

Robust Transmission Design for Multi-Cell D2D Underlaid Cellular Networks

Hao Xu, Gordon L. Stüber, *Fellow, IEEE*, Wei Xu, *Senior Member, IEEE*, Cunhua Pan, Jianfeng Shi, Zhaohui Yang, and Ming Chen

Abstract—This paper investigates the robust transmission design (RTD) of a multi-cell device-to-device (D2D) underlaid cellular network with imperfect channel state information (CSI). The bounded model is adopted to characterize the CSI impairment and the aim is to maximize the worst-case sum rate of the system. To protect cellular communications, it is assumed that the interference from all D2D transmitters to each base station (BS) is power-limited. It is first shown that the worst-case signal-to-interference-plus-noise ratio (SINR) of each D2D link can be obtained directly, while that of cellular links cannot be similarly found since the channel estimation error vectors of cellular links are coupled in the SINR expressions. To solve the nonconvex problem, the objective function of the original problem is replaced with its lower bound, and the resulted problem is decomposed into multiple semidefinite programming (SDP) subproblems which are convex and have computationally efficient solutions. An iterative RTD algorithm is then proposed to obtain a suboptimal solution. Simulation results show that D2D communication can significantly increase the performance of the conventional cellular systems while causing tolerable interference to cellular users. In addition, the proposed RTD algorithm outperforms the conventional non-robust transmission design greatly in terms of network spectral efficiency.

I. INTRODUCTION

With the daily increasing of wireless communication demand, the problem of spectrum insufficiency has become a major factor limiting the wireless system performance [1], [2]. Device-to-device (D2D) communication is a promising method for enhancing the spectral efficiency (SE) of traditional cellular systems and has drawn great attention recently [3]–[6]. Different from the conventional cellular communication where

all traffic is routed via base stations (BSs), D2D communication allows two closely located users to communicate directly and, thus, has distinct advantages such as high SE, short packet delay, low energy consumption and increased safety [7], [8]. In a D2D underlaid cellular network, D2D users (DUs) reuse the resource blocks (RBs) of cellular users (CUs) for communication, leading to cochannel interference. Therefore, efficient resource allocation and power control algorithms play an important role in reaping the potential benefits of D2D communication.

Thus far, there have been a flurry of works studying interference mitigation and sum SE maximization problems of D2D underlaid systems [9]–[15]. Reference [9] considered a multiple-input multiple-output (MIMO) D2D underlaid system and aimed to maximize the sum SE of DUs by optimizing the precoding matrix. However, the quality-of-service (QoS) of CUs was not guaranteed. Both [10] and [11] considered resource allocation and power control problems to maximize the sum SE of DUs with all CUs protected by the QoS constraints. It was shown in [10] and [11] that the minimum QoS constraints for cellular links always hold with equality under the optimal conditions. Since CUs usually possess higher priorities compared to DUs, this can be unfair for CUs. In [12]–[15], the problem of maximizing the sum rate of both cellular and D2D links was studied, and the minimum QoS constraints of CUs were imposed to protect cellular communication.

References [9]–[15] all considered a simple single-cell scenario, while ignoring the cumulated interference from neighbor cells. According to our survey, the performance of multi-cell D2D underlaid systems has less well been studied [16]–[18]. In [16], the authors established a tractable model for multi-cell D2D underlaid cellular networks and adopted Exclusion Regions around the BSs to mitigate cochannel interference. In [17], the subcarrier allocation problem for a multi-cell D2D underlaid system was characterized as a potential game, and an iterative algorithm was proposed to obtain a Nash equilibrium. Reference [18] considered a multi-cell D2D underlaid massive MIMO system, and investigated the SE of cellular as well as D2D links under both perfect and imperfect channel state information (CSI).

Most of the aforementioned works assumed that perfect CSI was available for system performance analysis and optimization [9]–[17]. However, in practice, it is difficult to obtain perfect CSI of all links due to channel estimation errors and quantization errors, especially in multi-cell systems. Therefore, enhancing the robustness of network performance under partial

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org. This work was in part supported by the NSFC (Nos. 61372106, 61471114, & 61221002), NSTMP under 2016ZX03001016-003, the Six Talent Peaks project in Jiangsu Province under GDZB-005, Science and Technology Project of Guangdong Province under Grant 2014B010119001, the Scholarship from the China Scholarship Council (No. 201606090039), Program Sponsored for Scientific Innovation Research of College Graduate in Jiangsu Province under Grant KYLX16_0221, and the Scientific Research Foundation of Graduate School of Southeast University under Grant YBJJ1651. (*Corresponding author: Hao Xu; Wei Xu.*)

H. Xu, W. Xu, J. Shi, Z. Yang and M. Chen are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: xuhao2013@seu.edu.cn; wxu@seu.edu.cn; shijianfeng@seu.edu.cn; yangzhaohui@seu.edu.cn; chenming@seu.edu.cn).

G. L. Stüber is with the Faculty of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: stuber@ece.gatech.edu).

C. Pan is with School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: c.pan@qmul.ac.uk).

or imperfect CSI has become an important issue [19]–[24]. According to [19] and [21], imperfect CSI can be described by two approaches: probability model and bounded model. When the CSI impairment is dominated by channel estimation errors, the probability model applies. On the other hand, the bounded model is more suitable when quantization errors are the dominant source of CSI uncertainty. In [19] and [20], transceiver designs were investigated under the probability model for CSI impairment, while in [21]–[24], the bounded CSI impairment model was adopted. Specifically, [21] studied the mean-squared error (MSE) optimization problem for a single-cell MIMO system, [22], [23] aimed to maximize the worst-case sum SE of a multi-cell network via robust beamforming, and an extended worst-case sum energy efficiency (EE) maximization problem was investigated in [24].

All references [19]–[24] considered robust transmission design (RTD) problems for the conventional cellular networks. According to our survey, only a few works have studied the RTD problem for single-cell D2D underlaid cellular systems [25]–[28]. In [25] and [26], the probability model was adopted to characterize the CSI impairment. Specifically, [25] aimed to maximize the signal-to-interference-plus-noise ratio (SINR) of the D2D link while guaranteeing the outage probability of the cellular link not exceeding a threshold. However, the considered network was simply composed of a cellular link and a D2D link. In [26], a network with one single-antenna CU and several single-antenna D2D pairs was considered. Centralized and distributed algorithms were proposed to maximize the coverage probability. In [27] and [28], the bounded CSI impairment model was adopted. In particular, [27] considered a similar simple network as [25] and designed a null-space based robust interference avoiding strategy. Reference [28] considered a D2D underlaid system with one multi-antenna CU and several multi-antenna D2D pairs, and aimed for SINR fairness among D2D users by designing robust transceivers.

To the best of the authors' knowledge, the RTD for a multi-cell D2D underlaid system with multiple CUs and multiple D2D pairs in each cell has not been studied. Hence, it is considered in this paper. In order to increase system SE, it is assumed that all BSs are equipped with multiple antennas [29], [30]. Since uplink spectrum is often underutilized comparing to that of downlink spectrum in cellular systems [31], [32], uplink resource sharing is thus assumed. According to [33], [34], in the next-generation cellular communication system with advanced estimation schemes, quantization errors will become the main source of CSI uncertainties. Therefore, the bounded model is adopted to characterize the CSI impairment. The main contributions of this paper are summarized as follows:

- Different from [25], [27] and [28], which focused on optimizing the performance of D2D links, considering the higher priorities of CUs, this paper aims to maximize the worst-case sum rate of both cellular and D2D links under CSI impairment, and the interference from all D2D transmitters (D2D-Txs) to each BS is assumed to be power-limited to protect cellular communication. Since a multi-cell network is considered with each cell consisting of a multi-antenna BS, multiple CUs and D2D pairs, joint

optimization of the transmit powers of all transmitters and receive filters of all BSs is required, which is more complex. Hence, the algorithms developed in the existing literature cannot be applied directly.

- To solve the formulated non-convex problem, the objective function is first analyzed. It is shown that the worst-case SINR of each D2D link can be obtained by independently finding the worst-case (smallest) numerator term and the worst-case (largest) denominator terms, while that of cellular links cannot be analogously found since channel estimation error vectors of cellular links are coupled in SINR expressions. Hence, a precise expression of the objective function cannot be obtained, and it is difficult to solve the original problem.
- In order to make the complicated problem tractable, a lower bound on the original objective function is derived by applying the relationship between the minimum mean square error (MMSE) and the SINR of received signals. This lower bound is then maximized subject to the same constraints instead of solving the primal problem. By using the definition of MSE and reformulating the objective function, the resulted problem can be decomposed into multiple semidefinite programming (SDP) subproblems, which are convex and have computationally efficient solutions. An iterative RTD algorithm is then proposed to obtain a suboptimal solution.
- In the simulation part, the performance of the proposed RTD algorithm is illustrated and compared in terms of sum SE. It is shown that D2D communication can significantly increase the performance of the conventional cellular system while causing tolerable interference to CUs. In addition, compared with the non-robust transmission design scheme, which takes the estimated channel as the true channel, the sum SE of the system can be greatly increased by the proposed RTD algorithm.

Note that for the sake of simplicity, it is assumed that the RB allocation, i.e., the matching between CUs and DUs, has been predetermined. This paper mainly focuses on the RTD for mobile equipments using the same RB. Obviously, the results obtained in this paper can be applied straightforwardly to a D2D underlaid massive MIMO system where all users use the same RB for transmission [18], [35].

The rest of this paper is organized as follows. In Section II, a multi-cell D2D underlaid cellular system and problem formulation are presented. In Section III, the considered sum SE maximization problem under CSI impairment is first analyzed and an RTD algorithm is then proposed to solve the problem. Finally, numerical verifications are presented in Section IV before concluding remarks in Section V.

This paper follows commonly used notations. \mathbb{R} and \mathbb{C} denote the real space and the complex space, respectively. The boldface upper (lower) case letters are used to denote matrices (vectors). \mathbf{I}_M stands for the $M \times M$ dimensional identity matrix and $\mathbf{0}$ denotes the all-zero vector or matrix. “ \setminus ” represents the set subtraction operation. Superscript $(\cdot)^H$ denotes the conjugated-transpose operation and $\mathbb{E}\{\cdot\}$ denotes the expectation operation. $\|a\|$ is used to denote the Euclidean norm of vector a .

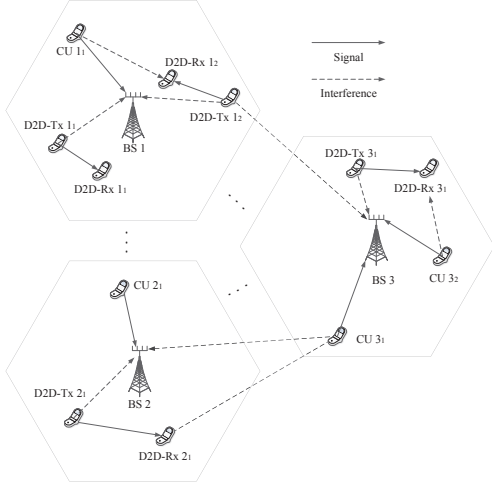


Fig. 1. An example of the considered system diagram.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

As illustrated in Fig. 1, this paper considers an L -cell interference network with a BS, M CUs and N D2D pairs in each cell while sharing the same uplink RB for transmission.¹ To increase the system SE, each BS is assumed to have B antennas to exploit higher spatial degrees of freedom. Each mobile user is equipped with one antenna. For the sake of clarity, only part of the interference signals are depicted in Fig. 1. In order to simultaneously communicate with the CUs in each cell, it is assumed that the number of antennas at each BS is not less than that of CUs in the corresponding cell, i.e., $B \geq M$. Let \mathcal{L} , \mathcal{C} and \mathcal{D} represent the sets of all BSs, CUs and D2D pairs, respectively. Denote the m th cellular user in cell l by CU l_m and the n th D2D pair in cell l by l_n . Then, the $B \times 1$ dimensional received data vector of BS l can be written as

$$\mathbf{y}_l^{(c)} = \sum_{k_t \in \mathcal{C}} \sqrt{p_{k_t}^{(c)}} \mathbf{h}_{k_t, l}^{(c)} x_{k_t}^{(c)} + \sum_{k_s \in \mathcal{D}} \sqrt{p_{k_s}^{(d)}} \mathbf{h}_{k_s, l}^{(d)} x_{k_s}^{(d)} + \mathbf{z}_l^{(c)}, \quad (1)$$

where $p_{k_t}^{(c)}$ and $x_{k_t}^{(c)}$ denote the transmit power and the zero-mean unit-variance data symbol of CU k_t , respectively. $\mathbf{h}_{k_t, l}^{(c)} \in \mathbb{C}^{B \times 1}$ is the channel vector from CU k_t to BS l , which accounts for large-scale fading (including path loss and shadow fading) and small-scale fading. $p_{k_s}^{(d)}$, $x_{k_s}^{(d)}$ and $\mathbf{h}_{k_s, l}^{(d)}$ are similarly defined for D2D-Tx k_s . $\mathbf{z}_l^{(c)} \in \mathbb{C}^{B \times 1}$ is the zero-mean circularly symmetric complex white Gaussian noise with covariance $N_0 \mathbf{I}_B$, i.e., $\mathbf{z}_l^{(c)} \sim \mathcal{CN}(0, N_0 \mathbf{I}_B)$. Similarly, the received signal of D2D receiver (D2D-Rx) l_n is given by

$$\mathbf{y}_{l_n}^{(d)} = \sum_{k_t \in \mathcal{C}} \sqrt{p_{k_t}^{(c)}} g_{k_t, l_n}^{(c)} x_{k_t}^{(c)} + \sum_{k_s \in \mathcal{D}} \sqrt{p_{k_s}^{(d)}} g_{k_s, l_n}^{(d)} x_{k_s}^{(d)} + \mathbf{z}_{l_n}^{(d)}, \quad (2)$$

where $g_{k_t, l_n}^{(c)}$, $g_{k_s, l_n}^{(d)}$ respectively denote the channel coefficient from CU k_t and D2D-Tx k_s to D2D-Rx l_n , and $\mathbf{z}_{l_n}^{(d)}$ is the

zero-mean circularly symmetric complex white Gaussian noise with variance N_0 .

