Joint Subchannel and Power Allocation Algorithm for NOMA Heterogeneous Cloud Radio Access Networks (Postprint)

Authors: Gu Zhaowei, Jiang Lingyun

Date: 2022-05-10T00:00:00+00:00

Abstract

To address the problem of low energy efficiency caused by the trade-off between increased network utility and higher energy consumption in NOMA-based heterogeneous cloud radio access networks, a joint subchannel and power allocation scheme is proposed. First, the scheme defines the difference between network utility and grid energy cost as system profit, and formulates an optimization problem with the objective of maximizing system profit, considering constraints such as maximum transmission power, energy harvesting (EH) battery capacity, users' minimum data rate requirements, and cross-layer interference threshold. Then, a greedy algorithm is adopted to pair users and allocate subchannels, achieving a low-complexity suboptimal solution; finally, power allocation is optimized using a Lagrangian maximization method based on the alternating direction method of multipliers (ADMM). Simulation results show that, compared with schemes without EH units in NOMA systems, system profit improves by approximately 18.8%; compared with schemes with EH units in OFDMA systems, system profit improves by approximately 11.8%.

Full Text

Abstract

In heterogeneous cloud radio access networks based on NOMA, increasing network utility comes at the cost of higher energy consumption, leading to low energy efficiency. To address this problem, this paper proposes a joint subchannel and power allocation scheme. First, the scheme defines the difference between network utility and grid energy cost as system revenue, and establishes an optimization problem to maximize system revenue subject to constraints including maximum transmission power, energy harvesting (EH) battery capacity, users' minimum data rate requirements, and cross-layer interference thresholds.

Then, a greedy algorithm is employed to pair users and allocate sub-channels, achieving a low-complexity suboptimal solution. Finally, power allocation is optimized using a Lagrangian maximization method based on the alternating direction method of multipliers. Simulation results show that compared with the scheme without EH units in NOMA systems, the system revenue is improved by about 18.8%; compared with the scheme with EH units in OFDMA systems, the system revenue is improved by about 11.8%.

Keywords: Heterogeneous cloud RAN; NOMA; hybrid energy; sub-channel allocation; power allocation

0 Introduction

Due to the proliferation of smart devices and emerging social network services, the number of mobile terminals and mobile data traffic has grown dramatically [?]. To meet the demands for higher data rates and massive wireless connectivity, heterogeneous networks composed of macro cells and small cells such as micro, pico, and femto cells have become a key technology. Among them, Heterogeneous Cloud Radio Access Network (HCRAN) has received extensive research attention due to its scalability, flexibility, and compatibility. HCRAN consists of a High Power Node (HPN), Low Power Nodes (LPN) (i.e., Remote Radio Heads, RRH), and a pool of Base Band-processing Units (BBU). HCRAN achieves the separation of control and data planes, where the HPN is only responsible for seamless coverage and transmitting control signaling, while RRHs handle data transmission services, thereby avoiding frequent handovers and control signaling overhead [?].

Non-orthogonal multiple access (NOMA) has emerged as a promising technology to meet 5G requirements in addressing spectrum scarcity [?]. The main principle of NOMA is to allow multiple users to non-orthogonally share the same time, frequency, and symbol resources. Multi-user signals are superposition-coded at the transmitter where inter-user interference is permitted, while Successive Interference Cancellation (SIC) is performed at the receiver to accurately separate multi-user signals. However, the detection process at the receiver is complex due to the combination of multi-user signals. The most popular form is power-domain NOMA, where power is allocated to users based on their channel conditions, with users having strong channel gains assigned lower power [?]. It is foreseeable that the combination of HCRAN and NOMA can leverage the advantages of both to improve spectral efficiency and energy efficiency. However, this combination faces challenges of co-tier and cross-tier interference as well as effective resource allocation. Therefore, effective resource allocation and interference management are fundamental aspects to consider in NOMA-HCRAN systems.

Furthermore, the large-scale deployment of RRHs is accompanied by enormous energy consumption and greenhouse gas emissions. To address this, EH technology has been introduced into HCRAN, where each LPN is equipped with an EH unit to harvest renewable energy as a supplement to grid energy supply, thereby achieving the goal of green communications [?]. However, unlike traditional stable and reliable grid power, renewable energy is characterized by uneven spatiotemporal distribution, randomness, and intermittent arrival, depending on weather conditions and RRH deployment locations [?]. This paper must consider when and for how long RRHs with hybrid energy supply can schedule renewable energy. Therefore, effective management of renewable energy battery storage and joint scheduling of grid and renewable energy is of practical significance. Finally, due to the uneven distribution of mobile users, traffic loads may not match the energy harvesting status and spectrum resources of each small cell base station, necessitating energy scheduling based on the status of each small cell RRH. To address these challenges, most previous works have focused on achieving grid energy savings through inter-cell traffic offloading (e.g., [?]) or on energy scheduling for network utility maximization (e.g., [?]). Notably, network utility increases with the consumption of more grid energy, which must be purchased at a certain electricity price. To ensure green communications brought by hybrid energy supply, a careful balance between network utility and grid energy cost must be evaluated. Therefore, this paper investigates the aforementioned problems.

1 Related Work

Effective resource allocation is crucial for achieving high performance in power-domain NOMA systems. Most current research is based on small cell heterogeneous networks (HetNets). However, since the transmission rate calculation methods and power consumption models in HCRAN are not significantly different from those in HetNets, relevant research results can be referenced and compared. First, to solve resource allocation problems in multi-cell scenarios, studies have investigated total data rate and energy efficiency maximization. [?] studied power allocation using a Stackelberg game approach to achieve Stackelberg equilibrium through a distributed power allocation algorithm. [?] considered base station power and user transmission rate constraints in multi-cell multi-user NOMA systems, then achieved effective power allocation based on KKT conditions. However, these studies did not address effective sub-channel allocation and user pairing in NOMA. To maximize overall system throughput, [?] proposed user scheduling and iterative distributed power control algorithms to obtain suboptimal intra-cell resource allocation, achieving high spectral efficiency but with high system complexity. [?] proposed initialization and SOEMA algorithms to pair users with optimal sub-channels to maximize small cell user throughput, with low complexity but without optimal power allocation. However, these studies did not consider the increased grid energy cost when maximizing data rates, nor did they consider system energy efficiency and green communications.

