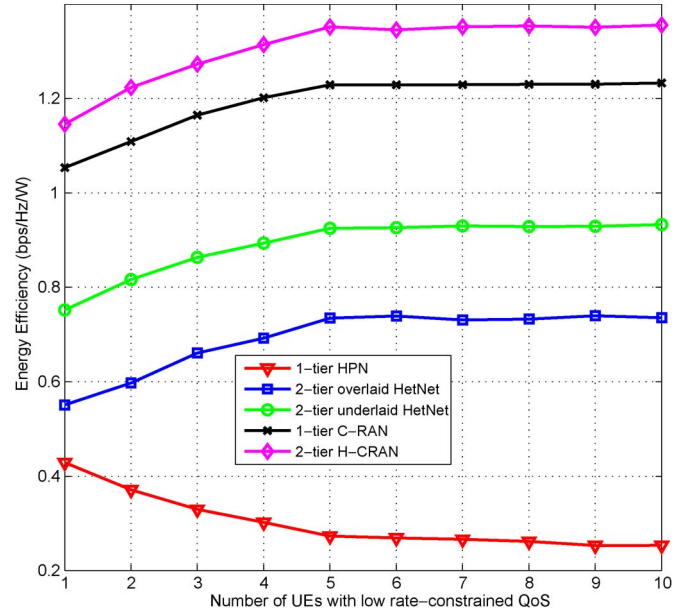


Fig. 3. Performance comparisons among one-tier HPN, two-tier HetNet, one-tier C-RAN, and two-tier H-CRAN.

and number of UEs served by the HPN and PBS are same as those in the two-tier overlaid HetNet scenario.

- One-tier C-RAN: All RBs in the set Ω_1 and Ω_2 are fully shared by RUEs. There are ten cell-center-zone RUEs and M cell-edge-zone RUEs in each RRH. Each RRH covers the same area coverage of the PBS in the two-tier HetNet scenario. The optimal power and RB allocations are based on the classical waterfilling and maximum SINR algorithms, respectively.
- Two-tier H-CRAN: The orthogonal RBs in set Ω_1 are only allocated to RUEs, whereas RBs in Ω_2 are shared by cell-edge-zone RUEs and HUEs, where the proposed optimal solution subject to the interference mitigation and RRH/HPN association is used. The number of cell-center-zone RUEs keeps constant and is set at 10, and the number of cell-edge-zone RUEs is varied to evaluate EE performances.

As shown in Fig. 3, EE performances are compared among the one-tier HPN, two-tier underlaid HetNet, two-tier overlaid HetNet, one-tier C-RAN, and two-tier H-CRAN. The EE performance decreases with the number of accessing UEs in the one-tier HPN scenario because more power is consumed to make cell-edge-zone UEs meet the low-rate-constrained QoS requirements. The EE performance becomes better in the two-tier HetNet than that in the one-tier HPN because lower transmit power is needed and higher transmission bit rate is achieved. Due to spectrum reuse and the fact that the interference is alleviated by the optimal solution in [9], the EE performance in the two-tier underlaid HetNet scenario is better than that in the two-tier overlaid HetNet scenario. Furthermore, the EE performance in the two-tier H-CRAN scenario is the best due to the gains from both the H-CRAN architecture and the corresponding optimal radio resource allocation solution. Note that the EE performance of one-tier C-RAN is a little worse than

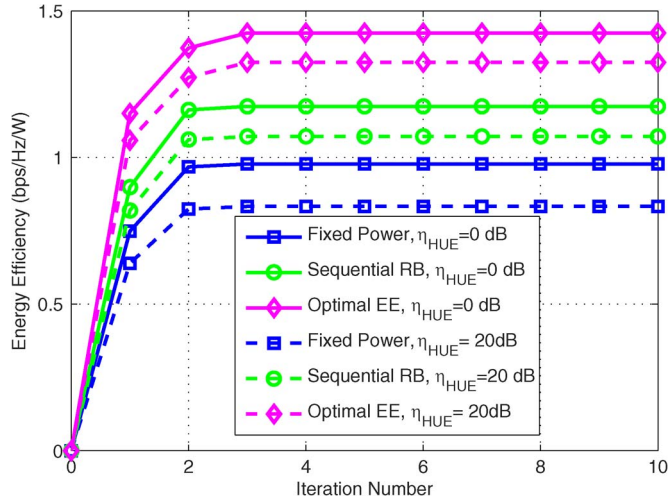


Fig. 4. Convergence evolution under different optimization solutions.

that of H-CRAN because it suffers from the coverage limitation when RRH covers the area that the HPN serves in the two-tier H-CRAN. Meanwhile, the EE performance of one-tier C-RAN is better than that of both one-tier HPN and two-tier HetNet due to the advantages of centralized cooperative processing and energy consumption savings.