Let $\mathbf{w}_{l_m} \in \mathbb{C}^{B \times 1}$ denote the unit-norm receive beamforming vector adopted by BS l for detecting signal $x_{l_m}^{(c)}$. Then, according to (1) and [36], the post-processing SINR of CU l_m can be written as

$$\text{SINR}_{l_m}^{(c)} = \frac{p_{l_m}^{(c)} \left| \mathbf{w}_{l_m}^H \mathbf{h}_{l_m, l}^{(c)} \right|^2}{\mathbf{w}_{l_m}^H \mathbf{G}_{l_m} \mathbf{w}_{l_m}}, \quad (3)$$

where \mathbf{G}_{l_m} denotes the interference plus noise covariance matrix and is give by

$$\begin{aligned} \mathbf{G}_{l_m} &= \mathbb{E} \left\{ \left(\mathbf{y}_l^{(c)} - \sqrt{p_{l_m}^{(c)}} \mathbf{h}_{l_m, l}^{(c)} x_{l_m}^{(c)} \right)^H \left(\mathbf{y}_l^{(c)} - \sqrt{p_{l_m}^{(c)}} \mathbf{h}_{l_m, l}^{(c)} x_{l_m}^{(c)} \right) \right\} \\ &= \sum_{k_t \in \mathcal{C} \setminus l_m} p_{k_t}^{(c)} \mathbf{h}_{k_t, l}^{(c)} \left(\mathbf{h}_{k_t, l}^{(c)} \right)^H + \sum_{k_s \in \mathcal{D}} p_{k_s}^{(d)} \mathbf{h}_{k_s, l}^{(d)} \left(\mathbf{h}_{k_s, l}^{(d)} \right)^H + N_0 \mathbf{I}_B. \end{aligned} \quad (4)$$

The second equality of (4) holds because it is assumed that each transmitter independently sends zero-mean unit-variance data symbols. Analogously, from (2), the received SINR of D2D-Rx l_n can be expressed as

$$\text{SINR}_{l_n}^{(d)} = \frac{p_{l_n}^{(d)} \left| g_{l_n, l_n}^{(d)} \right|^2}{\sum_{k_t \in \mathcal{C}} p_{k_t}^{(c)} \left| g_{k_t, l_n}^{(c)} \right|^2 + \sum_{k_s \in \mathcal{D} \setminus l_n} p_{k_s}^{(d)} \left| g_{k_s, l_n}^{(d)} \right|^2 + N_0}. \quad (5)$$

As mentioned in Section I, it is difficult for the BSs to obtain the perfect CSI of all channels, and with advanced estimation schemes, quantization errors will become the main source of CSI uncertainties in the next-generation cellular communication system [33], [34]. Therefore, it is assumed that the BSs only have partial CSI, and the bounded model is adopted to characterize the CSI impairment. Denote the imperfect CSI of $\mathbf{h}_{k_t, l}^{(c)}$, $\mathbf{h}_{k_s, l}^{(d)}$, $g_{k_t, l_n}^{(c)}$ and $g_{k_s, l_n}^{(d)}$ by $\tilde{\mathbf{h}}_{k_t, l}^{(c)}$, $\tilde{\mathbf{h}}_{k_s, l}^{(d)}$, $\tilde{g}_{k_t, l_n}^{(c)}$ and $\tilde{g}_{k_s, l_n}^{(d)}$, and the corresponding CSI errors by $\Delta_{k_t, l}^{(c)}$, $\Delta_{k_s, l}^{(d)}$, $\delta_{k_t, l_n}^{(c)}$ and $\delta_{k_s, l_n}^{(d)}$, i.e.,

$$\begin{aligned} \Delta_{k_t, l}^{(c)} &= \mathbf{h}_{k_t, l}^{(c)} - \tilde{\mathbf{h}}_{k_t, l}^{(c)}, \quad \forall k_t \in \mathcal{C}, l \in \mathcal{L}, \\ \Delta_{k_s, l}^{(d)} &= \mathbf{h}_{k_s, l}^{(d)} - \tilde{\mathbf{h}}_{k_s, l}^{(d)}, \quad \forall k_s \in \mathcal{D}, l \in \mathcal{L}, \\ \delta_{k_t, l_n}^{(c)} &= g_{k_t, l_n}^{(c)} - \tilde{g}_{k_t, l_n}^{(c)}, \quad \forall k_t \in \mathcal{C}, l_n \in \mathcal{D}, \\ \delta_{k_s, l_n}^{(d)} &= g_{k_s, l_n}^{(d)} - \tilde{g}_{k_s, l_n}^{(d)}, \quad \forall k_s \in \mathcal{D}, l_n \in \mathcal{D}. \end{aligned} \quad (6)$$

For the bounded model, the errors in (6) satisfy

$$\left\| \Delta_{k_t, l}^{(c)} \right\| \leq \epsilon_{k_t, l}^{(c)}, \quad \forall k_t \in \mathcal{C}, l \in \mathcal{L}, \quad (7a)$$

$$\left\| \Delta_{k_s, l}^{(d)} \right\| \leq \epsilon_{k_s, l}^{(d)}, \quad \forall k_s \in \mathcal{D}, l \in \mathcal{L}, \quad (7b)$$

$$\left| \delta_{k_t, l_n}^{(c)} \right| \leq \epsilon_{k_t, l_n}^{(c)}, \quad \forall k_t \in \mathcal{C}, l_n \in \mathcal{D}, \quad (7c)$$

$$\left| \delta_{k_s, l_n}^{(d)} \right| \leq \epsilon_{k_s, l_n}^{(d)}, \quad \forall k_s \in \mathcal{D}, l_n \in \mathcal{D}, \quad (7d)$$

where $\epsilon_{k_t, l}^{(c)}$, $\epsilon_{k_s, l}^{(d)}$, $\epsilon_{k_t, l_n}^{(c)}$ and $\epsilon_{k_s, l_n}^{(d)}$ represent the corresponding CSI error bounds.

¹Note that the number of CUs and D2D pairs in each cell can be different although the same M and N are used for notational brevity.

B. Problem Formulation

This paper aims to maximize the worst-case sum SE of the cooperative multi-cell network under imperfect CSI. From (3) and (5), the network throughput can be expressed as

$$\begin{aligned} R &= \sum_{l_m \in \mathcal{C}} R_{l_m}^{(c)} + \sum_{l_n \in \mathcal{D}} R_{l_n}^{(d)} \\ &= \sum_{l_m \in \mathcal{C}} \log_2 \left(1 + \text{SINR}_{l_m}^{(c)} \right) + \sum_{l_n \in \mathcal{D}} \log_2 \left(1 + \text{SINR}_{l_n}^{(d)} \right), \end{aligned} \quad (8)$$

where $R_{l_m}^{(c)}$ and $R_{l_n}^{(d)}$ respectively denote the throughput of cellular link l_m and D2D link l_n . In order to guarantee the QoS of cellular links, assume that the interference signal from all D2D-Txs to each BS is power-limited [37], i.e.,

$$\sum_{k_s \in \mathcal{D}} p_{k_s}^{(d)} \left\| \mathbf{h}_{k_s, l}^{(d)} \right\|^2 \leq a_l, \quad \forall l \in \mathcal{L}, \quad (9)$$

where a_l is the maximum interference threshold to protect CUs in the l th cell. Since the imperfect CSI is considered, an accurate value of the left-hand side term of (9) becomes intractable. Therefore, (9) is rewritten in a stricter form in the sequel. Beforehand, Lemma 1 is first given which is a simple extension of a result in [38] and can be readily proven.

Lemma 1: For any given $\tilde{\mathbf{h}} \in \mathbb{C}^{B \times 1}$, $\varepsilon \in R_+$ and the uncertainty region $\Omega = \{ \Delta | \Delta \in \mathbb{C}^{B \times 1}, \|\Delta\| \leq \varepsilon \}$, the following results hold

$$\begin{aligned} \min_{\Delta \in \Omega} \left\| \tilde{\mathbf{h}} + \Delta \right\|^2 &= \left[\left(\left\| \tilde{\mathbf{h}} \right\| - \varepsilon \right)^+ \right]^2, \\ \max_{\Delta \in \Omega} \left\| \tilde{\mathbf{h}} + \Delta \right\|^2 &= \left(\left\| \tilde{\mathbf{h}} \right\| + \varepsilon \right)^2, \end{aligned} \quad (10)$$

whose optimal solutions are respectively given by

$$\begin{aligned} \arg \min_{\Delta \in \Omega} \left\| \tilde{\mathbf{h}} + \Delta \right\|^2 &= -\tilde{\mathbf{h}} \times \min \left\{ 1, \frac{\varepsilon}{\left\| \tilde{\mathbf{h}} \right\|} \right\}, \\ \arg \max_{\Delta \in \Omega} \left\| \tilde{\mathbf{h}} + \Delta \right\|^2 &= \frac{\varepsilon \tilde{\mathbf{h}}}{\left\| \tilde{\mathbf{h}} \right\|}. \end{aligned} \quad (11)$$

Now by recalling (6) and (7), and invoking the result of the lemma above, a stricter form of (9) is given by

$$\sum_{k_s \in \mathcal{D}} p_{k_s}^{(d)} \rho_{k_s, l} \leq a_l, \quad \forall l \in \mathcal{L}, \quad (12)$$

where $\rho_{k_s, l} = \left(\left\| \tilde{\mathbf{h}}_{k_s, l}^{(d)} \right\| + \varepsilon_{k_s, l}^{(d)} \right)^2$.

Let Δ and δ respectively denote the sets of CSI errors from all transmitters to all BSs and D2D-Rxs, i.e.,

$$\begin{aligned} \Delta &= [\Delta_1, \dots, \Delta_L], \\ \delta &= [\delta_{1_1}, \dots, \delta_{1_N}, \dots, \delta_{L_N}], \end{aligned} \quad (13)$$

where

$$\begin{aligned} \Delta_l &= \left[\Delta_{1_1, l}^{(c)}, \dots, \Delta_{1_M, l}^{(c)}, \dots, \Delta_{L_M, l}^{(c)}, \Delta_{1_1, l}^{(d)}, \dots, \right. \\ &\quad \left. \Delta_{1_N, l}^{(d)}, \dots, \Delta_{L_N, l}^{(d)} \right], \quad \forall l \in \mathcal{L}, \\ \delta_{l_n} &= \left(\delta_{1_1, l_n}^{(c)}, \dots, \delta_{1_M, l_n}^{(c)}, \dots, \delta_{L_M, l_n}^{(c)}, \delta_{1_1, l_n}^{(d)}, \dots, \right. \\ &\quad \left. \delta_{1_N, l_n}^{(d)}, \dots, \delta_{L_N, l_n}^{(d)} \right)^T, \quad \forall l_n \in \mathcal{D}. \end{aligned} \quad (14)$$

Since CSI errors are unknown and norm-bounded by (7), there always exists a worst case (denote the corresponding CSI errors as Δ^* and δ^*) such that Δ^* and δ^* satisfy (7), and the system outputs the worst-case SE R^* . Then, the worst-case SE maximization problem can be formulated as

$$\max_{\mathbf{W}, \mathbf{p}} R^* \quad (15a)$$

$$\text{s.t. } 0 \leq p_{l_m}^{(c)} \leq P_{l_m}^{(c)}, \quad \forall l_m \in \mathcal{C}, \quad (15b)$$

$$0 \leq p_{l_n}^{(d)} \leq P_{l_n}^{(d)}, \quad \forall l_n \in \mathcal{D}, \quad (15c)$$

$$\sum_{k_s \in \mathcal{D}} p_{k_s}^{(d)} \rho_{k_s, l} \leq a_l, \quad \forall l \in \mathcal{L}, \quad (15d)$$

where $P_{l_m}^{(c)}$ and $P_{l_n}^{(d)}$ denote the maximum transmit power of CU l_m and D2D-Tx l_n , respectively. \mathbf{W} and \mathbf{p} are defined as follows

$$\begin{aligned} \mathbf{W} &= [\mathbf{W}_1, \dots, \mathbf{W}_L], \\ \mathbf{p} &= \left[\mathbf{p}^{(c)}; \mathbf{p}^{(d)} \right], \end{aligned} \quad (16)$$

where

$$\begin{aligned} \mathbf{W}_l &= [\mathbf{w}_{l_1}, \dots, \mathbf{w}_{l_M}], \quad \forall l \in \mathcal{L}, \\ \mathbf{p}^{(c)} &= \left(p_{1_1}^{(c)}, \dots, p_{1_M}^{(c)}, \dots, p_{L_M}^{(c)} \right)^T, \\ \mathbf{p}^{(d)} &= \left(p_{1_1}^{(d)}, \dots, p_{1_N}^{(d)}, \dots, p_{L_N}^{(d)} \right)^T. \end{aligned} \quad (17)$$

Note that the actual CSI errors are not variables and can not be optimized. However, by using the transmit power vector and receive beamforming vectors designed for the worst case, a better system throughput can be obtained than the nonrobust design scheme, which simply takes the estimated channel as the true channel. Due to the following two challenges, it is difficult to directly solve problem (15): First, CSI error vectors are coupled in the SINR calculation, making it hard to obtain an explicit expression of the objective function R^* ; Second, the fractional form of the SINR expressions and the $\log(\cdot)$ operation make the objective function nonconvex. To handle this problem, an alternative algorithm will be proposed in the following section.

III. PROBLEM ANALYSIS AND ROBUST OPTIMIZATION

In this section, the worst-case SE maximization problem (15) is investigated. As stated above, it is difficult to directly solve (15). Therefore, in the following of this section, problem (15) is first analyzed and transformed to a more tractable form. Then, an alternative algorithm is proposed to solve it.

A. Problem Analysis

From (5), it can be found that each channel coefficient from a mobile transmitter to a D2D-Rx appears only once in either the numerator or denominator of $\text{SINR}_{l_n}^{(d)}$, $\forall l_n \in \mathcal{D}$, and all channel coefficients $g_{k_t, l_n}^{(c)}, g_{k_s, l_n}^{(d)}$, $\forall k_t \in \mathcal{C}, k_s, l_n \in \mathcal{D}$ are independent of each other. As a result, the worst-case $\text{SINR}_{l_n}^{(d)}$ can be obtained by decoupling it into finding the worst-case (smallest) numerator term and the worst-case (largest) denominator terms, which can be implemented based on Lemma 1.

