Optimal ADMM-Based Spectrum and Power Allocation for Heterogeneous Small-Cell Networks with Hybrid Energy Supplies

Li Ping Qian, *Senior Member, IEEE*, Yuan Wu, *Senior Member, IEEE*, Bo Ji, *Senior Member, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

Abstract—Powering cellular networks with hybrid energy supplies is not only environment-friendly but can also reduce the on-grid energy consumption, thus being emerging as a promising solution for green networking. Intelligent management of spectrum and power can increase the network utility in cellular networks with hybrid energy supplies, usually at the cost of higher energy consumption. Unlike prior studies on either the network utility maximization or on-grid energy cost minimization, this paper studies the joint spectrum and power allocation problem that maximizes the system revenue in a heterogeneous small-cell network with hybrid energy supplies. Specifically, the system revenue is considered as the difference between the network utility and on-grid energy cost. By developing the convexity of the optimization problem through transformation and reparameterization, we propose a joint spectrum and power allocation algorithm based on the primal-dual arguments to obtain the optimal solution by iteratively solving the primal and dual sub-problems of the convex optimization problem. To solve the primal sub-problem, we further propose the Lagrangian maximization based on the alternating direction method of multipliers (ADMM), and derive the optimal solution in the closed-form expression at each iteration. It is shown that the proposed joint spectrum and power allocation algorithm approaches the global optimality at the rate of $1/n$ with n being the number of iterations. Also, the proposed ADMM-based Lagrangian maximization algorithm approaches the primal optimal solution with the time complexity of $O(1/\epsilon_r)$ iterations with ϵ_r being the termination parameter. Simulation results show that in comparison with the power control with equal frequency allocation algorithm and frequency allocation with equal power allocation algorithms the proposed algorithm increases the system revenue by over 20 and 60 percent without consuming more on-grid energy when the proportional fairness utility and the weighted sum rate utility are considered with the approximate system parameter settings, respectively. Meanwhile, in comparison with the full frequency reuse case, the proposed algorithm increases the system revenue by 20 percent at least in terms of the weighted sum rate utility, although it achieves the similar system revenue when considering the proportional fairness utility. Simulation results also show that our proposed algorithm can perform well under the realistic fast fading channel conditions.

Index Terms—Resource allocation, hybrid energy supplies, heterogeneous small-cell networks, network utility maximization, energy cost minimization

1 Introduction

Mobile data traffic has been experiencing unprecedented growth due to the proliferation of smart devices and emerging social network services [1]. To accommodate the over-increasing demands for mobile traffic, various kinds of small-cell base stations, such as picocells, femtocells, and perhaps relay base stations, have been densely underlaid in a macro-cellular network to offload data from conventional macro-cell base stations [2], [3], [4]. Despite the improved spectrum efficiency, the massive deployment of small-cell base stations has been accompanied by huge energy consumption and considerable greenhouse gas (such as carbon) emissions [5]. On the other hand, due to the remote locations, some small-cell base stations in rural areas cannot be powered by on-grid energy [6]. To this end, energy harvesting (EH) technology has been introduced into small-cell networks for improving spectrum and energy efficiency [7], [8], [9], [10]. Specifically, besides the conventional on-grid energy supply, small-cell base stations are equipped with EH devices (like solar panels or wind turbines) and harvest renewable energy (such as solar, wind, thermoelectric, electro-mechanical, and ambient radio frequency energy) as supplementary or alternative power sources [11].

Renewable energy, if properly managed, can be a sustainable and economical alternative. However, its scheduling poses three significant challenges for network operation and management. First, unlike the conventional on-grid power supply, renewable energy arrives randomly and intermittently in spatio-temporal domains, depending on the weather conditions and the locations of EH small-cell base stations

- *L.P. Qian is with the College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China. E-mail: lpqian@zjut.edu.cn.*
- *Y. Wu is with the State Key Laboratory of Internet of Things for Smart City, University of Macau, Macao, China, and also with the Department of Computer and Information Science, University of Macau, Macao, China. E-mail: yuanwu@um.edu.mo.*
- *B. Ji is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122 USA. E-mail: boji@temple.edu.*
- *X. Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada. E-mail: xshen@bbcr.uwaterloo.ca.*

Manuscript received 4 June 2018; revised 25 Sept. 2019; accepted 13 Oct. 2019. Date of publication 17 Oct. 2019; date of current version 7 Jan. 2021. (Corresponding author: Yuan Wu.)
Digital Object Identifier no. 10.1109/TMC.2019.2948014

[13]. Thus, it is of practical importance to efficiently manage energy utilization adaptive to the battery storage of renewable energy. Second, a heterogeneous small-cell network consists of various base stations with hybrid energy supplies. There exist conventional on-grid small-cell base stations (CSBSs), off-grid small-cell base stations powered solely by harvested renewable energy (RSBSs), and hybrid small-cell base stations (HSBSs) jointly powered by harvested renewable energy and on-grid energy. This implies that we have to take into account the individual operating characteristics of each small-cell base station when designing the efficient energy management strategy. Finally, due to the non-uniform distribution of mobile users, the traffic load might not be in accordance with the harvested energy status and spectrum resource of each small-cell station. It implies that the energy scheduling needs to be carried out according to the status of individual small-cell base stations (BSs). To address these challenges, on one hand, most previous works in the literature focus on the energy scheduling for on-grid energy saving [14], [15], [16], [17], [18] or network utility maximization [19], [20], [21], [22], [23], [24] through offloading data traffic among distinct small-cell BSs. It is worth noting that the network utility increases along with consuming more on-grid energy, which needs to be purchased at a certain electricity price. To guarantee the greenness brought by hybrid energy supplies, it is desirable to carefully evaluate the balance between the network utility and the cost of on-grid energy purchase. On the other hand, most previous works focus on the renewable energy scheduling, rather than the joint scheduling of on-grid energy and renewable energy. Apparently, it is of more practical meaning for the joint scheduling in the heterogeneous small-cell network with hybrid energy supplies. However, due to the intermittent nature of renewable energy, we have to consider when and how long the HSBSs can schedule the renewable energy, which renders the scheduling problem non-convex. This implies the joint scheduling problem is much more complicated due to the additional freedom in the time domain.

In this paper, we study the optimal management of energy and spectrum resource among base stations for the heterogeneous small-cell network with hybrid energy supplies, which consists of multiple mobile users, multiple diverse small-cell base stations (SBSs), and a macro-cell base station (MBS). The goal is to strike the balance between the network utility and the on-grid energy cost while meeting the minimum data rate requirements of individual mobile users. To this end, the optimal spectrum and power allocation can be obtained through the convexification and transformation of the optimization problem. Furthermore, all BSs can iteratively update the spectrum and power allocation of all served MUs in a distributed manner. Consequently, the data traffic can be dynamically offloaded among MBS and diverse SBSs based on the real-time information of traffic intensity, renewable energy availability, and electricity price. In particular, the main contributions of this paper are summarized as follows:

- *Problem formulation:* We formulate the joint spectrum and power allocation problem to maximize the system revenue (equal to the difference between the network utility and on-grid energy cost) under the constraints of minimum data rate requirements of individual mobile users, as well as limited renewable energy storage and spectrum provisioning at each base station. In doing this, we can strike a balance between the network utility and on-grid energy cost through adjusting the trade-off coefficient.
- *Efficient algorithm design:* We first explore the convex nature of the original optimization problem through reparameterization, and transform it into a linear-equality-constrained convex optimization problem through introducing auxiliary variables. Then, we propose the optimal spectrum and power allocation algorithm based on the primal-dual arguments to obtain the optimal solution by iteratively solving the primal and dual sub-problems of the equivalent transformed optimization problem. With the introduction of auxiliary variables, the primal sub-problem is a linear-equality-constrained convex optimization problem, and thus we propose the alternating direction method of multipliers (ADMM) based Lagrangian maximization algorithm, and derive the optimal solution in the closed-form expression at each iteration. In doing this, all BSs can update the spectrum and power allocation for all served MUs in a distributed manner.
- *Convergence and complexity analysis:* We theoretically prove that the proposed joint spectrum and power allocation algorithm approaches the global optimality at the rate of $1/n$ with n being the number of iterations. Also, we show that the proposed ADMM-based Lagrangian maximization algorithm can approach the primal optimal solution with the time complexity of $O(1/\epsilon_r)$ iterations, where ϵ_r is the termination parameter. The simulation results further show that the tradeoff between system performance and convergence time is adjustable through tuning the error tolerance of algorithms.

The remainder of this paper is organized as follows. In Section 2, we present related works in the existing literature. In Section 3, we introduce the hybrid energy supplying small-cell network model and present the problem formulation. In Section 4, we convexify the optimization problem through transformation and reparameterization, and then propose an efficient algorithm to obtain the optimal spectrum and power allocation based on the primal-dual arguments. In Section 5, the Lagrangian maximization based on ADMM is presented to solve the primal sub-problem of the optimization problem. In Section 6, the performance of the proposed algorithm is evaluated through extensive simulations. Finally, we conclude the paper in Section 7, and provide detailed proofs in Section 8.

Throughout the paper, vectors are denoted in bold small letters, e.g., $\boldsymbol{x}$, with x_i being its ith component. Sets are denoted by Euler letters, e.g., $\mathcal{S}$. The operation $||\cdot||_2$ means the euclidean norm, $(\cdot)^T$ denotes transpose, and $[x]^+$ means $\max\{x, 0\}$.

2 Related Work

In this section, we elaborate on the efforts spent on the data offloading for hybrid energy supplying small-cell networks, in terms of on-grid energy saving and utility maximization, respectively.

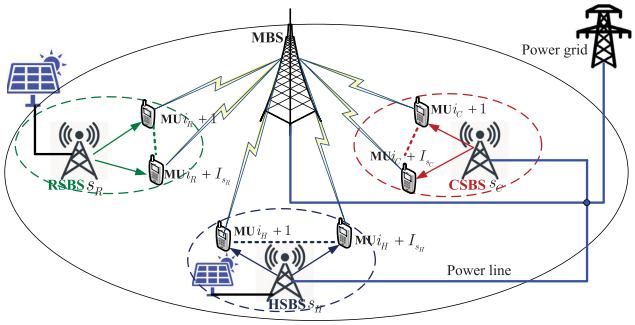


Fig. 1. A downlink two-tier heterogeneous small-cell network with hybrid energy supplies. Here, we assume that I_{s_C} MUs, I_{s_R} MUs and I_{s_H} MUs are served by CSBS s_C, RSBS s_R, and HSBS s_H, respectively.

2.1 On-Grid-Energy-Saving Data Offloading

In the literature, a large body of research works has been dedicated to the on-grid energy saving due to the data offloading [14], [15], [16], [17], [18]. In [14], the optimal utilization of harvested energy among BSs jointly powered by harvested energy and on-grid energy was studied under the constraint of minimum SINR (Signal-to-Interference-plus-Noise Ratio) requirements. The management of the on-grid energy saving and network delay was performed by downlink power control and user association reconfiguration for the cellular network with hybrid energy supplies in [15]. The authors of [16] proposed the energy-aware traffic offloading scheme to maximize the on-grid power saving of hybrid energy supplying heterogeneous network by optimizing user associations, ON-OFF states of small-cell base stations, and power control based on the statistical information of energy arrival and traffic load. Online and offline power allocation policies were studied for the throughput maximization and the grid energy minimization of a multiple-input-multiple-output wireless link with hybrid energy supplies in [17]. The authors in [18] proposed the Lyapunov optimization-based base station assignment and power control algorithm to minimize the sum of the on-grid energy cost and the packet drop cost for hybrid energy supplying networks. However, these works mainly focus on the grid energy saving, rather than the network utility maximization.

2.2 Utility-Maximization Data Offloading

There are also some work concerning the network-utility-maximization solution to improve the spectrum-efficiency (SE) and energy-efficiency (EE) [19], [20], [21], [22], [23], [24], [25]. In [19], the authors proposed a fully distributed online algorithm based on the primal decomposition theory to obtain the optimal user association and resource allocation that maximizes the proportional fairness utility for hybrid-energy-powered heterogeneous networks. The authors of [20] proposed a network utility aware traffic load balancing scheme to strike a tradeoff between the green power utilization and the traffic delivery latency by optimizing the user association for backhaul-constrained cache-enabled small cell networks powered by hybrid energy sources. In [21], the network layer delay was minimized with the optimal physical layer power allocation for energy harvesting cognitive radio networks. In addition, the energy efficiency was considered as the utility metric in [22], [23], [24], [25]. For example, the authors of [22] proposed online and offline resource allocation schemes to maximize the energy efficiency for the OFDMA system with hybrid energy harvesting base stations. The authors of [23] proposed a policy-gradient-based actor-critic reinforcement learning algorithm to maximize the energy efficiency of the hybrid energy supplying heterogeneous network. The energy efficiency maximization was also considered in [24]. The authors of [25] minimized the non-renewable energy usage through the joint design of transmit power allocation matrices and the receiver decoding policy. These work mainly studied how to perform the resource management for network utility maximization in wireless networks with hybrid energy supplies.

However, towards green communications, it is of practical importance to evaluate the trade-off between the network utility and on-grid energy cost. Also, existing works focused on the renewable energy scheduling, rather than the joint scheduling of on-grid energy and renewable energy. Apparently, it is of more practical meaning to address the joint scheduling in the heterogeneous small-cell network with hybrid energy supplies.

