Energy-Efficient Task Offloading in RIS-Aided HetNets With Wireless Backhaul

Li Zhang, Jinping Niu, Yiyao Wang, Gaojie Chen, *Senior Member, IEEE*, Tao Gu, *Senior Member, IEEE*, and Yanyan Li

***Abstract*— This letter exploits reconfigurable intelligent surface (RIS) in the heterogeneous networks (HetNets) with wireless backhaul to improve the task computing energy efficiency. To minimize the overall energy consumption of all users, we formulate the problem as a combinatorial optimization problem by jointly optimizing the transmit power of users and small cell base stations, computational capacity of users and base stations, task offloading rate of users, and passive beamforming matrix at RIS. The problem is NP-hard and difficult to solve directly. We decompose the problem into two sub-problems of energy consumption optimization at users and base stations, respectively, taking into account of wireless backhaul capacity. The two sub-problems are solved iteratively to find the final solution. Simulation results demonstrate that the proposed scheme can greatly reduce the energy consumption compared with the baselines.**

***Index Terms*— Heterogeneous networks (HetNets), task offloading, reconfigurable intelligent surface (RIS).**

I. Introduction

HETEROGENEOUS network (HetNet) with wireless backhaul is a key technology for new generation of wireless communications [1]. In the HetNet, small base stations (SBSs) are implemented within the traditional macrocell to form small cells, where small cells and macrocell can reuse the rare wireless spectrum resources, thus improving the system spectrum efficiency. However, the computing performance of mobile edge computing (MEC) in wireless backhaul HetNets is severely limited by the wireless interference and backhaul capacity [2]. Reconfigurable intelligent surface (RIS) is emerging recently as an efficient way to significantly improve the performance of data transmission in wireless networks [3]. It uses software control to re-modulate incident signals into desired signals with low cost and low power consumption, and has the potential in improving the computing performance of task offloading in MEC HetNets with wireless backhaul.

Several studies on task offloading have been done in networks leveraging RIS or wireless backhaul HetNets. Specifically, the authors in [3] analyzed the auxiliary role of RIS to MEC in shortening task processing delay and improving task completion rate, etc., through potential application cases. The authors in [4] proposed a scheme to enhance user's energy consumption and latency performance in a RIS-assisted MEC system. However, these studies cannot be applied to wireless backhaul HetNets, where the additional interference introduced by wireless backhaul link and the backhaul capacity restriction issue cannot be ignored. The energy saving of multi-user task offloading in HetNets was studied in [5]. However, it has a strong assumption that an ideal wired backhaul is used in HetNets. To release this assumption, the authors in [2] studied task offloading via MEC in wireless backhaul HetNet with only one small cell. However, the computing capability of SBS was not involved in their system and RIS was not exploited, either. Different from aforementioned works, this letter exploits RIS in HetNets with wireless backhaul to improve the energy consumption performance for partial offloading. To minimize the overall energy consumption for users' task computation, we formulate the problem by jointly optimizing computing resources of macrocell users (MUEs), small cell users (SUEs), SBSs and macarocell base station (MBS), and the power control scheme and task offloading scheme of users and SBSs. The difference between the aforementioned works and the proposed scheme is summarized in Table I.

TABLE I

Brief Summary of the Related Work

Research work	RIS	HetNets	Wireless Backhaul	Computing capability of SBS
[2]	×	✓	✓	×
[3]	✓	×	×	×
[4]	✓	×	×	×
[5]	×	✓	×	×
The proposed work	✓	✓	✓	✓

The main contributions of this letter can be summarized as: (1) we exploit the RIS in HetNet with wireless backhaul to facilitate energy-efficient task offloading for users; (2) The energy minimization problem for task offloading with the consideration of computing and communication resources optimization is non-convex. We develop an efficient two-step scheme to optimize the energy consumption at users and SBSs individually, and iterate the step to reach an effective solution. We further decompose each sub-problem by jointly optimizing computational capacity, task offloading rate, passive beamforming at RIS, transmit power and computation resource. We also provide the complexity analysis of the proposed scheme; (3) we conduct simulation studies and results show that the proposed scheme achieves up to 21% energy saving compared with the baseline.

II. System Model

We consider a RIS-aided MEC HetNet as in Fig. 1, where one macrocell is overlaid with M small cells. In macrocell, there is one MBS configured with Q antennas and K MUEs.

