

Joint User Pairing, Mode Selection, and Power Control for D2D-Capable Cellular Networks Enhanced by Non-Orthogonal Multiple Access

Daosen Zhai, *Member, IEEE*, Ruonan Zhang, *Member, IEEE*, Yutong Wang, Huakui Sun, Lin Cai, *Senior Member, IEEE*, and Zhiguo Ding, *Senior Member, IEEE*

Abstract—Non-orthogonal multiple access (NOMA) and device-to-device (D2D) are two promising technologies that have great potential in improving user connectivity. In this paper, we incorporate NOMA into the D2D-capable cellular networks and propose a new NOMA-aided D2D access scheme. In the proposed scheme, the D2D users can operate in four spectrum-sharing modes, which are the extension of the traditional underlay mode. To fully exploit the advantages of the NOMA-and-D2D integrated framework, we formulate a connectivity-maximization problem by jointly considering user pairing, mode selection, and power control under the constraints of the decoding thresholds of cellular users and D2D users. Based on the graph theory, we devise an efficient algorithm with polynomial complexity to solve the formulated problem optimally. We first analytically obtain the optimal transmission power and spectrum-sharing mode for every possible user pair through a graphical method. Based on the power control and mode selection policies, we transform the user pairing problem into a min-cost max-flow problem which can be tackled by the Ford-Fulkerson algorithm. Finally, simulation results indicate that the NOMA-aided D2D access scheme outperforms the traditional underlay mode, and the proposed algorithm yields a large performance gain in comparison with other schemes in terms of user connectivity and power consumption.

I. INTRODUCTION

As a spectrum-and-energy efficiency technology, device-to-device (D2D) has a wide range of applications in mobile communications, such as the vehicle-to-vehicle (V2V) commu-

This work was supported in part by China Postdoctoral Science Foundation under Grant BX20180262 and Grant 2018M641019, in part by the National Natural Science Foundation of China under Grant 61571370, Grant 61601365, and Grant 61801387, in part by the Fundamental Research Funds for the Central Universities under Grant 3102017OQD091, in part by the Industrial Innovation Chain Project of Shaanxi Province under Grant 2018ZDCXLYG03-04 and Grant 2019ZDLGY07-10, and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). The work of Z. Ding was supported by the UK EPSRC under grant number EP/P009719/2, NSFC under grant number 61728101, and H2020-MSCA-RISE-2015 under grant number 690750. (*Corresponding author: Ruonan Zhang.*)

Daosen Zhai, Ruonan Zhang, and Yutong Wang are with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, China (e-mail: zhaidaosen@nwpu.edu.cn; rzhang@nwpu.edu.cn; wangyutong@mail.nwpu.edu.cn).

Huakui Sun is with the School of Sergeancy, Weifang University of Science and Technology, Weifang, Shandong, 262700, China (email: sunhuakui@wfust.edu.cn).

Lin Cai is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, V8W 3P6 (e-mail: cai@ece.uvic.ca).

Zhiguo Ding is with the School of Electrical and Electronic Engineering, University of Manchester, M13 9PL, UK (email: zhiguo.ding@manchester.ac.uk).

nications, multi-user cooperative communications, and data-sharing and low-cost Internet of Things (IoT) [1]. Different from the traditional base-station-centered cellular networks, D2D is a more flexible point-to-point transmission paradigm, where the proximity devices can communicate directly with each other without going through the base station (BS) [2], [3]. The direct local transmission characteristic of D2D leads to the advantages of proximity gain and hop gain, and therefore the transmission power of the mobile devices can be greatly reduced. Furthermore, under a certain level of interference, the D2D devices can reuse the spectrum of the cellular networks. Thus, D2D also has the spatial multiplexing gain which enhances the spectrum efficiency. Owing to the various advantages, D2D has been proposed for the long term evolution (LTE) system in the third generation partnership project (3GPP) release-12 [4] and envisioned to be a key component of the fifth-generation (5G) mobile communication system [1], [5].

In the D2D-capable cellular networks, two spectrum-sharing modes have been proposed, namely the overlay mode and the underlay mode [6], [7]. In the overlay mode, the D2D users (DUEs) and the cellular users (CUEs) utilize orthogonal spectrum, such that there is no interference among the users. However, the number of users accommodated in the network is strictly limited by the number of orthogonal resources. Additionally, it is anticipated that the 5G wireless networks may connect 100 billion devices by 2030 [8]. Thus, the overlay mode is not suitable for the future wireless networks, especially for the IoT with massive connections. In the underlay mode, the CUEs can be regarded as the primary users in the cognitive networks, and the DUEs can be regarded as the secondary users. As such, the DUEs can reuse the spectrum of the CUEs under certain interference constraint. Compared with the overlay mode, the underlay mode can support more connections and hence receives significant research attention. Abundant researches have been conducted in the D2D underlying cellular networks (DUCN) [9]–[16]. In these networks, the cross-tier interference among DUEs and CUEs is the main source that limits the system performance [17], and how to coordinate the cross-tier interference is the key.

A. Motivations

The early study in the DUCN [9]–[16] mainly concentrated in the networks with orthogonal multiple access (OMA)

techniques, e.g., orthogonal frequency-division multiple access (OFDMA) [10], [13] and time-division multiple access (TDMA) [12]. In recent years, the non-orthogonal multiple access (NOMA) has attracted extensive attention and inspired thorough research in both academia [18], [19] and industry [20], [21]. Different from OMA, NOMA exploits the user diversities in the power domain and thereby is capable of supporting multiple users on the same spectrum. Due to this feature, NOMA targets the application scenarios with massive connections [19], [21], such as the IoT [22]–[25]. Specifically, the receiver in NOMA transmission can first detect the strong interference and then decode the desired signals by adopting the successive interference cancellation (SIC) technique. As such, NOMA can also be regarded as an efficient interference coordination technique. The potential of NOMA in user connectivity and interference coordination is much needed by the D2D-capable cellular networks.

Different from the OMA networks [9]–[16], the receivers in the NOMA networks can use the SIC technique to handle the strong cross-tier interference between DUEs and CUEs and thereby, more spectrum-sharing modes can be exploited. Thanks to the expanding spectrum-sharing modes, more D2D links can be activated simultaneously. Meanwhile, the resource management problem becomes more challenging. Specifically, when designing the power control policy, the decoding order of SIC should be taken into account according to different interference status. Furthermore, in order to maximize the user connectivity, the D2D links must be carefully selected to construct appropriate NOMA pairs in the admission control stage. Moreover, since more spectrum-sharing modes can be used by the DUEs and CUEs, the mode selection problem becomes more difficult to solve. Therefore, how to incorporate NOMA into the D2D-capable cellular networks and appropriately configure the resource among the DUEs and CUEs is an important and challenging problem.

In order to fully exploit the advantages of NOMA in D2D communications, some works related to the resource management and performance analysis have been conducted [26]–[35]. However, the existing works mainly concerned the system throughput or outage probability, none of which paid attention to the user connectivity. However, massive connectivity is a key performance indicator of the 5G wireless networks. Although there have been some works studying the connectivity-maximization problem in NOMA networks [23], [36], they are not suitable for the D2D-capable cellular networks. Furthermore, since both NOMA and D2D have potential in user connectivity, it is meaningful to incorporate these two technologies into the same network to improve the user connectivity. Moreover, in order to maximize the user connectivity in the NOMA-and-D2D integrated network, the communication resources should be appropriately allocated among the CUEs and DUEs.

B. Contributions

Motivated by the above reasons, we investigate the connectivity maximization problem for the NOMA enhanced D2D-capable cellular networks. Specifically, we propose a NOMA-aided D2D access scheme and optimize the resource allocation

for DUEs and CUEs by jointly considering user pairing, mode selection, and power control. The main contributions of this paper are summarized as follows.

- 1) We propose a new D2D access scheme by applying NOMA into the D2D-capable cellular networks. Different from the existing works, NOMA is utilized as the multiple access technique of DUEs into the cellular networks in our proposed scheme. Under this framework, the DUEs can operate in four spectrum-sharing modes, three of which are different from the traditional underlay and overlay modes. In particular, the additional spectrum-sharing modes can coordinate the cross-tier interference between the DUEs and CUEs and thereby improve the D2D connections.
- 2) We jointly optimize the user pairing, mode selection, and power control to maximize the accessed D2D links and meanwhile minimize the total power consumption under the constraints of the decoding thresholds of CUEs and DUEs. Specifically, this joint optimization problem is formulated as a mixed integer programming problem (MIP). To tackle the MIP, we first analytically obtain the optimal transmission power and spectrum-sharing mode for every possible DUE-CUE pair through a graphical method. Based on the power control and mode selection policies, we remodel the user pairing problem as a min-cost max-flow problem in graph theory and then solve it by adopting the Ford-Fulkerson algorithm. Our designed algorithm can solve the MIP optimally in complexity of $O(NM^2)$, where N and M are the numbers of CUEs and DUEs respectively.
- 3) We conduct extensive simulations to evaluate the performance of the proposed D2D access scheme and the resource allocation algorithm. Firstly, it shows that the NOMA-aided D2D access scheme can greatly improve the accessed D2D links with respect to the traditional underlay mode, which verifies the necessity of applying NOMA into the D2D-capable cellular networks. Furthermore, it demonstrates that the proposed resource allocation algorithm can achieve the optimal solution of the formulated problem. Moreover, comparing with the other schemes, our algorithm can yield a significant performance gain in terms of user connectivity and power consumption.

C. Organization

The rest of the paper is organized as follows. In section II, we overview the related works on the NOMA-and-D2D integrated networks. The considered network model and the proposed NOMA-aided spectrum-sharing modes are introduced in Section III. In Section IV, we present the problem formulation. Section V provides a detailed illustration for the proposed user pairing, mode selection, and power control algorithm. Simulation results are given in Section VI. Finally, we conclude our paper in Section VII.

II. RELATED WORKS

In recent years, there have been some works conducted in the NOMA enhanced D2D-capable cellular networks [26]–

[35]. Specifically, the authors in [26] studied the integration of a D2D connection to a downlink two-user NOMA system. The analysis shows that the cellular users achieve a higher sum-rate and individual rates in a NOMA-D2D system than in an OMA-D2D system. In a more general D2D underlying NOMA based cellular network, [27] jointly optimized the power control and channel assignment to maximize the sum-rate of D2D pairs under the constraint of the minimum rate requirements of CUEs. In the same scenario with [27], the authors in [28] and [29] investigated the power allocation and user clustering problem, where [28] was concerned about the resource allocation for NOMA users, and [29] focused on the joint optimization of NOMA-based CUEs and traditional DUEs. In [30], the outage probability and average link throughput were analyzed for NOMA based two-tier-cellular networks underlying inband D2D communications. As a common feature of [26]–[30], NOMA is utilized as a new multiple access technique of cellular networks. Unlike [26]–[30], the concept of D2D group was introduced in [31], where the D2D transmitter can send the superimposed messages to multiple receivers by applying the NOMA transmission protocol. In [31], a joint subchannel and power allocation scheme based on the matching theory was proposed for the NOMA enhanced D2D communications.