On the other hand, [?]-[?] studied energy efficiency maximization and user scheduling optimization in NOMA-HetNets downlinks. [?] investigated fair power allocation with given base station power constraints, while [?] studied sub-channel and power allocation in small cells, but these studies neglected user pairing. [?] simultaneously considered user pairing and access in non-uniform NOMA cells to achieve higher data rates and energy efficiency. [?] considered perfect and imperfect Channel State Information (CSI), used matching theory to pair two users in a sub-channel, and employed fractional transmit power allocation to allocate each user' s power. However, these studies did not consider using EH units at small base stations to reduce total network energy consumption. [?] studied sub-channel allocation and power optimization to maximize total cell energy efficiency under simultaneous wireless information and power transmission with EH units. However, this ignored co-tier interference between small cells reusing the same sub-channel, which would increase system energy consumption if not well mitigated [?]. [?] studied sub-channel and power allocation considering cross-tier/co-tier interference constraints, EH constraints, and imperfect CSI. However, this did not consider the renewable energy battery capacity constraint of small base stations to ensure they can obtain the minimum required energy to improve system energy efficiency.

Based on the above research, this paper proposes a joint sub-channel and power allocation scheme for NOMA-HCRAN that considers RRH EH battery capacity constraints and cross-tier/co-tier interference, while maximizing system revenue (defined as the difference between network utility and grid energy cost) and meeting individual mobile users' minimum data rate requirements.

2 System Model

This section describes the proposed system model and parameters for a downlink NOMA-based two-tier heterogeneous cloud radio access network (NOMA-HCRAN), along with the power consumption model used in this network.

2.1 System Model

We consider a downlink NOMA-HCRAN network as shown in [Figure 1: see original paper], consisting of a High Power Node, i.e., a Macro Base Station (MBS), and a set of Low Power Nodes (RRHs) located within its coverage area. All base stations and users are assumed to be equipped with a single antenna, i.e., a Single-Input Single-Output system. Since the MBS is mainly responsible for seamless coverage and control signaling transmission, it requires stable grid energy supply. Therefore, only each RRH is equipped with renewable energy harvesting devices and can be powered by both grid energy and renewable energy (such as solar, wind, thermal, and mechanical energy). Each RRH first stores the harvested energy in its battery and then consumes it in subsequent time slots.

Let $\mathcal{F} = \{1, 2, \ldots, F\}$ denote the set of RRHs, $\mathcal{M} = \{1, 2, \ldots, M\}$ denote the set of MBSs (for simplicity, we assume $M = 1$), $\mathcal{R} = \{1, 2, \ldots, R\}$ denote the set of macro cell users (mCUEs), and $\mathcal{W} = \{1, 2, \ldots, W\}$ denote the set of RRH users (fCUEs). The system employs NOMA technology to serve users in this heterogeneous network using S sub-channels. The total system bandwidth is denoted by B, and each sub-channel's bandwidth is obtained by equally dividing B among S sub-channels, i.e., $B_{\text{sch}} = B/S$. Let $P_{\max}^{\text{MBS}}$ and $P_{\max}^{\text{RRH}}$ denote the maximum transmit power of the MBS and RRH, respectively. A Rayleigh fading model depending on distance and path loss is adopted. Let $h_{f,w,s}$ denote the channel fading coefficient for the w-th fCUE served by RRH f on sub-channel s, and let $h_{m,r,s}$ denote the channel fading coefficient for the r-th mCUE served by MBS m on sub-channel s.

Assume all users' channel coefficients are ordered as in (1):

$$|h_{f,1,s}|^2 \leq |h_{f,2,s}|^2 \leq \cdots \leq |h_{f,W,s}|^2 \leq |h_{m,1,s}|^2 \leq \cdots \leq |h_{m,R,s}|^2$$

where $h_{f,w,s} = \sqrt{g_{f,w,s}} \cdot \chi_{f,w,s}$, $g_{f,w,s}$ represents the shadowing and path loss between the w-th fCUE and its serving RRH f, and $\chi_{f,w,s}$ is a parameter representing Rayleigh fading.

According to the NOMA principle, base stations transmit multi-user signals with different power levels using superposition coding. At the receiver, each user receives the desired signal from its serving base station and interference signals from other base stations. The proposed model considers interference signals from the same type of base stations (co-tier interference) as well as from different types of base stations (cross-tier interference). Since this model can capture the actual performance of fCUEs and mCUEs, the received signal at user $U_{f,w,s}$ can be mathematically expressed as (2):

$$y_{f,w,s} = \underbrace{h_{f,w,s}\sqrt{p_{f,w,s}}x_{f,w,s}}_{\text{desired signal}} + \underbrace{\sum_{q=1}^{w-1} h_{f,w,s}\sqrt{p_{f,q,s}}x_{f,q,s}}_{\text{intra-user interference}} + \underbrace{\sum_{q=w+1}^{W} h_{f,w,s}\sqrt{p_{f,q,s}}x_{f,q,s}}_{\text{intra-user interference}} + \underbrace{\sum_{f' \neq f}\sum_{q=1}^{W} h_{f,w,s}\sqrt{p_{f',q,s}}x_{f',q,}}_{\text{co-tier interference}}$$

where $x_{f,w,s}$ and $p_{f,w,s}$ denote the desired transmission symbol and power allocated to user $U_{f,w,s}$, respectively. In the cross-tier interference term, $x_{m,r,s}$, $h_{m,r,s}$ denote the desired transmission symbol and channel fading coefficient (between MBS m and user $U_{m,r,s}$), and $p_{m,r,s}$ denotes the power allocated to user $U_{m,r,s}$. Additionally, $z_{f,w,s} \sim \mathcal{N}(0, \sigma_{f,w,s}^2)$ represents the additive white Gaussian noise at $U_{f,w,s}$. The parameters $\alpha_{f,w,s} \in \{0,1\}$ and $\alpha_{m,r,s} \in \{0,1\}$ are binary variables indicating sub-channel allocation for fCUEs and mCUEs, respectively.

