



HAL
open science

Global Optimal Resource Allocation for Efficient FD-D2D Enabled Cellular Network

Hussein Chour, Eduard Jorswieck, Faouzi Bader, Youssef Nasser, Oussama
Bazzi

► **To cite this version:**

Hussein Chour, Eduard Jorswieck, Faouzi Bader, Youssef Nasser, Oussama Bazzi. Global Optimal Resource Allocation for Efficient FD-D2D Enabled Cellular Network. IEEE Access, 2019, 7, pp.59690 - 59707. 10.1109/ACCESS.2019.2914973 . hal-02122437

HAL Id: hal-02122437

<https://hal.science/hal-02122437>

Submitted on 7 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2019.DOI

Global Optimal Resource Allocation for Efficient FD-D2D Enabled Cellular Network

HUSSEIN CHOUR^{1,4}, (Student Member, IEEE), EDUARD A. JORSWIECK²,
(Senior Member, IEEE), FAOUZI BADER¹, (Senior Member, IEEE),
YOUSSEF NASSER³, (Senior Member, IEEE), AND OUSSAMA BAZZI⁴, (Member, IEEE)

¹SCEE Research Team, CentraleSupélec/IETR, 35576 Rennes, France

²Faculty of Electrical and Computer Engineering, Communications Laboratory, TU Dresden, Dresden 01062, Germany

³Department of Electrical and Computer Engineering, American University of Beirut, Beirut 11-0236, Lebanon

⁴Faculty of Sciences 1, Department of Physics and Electronics, Lebanese University, Beirut 6573, Lebanon

Corresponding author: Hussein Chour (e-mail: hussein.chour@supelec.fr).

This work was supported in part by CentraleSupélec/IETR research Lab. A part of this work was also carried out thanks to the STSM grant of COST Action CA15104 (IRACON) and CNRS (GdR-ISIS).

ABSTRACT This paper tackles the resource allocation (RA) problem of a Full-Duplex (FD) device-to-device (D2D) communications enabled cellular network. In the considered scenario, multiple FD-D2D pairs share the uplink channels of the regular cellular users (CUs) which leads to mutual interference between the two communication types. Within this interference environment, this work aims to properly allocate the network's resources, such as the transmit power and the channels, to maximize a network-centric metric like the weighted-sum rate (WSR) and the global energy efficiency (GEE). The complex coupling between the mutual interference of the different links, as well as the flexibility of assigning the channels to the users, results in a non-convex RA optimization problem, for which the global optimal solution is hard to obtain. This paper is a first and innovative approach that globally solves the RA problem of an FD-D2D based cellular network. In particular, we show that the global optimal solution can be achieved by decoupling the original problem into two sub-problems as power allocation (PA) and channel assignment (CA). The PA sub-problem is solved by means of monotonic optimization theory. Precisely, we propose a new polyblock-based algorithm, MARIO, which efficiently converges to the global solution of the PA problem. Then, based on the optimal PA solution, the CA problem reduces to an assignment problem, which can be solved by Khun-Munkers algorithm. Further, we propose a sub-optimal solution by solving the original RA problem in the reverse order, i.e., first assigning the channel and then allocating the power. Simulation results show the effectiveness of the proposed algorithms and provide important insights on the solution design parameters such as the proximity distance and the self-interference cancellation capability.

INDEX TERMS Channel assignment, device-to-device (D2D), energy efficiency, full-duplex (FD), power allocation, resource allocation, weighted sum-rate.

I. INTRODUCTION

THE proliferation of mobile devices and data-hungry applications imposes severe challenges to current wireless communication networks. For example, it is expected that the number of connected devices will surpass 50 billion by 2020 and that a 1000-fold increase in data rate is required to accommodate such an enormous number of devices [1]. Given the strict requirements defined for the future wireless networks, the increasing data rate should not lead to crush-

ing the available cellular spectrum as well as consuming more energy. Full-Duplex communication (FD) and device-to-device communication (D2D) are emerging as two candidates technologies for the next cellular network (5G) [2], [3].

D2D allows two devices in proximity of each other to initiate a direct link in a peer-to-peer fashion without the base station (BS) participation or with limited participation [2]. In literature, two different D2D scenarios are proposed named as D2D overlay communication and D2D underlay

communication [4]. The former prevents the co-channel interference between the D2D and the traditional cellular signals by allocating dedicated cellular resources for the D2D links. On the contrary, in the underlying scenario the D2D devices reuse the spectrum resources of the cellular users (CUs) [5]. Thus in the underlying scenario, rigorous interference management is required to improve the network performance. Due to the spectrum scarcity problem, the underlay D2D communication has gained much attention [4], [5], and several resource allocation techniques, including channel assignment, mode selection, and power control have been extensively explored [6]–[8].

On the other hand, FD technology enables a network node to simultaneously transmit and receive in the same time slot at the same frequency band [3]. However, FD introduces a new interference type to the cellular network named as self-interference (SI) [9]. This is because the receiver of an FD device may receive the transmitted signal by its own transmitter. Thus, the FD device may not always outperform the traditional half-duplex (HD) device. Fortunately, the recent antenna architecture of FD devices can achieve a high level of SI cancellation especially for the devices with low transmission power [10], [11]. Thus, the FD technology is getting closer to being applied in the 5G network.

Given the short distance property of the D2D communication, the transmission power of the D2D devices is relatively lower than the traditional CUs. Hence, integrating FD with D2D would be an excellent choice to further improve the spectrum efficiency and increase the data rate [12]. However, the practical FD devices cannot perfectly eliminate the SI. Thus, the residual SI (RSI) highly affects the performance of FD transmission, and in some situation, it might be worse than the HD technology. In addition, applying FD into D2D technology invokes additional co-channel interference in the network which may degrade its performance when compared with the traditional HD-D2D communication. Despite the complicated interference environment of an FD-D2D underlying cellular network, a non-trivial gain can be earned if efficient interference management between the FD-D2D devices and the CUs is adopted [13]. An overview of the different resource allocation techniques is provided below.

A. RELATED WORK AND CONTRIBUTION

Similar to the traditional HD-D2D network, the envisioned FD-D2D network can be classified into FD-D2D overlay/underlay cellular network. The works in [14]–[16] analyze the overlay FD-D2D scenario. Specifically, the work in [14] proposed a simple protocol to improve the rate of a single FD-D2D pair which is allocated a dedicated spectrum. Moreover, in [15] a power allocation scheme is developed to maximize the ergodic bitrate of an FD-D2D pair overlay cellular network. Besides, the authors of [16] aimed to maximize the effective capacity of an FD-D2D network while satisfying the statistical delay-bound QoS requirements. However, allocating a dedicated spectrum for the FD-D2D pairs as in the overlay scenario would lead to a low spectrum efficiency and

diminish the gain of FD transmission. That is why the focus is more about FD-D2D underlay cellular network [17]–[25].

The authors of [17] investigated the performance of an underlay cellular FD-D2D network by using the stochastic geometry. Simulation results showed that an FD-D2D network has a significant gain over its counterpart HD-D2D especially when the SI is low. However, in their work, only a simple on- off power allocation was used. To limit the D2D interference at the cellular link, an upper bound transmit power was derived in [18]. Moreover, an interference limited area method was used to mitigate the interference from the cellular users to the D2D link. The ergodic sum rate of a single FD-D2D pair which coexist with a cellular user was derived and analyzed in [20], [21]. In [22] a convex optimization problem that maximizes the FD-D2D link rate while satisfying the data rate requirement of the CU was derived. Aiming to maximize the energy efficiency of an FD-D2D link, the authors of [23] proposed an energy-efficient power control algorithm for an FD-D2D network. However, all the above mentioned works only focus on the Power Allocation (PA) problem of an FD-D2D network with single D2D pair and single cellular user. In addition, and due to the non-convexity feature of the PA problem, they only provide a sub-optimal power allocation solution by approximating the original PA problem with a more simpler convex problem.

The resource allocation problem of an FD-D2D network when multiple FD-D2D pairs coexist with multiple CUs has been addressed only in [24] and [25]. The authors of [24] aimed to maximize the sum-rate of an FD-D2D network by considering both the power allocation PA and the channel assignment (CA) problems. Besides the sum-rate, the energy consumption became an important criteria design for the 5G network. In [25] an energy-efficient resource allocation for FD-D2D based cellular network was provided. However, both works provided the sub-optimal solution of the RA optimization problem. In particular, both [24] and [25] rely on the sequential convex optimization (SCO) theory to solve the PA problem of an FD-D2D based cellular network. The latter is guaranteed to provide only the first-order optimal solution. Thus, both works do not find the global optimal gain of an FD-D2D network, and the gap to the optimal solution is still unknown. To fill this blank, in this paper we derive the maximum gain of an FD-D2D network by providing an efficient global optimization framework. To that end, first, we formulate the RA problem as a maximization problem of a general signal to interference plus noise ratio (SINR) based network-centric metrics such as the weighted sum-rate (WSR) and the global energy efficiency (GEE). Next, we show that both the WSR maximization problem and the GEE optimization problem can be decomposed into two sub-problems as PA and CA. Hence, the global solution of the original RA problem can be found by globally solving the PA and CA sub-problems sequentially. The PA sub-problem is known to be a non-convex problem in an interference limited scenario, and thus it is hard to achieve the global optimal solution. To overcome this issue, we propose to use

the Monotonic Optimization theory (MO) to globally solve the PA problem. The MO theory can globally solve the optimization problems where the utility and the constraints are monotonically increasing functions of the optimization variables, even if the optimization problem is non-convex [26], [27]. In general, both WSR and GEE are non-monotone functions of the transmit power. However, they are monotone functions of SINR, and thus they exhibit hidden monotonicity structure that allows us to find the global solution by means of MO framework.

The MO problem can be solved by using the existing Outer Polyblock Approximation algorithm (OPA) at the expense of high complexity [26]. In particular, the OPA tries to approximate the outer boundary of the feasible set by a tight polyblock, and thus the global optimal point can be found by searching the set of the polyblock vertices. Using the existing OPA leads to a rapid growth in the number of vertices, and thus to high complexity. To solve this issue, we propose in this work a new power allocation algorithm which aims to decline the complexity of the OPA by trimming down the unnecessary vertices. For the sake of clarity, we denote this new algorithm as Polyblock vertices triMming bAsed power allocatION (MARIO). Then, based on the optimal PA solution, the CA sub-problem becomes an assignment problem which can be well addressed by Khun-Munkers algorithm (Hungarian algorithm) [28]. Hence, the optimal RA solution can be obtained by finding the global optimal transmit powers by means of MO theory for all the possible couplings between the FD-D2D pairs and the CUs and then assigning the best CU reuse partner for each FD-D2D pair by means of Khun-Munkers algorithm. Hence even with the proposed MARIO algorithm, the global optimal solution still has high complexity. Motivated by the need of an efficient RA solution, we propose a novel algorithm, CATPA, which first assigns the channels for the different FD-D2D pairs and then allocates the transmit power for each FD-D2D pair and its reuse partner. In particular, we propose a new metric, denoted as *Profit*, that quantifies the profit of an FD-D2D pair when reusing the channel of a CU, and then we assign the CU reuse partner for each D2D pair based on the highest profit values. After that, we allocate the transmit power for each D2D pair and its CU reuse partner by means of the SCO framework. This new algorithm highly reduces the complexity of the solution, since it avoids solving the PA problem for all the possible couplings between the FD-D2D pairs and the CUs. An outline of the conventional RA solution adopted in [24], [25] as well as the proposed RA solutions for an FD-D2D network is depicted in Fig.1. In this figure, $N \times M$ counts the number of all the possible couplings between M D2D pairs that coexist with N CUs. Besides, to ease following the paper, we show in Fig.1 the sections or the references where every step is developed.

The main contribution of this work can be summarized as follows:

- Deriving the ultimate gain of an FD-D2D network by globally solving the RA problem of an FD-D2D based

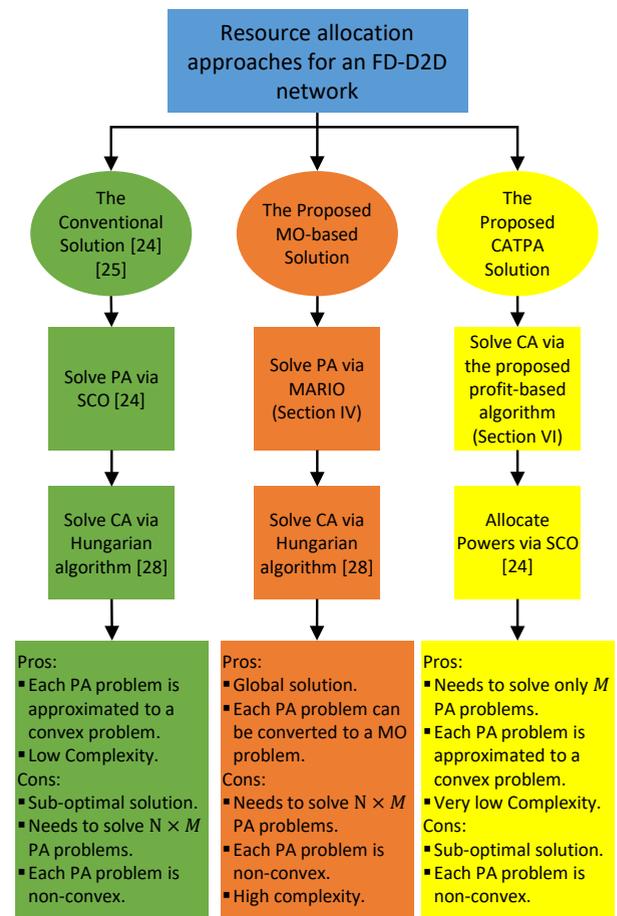


FIGURE 1. A comparison between the existing RA solution and the proposed RA solutions for an FD-D2D network.

cellular network, unlike the previous related works that only provide a sub-optimal RA scheme. The global solution is found by first decomposing the original RA problem into two sub-problems as PA and CA, and then globally solving each sub-problem individually.

