Resource Allocation for Uplink NOMA-Based D2D Communication in Energy Harvesting Scenario: A Two-Stage Game Approach

Runzhou Li[®], Student Member, IEEE, Peilin Hong[®], Kaiping Xue[®], Senior Member, IEEE, Ming Zhang, Student Member, IEEE, and Te Yang

Abstract-Energy harvesting (EH) endows device-to-device (D2D) communication and cellular equipment with the ability of continuous communication to provide internet-of-things (IoT) services in natural areas. While the available energy, which relies on EH, becomes an extra nonnegligible factor in resource allocation. Besides, we integrate uplink non-orthogonal multiple access (NOMA) with D2D communication to provide multiple access for D2D transmitters for more efficient IoT service and more efficient utilization of limited spectrum. In this scenario, ingenious resource allocation approach is a key focus for utilizing the advantages in energy and spectral efficiency. Aiming to investigate the inherent resource allocation issue, we set our goal as maximizing the energy efficiency for both NOMA-based D2D groups and cellular users (CUs), where the power and spectrum allocation are both considered. Then we propose a two-stage game approach, which is theoretically proved to be capable of obtaining the equilibrium and a stable result, to solve the unilateral energy efficiency maximization problems. Besides, an energy-aware screening method is proposed to reduce the computations based on the available energy of user equipment. Finally, the effectiveness of our proposed method is verified through elaborated simulation results.

Index Terms—D2D communications, uplink NOMA, distributed resource allocation, game theory, energy harvesting.

I. INTRODUCTION

WITH the trend of connecting an ever-increasing number of devices in the future internet of things (IoT) scenarios, there will exist a large number of devices that are widely deployed in natural environments, who have demand for data uploading, information sharing, etc [1]. In this scenario, energy harvesting (EH) endows cellular communication and device-to-device (D2D) communication with the ability to

The authors are with the Key Laboratory of Wireless-Optical Communications, School of Information Science and Technology, University of Science and Technology of China, Hefei 230026, China (e-mail: lrzfff@mail.ustc.edu.cn; plhong@ustc.edu.cn; kpxue@ustc.edu.cn; mzhang95@mail.ustc.edu.cn; yangte@mail.ustc.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3100567.

Digital Object Identifier 10.1109/TWC.2021.3100567

continuously operate for data uploading, information sharing in areas where the power line is infeasible [2]-[4], thus this reveals a prominent application prospect in practice. Through harvesting energy from the ambient environment, like solar and wind energy [5], [6], cellular communication, information sharing through device-to-device (D2D) communication [2]-[4] can play a crucial part in IoT or other similar purposes in natural environments. In particular, with D2D communication, information can be shared between user equipment (UE) without traversing the base station (BS), where D2D links can reuse the spectrum of cellular users (CU) to improve the spectrum efficiency [7], [8]. For the prominent advantages in easy deployment and spectral and energy efficiency, EH-powered D2D communication has been invoked into various communication scenarios for IoT services, machinetype communications, etc [4], [9]–[13]. While the available energy, which relies on EH, becomes an extra nonnegligible factor in resource allocation. Except for the transmit power in resource allocation, the spectrum assignment can also be different under various energy levels. Besides, most existing researches that involve D2D communications only consider the D2D communication model with one D2D transmitter (DT) transmitting to the D2D receiver (DR) [4], [13], which is limited in practical implementation for IoT services. In real scenarios, IoT services or other applications like machinetype communications may have substantial demand for multiple access for information collection purposes. Therefore, the limitation of existing literatures in providing more effective multiple connectivity drives us to come up with an effective approach to tackle the multiple access issue for providing IoT services.

Attracted by the great potential of non-orthogonal multiple access (NOMA) technology in improving spectral efficiency as well as providing massive connectivity [14], we come up with the idea of integrating uplink NOMA into D2D communication for providing IoT services. By employing the successive interference cancellation (SIC) technique in the receiver, the NOMA receiver can decode the superimposed signal of multiple users in the same time and frequency following a certain decoding order. For these prominent advantages in spectral efficiency and multiple connectivity, some works begin to implement trials on integrating NOMA with D2D communications [15]–[18]. In existing literatures on

1536-1276 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Manuscript received August 27, 2020; revised January 21, 2021 and May 7, 2021; accepted July 20, 2021. Date of publication August 4, 2021; date of current version February 14, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61671420 and Grant 61972371 and in part by the Youth Innovation Promotion Association of Chinese Academy of Sciences (CAS) under Grant Y202093. The associate editor coordinating the review of this article and approving it for publication was M. Payaró. (*Corresponding author: Peilin Hong.*)

NOMA-based D2D communications, one DT and multiple DRs jointly form a NOMA group, which is generally called a D2D group. Nevertheless, all those existing researches only consider downlink NOMA in D2D communication, which cannot meet the many-to-one transmission in IoT scenarios. In IoT scenarios, there may exist multiple DTs transmitting to one DR for the purpose of data collection or other IoT services. The cruel reality is that there still lacks work on uplink NOMA-based D2D communications till now. Different from downlink NOMA, where a transmitter transmits to multiple receivers, uplink NOMA enables multiple transmitters to transmit to one receiver at the same time and frequency by employing SIC on the receiver side. In this paper, the powerdomain NOMA is adopted [19]. Thus how to efficiently accomplish power allocation for multiple uplink NOMA-based D2D users and coordinate the interference in spectrum reusing becomes a new crucial issue.

In this paper, we propose to investigate the resource allocation issue of EH-powered cellular users (CUs) and D2D groups in the cellular network. In a D2D group, since transmitted signals are superimposed on the same spectrum, there may exist serious internal interference between transmitted signals. Thus how to carefully control the power of DTs in a D2D group, especially in the EH scenario where energy levels can be time-variant, is a fatal issue that determines the gain of NOMA compared with OMA schemes. Besides, how to deal with the interference between D2D and cellular communications is also a crucial issue, and taking the available energy of UEs into consideration to realize more effective resource allocations should also be discussed.

To resolve the above issues, we propose to optimize the unilateral energy efficiency of both D2D groups and CUs through resource allocation. With this goal, we put forward a two-stage game approach to tackle the energy efficiency maximization problem in a distributed manner. In the first stage, we introduce an approximation method to model the power allocation between each potential pair of D2D group and CU as a noncooperative game. In this game, the D2D group and CU are both seen as players, where the D2D group that contains multiple D2D links is approximated as an entity. Through this method, we skillfully avoid the complicated combinatorial game and greatly reduce the computational complexity. After obtaining equilibrium in the first stage, we use a matching game to obtain a global stable matching result. To sum up, our contributions can be concluded as follows:

- We propose to integrate uplink NOMA with D2D communication to provide access for multiple DTs for IoT service. To investigate this issue, we set a target of accomplishing the joint resource allocation of uplink NOMA-based D2D groups and CUs. To our knowledge, the uplink NOMA-based D2D communication issue has been rarely investigated before.
- 2) Different from most existing works that solve optimization problems in a centralized manner, the optimization of UEs' energy efficiency occurs in a distributed way in our method. To achieve this goal, we put forward a two-stage game approach to deal with the joint power

and spectrum allocation problem, where the computation separately occurs in D2D groups and CUs, where the available energy of UEs is considered during the game.

- 3) We introduce an approximation method to formulate the first-stage game as a noncooperative game, instead of using a complicated coalitional game with high computational overhead, during the power allocation. With this approach, the computational complexity and signaling overhead can be greatly reduced.
- 4) Moreover, we present an intuitional analysis of the influence of the energy levels under EH on resource allocation, and then propose an energy-aware screening approach on the establishment of preference lists to further reduce computational complexity. Finally, elaborated simulation results are presented to verify the effectiveness of our proposed method.

The remainder of this paper is organized as follows. In Section II, we present a brief review of related works. In Section III, the system model is depicted, as well as the problem formulation. Section IV focuses on the proposed two-stage game-theoretical approach, and Section V provides elaborated analysis on the influence of EH on resource allocation and puts forward additional improvements on the proposed method. In Section VI, detailed simulation results are presented. Finally, we conclude our paper in Section VII.

II. RELATED WORK

With the popular trend of green communications and IoT, D2D communication and its application in EH scenario has received much attention [3], [4], [9], [10], [20]. Due to the time-variant characteristic of EH, considering the resource allocation of EH-powered communication is a realistic and crucial issue. Zhou et al. [3] proposed to utilize wireless EH-aided D2D communication for traffic offloading to alleviate the heavy burden on the fronthaul. Kuang et al. [9] investigated the energy-efficient resource allocation in EH-based D2D heterogeneous networks using a convex optimizationbased method. In [4], the authors took the randomness of EH capacity of DTs into consideration during resource allocation, and then proposed an iterative as well as a low-complexity algorithm to resolve this issue. Besides, Saleem *et al.* [10] also investigated the resource allocation issue in EH-powered D2D communications, the system sum rate is maximized through a low-complexity gain-based algorithm. Salim et al. [20] focused on the rate and energy tradeoff of EH-based D2D communication for IoT services. However, most of the above literatures concentrate on maximizing communication rates. Another fatal issue is how to energy efficiently transmit in the EH scenario, where the energy levels of UEs can be variant. This issue remains to be further investigated based on the additional energy state information. Moreover, existing works only considered the D2D communication between one DT and one DR, while it does not work well when facing the requirement of information collection or other similar IoT services between multiple DTs and one DR.

In addition, with the emerging of NOMA technology, researchers start trials on integrating NOMA with D2D communications for multiple access and higher spectral efficiency [15]-[17], [21]. Budhiraja et al. [16] aimed to deal with the delay and interference issue in NOMA-based D2D cooperative communication, and they proposed a twophased tactile internet-driven scheme to resolve the problem. In [15], the authors formulated a sum-rate maximization problem to accomplish the subchannel and power allocation in NOMA-enhanced D2D communication. Besides, Baidas et al. [21] considered the NOMA-enhanced D2D association and channel assignment issue in multi-cell uplink NOMA networks. Asides from using NOMA in D2D communication for transmission, the authors in [17] creatively proposed to exploit NOMA technique to cancel the received interference signal in D2D communications. Although NOMA-based D2D communication is expected to bring improvements on system spectral efficiency performance, the price is the careful design of power allocation schemes. However, rare literatures considered the energy efficiency during power allocation in NOMA-based D2D communication scenarios. Besides, downlink NOMA is mostly adopted in the above literatures, but integrating uplink NOMA with D2D communication to provide access for multiple DTs is still an unexploited issue.

III. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we first describe the system model, where EH-powered NOMA-based D2D groups and CUs coexist in a cellular area. Then, the EH model is used to express the available energy of UEs, and the mathematical expressions of the communication rates are introduced.

A. System Model

We consider a cellular area with EH-powered DTs, DRs and EH-powered CUs in Fig. 1. Due to the low-power property of the decoding process, the power consumption of a DR is generally omitted. The EH-powered CUs are distributed in the cellular area, and the set of CUs is denoted as C = $\{k | k = 1, \dots, K\}$, where K is the number of CUs. Each CU is allocated with an orthogonal spectrum in advance for uplink transmission. Under this premise, the D2D links transmit through reusing the uplink spectrum of CUs. In a NOMA-based D2D group, which consists of a number of DTs that are connected to one DR, the EH-powered DTs transmit to the DR simultaneously over the reused uplink spectrum of a certain CU. Herein, we assume that there is only one DT cluster [22] in a NOMA-based D2D group for brevity in mathematical analysis.¹ By taking advantage of the equipped SIC techniques in receiving modules, DRs can decode the superimposed signals of multiple DTs in the same time and frequency resource. The set of DRs is denoted as \mathcal{R} = $\{i | i = 1, \dots, M\}$, where M is the number of DRs. Besides,



Fig. 1. System model of underlay D2D groups reusing the uplink spectrum of CUs in a cell, where CUs transmit in cellular uplink mode.

we denote the set of DTs as $\mathcal{T} = \{j | j \in \mathcal{G}_i, \forall i \in \mathcal{R}\}$, where \mathcal{G}_i stands for the set of DTs that are connected to DR *i*. When the DTs are associated with a DR *i*, we call this uplink NOMA-based D2D group as D2D group *i*. For brevity, we use "D2D group" to stand for "uplink NOMA-based D2D group" in the following.

To describe the energy level of DTs, the energy model of DT j in t-th time slot is expressed in the following. Without loss of generality, the length of a time slot Δt is normalized as 1. Since the resource allocation is conducted at the beginning of each time slot [15], [23], thus we have the transmit power satisfying the following energy constraint:

$$p_j^{D,t}/\eta \le \phi_j^{D,t}, (\phi_j^{D,t} = [B_j^{D,t} - p_{cir}^t]^+\}), \tag{1}$$

where $[x]^+ = \max\{x, 0\}, B_j^{D,t} = B_j^{D,t-1} + E_j^{D,t-1} - p_j^{D,t-1}/\eta - p_{cir}^t, p_j^{D,t}$ is the transmit power of DT j in t-th time slot, η is the power amplifier efficiency, and p_{cir}^t values zero when $\phi_j^{D,t} = 0$ because the transmission of DT j is infeasible.² For the DTs, we use $B_j^{D,t}$ to denote the battery level of DT j in time slot t, and $E_j^{D,t}$ is the harvested energy in the t-th time slot. Note that, the harvested energy can be fully used at the time of power allocation. The battery capacity is large enough for every quanta of incoming energy can be stored in the battery, and this assumption is practically valid for the current circumstance of the technology where batteries have much larger capacities compared to the efficiency of harvested energy flow [4], [6]. Note that since we normalize the expression of time length, the above energy constraint is actually established in the aspect of power. As for $E_j^{D,t-1}$, it is modeled as $E_j^{D,t-1} = \sum_{n=1}^{N(t-1)} U_j^n$ [24], where U_j^n is the size of the n-th energy packet in N(t-1), and the number of energy arrivals N(t) is assumed to be independent identically distributed, obeying a homogeneous Poisson process with parameter λ_j^D . Since we focus on the resource allocation

¹For brevity, we assume that a D2D group only contains one NOMA-based DT cluster in the following analysis, where the DTs in a DT cluster perform NOMA through reusing the spectrum of a certain CU. This can be easily extended to the scenario where a D2D group contains multiple DT clusters, and those DT clusters transmit to the same DR by reusing the spectrum of different CUs [22]. Then, a D2D group that contains multiple DT clusters is mathematically equivalent to multiple D2D groups that each contains one DT cluster [23].

 $^{^{2}}$ For conciseness, the following analysis assumes that all DTs are feasible in transmission, this assumption is obtained by excluding the infeasible DT from the D2D group.

in each time slot, we omit superscript t for brevity in the following, for example, p_j^D and B_j^D denote DT j's transmit power and battery level.

Similarly, the available energy for transmit power of CU k can be depicted as:

$$p_k^{C,t}/\eta \le \phi_k^{C,t}, (\phi_k^{C,t} = [B_k^{C,t} - p_{cir}^t]^+),$$
 (2)

where $B_{kC,t}^{C,t} = B_{kC,t-1}^{C,t-1} + E_{k}^{C,t-1} - p_{k}^{C,t-1}/\eta - p_{cir}^{t}$. For brevity, $p_{k}^{C,t}$ and $B_{k}^{C,t}$ are expressed as p_{k}^{C} , B_{k}^{C} for brevity in the following. Note that, the power allocation algorithm is conducted at the beginning of each time slot in D2D communications [15], [23], the power allocation result is obtained under the current energy level of the battery, which is jointly determined by the battery level at the beginning of the last time slot, consumed power and harvested energy in the last time slot.

In this paper, the underlay D2D mode is adopted, i.e., D2D communications reuse the uplink spectrum of CUs. To the interest of D2D groups, reusing the uplink spectrum of which CU to be less interfered remains to be an unresolved issue. In addition, the CUs, whose uplink spectrum is reused, also want to choose the D2D groups that bring about less interference to themselves. To formulate the expressions of the spectrum choices of D2D groups, we introduce another binary integer value $y_{i,k} = \{0,1\}$. When $y_{i,k} = 1$, it means that D2D group i reuses the uplink spectrum of CU k, and the value 0 means not choosing. To avoid serious interference between the D2D groups, one CU's uplink spectrum is only allowed to be reused by one D2D group. Besides, we also assume that one D2D group is only allowed to reuse one CU's spectrum [10]. Thus, the mathematical constraints of spectrum reusing relations can be depicted as;

$$\sum_{k\in\mathcal{C}} y_{i,k} \le 1,\tag{3}$$

$$\sum_{i \in \mathcal{R}} y_{i,k} \le 1. \tag{4}$$

Based on the above descriptions of spectrum reusing relations, we can proceed to formulate the communication rates of links. The Shannon formula is adopted to describe the achievable rates of communication links, thus the achievable rate of CU k can be expressed as

$$r_k^C = \log_2(1 + \frac{p_k^C g_k}{I_k^D + n}),$$
(5)

where $I_k^D = \sum_{\substack{j:j \in \mathcal{G}_i \\ y_{i,k}=1}} p_j^D g_{j,B}$ is the received superimposed interference on CU k's signal in the BS and n is the Gaussian additional noise. Besides, $g_{j,B} = L_{j,B}^{-\alpha} |h_{j,B}|^2$ denotes the interference channel gain from DT j to the BS, where $L_{j,B}$ is the distance from DT j to the BS, α is the path-loss exponent and $h_{j,B} \sim \mathcal{CN}(0,1)$ is the small-scale fading. Similarly, the cellular channel gain of CU k can be expressed as $g_k = L_k^{-\alpha} |h_k|^2$.

Besides, for a DT j in a D2D group, its achievable D2D communication rate can be formulated as follows:

$$r_j^D = \log_2(1 + \frac{p_j^D g_{j,\mathcal{I}(j)}}{I_j^D + I_{\mathcal{I}(j)}^C + n}),$$
(6)

where $\mathcal{I}(j)$ is a function that returns the associated DR of DT $j, g_{j,\mathcal{I}(j)} = L_{j,\mathcal{I}(j)}^{-\alpha} |h_{j,\mathcal{I}(j)}|^2$. In uplink NOMA-based D2D groups, decoding the signal of a DT will receive interference from the DTs whose signals are to be decoded latter. Thus, $I_j^D = \sum_{\substack{j':j' \in \mathcal{G}_{\mathcal{I}(j)} \\ \mathcal{S}(j) < \mathcal{S}(j')}} p_{j'}^D g_{j',\mathcal{I}(j)}$ is adopted to represent the received intra-D2D group interference when decoding DT j's signal at its associated DR *i*, wherein $\mathcal{S}(j)$ returns the decoding sequence number. Besides, $I_{\mathcal{I}(j)}^C = \sum_{k \in \mathcal{C}} y_{\mathcal{I}(j),k} p_k^C g_{k,\mathcal{I}(j)}$ stands for the received interference from CUs. Similarly, $g_{k,\mathcal{I}(j)} = L_{k,\mathcal{I}(j)}^{-\alpha} |h_{k,\mathcal{I}(j)}|^2$ is the interference channel gain from CU k to DR $\mathcal{I}(j)$, where $h_{k,\mathcal{I}(j)} \sim \mathcal{CN}(0, 1)$.

B. Problem Formulation

From each D2D group's and CU's perspective, due to their selfish nature, each of them wants to energy efficiently transmit information. Thus, we propose to separately maximize the energy efficiency for each D2D group and CU to obtain an energy-efficient resource allocation result.

