This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2947021, IEEE Transactions on Vehicular Technology

Performance Optimization for D2D Communications with Opportunistic Relay and Physical-layer Network Coding

Shijun Lin, Yong Li, *Senior Member, IEEE*, Haichuan Ding, Yuguang Fang, *Fellow, IEEE*, Jianghong Shi

***Abstract*—In this paper, we investigate the joint signal to interference plus noise ratio (SINR) thresholds optimization and resource allocation to maximize the sum-rate of Device-to-Device (D2D) communications while still retaining the rate requirements for active cellular users (CUs), when the inactive CUs are used as opportunistic relays under three operational modes: without using network-coding (NNC), using traditional high-layer network-coding (HNC), and using physical-layer network-coding (PNC). Under Rayleigh fading, we show that, given the selections of relays, this sum-rate maximization in no-relay scheme, NNC, HNC, and PNC opportunistic relay schemes can be formulated as a mixed integer non-linear programming (MINLP), which is NP-hard in general. To find the solution to the MINLP, we propose a two-step approach to solve the problem: 1) for each possible pairing of a D2D pair and a CU, we derive the optimal SINR thresholds to obtain the maximum transmission rate of the D2D pair while satisfying the rate requirement of the CU; 2) based on the maximum transmission rates of D2D pairs for each possible pairing in the first step, we develop a bipartite-matching method to find the optimal pairing CUs for D2D pairs. Finally, according to the solution to the MINLP, we propose an iterative relay selection algorithm to find out the relays that can further improve the sum-rate of D2D communications. Extensive simulation results demonstrate that, compared with the scenario without relaying, the NNC, HNC, and PNC opportunistic relay schemes achieve a maximum performance enhancement of 106%, 138%, and 168%, respectively.**

***Index Terms*—Device-to-Device communication, opportunistic relay, physical-layer network coding, Rayleigh fading channel.**

I. INTRODUCTION

As smart mobile phones become more and more popular, mobile data traffic grows exponentially fast [1]. To meet such a massive consumer demand for mobile data access, Device-to-Device (D2D) communications, which allow devices to communicate with each other directly without having to go through base stations (BSs), have been proposed to further enhance the capacity of cellular networks [2]–[6]. Moreover, physical-layer network coding (PNC) has been considered as one of the promising physical-layer technologies that can greatly improve the transmission efficiency [7]. An interesting idea emerges: what happens when D2D transmissions meet PNC? Fig. 1 shows the cases of D2D transmissions with and without adopting relay and PNC. From Fig. 1, we can see that when we introduce a relay node in a D2D transmission and adopt PNC, only two stages of transmission are needed. In the first stage, the two D2D nodes transmit packets concurrently to the relay node, and the relay node extracts network coding packets from the superimposed EM waves. In the second stage, the relay node broadcasts the network coding (NC) packets to the two D2D nodes. Obviously, due to the shortening of the transmission distance, the data rate of each transmission stage with relay and PNC is much higher than the case without relay and PNC. Therefore, the average transmission rate of D2D transmissions can be increased by adopting relays and PNC.

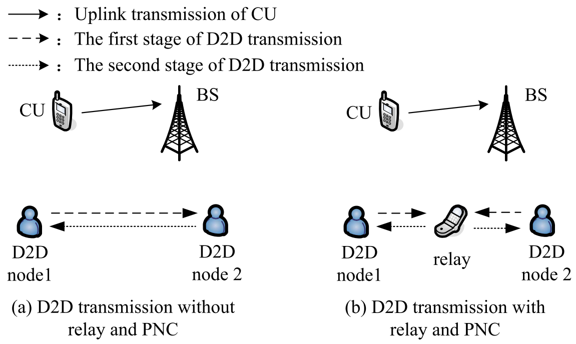


Fig. 1. D2D transmissions with and without relay and PNC.

Considering that in the scenarios with ultra-dense cellular users (CUs) [12], [13], i.e., shopping malls, amusement parks, bus/train stations, railway carriages, congested roads, office buildings, libraries, classrooms, or dormitories, et al., there may exist a large amount of inactive CUs, and part of these inactive CUs may be willing to serve as relay for other CUs if they can earn some rewards (i.e., given amount of free data plan according to their serving time) from the service provider (SP). In this case, we can introduce the inactive CUs as relays and adopt PNC for some D2D transmissions to further improve the end-to-end transmission rate.

In the literature, most of the existing works about D2D

This work was supported in part by the National Natural Science Foundation of China under Grant 61401380, and in part by the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20110121120019. Corresponding author: Shijun Lin.

Shijun Lin and Jianghong Shi are with the Department of Communication Engineering, Xiamen University, Xiamen 361000, Fujian, China, e-mail: linsj@xmu.edu.cn, shijh@xmu.edu.cn. Yong Li is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, e-mail: liyong07@tsinghua.edu.cn. H. Ding is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA, e-mail: dhcbit@gmail.com. Yuguang Fang are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville 32611, FL, USA, e-mail: fang@ece.ufl.edu.

Manuscript received XXX, XX, 2018; revised XXX, XX, 2018.

resource allocations focus on the scenarios without cellular relaying and network coding (i.e. [8]–[11]). In [8], Li et al. investigated the resource allocation of D2D networks by a coalition game. In [9], [10], the authors focused on the sum-rate maximization in the resource-abundant scenario where the number of CUs is more than that of D2D pairs. In [11], Liu et al. studied the outage probability of D2D-enabled multi-channel cellular networks from a general threshold-based perspective. However, what is less understood is how much performance gain we can obtain by introducing relays and adopting PNC in the centralized D2D underlaying cellular networks, especially under a time-varying channel.

Recently, some works started to consider optimizing the performance of D2D networks under the assist of relays without adopting NC schemes [14]–[20]. In [14], Hasan et al. studied the optimization of network throughput when D2D pairs and CUs share some common relays. In [15], [16], the authors discussed the relay selection schemes in D2D networks. In [17], Zhang et al. proposed a source-relay joint power allocation scheme for the relay aided D2D networks. In [18], Ebrahimi et al. investigated the D2D data transfer through multihop relay links. In [19], [20], the authors proposed stochastic-geometry-based analytical frameworks for the relay-assisted D2D overlaying multi-channel cellular network. The performance of D2D communications with relay and NC was discussed in [21]–[24]. However, in [21], Bai et al. considered a D2D overlaid network scenario with multiple BSs. In [22], Wei et al. discussed the energy efficiency and spectrum efficiency of multihop D2D networks with PNC. In [23] and [24], the authors focused on the time-invariant channel. Furthermore, in [21]–[24], the authors only considered how to analyze or optimize the performance of the D2D networks when the signal to interference plus noise ratio (SINR) threshold is given. They neglected the fact that if the SINR threshold for each transmission link can be properly selected, the overall network performance can be further improved.

In this paper, we maximize the sum-rate of D2D pairs for the D2D communications underlaying cellular networks while still retaining the rate requirements of active CUs. The inactive CUs are used as opportunistic relays under three operational modes: without using network-coding (NNC), using traditional high-layer network-coding (HNC), and using PNC. We adopt the Rayleigh fading channel model where the powers of the signals and interferences at the receiving nodes are exponentially distributed [25], and consider the joint optimization of the SINR threshold for each transmission link, the cellular resource allocation, and opportunistic relay selection. The proposed solution in this paper is fast since we theoretically derive the optimal SINR thresholds, adopt the Hungarian (Kuhn-Munkres) algorithm for the bipartite-matching problem, and propose a low-complexity algorithm for relay selection. That is, it can be used in the scenarios with user mobility, i.e. internet of vehicles (IOV) in urban environments.

The main contributions of this paper are summarized as follows:

1) We show that, given the selections of relays, this sum-rate maximization under Rayleigh fading in no-relay scheme, NNC, HNC, and PNC opportunistic relay schemes can be formulated as a mixed integer non-linear programming (MINLP). We propose a transmit power adjusting method to make sure that in PNC scheme, the relay node can extract an NC packet from the superposition of two received signals in time-varying channel.
2) We propose a two-step approach to obtain the solution to the formulated MINLP by first deriving the optimal SINR thresholds to maximize the transmission rates under different transmission schemes for each possible pairing of D2D pair and CU. Based on the maximum transmission rates of D2D pairs for each possible pairing in the first step, a bipartite-matching method is further proposed to optimize the CU-D2D pairing.
3) According to the solution to the MINLP, we develop an iterative relay selection algorithm to find out the relays that can further improve the sum-rate of D2D communications. We show that compared with the no-relay scheme, the NNC, HNC, and PNC opportunistic relay schemes achieve a maximum performance enhancement of 106%, 138%, and 168%, respectively. Furthermore, when three percent of time is used for the transmit power adjustment, the performance gains of the PNC opportunistic relay scheme versus the no-relay scheme, the NNC and HNC opportunistic relay schemes reach 241%, 150%, and 126%, respectively.

The rest of this paper is organized as follows. Section II presents the system model. In Section III, we formulate the sum-rate maximization problem of D2D pairs in different schemes as an MINLP, given the selections of relays. In Section IV, we propose a two-step solution to the formulated MINLP. In Section V, we develop an iterative relay selection algorithm to find out the relays that can further improve the sum-rate of D2D communications. In Section VI, we carry out simulations to evaluate the performance of the proposed solution. Finally, Section VII concludes this paper.

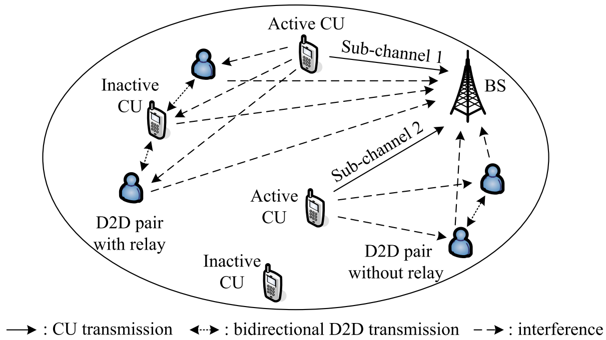


Fig. 2. The considered D2D communications underlaying cellular networks.

II. System description

A. Network model

We consider a cellular network that contains a BS, a set of active CUs, inactive CUs, and two-way D2D pairs randomly located in the coverage area of the BS, as shown in Fig. 2.

The roles of the mobile devices have been defined according to some rules when they join in the network. Considering that the relay nodes do not always exist, two devices in the same cell can be defined as one D2D pair only when they have data transmission requirement and are in the transmission range of each other. The devices that communicate with the devices in another cell or the devices in the same cell but out of the transmission range must be defined as active CUs since they need the help of BS. The idle devices that are willing to serve as relays for the D2D pairs are defined as inactive CUs. We neglect the idle devices that are not willing to serve as relays since they are irrelevant to this work. From the analysis above, we know that the numbers of D2D pairs and active CUs in the network are indeed determined by the communication requirements and locations of the devices. Thus, there exist some scenarios in which the number of D2D pairs is less than the number of active CUs. Usually, these scenarios are called "resource-abundant" scenarios and have been investigated in many existing works, i.e. [9], [10], [26]–[28]. In this paper, we focus on the resource-abundant scenarios where the D2D pairs share the uplink resource of the active CUs. Similar to [9], [10], [26]–[28], to avoid mutual interferences between D2D pairs, reduce the impact of D2D transmissions on the cellular transmissions, and simplify the theoretical analysis, we assume that the uplink resource of each active CU can be shared by at most one D2D pair, and each D2D pair can only share the resource of one active CU. For each D2D pair, it can choose a relay node from the inactive CUs to assist its communication if its average end-to-end transmission rate can be further improved; otherwise, it communicates directly without going through a relay node. When a D2D pair communicates via a relay node, we consider three transmission schemes: NNC scheme, HNC scheme, and PNC scheme. To protect privacy, the data packets of D2D pairs via the relay nodes need to be encrypted. When a relay node receives data signals, it only decodes the signals and corrects the errors caused by wireless transmissions, but can not obtain the content of the data since it does not know the encryption key.

The channel is divided into K sub-channels in frequency domain. Let W_{total} denote the total frequency bandwidth. Then, the bandwidth of each sub-channel, W, equals W_{total}/K. Each active CU is allocated one sub-channel, and the unused sub-channels are reserved for the newly-arrived CUs. Our goal is to maximize the sum-rate of D2D pairs while satisfying the rate requirements of active CUs by properly allocating the resource of active CUs, choosing the SINR thresholds, and the relay nodes.