- Providing insights into the importance of monotonic optimization theory in globally solving the PA optimization problem for an FD-D2D networks and more generally for any wireless cellular network. To the best of our knowledge, this is the first time in which the MO theory is applied to find the global RA solution in an FD-D2D based cellular network.
- Proposing a novel algorithm, referred to as MAPEL, to obtain the global optimal solution of the PA problem. The proposed algorithm is designed based on the existing OPA algorithm, and thus it obtains the solution by constructing a series of shrinking polyblocks that eventually closely approximate the boundary of the feasible set of the PA problem around the global solution. On the contrary to OPA, MARIO has a faster convergence time, since it trims down the unnecessary vertices of the generated polyblocks. The effectiveness of MARIO is

validated through numerical simulation which employs OPA as a benchmark. The results show that MARIO can achieve the same optimal point as OPA but within less number of iterations.

- Proposing an effective heuristic RA algorithm, referred to as CATPA, which first assigns the channels for the different users and then allocates the powers. We believe that this is the first work that provides such a methodology to solve the RA problem. To validate the proposed CATPA algorithm a numerical simulation is applied that employs the global optimal solution as a benchmark. Simulation results show that CATPA can achieve a solution within the 90% of the global optimal within a much lower number of iterations.
- Providing a comprehensive analysis of the FD-D2D network, by comparing the performance of the FD-D2D network with its counterpart HD-D2D network. Moreover, the effect of the different network parameters such as the QoS of the users, the maximum transmit powers of the devices, the D2D proximity distances, and the SI cancellation capability of the FD nodes, on the FD-D2D network performance have also been addressed.

B. OUTLINE AND NOTATION

The remainder of this paper is organized as follows. Section II presents the system model considered in this work while in Section III the resource allocation problem is formulated. Sections IV and V develop the MO-based solution of the derived RA problem, and Section VII presents the proposed CATPA solution. The numerical results are shown in Section VIII whereas the conclusion is made in Section VIII.

The following notation is used throughout this paper. Bold-face lower case letters stand for vectors, while lower case letter denotes scalars. \mathbb{R}^N is the $N \times 1$ real space vector and \mathbb{R}_+^N represents its non-negative orthant. The superscript T denotes the transpose of a matrix, and 0_N is the N dimensional all zeros vector. The gradient of a function $f(x)$ with respect to (w.r.t) x is represented by $\nabla_x f$. Moreover, we use $x \succeq y$ ($x \preceq y$) to indicate that $x \in \mathbb{R}^N$ is greater than or equal (less than or equal) to $y \in \mathbb{R}^N$ in a component-wise manner.

II. SYSTEM MODEL

This work tackles the resource allocation problem for an FD-D2D enabled cellular network. As depicted in Fig. 2 (a), we consider an uplink resource allocation scenario where N cellular users (CUs) share the uplink spectrum with M D2D pairs in a single cell system. In particular, uplink (UL) spectrum sharing is considered since UL resources are underutilized comparing to that of downlink (DL) [6], [29]. Furthermore, reusing UL resources in D2D links only affects the BS and incurred interference can be handled by BS coordination. The N CUs and the M D2D pairs are respectively organized in the sets $\mathcal{C} = \{CU_1, \dots, CU_N\}$ and $\mathcal{D} = \{D2D_1, \dots, D2D_M\}$, where CU_i and $D2D_j$ stand for the i th cellular user and the j th D2D pair respectively. Each $D2D_j$ consists of two nearby devices denoted as D_1^j

and D_2^j , and they have the capability to operate in FD mode while the CUs are assumed to operate only in HD mode.

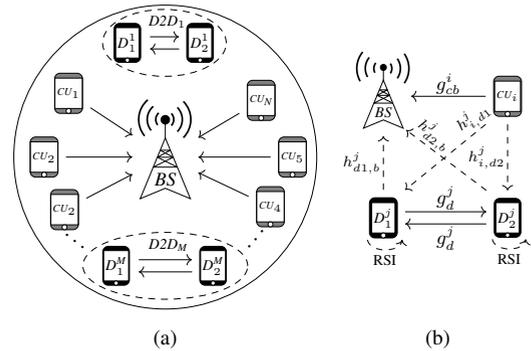


FIGURE 2. The adopted system model in which (a) multiple FD-D2D pairs coexist with multiple cellular users, and (b) the interference model when the cellular user CU_i shares its uplink resources with $D2D_j$.

The available UL bandwidth is composed of N orthogonal channels, and it is organized in the index set $CH = \{1, \dots, N\}$. Without loss of generality, the i th channel is assumed to be pre-assigned to CU_i . Thus, in this paper, we aim to allocate these channels to the M D2D pairs efficiently. Sharing the UL spectrum between the D2D users (DUs) and the CUs generates co-channel interference between the users sharing the same channel. Therefore, to limit the interference from the D2D pairs on the cellular links and avoid the interference between the D2D pairs, similar to [6], [7], [24], [25], we require that each UL channel is allowed to be shared with at most one D2D pair, and each $D2D_j$ can reuse at most one UL channel.

With the above channel assignment constraints and assuming imperfect SI cancellation, the interference model when $D2D_j$ and CU_i transmit on the i th channel can be depicted as in Fig. 2 (b). In this Figure, we define the channel gains as follows:

- g_{cb}^i denotes the direct channel gain between CU_i and BS.
- g_d^j stands for the direct channel gain between the D2D users of $D2D_j$. Here, the D2D channel is assumed reciprocal, since both D_1^j and D_2^j are using the same UL channel and they are close to each other.
- $h_{d1,b}^j$ and $h_{d2,b}^j$ respectively stand for the interference channel gain from D_1^j to BS and from D_2^j to BS.
- $h_{i,d1}^j$ and $h_{i,d2}^j$ denote the interference channel gain from CU_i to D_1^j and D_2^j respectively.
- RSI is the residual self interference due to the imperfect SI cancellation at the FD devices.

All direct/interference channels are assumed to be zero-mean complex Gaussian random variables (i.e., channels are expressing Rayleigh fading) with variance $l_{ij}^{-\alpha}$, where $i \in \{c; d1; d2\}$, $j \in \{bs; d1; d2\}$, $i \neq j$, l_{ij} denotes the distance between the nodes i and j , and α denotes the path loss exponent. The RSI can be modeled as complex Gaussian random variables with zero-mean and variance ηP_{tx} [15],

[22], [30]–[33], where P_{tx} is the transmission power of the FD device, η denotes the SI cancellation capability of the FD transmitter. Without loss of generality, assuming that all D2D users have the same SI cancellation capability, the power of RSI at D_1^j and D_2^j are respectively expressed by ηP_{d1}^j and ηP_{d2}^j , with P_{d1}^j , and P_{d2}^j being respectively the transmission powers of D_1^j , and D_2^j .

Now, denote by P_c^i the transmit power of CU_i , and let σ_N^2 represents the power of white Gaussian noise on each UL channel. Then, the Signal-to-Interference plus noise ratios (SINR) of CU_i , D_1^j , and D_2^j , when CU_i shares its channel with $D2D_j$, can be respectively expressed as

$$\Gamma^i(p_{ij}) = \frac{\Gamma_i^+}{\Gamma_i^-} = \frac{P_c^i g_{cb}^i}{P_{d1}^j h_{d1,b}^j + P_{d2}^j h_{d2,b}^j + \sigma_N^2}, \quad (1)$$

$$\Gamma^{j_1}(p_{ij}) = \frac{\Gamma_{j_1}^+}{\Gamma_{j_1}^-} = \frac{P_{d2}^j g_d^j}{P_c^i h_{i,d1}^j + \eta P_{d1}^j + \sigma_N^2}, \quad (2)$$

$$\Gamma^{j_2}(p_{ij}) = \frac{\Gamma_{j_2}^+}{\Gamma_{j_2}^-} = \frac{P_{d1}^j g_d^j}{P_c^i h_{i,d2}^j + \eta P_{d2}^j + \sigma_N^2}. \quad (3)$$

where $p_{ij} = [P_{d1}^j, P_{d2}^j, P_c^i]$ is the power allocation vector for $D2D_j$ and CU_i on the i th channel. Moreover, in the above equations, the SINR function Γ^k , $k \in \{i, j_1, j_2\}$, is expressed as a fraction of two non-negative functions denoted as $\Gamma_k^+(p_{ij})$ and $\Gamma_k^-(p_{ij})$. The benefits of these functions will appear later on when deriving the optimal RA scheme.

By applying Shannon's theorem, the data rates (in bits/s) of CU_i , D_1^j , and D_2^j can be respectively expressed as

$$\begin{aligned} R^i(p_{ij}) &= B_i \log_2(1 + \Gamma^i), i \in \{\mathcal{C}\} \\ &= B_i (\log_2(\Gamma_i^+ + \Gamma_i^-) - \log_2(\Gamma_i^-)) \\ &= r_i^+(p_{ij}) - r_i^-(p_{ij}) \end{aligned} \quad (4)$$

$$\begin{aligned} R^{j_1}(p_{ij}) &= B_i \log_2(1 + \Gamma^{j_1}), j \in \mathcal{D} \\ &= B_i \log_2(\Gamma_{j_1}^+ + \Gamma_{j_1}^-) - B_i \log_2(\Gamma_{j_1}^-) \\ &= r_{j_1}^+(p_{ij}) - r_{j_1}^-(p_{ij}) \end{aligned} \quad (4a)$$

$$\begin{aligned} R^{j_2}(p_{ij}) &= B_i \log_2(1 + \Gamma^{j_2}), j \in \mathcal{D} \\ &= B_i \log_2(\Gamma_{j_2}^+ + \Gamma_{j_2}^-) - B_i \log_2(\Gamma_{j_2}^-) \\ &= r_{j_2}^+(p_{ij}) - r_{j_2}^-(p_{ij}) \end{aligned} \quad (4b)$$

where B_i is the bandwidth of the i th UL channel. Since we are assuming that all the available channels are equally shared by the active CUs, in the sequel we refer to B_i by B . Besides, in the above equations, the achievable rate R^k , $k \in \{i, j_1, j_2\}$, is expressed as the difference of two non-negative functions $r_k^+ \triangleq B \log_2(\Gamma_k^+ + \Gamma_k^-)$ and $r_k^- \triangleq B \log_2(\Gamma_k^-)$. The application of this form will appear in Section IV when deriving the optimal power allocation scheme. Now, before terminating this section and started formulating the problem, we introduce a useful remark to be used later to derive the optimal RA scheme.

Remark 1: The functions r_k^+ , r_k^- , Γ_k^+ , Γ_k^- , $k \in \{i, j_1, j_2\}$ are all increasing functions with p_{ij} . However, the rate functions $R^i(p_{ij})$, $R^{j_1}(p_{ij})$, and $R^{j_2}(p_{ij})$, and the SINRs Γ^i , Γ^{j_1} , and Γ^{j_2} are in general non-increasing functions with p_{ij} .

A. PROBLEM FORMULATION

In a fully loaded UL network, the FD-D2D communication can be used to accommodate more users, and thus to enhance the performance of the cellular network. However, since the D2D pairs are reusing the cellular spectrum, a proper RA is required to maintain the QoS of the D2D pairs and their reuse CUs partners and mitigate the mutual interference between the different type of links. In this work, we aim to maximize the rate gain that can be brought by the D2D communication and minimize the energy cost of the D2D communication. Thus, we define two objective functions to be maximized while guaranteeing the QoS of all the links and respect the maximum allowed powers. The first utility function (U_1) is the weighted sum rate (WSR) of the admitted D2D pairs and their corresponding CUs reuse partners, and it is defined as

$$U_1(\rho, P) = \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} \varrho_{ij} [\omega_i R^i + \omega_{j_1} R^{j_1} + \omega_{j_2} R^{j_2}]. \quad (5)$$

where ω_i , ω_{j_1} , and ω_{j_2} are positive weights used to control the individual rates of CU_i , D_1^j , and D_2^j respectively. ϱ_{ij} is the resource reuse indicator for $D2D_j$ and CU_i , when $D2D_j$ shares the CU_i 's resources $\varrho_{ij} = 1$; otherwise $\varrho_{ij} = 0$. $P = [p_{ij}]$ is the power allocation matrix for all CUs and DUs and $\rho = [\varrho_{ij}]$ denotes the channel assignment matrix of the D2D pairs.

The second utility function (U_2) is the global energy efficiency (GEE) of the D2D pairs and their reuse partners. It is defined as the energy cost of the accepted D2D links and their corresponding cellular links, and it can be expressed as

$$U_2(\rho, P) = \frac{U_1(\rho, P)}{\sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} \rho_{ij} \mu (p_c^i + p_{d1}^j + p_{d2}^j) + 3MP_{cir}}. \quad (6)$$

wherein $\mu \geq 1$ is the inverse of the power amplifier efficiency at each transmitter. P_{cir} denotes the fixed circuit power consumption at each device, accounting for the dissipation in analog hardware and digital signal processing [34].