Manuscript received 4 July 2023; accepted 20 July 2023. Date of publication 27 July 2023; date of current version 12 September 2023. This work was supported in part by the National Natural Science Foundation of China (Grant No. 62072373, 61901372, 62271392), by the International Cooperation Foundation of Shaanxi Province (Grant No. 2019KW-012), by the Natural Science Research Program of Shaanxi Province (Grant No. 2020JQ-599), by China Postdoctoral Science Foundation (Grant No. 2020M683541), and by the U.K. Engineering and Physical Sciences Research Council Grant EP/X013162/1. The associate editor coordinating the review of this letter and approving it for publication was A. Papazafeiropoulos. *(Corresponding author: Jinping Niu.)*

Li Zhang, Jinping Niu, Yiyao Wang, and Yanyan Li are with the School of Information Science and Technology, Northwest University, Xi'an 710069, China (e-mail: zhangli1101@stumail.nwu.edu.cn; jinpingniu@nwu.edu.cn; wangyiyao@stumail.nwu.edu.cn; liyanyan_xd@163.com).

Gaojie Chen is with the Institute for Communication Systems (ICS), 5GIC and 6GIC, University of Surrey, Guildford, GU2 7XH Surrey, U.K. (e-mail: gaojie.chen@surrey.ac.uk).

Tao Gu is with the School of Computing, Macquarie University, Sydney, NSW 2109, Australia (e-mail: tao.gu@mq.edu.au).

Digital Object Identifier 10.1109/LCOMM.2023.3299380

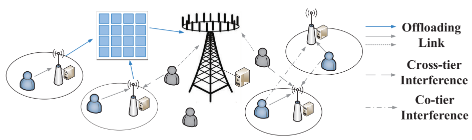


Fig. 1. Schematic diagram of a RIS-aided MEC HetNet.

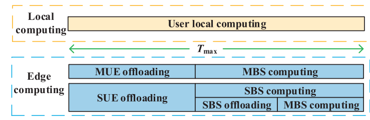


Fig. 2. A illustration of the computing time.

In each small cell, there is one SUE and one SBS.[1] Each of the SBSs and users is equipped with one antenna. The SBS helps to forward the SUE's data to the MBS when necessary. There is a RIS with N reflective elements in HetNet to improve data transmission in wireless backhaul links, which are the transmission links between the SBSs and the MBS. Each base station (BS), including the MBS and the SBS, is equipped with a server as a computing node to provide computing resources for its served users. In HetNet, a SBS can communicate with its associated SUE and the MBS simultaneously, introducing self-interference at the SBS, which limits system performance. To mitigate self-interference at SBSs, equally divided orthogonal frequency bands are exploited by users and backhaul links, which occupy half of the bandwidth, respectively. Besides, zero-force beamforming (ZFBF) is performed at the MBS to mitigate inter-user interference. In HetNet, a user can perform task computing locally or remotely by its associated BS via offloading partial of the user's task bits to the associated BS. For a SUE, its task bits can be further computed at the MBS, which is realized by offloading the SUE's task bits from its associated SBS to the MBS, after the SBS receiving all offloading task bits of the SUE. A BS start to perform task computing for a user after it receives all the user's offloaded task bits as shown in Fig. 2.

Denote $\mathcal{M}=\{1,2,\ldots,M\}$ to be the index set of SBSs and $\mathcal{M}_{\text{all}}=\{0\}\cup\mathcal{M}$ to be the index set of all BSs, where index 0 represents the MBS. Assume $\mathcal{U}_{\text{all}}=\cup_{m=0}^{M}\mathcal{U}_{\text{m}}$ is the index set of all users, with $\mathcal{U}_0=\{1,2,\ldots,K\}$ being the index of MUEs and $\mathcal{U}_{\text{m}}=\{1\}$, $m\neq 0$, being the index of the SUE in the mth small cell. Denote the uplink transmission data rates for MUE k, SBS m, and SUE $u\in\mathcal{U}_m$, $m\neq 0$, as $R_{k,0}=\frac{B}{2}\log_2\left(1+\frac{p_{k,0}}{\sigma^2\|\boldsymbol{w}_{1,k}\|^2}\right)$, $R_m=\frac{B}{2}\log_2\left(1+\frac{p_m}{\sigma^2\|\boldsymbol{w}_{2,m}\|^2}\right)$, and $R_{u,m}=\frac{B}{2}\log_2\left(1+\gamma_{u,m}\right)$, respectively. B is the system bandwidth, σ^2 is the noise power, $\boldsymbol{w}_{1,k}$ is the ZFBF vector for MUE k, and $p_{k,0}$ is the transmit power of the kth MUE. $\boldsymbol{w}_{2,m}$ is the ZFBF vector for SBS m, which is a function of the RIS passive beamforming matrix, and p_m is the transmit power of SBS m. $\gamma_{u,m}=\frac{p_{u,m}g_{u,m}}{I_0+I_1+\sigma^2}$ is the signal-to-interference-plus-noise ratio for the SUE, where $p_{u,m}$ is transmit power of the SUE, $g_{u,m}$ is the channel gain between the SUE and its associated SBS, $I_0=\sum_{k=1}^{K}p_{k,0}g_{k,m,0}$ is the cross-layer interference from MUEs, $g_{k,m,0}$ is the channel gain from MUE k to SBS m, $I_1=\sum_{j=1,j\neq m}^{M}p_{1,j}g_{j,m,1}$ is the co-layer interference from SUEs in other small cells, and $g_{j,m,1}$ is the channel gain from the SUE in the jth small cell to SBS m