Define the worst-case $\text{SINR}_{l_n}^{(d)}$ by $\text{sinr}_{l_n}^{(d)} \triangleq \min_{\delta_{l_n}} \text{SINR}_{l_n}^{(d)}$. It follows that

$$\begin{aligned} \text{sinr}_{l_n}^{(d)} &= \min_{\delta_{l_n}} \frac{p_{l_n}^{(d)} |g_{l_n, l_n}^{(d)}|^2}{\sum_{k_t \in \mathcal{C}} p_{k_t}^{(c)} |g_{k_t, l_n}^{(c)}|^2 + \sum_{k_s \in \mathcal{D} \setminus l_n} p_{k_s}^{(d)} |g_{k_s, l_n}^{(d)}|^2 + N_0} \\ &= \frac{p_{l_n}^{(d)} \min_{\delta_{l_n, l_n}^{(d)}} |g_{l_n, l_n}^{(d)}|^2}{\sum_{k_t \in \mathcal{C}} p_{k_t}^{(c)} \max_{\delta_{k_t, l_n}^{(c)}} |g_{k_t, l_n}^{(c)}|^2 + \sum_{k_s \in \mathcal{D} \setminus l_n} p_{k_s}^{(d)} \max_{\delta_{k_s, l_n}^{(d)}} |g_{k_s, l_n}^{(d)}|^2 + N_0} \\ &= \frac{p_{l_n}^{(d)} |\bar{g}_{l_n, l_n}^{(d)}|^2}{\sum_{k_t \in \mathcal{C}} p_{k_t}^{(c)} |\bar{g}_{k_t, l_n}^{(c)}|^2 + \sum_{k_s \in \mathcal{D} \setminus l_n} p_{k_s}^{(d)} |\bar{g}_{k_s, l_n}^{(d)}|^2 + N_0}, \end{aligned} \quad (18)$$

where Lemma 1 is used in the last equality, and $\bar{g}_{l_n, l_n}^{(d)}$, $\bar{g}_{k_t, l_n}^{(c)}$ and $\bar{g}_{k_s, l_n}^{(d)}$ are given by

$$\begin{aligned} \bar{g}_{l_n, l_n}^{(d)} &= \tilde{g}_{l_n, l_n}^{(d)} - \tilde{g}_{l_n, l_n}^{(d)} \times \min \left\{ 1, \frac{\epsilon_{l_n, l_n}^{(d)}}{|\tilde{g}_{l_n, l_n}^{(d)}|} \right\}, \quad \forall l_n \in \mathcal{D}, \\ \bar{g}_{k_t, l_n}^{(c)} &= \tilde{g}_{k_t, l_n}^{(c)} + \frac{\epsilon_{k_t, l_n}^{(c)} \tilde{g}_{k_t, l_n}^{(c)}}{|\tilde{g}_{k_t, l_n}^{(c)}|}, \quad \forall k_t \in \mathcal{C}, l_n \in \mathcal{D}, \\ \bar{g}_{k_s, l_n}^{(d)} &= \tilde{g}_{k_s, l_n}^{(d)} + \frac{\epsilon_{k_s, l_n}^{(d)} \tilde{g}_{k_s, l_n}^{(d)}}{|\tilde{g}_{k_s, l_n}^{(d)}|}, \quad \forall k_s, l_n \in \mathcal{D}, k_s \neq l_n. \end{aligned} \quad (19)$$

In contrast, for cellular links, since the CUs in each cell simultaneously communicate with the corresponding BS on the same RB, it can be found from (3) and (4) that the channel vector $\mathbf{h}_{l_m, l}^{(c)}$ appears not only in the numerator term of $\text{SINR}_{l_m}^{(c)}$ but also in the denominator terms of $\text{SINR}_{k_t}^{(c)}$, $\forall k_t \in \mathcal{C}_l \setminus l_m$, where \mathcal{C}_l denotes the set of all CUs in cell l . Therefore, the worst-case $\text{SINR}_{l_m}^{(c)}$ cannot be obtained by independently finding the worst-case numerator and the worst-case terms in the denominator. For example, in cell 3 of Fig. 1, CU 3_1 and CU 3_2 simultaneously communicate with BS 3. Hence, $\mathbf{h}_{3_1, 3}^{(c)}$ ($\mathbf{h}_{3_2, 3}^{(c)}$) appears both in the numerator term of $\text{SINR}_{3_1}^{(c)}$ ($\text{SINR}_{3_2}^{(c)}$) and in the denominator terms of $\text{SINR}_{3_2}^{(c)}$ ($\text{SINR}_{3_1}^{(c)}$) due to co-channel interference. It is thus difficult to directly obtain the worst-case $\text{SINR}_{3_1}^{(c)}$ and $\text{SINR}_{3_2}^{(c)}$.

To deal with the SINR expressions in fractional form, the relationship between the MMSE and the pre-processing SINR was applied in [22] and [24], which considered a downlink multi-cell system. In the following, it is shown that a similar relationship also holds for a multi-cell uplink system.

Theorem 1: In a D2D underlaid multi-cell uplink system, if the MMSE filter $\mathbf{w}_{l_m}^{\text{MMSE}}$ is adopted by BS l for detecting $x_{l_m}^{(c)}$, denote the MMSE of cellular link l_m by $\text{MMSE}_{l_m}^{(c)}$. Then, the following relationship exists

$$\text{MMSE}_{l_m}^{(c)} = \frac{1}{1 + \text{SINR}_{l_m}^{(c)}}. \quad (20)$$

As for D2D link l_n , denote the worst-case MMSE by $\text{MMSE}_{l_n}^{(d)}$. Then, an analogous relationship between $\text{MMSE}_{l_n}^{(d)}$ and $\text{sinr}_{l_n}^{(d)}$ can also be obtained, i.e.,

$$\text{MMSE}_{l_n}^{(d)} = \frac{1}{1 + \text{sinr}_{l_n}^{(d)}}. \quad (21)$$

Proof: See Appendix A. \square

Based on (18) and applying Theorem 1, the objective function of (15), i.e., R^* can be equivalently rewritten as²

$$\begin{aligned} &\min_{\Delta} \sum_{l_m \in \mathcal{C}} \ln \left(1 + \text{SINR}_{l_m}^{(c)} \right) + \sum_{l_n \in \mathcal{D}} \ln \left(1 + \text{sinr}_{l_n}^{(d)} \right) \\ &= - \sum_{l=1}^L \max_{\Delta_l} \sum_{m=1}^M \ln \text{MMSE}_{l_m}^{(c)} - \sum_{l_n \in \mathcal{D}} \ln \text{MMSE}_{l_n}^{(d)} \\ &= - \sum_{l=1}^L \max_{\Delta_l} \min_{\mathbf{W}_l} \sum_{m=1}^M \ln \text{MSE}_{l_m}^{(c)} - \sum_{l_n \in \mathcal{D}} \min_{f_{l_n}} \ln \text{MSE}_{l_n}^{(d)}, \end{aligned} \quad (22)$$

where $f_{l_n} \in \mathbb{C} \setminus \{0\}$ is the single-tap receive equalizer at D2D-Rx l_n . $\text{MSE}_{l_m}^{(c)}$ and $\text{MSE}_{l_n}^{(d)}$ respectively denote the MSEs of cellular link l_m and D2D link l_n . The first equality of (22) holds because Δ_l , $\forall l \in \mathcal{L}$ have independent uncertainties. According to (22), the worst-case SE maximization problem (15) can be equivalently rewritten as

$$\begin{aligned} &\min_{\mathbf{P}} \left(\sum_{l=1}^L \max_{\Delta_l} \min_{\mathbf{W}_l} \sum_{m=1}^M \ln \text{MSE}_{l_m}^{(c)} + \sum_{l_n \in \mathcal{D}} \min_{f_{l_n}} \ln \text{MSE}_{l_n}^{(d)} \right) \\ &\text{s.t.} \quad (15b) \sim (15d), (7a), (7b). \end{aligned} \quad (23)$$

From (1) and the definition of MSE, it is known that $\mathbf{h}_{l_m, l}^{(c)}$ appears in all $\text{MSE}_{l_m}^{(c)}$, $\forall l_m \in \mathcal{C}_l$, making it difficult to obtain a tractable expression of $\sum_{m=1}^M \ln \text{MSE}_{l_m}^{(c)}$ due to the $\ln(\cdot)$ operation. In addition, the composite min-max-min optimization makes it more difficult to solve (23).

To simplify problem (23), recall that for any function $f(x, y)$, the inequality $\min_x \max_y f(x, y) \geq \max_y \min_x f(x, y)$ always holds [22]. By exchanging the positions of max and min, an upper bound on the objective function of (23) can be obtained and problem (23) can be simplified as follows

$$\begin{aligned} &\min_{\mathbf{P}} \left(\sum_{l=1}^L \min_{\mathbf{W}_l} \max_{\Delta_l} \sum_{m=1}^M \ln \text{MSE}_{l_m}^{(c)} + \sum_{l_n \in \mathcal{D}} \min_{f_{l_n}} \ln \text{MSE}_{l_n}^{(d)} \right) \\ &\text{s.t.} \quad (15b) \sim (15d), (7a), (7b). \end{aligned} \quad (24)$$

For further simplification, motivated by [39], the following auxiliary function is introduced to remove the $\ln(\cdot)$ operation in (24)

$$S_l(\mathbf{u}_l^{(c)}) = \max_{\Delta_l} \sum_{m=1}^M \left\{ \exp \left(u_{l_m}^{(c)} - 1 \right) \text{MSE}_{l_m}^{(c)} - u_{l_m}^{(c)} \right\}, \quad \forall l \in \mathcal{L}, \quad (25)$$

²Note that for the convenience of the following analysis, $\log(\cdot)$ is replaced with $\ln(\cdot)$ in (22).

where $\mathbf{u}_l^{(c)} = (u_{l_1}^{(c)}, \dots, u_{l_M}^{(c)})^T \in \mathbb{R}^{M \times 1}$ is a newly introduced auxiliary vector variable. By checking the first-order optimality condition of (25),

$$\min_{\mathbf{u}_l^{(c)}} S_l(\mathbf{u}_l^{(c)}) = \max_{\Delta_l} \sum_{m=1}^M \ln \text{MSE}_{l_m}^{(c)}, \quad \forall l \in \mathcal{L}. \quad (26)$$

Similarly, for D2D links, an auxiliary function is introduced, viz.,

$$T_{l_n}(u_{l_n}^{(d)}) = \exp(u_{l_n}^{(d)} - 1) \text{MSE}_{l_n}^{(d)} - u_{l_n}^{(d)}, \quad \forall l_n \in \mathcal{D}. \quad (27)$$

It follows

$$\min_{u_{l_n}^{(d)}} T_{l_n}(u_{l_n}^{(d)}) = \ln \text{MSE}_{l_n}^{(d)}, \quad \forall l_n \in \mathcal{D}, \quad (28)$$

and the corresponding optimal $u_{l_n}^{(d)*}$

$$u_{l_n}^{(d)*} = 1 - \ln \text{MSE}_{l_n}^{(d)}, \quad \forall l_n \in \mathcal{D}. \quad (29)$$

Substituting (26) and (28) into (24), and using the independence of \mathbf{W}_l , f_{l_n} , $\mathbf{u}_l^{(c)}$ as well as $u_{l_n}^{(d)}$, the objective function of problem (24) can be rewritten as

$$\begin{aligned} & \sum_{l=1}^L \min_{\mathbf{W}_l} \min_{\mathbf{u}_l^{(c)}} S_l(\mathbf{u}_l^{(c)}) + \sum_{l_n \in \mathcal{D}} \min_{f_{l_n}} \min_{u_{l_n}^{(d)}} T_{l_n}(u_{l_n}^{(d)}) \\ &= \min_{\mathbf{W}, \mathbf{f}, \mathbf{u}} \left\{ \sum_{l=1}^L S_l(\mathbf{u}_l^{(c)}) + \sum_{l_n \in \mathcal{D}} T_{l_n}(u_{l_n}^{(d)}) \right\}, \quad (30) \end{aligned}$$

where \mathbf{f} and \mathbf{u} are defined as follows

$$\begin{aligned} \mathbf{f} &= (f_{1_1}, \dots, f_{1_N}, \dots, f_{L_N})^T, \\ \mathbf{u} &= [\mathbf{u}^{(c)}; \mathbf{u}^{(d)}], \\ \mathbf{u}^{(c)} &= [u_1^{(c)}; \dots; u_L^{(c)}], \\ \mathbf{u}^{(d)} &= (u_{1_1}^{(d)}, \dots, u_{1_N}^{(d)}, \dots, u_{L_N}^{(d)})^T. \quad (31) \end{aligned}$$

Substituting (25) and (27) in (30), problem (24) can be reformulated as

$$\begin{aligned} \min_{\mathbf{p}, \mathbf{W}, \mathbf{f}, \mathbf{u}} & \left\{ \sum_{l=1}^L \max_{\Delta_l} \sum_{m=1}^M \left[\exp(u_{l_m}^{(c)} - 1) \text{MSE}_{l_m}^{(c)} - u_{l_m}^{(c)} \right] \right. \\ & \left. + \sum_{l_n \in \mathcal{D}} \left[\exp(u_{l_n}^{(d)} - 1) \text{MSE}_{l_n}^{(d)} - u_{l_n}^{(d)} \right] \right\} \\ \text{s.t.} & \quad (15b) \sim (15d), (7a), (7b). \quad (32) \end{aligned}$$

Since an explicit expression of the objective function of (32) is unavailable, it is difficult to solve this problem. In the following subsection, problem (32) is divided into two consecutive parts and an alternative algorithm is proposed to solve it. In the first part, \mathbf{W} , \mathbf{f} and \mathbf{u} are optimized for fixed \mathbf{p} , and vice versa in the second part.

