Subchannel Allocation with Low Computational and Signaling Complexity in 5G D2D Networks

Bala Venkata Ramulu Gorantla, Student Member, IEEE

Neelesh B. Mehta, Fellow, IEEE

Abstract-Device-to-device (D2D) communication enables novel proximity services based applications in 5G networks. In underlay D2D, cellular users share subchannels with D2D users leading to interference between them. To efficiently manage the interference and assign D2D pairs to subchannels, we propose a novel relaxation-pruning algorithm (RPA). It allocates at most K D2D pairs per subchannel, where K is a system parameter that controls the trade-off between spatial reuse and inter-D2D interference. RPA is designed for a low signaling overhead scenario. In it, a D2D user feeds back a quantized rate to the base station that meets an outage probability constraint even though the user has only statistical knowledge of the inter-D2D and inter-cell interferences. RPA has polynomial-time complexity. It provably guarantees a D2D sum rate that is at least half of the optimal value, achieving which requires exponential complexity. This is unlike conventional approaches that offer no such performance guarantees or a weaker guarantee. Numerical results show that the D2D sum throughput of RPA is better than conventional algorithms and is within 1% of the optimal value.

I. INTRODUCTION

5G promises to support a wide variety of applications that connect a large number of devices with high data rates. Proximity services (ProSe) based applications, such as first-responder communications, social networking, advertising, vehicle-to-vehicle communication, and video caching, are an important class of 5G services [1], [2]. These are enabled by device-to-device (D2D) communication, in which the devices communicate directly but under the supervision of the base station (BS). D2D communication offloads traffic from the BS, reduces latency, and improves energy and spectral efficiencies [3]. Standardization activities on D2D protocols are apace in the third generation partnership project (3GPP).

In underlay D2D, cellular users (CUs) share subchannels with the D2D users. While this improves spatial reuse, it also leads to additional interference. For example, in the uplink, a CU causes interference to the D2D receiver (DRx), and the D2D transmitter (DTx) causes interference to the BS. Therefore, to assign D2D users to subchannels, the BS needs to take into account the channel state information (CSI) of the direct D2D links and CU-to-BS links, and the interference between the CUs and D2D users.

Three fundamental issues arise in practical D2D deployments. First, the CSI about the DTx-to-DRx and CU-to-DRx links is available to the DRx but not the BS because it is not a receiver in them. Therefore, this CSI must be fed back to the BS by the DRx. Second, the inter-D2D interference is not known to the DRx at the time of feedback because it does not yet know which other D2D pairs will share a subchannel with it. Third, neither the BS nor the DRx know the instantaneous inter-cell interference a priori. This is because they do not know which users will be scheduled in the neighboring cells and the channel gains from those users to them.

These issues have been tackled to a limited extent in the underlay D2D literature. Full CSI at the BS is assumed in [4]–[7]. This is practically infeasible as the BS is required to know NM DTx-to-DRx and NM CU-to-DRx channel gains, and $N\binom{M}{2}$ inter-D2D channel gains in a cell with N subchannels and M D2D pairs. This issue is addressed more realistically in [8], [9], which assume that only partial or statistical CSI about the CU-to-DRx, DTx-to-DRx, and inter-D2D links is available at the BS. In [8], at most one D2D pair is allocated to a subchannel while guaranteeing a minimum rate with a pre-specified outage probability for the CUs and D2D pairs. Multiple D2D pairs are allocated to a subchannel in [9]. However, the clustering-based ad hoc algorithm proposed in it offers no performance guarantees.

A fourth operational issue is the complexity and efficacy of the algorithm employed at the BS that assigns D2D pairs to subchannels. Its design is intimately coupled to the CSI model. In [4], [5], [8], only one D2D pair is assigned to a subchannel, which limits the spatial reuse gains. Assigning multiple D2D pairs to a subchannel in [6], [9], [10] is NP-hard. In [10], a polynomial-time cardinality-constrained subchannel assignment algorithm (CCSAA) is proposed when the BS has partial CSI. However, CCSAA is only guaranteed to achieve at least 1/3rd of the optimal D2D sum rate.

A. Contributions

We consider a low CSI feedback model and propose an algorithm to allocate multiple D2D pairs to multiple subchannels that address all the four issues discussed above. In our low signaling complexity model, a DRx feeds back only q bits per subchannel. The feedback is such that the quantized D2D data rate conveyed to the BS has an outage probability of at most ϵ_d . Here, q and ϵ_d are system parameters.

We propose a novel relaxation-pruning algorithm (RPA) for allocating multiple D2D pairs to subchannels. It has a polynomial-time complexity and is an adaptation of the Shmoys and Tardos algorithm [11]. It allocates at most K

B. V. R. Gorantla and N. B. Mehta are with the Dept. of Electrical Communication Eng., Indian Institute of Science, Bangalore, India. This work was supported by the Indigenous 5G Test Bed Project, Dept. of Telecommunications, India and the DST-Swarnajayanti Fellowship DST/SJF/ETA-01/2014-15. Emails: balavenkataramulu@gmail.com, nbmehta@iisc.ac.in.