We adopt the Rayleigh fading channel model with channel gain g following the exponential distribution [25]. Let $h(\cdot)$ and $E[\cdot]$ respectively denote the probability density function (PDF) and the expected value. Then, we have $h(g) = \frac{1}{E[g]}\exp\left(-\frac{g}{E[g]}\right)$. Furthermore, we adopt the block fading model, in which the channel gain does not change during a packet transmission but independently varies in different packet transmissions [29].

B. Transmission rate of wireless devices in LTE standard

In LTE standard, each device has a rate adaption module, which changes the modulation scheme and coding rate according to the channel condition [30]. Each configuration of the modulation scheme and coding rate corresponds to a required minimum SINR for ensuring a given bit error rate, which is called the SINR threshold. That is, the transmission is successful only when the SINR at the receiving device is bigger than the specified SINR threshold; otherwise, transmission failure occurs and the current packet needs to be retransmitted [31]. When the channel condition is good, the transmitting device will use a higher-order modulation scheme and a higher coding rate to achieve higher transmission rate, which corresponds to a higher SINR threshold. In this paper, we use Shannon's capacity formula to approximate the relationship between the instantaneous transmission rate $x_{\max}$ and the specified SINR threshold γ_0. That is, $x_{\max} = W\log(1+\gamma_0)$. Here, the unit of $x_{\max}$ is "nats/s" since natural logarithm is used. Considering that the channel is time-varying, the SINR at the receiving node might be lower than the specified SINR threshold, which leads to transmission failure. Therefore, statistically, the average rate of a transmission x_{avg} equals

$$x_{avg} = x_{\max}Q = QW\log(1+\gamma_0), \tag{1}$$

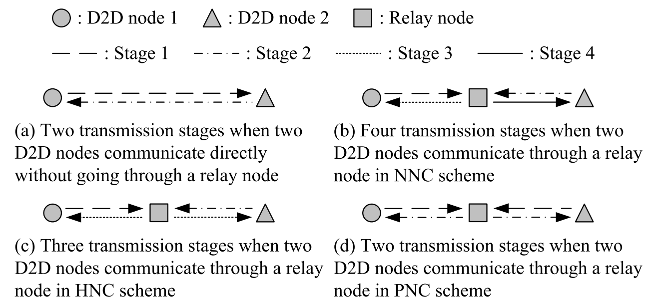
where Q is the successful probability of the considered transmission.

C. Transmission schemes of D2D pairs

To identify the performance gain achieved by the assist of relays and PNC scheme, we consider four kinds of transmission schemes for D2D pairs, namely, no-relay scheme, NNC opportunistic relay scheme, HNC opportunistic relay scheme, and PNC opportunistic relay scheme. In no-relay scheme, the two nodes of each D2D pair communicate directly without going through a relay node; while in NNC, HNC, and PNC opportunistic relay scheme, the two nodes of each D2D pair communicate through a relay node respectively by NNC, HNC, and PNC scheme if the average transmission rate can be improved; otherwise, they communicate directly without going through a relay node. The detailed transmission processes of the above schemes are shown in Fig. 3. From Fig. 3, we can see that when two D2D nodes communicate directly without going through a relay node, two transmission stages are needed; and when two D2D nodes communicate through a relay node in NNC, HNC, and PNC schemes, we need four, three, and two transmission stages, respectively.

Furthermore, we know that in the first transmission stage of PNC scheme, the relay node extracts NC packets from the superimposed EM waves of two D2D nodes. According to [7], to make sure that the extraction process is successful, the two signals from D2D nodes should have similar average power. Since the channel gains vary in different packet transmissions, we need a short transmit power adjusting process at the beginning of each packet transmission, as shown in Fig. 4. First, D2D node 1 and D2D node 2 transmit sequent test bits to the relay node. According to the average powers of the received signals from D2D node 1 and D2D node 2,

P^1_{rcv} and P^2_{rcv}, the relay node broadcasts several bits to the two D2D nodes with the information of the transmit power adjustment. In particular, if $P^1_{rcv} > P^2_{rcv}$, the relay node tells D2D node 1 to decrease its transmit power by a ratio $\left(1-\frac{P^2_{rcv}}{P^1_{rcv}}\right)$; if $P^1_{rcv} < P^2_{rcv}$, the relay node tells D2D node 2 to decrease its transmit power by a ratio $\left(1-\frac{P^1_{rcv}}{P^2_{rcv}}\right)$; if $P^1_{rcv} = P^2_{rcv}$, the relay node tells both the two D2D nodes that the transmit powers do not need to change. After the transmit power adjustment, the average powers of the received signals from D2D node 1 and D2D node 2 at the relay node both equal $\min\left(P^1_{rcv}, P^2_{rcv}\right)$.



(a) Two transmission stages when two D2D nodes communicate directly without going through a relay node

(b) Four transmission stages when two D2D nodes communicate through a relay node in NNC scheme

(c) Three transmission stages when two D2D nodes communicate through a relay node in HNC scheme

(d) Two transmission stages when two D2D nodes communicate through a relay node in PNC scheme

Fig. 3. Transmission processes of D2D pairs.

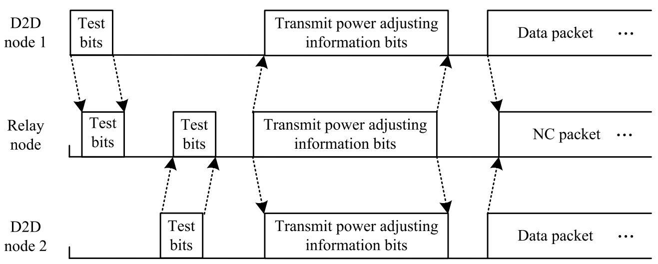


Fig. 4. The transmit power adjusting process in the first transmission stage of the PNC scheme.

Next, we show the implementation details of the considered scenario. In cellular networks, centralized control architecture is adopted, and the BS is used to direct the transmissions of all links via control channel. In particular, the active CUs and D2D nodes first send the BS the transmission requests and the information needed in the optimization process. Then, based on the available relay nodes, the BS starts the optimization process, and broadcasts the optimization results (including the optimal SINR thresholds, the optimal CU-D2D pairing, and the selected relay nodes) to all the devices. According to the optimal SINR thresholds, each device can select proper modulation scheme and coding rate for the transmission [30]. Also, from the optimization results, each device knows when to transmit and receive data packets, and which sub-channel the transmission will use. When the transmission requests or the availability of relay nodes change, the BS restarts the optimization process and renews the resource allocations according to the new optimization results. From the control process above, we know that the overhead is only the exchange of some additional information between the BS and mobile devices.

III. Problem Formulation

Let $\mathcal{C}_{act} = \{c^i_{act}, 1 \le i \le |\mathcal{C}_{act}|\}$, $\mathcal{C}_{nact} = \{c^n_{nact}, 1 \le n \le |\mathcal{C}_{nact}|\}$, and $\mathcal{D} = \{d_j, 1 \le j \le |\mathcal{D}|\}$ respectively denote the set of active CUs, inactive CUs, and D2D pairs in the system. Let d^1_j and d^2_j denote the two nodes of D2D pair d_j. If d_j communicates via a relay node, its relay node is denoted by r_j. We define $y_{i,j} \in \{0,1\}$ as the indicator which indicates whether CU c^i_{act} shares the uplink resource with D2D pair d_j. That is, $y_{i,j} = 1$ when CU c^i_{act} shares the uplink resource with D2D pair d_j; otherwise, $y_{i,j} = 0$. According to the descriptions in Section II-A, we have $\sum\limits_{c^i_{act} \in \mathcal{C}_{act}} y_{i,j} = 1, \forall d_j \in \mathcal{D}$ and $\sum\limits_{d_j \in \mathcal{D}} y_{i,j} \le 1, \forall c^i_{act} \in \mathcal{C}_{act}$. Next we calculate the average transmission rate of active CUs and D2D pairs based on the method described in Section II-B, respectively.

A. Average transmission rate of active CUs

When CU c^i_{act} does not share its resource with any D2D pair, its transmission rate obviously satisfies the rate requirement. Thus, we only need to calculate the transmission rate of CU c^i_{act} when it shares its resource with a D2D pair. According to the transmission processes of D2D pairs in Section II-C, there are two interference cases. The first case is CU c^i_{act} transmits under the interference of another device. The second case is CU c^i_{act} transmits under the interference of two concurrently transmitting D2D nodes. Then, based on the definition and calculation method of the average transmission rates for wireless devices in Section II-B, the average transmission rates of CU c^i_{act} under these two interference cases are derived in Proposition 1 and Proposition 2, respectively.

Proposition 1: Under the interference of device f_1, the average transmission rate from c^i_{act} to BS, $x^{c^i_{act},BS}_{f_1}$, equals

$$x^{c^i_{act},BS}_{f_1} = W\log\left(1+\gamma_0^{c^i_{act},BS}\right)\exp\left(-\frac{\gamma_0^{c^i_{act},BS}N_0W}{P_{tr}^{c^i_{act}}E\left[g^{BS}_{c^i_{act}}\right]}\right) \frac{P_{tr}^{c^i_{act}}E\left[g^{BS}_{c^i_{act}}\right]}{\gamma_0^{c^i_{act},BS}P^{f_1}_{tr}E\left[g^{BS}_{f_1}\right]+P_{tr}^{c^i_{act}}E\left[g^{BS}_{c^i_{act}}\right]}, \tag{2}$$

where $\gamma_0^{u,v}$, g^v_u, P^u_{tr}, and N_0 denote the SINR threshold for the successful transmission from device u to device v, the channel gain from device u to device v, the transmit power of device u, and the power spectral density of additive Gaussian white noise, respectively.

Proof: Let I^v_u and $\gamma^{u,v}$ denote the interference from device u to device v, and the SINR at device v when device u transmits, respectively. Then, we have $\gamma^{c^i_{act},BS} = \frac{P_{tr}^{c^i_{act}}g^{BS}_{c^i_{act}}}{N_0W+I^{BS}_{f_1}}$. Given $I^{BS}_{f_1}$, $\gamma^{c^i_{act},BS}$ follows the exponential distribution with expected value of $\frac{P_{tr}^{c^i_{act}}E\left[g^{BS}_{c^i_{act}}\right]}{N_0W+I^{BS}_{f_1}}$. That is, $h\left(\gamma^{c^i_{act},BS}\right) = \frac{N_0W+I^{BS}_{f_1}}{P_{tr}^{c^i_{act}}E\left[g^{BS}_{c^i_{act}}\right]}\exp\left(-\frac{\gamma^{c^i_{act},BS}\left(N_0W+I^{BS}_{f_1}\right)}{P_{tr}^{c^i_{act}}E\left[g^{BS}_{c^i_{act}}\right]}\right)$. Thus, the successful transmission probability from c^i_{act} to BS when $I^{BS}_{f_1}$

is given, $s_{f_1}^{c_{act}^i,BS}\left(I_{f_1}^{BS}\right)$, equals

$$s_{f_1}^{c_{act}^i,BS}\left(I_{f_1}^{BS}\right)=\int_{\gamma_0^{c_{act}^i,BS}}^{+\infty} h\left(\gamma^{c_{act}^i,BS}\right)d\gamma^{c_{act}^i,BS}$$
$$=\exp\left(-\frac{\gamma_0^{c_{act}^i,BS}(N_0W+I_{f_1}^{BS})}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right).$$

Then, according to equation (1), the average transmission rate from c_{act}^i to BS when $I_{f_1}^{BS}$ is given, $x_{f_1}^{c_{act}^i,BS}\left(I_{f_1}^{BS}\right)$, equals $W\log\left(1+\gamma_0^{c_{act}^i,BS}\right)\exp\left(-\frac{\gamma_0^{c_{act}^i,BS}(N_0W+I_{f_1}^{BS})}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right)$. Since $I_{f_1}^{BS}=P_{tr}^{f_1}g_{f_1}^{BS}$, $I_{f_1}^{BS}$ follows the exponential distribution with expected value $P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]$. That is, $h\left(I_{f_1}^{BS}\right)=\frac{1}{P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]}\exp\left(-\frac{I_{f_1}^{BS}}{P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]}\right)$. Then, $x_{f_1}^{c_{act}^i,BS}$ equals

$$x_{f_1}^{c_{act}^i,BS}=\int_0^{+\infty}x_{f_1}^{c_{act}^i,BS}\left(I_{f_1}^{BS}\right)h\left(I_{f_1}^{BS}\right)dI_{f_1}^{BS}$$
$$=W\log\left(1+\gamma_0^{c_{act}^i,BS}\right)\exp\left(-\frac{\gamma_0^{c_{act}^i,BS}N_0W}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right)$$
$$\frac{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}{\gamma_0^{c_{act}^i,BS}P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}.$$