3 System Model and Problem Formulation

3.1 System Model

We consider a downlink two-tier heterogeneous small-cell network as shown in Fig. 1, which consists of one macro-cell base station, a set $\mathcal{S} = \{1, 2, \ldots, S\}$ of small-cell base stations, and a set $\mathcal{I} = \{1, 2, \ldots, I\}$ of mobile users (MUs). The set of SBSs can be divided into three subsets based on the energy supply. In particular, we use subsets $\mathcal{S}_C = \{1, 2, \ldots, S_C\}$, $\mathcal{S}_R = \{S_C + 1, S_C + 2, \ldots, S_C + S_R\}$, and $\mathcal{S}_H = \{S_C + S_R + 1, S_C + S_R + 2, \ldots, S\}$ to denote the set of conventional on-grid SBSs (CSBSs) powered solely by on-grid power, the set of off-grid SBSs (RSBSs) powered solely by harvesting renewable ambient energy (such as solar, wind, thermoelectric, electromechanical, and ambient radio frequency energy), and the set of hybrid SBSs (HSBSs) powered jointly by off-grid power and on-grid power, respectively.

Each base station individually allocates its spectrum to all MUs in service in the orthogonal manner. The spectrum resource block assigned to MBS is orthogonal to that of each SBS in each time slot. The SBSs are assigned individual orthogonal spectrum resource blocks in a time slot, and thus they do not interfere with each other during the transmission. In the downlink communication, each mobile user is allowed to download data symbols from MBS and SBSs simultaneously in length-T time slots.[1] Each RSBS or HSBS first stores the harvested energy in its battery, and then consumes it in the following time slots. We assume that in each time slot t, $E_{s,t}$ joules of harvested energy[2] stored in the battery of RSBS s or HSBS s can be used for the offset of site power consumed by the baseband processor, the cooling system and etc, and the data transmission [16]. Since we consider the optimal resource management in each time slot

1. Considering the block-fading channels, every channel state remains constant during every time slot, and thus the slot length T is generally set to be less than the coherence time of channel fading.

2. In time slot t, the amount of harvested energy stored in the battery (i.e., $E_{s,t}$) is equal to the total amount of energy harvested in time slot $t-1$ and battery energy not used in time slot $t-1$.

individually, we remove the subscript index t for the notational brevity in the following.

3.2 Problem Formulation

We use W_B and W_s to denote the total bandwidth occupied by the MBS and SBS s, respectively. We assume that mobile users communicate with the base station in the frequency division duplex (FDD) mode. Specifically, when mobile user i is served by the MBS (or SBS s), it is assigned a fraction f_{Bi} (or f_{si}) of bandwidth. It is clear that f_{Bi}'s and f_{si}'s need to satisfy the bandwidth limitation, respectively, i.e.,

$$0 \leq \sum_{\forall i \in \mathcal{I}} f_{Bi} \leq 1, \tag{1}$$

$$0 \leq \sum_{\forall i \in \mathcal{I}} f_{si} \leq 1, \forall s \in \mathcal{S}. \tag{2}$$

Let g_{Bi} and g_{si} denote the channel gain between the MBS and mobile user i and the channel gain between SBS s and mobile user i, respectively, which are determined by various factors such as path loss and fading effects. We use p_{Bi} and p_{si} to respectively denote the transmit power used by the MBS and SBS s for the transmission towards mobile user i, with $P_{B,\max}$ and $P_{s,\max}$ being the maximum allowable value. By the Shannon-capacity formula, mobile user u can obtain the throughput of R_{Bi} from the MBS satisfying

$$R_{Bi} = W_B f_{Bi} \log\left(1 + \frac{p_{Bi} g_{Bi}}{W_B f_{Bi} n_0}\right), \tag{3}$$

and the throughput of R_{si} from SBS s satisfying

$$R_{si} = W_s f_{si} \log\left(1 + \frac{p_{si} g_{si}}{W_s f_{si} n_0}\right), \tag{4}$$

where n_0 is the noise power density at the side of mobile user i. Assume that all MUs reside in the coverage area of MBS, and all MUs can be served by MBS. Due to the limited coverage area of each SBS, let $\mathcal{S}_i$ denote the possible set of all SBSs serving MU i, where each MU selects SBS s as a candidate of association station according to the strength of pilot signals from the SBS. As a result, the total throughput obtained by mobile user i can be expressed as

$$R_i = R_{Bi} + \sum_{\forall s \in \mathcal{S}_i} R_{si}. \tag{5}$$

Since individual mobile users have different levels of quality-of-experience (satisfaction) for the obtained throughput, we use the utility function $U_i(R_i)$ to evaluate the satisfaction of mobile user u for the throughput R_i. Correspondingly, the network utility can be expressed as $\sum_{\forall i \in \mathcal{I}} U_i(R_i)$. Without loss of generality, the function $U_i(R_i)$ can be set to be a twice-differentiable, increasing and strictly concave function of R_i for elastic services [4].

To strike the balance between the network utility and system cost, we need to take into account the energy expense due to the on-grid power supply. Let π_B and π_s denote the unit energy price set by the power utility company for the MBS and SBS s, respectively. Since the MBS is on-grid powered, its energy expense, denoted by Q_B, is equal to

$$Q_B = \pi_B T\left(\sum_{\forall i \in \mathcal{I}} p_{Bi} + P_B\right), \tag{6}$$

where P_B denotes the site power of MBS consumed by the baseband processor, the cooling system and etc. Every HSBS is jointly powered by both on-grid and off-grid energy, and thus we assume that HSBS s first uses the stored renewable energy during the time of $\tau_s T$ and then uses the on-grid power during the left time of $(1-\tau_s)T$. Without loss of generality, CSBSs and RSBSs can be considered as special HSBSs. RSBSs always consume the stored renewable energy, and thus we have $\tau_s = 1$ for every RSBS s. Oppositely, CSBSs always consume the on-grid energy, and thus we have $\tau_s = 0$ for every CSBS s. Therefore, the energy expense of SBS s paid to the power utility company for on-grid energy, denoted by Q_s, can be expressed as

$$Q_s = \begin{cases} \pi_s T(\sum_{\forall i \in \mathcal{I}_s} p_{si} + P_s), \forall s \in \mathcal{S}_C, \\ \pi_s T(1-\tau_s)(\sum_{\forall i \in \mathcal{I}_s} p_{si} + P_s), \forall s \in \mathcal{S}_H, \\ 0, \forall s \in \mathcal{S}_R, \end{cases} \tag{7}$$

where P_s denotes the site power consumption of SBS s, and the set $\mathcal{I}_s$ means the set of all MUs residing in the coverage area of SBS s. Accordingly, the on-grid energy cost is calculated as

$$Co = Q_B + \sum_{\forall s \in \mathcal{S}} Q_s. \tag{8}$$

On the other hand, all HSBSs and RSBSs are powered by the renewable energy stored in the individual batteries. Thus, the transmit power of RSBS s and HSBS s is subject to the amount of renewable energy stored in the battery, say E_s, namely,

$$\begin{cases} \tau_s(\sum_{\forall i \in \mathcal{I}_s} p_{si} + P_s) \leq \frac{E_s}{T}, \forall s \in \mathcal{S}_H, \\ \sum_{\forall i \in \mathcal{I}_s} p_{si} + P_s \leq \frac{E_s}{T}, \forall s \in \mathcal{S}_R. \end{cases} \tag{9}$$

We want to balance the network utility and on-grid energy cost subject to individual users' QoS requirements. Mathematically, we formulate the trade-off problem in the following form:

$$\max_{\boldsymbol{f},\boldsymbol{p},\boldsymbol{\tau}} \quad \sum_{\forall i \in \mathcal{I}} U_i(R_i) - \alpha Co \tag{10}$$

$$\text{s. t.} \quad R_i \geq R_{i,\min}, \forall i \in \mathcal{I}, \tag{10-1}$$

$$\text{constraints } (1), (2), (9), \qquad f_{Bi} \geq 0, \tag{10-2}$$

$$f_{si} \geq 0, \forall i \in \mathcal{I}_s, \forall s \in \mathcal{S}, \tag{10-3}$$

$$f_{si} = 0, \forall i \notin \mathcal{I}_s, \forall s \in \mathcal{S}, \tag{10-4}$$

$$0 \leq p_{Bi} \leq P_{B,\max}, \forall i \in \mathcal{I}, \tag{10-5}$$

$$0 \leq p_{si} \leq P_{s,\max}, \forall i \in \mathcal{I}_s, \forall s \in \mathcal{S}, \tag{10-6}$$

$$p_{si} = 0, \forall i \notin \mathcal{I}_s, \forall s \in \mathcal{S}, \tag{10-7}$$

$$0 \leq \tau_s \leq 1, \forall s \in \mathcal{S}_H, , \tag{10-8}$$

where the non-negative coefficient α means the weight of on-grid energy cost reflecting the resource allocation. Here, $\boldsymbol{f}$, $\boldsymbol{p}$, and $\boldsymbol{\tau}$ denote the vector of f_{Bi}'s and f_{si}'s, the vector of p_{Bi}'s and p_{si}'s, and the vector of τ_s's, respectively. The constraint (10-1) means that to satisfy the QoS requirement,

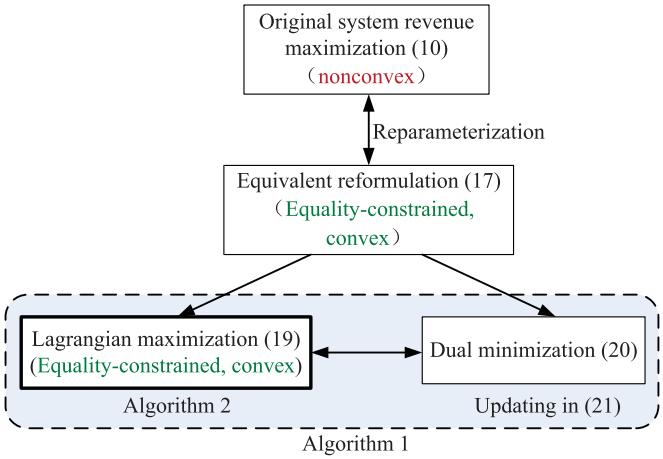


Fig. 2. Overview of our proposed approach to system revenue maximization. The equality-constrained convex reformulation is obtained by reparameterization. The box with dashed boundary means that Algorithm 1 is proposed to solve the reformulation problem by the primal-dual arguments. The box with bold boundary means that Algorithm 2 is proposed to solve the Lagrangian maximization problem.

each mobile user must meet the minimum data rate requirement (i.e., $R_{i,\min}$) in service. Constraints (10-5) and (10-6) mean that the transmit power of each MU served by the MBS and SBS s is subject to the maximum allowable values $P_{B,\max}$ and $P_{s,\max}$, respectively. Noticeably, we can strike different balances between the network utility and on-grid energy cost through adjusting α.

Note that due to the product of τ_s and p_{si}, the optimization problem (10) is inequality-constrained and non-convex at first glance. Therefore, common optimization methods such as ADMM and Lagrangian maximization cannot be directly applied to solve this optimization problem. In the following sections, we convexify problem (10) through transformation and reparameterization and propose an efficient algorithm to solve it.

4 Optimal Algorithm Design

In this section, we aim to find the optimal spectrum and power allocation by optimizing multiple variables in (10) sequentially. In particular, we first transform the problem (10) into a convex optimization problem through reparameterization. Then, we develop an efficient optimal algorithm to solve the equivalent problem based on the primal-dual arguments [27]. Fig. 2 gives an overview of the development in this paper, particularly the convexification, the algorithm design, and the connections between key optimization problems.

4.1 Convexification

To make problem (10) tractable, we first transform it into a convex optimization problem. To avoid the product of τ_s and p_{si}, we introduce the renewable power consumption variable $\hat{p}_{si}$ and the on-grid power consumption variable $\check{p}_{si}$ for each mobile user served by HSBS s, i.e.,

$$\begin{aligned}\hat{p}_{si} &= \tau_s\left(p_{si}+\frac{P_s}{I_s}\right),\\ \check{p}_{si} &= (1-\tau_s)\left(p_{si}+\frac{P_s}{I_s}\right), \forall i\in\mathcal{I}_s, \forall s\in\mathcal{S}_H,\end{aligned} \tag{11}$$

where I_s is the cardinality of $\mathcal{I}_s$. Each mobile user served by the HSBS consumes the renewable energy and on-grid power together, its total power consumption can be expressed as the summation of renewable power consumption and on-grid power consumption, i.e., (12).

$$p_{si}+\frac{P_s}{I_s}=\hat{p}_{si}+\check{p}_{si}, \forall i\in\mathcal{I}_s, \forall s\in\mathcal{S}_H. \tag{12}$$

Due to the fact that the function $U_i(R_i)$ is increasing with f_{Bi} and f_{si}'s, the constraints (1) and (2) are active when obtaining the optimal solution. Correspondingly, the constraints (1) and (2) are rewritten as

$$\sum_{\forall i\in\mathcal{I}} f_{Bi}=1, \tag{13}$$

$$\sum_{\forall i\in\mathcal{I}_s} f_{si}=1, \forall s\in\mathcal{S}. \tag{14}$$

To transfer the storage constraint (9) into an equality constraint, we further introduce auxiliary non-negative variables $\bar{p}_s$'s for all $s\in\mathcal{S}_H\cup\mathcal{S}_R$. Correspondingly, the constraint (9) is rewritten as

$$\sum_{\forall i\in\mathcal{I}_s}\hat{p}_{si}+\bar{p}_s=\frac{E_s}{T}, \forall s\in\mathcal{S}_H, \tag{15}$$

$$\sum_{\forall i\in\mathcal{I}_s} p_{si}+\bar{p}_s=\frac{E_s}{T}-P_s, \forall s\in\mathcal{S}_R, \tag{16}$$

by (11). As a result, we can rewrite problem (10) as

$$\max_{\boldsymbol{v}} \sum_{\forall i\in\mathcal{I}} U_i(R_i)-\alpha\left(Q_B+\sum_{\forall s\in\mathcal{S}_C} Q_s+\sum_{\forall s\in\mathcal{S}_H}\pi_s T\left(\sum_{\forall i\in\mathcal{I}_s}\check{p}_{si}\right)\right) \tag{17}$$

$$\text{s. t.}\quad \begin{aligned}&\text{constraints } (10-1)\sim(10-6),(12)\sim(16),\\ &\hat{p}_{si}\geq 0,\check{p}_{si}\geq 0,\forall i\in\mathcal{I}_s,\forall s\in\mathcal{S}_H,\end{aligned} \tag{17-1}$$

$$\bar{p}_s\geq 0, \forall s\in\mathcal{S}_H\cup\mathcal{S}_R, \tag{17-2}$$

where $\boldsymbol{v}=((f_{Bi},p_{Bi})_{\forall i\in\mathcal{I}},(f_{si},p_{si})_{\forall i\in\mathcal{I},\forall s\in\mathcal{S}},(\bar{p}_s)_{\forall s\in\mathcal{S}_H\cup\mathcal{S}_R},(\hat{p}_{si},\check{p}_{si})_{\forall i\in\mathcal{I}_s,\forall s\in\mathcal{S}_H})^T$.