D2D pairs to a subchannel. Here, K is a system parameter that trades off between the spectral efficiency gains from assigning more D2D pairs to a subchannel and the increased inter-D2D interference that this begets. RPA also guarantees each CU a minimum rate with a pre-specified probability of outage. RPA combines a reasonable level of CSI knowledge with a performance guarantee on the D2D sum rate.

An important property of RPA, which we prove in this paper, is that its D2D sum rate is at least half of the optimal D2D sum rate. Such a result is noteworthy because several algorithms in the literature provide no such guarantee [6]–[9]. Our numerical results show that the D2D sum throughput of RPA is indistinguishable from that of exhaustive search, which requires exponential complexity. It also outperforms several algorithms proposed in the literature [4], [5], [8], [10].

Comparison with Literature: Our work differs from the literature in several respects. First, we do not assume full CSI, which causes a high signaling complexity and is assumed in [4]–[7]. Second, inter-cell interference is not considered in [4]–[9]. Third, we consider multiple subchannels and allow multiple D2D pairs to share a subchannel. This problem is technically more challenging than assigning at most one D2D pair to a subchannel [4], [5], [8]. Lastly, RPA provides a much stronger theoretical performance guarantee than CCSAA [10].

B. Outline

Section II discusses the system model. Section III presents RPA. Section IV presents numerical results. Our conclusions follow in Section V.

II. UNDERLAY D2D SYSTEM MODEL

We consider N subchannels, N CUs and M D2D pairs in a cell. Let $\mathcal{D} = \{1, 2, ..., M\}$ be the set of D2D pairs and $\mathcal{S} = \{1, 2, ..., N\}$ be the set of orthogonal uplink subchannels. Each subchannel is assigned to a CU. Without loss of generality, let CU *i* be allocated to subchannel *i*. The D2D pairs share the uplink subchannels with the CUs.

The transmit power of the CU is P_c and that of the DTx of a D2D pair is P_d . The channel power gain from CU *i* to the DRx of D2D pair *j* on subchannel *i* is $g_{ji}(i)$. The uplink channel power gain from CU *i* to the BS on subchannel *i* is $h_{bi}(i)$. The channel power gain between the DTx and DRx of D2D pair *j* on subchannel *i* is $h_{jj}(i)$, and from the DTx of D2D pair *j* to the BS is $g_{bj}(i)$. The channel power gain from the DTx of D2D pair *k* to the DRx of D2D pair *j* on subchannel *i* is $g_{ik}^d(i)$. The system model is shown in Fig. 1.

A. CSI and Interference Model

The CSI available at the BS is inherently different from that at the DRx. The BS knows $h_{bi}(i)$ and $g_{bj}(i)$, $\forall i \in S, j \in D$, as it can estimate them using the reference signals transmitted by CU *i* and the DTx of D2D pair *j*. Similarly, the DRx of D2D pair *j* knows $h_{jj}(i)$ and $g_{ji}(i)$, $\forall i \in S, j \in D$, as it can estimate them using the reference signals transmitted by the DTx of D2D pair *j* and CU *i* [1].



Fig. 1. System model that illustrates the links from DTx to DRx, CU to BS, CU to DRx, and DTx to BS, and inter-cell interference.

Inter-D2D Interference: The inter-D2D interference power $I_{jk}(i)$ from the DTx of D2D pair k to the DRx of D2D pair j is $I_{jk}(i) = P_d g_{jk}^d(i)$. The channel gain $g_{jk}^d(i)$ is not known to D2D pair j as it requires considerable coordination among the D2D pairs. Therefore, only the statistics of $I_{jk}(i)$, $\forall k$, is assumed to be known to the D2D pair j.

Inter-Cell Interference: The interference powers from the neighboring-cell users to the BS and DRx of D2D pair j are $I_{BS}(i)$ and $I_j(i)$, respectively. These are not known to the BS and the DRx because they require a priori knowledge of the users scheduled in the neighboring cells and channel gains from those users. Therefore, only the statistics of $I_{BS}(i)$ and $I_j(i)$, $\forall j \in \mathcal{D}$, are assumed to be known at the BS and the DRxs. This is practically easier since their statistics change at a much slower time scale.

Let x_{ij} be the assignment variable that is 1 if subchannel *i* is assigned to D2D pair *j*, and is 0 otherwise. The signal-tointerference-plus-noise ratio (SINR) $\xi_j^d(i)$ of D2D pair *j* on subchannel *i* is

$$\xi_j^d(i) = \frac{P_d h_{jj}(i)}{P_c g_{ji}(i) + I_j^d(i) + \sigma^2},$$
(1)

where $I_j^d(i) = \sum_{k=1, k \neq j}^M x_{ik} I_{jk}(i) + I_j(i)$ is the sum of inter-D2D and inter-cell interferences, and σ^2 is the noise power.

D2D Assignment Limit: Assigning multiple D2D pairs to a subchannel can improve spatial reuse. However, it also increases the inter-D2D interference, which can decrease the rate of the D2D pairs. To address this trade-off, we allow at most K D2D pairs to share a subchannel, where K is a system parameter. We shall refer to it as the D2D assignment limit. Therefore, $\forall i \in S$, we have $\sum_{j=1}^{M} x_{ij} \leq K$.