B. Convergence of the Proposed Iterative Algorithm

To evaluate EE performances of the proposed optimal resource allocation solution (denoted by “optimal EE solution”), two algorithms are presented as baselines. The first baseline algorithm is based on the fixed power allocation (denoted by “fixed power”), i.e., the same and fixed transmit power is set for different RBs, and the optimal power allocation derived in (23) is not utilized. The second baseline algorithm is based on the sequential RB allocation (denoted by “sequential RB”), where the RB is allocated to RUEs sequentially. The presented two baseline algorithms use the same constraints of (9)–(13) as the optimal EE solution does. There are 12 RRHs (i.e., $L = 12$) uniformly spaced around the reference HPN.

The EE performances of these three algorithms with different allowed interference threshold under varied iteration numbers are shown in Fig. 4. The number of RUEs with low-rate-constrained QoS requirement is assumed as $M = 3$. It can be generally observed that the plotted EE performances are converged within three iteration numbers for different optimization algorithms. On the other hand, these two baseline algorithms are converged more quickly than the proposal does because the proposal has higher computing complexity. Furthermore, the maximum allowed intertier interference, which is defined as δ_0 in (12), should be constrained to make the HUE work efficiently. To evaluate δ_0 impacting on the EE performance, η_{HUE} described in (14) denoting the maximum allowed intertier interference is simulated. The large η_{HUE} indicates that the constraint of δ_0 should be controlled to a low level, which suggests decreasing the transmit power and even forbidding the RB to be allocated to RUEs. Consequently, the EE performance when η_{HUE} is 0 dB is better than that when it is 20 dB. The

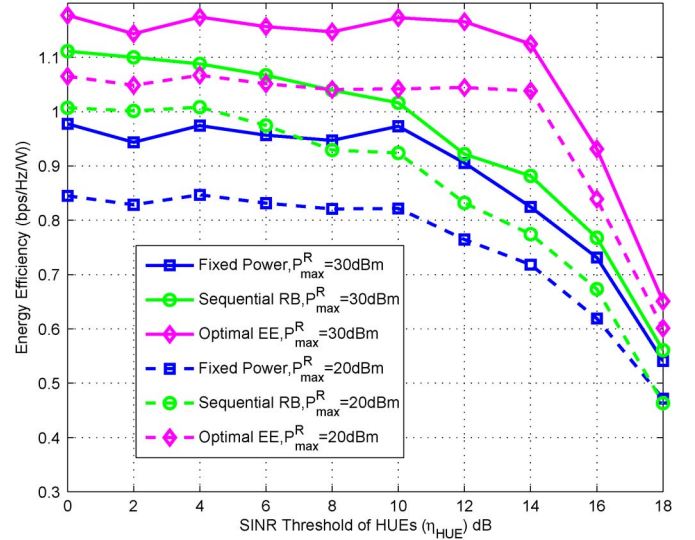
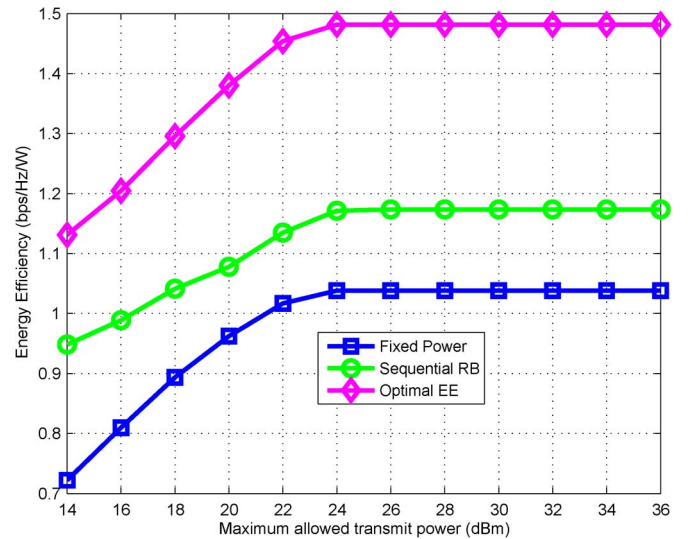

 Fig. 5. EE performance comparisons under different SINR thresholds of HUEs η_{HUE} .


Fig. 6. EE performance comparisons under different maximum allowed transmit power of RRHs.

EE performance decreases with the increasing SINR threshold η_{HUE} for these three algorithms.

C. EE Performances of the Proposed Solutions

Here, key factors impacting the EE performances of the proposed radio resource allocation solution are evaluated. Since EE performances are closely related to constraints (12)–(13), the variables η_{HUE} and P_{\max}^R are two key factors to be evaluated in Figs. 5 and 6, respectively. The ratio of Ω_1 to Ω_T impacting on EE performances is evaluated in Fig. 7 to show performance gains of the enhanced S-FFR.