The following theorem 1 shows the convexity of optimization problem (17).

Theorem 1. *The optimization problem (17) is a convex optimization problem.*

Proof. Since R_i is concave with respect to f_{Bi}, p_{Bi}, f_{si}'s and p_{si}'s, the minimum data rate constraint in (17-1) is convex. Together with the rest of linear constraints, it follows that the feasible set of the problem (17) is convex. Due to the concave nature of R_i, the concavity and increasing monotonicity of $U_i(R_i)$ in R_i implies that $U_i(R_i)$ is concave with respect to f_{Bi}, p_{Bi}, f_{si}'s and p_{si}'s. The objective function in (17) is the difference between the summation of concave functions and the summation of linear functions, and thus it is concave. Together with the convex feasible set, it follows that the problem (17) is a convex optimization problem. □

4.2 Algorithm Design

Due to the convex nature, we can obtain the optimal solution to the problem (17) by the primal-dual arguments [27]. In particular, we can solve the problem (17) through maximizing its Lagrangian and minimizing the corresponding dual function. Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_I)$ be the multiplier vector corresponding to all minimum data rate constraints in (10-1). We define the Lagrangian $L(\boldsymbol{v}, \boldsymbol{\eta})$ associated with the problem (17) in the following form:

$$L(\boldsymbol{v}, \boldsymbol{\eta}) = \sum_{\forall i \in \mathcal{I}} U_i(R_i) + \sum_{\forall i \in \mathcal{I}} \eta_i (R_i - R_{i,\min}) - \alpha \left(Q_B + \sum_{\forall s \in \mathcal{S}_C} Q_s + \sum_{\forall s \in \mathcal{S}_H} \pi_s T(\sum_{\forall i \in \mathcal{I}_s} \check{p}_{si}) \right). \tag{18}$$

Mathematically, we maximize the Lagrangian and minimize the dual problem as follows:

$$\begin{aligned} g(\boldsymbol{\eta}) = \max_{\boldsymbol{v}} \quad & L(\boldsymbol{v}, \boldsymbol{\eta}) \\ \text{s. t.} \quad & \text{constraints } (12) \sim (16), \\ & (10-2) \sim (10-6), \\ & (17-1) \sim (17-2), \end{aligned} \tag{19}$$

$$\begin{aligned} \text{Dual problem: } \min_{\boldsymbol{\eta}} \quad & g(\boldsymbol{\eta}) \\ \text{s. t.} \quad & \eta_i \geq 0, \forall i \in \mathcal{I}, \end{aligned} \tag{20}$$

Note that due to the mutual coupling of spectrum and power allocation among BSs, it is difficult to obtain the optimal solution to (19) in the closed-form expression. By the primal-dual arguments, we can obtain the optimal solution to (17) by solving (19) and (20) alternatively. Let $\boldsymbol{\eta}^-$ be the solution of (20) at the previous iteration. At the current iteration, the optimal solution to (19) is updated as $\boldsymbol{v}^+$ through maximizing (19) via Algorithm 2 proposed in Section 4, and the solution of (20) is updated as $\boldsymbol{\eta}^+$ by the subgradient method, i.e.,

$$\boldsymbol{\eta}^+ = \max\{0, \boldsymbol{\eta}^- - \beta(\boldsymbol{R}(\boldsymbol{v}^+) - \boldsymbol{R}_{\min})\}, \tag{21}$$

where $\boldsymbol{R}(\boldsymbol{v}^+)$ is the vector of R_i's obtained when all variables are equal to $\boldsymbol{v}^+$, $\boldsymbol{R}_{\min}$ is the vector $[R_{1,\min}, \ldots, R_{I,\min}]^T$, and β is the step-size. In particular, we solve (17) by performing Algorithm 1, referred to as the joint spectrum and power allocation algorithm.

Algorithm 1. The Joint Spectrum and Power Allocation (JSPA) Algorithm Proposed for the Optimal Solution to (17)

1: **Initialization:** Randomly choose $\boldsymbol{\eta}^+ \geq 0$.
2: **repeat**
3: Let $\boldsymbol{\eta}^- = \boldsymbol{\eta}^+$.
4: Update $\boldsymbol{v}^+$ by Algorithm 2 proposed in Section 4.
5: Update $\boldsymbol{\eta}^+$ by (21).
6: **until** Meet the stopping criterion, i.e., $||\boldsymbol{v}^+ - \boldsymbol{v}^-|| \leq \epsilon$.
7: Obtain the optimal solution to (17), i.e., $\boldsymbol{v}^+$.
* Here, $\boldsymbol{v}^-$ means the optimal solution to (19) obtained at the previous iteration, and $\epsilon > 0$ is the convergence tolerance.

4.3 Convergence and Optimality

The following theorem shows the convergence rate and optimality of the proposed joint spectrum and power allocation algorithm.

Theorem 2. *As $n \to \infty$, the joint spectrum and power allocation algorithm approaches a feasible solution of (17) at the rate of $1/n$, and the corresponding objective value reaches the optimal objective value of (17) with the gap less than $\frac{\beta L^2}{2}$, where n is the number of iterations and L is defined in (23).*

Proof. Let $\hat{\boldsymbol{v}}^{(n)}$ denote the optimal solution obtained by solving (19) at the nth iteration. We define $\bar{\boldsymbol{v}}^{(n)}$ as the average of the vectors $\hat{\boldsymbol{v}}^{(0)}, \ldots, \hat{\boldsymbol{v}}^{(n-1)}$, i.e.,

$$\bar{\boldsymbol{v}}^{(n)} = \frac{1}{n} \sum_{j=0}^{n-1} \hat{\boldsymbol{v}}^{(i)}. \tag{22}$$

For the notational simplicity, we use $F(\boldsymbol{v})$ to represent the objective function in (17). Let $\mathcal{V}$ be the feasible set of $\boldsymbol{v}$, which is determined by all equality and inequality constraints in (17) except the inequality constraint (17-1). Since the entries in $\boldsymbol{R}(\boldsymbol{v})$ are concave and continuous with $\boldsymbol{v}$, $\max_{\boldsymbol{v} \in \mathcal{V}} ||\boldsymbol{R}(\boldsymbol{v}) - \boldsymbol{R}_{\min}||_2$ is finite, and we define L as

$$L = \max_{\boldsymbol{v} \in \mathcal{V}} ||\boldsymbol{R}(\boldsymbol{v}) - \boldsymbol{R}_{\min}||_2. \tag{23}$$

By Proposition 1 in [26], we have

$$F^* - \frac{||\boldsymbol{\eta}^{(0)}||_2}{2n\beta} - \frac{\beta L^2}{2} \leq F(\bar{\boldsymbol{v}}^{(n)}) \leq F^* + ||\boldsymbol{\eta}^*||_2 ||[\boldsymbol{R}_{\min} - \boldsymbol{R}(\bar{\boldsymbol{v}}^{(n)})]^+||_2, \tag{24}$$

and

$$||[\boldsymbol{R}_{\min} - \boldsymbol{R}(\bar{\boldsymbol{v}}^{(n)})]^+||_2 \leq \frac{||\boldsymbol{\eta}^{(n)}||_2}{n\beta}, \tag{25}$$

where F^*, $\boldsymbol{\eta}^*$, $\boldsymbol{\eta}^{(0)}$, and $\boldsymbol{\eta}^{(n)}$ mean the optimal objective value of (17), the optimal dual solution of (20), the initial multiplier vector chosen for running Algorithm 1, and the multiplier vector obtained at the nth iteration, respectively. Due to the strict convexity of (17), the euclidean norm of multiplier vector $\boldsymbol{\eta}^{(n)}$ is bounded. Together with the inequality (25), it follows that $||[\boldsymbol{R}_{\min} - \boldsymbol{R}(\bar{\boldsymbol{v}}^{(n)})]^+||_2 \to 0$ at the rate of $1/n$ as $n \to \infty$. It implies that $\bar{\boldsymbol{v}}^{(n)}$ approaches a feasible solution of (17) (say $\bar{\boldsymbol{v}}^{\infty}$) at the rate of $1/n$ as $n \to \infty$. From (24), we further have that $F(\bar{\boldsymbol{v}}^{(n)})$ approaches the objective value of (17) (i.e., F^*) with the gap less than $\frac{\beta L^2}{2}$ for $n \to \infty$. By (22), we have $\hat{\boldsymbol{v}}^{(n)} = (n+1)\bar{\boldsymbol{v}}^{(n+1)} - n\bar{\boldsymbol{v}}^{(n)}$. Together with the fact that $\bar{\boldsymbol{v}}^{(n)} \to \bar{\boldsymbol{v}}^{\infty}$ for $n \to \infty$, we can get $\hat{\boldsymbol{v}}^{(n)} \to \bar{\boldsymbol{v}}^{(n)} \to \bar{\boldsymbol{v}}^{\infty}$ for $n \to \infty$, and thus $F(\hat{\boldsymbol{v}}^{(n)})$ approaches the objective value of (17) with the gap less than $\frac{\beta L^2}{2}$ for $n \to \infty$. Therefore, Theorem 2 follows. □

5 Lagrangian Maximization Based on ADMM

In this section, we consider how to solve the Lagrangian problem (19) in a distributed manner. Due to the lack of centralized infrastructure in the heterogeneous small-cell network,

all base stations should distributively decide the optimal spectrum and power allocation of each MU through limited information exchange with MUs. As we know, the novelty of the alternating direction method of multipliers is to effectively solve equality-constrained convex optimization problems suffering "big data" in a distributed manner [28], [29]. Note that the Lagrangian maximization (19) is an equality-constrained convex optimization problem. Therefore, we develop the ADMM-based Lagrangian maximization algorithm (i.e., Algorithm 2) to obtain the optimal solution to (19) in a distributed manner. That is, every BS can distributively update the spectrum and power allocation of every MU in service according to its data rate obtained from all associated BSs. In this section, we not only design the distributed algorithm, but also prove the optimality and time complexity of the proposed algorithm.

5.1 Algorithm Design

The proposed algorithm works as follows at each iteration. *First*, the MBS and each SBS update their spectrum and power allocation distributively by maximizing the augmented Lagrangian associated with (19). *Then*, the multipliers with respect to (12), (13), (14), (15), (16) are updated according to the updated variables. In what follows, we shall introduce every key ingredient in details.

5.1.1 Variable Updating

Let μ_B, $(\mu_s)_{s\in\mathcal{S}}$, $(\lambda_{si})_{\forall i\in\mathcal{I}_s,s\in\mathcal{S}_H}$, $(\phi_s)_{s\in\mathcal{S}_H}$, and $(\varphi_s)_{s\in\mathcal{S}_R}$ denote the multipliers with respect to constraints (12), (13), (14), (15), (16), respectively. For the notational brevity, we define $\boldsymbol{\theta}=(\mu_B,(\mu_s)_{s\in\mathcal{S}},(\lambda_{si})_{\forall i\in\mathcal{I}_s,s\in\mathcal{S}_H},(\phi_s)_{s\in\mathcal{S}_H},(\varphi_s)_{s\in\mathcal{S}_R})^T$. We use the matrix $\boldsymbol{A}$ and the vector $\boldsymbol{c}$ to denote the coefficient matrix with respect to variables in constraints (12), (13), (14), (15), (16), and the vector of coefficients at the right-hand side of constraints (12), (13), (14), (15), (16), respectively. Thus, we have the augmented Lagrangian of (19), denoted by $\tilde{L}_\rho(\boldsymbol{v},\boldsymbol{\theta},\boldsymbol{\eta})$, satisfying

$$\tilde{L}_\rho(\boldsymbol{v},\boldsymbol{\theta},\boldsymbol{\eta})=L(\boldsymbol{v},\boldsymbol{\eta})-\boldsymbol{\theta}^T(\boldsymbol{A}\boldsymbol{v}-\boldsymbol{c})-\frac{\rho}{2}||\boldsymbol{A}\boldsymbol{v}-\boldsymbol{c}||_2^2, \tag{26}$$

where $\rho>0$ is the penalty parameter. Based on the idea of ADMM, the augmented Lagrangian maximization and the multiplier update can be performed iteratively. In particular, at the kth iteration, the MBS and each SBS distributively update their spectrum and power allocation across MUs by maximizing the augmented Lagrangian (26), and obtain $\boldsymbol{v}^{(k)}$ that satisfies

MBS:

$$\begin{aligned} f_{Bi}^{(k)} &= \underset{f_{Bi}\geq 0}{\operatorname{argmax}} \quad \tilde{L}_\rho\Big(f_{Bi},\boldsymbol{v}_{-f_{Bi}}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}\Big)\\ p_{Bi}^{(k)} &= \underset{0\leq p_{Bi}\leq P_{B,\max}}{\operatorname{argmax}} \quad \tilde{L}_\rho\Big(p_{Bi},\boldsymbol{v}_{-p_{Bi}}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}\Big)\end{aligned} \tag{27}$$