Therefore, the resource allocation problem can be formulated as follows

$$P1_l : \max_{(\rho, P) \in \Omega} U_l(\rho, P), l = 1, 2. \quad (7)$$

$$\Omega = \left\{ \Gamma^i \geq \gamma_{\min}^i = 2^{r_{\min}^i} - 1, \forall i \in \mathcal{C}, \quad (7a) \right.$$

$$\Gamma^{j_1} \geq \gamma_{\min}^{j_1} = 2^{r_{\min}^{j_1}} - 1, \forall j \in \mathcal{D}, \quad (7b)$$

$$\Gamma^{j_2} \geq \gamma_{\min}^{j_2} = 2^{r_{\min}^{j_2}} - 1, \forall j \in \mathcal{D}, \quad (7c)$$

$$\sum_i \varrho_{ij} \leq 1, \varrho_{ij} \in \{0, 1\}, \forall j \in \mathcal{D}, \quad (7d)$$

$$\sum_j \varrho_{ij} \leq 1, \varrho_{ij} \in \{0, 1\}, \forall i \in \mathcal{C}, \quad (7e)$$

$$0_3 \preceq p_{ij} \preceq p_{\max}, \forall j \in \mathcal{D}, i \in \mathcal{C} \left. \right\} \quad (7f)$$

where Ω is the feasible set of $P1$, $p_{\max} = [P_{\max}^{d1}, P_{\max}^{d2}, P_{\max}^c]$ is the maximum power vector for any possible pair ($D2D_j, CU_i$), and $0_3 = [0, 0, 0]$ is the minimum power vector. γ_{\min}^i , $\gamma_{\min}^{j_1}$, $\gamma_{\min}^{j_2}$ denote the minimum required SINR

to achieve the minimum rate requirements r_{\min}^i , $r_{\min}^{j_1}$, and $r_{\min}^{j_2}$ for CU_i , D_1^j , and D_2^j respectively. Hence, constraints (7a)-(7c) represent the QoS requirements (rate requirements) for CU_i , D_1^j , and D_2^j respectively. Constraint (7d) ensures that a D2D pair reuses at most the resource of one CU. Constraint (7e) indicates that a CU_i can share its resources with at most one D2D pair. Constraint (7f) ensures that the transmit power of any user is within the maximum limit.

III. PROBLEM DECOMPOSITION

The authors of [35] have proved that the power allocation problem by itself is an NP problem in an interference limited system. The resource allocation problem (P1) defined in the previous section contains both power allocation and channel assignment (CA) problems, and thus it is more complex than the PA problem which is already NP-hard. In this section, we show that the solution of both $P1_1$ and $P1_2$ can be found by decoupling the original problem into two sub-problems. The first one is the power allocation and the second is the channel assignment problem. The decomposition technique reduces the complexity of the original problem, but the NP-hardness of the PA still exists. In the next sections, we present efficient algorithms to globally solve the PA and the channel assignment problems individually.

A. DECOMPOSITION OF $P1_1$

Observe that in the utility function U_1 the variables q_{ij} and p_{ij} are disjoint and thus the maximization problem $P1_1$ can be rewritten as follows:

$$P1_1 : \underset{\rho_{ij} \in \{0,1\}}{\text{maximize}} \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} q_{ij} \underset{P_{ij}}{\text{maximize}} U_1^{ij}, \text{ s.t. } (7a) - (7f), \quad (8)$$

wherein $U_1^{ij} = \omega_i R_c^i + \omega_{j_1} R^{j_1} + \omega_{j_2} R_{d2}^j$ is the WSR of the pair $(D2D_j, UE_i)$. Accordingly, $P1_1$ can be decomposed to the following two sub-problems:

$$P2_1 : \underset{P_{ij} \in \Phi, i \in \mathcal{C}, j \in \mathcal{D}}{\text{maximize}} U_1^{ij}, \text{ s.t. } \Phi = \{(7a), (7b), (7c), (7f)\}, \quad (9)$$

$$P3_1 : \underset{\rho_{ij} \in \{0,1\}}{\text{maximize}} \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} q_{ij} U_1^{ij*}, \text{ s.t. } (7d), (7e), \quad (10)$$

where U_1^{ij*} is the optimal solution of the PA sub-problem ($P2_1$). Thus, the optimal solution of $P1_1$ can be found by finding the optimal power allocation of all the possible pairs $(D2D_j, UE_i)$, and then selecting the optimal CU reuse partner for each D2D pair.

B. DECOMPOSITION OF $P1_2$

Giving the fractional nature of the GEE utility function, it is clear that $P1_2$ belongs to the class of fractional programming theory. Thus, the solution of $P1_2$ can be found by means of fractional programming tools [36], such as Dinkelbach's

algorithm. The latter is an iterative process which requires to solve the following auxiliary problem at iteration k :

$$\begin{aligned} F(\lambda_k) &= \\ &\underset{(\rho, P) \in \Omega}{\text{maximize}} \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} q_{ij} [\omega_i R_c^i + \omega_{j_1} R^{j_1} + \omega_{j_2} R_{d2}^j] \\ &\quad - \lambda_k \left(\sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} \rho_{ij} \mu (p_c^i + p_{d1}^j + p_{d2}^j) + 3MP_{cir} \right) \\ &= \underset{(\rho, P) \in \Omega}{\text{maximize}} \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} q_{ij} \underbrace{\left(\omega_i R_c^i + \omega_{j_1} R^{j_1} + \omega_{j_2} R_{d2}^j - \lambda_k \mu (p_c^i + p_{d1}^j + p_{d2}^j) \right)}_{U_2^{ij}} - P_{cons} \\ &= \underset{\rho \in \{0,1\}}{\text{maximize}} \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} q_{ij} \underset{p_{ij}}{\text{maximize}} U_2^{ij}(\lambda_k, p_{ij}) \end{aligned} \quad (11)$$

wherein $P_{cons} = \lambda_k 3MP_{cir}$. The last step in (11) is because the variable q_{ij} is disjoint than the variables (λ_k, p_{ij}) , and P_{cons} is a constant. Then, and similar to $P1_1$, the maximization of GEE can be divided into PA and CA sub-problems. The only difference is that PA has to be iteratively solved until convergence. Mathematically, the PA and CA problems can be expressed respectively as follows.

$$P2_2 : \underset{P_{ij}, i \in \mathcal{C}, j \in \mathcal{D}}{\text{maximize}} U_2^{ij}, \text{ s.t. } (7a), (7b), (7c), (7f), \quad (12)$$

$$P3_2 : \underset{\rho_{ij} \in \{0,1\}}{\text{maximize}} \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{D}} q_{ij} U_2^{ij*}, \text{ s.t. } (7d), (7e), \quad (13)$$

where U_2^{ij*} is the optimal power allocation of problem $P2_2$. For the reader's convenience, we report in Algorithm 1 the Dinkelbach's algorithm which solves $P2_2$. Despite the complexity of solving $U_2^{ij}(\lambda_k, p_{ij})$, the Dinkelbach's algorithm exhibits a *super-linear* convergence rate, since λ_k is updated according to Newton's method [36]. For more detail about the fractional programming and Dinkelbach's algorithm, the reader may refer to [37]. Solving efficiently $U_2^{ij}(\lambda_k, p_{ij})$ at each iteration is the target of the next section.

Remark 2: *In the first iteration of Dinkelbach's algorithm, i.e., when $\lambda_0 = 0$, Algorithm 1 finds the solution of the WSR problem. Hence, it is highly desirable to develop an algorithm which can solve both WSR and GEE problems.*

Here it is worthy to note that the PA sub-problem ($P2_l$) and the CA sub-problem ($P3_l$) must be sequentially solved to obtain the global solution of the original RA problem $P1_l$, $l \in \{1, 2\}$. As indicated in the second column of Fig. 1, solving the PA sub-problem is the target of the next section while the CA is the task of Section V.

IV. POWER ALLOCATION

The complicated interference coupling between the D2D pairs and their reuse partners turns the PA problem to be a non-convex problem. Thus, the global optimal solution of PA is hard to obtain because it may require examining every point in the feasible set. The monotonic optimization (MO) theory provides a guaranteed convergence to the global optimal solution and reduces the computational complexity by

Algorithm 1 Dinkelbach's algorithm

- 1: Set $\epsilon > 0$, $\lambda_0=0$, and $k = 0$
- 2: **repeat**
- 3: Solve $P2_2$ defined in (12) and denote the optimal power vector as $p_{k,ij}^*$.
- 4: $k = k + 1$
- 5: $\lambda_k = \frac{\omega_i R_c^k(p_{ij}^{k*}) + \omega_{j_1} R_{j_1}^{k*}(p_{ij}^{k*}) + \omega_{j_2} R_{j_2}^k(p_{ij}^{k*})}{\mu(p_c^{k*} + p_{d1}^{k*} + p_{d2}^{k*}) + 3P_{cir}}$
- 6: **until** $U_2^{ij}(\lambda_{k-1}, p_{k-1,ij}^*) \leq \epsilon$

exploiting the monotonicity property of the utility function and the constraints to solve the optimization problem. The key idea behind MO is to search for the global solution on the outer boundary of the feasible set instead of exploring the complete feasible set of the problem. Nevertheless, we note that the complexity of MO increases exponentially with the number of optimization variables. However, it is much lower than general global optimization methods, which do not exploit the monotonicity property of the problem. Naturally, the usage of MO is limited to the problems enjoying a monotonicity structure as indicated in the following definition.

Definition 1. An optimization problem belongs to the class of MO if it can be written in the following form:

$$\underset{x}{\text{maximize}} f(x) \text{ s.t. } x \in \mathcal{G} \cap \mathcal{H}, \quad (14)$$

where sets \mathcal{G} and \mathcal{H} are respectively normal and co-normal closed sets and function, $f(x)$ is an increasing function on \mathbb{R}_+^N .

Recall that a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is increasing if $f(x) \leq f(y)$ when $0 \leq x \leq y$. A set $\mathcal{G} \subset \mathbb{R}^n$ is normal set, if for any point $x \in \mathcal{G}$, all other points x' such that $x' \leq x$ are also in set \mathcal{G} . A set $\mathcal{H} \subset \mathbb{R}^n$ is co-normal set, if for any point $x \in \mathcal{H}$, all other points x' such that $x' \geq x$ are also in set \mathcal{H} .

At a first sight and based on Definition 1 and Remark 1, the monotonic optimization theory appears to be non-applicable for $P2_1$ and $P2_2$. Thus, in an FD-D2D based cellular network, the first step to efficiently obtain an optimal PA scheme is to transform the non-convex PA problem to an MO problem as shown in Subsection IV-A.

A. PROBLEM TRANSFORMATION

The aim of this subsection is to convert the PA problem in an FD-D2D network to an MO problem. From **Remark 2**, we know that Dinkelbach's algorithm (Algorithm 1), which is proposed to solve the GEE power allocation problem, inherently solves the WSR problem at the first iteration when $\lambda = 0$. Hence, in this subsection, we extract the hidden monotonicity property only for $P2_2$ (the GEE PA problem). The monotonicity structure of $P1_1$ can be found in the same manner by setting $\lambda = 0$.

To that end, first we reshaped the objective function of $P1_2$, i.e., U_2^{ij} defined in (11), as follows:

$$\begin{aligned} U_2^{ij}(p_{ij}, \lambda_k) &= \sum_{l \in \{i, j_1, j_2\}} \omega_l r_l^+(p_{ij}) - \sum_{l \in \{i, j_1, j_2\}} \omega_l r_l^-(p_{ij}) \\ &\quad - \lambda_k \mu(p_c^i + p_{d1}^j + p_{d2}^j) \\ &= u^+(p_{ij}) - u^-(p_{ij}, \lambda_k) \end{aligned} \quad (15)$$

wherein r_l^+ and r_l^- with $l \in \{i, j_1, j_2\}$ are defined in (4)-(4b), and they are increasing functions as indicated in remark 1. The functions $u^+(p_{ij})$ and $u^-(p_{ij}, \lambda_k)$ are also increasing functions because each one is a summation of several increasing functions (Remark 1). Next, we introduce the slack variable $t = u^-(p_{\max}, \lambda_k) - u^-(p_{ij}, \lambda_k)$, and we reformulate $P2_2$ as follows:

$$\underset{(p_{ij}, t) \in \Phi}{\text{maximize}} U_2^{ij}(p_{ij}, t) = u^+(p_{ij}) + t \quad (16)$$

$$\Phi = \left\{ (t, p); 0 \leq t + u^-(p) \leq u^-(p_{\max}, \lambda_k) \right.$$

$$\left. \frac{\Gamma_l^+(p_{ij})}{c_l^+(p_{ij})} - \frac{\gamma_{\min}^l \Gamma_l^-(p_{ij})}{c_l^-(p_{ij})} \geq 0, l \in \{i, j_1, j_2\} \right. \quad (17)$$

$$\left. 0_3 \preceq p_{ij} \preceq p_{\max} \forall j \in \mathcal{D}, i \in \mathcal{C} \right\}. \quad (18)$$

The objective function of the above problem is now an increasing function, and thus to show that (16) is a monotonic optimization problem, it remains to verify that the feasible set Φ is an intersection of a normal and a co-normal set.

