Distributed Channel Allocation for D2D-Enabled 5G Networks Using Potential Games

DEMIA DELLA PENDA [1], **ANDREA ABRARDO** [2,3], **MARCO MORETTI** [3,4], **(Member, IEEE), AND MIKAEL JOHANSSON** [5]

[1]Ericsson AB, 16480 Kista, Sweden
[2]Department of Information Engineering, University of Siena, 53100 Siena, Italy
[3]Consorzio Nazionale Interuniversitario per le Telecomunicazioni, 43124 Parma, Italy
[4]Department of Information Engineering, University of Pisa, 56122 Pisa, Italy
[5]Department of Automatic Control, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 10044 Stockholm, Sweden

Corresponding author: Demia Della Penda (demiadp@kth.se)

ABSTRACT Frequency channel allocation is a key technique for improving the performance of cellular networks. In this paper, we address the channel allocation problem for a 5G multi-cell system. We consider a heterogeneous network in which cellular users, micro-cell users, and device-to-device (D2D) communications coexist within the radio footprint of the macro cell. We maximize the aggregate transmission rate, exploiting channel diversity and managing both the inter-cell interference, typical of cellular networks and the intra-cell interference generated by the nonorthogonal transmissions of the small-cell and D2D users. By modeling the allocation problem as a potential game, whose Nash equilibria correspond to the local optima of the objective function, we propose a new decentralized solution. The convergence of our scheme is enforced by using a better response dynamic based on a message passing approach. The simulation results assess the validity of the proposed scheme in terms of convergence time and achievable rate under different settings.

INDEX TERMS Distributed allocation, 5G system, optimization, OFDMA, device-to-device, potential games, message passing.

I. INTRODUCTION

Device-to-device (D2D) communication is a promising technology for enhancing the performance of next generation (5G) cellular networks. The basic idea of D2D communication is to allow mobile devices in close proximity to communicate directly, bypassing the base station (BS). D2D technology leads to several advantages compared to the conventional cellular communications, such as offloading the BS [1], improving cell coverage, throughput and transmission latency [2], and enhancing both the spectral and the energy efficiency of the system [3].

Orthogonal Frequency Division Multiple Access (OFDMA) is the current LTE technology and the main candidate for 5G cellular networks [4]. Thus, OFDMA systems where several D2D pairs can share the uplink resources with multiple cellular users is a scenario of great interest.[1] Due to the cost and scarcity of licensed spectrum, a popular choice is to implement D2D communication in *underlay in-band* mode: the D2D users reuse the frequency resources assigned to the traditional cellular users. This may generate new and more critical interference situations compared to current LTE systems, in which transmissions within the same cell are orthogonal. The orthogonality constraint, in fact, limit the distance between the interfering nodes to be greater than the cell radius. Differently, in underlay D2D-enabled networks, interfering users can operate at any distance, potentially jeopardizing the communications performance.

In dense heterogeneous 5G networks, D2D communications can be deployed as an enhanced solution to face the intense demand for high data rates and quality of service. In such scenarios, if micro cells also reuse the macro cell spectrum, interference management between all different types of transmissions involved (i.e., D2D, small-cell and macro-cell transmissions) definitely becomes one of the most critical challenge. Therefore, the design of appropriate resource allocation methods is of crucial importance.

A. PRIOR AND RELATED WORKS

The resource allocation problem for D2D-enabled networks has attracted a lot of attention in the research community [6], [7]. Considering the large computational

---

[1]The choice of sharing uplink is mainly motivated by the asymmetric traffic loads in the two directions, and by the capacity of the BS to better handle the interference, compared with the mobile devices [5].