With the above goal, the mathematical problems are formulated in the following contents. For CU $k \in C$ that transmits to the BS via cellular uplink, its energy efficiency maximization problem can be expressed as:

$$\mathcal{P}_k^C: \max_{\mathbf{y}_k, p_k^C} \frac{r_k^C}{\frac{p_k^C}{2} + p_{cir}},\tag{7a}$$

$$r_k^{\mathcal{C}} \ge r_{k,th}^{\mathcal{C}}, \quad k \in \mathcal{C},$$
 (7c)

$$\sum_{i \in \mathcal{R}} y_{i,k} \le 1, \quad k \in \mathcal{C}, \tag{7d}$$

$$0 \le p_k^C \le p_{k,\max}^C, \quad k \in \mathcal{C}, \tag{7e}$$

where $\mathbf{y}_k = \{y_{i,k} | i \in \mathcal{R}\}, p_{cir}$ is the static circuit power, $r_{k,th}^C$ is the communication rate threshold of CU k and $p_{k,\max}^C$ is the maximum transmit power of CU k. By observing the transmit power constraints in \mathcal{P}_k^C , it is easy to merge (2) and (7e) into one constraint as:

$$0 \le p_k^C \le \min\{\eta \phi_k^C, p_{k,\max}^C\}, \quad k \in \mathcal{C}.$$
(8)

Besides, for D2D group *i*, the energy efficiency maximization problem can also be formulated as:

s.

$$\mathcal{P}_{i}^{D}: \max_{\mathbf{y}_{i}, \{p_{j}^{D}|j \in \mathcal{G}_{i}\}} \frac{\sum_{j \in \mathcal{G}_{i}} r_{j}^{D}}{\sum_{j \in \mathcal{G}_{i}} \left(\frac{p_{j}^{D}}{\eta} + p_{cir}\right)},$$
(9a)

$$t. (1),$$
 (9b)

$$r_j^D \ge r_{j,th}^D, \quad j \in \mathcal{G}_i,$$
 (9c)

$$\sum_{k \in \mathcal{C}} y_{i,k} \le 1, \quad j \in \mathcal{G}_i, \tag{9d}$$

$$0 \le p_j^D \le p_{j,\max}^D, \quad j \in \mathcal{G}_i, \qquad (9e)$$

where $\mathbf{y}_i = \{y_{i,k} | k \in \mathcal{C}\}, r_{j,th}^D$ is the communication rate threshold of DT $j, p_{j,\max}^D$ is the maximum transmit power of DT j. Similarly, we can also merge (1) and (9e) into one constraint:

$$0 \le p_j^D \le \min\{\eta \phi_j^D, p_{j,\max}^D\}, \quad j \in \mathcal{G}_i.$$
(10)

Based on the mathematical problems formulated above, it is easy to find that there exist contradictions between maximizing the energy efficiency of D2D groups and CUs. For example, if D2D group i is arranged to reuse the uplink spectrum of CU k, then D2D group i and CU k will both compete for higher energy efficiency for their own benefit. Once D2D group iwants to increase the transmit power of its DTs to obtain higher energy efficiency, more interference will be received on CU k's signal, and vice versa. Therefore, we will try to construct the aforementioned relation in resource allocation as a game in the following, where the D2D groups and CUs will be seen as "players".

IV. A TWO-STAGE GAME-THEORETICAL APPROACH

In this section, we focus on solving the spectrum and power allocation problem. We first introduce an approximation approach to construct the competition between a pair of D2D group and CU as a noncooperative game. Thus, by applying noncooperative games in each potential spectrum reusing pair of D2D group and CU, each player can obtain a preference level between each other based on the noncooperative game results. Based on the preference results in the first stage, a matching game is applied to accomplish the spectrum allocation in the second stage.

A. An Approximation-Based Approach for Applying Noncooperative Game

Let us first concentrate on the formulation of problem \mathcal{P}_i . By substituting (1) and (9e) with (10), problem \mathcal{P}_i can be contracted as:

$$\mathcal{P}_{i}^{D}: \max_{\mathbf{y}_{i}, \{p_{j}^{D} \mid j \in \mathcal{G}_{i}\}} \frac{\sum\limits_{j \in \mathcal{G}_{i}} r_{j}^{D}}{\sum\limits_{j \in \mathcal{G}_{i}} \left(\frac{p_{j}^{D}}{\eta} + p_{cir}\right)}, \qquad (11a)$$

$$s.t. (9c), (9d), (10)$$
 (11b)

Remark 1: By observing problem \mathcal{P}_i^D $(i \in \mathcal{R})$ and \mathcal{P}_k^C $(k \in \mathcal{C})$, it is easy to find that the intrinsic joint spectrum and power allocation problem is a mixed-integer nonlinear programming (MINLP) problem, which is NP-hard. Besides, the D2D groups and CUs all want to maximize their own energy efficiency, thus making this issue a multi-objective problem, where only Pareto optimum can be achieved [25].

In an uplink NOMA-based D2D group, after receiving the superimposed signals of multiple DTs, the DR will apply the SIC technique to sequentially decode the DTs' signals. Same as [22], [26], the decoding order of signals from DTs is set as the descending order of DTs' channel gain to the DR.³

In the following, let us transfer our attention to the transmit power control strategies. For each paired D2D group and CU in spectrum reusing, they both want to contend for higher energy efficiency for their own benefit. However, the cruel reality is that their transmitted signals will interfere with the useful signal of each other. Once one side of them wants to improve its transmit power to obtain higher energy efficiency, more interference will be simultaneously received on the other side. Facing this contradictory situation, we come up with the idea of depicting the relationship between the paired D2D group and CU with game theory.

In the traditional noncooperative games, each action is taken autonomously by a single player [27]. However, in our scenario, a D2D group is constituted by multiple DTs, where each DT will transmit with a certain power. From the perspective of a CU, the observed action of a D2D group is actually a vector of the transmit power of DTs. In terms of this vector, it outstands in intractability compared with a traditional convex scalar, where the intractability is reflected in its infinite number of combinations and non-continuity [28]. As a result, the above characteristics make directly employing the noncooperative game approach infeasible in this scenario. Hence, how to resolve this serious obstacle becomes the primary issue. In the following, we will introduce an approximation-based approach to formulate this problem as a noncooperative problem.

To further analyze the energy efficiency optimization problem in \mathcal{P}_i^D , we first expand the objective function through substituting the communication rate with the expanded expression in (6). Thus, we have:

$$\max_{\mathbf{y}_{i},\{p_{j}^{D}|j\in\mathcal{G}_{i}\}} \frac{\sum_{j\in\mathcal{G}_{i}} r_{j}^{D}}{\sum_{j\in\mathcal{G}_{i}} \left(\frac{p_{j}^{D}}{\eta} + p_{cir}\right)}$$
$$= \max_{\mathbf{y}_{i},\{p_{j}^{D}|j\in\mathcal{G}_{i}\}} \frac{\log_{2}\left(\frac{\sum_{j\in\mathcal{G}_{i}} p_{j}^{D}g_{j,i}+I_{i}^{C}+n}{I_{i}^{C}+n}\right)}{\sum_{j\in\mathcal{G}_{i}} \left(\frac{p_{j}^{D}}{\eta} + p_{cir}\right)}.$$
(12)

Remark 2: Actually, for the *m*-th decoded DT (denoted as j_m here) in the D2D group, its achievable rate is denoted as $r_{j_m}^D$, thus the total achievable rate of the D2D group can be expressed as:

$$\begin{split} \sum_{n=1}^{|\mathcal{G}_i|} r_{j_m}^D &= \sum_{m=1}^{|\mathcal{G}_i|} \log_2(1 + \frac{p_{j_m}^D g_{j_m,i}}{\sum_{l=m+1}^{|\mathcal{G}_i|} p_{j_m}^D g_{j_m,i} + I_i^C + n}) \\ &= \sum_{m=1}^{|\mathcal{G}_i|} \log_2(\frac{p_{j_m}^D g_{j_m,i} + \sum_{l=m+1}^{|\mathcal{G}_i|} p_{j_l}^D g_{j_l,i} + I_i^C + n}{\sum_{l=m+1}^{|\mathcal{G}_i|} p_{j_l}^D g_{j_l,i} + I_i^C + n}}) \\ &= \sum_{m=1}^{|\mathcal{G}_i|} \log_2(\frac{\sum_{l=m}^{|\mathcal{G}_i|} p_{j_l}^D g_{j_l,i} + I_i^C + n}{\sum_{l=m+1}^{|\mathcal{G}_i|} p_{j_l}^D g_{j_l,i} + I_i^C + n}}) \\ &= \log_2(\frac{\sum_{l=1}^{|\mathcal{G}_i|} p_{j_l}^D g_{j_l,i} + I_i^C + n}{I_i^C + n}}). \end{split}$$

Therefore, we obtain the energy efficiency expression shown in (12).

Based on the structure of the above objective function, we have a theorem on the power allocation strategy of D2D groups as below.

Theorem 1: For a D2D group, with a given total amount of power of the DTs, it is optimal to preferentially allocate power to the DT that has the highest channel gain to the DR to achieve higher energy efficiency.

³Note that our main focus in this paper is the resource allocation, including spectrum and power allocation, for uplink NOMA-based D2D communications underlaying cellular network, and the decoding order is out of the scope of resource allocation issue and will be investigated in our future work.

Proof: With the energy efficiency expression of the D2D group in the DR, i.e., (12), it is easy to find that under a given total amount of power of the DTs, only giving priority to the DT that has the highest channel gain to DR to increase its transmit power first can make the nominator of the objective function increase most. Herein, we have to mention the premise of this is the power allocation should be constrained in its feasible domain, which is determined by the constraints in (11b).

Based on Theorem 1, we introduce a lemma on the objective function of \mathcal{P}_i^D ($i \in \mathcal{R}$).

Lemma 1: The objective function of \mathcal{P}_i^D can be expressed as a quasi-concave function about the total transmit power of a D2D group $p_{\mathcal{G}_i}^D = \sum_{j=1}^{n} p_j^D$.

