Energy-Efficient Resources Allocation With Millimeter-Wave Massive MIMO in Ultra Dense HetNets by SWIPT and CoMP

Bin Li, Yonghong Dai, Zhicheng Dong, *Member, IEEE*, Erdal Panayirci, *Life Fellow, IEEE*, Huilin Jiang, *Member, IEEE*, and Hao Jiang

***Abstract*— Ultra-dense HetNets (UDN)-based Millimeter-Wave (mmWave) massive MIMO is considered a promising technology for 5th generation (5G) wireless communications systems since it can offer massively available bandwidth and improve energy efficiency (EE) substantially. However, in UDN, the power consumption of the system increases sharply with the increase of network density. In this paper, we investigate the optimization of the EE in the mmWave massive MIMO systems with UDN. To develop the functions of massive MIMO, we first propose a system model where the massive MIMO harvests electromagnetic energy from the environment employing simultaneous wireless information and power transfer (SWIPT) technology, implemented at the base station (BS). Then, the EE optimization problem is formulated for 5G mmWave massive MIMO systems within the UDN. Considering the nonconcave feature of the objective function, an iterative EE algorithm is developed, based on Dinkelbach method. To utilize the role of coordinated multi-point transmission and reception (CoMP) for improving the EE, a coordinated user(UE)-BS association algorithm-based CoMP with maximum energy efficiency (MaxEE) is proposed. The simulation results demonstrate that the proposed algorithm has a substantially faster convergence rate and is very effective, compared with existing methods.**

***Index Terms*— Energy efficiency, millimeter-wave communcation, massive MIMO, ultra dense HetNet, CoMP, SWIPT BS.**

Manuscript received November 30, 2019; revised October 15, 2020; accepted February 2, 2021. Date of publication February 19, 2021; date of current version July 12, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61801227, Grant 61561046, and Grant 61273053; in part by the National Natural Science Foundation of Jiangsu Higher Education Institutions under Grant 18KJB510022; in part by the Key Research & Development and Transformation Plan of Science and Technology Program for Tibet Autonomous Region under Grant XZ201901-GB-16; in part by the Central government supports the reform and development of local universities of Tibet University in 2019 and 2020; and in part by Special fund for the development of local universities supported by the central finance of Tibet University in 2018. The associate editor coordinating the review of this article and approving it for publication was A. Zaidi. *(Corresponding author: Yonghong Dai.)*

Bin Li and Yonghong Dai are with the School of Electronic Information, Wuhan University, Wuhan 430072, China (e-mail: joy2bonbon@163.com; yhdai@whu.edu.cn).

Zhicheng Dong is with the School of Electrical Engineering, Tibet University, Lhasa 850000, China (e-mail: dongzc666@163.com).

Erdal Panayirci is with the Department of Electrical and Electronics Engineering, Kadir Has University, 34230 İstanbul, Turkey, and also with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: eepanay@princeton.edu).

Huilin Jiang is with the School of Electronic Engineering, Nanjing Xiaozhuang University, Nanjing 210096, China (e-mail: huilin.jiang@njxzc.edu.cn).

Hao Jiang is with the School of Electronic Information, Wuhan University, Wuhan 430072, China, and also with the Collaborative Innovation Center for Geospatial Technology, Wuhan 430079, China (e-mail: jh@whu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3058776.

Digital Object Identifier 10.1109/TWC.2021.3058776

I. Introduction

ULTRA-DENSE HetNets (UDN)-based millimeter-wave (MmWave) massive MIMO is considered a promising technology for 5th generation (5G) wireless communications systems since it can offer massive bandwidth availability and substantially improve energy efficiency (EE). However, in UDNs, the power consumption of the system increases sharply with increases in network density. In this paper, we investigate the optimization of EE in mmWave massive MIMO systems with UDNs. To develop the functions of massive MIMO, we first propose a system model to improve the EE of a communication system, which is one of the key problems in the research and development of next-generation wireless communication networks. One reason why MmWave massive MIMO is considered a promising technology for 5G communication systems [1] is that the short wavelength of mmWave makes it feasible to deploy massive MIMO centralized or distributed [43], [44] in small base stations (BSs). It can also provide sufficient antenna gain to compensate for severe path loss caused by mmWave propagation [2]. One of the key advantages of massive MIMO technology lies in the potential gains it can achieve in terms of EE [3]. Recently, UDNs have received considerable research attention since they can drastically improve the spectral efficiency of a network [4]. However, with the rapid increase in density of the deployed small cells, network energy consumption has also increased significantly. Hence, improving EE in communications networks by employing mmWave massive MIMO technology has become an important research area in the design and development of UDN-based 5G communications systems [5], [6].

Although mmWave massive MIMO and UDNs can improve system performance, they also introduce some new challenges, including the problem of EE. As mentioned above, while the utilization of mmWave massive MIMO in UDNs can substantially improve system performance, it can also significantly increase a system's power consumption due to the sharp

increase in network density. Hence, it is urgent to solve the problems of how to reduce a system's power consumption and improve EE in such networking scenarios. Recovering energy from the environment is a promising technology for improving the EE of wireless networks [7], [8]. In addition to the well-known renewable energy sources (such as solar energy, wind energy, etc.), radio frequency (RF) signals generated by RF transmitters have become a new kind of renewable energy. Since electromagnetic waves cannot only transmit information but also transmit energy, simultaneous wireless information and power transfer (SWIPT) has become an attractive technology and received widespread interest [9].

It is well known that coordinated multi-point transmission and reception (CoMP) is an effective technology for improving EE. Compared with the non-cooperative transmission mode, the EE of cell edge UEs increases substantially because the multiple BSs cooperate to transmit the UE data as well as with the channel state information (CSI). Furthermore, the coverage area of BSs can be expanded by adopting multicell cooperation, thereby reducing the number of active BSs required to cover a certain area [11].

A. Related Work

UE association and load balancing have been investigated extensively [12]–[15]. In [12], a load-aware UE association scheme and maximum energy efficiency (MaxEE) were proposed. In [13], the joint UE association and power allocation-based MaxEE in two-tier heterogeneous networks was investigated. The power allocation and UE association in massive MIMO multi-cell scenarios were studied in [14]. In [15], the problem of UE association and resource allocation in energy-constrained HetNets were investigated and a distributed backhaul-aware UE association algorithm was proposed. In [42], C. Wang *et al.* investigated the achievable performance levels of downlink cloud radio access networks with BS cooperation, and two coding schemes for this scenario were developed and analyzed. The main difference between this paper and the works [12]-[15] is that here we adopt a setting in which the UEs can access to multiple BSs at the same time, namely, CoMP. Although the adoption of CoMP technology will make the system more complex, the work [11] shows that CoMP is an effective technology for improving system EE.

There have also been several studies on SWIPT, including [16]–[19]. In [16], the SWIPT and EE problems in 5G HetNet were studied. In [17], the concept of energy-pattern-aided SWIPT was proposed. In [18], the near-far problem of power transfer and information transfer in SWIPT was investigated. In [19], the relay problem of SWIPT in multi-user MIMO decode-and-forward relay broadcasting channels was considered. However, these works do not take into account the case whereby electromagnetic energy from an environment is recovered by SWIPT technology at massive MIMO BSs. The fundamental difference between this paper and the above-mentioned works [16]–[19] is that this paper considers an energy recovery system installed in the Massive MIMO BS, while the above-mentioned works [16]–[19] are concerned with recovering electromagnetic wave energy in the environment through UE equipment or the deployment of energy recovery nodes in the macrocell. The purpose of considering an energy recovery system for a Massive MIMO BS is to take full advantage of Massive MIMO BSs' large number of antennas and to reduce their power consumption. Given that the importance of massive MIMO for 5G systems has been pointed out, as well as the feasibility and necessity of deploying massive MIMO antennas in small cells (as highlighted in [1], [2]), it has become an important issue today to investigate the problem of deploying massive MIMO antennas in all BSs in UDNs and adopting SWIPT technology to harvest energy.

There have also been a number of studies on EE and resource allocation, including [20]–[24], [34], [36]–[38]. In [20], the EE problem in 5G networks was studied, and a power efficiency framework with quality of service (QoS)-driven green power allocation schemes was proposed. In [21], three hybrid precoding algorithms based on EE were introduced in 5G systems. In [22], throughput performance analysis of the chunk-based subcarrier allocation was presented by considering the average bit-error-rate constraint over a chunk in downlink multiuser orthogonal frequency division multiplexing transmission. In [23], a chunk-based resource allocation scheme was proposed and analyzed by maximizing the throughput under a total transmit power constraint. In [24], H. Zhu proposed a novel resource allocation scheme by considering multiple bit error rate requirements for different types of packets in one data stream. In [34], the interference problem between femtocell UEs and macrocell UEs was investigated in UDNs. In [36], a power and rate adaptation scheme based on multiple quadrature amplitude modulation was investigated. In [37], a power adaptation scheme based on the terminal velocity was proposed, while in [38], a conditional power and rate adaptation technique for OFDM [22]–[24] systems was proposed. However, these works did not take into account the problem of EE in the mmWave massive MIMO UDN scenario. The work [6] shows that mmWave massive MIMO UDN has become an important issue in 5G systems. Therefore, the rest of this paper will investigate the EE of mmWave Massive MIMO under UDN scenario, and examine the differences between this paper and the works [20]–[24] and [34], [36]–[38].

B. Motivations

The feasibility and necessity of deploying massive MIMO BS in small cells have been highlighted in [2]. Also, there are only a few investigations in the open literature that involve deploying massive MIMO BS in small cells. Furthermore, it is shown in [6] that the mmWave massive MIMO UDN constitutes an important network structure of 5G. However, there are only a few studies in the open literature regarding EE in mmWave massive MIMO with UDN scenarios, particularly, in small cell deploys with massive MIMO. Given the large number of exploitable antennas in the system, using these antennas to recover electromagnetic energy from the environment can certainly improve a system's EE. To the best of the

authors' knowledge, there have been no previous studies of the optimization problem of EE in the BSs with mmWave massive MIMO-based UDN, equipped with an energy harvesting system (EHS) to harvest the electromagnetic energy from the environment. Motivated by this fact, the present work focuses on the EE problem of mmWave massive MIMO in UDNs by utilizing the large number of antennas that exist in the system to recover the electromagnetic energy from the environment, thereby fully exploiting the EHS.

C. Main Contribution

The main contributions of this paper can be summarized as follows.

- A new system model that the massive MIMO BS retrieves electromagnetic energy from the environment is proposed.
- An EE optimization model of mmWave massive MIMO under the UDN scenario is formulated based on a fractional programming problem. To solve the resulting optimization problem, we first adopt the Dinkerbach method to transform the original problem with fractional form into the form of a difference. The Lagrange dual decoupling method is then employed to further transform the problem into a quasi-convex form. Finally, to remove the coupling between variables and to solve the resulting problems, we adopt the original decoupling method to transform the quasi-convex problem into five sub-optimizations problems.
- To determine the best UE-BS connection scheme, we propose an UE and BS association algorithm based CoMP and MaxEE in mmWave massive MIMO UDN scenario.
- To optimize the transmission power of massive MIMO BS and improve the EE of the system, as well as to ensure the QoS requirements of each UE, we propose a power optimization algorithm for massive MIMO BS-based PRP conjugate gradient method.
- The convergence of the proposed algorithms is analyzed and a fast convergence of the proposed algorithm is verified by computer simulations. At the same time, the effectiveness of the proposed algorithm is demonstrated numerically compared to other algorithms, presented in the literature.

The remainder of the paper is organized as follows: Section II presents the system model, the power consumption model, and problem formulations. Section III presents the transformation of the objective function and iterative algorithm for EE maximization to find an optimal solution to the proposed problem. Section IV includes an analysis of the performance of our proposed algorithms via computer simulations. Section V concludes the paper.

II. System Model and Problem Formulation

A. System Model

We consider a CoMP downlink transmission model-based mmWave massive MIMO UDN, with joint processing/joint transmission (JP/JT) scheme. The UDN adopts double-layer macro-pico network, which consists of ultra-dense picocells overlaid on one macrocell. Moreover, all cells adopt full-frequency multiplexing mode, and the total bandwidth of the system is W. The macro base station (MBS) and all of the pico base stations (PBSs) are equipped with N_t transmission antennas [2], and all UEs are equipped with single antenna; the antennas of all BSs can be dynamically activated and deactivated [45]. They are equipped with EHS to harvest the energy through electromagnetic waves in the environment [12]. The set of all BSs is denoted by $\mathcal{B} = \{1, 2, \ldots, B\}$, $\forall j \in \mathcal{B}$. The set of antennas of j-th BS is denoted by $\mathcal{C}_A^j = \{1, 2, \ldots, N_t\}$, $\forall n \in \mathcal{C}_A^j$; and the set of all UEs is denoted by $\mathcal{U} = \{1, 2, \ldots, U\}$, $\forall i \in \mathcal{U}$, as illustrated in Fig.1.