complexity of allocating resources in a D2D-enabled cellular setting, several works limit their theoretical analysis to a single-cell scenario, assuming that an advanced inter-cell interference mitigation scheme works on top of the per-cell allocation algorithms [8]–[13]. Other works incorporate the inter-cell interference into the noise level at the receivers [14]. As of today, D2D resource sharing in the more realistic multi-cell scenario has received less coverage in the literature and most of the existing works focus on suboptimal solutions. For instance, Belleschi *et al.* [15] propose a centralized resource allocation scheme that minimizes the total power consumption. However, because the solution is computationally demanding and not implementable in a real system, they also provide a suboptimal approach based on per-cell allocation. Gu textitet al. [16], rather than trying to solve the optimal allocation problem, consider a heuristic proportional fair scheduling, with at most one cellular and one D2D communication per channel. In [17], Feng *et al.* establish a suboptimal tractable model for multi-cell D2D underlaid cellular networks, and adopt exclusion regions around the BSs to mitigate cochannel interference. A robust transmission design in the case of imperfect channel state information (CSI) for D2D-enabled cellular network is studied in [18]. Here, the BSs are assumed to have multiple antennas while mobile user are equipped with single-antenna transceivers. The original optimization problem is decoupled into two subproblems, each of which is solved sequentially.

A different line of work spans from a game theoretic approach to the problem of D2D resource allocation. For this application, game theory proves to be a powerful tool. Indeed, although most games do not guarantee optimality, they are amenable for distributed implementation, which is a desirable feature for such complex problems and for dense networks. A game model for spectrum sharing has been proposed in [19], where D2D pairs reusing the spectrum of the same operator are regarded as coalitions, and the subchannel allocation is performed according to a matching algorithm. To simplify the analysis, the authors limit the channel reuse to at most two transmissions per channel (either two D2D pairs or one D2D and one cellular user). Recently, the coalition formation approach has been combined with Bayesian reinforcement learning in [20], to design a distributed scheme in which the D2D pairs make autonomous choices only on the basis of locally-observable information. In [21], coalition formation has been employed in cloud heterogenous networks to solve the problem of assigning subchannels of different bandwidth to multiple D2D pairs and remote radio head users. The work in [22] studies a game theoretic approach for resource allocation in a 5G network where multi-cell D2D communications interfere with small cells and macro cells. D2D links utilize common resources of multiple cells and each player's transmission parameter is unknown to the other players. The problem is formulated as an utility maximization where the utility is a function of the bandwidths allocated to D2D users. In this scenario, the effect of inter-cell interference is not considered and the channel is assumed flat over all the allocated bandwidth; that is, the subchannels allocation problem is not addressed.

An example of joint power control and subchannel allocation can be found in [23], for a scenario where D2D and small cells coexist with just *one* macro BS. Being the optimal solution too complex for practical implementation, the authors reformulate the problem as an exact potential game solved by a distributed learning algorithm capable to find a good Nash equilibrium. At most one subchannel is assigned to each user and the channel is assumed flat over all the allocated bandwidth. At the same time, only one D2D user and one small cellular user can be active at the same time on the same subchannel. This assumption allows to simplify the interference management but, at the same time, it risks to strongly limit the resource reuse in densely populated networks.

As highlighted in [22] and [23], the coexistence of D2D pairs, small cells, and macro cells is going to be a typical choice for 5G networks, where macro cells alone will not be able to provide the high rates and reliable connectivity required by future services. Note that in densely deployed 5G scenarios, small cells and D2D pairs will behave similarly in reusing the macro cell spectrum to provide service to short-distance communications. Therefore, there are many parallelisms in the allocation process of these two types of user.

B. CHALLENGES AND CONTRIBUTIONS

In future 5G networks, the number of heterogeneous transmitters is foreseen to grow massively, leading to allocation problems with prohibitive complexity for centralized schedulers. Furthermore, when the allocation is centralized at the BS, the signaling overhead caused by CSI feedback from the devices to the BS will scale linearly with the number of D2D users and small cells. For these reasons, one of the main challenges for 5G networks is to provide distributed schemes capable to handle the growing number of devices and interference. Moreover, in emerging applications like Intelligent Transportation System services [24], it is desirable that communications are reliable also outside the cellular network coverage, calling for solutions where the devices take autonomous decisions about the resources to use.

In this work, we address the problem of channel allocation for 5G heterogeneous networks, where macro cells coexist on the same spectrum with multiple D2D pairs and several small cells. Because of the interference and the integer nature of the allocation variable, the considered optimization problem is nonconvex and hard to solve to optimality. However, extending our previous work [25], we use the concept of potential games [26] to develop a tractable solution based on better response dynamic. This distributed game-theoretic solution is particularly suitable to be implemented with a *max-sum message passing* (MP) approach [27], which is definitely a novelty in this context. MP allocation is based on the distributed solution of simple local optimization problems, whose results are exchanged iteratively through nodes until convergence is reached [28], [29]. In a different

scenario, [30] proposes a MP strategy for user association and resource blanking in heterogeneous networks, because of its efficiency and distributed nature. Nevertheless, there are very few papers where MP schemes are employed in D2D-enabled networks: [31] and [32] are focused on cooperative aspects of D2D communications but in simplified single-cell scenarios.

