Optimal Resource Allocation for Full-Duplex IoT Systems Underlaying Cellular Networks with Mutual SIC NOMA

Antoine Kilzi, Joumana Farah, Charbel Abdel Nour, Catherine Douillard

Optimal Resource Allocation for Full-Duplex IoT Systems Underlaying Cellular Networks with Mutual SIC NOMA

Antoine Kilzi, Joumana Farah, Charbel Abdel Nour, Catherine Douillard

***Abstract*—Device-to-device (D2D) and non-orthogonal multiple access (NOMA) are promising technologies to meet the challenges of the next generations of mobile communications in terms of network density and diversity for internet of things (IoT) services. This paper tackles the problem of maximizing the D2D sum-throughput in an IoT system underlaying a cellular network, through optimal channel and power allocation. NOMA is used to manage the interference between cellular users and full-duplex (FD) IoT devices. To this aim, mutual successive interference cancellation (SIC) conditions are identified to allow simultaneously the removal of the D2D devices interference at the level of the base station and the removal of the cellular users (CU) interference at the level of D2D devices. To optimally solve the joint channel and power allocation (PA) problem, a time-efficient solution of the PA problem in the FD context is elaborated. By means of graphical representation, the complex non-convex PA problem is efficiently solved in constant time complexity. This enables the global optimal resolution by successively solving the separate PA and channel assignment problems. The performance of the proposed strategy is compared against the classical state-of-the-art FD and HD scenarios, where SIC is not applied between CUs and IoT devices. The results show that important gains can be achieved by applying mutual SIC NOMA in the IoT-cellular context, in either HD or FD scenarios.**

***Index Terms*—Non-orthogonal multiple access, D2D, IoT, mutual SIC, full-duplex, half-duplex, residual self-interference.**

I. Introduction

FOLLOWING the growth in the number of connected devices in recent years, unprecedented highs are expected for the near future [1]. Coupled with the expected increase in data traffic and the limited available spectrum, the corresponding network densification will require novel efficient solutions to supply the ever increasing demand. Full-duplex (FD) communication combined with device-to-device (D2D) communication represent an attractive solution to leverage the challenges of future generation networks.

By enabling direct communication between nearby devices avoiding the transit through base stations (BS) or gateways [2], D2D communication leverages network resources and enables increasing the number of connected devices. In FD communication, a node can send and receive simultaneously using the same frequency resource. The achieved gain can go up to a virtual two-fold increase in spectral efficiency (SE) compared to half-duplex (HD) send-then-receive systems. In return, a self-interference (SI) is incurred due to the transmitted signal looping back into the receiver, thus limiting its appeal compared to HD to the point where this latter may even outperform FD in some cases. Nonetheless, the improvement in antenna architecture and in SI cancellation circuitry dramatically reduces the residual self-interference (RSI) [3–5], advocating for the use of FD for future communication standards.

The increasing demand for higher spectral efficiency and massive connectivity for internet of things (IoT) steered research towards non-orthogonal multiple access (NOMA) techniques. The sharing of multiple devices over the same time and frequency resource enables important SE gains, lower latency communications, and increased number of connected IoT devices [6–12]. In power-domain NOMA, signals are differentiated in the power dimension [13], where superposition coding of users signals is used at the transmitter, and successive interference cancellation (SIC) is performed at the receiver side. At the level of a receiver, the message with the highest power is decoded first and then subtracted from the total received signal, then the message with the second highest power is extracted and so on until the user decodes its own message. Works such [14–16] used NOMA and FD for cooperative relaying as well as secrecy provision, calling on to virtual MIMO as physical layer security enabler for fifth generation centric IoT applications. In this paper, we study the resource allocation problem for D2D communications systems underlaying cellular networks using the NOMA technique coupled with FD transmission scenarios. This being said, the broader problem of resource optimization in IoT could be tackled via multiple other tools such as next generation reconfigurable intelligent surfaces (RIS) [17], quantum computing inspired metaheuristics [18], and much more.

A. Related Works

Recently, attention was focused on the combination of NOMA with D2D communications in underlay mode. The study in [19] considers resource block assignment and power allocation (PA) in a downlink NOMA system with D2D. HD

J. Farah is with the Department of Electricity and Electronics, Faculty of Engineering, Lebanese University, Roumieh, Lebanon (joumana-farah@ul.edu.lb).

A. Kilzi, C. Abdel Nour and C. Douillard are with IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238 Brest, France, (email: antoine.kilzi@imt-atlantique.fr; charbel.abdelnour@imt-atlantique.fr; catherine.douillard@imt-atlantique.fr).

This work has been funded with support from IMT Atlantique and the Lebanese University Research Support Program.

is used in the D2D pairs, and CUs are grouped in NOMA clusters. The influence of the HD-D2D users over the SIC decoding orders of CUs is accounted for in both the block assignment and the PA phase, because the interference they generate may change the decoding order. However, NOMA SIC is not used to decode the interfering signals of the collocated D2D pairs. The same is true for [20], but additional power constraints are introduced on the D2D pairs to maintain the same SIC decoding orders at CUs as for the case of D2D-disabled systems. The work in [21] introduces the concept of D2D group, where a D2D transmitter communicates with multiple D2D receivers via NOMA. To maximize the network sum-throughput, sub-channel allocation is conducted using many-to-one matching for CU-D2D grouping and optimal PA is approximated iteratively via successive convex approximation. When limiting the number of multiplexed D2Ds to one per CU user, the work in [22] provides a joint D2D-CU grouping and PA strategy for energy efficiency maximization: the Kuhn-Munkres technique is applied successively for channel allocation, while optimal PA is obtained using the Karush-Kuhn-Tucker conditions. In all the preceding studies, NOMA is applied either between the CU users [20], or between users of the same D2D group [21], [22], but the interference cancellation of the D2D signals at the level of CU users (and inversely) is not considered. At most, attention is given towards managing the SIC decoding order at the level of the CUs in [19], [20], or at the level of the D2D receivers in [21], [22].

The work in [23] tackles the problem of HD-D2D throughput maximization in an uplink system where NOMA is used between D2D and CU users. If the D2D causes strong interference on the BS, its signal can be decoded then subtracted before retrieving the CU signal. However, FD-D2D is not studied and SIC occurs only at the level of the BS, i.e. not at the devices levels. Besides, the information-theoretic conditions for SIC feasibility are not considered in the study. In [24], an efficient graph-based scheme is proposed to maximize the D2D sum-rate of an uplink system. To that end, an interlay mode is introduced to HD-D2D communication where a D2D pair can join a NOMA group to remove the interference between it and the cellular NOMA users. However, the conditions for applying SIC - and thus for determining the SIC decoding order - are only conditioned by the ascending order of channel gains between the senders and the receivers. In other words, the interfering signals that can be canceled are the ones that are attributed channel gains better than that of the useful signal, regardless of their power level at reception. This may lead to outage probabilities of one if no PA measures are taken to guarantee SIC stability as shown in [10]. The work in [25] incorporates NOMA into D2D cellular networks to maximize system connectivity. Unlike [24], the D2D NOMA-aided modes are defined according to the SIC orders at the level of the D2D and the BS. The SIC decoding orders are governed by the strong interfering signal which is bound to the channel conditions as well as the used PA. The optimal PA and mode selection are solved in the presence of decoding signal-to-interference-plus-noise ratios (SINRs) threshold constraints, then the user pairing problem is turned into a min-cost max-flow problem which is solved by the Ford–Fulkerson algorithm. However, the case of FD-D2D NOMA-aided network was not addressed neither in this study, nor in the entire literature combining NOMA and D2D.

B. Contributions

In this paper, we study the combination of NOMA with HD-D2D and FD-D2D systems using mutual SIC. We introduced the concept of mutual SIC in our previous works [26–28] where we showed that the signals of two or more users multiplexed in NOMA, and powered by distributed antennas, can be decoded and removed at the level of every user in the NOMA cluster. In a D2D cellular scenario, this translates into removing the interference of the D2D devices at the level of the BS, and the interference of CU users at the level of D2D pairs. The objective of the paper is to maximize the D2D sum-throughput, through joint optimal channel and power allocation, while maintaining the quality of service (QoS) requirement for all CU users. The main contributions of the paper are summarized as follows:

- We derive the power multiplexing conditions (PMC) and SIC conditions allowing for the interference cancellation between D2D and CU users. The SIC constraints are the set of conditions that make mutual SIC feasible from the information theory perspective, i.e. the conditions on achievable rates at the respective levels of the users. The PMCs are the set of conditions that make the mutual SIC technique feasible from a practical implementation perspective, i.e. they guarantee that the power level of the message to be decoded, m_d, at a given SIC iteration, is greater than the sum of the remaining messages, so that m_d can be detected and discerned from the noise plus background interference. This guarantees SIC stability, since every signal is ensured to be the dominant signal during its decoding [9][10].
- We show that the PMCs imply the SIC conditions for both HD and FD transmission modes, which greatly reduces the PA problem complexity for the case of FD-SIC.
- We solve analytically the PA problem for all the transmission strategies, especially for the case of FD-SIC where an efficient procedure is provided to optimally solve the D2D rate maximization problem with constant time complexity.
- We show that the optimal solution of the joint PA and channel allocation problem can be achieved by successively resolving the PA problem and then the channel allocation problem.

The remainder of the paper is organized as follows: section II presents the system model and formulates the joint channel and power allocation problem, decomposing the resource allocation into separate PA and channel allocation problems. The PA problems of FD and HD without SIC (FD-NoSIC, HD-NoSIC) are solved in section III, while the PA problem with SIC is reformulated for HD and FD (HD-SIC, FD-SIC) in section IV. Mutual SIC PA is solved for the case of HD transmission in section V. In sections VI, VII, and VIII the conditions of mutual SIC for FD-D2D are derived, the problem constraint

reduction is performed, and then the proposed geometrical resolution is exposed, allowing for a cost-effective resolution of the FD-SIC PA problem. The channel allocation is discussed in section IX. Simulation results are presented in section X, and conclusions are drawn in section XI.

II. System Model

In this work, we consider the integration of a D2D underlay IoT system into a pre-established cellular network. The base network consists of K CU users transmitting over their assigned uplink (UL) channels separately, with a maximum of one channel per CU user. The system bandwidth is divided into $N \geq K$ channels, and the IoT system is constituted of a total of D D2D pairs ($D \leq K$) exchanging data over a subset of the K UL cellular channels, with a single D2D pair per cellular channel. The D2D pairs can exchange data either in HD or in FD mode, while the CUs are always in HD. The objective of the study is to perform optimal D2D channel allocation and power control, such that the obtained D2D-CU pairs yield maximum D2D sum-throughput while guaranteeing the required rates of the collocated CU users. To that end, let $\mathbb{C}$ denote the set of cellular users, $\mathbb{C} = \{u_1, u_2, \ldots, u_K\}$, and $\mathbb{D}$ the set of D2D pairs, $\mathbb{D} = \{(d_{1,1}, d_{1,2}), (d_{2,1}, d_{2,2}), \ldots, (d_{D,1}, d_{D,2})\}$. A schematic of the network is presented in Fig. 1, where $d_{n,1}$ and $d_{n,2}$ are the n^{th} IoT device pair transmitting in FD mode.

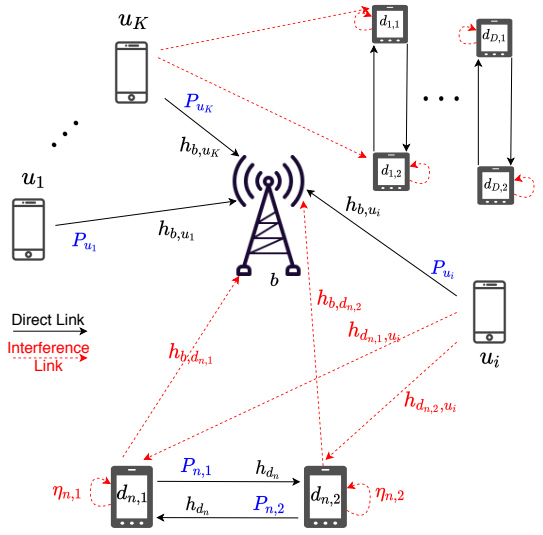


Figure 1: IoT system with D D2D pairs underlaying a cellular network with K CU users.

The interference channel gains between a CU u_i, on the one hand, and $d_{n,1}$ and $d_{n,2}$ on the other hand, are denoted by $h_{d_{n,1},u_i}$ and $h_{d_{n,2},u_i}$ respectively. The direct link between the CU u_i and the BS b has a squared channel gain denoted by h_{b,u_i}. The message m_{u_i}, transmitted by u_i with power P_{u_i}, reaches the BS with a power level $P_{u_i} h_{b,u_i}$, and causes an interference level of $P_{u_i} h_{d_{n,1},u_i}$ and $P_{u_i} h_{d_{n,2},u_i}$ at $d_{n,1}$ and $d_{n,2}$ respectively. Each device $d_{n,j}$ ($j \in \{1,2\}$) of the n^{th} D2D pair can transmit a message $m_{n,j}$ of power $P_{n,j}$ to the other D2D user and suffers from both the interference of user u_i and its RSI power $\eta_{n,j} P_{n,j}$, with $\eta_{n,j}$ denoting the SI cancellation capability. The D2D inter-user channel gain is denoted by h_{d_n} and the interference channel gains from $d_{n,1}$ and $d_{n,2}$ to the BS are denoted by $h_{b,d_{n,1}}$ and $h_{b,d_{n,2}}$ respectively. In this study, a frequency-non-selective channel is assumed, so that the channel gains are independent from the sub-band frequency and account only for large-scale fading including path-loss and shadowing. Table I contains the main notations used in the paper.

Table I: Notation Table

K	Total number of CUs	d_{max}	Maximum distance between D2Ds
N	Number of uplink channels	$R_{u,min}$, $P_{u,min}$	Minimum required rate and power for CU u
$\mathbb{C}$	Set of CUs	η	SI cancelation factor
$\mathbb{D}$	Set of D2D pairs	σ^2	White noise power
O	D2D chanel allocation matrix	$P_{1,M}$, $P_{2,M}$, $P_{u,M}$	Maximum transmit power of d_1, d_2 and u resp.
$\mathcal{R}_{D2D}$	Maximum achievable D2D rate for every D2D-CU couple	$\mathcal{PL}_1, \mathcal{PL}_2$ $\mathcal{PL}_3, \mathcal{PL}_4$	Planes associated to PMC_1, PMC_2, PMC_3, and PMC_4 resp.

In this work, it is assumed that, prior to resource allocation and data exchange, a D2D discovery phase [29] takes place in the system, during which the D2D devices inform the BS about their desire to initiate a D2D link, and forward to the BS their estimates of the D2D-CU links ($h_{d_{n,1},u_i}, h_{d_{n,2},u_i}$), as well as the D2D links (h_{d_n}). Therefore, the BS is assumed to have perfect knowledge of the long-term evolution of the different channel gains, through signaling exchange between the different entities. The BS then performs resource allocation based on these estimated channel gains to optimally pair the D2Ds to CUs and to instruct D2D-CU pairs of the required transmit powers on their collocated channels, according to the selected transmission scenario.