NOMA technology allows more than one user to share the same sub-channel, with SIC performed at the receiver for correct demodulation. SIC can eliminate interference by managing users based on their allocated power levels [**?**]. Based on the channel coefficient ordering in (1), user $U_{f,w,s}$ can effectively decode all signals of the $(w-1)$-th fCUE and treats all signals of $(w+1)$-th fCUEs as noise. Assuming all users have perfect knowledge of CSI, the Signal-to-Interference-plus-Noise Ratio (SINR) received by user $U_{f,w,s}$ can be expressed as (5), which can be simplified to (6):

$$\gamma_{f,w,s}^{\text{NOMA}} = \frac{|h_{f,w,s}|^2 p_{f,w,s}}{|h_{f,w,s}|^2 \sum_{q=1}^{w-1} p_{f,q,s} + \sum_{f' \neq f} \sum_{q=1}^{W} |h_{f,w,s}|^2 p_{f',q,s} + \sum_{m=1}^{M} \sum_{r=1}^{R} |h_{f,w,s}|^2 p_{m,r,s} + \sigma_{f,w,s}^2}$$

The achievable data rate in bits per second (bps) for user $U_{f,w,s}$ is given by (7):

$$R_{f,w,s} = B_{\text{sch}} \log_2(1 + \gamma_{f,w,s}^{\text{NOMA}})$$

Let $\alpha = \{\alpha_{f,w,s} | f \in \mathcal{F}, w \in \mathcal{W}, s \in \mathcal{S}\}$ denote the set of sub-channel allocation parameters, and $\mathbf{p} = \{p_{f,w,s} | f \in \mathcal{F}, w \in \mathcal{W}, s \in \mathcal{S}\}$ denote the power allocation values. The total data rate that user w can obtain from RRH f is given by (8):

$$R_{f,w}^{\text{sum}} = \sum_{s=1}^{S} \alpha_{f,w,s} R_{f,w,s}$$

Since individual mobile users have different satisfaction levels with obtained throughput, we use a utility function $U_w(R_w)$ to evaluate mobile user w' s satisfaction with throughput R_w. The overall network utility can then be expressed as $\sum_{w \in \mathcal{W}} U_w(R_w)$. Without loss of generality, for elastic services, the function can be set as a twice-differentiable, increasing, and strictly concave function.

2.2 Power Consumption Model

The power consumption parameter model proposed in this section is based on [?] and considers power consumption at both the base station transmitter and user receiver. The power consumption of RRH f is expressed as:

$$P_f^{\text{RRH}} = \underbrace{\frac{1}{\xi_f} \sum_{w=1}^{W} \sum_{s=1}^{S} \alpha_{f,w,s} p_{f,w,s}}_{\text{dynamic power}} + \underbrace{P_f^{\text{sta}}}_{\text{static power}}$$

where the first term is dynamic power, consisting of transmission power from the base station to users, i.e., the power consumed by the RF signal emitted from the power amplifier. ξ_f represents the efficiency of the power amplifier at base station f. The second term is static power, composed of parameters $P_{f,\text{cir}}^{\text{sta}}$ and $P_{f,\text{rec}}^{\text{sta}}$ as shown in (10). $P_{f,\text{cir}}^{\text{sta}}$ corresponds to power consumed by operational circuits and signal transmission systems (such as cooling systems, filters, etc.) at f. $P_{f,\text{rec}}^{\text{sta}}$ corresponds to power required for signal reception at the receiver (i.e., fCUE). Similarly, the power consumption of the macro base station is given by (11), with all relevant parameters similar to those for RRH.

To balance network utility and system cost, we need to consider the cost of grid power supply. Let π_m and π_f denote the unit energy prices set by the power company for MBS and RRH, respectively. Since the MBS is grid-powered, its energy cost is given by $C_m = \pi_m \sum_{t=1}^{T} P_m^{\text{MBS}}$. Each RRH is powered by both grid and renewable energy. Assuming T is the operation time, RRH f first uses stored renewable energy during time τ_f and then uses grid power during the remaining $(1-\tau_f)T$. The grid energy cost paid by RRH f to the power company is given by $C_f = \pi_f(1-\tau_f)\sum_{t=1}^{T} P_f^{\text{RRH}}$. The total grid energy cost can be calculated as:

$$C_{\text{total}} = \sum_{f=1}^{F} C_f + C_m$$

On the other hand, all RRHs are powered by renewable energy stored in their respective batteries. Therefore, the transmit power of RRH depends on the amount of renewable energy stored in the battery (i.e., E_f), leading to (16):

$$\sum_{w=1}^{W}\sum_{s=1}^{S} \alpha_{f,w,s} p_{f,w,s} \leq E_f, \quad \forall f \in \mathcal{F}$$

2.3 Energy Efficiency Metric

Energy efficiency is an important parameter for evaluating the performance of cellular systems designed to reduce energy consumption. By definition, energy efficiency E_E is the ratio of throughput (achievable data rate) to total power consumption [?]. It is measured in bits/Joule as shown in (17):

$$E_E = \frac{\text{Throughput (bps)}}{\text{Total Power Consumption (Joule/s)}} = \frac{\sum_{f=1}^{F}\sum_{w=1}^{W}\sum_{s=1}^{S} \alpha_{f,w,s} R_{f,w,s}}{\sum_{f=1}^{F} P_f^{\text{RRH}} + P_m^{\text{MBS}}}$$