SBS s :

$$f_{si}^{(k)}=\begin{cases}\underset{f_{si}\geq 0}{\operatorname{argmax}} \quad \tilde{L}_\rho(f_{si},\boldsymbol{v}_{-f_{si}}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}), \text{ if } i\in\mathcal{I}_s\\ 0, \text{ otherwise},\end{cases} \tag{28}$$

$$p_{si}^{(k)}=\begin{cases}\underset{0\leq p_{si}\leq P_{s,\max}}{\operatorname{argmax}} \quad \tilde{L}_\rho(p_{si},\boldsymbol{v}_{-p_{si}}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}), \text{ if } i\in\mathcal{I}_s\\ 0, \text{ otherwise},\end{cases} \tag{29}$$

$$\begin{aligned}\hat{p}_{si}^{(k)}&=\begin{cases}\underset{\hat{p}_{si}\geq 0}{\operatorname{argmax}} \quad \tilde{L}_\rho(\hat{p}_{si},\boldsymbol{v}_{-\hat{p}_{si}}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}), \text{ if } s\in\mathcal{S}_H\\ 0, \text{ otherwise},\end{cases}\\ \check{p}_{si}^{(k)}&=\begin{cases}\underset{\check{p}_{si}\geq 0}{\operatorname{argmax}} \quad \tilde{L}_\rho(\check{p}_{si},\boldsymbol{v}_{-\check{p}_{si}}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}), \text{ if } s\in\mathcal{S}_H\\ 0, \text{ otherwise},\end{cases}\\ \bar{p}_s^{(k)}&=\begin{cases}\underset{\bar{p}_s\geq 0}{\operatorname{argmax}} \quad \tilde{L}_\rho(\bar{p}_s,\boldsymbol{v}_{-\bar{p}_s}^{(k-1)},\boldsymbol{\theta}^{(k-1)},\boldsymbol{\eta}), \text{ if } s\in\mathcal{S}_H\cup\mathcal{S}_R,\\ 0, \text{ otherwise}.\end{cases}\end{aligned} \tag{30}$$

Here, $\boldsymbol{\theta}^{(k-1)}$ denotes the multipliers updated at the $(k-1)$th iteration. For example of $\boldsymbol{v}_{-f_{Bi}}^{(k-1)}$, it means the vector including all elements in $\boldsymbol{v}^{(k-1)}$ except $f_{Bi}^{(k-1)}$.

Since the augmented Lagrangian $\tilde{L}_\rho(\boldsymbol{v},\boldsymbol{\theta},\boldsymbol{\eta})$ is concave with every variable in $\boldsymbol{v}$, we can further obtain the optimal solutions to (27) and (28) in closed-form expressions. By calculating (27), we have $f_{Bi}^{(k)}$ and $p_{Bi}^{(k)}$ for all $i\in\mathcal{I}$ as shown in (31) and (32), respectively.

$$f_{Bi}^{(k)}=\begin{cases}0, & \text{if } \nabla_{Bi}\big|_{f_{Bi}=0,p_{Bi}^{(k-1)}}+\eta_i\leq\frac{F_{Bi}}{D_{Bi}},\\ 1, & \text{if } \nabla_{Bi}\big|_{f_{Bi}=1,p_{Bi}^{(k-1)}}+\eta_i\geq\frac{F_{Bi}+\rho}{D_{Bi}},\\ \tilde{f}_{Bi}, & \text{otherwise},\end{cases} \tag{31}$$

$$p_{Bi}^{(k)}=\begin{cases}0, & \text{if } \nabla_{Bi}\big|_{p_{Bu}=0,f_{Bi}^{(k-1)}}+\eta_i\leq\frac{\alpha\pi_B TN_0}{g_{Bi}},\\ P_{B,\max}, & \text{if } \nabla_{Bi}\big|_{p_{Bu}=P_{B,\max},f_{Bi}^{(k-1)}}+\eta_i\geq\frac{\alpha\pi_B T}{O_{Bi}},\\ \tilde{p}_{Bi}, & \text{otherwise},\end{cases} \tag{32}$$

with

$$\begin{aligned}\nabla_{Bi}&=\frac{\partial U_i(\sum_{\forall s\in\mathcal{S}_i}R_{si}^{(k-1)}+R_{Bi})}{\partial R_{Bi}},\\ D_{Bi}&=W_B\log\left(1+\frac{p_{Bi}^{(k-1)}g_{Bi}}{W_BN_0}\right)-\frac{W_Bp_{Bi}^{(k-1)}g_{Bi}}{W_BN_0+p_{Bi}^{(k-1)}g_{Bi}},\\ F_{Bi}&=\mu_B^{(k-1)}+\rho\left(\sum_{\forall i'\neq i}f_{Bi'}^{(k-1)}-1\right),\\ O_{Bi}&=\frac{W_Bf_{Bi}^{(k)}g_{Bi}}{W_Bf_{Bi}^{(k)}N_0+P_{B,\max}g_{Bi}},\end{aligned}$$

where $\tilde{f}_{Bi}$ is the root with respect to $(\nabla_{Bi}\big|_{f_{Bi},p_i^{(k-1)}}+\eta_i)(W_B\log(1+\frac{p_{Bi}^{(k-1)}g_{Bi}}{W_Bf_{Bi}N_0})-\frac{W_Bp_{Bi}^{(k-1)}g_{Bi}}{W_Bf_{Bi}N_0+p_{Bi}^{(k-1)}g_{Bi}})-\rho f_{Bi}=F_{Bi}$, and $\tilde{p}_{Bi}$ is the root satisfying $(\nabla_{Bi}\big|_{p_{Bi},f_{Bi}^{(k-1)}}+\eta_i)/(p_{Bi}g_{Bi}+W_Bf_{Bi}^{(k)}N_0)=\frac{\alpha\pi_BT}{W_Bf_{Bi}^{(k)}g_{Bi}}$.

By calculating (28), we have $f_{si}^{(k)}$, $p_{si}^{(k)}$, $\hat{p}_{si}^{(k)}$, $\check{p}_{si}^{(k)}$, and $\bar{p}_s^{(k)}$ for all $i\in\mathcal{I}_s$ in the following closed-form expressions. In particular, we have

$$f_{si}^{(k)} = \begin{cases} 0, \text{ if } \nabla_{si}\big|_{f_{si}=0,p_{si}^{(k-1)}} + \eta_i \le \frac{F_{si}}{D_{si}}, \\ 1, \text{ if } \nabla_{si}\big|_{f_{si}=1,p_{si}^{(k-1)}} + \eta_i \ge \frac{F_{si}+\rho}{D_{si}}, \\ \tilde{f}_{si}, \text{ otherwise.} \end{cases} \tag{33}$$

and

$$p_{si}^{(k)} = \begin{cases} 0, & \text{if } \nabla_{si}\big|_{p_{si}=0,f_{si}^{(k-1)}} + \eta_i \le \frac{H_{si}N_0}{g_{si}}, \\ P_{s,\max}, & \text{if } \nabla_{si}\big|_{p_{si}=P_{s,\max},f_{si}^{(k-1)}} + \eta_i \\ & \quad \ge \frac{H_{si}+[\rho P_{s,\max}]_{s\notin\mathcal{S}_C}}{O_{si}}, \\ \tilde{p}_{si}, & \text{otherwise,} \end{cases} \tag{34}$$

for all $i \in \mathcal{I}_s$, with

$$\nabla_{si} = \frac{\partial U_i(\sum_{\forall s'\in\mathcal{S}_i\setminus\{s\}} R_{s'i}^{(k-1)} + R_{Bi}^{(k-1)} + R_{si})}{\partial R_{si}},$$

$$D_{si} = W_s \log\left(1 + \frac{p_{si}^{(k-1)} g_{si}}{W_s N_0}\right) - \frac{W_s p_{si}^{(k-1)} g_{si}}{W_s N_0 + p_{si}^{(k-1)} g_{si}},$$

$$F_{si} = \mu_s^{(k-1)} + \rho\left(\sum_{\forall i'\neq i} f_{si'}^{(k-1)} - 1\right),$$

$$O_{si} = \frac{W_s f_{si}^{(k)} g_{si}}{W_s f_{si}^{(k)} N_0 + P_{s,\max} g_{si}},$$

$$H_{si} = \begin{cases} \alpha\pi_s T, \forall s \in \mathcal{S}_C, \\ \lambda_{si}^{(k-1)} - \rho(\hat{p}_{si}^{(k-1)} + \check{p}_{si}^{(k-1)} - \frac{P_s}{I_s}), \forall s \in \mathcal{S}_H, \\ \varphi_s^{(k-1)} + \rho(\sum_{\forall i'\neq i} p_{si}^{(k-1)} + \bar{p}_s^{(k-1)} - \frac{E_s}{T} + P_s), \\ \qquad\qquad \forall s \in \mathcal{S}_R, \end{cases}$$

$$[\rho P_{s,\max}]_{s\notin\mathcal{S}_C} = \begin{cases} 0, & \text{if } s \notin \mathcal{S}_C, \\ \rho P_{s,\max}, & \text{otherwise} \end{cases},$$

where $\tilde{f}_{si}$ is the root with respect to $(\nabla_{si}\big|_{f_{si},p_{si}^{(k-1)}} + \eta_i)(W_s \log(1 + \frac{p_{si}^{(k-1)} g_{si}}{W_s f_{si} N_0}) - \frac{W_s p_{si}^{(k-1)} g_{si}}{W_s f_{si} N_0 + p_{si}^{(k-1)} g_{si}}) - \rho f_{si} = F_{si}$, and $\tilde{p}_{si}$ is the root satisfying $(\nabla_{si}\big|_{p_{si},f_{si}^{(k-1)}} + \eta_u)/(p_{si} g_{si} + W_s f_{si}^{(k)} N_0) = \frac{H_{si}+[\rho p_{si}]_{s\notin\mathcal{S}_C}}{W_s f_{si}^{(k)} g_{si}}$. Besides, we have

$$\begin{aligned} \hat{p}_{si}^{(k)} &= \max\{0, \hat{H}_{si}\} \\ \check{p}_{si}^{(k)} &= \max\left\{0, p_{si}^{(k)} - \hat{p}_{si}^{(k)} + \frac{\lambda_{si}^{(k-1)} - \alpha\pi_s T}{\rho} + \frac{P_s}{I_s}\right\} \end{aligned} \tag{35}$$

for all $i \in \mathcal{I}_s$ and $s \in \mathcal{S}_H$, and

$$\bar{p}_s^{(k)} = \begin{cases} \max\{0, -\frac{\phi_s^{(k-1)}}{\rho} + (\frac{E_s}{T} - \sum_{\forall i\in\mathcal{I}_s} \hat{p}_{si}^{(k-1)})\}, & \text{if } s \in \mathcal{S}_H, \\ \max\{0, -\frac{\varphi_s^{(k-1)}}{\rho} + (\frac{E_s}{T} - P_s - \sum_{\forall i\in\mathcal{I}_s} p_{si}^{(k-1)})\}, & \text{if } s \in \mathcal{S}_R, \end{cases} \tag{36}$$

with $\hat{H}_{si} = \frac{\lambda_{si}^{(k-1)} - \phi_s^{(k-1)}}{2\rho} + \frac{1}{2}(p_{si}^{(k-1)} - \check{p}_{si}^{(k-1)} + P_s/I_s - \sum_{\forall i'\neq i} \hat{p}_{si}^{(k-1)} - \bar{p}_s^{(k-1)} + E_s/T)$.

Remark 1. Due to the concave nature of $\tilde{L}_\rho(v, \boldsymbol{\theta}, \boldsymbol{\eta})$, all roots can be efficiently obtained via either the bisection searching or Newton's method.

Note from (31), (32), (33), (34), (35), (36) that when every BS updates the spectrum and power allocation of MU i, it only needs the information about the data rates provided by all other BSs for MU i at the $(k-1)$ the iteration (i.e., $R_{si}^{(k-1)}$ or $R_{Bi}^{(k-1)}$) and the form of $U_i(\cdot)$. This implies that the updating of spectrum and power allocation can be performed at each BS in the distributed manner.

5.1.2 Multiplier Updating

After updating all variables in (27) and (28), we perform the update of multipliers according to $\boldsymbol{v}^{(k)}$ and $\boldsymbol{\theta}^{(k-1)}$. Let $\boldsymbol{r}^{(k)}$ denote the primal residual in the kth iteration, satisfying

$$\boldsymbol{r}^{(k)} = \boldsymbol{A}\boldsymbol{v}^{(k)} - \boldsymbol{c}. \tag{37}$$

It follows that the updated multipliers in the kth iteration, say $\boldsymbol{\theta}^{(k)}$, satisfies

$$\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + \rho\boldsymbol{r}^{(k)}. \tag{38}$$

When alternatively updating $\boldsymbol{v}^{(k)}$ and $\boldsymbol{\theta}^{(k)}$ until the stopping criterion is satisfied, the ADMM-based Lagrangian maximization algorithm terminates and the optimal solution to (19) is obtained. Specifically, we adopt the stopping criterion that the primal residual must be small, i.e.,

$$||\boldsymbol{r}^{(k)}||_2^2 \le \epsilon_r, \tag{39}$$

where $\epsilon_r > 0$ is the feasibility tolerance for the primal feasibility condition of (19).

Having introducing the main operations of the ADMM-based Lagrangian maximization algorithm, we present this algorithm in Algorithm 2.

Algorithm 2. The ADMM-Based Lagrangian Maximization Algorithm Proposed for the Optimal Solution to (19)

1: **Initialization:** Randomly choose $\boldsymbol{\theta}^{(0)}$, set $\boldsymbol{v}^{(0)} = \boldsymbol{v}^{-}$, and set $k = 1$.
2: **repeat**
3: All BSs update $\boldsymbol{v}^{(k-1)}$ as $\boldsymbol{v}^{(k)}$ by (31)-(36) in the distributed manner.
4: All BSs update $\boldsymbol{\theta}^{(k-1)}$ as $\boldsymbol{\theta}^{(k)}$ by (38).
5: Update $\boldsymbol{r}^{(k)}$ by (37).
6: $k = k + 1$,
7: **until** Meet the stopping criterion, i.e., $||\boldsymbol{r}^{(k-1)}||_2^2 \le \epsilon_r$.
8: Obtain the optimal solution to (19), i.e., $\boldsymbol{v}^{(k-1)}$.