According to [26], [27], for any continuous increasing functions $h(x)$ and $g(x)$ on \mathbb{R}_+^n the sets $\mathcal{H} = \{x \in \mathbb{R}_+^n | h(x) \leq 0\}$ and $\mathcal{G} = \{x \in \mathbb{R}_+^n | g(x) \geq 0\}$ are respectively normal and co-normal sets. Hence, to extract the hidden monotonicity property of $P2$ it only remains to verify that all the constraints in the feasible sets are continuous and increasing functions. It is easily verified that in the feasible set Φ only the QoS constraints defined in (17) are in general non-increasing function. By using the same technique as before, i.e., introducing a slack variable, (17) can be transformed to an increasing function. However, since the complexity of monotonic optimization problem exponentially increases with the number of optimization variables, we first grouped the QoS constraints defined in (17) under a single constraint as shown in (19), and then we introduce an auxiliary variable.

$$\begin{aligned} (17) \equiv \min_{v, l \in \{i, j_1, j_2\}} & \left[c_v^+(p_{ij}) - \left(\sum_v c_v^-(p_{ij}) - \sum_{l, l \neq v} c_l^-(p_{ij}) \right) \right] = \\ & \underbrace{\min_{v, l \in \{i, j_1, j_2\}} \left[c_k^+(p_{ij}) + \sum_{l, l \neq v} c_l^-(p_{ij}) \right]}_{c^+(p_{ij})} - \underbrace{\sum_l c_v^-(p_{ij})}_{c^-(p_{ij})} \geq 0 \end{aligned} \quad (19)$$

Now, by introducing the auxiliary variable $s = c^-(p_{\max}) - c^-(p)$ the problem defined in (16) can be rewritten as follows,

$$P2_{mo} : \text{maximize}_{v=[p_{ij}, t, s]} U_2^{ij}(v) = u^+(p_{ij}) + t, \text{ s.t. } v \in \mathcal{S}_n \cap \mathcal{S}_c \quad (20)$$

$$\mathcal{S}_n = \left\{ (t, p_{ij}, s) : t + u^-(p_{ij}, \lambda_k) \leq u^-(p_{\max}, \lambda_k), \right. \\ \left. s + c^-(p_{ij}) \leq c^-(p_{\max}), p_{ij} \preceq p_{\max} \right\}$$

$$\mathcal{S}_c = \left\{ (t, p_{ij}, s) : t \geq 0, s + c^+(p_{ij}) \geq c^-(p_{\max}), p_{ij} \succeq 0_3 \right\}.$$

It is clear that all the functions in \mathcal{S}_n and \mathcal{S}_c are increasing functions. Thus, \mathcal{S}_n and \mathcal{S}_c are respectively normal and co-normal sets. Therefore, $P2_{mo}$ is an MO optimization problem, and thus by using the MO tools such that the outer Polyblock algorithm (OPA) [26, Algorithm 3], the global optimal power allocation for any pair (CU_i, D2D_j) can be obtained. However, the OPA algorithm appears to have high complexity [26]. Moreover, in our system model, we assume that M D2D pairs coexist with N CUs, and thus OPA must be applied $N \times M$ times to calculate the global optimal power allocation for all the possible (CU_i, D2D_j) couples. This makes OPA intractable for deriving the optimal RA scheme of an FD-D2D network. To overcome this issue, we develop a new polyblock-based algorithm, referred to as MARIO, which has a much lower complexity compared to OPA. To easily understand MARIO, first, we present in the next subsection the conventional polyblock-based algorithm (OPA), then we develop in Subsection IV-C the new MARIO algorithm.

B. OUTER POLYBLOCK APPROXIMATION ALGORITHM

The main idea of OPA is to construct a sequence of shrinking polyblocks that eventually closely approximate the global solution which is located on the upper boundary of the feasible region. However, as mentioned earlier, OPA has high complexity feature. Thus, we propose a novel polyblock-based algorithm called MARIO. To understand MARIO the OPA concept has to be very clear for the reader. Thus, in the remaining of this subsection, the OPA process is made formal and clarified.

Definition 2. (Upper boundary): A point $x \in \mathbb{R}_+^N$ is called an upper boundary point of a normal closed set \mathcal{F} if $\mathcal{F} \cap \{x' \in \mathbb{R}_+^N | x' > x\} = \emptyset$. The set of all upper boundary points of \mathcal{F} is the upper boundary of \mathcal{F} and it is denoted as $\partial^+ \mathcal{F}$.

Definition 3. (Box): Given any two vectors a and $b \in \mathbb{R}_+^N$, the hyper rectangle $[a, b] = \{x | a \preceq x \preceq b\}$ is referred to as a box.

Definition 4. (Polyblock): Given any finite set $\mathcal{T} \in \mathbb{R}_+^N$ with elements v_i , the union of all the boxes $[0, v]$ is referred to as a polyblock with vertex set \mathcal{T} . A polyblock is clearly a normal set.

Definition 5. (Proper vertices): A vertex $v \in \mathcal{T}$ is said to be proper if there is no $v' \in \mathcal{T}$ such that $v' \neq v$ and $v' \succeq v$. A vertex is said to be improper if it is not proper, and improper

vertices can be removed from \mathcal{T} without affecting the shape of the polyblock.

Definition 6. (Projection): Given any nonempty compact normal set $\mathcal{G} \subset \mathbb{R}_+^N$, and any $v \in \mathbb{R}_+^N \setminus \mathcal{G}$, $\pi_{\mathcal{F}}(z)$ is a projection of z on \mathcal{F} if $\pi_{\mathcal{F}}(z) = \alpha z$ with $\alpha = \max\{\kappa > 0 | \kappa z \in \mathcal{F}\}$. In other words, $\pi_{\mathcal{F}}(z)$ is the unique point where the line segment joining 0 to z meets the upper boundary of \mathcal{F} .

Remark 3: Due to the normality of \mathcal{G} , the projection point, i.e., α , can be found by the well known bisection algorithm [26, Algorithm 1].

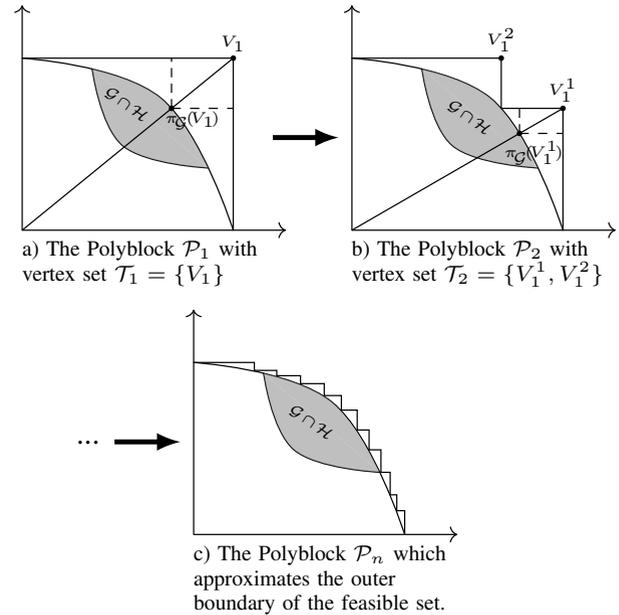


FIGURE 3. Illustration of the OPA process.

Based on the above concepts, the OPA which solves the MO problem $P2_{mo}$ works as follows. First, a polyblock \mathcal{P}_1 that contains the feasible set $\mathcal{G} \cap \mathcal{H}$ of an MO problem is constructed (see Fig.3 a). Let \mathcal{T}_1 be the proper vertex set of \mathcal{P}_1 . By Proposition 2 of [26], the maximum of the utility function of an MO problem over the polyblock \mathcal{P}_1 occurs at some proper vertex $v_1 \in \mathcal{T}_1$. If the projection of v_1 on $\partial^+ \mathcal{G}$ happens to reside on \mathcal{G} , i.e., $\pi_{\mathcal{G}}(v_1) \in \mathcal{G}$, then $v_1 = \pi_{\mathcal{G}}(v_1)$ solve the problem and the optimal solution is equal to v_1 . Otherwise, based on Proposition 3 of [26], a smaller polyblock $\mathcal{P}_2 \subset \mathcal{P}_1$ which still contains $\mathcal{G} \cap \mathcal{H}$ but excludes v_1 can be constructed (See Fig.3(b)). This is achieved by constructing the vertex set \mathcal{T}_2 by replacing v_1 in \mathcal{T}_1 with N new vertices $\{v_1^1, \dots, v_1^N\}$ and then removing the improper vertices, where $v_1^j = v_1 - (v_1^j - \pi_{\mathcal{G}}(v_1))e_j$ ¹. This procedure is repeated until an optimal solution is found. This leads to a sequence of polyblocks containing $\mathcal{G} \cap \mathcal{H}$: $\mathcal{P}_1 \supset \mathcal{P}_2 \supset \dots \supset \mathcal{G} \cap \mathcal{H}$ (See Fig.3). The algorithm terminates at n th iteration when $\pi_{\mathcal{G}}(v_n) \in \mathcal{G}$. According to [26] the OPA algorithm is guaranteed to converge to the global optimal

¹In this paper, e_j denotes the j th unit vector of \mathbb{R}^N .

solution *only* if the optimal solution v^* has a positive lower bound $v_l > 0$. Therefore, when at least one of the involved user has $\gamma_{\min} = 0$ the origin has to be shifted to the negative orthant to guarantee the convergence of the OPA.

Complexity analysis: The complexity and the convergence speed of OPA depends on the number of iterations required to search the outer boundary of the feasible set. The accuracy requirement of the bisection method for finding the projection point of the optimal vertex on the boundary of the feasible set also affects the complexity of the OAP algorithm. Another complexity key factor is the number of the optimization variables (N_{var}), i.e., the dimensionality of the optimization problem. The latter highly affects the computational time of each iteration. Indeed, at the n th iteration, the optimal vertex over \mathcal{P}_n needs to be found from $(nN_{var} - n - N_{var} + 2)$ vertices. This is because the size of the vertex set \mathcal{T}_n increases by $(N_{var} - 1)$ after each iteration. Thus, the computational time of each iteration linearly increases with N_{var} . Let N_{iter} be the required number of iterations to find the optimal solution by using the OPA algorithm, the complexity for obtaining the solution is in the order of $O(N_{var}N_{iter}^2)$. The rapid growth of the size of the vertex set motivates us to develop a more efficient algorithm based on the existing OPA algorithm. For ease of notation, we denote the proposed algorithm by MARIO (Polyblock vertices triMming bAsed powerR allocatIOn).

C. MARIO ALGORITHM

In this section, we propose the MARIO algorithm to reduce the complexity of OPA. The key idea of MARIO is to delete the unnecessary vertices that mislead the PA algorithm. Like the grass trimmer which trim down the quack grass to improve the growth of a meadow, MARIO accelerates the convergence time of PA by trimming down the hurtful vertices which slow down the PA algorithm. To make the last concept formal, first, we introduce the following proposition.

Proposition 1. *In the first step of Dinkelbach's algorithm, the optimal power vector p_{ij}^* is achieved only if at least one user is transmitting with its maximum allowed power.*

Proof: For a scaling factor $\alpha > 1$ and a power allocation vector p_{ij} satisfies the powers constraints we have the following,

$$U_1^{ij}(\alpha p_{ij}) = \omega_i \log_2 \left(1 + \frac{P_c^i g_{cb}^i}{P_{d1}^j + h_{d1,b}^j + P_{d2}^j h_{d2,b}^j + \frac{\sigma_N^2}{\alpha}} \right) \quad (21)$$

$$+ \omega_{j1} \log_2 \left(1 + \frac{P_{d2}^j g_d^j}{P_c^i h_{i,d1}^j + \eta P_{d1}^j + \frac{\sigma_N^2}{\alpha}} \right)$$

$$+ \omega_{j2} \log_2 \left(1 + \frac{P_{d1}^j g_d^j}{P_c^i h_{i,d2}^j + \eta P_{d2}^j + \frac{\sigma_N^2}{\alpha}} \right) > U_1^{ij}(p_{ij})$$

The power constraints imply that the maximum of αP_{d1}^j , αP_{d2}^j is P_{\max}^d and the maximum of αP_c^i is P_{\max}^c . This means that the optima p_{ij}^* has at least one power bounded by the peak power constraint. ■

By virtue of Proposition 1, we know that the non-misleading vertices of a polyblock \mathcal{P} are the vertices which have at least one abscissa above or equal to the maximum allowed power. Needless to say, these vertices must also satisfy the QoS of the users defined in (7a)-(7c). Observe that the extra auxiliary variables t and s defined in (20) may lead to a vertex which does not meet the above requirements. Hence, the misleading vertex is defined as follows.

Definition 7. *A vertex $v = [v_1, v_2, \dots, v_N]$ is said to be misleading if it does not have any power equal to or larger than the maximum power or if it cannot meet the QoS of the users or both. A misleading vertex can be deleted without affecting the PA process.*

Fig.4 illustrates the above concept. In this figure x and y are the optimization variables, the lines l_1 and l_2 represent the QoS, and the shaded area denotes the feasible set. The red triangle points stand for the vertices that do not have any power at least bounded to the peak power. The blue square points are the vertices that can not meet the QoS of all users. The green diamond points are the vertices that cannot meet the QoS and do not have any user at least transmitting with the maximum allowed power. The black points represent the non-misleading vertices which can be kept to generate the new polyblock.

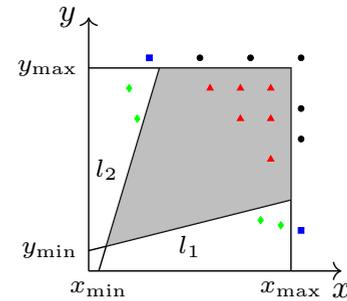


FIGURE 4. Illustration of the misleading vertices that can be generated when auxiliary variables have been added.

The above concept is valid for the WSR power allocation problem. To apply it to the GEE power allocation problem, the maximum transmission power of the users has to be updated after each Dinkelbalch's step as indicated in the following proposition.