B. Robust Optimization

1) *Solving (32) for Fixed \mathbf{p}* : For notational convenience, denote

$$\begin{aligned} \gamma_{l_m}^{(c)} &= \exp\left(\frac{u_{l_m}^{(c)} - 1}{2}\right), \quad q_{l_m}^{(c)} = \sqrt{p_{l_m}^{(c)}}, \quad \forall l_m \in \mathcal{C} \\ \gamma_{l_n}^{(d)} &= \exp\left(\frac{u_{l_n}^{(d)} - 1}{2}\right), \quad q_{l_n}^{(d)} = \sqrt{p_{l_n}^{(d)}}, \quad \forall l_n \in \mathcal{D}. \quad (33) \end{aligned}$$

Then, for fixed \mathbf{p} , problem (32) becomes

$$\begin{aligned} \min_{\mathbf{W}, \mathbf{f}, \Gamma, \mathbf{u}^{(d)}} & \left\{ \sum_{l=1}^L \max_{\Delta_l} \sum_{m=1}^M \left[\left(\gamma_{l_m}^{(c)}\right)^2 \text{MSE}_{l_m}^{(c)} - 2 \ln \gamma_{l_m}^{(c)} - 1 \right] \right. \\ & \left. + \sum_{l_n \in \mathcal{D}} \left[\exp(u_{l_n}^{(d)} - 1) \text{MSE}_{l_n}^{(d)} - u_{l_n}^{(d)} \right] \right\} \\ \text{s.t.} & \quad (7a), (7b), \\ & \quad \gamma_{l_m}^{(c)} > 0, \quad \forall l_m \in \mathcal{C}, \quad (34) \end{aligned}$$

where Γ is defined as

$$\begin{aligned} \Gamma &= [\Gamma_1, \dots, \Gamma_L], \\ \Gamma_l &= \text{diag}(\gamma_{l_1}^{(c)}, \dots, \gamma_{l_M}^{(c)}), \quad \forall l \in \mathcal{L}. \quad (35) \end{aligned}$$

Since \mathbf{W}_l , Γ_l , f_{l_n} and $u_{l_n}^{(d)}$, $\forall l \in \mathcal{L}$, $l_n \in \mathcal{D}$ are independent, and Δ_l , $\forall l \in \mathcal{L}$ have independent uncertainties, problem (34) can be divided into L subproblems as (36) for any cell $l \in \mathcal{L}$ and $L \times N$ subproblems as (37) for any D2D link $l_n \in \mathcal{D}$

$$\min_{\mathbf{W}_l, \Gamma_l} \max_{\Delta_l} \sum_{m=1}^M \left[\left(\gamma_{l_m}^{(c)}\right)^2 \text{MSE}_{l_m}^{(c)} - 2 \ln \gamma_{l_m}^{(c)} - 1 \right] \quad (36a)$$

$$\text{s.t.} \quad \left\| \Delta_{k_t, l}^{(c)} \right\| \leq \varepsilon_{k_t, l}^{(c)}, \quad \forall k_t \in \mathcal{C}, \quad (36b)$$

$$\left\| \Delta_{k_s, l}^{(d)} \right\| \leq \varepsilon_{k_s, l}^{(d)}, \quad \forall k_s \in \mathcal{D}, \quad (36c)$$

$$\gamma_{l_m}^{(c)} > 0, \quad \forall l_m \in \mathcal{C}_l. \quad (36d)$$

$$\min_{f_{l_n}^{(d)}, u_{l_n}^{(d)}} \left[\exp(u_{l_n}^{(d)} - 1) \text{MSE}_{l_n}^{(d)} - u_{l_n}^{(d)} \right]. \quad (37)$$

Since δ has been determined by (19), problem (37) can be optimally solved by respectively obtaining the optimal $f_{l_n}^{(d)}$ and $u_{l_n}^{(d)}$ based on (A.6) and (29). In contrast, Δ_l cannot be obtained directly due to the coupling of the cellular channel estimation error vectors. Therefore, it would be more difficult to solve (36). In the following theorem, it is shown that problem (36) can be cast to an SDP.

Theorem 2: For given \mathbf{p} , problem (36) can be optimized over \mathbf{W}_l and Γ_l efficiently as an SDP.

Proof: See Appendix B. \square

2) *Solving (32) for Fixed \mathbf{W} , \mathbf{f} and \mathbf{u}* : When \mathbf{W} , \mathbf{f} and \mathbf{u} have been determined, based on the definition of MSE, the objective function of (32) can be rewritten as (38), shown as the bottom of the next page, where \mathbf{J}_l , $\mathbf{J}_{l,m}$ and e_m are

defined in (B.3). From (38), it is seen that $q_{l_m}^{(c)}, \forall l_m \in \mathcal{C}$ can be obtained by separately solving the following subproblems

$$\min_{q_{l_m}^{(c)}} \max_{\Delta_{l_m}^{(c)}} \left\| q_{l_m}^{(c)} \mathbf{J}_l^H \mathbf{h}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m \right\|^2 + \sum_{k \neq l} \left\| q_{l_m}^{(c)} \mathbf{J}_k^H \mathbf{h}_{l_m,k}^{(c)} \right\|^2 + \left(q_{l_m}^{(c)} \right)^2 \sum_{k_s \in \mathcal{D}} \left| \gamma_{k_s}^{(d)} f_{k_s}^H \bar{g}_{l_m,k_s}^{(c)} \right|^2 \quad (39a)$$

$$\text{s.t.} \quad 0 \leq q_{l_m}^{(c)} \leq \sqrt{P_{l_m}^{(c)}}, \quad (39b)$$

$$\left\| \Delta_{l_m,k}^{(c)} \right\| \leq \varepsilon_{l_m,k}^{(c)}, \quad \forall k \in \mathcal{L}, \quad (39c)$$

where $\Delta_{l_m}^{(c)} = [\Delta_{l_m,1}^{(c)}, \dots, \Delta_{l_m,L}^{(c)}]$.

Due to the maximum interference constraints (15d), $\mathbf{q}^{(d)}$ cannot be obtained by separately finding $q_{l_n}^{(d)}, \forall l_n \in \mathcal{D}$. Therefore, the following problem is solved to obtain the optimal $\mathbf{q}^{(d)}$

$$\min_{\mathbf{q}^{(d)}} \max_{\Delta^{(d)}} \left\{ \sum_{l_n \in \mathcal{D}} \left\{ \sum_{k=1}^L \left\| q_{l_n}^{(d)} \mathbf{J}_k^H \mathbf{h}_{l_n,k}^{(d)} \right\|^2 + \theta_{l_n} \left(q_{l_n}^{(d)} \right)^2 - 2\phi_{l_n} q_{l_n}^{(d)} \right\} \right\} \quad (40a)$$

$$\text{s.t.} \quad 0 \leq q_{l_n}^{(d)} \leq \sqrt{P_{l_n}^{(d)}}, \quad \forall l_n \in \mathcal{D}, \quad (40b)$$

$$\sum_{k_s \in \mathcal{D}} \rho_{k_s,l} \left(q_{k_s}^{(d)} \right)^2 \leq a_l, \quad \forall l \in \mathcal{L}, \quad (40c)$$

$$\left\| \Delta_{l_n,k}^{(d)} \right\| \leq \varepsilon_{l_n,k}^{(d)}, \quad \forall l_n \in \mathcal{D}, k \in \mathcal{L}, \quad (40d)$$

where

$$\begin{aligned} \mathbf{q}^{(d)} &= \left(q_{1_1}^{(d)}, \dots, q_{1_N}^{(d)}, \dots, q_{L_N}^{(d)} \right)^T, \\ \Delta^{(d)} &= \left[\Delta_{1_1,1}^{(d)}, \dots, \Delta_{1_1,L}^{(d)}, \dots, \Delta_{L_N,L}^{(d)} \right], \\ \theta_{l_n}^{(d)} &= \sum_{k_s \in \mathcal{D}} \left| \gamma_{k_s}^{(d)} f_{k_s}^H \bar{g}_{l_n,k_s}^{(d)} \right|^2, \quad \forall l_n \in \mathcal{D}, \\ \phi_{l_n}^{(d)} &= \left(\gamma_{l_n}^{(d)} \right)^2 \text{Re} \left(f_{l_n}^H \bar{g}_{l_n,l_n}^{(d)} \right), \quad \forall l_n \in \mathcal{D}. \end{aligned} \quad (41)$$

To solve the robust optimization problems (39) and (40), the following theorem is given, which shows that they have computationally efficient solutions.

Theorem 3: For given \mathbf{W} , \mathbf{f} and \mathbf{u} , either problem (39) or problem (40) can be posed as an SDP.

Proof: See Appendix C. \square

Based on the above analysis, problem (32) can be solved by iteratively optimizing \mathbf{W} , \mathbf{f} , \mathbf{u} and \mathbf{p} . The detailed steps are summarized in Algorithm 1.

Algorithm 1 Robust Transmission Design (RTD)

- 1: Set $j = 0$, initialize $\mathbf{p}(j)$. $q_{l_m}^{(c)}(j) = \sqrt{p_{l_m}^{(c)}(j)}, \forall l_m \in \mathcal{C}$, $q_{l_n}^{(d)}(j) = \sqrt{p_{l_n}^{(d)}(j)}, \forall l_n \in \mathcal{D}$.
 - 2: **repeat**
 - 3: Obtain $\mathbf{J}_l(j+1)$ and $\Gamma_l(j+1), \forall l \in \mathcal{L}$ by solving SDP (B.9).
 - 4: $\mathbf{W}_l(j+1) = \mathbf{J}_l(j+1)\Gamma_l(j+1)^{-1}$, $u_{l_m}^{(c)}(j+1) = 1 + 2 \ln \gamma_{l_m}^{(c)}(j+1), \forall l \in \mathcal{L}, l_m \in \mathcal{C}$.
 - 5: Obtain $f_{l_n}^{(d)}(j+1)$ and $u_{l_n}^{(d)}(j+1), \forall l_n \in \mathcal{D}$ directly from (A.6) and (29).
 - 6: Obtain $q_{l_m}^{(c)}(j+1), \forall l_m \in \mathcal{C}$ and $\mathbf{q}^{(d)}(j+1)$ by solving SDP (C.2) and (C.5), respectively.
 - 7: $p_{l_m}^{(c)}(j+1) = \left[q_{l_m}^{(c)}(j+1) \right]^2, \forall l_m \in \mathcal{C}$, $p_{l_n}^{(d)}(j+1) = \left[q_{l_n}^{(d)}(j+1) \right]^2, \forall l_n \in \mathcal{D}$.
 - 8: Let $j = j + 1$.
 - 9: **until** convergence
-

C. Convergence and Complexity Analysis

Since an alternative algorithm is proposed to solve (32), it is necessary to characterize the convergence behavior of Algorithm 1.

Theorem 4: The iterative robust transmission design given in Algorithm 1 converges to a suboptimal solution of problem (32).

Proof: See Appendix D. \square

Besides the convergence behavior, it is also necessary to analyze the computational complexity of the proposed RTD algorithm. In the following, order notation $\mathcal{O}(\cdot)$ is adopted to characterize the computational complexity of Algorithm 1. According to [40]–[42], solving an SDP involves a complexity of $\mathcal{O}(i^{6.5} \log \lambda)$ for accuracy λ , where i denotes the dimension of matrix variables. Since $\mathbf{F}_{k_t,l}^{(c)}, \mathbf{F}_{k_s,l}^{(d)}, \forall l \in \mathcal{L}, k_t \in$

$$\begin{aligned} & \sum_{l_m \in \mathcal{C}} \left[\left(\gamma_{l_m}^{(c)} \right)^2 \text{MSE}_{l_m}^{(c)} - 2 \ln \gamma_{l_m}^{(c)} - 1 \right] + \sum_{l_n \in \mathcal{D}} \left[\left(\gamma_{l_n}^{(d)} \right)^2 \text{MSE}_{l_n}^{(d)} - 2 \ln \gamma_{l_n}^{(d)} - 1 \right] \\ &= \sum_{l_m \in \mathcal{C}} \left\{ \left\| q_{l_m}^{(c)} \mathbf{J}_l^H \mathbf{h}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m \right\|^2 + \sum_{k \neq l} \left\| q_{l_m}^{(c)} \mathbf{J}_k^H \mathbf{h}_{l_m,k}^{(c)} \right\|^2 + \left(q_{l_m}^{(c)} \right)^2 \sum_{k_s \in \mathcal{D}} \left| \gamma_{k_s}^{(d)} f_{k_s}^H \bar{g}_{l_m,k_s}^{(c)} \right|^2 \right\} \\ &+ \sum_{l_n \in \mathcal{D}} \left\{ \left(q_{l_n}^{(d)} \right)^2 \left[\sum_{k=1}^L \left\| \mathbf{J}_k^H \mathbf{h}_{l_n,k}^{(d)} \right\|^2 + \sum_{k_s \in \mathcal{D}} \left| \gamma_{k_s}^{(d)} f_{k_s}^H \bar{g}_{l_n,k_s}^{(d)} \right|^2 \right] - 2q_{l_n}^{(d)} \left(\gamma_{l_n}^{(d)} \right)^2 \text{Re} \left(f_{l_n}^H \bar{g}_{l_n,l_n}^{(d)} \right) \right\} \\ &+ \sum_{l_m \in \mathcal{C}} \left(N_0 \left\| \mathbf{J}_{l,m} \right\|^2 - 2 \ln \gamma_{l_m}^{(c)} - 1 \right) + \sum_{l_n \in \mathcal{D}} \left[\left(\gamma_{l_n}^{(d)} \right)^2 + N_0 \left| \gamma_{l_n}^{(d)} f_{l_n} \right|^2 - 2 \ln \gamma_{l_n}^{(d)} - 1 \right], \end{aligned} \quad (38)$$

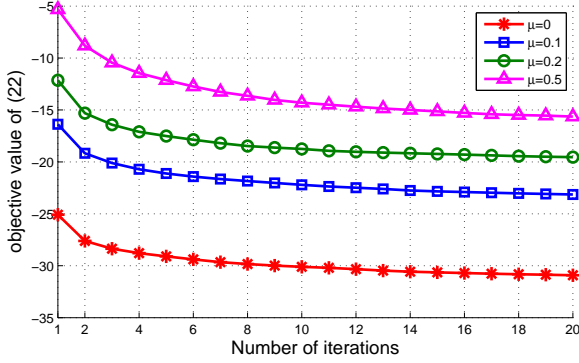


Fig. 2. Convergence behaviors of the proposed RTD algorithm with $L = 2$, $M = 2$, $N = 3$, $B = 4$, $P = 20$ dBm, $a = -80$ dBm and $D_{\max} = 100$.

\mathcal{C} , $\forall k_s \in \mathcal{D}$ are all $(1 + M + B) \times (1 + M + B)$ dimensional matrices, the complexity of solving SDP (B.9) is $\mathcal{O}([L(M + N)(1 + M + B)]^{6.5} \log \lambda)$. Similarly, obtaining \mathbf{q} by solving SDP (C.2) and (C.5) also requires complexity of $\mathcal{O}([L(M + N)(1 + M + B)]^{6.5} \log \lambda)$. Let I_{ite} denote the iteration numbers of the proposed RTD algorithm, then, the total complexity for solving (32) is $\mathcal{O}(I_{\text{ite}}[L(M + N)(1 + M + B)]^{6.5} \log \lambda)$. It is observed from simulations that the RTD algorithm converges after a few iterations, so the complexity is low and acceptable.

IV. SIMULATION RESULTS

In this section, simulation results are presented to evaluate the performance of the proposed RTD algorithm. Based on Theorem 2, Theorem 3 and Algorithm 1, it is required to solve SDPs (B.9), (C.2) and (C.5) to obtain a suboptimal solution of (32). CVX, a toolbox developed in MATLAB for solving convex problems [43], is used to solve SDPs. All simulation results are obtained by averaging over 1000 channel realizations, and each channel realization is obtained by generating a random user distribution as well as a random set of fading coefficients.