2.4 Problem Formulation

This paper aims to balance network utility and grid energy cost according to each user’ s data rate requirement. Mathematically, we formulate the trade-off problem as follows:

$$\begin{aligned}
\max_{\alpha,\mathbf{p},\tau} \quad & \sum_{w\in\mathcal{W}} U_w(R_w) - C_{\text{total}} && (1)\\
\text{s.t.} \quad & R_{f,w}^{\text{sum}} \geq R_{\text{req}}, \quad \forall w \in \mathcal{W} && (19\text{-}1)\\
& \sum_{w=1}^{W}\sum_{s=1}^{S} \alpha_{f,w,s} p_{f,w,s} \leq E_f, \quad \forall f \in \mathcal{F} && (19\text{-}2)\\
& \sum_{w=1}^{W}\sum_{s=1}^{S} \alpha_{f,w,s} p_{f,w,s} \leq P_{\max}^{\text{RRH}}, \quad \forall f \in \mathcal{F} && (19\text{-}3)\\
& \sum_{f=1}^{F}\sum_{w=1}^{W} \alpha_{f,w,s} |h_{m,f,w,s}|^2 p_{f,w,s} \leq I_{\text{th}}, \quad \forall s \in \mathcal{S} && (19\text{-}4)\\
& p_{f,w,s} \geq 0, \quad \forall f, w, s && (19\text{-}5)\\
& \sum_{s=1}^{S} \alpha_{f,w,s} \leq 1, \quad \forall w \in \mathcal{W} && (19\text{-}6)\\
& \alpha_{f,w,s} \in \{0,1\}, \quad \forall f, w, s && (19\text{-}7)\\
& 0 \leq \tau_f \leq 1, \quad \forall f \in \mathcal{F} && (19\text{-}8)
\end{aligned}$$

where $U_w(R_w)$ is the network utility function, which can take different forms. Here we select the proportional fair utility $\log(R_w)$ to obtain fair power allocation among users, and non-negative coefficient α represents the weight of grid energy consumption impact on resource allocation. The vectors α, $\mathbf{p}$, and τ represent the sub-channel allocation matrix, power allocation values, and renewable energy usage time, respectively. Constraint (19-1) ensures QoS requirements by enforcing a minimum data rate for each user. Constraint (19-2) indicates that the renewable energy used to supplement RRHs is limited by battery capacity. Constraint (19-3) represents the maximum transmit power limit for RRH f. Equation (19-4) ensures that the maximum cross-layer interference caused by fCUEs to the macro cell on each sub-channel does not exceed threshold I_{th}. Equation (19-5) ensures that power allocated to users is always positive, while (19-6) and (19-7) ensure each user can only access one sub-channel. Notably, we can adjust α to achieve different balances between network utility and grid energy cost.

Note that due to the product of $\alpha_{f,w,s}$ and $p_{f,w,s}$, optimization problem (19) is non-convex with inequality constraints at first glance. Therefore, common optimization methods such as ADMM and Lagrangian maximization cannot be directly applied. In the following sections, we convexify the problem through transformation and reparameterization, and propose an effective algorithm to solve it.

3 Joint Sub-channel and Power Allocation Algorithm (JCPA)

In the proposed system model, we initially assume that an advanced algorithm is used in the user access process, such as the algorithm used in [?]. Note that the proposed user pairing and power allocation solution is independent of the user access algorithm used, while the evaluation of user pairing and power allocation problems is the focus of this paper. The proposed JCPA applies a greedy algorithm (GA) for user sub-channel matching, followed by an Alternating Direction Method of Multipliers (ADMM)-based power allocation algorithm to obtain optimal system revenue in the NOMA-based hybrid energy supply heterogeneous cloud radio access network.

User pairing and power allocation form a joint optimization problem that affects NOMA system performance. To reduce the computational complexity of solving this joint optimization problem, it is decomposed into two sub-problems. The first part uses GA to solve the user pairing problem of assigning two users to the same sub-channel. The second part uses an ADMM-based optimization algorithm to solve the power allocation problem to obtain optimal power allocation for each user and optimal system revenue.

3.1 GA-Based User Pairing Algorithm

The user pairing process significantly impacts NOMA network performance. Optimizing the user pairing scheme is necessary to improve NOMA network performance. An exhaustive search method could be used to achieve optimal system performance by considering all possible user pairing combinations. However, computational complexity multiplies as the number of users increases. Therefore, this paper adopts a GA method with lower computational complexity that can achieve suboptimal user pairing for NOMA-HCRAN.

Assuming each sub-channel is allocated to only two users, equal power is allocated to each sub-channel, and fair power is allocated to each paired user in each sub-channel [?]. The power allocated to the first user $U_{f,1,s}$ is given by (20):

$$p_{f,1,s} = \frac{\delta_{f,1}}{\delta_{f,1} + \delta_{f,2}} p_{f,s}^{\text{total}}$$

where $\delta_{f,1}$ and $\delta_{f,2}$ are the channel coefficients between the first and second users, respectively, and $p_{f,s}^{\text{total}}$ is the total power allocated to sub-channel s. The power allocated to the other paired user $U_{f,2,s}$ is given by (23):

$$p_{f,2,s} = p_{f,s}^{\text{total}} - p_{f,1,s}$$

The main idea of using GA is to select the best possible choice at each stage of the problem to reach the optimal solution. The GA focuses on assigning two best users that can maximize energy efficiency.