Except for the independent application of NOMA and D2D, these two technologies can also be incorporated, which leads to a new communication protocol, namely the cooperative NOMA. In this protocol, the NOMA-strong user can help the NOMA-weak user via D2D communications to reduce the outage probability of cell-edge users or enhance the system throughput. In [32], the authors proposed a full-duplex D2D-aided cooperative NOMA scheme and found that the cooperative NOMA scheme achieves superior outage performance in comparison with the conventional NOMA and OMA. The capacity scaling law of the D2D aided cooperative relaying system (CRS) using NOMA was analyzed in [33], which indicated that the D2D-and-NOMA aided CRS upgrades the data rate compared to the conventional CRSs. In the heterogeneous networks with cooperative NOMA, [34] studied the min-rate and sum-rate maximization problems by jointly optimizing the power allocation, access point selection, and transmission mode switching. By applying the simultaneous wireless information and power transfer (SWIPT) technique into the cellular networks, the authors in [35] proposed a SWIPT-NOMA cooperative protocol, in which the near NOMA users act as energy harvesting relays to help the far NOMA users.

The existing works in [26]–[35] promote the application of NOMA into the D2D communications. Nevertheless, there are still some open problems remained to be further investigated. Specifically, the existing works mainly investigated the throughput maximization or outage probability minimization problems. None of them focused on the user connectivity problem. As previously mentioned, the 5G wireless networks is faced with the challenge of 100 billion connections [8]. There have been some works studying the connectivity-maximization problem in NOMA networks [23], [36]. In [23], a joint subcarrier-and-power allocation algorithm was proposed to maximize the connectivity of the machine-type communications devices in a NB-IoT system. In [36], the

authors adopted a graph-based method to jointly optimize the admission control, power control, user clustering, and channel assignment. However, [23], [36] focus on the cellular-type networks, and their proposed algorithms cannot be applied for the D2D-capable cellular networks. This is because there are cross-tier interference among the DUEs and CUEs in the D2D-capable cellular networks, which is different from the cellular-type networks. On the other hand, the D2D and NOMA can improve the user connectivity by exploiting the degree of freedom of users in spatial and power domains, respectively. To fully exploit the potential of D2D and NOMA in improving user connectivity, a new D2D-and-NOMA integrated framework with appropriate resource management algorithms should be investigated.

III. NETWORK MODEL

In this section, we first introduce the considered network model, followed by the proposed four NOMA-aided spectrum-sharing modes.

A. Network Model

In this paper, we consider a D2D-capable cellular network, which consists of a base station (BS), M DUEs¹, and N CUEs. The sets of the DUEs and the CUEs are denoted as $\mathcal{D} = \{1, 2, \dots, M\}$ and $\mathcal{C} = \{1, 2, \dots, N\}$, respectively. For notational simplicity, m is used as the index of a DUE and n is used as the index of a CUE. The cellular network is an uplink transmission scenario. To describe the wireless channels between the CUEs and the BS, we define $G_{n,B}$ as the channel power gain (CPG) from the CUE n to the BS, which accounts for the path loss, shadowing, and small-scale fading. Similarly, let $G_{m,B}$ denote the CPG from the D2D transmitter m to the BS. Besides, the CPG from the CUE n to the D2D receiver m is defined as $G_{n,m}$, and the CPG from the D2D transmitter m to its receiver is defined as $G_{m,m}$.

In this hybrid network, the DUEs can share the same spectrum with the CUEs under the condition that the normal communications of the CUEs cannot be interrupted by the inserted D2D links. In addition, it is assumed that each CUE has been assigned with one channel in advance. As such, a CUE in our concerned network is also equivalent to an orthogonal channel. To transmit data, a DUE must reuse the channel of a DUE. Specifically, we define $\mathbf{X} = \{x_{n,m} | n \in \mathcal{C}, m \in \mathcal{D}\}$ as the user pairing policy, where $x_{n,m} = 1$ denotes that CUE n is paired with DUE m , otherwise $x_{n,m} = 0$. Furthermore, we define $\mathbf{P}_C = \{P_n | n \in \mathcal{C}\}$ and $\mathbf{P}_D = \{P_m | m \in \mathcal{D}\}$ as the power control policy, where P_n and P_m represent the transmission power of CUE n and DUE m , respectively.

B. Four NOMA-aided Spectrum-Sharing Modes

Different from the traditional D2D underlying cellular networks, the DUEs and the CUEs in our considered network can share the same channel via NOMA. In particular, the BS and the D2D receiver can utilize the SIC technique to eliminate

¹A DUE in this paper represents a D2D pair including a D2D transmitter and a D2D receiver.

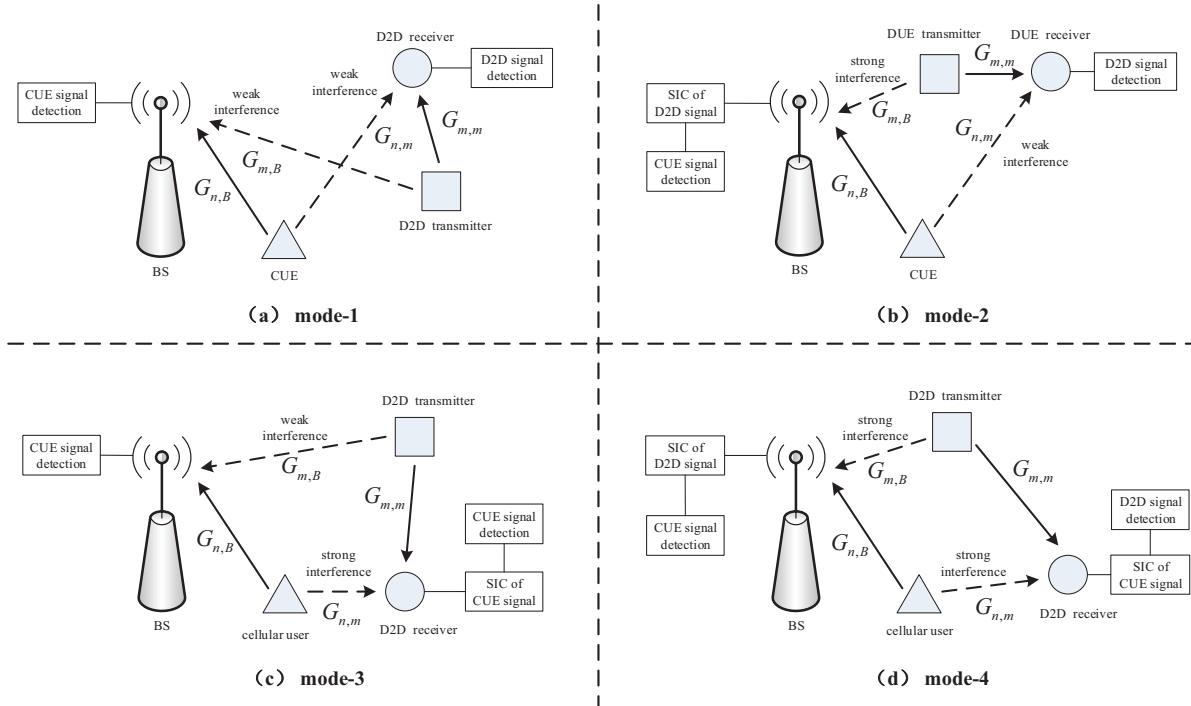


Fig. 1. Four spectrum-sharing modes in D2D-capable cellular networks enhanced by NOMA.

the strong interference and then decode their desired signals. According to different interference status, we propose four NOMA-aided spectrum-sharing modes (i.e., the D2D access scheme) for the paired DUE and CUE², shown in Fig. 1.

The first mode shown in Fig. 1(a) is the same with the traditional underlay mode, that is, both the CUE and the D2D transmitter cause weak interference to their victims (i.e., D2D receiver and the BS). As such, the BS and the D2D receiver can decode their signals directly. In this case, the signal-to-interference-plus-noise ratio (SINR) of the paired CUE and DUE ($x_{n,m} = 1$) must satisfy their respective decoding thresholds, which are specified as

$$\begin{cases} \frac{P_n G_{n,B}}{P_m G_{m,B} + \sigma^2} \geq x_{n,m} \gamma_n \\ \frac{P_m G_{m,m}}{P_n G_{n,m} + \sigma^2} \geq x_{n,m} \gamma_m \end{cases}, \quad (1)$$

where σ^2 denotes the noise power, and γ_n and γ_m represent the decoding thresholds of CUE n and DUE m .

In the second mode as shown in Fig. 1(b), the D2D transmitter causes strong interference to the BS, while the CUE

causes weak interference to the D2D receiver. In this case, the BS can first eliminate the strong interference by utilizing the SIC technique, and then decode its expected signal. On the contrary, the D2D receiver can decode its expected signal directly. To operate in this mode, the following constraints must be satisfied.

$$\begin{cases} \frac{P_m G_{m,B}}{P_n G_{n,B} + \sigma^2} \geq x_{n,m} \gamma_{SIC} \\ \frac{P_n G_{n,B}}{\sigma^2} \geq x_{n,m} \gamma_n \\ \frac{P_m G_{m,m}}{P_n G_{n,m} + \sigma^2} \geq x_{n,m} \gamma_{D_m} \end{cases}, \quad (2)$$

where γ_{SIC} represents the decoding threshold of the SIC.

In the third mode as shown in Fig. 1(c), the CUE causes strong interference to the D2D receiver, while the D2D transmitter causes weak interference to the BS. Similar with the second mode, the D2D receiver can first eliminate the interference and then decode its expected signal, while the BS can successfully decode its signal directly. Therefore, for this mode, the following constraints must be satisfied.

$$\begin{cases} \frac{P_n G_{n,B}}{P_m G_{m,B} + \sigma^2} \geq x_{n,m} \gamma_n \\ \frac{P_n G_{n,m}}{P_m G_{m,m} + \sigma^2} \geq x_{n,m} \gamma_{SIC} \\ \frac{P_m G_{m,m}}{\sigma^2} \geq x_{n,m} \gamma_{D_m} \end{cases}. \quad (3)$$

²In this paper, we consider the two-user NOMA model due to two reasons. On one hand, it is indicated that the performance gain of the NOMA networks with respect to the OMA networks mainly comes from the two-user NOMA transmission [18], [37]. By contrast, the performance gain in the multi-user NOMA transmission is not obvious in comparison with the two-user NOMA transmission, while the system complexity increases dramatically. In consideration of these, the two-user NOMA protocol is more applicable in a practical communication system. On the other hand, the multi-user NOMA model makes our concerned problem too complicated, and no useful insights can be provided.