5.2 Convergence and Complexity

The following theorem shows the global convergence of the ADMM-based Lagrangian maximization algorithm. The proof is given in Section 8.2.

Theorem 3. *The ADMM-based Lagrangian maximization algorithm achieves the primal residual convergence (i.e., $\boldsymbol{r}^{(k)} \to \boldsymbol{0}$ as $k \to \infty$), and hence the optimal solution for $k \to \infty$.*

The following theorem shows the convergence rate of the ADMM-based Lagrangian maximization algorithm. The proof is given in Section 8.2.

TABLE 1
Simulation Parameters

Simulation parameters	Value chosen
Carrier frequency	2000 MHz
Path loss model	$(128.1 + 37.6\log_{10}(d))$ dB (d in km)
Channel fading	Shadow fading $lognormal(0, 10^{1.6})$
Bandwidth W_B	20 MHz
Bandwidth W_s	5 MHz
Noise power spectral density N_0	-174 dBm/Hz
Maximum transmit power of MBS (i.e., $P_{B,\max}$)	0.5 W
Maximum transmit power of SBS (i.e.,$P_{s,\max}$)	0.1 W
Site power consumption of MBS (i.e., P_B)	0.2 W
Site power consumption of SBS (i.e., P_s)	0.1 W
Unit energy price π_B, π_s	0.6 $
Length of time slot (i.e., T)	60 sec

Theorem 4. *The ADMM-based Lagrangian maximization algorithm converges to the optimal solution to (19) at a sublinear rate.*

Let ϵ_p and ϵ_f represent termination parameters of updating the frequency and transmit power in the bisection searching, respectively. The following Theorem 5 shows the time complexity of the ADMM-based Lagrangian maximization algorithm when the bisection searching is applied to obtain the roots of variables in each iteration. The proof is given in Section 8.3.

Theorem 5. *When the bisection searching is applied to obtain the roots of variables in each iteration, the time complexity of the proposed ADMM-based Lagrangian maximization algorithm is* $O(IS(\log(\frac{\bar{P}}{\epsilon_p}) + \log(\frac{1}{\epsilon_f}))/\epsilon_r)$, *where* $\bar{P} = \max\{P_{B,\max}, \max_s P_{s,\max}\}$.

6 Numerical Results

In this section, we will evaluate the algorithm performance, the maximum network utility, the optimal energy consumption, and the maximum system revenue obtained when using the proposed algorithms. Referring to the Pico-cell parameter settings in [31], we set the simulation parameters in Table 1. All MUs and SBSs are uniformly distributed on a 1000m-by-1000m area, and the MBS is located at (500m, 500m). In the following simulation examples, we consider two types of network utilities: the proportional fairness utility $\sum_{\forall i \in \mathcal{I}} \log(R_i)$ and the weighted sum rate $\sum_{\forall i \in \mathcal{I}} \omega_i R_i$, where each w_i is uniformly distributed between 0 and 1.

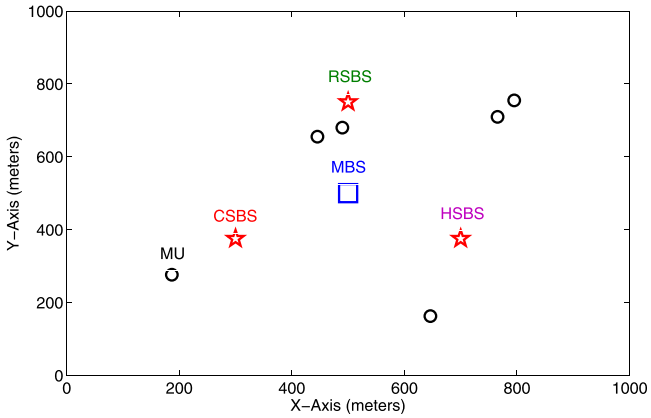
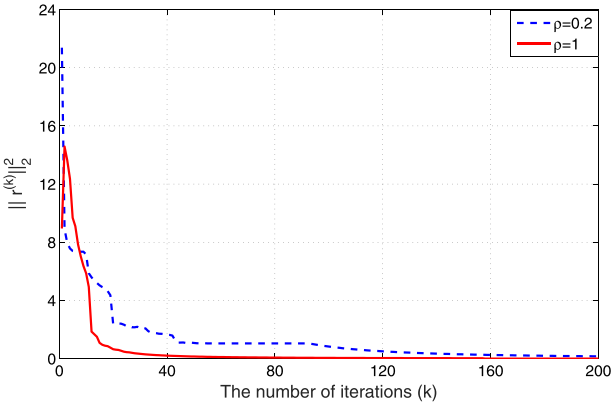
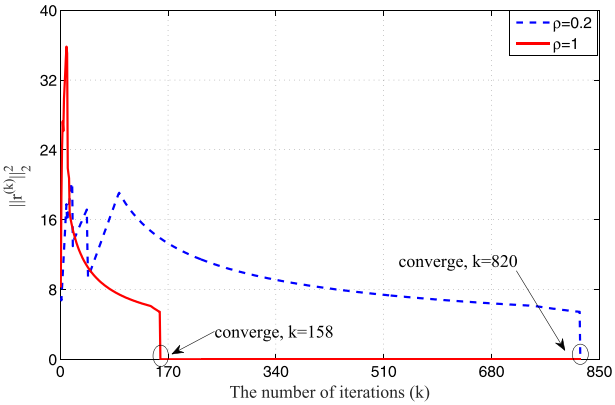


Fig. 3. The network topology used for Example 1.



(a) $U_u = \log(R_u)$



(b) $U_u = \omega_u R_u$

Fig. 4. The time complexity of the ADMM-based Lagrangian maximization algorithm for different network utilities in the different settings of penalty parameter.

6.1 Global Optimality and Time Complexity

Example 1. In this simulation example, we want to verify the global optimality and time complexity of the proposed ADMM-based Lagrangian maximization algorithm and the proposed JSPA algorithm when running them for the heterogeneous small-cell network with hybrid energy supplies as illustrated in Fig. 3. We set the minimum data rate requirement of each MU to be 2 Mbps, and set the battery storage of each HSBS and RSBS to be 10 Joules. The trade-off coefficient α is set to be 0.5.

Fig. 4 reveals that $\boldsymbol{r}^{(k)}$ approaches zero with the increase of the number of iterations. It implies that the ADMM-based Lagrangian maximization algorithm achieves the primal residual convergence, which is consistent with Theorem 3. Further, since $\boldsymbol{r}^{(k)}$ approaches zeros, we have that the sequence $\{\boldsymbol{\theta}^{(k)}\}$ sublinearly converges by (38), which implies that the ADMM-based Lagrangian maximization algorithm has the sublinear convergence rate, and hence Theorem 4. Also, we can find from Fig. 4 that the number of iterations can be fitted to a linear curve of $1/||\boldsymbol{r}^{(k)}||^2$, which implies that the ADMM-based Lagrangian maximization algorithm takes $O(1/\epsilon_r)$ iterations until the stopping criterion of $||\boldsymbol{r}^{(k-1)}||_2^2 \leq \epsilon_r$ is met. In addition, Fig. 4 shows that the time complexity of the ADMM-based Lagrangian maximization algorithm depends on the form of network utility and the setting of penalty parameter ρ. Specifically, it can be seen that the time

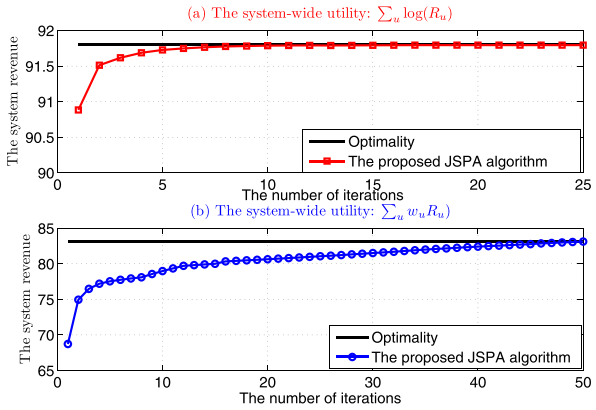


Fig. 5. The optimality and time complexity of the proposed JSPA algorithm for different network utilities.

complexity of the ADMM-based Lagrangian maximization algorithm is increasing with the decrease of ρ. Moreover, the ADMM-based Lagrangian maximization algorithm has the lower time complexity for the proportional fairness utility compared to that used for the weighted sum rate.

Fig. 5 shows that the optimality and convergence of the proposed JSPA algorithm with the step-size β being $10^{-6}/Num\_iterations$ for different network utilities. It can be seen that the proposed algorithm can always converge to the global optimality quickly. We can also find that running the proposed algorithm for the proportional network utility needs fewer iterations than that for the weighted sum rate.

6.2 Performance Comparison

To the best of our knowledge, there is no algorithm proposed for the same target in the literature. For the comparison with our proposed JSPA algorithm, we therefore introduce three baseline schemes: power control with equal frequency allocation, frequency allocation with equal power allocation, and full frequency reuse. The first baseline scheme means that each BS equally allocates its bandwidth to all associated MUs, and then all BSs cooperatively optimize the transmit power of all MUs to maximize the system revenue under the minimum data rate requirements of individual MUs. The second baseline scheme means that each on-grid/hybird-power-supplying BS utilizes the maximum transmit power to each associated MU and each off-grid BS equally allocates the battery energy to the transmission of each associated MU, and then all BSs optimally allocate the bandwidth to all MUs for the maximum system revenue subject to the minimum data rate requirements of individual MUs. The last baseline scheme means that the frequency is fully reused between small cells where each SBS equally allocates the frequency to its associated MUs, and then all SBSs cooperatively optimize the transmit power of all MUs with taking into the inter-cell interference. In the simulation, we evaluate the obtained network utility, total on-grid energy consumption, and system revenue as three performance criteria of the three algorithms. In the following performance comparison, we first compare our proposed algorithm with the former two baseline schemes which concern the orthogonal frequency allocation between small cells. Then, we compare our proposed algorithm with the full frequency reuse scheme.

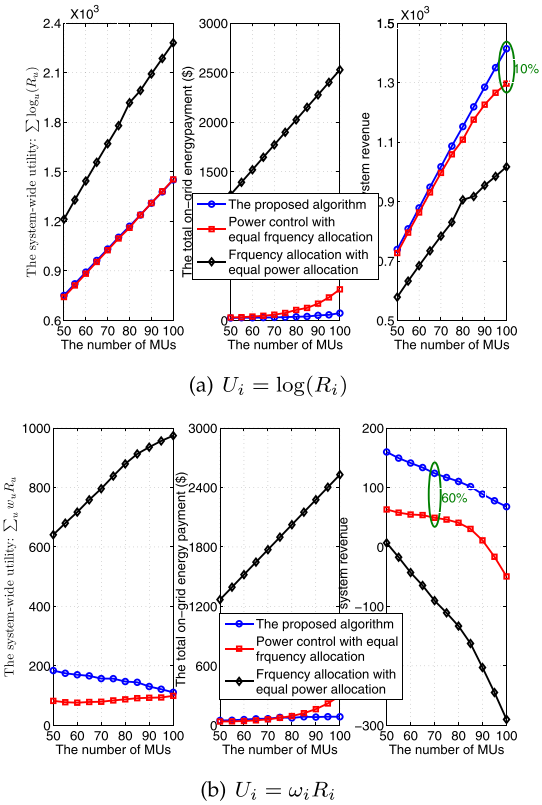
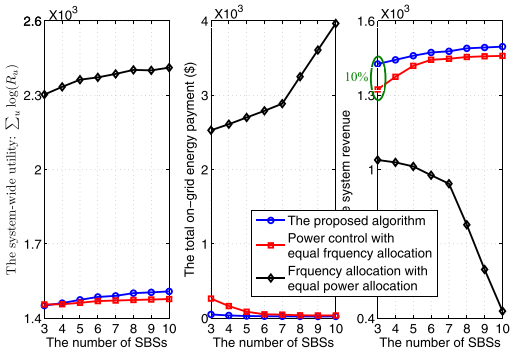


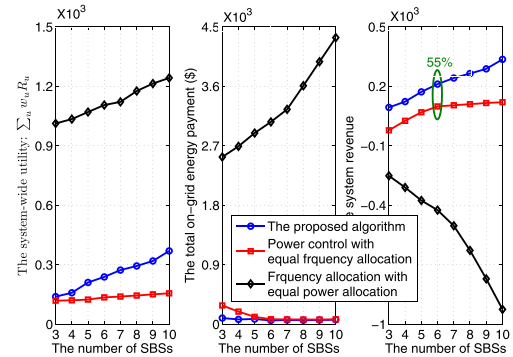
Fig. 6. The performance of three algorithms for different network utilities at different densities of MUs.

Example 2 (Performance comparison at different densities of MUs). We consider a set of hybrid energy supplying small-cell networks, where we fix the locations of four BSs as shown in Fig. 3, and randomly deploy $50 \sim 100$ MUs in the 1000m-by-1000m area. The parameter ϵ is set to be 10^{-4}. Each point in Fig. 6 is obtained by averaging over 100 different topologies with the same density of user nodes. Other simulation parameters are the same as those in Example 1.