A. Formulation of the Joint Channel and Power Allocation Problem

Let O be the channel allocation matrix, with the element $o(n,i)$ at the n^{th} row and i^{th} column equaling one if D2D pair n is collocated with CU u_i and zero otherwise. Also, let $\mathcal{R}_{D2D}(n,i)$ be the maximum achievable D2D rate of the pair n when collocated with u_i. Channel allocation is performed such that a D2D pair is multiplexed over a single UL channel, on the one hand, and such that a maximum of one D2D pair is multiplexed over a UL channel, on the other. The joint channel and power allocation problem for the maximization of the total D2D throughput can be cast as:

$$\max_{\{O,P_{n,1},P_{n,2},P_{u_i}\}} \left(\sum_{i=1}^{K} \sum_{n=1}^{D} o(n,i) \times \mathcal{R}_{D2D}(n,i) \right)$$
$$\text{s.t.} \sum_{i=1}^{K} o(n,i) = 1, \forall n \in \{1,\ldots,D\}, \sum_{n=1}^{D} o(n,i) \leq 1, \forall i \in \{1,\ldots,K\} \quad (1)$$

with $\mathcal{R}_{D2D}(n,i)$ the solution to:

$$\max_{\{P_{n,1},P_{n,2},P_{u_i}\}} R_{D2D}(n,i), \quad (2)$$

$$\text{s.t.} \quad R_{u_i} \geq R_{u_i,min}, \quad (2a) \qquad P_{u_i} \leq P_{u_i,M}, \quad (2c)$$
$$P_{n,1} \leq P_{n,1,M}, \quad (2b) \qquad P_{n,2} \leq P_{n,2,M}, \quad (2d)$$

with $R_{u_i,min}$ the minimum target rate of u_i, R_{u_i} its achieved rate, and $R_{D2D}(n,i)$ the D2D rate, sum of achieved rates by $d_{n,1}$ ($R_{d_{n,1}}$) and $d_{n,2}$ ($R_{d_{n,2}}$). The M in the subscripts of $P_{u_i,M}, P_{n,1,M}, P_{n,2,M}$ refers to the maximum transmit powers of $u_i, d_{n,1}$ and $d_{n,2}$ respectively.

From the structure of Problem (1), and since CU users are allocated orthogonal channels, the performance of a given D2D-CU pair is independent from the network activity over the remaining channels in the system. Therefore, one can optimize the throughput of all possible D2D-CU pairs, constructing a $D \times K$ table of achievable rates, and then proceed to the optimal allocation of channels to D2Ds in a second phase, thus pairing D2Ds to CUs based on their achievable rate to maximize the D2D sum-throughput in the system. The aim of the following sections is to obtain the optimal PAs of the four transmission methods FD-NoSIC, HD-NoSIC, HD-SIC and FD-SIC in order to build their corresponding tables of achievable rates $\mathcal{R}_{D2D}^{FD-NoSIC}, \mathcal{R}_{D2D}^{HD-NoSIC}, \mathcal{R}_{D2D}^{HD-SIC}$, and $\mathcal{R}_{D2D}^{FD-SIC}$ respectively. Based on these tables, the optimal channel allocation is conducted in Section IX. The overall procedure spanning from the PA problems formulations in section III, to the optimal channel assignment in section IX, is depicted in Fig. 2.

III. Power Allocation for No-SIC Scenarios

From hereinafter, since the optimal D2D rate of all (n,i) couples is to be computed and because the resolution of the PAs is independent of the elected D2D-CU couple, we drop the indices relative to a specific D2D pair and CU user. Hence, user u designates the CU user at hand, and d_1 and d_2 are the corresponding D2D pair. The involved channels gains are therefore denoted as h_d, h_{b,d_1}, h_{b,d_2}, $h_{d_1,u}$, $h_{d_2,u}$ and $h_{b,u}$, and the transmit powers of d_1, d_2 and u are P_1, P_2, P_u, with their power limits $P_{1,M}, P_{2,M}, P_{u,M}$.

A. FD-NoSIC

In FD, d_1 and d_2 transmit simultaneously, thus they both suffer from RSI. Since, in this method, SIC is not attempted at the levels of d_1, d_2 and the BS, the SINRs at the level of the BS and the D2D users are given by:

$$SINR_b = \frac{P_u h_{b,u}}{P_1 h_{b,d_1} + P_2 h_{b,d_2} + \sigma^2}$$
$$SINR_{d_1} = \frac{P_2 h_d}{P_u h_{d_1,u} + \eta_1 P_1 + \sigma^2}, SINR_{d_2} = \frac{P_1 h_d}{P_u h_{d_2,u} + \eta_2 P_2 + \sigma^2}, \quad (3)$$

with σ^2 the additive Gaussian noise power. The achieved rates are expressed according to the Shannon capacity theorem:

$$R_u = B \log_2(1 + SINR_b), \quad (4)$$
$$R_{d_1} = B \log_2(1 + SINR_{d_1}), \quad R_{d_2} = B \log_2(1 + SINR_{d_2}), \quad (5)$$

with B the bandwidth of each UL channel resource. Due to the interference terms in (3), Problem (2) is non-convex. To solve it, a geometrical representation can be used, leading to the analytical global solution in [30]. This method is adopted in our work to derive the results of the FD-NoSIC scenario in the performance assessment section.

B. HD-NoSIC

The time slot is now divided into two equal half-time slots where d_1 and d_2 alternately transmit and receive information. To maximize the total D2D rate, the optimization is conducted separately in the two half-time slots. In the first half, d_1 transmits information ($P_2 = 0$). In Problem (2), the objective function and CU rate are now:

$$R_{D2D,1} = R_{d_2} = B \log_2(1 + \frac{P_1 h_d}{P_{u,1} h_{d_2,u} + \sigma^2}),$$
$$R_{u,1} = B \log_2(1 + \frac{P_{u,1} h_{b,u}}{P_1 h_{b,d_1} + \sigma^2}).$$

Also, Problem (2) is constrained only by eqs. (2a) to (2c). Note that $P_{u,1}$ is the transmit power of u during the first half-time slot. $R_{D2D,1}$ is strictly increasing with P_1 and decreasing with $P_{u,1}$; therefore, to maximize $R_{D2D,1} = R_{d_2}$, P_1 should be increased and $P_{u,1}$ decreased as long as $R_{u,1}$ satisfies the minimum rate condition of the CU. Consequently, P_1 should be increased as much as possible and then $P_{u,1}$ is obtained as a function of P_1 ($P_{u,1} = f(P_1)$) by enforcing an equality between $R_{u,1}$ and $R_{u,min}$. If for $P_1 = P_{1,M}$, $f(P_{1,M}) \leq P_{u,M}$, the couple $(P_{1,M}, f(P_{1,M}))$ is retained as the $(P_1,\ P_{u,1})$ solution; otherwise, the couple $(f^{-1}(P_{u,M}), P_{u,M})$ delivers the best solution. The same reasoning is applied for the second half-time slot (where $P_1 = 0$) to maximize $R_{D2D,2} = R_{d_1}$. The total user u and D2D rates are given by:

$$R_u = \frac{1}{2} R_{u,1} + \frac{1}{2} R_{u,2}, \quad R_{D2D} = \frac{1}{2} R_{d_1} + \frac{1}{2} R_{d_2}.$$

IV. PA Problem Modification for HD and FD with mutual SIC (HD-SIC and FD-SIC)

Using a SIC receiver at the level of the BS and the D2D users, interfering messages can be decoded then subtracted from the received message, canceling thereby the interference in both FD and HD scenarios. Let m_1 and m_2 be the messages transmitted by the devices d_1 and d_2. In the case of FD, the BS can decode and subtract successively m_1 then m_2, or inversely, before proceeding to the decoding of m_u (the message transmitted by the CU); hence, two decoding orders are possible. Users d_1 and d_2 can also remove the interference of u, leading to the following SINR expressions:

$$SINR_{d_1} = \frac{P_2 h_d}{\eta_1 P_1 + \sigma^2}, SINR_{d_2} = \frac{P_1 h_d}{\eta_2 P_2 + \sigma^2}, SINR_b = \frac{P_u h_{b,u}}{\sigma^2}.$$

The SINRs are replaced in eqs. (4) and (5) to obtain R_u and $R_{D2D} = R_{d_1} + R_{d_2}$ that will be used in Problem (2). For the case of HD, the SINRs in the first half-time slot are:

$$SINR_{d_2} = P_1 h_d / \sigma^2, \qquad SINR_b = P_u h_{b,u} / \sigma^2.$$

In the second half-time slot, $SINR_b$ is the same and $SINR_{d_1} = P_2 h_d / \sigma^2$. Problem (2) is now reformulated in each time slot by expressing the rates using the present SINRs. However,

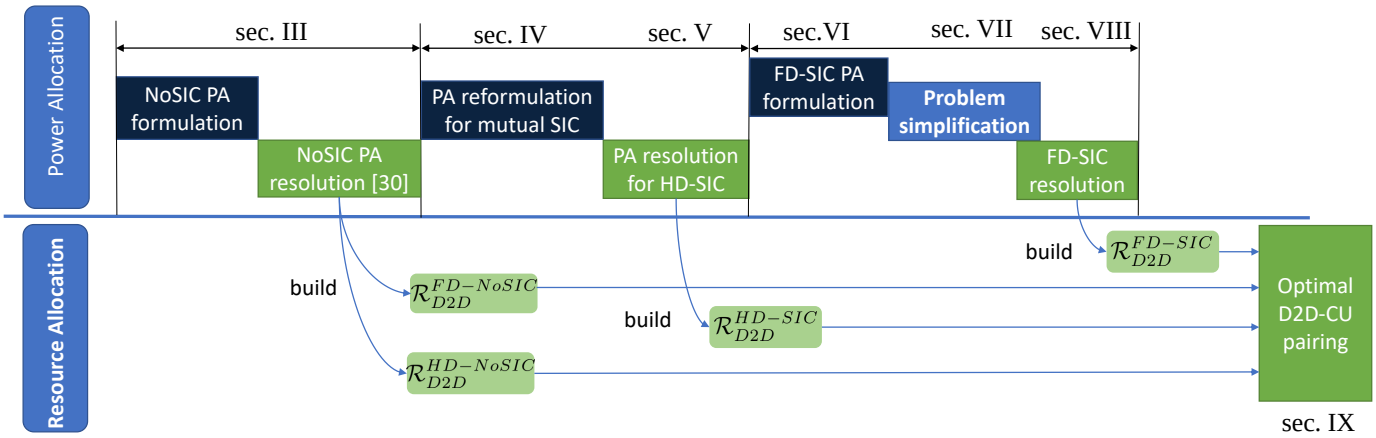


Figure 2: Paper workflow for the resolution of the D2D channel and power allocation problem.

additional constraints relative to the SIC feasibility must be added to the problem. They are derived in the next sections for HD-SIC, then for FD-SIC. Afterwards, Problem (2) is solved for the two SIC scenarios.

V. PA for HD-SIC scenario

Consider the first half-time slot, where u and d_1 are transmitting and b and d_2 are receiving. Hereafter, we develop the mutual SIC constraints between b and d_2 (as a receiver). Let $SINR_i^{m_j}$ be the SINR of the message m_j at the level of user i (i is either d_1, d_2 or b, and j is either 1, 2 or u). For b to successfully decode the message m_1 transmitted by d_1 to d_2, the received rate of m_1 at the level of b must be greater than the rate of m_1 at the level of d_2. Thus, we must have: $SINR_b^{m_1} > SINR_{d_2}^{m_1}$. Similarly, the rate condition for the decoding of m_u at the level of d_2 is derived from the condition $SINR_{d_2}^{m_u} > SINR_b^{m_u}$. This situation is equivalent to the case of two different radio resource heads (RRHs) transmitting both messages to two separate receivers, which was studied in [26]. It was shown that the SINR conditions lead to:

$$h_{b,d_1} h_{d_2,u} > h_d h_{b,u}. \tag{6}$$

In addition to condition (6), the PMCs must be verified, in order to ensure that the message to be decoded first at the level of a receiver has a higher power level than that of the remaining message [28]. The PMCs for the decoding of m_u and m_1 at the level of d_2 and b are given by:

$$\left.\begin{aligned} P_{u,1} h_{d_2,u} > P_1 h_d \\ P_1 h_{b,d_1} > P_{u,1} h_{b,u} \end{aligned}\right\} \Rightarrow A = \frac{h_d}{h_{d_2,u}} < \frac{P_{u,1}}{P_1} < \frac{h_{b,d_1}}{h_{b,u}} = B. \tag{7}$$

Note that (6) is satisfied if (7) is satisfied, since (6) is equivalent to $A < B$. Therefore, the PMCs encompass the rate conditions while being more restrictive. Problem (2) now only includes the additional constraint (7) for the first time slot. The HD-SIC rate expressions are as follows:

$$R_{D2D,1} = R_{d_2} = B \log_2(1+\frac{P_1 h_d}{\sigma^2}), R_{u,1} = B \log_2(1+\frac{P_{u,1} h_{b,u}}{\sigma^2}).$$

Maximizing R_{d_2} lies in the increase of P_1. Also, guaranteeing the CU rate $R_{u,min}$ can be achieved by setting $P_{u,1}$ to $P_{u,m} = (2^{\frac{R_{u,min}}{B}} - 1)\sigma^2 / h_{b,u}$. However, due to the PMCs, the increase in P_1 is very likely to increase $P_{u,1}$ according to the range of allowed values in (7), leading to an excess of CU rate. Since maximization of network throughput (i.e. sum of D2D and CU rates) is not the objective of this study, we select from the range of admissible $R_{u,1}$ values, the one closest to $R_{u,min}$. With that criterion in mind, the power allocation problem for D2D rate maximization is solved by increasing P_1 as much as possible (possibly until $P_{1,M}$) and adjusting $P_{u,1}$ accordingly. The proposed PA procedure, illustrated in Fig.

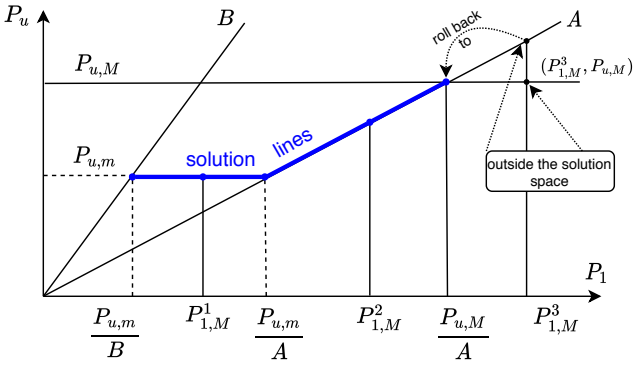


Figure 3: Schematic of the solution space to the HD-SIC PA problem, for different $P_{1,M}$ values.

3, operates as follows: if $P_{1,M} < P_{u,m}/A$, keep the couple $(P_1 = P_{1,M}, P_{u,1} = P_{u,m})$. This case is represented by the example $P^1_{1,M}$ on the horizontal blue line in Fig. 3. If this is not the case, check if $AP_{1,M} > P_{u,M}$. If yes (cf. example $P^3_{1,M}$ in Fig. 3), the solution is $(P_{u,M}/A, P_{u,M})$; if not (cf. example $P^2_{1,M}$), the solution is $(P_{1,M}, AP_{1,M})$. Restricting the solution space to the blue lines in Fig. 3 guarantees that the CU always transmits at the minimum necessary power that respects the problem constraints. Note that if $P_{1,M}$ is too low ($< P_{u,m}/B$), the problem is not feasible even when (6) is verified.