Proof: Firstly, we introduce an equivalent expression of the objective function in (11a). Assume a certain given power set of DTs as $(\hat{p}_j^D), j \in \mathcal{G}_i$, thus we have $\hat{p}_{\mathcal{G}_i}^D = \sum_{j \in \mathcal{G}_i} \hat{p}_j^D$. According to Theorem 1, if we introduce increment Δp on the transmit power of DTs, the optimal strategy is allocating Δp to the one with the highest channel gain to the DR among the feasible DTs. To dig into the relation between the objective function of (11a) and $p_{\mathcal{G}_i}^D$, we define a function about $p_{\mathcal{G}_i}^D$ as

$$\varphi(\hat{p}_{\mathcal{G}_{i}}^{D} + \Delta p) \stackrel{\Delta}{=} \frac{\xi(\hat{p}_{\mathcal{G}_{i}}^{D} + \Delta p)}{\frac{\hat{p}_{\mathcal{G}_{i}}^{D} + \Delta p}{\eta} + |\mathcal{G}_{i}| p_{cir}} \\
\stackrel{\Delta}{=} \frac{\log_{2}(1 + \frac{j \in \mathcal{G}_{i}, p_{j}^{D} \in D(j)}{I_{i}^{C} + n})}{\sum_{j \in \mathcal{G}_{i}} \left(\frac{p_{j}^{D}}{\eta} + p_{cir}\right)} \bigg|_{p_{\mathcal{G}_{i}}^{D} = \hat{p}_{\mathcal{G}_{i}}^{D} + \Delta p} \\
\stackrel{(*)}{=} \frac{\log_{2}(1 + \frac{\sum_{j \in \mathcal{G}_{i}} \hat{p}_{j}^{D} g_{j,i} + \max_{j \in \mathcal{G}_{i}, p_{j}^{D} \in D(j)} g_{j,i} \Delta p}{I_{i}^{C} + n})}{\sum_{j \in \mathcal{G}_{i}} \left(\frac{\hat{p}_{j}^{D}}{\eta} + p_{cir}\right) + \frac{\Delta p}{\eta}}, \tag{13}$$

where $\forall j \in \mathcal{G}_i, \ \hat{p}_j^D \in D(j), \ D(j)$ stands for the feasible domain of p_j^D . Let us introduce another power increment $\Delta p'$ on $\hat{p}_{\mathcal{G}_i}^D$ satisfying $\Delta p' \geq \Delta p$.

Assume that the maximization operation in (*) is $g_{j,i}^* \Delta p$. Thus, with a D2D group power of $p_{G_i}^D = \hat{p}_{G_i}^D + \Delta p'$, the above maximization returns a value of $g_{j,i}^* \Delta p + g_{j',i}^* (\Delta p' - \Delta p)$. Based on Theorem 1, we have $g_{j',i}^* \leq g_{j,i}^*$. Thus, $\xi(\hat{p}_{G_i}^D + \frac{\Delta p'}{\eta}) \leq \xi(\hat{p}_{G_i}^D + \frac{\Delta p}{\eta}) + (\Delta p' - \Delta p) \frac{d\xi(p_{G_i}^D)}{p_{G_i}^D} \bigg|_{p_{G_i}^D = \hat{p}_{G_i}^D + \Delta p}$ holds,

where equality holds when $\Delta p' = \Delta p$. Besides, the feasible domain of $\hat{p}_{\mathcal{G}_i}^D$ is obviously convex. Thus, $\xi(\hat{p}_{\mathcal{G}_i}^D)$ is concave about $\hat{p}_{\mathcal{G}_i}^D$ [29]. Based on this, by referring to Problem 4.7 in [29], the fractional form function $\varphi(p_{\mathcal{G}_i}^D)$ in (13) is quasi-concave.

Based on the above analysis, we obtain an equivalent problem to \mathcal{P}_i^D that can be depicted as:

$$\hat{\mathcal{P}}_{i}^{D}: \max_{\mathbf{y}_{i}, p_{\mathcal{G}_{i}}^{D}} \varphi(p_{\mathcal{G}_{i}}^{D}) = \max_{\mathbf{y}_{i}, p_{\mathcal{G}_{i}}^{D}} \frac{\xi(p_{\mathcal{G}_{i}}^{D})}{\frac{p_{\mathcal{G}_{i}}^{D}}{\eta} + |\mathcal{G}_{i}| p_{cir}}, \quad (14a)$$

s.t. (9d), (14b)

$$p_{\mathcal{G}_i}^D \in D(\mathcal{G}_i),\tag{14c}$$

where $D(\mathcal{G}_i)$ is the feasible domain of $p_{\mathcal{G}_i}$. Note that, for a certain $p_{\mathcal{G}_i}^D$ in $\varphi(p_{\mathcal{G}_i}^D)$, there only exists a unique feasible transmit power vector of DTs.

Remark 3: In terms of $D(\mathcal{G}_i)$, assume that the spectrum allocation result of \mathbf{y}_i is predetermined, then $D(\mathcal{G}_i)$ can be obtained according to constraint (9c) and (10). For the m-th decoded DT $j_m \in \mathcal{G}_i(0 \le m \le |\mathcal{G}_i|)$, constraint (9c) can be expanded as a constrain on its transmit power as

$$\frac{2^{r_{j_m,th}^D} - 1}{g_{j_m,i}} \left(\sum_{l=m+1}^{|G_i|} p_{j_l}^D g_{j_l,i} + I_i^C + n\right) \le p_{j_m}^D \\
\le \min_{l_0=1,\dots,m} \left(\frac{p_{j_{l_0-1}}^D g_{j_{l_0-1},i}}{2^{r_{j_{l_0-1},th}} - 1} - \sum_{\substack{l=l_0,\\l \neq m}}^{|G_i|} p_{j_l}^D g_{j_l,i} - I_i^C - n\right) / g_{j_m,i},$$
(15)

then combine power constraint in (10), the range of DTs in D2D group *i* can be obtained. Since $p_{\mathcal{G}_i}^D = \sum_{j \in \mathcal{G}_i} p_j^D$, thus $D(\mathcal{G}_i)$ can be obtained. Therefore, $D(\mathcal{G}_i)$ is actually jointly determined by the energy level of EH-powered DTs, maximum transmit power and the above decoding constraints.

As mentioned before, the vector form of the transmit power of the DTs in a D2D group is a significant obstacle in employing a noncooperative game-theoretical approach. In order to tackle this issue, we propose to use an approximated channel gain to describe the interference channel gain from the DTs in a D2D group to CUs.

Recall the aforementioned system model, a D2D group is generally constrained in a relatively small area compared with the cellular area, this is because D2D communication range is generally within tens of meters. Besides, the DTs are normally much further from the BS compared with the D2D communications distance [2]. Thus, we introduce an approximated interference channel gain from the D2D group to the BS by referring to [30]. The multiple interference channel gains from DTs in D2D group *i* to the uplink communication of CU k can be approximately formulated as:

$$\tilde{g}_{i,B} \stackrel{\Delta}{=} \frac{I_k^D}{p_{\mathcal{G}_i}^D} = \frac{\sum_{j \in \mathcal{G}_i} p_j^D g_{j,B}}{\sum_{j \in \mathcal{G}_i} p_j^D}.$$
(16)

Remark 4: For the approximated interference channel gain above, the value of $\tilde{g}_{i,B}$ can be adjusted periodically after a certain number of adjustments of $p_{G_i}^D$.

With this approximation method, the approximated interference to cellular uplink communication can be derived as $\tilde{I}_k^D = \tilde{g}_{i,B} p_{\mathcal{G}_i}^D$. Thus, for CU k, the original problem \mathcal{P}_k^C can be formulated as:

$$\hat{\mathcal{P}}_{k}^{C}: \max_{\mathbf{y}_{k}, p_{k}^{C}} \frac{\tilde{r}_{k}^{C}}{\frac{p_{k}^{C}}{\eta} + p_{cir}} \stackrel{\Delta}{=} \max_{\mathbf{y}_{k}, p_{k}^{C}} \frac{\log_{2}(1 + \frac{p_{k}^{C} g_{k}}{\tilde{I}_{k}^{D} + n})}{\frac{p_{k}^{C}}{\eta} + p_{cir}}, \quad (17a)$$

$$s.t. (7d), (8),$$
 (17b)

$$\tilde{r}_k^C \ge r_{k,th}^C, \quad k \in \mathcal{C}.$$
(17c)

Theorem 2: If D2D group *i* reuses the spectrum of CU k, the above problem \mathcal{P}_i' and \mathcal{P}_k' jointly form a noncooperative game.

Proof: According to Lemma 1, and assume D2D group i reuses the spectrum of CU k, i.e., $y_{i,k} = y_{k,i} = 1$, then the objective function of $\hat{\mathcal{P}}_i^D|_{y_{i,k}=1}$ is quasi-concave, and the feasible domain of $p_{\mathcal{G}_i}^D$ is convex. Similarly, $\hat{\mathcal{P}}_k^C|_{y_{k,i}=1}$ also satisfies the above conditions. Thus, according to [31], $\hat{\mathcal{P}}_i^D|_{y_{i,k}=1}$ and $\hat{\mathcal{P}}_k^C|_{y_{k,i}=1}$ jointly form a noncooperative game.

B. The First Stage: An Approximation-Based Noncooperative Game Approach

Based on the transformations in the above subsection, we finally obtain an equivalent optimization problem $\hat{\mathcal{P}}_i^D$, where the D2D group takes action in the form of a scalar rather than a vector of power. Therefore, a noncooperative game can be exploited to depict the power allocation process between each pair of D2D group and CU. Firstly, let us assume D2D group *i* reuses the uplink spectrum of CU *k*, i.e., $y_{i,k} = 1$. Thus we can define the noncooperative game between them in strategic form as $G = (\mathcal{N}, (s_\tau)_{\tau \in \mathcal{N}}, (u_\tau)_{\tau \in \mathcal{N}})$, where $\mathcal{N} = \{\mathcal{G}_i, k\}$ is the set of players, $(s_\tau)_{\tau \in \mathcal{N}}$ is the set of strategies of players and $(u_\tau)_{\tau \in \mathcal{N}}$ is the set of utilities of players. Herein, the strategy s_l of each player corresponds to their transmit power, i.e., $s_k \triangleq p_k^C$ and $s_{\mathcal{G}_i} \triangleq p_{\mathcal{G}_i}^D$, where $s_\tau \in D(\tau)(\tau = \mathcal{G}_i, k)$. And $u_\tau(\tau = \mathcal{G}_i, k)$ is the energy efficiency function in (14a) and (17a), i.e., $u_{\mathcal{G}_i} \triangleq \varphi(p_{\mathcal{G}_i}^D)$, $u_k \triangleq \frac{\frac{p_k^C}{p_k}}{\frac{p_k^C}{\eta} + p_{cir}}$.