In Fig. 5, EE performances under the varied SINR thresholds of HUEs η_{HUE} are compared among different algorithms when the maximum allowed transmit power of RRH is 20 or 30 dBm. When η_{HUE} is not sufficiently large, the EE performance

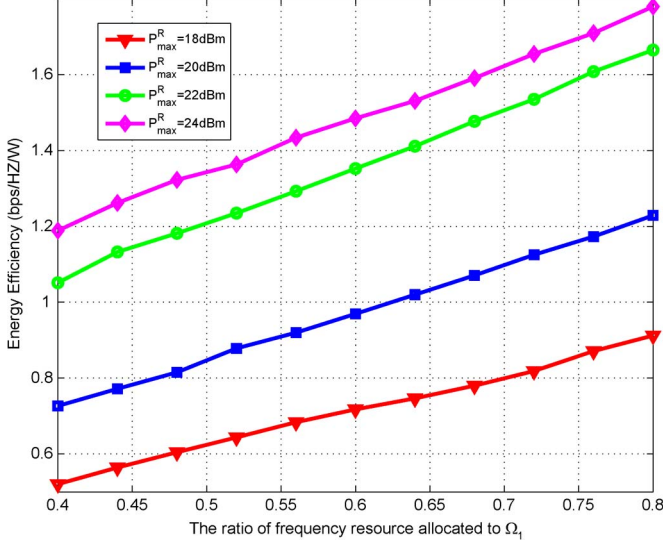


Fig. 7. EE performance comparisons for different ratios of Ω_1 to Ω_T .

almost keeps stable with the increasing η_{HUE} because the intertier interference is not severe due to adoption of the enhanced S-FFR. However, the EE performance declines for these two baseline algorithms when the SINR threshold η_{HUE} is over 10 dB. While the proposed optimal EE solution can sustain more intertier interferences, and the EE performance deteriorates after η_{HUE} is over 14 dB, indicating that the proposed solution can mitigate more intertier interferences and provide higher bit rates for HUEs than the other two baselines do. Summarily, the proposed optimal EE solution can achieve the best EE performance, whereas the “fixed power” algorithm has the worst EE performance, and the “sequential RB” algorithm is in-between. Furthermore, the EE performances for these three algorithms are strictly related to the maximum allowed transmit power of the RRH, and the EE performance with $P_{\max}^R = 30$ dBm is better than that with $P_{\max}^R = 20$ dBm.

To further evaluate the impact of the maximum allowed transmit power of RRHs P_{\max}^R on EE performances, Fig. 6 compares the EE performance of different algorithms in terms of P_{\max}^R . In this simulation case, η_{HUE} is set at 0 dB, the number of RUEs with low-rate-constrained QoS requirements is set at 4, and the iteration number is set at 5. The maximum allowed transmit power of the HPN is set at 43 dBm, and the maximum allowed transmit power of RRH varies within [14 dBm, 36 dBm] with the step size of 2 dBm. When P_{\max}^R is not large, EE performances increase almost linearly with the rising P_{\max}^R for all three algorithms. When $P_{\max}^R \geq 22$ dBm, both the SE performance and the total power consumption increases almost linearly with the rising P_{\max}^R . Therefore, the EE performance almost keeps stable. As shown in Fig. 6, EE performances of the “sequential RB” algorithm are often better than those of the “fixed power” algorithm due to the waterfilling power allocation gains indicated in (23). Furthermore, the proposed solution can achieve the best EE performance due to gains of RB assignment and power allocation.

In addition to η_{HUE} and P_{\max}^R , the ratio between Ω_1 and Ω_2 has a significant impact on the EE performance of H-CRANs.

The EE performances of the H-CRAN under the enhanced S-FFR with different ratios of Ω_1 to Ω_T are evaluated in Fig. 7, where the number of RUEs with low- and high-rate-constrained QoS requirements is set at $M = 5$ and $M = 10$, respectively. Meanwhile, the number of total RBs is $K = 25$, and the system bandwidth is 5 MHz. Each UE is allocated by at least one RB to guarantee the minimal QoS requirement. Fig. 7 suggests that the EE performance increases almost linearly with the ratio of Ω_1 to Ω_T . With more available exclusive RBs for RRHs, there are fewer intertier interferences because the shared RBs specified for RUEs and HUEs decrease. The increasing ratio of Ω_1 to Ω_T results in the drastically increasing of the expectation of SINRs for RUEs. Furthermore, since the intertier interference is mitigated by configuring few shared RBs, the transmission power of RUEs increases due to the proposed waterfilling algorithm in (23), which further improves the overall EE performances of H-CRANs. Moreover, the EE performance increases with the rising maximum allowed transmit power of RRH P_{\max}^R , which verifies the simulation results in Fig. 6 again. These results indicate that more radio resources should be configured for Ω_1 if only the EE performance optimization is pursued. However, the fairness of UEs should be considered jointly, and some necessary RBs should be fixed for both HUEs and cell-edge-zone RUEs to guarantee the seamless coverage and successful delivery of the control signalings to all UEs in the real H-CRANs, which is a challenging open issue for future research.