Consider a multi-cell D2D underlaid cellular system. All users are distributed uniformly and it is assumed that no user is closer to a BS than 20 meters. The distance between a D2D-Tx and its associated receiver is uniformly distributed in the range of $[0 \text{ m}, D_{\max} \text{ m}]$. The pathloss exponent and the standard deviation of log-normal shadowing fading are respectively set to be 3.7 and 8 dB [44]. The noise power is $N_0 = -100$ dBm. For brevity, equal maximum interference threshold at all BSs and equal maximum transmit power for all mobile transmitters are assumed, i.e., $a_l = a$, $\forall l \in \mathcal{L}$ and $P_{l_m}^{(c)} = P_{l_n}^{(d)} = P$, $\forall l_m \in \mathcal{C}$, $l_n \in \mathcal{D}$. Since this paper adopts the bounded CSI error model to characterize CSI impairment and quantization errors are the main source of CSI uncertainty for this model, a channel estimation error vector would be closely related to the corresponding channel vector estimate. Therefore, it is assumed that $\varepsilon_{k_t, l}^{(c)} = \mu \|\tilde{\mathbf{h}}_{k_t, l}^{(c)}\|$, $\varepsilon_{k_s, l}^{(d)} = \mu \|\tilde{\mathbf{h}}_{k_s, l}^{(d)}\|$, $\epsilon_{k_t, l_n}^{(c)} = \mu \|\tilde{\mathbf{g}}_{k_t, l_n}^{(c)}\|$ and $\epsilon_{k_s, l_n}^{(d)} = \mu \|\tilde{\mathbf{g}}_{k_s, l_n}^{(d)}\|$, $\forall l \in \mathcal{L}$, $k_t \in \mathcal{C}$, $l_n, k_s \in \mathcal{D}$, where $\mu \in [0, 1)$ is a metric used for evaluating the CSI error level [24].

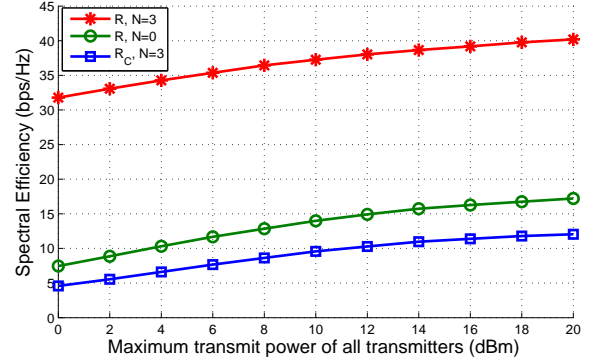


Fig. 3. Sum SE comparison between D2D underlaid cellular system and the conventional cellular system with $L = 2$, $M = 2$, $B = 4$, $a = -80$ dBm, $\mu = 0.3$ and $D_{\max} = 100$.

A. Convergence Behaviors of the Proposed RTD Algorithm

Based on the analysis in Section III, the original worst-case SE maximization problem (15) is first transformed into (23) using Theorem 1. By exchanging the positions of max and min, an upper bound on the objective function of problem (23) is obtained and (23) is further simplified as (24). Then, the proposed RTD algorithm is adopted to obtain a suboptimal solution of (24). Fig. 2 depicts the convergence behaviors of the proposed RTD algorithm under different values of channel error level μ . It can be seen from this figure that the objective value of (24) monotonically decreases during the iterative procedure and converges well in about 20 iterations for all considered configurations. Moreover, Fig. 2 also shows that the objective value of (24) grows with respect to (w.r.t.) μ . This is consistent with intuition since a larger μ usually results in more severe CSI uncertainties and, hence, increases the MSE of all links. Note that when $\mu = 0$, i.e., $\Delta = \mathbf{0}$ and $\delta = \mathbf{0}$, problem (24) becomes (23). In this case, a suboptimal solution of the original problem (15) can be obtained by using the proposed RTD algorithm.

B. Comparison to the Conventional Cellular Communication

To characterize the effect of D2D communication to the conventional cellular system (i.e., $N = 0$), the sum SE gains of the system and the performance loss of CUs resulted from D2D communication are investigated. Let R and R_C denote the sum SE of all links and the sum SE of all CUs, respectively.

Fig. 3 depicts the sum SE versus the maximum power of mobile transmitters. As expected, both R and R_C increase w.r.t. the maximum transmit power P . Fig. 3 also shows that D2D underlaid communication can provide significant performance gains over the conventional cellular communication. Specifically, for the case with $P = 10$ dBm, the sum SE of the conventional cellular system is increased by over 150% by D2D communication with $N = 3$. On the other hand, the performance of CUs may be affected by underlaid D2D communication due to cochannel interference. Since it is assumed in this paper that the interference signal from all D2D-Txs to each BS is upper bounded by a , the performance loss of CUs is much smaller than the performance gains

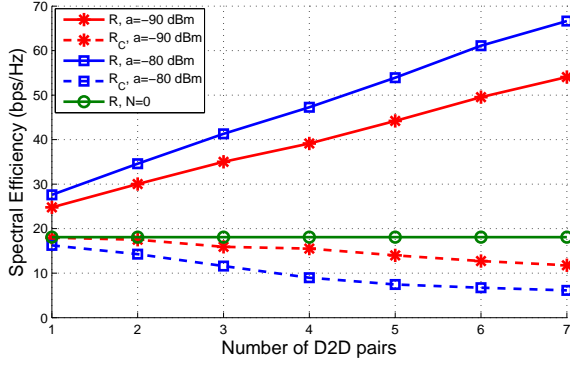


Fig. 4. Sum SE of all CUs and Sum SE of the system versus N with $L = 2$, $M = 2$, $B = 4$, $P = 20$ dBm, $\mu = 0.5$ and $D_{\max} = 100$.

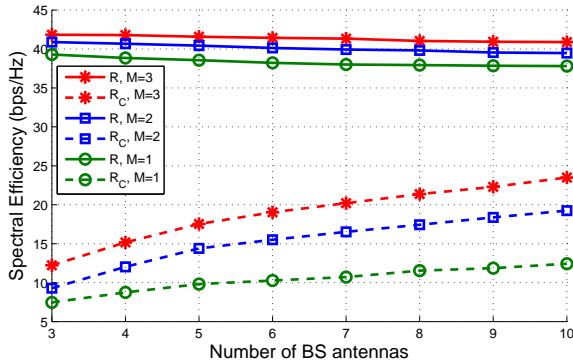


Fig. 5. Sum SE of all CUs and Sum SE of the system versus B with $L = 2$, $N = 3$, $P = 20$ dBm, $a = -80$ dBm, $\mu = 0.5$ and $D_{\max} = 100$.

brought by D2D communication, which can be verified by Fig. 3.

Fig. 4 plots the sum SE of all CUs and sum SE of the system versus the number of D2D pairs under different values of a . As a benchmark, the sum SE of the system without D2D communication is also depicted. From this figure, it is observed that the sum SE of the system, R , increases almost linearly with N , whereas R_C decreases with N due to cochannel interference resulted from D2D communication. When $a = -90$ dBm, the performance loss of CUs is small. For $a = -80$ dBm with small N , the performance loss of CUs appears also marginal, while as N grows larger than 4, noticeable cellular SE loss can be observed. This is because when aiming to maximize R , though the power of the interference signal from all D2D-Txs to each BS is upper bounded by a , the transmit power of CUs may be suppressed to increase the SE of DUs since DUs usually possess better channel conditions than CUs due to short transmission distance, especially when N is large. In this case, one can always reduce the performance loss of CUs by decreasing a to further protect cellular communication. Combining Fig. 3 and Fig. 4, one can conclude that it is an effective way to increase the system throughput by introducing D2D communication to a conventional cellular communication system while causing tolerable interference to CUs.

In Fig. 5, the effects of the number of BS antennas and the number of CUs are investigated. From this figure it can

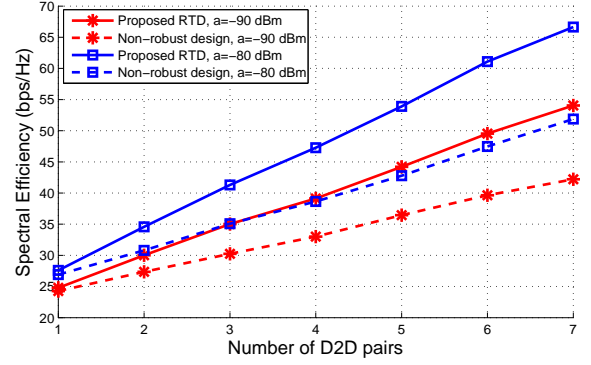


Fig. 6. Sum SE of the system versus N with $L = 2$, $M = 2$, $B = 4$, $P = 20$ dBm, $\mu = 0.5$ and $D_{\max} = 100$.

be found that both R and R_C increase with M , which is consistent with intuition. However, the increasing range of R is smaller than that of R_C . This is because as M grows, the co-channel interference experienced by DUs increases, leading to the reduction of D2D communication throughput. In addition, Fig. 5 also shows that for all considered cases, as B grows, R_C prominently increases, while R experiences a very slight decline. This can be explained by constraints (15d). From (12) and the assumption that $\varepsilon_{k_s,l}^{(d)} = \mu \|\tilde{\mathbf{h}}_{k_s,l}^{(d)}\|$, it follows that

$$\rho_{k_s,l} = (1 + \mu)^2 \|\tilde{\mathbf{h}}_{k_s,l}^{(d)}\|^2.$$

Since the elements in $\tilde{\mathbf{h}}_{k_s,l}^{(d)}$ are independent and follow the same distribution, as B grows, stronger constraints are imposed to DUs, i.e., DUs have to transmit in a relatively low power. Therefore, the sum SE of DUs decreases with B . In this case, one can always increase R by increasing a to relax constraints (15d).

C. Performance of the Proposed RTD Algorithm

For a single-cell single-input single-output (SISO) network, the problem of maximizing the sum system SE under perfect CSI was studied in [14]. However, for a multi-cell MIMO system with several CUs and several D2D pairs in each cell, as stated in Section I, the considered RTD problem under CSI impairment has not been studied. Therefore, to evaluate the performance of the proposed RTD algorithm, as in [22], [24], the simulation results obtained by using the ‘non-robust’ design are used as the benchmark, and compared with that obtained by the proposed RTD algorithm. In particular, the non-robust scheme tries to jointly optimize \mathbf{p} and \mathbf{W} by simply treating the estimated channel coefficients as the true channel, and the non-robust benchmark is obtained by using Algorithm 1 with μ set to be 0.

Fig. 6 compares the proposed RTD algorithm with the non-robust design versus the number of D2D pairs under different values of a . It can be seen that in the small N regime, compared with the non-robust design, a small SE gain is obtained by Algorithm 1, while as N grows large, the SE gap increases greatly. In particular, when $N = 5$, compared with the non-robust design scheme, the sum SE of the system can be respectively increased by about 25% and 20% for the cases with $a = -80$ dBm and $a = -90$ dBm via adopting the

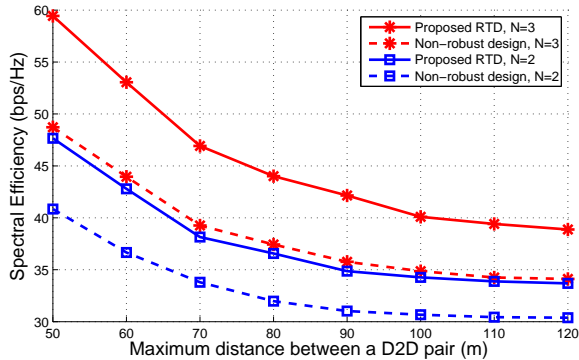


Fig. 7. Sum SE of the system versus D_{\max} with $L = 2$, $M = 2$, $B = 4$, $P = 20$ dBm, $\alpha = -80$ dBm and $\mu = 0.5$.

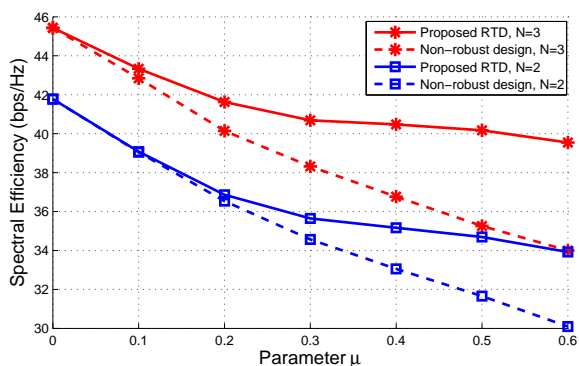


Fig. 8. Sum SE of the system versus μ with $L = 2$, $M = 2$, $B = 4$, $P = 20$ dBm, $\alpha = -80$ dBm and $D_{\max} = 100$.

proposed RTD algorithm. The effect of the maximum distance between a D2D pair is investigated in Fig. 7. As expected, the sum SE of the system decreases with D_{\max} for all considered cases.

Fig. 8 depicts the sum SE comparison versus channel error level μ under different values of N . It shows that the sum SE of the system monotonically decreases with μ , which is consistent with intuition. In addition, it can also be seen that the performance gains brought by Algorithm 1 over the non-robust design increase prominently with μ , which indicates that Algorithm 1 is more efficient in obtaining a higher throughput of the system when channel estimation suffers from serious uncertainties. The effect of the maximum interference from all D2D-Txs to each BS is investigated in Fig. 9 for different network sizes. As expected, the sum SE of the system increases with a and the network size. It also shows that in contrast to the non-robust design scheme, the sum SE gains brought by the proposed RTD algorithm grows with L . Hence, Algorithm 1 is suitable for networks with different sizes.

V. CONCLUSIONS

In this paper, the robust transmission design for a multi-cell D2D underlaid cellular system when BSs only have imperfect CSI of all links has been studied. To account for CSI uncertainties, this paper aims to maximize the worst-case sum rate of the system while guaranteeing that the interference

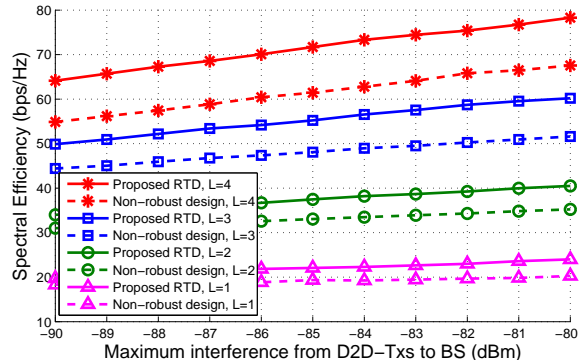


Fig. 9. Sum SE of the system versus a with $M = 2$, $N = 3$, $B = 4$, $P = 20$ dBm and $D_{\max} = 100$.

signal from all D2D-Txs to each BS is power-limited. To solve the nonconvex problem, it is first transformed into a more tractable form by replacing the objective function of the original problem with its lower bound. Then, the resulted problem is decomposed into several convex SDP subproblems, and an iterative algorithm is proposed to obtain a suboptimal solution. Simulation results show that the performance of the conventional cellular systems can be significantly improved by D2D communication while endurable interference is caused to CUs. In addition, the proposed robust transmission design algorithm greatly outperforms the non-robust transmission design in terms of system SE.