For the second time slot, the same methodology is followed, where the PMCs and the new necessary and sufficient channel conditions are given by:

$$h_{d_1,u} h_{b,d_2} > h_{b,u} h_d, \tag{8}$$

$$A^{'} = \frac{h_d}{h_{d_1,u}} < \frac{P_{u,2}}{P_2} < \frac{h_{b,d_2}}{h_{b,u}} = B^{'}. \quad (9)$$

As a conclusion, in the HD-SIC scenario, the system checks for the validity of the channel condition corresponding to the half-time slot before going through the procedure described above. If the channel condition is not favorable or if no solution exists (i.e. $P_{u,m} > P_{u,M}$ or $P_{1,M} < P_{u,m}/B$ for the first half, and $P_{2,M} < P_{u,m}/B^{'}$ for the second half), the system reverts to the HD-NoSIC solution of section III-B. This leads to four combinations of SIC/NoSIC procedures, two for every half-time slot, and they are all included in the HD-SIC algorithm.

VI. Derivation of the SIC conditions for FD mutual SIC

In this scenario, we are looking for the conditions that allow d_1 to decode m_u, d_2 to decode m_u, and b to decode m_1 and m_2. As already mentioned, two decoding orders are possible at the level of b.

A. First decoding order: b decodes m_2 then m_1

We first start by studying the mutual SIC constraints between b and d_1 (as a receiver). For b to successfully decode the message m_2 transmitted by d_2 to d_1, we must have:

$$SINR_b^{m_2} > SINR_{d_1}^{m_2},$$
$$\frac{P_2 h_{b,d_2}}{\sigma^2 + P_1 h_{b,d_1} + P_u h_{b,u}} > \frac{P_2 h_d}{\sigma^2 + P_1 \eta_1 + P_u h_{d_1,u}}.$$

Since practical systems are interference-limited [31], [32], the noise power is negligible compared to the interfering terms which yields the SIC condition:

$$P_1(h_{b,d_2}\eta_1 - h_d h_{b,d_1}) + P_u(h_{d_1,u}h_{b,d_2} - h_d h_{b,u}) > 0. \quad (10)$$

In addition to condition (10), the PMCs must be verified. Since b decodes m_2 first, then we have the following PMC for the decoding of m_2:

$$P_2 h_{b,d_2} > P_1 h_{b,d_1} + P_u h_{b,u}. \quad (11)$$

For d_1 to be able to remove the interference of m_u prior to retrieving m_2, we must have $SINR_{d_1}^{m_u} > SINR_b^{m_u}$, which leads to:

$$P_1(h_{d_1,u}h_{b,d_1} - h_{b,u}\eta_1) + P_2(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > 0, \quad (12)$$

and the corresponding PMC is:

$$P_u h_{d_1,u} > P_2 h_d + P_1 \eta_1. \quad (13)$$

Regarding the mutual SIC between the receivers b and d_2, the decoding of m_1 at the level of b requires $SINR_b^{m_1}$ to be greater than $SINR_{d_2}^{m_1}$:

$$\frac{P_1 h_{b,d_1}}{\sigma^2 + P_u h_{b,u}} > \frac{P_1 h_d}{\sigma^2 + P_2\eta_2 + P_u h_{d_2,u}},$$
$$P_2 h_{b,d_1}\eta_2 > P_u(h_{b,u}h_d - h_{d_2,u}h_{b,d_1}). \quad (14)$$

Note that $SINR_b^{m_1}$ does not include P_2 since m_2 is decoded and canceled prior to m_1. The corresponding PMC is given by:

$$P_1 h_{b,d_1} > P_u h_{b,u}. \quad (15)$$

At the level of d_2, $SINR_{d_2}^{m_u}$ must be greater than $SINR_b^{m_u}$ to decode and subtract m_u before retrieving m_1. This yields the following condition:

$$P_1(h_{b,d_1}h_{d_2,u} - h_d h_{b,u}) > P_2\eta_2 h_{b,u} \quad (16)$$

Finally, the PMC at the level of d_2 is given by:

$$P_u h_{d_2,u} > P_1 h_d + P_2\eta_2 \quad (17)$$

B. Second decoding order: b decodes m_1 then m_2

Following the same reasoning as in section VI-A, for the case where m_1 is decoded before m_2 at the level of b, the PMC and rate constraints for a full SIC between d_1 and b, and d_2 and b, are obtained and listed below:

$$P_1\eta_1 h_{b,d_2} > P_u(h_{b,u}h_d - h_{b,d_2}h_{b,d_1}) \quad (18)$$
$$P_2(h_{d_1,u}h_{b,d_2} - h_{u,b}h_d) > h_{u,b}\eta_1 P_1 \quad (19)$$
$$P_2(h_{b,d_1}\eta_2 - h_{b,d_2}h_d) + P_u(h_{b,d_2}h_{b,d_1} - h_{b,u}h_d) > 0 \quad (20)$$
$$P_1(h_{d_2,u}h_{b,d_1} - h_d h_{u,b}) + P_2(h_{d_2b}h_{d_2,u} - h_{u,b}\eta_2) > 0 \quad (21)$$
$$P_2 h_{b,d_2} > P_u h_{b,u} \quad (22)$$
$$P_u h_{d_1,u} > P_2 h_d + P_1\eta_1 \quad (23)$$
$$P_1 h_{b,d_1} > P_u h_{b,u} + P_2 h_{b,d_2} \quad (24)$$
$$P_u h_{d_2,u} > P_1 h_d + P_2\eta_2 \quad (25)$$

In addition to constraints eqs. (2a) to (2d), Problem (2) now includes eight new constraints that express the full SIC feasibility (either equations (10) to (17) or (18) to (25), depending on the decoding order). Solving this optimization problem with inequality constraints by means of the standard Karush–Kuhn–Tucker conditions implies exploring all the possible combinations of active/inactive constraints (an inequality constraint is active if it is verified with equality). This results in a total of $2^{12} - 1$ combinations to be considered. To reduce this exorbitant complexity, the interplay between SIC rate conditions and PMCs is analyzed in the next section, targeting the removal of redundant constraints.

VII. PA Problem Simplification of FD-SIC by Constraint Reduction

Consider the first decoding order at the level of b where m_2 is decoded before m_1. The PMCs for the decoding of m_1 at the level of b and of m_u at the level of d_2 are given by (15) and (17). By multiplying (15) by $h_{d_2,u}$ and adding it to (17) multiplied by $h_{b,u}$, one can eliminate P_u to obtain:

$$P_1(h_{b,d_1}h_{d_2,u} - h_d h_{b,u}) > P_2\eta_2 h_{b,u},$$

which is the SIC condition (16) introduced to remove m_u at the level of d_2. Also, eliminating P_1 from the two PMCs by means of adding (15) multiplied by h_d to (17) multiplied by h_{b,d_1} yields (14). Consequently, the PMCs for the decoding of m_1 at the level of b, and m_u at the level of d_2 imply their

counterpart rate conditions. Moreover, it is noted from (16) that the same necessary condition (6) that is found in HD-SIC between b and d_2 as receivers, is obtained for the application of FD-SIC between d_2 and b:

$$h_{b,d_1}h_{d_2,u} > h_d h_{b,u}. \tag{6}$$

Note that if (6) is not true, (16) becomes impossible to satisfy no matter P_1 and P_2; however, when (6) is true, (16) can be satisfied under an adequate power play between P_1 and P_2.

We now move to the PMC and SIC conditions for the decoding of m_2 and m_u at the level of b and d_1 respectively, i.e. (11), (13), (10) and (12). By adding (11) multiplied by h_d to (13) multiplied by h_{b,d_2}, P_2 is eliminated to yield:

$$P_u(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > P_1(h_{b,d_1}h_d + \eta_1 h_{b,d_2}), \tag{26}$$

which can be further transformed into:

$$\begin{aligned} &P_1(\eta_1 h_{b,d_2} - h_{b,d_1}h_d) + P_u(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > 2P_1\eta_1 h_{b,d_2} \\ &\Rightarrow P_1(\eta_1 h_{b,d_2} - h_{b,d_1}h_d) + P_u(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) > 0. \end{aligned}$$

Thus, the PMCs (11) and (13) imply (10). In fact, not only do they imply the rate condition, but it is clear that the PMCs represent more restrictive constraints than rate conditions. Finally, eliminating P_u from the PMCs through the combination of (11) multiplied by $h_{d_1,u}$ with (13) multiplied by $h_{b,u}$ yields:

$$P_2(h_{b,d_2}h_{d_1,u} - h_d h_{b,u}) > P_1(h_{b,d_1}h_{d_1,u} + \eta_1 h_{b,u}), \tag{27}$$

which can be rearranged into:

$$\begin{aligned} &P_2(h_{d_1,u}h_{b,d_2} - h_{b,u}h_d) + P_1(h_{d_1,u}h_{b,d_1} - h_{b,u}\eta_1) > 2P_1 h_{b,d_1}h_{d_1,u} \\ &\Rightarrow (12). \end{aligned}$$

Once again, the PMCs for the decoding of m_2 and m_u at b and d_1 imply their rate condition counterparts. Note that the necessary channel condition that appears from eqs. (26) and (27) is the same as in the case of HD-SIC in the second half-time slot:

$$h_{d_1,u}h_{b,d_2} > h_{b,u}h_d. \tag{8}$$

Also, the combinations of (15) with (13), and (17) with (11), while eliminating P_u, give the following conditions:

$$\begin{aligned} P_1(h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}) &> P_2 h_d h_{b,u}, \\ P_2(h_{b,d_2}h_{d_2,u} - \eta_2 h_{b,u}) &> P_1(h_{b,d_1}h_{d_2,u} + h_d h_{b,u}). \end{aligned}$$

These inequalities yield two other necessary, but not sufficient, channel conditions for the application of full SIC to the system:

$$h_{b,d_1}h_{d_1,u} > \eta_1 h_{b,u}, \tag{28}$$

$$h_{b,d_2}h_{d_2,u} > \eta_2 h_{b,u}. \tag{29}$$

Repeating the same procedure for the second decoding order delivers the same results: 1) the PMCs encompass the rate conditions, 2) the same necessary four channel conditions (6), (8), (28), and (29) are obtained.

Therefore, in the FD-SIC scenario, the system checks the validity of eqs. (6), (8), (28) and (29) prior to solving the PA problem for each decoding order. If the channel conditions are not valid or no solution is obtained for (2), the FD-SIC algorithm reverts to the FD-NoSIC procedure described in section III-A. As a conclusion for this section, Problem (2) is now only equipped with the PMC set corresponding to the decoding order (i.e. eqs. (11), (13), (15) and (17), or eqs. (22) to (25)), in addition to constraints eqs. (2a) to (2d). This reduces the number of combinations of active/inactive constraints from $2^{12} - 1$ to $2^8 - 1$ which is still considerable. The aim of the next section is to workaround the need of a full search over the corresponding 255 cases for determining the optimal PA. This is done by efficiently determining the meaningful constraint combinations, based on the geometrical interpretation of the FD-SIC PA problem. Considerable complexity reductions arise from this approach as shown next.

VIII. Solution for FD-SIC Optimal PA

The proposed geometrical resolution of the FD-SIC D2D rate maximization PA problem is presented in detail for the first decoding order. First, the geometrical representation of the solution space satisfying the PMC and power limit constraints is provided. Then, a procedure is elaborated leading to the reduction of the search space to the minimum required. Afterwards, the optimization is conducted on the resulting reduced search space. At last, a quick summary of the optimal PA procedure is presented including the required changes to obtain the optimal PA for the second decoding order.

A. 3D Solution Space Representation

The four PMCs that must be satisfied in eq. (11), (13), (15) and (17) are written in the following form:

$$\begin{aligned} P_u h_{b,u} &< P_2 h_{b,d_2} - P_1 h_{b,d_1} && (PMC_1) \\ P_u h_{d_1,u} &> P_2 h_d + P_1 \eta_1 && (PMC_2) \\ P_u h_{b,u} &< P_1 h_{b,d_1} && (PMC_3) \\ P_u h_{d_2,u} &> P_1 h_d + P_2 \eta_2 && (PMC_4) \end{aligned}$$

In the 3D space of axes x, y, z representing variables P_1, P_2 and P_u respectively, we introduce the planes $\mathcal{PL}_1, \mathcal{PL}_2, \mathcal{PL}_3$ and $\mathcal{PL}_4$ whose equation is given by the PMCs 1, 2, 3 and 4 when the conditions are met with equality. In the following, we refer to $\mathcal{PL}_i$ as the plane *derived from*, or equivalently, *corresponding to*, or simply, as *the plane of* PMC_i. Each PMC restricts the search space either to the half space below its corresponding plane like for PMC_1 and PMC_3, or to the half space above its corresponding plane as for PMC_2 and PMC_4. On the other hand, the transmit power limits restrict the search space to the region within the parallelepiped defined by the sides $x = P_{1,M}$, $x = 0$, $y = P_{2,M}$, $y = 0$, $z = P_{u,M}$, $z = P_{u,m}$. To have a non-empty search space (i.e. FD-SIC is feasible), the pentahedron defined by the space region above $\mathcal{PL}_2$ and $\mathcal{PL}_4$ and below $\mathcal{PL}_1$ and $\mathcal{PL}_3$ must be non-empty, and it must have a common region with the parallelepiped.

- Non-empty pentahedron: The pentahedron is non-empty if the intersection lines of $\mathcal{PL}_1$ with $\mathcal{PL}_2$ ($\triangleq W_{1,2}$) and $\mathcal{PL}_4$ ($\triangleq W_{1,4}$) are below the intersection line of $\mathcal{PL}_1$

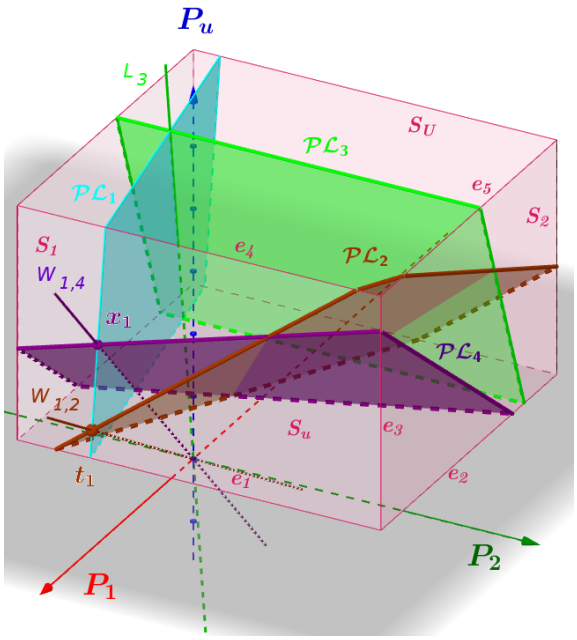


Figure 4: Schematic of the search space formed inside the intersection of the PMC planes with the parallelepiped of power limits.

with $\mathcal{PL}_3$ ($\triangleq L_3$), as shown in Fig. 4. Let $\vec{u}$ be the direction vector of $W_{1,2}$; we must have $\frac{z(\vec{u})}{x(\vec{u})} < h_{b,d_1}/h_{b,u}$, which yields:

$$\vec{u} = \begin{pmatrix} h_{d_1,u}h_{b,d_2} - h_{bu}h_d \\ h_{bu}\eta_1 + h_{d_1,u}h_{b,d_1} \\ \eta_1 h_{b,d_2} + h_{b,d_1}h_d \end{pmatrix} \Rightarrow \frac{\eta_1 h_{b,d_2} + h_{b,d_1}h_d}{h_{d_1,u}h_{b,d_2} - h_{bu}h_d} < \frac{h_{b,d_1}}{h_{b,u}},$$

leading to the following channel condition:

$$h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u} > 2h_{b,u}h_d\frac{h_{b,d_1}}{h_{b,d_2}}. \quad (30)$$

Doing the same for $W_{1,4}$ with respect to L_3, we get the following channel condition:

$$h_{b,d_1}h_{d_2,u} - h_{b,u}h_d > 2h_{b,u}\eta_2\frac{h_{b,d_1}}{h_{b,d_2}}. \quad (31)$$

Note that (30) and (31) are more restrictive than the necessary conditions of eq. (28) and (29), which is normal since they turn them into sufficient channel conditions.