V. CONCLUSION

In this paper, the EE performance optimization for H-CRANs has been analyzed. In particular, the RB assignment and power allocation subject to the intertier interference mitigation and the RRH/HPN association have been jointly optimized. To deal with the optimization of resource allocations, a nonconvex fractional programming optimization problem has been formulated, and the corresponding Lagrange dual decomposition method has been proposed. Simulation results have demonstrated that performance gains of H-CRANs over the traditional HetNet and C-RAN are significant. Furthermore, the proposed optimal energy-efficient resource allocation solution outperforms the other two baseline algorithms. To maximize EE performances further, the advanced S-FFR schemes with more RB sets should be researched, and the corresponding optimal ratio of different RB sets should be designed in the future.

APPENDIX A PROOF OF THEOREM 1

By following a similar approach presented in [29], we prove **Theorem 1** with two septated steps.

First, the sufficient condition of **Theorem 1** should be proved. We define the maximal EE performance of *Problem 1* as $\gamma^* = C(\mathbf{a}^*, \mathbf{p}^*)/P(\mathbf{a}^*, \mathbf{p}^*)$, where \mathbf{a}^* and \mathbf{p}^* are the optimal RB and power allocation policies, respectively. It is obvious that γ^* holds, i.e.,

$$\gamma^* = \frac{C(\mathbf{a}^*, \mathbf{p}^*)}{P(\mathbf{a}^*, \mathbf{p}^*)} \geq \frac{C(\mathbf{a}, \mathbf{p})}{P(\mathbf{a}, \mathbf{p})} \quad (35)$$

where \mathbf{a} and \mathbf{p} are the feasible RB and power allocation policies for solving *Problem 1*. Then, according to (35), we can derive the following formulas:

$$\begin{cases} C(\mathbf{a}, \mathbf{p}) - \gamma^* P(\mathbf{a}, \mathbf{p}) \leq 0 \\ C(\mathbf{a}^*, \mathbf{p}^*) - \gamma^* P(\mathbf{a}^*, \mathbf{p}^*) = 0. \end{cases} \quad (36)$$

Consequently, we can conclude that $\max_{\{\mathbf{a}, \mathbf{p}\}} C(\mathbf{a}, \mathbf{p}) - \gamma^* P(\mathbf{a}, \mathbf{p}) = 0$, and it is achievable by the optimal resource allocation policies \mathbf{a}^* and \mathbf{p}^* . The sufficient condition is proved.

Second, the necessary condition should be proved. Supposed that $\hat{\mathbf{a}}^*$ and $\hat{\mathbf{p}}^*$ are the optimal RB and power allocation policies of the transformed objective function, respectively, we can have $C(\hat{\mathbf{a}}^*, \hat{\mathbf{p}}^*) - \gamma^* P(\hat{\mathbf{a}}^*, \hat{\mathbf{p}}^*) = 0$. For any feasible RB and power allocation policies \mathbf{a} and \mathbf{p} , they can be expressed as

$$C(\mathbf{a}, \mathbf{p}) - \gamma^* P(\mathbf{a}, \mathbf{p}) \leq C(\hat{\mathbf{a}}^*, \hat{\mathbf{p}}^*) - \gamma^* P(\hat{\mathbf{a}}^*, \hat{\mathbf{p}}^*) = 0. \quad (37)$$

The above inequality can be derived as

$$\frac{C(\mathbf{a}, \mathbf{p})}{P(\mathbf{a}, \mathbf{p})} \leq \gamma^* \text{ and } \frac{C(\hat{\mathbf{a}}^*, \hat{\mathbf{p}}^*)}{P(\hat{\mathbf{a}}^*, \hat{\mathbf{p}}^*)} = \gamma^*. \quad (38)$$

Therefore, the optimal resource allocation policies $\hat{\mathbf{a}}^*$ and $\hat{\mathbf{p}}^*$ for the transformed objective function are also the optimal resource allocation policies for the original objective function. The necessary condition of **Theorem 1** is proved.