Algorithm 1 describes the GA algorithm for assigning two users to a valid sub-channel. At each stage, two users are selected and assigned to the best sub-channel that maximizes their energy efficiency. The process terminates after all users are assigned to their most utilized sub-channels. Algorithm 1 consists of two procedures: initialization and sub-channel allocation.

a) Initialization: Algorithm 1 first initializes user sets, sub-channel sets, allocation variables, and sets of selected users. Each user is initialized with zero data rate and each sub-channel with power allocation. Then, each user' s SINR is calculated and sorted in descending order. Power is allocated to each user through equations (20) and (23), thereby updating each user' s data rate.

b) Sub-channel allocation: If the first user has higher SINR compared to other users and cannot obtain the required data rate, it is assigned to the selected sub-channel. The first user must also satisfy the condition that co-channel interference is below the threshold of the assigned sub-channel. After selecting the first user, the set of selected users and user sub-channel allocation parameters are updated, and the user is removed from the user set.

On the other hand, if the second user's energy efficiency is lower than other users and its energy efficiency can be improved by adding the energy efficiency of the first selected user, then the second user is selected. After selecting the second user, both the set of selected users and user sub-channel allocation parameters are updated, and the user is removed from the user set.

This achieves optimal pairing and removes the optimal sub-channel from the sub-channel set. This process continues until all users are assigned to their most fully utilized sub-channels. The flow of Algorithm 1 is as follows.

Algorithm 1: GA-Based User Pairing Algorithm 1. Initialize: a) User set $\mathcal{W} = \{1, 2, \ldots, W\}$ b) Sub-channel set $\mathcal{S} = \{1, 2, \ldots, S\}$ c) Set sub-channel allocation parameters $\alpha_{f,w,s} = 0, \forall f, w, s$ d) Set of users already assigned to channels $\Omega_s = \emptyset, \forall s$ e) Each user' s initial data rate $R_{f,w} = 0, \forall f, w$ f) Each sub-channel's power allocation $p_{f,s}^{\text{total}} = P_{\max}^{\text{RRH}}/S, \forall f, s$ 2. Calculate each user's SINR $\gamma_{f,w,s}, \forall f, w, s$ and sort them in descending order 3. Obtain $p_{f,1,s}$ and $p_{f,2,s}$ through equations (20) and (23) 4. While $\mathcal{W} \neq \emptyset$ do 5. For each sub-channel $s \in \mathcal{S}$ do 6. Select the first user w^* with maximum SINR: $w^* = \arg\max_{w\in\mathcal{W}} \gamma_{f,w,s}$ 7. If $R_{f,w^*} < R_{\text{req}}$ and $I_{\text{cross}} \leq I_{\text{th}}$ then 8. Assign user w^* to sub-channel s: $\alpha_{f,w^*,s} = 1$ 9. Update $\Omega_s = \Omega_s \cup \{w^*\}$, $\mathcal{W} = \mathcal{W} \setminus \{w^*\}$ 10. End if 11. For each remaining user $j \in \mathcal{W}$ do 12. Calculate energy efficiency EE_j 13. If $EE_j < EE_{w^*}$ and $EE_j + EE_{w^*} > EE_{w^*}$ then 14. Select second user $j^* = \arg\max_{j\in\mathcal{W}}(EE_j + EE_{w^*})$ 15. Assign user j^* to sub-channel s: $\alpha_{f,j^*,s} = 1$ 16. Update $\Omega_s = \Omega_s \cup \{j^*\}$, $\mathcal{W} = \mathcal{W} \setminus \{j^*\}$ 17. End if 18. End for 19. End for 20. Until $\mathcal{W} = \emptyset$ 21. End while

3.1.1 Time Complexity The complexity analysis of the user pairing algorithm: Let S be the number of sub-channels and W represent the available users in the network, i.e., $W = 2S$. Applying exhaustive search helps achieve higher system performance but requires considering all possible user access combinations, resulting in $\mathcal{O}(2^W)$ computational complexity. The computational complexity of the GA-based solution (Algorithm 1) mainly depends on the initialization process of sorting users by SINR and the allocation process of assigning users to sub-channels. The initialization process requires $\mathcal{O}(W \log W)$ operations, and the allocation process requires $\mathcal{O}(W)$ operations. Therefore, the total complexity of GA can be expressed as $\mathcal{O}(W \log W + W) = \mathcal{O}(W \log W)$. Thus, the computational complexity of GA (Algorithm 1) is $\mathcal{O}(W \log W)$.

3.2 Optimization Algorithm Design

In this section, we find the optimal energy and power allocation by sequentially optimizing multiple variables in (19). First, we transform the problem into a convex optimization problem through reparameterization. Then, we propose an effective optimization algorithm based on primal-dual parameters to solve the equivalent problem. [Figure 2: see original paper] shows the architecture of the proposed algorithm, particularly the connections between convexification, algorithm design, and key optimization problems.

3.2.1 Convexification To make problem (19) tractable, we first convert it into a convex optimization problem. To avoid the product of $\alpha_{f,w,s}$ and $p_{f,w,s}$, we introduce renewable energy power consumption variables $\hat{p}_{f,w,s}$ and grid power consumption variables $\tilde{p}_{f,w,s}$, i.e., $p_{f,w,s} = \hat{p}_{f,w,s} + \tilde{p}_{f,w,s}$. To simplify, let $\mathbf{v}$ denote the vector of variables in (27). Then problem (19) can be rewritten as:

$$\max_{\mathbf{v}} \quad L(\mathbf{v}) = \sum_{w \in \mathcal{W}} U_w(R_w) - \sum_{f=1}^{F} \pi_f \sum_{w=1}^{W} \sum_{s=1}^{S} \tilde{p}_{f,w,s} - \pi_m P_m^{\mathrm{MBS}} \tag{2}$$

$$\text{s.t.} \quad R_{f,w}^{\mathrm{sum}} \geq R_{\mathrm{req}}, \quad \forall w \in \mathcal{W} \tag{27-1}$$