APPENDIX A PROOF OF THEOREM 1

From (1), the MMSE receive filter at BS l for detecting $x_{l_m}^{(c)}$ is given by ³

$$\begin{aligned} \mathbf{w}_{l_m}^{\text{MMSE}} &= \arg \min_{\mathbf{w}_{l_m}} \mathbb{E} \left\{ \left| \mathbf{w}_{l_m}^H \mathbf{y}_l^{(c)} - x_{l_m}^{(c)} \right|^2 \right\} \\ &= \sqrt{p_{l_m}^{(c)}} (\mathbf{A}_{l_m} + \mathbf{G}_{l_m})^{-1} \mathbf{h}_{l_m,l}^{(c)}, \end{aligned} \quad (\text{A.1})$$

where $\mathbf{A}_{l_m} = p_{l_m}^{(c)} \mathbf{h}_{l_m,l}^{(c)} (\mathbf{h}_{l_m,l}^{(c)})^H$. Using (A.1), the MMSE of cellular link l_m can be written as

$$\begin{aligned} \text{MMSE}_{l_m}^{(c)} &= \mathbb{E} \left\{ \left| (\mathbf{w}_{l_m}^{\text{MMSE}})^H \mathbf{y}_l^{(c)} - x_{l_m}^{(c)} \right|^2 \right\} \\ &= (\mathbf{w}_{l_m}^{\text{MMSE}})^H (\mathbf{A}_{l_m} + \mathbf{G}_{l_m}) \mathbf{w}_{l_m}^{\text{MMSE}} + 1 \\ &\quad - \sqrt{p_{l_m}^{(c)}} (\mathbf{w}_{l_m}^{\text{MMSE}})^H \mathbf{h}_{l_m,l}^{(c)} - \sqrt{p_{l_m}^{(c)}} (\mathbf{h}_{l_m,l}^{(c)})^H \mathbf{w}_{l_m}^{\text{MMSE}} \\ &= 1 - p_{l_m}^{(c)} (\mathbf{h}_{l_m,l}^{(c)})^H (\mathbf{A}_{l_m} + \mathbf{G}_{l_m})^{-1} \mathbf{h}_{l_m,l}^{(c)} \\ &= \frac{1}{1 + p_{l_m}^{(c)} (\mathbf{h}_{l_m,l}^{(c)})^H \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)}}. \end{aligned} \quad (\text{A.2})$$

³From the expressions of \mathbf{G}_{l_m} and \mathbf{A}_{l_m} , it is known that \mathbf{G}_{l_m} and $\mathbf{A}_{l_m} + \mathbf{G}_{l_m}$ are both positive definite matrices, and are thus reversible.

To prove equation (20), in the following, it is proven that the post-processing SINR of CU l_m satisfies $\text{SINR}_{l_m}^{(c)} = p_{l_m}^{(c)} \left(\mathbf{h}_{l_m,l}^{(c)} \right)^H \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)}$ by checking the following term

$$\begin{aligned} & \frac{\text{SINR}_{l_m}^{(c)}}{p_{l_m}^{(c)} \left(\mathbf{h}_{l_m,l}^{(c)} \right)^H \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)}} \\ \stackrel{(a)}{=} & \frac{\left(\mathbf{w}_{l_m}^{\text{MMSE}} \right)^H \mathbf{A}_{l_m} \mathbf{w}_{l_m}^{\text{MMSE}}}{\left(\mathbf{w}_{l_m}^{\text{MMSE}} \right)^H \mathbf{G}_{l_m} \mathbf{w}_{l_m}^{\text{MMSE}} p_{l_m}^{(c)} \left(\mathbf{h}_{l_m,l}^{(c)} \right)^H \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)}} \\ \stackrel{(b)}{=} & \frac{\left(\mathbf{w}_{l_m}^{\text{MMSE}} \right)^H \mathbf{A}_{l_m} \mathbf{w}_{l_m}^{\text{MMSE}}}{\left(\mathbf{w}_{l_m}^{\text{MMSE}} \right)^H \mathbf{G}_{l_m} \sqrt{p_{l_m}^{(c)}} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right)^{-1} \mathbf{A}_{l_m} \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)}}, \end{aligned} \quad (\text{A.3})$$

where (a) uses (3) and (b) follows by substituting (A.1). Since

$$\begin{aligned} & \mathbf{G}_{l_m} \sqrt{p_{l_m}^{(c)}} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right)^{-1} \mathbf{A}_{l_m} \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)} \\ = & \mathbf{G}_{l_m} \sqrt{p_{l_m}^{(c)}} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right)^{-1} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} - \mathbf{G}_{l_m} \right) \mathbf{G}_{l_m}^{-1} \mathbf{h}_{l_m,l}^{(c)} \\ = & \sqrt{p_{l_m}^{(c)}} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right) \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right)^{-1} \mathbf{h}_{l_m,l}^{(c)} \\ & - \sqrt{p_{l_m}^{(c)}} \mathbf{G}_{l_m} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right)^{-1} \mathbf{h}_{l_m,l}^{(c)} \\ = & \sqrt{p_{l_m}^{(c)}} \mathbf{A}_{l_m} \left(\mathbf{A}_{l_m} + \mathbf{G}_{l_m} \right)^{-1} \mathbf{h}_{l_m,l}^{(c)} \\ = & \mathbf{A}_{l_m} \mathbf{w}_{l_m}^{\text{MMSE}}, \end{aligned} \quad (\text{A.4})$$

combining (A.2), (A.3) and (A.4), it can be concluded that

$$\text{MMSE}_{l_m}^{(c)} = \frac{1}{1 + \text{SINR}_{l_m}^{(c)}}. \quad (\text{A.5})$$

Similarly, as for D2D link l_n , in the worst-case, the MMSE receive filter at D2D-Rx l_n is given by

$$\begin{aligned} f_{l_n}^{\text{MMSE}} &= \arg \min_{f_{l_n}} \mathbb{E} \left\{ \left| f_{l_n}^H \bar{y}_{l_n}^{(d)} - x_{l_n}^{(d)} \right|^2 \right\} \\ &= \frac{\sqrt{p_{l_n}^{(d)}} \bar{g}_{l_n,l_n}^{(d)}}{\sum_{k_t \in \mathcal{C}} p_{k_t}^{(c)} \left| \bar{g}_{k_t,l_n}^{(c)} \right|^2 + \sum_{k_s \in \mathcal{D}} p_{k_s}^{(d)} \left| \bar{g}_{k_s,l_n}^{(d)} \right|^2 + N_0}, \end{aligned} \quad (\text{A.6})$$

where $\bar{y}_{l_n}^{(d)}$ is obtained by replacing the channel coefficients in (2) with (19), i.e.,

$$\bar{y}_{l_n}^{(d)} = \sum_{k_t \in \mathcal{C}} \sqrt{p_{k_t}^{(c)}} \bar{g}_{k_t,l_n}^{(c)} x_{k_t}^{(c)} + \sum_{k_s \in \mathcal{D}} \sqrt{p_{k_s}^{(d)}} \bar{g}_{k_s,l_n}^{(d)} x_{k_s}^{(d)} + z_{l_n}^{(d)}. \quad (\text{A.7})$$

Then, the worst-case $\text{MMSE}_{l_n}^{(d)}$ can be written as

$$\begin{aligned} \text{MMSE}_{l_n}^{(d)} &= \mathbb{E} \left\{ \left| \left(f_{l_n}^{\text{MMSE}} \right)^H \bar{y}_{l_n}^{(d)} - x_{l_n}^{(d)} \right|^2 \right\} \\ &= 1 - \frac{p_{l_n}^{(d)} \left| \bar{g}_{l_n,l_n}^{(d)} \right|^2}{\sum_{k_t \in \mathcal{C}} p_{k_t}^{(c)} \left| \bar{g}_{k_t,l_n}^{(c)} \right|^2 + \sum_{k_s \in \mathcal{D}} p_{k_s}^{(d)} \left| \bar{g}_{k_s,l_n}^{(d)} \right|^2 + N_0}. \end{aligned} \quad (\text{A.8})$$

Combining (18) and (A.8), it can be easily seen that

$$\text{MMSE}_{l_n}^{(d)} = \frac{1}{1 + \text{sinr}_{l_n}^{(d)}}. \quad (\text{A.9})$$

Thus, Theorem 1 is proven.

APPENDIX B PROOF OF THEOREM 2

Before proving Theorem 2, the following useful result from [45], [46] is first given.

Lemma 2: For any given matrices \mathbf{X} , \mathbf{Y} and \mathbf{Z} , and a Hermitian matrix $\mathbf{\Lambda} = \mathbf{\Lambda}^H$, the inequality

$$\mathbf{\Lambda} \succeq \mathbf{X}^H \mathbf{Z} \mathbf{Y} + \mathbf{Y}^H \mathbf{Z}^H \mathbf{X}, \quad \forall \mathbf{Z} : \|\mathbf{Z}\| \leq \varepsilon \quad (\text{B.1})$$

holds if and only if

$$\exists \eta \geq 0 \text{ such that } \begin{bmatrix} \mathbf{\Lambda} - \eta \mathbf{X}^H \mathbf{X} & -\varepsilon \mathbf{Y}^H \\ -\varepsilon \mathbf{Y} & \eta \mathbf{I} \end{bmatrix} \succeq 0. \quad (\text{B.2})$$

Combining (1) and the definition of MSE, the objective function of problem (36) can be rewritten as

$$\begin{aligned} & \sum_{m=1}^M \left[\left(\gamma_{l_m}^{(c)} \right)^2 \text{MSE}_{l_m}^{(c)} - 2 \ln \gamma_{l_m}^{(c)} - 1 \right] \\ = & \sum_{m=1}^M \left\{ \left(\gamma_{l_m}^{(c)} \right)^2 \mathbb{E} \left\{ \left| \mathbf{w}_{l_m}^H \mathbf{y}_l^{(c)} - x_{l_m}^{(c)} \right|^2 \right\} - 2 \ln \gamma_{l_m}^{(c)} - 1 \right\} \\ = & \sum_{m=1}^M \left\| q_{l_m}^{(c)} \mathbf{J}_l^H \mathbf{h}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m \right\|^2 \\ & + \sum_{k_t \in \mathcal{C} \setminus \mathcal{C}_l} \left\| q_{k_t}^{(c)} \mathbf{J}_l^H \mathbf{h}_{k_t,l}^{(c)} \right\|^2 + \sum_{k_s \in \mathcal{D}} \left\| q_{k_s}^{(d)} \mathbf{J}_l^H \mathbf{h}_{k_s,l}^{(d)} \right\|^2 \\ & + \sum_{m=1}^M N_0 \|\mathbf{J}_{l,m}\|^2 - 2 \sum_{m=1}^M \ln \gamma_{l_m}^{(c)} - M, \end{aligned} \quad (\text{B.3})$$

where $\mathbf{J}_l = \mathbf{W}_l \mathbf{\Gamma}_l$ and $\mathbf{J}_{l,m}$ denotes the m th column of \mathbf{J}_l , i.e., $\mathbf{J}_{l,m} = \gamma_{l_m}^{(c)} \mathbf{w}_{l_m}$. \mathbf{e}_m is an M dimensional vector with one in the m th position and zeros elsewhere. By introducing a slack scalar variable r_l to bound (B.3), (36) can be reformulated in the following form

$$\min_{\mathbf{J}_l, \mathbf{\Gamma}_l, r_l} r_l \quad (\text{B.4a})$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{m=1}^M \left\| q_{l_m}^{(c)} \mathbf{J}_l^H \mathbf{h}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m \right\|^2 \\ & + \sum_{k_t \in \mathcal{C} \setminus \mathcal{C}_l} \left\| q_{k_t}^{(c)} \mathbf{J}_l^H \mathbf{h}_{k_t,l}^{(c)} \right\|^2 + \sum_{k_s \in \mathcal{D}} \left\| q_{k_s}^{(d)} \mathbf{J}_l^H \mathbf{h}_{k_s,l}^{(d)} \right\|^2 \\ & + \sum_{m=1}^M N_0 \|\mathbf{J}_{l,m}\|^2 - 2 \sum_{m=1}^M \ln \gamma_{l_m}^{(c)} - M \leq r_l, \end{aligned} \quad (\text{B.4b})$$

$$\left\| \mathbf{\Delta}_{k_t,l}^{(c)} \right\| \leq \varepsilon_{k_t,l}^{(c)}, \quad \forall k_t \in \mathcal{C}, \quad (\text{B.4c})$$

$$\left\| \mathbf{\Delta}_{k_s,l}^{(d)} \right\| \leq \varepsilon_{k_s,l}^{(d)}, \quad \forall k_s \in \mathcal{D}, \quad (\text{B.4d})$$

$$\gamma_{l_m}^{(c)} > 0, \quad \forall l_m \in \mathcal{C}_l. \quad (\text{B.4e})$$

To further simplify constraint (B.4b), additional auxiliary variables $\mathbf{b}_l = \left(b_{11,l}^{(c)}, \dots, b_{1M,l}^{(c)}, \dots, b_{LM,l}^{(c)}, b_{11,l}^{(d)}, \dots, b_{1N,l}^{(d)}, \dots \right)$