Assume that each user has a latency-bound task. Denote that $L_{u,m}$ (in bits) is the task input data size for user $u\in\mathcal{U}_m$, $m\in\mathcal{M}_{\text{all}}$, $\rho_{u,m}$ (in cycles/bit) is the number of CPU cycles required to complete a one-bit task for the user, $T_{\max}$ (in seconds) is the maximum allowed time latency to finish the task computing for each user. Then the total number of CPU cycles required to complete the user's task is $C_{u,m}=L_{u,m}\rho_{u,m}$. Assume $v_{u,m}^{\text{I}}\in[0,1]$ is the proportion of task bits user u offloaded to its associated BS, called task offloading rate here. The user can either perform its $(1-v_{u,m}^{\text{I}})L_{u,m}$ bits task locally or offload its $v_{u,m}^{\text{I}}L_{u,m}$ bits task to the associated BS for computing. Let $f_{u,m}$ be the computation capability of user $u\in\mathcal{U}_m$, $m\in\mathcal{M}_{\text{all}}$. Assume that v_m^{II} is the task offloading rate at SBS m and f_m is the computation capability of the SBS.

For local computing of user u, its task execution latency and energy consumption can be expressed as $T_{u,m}^{\text{loc}}=\frac{(1-v_{u,m}^{\text{I}})C_{u,m}}{f_{u,m}}$ and $E_{u,m}^{\text{loc}}=P_{u,m}^{\text{loc}}T_{u,m}^{\text{loc}}=\delta(1-v_{u,m}^{\text{I}})C_{u,m}(f_{u,m})^2$, respectively, where $P_{u,m}^{\text{loc}}=\delta(f_{u,m})^3$ is the user's computing power consumption with δ being the power consumption coefficient that depends on the chip hardware architecture [7]. For edge computing in HetNets, two stages are involved. The first stage relates to task offloading from users to their associated BSs, where all users are involved. The second stage is relevant to task offloading from SBSs to the MBS, where only SUEs are involved.

For edge computing corresponding to the MUEs, the overall time consumption to upload and compute the $v_{u,m}^{\text{I}}$ proportion of task bits of MUE $u\in\mathcal{U}_0$ can be computed by $T_{u,0}^{\text{mec}}=T_{u,0}^{\text{off}}+T_{u,0}^{\text{cp1}}$, and the corresponding overall energy consumption without considering the MBS's energy consumption can be given by $E_{u,0}^{\text{mec}}=E_{u,0}^{\text{off}}$. And $T_{u,m}^{\text{off}}=\frac{v_{u,m}^{\text{I}}L_{u,m}}{R_{u,m}}$ and $E_{u,m}^{\text{off}}=p_{u,m}T_{u,m}^{\text{off}}$ are the uplink transmission time for the user $u\in\mathcal{U}_m$, $m\in\mathcal{M}_{\text{all}}$, and the corresponding energy consumption at the user to offload its task, respectively. $T_{u,0}^{\text{cp1}}=\frac{v_{u,0}^{\text{I}}C_{u,0}}{F_{u,0}}$ is the task computation execution time at the BS server for MUE $u\in\mathcal{U}_0$. $F_{u,m}$ is the computing resource allocated by the MBS to computing the task bits of user $u\in\mathcal{U}_m$, and $m=0$ for the MUEs. The time and energy consumption for transmitting the task calculation result from the server to the user is relatively small and will be ignored as in [8].

For edge computing corresponding to SUEs, the overall time consumption for uploading and computing the $v_{u,m}^{\text{I}}$ proportion of task bits of SUE $u\in\mathcal{U}_m$, $m\neq 0$, during edge computing can be given as $T_{u,m}^{\text{mec}}=T_{u,m}^{\text{off}}+\max\{T_m^{\text{cp1}},T_m^{\text{off}}+T_m^{\text{cp2}}\}$, and the corresponding overall energy consumption for the SUE ignoring the MBS's energy consumption can be written as $E_{u,m}^{\text{mec}}=E_{u,m}^{\text{off}}+E_m^{\text{cp1}}+E_m^{\text{off}}$. $T_m^{\text{cp1}}=\frac{(1-v_m^{\text{II}})v_{u,m}^{\text{I}}C_{u,m}}{f_m}$ and $E_m^{\text{cp1}}=\delta(1-v_m^{\text{II}})v_{u,m}^{\text{I}}C_{u,m}(f_m)^2$ are the task computing time and energy consumption at SBS m, respectively. $T_m^{\text{off}}=\frac{v_m^{\text{II}}v_{u,m}^{\text{I}}L_{u,m}}{R_m}$ is the uplink transmission time regarding to task offloading of SBS m, $E_m^{\text{off}}=p_mT_m^{\text{off}}$ is the energy consumption of SBS m for task offloading, and $T_m^{\text{cp2}}=\frac{v_m^{\text{II}}v_{u,m}^{\text{I}}C_{u,m}}{F_{u,m}}$ is the task execution time at the MBS for the offloaded task bits from SBS m.