Proposition 2. *Let $p_{k,ij}^*$ and p_{\max}^k be the optimal transmission power and the maximum transmission power at the k th iteration of Dinkelbach's algorithm. The maximum transmission power in the $k + 1$ th iteration is equal to the optimal transmission power in the k th iteration, i.e., $p_{\max}^{k+1} = p_{k,ij}^*$.*

Proof: Recall that maximizing a general fraction $\frac{f(x)}{g(x)}$ by means of Dinkelback's algorithm requires to maximize $F(\lambda_k) = \max_{x \in \mathcal{S}} f(x) - \lambda_k g(x)$ at the k th iteration, with \mathcal{S} being the set of constraints of the problem (in our case $\mathcal{S} = \mathcal{S} \cap \mathcal{S}_c$). Moreover, at each iteration k and for any $x \in \mathcal{S}$, $F(\lambda_k) \geq 0$. In addition, at the next iteration, the value of λ

is updated as follows $\lambda_{k+1} = \frac{f(\mathbf{p}_{k,ij}^*)}{g(\mathbf{p}_{k,ij}^*)}$. Hence, at the $(k + 1)$ th iteration we have $F(\lambda_{k+1}) = \max_x f(x) - \frac{f(\mathbf{p}_{k,ij}^*)}{g(\mathbf{p}_{k,ij}^*)}g(x)$, which in turn implies that the maximum power at the $k + 1$ th iteration must be equal to the optimal power of the k th iteration. ■

Now, since the utility function of the MO problem is monotonically increasing function, the optimal point at the k th iteration will have at least one abscissa bounded to the maximum constraint. Thus, Definition 7 can also be used to reduce the complexity of globally solving the power allocation problem. This concept is illustrated in Fig.5, in which $\Psi_k (k \in \{1, 2, 3\})$ denotes the feasible set of the optimization problem at the k th iteration and the magenta star point is the optimal value. As it can be seen, the feasible set is getting smaller at each iteration and thus by updating the maximum power in PA algorithm the convergence time of PA will be significantly decreased.

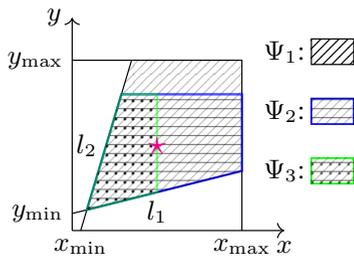


FIGURE 5. The update process of the maximum powers.

Combining the OPA algorithm with Definition 7, Proposition 2 and Algorithm 1, the proposed MARIO algorithm can be presented as in Algorithm 2.

Now, recall that one of the work's targets is to derive the global optimal RA scheme of an FD-D2D based cellular network (see the second column of Fig.1). Up to now, we have shown that the global RA scheme can be obtained by decomposing the original RA problem to two sub-problems as PA and CA and then globally solving each sub-problem individually. Then, we proposed MARIO to derive the global optimal PA scheme. Thus, to achieve the optimal RA scheme, it only remains to globally solve the CA sub-problem. The latter is the target of the following section.

V. CHANNEL ASSIGNMENT

So far, we find the global optimal power allocation for all the possible pairs $(D2D_j, UE_i)$. In this section, we aim at finding the optimal reuse candidate for each pair $D2D_j$, i.e., solving the channel assignment problems $P3_1$ and $P3_2$ defined in (10) and (13) respectively. Both problems $P3_1$ and $P3_2$ consist only of the channel reuse indicator variable ρ_{ij} . Moreover, according to the channel assignment constraints defined in (7d) and (7e), the D2D pairs can share at most one CU's resources and the resources of a CU can be allocated at most to one D2D pair. Therefore, the channel assignment problems $P3_1$ and $P3_2$ can be seen as a maximum weight bipartite matching problem. Fig.6 illustrates the maximum

Algorithm 2 The proposed MARIO algorithm

- 1: **Input:** The increasing utility function of $P2_{mo}$ defined in (20), the compact normal set \mathcal{S}_n and the closed conormal set \mathcal{S}_c such that $\mathcal{S}_n \cap \mathcal{S}_c \neq \emptyset$.
- 2: **Output:** The optimal solutions of $P2_1$ and $P2_2$.
- 3: Set the accuracy of Dinkelbach's algorithm $\epsilon_1 > 0$, the counter of Dinkelbach's algorithm $k = 0$, and $\lambda_0 = 0$.
- 4: **while** $F(\lambda_k) \geq \epsilon_1$ **do**
- 5: Solve (20) as shown below:
- 6: Initialization: Define an accuracy $\epsilon_2 > 0$, a counter $n = 0$, and an initial vertex $v_0 = [p_{\max}^k, s_{\max}, t_{\max}]$, with s_{\max} and t_{\max} being the maximum values of the introduced auxiliary variables in $P2_{mo}$. Let the initial polyblock \mathcal{P}_1 be box $[0, v_0]$ which contains $\mathcal{S}_n \cap \mathcal{S}_c$. The vertex set $\mathcal{T}_1 = \{v_0\}$. Denote by v_n^* the optimal vertex at the n th iteration.
- 7: **repeat**
- 8: $n = n + 1$.
- 9: Find the best vertex v_n^* in \mathcal{P}_n , i.e., $v_n^* = \arg \max \{f(v) | v \in \mathcal{T}_n\}$.
- 10: Obtain $\pi_{\mathcal{G}}(v_n^*)$, the projection of v_n^* on the upper boundary of \mathcal{G} .
- 11: **if** $\pi_{\mathcal{G}}(v_n^*) = v_n^*$, i.e., $v_n^* \in \mathcal{G}$ **then**
- 12: $v^* = v_n^*$, **break**.
- 13: **else**
- 14: Replace v_n^* with N new vertices $\{v_{11}, \dots, v_{1N}\}$.
- 15: Delete the improper and the misleading vertices.
- 16: **end if**
- 17: **until** $|f(v_n^*) - f(\pi_{\mathcal{G}}(v_n^*))| \leq \epsilon_2$
- 18: $v_n^* = v^* = [p_{k,ij}^*, s^*, t^*]$
- 19: $k = k + 1$
- 20: set $p_{\max}^k = p_{k-1,ij}^*$
- 21: update $\lambda_k = \frac{\omega_i R_i^k(p_{ij}^{k*}) + \omega_{j_1} R_{j_1}^{j_1}(p_{ij}^{k*}) + \omega_{j_2} R_{j_2}^{j_2}(p_{ij}^{k*})}{\mu(p_c^{i*} + p_{d_1}^{j_1*} + p_{d_2}^{j_2*}) + 3P_{cir}}$.
- 22: **end while**

weight bipartite problem in $P3_1$ and $P3_2$, where the set of D2D pairs \mathcal{D} , and the set of the cellular users \mathcal{C} are assumed to be the two disjoint groups of vertices in the bipartite graph. The pair $(D2D_j, UE_i)$ is joined by an edge ij when CU_i is a reuse candidate of $D2D_j$. The weight of edge ij is assumed to be the optimal objective function of $P3_l$, and it is denoted by U_l^{ij*} with $l \in \{1, 2\}$. The well known Khun-Munkres algorithm (Hungarian algorithm) [28] can be used to solve such assignment problems.

A. RESOURCE ALLOCATION IN HD-D2D NETWORK

As a complement, we introduce the resource allocation problem in an HD-D2D network. For a fair comparison with the FD-D2D network, we assume that when $D2D_j$ reuses the bandwidth of CU_i (B_i), the users D_1^j and D_2^j will operate on two equally orthogonal portions of the bandwidth B_i . Hence, when $D2D_j$ shares the resources of CU_i in HD mode, the SINR of the CU_i will remain the same as in (1). However, the SINRs of D_1^j and D_2^j do not longer contain the self

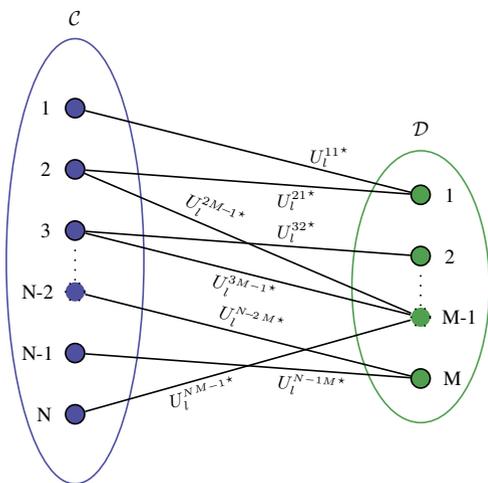


FIGURE 6. The bipartite graph representation of the channel assignment problem.

interference part. Consequently, the solution of the resource allocation problem of an HD-D2D based cellular network is similar to the FD-D2D network, with the only difference that in HD mode we have $\eta = 0$ and the bandwidth of D_1^j and D_2^j is equal to half of the CU_i bandwidth.

The optimal solution of both FD-D2D and HD-D2D resource allocation problem is reported in Algorithm 3.

Algorithm 3 Optimal Resource allocation algorithm

- 1: \mathcal{C} : The set of existing cellular users
- 2: \mathcal{D} : The set of D2D pairs
- 3: $U_i(\rho, P)$: The function to be maximized
- 4: **if** HD mode **then**
- 5: Set $\eta = 0$
- 6: Set the bandwidth of D_1^j and D_2^j to $B/2$
- 7: **end if**
- 8: Step 1: Power Control
- 9: **for all** $j \in \mathcal{D}$ and $i \in \mathcal{C}$ **do**
- 10: Solve $P2_i$ for the pair $(D2D_j, CU_i)$ by using MARIO algorithm.
- 11: **end for**
- 12: Step 2: Channel Assignment
- 13: Get the optimal CU_i for each $D2D_j$ by using Khun-Munkres algorithm

B. COMPLEXITY ANALYSIS

As shown earlier, solving the resource allocation problem requires sequentially solving the power allocation problem ($P2_i$) and the channel assignment ($P3_i$) problems, with $l \in \{1, 2\}$. Hence, the complexity of Algorithm 3 depends on the complexity of the power allocation problem and the complexity of the channel assignment problem. The complexity of finding the global optimal power allocation for one couple $(D2D_j, CU_i)$ has been derived in the previous section, and it is in the order of $\mathcal{O}(N_{iter}^2 N_{var})$. Since we have M D2D pairs and N CUs, the total number of possible

couples is MN , and thus the complexity of ($P2_i$) is in the order of $MNO(N_{iter}^2 N_{var})$. On the other hand, the Khun-Munkres algorithm [28] requires at most $\max^3(M, N)$ iteration to finish the assignment problem. Therefore, the overall complexity of the resource allocation problem can be described as $MNO(N_{iter}^2 N_{var}) + \mathcal{O}(\max^3\{M, N\})$. Due to the high complexity of this problem, in the following, we present a heuristic algorithm to efficiently solve the resource allocation problem. The proposed algorithm follows a reverse path, i.e., first, it assigns the channels to the different D2D pairs and then it allocates the power for the different couples $(D2D_j, CU_i)$. We denote this algorithm by Channel Assignment Then Power Allocation algorithm (CATPA) and it is described thereafter.

VI. THE PROPOSED CATPA ALGORITHM

From the above description, the global optimal solution of the resource allocation problem can be found by first deriving the optimal power allocation for all the possible couples $(D2D_j, CU_i)$ and then assigning the channels to the users by using the well known Hungarian algorithm. In the considered scenario, this will lead to solve $N \times M$ PA problems via the proposed MARIO algorithm and then apply the Hungarian algorithm, with $N \times M$ being the number of all the possible couplings between the M D2D pairs that coexist with the N CUs. The conventional RA sub-optimal solutions presented in [24] [25] have the same structure of the proposed MO based solution, and thus they also require solving $N \times M$ PA problems (refer to the first column of Fig.1). Hence, both the conventional sub-optimal solution and the MO based solution will be inapplicable for deriving the optimal RA scheme in a real FD-D2D network especially when N or M increases.

To reduce the complexity we develop in this section, the CATPA algorithm that first assigns the channels to the D2D pairs and then allocates the powers for all the involved users as shown in the third column of Fig.1. Hence, by using CATPA we will only need to solve M instead of $N \times M$ PA problems to achieve a solution for the RA problem.

First focusing on the channel assignment step, we propose to assign the available channels for the D2D pairs that are expressing good channel conditions. To that end, we define a profit for each possible couple $(D2D_j, CU_i)$ as shown below:

$$Profit^{ij} = \frac{g_{cb}^i + g_d^j + g_d^j}{I_i + I_{j1} + I_{j2}} \tag{22}$$

$$I_i = \gamma_{\min}^i (h_{d1,b}^j + h_{d2,b}^j + \sigma_N^2)$$

$$I_{j1} = \gamma_{\min}^{j1} (h_{i,d1}^j + \eta + \sigma_N^2), \text{ and } I_{j2} = \gamma_{\min}^{j2} (h_{i,d2}^j + \eta + \sigma_N^2).$$

The numerator of (22) can be seen as a measure of the possible gain when allocating the channel CH_i to $D2D_j$ while the denominator of (22) can be observed as the cost of reusing the bandwidth of CU_i . Hence, $Profit^{ij}$ is a metric that quantifies the profit of each couple $(D2D_j, CU_i)$. Since we are assuming that the BS has full-knowledge about the channel state information of the users, the $Profit^{ij}$ of all

the possible couples $(D2D_j, CU_i)$ can be easily computed at the BS. Having the profit table of all the possible couples in the network, we propose to allocate the channels to the couples with high profits. Moreover, to respect the channel allocation constraints defined in (7a)-(7c), we first assign the i th channel to the D2D pair that has the highest profit, and then we remove the assigned channel and the assigned D2D pair after each assignment. The following example clearly explains the proposed channel assignment method.