$\dots, b_{LN,l}^{(d)})^T$, $\mathbf{o}_l = (o_{l_1}, \dots, o_{l_M})^T$ are introduced, and (B.4b) with (B.4c), (B.4d) are rewritten as follows

$$\begin{aligned} & \left\| q_{l_m}^{(c)} \mathbf{J}_l^H \mathbf{h}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m \right\| \leq b_{l_m,l}^{(c)}, \\ & \forall \left\| \Delta_{l_m,l}^{(c)} \right\| \leq \varepsilon_{l_m,l}^{(c)}, \quad \forall l_m \in \mathcal{C}_l, \end{aligned} \quad (\text{B.5a})$$

$$\left\| q_{k_t}^{(c)} \mathbf{J}_l^H \mathbf{h}_{k_t,l}^{(c)} \right\| \leq b_{k_t,l}^{(c)}, \quad \forall \left\| \Delta_{k_t,l}^{(c)} \right\| \leq \varepsilon_{k_t,l}^{(c)}, \quad \forall k_t \in \mathcal{C} \setminus \mathcal{C}_l, \quad (\text{B.5b})$$

$$\left\| q_{k_s}^{(d)} \mathbf{J}_l^H \mathbf{h}_{k_s,l}^{(d)} \right\| \leq b_{k_s,l}^{(d)}, \quad \forall \left\| \Delta_{k_s,l}^{(d)} \right\| \leq \varepsilon_{k_s,l}^{(d)}, \quad \forall k_s \in \mathcal{D}, \quad (\text{B.5c})$$

$$\sqrt{N_0} \|\mathbf{J}_{l,m}\| \leq o_{l_m}, \quad \forall l_m \in \mathcal{C}_l, \quad (\text{B.5d})$$

$$\begin{aligned} & \sum_{k_t \in \mathcal{C}} \left(b_{k_t,l}^{(c)} \right)^2 + \sum_{k_s \in \mathcal{D}} \left(b_{k_s,l}^{(d)} \right)^2 + \sum_{m=1}^M o_{l_m}^2 \\ & - 2 \sum_{m=1}^M \ln \gamma_{l_m}^{(c)} - M \leq r_l. \end{aligned} \quad (\text{B.5e})$$

Obviously, (B.5d) is a second-order cone (SOC) constraint and (B.5e) is a convex constraint. As for constraints (B.5a)~(B.5c), they can be transformed into finite linear matrix inequalities (LMIs). First take (B.5a) for instance. By applying the Schur Complement Lemma [47], constraint (B.5a) can be equivalently stated as

$$\begin{bmatrix} b_{l_m,l}^{(c)} & q_{l_m}^{(c)} \left(\tilde{\mathbf{h}}_{l_m,l}^{(c)} + \Delta_{l_m,l}^{(c)} \right)^H \mathbf{J}_l - \gamma_{l_m}^{(c)} \mathbf{e}_m^H \\ q_{l_m}^{(c)} \mathbf{J}_l^H \left(\tilde{\mathbf{h}}_{l_m,l}^{(c)} + \Delta_{l_m,l}^{(c)} \right) - \gamma_{l_m}^{(c)} \mathbf{e}_m & b_{l_m,l}^{(c)} \mathbf{I} \end{bmatrix} \succeq 0, \quad \forall \left\| \Delta_{l_m,l}^{(c)} \right\| \leq \varepsilon_{l_m,l}^{(c)}, \quad \forall l_m \in \mathcal{C}_l, \quad (\text{B.6})$$

Denote $\mathbf{X} = [-1 \ 0]$, $\mathbf{Y} = \begin{bmatrix} \mathbf{0} & q_{l_m}^{(c)} \mathbf{J}_l \end{bmatrix}$, $\mathbf{Z} = \left(\Delta_{l_m,l}^{(c)} \right)^H$ and

$$\mathbf{A} = \begin{bmatrix} b_{l_m,l}^{(c)} & q_{l_m}^{(c)} \left(\tilde{\mathbf{h}}_{l_m,l}^{(c)} \right)^H \mathbf{J}_l - \gamma_{l_m}^{(c)} \mathbf{e}_m^H \\ q_{l_m}^{(c)} \mathbf{J}_l^H \tilde{\mathbf{h}}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m & b_{l_m,l}^{(c)} \mathbf{I} \end{bmatrix}. \quad (\text{B.7})$$

Then, based on Lemma 2, constraint (B.5a) can be equivalently represented by

$$\mathbf{F}_{l_m,l}^{(c)} = \begin{bmatrix} b_{l_m,l}^{(c)} - \eta_{l_m,l}^{(c)} & q_{l_m}^{(c)} \left(\tilde{\mathbf{h}}_{l_m,l}^{(c)} \right)^H \mathbf{J}_l - \gamma_{l_m}^{(c)} \mathbf{e}_m^H & \mathbf{0} \\ q_{l_m}^{(c)} \mathbf{J}_l^H \tilde{\mathbf{h}}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m & b_{l_m,l}^{(c)} \mathbf{I} & -\varepsilon_{l_m,l}^{(c)} q_{l_m}^{(c)} \mathbf{J}_l^H \\ \mathbf{0} & -\varepsilon_{l_m,l}^{(c)} q_{l_m}^{(c)} \mathbf{J}_l & \eta_{l_m,l}^{(c)} \mathbf{I} \end{bmatrix} \succeq 0, \quad \forall l_m \in \mathcal{C}_l, \quad (\text{B.8})$$

which is a LMI and can be surely satisfied by finding a proper $\eta_{l_m,l}^{(c)} \geq 0$. Similarly, constraints (B.5b) and (B.5c) can also be

transformed into LMIs. Thus, problem (36) can be cast to an SDP as follows

$$\min_{\mathbf{J}_l, \Gamma_l, r_l, \mathbf{b}_l, \mathbf{o}_l, \boldsymbol{\eta}_l} r_l \quad (\text{B.9a})$$

$$\text{s.t.} \quad \mathbf{F}_{k_t,l}^{(c)} \succeq 0, \quad \forall k_t \in \mathcal{C}, \quad (\text{B.9b})$$

$$\mathbf{F}_{k_s,l}^{(d)} \succeq 0, \quad \forall k_s \in \mathcal{D}, \quad (\text{B.9c})$$

$$\eta_{k_t,l}^{(c)} \geq 0, \quad \forall k_t \in \mathcal{C}, \quad (\text{B.9d})$$

$$\eta_{k_s,l}^{(d)} \geq 0, \quad \forall k_s \in \mathcal{D}, \quad (\text{B.9e})$$

$$(\text{B.5d}), (\text{B.5e}), (\text{B.4e}), \quad (\text{B.9f})$$

where $\boldsymbol{\eta}_l = (\eta_{11,l}^{(c)}, \dots, \eta_{1M,l}^{(c)}, \dots, \eta_{LM,l}^{(c)}, \eta_{11,l}^{(d)}, \dots, \eta_{1N,l}^{(d)}, \dots, \eta_{LN,l}^{(d)})^T$, $\mathbf{F}_{k_t,l}^{(c)}$, $\forall k_t \in \mathcal{C} \setminus \mathcal{C}_l$ and $\mathbf{F}_{k_s,l}^{(d)}$ are given by

$$\mathbf{F}_{k_t,l}^{(c)} = \begin{bmatrix} b_{k_t,l}^{(c)} - \eta_{k_t,l}^{(c)} & q_{k_t}^{(c)} \left(\tilde{\mathbf{h}}_{k_t,l}^{(c)} \right)^H \mathbf{J}_l & \mathbf{0} \\ q_{k_t}^{(c)} \mathbf{J}_l^H \tilde{\mathbf{h}}_{k_t,l}^{(c)} & b_{k_t,l}^{(c)} \mathbf{I} & -\varepsilon_{k_t,l}^{(c)} q_{k_t}^{(c)} \mathbf{J}_l^H \\ \mathbf{0} & -\varepsilon_{k_t,l}^{(c)} q_{k_t}^{(c)} \mathbf{J}_l & \eta_{k_t,l}^{(c)} \mathbf{I} \end{bmatrix}, \quad \forall k_t \in \mathcal{C} \setminus \mathcal{C}_l, \quad (\text{B.10})$$

and

$$\mathbf{F}_{k_s,l}^{(d)} = \begin{bmatrix} b_{k_s,l}^{(d)} - \eta_{k_s,l}^{(d)} & q_{k_s}^{(d)} \left(\tilde{\mathbf{h}}_{k_s,l}^{(d)} \right)^H \mathbf{J}_l & \mathbf{0} \\ q_{k_s}^{(d)} \mathbf{J}_l^H \tilde{\mathbf{h}}_{k_s,l}^{(d)} & b_{k_s,l}^{(d)} \mathbf{I} & -\varepsilon_{k_s,l}^{(d)} q_{k_s}^{(d)} \mathbf{J}_l^H \\ \mathbf{0} & -\varepsilon_{k_s,l}^{(d)} q_{k_s}^{(d)} \mathbf{J}_l & \eta_{k_s,l}^{(d)} \mathbf{I} \end{bmatrix}, \quad \forall k_s \in \mathcal{D}. \quad (\text{B.11})$$

Thus far, problem (B.9) has been recognized as an SDP with linear objective and LMI or SOC constraints. Once \mathbf{J}_l and Γ_l have been obtained by solving (B.9), one can readily get \mathbf{W}_l based on the relationship $\mathbf{W}_l = \mathbf{J}_l \Gamma_l^{-1}$.

APPENDIX C

PROOF OF THEOREM 3

A. Casting (39) to an SDP

Similar as the proof in Appendix B, (39) can be transformed to the following form by introducing auxiliary variables β_{l_m} and $\boldsymbol{\xi}_{l_m}^{(c)} = (\xi_{l_m,1}^{(c)}, \dots, \xi_{l_m,L}^{(c)})^T$

$$\min_{q_{l_m}^{(c)}, \beta_{l_m}, \boldsymbol{\xi}_{l_m}^{(c)}} \beta_{l_m} \quad (\text{C.1a})$$

$$\text{s.t.} \quad \left\| q_{l_m}^{(c)} \mathbf{J}_l^H \mathbf{h}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m \right\| \leq \xi_{l_m,l}^{(c)}, \quad (\text{C.1b})$$

$$\left\| q_{l_m}^{(c)} \mathbf{J}_k^H \mathbf{h}_{l_m,k}^{(c)} \right\| \leq \xi_{l_m,k}^{(c)}, \quad \forall k \in \mathcal{L} \setminus l, \quad (\text{C.1c})$$

$$\sum_{k \in \mathcal{L}} \left(\xi_{l_m,k}^{(c)} \right)^2 + \left(q_{l_m}^{(c)} \right)^2 \sum_{k_s \in \mathcal{D}} \left| \gamma_{k_s}^{(d)} f_{k_s}^H g_{l_m,k_s}^{(c)} \right|^2 \leq \beta_{l_m}^2, \quad (\text{C.1d})$$

$$\beta_{l_m} \geq 0, \quad (\text{C.1e})$$

$$0 \leq q_{l_m}^{(c)} \leq \sqrt{P_{l_m}^{(c)}}, \quad (\text{C.1f})$$

$$\left\| \Delta_{l_m,k}^{(c)} \right\| \leq \varepsilon_{l_m,k}^{(c)}, \quad \forall k \in \mathcal{L}. \quad (\text{C.1g})$$

Applying the Schur Complement Lemma and Lemma 2, constraints (C.1b), (C.1c) and (C.1g) can be transformed into LMIs by further introducing auxiliary variable $\varphi_{l_m}^{(c)} = (\varphi_{l_m,1}^{(c)}, \dots, \varphi_{l_m,L}^{(c)})^T$. Accordingly, problem (C.1) can be rewritten as

$$\min_{q_{l_m}^{(c)}, \beta_{l_m}, \xi_{l_m}^{(c)}, \varphi_{l_m}^{(c)}} \beta_{l_m} \quad (\text{C.2a})$$

$$\text{s.t.} \quad \mathbf{U}_{l_m,k}^{(c)} \succeq 0, \quad \forall k \in \mathcal{L}, \quad (\text{C.2b})$$

$$\varphi_{l_m,k}^{(c)} \geq 0, \quad \forall k \in \mathcal{L}, \quad (\text{C.2c})$$

$$(\text{C.1d}) \sim (\text{C.1f}), \quad (\text{C.2d})$$

where $\mathbf{U}_{l_m,k}$ is given by

$$\mathbf{U}_{l_m,l}^{(c)} = \begin{bmatrix} \xi_{l_m,l}^{(c)} - \varphi_{l_m,l}^{(c)} & q_{l_m}^{(c)} (\tilde{\mathbf{h}}_{l_m,l}^{(c)})^H \mathbf{J}_l - \gamma_{l_m}^{(c)} \mathbf{e}_m^H & \mathbf{0} \\ q_{l_m}^{(c)} \mathbf{J}_l^H \tilde{\mathbf{h}}_{l_m,l}^{(c)} - \gamma_{l_m}^{(c)} \mathbf{e}_m & \xi_{l_m,l}^{(c)} \mathbf{I} & -\varepsilon_{l_m,l}^{(c)} q_{l_m}^{(c)} \mathbf{J}_l^H \\ \mathbf{0} & -\varepsilon_{l_m,l}^{(c)} q_{l_m}^{(c)} \mathbf{J}_l & \varphi_{l_m,l}^{(c)} \mathbf{I} \end{bmatrix}, \quad (\text{C.3})$$

and

$$\mathbf{U}_{l_m,k}^{(c)} = \begin{bmatrix} \xi_{l_m,k}^{(c)} - \varphi_{l_m,k}^{(c)} & q_{l_m}^{(c)} (\tilde{\mathbf{h}}_{l_m,k}^{(c)})^H \mathbf{J}_k & \mathbf{0} \\ q_{l_m}^{(c)} \mathbf{J}_k^H \tilde{\mathbf{h}}_{l_m,k}^{(c)} & \xi_{l_m,k}^{(c)} \mathbf{I} & -\varepsilon_{l_m,k}^{(c)} q_{l_m}^{(c)} \mathbf{J}_k^H \\ \mathbf{0} & -\varepsilon_{l_m,k}^{(c)} q_{l_m}^{(c)} \mathbf{J}_k & \varphi_{l_m,k}^{(c)} \mathbf{I} \end{bmatrix}, \quad \forall k \in \mathcal{L} \setminus l. \quad (\text{C.4})$$

Obviously, problem (C.2) is an SDP.

