

# Backscatter-Assisted Data Offloading in OFDMA-based Wireless Powered Mobile Edge Computing for IoT Networks

Phu X. Nguyen, Dinh-Hieu Tran, *Student Member, IEEE*, O. Onireti, *Member, IEEE*, Phu Tran Tin\*, Sang Quang Nguyen, Symeon Chatzinotas, *Senior Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*,

\*Corresponding author: (Email: phutrantin@tdtu.edu.vn)

**Abstract**—Mobile edge computing (MEC) has emerged as a prominent technology to overcome sudden demands on computation-intensive applications of the Internet of Things (IoT) with finite processing capabilities. Nevertheless, the limited energy resources also seriously hinders IoT devices from offloading tasks that consume high power in active RF communications. Despite the development of energy harvesting (EH) techniques, the harvested energy from surrounding environments could be inadequate for power-hungry tasks. Fortunately, Backscatter communications (Backcom) is an intriguing technology to narrow the gap between the power needed for communication and harvested power. Motivated by these considerations, this paper investigates a backscatter-assisted data offloading in OFDMA-based wireless-powered (WP) MEC for IoT systems. Specifically, we aim at maximizing the sum compu-

tation rate by jointly optimizing the transmit power at the gateway (GW), backscatter coefficient, time-splitting (TS) ratio, and binary decision-making matrices. This problem is challenging to solve due to its non-convexity. To find solutions, we first simplify the problem by determining the optimal values of transmit power of the GW and backscatter coefficient. Then, the original problem is decomposed into two sub-problems, namely, TS ratio optimization with given offloading decision matrices and offloading decision optimization with given TS ratio. Especially, a closed-form expression for the TS ratio is obtained which greatly enhances the CPU execution time. Based on the solutions of the two sub-problems, an efficient algorithm, termed the fast-efficient algorithm (FEA), is proposed by leveraging the block coordinate descent method. Then, it is compared with exhaustive search (ES), bisection-based algorithm (BA), edge computing (EC), and local computing (LC) used as reference methods. As a result, the FEA is the best solution which results in a near-globally-optimal solution at a much lower complexity as compared to benchmark schemes. For instance, the CPU execution time of FEA is about 0.029 second in a 50-user network, which is tailored for ultra-low latency applications of IoT networks.

**Index Terms**—Backscatter communication, Internet of Things (IoT), mobile edge computing (MEC), OFDMA, wireless power transfer (WPT).

## I. INTRODUCTION

Internet of Things (IoT) can play a key role in improving quality of life through applications such as home automation, smart cars, smart city, health care, industrial or agriculture monitoring, augmented reality, and smart grid [1]. The number of IoT devices is estimated to reach 25 billion by 2025 [2], [3]. Moreover, there is an explosive surge of resource-intensive IoT applications such as interactive gaming, multi-view video construction, augmented reality (AR), virtual reality (VR), and face recognition, that impose stringent demands on high computation capability, and low-latency processing [4]. The restricted stor-

The work of Dinh-Hieu Tran and Symeon Chatzinotas was supported in part by the Luxembourg National Research Fund (FNR) in the framework of the FNR-FNRS bilateral project "InWIP-NETs: Integrated Wireless Information and Power Networks." The work of H. V. Poor was supported in part by the U.S. National Science Foundation under Grant CCF-1908308.

Phu X. Nguyen is with Department of Computer Fundamentals, FPT University, Ho Chi Minh City 700000, Vietnam (email: phunx4@fpt.edu.vn, phunx4@fe.edu.vn).

O. Onireti is with James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, UK.(email: Oluwakayo.Onireti@glasgow.ac.uk).

Dinh-Hieu Tran and Symeon Chatzinotas are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), the University of Luxembourg, Luxembourg, (e-mail: {hieu.tran-dinh, symeon.chatzinotas}@uni.lu).

Phu Tran Tin is with Wireless Communications Research Group, Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam, (email: phutrantin@tdtu.edu.vn).

Sang Quang Nguyen is with Institute of Fundamental and Applied Sciences, Duy Tan University, Ho Chi Minh City 700000, Vietnam and Faculty of Electrical-Electronic Engineering, Duy Tan University, Da Nang 550000, Vietnam, (email: nguyenguangsang3@duytan.edu.vn).

H. V. Poor is with the Electrical Engineering Department at Princeton University, NJ 08544, (email: poor@princeton.edu).

Corresponding author: Phu Tran Tin, (email: phutrantin@tdtu.edu.vn).

age and computation resources of IoT users besides their constrained battery capacity make them unsuitable for processing resource-hungry applications. Mobile cloud computing (MCC) is a promising solution to overcome the above issues in IoT networks [5]. Specifically, offloading the heavy computation tasks to a central cloud (CC) mitigates the IoT devices' workload responding to low latency and highly accurate applications. Nevertheless, the MCC faces many problems such as network congestion and high transmission delay due to many users connect to the CC simultaneously and distant from users to CC, respectively, which reduces the system performance. To this end, Mobile Edge Computing (MEC) was recently proposed to eliminate the shortcomings of the traditional MCC due to its compute-intensive capability and the close proximity to end-users. Consequently, MEC provides high bandwidth connectivity and comes hand-in-hand with ultra-low latency in computational offloading tasks [6]–[9].

Extensive studies have investigated the MEC in IoT networks [10]–[14]. The works in [10] and [11] investigated the scheduling of tasks for IoT services in MEC. Specifically, Hu et al. [11] studied the resource scheduling problem in ultra-dense edge computing (UDEEC) networks. The authors in [10] solved the problem of maximizing the number of admitted tasks based on their deadline requirements, wherein they mathematically formulated the dynamic task offloading and scheduling (DTOS) problem by jointly optimizing the computing offloading task, resource allocation, and task scheduling. Then, they applied the Logic-based Benders Decomposition approach to solve it efficiently. Different from works in [10], [11] that only focused on modeling the offloading decision, Yousafzai et al. [12] investigated the process-level computational offloading. Specifically, the authors proposed a process migration-based computational offloading (PMCO) framework that offloads a computation-intensive task from resource-limited IoT devices to a mobile cloud/edge computing. Facilitated by machine learning, Cui et al. [13] studied the predicted online learning method to improve the network association by forecasting the mobility of IoT mobile machines for ultra-low latency tasks in MEC networks. The concept of a decentralized and revised content-centric networking (CCN)-based MEC was presented in [14]. Specifically, they proposed a novel three-tiers hierarchical MEC network topology encompassing core cloud, regional cloud, and MEC nodes which outperforms the conventional two-tiers architecture in term of the average

service discovery time. Despite remarkable achievement, none of those works [10]–[14] took energy harvesting into consideration.

Especially, the data offloading process for low-powered IoT users can be expensive due to high power consumption in conventional RF communications which limited them to utilize MEC systems [15]. Moreover, since the IoT's battery stores a finite amount of energy, it needs to be replaced or charged, which is infeasible and costly in the hazardous environments, e.g., in toxic (i.e., gas or chemical) environments. This is the reason why energy harvesting (EH) has attracted much attention in the last decade [16], [17], [18]. Particularly, the integration of MEC and wireless power transfer (WPT) creates a novel paradigm called WPMEC which has emerged as a potential solution to tackle the limitation of energy budget and computation capability of low-power IoT devices [19]–[21]. Reference [19] considered a single-user WPMEC system wherein a multi-antenna source utilized energy beamforming to power a single-antenna user. Based on the harvested energy, the user will decide to compute locally a portion of tasks and offload the rest of tasks to MEC server. The authors aimed at minimizing the total transmission power of the source while guaranteeing the accomplished user's task operation. In a two-user WPMEC system [20], the authors studied the maximization of minimum energy-efficiency by optimizing time allocation under minimum rate constraint and maximum harvested energy requirement. Reference [21] investigated a multi-user WPMEC system with binary offloading, wherein the authors aimed at maximizing the total computation rate of all users by jointly optimizing the time allocation and binary decision-making for each user.

Nevertheless, the harvested energy from surrounding environments may be insufficient for power-hungry tasks. This leads to the mismatch between the EH from the harvester and the required energy by the IoT devices. To overcome this issue, the authors in [17] designed a system to push up the amount of harvested energy by scavenging from multiple sources. However, it required a more complicated hardware that may not be suitable for mobile IoT devices. Recently, backscatter communications (Backcom) has received significant attention from researchers due to its ability to bridge the gap between the required power for communication and the supplied power from the energy source [22], [23]. This is because the power consumption of a typical backscatter transmitter is less than  $1 \mu W$  [24] which is well suitable for RF-power energy ranging from  $1 \mu W$  to  $10 \mu W$  [25]. This motivates us to apply backscatter-assisted data offloading

in MEC for IoT networks. There are few existing studies on the backscatter-assisted MEC system [15], [26]. In [15], a deep reinforcement learning (DRL) framework was adopted to learn the optimal offloading schedule in different network scenarios. [26] was the first work that considers a backscatter-assisted hybrid offloading scheme in MEC. Specifically, based on the different required workload, the authors aimed to minimize the entire energy consumption by jointly optimizing the power allocation and offloading schemes for each device. Nevertheless, [15] and [26] only considered time-division multiple access (TDMA) model which was only tailored for delay-tolerant applications.