- Pentahedron ∩ parallelepiped: For the pentahedron to have a non-empty intersection with the parallelepiped, it is sufficient to make sure that the intersection line of $\mathcal{PL}_1$ with $\mathcal{PL}_3$ (L_3) intersects the plane $z = P_{u,m}$ within the $P_{1,M}$ and $P_{2,M}$ limits. These conditions on the x, y coordinates of $\mathcal{PL}_3 \cap \mathcal{PL}_1 \cap P_{u,m}$ yield the constraints:

$$P_{u,m}\frac{h_{b,u}}{h_{b,d_1}} < P_{1,M} \quad \&\& \quad 2P_{u,m}\frac{h_{b,u}}{h_{b,d_2}} < P_{2,M}. \quad (32)$$

Conditions (30), (31) and (32) form the necessary and sufficient constraints for the existence of a solution to the FD-SIC PA problem according to the first decoding order.

B. Search Space Reduction

We prove in this section that the optimal solution lies on the intersection line of $\mathcal{PL}_2, \mathcal{PL}_4$ or the lower side of the parallelepiped S_u, with one of the outer sides of the parallelepiped S_1, S_2, S_U (cf. Fig. 4), respectively defined by: $x = P_{1,M}$ for $(y,z) \in [0, P_{2,M}] \times [P_{u,m}, P_{u,M}]$, $y = P_{2,M}$ for $(x,z) \in [0, P_{1,M}] \times [P_{u,m}, P_{u,M}]$, and $z = P_{u,M}$ for $(x,y) \in [0, P_{1,M}] \times [0, P_{2,M}]$.

Proposition 1. *The optimal solution lies on one of the outer sides of the parallelepiped.*

Proof. The D2D rate is given by:

$$R_{D2D}(P_1, P_2) = B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2})$$

For any couple (P_1, P_2), and $\forall\beta > 1$, the throughput of $(\beta P_1, \beta P_2)$ is greater than $R_{D2D}(P_1, P_2)$ since:

$$\begin{aligned} R_{D2D}(\beta P_1, \beta P_2) = & \\ & B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2/\beta}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2/\beta}) \\ & > B\log_2(1 + \frac{P_1 h_d}{P_2\eta_2 + \sigma^2}) + B\log_2(1 + \frac{P_2 h_d}{P_1\eta_1 + \sigma^2}) \\ & = R_{D2D}(P_1, P_2) \end{aligned}$$

Therefore, given an initial triplet (P_1, P_2, P_u), a higher throughput-achieving triplet can be obtained by simply multiplying the components by a factor larger than 1. The higher the β, the higher the throughput, meaning that β should be increased until reaching the boundaries of the region, which can be either $P_{1,M}, P_{2,M}$ or $P_{u,M}$. □

Moreover, it is clear that the D2D rate is independent of P_u. This means that when moving on a vertical line in the solution space, R_{D2D} is constant and P_u only affects the CU rate. To keep the CU rate as close as possible to $R_{u,min}$, we select the smallest P_u value from the range of admissible values for a given (P_1, P_2) couple. Since every point in the solution space must be on top of $\mathcal{PL}_2$ and $\mathcal{PL}_4$, the minimum allowed value of P_u is given by forcing the equality either on PMC_2 or on PMC_4, according to the one that delivers the higher minimum value of P_u for the considered (P_1, P_2) couple.

As a conclusion, the optimal solution lies on the intersection segment of one of the outer sides of the parallelepiped S_1, S_2, or S_U, with one of the planes $\mathcal{PL}_2$, $\mathcal{PL}_4$, or S_u, resulting in a total of eight possible combinations (eight segments). Given the shape of the solution space, some of the combinations are mutually exclusive. The aim of the next section is to determine which subset of segments should be accounted for in the power optimization process, depending on the channel conditions of the D2D-CU couple.

C. Selection of the Useful Intersections

As can be seen from Fig. 4, some of the eight intersections can be discarded. For example, the intersection of S_u with S_1 and S_2 is not relevant, since the value of P_u is decided by PMC_2 and PMC_4, whose planes are on top of S_u near the sides S_1 and S_2. Fig. 5 shows the projection on the (P_1, P_2)

plane of the partition of the space into two vertical regions where PMC_4 encompasses PMC_2 for region 1, and PMC_2 encompasses PMC_4 for region 2. The plane separating the two regions is the vertical plane passing through the straight line $L_\lambda \triangleq \mathcal{PL}_4 \cap \mathcal{PL}_2$. Therefore, for the case of Figs. 4

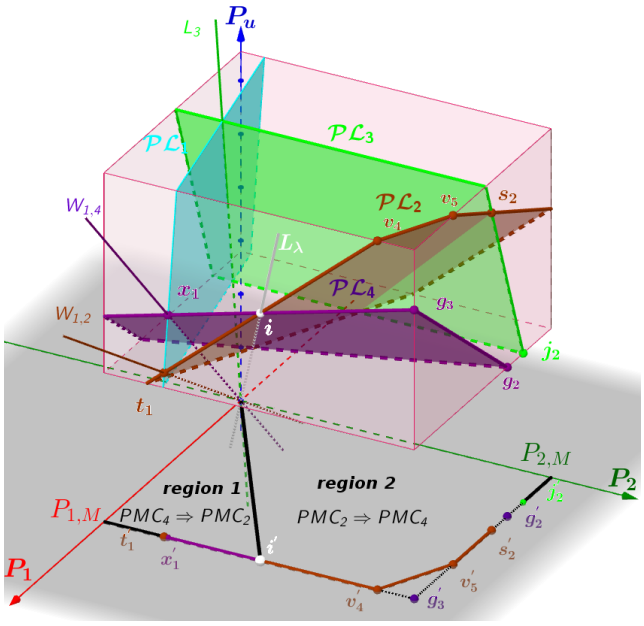
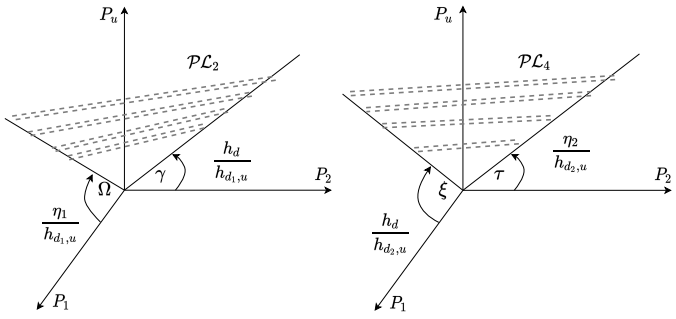


Figure 5: Schematic of the solution space showing the regions of dominance of PMC_4 over PMC_2 and vice-versa.

and 5, the D2D rate optimization is to be conducted over the segment $\overline{x_1 i} \cup \overline{i v_4}$ which is included in S_1, over the segment $\overline{v_4 v_5}$ included in S_U, and on the segment $\overline{v_5 s_2}$ over S_2. By doing so, the optimization over the segments $\overline{t_1 i}$, $\overline{i g_3}$, $\overline{g_3 g_2}$ and $\overline{g_2 j_2}$ is avoided.

Therefore, the first step in reducing the number of intersections to be considered lies in determining which of PMC_4 and PMC_2 encompasses the other, and for which region of the space. To that end, a schematic of $\mathcal{PL}_2$ and $\mathcal{PL}_4$ is presented in Figs. 6a and 6b, showing their intersection with the planes defined by $P_1 = 0$, and $P_2 = 0$. The angles of these intersection lines and their slopes are shown in Fig. 6.



(a) PMC_2 (b) PMC_4

Figure 6: Isolated schematics of $\mathcal{PL}_2$ and $\mathcal{PL}_4$ in the 3D space.

*1) **Interplay between PMC_2 and PMC_4**:* Depending on the angles Ω, γ, ξ and τ, four cases are identified to determine the interplay between PMC_2 and PMC_4:

1) $\Omega > \xi$, $\gamma > \tau$: PMC_2 encompasses PMC_4 ($PMC_2 \Rightarrow PMC_4$) over all the positive (P_1, P_2) plane.
2) $\Omega < \xi$, $\gamma < \tau$: PMC_4 encompasses PMC_2 ($PMC_4 \Rightarrow PMC_2$) over all the positive (P_1, P_2) plane.
3) $\Omega < \xi$, $\gamma > \tau$: PMC_4 encompasses PMC_2 in region 1 and PMC_2 encompasses PMC_4 in region 2, (cf. Fig. 5).
4) $\Omega > \xi$, $\gamma < \tau$: PMC_4 encompasses PMC_2 in region 2 and PMC_2 encompasses PMC_4 in region 1.

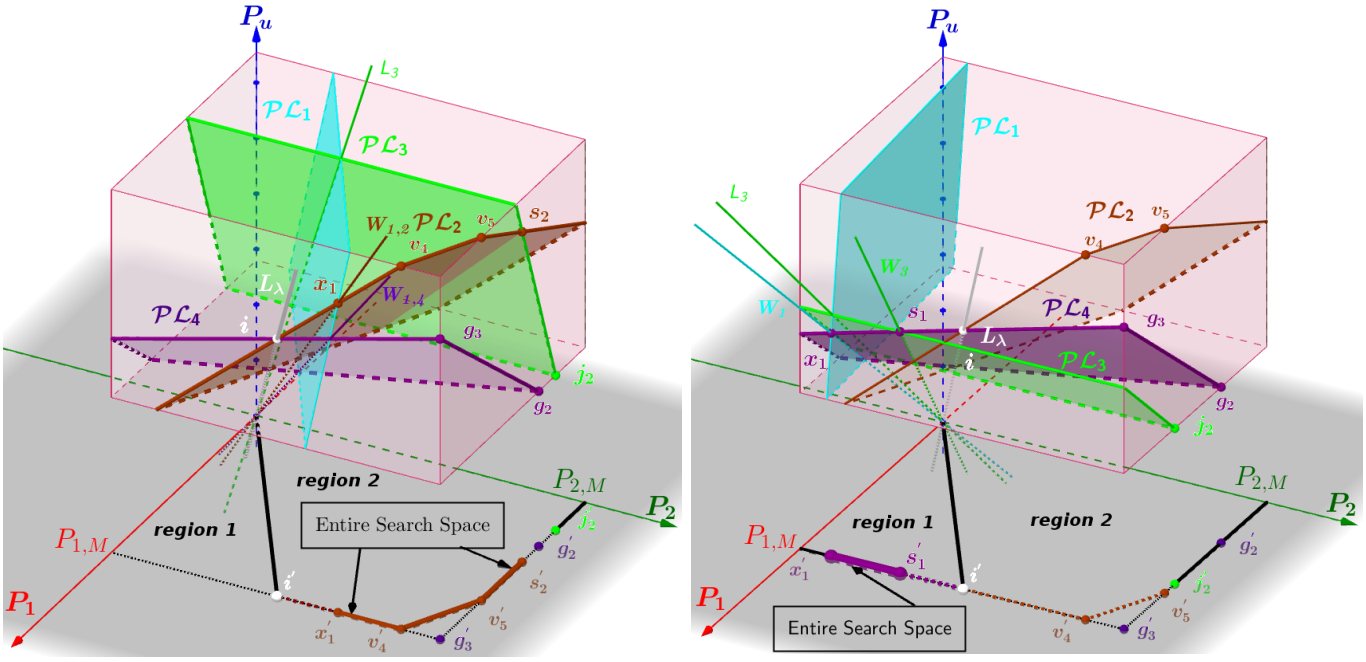
Before proceeding, it must be noted that even for cases 3) and 4), a PMC may encompass the other on the entire search space if the actual search space is included either in region 1 or 2. This is depicted in the examples of Fig. 7 which take back the conditions of Fig. 5 with some modifications. In Fig. 7a, $\mathcal{PL}_1$ is such that $W_{1,4}$ is at the right side of L_λ ($W_{1,4}$ is in region 2), then the search space is included in region 2 and only PMC_2 needs to be accounted for. The other scenario is represented in Fig. 7b where $\mathcal{PL}_3 \cap \mathcal{PL}_2$ is at the left side of L_λ (in region 1), hence PMC_4 encompasses PMC_2 over the entirety of the search space. The first scenario occurs when L_λ is on top of $\mathcal{PL}_1$, and the second one occurs when L_λ is on top of $\mathcal{PL}_3$. The explicit channel conditions enabling each scenario are derived in detail in Appendix A, where it is shown that two simple tests are required to determine if the search space is included in any region.

As a conclusion, by comparing Ω to ξ and γ to τ, and following the discussion in Appendix A, the number of intersections to be considered is reduced by selecting the appropriate PMC between PMC_2 and PMC_4 in the corresponding space region. For the sake of clarity, we introduce $PMC_{2,4}$ as the efficient combination of PMC_2 and PMC_4, given by:

$$P_u \geq \begin{cases} \dfrac{P_2 h_d + P_1 \eta_1}{h_{d_1,u}}, & \text{if } P_2\left(\frac{h_d}{h_{d_1,u}} - \frac{\eta_2}{h_{d_2,u}}\right) > P_1\left(\frac{h_d}{h_{d_2,u}} - \frac{\eta_1}{h_{d_1,u}}\right) \\ \dfrac{P_1 h_d + P_2 \eta_2}{h_{d_2,u}}, & \text{elsewhere.} \end{cases}$$

*2) **Selection of the useful parallelepiped sides**:* With $PMC_{2,4}$ at hand, the next step is to reduce the unnecessary sides of the parallelepiped, which do not intersect with $\mathcal{PL}_{2,4}$, or that do intersect with $\mathcal{PL}_{2,4}$ but not inside the range allowed between $\mathcal{PL}_1$ and $\mathcal{PL}_3$. To that end, we study PMC_1 and PMC_3 which do not affect the intersection segments (of $\mathcal{PL}_{2,4}$ with the parallelepiped sides) as such, but rather the end points of these intersection segments. A typical example is given in Fig. 7a where PMC_1 sets the end point x_1 from the side S_1, and PMC_3 sets the end point s_2 from the side S_2.