The fourth mode is depicted in Fig. 1(d), in which both the BS and the D2D receiver undergo strong interference. To decode their desired signals, both the BS and the D2D receiver should adopt the SIC technique to handle the strong interference, such that the following constraints must be satisfied.

$$\left\{ \begin{array}{l} \frac{P_m G_{m,B}}{P_n G_{n,B} + \sigma^2} \geq x_{n,m} \gamma_{SIC} \\ \frac{P_n G_{n,B}}{\sigma^2} \geq x_{n,m} \gamma_{C_n} \\ \frac{P_n G_{n,m}}{P_m G_{m,m} + \sigma^2} \geq x_{n,m} \gamma_{SIC} \\ \frac{P_m G_{m,m}}{\sigma^2} \geq x_{n,m} \gamma_{D_m} \end{array} \right. \quad (4)$$

It is noted that only if one of the above four conditions (1)-(4) is met, the DUE and the CUE can share the same channel without affecting their normal communications. Since the mode-2, mode-3, and mode-4 are the extensions of the traditional underlay mode, the NOMA-and-D2D integrated network is capable of accommodating more D2D connections in comparison with the conventional D2D underlying cellular network, which caters for the requirement of the 5G wireless networks on massive connectivity.

IV. PROBLEM FORMULATION

As introduced above, the performance of the D2D-capable cellular network is highly dependent on the resource allocation schemes of the CUEs and DUEs. Thus, to fully exploit the advantages of the NOMA-and-D2D integrated networks, it is necessary to jointly optimize the user pairing, mode selection, and power control according to the specific system parameters. In this paper, we aim to maximize the accessed D2D links and meanwhile minimize the total power consumption of the paired users. The joint user pairing, mode selection, and power control problem can be mathematically formulated as

$$\begin{aligned} & \max_{\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D} A(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D) - \alpha P(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D) \\ & \text{s.t. C1: (1) or (2) or (3) or (4)} \\ & \text{C2: } \sum_{m=1}^M x_{n,m} \leq 1, \forall n \in \mathcal{C} \\ & \text{C3: } \sum_{n=1}^N x_{n,m} \leq 1, \forall m \in \mathcal{D} \\ & \text{C4: } x_{n,m} \in \{0, 1\}, \forall n \in \mathcal{C}, m \in \mathcal{D} \\ & \text{C5: } 0 \leq P_m \leq P_m^{\max}, \forall m \in \mathcal{D} \\ & \text{C6: } 0 \leq P_n \leq P_n^{\max}, \forall n \in \mathcal{C}, \end{aligned} \quad (5)$$

where $A(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D)$ denotes the number of accessed D2D links, and $P(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D)$ denotes the total power consumption of the paired CUEs and DUEs. The expressions of $A(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D)$ and $P(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D)$ are given by

$$A(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D) = \sum_{n=1}^N \sum_{m=1}^M x_{n,m}, \quad (6)$$

$$P(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D) = \sum_{n=1}^N \sum_{m=1}^M x_{n,m} (P_n + P_m). \quad (7)$$

Besides, α is an weight coefficient, which is set as

$$0 < \alpha < \frac{1}{\sum_{m=1}^M P_m^{\max} + \sum_{n=1}^N P_n^{\max}}. \quad (8)$$

According to (7) and (8), it can be easily get that $\alpha P(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D) < 1$ under any control policy $\{\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D\}$. Therefore, the primary goal of the problem in (5) is to maximize the accessed D2D links, and the secondary goal is to minimize the total power consumption of the paired users. It is noted that if set $A(\mathbf{X}, \mathbf{P}_C, \mathbf{P}_D)$ as the single objective function, the optimal solution of (5) may be non-unique. Furthermore, C1 is the mode selection constraint, which indicates that the paired DUE and CUE must satisfy the conditions of one spectrum-sharing mode shown in Fig. 1. C2 and C3 specify that each CUE can only be paired with one DUE and vice versa. C5 and C6 limit the maximum transmission power of the DUEs and the CUEs, where P_m^{\max} and P_n^{\max} are the power budget of DUE m and CUE n .

Remark 1. The problem in (5) is a mixed-integer programming problem (MIP). Besides, it is needed to determine the optimal spectrum-sharing mode for each paired DUE and CUE from the four possible alternatives. Therefore, it is very hard to directly solve the problem in (5) through the existing optimization algorithms. In order to get the optimal solution of (5) efficiently, we decompose the primal problem into the power control and mode selection subproblem and the user pairing subproblem, and then tackle them through theoretical analysis and graph-based methods. Even so, it is noted that the obtained control policy is still globally optimal, which will be illustrated in the following section.

V. THE OPTIMAL SOLUTION FOR THE FORMULATED PROBLEM

In this section, we first analytically solve the power control and mode selection problem (PMP) for each paired DUE and CUE. Then, we remodel the user pairing problem (UPP) as a min-cost max-flow (MCMF) problem in graph theory and solve it efficiently. Finally, we propose the overall algorithm for the formulated problem and analyze its complexity.

A. The Optimal Power Control and Mode Selection Policies

In (5), the power control variables and the user pairing variables are tightly coupled, which makes the problem very hard to tackle. To make it tractable, we first focus on the PMP for each paired CUE and DUE. Specifically, if DUE m and CUE n is paired, the PMP can be formulated as

$$\begin{aligned} & \min_{P_n, P_m} P_n + P_m \\ & \text{s.t. C1: (1) or (2) or (3) or (4)} \\ & \text{C2: } 0 \leq P_m \leq P_m^{\max} \\ & \text{C3: } 0 \leq P_n \leq P_n^{\max}. \end{aligned} \quad (9)$$

TABLE I
THE OPTIMAL TRANSMISSION POWER OF THE PAIRED DUE m AND CUE n

Mode	P_m^*	P_n^*
Mode-1	$\frac{\sigma^2 \gamma_n (G_{n,B} G_{m,m} + \gamma_m G_{m,B} G_{n,m})}{(G_{n,B})^2 G_{m,m} - \gamma_n \gamma_m G_{m,B} G_{n,m}}$	$\frac{G_{n,m} (\gamma_n \gamma_m + \sigma^2 \gamma_m)}{G_{n,B} G_{m,m} - \gamma_n \gamma_m G_{m,B} G_{n,m}}$
Mode-2	$\max \left\{ \frac{\sigma^2 \gamma_m (\gamma_m G_{n,m} + G_{n,B})}{G_{n,B} G_{m,m}}, \frac{\sigma^2 \gamma_{SIC} (\gamma_m + 1)}{G_{m,B}} \right\}$	$\frac{\sigma^2 \gamma_m}{G_{n,B}}$
Mode-3	$\frac{\sigma^2 \gamma_m}{G_{m,m}}$	$\max \left\{ \frac{\sigma^2 \gamma_m \gamma_n G_{m,B}}{G_{n,B} G_{m,m}} + \frac{\sigma^2 \gamma_m}{G_{n,B}}, \frac{\sigma^2 \gamma_{SIC} (\gamma_m + 1)}{G_{m,B} G_{n,m}} \right\}$
Mode-4	$\max \left\{ \frac{\sigma^2 \gamma_{SIC} (G_{n,m} + \gamma_{SIC} G_{n,B})}{G_{m,B} G_{n,m} - \gamma_{SIC}^2 G_{n,B} G_{m,m}}, \frac{\sigma^2 \gamma_m}{G_{m,m}}, \frac{\sigma^2 \gamma_{SIC} (\gamma_m + 1)}{G_{m,B}} \right\}$	$\max \left\{ \frac{\sigma^2 \gamma_{SIC} (\gamma_{SIC} G_{m,m} + G_{m,B})}{G_{m,B} G_{n,m} - \gamma_{SIC}^2 G_{n,B} G_{m,m}}, \frac{\sigma^2 \gamma_m}{G_{n,B}}, \frac{\sigma^2 \gamma_{SIC} (\gamma_m + 1)}{G_{n,m}} \right\}$