APPENDIX B PROOF OF LEMMA 1

We can define an equivalent function as

$$F(\gamma) = \max_{\{\mathbf{a}, \mathbf{p}\}} C(\mathbf{a}, \mathbf{p}) - \gamma P(\mathbf{a}, \mathbf{p}) \quad (39)$$

and we assume that γ^1 and γ^2 are the optimal value for these two optimal RB allocation solution $\{\mathbf{a}^1, \mathbf{p}^1\}$ and $\{\mathbf{a}^2, \mathbf{p}^2\}$, where $\gamma^1 > \gamma^2$. Then

$$\begin{aligned} F(\gamma^2) &= C(\mathbf{a}^2, \mathbf{p}^2) - \gamma^2 P(\mathbf{a}^2, \mathbf{p}^2) \\ &> C(\mathbf{a}^1, \mathbf{p}^1) - \gamma^2 P(\mathbf{a}^1, \mathbf{p}^1) \\ &> C(\mathbf{a}^1, \mathbf{p}^1) - \gamma^1 P(\mathbf{a}^1, \mathbf{p}^1) = F(\gamma^1). \end{aligned} \quad (40)$$

Hence, $F(\gamma)$ is a strictly monotonic decreasing function in terms of γ .

Meanwhile, let \mathbf{a}' and \mathbf{p}' be any feasible RB and power allocation policies, respectively. If we set $\gamma' = C(\mathbf{a}', \mathbf{p}')/P(\mathbf{a}', \mathbf{p}')$, then

$$\begin{aligned} F(\gamma') &= \max_{\{\mathbf{a}, \mathbf{p}\}} C(\mathbf{a}, \mathbf{p}) - \gamma' P(\mathbf{a}, \mathbf{p}) \\ &\geq C(\mathbf{a}', \mathbf{p}') - \gamma' P(\mathbf{a}', \mathbf{p}') = 0. \end{aligned} \quad (41)$$

Hence, $F(\gamma) \geq 0$.

APPENDIX C PROOF OF THEOREM 2

Supposed that $\gamma^{(i)}$ and $\gamma^{(i+1)}$ represent the EE performance of the reference RRH in the i th and $(i+1)$ th iteration of the outer loop, respectively, where $\gamma^{(i)} > 0$, and $\gamma^{(i+1)} > 0$, neither of them is the optimal value γ^* . Meanwhile, $\gamma^{(i+1)}$ is given by $\gamma^{(i+1)} = C(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})/P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})$, where $\mathbf{a}^{(i)}$ and $\mathbf{p}^{(i)}$ are the optimal RB and power solutions of *Problem 3* in the i th iteration of the outer loop, respectively. Note that γ^* is defined as the achieved maximum EE performance for all feasible RA solutions $\{\mathbf{a}, \mathbf{p}\}$, and thus, $\gamma^{(i+1)}$ cannot be larger than γ^* , e.g., $\gamma^{(i+1)} \leq \gamma^*$. It has been proved that $F(\gamma) > 0$ in **Lemma 1** when γ is not the optimal value to achieve the maximum EE performance. Thus, $F(\gamma^{(i)})$ can be written as

$$\begin{aligned} F(\gamma^{(i)}) &= C(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) - \gamma^{(i)} P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) \\ &= P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) (\gamma^{(i+1)} - \gamma^{(i)}) > 0. \end{aligned} \quad (42)$$

Equation (42) indicates that $\gamma^{(i+1)} > \gamma^{(i)}$ due to $P(\mathbf{a}^{(i)}, \mathbf{p}^{(i)}) > 0$, which suggests that γ increases in each iteration of the outer loop in **Algorithm 1**. According to **Lemma 1**, the value of $F(\gamma)$ decreases with the increasing number of iterations due to the incremental value of γ .

On the other hand, it has been proved that the optimal condition $F(\gamma^*) = 0$ holds in **Theorem 1**. **Algorithm 1** ensures that γ monotonically increases. When the updated γ increases to the achievable maximum value γ^* , *Problem 2* can be solved with γ^* and $F(\gamma^*) = 0$. Then, the global optimal solutions \mathbf{a}^* and \mathbf{p}^* are derived. We update γ in the iterative algorithm to find the optimal value γ^* . It can be demonstrated that $F(\gamma)$ converges to zero when the number of iterations is sufficiently large and the optimal condition as stated in **Theorem 1** is satisfied. Therefore, the global convergence of **Algorithm 1** is proved.

APPENDIX D PROOF OF THEOREM 3

We can rewrite (17) as

$$\begin{aligned} C(\mathbf{a}, \mathbf{p}) - \gamma^{(i)} P(\mathbf{a}, \mathbf{p}) &= \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) \\ &\quad - \gamma^{(i)} \left(\varphi_{\text{eff}} \sum_{n=1}^{N+M} \sum_{k=1}^K a_{n,k} p_{n,k} + P_c^R + P_{\text{bh}} \right) \\ &= \sum_{k=1}^K \left(\sum_{n=1}^{N+M} a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) \right. \\ &\quad \left. - \sum_{n=1}^{N+M} \gamma^{(i)} \varphi_{\text{eff}} a_{n,k} p_{n,k} - \frac{\gamma^{(i)}}{K} (P_c^R + P_{\text{bh}}) \right). \end{aligned} \quad (43)$$