Assume a cellular network with 4 CUs and 4 D2D pairs with a profit table as shown in Fig.7(a). From this figure, it is clear that the $Profit^{1,3}$ is the highest value. Thus we assign the third channel to the first D2D pair and then we remove the first row and the third column. After this deletion, the profit table will reduce to the matrix shown in Fig.7(b). We repeat this process until all the D2D pairs get assigned channel (See Fig.7(a)-(d)). It is clear that by following this process each channel will be assigned to only one D2D pair, and each D2D pair will reuse the channel of only one CU.

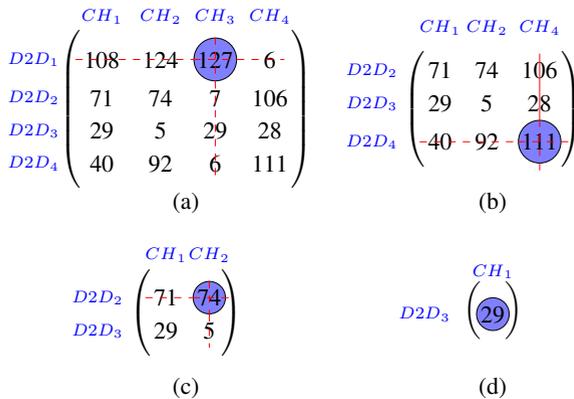


FIGURE 7. The process of the proposed CA.

The proposed algorithm requires at most $\min\{M, N\}$ iterations to finish the assignment problem, and thus it reduces the assignment complexity by $1/\min\{M, N\}$ comparing to the Hungarian algorithm. In addition, for that specific profit function, both the Hungarian algorithm and the proposed CA algorithm lead to the same value as shown in the next proposition.

Proposition 3. *For pairwise different profits, the proposed CA algorithm leads to the global optimal assignment solution.*

Proof: First recall that the Hungarian algorithm achieves the best assignment problem for a maximum-weight bipartite matching problem. In particular, The Hungarian algorithm assigns the channel to the users who give the higher WSR or GEE. Now, assume that all the weights of the assignment problem are different. Then, the best assignment will be achieved by finding the maximum weights that are locating in different rows and different columns. This is exactly the objective of our proposed CA algorithm. ■

By examining the profit equation defined in (22), we see that two profit values are identical only when the related D2D pairs and the CUs are locating at the same locations. Hence, in practice, the proposed CA algorithm leads to the optimal assignment solution as shown in the following Remark.

Remark 4: *In a real network, each node occupies a specific area in the cell, and thus two nodes cannot reside at the same point which in turns implies that all the profit values are different. Thus, the proposed CA algorithm will lead to an optimal assignment solution.*

Now, after assigning the channels for all the D2D pairs, the power allocation problem can be solved using the SCO method presented in [24], [25]. Algorithm 4 summarizes all the previous discussion and shows the different steps of CATPA.

Algorithm 4 The proposed CATPA algorithm

- 1: \mathcal{C} : The set of existing cellular users
- 2: \mathcal{D} : The set of D2D pairs
- 3: Step 1: Channel Assignment
- 4: **for all** $j \in \mathcal{D}$ and $i \in \mathcal{C}$ **do**
- 5: Compute $Profit_{i,j}$ by using (22).
- 6: **end for**
- 7: **for** $j = 1 : \min\{M, N\}$ **do**
- 8: Find the highest $Profit_{i,j}$ element.
- 9: Let I_{row}^j and I_{col}^j be the indices of the row and the column of the highest element at the j th iteration respectively.
- 10: Allow the I_{row}^j th D2D pair to reuse the channel of the I_{col}^j th CU.
- 11: Delete the I_{row}^j th row and the I_{col}^j th column of the profit table.
- 12: **end for**
- 13: Step 2: Power Control
- 14: Allocate the power for each $(D2D_j, CU_i)$ by using the SCO theory.

VII. NUMERICAL ASSESSMENT

We consider a single cell network, where CUs are uniformly distributed in the cell. The D2D users are usually within a short distance, and thus we distribute them in the cell according to the cluster distribution model [38]. In particular, the D2D users are uniformly distributed within a randomly located cluster. Moreover, in our simulation, we assume different D2D pairs are within different clusters. The CUs are assumed to share the total bandwidth equally. In addition, in our simulation, we only focused on the egalitarian solution, i.e., we set the weights ω_i , ω_{j_1} , and ω_{j_2} to one. Table 1 summarizes our simulation parameters.

A. MARIO VS OPA

The aim of this subsection is to compare the complexity of the proposed MARIO algorithm and the complexity of the traditional OPA algorithm. To that end, we consider a simple scenario in which a single D2D pair coexists with one CU.

TABLE 1. Simulation parameter

Cell radius (R)	0.5 Km
D2D cluster radius (r)	20, 30, 40, ..., 100 (m)
Uplink channel bandwidth	180 KHz
Noise power (σ_N^2)	-114 dBm
Path-loss exponent (α)	4
Maximum Tx power	24 dBm
Multiple-path fading	Exponential distribution with unit mean
Egaletarian Solution	$\omega_i = \omega_{j1} = \omega_{j2} = 1, \forall i \in \mathcal{C}, j \in \mathcal{D}$

The center of the D2D cluster is located at 300m from the BS, and the CU is located at 100m from the BS.

Giving this scenario, Fig.8 shows the behavior of both MARIO and OPA algorithms in the last iteration of Dinkelbach's algorithm, i.e., when λ_k equals the optimal GEE value. The y -axis of this figure is the difference between the GEE values of the best vertex point (v_n^*) and its projection point ($\pi_G(v_n^*)$), i.e., the line 17 in Algorithm 2. As can be seen in Fig.8, in both algorithms, the difference between the GEE values of v_n^* and $\pi_G(v_n^*)$ reaches zero after several numbers of iterations which in turn validates that both algorithms have already converged to the global maximum point. Fig.8 also shows the typical bounding behavior of the polyblock based algorithm [39]. For instance, both algorithms quickly find a feasible solution within a relatively small difference between the GEE values of v_n^* and $\pi_G(v_n^*)$, but many more iterations are required to squeeze the difference to zero. In addition, Fig.8 clearly shows the superiority of MARIO in term of the required number of iterations to achieve the optimal GEE value. For instance, the proposed MARIO algorithm achieved the optimal GEE value after 2.2×10^4 iterations while the regular OPA algorithm required five-times more iterations (11×10^4) to obtain the optimal GEE value. A similar behavior is obtained at the first iteration of Dinkelbach's algorithm, i.e., when maximizing WSR, but it is omitted for sake of clarity.

This is because MARIO deletes the misleading vertices that delay the convergence time. To illustrate this concept, we show in Fig. 9 the generated vertices of both algorithms after 400 iterations, in the first iteration of Dinkelbach's algorithm. As illustrated in Fig. 9(a)-(b), MARIO generates a much lower number of vertices compared to the regular OPA algorithm, and thus it has faster convergence time.

To emphasize the superiority of MARIO, we compare in Table 2 the complexity of both OPA and MARIO algorithms in terms of the average number of iterations and the average number of vertices to achieve the optimal GEE value for different cluster radii. As expected, MARIO requires less number of iterations and generate less number of vertices to attain the optimal GEE value.

B. THE OPTIMALITY GAP

As mentioned in the introduction, all the related RA works ([24], [25]) provide the solution of the RA problem by using the sequential convex optimization theory (SCO). The latter guarantees only a first-order optimal solution, and up to now

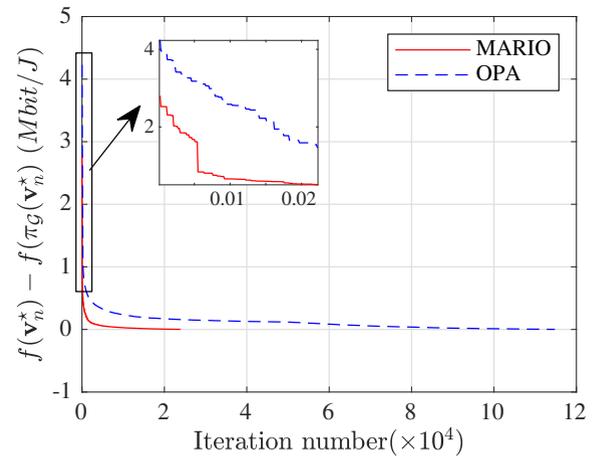
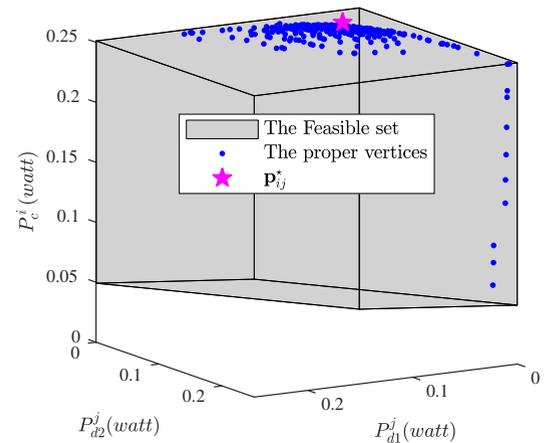
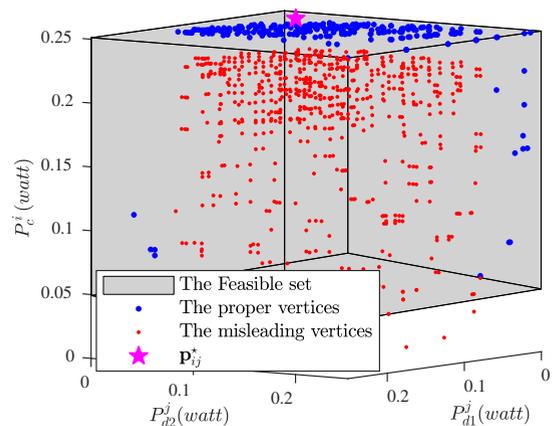


FIGURE 8. Convergence behavior of the proposed MARIO algorithm and the regular OPA algorithm, in the last iteration of Dinkelbach's algorithm. The behavior shows the convergence to the global GEE optimal solution and illustrates the superiority of the proposed MARIO algorithm.



(a) The proposed MARIO algorithm



(b) The regular OPA algorithm

FIGURE 9. The feasible set and the generated vertices of (a) the proposed MARIO algorithm and (b) the regular OPA algorithm after 400 iterations.

the optimality gap of the SCO method in the RA problem

TABLE 2. The average number of Iterations and Vertices for the proposed MARIO algorithm and the regular OPA algorithm.

r (m)	Number of Iterations		Number of vertices	
	MARIO	OPA	MARIO	OPA
10	80990	100000	388.0699	1253
20	74604	980000	413.4425	1258.9
40	77263	950000	400.2299	871.4
80	68395	100000	632.9584	800

of an FD-D2D network is not identified yet. Fig.10(a) and Fig.10(b) compare the achieved GEE and the achieved rate using the MO theory with that achieved by using the SCO for different self-interference cancellation factor (η). As can be seen, the optimality gap is less than 1%. This result is inline with the works in [26], [40]. Another remarkable observation is that both the achieved GEE and the achieved rate increase with the decline of the SI cancellation factor. This is expected because when η decreases the RSI power decreases and thus the total interference will also decrease.

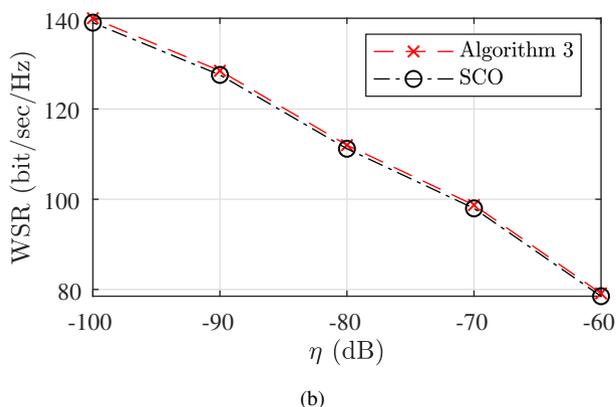
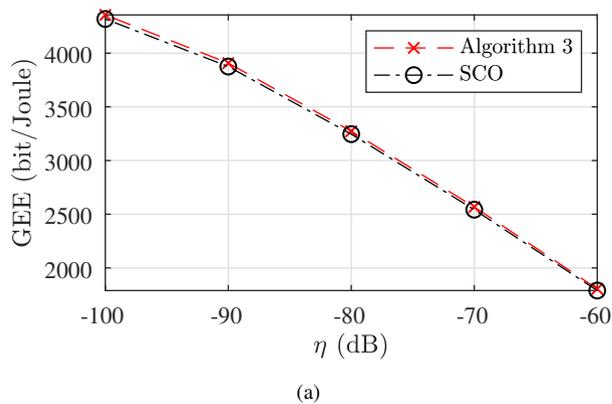


FIGURE 10. Achieved GEE versus SI cancellation factor, using 1) the proposed method, 2) SCO method ([24], [25]). ($r = 20m, \gamma_{\min}^i = \gamma_{\min}^j = \gamma_{\min}^k = 3dB$)

While in the above figure we have shown the optimality gap in terms of the achieved rate or the achieved GEE, in Table 3 we present the average number of iterations to achieve both the global optimal solution (MO based method) and the sub-optimal solution (SCO based method). As can be seen, the global optimal solution can be achieved at the price of a

high number of iterations. Hence, given the low optimality gap of the SCO method and the high complexity of the MO based method, the SCO method is an interesting candidate to analyze the FD-D2D network while the MO-based method is an interesting tool to benchmark the results. Thus, in the next sub-section, we will analyze the performance of our proposed sub-optimal RA algorithm, CAPTA, in the FD-D2D network as compared to the SCO optimization method applied in [24] and [25].