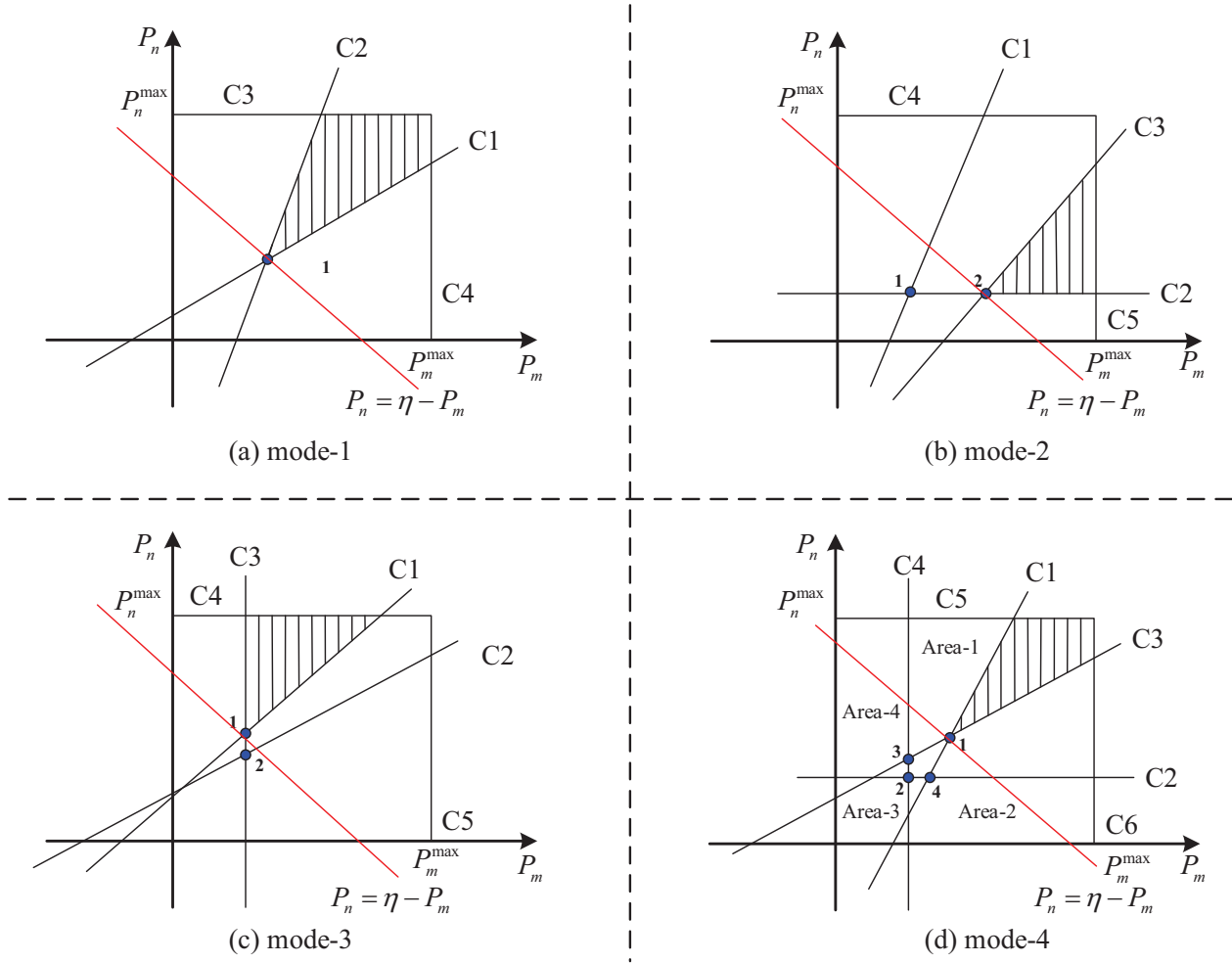


Fig. 2. Analysis for the PMP in the four spectrum-sharing modes.

To solve the problem in (9), we divide it into four independent subproblems, each of which corresponds to a spectrum-sharing mode. Then, we can get the following lemma.

Lemma 1. *If the power budget constraints in (9) (i.e., C2 and C3) are neglected, the explicit expressions of the optimal transmission power of the paired DUE m and CUE n can be obtained, which are summarized in Table I.*

Proof: Rearranging the constraints C1 in (9), we can transform the problem in (9) into a linear programming problem (LP) for each spectrum-sharing mode. Since there are only two

variables in the LP (i.e., P_m and P_n), we can solve it through a graphical method. The LPs corresponding to the four modes are specified as follows.

In mode-1, the power control problem is mathematically formulated as

$$\begin{aligned} \min_{P_n, P_m} & P_n + P_m \\ \text{s.t. C1: } & P_n \geq \frac{\gamma_n G_{m,B}}{G_{n,B}} P_m + \frac{\gamma_n \sigma^2}{G_{n,B}} \\ \text{C2: } & P_n \leq \frac{G_{m,m}}{\gamma_m G_{n,m}} P_m - \frac{\sigma^2}{G_{n,m}} \end{aligned}$$

$$\begin{aligned} \text{C3} : 0 &\leq P_m \leq P_m^{\max} \\ \text{C4} : 0 &\leq P_n \leq P_n^{\max}. \end{aligned} \quad (10)$$

The graphical representation of the above problem is shown in Fig. 2(a), wherein the shadow area denotes the feasible region, and the red line denotes the objective function. As can be seen, without considering the constraints C3 and C4 in (10), the optimal solution can be obtained at the intersection of lines C1 and C2 in (10). Combining the constraints C1 and C2, we can get the explicit expressions of P_m^* and P_n^* , which are given in the second row of Table I.

In mode-2, the power control problem is described in Fig. 2(b) and formulated as the following LP.

$$\begin{aligned} \min_{P_n, P_m} & P_n + P_m \\ \text{s.t. C1} : & P_n \leq \frac{G_{m,B}}{\gamma_{SIC} G_{n,B}} P_m - \frac{\sigma^2}{G_{n,B}} \\ \text{C2} : & P_n \geq \frac{\gamma_n \sigma^2}{G_{n,B}} \\ \text{C3} : & P_n \leq \frac{G_{m,m}}{\gamma_m G_{n,m}} P_m - \frac{\sigma^2}{G_{n,m}} \gamma_m \\ \text{C4} : & 0 \leq P_m \leq P_m^{\max} \\ \text{C5} : & 0 \leq P_n \leq P_n^{\max}. \end{aligned} \quad (11)$$

From Fig. 2(b), we can observe that the optimal solution of (11) is achieved at point-1 (i.e., the intersection of lines C1 and C2) or point-2 (i.e., the intersection of lines C3 and C2). If the horizontal coordinate of point-2 is larger than that of point-1, point-2 is the optimal solution, otherwise point-1 is the optimal solution. Solving the horizontal and vertical coordinates of point-1 and point-2 and then comparing their values, we can obtain the optimal solution of (11), which is summarized in the third row of Table I.

In mode-3, the power control problem is depicted in Fig. 2(c) and formulated as

$$\begin{aligned} \min_{P_n, P_m} & P_n + P_m \\ \text{s.t. C1} : & P_n \geq \frac{\gamma_n G_{m,B}}{G_{n,B}} P_m + \frac{\gamma_n \sigma^2}{G_{n,B}} \\ \text{C2} : & P_n \geq \frac{\gamma_{SIC} G_{m,m}}{G_{n,m}} P_m + \frac{\gamma_{SIC} \sigma^2}{G_{n,m}} \\ \text{C3} : & P_m \geq \frac{\gamma_m \sigma^2}{G_{m,m}} \\ \text{C4} : & 0 \leq P_m \leq P_m^{\max} \\ \text{C5} : & 0 \leq P_n \leq P_n^{\max}. \end{aligned} \quad (12)$$

Similar with mode-2, the optimal solution of (12) is achieved at point-1 (i.e., the intersection of lines C1 and C3) or point-2 (i.e., the intersection of lines C1 and C2), and the one with the larger vertical coordinate is the optimal solution. In the same way, we can get the optimal solution of (12), which is given in the fourth row of Table I.

In mode-4, the power control problem is shown in Fig. 2(d) and formulated as the following problem.

$$\begin{aligned} \min_{P_n, P_m} & P_n + P_m \\ \text{s.t. C1} : & P_n \leq \frac{G_{m,B}}{\gamma_{SIC} G_{n,B}} P_m - \frac{\sigma^2}{G_{n,B}} \\ \text{C2} : & P_n \geq \frac{\gamma_n \sigma^2}{G_{n,B}} \\ \text{C3} : & P_n \geq \frac{\gamma_{SIC} G_{m,m}}{G_{n,m}} P_m + \frac{\gamma_{SIC} \sigma^2}{G_{n,m}} \\ \text{C4} : & P_m \geq \frac{\gamma_m \sigma^2}{G_{m,m}} \\ \text{C5} : & 0 \leq P_m \leq P_m^{\max} \\ \text{C6} : & 0 \leq P_n \leq P_n^{\max}. \end{aligned} \quad (13)$$

For the above problem, there are four cases: 1) If point-1 (i.e., the intersection of lines C1 and C3) is located in area-1, the optimal solution is achieved at point-1, just as shown in Fig. 2(d). 2) If point-1 is located in area-2, the optimal solution is achieved at point-4 (i.e., the intersection of lines C1 and C2). 3) If point-1 is located in area-3, the optimal solution is achieved at point-2 (i.e., the intersection of lines C2 and C4). 4) If point-1 is located in area-4, the optimal solution is achieved at point-3 (i.e., the intersection of lines C3 and C4). To summarize, it finds that the optimal solution is achieved at the right upper one of the four points. According to this rule, we can get the explicit expressions of the optimal solution of (13), given in the last row of Table I. ■

Although the power budget is not considered in Lemma 1, we can easily acquire the optimal solution of (9) through a further judgment. More specifically, for a given spectrum-sharing mode, if $0 \leq P_m^* \leq P_m^{\max}$ and $0 \leq P_n^* \leq P_n^{\max}$, it means that this mode is feasible for the paired DUE and CUE. As previously mentioned, only if one of the four modes is feasible, the paired DUE and CUE can share the same channel.

In the power control and mode selection stage, our goal is to minimize the power consumption of each CUE-DUE pair. Firstly, we analytically obtain the optimal transmission power for each CUE-DUE pair in the four modes, such that we can know which modes are feasible. Among the feasible modes, we then select the one that leads to the minimum total transmission power as the optimal mode. By this way, the obtained power control and mode selection policies are optimal for the problem in (9).

B. The Optimal User Pairing Policy

In this subsection, we aim at solving the UPP based on the transmission power. In particular, let $P_{n,m}$ denote the optimal value of (9). If the problem in (9) is feasible, $P_{n,m}$ is set as

$$P_{n,m} = P_n^* + P_m^*. \quad (14)$$

If the problem in (9) is infeasible, $P_{n,m}$ is set as a very large value, so as to make $(1 - \alpha P_{n,m}) < 0$. By this way, the user pairing variable $x_{n,m}$ must be zero, that is, the CUE and DUE

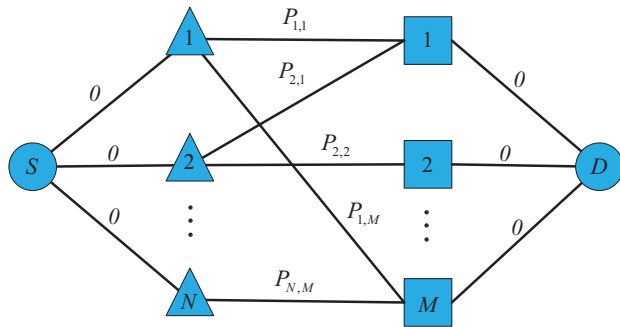


Fig. 3. A graph model for the user pairing problem.

cannot be paired. Afterward, the UPP can be recast as the following problem.

$$\begin{aligned}
 & \max_{\mathbf{X}} \sum_{m=1}^M \sum_{n=1}^N x_{n,m} (1 - \alpha P_{n,m}) \\
 & \text{s.t. C1: } \sum_{m=1}^M x_{n,m} \leq 1, \forall n \in \mathcal{C} \\
 & \quad \text{C2: } \sum_{n=1}^N x_{n,m} \leq 1, \forall m \in \mathcal{D} \\
 & \quad \text{C3: } x_{n,m} \in \{0, 1\}, \forall n \in \mathcal{C}, m \in \mathcal{D}. \quad (15)
 \end{aligned}$$

Remark 2. The problem in (15) is an integer programming problem, which can be solved optimally by the Bender’s decomposition and Branch-and-Bound methods [38]. However, the computational complexity of these methods is still very high. To reduce the computational complexity, we remodel the problem in (15) as a MCMF problem in graph theory and then solve it efficiently.