Fig. 6 shows that as the number of MUs increases, the obtained network utility, total on-grid energy payment and system revenue all increase for the three orthogonal frequency allocation algorithms when the proportional fairness utility is adopted. On the contrary, when adopting the weighted sum rate across MUs, the system revenue of three algorithms decreases with the increase of the number of MUs. Unlike the proportional fairness utility that leads to fair spectrum and energy allocation among MUs, the weighted sum rate makes more spectrum and energy resource allocated to MUs in the better channel conditions. Since the spectrum and energy resource is first used to guarantee the minimum data rate requirements of all MUs, it implies that in the case of weighted sum rate, less idle resource can be greedily allocated to MUs in the better channel conditions with increasing the number of MUs. This further leads to the decrease of system revenue related to the weighted sum rate. It can also be seen that the proposed algorithm always outperforms the other two algorithms in terms of system revenue. Specifically, the improvement increases with the increase of the number of MUs for the proportional fairness

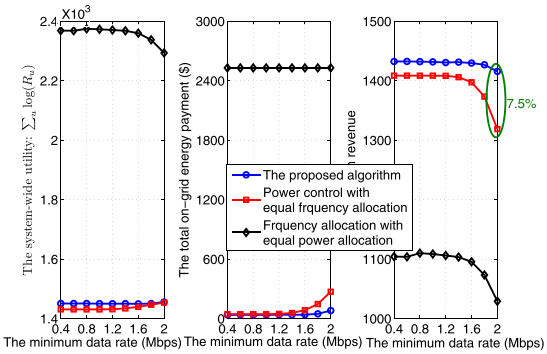


(a) $U_i = \log(R_i)$

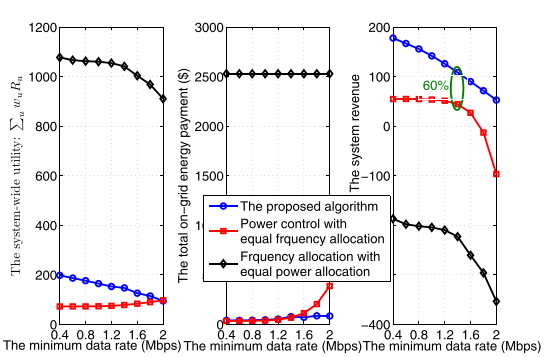


(b) $U_i = \omega_i R_i$

Fig. 7. The performance of three algorithms for different network utilities at different densities of SBSs.



(a) $U_i = \log(R_i)$



(b) $U_i = \omega_i R_i$

Fig. 8. The performance of three algorithms for different network utilities at different settings of $R_{i,\min}$.

utility, and the proposed algorithm can improve 10 percent system revenue compared to power control with equal frequency allocation when the number of MUs is 100. For the weighted sum rate, the proposed algorithm can increase 60 percent system revenue at least compared to power control with equal frequency allocation[3].

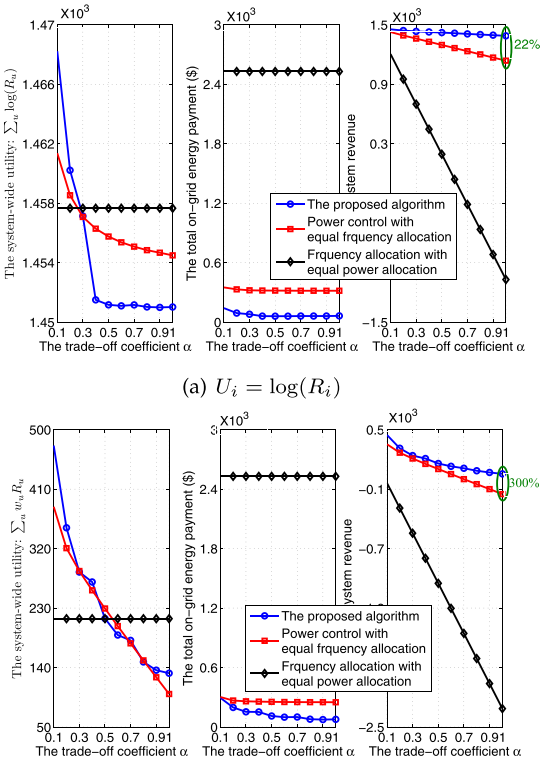
Example 3 (Performance comparison at different densities of SBSs). We consider a set of hybrid energy supplying small-cell networks, where we deploy the MBS at (500m, 500m), and randomly deploy $3 \sim 10$ SBSs and 100 MUs in the 1000m-by-1000m area. The type of each SBS is randomly set to be CSBS, HSBS, or RSBS. Each point in Fig. 7 is obtained by averaging over 100 different topologies with the same density of SBSs. Other simulation parameters are the same as those in Example 1.

Fig. 7 shows that except for the frequency allocation with equal power allocation, as the number of SBSs increases, the obtained network utility and system revenue of the other two orthogonal frequency allocation algorithms increases, and their total on-grid energy payment decreases for any type of network utility. We see that the proposed algorithm always outperforms the two baseline algorithms. Specifically, the minimum improvement of system revenue is 5 and 55 percent for the proportional fairness utility and the weighted sum rate, respectively. Furthermore, it can be seen that the power allocation has a larger impact in the system revenue compared to the frequency allocation. It implies that for the practical implementation, we only need to optimally control the transmit power used for the transmission of each MUs in addition to equal frequency allocation among MUs.

Example 4 (Performance comparison at different settings of $R_{i,\min}$). We consider a set of hybrid energy supplying small-cell networks, where we fix the locations of four BSs as shown in Fig. 3, and randomly deploy 100 MUs in the 1000m-by-1000m area. We vary the minimum data rate requirement $R_{i,\min}$ from 0.4 Mbps to 2 Mbps. Each point in Fig. 8 is obtained by averaging over 100 different topologies when giving $R_{i,\min}$. Other simulation parameters are the same as those in Example 1.

Fig. 8 shows that in comparison with the other two orthogonal frequency allocation algorithms, the frequency allocation with equal power allocation can attain the maximum network utility at the cost of more energy payment, which leads to the lowest system revenue. It can be seen that with increasing the minimum data rate requirements of all MUs, the power control with equal frequency allocation can increase the network utility at the cost of more energy

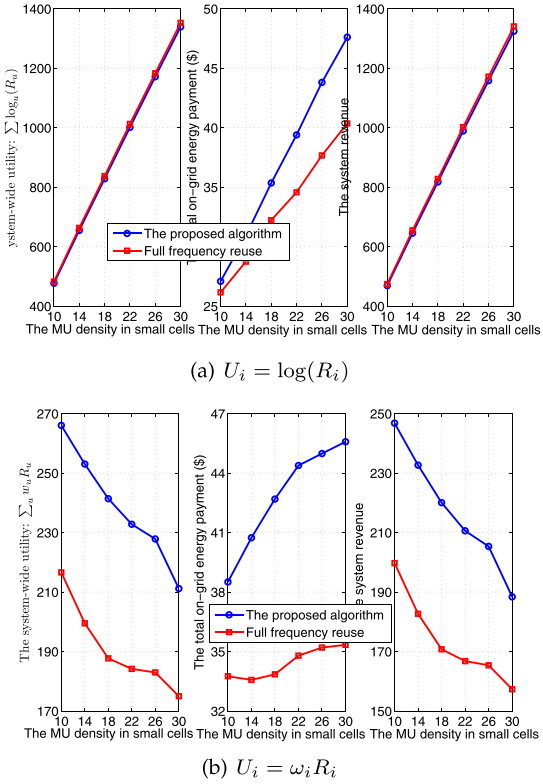
3. In this work, the system revenue is defined as the difference between the system utility and the weighted on-grid energy payment. When using the frequency allocation with equal power allocation, the total on-grig energy payment is equal to $(P_{B,\max} + P_B) \times I \times T \times \pi_B + (P_{s,\max} + P_s) \times I \times T \times \pi_s + (P_{s,\max} + \ P_s - E_s/T) \times T \times \pi_s$, where I means the number of MUs deployed in the network. Given the parameter settings in Table 1, we have the result that the total on-grid energy payment increases linearly with the number of MUs. On the other hand, when using the frequency allocation with equal power allocation, there is no power control and the maximum on-grid energy payment is needed, which leads to the negative system revenue with the electricity price being 0.6$. Considering that the energy consumption is fixed for any MU density, the energy price π has substantial role in the system revenue when using the frequency allocation with equal power allocation. It implies that in order to obtain the positive system revenue, we can set the electricity price to be a low price, e.g., 0.1$.



(a) $U_i = \log(R_i)$

(b) $U_i = \omega_i R_i$

Fig. 9. The performance of three algorithms for different network utilities at different setting of trade-off coefficient α.



(a) $U_i = \log(R_i)$

(b) $U_i = \omega_i R_i$

Fig. 10. The performance of the proposed algorithm and full frequency reuse scheme for different network utilities at different MU densities.

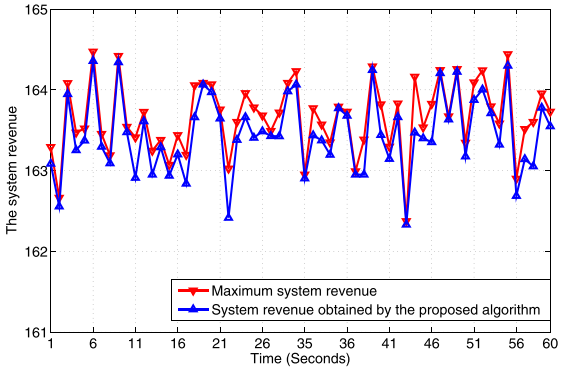
payment for any type of network utility. On the contrary, as the minimum data rate requirement increases, the proposed algorithm can yield an increasing utility in the case of proportional fairness utility and a decreasing utility in the case of weighted sum utility with the increase of energy payment. This is because that to meet the minimum data rate requirement of each MU, the greedy property of weighted sum rate leads to less spectrum resource allocated to MUs in the better channel conditions when increasing the minimum data rate requirement. We can also see that although the system revenue of three algorithms decreases with the increase of minimum data rate requirement $R_{i,\min}$, the proposed algorithm always outperforms the two baseline algorithms. For example of the weighted sum rate, the proposed algorithm can increase 60 percent system revenue at least compared to power control with equal frequency allocation.

Example 5 (Performance comparison at different setting of trade-off coefficient α). We consider a set of hybrid energy supplying small-cell networks, where we fix the locations of four BSs as shown in Fig. 3, and randomly deploy 100 MUs in the 1000m-by-1000m area. We vary the trade-off coefficient α from 0 to 1. Each point in Fig. 9 is obtained by averaging over 100 different topologies when fixing α. Other simulation parameters are the same as those in Example 1.

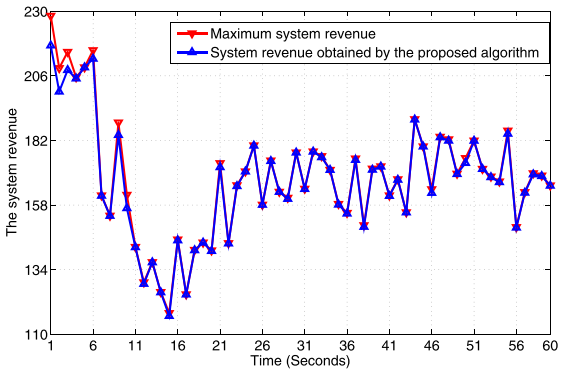
Fig. 9 shows that varying α cannot affect the network utility and total energy payment when using the frequency allocation with equal power allocation. This is because that given power allocation, the system revenue maximization is reduced to the network utility maximization through frequency allocation, which is independent of the setting of α. On the contrary, the obtained network utility, total energy payment, and system revenue of the other two orthogonal frequency allocation algorithms decrease with the increase of α. Since the trade-off coefficient α reflects the negative influence of energy payment to the system revenue, the increase of α implies that we need to reduce the energy payment to maximize the system revenue. Correspondingly, the decrease of energy payment further leads to the decrease of network utility. Also, it can be seen that the proposed algorithm always outperforms the two baseline algorithms. For example of the weighted sum rate, the proposed algorithm can increase 300 percent system revenue at least compared to power control with equal frequency allocation.

Example 6 (Performance comparison with the full frequency reuse scheme). In this simulation example, we want to evaluate the performance of our proposed algorithm through comparing it with the full frequency reuse scheme. We consider a set of hybrid energy supplying small-cell networks, where we fix the locations of four BSs as shown in Fig. 3, and randomly deploy 10∼30 MUs in the coverage area of each small cell. Each point in Fig. 10 is obtained by averaging over 100 different topologies with the same density of MUs. Other simulation parameters are the same as those in Example 1.

Fig. 10 shows that as the number of MUs in each small cell increases, the obtained network utility, total on-grid energy payment and system revenue all increase for the two algorithms when the proportional fairness utility is adopted. On the contrary, when adopting the weighted sum rate across MUs, only the obtained network utility and system revenue



(a) $U_i = \log(R_i)$



(b) $U_i = \omega_i R_i$

Fig. 11. The impact of Rayleigh fading on the algorithm performance with the increase of time.

decrease with the increase of the number of MUs for the two algorithms. Also, we can find that in comparison with the full frequency reuse scheme, the performance gain of our proposed algorithm depends on the form of selected system utility. In particular, when considering the proportional faintness utility, the system revenue of our proposed algorithm is close to that of the full frequency reuse scheme at the cost of more on-grid energy payment. On the contrary, when adopting the weighted sum rate across MUs, our proposed algorithm outperforms the full frequency reuse scheme. Specifically, the improvement increases with the increase of the number of MUs, and the proposed algorithm can improve 20 percent at least. For the weighted sum rate, the full frequency reuse scheme has to mitigate the co-channel interference through the optimal power allocation. Therefore, although the full frequency reuse scheme prefers to allocate more transmit energy to the MUs with the better channel conditions, less weighted sum rate (and hence less system revenue) is obtained due to the co-channel interference.