B. Low Signaling Complexity Feedback Model

The BS does not know $h_{jj}(i)$ and $g_{ji}(i)$, $\forall i \in S, j \in D$. Therefore, the DRx has to feed back its SINR to the BS. Even the DRx of D2D pair j cannot exactly know its SINR on subchannel i because of the uncertainty due to the inter-cell and inter-D2D interferences. However, as the following calculations show, it can still compute an SINR estimate $T_{ij}(\epsilon_d)$ such that when it transmits with the rate $\log_2(1 + T_{ij}(\epsilon_d))$ bps/Hz, its outage probability is ϵ_d . This is equivalent to

$$\Pr\left(\frac{P_d h_{jj}(i)}{P_c g_{ji}(i) + I_j^d(i) + \sigma^2} \ge T_{ij}(\epsilon_d)\right) = 1 - \epsilon_d.$$
(2)

Rearranging and writing in terms of the cumulative distribution function (CDF) $F_j(\cdot)$ of $I_i^d(i)$, we get

$$F_j\left(\frac{P_d h_{jj}(i)}{T_{ij}(\epsilon_d)} - P_c g_{ji}(i) - \sigma^2\right) = 1 - \epsilon_d.$$
 (3)

Rearranging terms again, we get

$$T_{ij}(\epsilon_d) = \frac{P_d h_{jj}(i)}{P_c g_{ji}(i) + F_j^{-1}(1 - \epsilon_d) + \sigma^2},$$
 (4)

where $F_i^{-1}(\cdot)$ is the inverse CDF of $I_i^d(i)$.

At the time of generating feedback, a D2D pair does not know which other D2D pairs will interfere with it. Therefore, to evaluate $F_j^{-1}(\cdot)$, we conservatively assume that the (K-1)closest D2D pairs interfere with it. Hence, in $I_j^d(i)$ we replace $\sum_{k=1,k\neq j}^M x_{ik}I_{jk}(i)$ with $\sum_{k=1}^{K-1} I_{j(k)}(i)$, where (k) denotes the k^{th} closest DTx to the DRx of D2D pair j. This ensures that the outage probability does not exceed ϵ_d regardless of which D2D pairs are assigned to the subchannel.

The DRx sends a q-bit feedback δ_{ij} about $T_{ij}(\epsilon_d)$ to the BS as follows. Let the $L = 2^q$ quantization thresholds be $0 = \Psi_0 < \Psi_1 < \cdots < \Psi_{L-1} < \infty$. The DRx quantizes $T_{ij}(\epsilon_d)$ and feeds back δ_{ij} , which is given by

$$\delta_{ij} = l, \quad \text{if } \Psi_l \le T_{ij}(\epsilon_d) < \Psi_{l+1}. \tag{5}$$

Given δ_{ij} , the BS determines the rate C_{ij} of D2D pair j on subchannel i as

$$C_{ij} = \log_2(1 + \Psi_{\delta_{ij}}). \tag{6}$$

This is the only information the BS has about the D2D rates.

Signaling Complexity: As each DRx feeds back q bits per subchannel, its signaling overhead is Nq bits. With M D2D pairs, the total signaling overhead in the cell is MNq bits.

C. Quality-of Service (QoS) Guarantee for All CUs

The SINR $\xi_i^c(i)$ of CU *i* on its allocated subchannel *i* is

$$\xi_i^c(i) = \frac{P_c h_{bi}(i)}{\sum_{j=1}^M x_{ij} P_d g_{bj}(i) + I_{\text{BS}}(i) + \sigma^2}.$$
 (7)

We require that it must be able to transmit at a minimum rate $R_{\min}^{(i)}$ with an outage probability that is at most ϵ_c , which is a system parameter and depends on the type of data traffic at the CU. Therefore,

$$\Pr\left(\log_2\left(1+\xi_i^c(i)\right) \ge R_{\min}^{(i)}\right) \ge 1-\epsilon_c.$$
(8)

Substituting $\xi_i^c(i)$ from (7) and rearranging terms, we get

$$\sum_{j=1}^{M} x_{ij} w_{ij} \le b_i, \tag{9}$$

where $w_{ij} = P_d g_{bj}(i)$ is the interference power at the BS on subchannel *i* due to the DTx of D2D pair *j*, $b_i = \frac{P_c h_{bi}(i)}{2^{R_{\min}^{(i)}} -} - \sigma^2 - F_{BS}^{-1}(1 - \epsilon_c)$, and $F_{BS}^{-1}(\cdot)$ is the inverse of the CDF of $I_{BS}(i)$. Thus, the cumulative interference at the BS from the D2D pairs assigned to subchannel *i* should not exceed b_i .