As shown in Fig. 3, we first construct an undirected graph $G_I(V_I, E_I)$, wherein V_I denotes the set of nodes, and E_I denotes the set of edges. In V_I , S and D are two virtual nodes, representing the source and the destination of a path. The intermediate nodes $\{1, 2, \dots, N\}$ and $\{1, 2, \dots, M\}$ correspond to the CUEs and DUEs in the network, respectively. Between the source node S and each node in $\{1, 2, \dots, N\}$, there exists an undirected edge. Similarly, there is an undirected edge between the destination node D and each node in $\{1, 2, \dots, M\}$. Different from the former, the existence of the edges between the nodes in $\{1, 2, \dots, N\}$ and the nodes in $\{1, 2, \dots, M\}$ are related to the power control solutions of (9). For instance, if CUE 1 and DUE 1 can be successfully paired, an edge will exist between these two nodes, as shown in Fig. 3. On the contrary, since CUE 1 and DUE 2 cannot share the same channel in any modes, there is no edge between these two nodes. Besides, we define a cost for each edge in $G_I(V_I, E_I)$. As depicted in Fig. 3, if the edge $(1, M)$ is utilized, it will consume the cost $P_{1,M}$, i.e., the minimum total transmission power of CUE 1 and DUE M .

Before explaining the relationship between the problem in (15) and the constructed graph $G_I(V_I, E_I)$, we first present

some concepts in graph theory.

Definition 1. (Edge-Disjoint Paths) If two paths have no common edges, they are called as edge-disjoint paths.

Definition 2. (Node-Disjoint Paths) If two paths have no common nodes in addition to the source and the destination, they are called as node-disjoint paths.

Definition 3. (Connectivity Degree) The connectivity degree of two nodes in a graph is defined as the maximum number of node-disjoint paths between the two nodes.

Definition 4. (Cost of A Path) The cost of a path is equal to the sum of the cost of all edges in the path.

Definition 5. (Cut-Set) A set (node or edge set) is called as a cut-set if removing it leads to the disconnection of the graph.

Lemma 2. The problem in (15) is equivalent to finding K node-disjoint paths from S to D in $G_I(V_I, E_I)$ and making the total cost of the paths minimized, where K is the connectivity degree of S and D in $G_I(V_I, E_I)$.

Proof: As shown in Fig. 3, the intermediate nodes $\{1, 2, \dots, N\}$ and $\{1, 2, \dots, M\}$ correspond to the CUEs and DUEs in the network respectively. As such, the edges between $\{1, 2, \dots, N\}$ and $\{1, 2, \dots, M\}$ can be regarded as the binary variables in (15). Specifically, if the link (m, n) in $G_I(V_I, E_I)$ is selected, the corresponding variable $x_{m,n}$ in (15) is equal to one, otherwise $x_{m,n} = 0$. Besides, the selected paths from S to D are required to be node-disjoint, which confirms that the constraints C1 and C2 in (15) are met. Thus, the path selection problem satisfies all of the constraints in (15).

On the other hand, the connectivity degree of S and D in $G_I(V_I, E_I)$ is determined by the adjacency between $\{1, 2, \dots, N\}$ and $\{1, 2, \dots, M\}$. As previously mentioned, the adjacency between $\{1, 2, \dots, N\}$ and $\{1, 2, \dots, M\}$ is relevant to the power control solution of (9), which reflects the potential user pairing policy. The connectivity degree of S and D is just equal to the maximum number of accessed D2D links. In addition, it requires that the K selected paths should consume the minimum total cost, that is, the total transmission power of the paired DUEs and CUEs is minimized. Hence, the path selection problem can make the objective function in (15) maximized.

Therefore, the path selection problem in $G_I(V_I, E_I)$ is equivalent to the user pairing problem in (15). ■

The path selection problem given in Lemma 2 is still an integer programming problem, which is hard to tackle. In order to efficiently solve it, we transform the initial graph $G_I(V_I, E_I)$ into another form $G_T(V_T, E_T)$, which is shown in Fig. 4. In $G_T(V_T, E_T)$, we add $M + N$ nodes $\{1', 2', \dots, N'\}$ and $\{1', 2', \dots, M'\}$, each of which is a copy of the nodes $\{1, 2, \dots, N\}$ and $\{1, 2, \dots, M\}$ in $G_I(V_I, E_I)$. Different from the initial graph $G_I(V_I, E_I)$, the transformed graph $G_T(V_T, E_T)$ is directional. Besides, we define the capacity of each edge in $G_T(V_T, E_T)$ as one, such that the cost of an edge becomes the expense when a unit flow passes through this edge. For $G_T(V_T, E_T)$, we have the following Lemma.

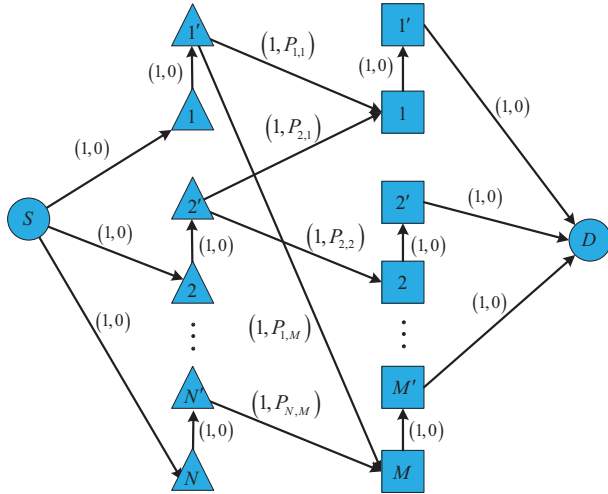


Fig. 4. A min-cost max-flow model for the user pairing problem.

Lemma 3. *The path selection problem given in Lemma 2 is equivalent to the MCMF problem in $G_T(V_T, E_T)$, where the source and the sink are S and D , respectively.*

Proof: The connectivity degree of S and D in $G_I(V_I, E_I)$ is K . Accordingly, the size of the smallest node-cut-set of S and D in $G_I(V_I, E_I)$ is also equal to K . Denote \mathcal{U} as the smallest node-cut-set, where $\mathcal{U} \subseteq \{1, 2, \dots, N\} \cup \{1', 2', \dots, M'\}$ and $|\mathcal{U}| = K$. It can be observed from Fig. 3 and Fig. 4 that there is a one-to-one correspondence between the new added edges in $G_T(V_T, E_T)$ (i.e., (n, n') or (m, m')) and the intermediate nodes in $G_I(V_I, E_I)$ (i.e., n or m). As such, the edges corresponding to the nodes in \mathcal{U} constitute a minimum edge-cut-set of S and D in $G_T(V_T, E_T)$. The Max-Flow Min-Cut Theorem in graph theory indicates that the maximum flow is equal to the minimum edge-cut-set [39]. Hence, the maximum flow with source as S and sink as D in $G_T(V_T, E_T)$ is equal to K .

On the other hand, the network flow theory indicates that if the capacity of each edge in a graph is one, the most edge-disjoint paths between two nodes can be found by solving the corresponding maximum flow problem. Comparing the features of $G_I(V_I, E_I)$ and $G_T(V_T, E_T)$, we can find that each edge-disjoint path from S to D in $G_T(V_T, E_T)$ corresponds to a node-disjoint path from S to D in $G_I(V_I, E_I)$, and their costs are identical. Therefore, the path selection problem given in Lemma 2 can be tackled by solving the MCMF problem in $G_T(V_T, E_T)$.

To this end, we have proofed Lemma 3. \blacksquare

The MCMF problem has been well studied in graph theory, and there are a lot of algorithms with low complexity that can solve it optimally. By this equivalent transformation, the computing complexity of finding the optimal solution of (15) is reduced significantly.

C. The Overall Algorithm

The overall algorithm is summarized in Algorithm 1. In this algorithm, we first calculate the optimal transmission power for each paired CUE and DUE in each spectrum-sharing mode (step-4). Then, we can get the optimal mode for the paired CUE and DUE (step-5). Based on the obtained transmission power and mode selection solutions, we construct the initial graph $G_I(V_I, E_I)$ and then generate the transformed graph $G_T(V_T, E_T)$. Through solving the MCMF problem in $G_T(V_T, E_T)$, we can get the optimal user pairing policy \mathbf{X}^* (step-8 to step-10). It's worth pointing out that although the PMP and the USP are tackled separately, the obtained control policy $\{\mathbf{X}^*, \mathbf{P}_C^*, \mathbf{P}_D^*\}$ is still globally optimal for the problem in (5). This is because every possible power control policy and user pairing policy are taken into account in our proposed algorithm, and the final solution is selected from them.

The computational complexity of Algorithm 1 is dominated by step-4 and step-10. In step-4, we should calculate P_n^* and P_m^* for each possible combination of DUE and CUE in four modes. The total number of DUE-CUE combinations is MN , thereby the complexity of step-4 is $O(8MN)$. In step-10, we can adopt the Ford-Fulkerson algorithm [40] to solve the MCMF problem in $G_T(V_T, E_T)$. The complexity of the Ford-Fulkerson algorithm is $O(FE)$, where F denotes the max-flow, and E denotes the number of edges in the graph. In $G_T(V_T, E_T)$, the maximum value of the max-flow from S to D is equal to the number of DUEs (i.e., M), and the maximum number of edges is $2 + 2N + 2M + NM$. Thus, the complexity of step-10 is $O(2M + 2NM + 2M^2 + NM^2)$. In a practical system, the number of DUEs and CUEs are usually much larger than 2, thus the complexity of step-10 is approximate to $O(NM^2)$. Therefore, the total computational complexity of Algorithm 1 is $O((8 + M)MN)$. Furthermore, if $M \gg 8$ (the scenario with massive connections), the total complexity of Algorithm 1 is in the order of $O(NM^2)$.

Algorithm 1 The Overall Algorithm for the Problem in (5).