Example 7 (Impact of fast fading on the algorithm performance). As mentioned in the previous simulation examples, we evaluate the algorithm performance in the context of slow fading. In this simulation example, we want to evaluate how the fast fading (i.e., Rayleigh fading) influences the algorithm performance in the practical algorithm implementation. Thus, the fading effects include the slow fading (i.e., shadow fading) and fast fading (i.e., Rayleigh fading) for evaluation purposes. We consider a hybrid energy supplying small-cell network, where we fix the locations of four BSs as shown in Fig. 3, and randomly deploy 10 MUs in the coverage area of each small cell. We use the Clarke's model to generate the Rayleigh fading level with the approximal coherence time being 1 second. Thus, we set the length of the time slot to be 1 second. Other simulation parameters are the same as those in Example 1.

The system revenue obtained per time slot (i.e., per second) is plotted in Fig. 11. In the algorithm implementation with fast fading, the fading realization estimated in the previous time slot is used for running the algorithm in the current time slot. Considering the convergence time of the proposed algorithm, the proposed algorithm cannot converge in a time slot for the fast fading. Thus, the system revenue obtained in the last iteration in each time slot is considered as the system revenue obtained by the proposed algorithm, which is shown in the blue curve of Fig. 11. For the comparison purpose, we assume that the fading realizations are known for the optimization in the results with fast fading, which leads to the maximum system revenue shown in the red curve of Fig. 11. We can see from Fig. 11b that when adopting the weighted sum rate across MUs, the system revenue obtained by the proposed algorithm can approach the maximum system revenue with the increase of time. On the contrary, we can find from Fig. 11a that when the proportional fairness utility is considered, the system revenue obtained by the proposed algorithm cannot meet the maximum system revenue. Interestingly, the performance gap is very small, i.e., 4.11 percent at most in comparison with the maximum system revenue. These observations reveal that in the practical implementation of the proposed algorithm, the impact of realistic effects such as fast fading on the algorithm performance can be negligible.

7 Conclusion

In this paper, we have studied the joint optimization of spectrum and power that maximizes the system revenue defined for the heterogeneous small-cell network with hybrid energy supplies. By exploiting the hidden convexity of this optimization problem, we have proposed an efficient joint spectrum and power allocation algorithm to obtain the optimal solution. The proposed algorithm is computationally efficient, since it approaches the global optimality at the rate of $1/n$. Specifically, in our proposed algorithm, we have designed the ADMM-based Lagrangian maximization algorithm to obtain the optimal solution to the primal sub-problem in the closed-form expression at each iteration, and have applied the subgradient method to update the solution to the dual sub-problem. Our theoretical analysis shows that the proposed ADMM-based Lagrangian maximization algorithm approaches the primal optimal solution with the time complexity of $O(1/\epsilon_r)$ iterations. In the future work, we will study the dynamic optimization of spectrum and power in the long-term scheduling period, taking into account the intermittent nature of renewable energy arrival. Furthermore, the mobility of MUs and the frequency reuse among small cells will be considered when designing the joint spectrum and power allocation algorithm for hybrid energy supplying small-cell networks.

8 PROOFS

8.1 Proof of Theorem 3

Note that R_u is a continuous and concave function with respect to v. Together with the twice-differentiable and strictly concave nature of $U_u(R_u)$, it follows that the function $L(v, \boldsymbol{\eta})$ is closed, proper, differentiable, and concave in $\boldsymbol{v}$. Also, it follows that the unaugmented Lagrangian $\tilde{L}_0(v, \boldsymbol{\theta}, \boldsymbol{\eta})$ has a saddle point $(\boldsymbol{v}^*, \boldsymbol{\theta}^*)$ satisfying

$$\tilde{L}_0(\boldsymbol{v}^*, \boldsymbol{\theta}, \boldsymbol{\eta}) \geq L(\boldsymbol{v}^*, \boldsymbol{\eta}) = \tilde{L}_0(\boldsymbol{v}^*, \boldsymbol{\theta}^*, \boldsymbol{\eta}) \geq \tilde{L}_0(v, \boldsymbol{\theta}^*, \boldsymbol{\eta}) \tag{40}$$

$$\boldsymbol{A}\boldsymbol{v}^* = \boldsymbol{c} \tag{41}$$

for all feasible v, where $\boldsymbol{v}^*$ and $\boldsymbol{\theta}^*$ mean a pair of primal and dual optimal solutions to (19), respectively. Accordingly, $L(\boldsymbol{v}^*, \boldsymbol{\eta})$ is the optimal objective value of (19). By (40) and (41), we have

$$\begin{aligned} L(\boldsymbol{v}^*, \boldsymbol{\eta}) &= \tilde{L}_0(\boldsymbol{v}^*, \boldsymbol{\theta}^*, \boldsymbol{\eta}) \\ &\geq \tilde{L}_0(\boldsymbol{v}^{(k)}, \boldsymbol{\theta}^*, \boldsymbol{\eta}) \\ &= L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) - \boldsymbol{\theta}^{*T}\boldsymbol{r}^{(k)}. \end{aligned} \tag{42}$$

Since $L(\boldsymbol{v}, \boldsymbol{\eta})$ is closed, proper, differentiable, and concave, so is $\tilde{L}_\rho(v, \boldsymbol{\theta}, \boldsymbol{\eta})$. Thus, the optimality condition at the kth iteration satisfies

$$\begin{aligned} 0 =& \nabla \tilde{L}_\rho(\boldsymbol{v}^{(k)}, \boldsymbol{\theta}, \boldsymbol{\eta}) \\ =& \nabla L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) - \boldsymbol{A}^T\boldsymbol{\theta}^{(k-1)} - \rho\boldsymbol{A}^T\boldsymbol{r}^{(k)}, \end{aligned} \tag{43}$$

where ∇ denotes the differential operator. Together with the fact that $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + \rho\boldsymbol{r}^{(k)}$ in (38), it follows that

$$\nabla L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) - \boldsymbol{A}^T\boldsymbol{\theta}^{(k)} = 0. \tag{44}$$

This implies that $\boldsymbol{v}^{(k)}$ maximizes

$$L(v, \boldsymbol{\eta}) - (\boldsymbol{\theta}^{(k)})^T\boldsymbol{A}v. \tag{45}$$

Thus, we have

$$L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) - (\boldsymbol{\theta}^{(k)})^T\boldsymbol{A}\boldsymbol{v}^{(k)} \geq L(\boldsymbol{v}^*, \boldsymbol{\eta}) - (\boldsymbol{\theta}^{(k)})^T\boldsymbol{A}\boldsymbol{v}^*. \tag{46}$$

Adding (42) and (46), and using $\boldsymbol{A}\boldsymbol{v}^* = \boldsymbol{c}$, we have

$$(\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*)^T\boldsymbol{r}^{(k)} \leq 0. \tag{47}$$

Substituting $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + \rho\boldsymbol{r}^{(k)}$, we get

$$\begin{aligned} & 2(\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*)^T\boldsymbol{r}^{(k)} \\ =& \, 2(\boldsymbol{\theta}^{(k-1)} - \boldsymbol{\theta}^*)^T\boldsymbol{r}^{(k)} + 2\rho||\boldsymbol{r}^{(k)}||_2^2 \\ =& \, \frac{2}{\rho}(\boldsymbol{\theta}^{(k-1)} - \boldsymbol{\theta}^*)^T(\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}) \\ & + \frac{1}{\rho}||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}||_2^2 + \rho||\boldsymbol{r}^{(k)}||_2^2. \end{aligned} \tag{48}$$

Since $\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)} = (\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*) - (\boldsymbol{\theta}^{(k-1)} - \boldsymbol{\theta}^*)$, Eq. (48) can be rewritten as

$$\frac{1}{\rho}(||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*||_2^2 - ||\boldsymbol{\theta}^{(k-1)} - \boldsymbol{\theta}^*||_2^2) + \rho||\boldsymbol{r}^{(k)}||_2^2. \tag{49}$$

We define $V^{(k)} = \frac{1}{\rho}||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*||_2^2$, and it follows from (47) and (49) that

$$V^{(k)} \leq V^{(k-1)} - \rho||\boldsymbol{r}^{(k)}||_2^2. \tag{50}$$

This implies that $V^{(k)}$ decreases in each iteration. Iterating (50), we have

$$\rho\sum_{k=1}^{\infty}||\boldsymbol{r}^{(k)}||_2^2 \leq V^{(0)}. \tag{51}$$

It follows that $\boldsymbol{r}^{(k)} \to \boldsymbol{0}$ as $k \to \infty$, which implies the primal residual convergence.

By (41) and (37), we can get from (42) and (46) that

$$L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) - L(\boldsymbol{v}^*, \boldsymbol{\eta}) \leq (\boldsymbol{\theta}^*)^T\boldsymbol{r}^{(k)}, \tag{52}$$

and

$$L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) - L(\boldsymbol{v}^*, \boldsymbol{\eta}) \geq (\boldsymbol{\theta}^{(k)})^T\boldsymbol{r}^{(k)}, \tag{53}$$

respectively. Along with the fact that $\boldsymbol{r}^{(k)} \to \boldsymbol{0}$ as $k \to \infty$, it follows that $L(\boldsymbol{v}^{(k)}, \boldsymbol{\eta}) \to L(\boldsymbol{v}^*, \boldsymbol{\eta})$ and $\boldsymbol{A}\boldsymbol{v}^{(k)} \to \boldsymbol{c}$ as $k \to \infty$. It implies that $\boldsymbol{v}^{(k)}$ approaches one optimal solution to (19) for $k \to \infty$. Therefore, Theorem 3 follows.

8.2 Proof of Theorem 4

The inequality (50) shows that $||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*||_2^2$ decreases in each iteration, and thus the sequence $\{\boldsymbol{\theta}^{(k)}\}$ is bounded. Thus, we have

$$\lim_{k\to\infty}||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}||_2^2 = 0. \tag{54}$$

By (38), the inequality (50) can be rewritten as

$$||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*||_2^2 - ||\boldsymbol{\theta}^{(k-1)} - \boldsymbol{\theta}^*||_2^2 \leq -||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}||_2^2. \tag{55}$$

By (54), the inequality (55) implies that

$$\lim_{k\to\infty}(||\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^*||_2^2 - ||\boldsymbol{\theta}^{(k-1)} - \boldsymbol{\theta}^*||_2^2) \leq 0. \tag{56}$$

Let $\boldsymbol{\theta}^\infty$ be the cluster point of the sequence $\{\boldsymbol{\theta}^{(k)}\}$. It follows from (56) that the sequence $\{\boldsymbol{\theta}^{(k)}\}$ sublinearly converges to $\boldsymbol{\theta}^\infty$. As shown in the proof of Theorem 3, the sequence $\{\boldsymbol{v}^{(k)}\}$ with respect to the sequence $\{\boldsymbol{\theta}^{(k)}\}$ converges to one primal optimal solution to (19). Due to the strict convexity of (19), the multiplier sequence $\{\boldsymbol{\theta}^{(k)}\}$ converges to one dual optimal solution (and hence $\boldsymbol{\theta}^\infty$) accordingly. Therefore, it follows that the sequence $\{\boldsymbol{v}^{(k)}\}$ converges to the optimal solution to (19) at a sublinear rate, and thus Theorem 4 follows.

8.3 Proof of Theorem 5

Let $\tilde{\boldsymbol{v}}^{(k)} = \frac{1}{k}\sum_{n=0}^{k-1}\boldsymbol{v}^{(n)}$, and $(\boldsymbol{v}^*, \boldsymbol{\theta}^*)$ be a saddle point of the unaugmented Lagrangian $\tilde{L}_0(v, \boldsymbol{\theta}, \boldsymbol{\eta})$. We denote the initial Lagrange multipliers and initial primal variables as $\boldsymbol{\theta}^{(0)}$ and $\boldsymbol{v}^{(0)}$, respectively. Due to the strict convex nature of (19), we have

$$L(\tilde{\boldsymbol{v}}^{(k)}, \boldsymbol{\eta}) - L(\boldsymbol{v}^*, \boldsymbol{\eta}) - \boldsymbol{\theta}^T(\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)} - \boldsymbol{c}) \geq -\frac{1}{2k}\psi_0(\boldsymbol{\theta}) \tag{57}$$

for any feasible $\boldsymbol{\theta}$ by Theorem 1 in [30], where $\psi_0(\boldsymbol{\theta}) = \rho||\boldsymbol{A}(\boldsymbol{v}^{(0)} - \boldsymbol{v}^*)||_2^2 + \frac{1}{\rho}||\boldsymbol{\theta}^{(0)} - \boldsymbol{\theta}||_2^2$. Using $\boldsymbol{\theta} = \boldsymbol{\theta}^* + \frac{\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)} - \boldsymbol{c}}{||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)} - \boldsymbol{c}||_2}$, we have

$$L(\tilde{\boldsymbol{v}}^{(k)},\boldsymbol{\eta})-L(\boldsymbol{v}^*,\boldsymbol{\eta})-(\boldsymbol{\theta}^*)^T(\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c})-||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2 \geq -\frac{1}{2k}\psi_0\left(\boldsymbol{\theta}^*+\frac{\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}}{||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2}\right) \tag{58}$$

by (57). With the saddle point inequality (40), we get

$$L(\tilde{\boldsymbol{v}}^{(k)},\boldsymbol{\eta})-(\boldsymbol{\theta}^*)^T(\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}) \leq L(\boldsymbol{v}^*,\boldsymbol{\eta})-(\boldsymbol{\theta}^*)^T(\boldsymbol{A}\boldsymbol{v}^*-\boldsymbol{c}). \tag{59}$$

It implies that

$$L(\tilde{\boldsymbol{v}}^{(k)},\boldsymbol{\eta})-L(\boldsymbol{v}^*,\boldsymbol{\eta})-(\boldsymbol{\theta}^*)^T(\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}) \leq 0. \tag{60}$$