Let W_1 regroup the intersection lines $W_{1,2}$ and $W_{1,4}$ such that $W_1 = \mathcal{PL}_{2,4} \cap \mathcal{PL}_1$, and let W_3 be the intersection line of $\mathcal{PL}_3$ with $\mathcal{PL}_{2,4}$ (cf. Fig. 8). Each of W_1 and W_3 may intercept the sides S_1 or S_2 or S_U, yielding a total of nine potential combinations. Since each side S_i is a rectangular surface within the infinite plane $\mathcal{S}_i$ of equation $P_i = P_{i,M}$, then W_1 and W_3 can intercept only one side of the parallelepiped



(a) Search space included in region 2

(b) Search space included in region 1

Figure 7: Figures representing solution search spaces included in regions 1 or 2.

(S_u aside) for a given channel configuration. Let x_i and s_i be the intersection points of W_1 and W_3 with S_i, we have:

$$x_i = W_1 \cap S_i,\ s_i = W_3 \cap S_i, \forall i \in \{1, 2, U\}.$$

To determine which sides are hit by W_1 (resp. W_3), i.e. to determine if we have x_1, x_2 or x_U (resp. s_1, s_2 or s_U), we consider the points xl_i (resp. sl_i), intersections of W_1 (resp. W_3) with the planes $\mathcal{S}_i, \forall i \in \{1, 2, U\}$. The coordinates of xl_i are given by:

$$xl_1 = \begin{pmatrix} P_{1,M} \\ y(W_1 \cap \mathcal{S}_1) \\ z(W_1 \cap \mathcal{S}_1) \end{pmatrix}, xl_2 = \begin{pmatrix} x(W_1 \cap \mathcal{S}_2) \\ P_{2,M} \\ z(W_1 \cap \mathcal{S}_2) \end{pmatrix}, xl_U = \begin{pmatrix} y(W_1 \cap \mathcal{S}_U) \\ y(W_1 \cap \mathcal{S}_U) \\ z = P_{u,M} \end{pmatrix}$$

Then, two tests are needed to determine which of x_1, x_2 or x_U occurs for the given channel states (cf. Algorithm 1). Note that xl_i and sl_i have positive coordinates as shown in Appendix B-C. The same tests are replicated for s_i. From the nine possibilities, only six combinations are actually viable because the pairs (x_U, s_1), (x_2, s_U) and (x_2, s_1) cannot be achieved without violating (30) or (31) as can be seen in Fig. 8. Indeed, the three cases shown in Fig. 8 lead to empty search spaces. The six viable pairs are given in Table II with the correspondence between the pairs and the parallelepiped sides hosting the useful intersection segments. Note that if $\mathcal{PL}_1$ and $\mathcal{PL}_3$ intercept $\mathcal{PL}_{2,4}$ at the same side, then the search space can be reduced to a single segment as it is the case for the first, the fourth and the fifth rows in Table II. For the second and third rows, two segments are involved in the D2D rate optimization. Finally, in the case where W_1 intercepts S_1 and W_3 intercepts S_2 (as in Fig. 5), the segment $\overline{v_4 v_5}$ belonging to S_U is to be included in the D2D optimization process - in addition to the segments in S_1 and S_2 - if and only if the value of P_u obtained from $\mathcal{PL}_{2,4}$ at $P_1 = P_{1,M}$ and $P_2 = P_{2,M}$ is greater than $P_{u,M}$.

Algorithm 1: W_1 intersection with the parallalepiped

input : $P_{1,M}$, $P_{2,M}$, $P_{u,M}$, $P_{u,m}$, $h_{b,u}$, h_d, η_1, η_2, $h_{d1,u}$, $h_{d_2,u}$, $h_{d_1,b}$, $h_{d_2,b}$

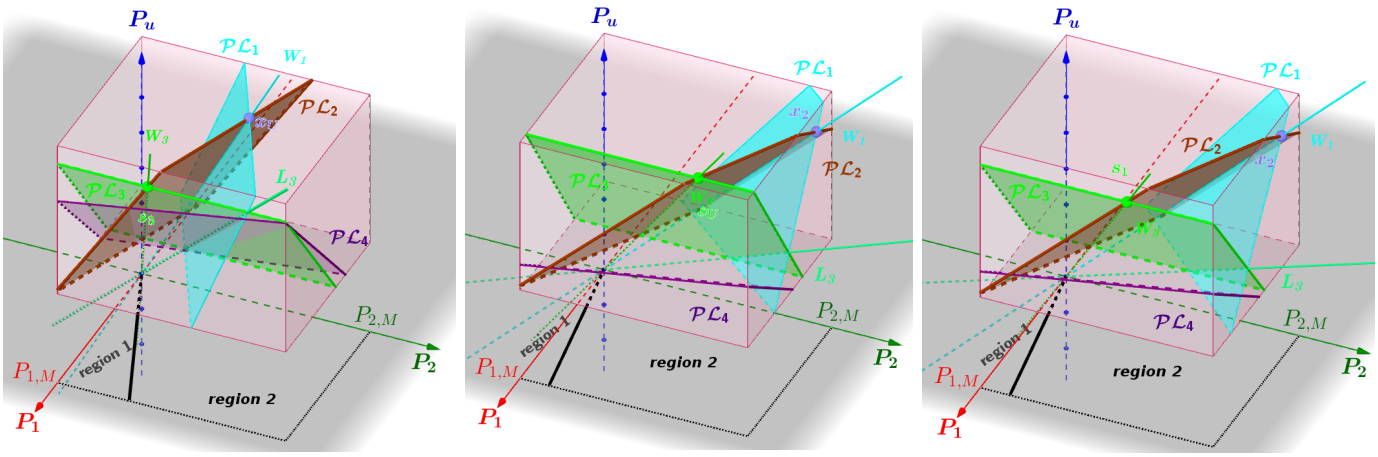
Result: Returns $i / W_i \cap S_i = x_i = xl_i \neq \emptyset$.

```
if y(xl_1) < P_{2,M} then
    if z(xl_1) ≤ P_{u,M} then
        i = 1, keep x_1
    else
        i = U, keep x_U
    end
else
    if z(xl_2) < P_{u,M} then
        i = 2, keep x_2
    else
        i = U, keep x_U
    end
end
```

	S_1	S_2	S_U
x_U and s_U			✓
x_1 and s_U	✓		✓
x_U and s_2		✓	✓
x_1 and s_1	✓		
x_2 and s_2		✓	
x_1 and s_2	✓	✓	Depends

Table II: Table showing the sides involved in the D2D rate optimization for each of the six (x_i, s_j) viable pairs due to PMC_1 and PMC_3.

*3) **Segments endpoints**:* Having determined the relevant intersection segments (a maximum of three segments) for



(a) W_1 intercepts S_U and W_3 intercepts S_1 (b) W_1 intercepts S_2 and W_3 intercepts S_U (c) W_1 intercepts S_2 and W_3 intercepts S_1

Figure 8: The three non-feasible combinations between x_i and s_i.

the D2D rate optimization using PMCs 1 and 3, we detail hereafter how the endpoints of every segment are determined for each side of the parallelepiped. For the sake of clarity, let e_1, e_2, e_3, e_4, e_5 be the edges of the parallelepiped (cf. Fig. 4) given by:

$$e_1 = S_u \cap S_1, e_2 = S_u \cap S_2, e_4 = S_U \cap S_1, e_5 = S_U \cap S_2, e_3 = S_2 \cap S_1$$

Also, let the three families of points v_i, g_i, and w_i be the intersections of $\mathcal{PL}_2, \mathcal{PL}_4$ and $\mathcal{PL}_{2,4}$ with e_i:

$$v_i = \mathcal{PL}_2 \cap e_i, \quad g_i = \mathcal{PL}_4 \cap e_i, \quad w_i = \mathcal{PL}_{2,4} \cap e_i$$

Examples of such points can be seen in Fig. 7b for v_4, v_5, g_3 and g_2. Note that points w_i are only used to designate the points v_i or g_i depending on whether we are in region 1 or 2. We can now efficiently designate the segment endpoints on each side.

a) Side S_2: The optimization over S_2 translates into an optimization over P_1, since P_2 is equal to $P_{2,M}$. It is clear that the minimal value of P_1 is bound to PMC_3. In Fig. 5 for example, the minimal value of P_1 is obtained for the point s_2, intersection of $\mathcal{PL}_3$ with $\mathcal{PL}_{2,4}$. However, if PMC_4 were the one encompassing PMC_2 in region 2, the considered segment over S_2 would have been $\overline{g_3 g_2} \cup \overline{g_2 j_2}$; then, the minimum P_1 value would have been given by the point $j_2 = \mathcal{PL}_3 \cap S_2 \cap S_u$. Thus, we can generalize by stating that the minimal value of P_1 in the segment over S_2 is given by:

$$\min P_1 = \max[x(\mathcal{PL}_3 \cap \mathcal{PL}_{2,4} \cap S_2), x(\mathcal{PL}_3 \cap S_u \cap S_2)],$$
$$\min P_1 = \max[x(s_2), x(j_2)].$$

Regarding the maximum value of P_1, it can be due to the intersection of $S_2 \cap \mathcal{PL}_{2,4}$ with either S_U (like for v_5 in Fig. 5), S_1 (like for g_3), or with $\mathcal{PL}_1$ (for the case of x_2 and s_2, in the fifth row of table II). Also, the maximum P_1 value may be simply set by k_2, the intersection of $S_2 \cap \mathcal{PL}_1$ with S_u. The maximum value of P_1 is given by:

$$\max P_1 = \min[x(\mathcal{PL}_{2,4} \cap e_2), x(\mathcal{PL}_{2,4} \cap e_3),$$
$$x(\mathcal{PL}_{2,4} \cap \mathcal{PL}_1 \cap S_2), x(\mathcal{PL}_1 \cap e_2)],$$
$$\max P_1 = \min[x(w_5), x(w_3), x(x_2), x(k_2)].$$

Note that for the side S_2, PMC_3 is involved in the minimum P_1 value, and PMC_1 in the maximum value.

b) Side S_1: Regarding the side S_1, PMC_3 is now involved in the maximum P_2 value, while PMC_1 settles the minimum P_2 value; its expression is given by:

$$\min P_2 = \max[y(\mathcal{PL}_1 \cap \mathcal{PL}_{2,4} \cap S_1), y(\mathcal{PL}_1 \cap S_u \cap S_1)],$$
$$\min P_2 = \max[y(x_1), y(k_1)].$$

The maximum value of P_2 depends on which plane intercepts first $\mathcal{PL}_{2,4}$ among the three candidates: $\mathcal{PL}_3$, S_U (cf. Fig. 5), or S_2. We have:

$$\max P_2 = \min[y(\mathcal{PL}_{2,4} \cap \mathcal{PL}_3 \cap S_1), y(\mathcal{PL}_{2,4} \cap S_2 \cap S_1),$$
$$y(\mathcal{PL}_{2,4} \cap S_U \cap S_1)],$$
$$\max P_2 = \min[y(s_1), y(w_3) = P_{2,M}, y(w_4)].$$

In the example of Fig. 5, the intersection segment starts at x_1 and ends at v_4 passing by i. Although $\overline{x_1 i} \cup \overline{i v_4}$ is a different segment from $\overline{x_1 v_4}$, their projections over the (P_1, P_2) plane are identical, thus we are only interested in segment ends over both sides S_1 and S_2 (the projection of $\overline{g_3 g_2} \cup \overline{g_2 j_2}$ is the same as that of the segment $\overline{g_3 j_2}$).

c) Side S_U: Unlike for the other sides, none of P_1 or P_2 is fixed, but P_2 can be expressed in terms of P_1; therefore, we evaluate the position of the endpoints of the segments on S_U in terms of maximum P_1 and minimum P_1.
When $L_\lambda = \mathcal{PL}_2 \cap \mathcal{PL}_4$ does not intercept S_U as it is the case for Fig. 5 for example, the intersection of $\mathcal{PL}_{2,4}$ with S_U yields a unique segment ($\overline{v_4 v_5}$ in case of Fig. 5). The endpoint corresponding to the minimum value of P_1 is due to the intersection of $\mathcal{PL}_{2,4}$ with either S_2 or $\mathcal{PL}_3$.

$$\min P_1 = \max[x(\mathcal{PL}_{2,4} \cap \mathcal{PL}_3 \cap S_U), x(\mathcal{PL}_{2,4} \cap S_2 \cap S_U)],$$
$$\min P_1 = \max[x(s_U), x(w_5)]. \quad (33)$$

The maximum value of P_1 is due to the intersection of $\mathcal{PL}_{2,4}$ with either $\mathcal{PL}_1$ or S_1.

$$\max P_1 = \min[x(\mathcal{PL}_{2,4} \cap \mathcal{PL}_1 \cap S_U), x(\mathcal{PL}_{2,4} \cap S_1 \cap S_U)],$$
$$\max P_1 = \min[x(x_U), x(w_4)]. \quad (34)$$

If i resides on S_U, then the intersection segment of $\mathcal{PL}_{2,4}$ with S_U is broken into two segments. In that case, if we let a and b be the points given by (33) and (34) respectively, then the optimization over S_U has to be conducted separately over $\overline{bi}$ from the side of region 1, and over $\overline{ia}$ from the side of region 2. In this case, i corresponds to the $\max P_1$ point in $\overline{ia}$ and to the $\min P_1$ point in $\overline{bi}$. Assuming the conditions of the last row in Table II, this is the only case where 4 segments in total have to be checked to find the optimal D2D throughput achieving point. The coordinates of all the points mentioned in this section are provided in Appendix B-C.

D. D2D Throughput Optimization

At last, given the segments locations and endpoints, the analytical power optimization can be conducted. The mathematical formulation varies according to the side the segment is included in.

*1) **Side S_1**:* The optimization variable is P_1 and the problem formulation is the following:

$$P_1^* = \arg\max_{P_1} B\log_2(1+\frac{P_1 h_d}{P_2\eta_2+\sigma^2}) + B\log_2(1+\frac{P_2 h_d}{P_1\eta_1+\sigma^2})$$
$$\text{s.t. } P_1 \in [\min P_1, \max P_1] \& \quad P_2 = P_{2,M}$$

Taking the derivative of $F(P_1) = R_{D2D}(P_1, P_{2,M})$ with respect to P_1, we get:

$$\frac{\partial F}{\partial P_1}\frac{\ln 2}{B} = \frac{h_d}{P_1 h_d + P_{2,M}\eta_2 + \sigma^2} + \frac{-\eta_1 P_{2,M} h_d}{(P_1\eta_1+\sigma^2)(P_1\eta_1 + P_{2,M}h_d + \sigma^2)}$$

The sign of $\partial F/\partial P_1$ is equal to the sign of the following second-degree polynomial of P_1:

$$P_1^2 \underbrace{\eta_1^2}_{a} + P_1 \underbrace{2\eta_1\sigma^2}_{b} + \underbrace{P_{2,M}(h_d-\eta_1)\sigma^2 - P_{2,M}^2\eta_2\eta_1 + \sigma^4}_{c}$$

It is shown in Appendix B-A that independently of the sign of the polynomial's discriminant, it is sufficient to test which of $\min P_1$ or $\max P_1$ delivers the best throughput and then select the corresponding segment endpoint. The coordinates of the endpoint form the optimal triplet (P_1^*, P_2^*, P_u^*) maximizing the D2D throughput over the side S_1. The endpoint coordinates are given in Appendix B-C.