It is obvious that for the given RB allocation scheme, if letting

$$f_k(\mathbf{p}_{n,k}) = \sum_{n=1}^{N+M} a_{n,k} B_0 \log_2(1 + \sigma_{n,k} p_{n,k}) - \sum_{n=1}^{N+M} \gamma^{(i)} \varphi_{\text{eff}} a_{n,k} p_{n,k} - \frac{\gamma^{(i)}}{K} (P_c^R + P_{\text{bh}}) \quad (44)$$

then (43) can be written as $C(\mathbf{a}, \mathbf{p}) - \gamma^{(i)} P(\mathbf{a}, \mathbf{p}) = \sum_{k=1}^K f_k(\mathbf{p}_{n,k})$, where $\mathbf{p}_{n,k} \in \mathbb{C}^N$, and $f_k(\cdot) : \mathbb{C}^N \rightarrow \mathbb{R}$ is not necessarily convex. Similarly, constraints (9)–(13) can be expressed as the function of $\mathbf{p}_{n,k}$ with $\sum_{k=1}^K \mathbf{h}_k(\mathbf{p}_{n,k}) \leq \mathbf{0}$, where $\mathbf{h}_k(\cdot) : \mathbb{C}^N \rightarrow \mathbb{R}^L$, and L represents the number of constraints. Thus, (17) can be expressed as

$$EE^* = \max \sum_{k=1}^K f_k(\mathbf{p}_{n,k}) \quad (45)$$

$$\text{s.t.} \quad \sum_{k=1}^K \mathbf{h}_k(\mathbf{p}_{n,k}) \leq \mathbf{0} \quad (46)$$

where $\mathbf{0} \in \mathbb{R}^L$. To prove the duality gap between (17) and the optimal value of its dual problem D^* is zero, a perturbation function $v(\mathbf{H})$ is defined and can be written as

$$v(\mathbf{H}) = \max \sum_{k=1}^K f_k(\mathbf{p}_{n,k}) \quad (47)$$

$$\text{s.t.} \quad \sum_{k=1}^K \mathbf{h}_k(\mathbf{p}_{n,k}) \leq \mathbf{H} \quad (48)$$

where $\mathbf{H} \in \mathbb{R}^L$ is the perturbation vector.

Following [25], if $v(\mathbf{H})$ is a concave function of \mathbf{H} , the duality gap between D^* and EE^* is zero. Therefore, to prove the concavity of $v(\mathbf{H})$, a time-sharing condition should be demonstrated as follows.

Definition 1 (Time-Sharing Condition): Let $\mathbf{p}_{n,k}^{1*}$ and $\mathbf{p}_{n,k}^{2*}$ be the optimal solutions of (47) to $v(\mathbf{H}_1)$ and $v(\mathbf{H}_2)$, respectively. Equation (45) satisfies the time-sharing condition if for any $v(\mathbf{H}_1)$ and $v(\mathbf{H}_2)$, there always exists a solution $\mathbf{p}_{n,k}^3$ to meet

$$\sum_{k=1}^K \mathbf{h}_k(\mathbf{p}_{n,k}^3) \leq \alpha \mathbf{H}_1 + (1 - \alpha) \mathbf{H}_2 \quad (49)$$

$$\sum_{k=1}^K f_k(\mathbf{p}_{n,k}^3) \geq \alpha f_k(\mathbf{p}_{n,k}^{1*}) + (1 - \alpha) f_k(\mathbf{p}_{n,k}^{2*}) \quad (50)$$

where $0 \leq \alpha \leq 1$.

Now, we should prove that $v(\mathbf{H})$ is a concave function of \mathbf{H} . For some α , it is easy to find \mathbf{H}_3 satisfying $\mathbf{H}_3 = \alpha \mathbf{H}_1 + (1 - \alpha) \mathbf{H}_2$. Let $\mathbf{p}_{n,k}^{1*}$, $\mathbf{p}_{n,k}^{2*}$, and $\mathbf{p}_{n,k}^{3*}$ be the optimal solutions with these constraints of $v(\mathbf{H}_1)$, $v(\mathbf{H}_2)$, and $v(\mathbf{H}_3)$, respectively. According to the definition of the time-sharing condition, there exists $\mathbf{p}_{n,k}^3$ satisfying $\sum_{k=1}^K \mathbf{h}_k(\mathbf{p}_{n,k}^3) \leq \alpha \mathbf{H}_1 + (1 - \alpha) \mathbf{H}_2$ and $\sum_{k=1}^K f_k(\mathbf{p}_{n,k}^3) \geq \alpha f_k(\mathbf{p}_{n,k}^{1*}) + (1 - \alpha) f_k(\mathbf{p}_{n,k}^{2*})$. Since $\mathbf{p}_{n,k}^{3*}$ is the optimal solution to $v(\mathbf{H}_3)$, it is obvious that $\sum_{k=1}^K f_k(\mathbf{p}_{n,k}^{3*}) \geq \sum_{k=1}^K f_k(\mathbf{p}_{n,k}^3) \geq \alpha f_k(\mathbf{p}_{n,k}^{1*}) + (1 - \alpha) f_k(\mathbf{p}_{n,k}^{2*})$ holds. Then, the concavity of $v(\mathbf{H})$ is proved.