For a pair of D2D group *i* and CU *k* in a game, their utilities can be expressed as the function of players' strategies as $u_{\tau}(s_{\tau}, s_{-\tau})$, where $s_{-\tau}$ is the power strategy of the player except player τ . Thus, for each player τ , its objective can be expressed as

$$\max u_{\tau}(s_{\tau}, s_{-\tau}), \quad \forall \tau \in \mathcal{N}.$$
(18)

Obviously, each player's utility depends not only on their own transmit power, but also on the other player's strategy. Thus, we introduce a best-response function of a player τ based on the profile of strategy $s_{-\tau}$, expressed as

$$b_{\tau}(s_{-\tau}) = \{s_{\tau} \in D(\tau) | u_{\tau}(s_{\tau}, s_{-\tau}) \ge u_{\tau}(s'_{\tau}, s_{-\tau}), \forall s'_{\tau} \in D(\tau)\}.$$
(19)

As for the best-response function, its objective is to obtain the optimal power allocation of a player under the strategy of the other player, i.e., $b_k(s_{-k}) = \underset{p_k^C}{\arg\max} \frac{\tilde{r}_k^C}{\frac{p_k^C}{\eta} + p_{cir}},$ $b_{\mathcal{G}_i}(s_{-\mathcal{G}_i}) = \underset{p_{\mathcal{G}_i}^D}{\arg\max} \varphi(p_{\mathcal{G}_i}^D)$. Obviously, both maximization objective functions are in fractional form and thus not concave. To solve the non-concave maximization problems, we intro-

To solve the non-concave maximization problems, we introduce variable q_k and $q_{\mathcal{G}_i}$ to transform the objective functions into concave ones [32]. For CU k, we have $q_k = \frac{\tilde{r}_k^C}{\frac{p_k^C}{\eta} + p_{cir}}$, thus its optimal result obeys the following theorem. Algorithm 1 Iterative Algorithm in CU's Best-Response Function

1 Initialize: for CU k, set $q_k = 0$, $F_{q_k} = \infty$, $\delta = 10^{-2}$ and a given strategy $p_{q_i}^D$ of D2D group i;

2 while
$$F_{q_k} \ge \delta$$
 do
3 $\left|\begin{array}{c} q_k = \frac{\tilde{r}_k^C}{\frac{p_k^C}{\eta} + p_{cir}}; \text{ calculate } F_{q_k} = \max_{p_k^C} \tilde{r}_k^C - q_k (\frac{p_k^C}{\eta} + p_{cir}) \\ \text{with constraint (21b);} \\ \text{5 end} \\ \text{6 Return } q_k^* = q_k; \end{array}\right|$

Algorithm 2 Iterative Algorithm in D2D Group's Best-Response Function

Theorem 3: The optimal result can be obtained if and only if:

$$\max_{p_k^C} = \tilde{r}_k^C - q_k (\frac{p_k^C}{\eta} + p_{cir}) = \tilde{r}_k^{C*} - q_k^* (\frac{p_k^{C*}}{\eta} + p_{cir}) = 0,$$
(20)

where the variables with superscript * means its value under the optimal argument.

Proof: It is similar to the proof in [32]. \Box Thus, for CU k, we have a concave optimization problem to obtain the best-response value as:

$$\hat{\mathcal{P}}_{b_{k}}^{C}: \ \arg\max_{p_{k}^{C}} \tilde{r}_{k}^{C} - q_{k}^{*}(\frac{p_{k}^{C}}{\eta} + p_{cir}),$$
(21a)

$$s.t.(8),(17c).$$
 (21b)

In terms of the above concave maximization problem, it can be quickly solved by using convex optimization tools, like CVX [33]. To obtain the optimal q_k^* , an iterative method is derived in **Algorithm 1**.

Similarly, for D2D group i, we can also derive a concave maximization problem to obtain its best-response value as:

$$\hat{\mathcal{P}}_{b_{\mathcal{G}_i}}^D: \ \underset{p_{\mathcal{G}_i}}{\operatorname{arg\,max}} \xi(p_{\mathcal{G}_i}^D) - q_{\mathcal{G}_i}^*(\frac{p_{\mathcal{G}_i}^D}{\eta} + |\mathcal{G}_i|p_{cir}), \ (22a)$$

s.t.
$$p_{\mathcal{G}_i}^D \in D(\mathcal{G}_i),$$
 (22b)

where $D(\mathcal{G}_i)$ is a convex set determined according to **Remark 3**. The solving procedures are elaborated in **Algorithm 2**.

With the results returned from best-response functions, which are calculated in each player, we derive an iterative algorithm to find the Nash equilibrium. The procedures is described as: (a)

- In the beginning, initialize every player's strategy with the optimal transmit power under no interference, expressed as s_{Gi} = p^D_{Gi,0} and s_k = p^C_{k,0};
 Calculate each player's new strategy s^{*}_τ = b_τ(s_{-τ})(τ =
- 2) Calculate each player's new strategy $s_{\tau}^* = b_{\tau}(s_{-\tau})(\tau = \mathcal{G}_i, k);$
- 3) Judge if equilibrium is achieved by a differential result, define ϵ as the differential threshold with small value.(c1)
 - a) If $|s_{\tau}^* s_{\tau}| \le \epsilon(\tau = \mathcal{G}_i, k)$, return value s_{τ}^* as the equilibrium result.
 - b) Else, let $s_{\tau} = s_{\tau}^*(\tau = \mathcal{G}_i, k)$, return to step (2);

With the above iterative algorithm, the players exchange the information of their power in the game process, then the Nash equilibrium of the noncooperative game can be found.

Theorem 4: The above noncooperative game has a Nash equilibrium.

Proof: According to Kakutani's fixed point theorem [28], we can conclude that the above game has the following two properties:

- for all players, the set of strategies is nonempty and convex;
- 2) their utility functions are continuous and quasi-concave.

Obviously, the transmit power of CU that satisfies $p_k^C \in [0, \min\{\eta \phi_k^C, p_{k,\max}^C\}]$ is obviously nonempty and convex. For $p_{G_i}^D$, by referring to **Remark 4**, it is also nonempty and convex. Besides, the objective function in (17a) is in fractional form, with a concave nominator and a linear denominator, thus it is quasi-concave [29]. In terms of the utility function of D2D groups, its continuous and quasi-concave characteristics are proved in the proof of Lemma 1.

Note that, as for an obtained total transmit power of the D2D group, we first make sure that the QoS constraints are satisfied, i.e., each DT is allocated with an amount of power that the rate threshold is reached. Then, the rule in Theorem 1 is applied to allocate the remainder of the total power. That is, we first allocate power to the DT that has the highest D2D channel gain within its feasible domain of power, then to the next DT that has the highest channel gain among the rest of DTs until the available quota is exhausted.

Based on the above procedures, if we traverse each potential pair of D2D group and CU in spectrum reusing through noncooperative games, each player can know about the energy efficiency of each spectrum reusing choice. However, how to make the choices in players to finish the spectrum and power allocation remains to be solved, thus a matching game is put forward to accomplish this issue based on the energy efficiency result of the noncooperative games in the following.

C. The Second Stage: A Matching Game Approach

In this section, we propose to use a matching approach to accomplish the spectrum and power allocation issue. Recall the system model in Section III-A, a D2D group can only reuse the uplink spectrum of one CU, and a CU's uplink spectrum can only be reused by one D2D group. Thus, we formulate a one-to-one matching game to obtain the spectrum matching result.

For this matching game, we can defined it as a bidirectional mapping μ between the set of D2D groups $\mathbf{G} \triangleq \{\mathcal{G}_i | \forall i \in \mathcal{R}\}$ and CUs \mathcal{C} :1)

1)
$$\forall i \in \mathcal{R}, \ \mu(\mathcal{G}_i) \in \mathcal{C}, \ \text{and} \ |\mu(\mathcal{G}_i)| = 1;$$

2) $\forall k \in \mathcal{C}, \ \mu(k) \in \mathbf{G} \bigcup \{\emptyset\}, \text{ and } |\mu(k)| = 1.$

Herein, $\mu(k) = \emptyset$ means CU k's uplink spectrum is not reused by any D2D group, and we recognize this as the least preferred choice of CU k. Because exclusively occupying the spectrum goes against the purpose of making full use of the limited frequency resource in the cellular area.

In this matching game, players may reveal different level of preference in matching with the opposite players. If D2D group *i* prefers CU *k* to CU *k'*, the preference relation of G_i over $k, k' \in C(k \neq k')$ can be defined as

$$k \succ_{\mathcal{G}_i} k' \Leftrightarrow q_{\mathcal{G}_i}^* \big|_{\mu(\mathcal{G}_i)=k} > q_{\mathcal{G}_i}^* \big|_{\mu(\mathcal{G}_i)=k'}, \qquad (23)$$

where $\succ_{\mathcal{G}_i}$ is called a strict preference, and this preference relation is a complete, reflective and transitive binary relation between the players. Aside from this, if \mathcal{G}_i likes k not less than k', we have $k \succeq_{\mathcal{G}_i} k' \Leftrightarrow q^*_{\mathcal{G}_i} |_{\mu(\mathcal{G}_i)=k} \ge q^*_{\mathcal{G}_i} |_{\mu(\mathcal{G}_i)=k'}$. Similarly, we can also define CU k's preference over D2D groups $\mathcal{G}_i, \mathcal{G}_{i'} \in \mathcal{R}(\mathcal{G}_i \neq \mathcal{G}_{i'})$ as

$$\mathcal{G}_i \succ_k \mathcal{G}_{i'} \Leftrightarrow q_k^* \big|_{\mu(k) = \mathcal{G}_i} > q_k \big|_{\mu(k) = \mathcal{G}_{i'}}.$$
 (24)

Moreover, $\mathcal{G}_i \succeq_k \mathcal{G}_{i'} \Leftrightarrow q_k^*|_{\mu(k) = \mathcal{G}_i} \ge q_k|_{\mu(k) = \mathcal{G}_{i'}}$ means k likes \mathcal{G}_i not less than $\mathcal{G}_{i'}$. Therefore, based on the energy efficiency result obtained in the above subsection, the preference list of D2D group i and CU k can be established and represented as $PL_{\mathcal{G}_i} = \{q_{\mathcal{G}_i}^*|_{\forall \mu(\mathcal{G}_i) \in \mathcal{C}}\}$ and $PL_k = \{q_k^*|_{\forall \mu(k) \in \mathbf{G}}\}$.

After obtaining the overall preference profile $PL = \{PL_k, PL_{\mathcal{G}_i} | k \in C, i \in \mathcal{R}\}$, the matching game between D2D groups and CUs can proceed. To judge if a matching μ is stable, we introduce the definition of swap-blocking pair. An unmatched pair (k, \mathcal{G}_i) is called a blocking pair if the following conditions are both satisfied:(a)

k ∈ C and G_i ∈ G. In matching μ, there exists μ(G_i) ≠ k and μ(k) ≠ G_i;

2)
$$k \succ_{\mathcal{G}_i} \mu(\mathcal{G}_i)$$
 and $\mathcal{G}_i \succ_k \mu(k)$.