By combining game theory and MP, we manage to implement an extremely flexible and low complexity solution. As a matter of fact, in our approach the players of the game can be indifferently macro cells, small cells, or D2D pairs. Moreover, unlike most of the existing works in the literature, we do not pose any limit to the number of cellular users (either macro or micro) and D2D terminals that are active at the same time on the same channel, not even in the number of interfering cells and resources to be allocated to each user. Additionally, for such a heterogeneous environment where different sources of interference coexist, we propose a resource allocation problem that converges to a local optimum of the system throughput in a distributed manner.

We summarize our main contributions as follows:

- We address the problem of channel allocation in an OFDMA 5G-like heterogeneous network, where multiple D2D pairs and small cells reuse the cellular spectrum. Unlike most existing literature, in formulating our problem we do not limit the number of users, nor the number of cells in the network. Moreover, the same cellular resource can be used by multiple D2D pairs, as well as by users of the multiple small cells within the macro cell.
- Leveraging the theory of potential games, we solve the nonconvex resource allocation problem by means of an iterative algorithm based on better response dynamics, in which the players' strategies are updated via distributed and low-complexity MP approach. Additionally, we design a customized branch-and-bound (B&B) solver to find the optimal solution that we use as benchmark.
- To reduce the run-time of the solution, as well as the overall signaling overhead, we investigate the performance of the MP scheme with a reduced number of message exchanges. We also provide guidelines for its practical implementation, showing that our proposed scheme requires limited complexity and a limited amount of exchanged information.

C. PAPER ORGANIZATION

The paper is organized as follows. Sect. II introduces the system model and the allocation problem we study. In Sect. III we formulate the resource allocation problem as a potential game. Sect. IV and Sect. V study how to solve the problem employing best response and better response dynamics, respectively. In Sect. VII we design a customized solver based on a B&B approach, with the objective of having a performance benchmark for the algorithm we propose. Sect. VIII presents the numerical results, while conclusions are drawn in Sect. IX.
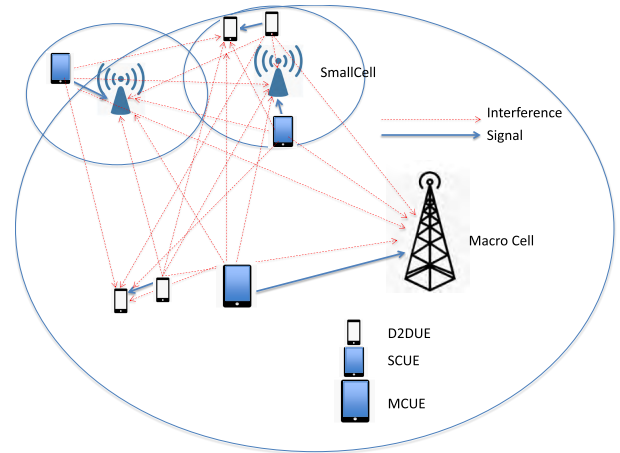
FIGURE 1. Illustrative scenario. Two D2D pairs, a macro cell user, and two micro cell users assigned to the same resource, and thus interfering with each other.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We focus on a 5G heterogeneous cellular network, where densely deployed small cells and D2D pairs coexist within the radio footprint of a macro cell. In this scenario, which is illustrated in Fig. 1, we consider *uplink* transmissions of an OFDMA system, where both D2D and small-cell communications are enabled to reuse the macro-cell radio spectrum.

For the sake of generalization, in our analysis we refer to the different BSs without distinguishing between macro or micro cells. We assume that the network is populated by a set $\mathcal{B}$ of BSs, each serving users within its own cell area. The receiver of each D2D pair can be seen as *virtual base station* serving a single uplink user and we denote with $\mathcal{D}$ the set of all D2D receivers. With the notation $\mathcal{K} = \mathcal{B} \cup \mathcal{D}$ we indicate the set of all BS receivers in the system. For each receiver $k \in \mathcal{K}$, we let $\mathcal{C}_k$ be the set of uplink users served by k. Thus, $\mathcal{C}_k$ is a singleton if $k \in \mathcal{D}$, while it represents multiple users if $k \in \mathcal{B}$.