In this network, we focus on the UE and BS association, the optimization of the number of active antennas, the UE and antenna association and the transmission power allocation for BSs on a downlink transmission scenario. We use x_{ij} to denote the association variable of the i-th UE and j-th BS. If i-th UE is associated with j-th BS, $x_{ij} = 1$, otherwise $x_{ij} = 0$, namely,

$$x_{ij} = \begin{cases} 1, & \text{if the UE } i \text{ associated with the BS } j, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

$\mathbf{X}$ denotes the association matrix between UEs and BSs, namely

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1B} \\ \vdots & \ddots & \vdots \\ x_{U1} & \cdots & x_{UB} \end{pmatrix}. \tag{2}$$

Since the CoMP JP/JT scheme is adopted, the UE can access to multiple BSs at the same time. Thus,

$$\sum\nolimits_j x_{ij} \leq M, \tag{3}$$

where M represents the maximum number of accessible BSs for the UE.

The notation $S = [s_{ijn}]_{U \times B \times N_t}$ denotes the association indication matrix between the UEs and antennas. That is, if the n-th antenna of the j-th BS is allocated to its associated i-th UE then $s_{ijn} = 1$, otherwise $s_{ijn} = 0$, namely

$$s_{ijn} = \begin{cases} 1, & \text{if the antenna } n \text{ of the BS } j \text{ allocates to} \\ & \text{the UE } i \text{ associated with the BS } j, \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where

$$x_{ij} = \left\lceil \frac{\sum_n s_{ijn}}{N_t} \right\rceil. \tag{5}$$

To save energy and improve the EE of the system, we consider switching the antennas of the BSs on and off (on/off) dynamically. When all antennas are deactivated at a BS, the BS is dormant. We use j_{jn} to denote the switching on/off indication of the antennas of a given BS, where $j \in \mathcal{B}$ and

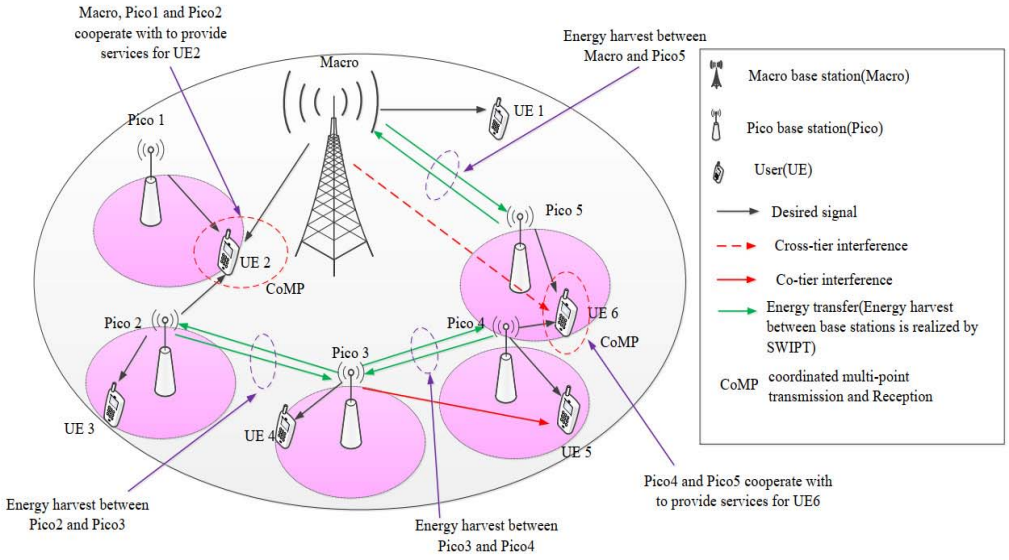


Fig. 1. System model.

$n \in \mathcal{C}_A^j$. If the n-th antenna of j-th BS is activated, then $j_{jn} = 1$, otherwise $j_{jn} = 0$, namely

$$j_{jn} = \begin{cases} 1, & \text{if the antenna } n \text{ of BS } j \text{ is activated,} \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

We use $\mathbf{J}$ to denote the switching off/on indication matrix of antennas of all BSs, namely

$$\mathbf{J} = \begin{pmatrix} j_{11} & \cdots & j_{1N_t} \\ \vdots & \ddots & \vdots \\ j_{B1} & \cdots & j_{BN_t} \end{pmatrix}, \tag{7}$$

where

$$j_{jn} = \sum\nolimits_i s_{ijn}. \tag{8}$$

Also b_j denotes the activation and dormant indicator of j-th BS. If j-th BS is not associated with the any i-th UE ($i \in \mathcal{U}$), then the j-th BS is dormant, $b_j = 0$; otherwise $b_j = 1$. That is,

$$b_j = \begin{cases} 1, & \text{otherwise,} \\ 0, & \text{if } \sum_i x_{ij} = 0, \end{cases} \tag{9}$$

where

$$b_j = \left\lceil \frac{\sum_i x_{ij}}{K_U^j} \right\rceil, \tag{10}$$

and K_U^j represents the maximum number of service UEs of the j-th BS.

The vector $\mathbf{b}$ denotes the activation and dormant indicator of all BSs, namely

$$\mathbf{b} = [b_1, \cdots, b_j, \cdots, b_B]^T. \tag{11}$$

$\mathbf{P}$ is the transmission power vector of all BSs, namely,

$$\mathbf{P} = \left[P_{TX}^1, \cdots, P_{TX}^j, \cdots, P_{TX}^B\right]^T. \tag{12}$$

We use $\mathbf{N}_{AA}$ to denote the activated antenna number vector of all BSs, that is,

$$\mathbf{N}_{AA} = \left[N_{AA}^1, \cdots, N_{AA}^j, \cdots, N_{AA}^B\right]^T, \tag{13}$$

where

$$N_{AA}^j \le N_t. \tag{14}$$

Similarly, $\mathbf{K}$ and $\mathbf{K}_U$ denote the vectors of the number of UEs accessed by all BSs and the maximum number of UEs served by all BSs, respectively. Namely,

$$\mathbf{K} = [K_1, \cdots, K_j, \cdots, K_B]^T, \tag{15}$$

$$\mathbf{K}_U = \left[K_U^1, \cdots, K_U^j, \cdots, K_U^B\right]^T, \tag{16}$$

where

$$K_j = \sum\nolimits_i x_{ij}, \quad \forall j \in \mathcal{B}, \quad 0 \le K_j \le K_U^j. \tag{17}$$

The received signal at the i-th UE can be expressed as:

$$y_i = \sum_{j \in \mathcal{C}_i} \mathbf{h}_{ij}^H \mathbf{w}_{ij} \mathbf{s}_{ij} + \sum_{k \in \mathcal{B}/\mathcal{C}_i} \mathbf{h}_{ik}^H \mathbf{w}_{ik} \mathbf{s}_{ik} + n_i, \tag{18}$$

where $\mathbf{s}_{ij} \in \mathcal{C}^{K_j \times 1}$ is the signal vector transmitted from the j-th BS to K_j UEs, $\mathbf{h}_{ij}^H \in \mathcal{C}^{1 \times N_t}$ is the downlink channel vector between j-th BS and i-th UE and $\mathbf{w}_{ij} \in \mathcal{C}^{N_t \times K_j}$ is the precoding matrix. $n_i \in \mathcal{CN}(0,1)$ is the additive white Gaussian noise (AWGN) and $\mathcal{C}_i$ denotes the coordinated set of the i-th UE. In this paper, since the CoMP JP/JT scheme is adopted, we have similar assumptions as in [39]. Namely, the backhaul between the macro-BS and the small-BSs is assumed to be capacity-sufficient. The signal transmitted at the j-th BS can be expressed as $\mathbf{x_j} = \sum_{i \in \mathcal{C}_j} \mathbf{w}_{ij} \mathbf{s}_{ij}$, where C_j is the set of service UEs of the j-th BS [40], [41].

It is known that mmWave channel does not obey the conventional Rayleigh fading model due to the limited scattering.

In this paper, we consider the *geometric Saleh-Valenzuela channel model* that is more appropriate for mmWave communications [25], [26]. In this model, the channel matrix $\mathbf{H}$ can be expressed as:

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{L\rho}} \sum_{l=0}^{L} \alpha_l \Lambda_r(\phi_l^r) \Lambda_t(\phi_l^t) \mathbf{f}_r(\phi_l^r) \mathbf{f}_t(\phi_l^t), \quad (19)$$

where L denotes the number of the multipath between the BSs and UEs. ρ is the path loss exponent. α_l denotes the complex gain of the l-th path, which follows the Rayleigh distribution. ϕ_l^t and ϕ_l^r represent azimuth angles of departure or arrival(AoDs/AoAs). $\Lambda_t(\phi_l^t)$ and $\Lambda_r(\phi_l^r)$ denote the transmit and receive antenna array gains at a specific AoD and AoA, respectively. $\mathbf{f}_t(\phi_l^t)$ and $\mathbf{f}_r(\phi_l^r)$ are the antenna array response vectors respectively.

In this paper, both the transmit antenna array and the receive antenna array have a widely used uniform linear arrays (ULA) structure. Then, $\mathbf{f}_t(\phi_l^t)$ and $\mathbf{f}_r(\phi_l^r)$ can be expressed as [27]:

$$\mathbf{f}_t(\phi_l^t) = \sqrt{\frac{1}{N_t}} \left[1, e^{j\frac{2\pi}{\lambda} d \sin(\phi_l^t)}, \cdots, e^{j(N_t-1)\frac{2\pi}{\lambda} d \sin(\phi_l^t)}\right]^T, \quad (20)$$

$$\mathbf{f}_r(\phi_l^r) = \sqrt{\frac{1}{N_r}} \left[1, e^{j\frac{2\pi}{\lambda} d \sin(\phi_l^r)}, \cdots, e^{j(N_r-1)\frac{2\pi}{\lambda} d \sin(\phi_l^r)}\right]^T, \quad (21)$$

where λ denotes the wavelength of the signal and d represents the distance between two adjacent antenna elements, $d = \frac{\lambda}{2}$. Note that only the azimuth angles of AoAs and AoDs are considered in this work.

When the i-th UE is associated with the j-th BS, based on the datarate formula in [28]–[35], the signal-to-interference-and-noise ratio ($SINR_{ij}$) of i-th UE can be expressed as:

$$SINR_{ij} = \frac{N_{AA}^j - K_j + 1}{K_j} \frac{P_{TX}^{ij} G_{ij}}{\sum_{k \in \mathcal{C}, k \neq j} P_{TX}^{ik} G_{ik} + \sigma_i^2}, \quad (22)$$

where ${\sigma_i}^2$ is the noise power. $G_{ij} = \sum_n s_{ijn} G_{ijn}$, $G_{ijn} = \left\| h_{ijn}^H w_{ijn} \right\|^2$ is the channel gain between the n-th antenna of j-th BS and i-th UE, h_{ijn} denotes channel coefficient between the n-th antenna of j-th BS and i-th UE. P_{TX}^{ij} is the power that the j-th BS allocates to the i-th UE, namely,

$$P_{TX}^{ij} = \sum_n s_{ijn} \frac{P_{TX}^j}{N_{AA}^j}, \quad (23)$$

where the antenna power adopts the equal allocation scheme.

The achievable datarate of the i-th UE can be defined as:

$$R_i = \sum\nolimits_j x_{ij} r_{ij}, \quad (24)$$

where r_{ij} denotes the achievable datarate of the i-th UE from the j-th BS, given by

$$r_{ij} = W_j \log_2\left(1 + SINR_{ij}\right). \quad (25)$$

W_j in (25) is the bandwidth allocated by the j-th BS to the i-th UE. That is, $W_j = \frac{W}{K_j}$ with K_j being the total number of UEs associated with the j-th BS. Hence each UE can receive $\frac{1}{K_j}$ of the total frequency band available. Consequently, R_i can be rewritten as:

$$R_i = \sum\nolimits_j x_{ij} W_j \log_2(1 + SINR_{ij}), \quad (26)$$

and the sum-rate of all UEs in this system can be expressed as:

$$U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) = \sum\nolimits_i R_i. \quad (27)$$

B. Power Consumption Model

In this paper, we consider that each BS is equipped with an EHS based SWIPT, receiving energy for replenishing from a rechargeable battery. Therefore, we model the power consumption under energy harvesting by BSs, represented by P_{sum} as [29]:

$$P_{\text{sum}} = \frac{P_{TX}}{\mu} + P_{BCP} + P_{SBB} + P_{SRF} + P_{SPS} - P_{EH}, \quad (28)$$

where $\frac{P_{TX}}{\mu}$ denotes the total transmission power and μ represents the power amplifier coefficient. P_{BCP} is the total circuit power of the BS, including power supply, cooling system, etc.. P_{SBB} is the total power of baseband processing unit and finally, P_{SRF}, P_{SPS}, P_{EH} denote total power of radio frequency link, the total power of phase shifter and total energy harvesting from the EHS, respectively. Total transmission power can be expressed as $P_{TX} = \sum_j b_j P_{TX}^j$, where P_{TX}^j is the transmission power of j-th BS. The remaining terms are described as follows:

$P_{BCP} = P_{SB} N_{SB} + P_{AB} N_{AB}$, where P_{SB} and N_{SB} are the power and number of sleeping BSs, respectively, P_{AB} and N_{AB} denote the circuit power and the number of activated BSs, respectively, where $N_{AB} = \sum_j b_j$. The BS only has two states (namely, sleep and activation), $N_{SB} = N_{\text{sum}B} - N_{AB}$, $N_{\text{sum}B} = B$ represents the total number of BSs in the system. $P_{SBB} = P_{BB} N_{AB}$, N_{AB} denotes the number of activated BSs, and P_{BB} represents the power of baseband processing unit. $P_{SRF} = N_{RF} P_{RF} N_{AB}$ with N_{RF} and P_{RF} representing the number of RF links and the power of the RF link, respectively. $P_{SPS} = N_{RF} N_{TX} P_{PS}$ with P_{PS} and N_{TX} denoting the phase shifter power and the number of transmit antennas, respectively. $N_{TX} = N_{AA}$ with N_{AA} being the total number of activated antennas. $N_{AA} = \sum_j b_j N_{AA}^j$ with N_{AA}^j representing the number of activated antennas for the j-th BS. P_{EH} is the total energy recovered from the EHS.