■

Proposition 2: Under the interference of two concurrently transmitting D2D nodes d_j^1 and d_j^2, the average transmission rate from c_{act}^i to BS, $x_{d_j^1,d_j^2}^{c_{act}^i,BS}$, equals

$$x_{d_j^1,d_j^2}^{c_{act}^i,BS}=\frac{W\log\left(1+\gamma_0^{c_{act}^i,BS}\right)P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}{\left(\gamma_0^{c_{act}^i,BS}P_{tr}^{d_j^1}E\left[g_{d_j^1}^{BS}\right]+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]\right)}\times$$
$$\frac{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]exp\left(-\frac{\gamma_0^{c_{act}^i,BS}N_0W}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right)}{\left(\gamma_0^{c_{act}^i,BS}P_{tr}^{d_j^2}E\left[g_{d_j^2}^{BS}\right]+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]\right)}. \quad (3)$$

Proof: The successful transmission probability from c_{act}^i to BS when $I_{d_j^1}^{BS}$ and $I_{d_j^2}^{BS}$ are given, $s_{d_j^1,d_j^2}^{c_{act}^i,BS}\left(I_{d_j^1}^{BS},I_{d_j^2}^{BS}\right)$, can be derived by similar methods used in the proof of Proposition 1, which equals

$$s_{d_j^1,d_j^2}^{c_{act}^i,BS}\left(I_{d_j^1}^{BS},I_{d_j^2}^{BS}\right)=\exp\left(-\frac{\gamma_0^{c_{act}^i,BS}\left(N_0W+I_{d_j^1}^{BS}+I_{d_j^2}^{BS}\right)}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right).$$

According to equation (1), the average transmission rate from c_{act}^i to BS when $I_{d_j^1}^{BS}$ and $I_{d_j^2}^{BS}$ are given, $x_{d_j^1,d_j^2}^{c_{act}^i,BS}\left(I_{d_j^1}^{BS},I_{d_j^2}^{BS}\right)$, equals $W\log\left(1+\gamma_0^{c_{act}^i,BS}\right)\exp\left(-\frac{\gamma_0^{c_{act}^i,BS}\left(N_0W+I_{d_j^1}^{BS}+I_{d_j^2}^{BS}\right)}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right)$. Considering that $I_{d_j^1}^{BS}$ and $I_{d_j^2}^{BS}$ are independent and respectively follow the exponential distribution with expected values $P_{tr}^{d_j^1}E\left[g_{d_j^1}^{BS}\right]$ and $P_{tr}^{d_j^2}E\left[g_{d_j^2}^{BS}\right]$, $x_{d_j^1,d_j^2}^{c_{act}^i,BS}$ equals

$$x_{d_j^1,d_j^2}^{c_{act}^i,BS}=\int_0^{+\infty}\int_0^{+\infty}x_{d_j^1,d_j^2}^{c_{act}^i,BS}\left(I_{d_j^1}^{BS},I_{d_j^2}^{BS}\right)h\left(I_{d_j^1}^{BS}\right)$$
$$h\left(I_{d_j^2}^{BS}\right)dI_{d_j^1}^{BS}dI_{d_j^2}^{BS}=\frac{W\log\left(1+\gamma_0^{c_{act}^i,BS}\right)P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}{\left(\gamma_0^{c_{act}^i,BS}P_{tr}^{d_j^1}E\left[g_{d_j^1}^{BS}\right]+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]\right)}$$
$$\times\frac{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]exp\left(-\frac{\gamma_0^{c_{act}^i,BS}N_0W}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right)}{\left(\gamma_0^{c_{act}^i,BS}P_{tr}^{d_j^2}E\left[g_{d_j^2}^{BS}\right]+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]\right)}.$$

■

Based on the Proposition 1 and 2, the transmission rate from CU c_{act}^i to BS under the interference of D2D pair d_j, $x_{d_j}^{c_{act}^i,BS}$, can be expressed as follows. When d_j^1 and d_j^2 communicate directly without going through a relay node, $x_{d_j}^{c_{act}^i,BS}$ equals

$$x_{d_j}^{c_{act}^i,BS}=\min\left(x_{d_j^1}^{c_{act}^i,BS},x_{d_j^2}^{c_{act}^i,BS}\right). \quad (4)$$

When d_j^1 and d_j^2 communicate via relay node r_j in NNC and NC schemes, $x_{d_j}^{c_{act}^i,BS}$ equals

$$x_{d_j}^{c_{act}^i,BS}=\min\left(x_{d_j^1}^{c_{act}^i,BS},x_{d_j^2}^{c_{act}^i,BS},x_{r_j}^{c_{act}^i,BS}\right). \quad (5)$$

When d_j^1 and d_j^2 communicate via relay node r_j in PNC scheme, $x_{d_j}^{c_{act}^i,BS}$ equals

$$x_{d_j}^{c_{act}^i,BS}=\min\left(x_{d_j^1}^{c_{act}^i,BS},x_{d_j^2}^{c_{act}^i,BS},x_{r_j}^{c_{act}^i,BS},x_{d_j^1,d_j^2}^{c_{act}^i,BS}\right). \quad (6)$$

According to equation (2), we know that $x_{f_1}^{c_{act}^i,BS}$ decreases as the average interference power $P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]$ increases. Let d_j^u denote the node in the transmission of D2D pair d_j that has the highest average interference power to device u. That is, if d_j^1 and d_j^2 communicate directly without going through a relay node, d_j^u is selected from the set $\left\{d_j^1,d_j^2\right\}$; if d_j^1 and d_j^2 communicate via relay node r_j, d_j^u is selected from the set $\left\{d_j^1,d_j^2,r_j\right\}$. Then, equations (4) and (5) can be both simply expressed as

$$x_{d_j}^{c_{act}^i,BS}=x_{d_j^{BS}}^{c_{act}^i,BS}. \quad (7)$$

And, equation (6) can be simply expressed as

$$x_{d_j}^{c_{act}^i,BS}=\min\left(x_{d_j^{BS}}^{c_{act}^i,BS},x_{d_j^1,d_j^2}^{c_{act}^i,BS}\right). \quad (8)$$

B. Average transmission rate of D2D pairs

According to the descriptions in Section II-C, the transmissions of D2D pairs in different stages can be classified into three types. In the first type, a node transmits packets to another node. In the second type, the relay node broadcasts NC packets to two D2D nodes. In the third type, two D2D nodes transmit packets concurrently to the relay node. In the following, we first derive the transmission rates of the three types of transmission of D2D pairs respectively in Proposition 3-5. The average rates of D2D pairs with and without relay nodes can then be calculated accordingly.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2947021, IEEE Transactions on Vehicular Technology

Proposition 3: In the first type of transmission of D2D pairs, the transmission rate from node f_1 to node f_2 under the interference of CU c_{act}^i, $x_{c_{act}^i}^{f_1,f_2}$, equals

$$x_{c_{act}^i}^{f_1,f_2} = W\log\left(1+\gamma_0^{f_1,f_2}\right)\exp\left(-\frac{\gamma_0^{f_1,f_2}N_0W}{P_{tr}^{f_1}E\left[g_{f_1}^{f_2}\right]}\right)\times \frac{P_{tr}^{f_1}E\left[g_{f_1}^{f_2}\right]}{\gamma_0^{f_1,f_2}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{f_2}\right]+P_{tr}^{f_1}E\left[g_{f_1}^{f_2}\right]}. \tag{9}$$

The proof of Proposition 3 is similar to the proof of Proposition 1 and thus is omitted here.

Proposition 4: In the second type of transmission of D2D pairs, the transmission rate from the relay node r_j to two D2D nodes d_j^1 and d_j^2 under the interference of CU c_{act}^i, $x_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}}$, equals

$$x_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}} = \frac{W\log\left(1+\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}\right)P_{tr}^{r_j}E\left[g_{r_j}^{d_j^1}\right]}{\left(\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{d_j^1}\right]+P_{tr}^{r_j}E\left[g_{r_j}^{d_j^1}\right]\right)}\times \frac{P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]exp\left(-\frac{\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}N_0W}{P_{tr}^{r_j}}\left(\frac{1}{E\left[g_{r_j}^{d_j^1}\right]}+\frac{1}{E\left[g_{r_j}^{d_j^2}\right]}\right)\right)}{\left(\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{d_j^2}\right]+P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]\right)}, \tag{10}$$

where $\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}$ is the SINR threshold for the successful transmission from r_j to d_j^1 and d_j^2.

Proof: Under the interference of CU c_{act}^i, the SINR at d_j^1 and d_j^2 when r_j transmits, $\gamma_1^{r_j,\left\{d_j^1,d_j^2\right\}}$ and $\gamma_2^{r_j,\left\{d_j^1,d_j^2\right\}}$, equal $\frac{P_{tr}^{r_j}g_{r_j}^{d_j^1}}{N_0W+I_{c_{act}^i}^{d_j^1}}$ and $\frac{P_{tr}^{r_j}g_{r_j}^{d_j^2}}{N_0W+I_{c_{act}^i}^{d_j^2}}$, respectively. Given $I_{c_{act}^i}^{d_j^1}$ and $I_{c_{act}^i}^{d_j^2}$, the PDFs of $\gamma_1^{r_j,\left\{d_j^1,d_j^2\right\}}$ and $\gamma_2^{r_j,\left\{d_j^1,d_j^2\right\}}$ respectively equal $\frac{N_0W+I_{c_{act}^i}^{d_j^1}}{P_{tr}^{r_j}E\left[g_{r_j}^{d_j^1}\right]}\exp\left(-\frac{\gamma_1^{r_j,\left\{d_j^1,d_j^2\right\}}\left(N_0W+I_{c_{act}^i}^{d_j^1}\right)}{P_{tr}^{r_j}E\left[g_{r_j}^{d_j^1}\right]}\right)$ and $\frac{N_0W+I_{c_{act}^i}^{d_j^2}}{P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]}\exp\left(-\frac{\gamma_2^{r_j,\left\{d_j^1,d_j^2\right\}}\left(N_0W+I_{c_{act}^i}^{d_j^2}\right)}{P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]}\right)$.

Since the transmission from r_j to d_j^1 and d_j^2 is successful only when both d_j^1 and d_j^2 can successfully receive the packets, the successful transmission probability from r_j to d_j^1 and d_j^2 when $I_{c_{act}^i}^{d_j^1}$ and $I_{c_{act}^i}^{d_j^2}$ are given, $s_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}}\left(I_{c_{act}^i}^{d_j^1},I_{c_{act}^i}^{d_j^2}\right)$, can be calculated as follows.

$$\begin{aligned} s_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}}\left(I_{c_{act}^i}^{d_j^1},I_{c_{act}^i}^{d_j^2}\right) &= \Pr\left(\gamma_1^{r_j,\left\{d_j^1,d_j^2\right\}}\geq\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}\right) \\ &\times\Pr\left(\gamma_2^{r_j,\left\{d_j^1,d_j^2\right\}}\geq\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}\right) \\ &=\exp\left(-\frac{\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}}{P_{tr}^{r_j}}\left(\frac{\left(N_0W+I_{c_{act}^i}^{d_j^1}\right)}{E\left[g_{r_j}^{d_j^1}\right]}+\frac{\left(N_0W+I_{c_{act}^i}^{d_j^2}\right)}{E\left[g_{r_j}^{d_j^2}\right]}\right)\right). \end{aligned}$$

Then, given $I_{c_{act}^i}^{d_j^1}$ and $I_{c_{act}^i}^{d_j^2}$, the average transmission rate from r_j to d_j^1 and d_j^2, $x_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}}\left(I_{c_{act}^i}^{d_j^1},I_{c_{act}^i}^{d_j^2}\right)$, equals $W\log\left(1+\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}\right)\exp\left(-\frac{\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}}{P_{tr}^{r_j}}\left(\frac{\left(N_0W+I_{c_{act}^i}^{d_j^1}\right)}{E\left[g_{r_j}^{d_j^1}\right]}+\frac{\left(N_0W+I_{c_{act}^i}^{d_j^2}\right)}{E\left[g_{r_j}^{d_j^2}\right]}\right)\right)$. Thus, $x_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}}$ equals

$$\begin{aligned} x_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}} &= \int_0^{+\infty}\int_0^{+\infty}x_{c_{act}^i}^{r_j,\left\{d_j^1,d_j^2\right\}}\left(I_{c_{act}^i}^{d_j^1},I_{c_{act}^i}^{d_j^2}\right)\times \\ &h\left(I_{c_{act}^i}^{d_j^1}\right)h\left(I_{c_{act}^i}^{d_j^2}\right)dI_{c_{act}^i}^{d_j^1}dI_{c_{act}^i}^{d_j^2} \\ &= \frac{W\log\left(1+\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}\right)P_{tr}^{r_j}E\left[g_{r_j}^{d_j^1}\right]}{\left(\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{d_j^1}\right]+P_{tr}^{r_j}E\left[g_{r_j}^{d_j^1}\right]\right)}\times \\ &\frac{P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]exp\left(-\frac{\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}N_0W}{P_{tr}^{r_j}}\left(\frac{1}{E\left[g_{r_j}^{d_j^1}\right]}+\frac{1}{E\left[g_{r_j}^{d_j^2}\right]}\right)\right)}{\left(\gamma_0^{r_j,\left\{d_j^1,d_j^2\right\}}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{d_j^2}\right]+P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]\right)}. \end{aligned}$$