The matching μ is said to be stable when there is no blocking pair, i.e., the matching result cannot be improved by any individual agent (i.e., D2D groups and CUs) [34]. Therefore, by using the deferred acceptance (DA) algorithm, a stable matching can be achieved [35].

In our method, D2D groups firstly send requests to CUs, then adjustments are made according to their preference lists. Therein, the DR in a D2D group sends requests to CUs for reusing the spectrum. A corresponding CU that has received requests will decide to accept a certain request and deny the others, where the feedback information can be broadcasted to D2D groups by the CU. The detailed process is described in **Algorithm 3**.

Theorem 5: Algorithm 3 outputs a stable matching result.

Algorithm	3	An	Algorithm	to	Obtain	Stable	Matching
Result							

1 Initialize: each D2D group $i \in \mathcal{R}$ and CU $k \in \mathcal{C}$ separately maintain their preference list $PL_{\mathcal{G}_i}$ and PL_k , where the preference values are sorted in descending order; define ξ as the set of unmatched D2D groups, set its initial value as $\xi = \mathbf{G}$;

2 while $\xi \neq \emptyset$ do

- for each $\mathcal{G}_i \in \xi$ do 3
- D2D group i sends a request to the CU that ranks 4 highest in PL_{G_i} ; ıd

for each $k \in C$ do 6

CU k compares the new requests with its current 7 matching result, then it chooses the one that ranks highest in its preference list PL_k , add the matching result to μ ; if CU k's choice is unchanged then 8 Reject all the new requests and add those D2D groups into ξ . For each \mathcal{G}_i in these rejected requests, delete CU k from their preference list $PL_{\mathcal{G}_i};$ else 10 Reject the previously matched D2D group and 11 the new requests except the chosen one. For each rejected \mathcal{G}_i , delete CU k from their

preference list $PL_{\mathcal{G}_i}$;

```
12
        end
```

13 end

```
14 end
15 Return matching result \mu^*.
```

Proof: Assume that there exists a blocking pair in matching result μ^* . Based on Algorithm 3, each player chooses a matching result which it likes most. For each CU, it has rejected all D2D groups that are not matched with itself. Similarly, for each D2D group, it must have been rejected by all the CUs who received its requests. If there exists a blocking pair (k, \mathcal{G}_i) in μ^* , the D2D group \mathcal{G}_i must have issued a request to CU k before requesting $\mu(\mathcal{G}_i)$, since $k \succ_{\mathcal{G}_i} \mu(\mathcal{G}_i)$. And we also know that CU k will not reject \mathcal{G}_i to choose $\mu(k)$ because $\mathcal{G}_i \succ_k \mu(k)$. Therefore, the assumption contradicts the procedures in Algorithm 3, the proof is complete.

Theorem 6: The matching result is weak Pareto-optimal.

Proof: In Algorithm 3, each D2D group and CU choose a matching result that ranks highest in its preference list. Thus, for any D2D group or CU, they cannot unilaterally obtain higher utility without making others' utility worse off, i.e., the others may have to change their choices from the original ones with higher preference. Therefore, the matching result is weak Pareto-optimal.

V. IMPROVEMENTS OF THE PROPOSED METHOD AND ANALYSIS

In this section, we first investigate into the influence of EH on the resource allocation, then we further propose an



Fig. 2. An example of the influence of EH on the resource allocation.

energy-aware screening method in the BS before applying the proposed two-stage game method to reduce computational complexity. Finally, the signaling and computational overhead of our proposed methods are analyzed.

A. The Influence of EH on Resource Allocation

Firstly, recall the transmit power constraint (10), we can find that the transmit power of every DT depends not only on its maximum transmit power p_i^{\max} , but also on its energy level. Thus, by incorporating Remark 3 into consideration, we can find that the transmit power of a D2D group $p_{G_i}^D$ also depends on the energy level of its DTs. Similarly, (2) also shows that a CU's transmit power is also related to their energy level. Apart from the transmit power, the spectrum allocation result can be affected under different energy levels. Since the matching game is a two-sided matching, we present the following analysis from the aspect of a D2D group. To depict the mechanism, we give the following example for detailed explanation.

As shown in Fig. 2, assume the cellular channel gain of CU 1 is similar or slightly higher than CU 2. Due to the energy level constraint, CU 2 cannot transmit with the optimal power, which is relatively high, to achieve its maximum energy efficiency, thus it only transmits with a relatively low power to improve its energy efficiency as much as it can. However, CU 1 has enough power for transmission, thus it can transmit within full power to maximize its energy efficiency. Therefore, with CU 1 transmitting with a much higher transmit power than CU 2, if the D2D group reuses the uplink spectrum of CU 1, it will suffer from more serious interference under the condition of the negligible difference of the interference channel gain from two CUs. In this scenario, the D2D group prefers to reuse the spectrum of CU 2 rather than CU 1. In contrast, when there exists no energy level limit introduced by EH, CU 2 will be able to transmit within full power, thus it will transmit with higher power than CU 1 to maximize its energy efficiency [36]. Under this condition, the D2D group will prefer CU 1 for less interference.

B. An Energy-Aware Screening Method for Reducing Complexity

In Section IV, we proposed a two-stage game approach to tackle the unilateral energy efficiency maximization problem of each player. However, we notice that each D2D group must compete with every CU to establish their preference lists, but it will only reuse one CU's uplink spectrum in the second stage. It is noteworthy that the number of CUs is generally much more than D2D groups in normal circumstances [16]. Thus, there exists substantial room for further reducing the complexity of the solving method. In the following, we first have a theorem on the resource allocation of a D2D group.

Theorem 7: For each D2D group, M available items in its preference list is adequate for its resource allocation.

Proof: In the worst case, a D2D group can be rejected by at most (M-1) CUs during the matching game, thus M available items in its preference list will be adequate for its resource allocation.

Based on Theorem 7 and the analysis in Section V-A, we come up with an energy-aware screening method for preliminarily narrowing down the scale of CUs that a D2D group competes with to establish preference before the first stage of our method.

As for a D2D group, it is apparent that preliminarily choosing M most potentially preferred CUs rather than all CUs is much more cost-efficient while establishing preference lists. Therefore, how to effectively implement the screening of CUs becomes a crucial issue. By observing the objective function of $\hat{\mathcal{P}}_k^C$ in (17a), it is easy to raise the following corollary.

Theorem 8: For a D2D group, it will prefer the CU who has higher cellular channel gain under the same transmit power limits and interference channel gain of CUs.

Proof: For a CU, it will transmit with lower power if we raise its channel gain to the BS while maximizing its energy efficiency [37]. Thus, less interference will be received in the D2D group. Therefore, the D2D group will prefer the CU with higher cellular channel gain.

Furthermore, we also find that the interference item I_i^C in $\hat{\mathcal{P}}_i^D$ is also a fatal factor in constructing a D2D group's preference over CUs.

Corollary 1: Similarly, a D2D group will prefer the CU who has lower interference channel gain under the same transmit power limits and cellular channel gain of CUs.

Therefore, based the above analysis and the consideration of energy level in the above subsection, we come up with an energy-aware screening parameter π_i^k to select *M* CUs for constructing preference list, defined as:

$$\pi_i^k \stackrel{\Delta}{=} \frac{g_k}{g_{k,i} \cdot \min\{\eta \phi_k^C, p_{k,\max}^C\}}$$
(25)

By calculating $\{\pi_i^k\}(\forall k \in C)$ for each D2D group *i* and selecting *M* CUs who have the highest value, the scale of CUs that a D2D group competes with can be greatly reduced. Note that the energy-aware screening parameter is computed by the BS in a centralized manner before applying the proposed twostage game approach in a distributed algorithm to allocate resources. Therein, the screening method only requires one transmission from each player to the BS, and the computational cost is only $O(K \cdot M)$.

C. Analysis of Signaling and Computational Overhead

With the proposed two-stage game approach, we can obtain an energy-efficient resource allocation. As for the channel gains, the channel gain from the D2D group to the BS can be obtained in the BS by estimating the received power from the DTs in a D2D group following (16), and CUs are informed about the interference channel gain. By broadcasting pilot signals from CUs, the uplink channel gain to the BS and interference channel gain from CUs to a D2D group can be both estimated, where a CU only needs to broadcast once. Assume that each noncooperative game requires O(S)times of signaling. Thus, establishing a preference list without screening CUs requires $O(K \cdot M \cdot S)$ times of signaling, where the signaling overhead of each D2D group and CU is $O(K \cdot S)$, $O(M \cdot S)$ respectively. While after using the screening method in Section V-B, the signaling overhead becomes $O(M^2 \cdot S)$, and the signaling overhead of each D2D group and CU is $O(M \cdot S), O(M \cdot S)$ respectively. Moreover, in the matching game, it costs $O(\beta M)$ times of signaling, where β is generally a small value bigger than 1. Thus, the total signaling overhead with and without the screening method in Section V-B is separately $O(M(\beta + M \cdot S)), O(M(\beta + K \cdot S)).$

In terms of the computational cost, first of all, solving each fractional programming in the best-response functions can be very fast, herein we denote its computational cost as O(C). In each D2D group and CU, their computational complexity with and without the screening method in Section V-B are separately $O(M \cdot S \cdot C)$, $O(M^2/K \cdot S \cdot C)$ and $O(K \cdot S \cdot C)$, $O(M \cdot S \cdot C)$. Therefore, from the systematic point of view, the total computational complexity with and without the screening method in Section V-B is $O(M^2 \cdot S \cdot C)$ and $O(K \cdot M \cdot S \cdot C)$.

Compared with the centralized algorithms with a signaling overhead of O(M+K), the signaling overhead in our method is higher but still within a relatively low level. In addition, the proposed distributed algorithm is much lower compared with other centralized algorithms (not including implementing the proposed method in a centralized way) with a computational overhead of $O(K^M \cdot C)$, where the combinatorial problem in spectrum allocation leads to the exponential computational cost, while the matching game in our method requires much less computations.