D. Subchannel Allocation Problem Formulation

Our problem of allocating subchannels to D2D pairs to maximize the sum of D2D rates is as follows:

$$\mathcal{P}: \max_{x_{ij}, \forall i \in \mathcal{S}, j \in \mathcal{D}} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij} C_{ij} \right\},$$
(10)

subject to
$$\sum_{i=1}^{N} x_{ij} \le 1, \quad \forall j \in \mathcal{D},$$
 (11)

$$\sum_{j=1}^{M} x_{ij} w_{ij} \le b_i, \quad \forall i \in \mathcal{S},$$
(12)

$$\sum_{j=1}^{M} x_{ij} \le K, \quad \forall i \in \mathcal{S},$$
(13)

$$x_{ij} \in \{0, 1\}, \quad \forall i \in \mathcal{S}, j \in \mathcal{D}.$$
 (14)

Constraint (11) ensures that at most one subchannel is assigned to a D2D pair, (12) represents the QoS guarantee for the CUs, and (13) mandates that at most K D2D pairs are assigned to a subchannel. \mathcal{P} is known to be NP-hard [11].

Generality of Approach: The above formulation applies to any model of the inter-D2D and inter-cell interferences. It also allows for $I_1^d(i), \ldots, I_M^d(i)$ being statistically nonidentical since the distances between co-channel interferers and the different DRxs are different. For example, with Rayleigh fading and lognormal shadowing, $I_j^d(i)$ is the sum of Suzuki random variables (RVs). It can be approximated as a lognormal RV with dB-mean μ_j and dB-standard deviation σ_j [12, Ch. 3]. The CDF $F_j(\cdot)$ of $I_j^d(i)$ can be shown to be $F_j(x) = 1 - Q\left(\frac{10 \log_{10}(x) - \mu_j}{\sigma_j}\right)$, for $x \ge 0$ [12, Ch. 3], where $Q(\cdot)$ is the Q-function. Therefore, the inverse CDF is $F_j^{-1}(x) = 10^{0.1(\mu_j + \sigma_j Q^{-1}(1-x))}$, for $0 \le x \le 1$, where $Q^{-1}(\cdot)$ is the inverse Q-function. Similarly, $I_{\rm BS}(i)$ can be approximated as a lognormal RV with dB-mean μ_B and dB-standard deviation σ_B . Its inverse CDF is $F_{\rm BS}^{-1}(x) = 10^{0.1(\mu_B + \sigma_B Q^{-1}(1-x))}$, for $0 \le x \le 1$.

III. RELAXATION-PRUNING ALGORITHM (RPA)

RPA uses an approach based on linear programming relaxation and rounding. It consists of four steps. The rationale behind them will come out in the proof of the performance guarantee in Result 1. The four steps are as follows.

1) For all $i \in S$, $j \in D$, we set $x_{ij} = 0$ if $w_{ij} > b_i$, since the D2D pair j will violate (12) and can never be assigned to subchannel i. For all other x_{ij} , the binary integer constraint in (14) is relaxed to $0 \le x_{ij} \le 1$. This changes \mathcal{P} to a linear program, which is optimally solved in polynomial time by using the dual simplex [13, Ch. 4] or interior-point

$$\begin{array}{c|ccccc} \text{Virtual} & & i_1 & \tilde{x}_{i_1[3]} + \tilde{x}_{i_2[3]} & i_2 & \tilde{x}_{i_2[5]} + \tilde{x}_{i_3[5]} & i_3 \\ & = \tilde{x}_{i[3]} & & \\ & & & & \\ & & & \\ & & & \\ & & & & \\ & & & \\ & & & \\ & & & & \\ & & & \\ & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & &$$

Fig. 2. Step 2 of RPA: Example showing the edges in the bipartite graph between the D2D pairs and the virtual subchannels for subchannel i.

methods [13, Ch. 5]. Let \tilde{x}_{ij} , $\forall i \in S, j \in D$, be the optimal solution to the linear program. We shall refer to $\tilde{x}_{ij}, \forall j \in D$, as the *fractional solution* for subchannel *i*.

2) For each subchannel *i*, compute $n_i = \lceil \sum_{j=1}^{M} \tilde{x}_{ij} \rceil$, where $\lceil . \rceil$ denotes the ceiling function. Clearly, $n_i \leq K$ since (13) implies that $\sum_{j=1}^{M} \tilde{x}_{ij} \leq K$. We construct a bipartite graph with $\sum_{i=1}^{N} n_i$ vertices on one side and M D2D pairs as vertices on the other side. For each subchannel *i*, the construction proceeds as follows:

- Create n_i copies of the subchannel *i*, which are denoted by $i_1, i_2, \ldots, i_{n_i}$. Henceforth, we shall refer to these as *virtual subchannels* of *i*.
- Consider the set D'_i = {j : x̃_{ij} ≠ 0, j ∈ D} of D2D pairs, whose fractional solution for subchannel i is non-zero. Arrange the D2D pairs in D'_i in the non-increasing order of their interference power to the BS:

$$w_{i[1]} \ge w_{i[2]} \ge \dots \ge w_{i[|\mathcal{D}'_i|]}.$$
 (15)

Here, using order statistics notation, [k] is the D2D pair in \mathcal{D}'_i that causes the k^{th} largest interference to the BS.