[1]Our scheme applies to the case of multiple SUEs per small cell after considering the orthogonal frequency resource allocation scheme in [6].

III. Problem Formulation

We formulate the problem to minimize the overall energy consumption as

$$\min_{\left\{\substack{v^{\mathrm{I}}_{u,m},f_{u,m},F_{u,m},p_{u,m},\forall u\in\mathcal{U}_m,m\in\mathcal{M}_{\text{all}};\\ v^{\mathrm{II}}_m,\mathbf{\Theta},f_m,p_m,\forall m\in\mathcal{M}_{\text{all}}}\right\}} E \tag{1}$$

$$\text{s.t. } 0\le\sum_{m\in\mathcal{M}_{\text{all}}}\sum_{u\in\mathcal{U}_m}F_{u,m}\le F_{\max}, \tag{1a}$$

$$F_{u,m}\ge 0,\ u\in\mathcal{U}_m,\ m\in\mathcal{M}_{\text{all}}, \tag{1b}$$

$$0\le f_m\le f_m^{\max},\quad m\in\mathcal{M}, \tag{1c}$$

$$0\le f_{u,m}\le f_{\max},\ u\in\mathcal{U}_m,\ m\in\mathcal{M}_{\text{all}}, \tag{1d}$$

$$T^{\text{mec}}_{u,m}\le T_{\max},\ u\in\mathcal{U}_m,\ m\in\mathcal{M}_{\text{all}}, \tag{1e}$$

$$T^{\text{loc}}_{u,m}\le T_{\max},\ u\in\mathcal{U}_m,\ m\in\mathcal{M}_{\text{all}}, \tag{1f}$$

$$0\le v^{\mathrm{I}}_{u,m}\le 1, u\in\mathcal{U}_m, m\in\mathcal{M}_{\text{all}}, \tag{1g}$$

$$0\le v^{\mathrm{II}}_m\le 1,\quad m\in\mathcal{M}, \tag{1h}$$

$$\theta_n\in\Psi,\quad n=1,2,\ldots,N, \tag{1i}$$

$$0\le p_{u,m}\le p_{\max},\ u\in\mathcal{U}_m,\ m\in\mathcal{M}_{\text{all}}, \tag{1j}$$

$$0\le p_m\le p_m^{\max},\quad m\in\mathcal{M}, \tag{1k}$$

$$R_m\ge R_{u,m},\quad m\in\mathcal{M}, \tag{1l}$$

where $E=\sum_{m\in\mathcal{M}_{\text{all}}}\sum_{u\in\mathcal{U}_m}\left(E^{\text{loc}}_{u,m}+E^{\text{mec}}_{u,m}\right)$ is the overall energy consumption to complete all users' task computation, $F_{\max}$, $f_m^{\max}$, and $f_{\max}$ is the maximum CPU computation capacity for the MBS, SBS m, and each user, respectively, $T_{\max}$ is the maximum time limit for task bits computation of each user, θ_n is the configurable phase of the nth RIS element, $\Psi=\left\{0,\frac{2\pi}{2^b},\ldots,\frac{2\pi\left(2^b-1\right)}{2^b}\right\}$ is the feasible phase set of each RIS element with b being the bit-resolution of phase at RIS, and $p_m^{\max}$ and $p_{\max}$ is the maximum transmit power of SBS m and each user, respectively. In (1), Constraints (1a) - (1c) are the computing capacity limitations for BSs, and (1d) is the computation capacity limitation for users. (1e) and (1f) require that the task computing time during edge computing and the local computing time should be not greater than the maximum allowable time delay, respectively. (1g) and (1h) are the task offloading rate at users and SBS m, respectively. (1i) is the RIS phase constraint. (1j) and (1k) are the maximum transmit power limit for each user and SBS m. (1l) is the wireless backhaul capacity limit.

The problem in (1) includes both continuous and discrete variables and variable coupling exists. It is NP-hard due to the passive beamforming matrix constraint in (1i) [9], and is difficult to solve directly. We decompose it into two easier tractable sub-problems to find the sub-optimal solution next.