Adding (58) and (60), we have

$$\begin{aligned}||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2 &\leq \frac{1}{2k}\psi_0\left(\boldsymbol{\theta}^*+\frac{\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}}{||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2}\right)\\ &=\frac{\rho}{2k}||\boldsymbol{A}(\boldsymbol{v}^{(0)}-\boldsymbol{v}^*)||_2^2\\ &\quad+\frac{1}{2k\rho}\left|\left|\boldsymbol{\theta}^{(0)}-\boldsymbol{\theta}^*-\frac{\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}}{||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2}\right|\right|_2^2\\ &\leq\frac{\rho}{2k}||\boldsymbol{A}(\boldsymbol{v}^{(0)}-\boldsymbol{v}^*)||_2^2\\ &\quad+\frac{1}{k\rho}(||\boldsymbol{\theta}^{(0)}-\boldsymbol{\theta}^*||_2^2+1).\end{aligned} \tag{61}$$

Recall $\tilde{\boldsymbol{v}}^{(k)}=\frac{1}{k}\sum_{n=0}^{k-1}\boldsymbol{v}^{(n)}$, and we have

$$||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2^2=\left|\left|\frac{1}{k}\sum_{n=0}^{k-1}\boldsymbol{r}^{(n)}\right|\right|_2^2\leq\frac{1}{k}\sum_{n=0}^{k-1}||\boldsymbol{r}^{(n)}||_2^2. \tag{62}$$

Since Theorem 3 shows that $\boldsymbol{A}\boldsymbol{v}^{(k)}-\boldsymbol{c}\to\boldsymbol{0}$ for $k\to\infty$, the inequality (62) implies that we have $||\boldsymbol{r}^{(k-1)}||_2^2\leq\epsilon_r$ (and hence Algorithm 2 terminates) if $||\boldsymbol{A}\tilde{\boldsymbol{v}}^{(k)}-\boldsymbol{c}||_2^2\leq\epsilon_r$. Thus, it follows from (61) that Algorithm 2 terminates when

$$\frac{\rho}{2k}||\boldsymbol{A}(\boldsymbol{v}^{(0)}-\boldsymbol{v}^*)||_2^2+\frac{1}{k\rho}(||\boldsymbol{\theta}^{(0)}-\boldsymbol{\theta}^*||_2^2+1)\leq\epsilon_r. \tag{63}$$

Accordingly, we have

$$k\geq\frac{\frac{\rho}{2}||\boldsymbol{A}(\boldsymbol{v}^{(0)}-\boldsymbol{v}^*)||_2^2+\frac{1}{\rho}(||\boldsymbol{\theta}^{(0)}-\boldsymbol{\theta}^*||_2^2+1)}{\epsilon_r}. \tag{64}$$

Therefore, the ADMM-based Lagrangian maximization algorithm takes $O(1/\epsilon_r)$ iterations until the convergence. When the bisection searching is applied, the time complexity of updating a frequency variable and a transmit power variable in each iteration is $O(\log(\frac{1}{\epsilon_f}))$ and $O(\log(\frac{P_{s,\max}}{\epsilon_p}))$ (or $O(\log(\frac{P_{B,\max}}{\epsilon_p}))$), respectively. By (31)-(34), we thus have that the time complexity of updating these variables is $O(IS\log(\frac{1}{\epsilon_f})+\ IS\log(\frac{\max\{P_{B,\max},\max_s P_{s,\max}\}}{\epsilon_p}))$ in each iteration. By (35)-(36), we have that the time complexity of updating these variables is $O(IS)$ in each iteration. Therefore, the time complexity of ADMM-based Lagrangian maximization algorithm is $O(IS(\log(\frac{1}{\epsilon_f})+\log(\frac{\max\{P_{B,\max},\max_s P_{s,\max}\}}{\epsilon_p}))/\epsilon_r)$, and Theorem 5 follows.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 61572440, 61379122, and 91638204, in part by the Zhejiang Provincial Natural Science Foundation of China under Grants LR17F010002 and LR16F010003, in part by the Open Research Fund of the National Mobile Communications Research Laboratory, Southeast University under Grant 2019D11, in part by the Research Grant of University of Macau under Grant SRG2019-00168-IOTSC, in part by the Natural Sciences and Engineering Research Council, Canada, and in part by the NSF under Grant CNS-1651947, USA.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2015–2020," Cisco Visual Networking Index White Paper, Feb. 2015 [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[2] Y. Wu, Y. He, L. P. Qian, J. Huang, and X. S. Shen, "Optimal resource allocations for mobile data offloading via dual-connectivity," *IEEE Trans. Mobile Comput.*, vol. 17, no. 10, pp. 2349–2365, Oct. 2018.

[3] M. G. Khoshkholgh, K. Navaie, K. G. Shin, and V. C. M. Leung, "Cell association in dense heterogeneous cellular networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, pp. 1019–1032, May 2018.

[4] L. P. Qian, Y. Wu, H. Zhou, and X. S. Shen, "Joint uplink base station association and power control for small-cell networks with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5567–5582, Sep. 2017.

[5] M. Ismail, W. Zhuang, E. Serpedin, and K. Qaraqe, "A survey on green mobile networking: from the perspectives of network operators and mobile users," *IEEE Commun. Surveys Tut.*, vol. 17, no. 3, pp. 1535–1556, 3Q 2014.

[6] R. Martin, "Nearly 400,000 off-grid mobile telecommunications base stations employing renewable or alternative energy sources will be deployed from 2012 to 2020," Navigant Research, Feb. 2013. [Online]. Available: http://www.navigantresearch.com/newsroom/nearly-400000-off-grid-mobile-telecommunications-basestations-employing-renewable-or-alternative-energy-sources-will-bedeployed-from-2012-to-2020, Accessed: Mar. 9, 2015.

[7] H.-J. Hung, T.-Y. Ho, S.-Y. Lee, C.-Y. Yang, and D.-N. Yang, "Relay selection for heterogeneous cellular networks with renewable green energy sources," *IEEE Trans. Mobile Comput.*, vol. 17, no. 3, pp. 661–674, Mar. 2018.

[8] H. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. Andrews, "Fundamentals of heterogeneous cellular networks with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2782–2797, May 2014.

[9] L. P. Qian, G. Feng, and V. C. M. Leung, "Optimal transmission policies for relay communication networks with ambient energy harvesting relays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3754–3768, Dec. 2016.

[10] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. M. Leung, and H. V. Poor, "Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1936–1947, Sep. 2017.

[11] X. Wang, T. Chen, X. Chen, X. Zhou, and G. B. Giannakis, "Dynamic resource allocation for smart-grid powered MIMO downlink transmissions," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3354–3365, Dec. 2016.

[12] Y. L. Che, L. Duan, and R. Zhang, "Dynamic base station operation in large-scale green cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3127–3141, Dec. 2016.

[13] D. Zhang, Z. Chen, L. X. Cai, H. Zhou, S. Duan, J. Ren, X. S. Shen, and Y. Zhang, "Resource allocation for green cloud radio access networks with hybrid energy supplies," *IEEE Trans. Veh. Tech.*, vol. 67, no. 2, pp. 1684–1697, Feb. 2018.

[14] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 2872–2882, Aug. 2013.

[15] V. Chanola, B. Sikdar, and B. Krishnamachari, "Delay aware resource management for grid energy savings in green cellular base stations with hybrid power supplies," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1092–1104, Mar. 2017.
[16] S. Zhang, N. Zhang, S. Zhou, J. Gong, Z. Niu, and X. S. Shen, "Energy-aware traffic offloading for green heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1116–1129, May 2016.
[17] C. Hu, J. Gong, X. Wang, S. Zhou, and Z. Niu, "Optimal green energy utilization in MIMO systems with hybrid energy supplies," *IEEE Trans. Veh. Tech.*, vol. 64, no. 8, pp. 3675–3688, Aug. 2015.
[18] Y. Mao, J. Zhang, and K. B. Lataief, "A Lyapunov optimization approach for green cellular networks with hybrid energy supplies," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2463–2477, Dec. 2015.
[19] Q. Han, B. Yang, G. Miao, C. Chen, X. Wang, and X. Guan, "Backhaul-aware user association and resource allocation for energy-constrained HetNets," *IEEE Trans. Veh. Tech.*, vol. 66, no. 1, pp. 580–593, Jan. 2017.
[20] T. Han and N. Ansari, "Network utility aware traffic load balancing in backhaul-constrained cache-enabled small cell networks with hybrid power supplies," *IEEE Trans. Mobile Comput.*, vol. 16, no. 10, pp. 2819–2832, Oct. 2017.
[21] T. Zhang, W. Chen, and F. Yang, "Balancing delay and energy efficiency in energy harvesting cognitive radio networks: A stochastic stackelberg game approach," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 2, pp. 201–216, Jun. 2017.
[22] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3412–3427, Jul. 2013.
[23] Y. Wei, F. R. Yu, M. Song, and Z. Han, "User scheduling and resource allocation in HetNets with hybrid energy supply: an actor-critic reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 680–692, Jan. 2018.
[24] H. M. Doost, N. Prasad, and S. Rangarajan, "Optimizing energy efficiency over energy-harvesting LTE cellular networks," *IEEE Trans. Green Commun. Netw.*, vol. 1, no. 3, pp. 320–332, Sep. 2017.
[25] A. Yadav, T. M. Nguyen, and W. Ajib, "Optimal energy management in hybrid energy small-cell access points," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5334–5348, Dec. 2016.
[26] A. Nedicand and A. OZdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1757–1780, Feb. 2009.
[27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[28] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.
[29] X. Lin and N. B. Shroff, "Utility maximization for communication networks with multipath routing," *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 766–781, May 2006.
[30] S. Lee, N. Chatzipanagiotis, and M. M. Zavlanos, "Complexity certification of a distributed augmented lagrangian method," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 827–834, Mar. 2018.
[31] 3GPP, TR36.931, V9.0.0, "LTE; Evolved universal terrestrial radio access (E-UTRA); Radio Frequency (RF) requirements for LTE Pico Node B," Release 9, May 2011. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/136900_136999/136931/09.00.00_60/tr_136931v090000p.pdf

Li Ping Qian (S'08-M'10-SM'16) received the PhD degree in information engineering from The Chinese University of Hong Hong, Hong Kong, in 2010. She worked as a postdoctoral research associate with The Chinese University of Hong Kong, Hong Kong, during 2010-2011. Since 2011, she has been with the College of Information Engineering, Zhejiang University of Technology, China, where she is currently a full professor. From 2016 to 2017, she was a visiting scholar with the Broadband Communications Research Group, ECE Department, University of Waterloo. Her research interests include wireless communication and networking, resource management in wireless networks, massive IoTs, mobile edge computing, emerging multiple access techniques, and machine learning oriented towards wireless communications. She was a co-recipient of the IEEE Marconi Prize Paper Award in Wireless Communications in 2011, the Best Paper Award from IEEE ICC 2016, and the Best Paper Award from IEEE Communication Society GCCTC 2017. She is currently on the editorial board of *IET Communications*. She is a senior member of the IEEE.

Yuan Wu (S'08-M'10-SM'16) received the PhD degree in electronic and computer engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2010. He is currently an associate professor with the State Key Laboratory of Internet of Things for Smart City, University of Macau, and also with the Department of Computer and Information Science, University of Macau. Prior to that, he was a full professor with the College of Information Engineering, Zhejiang University of Technology, China. During 2010-2011, he was the postdoctoral research associate with The Hong Kong University of Science and Technology. During 2016-2017, he was with the Broadband Communications Research (BBCR) Group, Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include resource management for wireless networks, green communications and computing, mobile edge computing, and smart grids. He was a recipient of the Best Paper Award from the IEEE International Conference on Communications, in 2016, and the IEEE Technical Committee on Green Communications and Computing, in 2017. He is a senior member of the IEEE.

Bo Ji (M'12-SM'18) received the PhD degree in electrical and computer engineering from The Ohio State University, Columbus, Ohio, in 2012. He joined the Department of Computer and Information Sciences (CIS), Temple University, in July 2014, where he is currently an assistant professor. He is also a faculty member of the Center for Networked Computing (CNC), Temple University. His research interests include modeling, analysis, control, optimization, and learning of computer and networking systems, such as communication networks, information-update systems, cloud/datacenter networks, and cyber-physical systems. He is a National Science Foundation (NSF) CAREER awardee (2017) and an NSF CISE Research Initiation Initiative (CRII) awardee (2017). He is a recipient of the IEEE INFOCOM 2019 Best Paper Award. He is a senior member of the IEEE.

Xuemin (Sherman) Shen (M'97-SM'02-F'09) is an university professor and an associate chair for graduate studies with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on wireless resource management, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is the IEEE ComSoc VP Publication, was an elected member of IEEE ComSoc Board of Governor, and the chair of Distinguished Lecturers Selection Committee. He served as the technical program committee chair/co-chair for IEEE Globecom'16, Infocom'14, IEEE VTC'10 Fall, and Globecom'07, the Symposia chair for IEEE ICC'10, the tutorial chair for IEEE VTC'11 Spring and IEEE ICC'08, the general co-chair for ACM Mobihoc'15, and the chair for IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as an editor-in-chief for the *IEEE Internet of Things Journal*, and the *IEEE Network*, a founding area editor for the *IEEE Transactions on Wireless Communications*, and an associate editor for the *IEEE Transactions on Vehicular Technology* and the *IEEE Wireless Communications*, etc. He received the IEEE ComSoc Education Award, the Joseph LoCicero Award for Exemplary Service to Publications, the Excellent Graduate Supervision Award in 2006, and the Premiers Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. He is a registered professional engineer of Ontario, Canada, a fellow of the IEEE, an Engineering Institute of Canada fellow, a Canadian Academy of Engineering fellow, a Royal Society of Canada fellow, and a distinguished lecturer of IEEE Vehicular Technology Society and Communications Society.