In this work, we investigate the backscatter-assisted data offloading in OFDMA-based wireless power MEC (WPMEC) for limited computational resources and low-power IoT devices. Notably, the integration of the latest wireless advances, e.g., backscatter, OFDMA, WPMEC, in IoT communication networks aims at designing a system that overcomes the restrictions of IoT users, i.e., low computational capacity and battery constraint, and suitable with ultra-low latency applications. So clearly, these technologies have intimate contact with each other, and combining them is promising. OFDMA can be considered as an extended version of OFDM which is optimized for multi-user scenarios [2], [27], [28]. In an OFDMA-based system, multi-users can operate at the same time at different sub-channels which is appropriate for ultra-low latency applications. Meanwhile, OFDMA-based systems outperforms their TDMA-based counterparts wherein each user is allocated a different time slot to prevent interference, which thus leads to much higher delay. Despite many advantages, subchannel allocation in OFDMA-based MEC systems impose a more involved mixed-integer programming problem which is challenging to solve [29], [30]. To this end, we design an efficient decision-making algorithm that efficiently utilizes limited MEC resources to maximize the overall network performance e.g., sum-rate maximization. In summary, our contributions are as follows:

- This is the first work that jointly considers the combination of OFDMA and backscatter-assisted WPMEC in IoT networks. Specifically, we propose a novel model in which an IoT user is able to harvest energy from the gateway (GW) and it then utilizes the harvested energy for backscattering its own data to the GW or computes locally.
- We aim at maximizing the sum computation rate subject to constraints on the maximum transmit power

at the GW, maximum backscatter coefficient, time-splitting (TS) ratio, and decision-making matrices. This optimization problem is a mixed-integer non-linear programming (MINLP) problem, which is NP-hard. Particularly, the binary nature of the decision-making variables in the objective function makes it more troublesome. Instead of applying existing approximation methods or an Inner Approximation framework [31]–[33] to solve the problem which is still highly complex, we design two-layer alternating algorithms to solve the non-convex optimization problems by adopting the block coordinate descent (BCD) method [34].

- We first simplify the problem by obtaining the optimal values of transmit power at the GW and backscatter coefficient. Then, the problem is decomposed into two sub-problems, namely, TS ratio optimization with fixed offloading decision matrices and offloading decision optimization with a fixed value of TS ratio. Particularly, a closed-form solution for TS ratio optimization is derived which dramatically pushes up the CPU execution time. Based on the solutions of these two sub-problems, a BCD-based algorithm, termed the *fast-efficient algorithm* (FEA), is proposed for alternately optimizing the TS ratio and offloading decision to maximize the sum computation rate. Then, it is compared with benchmark schemes such as exhaustive search (ES), bisection-based algorithm (BA), edge computing (EC), and local computing (LC). Although the ES provides a globally optimal solution, this method is limited to a small-scale network because its complexity is exponentially increasing with the number of IoT devices. Besides, the EC and LC methods have significantly degraded performance. Furthermore, BA also achieves a near-globally optimal solution but this scheme consumes a higher CPU execution time as compared to the FEA method. Thus, the FEA is adopted to significantly reduce the complexity while guaranteeing that the performance converges to that of the exhaustive method.
- The effectiveness of the proposed algorithm is verified through simulation results, which show significant improvements in the maximum sum computation rate in comparison with the benchmark schemes. Specifically, the benchmark ES algorithm takes all offloading decisions into consideration, the benchmark BA algorithm is designed similar to FEA but it adopts the bisection method to solve the

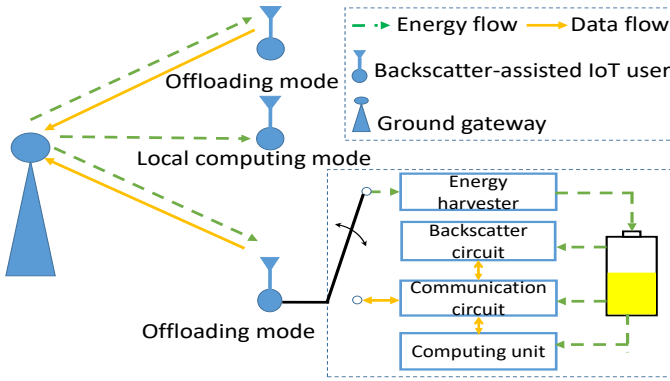


Fig. 1: System model.

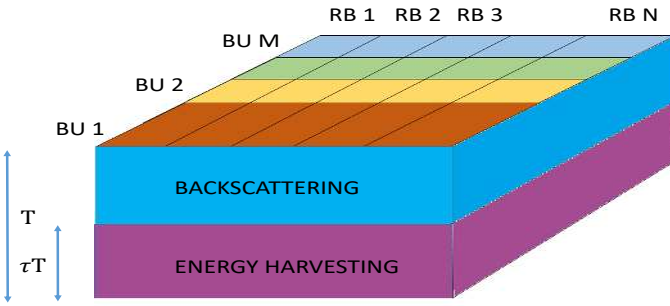


Fig. 2: Frame structure of OFDMA-based MEC system.

TS ratio optimization problem instead of using a closed-form expression. Moreover, the benchmark EC and LC schemes are designed such that each IoT user tries to offload all computation tasks to the MEC server or execute locally without considering offloading decision-making.

The rest of this work is organized as follows. Section II introduces the system model and problem formulation. The sum-rate maximization problem is analyzed in Section III. The proposed iterative algorithms is proposed in Section IV. Section V shows the simulation results and Section VI concludes the paper.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

As shown in Fig. 1, we consider an OFDMA-based WPMEC system with  $M$  single-antenna backscatter-assisted IoT users (BUs) denoted by  $\mathcal{M} = \{1, \dots, m, \dots, M\}$ , which are served by a single-antenna GW integrated with MEC server and a stable power supply. We assume that the EH and communication operate in the same frequency band. Without loss of generality, the system bandwidth is partitioned into  $N$  resource blocks (RBs) denoted by  $\mathcal{N} = \{1, \dots, n, \dots, N\}$ , where each RB has a bandwidth of 180 kHz [2], [28]. Besides, let us define the offloading decision-making (i.e., offloading or local computing) as

a matrix  $\mathbf{D}_{M \times N} \triangleq \{d_{mn} | m \in \mathcal{M}, n \in \mathcal{N}\}$ , then we have decision-making constraints as follows

$$d_{mn} \in \{0, 1\}, \quad (1)$$

$$\sum_{m \in \mathcal{M}} d_{mn} \leq 1, \forall n \in \mathcal{N}, \quad (2)$$

where  $d_{mn} = 1$  or  $d_{mn} = 0$  imply that the BU  $m$  will offload the data to MEC server through RB  $n$  or compute locally, respectively. Constraint (2) means that one RB  $n$  is only allocated to one BU  $m$  during the time period  $T$ . During each time frame  $T$ , the wireless channel coefficient between GW and  $m$ -th BU operating on RB  $n$  is denoted by  $h_{mn}$ , which is assumed to be reciprocal for the uplink (UL) and downlink (DL) [21]. Moreover,  $T$  is assumed to be sufficiently small such that wireless channels are static within each time frame  $T$  but it may change across different time frames.

To avoid mutual interference, we apply the time-splitting (TS) illustrated as in Fig. 2 to optimize the fraction of time  $\tau T$  and  $(1 - \tau)T$  for the EH and data offloading, respectively. Whereas  $\tau \in (0, 1]$  is the TS ratio, with  $\tau = 1$  indicates that all BUs locally perform their computation tasks during  $T$  and  $\tau < 1$  means that a portion of time  $\tau T$  is used for EH and  $(1 - \tau)T$  is used for backscattering its own signal to MEC server. Thus, the harvested energy at the BU  $m$  is  $E_m^h = \chi \tau T P_w \sum_{n \in \mathcal{N}} |h_{mn}|^2$ , where  $0 \leq \chi \leq 1$  is the energy harvesting efficiency,  $P_w$  is the transmit power of the GW, and  $|h_{mn}|^2$  is the channel gain from GW to BU  $m$  using RB  $n$ .