Based on previous studies [12], we propose an energy recovery model that massive MIMO BS harvests electromagnetic energy from the environment. For information receiving and energy recovery, we adopt the Time-Switching model on each antenna of the massive MIMO BS. The specific content of Time-Switching model is beyond the scope of this article. Although the energy recovery model proposed in this paper is based on the one introduced in [12], that is fundamentally different from [12]. In our proposed model, the activation of each antenna in the Massive MIMO BS is described in detail. Because the information and energy are received via antennas, it is important to consider the state (Switching on or off) of each antenna in detail in the model. Therefore, the proposed

model is more detailed and practical in real applications. The total energy recovered from the EHS is modeled as:

$$P_{EH} = \eta \sum_{j} b_j \Big(\sum_{k,k\neq j} b_k \sum_{n} j_{jn} \sum_{k} j_{km} P_{TX}^{k,j,m,n} |G_{k,j}^{m,n}|^2 \Big), \tag{29}$$

where η represents the energy recovery efficiency coefficient; b_j is the activation and dormant indicator of the j-th BS. j_{jn} represents the switching on/off indication of BS antennas. $P_{TX}^{k,j,m,n}$ represents the transmission power of the m-th antenna of the k-th BS to the n-th antenna of the j-th BS. $P_{TX}^{k,j,m,n} = \frac{P_{TX}^{k}}{N_{AA}^{k}}$, P_{TX}^{k} is the transmission power of the k-th BS, N_{AA}^{k} denotes the number of active antennas for the k-th BS. $G_{k,j}^{m,n} = \|\mathbf{h}_{k,j,m,n}^{H} \mathbf{w}_{k,j,m,n}\|^2$ is the channel gain from the m-th antenna of the k-th BS to the n-th antenna of the j-th BS.

Total power consumption of the system can be expressed as:

$$\begin{aligned} &P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \\ &= \frac{P_{TX}}{\mu} + P_{BCP} + P_{SBB} + P_{SRF} \\ &\quad + P_{SPS} - \eta \sum_{j} b_j \Big(\sum_{k,k\neq j} b_k \sum_{n} j_{jn} \\ &\quad \times \sum_{m} j_{km} P_{TX}^{k,j,m,n} |G_{k,j}^{m,n}|^2 \Big). \end{aligned} \tag{30}$$

C. Problem Formulation

The downlink EE of the mmWave massive MIMO system is defined as:

$$EE = \frac{\sum_{i=1}^{U} R_i}{P_{\text{sum}}} \text{bits/Joule/Hz}. \tag{31}$$

Since the logarithmic utility function can improve the load of BS, this paper adopts the logarithmic utility function to balance the load of the network, so the EE of the network can be rewritten as follows:

$$EE = \frac{\sum_{i=1}^{U} \log R_i}{P_{\text{sum}}} \text{bits/Joule/Hz}. \tag{32}$$

Consequently, the problem of EE in mmWave massive MIMO can be modeled as follows:

$$\begin{aligned} \max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} \; & EE = \frac{\sum_i \sum_j x_{ij} \log(W_j \log_2(1 + SINR_{ij}))}{P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})} \\ s.t.\; & C1: x_{ij} \in \{0,1\}, \forall i,j, \\ & C2: \textstyle\sum_j x_{ij} \leq M, \forall i, \\ & C3: \textstyle\sum_i x_{ij} \leq K_U^j, \forall j, \\ & C4: 0 \leq P_{TX}^j \leq P_{TX}^{max}, \forall j, \\ & C5: R_{QoS} \leq R_i, \forall i, \\ & C6: s_{ijn} \in \{0,1\}, \forall i,j,n, \\ & C7: \textstyle\sum_i x_{ij} \leq N_{AA}^j \leq N_t, \forall j. \end{aligned} \tag{33}$$

where C1 and C2 are the restrictions on the number of UEs associated BSs. C3 is the constraint that each BS services a number of UEs. C4 is the constraint of transmission power of BSs. C5 is the constraint of quality of service for UEs. C6 is the constraint of UE and antenna association. C7 is the constraint of number of active antennas.

III. Proposed Algorithms

The objective function in (33) is a ratio of two functions, which is generally a nonconvex function. This problem is a mixed-integer optimization problem that has very high complexity due to the non-convexity of the objective function and the binary nature of the variables in $\mathbf{X}$ and $\mathbf{S}$. To address the latter problem, we introduce the following transformation.

A. Transformation of the Objective Function

To derive an efficient algorithm to solve the considered problem, we introduce a transformation to handle the objective function via nonlinear fractional programming [30]. Without loss of generality, the maximum weighted EE, q^*, is defined as:

$$\begin{aligned} q^* &= \frac{U(\mathbf{X}^*, \mathbf{N}_{AA}^*, \mathbf{S}^*, \mathbf{P}_{TX}^*)}{P_{\text{sum}}(\mathbf{X}^*, \mathbf{N}_{AA}^*, \mathbf{S}^*, \mathbf{P}_{TX}^*)} \\ &= \max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} \frac{U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})}{P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})}. \end{aligned} \tag{34}$$

The maximum weighted EE q^* is achieved if and only if Eq. (35) is satisfied,for $U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \geq 0$ and $P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) > 0$ [30]. That is,

$$\begin{aligned} &\max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \\ &\quad - q^* P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \\ &= U(\mathbf{X}^*, \mathbf{N}_{AA}^*, \mathbf{S}^*, \mathbf{P}_{TX}^*) \\ &\quad - q^* P_{\text{sum}}(\mathbf{X}^*, \mathbf{N}_{AA}^*, \mathbf{S}^*, \mathbf{P}_{TX}^*) = 0. \end{aligned} \tag{35}$$

For any objective function in fractional form, there exists an equivalent objective function in subtractive form, e.g. $U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) - q^* P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})$ that shares the same optimal resource allocation policy. Hence, we focus on the equivalent objective function for finding the optimal solution of proposed problem in the rest of the paper.

The optimization problem in (33) can be restated as:

$$\begin{aligned} \max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} \; & U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \\ & - q P_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \\ s.t.\; & C1 - C7, \end{aligned} \tag{36}$$

where q is the weight of EE.

Setting $L(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) = U(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) - qP_{\text{sum}}(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})$, $L(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})$ can then be rewritten as:

$$\begin{aligned} &L(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX}) \\ &\quad = \sum_i \sum_j x_{ij} \log(W_j \log_2(1 + SINR_{ij})) \\ &\qquad - q\Big(\frac{\sum_j b_j P_{TX}^j}{\mu} + C_0 + C_1 \sum_j b_j + C_2 \sum_j b_j N_{AA}^j\Big) \\ &\qquad + q\eta \sum_j b_j \Big(\sum_k b_k \sum_n j_{jn} \sum_m j_{km} \frac{P_{TX}^k}{N_{AA}^k} |G_{k,j}^{m,n}|^2\Big), \end{aligned} \tag{37}$$

where

$$C_0 = P_{SB} \times B, C_1 = P_{BB} - P_{SB} + P_{AB} + N_{RF} \times P_{RF},$$
$$C_2 = N_{RF} \times P_{PS}. \tag{38}$$

B. Iterative Algorithm for EE Maximization

As stated earlier, the problem (36) is a mixed-integer optimization problem, which has very high complexity due to the non-convexity of the objective function and the binary nature of optimization variables. To address the latter problem, we relax the binary-valued variables, yielding Eq. 36 expressed as:

$$\begin{aligned} \max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} \; & U(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}) \\ & -qP_{\text{sum}}(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}) \\ s.t.\; & C1: 0 \le x_{ij} \le 1, \forall i,j, \\ & C2: \textstyle\sum_j x_{ij} \le M, \forall i, \\ & C3: \textstyle\sum_i x_{ij} \le K_U^j, \forall j, \\ & C4: 0 \le P_{TX}^j \le P_{TX}^{max}, \forall j, \\ & C5: R_{QoS} \le R_i, \forall i, \\ & C6: 0 \le s_{ijn} \le 1, \forall i,j,n, \\ & C7: \textstyle\sum_i x_{ij} \le N_{AA}^j \le N_t, \forall j. \end{aligned} \tag{39}$$

Owing to the non-convexity of objective function (39), we use the Lagrangian dual decomposition method to solve the relaxed problem (39), which is given as:

$$\begin{aligned} & L(x_{ij}, b_j, N_{AA}^j, s_{ijn}, j_{jn}, P_{TX}^j, \boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}) \\ & = \textstyle\sum_i \sum_j x_{ij} \log(W_j \log_2(1 + SINR_{ij})) \\ & \quad -q\Big(\frac{\sum_j b_j P_{TX}^j}{\mu} + C_0 + C_1 \sum_j b_j + C_2 \sum_j b_j N_{AA}^j\Big) \\ & \quad +q\eta \sum_j b_j \Big(\sum_k b_k \sum_n j_{jn} \sum_m j_{km} \frac{P_{TX}^k}{N_{AA}^k} |G_{k,j}^{m,n}|^2\Big) \\ & \quad + \textstyle\sum_j \tau_j (K_U^j - \sum_i x_{ij}) + \sum_i \nu_i (M - \sum_j x_{ij}) \\ & \quad + \textstyle\sum_j \omega_j (P_{TX}^{max} - P_{TX}^j) + \sum_i \alpha_i (R_i - R_{QoS}) \\ & \quad + \textstyle\sum_j \gamma_j (N_{AA}^j - \sum_i x_{ij}) + \sum_j \rho_j (N_t - N_{AA}^j), \end{aligned} \tag{40}$$

where $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_B)^T$,$\boldsymbol{\nu} = (\nu_1, \nu_2, \cdots, \nu_U)^T$, $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_B)^T$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_U)^T$, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \cdots, \gamma_B)^T$, $\boldsymbol{\rho} = (\rho_1, \rho_2, \cdots, \rho_B)^T$ are the Lagrange multipliers used to relax the coupled constraint.

Therefore, the Lagrangian dual function may be determined from

$$\begin{aligned} & D(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}) \\ & \quad = \max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} L(x_{ij}, b_j, N_{AA}^j, s_{ijn}, j_{jn}, P_{TX}^j, \\ & \quad \times \boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}). \end{aligned} \tag{41}$$

Consequently, then, the Lagrangian dual problem of problem (41) can be formulated as:

$$\min_{\boldsymbol{\tau},\boldsymbol{\nu},\boldsymbol{\omega},\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\rho}} D(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}). \tag{42}$$

Since these optimization variables are coupled, we introduce the primal decomposition method to separate the original problem into five subproblems (namely: **Subproblem1:** the UE and BS association; **Subproblem2:** the optimization of the number of the activated antenna; **Subproblem3:** the UE and antenna association; **Subproblem4:** the transmit power allocation of BS; **Subproblem5:** the iterative algorithm for EE maximization.). For each subproblem, we fix the other variables to solve it. The specific solution process for each subproblem is as follows:

Subproblem 1: the UE and BS Association

For the UE and BS association in mmWave massive MIMO UDNs, except the variable x_{ij}, we fix the other variables. Thus, the Lagrangian function (40) can be written as:

$$\begin{aligned} L_1(x_{ij}) = & \sum_i \sum_j x_{ij} \log(W_j \log_2(1 + SINR_{ij})) \\ & + \sum_j \tau_j (K_U^j - \sum_i x_{ij}) + \sum_i \nu_i (M - \sum_j x_{ij}) \\ & + \sum_j \gamma_j (N_{AA}^j - \sum_i x_{ij}) + \sum_i \alpha_i (R_i - R_{QoS}). \end{aligned} \tag{43}$$

Then, the first-order partial derivative of the function (43) can be written as:

$$\begin{aligned} \frac{\partial L_1(x_{ij})}{\partial x_{ij}} = & \log(W_j \log_2(1 + SINR_{ij})) - \tau_j - \nu_i - \gamma_j \\ & + \alpha_i (W_j \log_2(1 + SINR_{ij})). \end{aligned} \tag{44}$$

In order to achieve the maximum of the problem (36), the maximizer x_{ij} of the subproblem 1 is defined as [12].

$$x_{ij} = \begin{cases} 1, & \text{if } j = j^*, \\ 0, & \text{otherwise } j \ne j^*, \end{cases} \tag{45}$$

where the expression of j^* is given by (46):

$$\begin{aligned} j^* = \arg\max_j \Big(& \log(W_j \log_2(1 + SINR_{ij})) - \tau_j - \nu_i - \gamma_j \\ & + \alpha_i (W_j \log_2(1 + SINR_{ij})) \Big). \end{aligned} \tag{46}$$

Note that, at the k-th inner iteration, the equation (46) can be considered to be a judgment criterion for UEs to determine the best servicing BS.