A. Sub-Problem 1: Energy Consumption for Users

We first optimize the energy consumption for all users with the assumption that all users' offloading task bits are fully computed by their associated BSs. Then the sub-problem to optimize energy consumption for users can be simplified as

$$\min_{\left\{\substack{v^{\mathrm{I}}_{u,m},f_{u,m},p_{u,m},\forall u\in\mathcal{U}_m,m\in\mathcal{M}_{\text{all}},\\ F_{u,m},\forall u\in\mathcal{U}_0,m=0}\right\}}\sum_{m\in\mathcal{M}_{\text{all}}}\sum_{u\in\mathcal{U}_m}E_{u,m} \tag{2}$$

$$\text{s.t. } 0\le\sum_{u\in\mathcal{U}_0}F_{u,0}\le F_{\max}-F_0,$$

$$\text{(1b) when } u\in\mathcal{U}_0, \text{(1d)}-\text{(1g)}, \text{(1j)}, \text{ and (1l)}. \tag{2a}$$

where $E_{u,m}=E^{\text{loc}}_{u,m}+E^{\text{off}}_{u,m}$ is the overall energy consumption for user u and F_0 is the reserved computation resource at the MBS for the offloaded task bits for all SBSs. We let $F_0=\sum_{m\in\mathcal{M}}\sum_{i\in\mathcal{U}_m}\sqrt{v^{\mathrm{I}}_{i,m}C_{i,m}}$ and update its value during optimization iteration. The problem in (2) is non-convex and with variable coupling, which is difficult to solve directly. We will solve it in Section IV-A.

B. Sub-Problem 2: Energy Consumption for SBSs

After each SBS receives the user's offloading task, its energy consumption can be computed by $E_m=E^{\text{cp1}}_m+E^{\text{off}}_m$. Then the problem of optimizing energy consumption for SBSs is formulated as

$$\min_{\{v^{\mathrm{II}}_m,\mathbf{\Theta},f_m,F_{u,m},p_m,\forall m\in\mathcal{M}\}}\sum_{m\in\mathcal{M}}E_m \tag{3}$$

$$\text{s.t. }\sum_{m\in\mathcal{M}}F_{u,m}\le F_{\max}-\sum_{u\in\mathcal{U}_0}F_{u,0}, \tag{3a}$$

$$F_{u,m}\ge 0,\quad m\in\mathcal{M}, \tag{3b}$$

$$T^{\text{mec}}_m\le T^{\text{cp1}}_m,\quad m\in\mathcal{M}, \tag{3c}$$

$$T^{\text{loc}}_m\le T^{\text{cp1}}_m,\quad m\in\mathcal{M},$$

$$\text{(1c)}, \text{(1h)}, \text{(1i)}, \text{(1k) and (1l)}. \tag{3d}$$

The optimization problem in (3) includes strong variable coupling and is non-convex. We will solve it in Section IV-B.

IV. Optimizing Energy Consumption

A. Optimizing Energy Consumption for Users

In this section, we propose an energy efficient scheme to solve Sub-problem 1 in four steps next.

Step 1 (Computation Capacity Optimization for Users): By fixing other optimization variables, the sub-problem of computation capacity optimization of a user can be simplified as

$$\min_{f_{u,m}}\sum_{m\in\mathcal{M}_{\text{all}}}\sum_{u\in\mathcal{U}_m}E^{\text{loc}}_{u,m}, \tag{4}$$

subject to (1d) and (1f). The problem is to find the minimum value of a quadratic function in the linear interval given by the delay constraint, which is convex. Thus, the optimal value of computation capacity of user u can be determined as $f^*_{u,m}=\frac{(1-v^{\mathrm{I}}_{u,m})C_{u,m}}{T_{\max}}$.

Step 2 (Task Offloading Rate Optimization for Users): Substituting $f^*_{u,m}$ into $E_{u,m}$ and fixing other optimization variables, the sub-problem of task offloading rate optimization for users can be written as

$$\min_{v^{\mathrm{I}}_{u,m}}\sum_{m\in\mathcal{M}_{\text{all}}}\sum_{u\in\mathcal{U}_m}E_{u,m}\left(f^*_{u,m}\right), \tag{5}$$

subject to (1e) and (1g). The problem is convex and the optimal value of $v^{\mathrm{I}}_{u,m}$ can be obtained as $v^{\mathrm{I}*}_{u,m}=\begin{cases}0, & \eta\in(-\infty,0]\\ \eta, & \eta\in(0,v_0]\\ v_0, & \eta\in[v_0,1]\end{cases}$, where $v_0=\min\left\{\frac{T_{\max}R_{u,m}}{L_{u,m}},1\right\}$ and $\eta=1-\sqrt{\frac{p_{u,m}L_{u,m}T^2_{\max}}{3\delta C^3_{u,m}R_{u,m}}}$.