**Local Computing:** The total processed bits at the BU  $m$  can be defined as  $S_m = f_m t_m / n_c$ , where  $f_m$ ,  $t_m$ , and  $n_c$  are the processor's computation speed (cycles/second), computation time, and the number of cycles used to process one bit. In local computing mode, we assume that the BU can perform local computing and energy harvesting at the same time [35]–[37]. Consequently, the local computing rate (in bps) at the BU  $m$  can be calculated as  $R_m^L = S_m / T = f_m t_m / (n_c T)$ .

Furthermore, the energy consumption of the BU  $m$  utilized for local computing must be less than or equal to the harvested energy, which is expressed as  $\beta_m f_m^3 t_m \leq E_m^h$ , with  $\beta_m$  denotes the computation energy efficiency coefficient [37]–[39]. In order to maximize the local computing rate, we assume that BU  $m$  utilizes all the harvested energy for computing task, thus we have

$$R_m^L(P_w, \tau) = \left( \frac{\chi \tau P_w \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\beta_m} \right)^{1/3} \frac{1}{n_c}, \quad (3)$$

where  $t_m = T$ ,  $f_m = (E_m^h / (\beta_m t_m))^{1/3}$ .

**Computing Offloading:** The BUs offloads/reflects their computing tasks to GW based on OFDMA RBs allocation as known in priori, i.e., the BUs have  $d_{mn} = 1$ . During this mode, the BUs receive the activation RF signal from GW and reflect back to GW during  $(1 - \tau)T$  period. Let  $P_c$  denotes the circuit power consumption of the BU, thus we have the following constraint [40, Eq. 6]

$$P_c(1 - \tau)T \leq E_m^h = \chi \tau T P_w \sum_{n \in \mathcal{N}} |h_{mn}|^2. \quad (4)$$

Constraint (4) means that the total harvested energy must be higher than or equal to the circuit power consumption of the BU in the case of offloading mode.

Let  $x_w$  denotes the broadcasting RF signal of the GW with unit power. Thus, the received signal at the BU  $n$  from GW is given by

$$y_m = \sqrt{P_w} \sum_{n \in \mathcal{N}} |h_{mn}|^2 x_w. \quad (5)$$

Note that the noise power at the BU is neglected due to the BU only consists of passive RF components [40]–[42].

The received signal at the GW by reflecting from BU  $m$  is expressed as [41]

$$y_w = \sqrt{\eta P_w} h_{mn} \sum_{n \in \mathcal{N}} h_{mn} x_s c_m + n_w, \quad (6)$$

where  $c_m$  denotes the information signal of BU  $m$ ,  $n_w \sim \mathcal{CN}(0, \sigma_w^2)$  is the additive white Gaussian noise (AWGN), and  $0 \leq \eta \leq 1$  is the backscatter coefficient of the BU. Due to the losses in Backscatter device, the value of  $\eta$  can not achieve 1 in practice [39], [43]. This motivates us to introduce a threshold for  $\eta$ , i.e.,  $\eta \leq \eta_{max}$  with  $0 < \eta_{max} < 1$ . Particularly, the processing delay can be considered as a very small value and is ignored in this work which is commonly used in [40]–[42].

Consequently, the achieved offloading rate (in bps) for BU  $m$  over RB  $n$  is expressed as

$$R_{mn}^O(P_w, \eta, \tau) = B_n(1 - \tau) \log_2 \left( 1 + \mathcal{R}_{mn}^O \right), \quad (7)$$

where  $\mathcal{R}_{mn}^O = \frac{\eta P_w |h_{mn}|^2 \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\sigma_w^2}$  denotes signal to noise ratio and  $B_n$  is the bandwidth of the RB  $n$ , i.e.,  $B_n = 180$  khz [28], [44].

### III. SUM-RATE MAXIMIZATION

In this section, we design an optimal decision-making for offloading the BU's data to GW utilizing OFDMA resource allocation. Specifically, we aim at maximizing a sum-rate local and offloading throughput by jointly optimizing the TS ratio, backscatter coefficient, power

allocation, and resource allocation. Specifically, the sum computation rate is represented as

$$S(\mathbf{D}, P_w, \eta, \tau) \triangleq \sum_{m \in \mathcal{M}} \left( 1 - \min \left( 1, \sum_{n \in \mathcal{N}} d_{mn} \right) \right) R_m^L(P_w, \tau) + \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} d_{mn} R_{mn}^O(P_w, \eta, \tau). \quad (8)$$

Then, the optimization problem can be formulated as

$$\mathcal{P}_1 : \max_{\mathbf{D}, P_w, \eta, \tau} S(\mathbf{D}, P_w, \eta, \tau) \quad (9a)$$

$$\text{s.t. } (1), (2), (4), \quad (9b)$$

$$\tau \in (0, 1], \quad (9c)$$

$$0 < \eta \leq \eta_{max}, \text{ with } 0 < \eta_{max} < 1, \quad (9d)$$

$$0 < P_w \leq P_{max}, \quad (9e)$$

where  $P_{max}$  is the power budget of the GW. Constraint (9e) implies that the transmit power of the GW should be less than the power budget  $P_{max}$ .

$\mathcal{P}_1$  is a highly non-convex optimization problem due to the binary nature of  $d_{mn}$  and the couple of  $\tau$  and  $\eta$  which is difficult to be directly solved. First, we transform  $\mathcal{P}_1$  into a more tractable form. Particularly, the following lemma holds.

**Lemma 1:** Given  $\mathbf{D}$ ,  $P_w$ , and  $\tau$ , the objective  $S(\eta)$  is a non-decreasing function w.r.t.  $\eta$  and it obtains the maximum value at  $\eta = \eta_{max}$ .

*Proof:* See Appendix A. ■

**Lemma 2:** Given  $\mathbf{D}$ ,  $\eta$ , and  $\tau$ , the objective  $S(P_w)$  is a non-decreasing function w.r.t.  $P_w$  and it obtains the maximum value at  $P_w = P_{max}$ .

*Proof:* See Appendix A. ■

Based on Lemmas 1 and 2, problem  $\mathcal{P}_1$  is simplified from four variables to two variables, which is expressed as follows

$$\mathcal{P}_2 : \max_{\mathbf{D}, \tau} S(\mathbf{D}, P_{max}, \eta_{max}, \tau) \quad (10a)$$

$$\text{s.t. } (1), (2), (4), \quad (10b)$$

$$\tau \in (0, 1]. \quad (10c)$$

Although  $\mathcal{P}_2$  is more tractable than  $\mathcal{P}_1$ , it is still non-convex due to the binary nature of  $\mathbf{D}$ . To solve it, we propose an efficient algorithm, namely, *fast-efficient algorithm* adopting block coordinate descent (BCD) technique. In this regard, we provide a corresponding solution in the next section.

### IV. PROPOSED ALTERNATING ALGORITHM FOR SOLVING $\mathcal{P}_1$

This section provides an iterative algorithm based on the BCD method to solve the design problem  $\mathcal{P}_1$ . We first present the ES scheme for the globally optimal solution. Then, the proposed FEA scheme is shown in details.

### A. Exhaustive Search

We first enumerate all  $(M + 1)^N$  possible offloading decision-making matrices denoted by  $\mathcal{D} \triangleq \{\mathbf{1}, \dots, \mathbf{D}^k, \dots, \mathbf{D}^{(M+1)^N}\}$  with  $1 \leq k \leq (M + 1)^N$  that satisfy constraints (1) and (2). For each given  $\mathbf{D}^k$ ,  $\mathcal{P}_2$  becomes a single-variable optimization problems that is given as follows:

$$\mathcal{P}_3 : \max_{\tau} S(\mathbf{D}^k, P_{\max}, \eta_{\max}, \tau) \quad (11a)$$

$$\text{s.t. (4),} \quad (11b)$$

$$\tau \in (0, 1]. \quad (11c)$$

We next specify the feasible set of  $\tau$  for given  $\mathbf{D}^k$ . From (4), we obtain  $\tau \geq \frac{P_c T}{P_c T + \chi T P_w \sum_{n \in \mathcal{N}} |h_{mn}|^2} \triangleq \zeta$  which means that  $\zeta \leq \tau \leq +\infty$ . By combining with constraint (9c), we obtain  $\zeta \leq \tau \leq 1$ . Fortunately, since  $\zeta$  always less than or equal to 1, thus the feasible set of  $\tau$  always exists. Consequently, for any given offloading decision matrix  $\mathbf{D}^k$ , the optimal time allocation  $\tau$  of  $\mathcal{P}_3$  can be obtained by addressing the following problem

$$\mathcal{P}_4 : \max_{\tau} S(\mathbf{D}^k, P_{\max}, \eta_{\max}, \tau) \quad (12a)$$

$$\text{s.t. } \zeta \leq \tau \leq 1. \quad (12b)$$

**Lemma 3:** Let us define  $\tau^*$  as a strict local optimal solution of  $\mathcal{P}_4$ . Then, the following closed-form expression of  $\tau^*$  can be expressed as

$$\tau^* = \left( \frac{\Omega_3}{3\Omega_4} \right)^{3/2}. \quad (13)$$

where

$$\Omega_3 \triangleq \sum_{m \in \mathcal{M}} \left( 1 - \min \left( 1, \sum_{n \in \mathcal{N}} d_{mn} \right) \right) \times \left( \frac{\chi \tau P_w \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\beta_m} \right)^{1/3} \frac{1}{n_c}, \quad (14)$$

$$\Omega_4 \triangleq \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} d_{mn} B_n \log_2 \left( 1 + \frac{\eta P_w |h_{mn}|^2 \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\sigma_w^2} \right) \quad (15)$$

*Proof:* See Appendix B. ■

The pseudo-code for solving  $\mathcal{P}_3$  is summarized as in Algorithm 1.