- 1: **Input:** The decoding thresholds and the CPGs of all users.
 - 2: **for** $n = 1 : 1 : N$ **do**
 - 3: **for** $m = 1 : 1 : M$ **do**
 - 4: Calculate the optimal transmission power of CUE n and DUE m in the four spectrum-sharing modes according to Lemma 1;
 - 5: Obtain the optimal power control $\{\mathbf{P}_C^*, \mathbf{P}_D^*\}$ and mode selection solutions of the problem in (9) and set the value of $P_{n,m}$;
 - 6: **end for**
 - 7: **end for**
 - 8: Construct the initial graph $G_I(V_I, E_I)$ based on the obtained power control and mode selection solutions.
 - 9: Generate the transformed graph $G_T(V_T, E_T)$ according to the initial graph $G_I(V_I, E_I)$.
 - 10: Obtain the user pairing policy \mathbf{X}^* by solving the MCMF problem in $G_T(V_T, E_T)$.
 - 11: **Output:** The optimal control policy $\{\mathbf{X}^*, \mathbf{P}_C^*, \mathbf{P}_D^*\}$.
-

TABLE II
SIMULATION PARAMETERS

Cell radius, R	500~1000 m
Path loss	$32.44+20\lg(d[km])+20\lg(f_c[MHz])$
Shadowing	Log normal as $\mathcal{N}(0, 4^2)$
Fading	Rayleigh fading with 1 variance
Noise Power, σ^2	-150 dBm/Hz
Power budget, P_n^{\max}, P_m^{\max}	2 Watt
Carrier frequency, f_c	3.5 GHz
Channel bandwidth, B_0	2 MHz
Simulation times	500

VI. SIMULATION RESULTS

In this section, we conduct extensive simulations to investigate the performance of our proposed algorithm. The detailed simulation parameters are listed in Table II. For notational simplicity, our proposed algorithm is denoted as the Hybrid_Mode. To demonstrate the advantages of our algorithm, we compare it with the other six benchmark schemes, namely the Max_Access, Rand_Access, NOMA_Mode1, NOMA_Mode2, NOMA_Mode3, and NOMA_Mode4. The detailed explanation of the six schemes is given as follows.

- **Max_Access:** This scheme adopts the four NOMA-aided spectrum-sharing modes shown in Fig. 1 and our designed power control and mode selection policy given in Subsection V-A. The goal of the Max_Access is to maximize the number of accessed D2D links but without taking into account the power consumption of devices. As such, the user pairing problem can be modeled as a maximum matching problem in graph theory and optimally solved by the Hungarian algorithm [40]. The Max_Access is utilized to verify whether our algorithm can achieve the optimal solution and demonstrate the performance gain of our algorithm in terms of power consumption.
- **Rand_Access.** The Rand_Access also adopts the four NOMA-aided spectrum-sharing modes and our designed power control and mode selection policy. Different from our algorithm, the DUEs are randomly paired with the CUEs in the Rand_Access. This scheme is used as a benchmark to evaluate the performance of our proposed user pairing policy given in Subsection V-B.
- **NOMA_Mode1~NOMA_Mode4:** The resource allocation algorithms of these four schemes are the same with our algorithm. The difference is that each of them only utilize one of the four spectrum-sharing modes. It is noted that the spectrum-sharing mode utilized by the NOMA_Mode1 is just the traditional underlay mode in the D2D underlying cellular networks. By comparing with it, we can investigate the advantages of the NOMA enhanced network architecture. Besides, the other three schemes can demonstrate the specific performance of each spectrum-sharing mode and give some insights for the system design.

A. Number of Accessed D2D Links

Fig. 5 shows the effect of the number of DUEs M on the number of accessed D2D links (NAD). From this figure, we

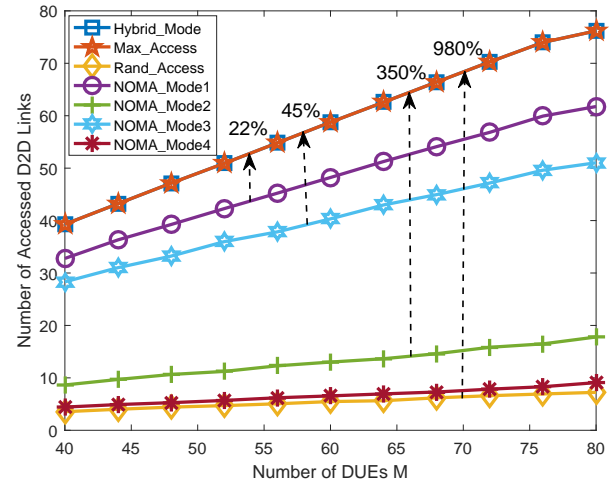


Fig. 5. The effect of the number of DUEs on the number of accessed D2D links ($N = 100$, $\gamma_n = 50$, $\gamma_m = 100$, $\gamma_{SIC} = 50$, and $R = 500$ m).

can observe that the NAD increases almost linearly with the number of DUEs. This is because with the four spectrum-sharing modes, almost all of the DUEs can access to the cellular network, as shown by the Max_Access and Hybrid_Mode. Furthermore, the Hybrid_Mode has the same performance with the Max_Access in terms of the NAD. Since the optimal power control policy is obtained by the graphical method, our proposed algorithm can achieve the optimal solution of the formulated problem in (5). Moreover, the simulation results indicate that the Hybrid_Mode can increase about 22% NAD with respect to the NOMA_Mode1 (i.e., the underlay mode). This result illustrates that the NOMA technique has great potential in improving D2D connections. However, if the DUEs and the CUEs are not paired appropriately, the performance of the NOMA enhanced network is still very poor, as shown by the Rand_Access. Thus, to fully exploit the advantages of NOMA, the user pairing policy between DUEs and CUEs must be carefully designed.

Fig. 6 plots the NAD versus the number of CUEs N . This figure also shows that the Hybrid_Mode greatly outperforms the other schemes with any number of CUEs, which verifies the effectiveness of our proposed algorithm. On the other hand, we can find that the performance of the NOMA_Mode3 is much better than the NOMA_Mode2 and NOMA_Mode4. The reason is that the distance between the D2D transmitter and the BS is relatively far, which results in that the decoding threshold of the SIC at the BS is hardly satisfied under the power budget constraint of DUEs. By contrast, there are multiple DUEs and CUEs in the network, such that there will always be some CUEs near the D2D receivers. As a consequence, the decoding threshold of the SIC at the D2D receivers can be easily satisfied especially when the number of CUEs is large. As shown in Fig. 6, with the increment of the number of CUEs, the NAD of the NOMA_Mode3 increases a lot, while the NAD of the NOMA_Mode2 and NOMA_Mode4 almost remains the same. These results tell us that the performance gain of our algorithm with respect to the NOMA_Mode1 mainly comes

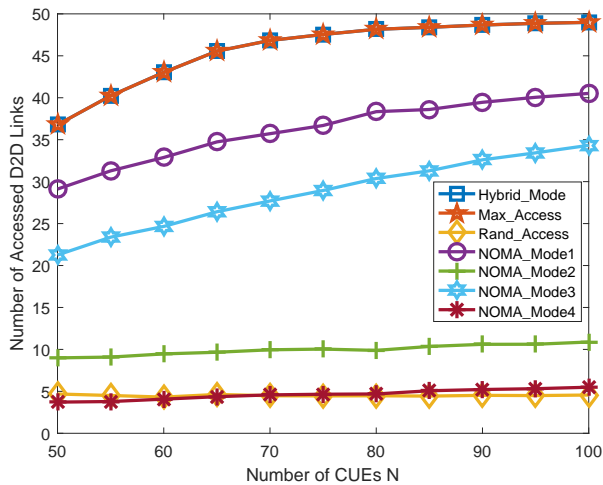


Fig. 6. The effect of the number of CUEs on the number of accessed D2D links ($M = 50$, $\gamma_n = 50$, $\gamma_m = 100$, $\gamma_{SIC} = 50$, and $R = 500$ m).

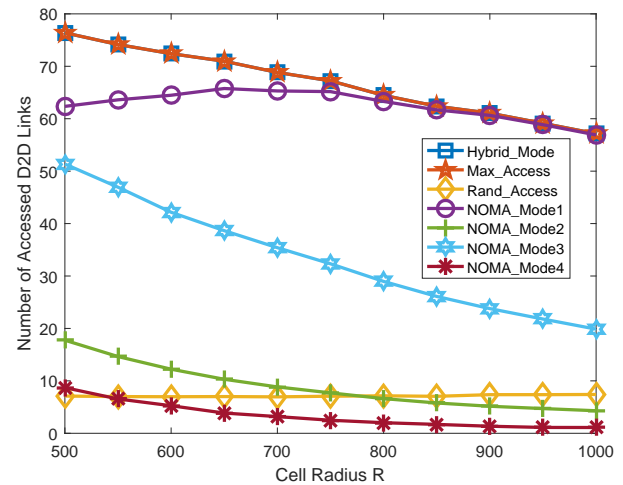


Fig. 8. The effect of the cell radius on the number of accessed D2D links ($N = 100$, $M = 80$, $\gamma_n = 50$, $\gamma_m = 100$, and $\gamma_{SIC} = 50$).

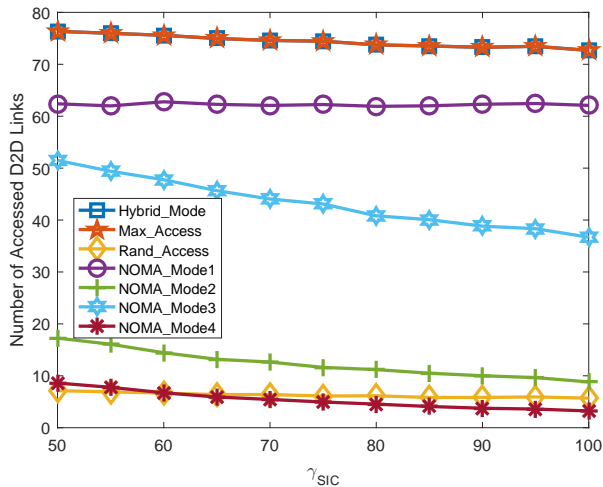


Fig. 7. The effect of the SIC decoding threshold on the number of accessed D2D links ($N = 100$, $M = 80$, $\gamma_n = 50$, and $\gamma_m = 100$).

from the utilization of the mode-3. Therefore, if the complexity of the network is limited, we can only employ the mode-1 and mode-3 for the DUEs, which can still achieve a good performance.

Fig. 7 illustrates the effect of the SIC decoding threshold γ_{SIC} on the NAD. With the increment of γ_{SIC} , the constraints of the mode-2, mode-3, and mode-4 become harder to be satisfied. In addition to the NOMA_Mode1, the other schemes adopt the SCI technique. Thus, the NAD of the schemes except for the NOMA_Mode1 decreases with γ_{SIC} . Since the Hybrid_Mode and the Max_Access also utilize the mode-1, γ_{SIC} has a relatively small impact on their performance. On the contrary, the successful operating probability of the NOMA_Mode2, NOMA_Mode3, and NOMA_Mode4 is highly dependent on the value of γ_{SIC} . As such, γ_{SIC} has a great influence on their performance. This figure shows that our algorithm has a stable performance even when the SIC decoding threshold is high, which is a good feature for its

application in practical systems.