■

Proposition 5: In the third type of transmission of D2D pairs, the transmission rate from two concurrently transmitting D2D nodes (d_j^1 and d_j^2) to the relay node r_j under the interference of CU c_{act}^i, $x_{c_{act}^i}^{\left\{d_j^1,d_j^2\right\},r_j}$, equals

$$x_{c_{act}^i}^{\left\{d_j^1,d_j^2\right\},r_j} = W\left(1-\beta\right)\log\left(1+\gamma_0^{\left\{d_j^1,d_j^2\right\},r_j}\right)\times \frac{exp\left(-\gamma_0^{\left\{d_j^1,d_j^2\right\},r_j}N_0W\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)\right)}{\gamma_0^{\left\{d_j^1,d_j^2\right\},r_j}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{r_j}\right]\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)+1}, \tag{11}$$

where β is the proportion of time that is used for the transmit power adjustment, and $\gamma_0^{\left\{d_j^1,d_j^2\right\},r_j}$ is the SINR threshold for the successful transmission from d_j^1 and d_j^2 to r_j.

Proof: From Section II-C, we know that after the transmit power adjustment, the average powers of the received signals from the two D2D nodes both equal $\min\left(P_{tr}^{d_j^1}g_{d_j^1}^{r_j},P_{tr}^{d_j^2}g_{d_j^2}^{r_j}\right)$. Let l_1 and l_2 denote the two D2D signals arriving at the relay node. Let b_1 and b_2 denote the baseband signal corresponding to l_1 and l_2, respectively. According to [7], the distance between two adjacent values of the superimposed baseband signal "b_1+b_2" is the same as that between two adjacent values of either b_1 or b_2. Thus, obtaining a packet from the superimposed signal has the same difficulty as obtaining a packet from l_1 or l_2. Therefore, in this case, the SINR at the relay node, $\gamma^{\left\{d_j^1,d_j^2\right\},r_j}$, equals $\frac{\min\left(P_{tr}^{d_j^1}g_{d_j^1}^{r_j},P_{tr}^{d_j^2}g_{d_j^2}^{r_j}\right)}{N_0W+I_{c_{act}^i}^{r_j}}$. Given $I_{c_{act}^i}^{r_j}$, the successful transmission probability in this case, $s_{c_{act}^i}^{\left\{d_j^1,d_j^2\right\},r_j}\left(I_{c_{act}^i}^{r_j}\right)$, can be calculated as follows.

$$
\begin{aligned}
&s_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}\left(I_{c_{act}^i}^{r_j}\right)=\Pr\left(\gamma^{\{d_j^1,d_j^2\},r_j}>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\\
&=\Pr\left(\frac{\min\left(P_{tr}^{d_j^1}g_{d_j^1}^{r_j},P_{tr}^{d_j^2}g_{d_j^2}^{r_j}\right)}{N_0W+I_{c_{act}^i}^{r_j}}>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\\
&=\Pr\left(\min\left(\frac{P_{tr}^{d_j^1}g_{d_j^1}^{r_j}}{N_0W+I_{c_{act}^i}^{r_j}},\frac{P_{tr}^{d_j^2}g_{d_j^2}^{r_j}}{N_0W+I_{c_{act}^i}^{r_j}}\right)>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\\
&\qquad=\Pr\left(\frac{P_{tr}^{d_j^1}g_{d_j^1}^{r_j}}{N_0W+I_{c_{act}^i}^{r_j}}>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\times\\
&\qquad\Pr\left(\frac{P_{tr}^{d_j^2}g_{d_j^2}^{r_j}}{N_0W+I_{c_{act}^i}^{r_j}}>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right).
\end{aligned}
$$

Let z_1 and z_2 denote $\frac{P_{tr}^{d_j^1}g_{d_j^1}^{r_j}}{N_0W+I_{c_{act}^i}^{r_j}}$ and $\frac{P_{tr}^{d_j^2}g_{d_j^2}^{r_j}}{N_0W+I_{c_{act}^i}^{r_j}}$, respectively. Given $I_{c_{act}^i}^{r_j}$, z_1 and z_2 follow the exponential distribution respectively with expected values of $\frac{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}{N_0W+I_{c_{act}^i}^{r_j}}$ and $\frac{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}{N_0W+I_{c_{act}^i}^{r_j}}$. Thus, $s_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}\left(I_{c_{act}^i}^{r_j}\right)$ can be expressed as

$$
\begin{aligned}
&s_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}\left(I_{c_{act}^i}^{r_j}\right)=\Pr\left(z_1>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\times\\
&\Pr\left(z_2>\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\\
&=\exp\left(-\gamma_0^{\{d_j^1,d_j^2\},r_j}\left(N_0W+I_{c_{act}^i}^{r_j}\right)\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\right.\right.\\
&\left.\left.\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)\right).
\end{aligned}
$$

Then, according to equation (1), the average transmission rate from d_j^1 and d_j^2 to r_j when $I_{c_{act}^i}^{r_j}$ is given, $x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}\left(I_{c_{act}^i}^{r_j}\right)$, equals $W\log\left(1+\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\times\exp\left(-\gamma_0^{\{d_j^1,d_j^2\},r_j}\left(N_0W+I_{c_{act}^i}^{r_j}\right)\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)\right)$. Thus, $x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}$ equals

$$
\begin{aligned}
&x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}=\int_0^{+\infty}x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}\left(I_{c_{act}^i}^{r_j}\right)h\left(I_{c_{act}^i}^{r_j}\right)dI_{c_{act}^i}^{r_j}\\
&=W\log\left(1+\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\times\\
&\frac{exp\left(-\gamma_0^{\{d_j^1,d_j^2\},r_j}N_0W\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)\right)}{\gamma_0^{\{d_j^1,d_j^2\},r_j}P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{r_j}\right]\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)+1}.
\end{aligned}
$$

Considering that the proportion of time that is used for data transmission equals $(1-\beta)$, equation (11) can be obtained. ■

The average transmission rate of a D2D pair is defined as the average number of nats exchanged between the two D2D nodes per second when the two D2D nodes transmit the same number of packets to each other. Let $x_{c_{act}^i}^{d_j}$ denote the average transmission rate of d_j under the interference of CU c_{act}^i. Then, based on Proposition 3-5, we can calculate $x_{c_{act}^i}^{d_j}$ as follows. In the case that the relay node is not used, $x_{c_{act}^i}^{d_j}$ equals

$$x_{c_{act}^i}^{d_j}=\frac{2}{\frac{1}{x_{c_{act}^i}^{d_j^1,d_j^2}}+\frac{1}{x_{c_{act}^i}^{d_j^2,d_j^1}}}. \tag{12}$$

In the case that the relay node is used, $x_{c_{act}^i}^{d_j}$ in NNC, NC, and PNC schemes respectively equals

$$x_{c_{act}^i}^{d_j}=\frac{2}{\frac{1}{x_{c_{act}^i}^{d_j^1,r_j}}+\frac{1}{x_{c_{act}^i}^{d_j^2,r_j}}+\frac{1}{x_{c_{act}^i}^{r_j,d_j^2}}+\frac{1}{x_{c_{act}^i}^{r_j,d_j^1}}}, \tag{13}$$

$$x_{c_{act}^i}^{d_j}=\frac{2}{\frac{1}{x_{c_{act}^i}^{d_j^1,r_j}}+\frac{1}{x_{c_{act}^i}^{d_j^2,r_j}}+\frac{1}{x_{c_{act}^i}^{r_j,\{d_j^1,d_j^2\}}}}, \tag{14}$$

$$x_{c_{act}^i}^{d_j}=\frac{2}{\frac{1}{x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}}+\frac{1}{x_{c_{act}^i}^{r_j,\{d_j^1,d_j^2\}}}}. \tag{15}$$

The sum-rate of all the D2D pairs in the system, $R_{sum}^{\mathcal{D}}$, can then be expressed as

$$R_{sum}^{\mathcal{D}}=\sum_{d_j\in\mathcal{D}}\sum_{c_{act}^i\in\mathcal{C}_{act}}y_{i,j}x_{c_{act}^i}^{d_j}. \tag{16}$$

Thus, given the selections of relays, the maximization problem of the sum-rate of all the D2D pairs can be formulated as follows.

$$
\begin{aligned}
&\max\ R_{sum}^{\mathcal{D}},\\
&\text{s.t.}\ \sum_{c_{act}^i\in\mathcal{C}_{act}}y_{i,j}=1,\forall d_j\in\mathcal{D}, &\text{(17a)}\\
&\qquad\sum_{d_j\in\mathcal{D}}y_{i,j}\le 1,\forall c_{act}^i\in\mathcal{C}_{act}, &\text{(17b)}\\
&\qquad x_{d_j}^{c_{act}^i,BS}>X_{\min}^{c_{act}^i}\ \text{if}\ y_{i,j}=1,\forall c_{act}^i\in\mathcal{C}_{act},\forall d_j\in\mathcal{D}. &\text{(17c)}
\end{aligned}
$$

Here, $X_{\min}^{c_{act}^i}$ is the transmission rate requirement of CU c_{act}^i. Constraint (17a) demonstrates that each D2D pair only shares the resource of one active CU. Constraint (17b) accounts the fact that each active CU at most shares its resource with one D2D pair. Constraint (17c) ensures that the transmission rates of CUs satisfy the rate requirements. Problem (17) is an MINLP, which is difficult to solve in general. In the next section, we discuss how Problem (17) is solved.

IV. Joint Resource Allocation and SINR Optimization

Obviously, Problem (17) can be treated as a maximum-weight bipartite-matching problem as follows.

$$\max \sum_{c_{act}^i \in \mathcal{C}_{act}} \sum_{d_j \in \mathcal{D}} y_{i,j} Q_{i,j},$$
$$\text{s.t.} \sum_{c_{act}^i \in \mathcal{C}_{act}} y_{i,j} = 1, \forall d_j \in \mathcal{D}, \tag{18a}$$
$$\sum_{d_j \in \mathcal{D}} y_{i,j} \leq 1, \forall c_{act}^i \in \mathcal{C}_{act}. \tag{18b}$$

Here, $Q_{i,j}$ is the maximum transmission rate of d_j when it shares the resource of c_{act}^i under the given relay selection (including the case that no relay node is selected), which can be obtained as follows. We first optimize the SINR thresholds to maximize $x_{c_{act}^i}^{d_j}$ and $x_{d_j}^{c_{act}^i,BS}$ according to the method given in Section IV-A. If the maximum value of $x_{d_j}^{c_{act}^i,BS}$ is larger than the transmission rate requirement of CU c_{act}^i, then, $Q_{i,j}$ equals the maximum value of $x_{c_{act}^i}^{d_j}$; otherwise, d_j with the current relay selection cannot share the resource of c_{act}^i and thus $Q_{i,j}$ is set to "$-\infty$". After the calculation of $Q_{i,j}$, we can use the classical Hungarian (Kuhn-Munkres) algorithm to obtain the optimal resource sharing between D2D pairs and CUs that maximizes the sum-rate of D2D pairs [32], [33]. Next, we derive the optimal SINR thresholds to maximize $x_{c_{act}^i}^{d_j}$ and $x_{d_j}^{c_{act}^i,BS}$, respectively.