**Complexity Analysis:** Algorithm 1 complexity is  $\mathcal{O}(M^N)$ .

**Convergence Analysis:** Firstly, a global optimal solution of  $\tau^*$  can be achieved by utilizing a closed-form expression as in Lemma 3. Secondly,  $\mathcal{P}_2$  is addressed by enumerating all possible offloading decision-making matrices  $\mathbf{D}^k \in \mathcal{D}$  satisfying (1) and (2). Combining with Lemmas 1 and 2, we obtain the global optimal solution

---

### Algorithm 1: Exhaustive Search Algorithm for Solving $\mathcal{P}_3$

---

- 1: Enumerate all  $(M + 1)^N$  possible offloading decision-makings satisfying (1) and (2),  $\mathcal{D} = [\mathbf{D}^1, \dots, \mathbf{D}^k, \dots, \mathbf{D}^{(M+1)^N}]$ ;
  - 2: **Initialization:** A zero vector  $\mathbf{f}$  with length  $(M + 1)^N$ .
  - 3: Set  $i := 1$ ;
  - 4: **for**  $k = 1 : (M + 1)^N$  **do**
  - 5:   Substitute  $\mathbf{D}^k$  to  $\mathcal{P}_3$ .
  - 6:   Solving  $\mathcal{P}_4$  as in Lemma 3:  $\tau^* = \left( \frac{\Omega_3}{3\Omega_4} \right)^{3/2}$ .
  - 7:   Calculate  $S(\mathbf{D}^k, P_{\max}, \eta_{\max}, \tau^*)$ .
  - 8:    $\mathbf{f}^k := S(\mathbf{D}^k, P_{\max}, \eta_{\max}, \tau^*)$ .
  - 9: **end for**
  - 10:  $\mathbf{f}^* \triangleq \arg \max_{\mathbf{D}^k}(\mathbf{f})$ .
- 

of the original optimization problem  $\mathcal{P}_1$ .

### B. Fast-Efficient Algorithm

Although  $\mathcal{P}_2$  obtain the global optimal solution by utilizing *exhaustive search* algorithm, checking all possible offloading decision-making matrices with a high number of user is a inefficient method due to its exponential complexity. In this section, we introduce a fast-efficient algorithm, which is tailored with large-scale networks. Moreover, our proposed algorithm obtains a near-global optimal solution and converges with a super-fast speed, it thus becomes a promising method for real-time applications under fast-fading channel. Inspired from Block coordinate descent (BCD) algorithm, we decompose the optimization problem  $\mathcal{P}_2$  into two sub-problems, i.e., TS ratio optimization with fixed  $\mathbf{D}^k$  and offloading decision optimization with fixed  $\tau$ . Based on the solutions of two sub-problems, a BCD method is proposed for alternatively optimizing the TS ratio and offloading decision. Specifically, we first initialize a offloading decision-making satisfying (1) and (2) and solve the TS ratio optimization problem to find the sub-optimal  $\tau$ , then  $\tau$  is updated and solve the offloading decision optimization problem to find the sub-optimal offloading decision-making. These steps are repeatedly performed until convergence. The details are presented as follows:

1) **TS Ratio Optimization:** For given value of offloading decision matrix  $\mathbf{D}^k \in \mathcal{D}$ , the optimal value of  $\tau$  can be achieved by adopting the closed-form expression in Lemma 3.

2) **Offloading Decision Optimization:** For the given TS ratio  $\tau$ , the optimal value of offloading decision  $\mathbf{D}^*$

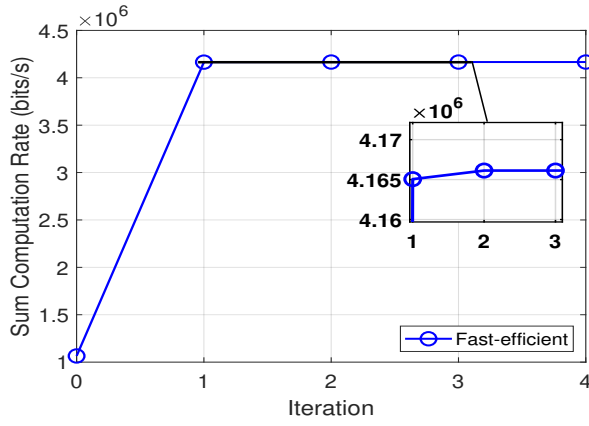


Fig. 3: Convergence behaviour of fast-efficient algorithm.

can be obtained by solving the following problem

$$\mathcal{P}_5: \max_{\mathbf{D}^k} S(\mathbf{D}^k, P_{\max}, \eta_{\max}, \tau) \quad (16a)$$

$$\text{s.t. (1), (2).} \quad (16b)$$

For ease of exploration the objective function, let us define two variables  $X$  and  $Y$  as follows

$$X \triangleq \sum_{m \in \mathcal{M}} \left( 1 - \min \left( 1, \sum_{n \in \mathcal{N}} d_{mn} \right) \right), \quad (17)$$

$$Y \triangleq \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} d_{mn}. \quad (18)$$

For given TS ratio  $\tau$ ,  $\mathcal{P}_5$  becomes a binary optimization problem. It is easy to recognize that  $X$  and  $Y$  are not dependent on  $\tau$ . Therefore,  $\mathcal{P}_5$  can be solved by allocating value "1" on the binary matrix  $\mathbf{D}^k$  satisfying (1) and (2) such that  $S(\mathbf{D}^k, P_{\max}, \eta_{\max}, \tau)$  reaches the optimal value. Notably, constraint (2) is only applied to individual columns on binary matrix  $\mathbf{D}^k$ , thus, we independently allocate value "1" for each column, wherein positions with "1" value are called by offloading positions. Specifically, the offload position at  $n$ -th column is determined via two following steps:

- Step 1: We define  $\mathcal{M}_n^{(c)}$  as a subset of all elements in  $n$ -th column such that  $R_m^L(\tau) \leq R_{mn}^O(\tau)$ . Otherwise, if  $R_m^L(\tau) > R_{mk}^O(\tau)$  for all elements in  $n$ -th column, we then have  $\mathcal{M}_n^{(c)} = \emptyset$ , this means that all BUs do not offload at  $n$ -th RB.
- Step 2: After obtaining all candidates  $\mathcal{M}_n^{(c)}$  at  $n$ -th column. The best offloading position can be selected as

$$m^* \triangleq \arg \max_{m \in \mathcal{M}_n^{(c)}} \{R_{mn}^O(\tau)\}, \forall n \in \mathcal{N}. \quad (19)$$

The pseudo-code for solving  $\mathcal{P}_2$  is summarized as in Algorithm 2.