B. Casting (40) as an SDP

Analogous to the procedure in the above subsection, problem (40) can be cast to an SDP as follows

$$\min_{q^{(d)}, v, \xi^{(d)}, \varphi^{(d)}} v \quad (\text{C.5a})$$

$$\text{s.t.} \quad \mathbf{U}_{l_n,k}^{(d)} \succeq 0, \quad \forall l_n \in \mathcal{D}, k \in \mathcal{L}, \quad (\text{C.5b})$$

$$\varphi_{l_n,k}^{(d)} \geq 0, \quad \forall l_n \in \mathcal{D}, k \in \mathcal{L}, \quad (\text{C.5c})$$

$$\sum_{l_n \in \mathcal{D}} \left\{ \sum_{k=1}^L (\xi_{l_n,k}^{(d)})^2 + \theta_{l_n} (q_{l_n}^{(d)})^2 - 2\phi_{l_n} q_{l_n}^{(d)} \right\} \leq v, \quad (\text{C.5d})$$

$$(40b), (40c), \quad (\text{C.5e})$$

where $v, \xi^{(d)} = (\xi_{1,1}^{(d)}, \dots, \xi_{1,L}^{(d)}, \dots, \xi_{L,N}^{(d)})^T, \varphi^{(d)} = (\varphi_{1,1}^{(d)}, \dots, \varphi_{1,L}^{(d)}, \dots, \varphi_{L,N}^{(d)})^T$ are auxiliary variables, and $\mathbf{U}_{l_n,k}^{(d)}$ is given by

$$\mathbf{U}_{l_n,k}^{(d)} = \begin{bmatrix} \xi_{l_n,k}^{(d)} - \varphi_{l_n,k}^{(d)} & q_{l_n}^{(d)} (\tilde{\mathbf{h}}_{l_n,k}^{(d)})^H \mathbf{J}_k & \mathbf{0} \\ q_{l_n}^{(d)} \mathbf{J}_k^H \tilde{\mathbf{h}}_{l_n,k}^{(d)} & \xi_{l_n,k}^{(d)} \mathbf{I} & -\varepsilon_{l_n,k}^{(d)} q_{l_n}^{(d)} \mathbf{J}_k^H \\ \mathbf{0} & -\varepsilon_{l_n,k}^{(d)} q_{l_n}^{(d)} \mathbf{J}_k & \varphi_{l_n,k}^{(d)} \mathbf{I} \end{bmatrix}, \quad \forall l_n \in \mathcal{D}, k \in \mathcal{L}. \quad (\text{C.6})$$

Then, Theorem 3 is proven.

APPENDIX D PROOF OF THEOREM 4

For notational convenience, denote the objective function of (32) by

$$\begin{aligned} V(\mathbf{p}, \mathbf{W}, \mathbf{f}, \mathbf{u}) &= \sum_{l=1}^L \max_{\Delta_l} \sum_{m=1}^M \left[\exp(u_{l_m}^{(c)} - 1) \text{MSE}_{l_m}^{(c)} - u_{l_m}^{(c)} \right] \\ &+ \sum_{l_n \in \mathcal{D}} \left[\exp(u_{l_n}^{(d)} - 1) \text{MSE}_{l_n}^{(d)} - u_{l_n}^{(d)} \right]. \end{aligned} \quad (\text{D.1})$$

To verify the convergence of Algorithm 1, it is first shown that (D.1) is lower bounded. By dropping the positive interference terms in the denominator of $\text{SINR}_{l_m}^{(c)}$ and using the Cauchy-Schwartz inequality, the SINR of CU l_m is upper bounded by

$$\begin{aligned} \text{SINR}_{l_m}^{(c)} &\leq \frac{p_{l_m}^{(c)} \left| (\mathbf{h}_{l_m,l}^{(c)})^H \mathbf{w}_{l_m} \right|^2}{N_0 \|\mathbf{w}_{l_m}\|^2} \leq \frac{p_{l_m}^{(c)} \|\mathbf{h}_{l_m,l}^{(c)}\|^2 \|\mathbf{w}_{l_m}\|^2}{N_0 \|\mathbf{w}_{l_m}\|^2} \\ &\leq \frac{P_{l_m}^{(c)}}{N_0} \left(\|\tilde{\mathbf{h}}_{l_m,l}^{(c)}\| + \varepsilon_{l_m,l}^{(c)} \right)^2 \triangleq \sigma_{l_m}^{(c)}, \quad \forall l_m \in \mathcal{C}. \end{aligned} \quad (\text{D.2})$$

Similarly, the SINR of D2D link l_n is upper bounded by

$$\text{SINR}_{l_n}^{(d)} \leq \frac{P_{l_n}^{(d)}}{N_0} \left| g_{l_n,l_n}^{(d)} \right|^2 \triangleq \sigma_{l_n}^{(d)}, \quad \forall l_n \in \mathcal{D}. \quad (\text{D.3})$$

Hence, a lower bound on the objective function of (23) is

$$\sum_{l_m \in \mathcal{C}} \ln \frac{1}{1 + \sigma_{l_m}^{(c)}} + \sum_{l_n \in \mathcal{D}} \ln \frac{1}{1 + \sigma_{l_n}^{(d)}}. \quad (\text{D.4})$$

As mentioned in Subsection III-A, the objective function of (23) is upper bounded by (D.1). Therefore, (D.4) is also a lower bound to (D.1).

Next, it is shown that the objective function of (32) is non-increasing in each iteration. Without loss of generality, denote the solution obtained in the j th iteration by $\{\mathbf{p}(j), \mathbf{W}(j), \mathbf{f}(j), \mathbf{u}(j)\}$. Then, in the $(j+1)$ th iteration, Algorithm 1 starts by solving (B.9), which is an SDP problem and can be optimally solved by applying CVX. So the optimal $\mathbf{W}(j+1)$ and $\mathbf{u}^{(c)}(j+1)$ can be obtained according to the relationships $\mathbf{W}_l(j+1) = \mathbf{J}_l(j+1) \Gamma_l(j+1)^{-1}$ and $u_{l_m}^{(c)}(j+1) = 1 + 2 \ln \gamma_{l_m}^{(c)}(j+1), \forall l \in \mathcal{L}, l_m \in \mathcal{C}_l$. Moreover, the optimal $f_{l_n}^{(d)}(j+1)$ and $u_{l_n}^{(d)}(j+1), \forall l_n \in \mathcal{D}$ can be calculated based on (A.6) and (29). This yields the following relationship

$$\begin{aligned} V(\mathbf{p}(j), \mathbf{W}(j), \mathbf{f}(j), \mathbf{u}(j)) &\geq V(\mathbf{p}(j), \mathbf{W}(j+1), \mathbf{f}(j+1), \mathbf{u}(j+1)). \end{aligned} \quad (\text{D.5})$$

For given $\mathbf{W}(j+1), \mathbf{f}(j+1)$ and $\mathbf{u}(j+1)$, the optimal $\mathbf{p}(j+1)$ can be obtained by solving SDPs (C.2) and (C.5), and using

the relationships $p_{l_m}^{(c)}(j+1) = [q_{l_m}^{(c)}(j+1)]^2$, $\forall l_m \in \mathcal{C}$, $p_{l_n}^{(d)}(j+1) = [q_{l_n}^{(d)}(j+1)]^2$, $\forall l_n \in \mathcal{D}$. Then,

$$\begin{aligned} & V(\mathbf{p}(j), \mathbf{W}(j+1), \mathbf{f}(j+1), \mathbf{u}(j+1)) \\ & \geq V(\mathbf{p}(j+1), \mathbf{W}(j+1), \mathbf{f}(j+1), \mathbf{u}(j+1)). \end{aligned} \quad (\text{D.6})$$

Incorporating (D.5) and (D.6) yields

$$\begin{aligned} & V(\mathbf{p}(j), \mathbf{W}(j), \mathbf{f}(j), \mathbf{u}(j)) \\ & \geq V(\mathbf{p}(j+1), \mathbf{W}(j+1), \mathbf{f}(j+1), \mathbf{u}(j+1)), \end{aligned} \quad (\text{D.7})$$

which indicates that the objective function of (32) is non-increasing in each iteration. Noting the fact that (D.1) is lower bounded by (D.4), it can be concluded that Algorithm 1 converges to a suboptimal solution of problem (32).

REFERENCES

- [1] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," 2002.
- [2] Cisco, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021 White Paper*, San Jose, CA, USA, Feb. 2017. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf>
- [3] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communication in LTE-advanced networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1923–1940, Dec. 2015.
- [4] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, Fourth Quart. 2014.
- [5] X. Lin, R. W. Heath, and J. G. Andrews, "Spectral efficiency of massive MIMO systems with D2D underlay," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, UK, Jun. 2015, pp. 4345–4350.
- [6] A. He, L. Wang, Y. Chen, K.-K. Wong, and M. El-kashlan, "SE and EE of Uplink D2D Underlaid Massive MIMO Cellular Networks with Power Control," in *Proc. IEEE WCNC*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.
- [7] X. Lin, J. 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.
- [8] T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "LTE for public safety networks," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 106–112, Feb. 2013.
- [9] L. Wei, R. Q. Hu, T. He, and Y. Qian, "Device-to-device (D2D) communications underlying MU-MIMO cellular networks," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Atlanta, GA, USA, Dec. 2013, pp. 4902–4907.
- [10] D. Zhu, J. Wang, A. L. Swindlehurst, and C. Zhao, "Downlink resource reuse for device-to-device communications underlying cellular networks," *IEEE Sig. Process. Lett.*, vol. 21, no. 5, pp. 531–534, May 2014.
- [11] H. Xu, Z. Yang, N. Huang, J.-Y. Wang, J. Shi, and M. Chen, "Channel allocation and power control in D2D uplink underlaid cellular networks," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Washington, DC, USA, Dec. 2016, pp. 1–6.
- [12] T. D. Hoang, L. B. Le, and T. Le-Ngoc, "Resource allocation for D2D communications under proportional fairness," in *Proc. IEEE GLOBECOM*, Austin, TX, USA, Dec. 2014, pp. 1259–1264.
- [13] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlying cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3541–3551, Aug. 2013.
- [14] W. Zhao and S. Wang, "Resource allocation for device-to-device communication underlying cellular networks: An alternating optimization method," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1398–1401, Aug. 2015.
- [15] B. Fang, Z. Qian, W. Zhong, W. Shao, and H. Xue, "Coordinated precoding for D2D communications underlay uplink MIMO cellular networks," in *Proc. IEEE International Conference on Communications in China (ICCC)*, Shenzhen, China, Nov. 2015, pp. 1–5.
- [16] H. Feng, H. Wang, X. Xu, and C. Xing, "A tractable model for device-to-device communication underlying multi-cell cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, 2014, pp. 587–591.
- [17] D. Della Penda, A. Abrardo, M. Moretti, and M. Johansson, "Potential games for subcarrier allocation in multi-cell networks with D2D communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, 2016, pp. 1–6.
- [18] X. Lin, R. W. Heath, and J. G. Andrews, "The interplay between massive MIMO and underlaid D2D networking," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3337–3351, June, 2015.
- [19] M. B. Shenouda and T. N. Davidson, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 1015–1024, Aug. 2008.
- [20] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Sig. Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.
- [21] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.
- [22] A. Tajer, N. Prasad, and X. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 235–251, Jan. 2011.
- [23] M. F. Hanif, L.-N. Tran, A. Tölli, M. Juntti, and S. Glisic, "Efficient solutions for weighted sum rate maximization in multicellular networks with channel uncertainties," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5659–5674, Nov. 2013.
- [24] W. Xu, Y. Cui, H. Zhang, G. Y. Li, and X. You, "Robust beamforming with partial channel state information for energy efficient networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2920–2935, Dec. 2015.
- [25] H. Tang, Z. Ding, S. B. Yoo, and J. Hamalainen, "Outage constrained joint precoding for D2D underlay cellular networks," in *Proc. IEEE GLOBECOM*, Atlanta, GA, USA, Dec. 2013, pp. 3540–3545.
- [26] A. Memmi, Z. Rezk, and M.-S. Alouini, "Power control for D2D underlay cellular networks with channel uncertainty," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1330–1343, Feb. 2017.
- [27] W. Fu, R. Yao, F. Gao, J. C. Li, and M. Lei, "Robust null-space based interference avoiding scheme for D2D communication underlying cellular networks," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 4158–4162.
- [28] M. J. Rahman and L. Lampe, "Robust transceiver optimization for underlay device-to-device communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, UK, Jun. 2015, pp. 7695–7700.
- [29] C. Pan, W. Xu, J. Wang, H. Ren, W. Zhang, N. Huang, and M. Chen, "Pricing-based distributed energy-efficient beamforming for MISO interference channels," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 710–722, Apr. 2016.
- [30] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint precoding and RRH selection for user-centric green MIMO C-RAN," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2891–2906, May 2017.
- [31] M. Wellens, J. Wu, and P. Mähönen, "Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio," in *2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications*. IEEE, 2007, pp. 420–427.
- [32] M. Zulfhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlying lte network," in *IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2010, pp. 368–375.
- [33] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Losow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, May 2014.
- [34] X. Rao, L. Ruan, and V. K. Lau, "CSI feedback reduction for MIMO interference alignment," *IEEE Trans. Sig. Process.*, vol. 61, no. 18, pp. 4428–4437, Sep. 2013.
- [35] H. Xu, N. Huang, Z. Yang, J. Shi, B. Wu, and M. Chen, "Pilot allocation and power control in D2D underlay massive MIMO systems," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 112–115, Jan. 2017.
- [36] H. Xu, W. Xu, Z. Yang, J. Shi, and M. Chen, "Pilot reuse among D2D users in D2D underlaid massive MIMO systems," *IEEE Trans. Veh. Tech.*, 2017. [Online]. Available: <http://ieeexplore.ieee.org/abstract/document/7983007/>
- [37] Y.-J. Kim, M.-G. Song, Y.-S. Cho, and G.-H. Im, "Decentralized beamformer design with limited multi-cell cooperation for interference channel of cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4142–4152, Aug. 2013.

- [38] S. A. Vorobyov, A. B. Gershman, Z.-Q. Luo, and N. Ma, "Adaptive beamforming with joint robustness against mismatched signal steering vector and interference nonstationarity," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 108–111, Feb. 2004.
- [39] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
- [40] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, 1996.
- [41] M. F. Hanif, L.-N. Tran, A. Tolli, and M. Juntti, "Computationally efficient robust beamforming for SINR balancing in multicell downlink with applications to large antenna array systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1908–1920, June 2014.
- [42] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1037–1052, Apr. 2004.
- [43] M. Grant, S. Boyd, and Y. Ye, *CVX: Matlab Software for Disciplined Convex Programming, Version 2.1*. [Online] Available: <http://cvxr.com/cvx/>.
- [44] E. U. T. R. Access, "Further advancements for E-UTRA physical layer aspects," 3GPP TR 36.814, Tech. Rep., 2010.
- [45] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, Jun. 1994, vol. 15.
- [46] A. Ben-Tal, A. Nemirovski, and C. Roos, "Extended matrix cube theorems with applications to μ -theory in control," *Math. Oper. Res.*, vol. 28, no. 3, pp. 497–523, 2003.
- [47] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge U.K.: Cambridge Univ. Press, 2012.