Fig. 8 shows the NAD versus the cell radius R . First, we can observe that the NAD decreases with R . The reason is that with the increment of R , the CUEs must improve their transmission power to satisfy the decoding threshold γ_n . Since the power budget of the CUEs is limited by P_n^{max} , less DUEs can be allowed to access to the cellular network in order to satisfy the transmission demand of CUEs. Furthermore, we can find that the gap between the Hybrid_Mode and the NOMA_Mode1 narrows when R becomes large. This is because when R is large, the distance between CUEs and DUEs becomes far. As a consequence, the cross-tier interference between CUEs and DUEs degrades. Then, the BS and the D2D receivers can successfully decode their desired signals directly. Accordingly, the gap between the Hybrid_Mode and NOMA_Mode1 narrows. However, it is noted that the performance of the Hybrid_Mode is not worse than that of the NOMA_Mode1, as the solutions of the Hybrid_Mode can cover those of the NOMA_Mode1. This result indicates that NOMA is more suitable for the dense wireless networks, as it can well coordinate the serious cross-tier interference. For the traditional macro cellular network, we can adopt the traditional underlay mode for the D2D devices to reduce operation overhead.

B. Power Consumption

Fig. 9 presents the total transmission power (TTP) of the paired DUEs and CUEs versus the number of DUEs M . As shown in Fig. 5, the NAD increases with the number of DUEs. Hence, the TTP also increases with the number of DUEs. This figure shows that the TTP of the Hybrid_Mode is much smaller than that of the Max_Access. This result demonstrates the necessity of power control in the D2D-capable cellular networks. Besides, the simulation results shown in Fig. 5 and Fig. 9 indicate that our algorithm can support massive connections and significantly reduce the power consumption of devices. The IoT with a massive number of mobile devices

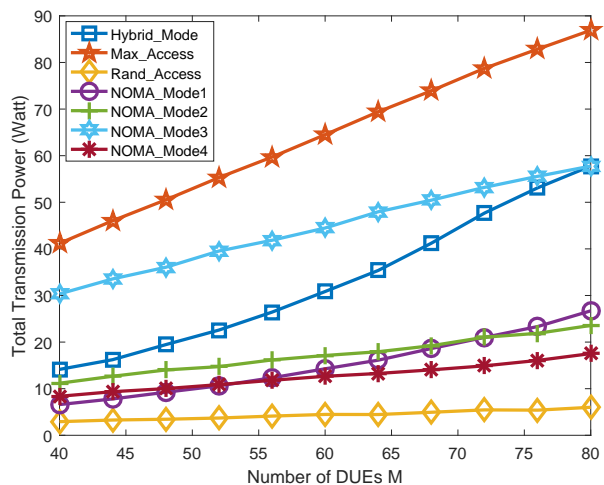


Fig. 9. Total transmission power versus the number of DUEs ($N = 100$, $\gamma_n = 50$, $\gamma_m = 100$, $\gamma_{SIC} = 50$, and $R = 500$).

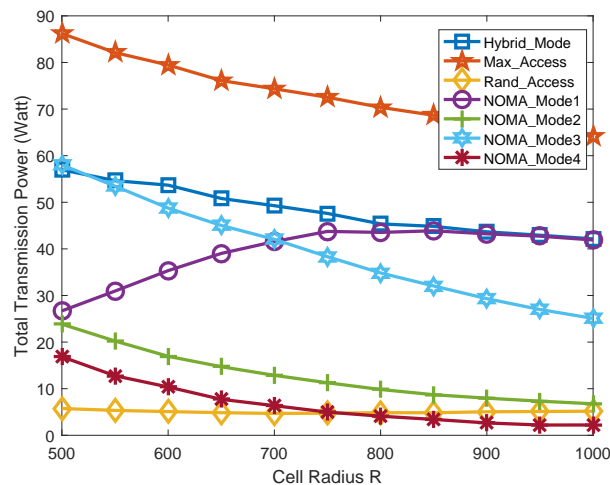


Fig. 11. Total transmission power versus the cell radius ($N = 100$, $M = 80$, $\gamma_n = 50$, $\gamma_m = 100$, and $\gamma_{SIC} = 50$).

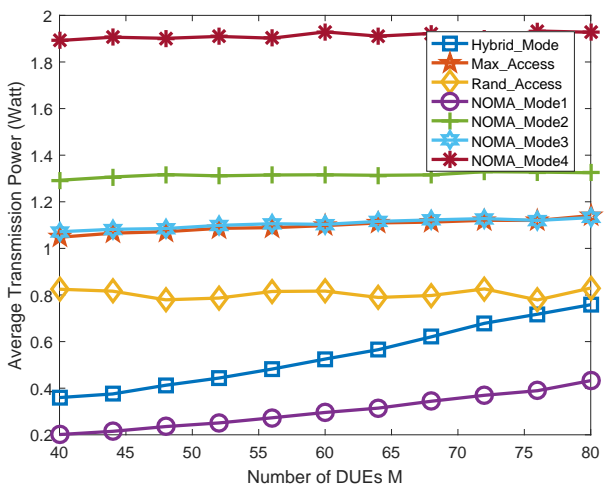


Fig. 10. Average transmission power versus the number of DUEs ($N = 100$, $\gamma_n = 50$, $\gamma_m = 100$, $\gamma_{SIC} = 50$, and $R = 500$).

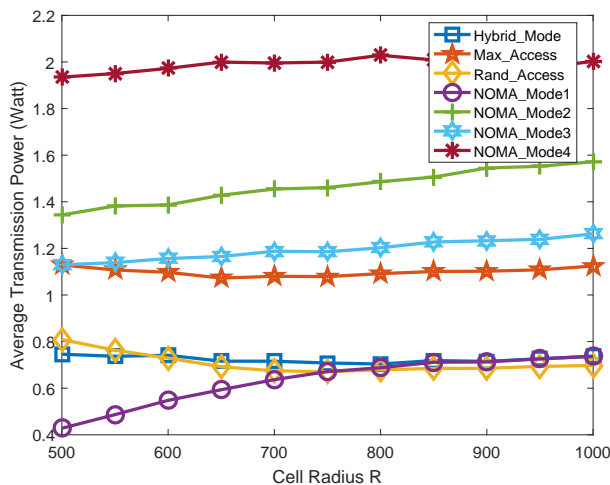


Fig. 12. Average transmission power versus the cell radius ($N = 100$, $M = 80$, $\gamma_n = 50$, $\gamma_m = 100$, and $\gamma_{SIC} = 50$).

is an important application of future wireless networks. Our algorithm is very suitable for the IoT scenario. On the other hand, we can observe from this figure that the TTP of the Hybrid_Mode is larger than the NOMA_Mode1, NOMA_Mode2, NOMA_Mode4, and Rand_Access. The reason is that the Hybrid_Mode accommodate more CUEs and DUEs than these schemes, such that the TTP is large accordingly. For the NOMA_Mode3, the SIC is implemented at the D2D receivers. To satisfy the SIC decoding threshold γ_{SIC} , the transmission power of the D2D transmitters should be improved greatly. Since the NAD of the NOMA_Mode3 is large (as shown in Fig. 5), the generated TTP is thus very high. This accounts for why the TTP of the NOMA_Mode3 is larger than the Hybrid_Mode.

Fig. 10 depicts the effect of the number of DUEs on the average transmission power (ATP) of the paired DUEs and CUEs. The ATP is equal to the TTP dividing the total number of paired DUEs and CUEs, which reflects the power

consumption of each device. Different from the results shown in Fig. 9, it can be observed from this figure that the ATP of the Hybrid_Mode is smaller than the other schemes except for the NOMA_Mode1. Therefore, our algorithm also exhibits good performance in energy efficiency, which is vital for the IoT devices with limited energy. Besides, the ATP of the NOMA_Mode4 is much larger than the others', as it should implement SIC at both the BS and the D2D receivers. The transmission power of the CUEs and the D2D transmitters should be greatly improved to enhance the interference signals. On the contrary, in the NOMA_Mode1, the CUEs are only paired with the DUEs that cause weak interference, thereby it consumes the less ATP.

Fig. 11 shows the TTP versus the cell radius R . In accordance with the NAD shown in Fig. 8, the TTP decreases with R . When R is large, the TTP of the Hybrid_Mode and the NOMA_Mode1 becomes the same. This reflects that our algorithm mainly adopts the mode-1 in a macro cellular

network. The effect of R on the ATP is illustrated in Fig. 12. This figure shows that our algorithm consumes very small power in comparison with the other schemes. Although the Max_Access yields the maximum NAD, it consumes the largest power as well. On the contrary, the NOMA_Model can reduce the power consumption at the expense of reducing the device connections. Our algorithm strikes a good balance between the device connections and the power consumption, that is, it can maximize the NAD and meanwhile lead to small power consumption.

VII. CONCLUSION

In this paper, we have investigated the connectivity-maximization problem for the NOMA enhanced D2D-capable cellular networks. Firstly, we have proposed a NOMA-aided D2D access scheme, where the D2D users can operate in four spectrum-sharing modes. Then, we have formulated a resource allocation problem by jointly considering user pairing, mode selection, and power control with the objective to maximize the accessed D2D links and meanwhile minimize the total power consumption. In order to solve the formulated problem efficiently, an algorithm with polynomial complexity has been devised based on the methods in graph theory. Finally, simulation results have been provided to verify the efficiency of our proposed D2D access scheme and the resource management algorithm.