- Let j_1 be such that $\tilde{x}_{i[1]} + \tilde{x}_{i[2]} + \cdots + \tilde{x}_{i[j_1-1]} < 1$ and $\tilde{x}_{i[1]} + \cdots + \tilde{x}_{i[j_1-1]} + \tilde{x}_{i[j_1]} \ge 1$. Then, construct edges between virtual subchannel i_1 and D2D pairs $[1], [2], \ldots, [j_1]$.
- An edge between i₂ and [j₁] is constructed only if x̃_{i[1]} + ... + X̃_{i[j1-1]} + X̃_{i[j1]} > 1. Let j₂ be such that X̃_{i[1]} + ... + X̃_{i[j1]} + ... + X̃_{i[j2-1]} < 2 and X̃_{i[1]} + ... + X̃_{i[j2-1]} + ... + X̃_{i[j2-1]} ≥ 2. Construct edges between i₂ and D2D pairs [j₁ + 1], [j₁ + 2], ..., [j₂].
 In general, let j_k be such that ∑^{j_k-1}_{j=1} X̃_{i[j]} < k and ∑^{j_k}
- In general, let j_k be such that $\sum_{j=1}^{j_k-1} \tilde{x}_{i[j]} < k$ and $\sum_{j=1}^{j_k} \tilde{x}_{i[j]} \ge k$, for $k = 1, 2, \ldots, n_i$. Edges are constructed between virtual subchannel i_{k+1} and D2D pairs $[j_k+1], [j_k+2], \ldots, [j_{k+1}]$. Also, an edge is constructed between i_{k+1} and $[j_k]$ only if $\sum_{j=1}^{j_k} \tilde{x}_{i[j]} > k$.
- The weight of the edge constructed between any virtual subchannel of i and D2D pair j is C_{ij} . The bipartite graph for subchannel i is illustrated in Fig. 2.

3) Run the Kuhn-Munkres algorithm [14, Ch. 3] to find the optimal maximum weighted matching for the above bipartite

graph. It solves the following optimization problem:

$$Q: \max_{y_{i_l j}, \forall i, l, j} \sum_{i=1}^{N} \sum_{l=1}^{n_i} \sum_{j=1}^{M} y_{i_l j} C_{ij},$$
(16)

subject to
$$\sum_{i=1}^{N} \sum_{l=1}^{n_i} y_{i_l j} \le 1, \forall j \in \mathcal{D},$$
 (17)

$$\sum_{j=1}^{M} y_{i_l j} \le 1, \forall i \in \mathcal{S}, \ l \in \{1, \dots, n_i\}, \quad (18)$$

$$y_{i_l j} \in \{0, 1\}, \forall i \in \mathcal{S}, \ l \in \{1, \dots, n_i\}, j \in \mathcal{D},$$
 (19)

where $y_{i_i j}$ is 1 if the edge between virtual subchannel i_l and D2D pair j is selected, and is 0 otherwise. The algorithm maximizes the sum of weights of the selected edges while ensuring that at most one virtual subchannel is connected to a D2D pair (cf. (17)) and at most one D2D pair is connected to a virtual subchannel (cf. (18)). When an edge between any virtual subchannel of i and a D2D pair j is selected, we say that D2D pair j is assigned to subchannel i. We shall refer to this assignment as the *integral matching solution*.

4) Prune the solution of Q to satisfy the constraint in (12) as follows. Let subchannel *i* be a subchannel for which (12) is not satisfied. Let the selected edges in the bipartite graph connect the virtual subchannels $i_1, i_2, \ldots, i_{n_i}$ to the D2D pairs $k_1, k_2, \ldots, k_{n_i}$, respectively, such that their interference powers to the BS are in the descending order: $w_{ik_1} \ge w_{ik_2} \ge \cdots \ge w_{ik_{n_i}}$. If $C_{ik_1} \ge C_{ik_2} + \cdots + C_{ik_{n_i}}$, then only D2D pair k_1 is allocated to subchannel *i*. Otherwise, the D2D pairs k_2, \ldots, k_{n_i} are all allocated to subchannel *i*. As shown in Lemma 1 below, this assignment satisfies (12). This yields the final allocation of subchannels to the D2D pairs.

Lemma *1*: The pruning in Step 4 ensures that the D2D pairs allocated to a subchannel satisfy the constraint in (12).

Proof: The proof is relegated to Appendix A.

We now prove that the four steps of RPA provide the following theoretical guarantee about its performance.

Result 1: The D2D sum rate of RPA is at least half of the optimal D2D sum rate.

Proof: The proof is given in Appendix B.

Computational Complexity: The linear program in Step 1 has a complexity of $\mathcal{O}(N^3M^3)$ [13, Ch. 5]. Step 2 has a complexity of $\mathcal{O}(NM \log M)$. The complexity of Step 3 is $\mathcal{O}((N+2M)^3)$ [14, Ch. 3] and Step 4 is $\mathcal{O}(NM)$. Combining these, the complexity of RPA is $\mathcal{O}(N^3M^3)$.