We use subgradient method to update the Lagrange multipliers which can be expressed as:

$$\tau_j(t+1) = \big(\tau_j(t) - \delta_1(t)(K_U^j - \textstyle\sum_i x_{ij}(t))\big)^+, \tag{47}$$
$$\nu_i(t+1) = \big(\nu_i(t) - \delta_2(t)(M - \textstyle\sum_j x_{ij}(t))\big)^+, \tag{48}$$
$$\gamma_j(t+1) = \big(\gamma_j(t) - \delta_3(t)(N_{AA}^j(t) - \textstyle\sum_i x_{ij}(t))\big)^+, \tag{49}$$
$$\alpha_i(t+1 = \big(\alpha_i(t) - \delta_4(t)(R_i(t) - R_{QoS})\big)^+, \tag{50}$$

where $[a]^+ = \max\{0, a\}$, and $\delta_1(t)$, $\delta_2(t)$, $\delta_3(t)$, $\delta_4(t)$ are step sizes. By updating the Lagrange multipliers $\tau_j(t)$, $\nu_i(t)$, $\gamma_j(t)$, and $\alpha_i(t)$ via (47)-(50), the dual problem will achieve the global optimum when the multipliers converge. An UE and

BS association algorithm based CoMP and MaxEE is proposed and summarized in **Algorithm 1**.

Algorithm 1: UE and BS Association Algorithm Based CoMP and MaxEE

1.Initialization: $N_{AA}^{j}=N_t$, $P_{TX}^{j}=P_{TX}^{max}$, $b_j=1$, $j_{jn}=1$, $S_{ijn}=1$; for $\forall j$;
2. Cooperative user association
3. **for** $i=1$ to U **do**
4. **for** $j=1$ to B **do**
5. Calculate the power consumption $P_{sum}(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX})$ according to (30);
6. Calculate $\frac{\partial L_1(x_{ij})}{\partial x_{ij}}$ according to (44);
7. **end for**
8. **end for**
9. **for** $i=1$ to U **do**
10. **for** $k=1$ to M **do**
11. Calculate j^* according to (46);
12. Use the j^* to update x_{ij} according to (45);
13. **for** $j=1$ to B **do**
14. Calculate K_j, according to (17);
15. Judge whether K_j exceeds the maximum value K_U^j;
16. **if** $K_j \geq K_U^j$ **do then**
17. $\frac{\partial L_1(x_{ij})}{\partial x_{i+1,j}}=0$;
18. **end if**
19. Set $\frac{\partial L_1(x_{ij})}{\partial x_{ij^*}}=0$, update $\frac{\partial L_1(x_{ij})}{\partial x_{ij}}$;
20. Update τ_j, ν_i, γ_j, and α_i according to (47)-(50).
21. **end for**
22. **end for**
23. **end for**

Subproblem 2: The Optimization of Number of Activated Antenna

In this part, we investigate the optimization problem regarding the number of the activated antenna with massive MIMO BS. We adopt the Newton-Raphson method to solve the subproblem. Since the number of active antennas is an integer, the optimization subproblem of the number of activated antennas is an integer programming problem. To tackle this problem, we first relax the number of activated antennas N_{AA}^{j} and transfer them into real variables.

Rewriting (40), the optimal function of activated antenna number is obtained as:

$$\begin{aligned}
&L_2(N_{AA}^{j})\\
&=\sum\nolimits_i\sum\nolimits_j x_{ij}\log(W_j\log_2(1+SINR_{ij}))\\
&\quad -q(\frac{\sum_j b_j P_{TX}^{j}}{\mu}+C_0+C_1\sum\nolimits_j b_j+C_2\sum\nolimits_j b_j N_{AA}^{j})\\
&\quad +q\eta\sum\nolimits_j b_j(\sum\nolimits_k b_k\sum\nolimits_n j_{jn}\sum\nolimits_m j_{km}\frac{P_{TX}^{k}}{N_{AA}^{k}}|G_{k,j}^{m,n}|^2)\\
&\quad +\sum\nolimits_j\gamma_j(N_{AA}^{j}-\sum\nolimits_i x_{ij})+\sum\nolimits_j\rho_j(N_t-N_{AA}^{j}). \quad (51)
\end{aligned}$$

Furthermore, for any j-th BS, (51) can be rewritten as:

$$\begin{aligned}
L_{21}(N_{AA}^{j}) &= \sum_i x_{ij}\log(W_j\log_2(1+SINR_{ij}))\\
&\quad -q(\frac{P_{TX}^{j}}{\mu}+\frac{C_0}{B}+C_1+C_2N_{AA}^{j})\\
&\quad +q\eta(\sum_k b_k\sum_n j_{jn}\sum_m j_{km}\frac{P_{TX}^{k}}{N_{AA}^{k}}|G_{k,j}^{m,n}|^2)\\
&\quad +\gamma_j(N_{AA}^{j}-\sum_i x_{ij})+\rho_j(N_t-N_{AA}^{j}). \quad (52)
\end{aligned}$$

Assuming the i-th UE associates with the j-th BS, the variable x_{ij} can be regarded as a constant with unit value. Then, the first-order partial derivative and the second-order derivative of the function of activated antenna number (52) can be expressed by (53) and (54):

$$\begin{aligned}
&\frac{\partial L_{21}(N_{AA}^{j})}{\partial N_{AA}^{j}}\\
&\quad=\sum\nolimits_i x_{ij}\frac{1}{\log_2(1+SINR_{ij})}\frac{1}{(1+SINR_{ij})\ln 2}\\
&\quad\times\frac{\partial SINR_{ij}}{\partial N_{AA}^{j}}-qC_2+\gamma_j-\rho_j, \quad (53)\\
&\frac{\partial^2 f(x)}{\partial^2 N_{AA}^{j}}=\frac{\partial g(x)}{\partial N_{AA}^{j}}\\
&\quad=\sum\nolimits_i x_{ij}(-1)(\log_2(1+SINR_{ij}))^{-2}\\
&\quad\times((1+SINR_{ij})\ln 2)^{-1}\\
&\quad\times\frac{\partial SINR_{ij}}{\partial N_{AA}^{j}}((1+SINR_{ij})\ln 2)^{-1}\frac{\partial SINR_{ij}}{\partial N_{AA}^{j}}\\
&\quad+\sum\nolimits_i x_{ij}(\log_2(1+SINR_{ij}))^{-1}\\
&\quad(-1)(\ln 2)\frac{\partial SINR_{ij}}{\partial N_{AA}^{j}}\\
&\quad\times((1+SINR_{ij})\ln 2)^{-2}\frac{\partial SINR_{ij}}{\partial N_{AA}^{j}}\\
&\quad+\sum\nolimits_i x_{ij}(\log_2(1+SINR_{ij}))^{-1}\\
&\quad\times((1+SINR_{ij})\ln 2)^{-1}\times\frac{\partial^2 SINR_{ij}}{\partial^2 N_{AA}^{j}}, \quad (54)
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial SINR_{ij}}{\partial N_{AA}^{j}} &= SINR_{ij}((N_{AA}^{j}-K_j+1)^{-1}\\
&\quad -(N_{AA}^{j})^{-1}), \quad (55)\\
\frac{\partial^2 SINR_{ij}}{\partial^2 N_{AA}^{j}} &= \frac{\partial SINR_{ij}}{\partial N_{AA}^{j}}\\
&\quad\times\left((N_{AA}^{j}-K_j+1)^{-1}-(N_{AA}^{j})^{-1}\right)\\
&\quad -SINR_{ij}\Big((N_{AA}^{j}-K_j+1)^{-2}\\
&\quad -(N_{AA}^{j})^{-2}\Big), \quad (56)
\end{aligned}$$

Due to the computational complexity of introducing the Hessian matrix, we focus only on the two-dimensional case

in this paper, where

$$\Delta N_{AA}^{j} = \left|\frac{\partial f(x)}{\partial N_{AA}^{j}}\right| / \left|\frac{\partial^2 f(x)}{\partial^2 N_{AA}^{j}}\right|. \quad (57)$$

Hence, the updating formula for the number of activated antennas can be written as:

$$N_{AA}^{j}(t+1) = N_{AA}^{j}(t) + \delta(t)\Delta N_{AA}^{j}, \quad (58)$$

$$\rho_j(t+1) = \left(\rho_j(t) - \delta_5(t)(N_t - N_{AA}^{j}(t))\right)^{+}, \quad (59)$$

where $\delta_5(t)$ is a step size.

Since the optimal number of activated antennas $N_{AA}^{j^*}$ obtained by the Newton-Raphson method is a real number, the actual optimal number of activated antennas can only be an integer. Therefore, the actual optimum number of activated antennas becomes $\widehat{N_{AA}^{j^*}} = \lceil N_{AA}^{j^*} \rceil$, where $\lceil . \rceil$ is the ceiling operator. We now propose an optimization algorithm for the number of activated antenna based Newton Raphson method. The algorithm is summarized in **Algorithm 2**.

Algorithm 2: Optimization Algorithm of the Number of Activated Antenna Based Newton Raphson Method

1.Initialization: $P_{TX}^{j} = P_{TX}^{max}$,$S_{ijn} = 1$, $j_{jn} = 1$; for $\forall j$; x_{ij}, b_j, and K_j obtained via Algorithm 1;
2. **for** j=1 to B **do**
3. Calculate $\frac{\partial f(x)}{\partial N_{AA}^{j}}$ and $\frac{\partial^2 f(x)}{\partial^2 N_{AA}^{j}}$, according to the formula (53) and (54);
4. Update ΔN_{AA}^{j}, according to the formula (57);
5. Update N_{AA}^{j}, according to the formula (58);
6. **if** $N_{AA}^{j} < K_j$ **do then**
7. $N_{AA}^{j} = K_j$;
8. **end if**
9. Update ρ_j, according to (59).
10.**end for**

Subproblem 3: the UE and Antenna Association

In this section, we investigate the UE and antenna association subproblem of massive MIMO BS. We employ the antenna selection algorithm based on the maximum Frobenius-norm (F-norm) since it has the advantages of low computational complexity with a simple calculation. The basic idea of the algorithm is as follows: firstly, we calculate the norm of all N_t column vectors of the channel matrix **H** as follows:

$$F_n = \sum_{i=1}^{K_j} |h_{ijn}|^2, \quad n = 1, 2, \cdots, N_t. \quad (60)$$

Then, we adopt the quick sort algorithm to sort F_n's in an ascending order, to obtain the new sequence $\{F_n^1 >, \cdots, > F_n^{N_t}\}$. Finally, we take the corresponding antenna of the first N_{AA}^{j} sequence numbers of the sequence $\{n_1, \cdots, n_{N_{AA}^{j}}\}$ as the transmit antenna set associated with the j-th BS.

For the UE and antenna association subproblem, we first fix the other variables except for the variable $\{s_{ijn}\}$. Then, the Lagrange function (40) is rewritten in the following form:

$$L_3(s_{ijn}) = \sum\nolimits_i \sum\nolimits_j x_{ij} \log(W_j \log_2(1 + SINR_{ij})) + \sum\nolimits_i \alpha_i (\sum\nolimits_j x_{ij} W_j \log_2 \times (1 + SINR_{ij}) - R_{QoS}). \quad (61)$$

The partial derivatives of equation (61) can be written as:

$$\frac{\partial L_3(s_{ijn})}{\partial s_{ijn}} = x_{ij} \frac{1}{\log_2(1+SINR_{ij})} \frac{1}{(1+SINR_{ij})\ln 2} \frac{A_0}{B_0} \frac{P_{TX}^{j}}{N_{AA}^{j}} G_{ijn} + \alpha_i x_{ij} W_j \frac{1}{(1+SINR_{ij})\ln 2} \frac{A_0}{B_0} \frac{P_{TX}^{j}}{N_{AA}^{j}} G_{ijn}, \quad (62)$$

where

$$A_0 = \frac{N_{AA}^{j} - K_j + 1}{K_j}, B_0 = \sum_k \sum_m s_{ikm} \frac{P_{TX}^{k}}{N_{AA}^{k}} G_{ikm} + \sigma_i^2. \quad (63)$$

The optimal solution of the subproblem 3, in (36), for $\{s_{ijn}\}$ can be obtained as follows:

$$s_{ijn} = \begin{cases} 1, & \text{if } x_{ij} = 1 \text{ and } n = n^*, \\ 0, & \text{otherwise}, \end{cases} \quad (64)$$

where the expression of n^* is given (65):

$$n^* = \arg\max_n \times \Big(x_{ij} \frac{1}{\log_2(1+SINR_{ij})} \frac{1}{(1+SINR_{ij})\ln 2} \frac{A_0}{B_0} \frac{P_{TX}^{j}}{N_{AA}^{j}} \times G_{ijn} + \alpha_i x_{ij} W_j \frac{1}{(1+SINR_{ij})\ln 2} \frac{A_0}{B_0} \frac{P_{TX}^{j}}{N_{AA}^{j}} G_{ijn} \Big). \quad (65)$$

The proposed algorithm for the UE and antenna association based F-norm and MaxEE is provided in **Algorithm 3**.