A. SINR thresholds optimization

From equations (7), (8), (12), (13), (14), and (15), we know that in all cases, if the relay selections of D2D pairs are given, $x_{c_{act}^i}^{d_j}$ and $x_{d_j}^{c_{act}^i,BS}$ are maximized when $x_{c_{act}^i}^{f_1,f_2}$, $x_{c_{act}^i}^{r_j,\{d_j^1,d_j^2\}}$, $x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}$, $x_{f_1}^{c_{act}^i,BS}$, and $x_{d_j^1,d_j^2}^{c_{act}^i,BS}$ are maximized. In the following, we first derive the optimal SINR thresholds that respectively maximize $x_{c_{act}^i}^{f_1,f_2}$, $x_{c_{act}^i}^{r_j,\{d_j^1,d_j^2\}}$, $x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}$, $x_{f_1}^{c_{act}^i,BS}$, and $x_{d_j^1,d_j^2}^{c_{act}^i,BS}$ in Theorem 1-5, when the average interference power at the receiving nodes is smaller than the given values. In the case that the average interference power at the receiving nodes is higher than the given values, the optimal SINR thresholds can be obtained by numerical search.

Theorem 1: If $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right] < P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]$, $x_{c_{act}^i}^{f_1,f_2}$ is maximized when $\gamma_0^{f_1,f_2}$ satisfies

$$1 - \frac{N_0 W\left(1+\gamma_0^{f_1,f_2}\right)\log\left(1+\gamma_0^{f_1,f_2}\right)}{P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]} - \frac{\left(1+\gamma_0^{f_1,f_2}\right)\log\left(1+\gamma_0^{f_1,f_2}\right)P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]} = 0. \tag{19}$$

Proof: According to equation (9), the first derivative of $x_{c_{act}^i}^{f_1,f_2}$ equals

$$\frac{\mathrm{d}x_{c_{act}^i}^{f_1,f_2}}{\mathrm{d}\gamma_0^{f_1,f_2}} = \frac{W P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]\exp\left(-\frac{\gamma_0^{f_1,f_2} N_0 W}{P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]}\right)}{\left(1+\gamma_0^{f_1,f_2}\right)\left(P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]\right)} \times \left(1 - \frac{N_0 W\left(1+\gamma_0^{f_1,f_2}\right)\log\left(1+\gamma_0^{f_1,f_2}\right)}{P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]} - \frac{\left(1+\gamma_0^{f_1,f_2}\right)\log\left(1+\gamma_0^{f_1,f_2}\right)P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]}\right).$$

Let $Q_1 = 1 - \frac{N_0 W\left(1+\gamma_0^{f_1,f_2}\right)\log\left(1+\gamma_0^{f_1,f_2}\right)}{P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]} - \frac{\left(1+\gamma_0^{f_1,f_2}\right)\log\left(1+\gamma_0^{f_1,f_2}\right)P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]}$. We next show that, if $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right] < P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]$, Q_1 decreases from a positive value to a negative value when $\gamma_0^{f_1,f_2}$ increases from 0. Let $U_1 \triangleq \frac{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\left(1+\gamma_0^{f_1,f_2}\right)}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]}$, which can be expressed as $U_1 = 1 - \frac{P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right] - P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]}$. If $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right] < P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]$, we have $P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right] - P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right] > 0$. Thus, U_1 increases as $\gamma_0^{f_1,f_2}$ increases. Therefore, Q_1 decreases as $\gamma_0^{f_1,f_2}$ increases. Furthermore, we have $\lim_{\gamma_0^{f_1,f_2} \to 0} Q_1 = 1$, and $\lim_{\gamma_0^{f_1,f_2} \to +\infty} Q_1 = -\infty$. Thus, Q_1 decreases from a positive value to a negative value as $\gamma_0^{f_1,f_2}$ increases from 0. Since $\frac{W P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]\exp\left(-\frac{\gamma_0^{f_1,f_2} N_0 W}{P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]}\right)}{\left(1+\gamma_0^{f_1,f_2}\right)\left(P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right]\gamma_0^{f_1,f_2} + P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]\right)} > 0$, $\frac{\mathrm{d}x_{c_{act}^i}^{f_1,f_2}}{\mathrm{d}\gamma_0^{f_1,f_2}}$ decreases from a positive value to a negative value as $\gamma_0^{f_1,f_2}$ increases from 0. That is, if $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{f_2}\right] < P_{tr}^{f_1} E\left[g_{f_1}^{f_2}\right]$, $x_{c_{act}^i}^{f_1,f_2}$ increases first and then decreases as $\gamma_0^{f_1,f_2}$ increases, and it reaches the maximum value when $\gamma_0^{f_1,f_2}$ satisfies equation (19). ■

Theorem 2: If $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{d_j^1}\right] < P_{tr}^{r_j} E\left[g_{r_j}^{d_j^1}\right]$ and $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{d_j^2}\right] < P_{tr}^{r_j} E\left[g_{r_j}^{d_j^2}\right]$, $x_{c_{act}^i}^{r_j,\{d_j^1,d_j^2\}}$ is maximized when $\gamma_0^{r_j,\{d_j^1,d_j^2\}}$ satisfies

$$1 - N_0 W\left(\frac{1}{P_{tr}^{r_j} E\left[g_{r_j}^{d_j^1}\right]} + \frac{1}{P_{tr}^{r_j} E\left[g_{r_j}^{d_j^2}\right]}\right) \times \left(1+\gamma_0^{r_j,\{d_j^1,d_j^2\}}\right)\log\left(1+\gamma_0^{r_j,\{d_j^1,d_j^2\}}\right) - \left(\frac{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{d_j^1}\right]\left(1+\gamma_0^{r_j,\{d_j^1,d_j^2\}}\right)\log\left(1+\gamma_0^{r_j,\{d_j^1,d_j^2\}}\right)}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{d_j^1}\right]\gamma_0^{r_j,\{d_j^1,d_j^2\}} + P_{tr}^{r_j} E\left[g_{r_j}^{d_j^1}\right]}\right.$$

$$+\frac{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{d_j^2}\right]\left(1+\gamma_0^{r_j,\{d_j^1,d_j^2\}}\right)\log\left(1+\gamma_0^{r_j,\{d_j^1,d_j^2\}}\right)}{P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{d_j^2}\right]\gamma_0^{r_j,\{d_j^1,d_j^2\}}+P_{tr}^{r_j}E\left[g_{r_j}^{d_j^2}\right]}\Bigg) = 0.$$

Theorem 3: If $P_{tr}^{c_{act}^i} E\left[g_{c_{act}^i}^{r_j}\right] < \frac{1}{\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)}$, $x_{c_{act}^i}^{\{d_j^1,d_j^2\},r_j}$ is maximized when $\gamma_0^{\{d_j^1,d_j^2\},r_j}$ satisfies

$$1-\left(1+\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\log\left(1+\gamma_0^{\{d_j^1,d_j^2\},r_j}\right)\times\left(N_0W\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)\right.$$
$$\left.+\frac{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{r_j}\right]\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{r_j}\right]\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)\gamma_0^{\{d_j^1,d_j^2\},r_j}+1}\right)=0.$$

Theorem 4: If $P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right] < P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]$, $x_{f_1}^{c_{act}^i,BS}$ is maximized when $\gamma_0^{c_{act}^i,BS}$ satisfies

$$1-\frac{N_0W\left(1+\gamma_0^{c_{act}^i,BS}\right)\log\left(1+\gamma_0^{c_{act}^i,BS}\right)}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}$$
$$-\frac{P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]\left(1+\gamma_0^{c_{act}^i,BS}\right)\log\left(1+\gamma_0^{c_{act}^i,BS}\right)}{P_{tr}^{f_1}E\left[g_{f_1}^{BS}\right]\gamma_0^{c_{act}^i,BS}+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}=0.$$

Theorem 5: If $P_{tr}^{d_j^1}E\left[g_{d_j^1}^{BS}\right] < P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]$ and $P_{tr}^{d_j^2}E\left[g_{d_j^2}^{BS}\right] < P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]$, $x_{d_j^1,d_j^2}^{c_{act}^i,BS}$ is maximized when $\gamma_0^{c_{act}^i,BS}$ satisfies

$$1-\frac{N_0W\left(1+\gamma_0^{c_{act}^i,BS}\right)\log\left(1+\gamma_0^{c_{act}^i,BS}\right)}{P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}-$$
$$\left(\frac{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{BS}\right]\left(1+\gamma_0^{c_{act}^i,BS}\right)\log\left(1+\gamma_0^{c_{act}^i,BS}\right)}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{BS}\right]\gamma_0^{c_{act}^i,BS}+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}+\right.$$
$$\left.\frac{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{BS}\right]\left(1+\gamma_0^{c_{act}^i,BS}\right)\log\left(1+\gamma_0^{c_{act}^i,BS}\right)}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{BS}\right]\gamma_0^{c_{act}^i,BS}+P_{tr}^{c_{act}^i}E\left[g_{c_{act}^i}^{BS}\right]}\right)=0.$$

The proofs of Theorem 2-5 are similar to the proof of Theorem 1 and thus are omitted here. Theorem 1-5 imply that, in all the transmission cases of CUs and D2D pairs, the optimal SINR threshold can be derived through theoretical analysis if the average interference power at the receiving node(s) is smaller than a given value. In Theorem 1, Theorem 2, Theorem 4, and Theorem 5, this given value equals the average signal power at the corresponding receiving node; while in Theorem 3, the given value equals $\frac{1}{\left(\frac{1}{P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right]}+\frac{1}{P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]}\right)}$, which is no less than half of the average power of the weaker signal at the relay node, $\frac{1}{2}\min\left(P_{tr}^{d_j^1}E\left[g_{d_j^1}^{r_j}\right], P_{tr}^{d_j^2}E\left[g_{d_j^2}^{r_j}\right]\right)$. That is, the conditions in all theorems can usually be satisfied in most of scenarios where the mutual interference is not too severe. For the other scenarios where the mutual interference is extremely severe, we can obtain the optimal SINR thresholds by numerical search.

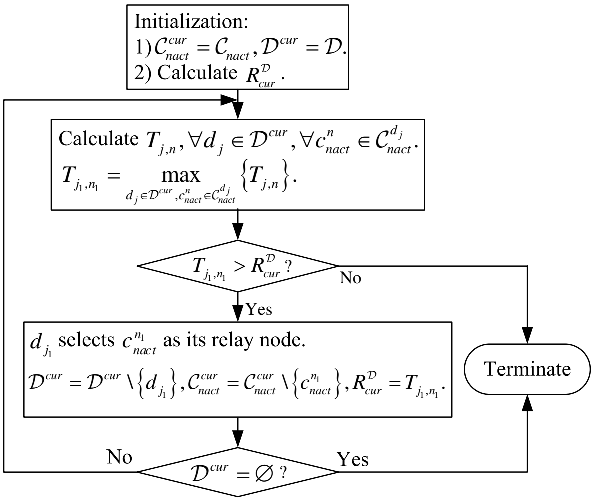


Fig. 5. The iterative relay selection algorithm.

V. The Iterative Relay Selection Algorithm

In this section, we propose an iterative relay selection algorithm to find out the relays that can further improve the sum-rate of D2D communications, based on the solution to Problem (17). Since all the D2D pairs transmit concurrently by using different frequency-domain sub-channels, an inactive CU can only be used as the relay node of one D2D pair. Thus, to avoid the case that two or more D2D pairs choose the same relay node, the relay selection processes of different D2D pairs cannot be treated independently. In the following, an iterative algorithm is proposed for the relay selection of all the D2D pairs to increase their sum-rate, as shown in Fig. 5. At the beginning of the iterative relay selection algorithm, no inactive CU is selected by the D2D pairs. That is, all the inactive CUs are unoccupied. Let $R_{cur}^{\mathcal{D}}$, $\mathcal{C}_{nact}^{cur}$, and $\mathcal{D}^{cur}$ denote the maximum sum-rate of the D2D pairs under the current relay selection, the set of inactive CUs that are unoccupied and the set of D2D pairs that have not selected a relay node at current time, respectively. Then, at the beginning of the iterative relay selection algorithm, $R_{cur}^{\mathcal{D}}$ is the maximum sum-rate of the D2D pairs when no D2D pairs communicates via a relay node, and, $\mathcal{C}_{nact}^{cur}$ and $\mathcal{D}^{cur}$ respectively equal $\mathcal{C}_{nact}$ and $\mathcal{D}$. Consider that adding relay nodes can improve the transmission rate of the D2D pairs only when the relay nodes can increase the transmission SINRs. Thus, in the relay selection of each D2D pair, we can only consider the inactive CUs that are close to the corresponding D2D nodes. In particular, for $d_j \in \mathcal{D}^{cur}$, we define $\mathcal{C}_{nact}^{d_j}$ as the set of inactive CUs in $\mathcal{C}_{nact}^{cur}$ whose distances from d_j^1

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2947021, IEEE Transactions on Vehicular Technology

and d_j^2 are both smaller than the distance between d_j^1 and d_j^2. When selecting a relay node for $d_j \in \mathcal{D}^{cur}$, we only select from $\mathcal{C}_{nact}^{d_j}$. In each round, based on the solution to Problem (17), we calculate $T_{j,n}$, which is the maximum sum-rate of the D2D pairs if $d_j \in \mathcal{D}^{cur}$ further uses $c_{nact}^n \in \mathcal{C}_{nact}^{d_j}$ as its relay node. Let T_{j_1,n_1} denote $\max\limits_{d_j \in \mathcal{D}^{cur}, c_{nact}^n \in \mathcal{C}_{nact}^{d_j}} \{T_{j,n}\}$. Then, if $T_{j_1,n_1} > R_{cur}^{\mathcal{D}}$, $d_{j_1} \in \mathcal{D}^{cur}$ select inactive CU $c_{nact}^{n_1}$ as its relay node, $\mathcal{D}^{cur} = \mathcal{D}^{cur} \backslash \{d_{j_1}\}$, $\mathcal{C}_{nact}^{cur} = \mathcal{C}_{nact}^{cur} \backslash \{c_{nact}^{n_1}\}$, and $R_{cur}^{\mathcal{D}} = T_{j_1,n_1}$. The iterative relay selection algorithm is terminated if $T_{j_1,n_1} \le R_{cur}^{\mathcal{D}}$ or $\mathcal{D}^{cur} = \emptyset$.