VI. NUMERICAL SIMULATIONS

In this section, numerical simulations are presented to demonstrate the performance of our proposed method in various aspects. In the following, we first give an introduction to the simulation setup, including parameters setting and comparison setup. Then, numerical and analytical results are presented.

To provide access to the DR for a number of DTs, we introduce multiple DT clusters in a D2D group [22], where each DT cluster in a D2D group reuses the uplink spectrum of a CU. Thus, a D2D group that consists of multiple DT clusters can be mathematically regarded as multiple D2D groups that have one DT cluster in each, but with the overlapped DR.

In the simulations, we measure the performance of the proposed method, i.e., the two-stage game approach with the screening method in Section V-B. For comparison, we introduce an optimal-power and no-screening (OPNS) method, where the optimal power allocation is obtained through centralized multi-objective programming without using the approximation method in Section V-B [23], then a matching game without the screening method in Section V-B is adopted for spectrum allocation. Besides, we also compare our proposed method with the results obtained by using OMA protocol, like time division multiple access (TDMA), where the power of OMA users are uniformly allocated to users [19], [26], and the resource allocation process is the same as the proposed method. Moreover, another channel gain-based spectrum allocation method (GM) is also adopted for comparison [10], where the power allocation part is substituted with our proposed method due to the difference in objective.

A. Simulation Setup

In the simulation settings, the BS locates in the center of the round cellular area with a radius of 500 meters, where CUs and D2D groups are randomly deployed in the cellular area [23]. The D2D distance from DTs in a DT cluster to the DR is discriminatively set to ensure the channel gain difference between NOMA users because the distance between users determines the large-scale path loss in channel gain [26]. For example, with a maximum D2D distance of 50m, the distance between 2 DTs to the DR is randomly set within the two evenly partitioned ranges within 50m, i.e., 0 - 25m and 25 - 50m. The number of D2D groups varies from 2 to 10, and the ratio of CU and D2D group numbers ranges from 5 to 10. Moreover, the EH parameter is set from 21 to 28 to investigate its influence on system performance. More practically, we assume that there are 2 or 3 DTs in a NOMA-based D2D group for ensuring effective SIC [19]. Based on the above setup, the simulation results are obtained by averaging 1000 repeats. Detailed simulation parameters are summarized in Table I.

B. Numerical Results

To demonstrate the performance of the proposed method under different value of parameters, we first elaborate the average energy efficiency performance of D2D groups under different number of D2D groups, ratio of CU and D2D group numbers and EH arrival parameters of DTs and CUs, separately corresponding to M, K/M, λ_d , λ_c . In the simulation results, we pay most of our attention to the energy efficiency performance of D2D groups, since the NOMA-based D2D groups are our major concern. Besides, for completeness, the energy efficiency performance of CUs is also analyzed in the end of this section.

Fig. 3 evaluates the energy efficiency of DT clusters in D2D groups versus the number of DT clusters in a D2D group. With the increase of the number of clusters, the energy efficiency decreases with an accelerating decreasing rate. This is because, with the increase of the number of DT clusters in

TABLE I Simulation Parameters Setup

Description	Value		
Cell radius	500 m		
Uplink bandwidth of each CU	180 KHz		
Path-loss exponent	3		
Noise density	-174 dBm/Hz		
Number of DTs in a D2D group	2, 3		
Number of D2D groups	[2:2:10]		
Number of DT clusters in a D2D group	[1:1:8]		
Ratio of CU and D2D group numbers	[5:1:10]		
Energy packet size	$U_j^n = 0.5 \text{ mW}$		
Power amplifier efficiency	$\eta = 40\%$		
Energy arrival parameter of DTs	$\lambda_d = [21:1:28]$		
Energy arrival parameter of CUs	$\lambda_c = [21:1:28]$		
Cellular rate threshold	1 bit/s/Hz		
D2D communication rate threshold	1 bit/s/Hz		
Maximum cellular transmit power	23 dBm		
Maximum D2D transmit power	20 dBm		
Static circuit power	10 mW		



Fig. 3. Avg. energy efficiency of DT clusters in D2D groups vs. the number of DT clusters in a D2D group under M = 4, K = 40, $\lambda_d = 24$, $\lambda_c = 26$.

a D2D group, those DT clusters will compete more intensely with each other for matching with their preferable CUs, where the competition will lead to some matchings that bring about serious interference between CUs and DT clusters, this situation happens especially when the number of DT clusters is large. Besides, it can be observed that 3 DTs in a DT cluster can lead to inferior performance compared with 2 DTs in a DT cluster, this is due to the more severe interference between DTs.

Fig. 4 plots the average energy efficiency of DT clusters in D2D groups under different EH arrival parameter λ_d of DTs with 2 and 3 DTs in each DT cluster in a D2D group. With an increasing λ_d , the average energy efficiency of D2D groups grows with a decreasing growth rate. Compared with the GM and the OMA scheme, the proposed method can both obtain significant superiority over them. We also find that the performance gap between GM algorithm and the proposed method is larger with a relatively small λ_d . This is because, the proposed method takes the energy level of players into consideration during the spectrum allocation game to make better spectrum



Fig. 4. Avg. energy efficiency of DT clusters in D2D groups vs. different energy arrival parameter λ_d under M = 4, K = 40 and $\lambda_c = 26$.

reusing choices instead of only during the power allocation. Besides, we also notice that the performance gap between the proposed method and the OPNS algorithm enlarges with the increase of λ_d . Because in the proposed method, we adopt an approximation-based approach to formulate a noncooperative game, the deviation from optimal result may enlarge under a larger transmit power. In addition, the screening of CUs may also lead to small performance loss compared with the OPNS algorithm.

Fig. 5 depicts the average energy efficiency of DT clusters in D2D groups under different number of D2D groups. It can be seen that, the average energy efficiency increases with the number of D2D groups. That is because, under the same ratio of CU and D2D group numbers, with the increase of D2D groups, they will have a larger range of spectrum reusing choices. Thus, this leads to the increase of the average energy efficiency of D2D groups. However, we notice that the growth rate distinctly slows down by observing the curve with 3 DTs in a DT cluster, this is because more DT clusters will compete with each other for their preferable CUs, and those D2D groups that are rejected in the matching game only have inferior choices in the matching game. In this circumstance, a NOMA-based DT cluster with 3 DTs is more sensitive to cellular interference, thus the growth rate slows down. Besides, we also notice that the proposed method is very close to the OPNS algorithm, and also superior to the other methods.

As shown in Fig. 6, the curves show an overview of the average energy efficiency of DT clusters in D2D groups under different ratios of CU and D2D group numbers. With the increase of K/M, the average energy efficiency of D2D clusters in D2D groups reveals an increasing tendency, but with a decreasing growth rate. It is obvious that the proposed method is pretty close to the OPNS algorithm, however, it is superior to the other two methods, including the OMA scheme. Besides, we also notice that with the increase of K/M, the performance gap between the clusters with 3 DTs and with 2 DTs narrows down, which can be seen by $\Delta h_1 > \Delta h_2$. This is because when K/M is small, there may exist more serious perceived cellular interference in D2D groups, and the NOMA



Fig. 5. Avg. energy efficiency of DT clusters in D2D groups vs. number of D2D groups with 4 DT clusters in each D2D group, K/M = 10, $\lambda_d = 24$, $\lambda_c = 26$.



Fig. 6. Avg. energy efficiency of DT clusters in D2D groups vs. the ratio of CU and D2D group numbers with 4 DT clusters in each D2D group under M = 4, $\lambda_d = 24$, $\lambda_c = 26$.

cluster with 3 DTs is more sensitive to the interference. When K/M increases, the DT clusters can make better choices in the matching game, thus the NOMA cluster with 3 DTs increases with a higher growth rate compared with the cluster with 2 DTs.

Fig. 7 provides an overview of the average energy efficiency of DT clusters in D2D groups under different number of D2D groups and different ratios of CU and D2D group numbers. Under a given number of D2D groups, with the increase of the ratio of CU and D2D group numbers, the average energy efficiency of D2D groups increases, since the D2D groups can choose more suitable CUs within a larger range to achieve higher energy efficiency. However, it is obvious that the growth rate slows down with the increase of K/M. This is because when the number of CUs reaches a relatively high level, D2D groups can already make choices within a fairly sufficient range of CU's spectrum without the increased number of CUs, thus increasing the number of CUs can only bring a marginal



Fig. 7. Avg. energy efficiency of DT clusters in D2D groups vs. different number of D2D groups and different ratios of CU and D2D group numbers under $\lambda_d = 24$, $\lambda_c = 26$.



Fig. 8. Avg. energy efficiency of DT clusters in D2D groups vs. different energy arrival parameter λ_d and different ratios of CU and D2D group numbers under M = 4, $\lambda_c = 26$.

increase in energy efficiency. Besides, with the increase of the number of D2D groups, the average energy efficiency is also increasing at a relatively slow speed. Because under the same K/M, the D2D groups have a wider choice range of CUs' spectrum when M increases, thus the D2D groups may make better choices in spectrum reusing to achieve higher energy efficiency.

In Fig. 8, the curves show an overview of the average energy efficiency performance of D2D groups under different energy arrival intensities of DTs and ratios of CU and D2D group numbers. With the increase of K/M, the growth of the average energy efficiency of D2D groups slows down. The reason is the same as that in the descriptions of Fig. 6. Besides, with the increase of the EH arrival parameter λ_d of D2D groups, the average energy efficiency of D2D groups also increases, but with a relatively high growth rate around the beginning and a low growth rate after that. That is because a relatively small EH parameter will become an obstacle when D2D groups increase their transmit power to obtain higher energy efficiency. Thus, when λ_d increases, the D2D groups will possibly have more available transmit power for obtaining



Fig. 9. Avg. energy efficiency of DT clusters in D2D groups vs. different D2D communication range with 4 DT clusters in each D2D group under $M = 4, K = 40, \lambda_d = 24, \lambda_c = 26.$



Fig. 10. Avg. energy efficiency of players vs. different energy arrival parameter λ_c of CUs under M = 4, K = 40, $\lambda_d = 24$.

higher energy efficiency. But when the EH capacity is enough for transmitting, the average energy efficiency only slightly increases.