**Complexity Analysis:** The complexity at each iteration of Algorithm 2 is  $\mathcal{O}(MN)$ .

### Algorithm 2: Fast-Efficient Algorithm for Solving $\mathcal{P}_2$

- 1: **Initialization:** Offloading decision-making  $\mathbf{D}^{(r)}$  satisfying (1) and (2), set  $r = 0$ .
- 2: **repeat**
- 3: Solve  $\mathcal{P}_4$  for the given offloading decision-making  $\mathbf{D}^{(r)}$ , and denote optimal solution as  $\tau^{(r+1)}$ .
- 4: **for**  $n = 1 : N$  **do**
- 5:     **Initialization:** Set  $\mathcal{M}_n^{(c)} = \emptyset$ ;
- 6:     **for**  $m = 1 : M$  **do**
- 7:         **if**  $R_m^L(\tau^{(r+1)}) > R_{mn}^O(\tau^{(r+1)})$  **then**
- 8:              $d_{mn} = 0$ ;
- 9:         **else**
- 10:              $d_{mn} = 1$ ;
- 11:              $\mathcal{M}_n^{(c)} = \mathcal{M}_n^{(c)} + \{m\}$ ;
- 12:         **end if**
- 13:     **end for**
- 14:      $m^* = \arg \max_{m \in \mathcal{M}_n^{(c)}} \{R_{mn}^O(\tau)\}$ ;
- 15:      $d_{m^*n} = 1$ ;
- 16:     **end for**
- 17: Obtain the optimal solution as  $\mathbf{D}^{(r+1)}$ ;
- 18:  $r \leftarrow r + 1$ ;
- 19: **until** The fractional increase of the objective value of  $\mathcal{P}_2$  is less than or equal to a small threshold  $\epsilon > 0$ .

TABLE I: Simulation Parameters

Parameter	Value
System bandwidth	1.25 MHz [44]
RB bandwidth	180 kHz [44]
The number of RBs	6 [44]
The number of BUs	4
Maximum reflection coefficient, $\eta_{\max}$	0.5
Circuit power consumption at each BU, $P_c$	0.1 mW
Path loss exponent, $\varphi$	2.8
Carrier frequency, $f_s$	915 MHz
Power budget at GW, $P_{\max}$	3 W
Antenna channel gain, $G$	4.11
The energy harvesting efficiency, $\chi$	0.7
The computing efficiency at each BU, $\beta_m$	$10^{-26}$
AWGN, $\sigma^2$	$10^{-10}$
The processor's computation speed, $n_c$	100 cycles/s
Error tolerance threshold, $\epsilon_{\text{tol}}$	$10^{-4}$
Time frame duration T	1 second

**Convergence Analysis:** First, since the optimal TS ratio  $\tau^{(r+1)}$  of optimization problem  $\mathcal{P}_2$  is achieved for the given offloading decision-making  $\mathbf{D}^{(r)}$ , we have

$$\begin{aligned} & S(\mathbf{D}^{(r)}, P_{\max}, \eta_{\max}, \tau^{(r)}) \\ & \leq S(\mathbf{D}^{(r)}, P_{\max}, \eta_{\max}, \tau^{(r+1)}). \end{aligned} \quad (20)$$

Second, for the given TS ratio  $\tau^{(r+1)}$ ,  $\mathcal{P}_2$  obtains the



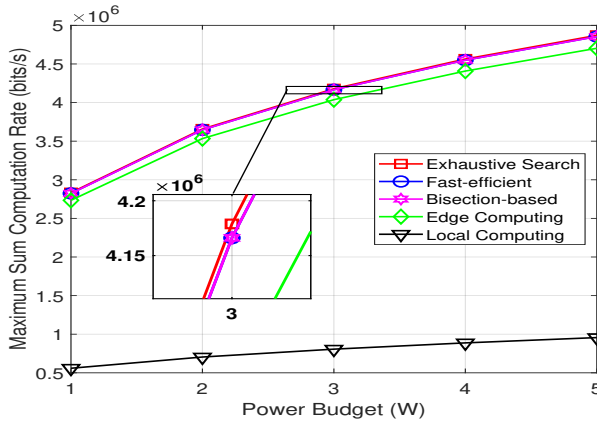


Fig. 4: The maximum computation rate versus power budget  $P_{max}$ .

optimal offloading decision-making  $\mathbf{D}^{(r+1)}$ . It is shown as

$$\begin{aligned} S(\mathbf{D}^{(r)}, P_{max}, \eta_{max}, \tau^{(r+1)}) \\ \leq S(\mathbf{D}^{(r+1)}, P_{max}, \eta_{max}, \tau^{(r+1)}). \end{aligned} \quad (21)$$

From (20) and (21), we obtain

$$\begin{aligned} S(\mathbf{D}^{(r)}, P_{max}, \eta_{max}, \tau^{(r)}) \\ \leq S(\mathbf{D}^{(r+1)}, P_{max}, \eta_{max}, \tau^{(r+1)}). \end{aligned} \quad (22)$$

Inequality (22) proves that the weighted sum computation rate  $S(\mathbf{D}, P_{max}, \eta_{max}, \tau)$  is always non-decreasing after each iteration as in Algorithm 2. Furthermore, section III shows that the optimization  $P_1$  exists a global optimal solution, this means that function  $S(\mathbf{D}, P_{max}, \eta_{max}, \tau)$  is upper bounded by a finite value. Therefore, it is guaranteed that Algorithm 2 converges.

## V. NUMERICAL RESULTS

This section provides some numerical results to present the comparative study of *exhaustive search* and *fast-efficient algorithm*. Two proposed methods are analyzed adopting Python 3.6 on a PC with an AMD Ryzen 7 3.7-4.3 GHz CPU and 16 GB memory, where simulation parameters are mainly referenced from the existing works in [37], [41], [44] and they are presented in Table I. We consider a system with four BUs and one GW, as illustrated in Fig. 1. Herein, the distance from the GW to  $m$ -th BU are uniformly distributed between 2.5 and 5.2 meters. The average channel gain between GW and  $m$ -th BU over  $n$ -th RB is assumed to follow the free-space path loss model which is given by  $h_{mn} = G \left( \frac{3.10^8}{4\pi f_s d_m} \right)^\varphi$ . At time frame  $T$ , the time-varying wireless channel gain of  $m$ -th BU over  $n$ -th RB, i.e.  $h_{mn}^T$ , is considered as a Rayleigh fading channel model  $h_{mn}^T = h_{mn} \gamma_{mn}^t$ , where  $\gamma_{mn}^t$  denotes the independent random channel fading

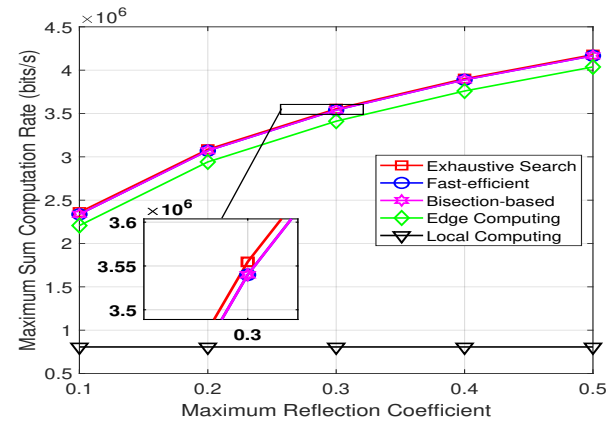


Fig. 5: The maximum computation rate versus maximum reflection coefficient  $\eta_{max}$ .

factor which is an exponential distribution with unit mean. Note that the numerical results in all the figures are averaged over 10000 different channel realizations. To show the advantages of designed algorithms, we compare our proposed methods with benchmark ones. Specifically, two benchmark schemes are described as follows:

- **Exhaustive search (ES) algorithm:** it considers all offloading decision-making matrices, hence, its complexity is proportional to the network size [45].
- **Bisection-based algorithm (BA):** This method is implemented similar to FEA but it adopts the Bisection method to solve the TS ratio optimization problem.
- **Edge computing (EC):** Each BU has to compete with others to maximize its offloading tasks to the edge server using OFDMA-based system without considering offloading decision-making as in proposed algorithms.
- **Local computing (LC):** All BUs locally execute their computation tasks without offloading to the edge server.

The convergence behaviour of the fast-efficient algorithm is illustrated in Fig. 3, where the error tolerance is given by  $\epsilon_{tol} = 10^{-4}$ . As can be observed from Fig. 3, the sum computation rate increases with the algorithm iterations. This can be explained by the fact that the TS ratio and the offloading decision-making are updated to better values after each iteration. Particularly, the fast-efficient algorithm converges to the saturation value only after 2 iterations. This highlights the superiority of our proposed fast-efficient method.