Step 3 (Computation Resource Allocation for MUEs' Task Bits at the MBS): Shorter task execution time consumed at the BS helps reduce the user's transmit power and the energy consumption during the offloading process. Thus we optimize the computational resource for MUEs at the MBS by

$$\min_{F_{u,m}} \sum_{u\in\mathcal{U}_0} T_{u,0}^{\mathrm{cp1}}, \tag{6}$$

subject to (2a) and (1b) when $u \in \mathcal{U}_0$. The problem is convex and can be solved directly based on the Lagrangian multiplier method. Then the optimal computing resource allocated by the MBS to compute each user's task is determined as $F_{u,m}^{*} = \frac{F_{\max}\sqrt{v_{u,m}^{\mathrm{I}}C_{u,m}}}{\sum\limits_{m1\in\mathcal{M}_{\mathrm{all}}}\sum\limits_{i\in\mathcal{U}_{m1}}\sqrt{v_{i,m1}^{\mathrm{I}}C_{i,m1}}}$, where the value corresponding to SUEs will be finally determined in Step 4 of next section. At this point, the minimum task offloading data rate for MUE u is $R_{u,0}^{\min} = \frac{v_{u,0}^{\mathrm{I}}L_{u,0}}{T_{\max}-v_{u,0}^{\mathrm{I}}C_{u,0}/F_{u,0}^{*}}$. Let $f_m = \frac{v_{u,m}^{\mathrm{I}}C_{u,m}}{v_{u,m}^{\mathrm{I}}L_{u,m}/R_m+v_{u,m}^{\mathrm{I}}L_{u,m}/F_{u,m}^{*}}$. Then the minimum task offloading data rate for SUE u is $R_{u,m}^{\min} = \frac{v_{u,m}^{\mathrm{I}}L_{u,m}}{T_{\max}-v_{u,m}^{\mathrm{I}}C_{u,m}/f_m}$.

Step 4 (Transmit Power Optimization for Users): Let $\mathcal{P} = \{p_{u,m}, m \in \mathcal{M}_{\mathrm{all}}, u \in \mathcal{U}_m\}$ be the set of transmit power of offloading users. We optimize the users' transmit power by $\min_{\mathcal{P}} \sum\limits_{m\in\mathcal{M}_{\mathrm{all}}}\sum\limits_{u\in\mathcal{U}_m} p_{u,m}$, subject to (1j) and (1l). The problem is non-convex because of non-convexity of (1l). We first convert it into a standard convex optimization problem and then adopt the *successive convex approximation* (SCA) method [10] for the solution. Specifically, Constraint (1l) is first tightened according to [11] as $R_{u,m} \geq \widetilde{R}_{u,m} = B\alpha_{u,m}\log_2(\gamma_{u,m})/2 + B\chi_{u,m}/2$, where $\alpha_{u,m} = \frac{\gamma_{u,m}}{1+\gamma_{u,m}}$ and $\chi_{u,m} = \log_2(1+\gamma_{u,m}) - \frac{\gamma_{u,m}}{1+\gamma_{u,m}}\log_2(\gamma_{u,m})$. Let $p_{u,m} = \mathrm{e}^{\beta_{u,m}}$ and $\beta = \{\beta_{u,m}, m \in \mathcal{M}_{\mathrm{all}}, u \in \mathcal{U}_m\}$. The users' transmit power optimization problem can be equivalently converted into a standard convex optimization problem as

$$\min_{\beta} \sum_{m\in\mathcal{M}_{\mathrm{all}}}\sum_{u\in\mathcal{U}_m} e^{\beta_{u,m}}, \tag{7}$$

subject to (1j) and $\widetilde{R}_{u,m} \geq R_{u,m}^{\min}, m \in \mathcal{M}_{\mathrm{all}}, u \in \mathcal{U}_m$. It can be solved iteratively by updating β using the SCA method, which obtains the transmit power of each user correspondingly.

We perform each step sequentially to solve the Sub-problem 1, where the solution of the subsequent step is based on the results obtained by the preceding steps. These steps will be iterated until converge.

B. Optimizing Energy Consumption for SBSs

We propose scheme to optimize the energy consumption for SBSs in Sub-problem 2 in five steps in this section.

Step 1 (Passive Beamforming): Generally, the energy consumption for a SBS's task offloading is monotonically decreasing with the increasing of a SBS's uplink transmission data rate. Thus we optimize the passive beamforming matrix by $\max_{\mathbf{\Theta}} \sum\limits_{m\in\mathcal{M}} R_m$, subject to (1i). We utilize the low-complexity local search scheme in [12] to find its sub-optimal solution.

Step 2 (Computation Capacity Optimization at SBSs): The sub-problem of computation capacity optimization at SBSs when fixing other optimization variables can be simplified as $\min_{\{f_m,\forall m\in\mathcal{M}\}} \sum\limits_{m\in\mathcal{M}} E_m$, subject to (1c), (3c) and (3d). The objective function increases monotonically with f_m. Therefore, the optimal CPU cycle frequency of SBS m and the corresponding energy consumption is $f_m^{*} = \frac{(1-v_m^{\mathrm{II}})v_{u,m}^{\mathrm{I}}C_{u,m}}{T_m^{\mathrm{cp1}}}$ and $E_m^{loc,*} = \delta(1-v_m^{\mathrm{II}})v_{u,m}^{\mathrm{I}}C_{u,m}(f_m^{*})^2$, respectively.