The system has a set $\mathcal{F}$ of F orthogonal time-frequency physical resource blocks (RBs), and each BS is responsible for assigning those RBs to its users. Channel conditions vary across both RBs and users. As customary, uplink communications towards the same BS are assigned to orthogonal RBs, meaning that there is no intra-cell interference within a macro or a small cell. However, frequency reuse among different cells give rise to inter-cell interference, and uplink transmissions in the small cells and underlay D2D communications add both intra-cell and inter-cell interference, as shown in Fig. 1.

Let $G_{n,k}^f$ denote the channel gain between transmitter n and receiver k on RB f, and let P_n^f be the transmission power used by transmitter n on RB f. The transmit power allocation follows the approach used in current LTE systems. In particular, the transmission power levels of the mobile users are computed with the Uplink Open-Loop with Fractional Path Loss Compensation (OFPC) scheme (as described in [33]), and updated on a slower time-scale than the RBs.

We indicate with $\mathbf{X}_k$ the RB assignment to all users transmitting to receiver $k \in \mathcal{K}$. Specifically, $\mathbf{X}_k$ is a matrix of dimensions $|\mathcal{C}_k| \times F$, whose nth row represents the user $n \in \mathcal{C}_k$ and the fth column represents the RB $f \in \mathcal{F}$. We label each element of $\mathbf{X}_k$ with x_n^f, which is 1 if user n is assigned to RB f, and 0 otherwise. We let $\mathbf{X}_{-k}$ be the set of $|\mathcal{K}| - 1$ allocation decisions taken by all receivers except the kth. Moreover, for every $k \in \mathcal{K}$, we indicate with $\mathbf{X} = (\mathbf{X}_k, \mathbf{X}_{-k})$ the overall system RB allocation.

A Shannon-like capacity is considered as a measure of the instantaneous achievable rate. Specifically, the normalized rate in bps/Hz that receiver k can achieve on RB f from transmitter n is

$$R_n^f(\mathbf{X}_{-k}) = \log_2 \left(1 + \frac{P_n^f G_{n,k}^f}{\sigma^2 + I_k^f(\mathbf{X}_{-k})} \right), \qquad (1)$$

where $I_k^f(\mathbf{X}_{-k}) = \sum_{q \in \mathcal{K} \setminus \{k\}} \sum_{m \in \mathcal{C}_q} P_m^f G_{m,k}^f x_m^f$ is the interference perceived at receiver k on RB f, and σ^2 is the noise power, assumed equal for all RBs. Because of the orthogonality constraint within each cell, the interference $I_k^f(\mathbf{X}_{-k})$ depends on the overall allocation except the one in the cell $k \in \mathcal{K}$, to which user n belongs.

We consider the problem of allocating RBs to all users in the network. The goal is to maximize the total rate by exploiting frequency diversity and properly managing the interference. To avoid situations in which users with favorable channel conditions take most of the available RBs, we limit the number of RBs that can be assigned to each user. Accordingly, the overall RB allocation problem can be formally stated as the following integer programming problem:

$$\max_{\{x_n^f\}} \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{C}_k} \sum_{f \in \mathcal{F}} \omega_k R_n^f(\mathbf{X}_{-k}) x_n^f \qquad (2)$$

$$\text{subject to } \sum_{n \in \mathcal{C}_k} x_n^f \leq 1, \quad \forall f \in \mathcal{F}, \ \forall k \in \mathcal{K}, \qquad (2.a)$$

$$\sum_{f \in \mathcal{F}} x_n^f \leq F_n, \quad \forall n \in \mathcal{C}_k, \ \forall k \in \mathcal{K}, \qquad (2.b)$$

$$x_n^f \in \{0, 1\}, \quad \forall f \in \mathcal{F}, \ \forall n \in \mathcal{C}_k, \ \forall k \in \mathcal{K}. \qquad (2.c)$$

Here, $\{\omega\}_{k \in \mathcal{B}}$ is a set of no-negative real numbers that can be tuned to differentiate the performance of the different types of users in the network. For example, by setting ω_k large, we prioritize the resource allocation for the users in cell k; that is, the algorithm will aim at maximizing the rate of cell k more than the rate of the other cells and D2D pairs. Constraints (2.a) are the orthogonality constraints within the cell k, they are active only for $k \in \mathcal{B}$, because if $k \in \mathcal{D}$ it holds $|\mathcal{C}_k| = 1$ and (2.a) is always verified. Finally, (2.b) limit the number of RBs that can be assigned to each user.