TABLE 3. The average number of Iterations for the proposed MARIO algorithm and SCO algorithm.

η (dB)	Number of Iterations	
	SCO	MARIO
-60	63.2	77378
-70	51.65	76628
-80	46.9	75828
-90	41.29	74721
-100	34.26	72934

C. THE PERFORMANCE OF THE SUB-OPTIMAL RESOURCE ALLOCATION CAPTA ALGORITHM

Here we aim to identify the performance of the proposed CAPTA algorithm. To that end, we assume a single cell network where $M = 4$ D2D pairs coexist with $N = 4$ CUs. The CUs are uniformly distributed in the cell, and the D2D users are uniformly distributed within a randomly located cluster. To cover all the possible location, we generate the D2D users and the CUs 500 times, and then we averaged the results. Fig. 11(a) and 11(b) respectively compare the achieved GEE and WSR values of the FD-D2D network with the achieved GEE and WSR values of the HD-D2D network. The GEE and the WSR values of the FD-D2D network are obtained through the proposed CAPTA algorithm and the SCO method adopted in [24], [25].

Fig. 11 shows that the GEE and WSR values of both FD-D2D and HD-D2D networks decline as the cluster radii increases. This is expected, since when r increases the proximity distance between the D2D users also increases, and thus more power is needed to maintain the same QOS. However, increasing the power also increases the interference, and thus the total rate and the GEE decrease. In addition, Fig. 11 also shows that the GEE and the WSR values of the FD-D2D network decline with the increasing of η which in turn re-validate the result of the previous sub-section. Hence, the FD-D2D network's performance highly depends on both the SI cancellation capability and the proximity distance of the D2D users. For instance, the HD-D2D network outperforms the FD-D2D network when $\eta = -60dB$ and the D2D users are within a $10m$ distance from each other. Thus, to achieve the maximum gain of a D2D based cellular network, the transmission mode must always alternate between HD and FD based on the channel situation, the SIC factor, and the proximity distance.

In addition, Fig. 11(a)-(b) clearly indicate that the proposed CAPTA algorithm can achieve GEE and WSR values that are

within the 91% – 98% of the optimal value obtained by SCO method which in turn validates the accuracy of the proposed CATPA algorithm. Now, to shed the light on the low complexity feature of the proposed CATPA algorithm, we present in Table 4 the required number of iterations to obtain the optimal GEE value for both CATPA and the SCO algorithm for different SIC cancellation factors and different cluster radii values. As expected, the proposed CATPA solution decreases the number of iterations by at least $\min\{M, N\}$. This is because CATPA does not compute the optimal power allocation for all the possible couples $(D2D_j, CU_i)$ as in the global optimal solution. For instance, the adopted scenario in

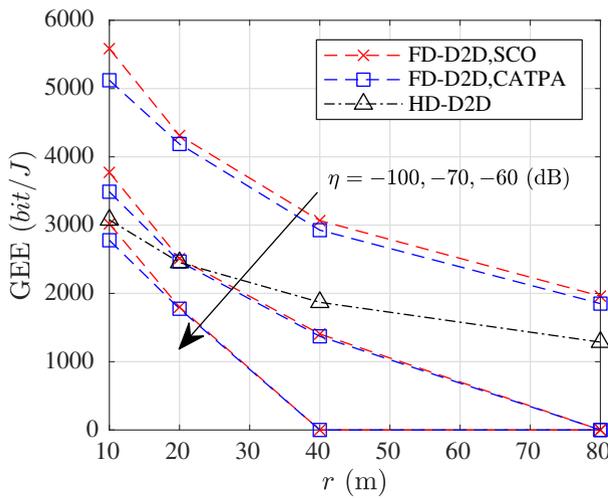
TABLE 4. The number of Iterations for the proposed CATPA algorithm and SCO algorithm.

η (dB)	$r = 10$ (m)		$r = 20$ (m)		$r = 40$ (m)	
	CATPA	SCO	CATPA	SCO	CATPA	SCO
-60	108.6	425.9	121.64	481.8	52	216
-70	95.63	362.7	116.43	444.6	119.6	472.9
-80	74.38	279.19	96.39	360.3	106.9	410.6
-90	53.05	203.8	68.58	257.9	78.6	302.1
-100	38.48	153.06	46.03	182.7	55.1	220.8

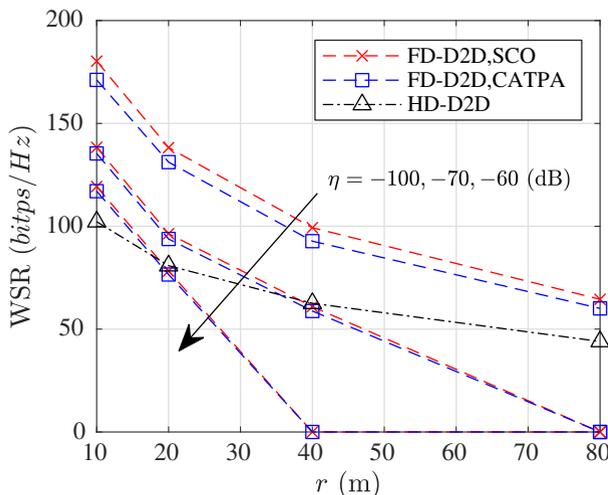
D. PERFORMANCE ANALYSIS OF THE FD-D2D NETWORK

In the previous subsections, we saw the effect of the RSI power and the D2D cluster radius on the performance of the FD-D2D network. Moreover, we observed that the FD-D2D network requires a small D2D proximity distance and high SI cancellation capability to achieve its ultimate gain. Hence, in this subsection, we aim to analyze the effect of the remaining network’s parameters, such as the maximum transmission power and the required QoS, on the performance of the FD-D2D network by assuming $\eta = -100$ dB and $r = 20$ m.

Fig. 12 presents the effect of the maximum allowed power on the achieved GEE value considering three power allocation strategies: 1) the obtained powers when maximizing GEE; 2) the obtained optimal powers when maximizing WSR; 3) the maximum allowed powers. In this figure, P_{\max}^d denotes the maximum power of D_1^j and D_2^j with $j \in \mathcal{D}$. From Fig. 12(a) it is seen that the GEE obtained by GEE maximizing first increases by the increasing of p_{\max} and then it saturates at large p_{\max} . This is because when p_{\max} is large enough to allow achieving the optimal GEE value, the excess power is no longer utilized. However, the GEE attained by WSR maximizing first increased by the increase of p_{\max} until $p_{\max} = -30$ dBw, and then it started to decline for larger p_{\max} . This indicates that the global optimal GEE value can be achieved if at least one user is transmitting at -30 dBw. This result also confirms our proposed MARIO algorithm in which we update the maximum power level at each step on Dinkelbach’s algorithm. In addition, transmitting with full power at all users highly degrades the GEE performance. Now, looking at Fig. 12(b), it can be observed that the achieved WSR when maximizing WSR increases as the maximum power increases. This result confirms our observation that the maximum achieved rate can be obtained only if at least one of the users is transmitting with the maximum allowed power. However, the achieved WSR when maximizing GEE saturate at large p_{\max} which is inline with the results of Fig.12(a). In addition, Fig. 12 also confirms that the full power transmission strategy is not suitable for the FD-D2D network since it failed to achieve good WSR or GEE values. Another interesting observation can be extracted from Fig.12 is that the difference between the achieved WSR when maximizing GEE or maximizing WSR is relatively smaller than the difference between the obtained GEE when maximizing GEE or maximizing WSR. Giving this result, in a resource allocation problem, maximizing the GEE is much



(a)



(b)

FIGURE 11. Comparison between the performance of FD-D2D network and the HD-D2D network when applying the proposed CATPA algorithm or the SCO solution adopted in [24], [25] for (a) maximizing GEE, and (b) maximizing WSR. ($\gamma_{\min}^i = \gamma_{\min}^{j1} = \gamma_{\min}^{j2} = 3$ dB)

Table 4 contains 4 D2D pairs and 4 CUs, and thus CATPA will lead to at least a 4-fold decrease in the number of iterations as shown in Table 4.

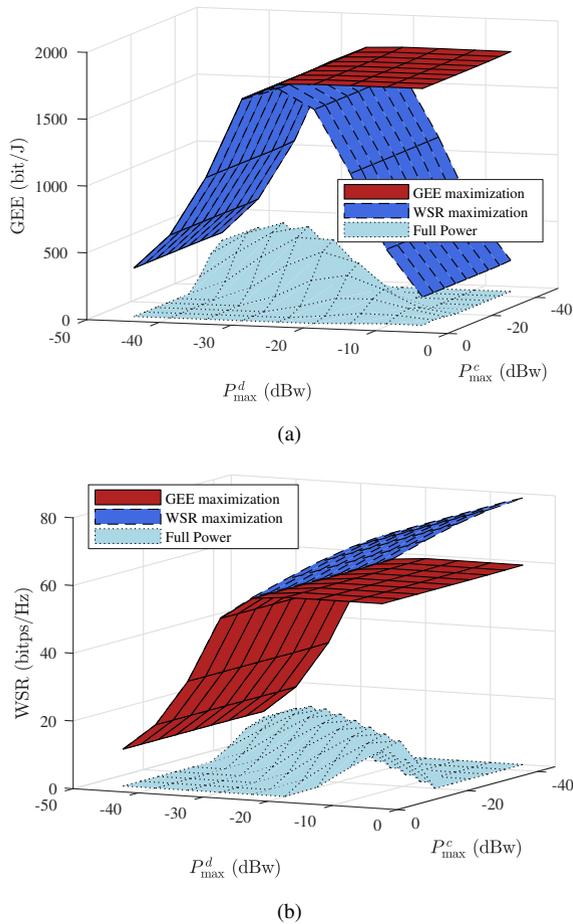


FIGURE 12. The effect of the maximum transmission power on the (a) achieved GEE value and (b) achieved WSR value considering three power allocation strategies: 1) the obtained powers when maximizing GEE; 2) the obtained optimal powers when maximizing WSR; 3) the maximum allowed powers. ($r = 20m, \gamma_{\min}^i = \gamma_{\min}^{j1} = \gamma_{\min}^{j2} = 3dB$)

attractive than maximizing the overall rate.

Next, to see the effect of the QOS on the performance of FD-D2D network, we set the maximum transmit power $p_{max} = [-6, -6, -6](dBw)$, and then we draw in Fig. 13(a)-(b) the variation of the achieved GEE and WSR w.r.t γ_{\min}^c and γ_{\min}^d with γ_{\min}^d being equal to γ_{\min}^{d1} and γ_{\min}^{d2} . Since the maximum power allocation strategy is not suitable for FD-D2D network (see Fig.12-(b)), in this figure we only consider the remaining two power allocation schemes: 1) the obtained powers when maximizing GEE; 2) the achieved powers when maximizing WSR. As expected the achieved GEE and WSR decline with the increase of the minimum required SINR for all power allocation strategies since at large SINR requirement less number of FD-D2D pairs can be admitted. In addition, Fig. 13 affirm the previous results obtained in Fig. 12 which says that in a FD-D2D network maximizing GEE will be more attractive than maximizing the overall rate. This is because the gap between the achieved GEE when maximizing GEE and maximizing WSR (see Fig. 13(a)) is much larger than the gap between the achieved WSR when maximizing GEE and maximizing WSR (see

Fig. 13(b)). To have full knowledge of the performance of an FD-D2D network when maximizing GEE or maximizing WSR a joint optimization framework must be developed which is left for future work.

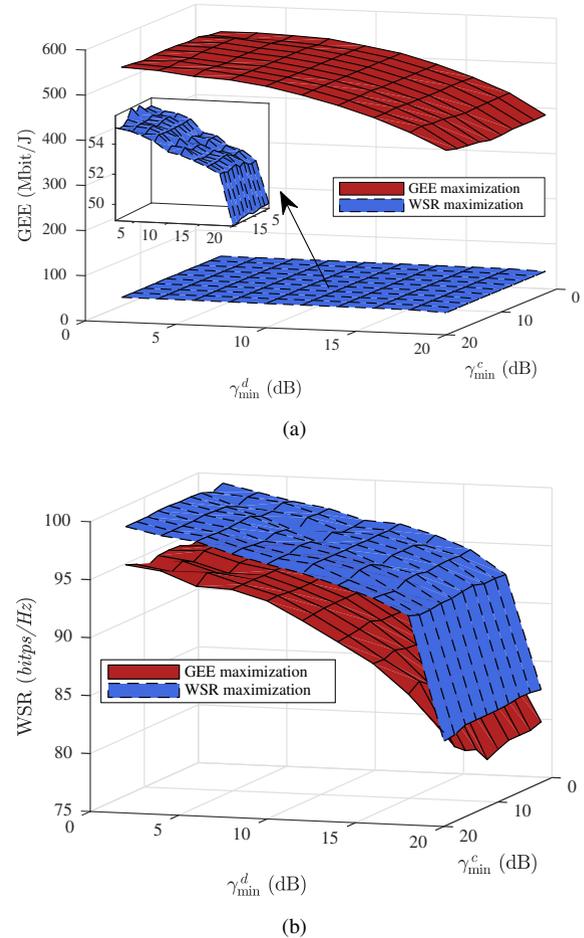


FIGURE 13. The effect of the QoS requirement (a) achieved GEE value and (b) achieved WSR value considering two power allocation strategies: 1) the obtained powers when maximizing GEE; 2) the obtained optimal powers when maximizing WSR. ($r = 10m, p_{max} = [-6, -6, -6](dBw)$)

VIII. CONCLUSION

In this paper, we have investigated the resource allocation problem for a full duplex D2D communication underlying cellular network. To maximize the overall throughput and the energy efficiency while guaranteeing the data rate requirement of both D2D users and CUS, first we formulate the resource allocation optimization problem, and then we find the global optimal solution through two steps: power control for all the possible couples ($D2D_j, CU_i$); maximum weight matching to obtain the optimal cellular user for each D2D pair. By means of monotonic optimization theory, we proposed a new polyblock-based algorithm denoted as MARIO that can efficiently find the global optimal power control for the involved users. The Khun-Munkers algorithm is used to solve the matching problem. Due to the high complexity of the global optimal solution, we proposed the CATPA

algorithm that can find an efficient sub-optimal solution by first assigning the channels and then controlling the users' powers. Simulation results verified the proposed algorithms and showed that the performance of the FD-D2D network highly depends on the self-interference capability and the proximity distance. In future work, we will investigate the effect of the imperfect channel state information on the optimal resource allocation. Meanwhile, the scenario when an FD-D2D pair can share the resources of multiple CUs, and when multiple FD-D2D pairs can reuse the same resource should be also studied.