Algorithm 3: User and Antenna Association Algorithm Based F-norm and MaxEE

1.Initialization: $P_{TX}^{j} = P_{TX}^{max}$, $j_{jn} = 1$; for $\forall j$; x_{ij}, N_{AA}^{j}, b_j, and K_j obtained via Algorithm 1 and 2;
2. **for** j=1 to B **do**
3. Select out N_{AA}^{j} antennas from the antenna set $\mathcal{C}_A^{j}$ of the j-th BS according to (60);
4. **repeat**
5. **for** $i = 1$ to K_j **do**
6. Calculate n^* according to (65);
7. Use the n^* to update s_{ijn} according to (64);
8. Update the antenna set $\mathcal{C}_A^{j}$ of the j-th BS;
9. **end for**
10.**until** $\mathcal{C}_A^{j} = \phi$(null set).
11.**end for**

After the UE and antenna association $\{s_{ijn}\}$ is completed, we determine the antenna switching off/on indication $\{j_{in}\}$ of BSs according to the formula (8).

Subproblem 4: the Transmit Power Allocation

In this section, we investigate the transmission power allocation subproblem of massive MIMO BS by adopting the

conjugate gradient method, which is a common algorithm for solving extremal problems of continuous functions. It has the advantages of being a simple algorithm and having small storage requirement, as well as solving large-scale optimization problems. Its iterative formula is given as:

$$x_{k+1} = x_k + t_k d_k, \tag{66}$$

where t_k is determine by a line search, d_k is search direction. The formula for d_k is given as:

$$d_k = \begin{cases} -g_k, & k = 1, \\ -g_k + \beta_k d_{k-1}, & k \geq 2, \end{cases} \tag{67}$$

where β_k is a parameter. $\nabla f(x_k)$ denotes the gradient of the function $f(x)$ at x_k, abbreviated as g_k. Different β_k's correspond to different conjugate gradient methods. The widely employed ones are β_k^{HS}, β_k^{FR}, and β_k^{PRP}. We use β_k^{PRP} here and hence the formula of β_k is given as:

$$\beta_k = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}. \tag{68}$$

To determine the step t_k, we employ the Wolfe-Powell criterion as follows: for given parameters $0 < \mu < \sigma < 1$, $t_k > 0$ is selected to satisfy the following equations:

$$\begin{cases} f(x_k + t_k d_k) \leq f(x_k) + \mu t_k g_k^T d_k, \\ \sigma g_k^T d_k \leq g(x_k + t_k d_k)^T d_k. \end{cases} \tag{69}$$

By rewriting the Lagrange function (40), the transmission power allocation function of BS can be obtained as:

$$\begin{aligned} L_4(P_{TX}^j) &= \sum\nolimits_j \sum\nolimits_i x_{ij} \log(W_j \log_2(1 + SINR_{ij})) \\ &\quad -q(\frac{\sum_j b_j P_{TX}^j}{\mu} + C_0 + C_1 \sum\nolimits_j b_j \\ &\quad +C_2 \sum\nolimits_j b_j N_{AA}^j) \\ &\quad + \sum\nolimits_j \omega_j (P_{TX}^{max} - P_{TX}^j) \\ &\quad + \sum\nolimits_i \alpha_i (R_i - R_{QoS}). \end{aligned} \tag{70}$$

For any j-th BS, the formula (70) can be expressed as:

$$\begin{aligned} L_{41}(P_{TX}^j) &= \sum\nolimits_i x_{ij} \log(W_j \log_2(1 + SINR_{ij})) \\ &\quad -q(\frac{b_j P_{TX}^j}{\mu} + \frac{C_0}{B} + C_1 b_j + C_2 b_j N_{AA}^j) \\ &\quad +\omega_j (P_{TX}^{max} - P_{TX}^j) \\ &\quad + \sum\nolimits_i \alpha_i (x_{ij} W_j \log_2(1 + SINR_{ij}) \\ &\quad -\frac{R_{QoS}}{B}). \end{aligned} \tag{71}$$

Algorithm 4: BS Transmission Power Allocation Algorithm Based PRP Conjugate Gradient Method

1.Initialization: Initialzation: x_{ij}, N_{AA}^j, b_j,s_{ijn}, j_{in}, K_j obtained via Algorithm 1, 2, and 3;
2. **for** $j = 1$ to B **do**
3. Initialzation: $P_{TX,1}^j \in R$, $\epsilon > 0$, I_{max} ;
4. Set $d_1 = -g$, $k = 1$;
5. **if** $\|g_1\| \leq \epsilon$ **do then**
6. Break;
7. **else**
8. Return 10;
9. **end if**
10. Calculate the step t_k according to (69);
11. Set $P_{TX,k+1}^j = P_{TX,k}^j + t_k d_k$, $g_{k+1} = g(x_{k+1})$;
12. **if** $\|g_{k+1}\| \leq \epsilon$ **do then**
13. break;
14. **else**
15. Return 17;
16. **end if**
17. Calculate the parameter β_k according to (68);
18. Calculate the search direction d_{k+1} according to (67);
19. Update ω_j according to (76);
20. Set $k = k + 1$, return 10;
21. **Until** $\|g_{k+1}\| \leq \epsilon$ or $k + 1 = I_{max}$;
22. **if** $P_{TX}^j < \widehat{P}_{TX}^j$, **do then**
23. $P_{TX}^j = \widehat{P}_{TX}^j$.
24. **end if**
25. **end for**

Taking the partial derivative of equation (71), it follows that:

$$\begin{aligned} &\frac{\partial L_{41}(P_{TX}^j)}{\partial P_{TX}^j} \\ &= \sum\nolimits_i x_{ij} \frac{1}{\log_2(1 + SINR_{ij})} \\ &\quad \frac{1}{(1 + SINR_{ij}) \ln 2} \frac{A_0}{B_0} \\ &\quad \times \sum\nolimits_n s_{ijn} \frac{1}{N_{AA}^j} G_{ijn} - \frac{q b_j}{\mu} - \omega_j \\ &\quad + \sum\nolimits_i \alpha_i x_{ij} W_j \frac{1}{(1 + SINR_{ij}) \ln 2} \frac{A_0}{B_0} \\ &\quad \sum\nolimits_n s_{ijn} \frac{1}{N_{AA}^j} G_{ijn}. \end{aligned} \tag{72}$$

Then, $f(x)$ and $g(x)$ can be expressed by the expressions:

$$f(x) = L_{41}(P_{TX}^j), \tag{73}$$

$$g(x) = \frac{\partial L_{41}(P_{TX}^j)}{\partial P_{TX}^j}. \tag{74}$$

The variable $P_{TX,k}^j$ is then updated as:

$$P_{TX,k+1}^j = P_{TX,k}^j + t_k d_k. \tag{75}$$

We use the subgradient method to update the Lagrange multiplier and the updating formula is as follows:

$$\omega_j(t+1) = \left(\omega_j(t) - \delta_6(t)(P_{TX}^{max} - P_{TX}^{j}(t))\right)^{+}, \quad (76)$$

where $\delta_6(t)$ is a step size.

From the QoS constraint (33C5), we can obtain the minimum transmit power as follows:

$$\widehat{P}_{TX}^{j} > (2^{\frac{R_{QoS}}{\sum_j x_{ij} W_{ij}}} - 1)\frac{B_0}{A_0}\frac{1}{\sum_n s_{ijn} G_{ijn} (N_{AA}^{j})^{-1}}. \quad (77)$$

If $P_{TX}^{j} < \widehat{P}_{TX}^{j}$, then $P_{TX}^{j} = \widehat{P}_{TX}^{j}$. We now propose a BS transmission power allocation algorithm based PRP conjugate gradient method, summarized in **Algorithm 4**.

Subproblem 5: Iterative Algorithm for EE Maximization

In this section, an iterative algorithm (known as the Dinkelbach method [33]) is proposed for obtaining the solution of the problem (36). The proposed algorithm is summarized in **algorithm 5**. Its convergence to optimal EE is guaranteed.

Proof: The proof of convergence of algorithm 5 see Appendix A.

After algorithm 5 achieves convergence or reaches to a given maximum iteration numbers I_{max}, the BS is put in sleep and switch off the redundant antennas to improve the EE of system according to the b_j and j_{in}.

Algorithm 5: Iterative EE Algorithm Based Dinkelbach Method

1. Initialize the maximum number of iterations I_{max} and the maximum tolerance Δ;
2. Set maximum energy efficiency $q = 0$ and iteration index $n = 0$.
3. **repeat** {Main Loop}
4. Calculate $\{X^n\}$, $\{S^n\}$, $\{N_{AA}^n\}$, $\{P_{TX}^n\}$ via Algorithm1, 2, 3 and 4;
5. **if** $U(\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n)$ $-qP_{sum}(\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n) < \Delta$, **then**
6. Convergence ==**true**;
7. Return $\{\mathbf{X}^*, \mathbf{N}_{AA}^*, \mathbf{S}^*, \mathbf{P}_{TX}^*\}$ $= \{\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n\}$, and $q^* = \frac{U(\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n)}{P_{sum}(\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n)}$;
8. **else**
9. Set $q = \frac{U(\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n)}{P_{sum}(\mathbf{X}^n, \mathbf{N}_{AA}^n, \mathbf{S}^n, \mathbf{P}_{TX}^n)}$, and $n = n + 1$;
10. Convergence ==**false**.
11. **end if**
12. **Until** Convergence == **true** or $n = I_{max}$

C. Duality Gap Bound

Since the variable x_{ij} and s_{ijn} is naturally discrete, there may exist a non-zero duality gap between the primal problem (36) and the dual problem (42). Nevertheless, the Lagrangian dual method often provides good solutions of the primal optimization problem [48], and the following proposition further illustrates that the duality gap is bounded.

Proposition : For the UE and BS association, the optimization of number of activated antennas, the UE and antenna association, and the transmit power allocation problem (36), the gap between the objective function $L(\mathbf{X}, \mathbf{N}_{AA}, \mathbf{S}, \mathbf{P}_{TX})$ obtained via the subgradient method and the global optimization problem in (36) is bounded by $\sum_j \tau_j(K_U^j - K_j) + \sum_i \nu_i(M - U_i) + \sum_j \omega_j(P_{TX}^{max} - P_{TX}^{*j}) + \sum_i \alpha_i(R_i^* - R_{QoS}) + \sum_j \gamma_j(N_{AA}^{*j} - K_j) + \sum_j \rho_j(N_t - N_{AA}^{*j})$, where $(\mathbf{X}^*, \mathbf{N}_{AA}^*, \mathbf{S}^*, \mathbf{P}_{TX}^*)$ represent the primal solution recovered from the dual variable $(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho})$ using (46), Algorithm 2, (65), and Algorithm 4. U_i represents the number of access BSs for the i-th UE, namely, $U_i = \sum_j x_{ij}$ and $R_i^* = \max_j W_j \log_2(1 + SINR_{ij}^*)$, where $SINR_{ij}^* = \frac{N_{AA}^j - K_j + 1}{K_j} \frac{\max_n \frac{P_{TX}^{*j}}{N_{AA}^{*j}} G_{ij}}{\sum_{k \in \mathcal{C}, k \neq j} \max_m \frac{P_{TX}^{*k}}{N_{AA}^{*k}} G_{ik} + \sigma_i^2}$.

Proof: See Appendix B.

D. Computational Complexity Analysis

In this subsection, we analyze the computational complexity of the proposed algorithm, where o(B) operations represent the time complexity of the algorithm. In algorithm 1, computing the partial derivative of x_{ij} with respect to $L_1(x_{ij})$ for all UEs requires $o(B \times U)$ operations, and $o(B \times M \times U)$ operations are required to determine the cooperative BS for all UEs. Therefore, Algorithm 1 needs $o(B \times U + B \times U \times M)$ operations to establish the associated relationship between UEs and BSs in each iteration [12]. Algorithm 2 needs $o(B)$ operations to obtain the optimal solution of the number of the activated antenna at each iteration. Algorithm 3 needs $o(B \times \max(K_U^j))$ operations to establish the associated relationship between UEs and antennas at each iteration. In Algorithm 4, we assume that the optimal step of the problem (71) requires I_1 iterations to converge, then Algorithm 4 needs $o(B \times I_1)$ operations to obtain the transmission power of each BS. In Algorithm 5, we assume that the optimal solution of problem (36) requires I_2 iterations to converge. Each of the Lagrange multipliers τ_j, γ_j, ρ_j, ω_j requires $o(B)$ operations, ν_i, α_i require $o(U)$ operations. So the complexity of Algorithm 5 is $o(I_2 \times (B \times U + B \times U \times M + B \times \max(K_U^j) + B \times I_1 + 5B + 2U))$.