$$\sum_{w=1}^{W} \sum_{s=1}^{S} \hat{p}_{f,w,s} \leq E_f, \quad \forall f \in \mathcal{F} \tag{27-2}$$

$$\sum_{w=1}^{W} \sum_{s=1}^{S} p_{f,w,s} \leq P_{\max}^{\mathrm{RRH}}, \quad \forall f \in \mathcal{F} \tag{27-3}$$

$$\text{Constraints (19-4) to (19-8)} \tag{27-4}$$

Proof: Since $R_{f,w,s}$ is concave in $p_{f,w,s}$, the minimum data rate constraint in (27-1) is convex. Together with the remaining linear constraints, the feasible set of problem (27) is convex. Due to the concave nature of $R_{f,w,s}$, the monotonic increasing function $U_w(R_w)$ is also concave. The objective function in (27) is the difference between a sum of concave functions and a sum of linear functions, thus it is concave. Together with its convex feasible set, we can conclude that (27) is a convex optimization problem.

3.2.2 Algorithm Design Due to convexity, we can obtain the optimal solution of problem (27) through primal-dual methods, i.e., by maximizing its Lagrangian function and minimizing its corresponding dual function. Let $\eta = \{\eta_w | w \in \mathcal{W}\}$ denote the multiplier vector for the minimum data rate constraint in (19-1). We define the Lagrangian form of problem (27) as:

$$L(\mathbf{v}, \eta) = \sum_{w \in \mathcal{W}} U_w(R_w) - \sum_{f=1}^{F} \pi_f \sum_{w=1}^{W} \sum_{s=1}^{S} \tilde{p}_{f,w,s} - \pi_m P_m^{\mathrm{MBS}} + \sum_{w \in \mathcal{W}} \eta_w \left(R_{f,w}^{\mathrm{sum}} - R_{\mathrm{req}} \right)$$

The dual problem is:

$$\min_{\eta \geq 0} \quad g(\eta) = \max_{\mathbf{v}} L(\mathbf{v}, \eta)$$

Due to coupling between variables, it is difficult to obtain a closed-form optimal solution. We can solve (29) and (30) alternately through primal-dual variables. Let $\eta^{(n)}$ be the solution from the previous iteration of (30). In the current iteration, the optimal solution is updated to $\mathbf{v}^{(n+1)}$ by maximizing (29) through Algorithm 3, while the solution of (30) is updated to $\eta^{(n+1)}$ through subgradient method, i.e.:

$$\eta_w^{(n+1)} = \max \left(0, \eta_w^{(n)} - \beta \left(R_{f,w}^{\mathrm{sum}} - R_{\mathrm{req}} \right) \right), \quad \forall w \in \mathcal{W}$$

where β is the step size. The power allocation algorithm for solving (27) is presented as Algorithm 2.

Algorithm 2: Joint Sub-channel and Power Allocation Algorithm (JCPA) 1. Obtain user pairing scheme through Algorithm 1 and update sub-channel allocation matrix α 2. Initialize: randomly select $\mathbf{v}^{(0)}$, let $\eta^{(0)} = \mathbf{0}$, set iteration counter $n = 0$ 3. Repeat 4. $n \leftarrow n + 1$ 5. Update $\mathbf{v}^{(n)}$ through Algorithm 3 6. Update $\eta^{(n)}$ through equation (31) 7. Until stopping criterion is met: $\|\mathbf{v}^{(n)} - \mathbf{v}^{(n-1)}\| \leq \epsilon$ 8. Obtain optimal solution of (27), i.e., $\mathbf{v}^* = \mathbf{v}^{(n)}$

3.2.3 Convergence and Optimality Analysis Let W denote the number of users in the network. Each mobile user served by RRH simultaneously consumes renewable and grid energy, with total power consumption expressed as (25). To convert battery capacity constraint (19-2) into an equality constraint, we introduce non-negative variables τ_f for all RRHs, then (19-2) can be reformulated as (26). Let $L(\mathbf{v})$ denote the objective function in (27) and $\mathcal{V}$ denote the feasible set of $\mathbf{v}$. Since $\mathcal{V}$ is finite and $L(\mathbf{v})$ is concave and continuous in $\mathbf{v}$, the power allocation algorithm approaches a feasible solution of (27) at a rate of $1/n$, achieving a target value within ϵ of the optimal value, where n is the number of iterations and ϵ is defined in (33).

Proof: Let $\mathbf{v}^{(n)}$ denote the solution at the n-th iteration of solving (29) and define $\hat{\mathbf{v}}^{(n)}$ as the average of vector $\mathbf{v}^{(n)}$, i.e., $\hat{\mathbf{v}}^{(n)} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{v}^{(i)}$. Since (27) is strictly convex, the Euclidean norm of the multiplier vector $\eta^{(n)}$ is bounded. Inequality (35) follows when $\|\eta^{(n)}\| \leq \eta_{\max}$. From (34), we can see that $\hat{\mathbf{v}}^{(n)}$

approaches a feasible solution at a rate of $1/n$. When $n \to \infty$, we further obtain that the difference is less than the optimal target value, i.e., $|L(\hat{\mathbf{v}}^{(n)}) - L(\mathbf{v}^*)| \le \epsilon$. According to (32), we can obtain the optimal target value of (27) when the difference is less than ϵ.

3.3 ADMM-Based Lagrangian Maximization

In this section, we consider solving (29) in a distributed manner, where all base stations determine the optimal power allocation for each user through limited information exchange with users. This is highly practical for large-scale HetNets associated with numerous femto cells due to low implementation complexity and signaling overhead. Therefore, we propose an ADMM-based Lagrangian maximization algorithm (i.e., Algorithm 3) for the equality-constrained convex optimization problem (29), and prove the optimality and time complexity of the proposed algorithm.