III. RESOURCE ALLOCATION BASED ON GAME THEORY
The non-linear integer programming problem in (2) is in general difficult to solve efficiently as the number of devices increases. In particular, when considering highly dense D2D-enabled networks, the running time for state-of-the

art nonlinear integer programming solvers quickly becomes impractical. Moreover, resorting to a centralized scheduler to decide the overall allocation requires a large signaling overhead to collect the CSI of all involved links. To overcome the above concerns, we leverage the theory of potential game to achieve distributed and computational efficient, although not necessarily globally optimal, solutions.

For completeness, in this section we first give some basic definitions and fundamental properties of potential games [26], [34]. Then, we present our strategic-game formulation for the multi-cell D2D resource allocation problem.

A. PRELIMINARIES ON POTENTIAL GAMES
The interaction among autonomous decision-makers can be modeled as a strategic game, in which each player chooses a strategy independently from the choices of the other players. A strategic game can be formally described by a triplet $\mathcal{G} = \langle \mathcal{K}, \{\mathcal{X}_k\}, \{U_k\} \rangle$, where $\mathcal{K}$ is the set of players, $\mathcal{X}_k$ is the set of all possible strategies for player k, and U_k is player k's utility function. As customary in the game-theoretic literature, we indicate with $\mathbf{X}_k$ a specific strategy of player k, and we write $U_k(\mathbf{X}_k, \mathbf{X}_{-k})$ to stress that the utility of player k depends both on her own strategy and on the strategies of all other players. Because any change of strategy of one of the players affects all other players, the game becomes a dynamic process where players iteratively update their own strategies as a reaction to the changes in the strategy of the other players. In this work, we will always refer to *pure strategy*. We now recall some useful definitions.

Definition 1 (Best- and Better-Response Dynamic): The best-response dynamic is a strategy update rule where each player selects the strategy that maximizes her utility, assuming that the other players do not change their current strategies. Specifically, given a strategy profile $\mathbf{X} = (\mathbf{X}_k, \mathbf{X}_{-k})$, *player k chooses its new strategy* $\mathbf{Y}_k \in \mathcal{X}_k$ *such that*

$$\mathbf{Y}_k \in \{\tilde{\mathbf{X}}_k \in \mathcal{X}_k : U_k(\tilde{\mathbf{X}}_k, \mathbf{X}_{-k}) \geq U_k(\mathbf{X}_k, \mathbf{X}_{-k}), \ \forall \mathbf{X}_k \in \mathcal{X}_k\}. \qquad (3)$$

In the less demanding better-response dynamic, *the strategy update of player k is defined by replacing condition* (3) *by*

$$\mathbf{Y}_k \in \{\tilde{\mathbf{X}}_k \in \mathcal{X}_k : U_k(\tilde{\mathbf{X}}_k, \mathbf{X}_{-k}) \geq U_k(\mathbf{X}_k, \mathbf{X}_{-k})\}. \qquad (4)$$

Thus, under a better response dynamic, the new strategy is only guaranteed to be better than the previous one, but it might not be the best among all possible strategies.

Definition 2 (Potential Game): A strategic game $\mathcal{G} = \langle \mathcal{K}, \{\mathcal{X}_k\}, \{U_k\} \rangle$ *is an (exact) potential game if there exists a function* $\Phi : \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \mathcal{X}_{|\mathcal{K}|} \rightarrow \mathcal{R}$ *such that for any* $k \in \mathcal{K}$

$$U_k(\mathbf{X}_k, \mathbf{X}_{-k}) - U_k(\mathbf{X}_k', \mathbf{X}_{-k}) = \Phi(\mathbf{X}_k, \mathbf{X}_{-k}) - \Phi(\mathbf{X}_k', \mathbf{X}_{-k}), \qquad (5)$$

where $\mathbf{X}_k$ *and* $\mathbf{X}_k'$ *are two strategies of player k, and* Φ *is called the potential function of* $\mathcal{G}$.