Next we discuss the computational complexity of the proposed solution. From the descriptions above, at each round of the iterative relay selection algorithm, we need to solve Problem (17) $\sum\limits_{d_j \in \mathcal{D}} \left|\mathcal{C}_{nact}^{d_j}\right|$ times at most. Since there are at most $|\mathcal{D}|$ rounds, the maximum computational complexity of the proposed solution is to solve Problem (17) $|\mathcal{D}| \cdot \sum\limits_{d_j \in \mathcal{D}} \left|\mathcal{C}_{nact}^{d_j}\right|$ times. According to [34], the computational complexity of the solution for Problem (17) by Hungarian algorithm is $O\left(|\mathcal{C}_{act}|^4\right)$. Therefore, the maximum computational complexity of the proposed solution is $|\mathcal{D}| \cdot \left(\sum\limits_{d_j \in \mathcal{D}} \left|\mathcal{C}_{nact}^{d_j}\right|\right) \cdot O\left(|\mathcal{C}_{act}|^4\right)$.

VI. Performance Evaluation

In this section, we carry out simulations to evaluate the performance of the proposed methods and compare the performance of different schemes. In the simulation, the active and inactive CUs and D2D pairs are uniformly distributed within the circular area centered at the BS. The expected value of channel gain $E[g]$ is obtained from log-distance path-loss model. We cite the system parameters adopted in [35]. In particular, the cellular radius and the path-loss exponent for the calculation of the expected value of channel gain are set to 300 m and 4, respectively. The transmit power of active CUs, the noise power density, the bandwidth of each sub-channel, and the rate requirements of active CUs are set to 23 dBm, -174 dBm/Hz, 20 kHz, and 20 $knats/s$, respectively.

A. Validation of Propositions for transmission rate calculations

To validate the average transmission rate calculations derived in Propositions 1-5, we compare them with simulations in which the real Rayleigh fading channel is adopted when the SINR threshold is varied in the typical range [0.5, 1000]. The average rate in the simulation is obtained from 30000 transmission attempts. We randomly generate five CU-D2D pairs. Each D2D pair is assigned a relay node located near the center of the D2D link. In the validation of Proposition 3, we consider the average transmission rate from d_j^1 to d_j^2. In the validation of Proposition 5, the proportion of time that is used for the transmit power adjusting process, β, is set to 0.1. From Fig. 6, we can see that the closed-form expressions derived in Proposition 1-5 are quite accurate since the analytical results (lines) exactly match the simulations (symbols).

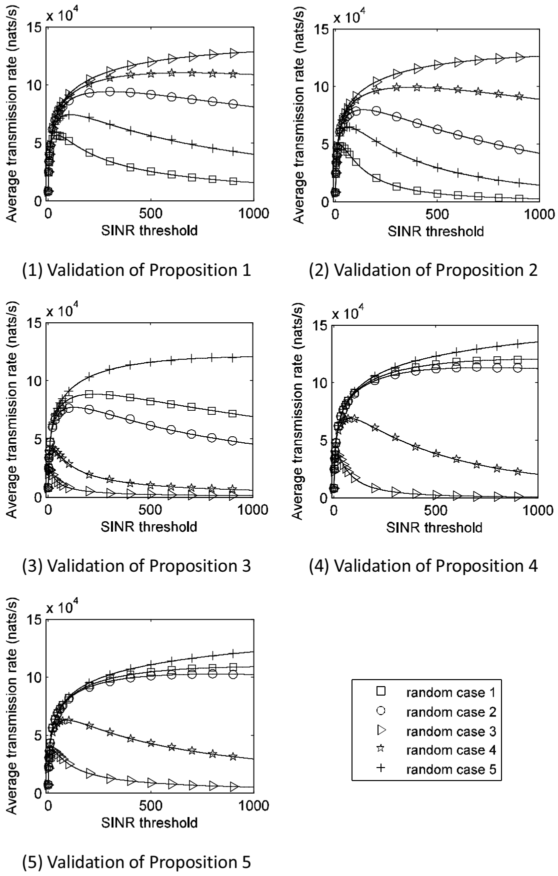


Fig. 6. Validation of Propositions.

B. Performance comparison with other policies

Fig. 7 compares the performance of the proposed approach with the optimal solution via exhaustive search when the distance between two D2D nodes in a D2D pair randomly varies from 20 m to 100 m. Since the running time of the exhaustive search increases exponentially as the numbers of active CUs and D2D pairs increase, we can only consider an extremely small network with two D2D pairs and three active CUs. The transmit power of D2D transmission is set to 10 dBm. The number of inactive CUs is set to 1000. In the PNC opportunistic relay scheme, the proportion of time that is used for the transmit power adjusting process, β, is set to 0.1. In each random case, we regenerate the positions of the CUs and D2D nodes. From Fig. 7, we can see that the performance of the proposed approach is close to the optimal solution. In the next part, we will further compare the performance of the proposed approach with other policies in big networks.

Fig. 8 shows the performance of the derived SINR thresholds, denoted as "TH", versus the optimal SINR thresholds from exhaustive-search method (ES), and the following SINR threshold setting policies respectively in no-relay scheme, NNC, HNC, and PNC opportunistic relay schemes:

1) Average SINR policy (AS). For the transmission from one device to another device, the SINR threshold is set to the ratio of the average signal power to the average interference and noise power at the receiving device. For

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2947021, IEEE Transactions on Vehicular Technology

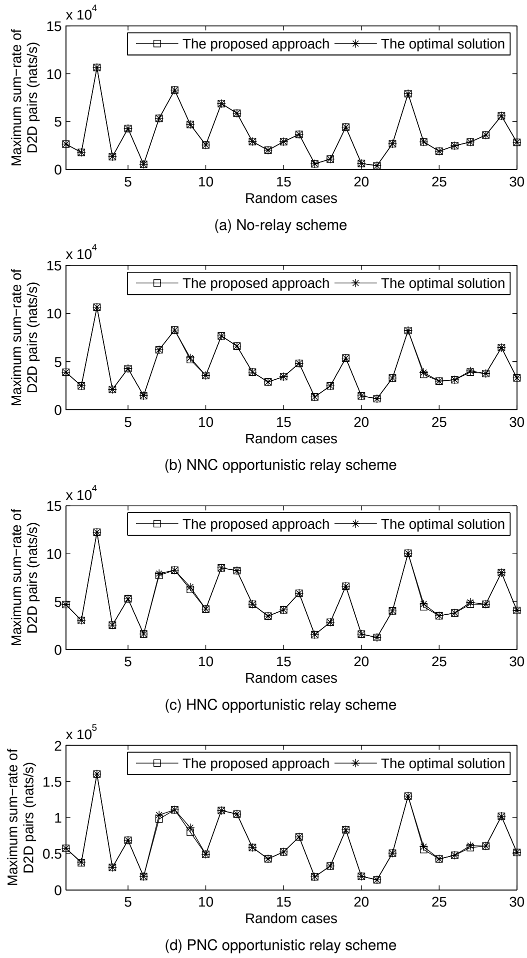


Fig. 7. The performance comparison with the optimal solution.

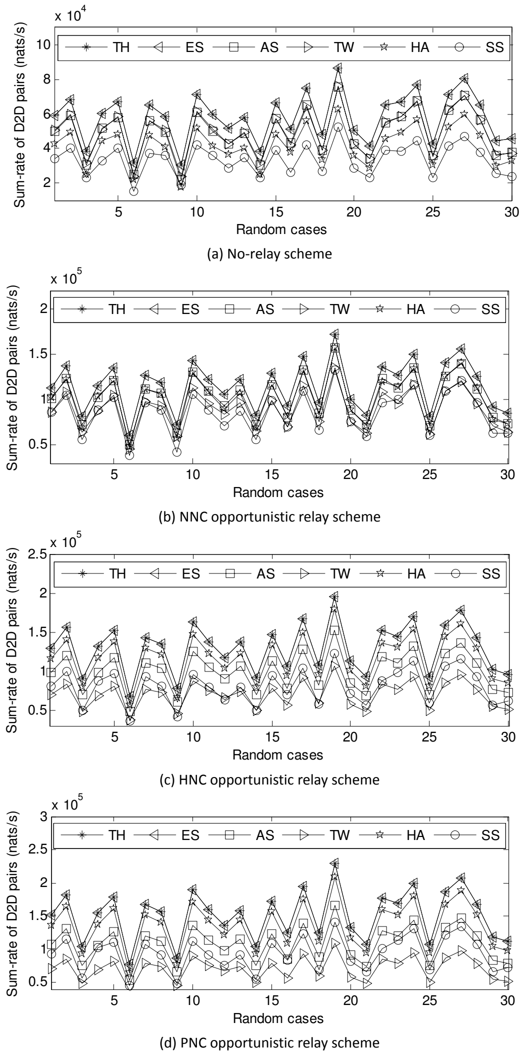


Fig. 8. The performance of the derived SINR threshold.

the transmission from the relay node to two D2D nodes, the SINR threshold is set to $\frac{1}{2}(SINR_1 + SINR_2)$, where $SINR_1$ and $SINR_2$ are the ratio of the average signal power to the average interference and noise power at the two D2D nodes, respectively. For the transmission from two D2D nodes to the relay node, the SINR threshold is set to the ratio of the smaller average power of the two signals from D2D nodes to the average interference and noise power at the relay node.

2) Twice Average SINR policy (TW), in which the SINR thresholds are set to twice of the ones in the AS policy.
3) Half Average SINR policy (HA), in which the SINR thresholds are set to half of the ones in the AS policy.
4) Single SINR policy (SS), in which the SINR thresholds for all the transmissions are the same and optimized by numerical search.

We consider the random cases in which the numbers of active CUs and D2D pairs are randomly selected from the range of [25, 35] and [10, 20], respectively, and each D2D pair is paired with a random active CU. The distance between two D2D nodes in a D2D pair and the transmit power of D2D transmission are set to 80 m and 10 dBm, respectively. Each D2D pair is allocated a given relay. In the PNC opportunistic relay scheme, β is set to 0.1.

From Fig. 8, we can see that: 1) The derived SINR thresholds exactly match the optimal SINR thresholds from exhaustive-search method. 2) Compared with the AS, TW, HA, and SS SINR threshold setting policies, the derived SINR threshold can obtain a performance enhancement of 9%-125%. The reason is that the four compared policies can not adjust the SINR threshold according to the channel condition of each transmission. Another observation is that the performance of the four compared SINR threshold setting policies depends on the adopted scheme. For example, in no-relay scheme, the AS and TW policies outperform the HA and the SS policies; while in HNC and PNC opportunistic relay schemes, the HA policy has better performance than the AS, TW, and SS policies.