3.3.1 Algorithm Design The proposed algorithm works as follows at each iteration: First, each RRH distributively updates power allocation by maximizing the augmented Lagrangian associated with (29). Then, multipliers regarding constraints (25)-(26) are updated based on the updated variables. Let θ denote the multiplier vector for constraints (25)-(26). For notational simplicity, define $\mathbf{A}$ and vector $\mathbf{c}$ as the coefficient matrix of variables and the vector of right-hand side coefficients in constraints (25)-(26), respectively. Thus, we denote the augmented Lagrangian of (29) as $L_\rho(\mathbf{v}, \theta)$, satisfying $L_\rho(\mathbf{v}, \theta) = L(\mathbf{v}) + \theta^T(\mathbf{A}\mathbf{v} - \mathbf{c}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{v} - \mathbf{c}\|^2$, where $\rho > 0$ is the penalty factor. After updating variables, multipliers are updated through $\theta^{(k+1)} = \theta^{(k)} + \rho(\mathbf{A}\mathbf{v}^{(k+1)} - \mathbf{c})$. Let $\mathbf{r}^{(k)} = \mathbf{A}\mathbf{v}^{(k)} - \mathbf{c}$ denote the primal space residual at iteration k, and the updated multiplier at iteration k is $\theta^{(k)}$, with $\theta^{(k+1)} = \theta^{(k)} + \rho\mathbf{r}^{(k+1)}$. The ADMM-based Lagrangian maximization algorithm terminates and obtains the optimal solution when alternating updates of $\mathbf{v}$ and θ meet the stopping criterion. Here we adopt the stopping criterion that the primal residual must be small, i.e., $\|\mathbf{r}^{(k)}\| \le \epsilon_{\text{pri}}$, where ϵ_{pri} is the feasible tolerance for the primal feasibility condition of (29).

Algorithm 3: ADMM-Based Lagrangian Maximization Algorithm for Solving (29) 1. Initialize: randomly select $\mathbf{v}^{(0)}$, let $\theta^{(0)} = \mathbf{0}$, set iteration counter $k = 0$ 2. Repeat 3. All RRHs distributively update power allocation for users through equations (38)-(40) 4. Update $\mathbf{v}^{(k+1)}$ by maximizing the augmented Lagrangian function (36) 5. Update $\theta^{(k+1)}$ through equation (41) 6. $k \leftarrow k+1$ 7. Until stopping criterion is met, i.e., $\|\mathbf{r}^{(k)}\| \le \epsilon_{\text{pri}}$ 8. Obtain optimal solution

Because the augmented Lagrangian function is concave in each variable in $\mathbf{v}$, we can further obtain the optimal solution of (37). Let $p_{f,w,s}^{(k+1)}$ be the root satisfying $\nabla_{p_{f,w,s}} L_\rho = 0$, and we can obtain $p_{f,w,s}^{(k+1)}$ as shown in (38)-(40). The multiplier update is given by (41), and the primal residual is calculated as (42).

3.3.2 Time Complexity Let ϵ denote the termination parameter for updating transmit power in the bisection search. From (29), the ADMM-based Lagrangian maximization algorithm requires $\mathcal{O}(\log(1/\epsilon))$ iterations to converge, and the complexity of updating variables in each iteration is $\mathcal{O}(W)$, where W denotes the number of users. Therefore, when applying bisection search to obtain the root of variables at each iteration, the time complexity of the proposed algorithm is $\mathcal{O}(W\log(1/\epsilon))$. Additionally, the algorithm' s time complexity also depends on the setting of penalty factor ρ, i.e., the complexity increases as ρ decreases.

4 Simulation and Analysis

For the proposed JCPA method, this section uses Matlab simulation tools to evaluate the performance of the proposed algorithm, maximum network utility, optimal energy consumption, and maximum system revenue. Since related literature has not proposed algorithms for the same optimization objective, numerical simulations are conducted by comparing with three other algorithms: equal power allocation method, the no-EH-JSPA algorithm from [**?**] that does not employ EH technology in NOMA systems, and the EH-OFDMA algorithm from [**?**] that employs EH technology in OFDMA systems. Equal power allocation means each hybrid-powered RRH uses maximum transmit power for each accessed user, then optimizes bandwidth allocation to each user. The no-EH-JSPA algorithm considers cross-tier/co-tier interference while using convex relaxation and dual decomposition techniques to optimize sub-channel and power allocation in NOMA HetNets to maximize overall system energy efficiency. The EH-OFDMA algorithm utilizes EH technology in small cell heterogeneous networks, introduces time-varying cross-tier/co-tier interference pricing, and uses a non-cooperative game approach to solve sub-channel and power allocation problems.

Following parameter settings from [**?**], the simulation parameters are set as shown in . All users and RRHs are uniformly distributed in a 1000m$×$1000m area, with the MBS located at coordinates (500m, 500m). Unless otherwise specified, each user' s minimum data rate requirement is set to 2Mbps, the battery capacity for storing renewable energy in hybrid-powered RRHs is 10 Joules, and the trade-off coefficient α is initially set to 0.5.

First, [Figure 3: see original paper] verifies the complexity and convergence of the proposed ADMM-based Lagrangian maximization algorithm under different penalty factor settings. The figure shows that as the number of iterations increases, $\|\mathbf{r}^{(k)}\|$ tends to zero, indicating that the ADMM-based Lagrangian maximization algorithm achieves primal residual convergence. Moreover, since $\|\mathbf{r}^{(k)}\|$ tends to zero, the sequence $\{\mathbf{v}^{(k)}\}$ can be obtained through (42). The curve can be fitted as $\|\mathbf{r}^{(k)}\| \approx \mathcal{O}(1/k)$, indicating that the ADMM-based Lagrangian maximization algorithm undergoes linear convergence to reach the stopping criterion. The time complexity also depends on the setting of penalty factor ρ, i.e., complexity increases as ρ decreases.