IV. NUMERICAL RESULTS

We present Monte Carlo simulation results for the following setting. The N CUs and DRxs of the M D2D pairs are dropped with uniform probability within a cell of radius 500 m. The DTx lies with uniform probability within a circle of radius 50 m around the DRx. This models the different distances possible between the DTx and DRx. As per 3GPP [15], the path-loss in dB for the DTx to DRx and CU to DRx links is $148 + 40 \log_{10}(d)$, and for the CU-to-BS and DTx-to-BS links is $128.1 + 37.6 \log_{10}(d)$, where d

 TABLE I

 QUANTIZATION THRESHOLDS FOR GENERATING FEEDBACK

q	Thresholds in dB
1	$\Psi_1 = 12$
2	$\Psi_1 = 4, \ \Psi_2 = 12, \ \Psi_3 = 20$
4	$\Psi_1 = 0, \Psi_2 = 2, \Psi_3 = 4, \Psi_4 = 6, \Psi_5 = 8, \Psi_6 = 10, \Psi_7 =$
	12, $\Psi_8 = 14$, $\Psi_9 = 16$, $\Psi_{10} = 18$, $\Psi_{11} = 20$, $\Psi_{12} = 22$,
	$\Psi_{13} = 24, \ \Psi_{14} = 26, \ \Psi_{15} = 28$

is the distance in km. We illustrate the results for Rayleigh fading, lognormal shadowing with dB-standard deviation of 6, $P_c = 10$ dBm, $P_d = 0$ dBm, $\sigma^2 = -114$ dBm, $\epsilon_c = 0.1$, and $R_{\min}^{(i)} = 1$ bps/Hz, $\forall i \in S.^1$ The quantization thresholds are as per Table I. These are centered around 12 dB and span a 16 dB range for q = 2 bits and a 28 dB range for q = 4 bits. We also consider $q = \infty$, which serves as an upper bound.

The statistics of $I_{BS}(i)$, $I_j(i)$, and $I_j^d(i)$ that are used in the simulations account for the randomness due to shadowing, small-scale fading, and user locations. $I_{BS}(i)$ and $I_j(i), \forall j \in \mathcal{D}$, are measured at the BS and the DRxs for 10,000 realizations, and their empirical CDFs are obtained. Given the locations of the D2D pairs j and k, the statistics of the inter-D2D interference $I_{jk}(i)$ are obtained similarly. Then, the CDF of $I_j^d(i) = \sum_{k=1, k \neq j}^{K-1} I_{j(k)}(i) + I_j(i)$ is determined. *Benchmarking:* We compare RPA with the following:

- Exhaustive Search (ES): In this, the optimal solution of \mathcal{P} is found by searching over all the 2^{MN} possible assignments of D2D pairs to subchannels.
- CCSAA: In this, \mathcal{P} is solved using CCSAA. It uses a locally greedy algorithm that maximizes a sub-modular function to assign D2D pairs to subchannels. The details are in [10], and are not repeated here to conserve space.
- Semi-Orthogonal Sharing Assignment (SSA) [4], [5], [8]: In this, at most one D2D pair is assigned to a subchannel.

A comparison with the approaches in [6], [7], [9] that consider multiple D2D pairs per subchannel is not possible because full CSI at the BS is assumed in [6], [7], and the QoS guarantees and objective functions are different in [7], [9].

Fig. 3 plots the D2D sum throughputs per subchannel of ES, RPA, CCSAA, and SSA as a function of the D2D assignment limit K for $q = 1, 2, \text{ and } \infty$. The throughput of a D2D pair is its assigned rate if the transmission is not in outage, otherwise it is zero. Thus, the throughput accounts for outages. Due to the exponential complexity of ES, we present results for a scenario in which M and N are small. The D2D sum throughput of RPA is within 1% of that of ES for all K. As q increases, the D2D sum throughput increases due to the higher feedback resolution. As K increases, the D2D sum throughputs of all the algorithms increase, reach a maximum value, and then decrease. This is because of the aforementioned trade-off between the spectral efficiency



Fig. 3. Small user count example: Comparison of RPA with ES, CCSAA, and SSA (N = 4, M = 6, and $\epsilon_d = 0.1$).



Fig. 4. Effect of number of D2D pairs M and feedback quantization q on the D2D sum throughput (N = 10, K = 3, and $\epsilon_d = 0.1$).

and the inter-D2D interference. RPA significantly outperforms SSA, whose sum throughput does not depend on K and is a flat line. The optimal K is 3 for ES, CCSAA, and RPA.

Fig. 4 plots the D2D sum throughput of RPA and CCSAA as a function of the number of D2D pairs M for K = 3. SSA is not shown to avoid clutter. RPA outperforms CCSAA for all M, and the gap between the two increases as q increases. As M increases, the D2D sum throughput increases due to multi-user diversity. Therefore, RPA is also scalable with M. Its sum throughput increases as q increases.

V. CONCLUSIONS

For an underlay D2D system with multiple subchannels and D2D pairs, we proposed a novel polynomial-time algorithm called RPA. It assigned up to K D2D pairs per subchannel, where K was optimized to control the trade-off between spatial reuse and inter-D2D interference. RPA was designed for a low signaling overhead model in which the DRx fed back to the BS a quantized rate whose outage probability was below a pre-specified value. RPA also guaranteed a minimum rate with a pre-specified outage probability for the CUs.