Fig. 9 compares the performance of the proposed bipartite-matching method, with the Nearest First (NF), Farthest First (FF), and Random Selection (RS) resource sharing policies, respectively in no-relay scheme, NNC, HNC, and PNC opportunistic relay schemes under the optimal SINR thresholds. In NF, FF, and RS resource sharing policies, for each D2D pair, we respectively select the nearest, the farthest, and a random active CU that satisfies: 1) the active CU has not been selected

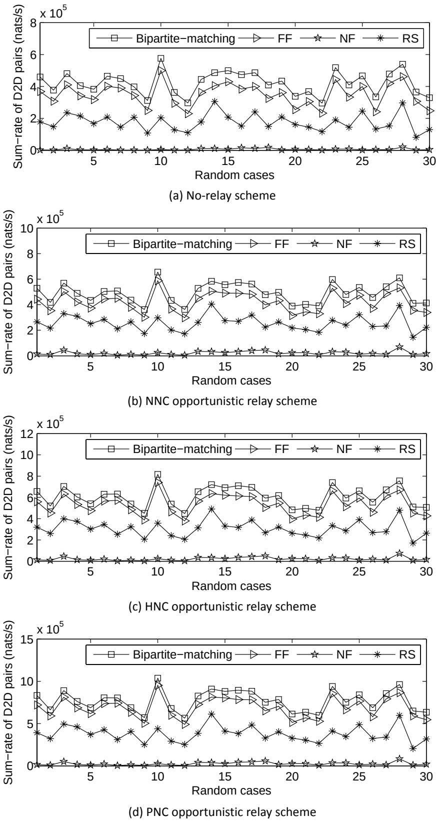


Fig. 9. The performance of the optimal bipartite-matching.

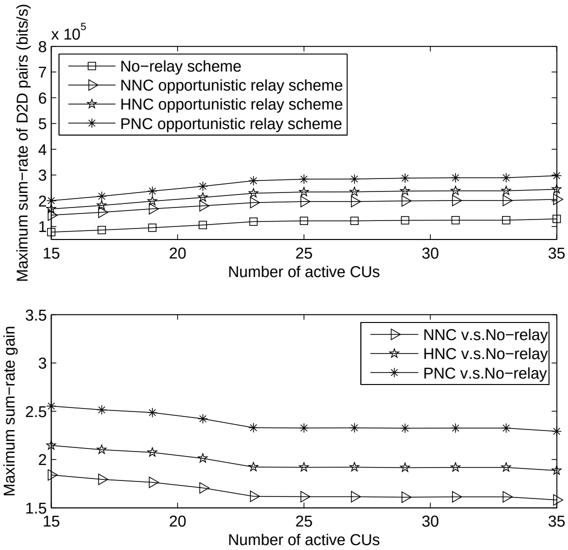


Fig. 10. Performance comparison of the four schemes under different numbers of active CUs.

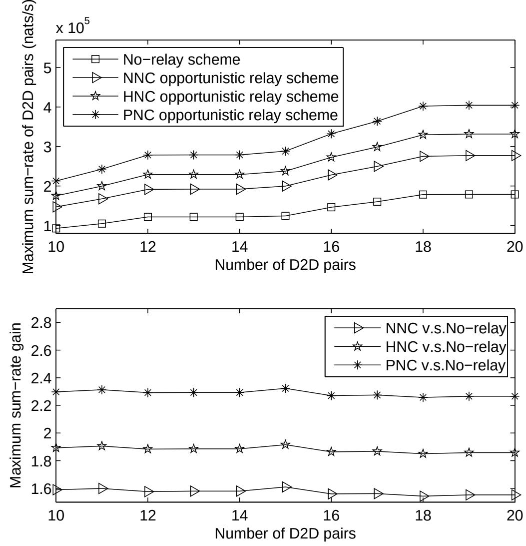


Fig. 11. Performance comparison of the four schemes under different numbers of D2D pairs.

by other D2D pairs; 2) the rate requirement of the active CU can be satisfied. Here, the distance between a CU and a D2D pair is defined as the average distance between the CU and the two D2D nodes. The distance between two D2D nodes in a D2D pair, the transmit power of D2D transmission, and the value of β in PNC opportunistic relay scheme are the same as those used in Fig. 8. Each D2D pair is allocated a given relay. We consider the random cases in which the numbers of active CUs and D2D pairs are randomly selected from the range of [25, 35] and [10, 20], respectively.

From Fig. 9, we can see that in all schemes: 1) The NF resource sharing policy has the worst performance since it has the biggest mutual interference between D2D pairs and their sharing CUs. 2) Compared with the FF and RS policies, the bipartite-matching method can achieve an average sum-rate enhancement of 16% and 115%, respectively.

C. Performance comparison of the considered schemes

Fig. 10 and Fig. 11 show the performance comparison of the four schemes under different numbers of active CUs when the number of D2D pairs is set to 15, and under different numbers of D2D pairs when the number of active CUs is set to 30, respectively. The distance between two D2D nodes in a D2D pair, the transmit power of D2D transmission, and the value of β in PNC opportunistic relay scheme are the same as those used in Fig. 8. The number of inactive CUs is set to 1000. From Fig. 10 and Fig. 11, we can see that in all schemes, the maximum sum-rate of D2D pairs increases as the numbers of D2D pairs and active CUs increase. The reason is as follows. Increasing the number of D2D pairs increases the sum-rate of D2D pairs directly; while increasing the number of active CUs reduces the interference to the D2D pairs. Another observation from these two figures is that the performance gains of NNC, HNC, and PNC opportunistic relay schemes versus the no-relay scheme change slowly as the numbers of active CUs and D2D pairs increase. That is, the performance gains are not sensitive to the network size.

Fig. 12 and Fig. 13 show the performance comparison of the four schemes under different settings of the distance between two D2D nodes in a D2D pair when the transmit power of D2D transmission is set to 10 dBm, and under different transmit powers of D2D transmission when the distance between two D2D nodes in a D2D pair is set to 80 m, respectively. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2947021, IEEE Transactions on Vehicular Technology

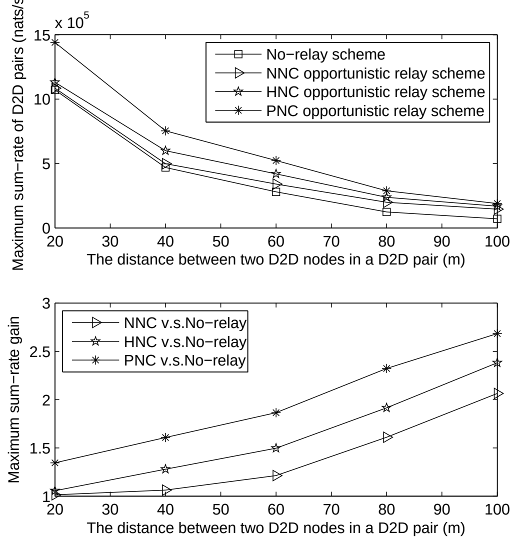


Fig. 12. Performance comparison of the four schemes under different settings of the distance between two D2D nodes in a D2D pair.

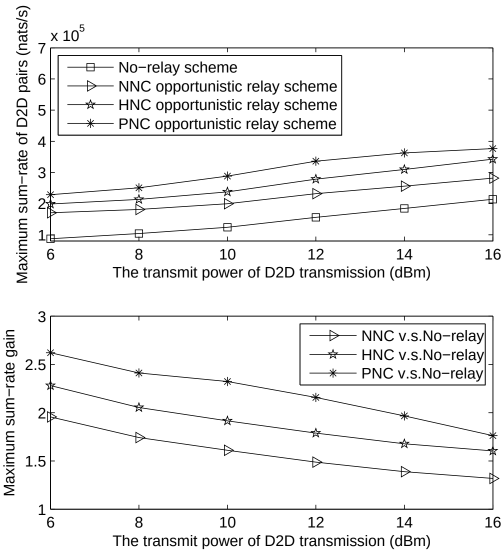


Fig. 13. Performance comparison of the four schemes under different transmit powers of D2D transmission.

number of active CUs, the number of D2D pairs, the number of inactive CUs, and the value of β in PNC opportunistic relay scheme are set to 30, 15, 1000, and 0.1, respectively. From Fig. 12 and Fig. 13, we can see that in all schemes, as the distance between two D2D nodes in a D2D pair increases or the transmit power of D2D transmission decreases, the maximum sum-rate of D2D pairs decreases and the performance gains of NNC, HNC, and PNC opportunistic relay schemes versus the no-relay scheme increase. The maximum performance gains of NNC, HNC, and PNC opportunistic relay schemes versus the no-relay scheme reach 206%, 238%, and 268%, respectively. The reason is that as the distance between two D2D nodes in a D2D pair increases or the transmit power of D2D transmission decreases, the SINR at each receiving device of D2D transmissions decreases, which leads to the decrease of transmission rate. Furthermore, according to the Shannon's capacity formula, the transmission rate is more sensitive to the SINR at the low SINR domain. Thus, when the SINR at each receiving device of D2D transmissions is lower, adding a relay node to increase the SINR can increase the average transmission rate of D2D pairs more effectively. Therefore, the performance gains of NNC, HNC, and PNC opportunistic relay schemes versus the no-relay scheme increase as the distance between two D2D nodes in a D2D pair increases or the transmit power of D2D transmission decreases.

Fig. 14 shows the effect of the number of inactive CUs on the system performance. The number of active CUs, the number of D2D pairs, the distance between two D2D nodes in a D2D pair, the transmit power of D2D transmission, and the value of β in PNC opportunistic relay scheme are set to 30, 15, 80 m, 10 dBm, and 0.1, respectively. From Fig. 14, we can see that the performance of NNC, HNC, PNC opportunistic relay scheme increases as the number of inactive CUs increases. The reason is that when the number of inactive CUs is bigger, we can select a properer relay for each D2D pair, which improves the system performance.

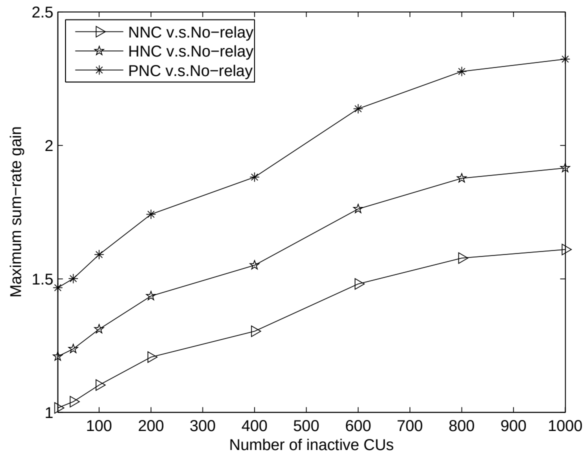


Fig. 14. Effect of the number of inactive CUs.

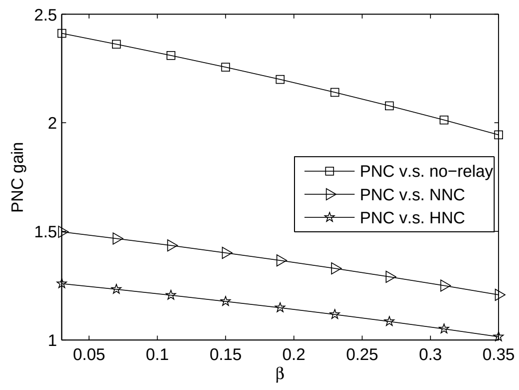


Fig. 15. Performance gain of PNC opportunistic relay scheme under different values of β.

Fig. 15 shows the performance gain of PNC opportunistic relay scheme under different values of β when the number of active CUs, the number of D2D pairs, the number of inactive CUs, the distance between two D2D nodes in a D2D pair, and the transmit power of D2D transmission are set to 30, 15, 1000, 80 m, and 10 dBm, respectively. From Fig. 15, we can see that the performance gains of the PNC opportunistic relay scheme versus the no-relay scheme, the NNC and HNC opportunistic relay schemes decrease as the value of β increases, and the performance of PNC opportunistic relay scheme outperforms the other schemes when β is smaller than 0.35. The maximum performance gains of the PNC opportunistic relay scheme versus the no-relay scheme, the NNC, and HNC opportunistic relay schemes reach 241%, 150%, and 126%, respectively.

VII. Conclusion

In this paper, we have investigated the sum-rate maximization problem of the D2D pairs while satisfying the rate requirements of active CUs in the D2D communications underlaying cellular network over the Rayleigh fading channel. For the D2D pairs, the no-relay scheme, the NNC, the HNC, and the PNC opportunistic relay schemes were considered. We found that in the four schemes, given the relay selections, the sum-rate maximization problem of D2D pairs can be formulated as an MINLP. To solve this problem, we proposed a two-step approach to obtain the solution to the formulated MINLP by first deriving the optimal SINR thresholds to maximize the transmission rates under different transmission schemes for each possible pairing of a D2D pair and a CU. Based on the maximum transmission rates of D2D pairs for each possible pairing in the first step, a bipartite-matching method was further proposed to optimize the CU-D2D pairing. According to the solution to the MINLP, an iterative relay selection algorithm was developed to find out the relays that can further improve the sum-rate of D2D communications. Extensive simulation results showed that: 1) compared with the no-relay scheme, the NNC, HNC, and PNC opportunistic relay schemes achieve a maximum performance enhancement of 106%, 138%, and 168%, respectively; 2) when three percent of time is used for the transmit power adjustment, the performance gains of the PNC opportunistic relay scheme versus the no-relay scheme, the NNC and HNC opportunistic relay schemes can reach 241%, 150%, and 126%, respectively.