[Figure 4: see original paper] shows the optimality and convergence of the proposed JCPA algorithm, with step size β set to 10^{-6}. It can be seen that the proposed algorithm can achieve fast convergence, reaching global optimality within 10 iterations.

Next, we compare algorithm performance under different Mobile User (MU) densities. We consider a NOMA-HCRAN network with one MBS and three hybrid energy supply RRHs, fix the positions of the four base stations as shown in [Figure 5: see original paper], and place 50-100 MUs in this area, with parameter ϵ set to 10^{-4}.

[Figure 6: see original paper] shows that when using the proportional fair utility function as the network utility function, the network utility, total grid energy cost, and system revenue obtained by the proposed JCPA and three comparison algorithms all increase with the number of MUs, indicating that all schemes can effectively improve cell energy efficiency in NOMA-HCRAN networks. On one hand, the equal power allocation algorithm achieves the lowest system revenue, far below other schemes, because each hybrid-powered RRH provides maximum transmit power to each accessed user and then optimizes bandwidth allocation, resulting in maximum network utility at the cost of maximum grid power consumption. When the number of users increases rapidly, grid power consumption rises sharply, leading to lower system revenue. On the other hand, when the number of users reaches 100, the proposed JCPA improves system revenue by about 18.8% compared to no-EH-JSPA without EH units, because systems with EH units can use renewable energy stored in EH unit batteries during transmission. Compared with EH-OFDMA with EH units, the proposed JCPA improves system revenue by about 11.8%, because NOMA systems have higher spectral efficiency than OFDMA and utilize successive interference cancellation technology to eliminate interference, thereby achieving higher network utility. EH-OFDMA achieves higher system revenue than no-EH-JSPA in NOMA systems due to its EH units, which can significantly reduce system grid energy cost and thus improve system revenue. In summary, the proposed JCPA can achieve more system revenue than other schemes while spending the least grid energy cost, effectively improving system energy efficiency.

[Figure 7: see original paper] shows the performance of several algorithms when changing the number of RRHs. Here, we deploy the MBS at (500m, 500m) and randomly deploy 3-10 hybrid energy supply RRHs in the 1000m$×$1000m area, with other simulation parameters as above. The figure shows that except for equal power allocation, as the number of RRHs increases, the network utility and system revenue obtained by other algorithms increase, while total grid energy cost decreases, because each RRH is equipped with an energy harvesting unit that can provide energy supplement. Specifically, when using the proportional fair utility function, the proposed JCPA improves system revenue by at least about 13.3% and 8.0% compared to no-EH-JSPA and EH-OFDMA, respectively. Additionally, for equal power allocation, grid energy cost increases sharply as the number of RRHs increases, because this algorithm defaults to each RRH providing maximum transmit power to its served users, leading to reduced system revenue. This demonstrates that optimizing power allocation has a greater impact on improving system revenue than only optimizing frequency allocation at the same power, and both the proposed JCPA and other resource allocation algorithms optimize the transmit power allocated for each MU' s data transmission.

[Figure 8: see original paper] shows the performance comparison of several algorithms when changing the minimum user data rate requirement from 0.4Mbps to 2Mbps. The figure shows that, on one hand, compared with other algorithms, frequency allocation with equal power allocation achieves maximum network utility at the cost of more energy, resulting in the lowest system revenue. On the other hand, no-EH-JCPA and EH-OFDMA algorithms can increase network utility at the cost of more energy as the minimum data rate requirement increases, which also causes the system revenue of both schemes to decrease as R_{req} increases. In contrast, as the minimum data rate requirement increases, the proposed JCPA meets the minimum rate requirement by increasing grid energy cost, but the overall network utility improvement is not significant, and system revenue decreases slightly. However, the proposed algorithm always outperforms other algorithms, improving system revenue by at least about 7.5% compared to no-EH-JSPA.

[Figure 9: see original paper] shows algorithm performance when changing the trade-off coefficient α from 0.1 to 0.9. When using equal power allocation, changing α does not affect network utility and total grid energy cost, because with given power allocation, system revenue is limited only by network utility maximized through frequency optimization, which is independent of α. In contrast, other algorithms including the proposed JCPA show decreasing network utility, total grid energy cost, and system revenue as α increases. This is because the trade-off coefficient α reflects the negative impact of grid energy cost on system revenue; increasing α means we need to reduce grid energy cost to maximize system revenue, which further leads to decreased network utility. Moreover, the proposed algorithm always outperforms other algorithms. For example, compared with equal power allocation, the proposed algorithm can increase system revenue by at least 19.9%; compared with EH-OFDMA, it can increase by up to 24.9%.

[Figure 10: see original paper] shows the performance comparison between the proposed JCPA algorithm and full frequency reuse scheme when randomly deploying 10-30 MUs within each hybrid energy supply RRH' s coverage area. The figure shows that when using proportional fair utility, both algorithms increase network utility, total grid energy cost, and system revenue as the number of MUs per small cell increases. Additionally, the proposed algorithm' s system revenue approaches that of the full frequency reuse scheme, indicating that the proposed algorithm achieves high spectral efficiency but at the cost of more grid energy.

5 Conclusion

This paper studied the joint optimization of sub-channel and power allocation in NOMA-HCRAN networks considering EH technology and inter-node interference, aiming to maximize system revenue. To obtain a low-complexity suboptimal solution, the problem was decoupled into user pairing and power allocation sub-problems. First, the GA algorithm was used to pair two users in each sub-channel. Next, power allocation was formulated as an optimization problem maximizing the difference between network utility and grid energy cost (i.e., system revenue). Then, an ADMM-based Lagrangian maximization algorithm was used to obtain the optimal power allocation. Simulation results show that the proposed algorithm converges within 10 iterations, proving its effectiveness. Moreover, compared with existing algorithms, the proposed algorithm demonstrates superior performance.

Note: Figure translations are in progress. See original paper for figures.