*2) **Side S_2**:* Following the same reasoning as for S_1 (with the only difference that the optimization variable is now P_2 instead of P_1, and $P_1 = P_{1,M}$), the same conclusion is reached, i.e. the maximum D2D throughput is delivered by the points corresponding either to $\min P_2$ or to $\max P_2$.

*3) **Side S_U**:* Since i is accounted for in the maximum and minimum values of P_1 for each intersection segment, the optimization can thus be conducted over each segment independently.

The D2D throughput maximization problem over the intersection segment of $\mathcal{PL}_2$ with S_U can be written as follows:

$$P_1^* = \arg\max_{P_1} F(P_1, P_2)$$
$$\text{s.t } P_{u,M} = \frac{P_1\eta_1 + P_2 h_d}{h_{d_1,u}}, P_1 \in \mathbb{U} = [\min P_1, \max P_1], P_u = P_{u,M}.$$

Replacing P_2 by $(P_{u,M}h_{d_1,u} - P_1\eta_1)/h_d$ in $F(P_1,P_2) = B\log_2(1+\frac{P_1 h_d}{P_2\eta_2+\sigma^2}) + B\log_2(1+\frac{P_2 h_d}{P_1\eta_1+\sigma^2})$, we get:

$$F(P_1) = B\log_2(1+\frac{P_1 h_d^2}{(P_{u,M}h_{d_1,u} - P_1\eta_1)\eta_2 + h_d\sigma^2}) + B\log_2(1+\frac{P_{u,M}h_{d_1,u} - P_1\eta_1}{P_1\eta_1+\sigma^2}).$$

Since $P_2 > 0$, we must have $P_1 < P_{u,M}h_{d_1,u}/\eta_1$ (which adds to the constraint of $\max P_1$). Taking the derivative of F with respect to P_1 leads to: $\ln(2)\partial F/\partial P_1 B =$

$$= \frac{h_d^2(h_{d_1,u}\eta_2 P_{u,M} + \sigma^2 h_d)/[P_{u,M}h_{d_1,u}\eta_2 - P_1\eta_1\eta_2 + h_d\sigma^2]}{(P_{u,M}h_{d_1,u}\eta_2 - P_1(\eta_1\eta_2 - h_d^2) + h_d\sigma^2)} - \frac{\eta_1}{(P_1\eta_1+\sigma^2)}.$$

Since $P_1 \in \mathbb{U}$, it can be easily verified that both denominators are positive; therefore, only the numerator is needed to evaluate the sign of $\partial F/\partial P_1$:

$$\operatorname{sgn}\frac{\partial F}{\partial P_1} = \operatorname{sgn}[h_d^2(h_{d_1,u}\eta_2 P_{u,M} + \sigma^2 h_d)(P_1\eta_1+\sigma^2) - \eta_1(P_{u,M}h_{d_1,u}\eta_2 - P_1\eta_1\eta_2 + h_d\sigma^2)(P_{u,M}h_{d_1,u}\eta_2 - P_1(\eta_1\eta_2 - h_d^2) + h_d\sigma^2)].$$

After some simplifications and re-arrangements, the sign of $\partial F/\partial P_1$ can be written as the sign of a second-degree polynomial of P_1 of the form $AP_1^2 + BP_1 + C$ with:

$$A = -(\eta_1\eta_2 - h_d^2)\eta_1^2\eta_2; \quad B = 2\eta_1^2\eta_2(P_{u,M}h_{d_1,u}\eta_2 + \sigma^2 h_d);$$
$$C = -P_{u,M}^2 h_{d_1,u}^2\eta_2^2\eta_1 + P_{u,M}\sigma^2 h_d h_{d_1,u}\eta_2(h_d - 2\eta_1) + \sigma^4 h_d^2(h_d - \eta_1).$$

Given the root $sol_1 = (-B - \sqrt{B^2 - 4AC})/2A$ of the polynomial, we show in Appendix B-B that P_1^* is either given by $\min P_1$, $\max P_1$, or sol_1 (when it is included in the interval $[\min P_1, \max P_1]$), according to the value delivering the highest throughput. Regarding the optimization over the intersection segment of $\mathcal{PL}_4$ with S_U, the same steps are followed to determine the optimal value of P_1: we start by writing the expression of $F(P_1)$ by replacing P_2 in $F(P_1,P_2)$ with $(P_{u,M}h_{d_2,u} - P_1 h_d)/\eta_2)$. Then, the study of the sign of $\partial F/\partial P_1$ turns into the study of the sign of another second-degree polynomial $A'P_1^2 + B'P_1 + C'$ with:

$$A' = (\eta_1\eta_2 - h_d^2)\eta_1; \quad B' = 2\eta_1(P_{u,M}h_{d_2,u}h_d + \sigma^2\eta_2);$$
$$C' = -P_{u,M}^2 h_{d_2,u}^2\eta_1 - \sigma^2\eta_1 h_{d_2,u}P_{u,M} + \sigma^4(\eta_2 - h_d).$$

Also, following the different channel conditions concerning $\operatorname{sgn}(\eta_1\eta_2 - h_d^2)$, and considering all the possible relative positions between $\max P_1, \min P_1$, and sol_1', the same result as previously is obtained, which can be cast as:

$$P_1^* = \arg\max[F(\min P_1), F(\max P_1), F(sol_1')].$$

As a conclusion, the optimization over the sides S_1 and S_2 resides in selecting the corresponding endpoint achieving the highest throughput. On the side S_U, a maximum of three additional points $(i, sol_1, sol_1^{'})$ may need to be considered to get the highest D2D throughput.

E. Summary of the PA Procedure and Extension to the Second Decoding Order

In this section, the geometrical representation of the FD-SIC PA problem allowing for a drastic reduction of the search space size was described. It was shown that the initial search volume in Sec. VIII-A can be reduced to a set of intersection segments (Sec. VIII-B) from which a subset is selected (VIII-C). These segments search spaces are then further reduced to become a finite set of points (Secs. VIII-C3, VIII-D). In the worst case scenario, the original PA problem, which had $2^{12}-1$ variants, is converted into the search for the maximum throughput of a list of seven elements: two elements from S_1, two from S_2 and three additional elements from S_U (w_4 is a common endpoint to S_1 and S_U, and w_5 is common to S_2 and S_U). The global PA procedure to determine the optimal D2D rate for the first decoding order of FD-SIC is summarized in algorithm 2.

Algorithm 2: Optimal PA procedure for FD-SIC

input : $P_{1,M}$, $P_{2,M}$, $P_{u,M}$, $P_{u,m}$, $h_{b,u}$, h_d, η_1, η_2, $h_{d1,u}$, $h_{d_2,u}$, $h_{d_1,b}$, $h_{d_2,b}$
Result: Optimal triplet (P_1^*, P_2^*, P_u^*).
if (30) $\wedge$ (31) $\wedge$ (32) **then**
  Test $\Omega, \xi, \gamma, \tau$ and build $PMC_{2,4}$;
  Execute Algorithm 1 to determine x_i and s_j;
  Follow Table II to keep the necessary segments;
  Compute R_{D2D} for the edges of each segment;
  Keep the point providing the highest throughput.
else
  Empty search space, no solution
end

Regarding the resolution for the second decoding order, the PA procedure itself is unchanged, but the changes in PMC_1 and PMC_3 lead to some modifications. Here is the list:

- Modification in the expressions of PMC_1 and PMC_3:

$$P_u h_{b,u} < P_1 h_{b,d_1} - P_2 h_{b,d_2} \qquad (PMC_1)$$
$$P_u h_{b,u} < P_2 h_{b,d_2} \qquad (PMC_3)$$

- The necessary and sufficient conditions (30), (31) and (32) become:

$$h_{d_1,u} h_{b,d_2} - h_d h_{b,u} > 2\frac{\eta_1 h_{b,u} h_{b,d_2}}{h_{b,d_1}} \qquad (30)$$

$$h_{d_2,u} h_{b,d_2} - \eta_2 h_{b,u} > 2\frac{h_{b,u} h_d h_{b,d_2}}{h_{b,d_1}} \qquad (31)$$

$$2P_{u,m}\frac{h_{b,u}}{h_{b,d_1}} < P_{1,M} \quad \&\& \quad P_{u,m}\frac{h_{b,u}}{h_{b,d_2}} < P_{2,M} \qquad (32)$$

- Concerning section VIII-C3, the roles of PMC_1 and PMC_3 are interchanged concerning the settlement of the segment endpoints.
- The three non-occuring (x_i, s_i) pairs of section VIII-C2 become: $(x_1, s_2), (x_1, s_U), (x_U, s_2)$.

Sections VIII-C1 and VIII-D are kept unchanged because building $PMC_{2,4}$ is independent of PMC_1 and PMC_3, and given the endpoints of the segments subset, the optimization of section VIII-D is not affected by the change in PMC_1 and PMC_3.

IX. Channel Allocation

In this section, the procedure for optimal channel allocation to D2D devices is conducted. Recalling that the D2D system is underlaying a pre-established CU network, D2D channel allocation is equivalently referred to as D2D-CU pairing.

Having determined the analytical PA solutions for all the transmission scenarios, their resolution cost is a constant-time operation. Therefore, filling the D2D rate tables $\mathcal{R}_{D2D}^{FD-NoSIC}, \mathcal{R}_{D2D}^{HD-NoSIC}, \mathcal{R}_{D2D}^{HD-SIC}$, and $\mathcal{R}_{D2D}^{FD-SIC}$ for every D2D-CU pair is accomplished with a complexity in $O(KD)$. In the case of FD-SIC, the channel links, required CU rate and transmit power limits of a D2D n and a CU u_i may be such that one of the conditions (30), (31), (32) is not valid. If this is the case for both decoding orders, then the PA of FD-SIC reverts to that of FD-NoSIC to fill the element $\mathcal{R}_{D2D}^{FD-SIC}(n,i)$ as explained in the end of section VII. Also, if both decoding orders are possible for this combination, $\mathcal{R}_{D2D}^{FD-SIC}(n,i)$ is filled with the highest rate among the two possible orders. When filling matrix $\mathcal{R}_{D2D}^{HD-SIC}$, and as explained in section V, HD-SIC reverts to HD-NoSIC in any of the two half-time slots, when conditions (6) or (8) are not valid. Given these rate tables, the optimal channel allocation tables $O_{FD-NoSIC}^*, O_{HD-NoSIC}^*, O_{FD-SIC}^*$, and O_{HD-SIC}^* corresponding to every transmission scenario are obtained by solving the channel assignment problem in a way to maximize the total D2D throughput. This problem takes the generic formulation given by:

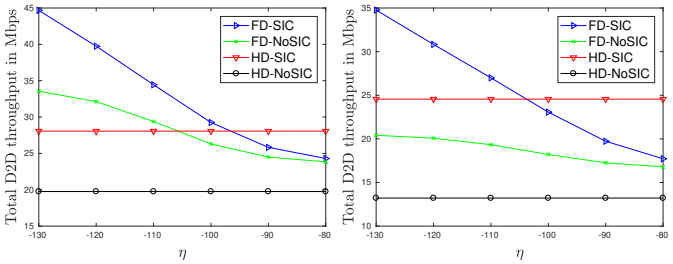
$$O^* = \underset{(i,j)\in[\![1..D]\!]\times[\![1..K]\!]}{\arg\max} \left(\sum_{i=1}^{D}\sum_{j=1}^{K} \mathcal{R}_{D2D}(i,j) \times o(i,j)\right)$$
$$\text{s.t. the constraints of (1) are verified}$$

This assignment problem is efficiently solved by the Kuhn-Munkres (KM) algorithm [33], also called the Hungarian method, with a complexity of $O(D^2K)$ [34]. Note that the global resource allocation complexity is now dominated by that of the channel assignment after the important PA complexity reduction. The KM can be directly applied in our study to yield the optimal channel assignment by rewriting the problem as a minimization of the opposite objective function ($-\mathcal{R}_{D2D}$). As a conclusion, the optimal PA procedures allowed for an efficient filling of the rate tables which are then fed to the KM solver. This delivers the global optimal solution of the joint channel and power allocation problem formulated in section II-A.

X. Numerical Results

In our simulation setup, the BS is positioned at the center of a hexagonal cell with an outermost radius of 300 m. The

D2D users and the CU users are randomly located within the cell. The distance between the D2D users of every pair is below a maximum value d_{max}. The propagation model includes large-scale fading with a path-loss exponent $\alpha = 3.76$, and an 8 dB zero mean lognormal shadowing. The maximum transmit power of the devices and CU is 24 dBm. The system bandwidth is 20 MHz, divided into $N = 64$ channels, leading to a UL bandwidth of $B = 312.5$ kHz, with a noise power of -119 dBm. The minimum required rate $R_{u,min}$ is the same for all the CU users, and the SI cancellation factor η is the same for all D2D pairs, its value being varied between -130 and -80 dB. The results are averaged over 1000 different realizations of the devices and CU positions. Unless specified otherwise, $R_{u,min}$ is set to 1.5 Mbps, K is set to 20 CUs, $d_{max} = 100$ m, and $D = 5$ D2D pairs. Figure 9 presents



(a) $R_{u,min} = 1.5$ Mbps (b) $R_{u,min} = 3$ Mbps

Figure 9: Total D2D throughput as a function of η for $K = 20$ CUs, $D = 5$ D2D pairs, and $d_{max} = 100$ m.

the total D2D throughput as a function of η, for two different values of $R_{u,min}$. At first, it can be noted that HD schemes are not affected by η unlike FD schemes. This was expected since self-interference occurs only for FD transmission. Secondly, the mutual SIC enabled schemes outperform their counterpart No-SIC schemes for both HD and FD transmission scenarios. Indeed, a 41 % rate increase is observed in Fig. 9a between HD-SIC and HD-NoSIC (going from 19.8 Mbps to 28.1 Mbps). The throughput enhancements due to mutual SIC for the case of FD transmission vary between a 2 % increase for $\eta = -80$ dB, to 33 % increase for $\eta = -130$ dB. The performance gains of FD-SIC with respect to FD-NoSIC increase with the SI cancellation capabilities of the devices because of two reasons: on the one hand, the decrease of η relaxes the constraints (28) and (29), thereby increasing the number of D2D-CU pairs that benefit from FD-SIC (from an average of 0.36 FD-SIC D2D pairs for $\eta = -80$ dB to 1.92 pairs for $\eta = -130$ dB, with $R_{u,min} = 1.5$ Mbps). On the other hand, the decrease of η reduces the interference terms in the D2D throughput expression, which translates into a higher achieved throughput.