RPA provably achieved at least half of the optimal D2D sum rate of the exponentially-complex exhaustive search. This

¹These correspond to a fading-averaged signal-to-noise ratio of 7.2 dB for a CU at the cell edge and 18 dB for a D2D pair with DTx-to-DRx distance of 50 m.

was unlike conventional algorithms, which either provided no performance guarantees or had a weaker guarantee. Numerically, the sum throughput of RPA was higher than that of conventional algorithms and indistinguishable from that of exhaustive search. An interesting avenue for future work is jointly optimizing subchannel and power allocation.

APPENDIX

A. Proof of Lemma 1

The following two cases arise:

1) When D2D Pair k_1 is Assigned to Subchannel i: This assignment satisfies (12) because D2D pair k_1 can be considered for assignment only if $w_{ik_1} \leq b_i$. Else, Step 1 of RPA would have set x_{ik_1} to 0.

2) When D2D Pairs k_2, \ldots, k_{n_i} are Assigned to Subchannel *i*: We first express the fractional solution $\{\tilde{x}_{ij}, j \in \mathcal{D}'_i\}$ of subchannel *i* in terms of its virtual subchannels i_1, \ldots, i_{n_i} in the bipartite graph as illustrated in Fig. 2. For i_1 , define $\tilde{x}_{i_1[1]} \triangleq$ $\tilde{x}_{i[1]}, \tilde{x}_{i_1[2]} \triangleq \tilde{x}_{i[2]}, \ldots, \tilde{x}_{i_1[j_1-1]} \triangleq \tilde{x}_{i[j_1-1]}$. The term $\tilde{x}_{i[j_1]}$ is split into $\tilde{x}_{i_1[j_1]}$ and $\tilde{x}_{i_2[j_1]}$ such that $\sum_{j=1}^{j_1} \tilde{x}_{i_1[j]} = 1$ and $\tilde{x}_{i_1[j_1]} + \tilde{x}_{i_2[j_1]} = \tilde{x}_{i[j_1]}$. For i_2 , in addition to $\tilde{x}_{i_2[j_1]}$, define $\tilde{x}_{i_2[j_1+1]} \triangleq \tilde{x}_{i[j_1+1]}, \ldots, \tilde{x}_{i_2[j_2-1]} \triangleq \tilde{x}_{i[j_2-1]}$. The term $\tilde{x}_{i[j_2]}$ is split into $\tilde{x}_{i_2[j_2]}$ and $\tilde{x}_{i_3[j_2]}$ such that $\sum_{j=j_1}^{j_2} \tilde{x}_{i_2[j]} = 1$ and $\tilde{x}_{i_2[j_2]} + \tilde{x}_{i_3[j_2]} = \tilde{x}_{i[j_2]}$. Note that if $\sum_{j=1}^{j_1} \tilde{x}_{i[j]} = 1$, then $\tilde{x}_{i_2[j_1]} = 0$. Also, if $\sum_{j=1}^{j_2} \tilde{x}_{i[j]} = 2$, then $\tilde{x}_{i_3[j_2]} = 0$.

In general, for $k = 1, 2, ..., n_i$, define $\tilde{x}_{i_k[j]} \triangleq \tilde{x}_{i[j]}$ for $j = j_{k-1} + 1, j_{k-1} + 2, ..., j_k - 1$, and $\tilde{x}_{i[j_k]}$ is split into $\tilde{x}_{i_k[j_k]}$ and $\tilde{x}_{i_{k+1}[j_k]}$ such that $\sum_{j=j_{k-1}}^{j_k} \tilde{x}_{i_k[j]} = 1$ and $\tilde{x}_{i_k[j_k]} + \tilde{x}_{i_{k+1}[j_k]} = \tilde{x}_{i[j_k]}$, where $j_0 \triangleq 1$ and $\tilde{x}_{i_1[j_0]} \triangleq \tilde{x}_{i[j_0]}$. Also, if $\sum_{j=1}^{j_k} \tilde{x}_{i[j]} = k$, then $\tilde{x}_{i_{k+1}[j_k]} = 0$.

For each subchannel *i*, the linear program in Step 1 of RPA satisfies (12). Hence, $b_i \geq \sum_{j=1}^{|\mathcal{D}'_i|} \tilde{x}_{i[j]} w_{i[j]}$. In terms of the notation above, $\sum_{j=1}^{|\mathcal{D}'_i|} \tilde{x}_{i[j]} w_{i[j]}$ can be expressed as

$$\sum_{j=1}^{|\mathcal{D}'_i|} \tilde{x}_{i[j]} w_{i[j]} = \sum_{k=1}^{n_i-1} \sum_{j=j_{k-1}}^{j_k} \tilde{x}_{i_k[j]} w_{i[j]} + \sum_{j=j_{n_i-1}}^{|\mathcal{D}'_i|} \tilde{x}_{i_{n_i}[j]} w_{i[j]},$$

$$\geq \sum_{k=1}^{n_i-1} \sum_{j=j_{k-1}}^{j_k} \tilde{x}_{i_k[j]} w_{i[j]}.$$
(20)