IV. Simulation Results

We now present numerical results to demonstrate the effectiveness of the proposed algorithm compared with other schemes as well as the conventional counterparts, in terms of the EE and energy harvesting. In computer simulations, a two-tier UDN is considered, where one massive MIMO macro BS is located in the center, and 50 massive MIMO pico BSs are randomly distributed over the macro cell. The UEs are also randomly distributed in one macrocell. For the path loss, the IEEE802.15.3c logarithmic distance path loss model is adopted [31], for which $PL(d) = 20\log_{10}(\frac{4\pi d_0}{\lambda}) + 10\eta_{pl}\log_{10}(\frac{d}{d_0}) + \sigma^2$, where $d_0 = 1m$ shows the reference distance, $\eta_{pl} = 6$ is the path loss index, and the standard derivation of shadow fading σ^2 is 18 dB. The other simulation parameters are shown in Table I.

Definition 1: Utilization rate of BS: We define the ratio of the product of the number of UEs (NUE) and the number of cooperative BSs (M_C) to the maximum number of services,

TABLE I
SIMULATION PARAMETERS SETTING

Parameters	Default value	Parameters	Default value
W	20MHz	P_{TX}^{max}	Macro:46dbm; pico:30dbm[32]
N_0	-174dBm/Hz	N_t	Macro:128; Pico:64
P_{AB}	Macro:65.8W; Pico:1.5W	P_{SB}	Macro:19.7w; Pico0.5W
P_{BB}	7.54W	N_{RF}	Macro:12; Pico:6
P_{PS}	40mW	P_{RF}	300mW
η	0.8	I_j	7.5×10^{-13}W
R_{QoS}	20Mbps	μ	Macro:0.388; Pico:0.08
f_c	60GHz	ϕ_l^t and ϕ_l^r	Uniformly distributed within $[0, 2\pi]$
α_l	$\alpha_l \in \mathcal{CN}(0,1)$[2]	$\Lambda_t(\phi_l^t)$ and $\Lambda_r(\phi_l^r)$	1
K_U	Macro:12; Pico:6		

UEs ($MSUE$), as the utilization rate of BS ($URBS$) that can be expressed as:

$$URBS = \frac{NUE \times M_C}{MSUE}. \tag{78}$$

A. Convergence of the Algorithm

We now investigate the convergence of the proposed iterative algorithm. In Fig. 2, EE vs. number of iterations are shown where a macro and 50 picos BSs serve 20 UEs, and the UEs adopt 3/5/6/7 BSs with CoMP as well as NCoMP. As can be seen from Fig. 2, the proposed algorithm converges within 15 iterations to the optimal value, regardless of whether UEs adopt the 3/5/6/7 BSs with CoMP or NCoMP. In other words, the MaxEE of the system can be achieved within a few iterations on average with a superlinear convergence rate [33]. It is concluded that fast convergence of the algorithm to the optimal value depends on the settings of simulation parameters. Namely, the trend of the curve's behavior is determined by the initial value of the simulation parameters as well as the proposed algorithm. Initially, the EE q is set to 0, the transmission power of BS is set to the maximum value, all BSs in the system are turned on and all antennas are activated. Since the simulations carried out in this paper employ the same initial value, the behavior of the curves in Fig.2, Fig.3, Fig.4, and Fig.5 have similar trends.That is, they all start from zero and increase monotonically to a larger value, then drop rapidly, and finally increase again slowly until they stabilize by reaching a steady-state value. As mentioned above, the main reason that the curves initially begin ascending from zero to a larger certain value is due to how the algorithm is initialized. In the beginning, the transmission power of BSs, the number of active antennas, and other parameters are all set to a predetermined maximum value except for the EE q. On the other hand, the rapid decline of the curves is due to the fact that some antennas of the BSs are deactivated during the operation of the algorithm, resulting in a decrease of the transmission power of the BSs, and the change in the relationship between the UEs and the BSs. Finally, the main reason for the curves slowly increasing again until they stabilize is due to the combination of the UE and BS association, the UE and antenna association, the optimization of the number of active antennas and the transmission power allocation.

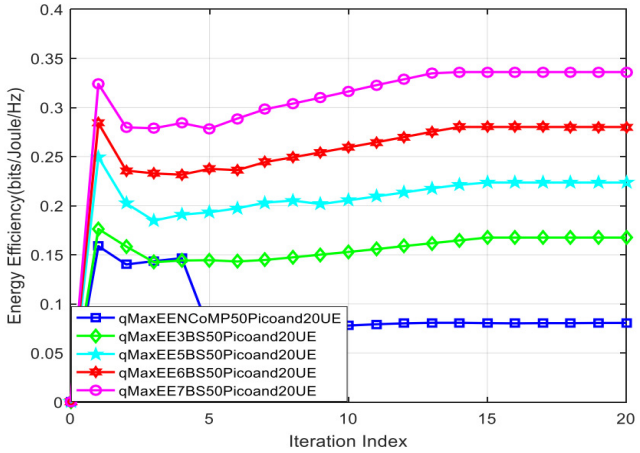


Fig. 2. The EE with cooperative BSs of different number.

B. Performance Comparisons

In Fig. 3, the EE performance of the proposed algorith and the maximum signal to interference plus noise ratio(MaxSINR) algorithm are compared for the scenarios where each EU can select six BSs to serve for themselves, macro BS can serve 12 UEs, and pico BSs can serve 6 UEs, at the same time. Fig. 3 shows that the EE of the proposed algorithm is higher than that of the MaxSINR algorithm since the proposed algorithm adopts the MaxEE association between UEs and BSs. In this article, the purpose of our investigation is to improve the EE of the system. Our optimization goal is to maximize EE, so the algorithm based on the MaxSINR cannot guarantee the maximum EE. This is the main reason why our proposed algorithm is superior to the MaxSINR algorithm in terms of performance. The reason for the curve behavior change has been explained in detail in the algorithm convergence part. In terms of convergence, it can be also observed from Fig. 3 that the convergence speed of the proposed algorithm is faster.

C. EE With Different Number of UEs and Cooperative BSs

In this part, we investigate the EE of the system with a different number of cooperative BSs and UEs. Fig. 2 shows that with the increase in the number of cooperative BSs, the EE of the system also increases gradually. When UEs cooperate with 7 BSs, the EE of the system reaches its highest

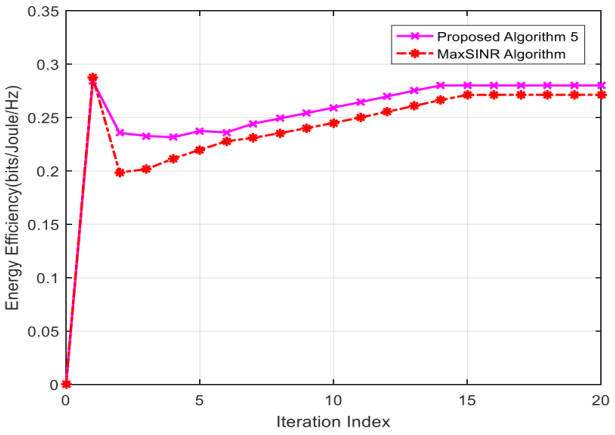


Fig. 3. The comparisons of EE with proposed algorithm vs. MaxSINR algorithm.

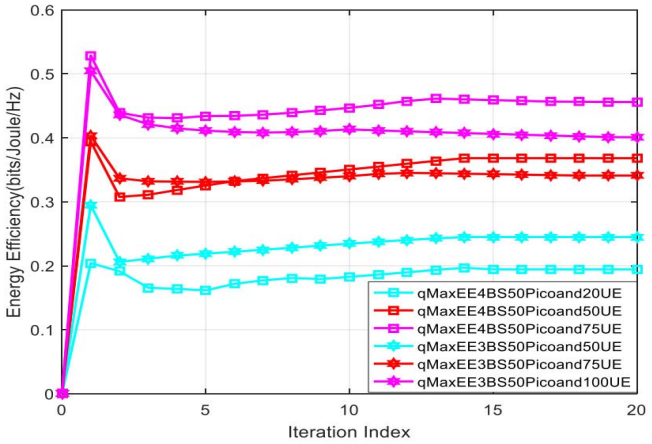


Fig. 4. The EE with UEs of different number.

value. Compared to the NCoMP, the system with 7 BSs cooperation can achieve more than triple in terms of EE gain, emphasizing that it is effective and necessary to adopt CoMP in scenarios where mmWave massive MIMO UDN is adopted. From Table II, we can see that the URBS gradually increases with the number of cooperative BSs selected by UEs when the number of picos and UEs are fixed. Table II and Fig. 2 further indicate that the EE of the system gradually increases with the URBS.

We next consider the scenarios where a macro and 50 picos BSs serve 20/50/75/100 UEs, and the UEs adopt two cooperative modes: 3 BSs cooperation(3BSsCoMP) and 4 BSs cooperation(4BSsCoMP). Fig. 4 shows that with the increase of the number of UEs in the macrocell, the EE of the system increases gradually, regardless of whether the UEs adopt 3BSsCoMP or 4BSsCoMP. As can be seen also from Table III and Fig. 4, the URBS and the EE of system increase gradually with the increase of the number of UEs, when the number of cooperative BSs is the same. For the system configuration 4BSsCoMP, except the cases where the number of service UEs is equal to 20, the EE of 4BSsCoMP is higher than that of 3BSsCoMP, since cooperation with more BSs can assist to improve EE. In the case of 4BSsCoMP to serve 20 EUs, EE is lower than that of 3BSsCoMP to serve 50 EUs because the URBS is too low. As can be seen from Table III, when 4BSsCoMP is to serve 20 UEs, the URBS is only 12.82 %, and when 3BSsCoMP is to serve 50 UEs, the URBS is 48.08%. Hence, the URBS of the latter is about 4 times greater than the former. On the other hand, as an extreme example, when the number of UEs served by a BS is zero, no matter how many BSs cooperate, the EE of the system also becomes zero, showing that the EE of the system increases with the increase of URBS. In the case of having the same URBS, 4BSsCoMP serves 75 UEs with higher EE than 3BSsCoMP serving 100 UEs. This indicates that when the URBS is the same, the cooperation with more BSs can further improve the EE of the system.

In Fig. 5, the simulation results are presented where one macro and 50 picos BSs serving 50/75/100 UEs, and the UEs adopt three cooperative modes: 3BSsCoMP, 4BssCoMP, and 6BSsCoMP. As can be seen from Table IV, the URBS

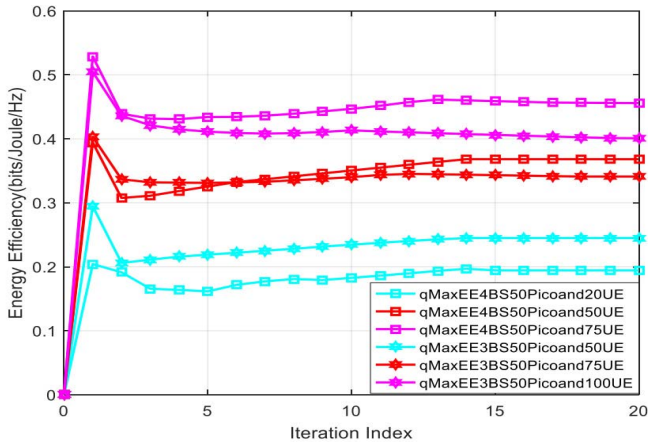


Fig. 5. The EE with XBSsCoMP servicing XUEs.

TABLE II
THE URBS WITH COOPERATIVE BS OF DIFFERENT NUMBER

	NCoMP	3BSsCoMP	5BSsCoMP	6 BSsCoMP	7BSsCoMP
URBS	6.41%	19.23%	32.05%	38.46%	44.87%

with three different UE numbers and three cooperative modes are the same. Fig. 5 shows that one macro and 50 picos BSs serve 50 UEs, and that the UEs with 6BSsCoMP can achieve higher EE than the other two settings (one macro and 50 picos serve 75 UEs and the UEs adopt 4BSsCoMP as well as one macro and 50 picos serve 100 UEs and the UEs adopt 3BSsCoMP). Fig. 5 and Fig. 2 further show that the EE of the system increases gradually with the increase of the number of cooperative BSs when the URBS is the same. Fig. 5 and Fig. 2 also show that CoMP is effective in the scenario of mmWave massive MIMO UDN. The reason for the curve behavior change has been explained in detail in the algorithm convergence part.