As expected, when comparing the performance for different required CU rates between Figs. 9a and 9b, the increase of R_u from 1.5 Mbps to 3 Mbps decreases the achieved D2D throughput for all proposed methods. However, the percentage gain in the performance of SIC procedures with respect to NoSIC increases from 41 % to 86 % for the HD case, and from 33 % to 70 % for the FD case (for $\eta = -130$ dB). The reason behind this gain increase is that NoSIC algorithms are highly affected by the value of P_u ($\geq P_{u,m}$) since they suffer from its interference, which is not the case of SIC techniques. In fact, even though the total number of FD-SIC enabled D2D-CU pairs decreases with $R_{u,min}$ (due to harsher mutual SIC constraints, cf. eq. (32)), the Munkres allocation yields an increasing number of selected D2D-CU pairs achieving FD-SIC (or HD-SIC) with $R_{u,min}$ (from an average of 0.8 for $R_{u,min} = 1.5$ Mbps to an average of 1.24 for $R_{u,min} = 3$ Mbps, with $\eta = -90$ dB). This corroborates the idea that the throughput decrease of No-SIC techniques with $R_{u,min}$ is more important than that of SIC techniques, to a point where the contribution of mutual SIC techniques in maximizing the throughput is more prominent when $R_{u,min}$ increases. This is verified by comparing the percentage decrease of D2D throughput for every algorithm when moving from $R_u = 1.5$ Mbps to $R_u = 3$ Mbps: a decrease of 39 %, 33 %, 22 %, and 13 % is observed for the algorithms FD-NoSIC, HD-NoSIC, FD-SIC, HD-SIC respectively. The greater decrease of FD-NoSIC performance compared to HD-NoSIC justifies the shift of the intersection point between FD-SIC and HD-SIC to the left when $R_{u,min}$ increases. Indeed, as explained in Section IX, FD-SIC and HD-SIC are applied when possible, on top of FD-NoSIC and HD-NoSIC respectively. If the performance gap between FD-NoSIC and HD-NoSIC diminishes, HD-SIC will outperform FD-SIC over a broader span of η values before FD-SIC eventually catches up and surpasses HD-SIC for smaller η values (i.e. for better SI cancellation capabilities of the devices). This evolution of FD-SIC and HD-SIC can

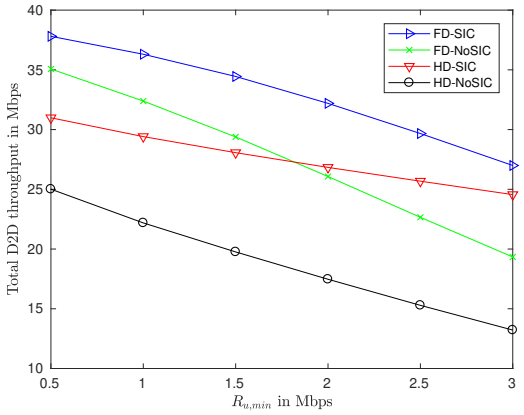


Figure 10: Total D2D throughput as a function of $R_{u,min}$ for $\eta = -110$ dB.

also be observed from another perspective in Fig. 10, where the total D2D throughput is presented as a function of the CU required rate. In the conditions of Fig. 10, the gap between FD-NoSIC and HD-NoSIC is large enough so that no intersection occurs between FD-SIC and HD-SIC. However, it can still be observed that the gap between FD-SIC and HD-SIC reduces as the CU required rate increases.

In Fig. 11, the variation of the total D2D throughput is presented as a function of the D2D maximum user distance d_{max}. The increase of d_{max} leads to a significant decrease in the performance of all proposed methods since h_d, the channel gain of the direct link between d_1 and d_2, is reduced on average. However, this increase of d_{max} is accompanied by

a greater percentage increase in performance due to mutual SIC for FD and HD transmission scenarios, with respect to No-SIC scenarios. Indeed, FD-SIC achieves a D2D throughput 128 % higher than FD-NoSIC for d_{max} = 100 m, compared to the 81 % increase achieved for d_{max} = 20 m. This is due to having more FD-SIC enabled D2D-CU pairs when distancing the D2D users further apart from one another, since an average of 1.96 pairs successfully apply FD-SIC for d_{max} = 20 m as opposed to 3.33 pairs for d_{max} = 100 m. The reason behind this increase is the decrease in h_d which relaxes the sufficient conditions (30) and (31), thereby enabling more FD-SIC cases.

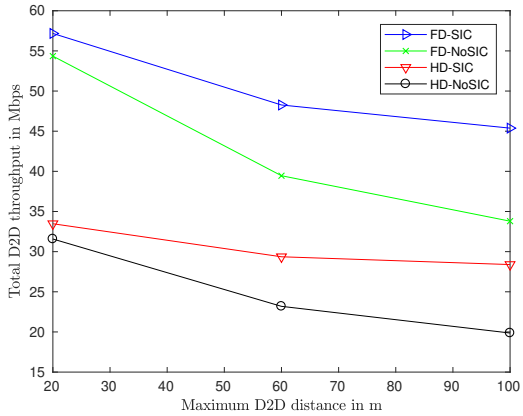


Figure 11: Total D2D throughput as a function of d_{max} for $\eta = -130$ dB.

Fig. 12 presents the evolution of the D2D throughput as a function of the number of CU users in the cell. Although the channel properties of the D2D users (i.e. h_d, $h_{d_1,u}$ and $h_{d_2,u}$) are unchanged, the total D2D throughput of all techniques benefits from the additional diversity provided by the greater number of CU users. This also favors the FD-SIC enabled pairs, as their average number grows from 1.98 for K = 20 to 2.46 for K = 50. We can therefore conclude that the important performance gain achieved by SIC methods, with respect to No-SIC methods, can be obtained without requiring the implementation of SIC at all D2D and CU receivers. Indeed, generally only 2 or 3 triplets need to perform SIC which is enough to boost the D2D system capacity, while the others can settle for the simple classical No-SIC receivers. Therefore, the additional complexity is localised at the level of the users performing SIC for which the major throughput increase is worth the incurred SIC complexity. Finally, the total and average throughput variations are presented in Fig. 13 as a function of the number of D2D pairs in the system, for a fixed value of K = 50. In Fig. 13a, the average throughput per D2D pair is shown to slightly decrease with the increasing number of D2D pairs. In a sense, this is the dual of the behavior observed in Fig. 12, since the ratio K/D decreases with D and thus the system diversity - in terms of the average number of possible CU channel choices for every D2D pair to be collocated on - decreases, thus reducing the achievable throughput per D2D pair. Nonetheless, the total throughput follows a quasi linear progression with the number of D2D pairs because the additional D2D pairs are

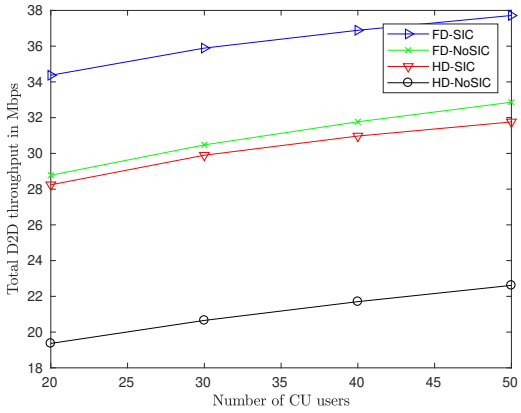
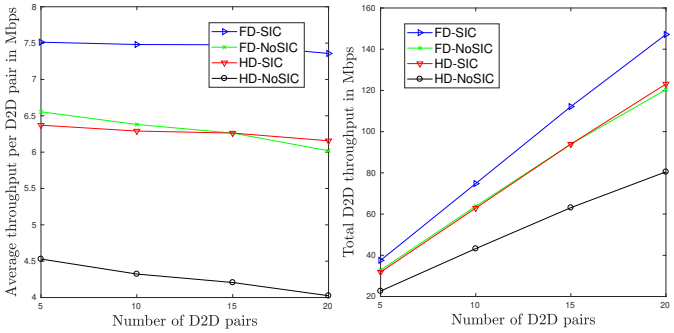


Figure 12: Total D2D throughput as a function of K for $\eta = -110$ dB.

allocated on orthogonal channels, therefore each D2D pair can be associated more or less to an additional D2D rate unit. Figs.



(a) Average D2D throughput (b) Total D2D throughput

Figure 13: Total and average D2D throughput as a function of the number of D2D pairs for K = 50 CUs and $\eta = -110$ dB.

12 and 13 indicate that, for a fixed number of $D2D$ users or CU users, the effect of the proportion K/D on the average D2D throughput per D2D pair is rather limited. The most dominant factors remain the distance between D2D users, the SI cancellation capabilities of the receivers (for FD-SIC), and the required CU rate.

XI. Conclusion

In this paper, the use of NOMA with mutual SIC was proposed for the first time between cellular users and FD-D2D devices underlaying the cellular channels. The necessary and sufficient conditions for applying FD-SIC were derived and a highly efficient PA procedure was elaborated to solve, in constant time operation, the throughput maximization problem of significant original complexity. The optimal, yet simple, PA resolution allowed for achieving global optimal resource allocation by conveniently combining the Kuhn-Munkres channel assignment with the proposed PA methods. The results show important performance gains obtained by applying SIC in D2D underlay systems in both HD and FD transmission schemes, promoting thereby the use of mutual SIC NOMA for D2D systems whenever possible. When applying mutual SIC, the comparison between HD and FD transmission scenarios

showed that FD-SIC is more efficient for average to high SI cancellation capabilities, moderate CU rate requirements and significant D2D distances, while HD-SIC performs better especially at low SI cancellation capabilities. Future work directions of the study could be to adapt the PA procedure for D2D underlay to downlink cellular systems which present different sets of challenges and networking paradigms. Moreover, the integration of our work in UAV-assisted NOMA networks for IoT could also be considered, as in the context of [35].

ACKNOWLEDGMENT

This work has been funded with support from IMT Atlantique and the Lebanese University Research Support Program.

APPENDIX A

To determine if the search space is totally included in region 1 or 2 for the cases 3) and 4), we introduce f_1, f_3 and f_λ, the functions of P_1, P_2 which yield the P_u value corresponding to the planes $\mathcal{PL}_1, \mathcal{PL}_3$ and to L_λ. A parametric equation of L_λ is given by:

$$L_\lambda = \begin{cases} x = (\frac{h_d}{h_{d_1,u}} - \frac{\eta_2}{h_{d_2,u}})m = (\tan(\gamma) - \tan(\tau))m \\ y = (\frac{h_d}{h_{d_2,u}} - \frac{\eta_1}{h_{d_1,u}})m = (\tan(\xi) - \tan(\Omega))m \\ z = \frac{h_d^2 - \eta_1\eta_2}{h_{d_1,u}h_{d_2,u}}m \end{cases}$$

In the case of Fig. 7a, the search space is included in region 2 if and only if L_λ is on top of $\mathcal{PL}_1$. For the case of Fig. 7b, the search space is included in region 1 if and only if L_λ is on top of $\mathcal{PL}_3$. To determine the conditions of each scenario, we first have to check if the conditions of case 3), where $\gamma > \tau$ and $\xi > \Omega$, or those of case 4), where $\gamma < \tau$ and $\xi < \Omega$, are met. To study the relative position of L_λ with respect to $\mathcal{PL}_1$ and $\mathcal{PL}_3$, m is chosen such that the comparison is conducted in the first octant. Since in case 3), $\gamma > \tau \Rightarrow \tan(\gamma) - \tan(\tau) > 0$, then m must be positive in case 3), and conversely, negative in case 4).

The search space is included in region 2 if:

$$f_\lambda(P_1, P_2) > f_1(P_1, P_2)$$
$$\Rightarrow \frac{h_d^2 - \eta_1\eta_2}{h_{d_1,u}h_{d_2,u}}m > \frac{P_2 h_{b,d_2} - P_1 h_{b,d_1}}{h_{b,u}}$$

Replacing P_1 by $(h_d/h_{d_1,u} - \eta_2/h_{d_2,u})m$, and P_2 by $(h_d/h_{d_2,u} - \eta_1/h_{d_1,u})m$, we get:

$$\frac{h_{b,u}(h_d^2 - \eta_1\eta_2)}{h_{d_1,u}h_{d_2,u}}m > (\frac{h_d}{h_{d_2,u}} - \frac{\eta_1}{h_{d_1,u}})m h_{b,d_2} - (\frac{h_d}{h_{d_1,u}} - \frac{\eta_2}{h_{d_2,u}})m h_{b,d_1}$$

Let Γ be the proposition

$$\frac{h_{b,u}(h_d^2 - \eta_1\eta_2)}{h_{d_1,u}h_{d_2,u}} > \frac{h_{b,d_2}h_d + h_{b,d_1}\eta_2}{h_{d_2,u}} - \frac{h_{b,d_2}\eta_1 + h_{b,d_1}h_d}{h_{d_1,u}}$$

Then, we conclude that:

- case 3) "Search Space included in region 2" $\Leftrightarrow \Gamma = 1$.
- case 4) "Search Space included in region 2" $\Leftrightarrow \Gamma = 0$

On the other hand, the search space is included in region 1 if:

$$f_\lambda(P_1, P_2) > f_3(P_1, P_2)$$
$$\Rightarrow (h_d^2 - \eta_1\eta_2)m > \frac{(h_d h_{d_2,u} - \eta_2 h_{d_1,u})h_{b,d_1}}{h_{b,u}}m$$

Let Ξ be the proposition:

$$(h_d^2 - \eta_1\eta_2)h_{b,u} > (h_d h_{d_2,u} - \eta_2 h_{d_1,u})h_{b,d_1} \quad (35)$$

Therefore, the search space is included in region 1 if $\Xi = 1$ for case 3), and $\Xi = 0$ for case 4).

Conclusion: to determine if the search space is completely included in one of the two regions, for case 3) and 4), we simply have to test the validity of Γ and Ξ and draw the corresponding conclusion to each case.

APPENDIX B

A. Optimal Throughput point over the Side S_1

The sign of $\partial F/\partial P_1$ is equal to the sign of the following second-degree polynomial of P_1: $aP_1^2 + bP_1 + c$, with $a = \eta_1^2$, $b = 2\eta_1\sigma^2$, and $c = P_{2,M}(h_d - \eta_1)\sigma^2 - P_{2,M}^2\eta_2\eta_1 + \sigma^4$. If $\Delta = b^2 - 4ac < 0$, the second-degree polynomial is positive, hence the throughput is increasing with P_1, and P_1^* is obtained by setting P_1 to $\max P_1$.

If $\Delta > 0$, the polynomial is negative inside the solutions interval, and positive elsewhere. The solutions are: $sol_1 = (-b - \sqrt{\Delta})/2a$, $sol_2 = (-b + \sqrt{\Delta})/2a$. Therefore, the throughput is decreasing between sol_1 and sol_2, then increasing for $P_1 > sol_2$. Since $sol_1 < 0$, three cases are identified depending on the location of sol_2 with respect to $\min P_1$ and $\max P_1$:

- $sol_2 < \min P_1$: the throughput increases with $P_1 \Rightarrow P_1^* = \max P_1$.
- $sol_2 > \max P_1$: the throughput decreases with $P_1 \Rightarrow P_1^* = \min P_1$.
- $sol_2 \in [\min P_1, \max P_1]$: as shown in the variation table of Fig. 14, the throughput is decreasing between $\min P_1$ and sol_2, and increasing between sol_2 and $\max P_1$. Therefore, we obtain $P_1^* = \arg\max[F(\min P_1), F(\max P_1)]$.

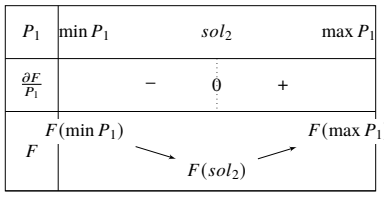


Figure 14: R_{D2D} variation table when $sol_2 \in [\min P_1, \max P_1]$

Finally, no matter if $\Delta > 0$ or < 0, it is sufficient to test which of $\min P_1$ or $\max P_1$ delivers the best throughput and then select the corresponding segment endpoint.