ACKNOWLEDGMENT

This work was partially funded through French ANR Agency projects ACCENT5 with grant agreement code: ANR-14-CE28-0026-02. A special thanks also for IRACON and CNRS GdR-ISIS for their mobility grants. Besides, Hussein Chour thanks Mr. Bho Matthiesen for helping him in understanding and developing the code of the Outer Polyblock Algorithm.

REFERENCES

- [1] V. Cisco Mobile, "Cisco visual networking index: Global mobile data traffic forecast update, 2016–2021 white paper," 2017.
- [2] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, May 2014.
- [3] X. Zhang, W. Cheng, and H. Zhang, "Full-duplex transmission in PHY and MAC layers for 5G mobile wireless networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 112–121, Oct. 2015.
- [4] S. Mumtaz and J. Rodriguez, *Smart device to smart device communication*. Springer, 2014.
- [5] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, vol. 47, no. 12, 2009.
- [6] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlying cellular networks," *IEEE Transactions on Communications*, vol. 61, no. 8, pp. 3541–3551, Aug. 2013.
- [7] H. Zhang, L. Song, and Z. Han, "Radio resource allocation for device-to-device underlay communication using hypergraph theory," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4852–4861, Jul. 2016.
- [8] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. K. Wong, and X. Chen, "Joint spectrum and power allocation for D2D communications underlying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2182–2195, Apr. 2016.
- [9] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, 2014.
- [10] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap rf canceller for in-band full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4321–4334, 2016.
- [11] 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 Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.
- [12] V. Tapio, "System scenarios and technical requirements for full-duplex concept," *DÜPLO, EU-Project Deliverable*, 2013.
- [13] K. S. Ali, H. ElSawy, and M. Alouini, "Modeling cellular networks with full-duplex D2D communication: A stochastic geometry approach," *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4409–4424, Oct. 2016.
- [14] S. Kim and W. Stark, "Full duplex device to device communication in cellular networks," in *Proc. Networking and Communications (ICNC) 2014 Int. Conf. Computing*, Feb. 2014, pp. 721–725.
- [15] W. Cheng, X. Zhang, and H. Zhang, "Optimal power allocation for full-duplex D2D communications over wireless cellular networks," in *2014 IEEE Global Communications Conference*, Dec 2014, pp. 4764–4769.
- [16] —, "Heterogeneous statistical QoS provisioning for full-duplex D2D communications over 5G wireless networks," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.
- [17] H. V. Vu, N. H. Tran, and T. Le-Ngoc, "On coverage probabilities and sum-rate of full-duplex device-to-device cellular networks," in *Proc. IEEE Int. Conf. Communications (ICC)*, May 2018, pp. 1–6.
- [18] S. Ali, N. Rajatheva, and M. Latva-aho, "Full duplex device-to-device communication in cellular networks," in *Proc. European Conf. Networks and Communications (EuCNC)*, Jun. 2014, pp. 1–5.
- [19] L. Wang, F. Tian, T. Svensson, D. Feng, M. Song, and S. Li, "Exploiting full duplex for device-to-device communications in heterogeneous networks," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 146–152, May 2015.
- [20] K. T. Hemachandra, N. Rajatheva, and M. Latva-aho, "Sum-rate analysis for full-duplex underlay device-to-device networks," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2014, pp. 514–519.
- [21] H. Chour, F. Bader, Y. Nasser, and O. Bazzi, "Full-duplex or half-duplex D2D mode? closed form expression of the optimal power allocation," in *25th International Conference on Telecommunication (ICT)*, 2018.
- [22] B. Zuo, L. Jiang, C. He, and Z. Lian, "Power allocation optimization for full-duplex D2D communications underlying cellular networks," in *2016 International Conference on Networking and Network Applications (NaNA)*, 2016, pp. 103–108.
- [23] J. Ding, L. Jiang, and C. He, "Energy-efficient power control for underlying full-duplex device-to-device communications," in *Proc. Asia Modelling Symp. (AMS)*, Dec. 2017, pp. 155–160.
- [24] S. Li, Q. Ni, Y. Sun, and G. Min, "Resource allocation for weighted sum-rate maximization in multi-user full-duplex device-to-device communications: Approaches for perfect and statistical csis," *IEEE Access*, vol. 5, pp. 27 229–27 241, 2017.
- [25] R. Tang, J. Zhao, H. Qu, and Z. Zhang, "Energy-efficient resource allocation for 5G full-duplex enabled device-to-device communication," in *Proc. IEEE Globecom Workshops*, Dec. 2016, pp. 1–7.
- [26] Y. J. A. Zhang, L. Qian, J. Huang et al., "Monotonic optimization in communication and networking systems," *Foundations and Trends® in Networking*, vol. 7, no. 1, pp. 1–75, 2013.
- [27] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.
- [28] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, pp. 83–97, 1955.
- [29] M. Wellens, J. Wu, and P. Mahonen, "Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio," in *Proc. 2nd Int. Conf. Cognitive Radio Oriented Wireless Networks and Communication (CROWNCOM)*, Aug. 2007, pp. 420–427.
- [30] C. Liu, T. Lv, W. Tang, H. Gao, and Y. Lu, "On the performance of mode selection for hybrid-duplex D2D communications," in *Proc. MILCOM 2017 - 2017 IEEE Military Communications Conf. (MILCOM)*, Oct. 2017, pp. 852–857.
- [31] L. Han, W. Zou, and G. Zhao, "Duplex mode selection for device-to-device communications underlying the cellular uplink," in *Proc. IEEE Conf. Computer Communications Workshops (INFOCOM WKSHPS)*, May 2017, pp. 157–162.
- [32] A. N. Kadhim, F. Hajiaghajani, and M. Rasti, "On selecting duplex-mode and resource allocation strategy in full duplex D2D communication," in *Proc. Iranian Conf. Electrical Engineering (ICEE)*, May 2017, pp. 1640–1645.
- [33] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5983–5993, 2011.
- [34] S. Buzzi, C. I. T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 697–709, Apr. 2016.
- [35] Z. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [36] J.-P. Crouzeix and J. A. Ferland, "Algorithms for generalized fractional programming," *Mathematical Programming*, vol. 52, no. 1-3, pp. 191–207, 1991.

- [37] A. Zappone, E. Jorswieck et al., "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends® in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.
- [38] B. Kaufman and B. Aazhang, "Cellular networks with an overlaid device to device network," in 2008 42nd Asilomar Conference on Signals, Systems and Computers (ACSSC), Oct. 2008, pp. 1537–1541.
- [39] E. Björnson and E. Jorswieck, *Optimal Resource Allocation in Coordinated Multi-Cell Systems*. Foundations and Trends® in Communications and Information Theory, 2013. [Online]. Available: <https://ieeexplore.ieee.org/https://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=8187586>
- [40] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.



HUSSEIN CHOUR (STM'18) was born in 1990 in Tyr, Lebanon. He received his B.E. degree in electrical and communications engineering from the Iran University of Science and Technology, Tehran, Iran, in 2013 and the M.E. degree in electrical and computer engineering from the American University of Beirut, Beirut, Lebanon, in 2016. He is currently pursuing his Ph.D. degree in communication engineering with a focus on the resource allocation techniques for a full duplex

D2D communication network. Hussein's research interests include Resource Allocation techniques, Optimization, wireless networking, device-to-device communication, Full-Duplex, cellular technologies, and VANET. He has been involved in the ACCENT5 National Project. He received two mobility grants from IRACON and CNRS GdR-ISIS for cooperating with TU Dresden- Communication Theory Lab in 2018.



EDUARD A. JORSWIECK (S'01-M'03-SM'08) received the Diplom-Ingenieur (M.Sc.) and Doktor-Ingenieur (Ph.D.) degrees from the Technische Universität Berlin, Berlin, Germany, in 2000 and 2004, respectively. Since 2008, he has been the Head of the Chair of Communications Theory and a Full Professor with the Technische Universität Dresden, Dresden, Germany. He is a Principal Investigator with the excellence cluster center for Tactile Internet with Human-in-the-

Loop and PI within the DFG collaborative research center HAEC. His main research interests include signal processing for communications and networks, applied information theory, and communications theory. He has authored more than 100 journal papers, 11 book chapters, 260 conference papers, and 3 monographs on these research topics. He was a corecipient of the IEEE Signal Processing Society Best Paper Award in 2006 and coauthored papers that won the Best Paper or Best Student Paper Awards at IEEE WPMC 2002, Chinacom 2010, IEEE CAMSAP 2011, IEEE SPAWC 2012, IEEE WCSP 2012, and IEEE ICUFN 2018. He was a member of the IEEE SPCM Technical Committee (2008-2013) and has been a member of the IEEE SAM Technical Committee since 2015. Since 2017, he serves as an Editor-in-Chief for the Springer EURASIP Journal on Wireless Communications and Networking. Since 2011, he has been an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. From 2008 to 2011, he has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, and from 2011 to 2013, as a Senior Associate Editor. Since 2013, he has served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.



FAOUZI BADER (SM'07) received the Ph.D. degree (Hons.) in telecommunications from the Universidad Politécnica de Madrid, Madrid, Spain, in 2002. He joined the Center Technologic de Telecomunicacions de Catalunya-CTTC, Barcelona, Spain, as an Associate Researcher in 2002, where he was nominated as a Senior Research Associate in 2006. Since 2013, he has been an Associate Professor with CentraleSupélec, Rennes, France. Since 2018 he is the head of the signal and communications department of the IETR lab in Rennes-France. His research

activities mainly focus on advanced multicarrier waveforms (OFDM(A), (non-) uniform multimode filter-based multicarrier schemes) and frequency allocation techniques in relay cognitive environments. He has been involved in several European projects from the 5th-7th EC research frameworks, and from 2012 to 2013, he was nominated as the General Coordinator and the Manager of the EC-funded research project ICT EMPhAtiC, focusing on enhanced multicarrier techniques for professional ad-hoc and cell-based communications. He has authored more than 150 papers in peer-reviewed journals and international conferences, more than 13 book chapters, and published four books. He served as Technical Program Committee Member in major IEEE ComSoc and VTS conferences, and as the General Chair of the eleventh edition of the ISWCS'2014 conference, and the Co-Chair of the ISWCS 2015 edition.



YOUSSEF NASSER obtained his Master and Ph.D. degrees in Signal Processing and Communications from the National Polytechnic Institute of Grenoble, France in 2003 and 2006 respectively and his 'Habilitation à Diriger la Recherche (HDR)' diploma, the highest education degree a scholar can achieve, in 2015. He is currently with the Department of Electrical and Computer Engineering at the American University of Beirut. From 2006 to 2010, he worked as a Senior Research Engineer with the Institute of Electronics and Telecommunications of Rennes (IETR), France. From 2003 to 2006, he worked with the Laboratory of Electronics and Information Technologies in Grenoble (Laboratoire d'Electronique et de Technologies de l'Information-LETI) as a R&D Engineer. Since 2007, Y. Nasser is involved in the standardization of European broadcasting systems including DVB-T2, DVB-T2 Lite and DVB-NHG. He coordinated tasks and/or packages of many European projects. Y. Nasser has published more than 100 papers in international peer-reviewed journals and conferences. Since 2011 He is a Senior Member of the IEEE.

From 2006 to 2010, he worked as a Senior Research Engineer with the Institute of Electronics and Telecommunications of Rennes (IETR), France. From 2003 to 2006, he worked with the Laboratory of Electronics and Information Technologies in Grenoble (Laboratoire d'Electronique et de Technologies de l'Information-LETI) as a R&D Engineer. Since 2007, Y. Nasser is involved in the standardization of European broadcasting systems including DVB-T2, DVB-T2 Lite and DVB-NHG. He coordinated tasks and/or packages of many European projects. Y. Nasser has published more than 100 papers in international peer-reviewed journals and conferences. Since 2011 He is a Senior Member of the IEEE.



OUSSAMA BAZZI received the Bachelor of Engineering degree in electrical engineering from the American University of Beirut in 1987, and the master's and Ph.D. degrees in electronics from the University of Valenciennes and Hainaut Cambresis, France, in 1988 and 1992, respectively. He joined Lebanese University in 1996, where he is currently a Full Professor. He has been the Head of the Research Group, Telecoms, Signal Processing and Images with Lebanese University since 2009.

His current research projects include radio resource management, broadcast techniques, cooperative communication techniques, and cognitive radio networks. His research interests are in the areas of signal processing and wireless and mobile radio communications.

...