Since $v(\mathbf{H})$ is concave, (45) could be proved to satisfy the time-sharing condition. The time-sharing condition is always satisfied for the multicarrier system when the number of carriers goes to infinity [26], such as the OFDMA-based H-CRAN system in this paper. We can let $\mathbf{p}_{n,k}^1$ and $\mathbf{p}_{n,k}^2$ be two feasible power allocation solutions. There are α percentages of the total carriers to be allocated the power with $\mathbf{p}_{n,k}^1$, whereas the remaining $(1 - \alpha)$ percentages of the total carriers are allocated the power with $\mathbf{p}_{n,k}^2$. Then, $\sum_{k=1}^K f_k(\mathbf{p}_{n,k})$ is a linear combination, which is expressed as $\alpha f_k(\mathbf{p}_{n,k}^1) + (1 - \alpha) f_k(\mathbf{p}_{n,k}^2)$. Therefore, the constraints are linear combinations, and it is proved that (45) satisfies the time-sharing condition. Consequently, $v(\mathbf{H})$ is a concave function of \mathbf{H} , and the duality gap between D^* and EE^* should be zero. This theorem is proved.

REFERENCES

- [1] C. I. C. Rowell *et al.*, "Toward green and soft: A 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.
- [2] M. Peng and W. Wang, "Technologies and standards for TD-SCDMA evolutions to IMT-Advanced," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 50–58, Dec. 2009.
- [3] S. Park, O. Simeone, O. Sahin, and S. Shamai, "Robust and efficient distributed compression for cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 692–703, Feb. 2013.
- [4] M. Peng, Y. Li, T. Q. S. Quek, and W. Chonggang, "Device-to-Device underlaid cellular networks under Rician fading channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4247–4259, Aug. 2014.
- [5] M. Peng, D. Liang, Y. Wei, J. Li, and H.-H. Chen, "Self-configuration and self-optimization in LTE-Advanced heterogeneous networks," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 36–45, May 2013.
- [6] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.* [Online]. Available: <http://arxiv.org/abs/1410.3028>
- [7] M. Peng, S. Yan, and H. V. Poor, "Ergodic capacity analysis of remote radio head associations in cloud radio access networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 365–368, Aug. 2014.
- [8] S. Atapattu, Y. Jing, H. Jing, and C. Tellambura, "Relay selection and performance analysis in multiple-user networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1517–1529, Aug. 2013.
- [9] L. Li, C. Xu, and M. Tao, "Resource allocation in open access OFDMA femtocell networks," *IEEE Wireless Commun. Lett.*, vol. 1, no. 6, pp. 625–628, Dec. 2012.
- [10] X. Gong, A. Ispas, G. Dartmann, and G. Ascheid, "Power allocation and performance analysis in spectrum sharing systems with statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1819–1831, Apr. 2013.
- [11] C. Huang, R. Zhang, and S. Cui, "Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 1074–1087, Feb. 2014.
- [12] D. Feng *et al.*, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, First Quart. 2013.
- [13] Z. Fei, C. Xing, N. Li, and J. Kuang, "Adaptive multi-objective optimization for energy efficient interference coordination in multi-cell networks," *IET Commun.*, vol. 8, no. 8, pp. 1374–1383, May 2014. [Online]. Available: <http://arxiv.org/abs/1308.4777>
- [14] C. Ho and C. Huang, "Energy efficient resource block-allocation and relay selection scheme for OFDMA-based cooperative relay networks," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, pp. 1–6.
- [15] C. Sun, Y. Cen, and C. Yang, "Energy efficient OFDM relay systems," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1797–1809, May 2013.
- [16] Y. Wang, W. Xu, K. Yang, and J. Lin, "Optimal energy-efficient power allocation for OFDM-cognitive radio networks," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1420–1423, Sep. 2012.
- [17] D. Lopez-Perez, X. Chu, A. V. Vasilakos, and H. Claussen, "Power minimization based resource allocation for interference mitigation in OFDMA femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 333–344, Feb. 2014.