B. Optimal Throughput point over the Side S_U

Consider the sign of the sign of the polynomial's discriminant $\Delta = B^2 - 4AC$. If $\Delta < 0$: $\operatorname{sgn}\partial F/\partial P_1 = \operatorname{sgn}(h_d^2 - \eta_1\eta_2)$.

- If $h_d^2 > \eta_1\eta_2$, F is increasing with $P_1 \Rightarrow$ Set P_1^* to $\max P_1$
- If $h_d^2 < \eta_1\eta_2$, F is decreasing with $P_1 \Rightarrow$ Set P_1^* to $\min P_1$

However, if $\Delta > 0$, then we have the two solutions $sol_1 = (-B - \sqrt{\Delta})/2A$ and $sol_2 = (-B + \sqrt{\Delta})/2A$, with the variation tables (Figs. 15 and 16) depending on the sign of $h_d^2 - \eta_1\eta_2$.

- If $h_d^2 > \eta_1\eta_2 \Rightarrow sol_1 < sol_2$ and $sol_1 < 0$. But not much can be said about the sign of sol_2 and how it compares to $\min P_1$ and $\max P_1$.

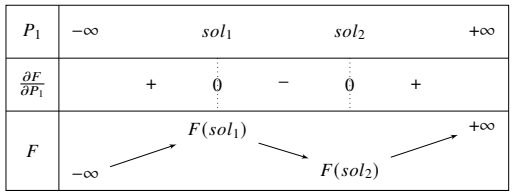


Figure 15: Variation table for $h_d^2 > \eta_1\eta_2$

However, we note that the right side of the variation table (where $P_1 > sol_1$) is similar to the variation table in Fig. 14. Therefore, we conclude that:

$$P_1^* = \arg\max[F(\min P_1), F(\max P_1)].$$

- if $h_d^2 < \eta_1\eta_2 \Rightarrow sol_2 < sol_1$ and $sol_1 > 0$.

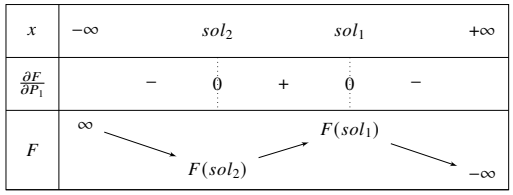


Figure 16: Variation table for $h_d^2 < \eta_1\eta_2$

Since sol_1 is a local maximum, $F(sol_1) > F(P_1), \forall P_1 > sol_2$. Then, the only values of P_1 which might give a better throughput than sol_1 are those at the left of sol_2. We can distinguish the following three cases:

- if $\max P_1 < sol_1$, set P_1^* to $\max P_1$.
- if $sol_1 < \min P_1$, set P_1^* to $\min P_1$.
- if $sol_1 \in [\min P_1, \max P_1]$, then:
  - if $\min P_1 > sol_2$, set P_1^* to sol_1.
  - if $\min P_1 < sol_2$, set $P_1^* = \arg\max[F(\min P_1), F(sol_1)]$.

To sum up, in the optimization over the intersection segment of $\mathcal{PL}_2$ with S_U, all the possible channel conditions lead at some point to choosing P_1^* from the values $\min P_1$, $\max P_1$, and sol_1 (when it is included in the interval $\mathbb{U}$) according to the one delivering the highest throughput.

C. Endpoint Coordinates

The coordinates of the segment endpoints $k_1, k_2, j_2, x_1, x_2, x_U, s_1, s_2, s_U, v_1, v_2, v_3, v_4, v_5, g_1, g_2, g_3, g_4, g_5$ are given below. Note that x_i and xl_i (resp. s_i and sl_i) have the same expressions with the difference that xl_i (resp. sl_i) is not defined outside of S_i. Moreover, x_i and s_i have strictly positive coordinates since $h_{d_1,u}h_{b,d_2} - h_{b,u}h_d > 0$ from eq. (8), and $h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u} > 0$ from eq (28).

$$
\begin{aligned}
k_1 &= (P_{1,M}, (P_{u,m}h_{bu} + P_{1,M}h_{b,d_1})/h_{b,d_2}, P_{u,m}),\\
k_2 &= ((P_{2,M}h_{b,d_2} - P_{u,m}h_{b,u})/h_{b,d_1}, P_{2,M}, P_{u,m}),\\
j_2 &= (P_{u,m}h_{b,u}/h_{b,d_1}, P_{2,M}, P_{u,m}),\\
x_1 &= (1, \frac{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}}{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d}, \frac{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d})P_{1,M},\\
x_2 &= (\frac{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d}{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}}, 1, \frac{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}})P_{2,M},\\
s_1 &= (1, \frac{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}}{h_{b,u}h_d}, \frac{h_{b,d_1}}{h_{b,u}})P_{1,M},\\
s_2 &= (\frac{h_{b,u}h_d}{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}}, 1, \frac{h_{b,d_1}h_d}{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}})P_{2,M},\\
s_u &= (\frac{h_{b,u}}{h_{b,d_1}}, \frac{h_{b,d_1}h_{d_1,u} - \eta_1 h_{b,u}}{h_{b,d_1}h_d}, 1)P_{u,M},\\
x_u &= (\frac{h_{d_1,u}h_{b,d_2} - h_{b,u}h_d}{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}, \frac{h_{b,u}\eta_1 + h_{d_1,u}h_{b,d_1}}{h_{b,d_2}\eta_1 + h_d h_{b,d_1}}, 1)P_{u,M}.
\end{aligned}
$$

The w_i family is obtained by combining v_i and g_i.

$$
\begin{aligned}
v_1 &= (P_{1,M}, (P_{u,m}h_{d_1,u} - P_{1,M}\eta_1)\tfrac{1}{h_d}, P_{u,m}),\\
v_2 &= ((P_{u,m}h_{d_1,u} - P_{2,M}h_d)\tfrac{1}{\eta_1}, P_{2,M}, P_{u,m}),\\
v_3 &= (P_{1,M}, P_{2,M}, (P_{1,M}\eta_1 + P_{2,M}h_d)\tfrac{1}{h_{d_1,u}}),\\
v_4 &= (P_{1,M}, (P_{u,M}h_{d_1,u} - P_{1,M}\eta_1)\tfrac{1}{h_d}, P_{u,M}),\\
v_5 &= ((P_{u,M}h_{d_1,u} - P_{2,M}h_d)\tfrac{1}{\eta_1}, P_{2,M}, P_{u,M}),\\
g_1 &= (P_{1,M}, (P_{u,m}h_{d_2,u} - P_{1,M}h_d)\tfrac{1}{\eta_2}, P_{u,m}),\\
g_2 &= ((P_{u,m}h_{d_2,u} - P_{2,M}\eta_2)\tfrac{1}{h_d}, P_{2,M}, P_{u,m}),\\
g_3 &= (P_{1,M}, P_{2,M}, (P_{1,M}h_d + P_{2,M}\eta_2)\tfrac{1}{h_{d_2,u}}),\\
g_4 &= (P_{1,M}, (P_{u,M}h_{d_2,u} - P_{1,M}h_d)\tfrac{1}{\eta_2}, P_{u,M}),\\
g_5 &= ((P_{u,M}h_{d_2,u} - P_{2,M}\eta_2)\tfrac{1}{h_d}, P_{2,M}, P_{u,M}).
\end{aligned}
$$

References

[1] Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2018-2023," Mar. 2020.

[2] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-Device Communication as an Underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec 2009.

[3] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF Canceller for In-Band Full-Duplex Wireless Communications," *IEEE Trans. on Wireless Commun.*, vol. 15, no. 6, pp. 4321–4334, June 2016.

[4] L. Laughlin, C. Zhang, M. A. Beach, K. A. Morris, and J. Haine, "A Widely Tunable Full Duplex Transceiver Combining Electrical Balance Isolation and Active Analog Cancellation," in *2015 IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–5.

[5] Y. Choi and H. Shirani-Mehr, "Simultaneous Transmission and Reception: Algorithm, Design and System Level Performance," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 12, pp. 5992–6010, 2013.

[6] D. Zhai, R. Zhang, L. Cai, B. Li, and Y. Jiang, "Energy-Efficient User Scheduling and Power Allocation for NOMA-Based Wireless Networks With Massive IoT Devices," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1857–1868, 2018.

[7] J. Farah, E. Sfeir, C. Abdel Nour, and C. Douillard, "New Resource Allocation Techniques for Base Station Power Reduction in Orthogonal and Non-Orthogonal Multiplexing Systems," in *2017 IEEE Int. Conf. on Commun. Workshops (ICC Workshops)*, May 2017, pp. 618–624.

[8] M. J. Youssef, J. Farah, C. Abdel Nour, and C. Douillard, "Resource Allocation for Mixed Traffic Types in Distributed Antenna Systems Using NOMA," in *2018 IEEE 77th Veh. Technol. Conf. (VTC fall)*, Aug. 2018, pp. 1–5.

[9] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On Optimal Power Allocation for Downlink Non-Orthogonal Multiple Access Systems," *IEEE J. on Sel. Areas in Commun.*, vol. 35, no. 12, pp. 2744–2757, 2017.

[10] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[11] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *2013 IEEE 77th Veh. Tech. Conf. (VTC Spring)*, June 2013, pp. 1–5.

[12] M. Youssef, J. Farah, C. A. Nour, and C. Douillard, "Resource Allocation in NOMA Systems for Centralized and Distributed Antennas With Mixed Traffic Using Matching Theory," *IEEE Trans. on Commun.*, vol. 68, no. 1, pp. 414–428, 2020.

[13] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-Level Performance Evaluation of Downlink Non-Orthogonal Multiple Access (NOMA)," in *2013 IEEE 24th Annual Int. Symp. on Personal, Indoor, and Mobile Radio Comm. (PIMRC)*, Sep 2013, pp. 611–615.

[14] B. Chen, Y. Chen, Y. Chen, Y. Cao, Z. Ding, N. Zhao, and X. Wang, "Secure primary transmission assisted by a secondary full-duplex noma relay," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 7214–7219, 2019.

[15] Y. Cao, N. Zhao, G. Pan, Y. Chen, L. Fan, M. Jin, and M. Alouini, "Secrecy Analysis for Cooperative NOMA Networks with Multi-Antenna Full-Duplex Relay," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5574–5587, 2019.

[16] A. Jaiswal, S. Kumar, O. Kaiwartya, N. Kumar, H. Song, and J. Lloret, "Secrecy Rate Maximization in Virtual-MIMO Enabled SWIPT for 5G Centric IoT Applications," *IEEE Syst. J.*, pp. 1–12, 2020.

[17] A. U. Makarfi, K. M. Rabie, O. Kaiwartya, O. S. Badarneh, X. Li, and R. Kharel, "Reconfigurable Intelligent Surface Enabled IoT Networks in Generalized Fading Channels," in *ICC 2020 - 2020 IEEE Int. Conf. Commun. (ICC)*, 2020, pp. 1–6.

[18] S. Kumar, O. Kaiwartya, M. Rathee, N. Kumar, and J. Lloret, "Toward Energy-Oriented Optimization for Green Communication in Sensor Enabled IoT Environments," *IEEE Systems J.*, 2020.

[19] J. Chen, J. Jia, Y. Liu, X. Wang, and A. H. Aghvami, "Optimal Resource Block Assignment and Power Allocation for D2D-Enabled NOMA Communication," *IEEE Access*, vol. 7, pp. 90 023–90 035, 2019.

[20] Y. Pan, C. Pan, Z. Yang, and M. Chen, "Resource Allocation for D2D Communications Underlaying a NOMA-Based Cellular Network," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 130–133, Feb 2018.

[21] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, and M. Elkashlan, "Joint Subchannel and Power Allocation for NOMA Enhanced D2D Communications," *IEEE Trans. on Commun.*, vol. 65, no. 11, pp. 5081–5094, Nov 2017.

[22] S. Alemaishat, O. A. Saraereh, I. Khan, and B. J. Choi, "An Efficient Resource Allocation Algorithm for D2D Communications Based on NOMA," *IEEE Access*, vol. 7, pp. 120 238–120 247, 2019.

[23] Y. Wang, D. Zhai, R. Zhang, and Z. Zhang, "Sum-Rate Maximization for D2D and Cellular Hybrid Networks Enhanced by NOMA," in *2019 IEEE 20th International Conf. on High Performance Switching and Routing (HPSR)*, 2019, pp. 1–5.

[24] Y. Dai, M. Sheng, J. Liu, N. Cheng, X. Shen, and Q. Yang, "Joint Mode Selection and Resource Allocation for D2D-Enabled NOMA Cellular Networks," *IEEE Trans. on Veh. Technol.*, vol. 68, no. 7, pp. 6721–6733, 2019.

[25] D. Zhai, R. Zhang, Y. Wang, H. Sun, L. Cai, and Z. Ding, "Joint User Pairing, Mode Selection, and Power Control for D2D-Capable Cellular Networks Enhanced by Nonorthogonal Multiple Access," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8919–8932, 2019.

[26] J. Farah, A. Kilzi, C. Abdel Nour, and C. Douillard, "Power Minimization in Distributed Antenna Systems Using Non-Orthogonal Multiple Access and Mutual Successive Interference Cancellation," *IEEE Trans. on Veh. Technol.*, vol. 67, no. 12, pp. 11 873–11 885, Dec. 2018.

[27] A. Kilzi, J. Farah, C. A. Nour, and C. Douillard, "New Power Minimization Techniques in Hybrid Distributed Antenna Systems With Orthogonal and Non-Orthogonal Multiple Access," *IEEE Trans. on Green Commun. and Netw.*, vol. 3, no. 3, pp. 679–690, Sep. 2019.

[28] A. Kilzi, J. Farah, C., and C. Douillard, "Mutual Successive Interference Cancellation Strategies in NOMA for Enhancing the Spectral Efficiency of CoMP Systems," *IEEE Trans. on Commun.*, vol. 68, no. 2, pp. 1213–1226, Feb 2020.

[29] P. Mach, Z. Becvar, and T. Vanek, "In-Band Device-to-Device Communication in OFDMA Cellular Networks: A Survey and Challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1885–1922, 2015.

[30] H. Chour, F. Bader, Y. Nasser, and O. Bazzi, "GALEN: A Geometric Framework for Global Optimal Power Allocation in a Full Duplex D2D Network," in *2019 IEEE Wireless Commun. and Netw. Conf. (WCNC)*, April 2019, pp. 1–7.

[31] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo, H. S. Dhillon, and T. D. Novlan, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, June 2012.

[32] H. Tabassum, E. Hossain, and J. Hossain, "Modeling and analysis of uplink non-orthogonal multiple access in large-scale cellular networks using poisson cluster processes," *IEEE Trans. on Commun.*, vol. 65, no. 8, pp. 3555–3570, Aug. 2017.

[33] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, no. 1, pp. 32–38, 1957. [Online]. Available: https://doi.org/10.1137/0105003

[34] H. Cui and J. Zhang and C. Cui and Q. Chen, "Solving large-scale assignment problems by Kuhn-Munkres algorithm," in *International Conference on Advances in Mechanical Engineering and Industrial Informatics. Hangzhou, Zhejiang*, 2016, pp. 822–827.

[35] M. Liu, J. Yang, and G. Gui, "DSF-NOMA: UAV-assisted Emergency Communication Technology in a Heterogeneous Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5508–5519, 2019.