Step 3 (Task Offloading Rate Optimization at SBSs): The sub-problem of task offloading rate optimization at SBSs can be written as

$$\min_{v_m^{\mathrm{II}}} \sum_{m\in\mathcal{M}} E_m\left(f_m^{*}\right), \tag{8}$$

subject to (1h), (3c) and (3d). The problem is convex and the closed-form optimal value of v_m^{II} can be obtained as $v_m^{\mathrm{II}*} = \begin{cases} 0, & \eta_2 \in (-\infty, 0] \\ \eta_2, & \eta_2 \in (0, \frac{T_m^{\mathrm{cp1}}R_m}{v_{u,m}^{\mathrm{I}}L_{u,m}}] \\ \frac{T_m^{\mathrm{cp1}}R_m}{L_{u,m}}, & \eta_2 \in [\frac{T_m^{\mathrm{cp1}}R_m}{v_{u,m}^{\mathrm{I}}L_{u,m}}, 1] \end{cases}$, where $\eta_2 = 1 - \sqrt{\frac{p_m(T_m^{\mathrm{cp1}})^2 v_{u,m}^{\mathrm{I}}L_{u,m}}{3\delta(v_{u,m}^{\mathrm{I}}C_{u,m})^3 R_m}}$.

Step 4 (Computation Resource Allocation for SUEs' Task Bits at the MBS): Similar as Step 3 in Section IV-A, we optimize the computation resource allocation for SUEs' task at the MBS by

$$\min_{F_{u,m}} \sum_{m\in\mathcal{M}} T_m^{\mathrm{cp2}}, \tag{9}$$

subject to (3a) and (3b). The problem is convex and can be solved based on the Lagrangian multiplier method. The optimal computing resource allocation of each small BS can be obtained as $F_{u,m}^{*} = \frac{(F_{\max}-\sum_{u\in\mathcal{U}_0}F_{u,0})\sqrt{v_m^{\mathrm{II}}v_{u,m}^{\mathrm{I}}C_{u,m}}}{\sum_{i\in\mathcal{M}}\sqrt{v_i^{\mathrm{II}}v_{u,i}^{\mathrm{I}}C_{u,i}}}$.

Step 5 (The SBSs' Transmit Power Optimization): The optimization problem of offloading all SBSs on the same bandwidth is expressed as $\min_{\mathcal{P}_1} \sum\limits_{m\in\mathcal{M}} p_m$, subject to (1k) and (1l), where $\mathcal{P}_1 = \{p_m, m \in \mathcal{M}\}$ is the set of transmit power offloaded by the SBS. The problem is non-convex. We utilize Step 4 of Section IV-A to solve it and thus omitted here.

Similar to Section IV-A, these five steps will be solved sequentially and iteratively until converge to solve the Sub-problem 2.

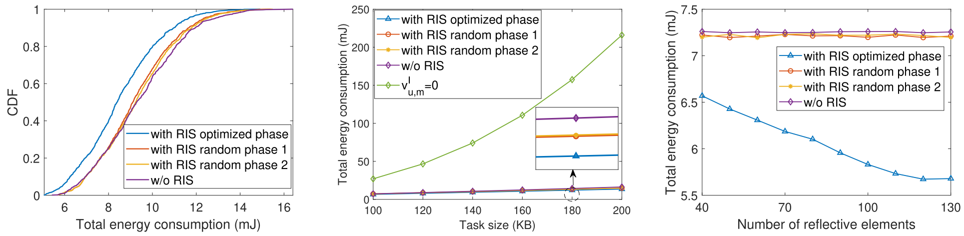
C. The Overall Scheme and Analysis

To solve the original problem in (1), the overall scheme performs iteration between sub-problem 1 and 2 until convergence. The energy consumption decreases during the steps iteration and the proposed scheme can converge.

To analyze the complexity of the proposed scheme, we assume that I_1, I_2, and I_3 are the numbers of iterations for Sub-problem 1, Sub-problem 2, and the original problem (1), respectively. The computational complexity for Sub-problem 1 is $\mathcal{O}_1\left(\left(3K+2M+\left(K+M\right)L_1\right)\times I_1\right)$. The computational complexity for Sub-problem 2 is $\mathcal{O}_2\left(\left(N\times 2^b+3M+ML_2\right)\times I_2\right)$. Thus the computational complexity for problem (1) is $\mathcal{O}((\mathcal{O}_1+\mathcal{O}_2)\times I_3)$.

V. Performance Evaluation

In this section, we conduct simulation to evaluate the performance of the proposed scheme and report the results. In this simulation setup, the radius of macrocell and each small cell is 300 m and 10 m, respectively. The RIS is 10 m away



(a) CDF performance. (b) Under different task sizes of user. (c) Under different RIS element No.