Fig. 4 depicts the maximum sum computation rate versus the power budget at GW. As illustrated, the maximum sum computation rate is enhanced by increasing the power budget of GW. This is due to the fact that the offloading and local computation rate is proportional to the transmit



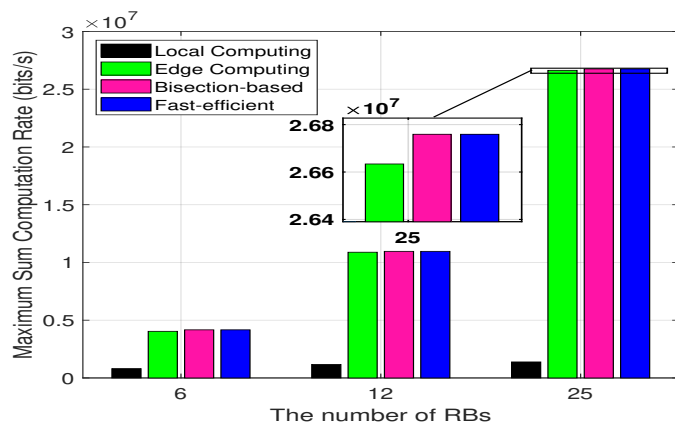


Fig. 6: The maximum sum computation rate versus the number of RBs.

power of GW as in Eqs. (3) and (7). Moreover, while the LC scheme always executes computation tasks locally and does not take offloading decision into account, it thus achieved the worst performance due to the limited computing resources of BUs, e.g., processing capability, energy budget. This also explains that the maximum computation rate of LC is only slightly increased with a higher value of power budget as compared with other methods. On the other hand, the EC scheme achieves better performance compared to that of LC method, this shows the benefits of edge server to overcome the limitation of low-power IoT users. Nevertheless, the EC always try to offload their tasks to the MEC server without considering local computing which is not a wise strategy due to the limited resources at the edge server, i.e., limited bandwidth and computation capability, thus MEC server can not always serve all users at the same time during peak hours. This can explain the reason why the performance of EC is always lower than that of fast-efficient and exhaustive search algorithms. Especially, the exhaustive search and fast-efficient methods are proposed to balance between offloading and local computing decision. Thus, they obtain the best results and have almost the identical performance with the gap is less than 0.3 % which shows the supremacy of our designed schemes.

In Fig. 5, we study the influence of maximum reflection coefficient on the maximum sum computation rate. Most of properties in Fig. 4 remains unchanged except for the local computing scheme. It is observed that the sum computation rate of LC is unaltered with different values of reflection coefficient. This is expected since the reflection coefficient does not impact on the local computation rate which is shown in Eq. (3). In contrast to that, since the reflection coefficient is proportional to

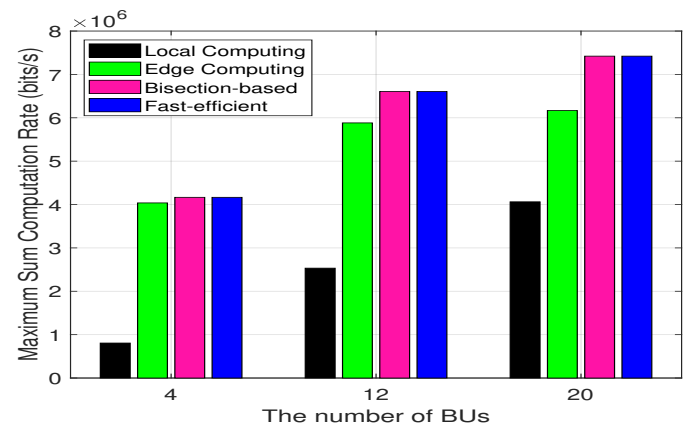


Fig. 7: The maximum sum computation rate versus the number of BUs.

the offloading rate as shown in Eq. (7). Therefore, the performance of exhaustive search, fast-efficient, and EC methods is significantly improved when we increase the maximum value of reflection coefficient.

In Figs. 6 and 7, we evaluate the maximum sum computation rate versus the number of RBs and the number of BUs, respectively. As we know, the exhaustive search's complexity tends to grow exponentially as the problem size increases. Thus, we do not study exhaustive search method in these figures when the number of RBs and the number of BUs increase up to 25 and 20, respectively.

In Figs. 6, we see that the sum computation rate of EC and fast-efficient methods are drastically increased corresponding to a higher number of RBs while it only slightly changes with respect to LC method. This is expected since a higher number of RBs leads to a larger allocated bandwidth for communication between GW and BUs. On the one hand, the RB's bandwidth does not have an effect on the local computing rate as in Eq. (3), it thus does not impact on LC scheme. On the other hand, RB's bandwidth has an influence directly on the offloading rate as in Eq. (7), thus it has a positive impact on EC and fast-efficient schemes. Furthermore, one more interesting thing is observed that when the number of RBs is large enough, the performance of EC and fast-efficient is no discernible difference. This is because the offloading rate is dominant as compared with local computing rate.

In Fig. 7, we compare the performance of different algorithms when the number of BUs varies from 4 to 20. The preeminence of the proposed fast-efficient algorithm is more perspicuous as the number of BUs is larger. As in practical scenario, with a limited number of RBs but the number of IoT users or the traffic demand is dramatically increasing, it leads to bottleneck problems in which the

TABLE II: Complexity analysis for different schemes

Scheme	Complexity
ES	$\mathcal{O}(M^N)$
BA	$\mathcal{O}(\Upsilon_1 (\log_2(1/\varepsilon) + MN))$
FEA	$\mathcal{O}(\Upsilon_2 MN)$
EC	$\mathcal{O}(\Upsilon_3 MN)$

TABLE III: Comparisons of CPU execution latency (second)

The number of BUs	EC	FEA	BA	ES
4	0.002	0.004	0.01	1.5
50	0.019	0.029	0.037	x
100	0.037	0.056	0.072	x

GW is unable to support all users at the same time. In these cases, the simple designs such as LC and EC method are clearly expose their disadvantages. To overcome these issues, the proposed fast-efficient method finds an optimal trade-off between the offloading and the local computing rate to keep the sum computation rate to be as best as possible.

The results in Figs. 4-7 show that there is no discernible difference of FEA and BA w.r.t. sum computation rate, it is because their solutions are both based on the BCD method. However, the TS ratio optimization problem of FEA is solved by deriving a closed-form expression while the Bisection method is adopted in BA. It leads to the execution time of FEA is better than that of BA. In the case of achieving closed-form expression is troublesome, the BA method may become a better solution.

In Table II, we provide the complexity analysis of FEA and other algorithms, i.e., ES, BA, and EC. Here,  $\Upsilon_1$ ,  $\Upsilon_2$ , and  $\Upsilon_3$  denote the number of iterations needed for BCD method converges,  $\varepsilon$  is the stopping criterion of bisection algorithm that used to solve the subproblem  $\mathcal{P}_4$ , i.e.,  $\varepsilon = 10^{-4}$ .

Finally, we study a new metric called CPU execution time as a function of the number of BUs shown in Table III. Specifically, CPU execution time is defined as the running time of algorithms (i.e., EC, FEA, BA, and ES), to find the offloading decision for BUs. For a fair comparison, these algorithms are alternately optimizing TS ratio and offloading decision. It can be seen that the exhaustive is only feasible in a small-scale network, i.e., a system consists of 4 BUs and 6 RBs. Moreover, it imposes a largest latency as compared with others as it tries all possible decision matrices, while The EC and fast-efficient consumes much less time. From practical aspects, the fast-efficient algorithm is preferred as it has low complexity and obtain near-global optimal solution. The EC algorithm is fastest method but it significantly

sacrifices the performance. The LC does not consume time for calculating offloading decision since it operate locally, but it has worst performance compared to other methods, as shown in Figs. 4, 5, 6, and 7.

## VI. CONCLUSION AND DISCUSSION

This paper has studied a sum-rate maximization in backscatter-assisted OFDMA-based WPMEC for IoT networks with binary offloading decision. In this context, we have formulated a problem of maximizing the total computation rate via joint optimization of transmit power of the GW, backscatter coefficient, TS ratio, and computing mode (i.e., local or offloading). Since the formulated problem is of MINLP type which is NP-hard, we have simplified the original problem into a tractable form by obtaining the optimal values of transmit power of the GW and backscatter coefficient. Then, we have proposed FEA to tackle the difficulty of offloading decisions based on the BCD method. Extensive numerical results have shown that the proposed method is able to achieve a near-globally optimal solution and dramatically outperform other benchmark schemes. In practical scenarios, the FEA is the best option due to its low complexity and high performance which is suitable for ultra-low latency applications in IoT networks such as virtual reality, augmented reality, remote surgery, and mission-critical applications. The results obtained in this work open future research directions of backscatter-assisted WPMEC in IoT networks. One problem of interest is to consider a GW with multiple antennas, which enhances the network performance, uses relays in blockage/peak hours, or uses a power splitting method instead of time-switching. Another promising problem is to jointly optimize the power allocation and other metrics in a mobility system model or investigate the intelligent reflecting surface-based MEC.