Moreover, we also investigate the average energy efficiency of DT clusters in D2D groups under different D2D communication ranges in Fig. 9. With a larger D2D communication range, the average energy efficiency of DT clusters decreases. We also find that the gap between the proposed method and the gain-based method enlarges with the D2D communication range, which can be seen from D_1 and D_2 . Because D2D communication within a smaller range will require less transmit power for achieving optimal energy efficiency compared with that within a larger D2D communication range. Under this circumstance, the interference between D2D groups and CUs becomes less severe, thus the performance gap between the gain-based and the proposed method is smaller.

In Fig. 10, we investigate the energy efficiency of DT clusters in D2D groups and matched CUs under different energy arrival parameters of CUs λ_c . With the growth of λ_c , the average energy efficiency of CUs increases, meanwhile,



Fig. 11. Energy efficiency of players vs. the number of iterations in a noncooperative game process under M = 4, K = 40, $\lambda_d = 24$, $\lambda_c = 26$.

the average energy efficiency of D2D groups decreases. Because with a larger λ_c , the CUs may have more power for transmitting to obtain higher energy efficiency. However, this will introduce more interference to D2D groups, thus the average energy efficiency decreases.

Lastly, we also demonstrate the game process of a pair of DT cluster in a D2D group and CU during the first-stage noncooperative game. In Fig. 11, the energy efficiency of players both goes through a significant decrease in the first two iterations. And after that, their energy efficiency changes marginally. This reflects the quick convergence rate to achieve Nash equilibrium during the game process. Besides, we also notice that the energy efficiency of the players is pretty high in the first iteration. This is because we ignore the interference between them in the first iteration for initialization. However, after the first iteration, the interference between the DT cluster and the CU becomes a nonnegligible factor, leading to the sharp decrease of the CU's energy efficiency as well as the DT cluster's.

VII. CONCLUSION

In this paper, we studied the energy-efficient resource allocation issue of uplink NOMA-based D2D communication in EH scenario, where the available energy of UEs is considered. We introduced an approximation method to formulate the intractable power allocation problem between the pair of a D2D group and a CU as a noncooperative game. Then a two-stage game approach was proposed to solve the unilateral energy efficiency maximization problem of each D2D group and CU in a distributed manner. Besides, the influence of EH was also discussed in the subsequent part, and an energyaware screening method is proposed to reduce the computational complexity. Finally, the simulation results showed the effectiveness of the proposed method.

REFERENCES

 A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

- [2] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communication in LTE-advanced networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1923–1940, 4th Quart., 2015.
- [3] Z. Zhou, M. Peng, and Z. Zhao, "Joint data-energy beamforming and traffic offloading in cloud radio access networks with energy harvestingaided D2D communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8094–8107, Dec. 2018.
- [4] Y. Luo, P. Hong, R. Su, and K. Xue, "Resource allocation for energy harvesting-powered D2D communication underlaying cellular networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10486–10498, Nov. 2017.
- [5] A. H. Sakr and E. Hossain, "Cognitive and energy harvesting-based D2D communication in cellular networks: Stochastic geometry modeling and analysis," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1867–1880, May 2015.
- [6] Y. He, X. Cheng, W. Peng, and G. L. Stuber, "A survey of energy harvesting communications: Models and offline optimal policies," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 79–85, Jun. 2015.
- [7] F. Jameel, Z. Hamid, F. Jabeen, S. Zeadally, and M. A. Javed, "A survey of device-to-device communications: Research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2133–2168, 3rd Quart., 2018.
- [8] R. Li, P. Hong, K. Xue, M. Zhang, and T. Yang, "Energy-efficient resource allocation for high-rate underlay D2D communications with statistical CSI: A one-to-many strategy," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4006–4018, Apr. 2020.
- [9] Z. Kuang, G. Liu, G. Li, and X. Deng, "Energy efficient resource allocation algorithm in energy harvesting-based D2D heterogeneous networks," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 557–567, Feb. 2019.
- [10] U. Saleem, S. Jangsher, H. K. Qureshi, and S. A. Hassan, "Joint subcarrier and power allocation in the energy-harvesting-aided D2D communication," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2608–2617, Jun. 2018.
- [11] J. Huang, C.-C. Xing, and M. Guizani, "Power allocation for D2D communications with SWIPT," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2308–2320, Apr. 2020.
- [12] C. Networks, R. Li, P. Hong, D. Li, and J. Pei, "Energy harvesting-based D2D relaying achieving energy cooperation underlaying," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7.
- [13] A. Y. Awan, M. Ali, M. Naeem, F. Qamar, and M. N. Sial, "Joint network admission control, mode assignment, and power allocation in energy harvesting aided D2D communication," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1914–1923, Mar. 2020.
- [14] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.
- [15] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, and M. Elkashlan, "Joint subchannel and power allocation for NOMA enhanced D2D communications," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 5081–5094, Nov. 2017.
- [16] I. Budhiraja, S. Tyagi, S. Tanwar, N. Kumar, and J. J. P. C. Rodrigues, "DIYA: Tactile internet driven delay assessment NOMA-based scheme for D2D communication," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6354–6366, Dec. 2019.
- [17] Y. Dai, M. Sheng, J. Liu, N. Cheng, X. Shen, and Q. Yang, "Joint mode selection and resource allocation for D2D-enabled NOMA cellular networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6721–6733, Jul. 2019.
- [18] L. Pei *et al.*, "Energy-efficient D2D communications underlaying NOMA-based networks with energy harvesting," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 914–917, May 2018.
- [19] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016.
- [20] M. M. Salim, D. Wang, H. A. El A. Elsayed, Y. Liu, and M. A. Elaziz, "Joint optimization of energy-harvesting-powered two-way relaying D2D communication for IoT: A rate–energy efficiency tradeoff," *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11735–11752, Dec. 2020.
- [21] M. W. Baidas, M. S. Bahbahani, E. Alsusa, K. A. Hamdi, and Z. Ding, "Joint D2D group association and channel assignment in uplink multicell NOMA networks: A matching-theoretic approach," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8771–8785, Dec. 2019.
- [22] B. Liu, C. Liu, and M. Peng, "Resource allocation for energy-efficient MEC in NOMA-enabled massive IoT networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1015–1027, Apr. 2021.

- [23] Y. Hao, Q. Ni, H. Li, and S. Hou, "Robust multi-objective optimization for EE-SE tradeoff in D2D communications underlaying heterogeneous networks," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4936–4949, Oct. 2018.
- [24] P. Mitran, "On optimal online policies in energy harvesting systems for compound Poisson energy arrivals," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 960–964.
- [25] M. Ehrgott, Multicriteria Optimization. Berlin, Germany: Springer, 2005.
- [26] M. Zeng, A. Yadav, O. Dobre, and H. V. Poor, "Energy-efficient joint user-RB association and power allocation for uplink hybrid NOMA-OMA," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5119–5131, Feb. 2019.
- [27] C. Pan et al., "Totally distributed energy-efficient transmission in MIMO interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6325–6338, Nov. 2015.
- [28] M. J. Osborne and A. Rubinstein, A Course Game Theory. Cambridge, MA, USA: MIT Press, 1994.
- [29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [30] C. Yang, J. Li, P. Semasinghe, E. Hossain, S. M. Perlaza, and Z. Han, "Distributed interference and energy-aware power control for ultra-dense D2D networks: A mean field game," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1205–1217, Feb. 2017.
- [31] H. Zhu, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, Game Theory in Wireless and Communication Networks: Theory, Models, and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [32] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
- [33] M. Grant and S. Boyd. (Mar. 2014). CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1. [Online]. Available: http://cvxr.com/cvx
- [34] A. E. Roth and M. Sotomayor, *Handbook of Game Theory With Economic Applications*, vol. 1. Amsterdam, The Netherlands: Elsevier, 1992.
- [35] A. E. Roth, "Deferred acceptance algorithms: History, theory, practice, and open questions," *Int. J. Game Theory*, vol. 36, pp. 537–569, Mar. 2008.
- [36] J. Xu and R. Zhang, "Throughput optimal policies for energy harvesting wireless transmitters with non-ideal circuit power," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 322–332, Feb. 2014.
- [37] A. Khalili, M. R. Mili, M. Rasti, S. Parsaeefard, and D. W. K. Ng, "Antenna selection strategy for energy efficiency maximization in uplink OFDMA networks: A multi-objective approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 595–609, Jan. 2020.



Peilin Hong received the B.S. and M.S. degrees from the Department of Electronic Engineering and Information Science (EEIS), University of Science and Technology of China (USTC), in 1983 and 1986. She is currently a Professor and an Advisor for Ph.D. candidates with the Department of EEIS, USTC. She has published two books and over 150 academic papers in several journals and conference proceedings. Her research interests include next-generation Internet, policy control, IP QoS, and information security.



Kaiping Xue (Senior Member, IEEE) received the bachelor's degree from the Department of Information Security, University of Science and Technology of China (USTC) in 2003, and the Ph.D. degree from the Department of Electronic Engineering and Information Science (EEIS), USTC, in 2007. From May 2012 to May 2013, he was a Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, University of Florida. He is currently a Professor with the School of Cyber Security, USTC. His research interests include next-

generation internet architecture design, transmission optimization, and network security. He is an fellow of IET. He serves on the Editorial Board of several journals, including the IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING (TDSC), the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (TWC), and the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT (TNSM). He has also served as a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC) and a Lead Guest Editor for *IEEE Communications Magazine* and *IEEE Network*.



Ming Zhang (Student Member, IEEE) received the B.S. degree from the Hefei University of Technology (HFUT), Hefei, China, in 2017. He is currently pursuing the Ph.D. degree in communication and information systems with the Department of Electronic Engineering and Information Science, University of Science and Technology of China (USTC), Hefei. His current research interests include wireless video transmission and multiple access networks.



Runzhou Li (Student Member, IEEE) received the B.S. degree from the Department of Electronic Engineering and Information Science, University of Science and Technology of China (USTC), Hefei, China, in 2018, where he is currently pursuing the Ph.D. degree. His current research interests include resource management in D2D communications and multiple access networks.



Te Yang received the B.S. degree from the Dalian University of Technology (DUT), Dalian, China, in 2018. He is currently pursuing the M.S. degree in communication and information systems with the Department of Electronic Engineering and Information Science, University of Science and Technology of China (USTC), Hefei, China. His current research interest includes resource allocation in device-todevice communication networks.