Fig. 3. Overall energy consumption of the proposed scheme.

TABLE II

SIMULATION PARAMETER SETTINGS

Parameter	Value	Parameter	Value
M	4	K	4
$F_{\max}$	25 G cycles/s	$P_{\max}$	23 dBm
$f_m^{\max}$	2.5 G cycles/s	σ^2	-80 dBm
$\rho_{u,m}$	150 G cycles/bit	$T_{\max}$	100 ms
δ	10^{-26} G cycles/s	$P_m^{\max}$	24 dBm
$f_{\max}$	1 G cycles/s	$L_{u,m}$	100 Kbits

from the MBS. The carrier frequency of the system is $f_c = 2$ GHz and $B = 20$ MHz. The computing power of all SBSs is set to be the same and the computing power of the MBS is stronger than that of the SBS. For computing tasks, the task size of all users for computing is set to be the same. $Q = 40$ and $N = 40$. The default value of other simulation parameters are summarized in Table II [13].

Fig. 3(a) shows the cumulative distribution function (CDF) of the total energy consumption of user task offloading, where "with RIS random phase 1" and "with RIS random phase 2" denote the proposed scheme with RIS configured with the continuous random phase at $[0, 2\pi)$ and the one with discrete random phase within feasible region, respectively. As can be seen in the figure, the proposed scheme with optimized RIS phase performs significantly better than other schemes. Although the performance of random phase with RIS is close to that without RIS, it is still better than that without RIS.

Fig. 3(b) shows the total energy consumption versus user task size. From the figure, when the amount of user tasks is small, the performance of the user's local computation is close to that of the optimization algorithm. However, there are some critical values of task size depending on specific applications. Compared with local computing, reducing the task size and using the optimized offloading strategy in this letter will not reduce the total energy consumption. This means that when the number of task bits is small, the local computing strategy is the most energy-efficient strategy. When the task volume is large, the total energy consumption of the optimization algorithm is lower than the local processing energy consumption. Fig. 3(c) shows the total energy consumption under different number of RIS reflective elements for four MUEs and SBSs. As the number of RIS components increases, the performance under the random RIS phase is close to that without RIS. The proposed scheme outperforms the other methods and dramatically reduces the total energy consumption. Specifically, the proposed scheme consumes 9.1% less energy than that under random phase at $N = 40$, while it consumes 21% less energy at $N = 130$. Increasing the number of reflective elements will not reduce the energy consumption of user offloading transmission because RIS is utilized to help SBSs offloading. As the number of reflective elements continues to increase, the total energy consumption will converge.

VI. CONCLUSION

This letter investigated energy efficient task offloading in HetNets with RIS-aided wireless backhaul. We formulated the joint task offloading and resource allocation problem to minimize the overall energy consumption considering the energy consumption of SBSs and implement a secondary offloading strategy for the SUEs. To facilitate energy efficient task offloading, we proposed a two-step scheme and demonstrate an optimal solution can be obtained within several finite iterations. Simulation results showed that the use of RIS significantly reduces the energy consumption of task offloading of the SBSs in HetNets with wireless backhaul. In the future work, we will investigate the effective resource allocation scheme and apply non-orthogonal multiple access to further improve the scenarios with multiple users.

REFERENCES

[1] B. Tezergil and E. Onur, "Wireless backhaul in 5G and beyond: Issues, challenges and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2579–2632, 4th Quart., 2022.

[2] Q.-V. Pham et al., "Mobile edge computing with wireless backhaul: Joint task offloading and resource allocation," *IEEE Access*, vol. 7, pp. 16444–16459, 2019.

[3] T. Bai et al., "Reconfigurable intelligent surface aided mobile edge computing," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 80–86, Dec. 2021.

[4] T. Bai et al., "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2666–2682, Nov. 2020.

[5] Y. Hao et al., "Energy-efficient multi-user mobile-edge computation offloading in massive MIMO enabled HetNets," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[6] X. Yuan et al., "Joint user, channel, modulation-coding selection, and RIS configuration for jamming resistance in multiuser OFDMA systems," *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1631–1645, Mar. 2023.

[7] W. Zhang et al., "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4569–4581, Sep. 2013.

[8] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, 3rd Quart., 2017.

[9] C. You et al., "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.

[10] M. Sami and J. N. Daigle, "User association and power control for UAV-enabled cellular networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 267–270, Mar. 2020.

[11] A. Zappone et al., "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668–1683, Apr. 2016.

[12] Y. Chen et al., "Reconfigurable intelligent surface assisted device-to-device communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 2792–2804, May 2021.

[13] N. Nouri et al., "Joint access and resource allocation in ultra-dense mmWave NOMA networks with mobile edge computing," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1531–1547, Feb. 2020.