### APPENDIX A: PROOF OF LEMMA 1 AND LEMMA 2

Given  $\mathbf{D}$  and  $\tau$ , the partial derivatives of  $S$  w.r.t.  $\eta$  is calculated as

$$\frac{\partial S}{\partial \eta} = \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} d_{mn} \frac{\Omega_1 \Omega_2}{(1 + \Omega_2 \eta) \ln 2}, 0 \leq \eta \leq \eta_{max}. \quad (\text{A.1})$$

where  $\Omega_1 \triangleq B_n(1 - \tau)$ ,  $\Omega_2 \triangleq \frac{P_w |h_{mn}|^2 \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\sigma_w^2}$ .

Since  $\frac{\partial S}{\partial \eta} \geq 0$  with  $0 \leq \eta \leq \eta_{max}$ , thus we can conclude that  $S$  is a non-decreasing function w.r.t.  $\eta$  and it can obtain the maximum value at  $\eta = \eta_{max}$ , which proves the Lemma 1.

Similarly, it is also proved that  $S$  is a non-decreasing function w.r.t.  $P_w$  and its maximum value obtains at  $P = P_{max}$ . Thus, Lemma 2 is proved.

#### APPENDIX B: PROOF OF LEMMA 3

The objective function  $S(\tau)$ ,  $\zeta \leq \tau^* \leq 1$ , can be rewritten as

$$S(\tau) \triangleq \Omega_3 \tau^{1/3} + \Omega_4(1 - \tau), \quad (\text{C.1})$$

with

$$\Omega_3 \triangleq \sum_{m \in \mathcal{M}} \left( 1 - \min \left( 1, \sum_{n \in \mathcal{N}} d_{mn} \right) \right) \times \left( \frac{\chi \tau P_w \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\beta_m} \right)^{1/3} \frac{1}{n_c}, \quad (\text{C.2})$$

$$\Omega_4 \triangleq \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{N}} d_{mn} B_n \log_2 \left( 1 + \frac{\eta P_w |h_{mn}|^2 \sum_{n \in \mathcal{N}} |h_{mn}|^2}{\sigma_w^2} \right). \quad (\text{C.3})$$

We consider two special cases:

- **Case 1:**  $d_{mn} = 1, \forall m \in \mathcal{M}, n \in \mathcal{N}$ . This means that  $\Omega_3 = 0$ . The optimal value of function  $S(\tau)$  is  $\tau^* = \zeta$ .
- **Case 2:**  $d_{mn} = 0, \forall m \in \mathcal{M}, n \in \mathcal{N}$ . This means that  $\Omega_4 = 0$ . The optimal value of function  $S(\tau)$  is  $\tau^* = 1$ .

More generally, let us define a value  $\tau^*$  satisfy [46]

$$\frac{\partial S(\tau^*)}{\partial \tau^*} = 0, \quad (\text{C.4})$$

$$\frac{\partial^2 S(\tau^*)}{\partial^2 \tau^*} < 0. \quad (\text{C.5})$$

Firstly, we calculate the Eq. (C.4) such as

$$\Omega_3 \frac{\tau^{-2/3}}{3} - \Omega_4 = 0 \iff \tau^* = \left( \frac{\Omega_3}{3\Omega_4} \right)^{3/2}. \quad (\text{C.6})$$

Secondly, we have  $\frac{\partial^2 S(\tau^*)}{\partial^2 \tau^*} = \frac{-2\Omega_3 \tau^{-5/3}}{9}$ , then by substituting  $\tau^*$  we can justify the Eq. (C.5).

If  $\zeta < \tau^* < 1$ , then we can conclude that  $\tau^* = \left( \frac{\Omega_3}{3\Omega_4} \right)^{3/2}$  obtained in (C.6) is a strict optimal solution [46]. In case of  $0 \leq \tau^* \leq \zeta$ , the optimal value in (C.6) is  $\tau^* = \zeta$ . Meanwhile, if  $1 \leq \tau^*$ , then the optimal in (C.6) is  $\tau^* = 1$ .

#### REFERENCES

- [1] D. Ma, G. Lan, M. Hassan, W. Hu, and S. K. Das, "Sensing, computing, and communications for energy harvesting IoTs: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1222–1250, 2019.
- [2] Ericsson, "Ericsson mobility report: November 2019," 2019.
- [3] D. H. Tran, V. D. Nguyen, S. Gautam, S. Chatzinotas, T. X. Vu, and B. Ottersten, "UAV relay-assisted emergency communications in IoT networks: Resource allocation and trajectory optimization," preprint arXiv:2008.00218.
- [4] T. X. Tran and D. Pompili, "Joint task offloading and resource allocation for multi-server mobile-edge computing networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 856–868, 2018.
- [5] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches," *Wirel. Commun. Mob. Com.*, vol. 13, no. 18, pp. 1587–1611, 2013.
- [6] N. Hassan, S. Gillani, E. Ahmed, I. Yaqoob, and M. Imran, "The role of edge computing in internet of things," *IEEE Commun. Mag.*, vol. 56, no. 11, pp. 110–115, 2018.
- [7] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 3, pp. 1628–1656, 2017.
- [8] J. Zhao, Q. Li, Y. Gong, and K. Zhang, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7944–7956, 2019.
- [9] T. Bai, C. Pan, Y. Deng, M. ElKashlan, A. Nallanathan, and L. Hanzo, "Latency Minimization for Intelligent Reflecting Surface Aided Mobile Edge Computing," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2666–2682, 2020.
- [10] H. A. Alameddine, S. Sharafeddine, S. Sebbah, S. Ayoubi, and C. Assi, "Dynamic task offloading and scheduling for low-latency iot services in multi-access edge computing," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 3, pp. 668–682, 2019.
- [11] S. Hu and G. Li, "Dynamic request scheduling optimization in mobile edge computing for IoT applications," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1426–1437, 2019.
- [12] A. Yousafzai, I. Yaqoob, M. Imran, A. Gani, and R. M. Noor, "Process migration-based computational offloading framework for IoT-supported mobile edge/cloud computing," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4171–4182, 2019.
- [13] Q. Cui, J. Zhang, X. Zhang, K.-C. Chen, X. Tao, and P. Zhang, "Online anticipatory proactive network association to mobile edge computing for iot," *IEEE Trans. Wireless Commun.*, 2020.
- [14] T.-D. Nguyen, E.-N. Huh, and M. Jo, "Decentralized and revised content-centric networking-based service deployment and discovery platform in mobile edge computing for IoT devices," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4162–4175, 2018.
- [15] S. Gong, Y. Xie, J. Xu, D. Niyato, and Y.-C. Liang, "Deep reinforcement learning for backscatter-aided data offloading in mobile edge computing," *IEEE Netw.*, pp. 1–8, 2020.
- [16] T. D. Hieu, T. T. Duy, and B.-S. Kim, "Performance enhancement for multihop harvest-to-transmit WSNs with path-selection methods in presence of eavesdroppers and hardware noises," *IEEE Sensors J.*, vol. 18, no. 12, pp. 5173–5186, 2018.
- [17] D. Altinel and G. K. Kurt, "Modeling of multiple energy sources for hybrid energy harvesting IoT systems," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10846–10854, 2019.
- [18] W. Wang, J. Tang, N. Zhao, X. Liu, X. Y. Zhang, Y. Chen, and Y. Qian, "Joint Precoding Optimization for Secure SWIPT in UAV-Aided NOMA Networks," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5028–5040, 2020.
- [19] F. Wang, J. Xu, and S. Cui, "Optimal energy allocation and task offloading policy for wireless powered mobile edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2443–2459, 2020.
- [20] L. Ji and S. Guo, "Energy-efficient cooperative resource allocation in wireless powered mobile edge computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4744–4754, 2018.