REFERENCES

- [1] IMT-2020 5G Promotion Group, "5G wireless technology architecture," White paper, May 2015.
- [2] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, S. Li, and G. Feng, "Device-to-device communications in cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 49–55, Apr. 2014.
- [3] F. Malandrino, C. Casetti, and C. Chiasserini, "Toward D2D-enhanced heterogeneous networks," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 94–100, Nov. 2014.
- [4] 3GPP TR36.843, "Study on LTE device-to-device proximity service [S]," Tech. Rep., Sep. 2014.
- [5] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [6] R. Ma, N. Xia, H. Chen, C. Chiu, and C. Yang, "Mode selection, radio resource allocation, and power coordination in D2D communications," *IEEE Commun. Mag.*, vol. 24, no. 3, pp. 112–121, June 2017.
- [7] D. Zhai, M. Sheng, X. Wang, Z. Sun, C. Xu, and J. Li, "Energy-saving resource management for D2D and cellular coexisting networks enhanced by hybrid multiple access technologies," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2678–2692, Apr. 2017.
- [8] IMT-2020 5G Promotion Group, "5G vision and requirements," White paper, May 2014.
- [9] C. H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.
- [10] W. Zhao, S. Wang, and J. Guo, "Efficient resource allocation for OFDMA-based device-to-device communication underlying cellular networks," in *Proc. IEEE ICC'15, Shenzhen, China*, Nov. 2015, pp. 1–6.
- [11] M. Banagar, B. Maham, P. Popovski, and F. Pantisano, "Power distribution of device-to-device communications in underlaid cellular networks," *IEEE Wireless Commun. Lett.*, vol. 5, no. 2, pp. 204–207, Apr. 2016.
- [12] Z. Uysal and R. Jantti, "Transmission-order optimization for bidirectional device-to-device (D2D) communications underlying cellular TDD networks—a graph theoretic approach," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 1–14, Jan. 2016.
- [13] A. Abrardo and M. Moretti, "Distributed power allocation for D2D communications underlying/overlying OFDMA cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1466–1479, Mar. 2017.
- [14] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlying cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3541–3551, Aug. 2013.
- [15] Y. Li, T. Jiang, M. Sheng, and Y. Zhu, "QoS-aware admission control and resource allocation in underlay device-to-device spectrum-sharing networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2874–2886, Nov. 2016.
- [16] M. Sheng, Y. Li, X. Wang, J. Li, and Y. Shi, "Energy efficiency and delay tradeoff in device-to-device communications underlying cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 92–106, Jan. 2016.
- [17] P. Mach, Z. Becvar, and T. Vanek, "In-band device-to-device communication in OFDMA cellular networks: A survey and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1885–1922, Fourthquarter 2015.
- [18] L. Dai, B. Wang, Y. Yuan, S. Han, C. I. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
- [19] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C. L. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.
- [20] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE VTC'13-spring, Dresden, Germany*, Jun. 2013, pp. 1–5.
- [21] Y. Chen, A. Bayesteh, Y. Wu, B. Ren, S. Kang, S. Sun, Q. Xiong, C. Qian, B. Yu, Z. Ding, S. Wang, S. Han, X. Hou, H. Lin, R. Visoz, and R. Razavi, "Toward the standardization of non-orthogonal multiple access for next generation wireless networks," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 19–27, Mar. 2018.
- [22] M. Shirvanimoghaddam, M. Dohler, and S. J. Johnson, "Massive non-orthogonal multiple access for cellular IoT: Potentials and limitations," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 55–61, Sep. 2017.
- [23] A. E. Mostafa, Y. Zhou, and V. W. S. Wong, "Connectivity maximization for narrowband IoT systems with NOMA," in *Proc. IEEE ICC'17, Paris, France*, May 2017, pp. 1–6.
- [24] D. Zhai, R. Zhang, L. Cai, B. Li, and Y. Jiang, "Energy-efficient user scheduling and power allocation for NOMA-based wireless networks with massive IoT devices," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1857–1868, June 2018.
- [25] D. Zhai, R. Zhang, L. Cai, and F. R. Yu, "Delay minimization for massive internet of things with non-orthogonal multiple access," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 553–566, June 2019.
- [26] N. Madani and S. Sodagari, "Performance analysis of non-orthogonal multiple access with underlaid device-to-device communications," *IEEE Access*, vol. 6, pp. 39820–39826, 2018.
- [27] Y. Pan, C. Pan, Z. Yang, and M. Chen, "Resource allocation for D2D communications underlying a NOMA-based cellular network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 130–133, Feb. 2018.
- [28] H. Zheng, S. Hou, H. Li, Z. Song, and Y. Hao, "Power allocation and user clustering for uplink MC-NOMA in D2D underlaid cellular networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 1030–1033, Dec. 2018.
- [29] S. M. A. Kazmi, N. H. Tran, T. M. Ho, A. Manzoor, D. Niyato, and C. S. Hong, "Coordinated device-to-device communication with non-orthogonal multiple access in future wireless cellular networks," *IEEE Access*, vol. 6, pp. 39860–39875, 2018.
- [30] A. Anwar, B. Seet, S. F. Hasan, X. J. Li, P. H. J. Chong, and M. Y. Chung, "An analytical framework for multi-tier NOMA networks with underlay D2D communications," *IEEE Access*, vol. 6, pp. 59221–59241, 2018.
- [31] 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.
- [32] Z. Zhang, Z. Ma, M. Xiao, Z. Ding, and P. Fan, "Full-duplex device-to-device-aided cooperative nonorthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4467–4471, May 2017.
- [33] J. Kim, I. Lee, and J. Lee, "Capacity scaling for D2D aided cooperative relaying systems using NOMA," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 42–45, Feb. 2018.

- [34] J. Liu, S. Xiao, X. Zhou, G. Y. Li, G. Wu, and S. Li, "Optimal mobile association and power allocation in device-to-device-enable heterogeneous networks with non-orthogonal multiple access protocol," in *Proc. IEEE ICC, Kansas City, MO, USA*, May 2018, pp. 1–6.
- [35] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.
- [36] D. Zhai and R. Zhang, "Joint admission control and resource allocation for multi-carrier uplink NOMA networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 922–925, Dec. 2018.
- [37] 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.
- [38] M. Sheng, D. Zhai, X. Wang, Y. Li, Y. Shi, and J. Li, "Intelligent energy and traffic coordination for green cellular networks with hybrid energy supply," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1631–1646, Feb. 2017.
- [39] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.
- [40] J. Bondy and U. Murty, *Graph Theory*. Germany: Springer, 2008.



Daosen Zhai (S'16-M'18) received his B.E. degree in telecommunication engineering from Shandong University, Weihai, China, in 2012, and his Ph.D. degree in communication and information systems from Xidian University, Xi'an, China, in 2017. He is currently an assistant professor with the School of Electronics and Information, Northwestern Polytechnical University. His research interests focus on radio resource management in 5G and energy harvesting networks, satellite communications, and convex optimization and graph theory and their applications in

wireless communications.



Ruonan Zhang (S'09-M'10) received his B.S. and M.Sc. degrees from the Xi'an Jiaotong University, Xi'an, China, in 2000 and 2003, respectively, and his Ph.D. degree from the University of Victoria, Victoria, BC, Canada, in 2010, all in electrical and electronics engineering.

He was an IC design engineer in the Motorola Inc. and Freescale Semiconductor Inc., Tianjin, China, from 2003 to 2006. Since 2010, he has been with the Department of Communication Engineering, Northwestern Polytechnical University, Xi'an, where he

is currently a Professor. His research interests include wireless channel measurement and modeling, architecture and protocol design of wireless networks, and satellite communications.

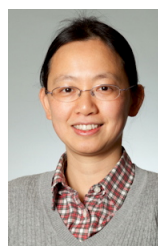
He is a recipient of the New Century Excellent Talent Grant from the Ministry of Education of China. He served as a Local Arrangement Co-Chair for the IEEE/CIC International Conference on Communications in China (ICCC) in 2013, Industry Track and Workshop Chair for the IEEE International Conference on High Performance Switching and Routing (HPSR) in 2019, and as an Associate Editor for the Journal of Communications and Networks.



Yutong Wang received the B.S. degree in communication engineering from the Xi'an University of Science and Technology, China, in 2016. She is currently pursuing the M.S degree in communication engineering with the Northwestern Polytechnical University. Her research interests focus on the optimization of NOMA networks and wireless channel measurement and modeling.



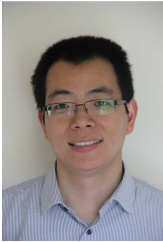
Huakui Sun received the B.E. degree in Telecommunication Engineering and M.E. degree in Electronics and Communications Engineering from Shandong University, Weihai, China, in 2015. He is currently an Assistant with the School of Sergeancy, Weifang University of Science and Technology, Weifang. His current research interests include radio resource management in NOMA networks, wireless sensor networks, and digital image processing.



Lin Cai (S'00-M'06-SM'10) received her M.A.Sc. and PhD degrees (awarded Outstanding Achievement in Graduate Studies) in electrical and computer engineering from the University of Waterloo, Waterloo, Canada, in 2002 and 2005, respectively. Since 2005, she has been with the Department of Electrical & Computer Engineering at the University of Victoria, and she is currently a Professor. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia

traffic and Internet of Things.

She was a recipient of the NSERC E.W.R. Steacie Memorial Fellowships in 2019, the NSERC Discovery Accelerator Supplement (DAS) Grants in 2010 and 2015, respectively, and the Best Paper Awards of IEEE ICC 2008 and IEEE WCNC 2011. She has founded and chaired IEEE Victoria Section Vehicular Technology and Communications Joint Societies Chapter. She has been elected to serve the IEEE Vehicular Technology Society Board of Governors, 2019-2021. She has served as an area editor for *IEEE Transactions on Vehicular Technology*, a member of the Steering Committee of the *IEEE Transactions on Big Data (TBD)* and *IEEE Transactions on Cloud Computing (TCC)*, an Associate Editor of the *IEEE Internet of Things Journal*, *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Vehicular Technology*, *EURASIP Journal on Wireless Communications and Networking*, *International Journal of Sensor Networks*, and *Journal of Communications and Networks (JCN)*, and as the Distinguished Lecturer of the IEEE VTS Society. She has served as a TPC symposium co-chair for IEEE Globecom'10 and Globecom'13. She is a registered professional engineer of British Columbia, Canada.



Zhiguo Ding (S'03-M'05-SM'17) received his B.Eng in Electrical Engineering from the Beijing University of Posts and Telecommunications in 2000, and the Ph.D degree in Electrical Engineering from Imperial College London in 2005. From Jul. 2005 to Apr. 2018, he was working in Queen's University Belfast, Imperial College, Newcastle University and Lancaster University. Since Apr. 2018, he has been with the University of Manchester as a Professor in Communications. From Oct. 2012 to Sept. 2018, he has also been an academic visitor in

Princeton University.

Dr Ding' research interests are 5G networks, game theory, cooperative and energy harvesting networks and statistical signal processing. He is serving as an Editor for *IEEE Transactions on Communications*, *IEEE Transactions on Vehicular Technology*, and *Journal of Wireless Communications and Mobile Computing*, and was an Editor for *IEEE Wireless Communication Letters*, *IEEE Communication Letters* from 2013 to 2016. He received the best paper award in IET ICWMC-2009 and IEEE WCSP-2014, the EU Marie Curie Fellowship 2012-2014, the Top IEEE TVT Editor 2017, IEEE Heinrich Hertz Award 2018 and the IEEE Jack Neubauer Memorial Award 2018.