Since $w_{i[j_{k-1}]} \ge w_{i[j_{k-1}+1]} \ge \cdots \ge w_{i[j_k]}$, it follows that

$$\sum_{k=1}^{n_i-1} \sum_{j=j_{k-1}}^{j_k} \tilde{x}_{i_k[j]} w_{i[j]} \ge \sum_{k=1}^{n_i-1} \left(\sum_{j=j_{k-1}}^{j_k} \tilde{x}_{i_k[j]} \right) w_{i[j_k]} = \sum_{k=1}^{n_i-1} w_{i[j_k]},$$
(21)

where the last equality follows because $\sum_{j=j_{k-1}}^{j_k} \tilde{x}_{i_k[j]} = 1$, for $k = 1, ..., n_i - 1$. From (20) and (21), we get

$$b_i \ge w_{i[j_1]} + w_{i[j_2]} + \dots + w_{i[j_{n_i-1}]}.$$
(22)

The virtual subchannel i_2 can be assigned to at most one D2D pair among the pairs $[j_1], [j_1 + 1], \ldots, [j_2]$, and the maximum interference possible is $w_{i[j_1]}$. Hence, $w_{i[j_1]} \ge$ w_{ik_2} . In general, we can show that $w_{i[j_1]} \ge w_{ik_2}, w_{i[j_2]} \ge$ $w_{ik_3}, \ldots, w_{i[j_{n_i-1}]} \ge w_{ik_{n_i}}$. Thus, from (22), we get $b_i \ge$ $w_{ik_2} + w_{ik_3} + \cdots + w_{ik_{n_i}}$. Hence, the set $\{k_2, k_3, \ldots, k_{n_i}\}$ of D2D pairs assigned to subchannel *i* satisfies (12).

B. Proof of Result 1

Let Z_{opt} and Z_{frac} be the optimal D2D sum rates for \mathcal{P} and the fractional solution of Step 1, respectively. Since the fractional solution is obtained by relaxing the integer constraint in (14), $Z_{frac} \geq Z_{opt}$. Let Z_{match} be the D2D sum rate obtained by the integral matching solution in Step 3. The integral matching solution ignores the constraint in (12). Hence, $Z_{match} \geq Z_{frac}$. After pruning the infeasible assignment for subchannel *i* in Step 4, the D2D sum rate achieved is max $\{C_{ik_1}, C_{ik_2} + \cdots + C_{ik_{n_i}}\} \geq \frac{1}{2} (C_{ik_1} + C_{ik_2} + \cdots + C_{ik_{n_i}})$. Summing over all subchannels, the D2D sum rate Z_{final} of the final allocation satisfies $Z_{final} \geq \frac{1}{2}Z_{match}$. From the above inequalities, we get $Z_{final} \geq \frac{1}{2}Z_{opt}$.

REFERENCES

- [1] R. I. Ansari, C. Chrysostomou, S. A. Hassan, M. Guizani, S. Mumtaz, J. Rodriguez, and J. J. P. C. Rodrigues, "5G D2D networks: Techniques, challenges, and future prospects," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3970–3984, Dec. 2018.
- [2] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, May 2014.
- [3] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 40–48, Apr. 2014.
- [4] D. Feng, L. Lu, Y. W. Yi, G. Y. Li, G. Feng, and S. Li, "Deviceto-device communications underlaying cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3541–3551, Aug. 2013.
- [5] T. D. Hoang, L. Le, and T. Le-Ngoc, "Resource allocation for D2D communication underlaid cellular networks using graph-based approach," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 7099–7113, Oct. 2016.
- [6] W. Zhao and S. Wang, "Resource sharing scheme for device-to-device communication underlaying cellular networks," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4838–4848, Dec. 2015.
- [7] A. Abdallah, M. M. Mansour, and A. Chehab, "Power control and channel allocation for D2D underlaid cellular networks," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 3217–3234, Jul. 2018.
- [8] L. Wang, H. Tang, H. Wu, and G. L. Stüber, "Resource allocation for D2D communications underlay in Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1159–1170, Feb. 2017.
- [9] R. Wang, J. Zhang, S. H. Song, and K. B. Letaief, "Optimal QoS-aware channel assignment in D2D communications with partial CSI," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7594–7609, Nov. 2016.
- [10] B. V. R. Gorantla and N. B. Mehta, "Resource and computationally efficient subchannel allocation for D2D in multi-cell scenarios with partial and asymmetric CSI," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5806–5817, Dec. 2019.
- [11] D. B. Shmoys and E. Tardos, "An approximation algorithm for the generalized assignment problem," *Math. Program.*, vol. 62, no. 3, pp. 461–474, Dec. 1993.
- [12] G. L. Stüber, *Principles of Mobile Communications*, 2nd ed. Springer, 2009.
- [13] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 3rd ed. Springer, 2008.
- [14] D. B. West, *Introduction to Graph Theory*, 2nd ed. Prentice Hall, 2000.
- [15] "Technical specification group radio access network; study on channel model for frequencies from 0.5 to 100 GHz (release 15)," 3rd Generation Partnership Project (3GPP), Tech. Rep. 38.901 (v15.0.0), Jun. 2018.