- [21] S. Bi and Y. J. Zhang, "Computation rate maximization for wireless powered mobile-edge computing with binary computation offloading," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4177–4190, 2018.
- [22] S. Gautam, T. D. Hieu, S. Chatzinotas, and B. Ottersten, "Hybrid backscatter and relaying scheme for 6G greencom IoT networks with SWIPT," *10.36227/techrxiv.12893750.v1*, 2020.
- [23] H. Tran-Dinh, S. Gautam, S. Chatzinotas, and B. Ottersten, "Throughput Maximization for Wireless Communication systems with Backscatter- and Cache-assisted UAV Technology," in *Arxiv*, preprint arXiv:2011.07955.
- [24] B. Kellogg, A. Parks, S. Gollakota, J. R. Smith, and D. Wetherall, "Wi-fi backscatter: Internet connectivity for RF-powered devices," in *Proc. 2014 ACM conf. SIGCOMM*, 2014, pp. 607–618.
- [25] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2014.
- [26] Y. e. a. Zou, "Backscatter-aided hybrid data offloading for wireless powered edge sensor networks," in *Proc. IEEE GLOBECOM 2019*, 2019, pp. 1–6.
- [27] P. X. Nguyen, T. H. Pham, T. Hoang, and O.-S. Shin, "An efficient spectral leakage filtering for IEEE 802.11 af in TV white space," in *Proc. 2018 2nd International Conf. Recent Advances in Sig. Proces., Telecom. & Comp. (SigTelCom)*. IEEE, 2018, pp. 219–223.
- [28] S. Sesia, I. Toufik, and M. Baker, *LTE-the UMTS Long Term Evolution: From Theory to Practice*. John Wiley & Sons, 2011.
- [29] Y. Wu, Y. Wang, F. Zhou, and R. Q. Hu, "Computation efficiency maximization in ofdma-based mobile edge computing networks," *IEEE Comm. Lett.*, vol. 24, no. 1, pp. 159–163, 2019.
- [30] M. Masoudi and C. Cavdar, "Device vs edge computing for mobile services: Delay-aware decision making to minimize power consumption," *IEEE Trans. Mobile. Comput.*, June 2020.
- [31] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.
- [32] P. X. Nguyen, H. V. Nguyen, V. Nguyen, and O. Shin, "UAV-enabled jamming noise for achieving secure communications in cognitive radio networks," in *Proc. IEEE Consumer Commun. & Network. Conf. (CCNC)*, Jan 2019, pp. 1–6.
- [33] P. X. Nguyen, V.-D. Nguyen, H. V. Nguyen, and O.-S. Shin, "UAV-assisted secure communications in terrestrial cognitive radio networks: Joint power control and 3D trajectory optimization," *arXiv preprint arXiv:2003.09677*, 2020.
- [34] N. Zhao, X. Pang, Z. Li, Y. Chen, F. Li, Z. Ding, and M. Alouini, "Joint Trajectory and Precoding Optimization for UAV-Assisted NOMA Networks," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3723–3735, 2019.
- [35] F. W. et al., "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," in *Proc. 2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.
- [36] F. W. et al., "Joint offloading and computing optimization in wireless powered mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1784–1797, 2018.
- [37] L. Huang, S. Bi, and Y. J. Zhang, "Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks," *IEEE Trans. Mobile Comput.*, pp. 1–1, 2019.
- [38] S. Guo, B. Xiao, Y. Yang, and Y. Yang, "Energy-efficient dynamic offloading and resource scheduling in mobile cloud computing," in *Proc. IEEE INFOCOM 2016*, 2016, pp. 1–9.
- [39] S. Guo, J. Liu, Y. Yang, B. Xiao, and Z. Li, "Energy-efficient dynamic computation offloading and cooperative task scheduling in mobile cloud computing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 2, pp. 319–333, 2018.
- [40] B. Lyu, C. You, Z. Yang, and G. Gui, "The optimal control policy for RF-powered backscatter communication networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2804–2808, 2017.
- [41] S. Xiao, H. Guo, and Y.-C. Liang, "Resource allocation for full-duplex-enabled cognitive backscatter networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3222–3235, 2019.
- [42] X. Kang, Y.-C. Liang, and J. Yang, "Riding on the primary: A new spectrum sharing paradigm for wireless-powered iot devices," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 6335–6347, 2018.
- [43] J. D. Griffin and G. D. Durgin, "Complete link budgets for backscatter-radio and RFID systems," *IEEE Antenn. Propag. Mag.*, vol. 51, no. 2, pp. 11–25, 2009.
- [44] O. Liberg, M. Sundberg, E. Wang, J. Bergman, and J. Sachs, *Cellular Internet of Things: Technologies, Standards, and Performance*. Academic Press, 2017.
- [45] D.-H. Tran, T. X. Vu, S. Chatzinotas, S. ShahbazPanahi, and B. Ottersten, "Coarse trajectory design for energy minimization in uav-enabled," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 9483–9496, 2020.
- [46] Y. Wotao, "Math 273a: Optimization basic concepts," [www.math.ucla.edu/wotaoyin/math273a/slides/Lec2basicsofoptimization2732015f.pdf](http://www.math.ucla.edu/wotaoyin/math273a/slides/Lec2basicsofoptimization2732015f.pdf).



**Phu X. Nguyen** was born and grew up in Dong Thap, Vietnam (1994). He received the B.E. degree in Electronics and Telecommunication Engineering Department from Ho Chi Minh City University of Technology, Vietnam, in 2017. In 2019, he received a Master's degree in Electronic Engineering from Soongsil University, South Korea. He is currently a lecturer at the Department of Computer Fundamentals, FPT University, Ho Chi Minh City, Vietnam. His research interests include machine learning, optimization, quantum computing and their applications in IoTs, wireless networks, and computer vision.



**Dinh-Hieu Tran** (S'20) was born and grew up in Gia Lai, Vietnam (1989). He received the B.E. degree in Electronics and Telecommunication Engineering Department from Ho Chi Minh City University of Technology, Vietnam, in 2012. In 2017, he finished the M.Sc degree in Electronics and Computer Engineering from Hongik University (Hons.), South Korea. He is currently pursuing the Ph.D. degree at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, under the supervision of Prof. Symeon Chatzinotas and Prof. Björn Ottersten. His research interests include UAVs, IoTs, Mobile Edge Computing, Caching, Backscatter, B5G for wireless communication networks. He was a recipient of the IS3C 2016 best paper award.



**Oluwakayode Onireti** (S'11, S'13) received the B.Eng. degree (Hons.) in electrical engineering from the University of Ilorin, Ilorin, Nigeria, in 2005, and the M.Sc. degree (Hons.) in mobile and satellite communications, and the Ph.D. degree in electronics engineering from the University of Surrey, Guildford, U.K., in 2009 and 2012, respectively. He is currently a lecturer at the University of Glasgow, U.K. He has been actively involved in projects such as ROCKET, EARTH, Greencom, QSON, DARE, and Energy proportional ENodeB for LTE-Advanced and Beyond. His main research interests include self-organizing cellular networks, energy efficient networks, wireless blockchain networks, millimeter wave communications, and cooperative communications.



**Phu Tran Tin** was born in Khanh Hoa, Viet Nam, in 1979. He received the Bachelor's degree (2002) and Master's degree (2008) from Ho Chi Minh City University of Science. He is currently a lecturer at the Faculty of Electronics Technology (FET), Industrial University of Ho Chi Minh City. In 2019, he received the Ph.D. degree in Faculty of Electrical Engineering and Computer Science, VSB – Technical University of Ostrava, Czech Republic. His major research interests are wireless communication in 5G, energy harvesting, performance of cognitive radio, physical layer security and NOMA.



**Symeon Chatzinotas**, (S'06-M'09-SM'13) is currently Full Professor / Chief Scientist I in Satellite Communications and Head of the SIG-COM Research Group at SnT, University of Luxembourg. He is coordinating the research activities on communications and networking, acting as a PI for more than 20 projects and main representative for 3GPP, ETSI, DVB. In the past, he has been a Visiting Professor at the University of Parma, Italy, lecturing on "5G Wireless Networks". He was involved in numerous RD projects for NCSR Demokritos, CERTH Hellas and CCSR, University of Surrey. He was the co-recipient of the 2014 IEEE Distinguished Contributions to Satellite Communications Award and Best Paper Awards at EURASIP JWCN, CROWNCOM, ICSSC. He has (co-)authored more than 450 technical papers in refereed international journals, conferences and scientific books. He is currently in the editorial board of the IEEE Transactions on Communications, IEEE Open Journal of Vehicular Technology and the International Journal of Satellite Communications and Networking.



**Nguyen Quang Sang** received the BE degree (2010) and M.E. degree (2013) in Ho Chi Minh City University of Transport and Ho Chi Minh City University of Technology, Vietnam, respectively. In 2017, he received Ph. D degree in Electrical Engineering from University of Ulsan, South Korea. Since January – June 2017, he was a post-doc research fellow at Queen's University Belfast. Since July 2017, he has been a lecturer at Duy Tan University, Vietnam. His major research interests are: Cooperative communication, cognitive radio network, physical layer security, energy harvesting, non-orthogonal multiple access,...



**H. Vincent Poor** (S'72, M'77, SM'82, F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Michael Henry Strater University Professor of Electrical Engineering. From 2006 until 2016 he also served as Dean of Princeton's School of Engineering and Applied Science. Dr. Poor's research interests are in the areas of information theory, machine learning and network science, and their applications in wireless networks, energy systems, and related fields. Among his publications in these areas is the forthcoming book *Machine Learning and Wireless Communications* (Cambridge University Press, 2021).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal and a D.Eng. honoris causa from the University of Waterloo, awarded in 2019.