D. Performance of Energy Harvesting

In this section, we investigate the role of EHS in improving EE. In Fig. 6, simulations are carried out for one macro and 50 pico BSs serve 20/50/75 UEs, and the UEs adopt two cooperative modes: 4BSsCoMP and 6BSsCoMP. It can be seen from Fig. 6 that better EE of the proposed algorithm

TABLE III
THE URBS WITH UES OF DIFFERENT NUMBER

	20UE &4BSsCoMP	50UE &4BSsCoMP	75UE &4BSsCoMP	50UE &3BSsCoMP	75UE &3BSsCoMP	100UE &3BSsCoMP
URBS	12.82%	64.1%	96.15%	48.08%	72.12%	96.15%

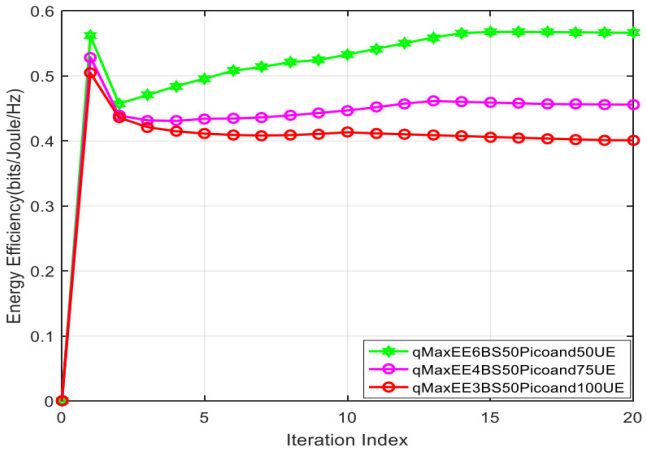


Fig. 6. The EE of XBSsCoMP servicing XUEs for proposed algorithm vs. no energy harvesting algorithm.

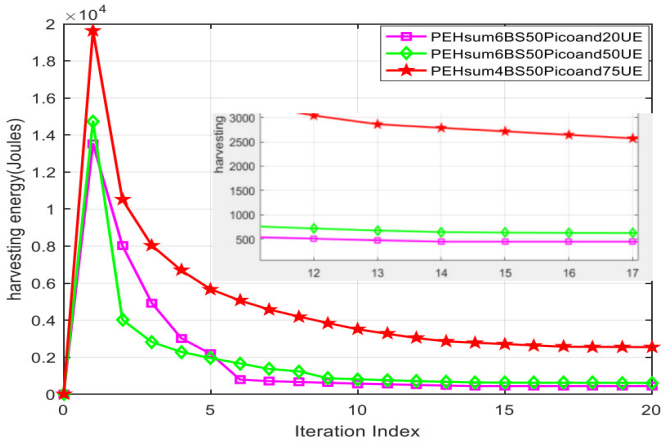


Fig. 7. The harvesting Energy for XBSsCoMP servicing XUEs.

TABLE IV
THE URBS WITH XBSCOMP SERVICING XUES

	50UE&6BSCoMP	75UE&4BSCoMP	100UE&3BSCoMP
URBS	96.15%	96.15%	96.15%

TABLE V
THE GAIN OF EE FOR PROPOSED ALGORITHM AND NO ENERGY HARVESTING ALGORITHM

	20UE&6BSsCoMP	50UE&6BSsCoMP	75UE&4BSsCoMP
GEE	1.0518	1.0547	1.1443

is achieved than that of the algorithm without adopting the EHS. The achieved URBSs of the three cases, namely, 20UE & 6BSsCoMP, 75UE & 4BSsCoMP, and 50UE & 6BSsCoMP are 38.46 %, 96.15 %, and 96.15 % respectively. The EE of the system 6BSsCoMP serving 20 UEs is lower than that of the 4BSsCoMP with 75 UEs, because the former URBS is too low. For the two situations where the URBS is the same, the EE of 50UE & 6BSsCoMP is higher than that of 75UE & 4BSsCoMP because the former has more cooperative BSs. On the other hand, the three curves Fig. 6, using the EHS have similar trends to Fig.2, Fig.3, Fig.4, and Fig.5, due to the reasons mentioned previously. As can be seen from Fig.6, the behavior of the three curves without the EHS is different than that of the curves given by Fig.2, Fig.3, Fig.4, and Fig.5. We concluded that this discrepancy is related to the EHS. In addition, the rapidly increasing behavior of the curves from zero is due to the selection of the initial value in the computer simulations, as explained previously in detail.

The EE gains and the number of activated antennas of the system obtained for the three cases, are given in Tables V and VI, respectively. From the model of EHS in (30), we can see that the number of activated antennas in the system is one of the main factors affecting the EE gain of EHS. Tables V and table VI show that the gain of EE of the

TABLE VI
THE NUMBER OF ACTIVATED ANTENNA (NUMAA) WITH THREE SCHEMES

	20UE&6BSsCoMP	50UE&6BSsCoMP	75UE&4BSsCoMP
NumAA	1599	1786	1983

EHS increases gradually with the increase of the number of activated antennas. Fig. 7 is an iterative diagram of the energy recovery of the system in those cases. From Fig. 7, we can see that 75UE&4BSCoMP (one macro and 50 picos BSs serve 75 UEs, and the UEs adopt 4 BSs cooperation) can harvest the most energy compared with the other two conditions because it has the largest number of activated antennas. When the energy recovery coefficient is fixed, the performance of the EHS depends mainly on the number of active antennas, the transmission power of BSs, as well as on the channel state information (CSI). Therefore, when the transmission power of BSs and CSI are the same, more energy can be recovered as the number of active antennas gets larger. It can be seen from Fig.7 that the energy recovery curves rapidly increase from 0 to a maximum, and then gradually decrease until it stabilizes. The behavior of the curves is determined by the initial value, set by during the simulations, as well as by the proposed algorithm. Namely, the initial value of the energy recovery is chosen to zero, and the transmission power of BSs, the number of active antennas, and other parameters are all set to a maximum value. As the redundant antennas in the system are gradually deactivated while running the algorithm, the transmission power of BSs is also optimized continuously and the energy recovery curves also tend to stabilize as the number of iterations increases.

V. CONCLUSION

In this paper, we investigated the EE of mmWave massive MIMO in UDN. To develop the functions of Massive MIMO,

the EHS based SWIPT is adopted at each BS, and a system model of energy recovery was proposed. To utilize the characteristics of UDN, the UE-centered CoMP technology was employed and the association algorithm between UE and BS based CoMP and MaxEE was designed. Also, to utilize the energy recovery model, a new association algorithm between UE and antenna-based F-norm and MaxEE was proposed. The convergence and optimality of the proposed algorithm have been verified via both theoretical analysis and computer simulations. It has been shown that our algorithm outperformed the MaxSINR algorithm. Compared with NCoMP, we concluded that utilizing CoMP can achieve higher EE, and the EE gain increases gradually with the increase in the number of cooperative BSs.

APPENDIX A

The proof of convergence of algorithm 5:

In this part, we prove the convergence of Algorithm 5 [30], [33], [46], [47]. To simplify the symbol, we define the equivalent objective function in (36) as:

$$\begin{aligned} F(\Xi) = &\max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} U(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}) \\ &-\Xi P_{\text{sum}}(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}). \end{aligned} \tag{79}$$

To facilitate the subsequent proof, we first introduce the following two propositions.

Proposition 1: $F(\Xi)$ is a strictly monotonic decreasing function in Ξ, i.e.$F(\Xi_2) > F(\Xi_1)$, if $\Xi_1 > \Xi_2$.

Proof: Let $\{\mathbf{X}^1, \mathbf{N}^1_{AA}, \mathbf{S}^1, \mathbf{P}^1_{TX}\} \in F$ and $\{\mathbf{X}^2, \mathbf{N}^2_{AA}, \mathbf{S}^2, \mathbf{P}^2_{TX}\} \in F$ be two distinct optimal resource allocation policies of $F(\Xi)$, i.e. $F(\Xi_1)$ and $F(\Xi_2)$.

$$\begin{aligned} F(\Xi_2) &= \max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} U(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}) \\ &\quad -\Xi_2\ P_{\text{sum}}(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}) \\ &= U(\mathbf{X}^2,\mathbf{N}^2_{AA},\mathbf{S}^2,\mathbf{P}^2_{TX}) \\ &\quad -\Xi_2\ P_{\text{sum}}(\mathbf{X}^2,\mathbf{N}^2_{AA},\mathbf{S}^2,\mathbf{P}^2_{TX}) \\ &> U(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX}) \\ &\quad -\Xi_2\ P_{\text{sum}}(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX}) \\ &\geq U(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX}) \\ &\quad -\Xi_1\ P_{\text{sum}}(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX}) \\ &= F(\Xi_1). \end{aligned} \tag{80}$$

Proposition 2: For $\forall\{\mathbf{X}^1, \mathbf{N}^1_{AA}, \mathbf{S}^1, \mathbf{P}^1_{TX}\} \in F$ and $\Xi = \dfrac{U(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX})}{P_{\text{sum}}(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX})}$, $F(\Xi) \geq 0$

Proof:

$$\begin{aligned} F(\Xi) = &\max_{\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}} U(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}) \\ &-\Xi P_{\text{sum}}(\mathbf{X},\mathbf{N}_{AA},\mathbf{S},\mathbf{P}_{TX}). \\ \geq\ &U(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX}) \\ &-\Xi P_{\text{sum}}(\mathbf{X}^1,\mathbf{N}^1_{AA},\mathbf{S}^1,\mathbf{P}^1_{TX}) = 0. \end{aligned} \tag{81}$$

Proof of Convergence: Let $\{\mathbf{X}_n, \mathbf{N}_{AA,n}, \mathbf{S}_n, \mathbf{P}_{TX,n}\}$ be the optimal resource allocation policies of $F(\Xi)$ in the n-th iteration. Suppose Ξ_n and Ξ_{n+1} represent the EE of the considered system in the n-th and $(n+1)$-th iterations, respectively, and $\Xi_n \neq \Xi^*$, $\Xi_{n+1} \neq \Xi^*$. Base on (35) and Proposition 2, we have $F(\Xi_n) > 0$ and $F(\Xi_{n+1}) > 0$. In Algorithm 5, we calculate Ξ_{n+1} as $\Xi_{n+1} = \dfrac{U(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n})}{P_{\text{sum}}(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n})}$ [46], [47]. Then we can express $F(\Xi_n)$ as:

$$\begin{aligned} F(\Xi_n) &= U(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n}) \\ &\quad -\Xi_n P_{\text{sum}}(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n}) \\ &= \Xi_{n+1} P_{\text{sum}}(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n}) \\ &\quad -\Xi_n P_{\text{sum}}(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n}) \\ &= (\Xi_{n+1}-\Xi_n) P_{\text{sum}}(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n}). \end{aligned} \tag{82}$$

$\because P_{\text{sum}}(\mathbf{X}_n,\mathbf{N}_{AA,n},\mathbf{S}_n,\mathbf{P}_{TX,n}) > 0$, $\therefore \Rightarrow \Xi_{n+1} > \Xi_n$. Therfore the EE Ξ increases in each iteration. When the number of iterations is sufficiently large, $F(\Xi_n)$ will eventually approach zero, i.e.$F(\Xi_n) = 0$, find the MaxEE Ξ^* [46].

APPENDIX B

The proof of the proposition:

Let $(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho})$ be the optimized Lagrange multipliers at the convergence of the subgradient method. Let $(\mathbf{X}^*, \mathbf{N}^*_{AA}, \mathbf{S}^*, \mathbf{P}^*_{TX})$ be the primal solution recovered from the dual variable $(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho})$ by means of Eq. 46, Algorithm 2, Eq.65, and Algorithm 4 [48]. We have then:

$$\begin{aligned} &L(\mathbf{X}^*, \mathbf{N}^*_{AA}, \mathbf{S}^*, \mathbf{P}^*_{TX}) \\ &= \sum\nolimits_i \max_j \log(W_j \log_2(1+SINR^*_{ij})) \\ &\quad -q(\frac{\sum_j b_j P^{*j}_{TX}}{\mu} + C_0 + C_1 \sum\nolimits_j b_j + C_2 \sum\nolimits_j b_j N^{*j}_{AA}) \\ &\quad +q\eta \sum\nolimits_j b_j (\sum\nolimits_k b_k \sum\nolimits_n j_{jn} \sum\nolimits_m j_{km} \frac{P^{*k}_{TX}}{N^{*k}_{AA}} |G^{m,n}_{k,j}|^2), \end{aligned} \tag{83}$$

$$\begin{aligned} &D(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}) \\ &= \sum\nolimits_i \max_j \log(W_j \log_2(1+SINR^*_{ij})) \\ &\quad -q(\frac{\sum_j b_j P^{*j}_{TX}}{\mu} + C_0 + C_1 \sum\nolimits_j b_j + C_2 \sum\nolimits_j b_j N^{*j}_{AA}) \\ &\quad +q\eta \sum\nolimits_j b_j (\sum\nolimits_k b_k \sum\nolimits_n j_{jn} \sum\nolimits_m j_{km} \frac{P^{*k}_{TX}}{N^{*k}_{AA}} |G^{m,n}_{k,j}|^2) \\ &\quad +\sum\nolimits_j \tau_j (K^j_U - K_j) + \sum\nolimits_i \nu_i (M - U_i) \\ &\quad +\sum\nolimits_j \omega_j (P^{max}_{TX} - P^{*j}_{TX}) + \sum\nolimits_i \alpha_i (R^*_i - R_{QoS}) \\ &\quad +\sum\nolimits_j \gamma_j (N^{*j}_{AA} - K_j) + \sum\nolimits_j \rho_j (N_t - N^{*j}_{AA}). \end{aligned} \tag{84}$$