In this paper, we only performed numerical simulations to validate the theoretical analysis and the proposed algorithms. In the future work, some practical experiments will be considered for further validations.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2013-2018," presented at the White Paper, San Jose, CA, USA, Feb. 2014.
[2] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-todevice communication as an underlay to LTE-advanced networks," IEEE Communications Magazine, vol. 47, no. 12, pp. 42-49, Dec. 2009.
[3] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-Device Communication in LTE-Advanced Networks: A Survey," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 1923-1940, 2015.
[4] H. Nishiyama, M. Ito, and N. Kato, "Relay-by-smartphone: realizing multihop device-to-device communications," IEEE Communications Magazine, vol. 52, no. 4, pp. 56-65, Apr. 2014.
[5] J. Liu, S. Zhang, N. Kato, H. Ujikawa, and K. Suzuki, "Device-to-Device Communications for Enhancing Quality of Experience in Software Defined Multi-Tier LTE-A Networks," IEEE Network Magazine, vol. 29, no. 4, pp. 46-52, Jul. 2015.
[6] J. Liu, Y. Kawamoto, H. Nishiyama, N. Kato, and N. Kadowaki, "Device-to-Device Communications Achieve Efficient Load Balancing in LTE-Advanced Networks," IEEE Wireless Communications Magazine, vol. 21, no. 2, pp.57-65, Apr. 2014.
[7] S. Zhang, S. C. Liew, and P. P. Lam, "Physical layer network coding," http://arxiv.org/ftp/arxiv/papers/0704/0704.2475.pdf.
[8] Y. Li, D. Jin, J. Yuan, and Z. Han, "Coalitional Games for Resource Allocation in the Device-to-Device Uplink Underlaying Cellular Networks," IEEE Transactions on Wireless Communications, vol. 13, no. 7, pp. 3965-3977, Jul. 2014.
[9] D. Feng, L. Lu, Y. W. Yi, G. Y. Li, G. Feng, and S. Li, "Device-to-Device Communications Underlaying Cellular Networks," IEEE Transactions on Communications, vol. 61, no. 8, pp. 3541-3551, Aug. 2013.
[10] L. Wang, H. Tang, H. Wu, and G. L. Stuber, "Resource Allocation for D2D Communications Underlay in Rayleigh Fading Channels," IEEE Transactions on Vehicular Technology, vol. 66, no. 2, pp. 1159-1170, Feb. 2017.
[11] J. Liu, H. Nishiyama, N. Kato, and J. Guo, "On the Outage Probability of Device-to-Device Communication Enabled Multi-Channel Cellular Networks: A RSS Threshold-Based Perspective," IEEE Journal on Selected Areas in Communications, vol. 34, no. 1, pp. 163-175, Jan. 2016.
[12] X. Ge, S. Tu, G. Mao, C. X. Wang, and T. Han, "5G Ultra-Dense Cellular Networks," IEEE Wireless Communications, vol. 23, no. 1, pp. 72-79, Feb. 2016.
[13] D. Calabuig, S. Barmpounakis, S. Gimenez, et al. "Resource and Mobility Management in the Network Layer of 5G Cellular Ultra-Dense Networks," IEEE Communications Magazine, vol. 55, no. 6, pp. 162-169, Jun. 2017.
[14] M. Hasan, E. Hossain, and D. I. Kim, "Resource Allocation Under Channel Uncertainties for Relay-Aided Device-to-Device Communication Underlaying LTE-A Cellular Networks," IEEE Transactions on Wireless Communications, vol. 13, no. 4, pp. 2322-2338, Apr. 2014.
[15] B. Zhou, H. Hu, S. Huang, and H. Chen, "Intracluster Device-to-Device Relay Algorithm With Optimal Resource Utilization," IEEE Transactions on Vehicular Technology, vol. 62, no. 5, pp. 2315-2326, Jun. 2013.
[16] R. Ma, Y. Chang, H. Chen, and C. Chiu, "On Relay Selection Schemes for Relay-Assisted D2D Communications in LTE-A Systems," IEEE Transactions on Vehicular Technology, vol. 66, no. 9, pp. 8303-8314, Sept. 2017.
[17] R. Zhang, Y. Li, C. X. Wang, Y. Ruan, and H. Zhang, "Performance Tradeoff in Relay Aided D2D-Cellular Networks," IEEE Transactions on Vehicular Technology, vol. 67, no. 10, pp. 10144-10149, Oct. 2018.
[18] D. Ebrahimi, H. Elbiaze, and W. Ajib, "Device-to-Device Data Transfer Through Multihop Relay Links Underlaying Cellular Networks," IEEE Transactions on Vehicular Technology, vol. 67, no. 10, pp. 9669-9680, Oct. 2018.
[19] J. Dai, J. Liu, Y. Shi, S. Zhang, and J. Ma, "Analytical Modeling of Resource Allocation in D2D Overlaying Multi-hop Multi-channel Uplink Cellular Networks," IEEE Transactions on Vehicular Technology, vol. 66, no. 8, pp. 6633-6644, Aug. 2017.
[20] J. Liu, S. Zhang, H. Nishiyama, N. Kato, and J. Guo, "A stochastic geometry analysis of D2D overlaying multi-channel downlink cellular networks," IEEE Conference on Computer Communications (INFOCOM), pp. 46-54, 2015.
[21] B. Bai, W. Chen, K. B. Letaief, and Z. Cao, "Distributed WRBG Matching Approach for Multiflow Two-Way D2D Networks," IEEE Transactions on Wireless Communications, vol. 15, no. 4, pp. 2925-2939, Apr. 2016.
[22] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Energy Efficiency and Spectrum Efficiency of Multihop Device-to-Device Communications Underlaying Cellular Networks," IEEE Transactions on Vehicular Technology, vol. 65, no. 1, pp. 367-380, Jan. 2016.
[23] C. Gao, Y. Li, Y. Zhao, and S. Chen, "A Two-Level Game Theory Approach for Joint Relay Selection and Resource Allocation in Network Coding Assisted D2D Communications" IEEE Transactions on Mobile Computing, vol. 16, no. 10, pp. 2697-2711, Oct. 2017.
[24] Y. Zhao, Y. Li, and N. Ge, "Physical Layer Network Coding Aided Two-Way Device-to-Device Communication Underlaying Cellular Net-

works," IEEE Global Communications Conference (GLOBECOM), pp. 1-6, 2015.

[25] J. G. Proakis, Digital Communications, 3rd ed. Singapore: McGraw-Hill, 1995.

[26] D. Zhu, J. Wang, A. L. Swindlehurst, C. Zhao, "Downlink Resource Reuse for Device-to-Device Communications Underlaying Cellular Networks," IEEE Signal Processing Letters, vol. 21, no. 5, pp. 531-534, May 2014.

[27] I. Mondal, A. Neogi, P. Chaporkar, A. Karandikar, "Bipartite Graph Based Proportional Fair Resource Allocation for D2D Communication," IEEE Wireless Communications and Networking Conference (WCNC), pp. 1-6, Mar. 2017.

[28] J. Hu, W. Heng, X. Li, J. Wu, "Energy-Efficient Resource Reuse Scheme for D2D Communications Underlaying Cellular Networks," IEEE Communications Letters, vol. 21, no. 9, pp. 2097-2100, Sept. 2017.

[29] A. J. Goldsmith, Wireless Communications. New York: Cambridge University Press, 2005.

[30] Physical Channels and Modulation. 3GPP TS 36.211. (Release 8).

[31] H. Song, X. Fang, L. Yan, and Y. Fang, "Control/User Plane Decoupled Architecture Utilizing Unlicensed Bands in LTE Systems," IEEE Wireless Communications, vol. 24, no. 5, pp. 132-142, Oct. 2017.

[32] D. West et al., Introduction to Graph Theory. Prentice Hall, 2001.

[33] H. Song, X. Fang, and Y. Fang, "Unlicensed Spectra Fusion and Interference Coordination for LTE Systems," IEEE Transactions on Mobile Computing, vol. 15, no. 12, pp. 3171-3184, Dec. 2016.

[34] B. Jackson, "The Optimal Assignment Problem," http://www.maths.qmul.ac.uk/ bill/MAS210/ch6.pdf.

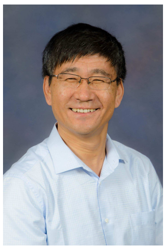
[35] L. Fu, H. Kim, J. Huang, S. C. Liew, and M. Chiang, "Energy conservation and interference mitigation: From decoupling property to win-win strategy," IEEE Transactions on Wireless Communications, vol. 10, no. 11, pp. 3943-3955, Nov. 2011.

Shijun Lin (M'12) is currently an Associate Professor with the Department of Communication Engineering at Xiamen University (Xiamen, China). He received the B.S. degree (honors) in Electrical Engineering from Xiamen University in 2005, and the Ph.D. Degree in Electrical Engineering from Tsinghua University (Beijing, China) in 2010. From 2016 to 2017, he was a visiting scholar with the Department of Electrical and Computer Engineering, University of Florida. His current research interests are in the area of wireless communications and networking, with focus on cognitive radio network, physical-layer network coding, Device-to-Device communication, wireless greening, resource allocation, and distributed protocol design. He has served as TPC Chair/TPC Member for several international conferences such as ICC, Globecom, ICCC, MLICOM. He is also on the editorial board of KSII Transactions on Internet and Information Systems.

Yong Li (M'09-SM'16) received the B.S. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007 and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2012. He is currently a Faculty Member of the Department of Electronic Engineering, Tsinghua University.

Dr. Li has served as General Chair, TPC Chair, SPC/TPC Member for several international workshops and conferences, and he is on the editorial board of two IEEE journals. His papers have total citations more than 6900. Among them, ten are ESI Highly Cited Papers in Computer Science, and four receive conference Best Paper (run-up) Awards. He received IEEE 2016 ComSoc Asia-Pacific Outstanding Young Researchers, Young Talent Program of China Association for Science and Technology, and the National Youth Talent Support Program.

Haichuan Ding received the B.Eng. and M.S. degrees in electrical engineering from Beijing Institute of Technology (BIT), Beijing, China, in 2011 and 2014, and the Ph.D. degree in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in 2018. From 2012 to 2014, he was with the Department of Electrical and Computer Engineering, the University of Macau, as a visiting student. Since 2019, he has been a postdoctoral research fellow with the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, MI, USA. His current research interests include mmWave and V2X communications.



Yuguang Fang (F'08) received an MS degree from Qufu Normal University, Shandong, China in 1987, a PhD degree from Case Western Reserve University in 1994, and a PhD degree from Boston University in 1997. He joined the Department of Electrical and Computer Engineering at University of Florida in 2000 as an assistant professor, then was promoted to an associate professor in 2003 and a full professor in 2005, and has been a distinguished professor since 2019. He holds a University of Florida Foundation Preeminence Term Professorship (2019-2022), a University of Florida Foundation Professorship (2017-2020, 2006-2009), a University of Florida Term Professorship (2017-2019, 2019-2021).

Dr. Fang received the US NSF Career Award in 2001, the US ONR Young Investigator Award in 2002, the 2018 IEEE Vehicular Technology Outstanding Service Award, the 2015 IEEE Communications Society CISTC Technical Recognition Award, the 2014 IEEE Communications Society WTC Recognition Award, the Best Paper Award from IEEE ICNP (2006), and a 2010-2011 UF Doctoral Dissertation Advisor/Mentoring Award. He was the Editor-in-Chief of IEEE Transactions on Vehicular Technology (2013-2017) and IEEE Wireless Communications (2009-2012), and serves/served on several editorial boards of premier journals. He also served as the Technical Program Co-Chair of IEEE INFOCOMąŕ2014. He is a fellow of IEEE and AAAS.

Jianghong Shi received his Ph.D. from Xiamen University, China, in 2002. He is currently a professor in the School of Information Science and Technology, Xiamen University. He is also the director of the West Straits Communications Engineering Center, Fujian Province, China. His research interests include wireless communication networks and satellite navigation.