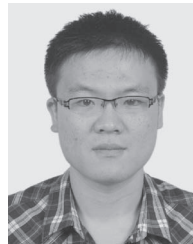
- [18] W. Pramudito and E. Alsusa, "A hybrid resource management technique for energy and QoS optimization in fractional frequency reuse based cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4948–4960, Dec. 2013.
- [19] K. Shen and W. Yu, "Downlink cell association optimization for heterogeneous networks via dual coordinate descent," in *Proc. IEEE ICASSP*, Vancouver, BC, Canada, May 2013, pp. 4779–4783.
- [20] Y. Soh, T. Quek, M. Kountouris, and S. Hyundong, "Energy efficient heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 840–850, May 2013.
- [21] N. Saquib, E. Hossain, and D. Kim, "Fractional frequency reuse for interference management in LTE-advanced hetnets," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 113–122, Apr. 2013.
- [22] "Technical report on UMTS heterogeneous networks; Study on UMTS heterogeneous networks (Release 12)," Third-Generation Partnership Project, Sophia-Antipolis Cedex, France, TR 25.800, V12.1.0, Dec. 2013. [Online]. Available: <http://www.3gpp.org/DynaReport/25800.htm>
- [23] C. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1982.
- [24] W. Dinkelbach, "On nonlinear fractional programming," *Manag. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
- [25] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE ISIT*, Jul. 2006, pp. 1394–1398.
- [26] W. Yu and R. Lui, "Dual methods of nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
- [27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [28] M. Imran and E. Ktranas, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," EARTH, Paris, France, Public Deliverable D2.3, Jan. 2012. [Online]. Available: <https://www.ict-earth.eu/publications/deliverables/deliverables.html>
- [29] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3204, Sep. 2012.



Mugen Peng (M'05–SM'11) received the B.E. degree in electronics engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2000 and the Ph.D. degree in communication and information system from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2005.

After completing the Ph.D. degree, he joined BUPT, where he has been a Full Professor with the School of Information and Communication Engineering since October 2012. In 2014, he became an Academic Visiting Fellow with Princeton University, Princeton, NJ, USA. He is leading a research group focusing on wireless transmission and networking technologies with the Key Laboratory of Universal Wireless Communications (Ministry of Education), BUPT. He has authored/coauthored over 40 refereed IEEE journal papers and over 200 conference proceeding papers. His main research areas include wireless communication theory, radio signal processing, and convex optimizations, with particular interest in cooperative communication, radio network coding, self-organization networking, heterogeneous networking, and cloud communication.

Dr. Peng is currently on the Editorial/Associate Editorial Board of the *IEEE Communications Magazine*, *IEEE Access*, the *International Journal of Antennas and Propagation (IJAP)*, *China Communication*, and the *International Journal of Communications System (IJCS)*. He has been the Guest Leading Editor for the Special Issues of *IEEE Wireless Communications*, *IJAP*, and the *International Journal of Distributed Sensor Networks*. He received the 2014 IEEE ComSoc AP Outstanding Young Researcher Award and the Best Paper Award at GameNets 2014, CIT 2014, ICCTA 2011, IC-BNMT 2010, and IET CCWMC 2009. He received the First Grade Award of Technological Invention Award from the Ministry of Education of China for the hierarchical cooperative communication theory and technologies and the Second Grade Award of Scientific and Technical Progress from the China Institute of Communications for the coexistence of multiradio access networks and 3G spectrum management.



Kecheng Zhang received the B.S. degree in telecommunication engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2012, where he is currently working toward the Ph.D. degree with the Key Laboratory of Universal Wireless Communications for the Ministry of Education.

His research interests include radio resource allocation optimization and cooperative communications in heterogeneous cloud radio access networks.



Jiamo Jiang received the Ph.D. degree in communication and information systems from Beijing University of Posts and Telecommunications, Beijing, China, in 2014.

He is currently with the Institute of Communication Standards Research, China Academy of Telecommunication Research of MIT, Beijing. His research focuses on self-organizing networks and energy-efficient resource allocation in heterogeneous networks.



Jiaheng Wang (S'08–M'10) received the B.E. and M.S. degrees from Southeast University, Nanjing, China, in 2001 and 2006, respectively, and the Ph.D. degree in electrical engineering from The Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2010.

He is currently an Associate Professor with the National Mobile Communications Research Laboratory, Southeast University. From 2010 to 2011, he was with the Signal Processing Laboratory, ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. He also held a visiting position with the Department of Computer and Information Science, University of Macau, Macau. His research interests mainly include optimization in signal processing, wireless communications, and networks.

Dr. Wang serves as an Associate Editor for the *IEEE SIGNAL PROCESSING LETTERS*. He received the Humboldt Fellowship for Experienced Researchers and the Best Paper Award at WCSP 2014.



Wenbo Wang (M'95) received the B.S., M.S., and Ph.D. degrees from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1986, 1989, and 1992, respectively.

He is currently the Dean of Telecommunication Engineering with BUPT. He is also the Assistant Director of the Academic Committee with the Key Laboratory of Universal Wireless Communication (Ministry of Education), BUPT. He has published more than 200 journal and international conference papers and six books and holds 12 patents. His research interests include radio transmission technology, wireless network theory, broadband wireless access, and software radio technology.