Comparing (83) and (84), it follows that:

$$\begin{aligned} &L(\mathbf{X}^*, \mathbf{N}^*_{AA}, \mathbf{S}^*, \mathbf{P}^*_{TX}) \\ &\quad = D(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}) - [\sum\nolimits_j \tau_j (K^j_U \\ &\qquad -K_j) + \sum\nolimits_i \nu_i (M - U_i) \\ &\qquad +\sum\nolimits_j \omega_j (P^{max}_{TX} - P^{*j}_{TX}) \\ &\qquad +\sum\nolimits_i \alpha_i (R^*_i - R_{QoS}) + \sum\nolimits_j \gamma_j (N^{*j}_{AA} - K_j) \\ &\qquad +\sum\nolimits_j \rho_j (N_t - N^{*j}_{AA})] \end{aligned} \tag{85}$$

Then, by assuming that $(\mathbf{X}^{**}, \mathbf{N}_{AA}^{**}, \mathbf{S}^{**}, \mathbf{P}_{TX}^{**})$ is the global optimal solution for the problem in Eq.36, by the weak duality, it always holds that $D(\boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\rho}) \geq L(\mathbf{X}^{**}, \mathbf{N}_{AA}^{**}, \mathbf{S}^{**}, \mathbf{P}_{TX}^{**})$. Combining this result with (85), we prove the claim

$$\begin{aligned} & L(\mathbf{X}^{*}, \mathbf{N}_{AA}^{*}, \mathbf{S}^{*}, \mathbf{P}_{TX}^{*}) \\ & \geq L(\mathbf{X}^{**}, \mathbf{N}_{AA}^{**}, \mathbf{S}^{**}, \mathbf{P}_{TX}^{**}) \\ & \quad -[\sum\nolimits_{j} \tau_j (K_U^j - K_j) \\ & \quad + \sum\nolimits_{i} \nu_i (M - U_i) \\ & \quad + \sum\nolimits_{j} \omega_j (P_{TX}^{max} - P_{TX}^{*j}) + \sum\nolimits_{i} \alpha_i (R_i^{*} - R_{QoS}) \\ & \quad + \sum\nolimits_{j} \gamma_j (N_{AA}^{*j} - K_j) \\ & \quad + \sum\nolimits_{j} \rho_j (N_t - N_{AA}^{*j})] \end{aligned} \tag{86}$$

ACKNOWLEDGMENT

The authors would like to thank the Editor and the anonymous reviewers for carefully reviewing this article and providing valuable comments for improving the manuscript. They would also like to thank to Richard Wilkinson for his diligent proofreading of this article.

REFERENCES

[1] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[2] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 13–21, Oct. 2015.

[3] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: The next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[4] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[5] S. Yunas, M. Valkama, and J. Niemelä, "Spectral and energy efficiency of ultra-dense networks under different deployment strategies," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 90–100, Jan. 2015.

[6] W. Feng, Y. Wang, D. Lin, N. Ge, J. Lu, and S. Li, "When mmWave communications meet network densification: A scalable interference coordination perspective," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1459–1471, Jul. 2017.

[7] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 104–110, Nov. 2014.

[8] Z. Ding *et al.*, "Application of smart antenna technologies in simultaneous wireless information and power transfer," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 86–93, Apr. 2015.

[9] J. Huang, C.-C. Xing, and C. Wang, "Simultaneous wireless information and power transfer: Technologies, applications, and research challenges," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 26–32, Nov. 2017.

[10] M. Feng, S. Mao, and T. Jiang, "Base station on-off switching in 5G wireless networks: Approaches and challenges," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 46–54, Aug. 2017.

[11] T. Han and N. Ansari, "On greening cellular networks via multicell cooperation," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 82–89, Feb. 2013.

[12] 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.

[13] G. Ye, H. Zhang, H. Liu, J. Cheng, and V. C. M. Leung, "Energy efficient joint user association and power allocation in a two-tier heterogeneous network," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–5.

[14] T. Van Chien, E. Bjornson, and E. G. Larsson, "Joint power allocation and user association optimization for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6384–6399, Sep. 2016.

[15] 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. Technol.*, vol. 66, no. 1, pp. 580–593, Jan. 2017.

[16] M. Sheng, L. Wang, X. Wang, Y. Zhang, C. Xu, and J. Li, "Energy efficient beamforming in MISO heterogeneous cellular networks with wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 954–968, Apr. 2016.

[17] R. Zhang, L.-L. Yang, and L. Hanzo, "Energy pattern aided simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1492–1504, Aug. 2015.

[18] G. Zhu and K. Huang, "Analog spatial cancellation for tackling the near-far problem in wirelessly powered communications," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3566–3576, Dec. 2016.

[19] F. Benkhelifa, A. S. Salem, and M.-S. Alouini, "Sum-rate enhancement in multiuser MIMO decode-and-forward relay broadcasting channel with energy harvesting relays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3675–3684, Dec. 2016.

[20] W. Cheng, X. Zhang, and H. Zhang, "Statistical-QoS driven energy-efficiency optimization over green 5G mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3092–3107, Dec. 2016.

[21] R. Zi, X. Ge, J. Thompson, C.-X. Wang, H. Wang, and T. Han, "Energy efficiency optimization of 5G radio frequency chain systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 758–771, Apr. 2016.

[22] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems—Part I: Chunk allocation," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2734–2744, Sep. 2009.

[23] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems—Part II: Joint chunk, power and bit allocation," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 499–509, Feb. 2012.

[24] H. Zhu, "Radio resource allocation for OFDMA systems in high speed environments," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 748–759, May 2012.

[25] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[26] X. Gao, L. Dai, Y. Sun, S. Han, and I. Chih-Lin, "Machine learning inspired energy-efficient hybrid precoding for mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[27] Z. Gao, L. Dai, and Z. Wang, "Channel estimation for mmWave massive MIMO based access and backhaul in ultra-dense network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[28] N. Wang, E. Hossain, and V. K. Bhargava, "Joint downlink cell association and bandwidth allocation for wireless backhauling in two-tier HetNets with large-scale antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3251–3268, May 2016.

[29] S. Han, C. Yang, and A. F. Molisch, "Spectrum and energy efficient cooperative base station doze," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 285–296, Feb. 2014.

[30] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation for secure OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2572–2585, Jul. 2012.

[31] A. Ghosh *et al.*, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, Jun. 2014.

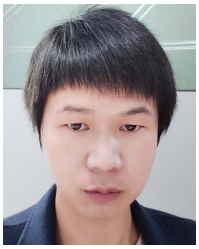
[32] *Technical Specification Group Radio Access Network; Evolved UniversalTerrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, Standard 3GPP TR 36.814 V9.0.0, 3rd Generation Partnership Project, Tech. Rep., 2010.

[33] S. Schaible, "Fractional programming. II, on dinkelbach's algorithm," *Manage. Sci.*, vol. 22, no. 8, pp. 868–873, Apr. 1976.

[34] M. T. Altabbaa, T. Arsan, and E. Panayirci, "Subchannel allocation and power control for uplink femtocell radio networks with imperfect channel state information," *Wireless Pers. Commun.*, vol. 108, no. 3, pp. 1345–1361, Oct. 2019, doi: 10.1007/s11277-019-06472-1.

[35] D. Bethanabhotla, O. Y. Bursalioglu, H. C. Papadopoulos, and G. Caire, "User association and load balancing for cellular massive MIMO," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2014, pp. 1–10.

[36] Z. Dong, P. Fan, W. Zhou, and E. Panayirci, "Power and rate adaptation for MQAM/OFDM systems under fast fading channels," in *Proc. IEEE 75th Veh. Technol. Conf. (VTC Spring)*, May 2012, pp. 1–5.
[37] Z. Dong, P. Fan, and X. Lei, "Power adaptation in OFDM systems based on velocity variation under rapidly time-varying channels," *IEEE Commun. Lett.*, vol. 19, no. 4, pp. 689–692, Apr. 2015.
[38] Z. Dong, P. Fan, E. Panayirci, and X. Lei, "Conditional power and rate adaptation for MQAM/OFDM systems under CFO with perfect and imperfect channel estimation errors," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5042–5055, Nov. 2015.
[39] J. Tang, A. Shojaeifard, D. K. C. So, K.-K. Wong, and N. Zhao, "Energy efficiency optimization for CoMP-SWIPT heterogeneous networks," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6368–6383, Dec. 2018.
[40] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Multi-objective optimization for robust power efficient and secure full-duplex wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5511–5526, Aug. 2016.
[41] C. Pan *et al.*, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, Aug. 2020.
[42] C.-Y. Wang, M. Wigger, and A. Zaidi, "On achievability for downlink cloud radio access networks with base station cooperation," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5726–5742, Aug. 2018.
[43] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1151–1163, Jun. 2011.
[44] J. Wang, H. Zhu, and N. J. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 675–683, May 2012.
[45] Q. Zhang, C. Yang, H. Haas, and J. S. Thompson, "Energy efficient downlink cooperative transmission with BS and antenna switching off," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 5183–5195, Sep. 2014.
[46] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. M. Leung, "Joint user scheduling and power allocation optimization for energy-efficient NOMA systems with imperfect CSI," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2874–2885, Dec. 2017.
[47] G. Ozcan, M. C. Gursoy, N. Tran, and J. Tang, "Energy-efficient power allocation in cognitive radio systems with imperfect spectrum sensing," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3466–3481, Dec. 2016.
[48] K. Shen and W. Yu, "Distributed pricing-based user association for downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1100–1113, Jun. 2014.



Bin Li is currently pursuing the Ph.D. degree with the School of Electronic Information, Wuhan University, Wuhan, China. His research interests include convex optimization, physical layer security, energy harvesting, massive MIMO, NOMA, terahertz communication, large intelligent surface, and 6G wireless communication.

Yonghong Dai received the B.E., M.S., and Ph.D. degrees from the School of Electronics and Information, Wuhan University, Wuhan, China, in 1993, 2002, and 2015, respectively. He is currently an Associate Professor with the School of Electronic Information, Wuhan University. His research interests include GPS satellite positioning and navigation, laser communications, visible light communications, and 5G mobile communications.

Zhicheng Dong (Member, IEEE) was born in Langzhong, China, in 1982. He received the B.E., M.S., and Ph.D. degrees from the School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China, in 2004, 2008, and 2016, respectively.

Since 2016, he has been an Associate Professor with the School of Engineering, Tibet University, Lhasa, China. He has served as a Technical Program Committee member for IEEE Globecom 2017, IEEE ICC (2017, 2016, and 2015), TENSYMP 2015, and Chinacom 2014. His research interests include adaptation technology, performance analysis, and signal processing for high mobility wireless communications.

Erdal Panayirci (Life Fellow, IEEE) received the Ph.D. degree in electrical engineering and system science from Michigan State University, USA, in 1971.

He is currently a Professor with the Electrical and Electronics Engineering Department, Kadir Has University, İstanbul, Turkey, and the Visiting Research Collaborator with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA. He spent the academic years 2008–2009 and 2017–2018 with the Department of Electrical Engineering, Princeton University. He has published extensively in leading scientific journals and international conferences and coauthored the book *Principles of Integrated Maritime Surveillance Systems* (Kluwer Academic, 2000). His research interests include communication theory, synchronization, advanced signal processing techniques, and their applications to wireless electrical, underwater, and optical communications. He has served as a member of IEEE Fellow Committee from 2005 to 2008 and from 2018 to 2020 and as a member of the IEEE GLOBECOM/ICC Management and Strategy Standing Committee from 2017 to 2020. He was the Technical Program Co-Chair of the IEEE International Conference on Communications (ICC2006) and the Technical Program Chair of the IEEE PIMRC, both held in İstanbul, Turkey, in 2006 and 2010, respectively. He was the Executive Vice Chairman of the IEEE Wireless Communications and Networking Conference, İstanbul, in April 2014. He was the General Co-Chair of the IEEE PIMRC held in İstanbul, in September 2019. He was an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS in the areas of Synchronization and Equalization from 1995 to 2000.

Huilin Jiang (Member, IEEE) received the Ph.D. degree in information and communication engineering from the National Mobile Communication Laboratory, Southeast University, Nanjing, in 2016. She is currently a Visiting Scholar with the University of Kent and an Associate Professor with the School of Electronic Engineering, Nanjing Xiaozhuang University, Nanjing. Her research interests include radio resource management, green wireless network, small cell and ultra-dense networks, cloud network, energy harvesting, and next-generation wireless communication systems.

Hao Jiang received the B.Eng. degree in communication engineering and the M.Eng. and Ph.D. degrees in communication and information systems from Wuhan University, China, in 1999, 2001, and 2004, respectively. He undertook his post-doctoral research work with LIMOS, Clermont-Ferrand, France, from 2004 to 2005. He was a Visiting Professor with the University of Calgary, Canada, and ISIMA, Blaise Pascal University, France. He is currently a Professor with Wuhan University. He has authored over 60 articles in different journals and conferences. His research interests include mobile ad hoc networks and mobile big data.