Definition 3 (Nash Equilibrium (NE)): Given a strategic game $\mathcal{G} = \langle \mathcal{K}, \{\mathcal{X}_k\}, \{U_k\}\rangle$, the K-tuple $(\mathbf{X}_1^\star, \mathbf{X}_2^\star, \cdots, \mathbf{X}_{|\mathcal{K}|}^\star) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \mathcal{X}_{|\mathcal{K}|}$ is a NE if $U_k(\mathbf{X}_k^\star, \mathbf{X}_{-k}^\star) \geq U_k(\mathbf{X}_k, \mathbf{X}_{-k}^\star)$, for all strategies $\mathbf{X}_k \neq \mathbf{X}_k^\star$, for all $k = 1, \cdots, |\mathcal{K}|$. Put in words: at the NE no player has an incentive to change her strategy.

Potential games possess the important property that the set of (pure-strategy) NE points correspond to the local optima of the potential function [26], [34]. Moreover, for finite potential games,[2] whose potential function is bounded from above, the iterative processes based on best or better response dynamics converge to the NE set of the game, see [35, Th. 19].

In the sequel, we will use these attractive properties to design a convergent noncooperative game that leads to the RB allocation of the overall network.

B. POTENTIAL GAMES FOR RESOURCE ALLOCATION

To solve Problem (2) in a distributed manner, we propose a strategic game between all receivers. The game is described by $\mathcal{G} = \langle \mathcal{K}, \{\mathcal{X}_k\}, \{U_k\}\rangle$, where $\mathcal{X}_k$ is the set of all possible RB-allocation decisions for transmissions to receiver $k \in \mathcal{K}$, while U_k is the utility function of the kth player, given by the sum of all the achievable rates in the system, that is:

$$U_k(\mathbf{X}_k, \mathbf{X}_{-k}) = \sum_{f \in \mathcal{F}} \sum_{q \in \mathcal{K}} \sum_{n \in \mathcal{C}_q} \omega_q R_n^f(\mathbf{X}_{-q}) x_n^f. \tag{6}$$

The function U_k in (6) is a scalar function corresponding to the objective in (2). Despite the fact that each player k aims at maximizing U_k with respect to her own strategy only (i.e., $\mathbf{X}_k$), all players' utilities in (6) are chosen to be the same, making game $\mathcal{G}$ an *identical interest game*. The choice of this utility function implies that all users share the same goal, which is the maximization of the weighted sum rate of the system. By defining the function $\Phi(\mathbf{X}) \triangleq U_k(\mathbf{X}_k, \mathbf{X}_{-k})$, for all $k \in \mathcal{K}$, the next result follows immediately from Definition 2:

Proposition 1: $\mathcal{G} = \langle \mathcal{K}, \{\mathbf{X}_k\}, \{U_k\}\rangle$ is an exact potential game with potential function $\Phi(\mathbf{X})$.

Thus, in this game the potential function coincides with the weighted sum rate of the system and each user selects its own strategy aiming at maximizing the potential function.

Because $\mathcal{G}$ is an exact potential game, the best and better response dynamics converge to one of its NE, which, in our case, corresponds to a local optimum of (2). Thus, we can design an iterative algorithm based on a receiver-alternating allocation, meaning that the allocation is performed by rotation among all cells and D2D pairs. In particular, given any initial resource allocation, the players (receivers) take turn in choosing their best/better response strategy until no player is willing to change her decision; that is, until when a NE is achieved. The iterative game among the receivers is summarized in Algorithm 1.

---

[2]A *finite game* is a game with finitely many players, each of which has a finite set of strategies.

Algorithm 1 Iterative Game

1 Compute $P_n^f, \forall n \in \cup_k \mathcal{C}_k, \forall f \in \mathcal{F}$ with the LTE OFPC algorithm;
2 Select an order of playing $\pi(\mathcal{K})$; and select an initial allocation $\mathbf{X}_k^0, \forall k \in \mathcal{K}$;
3 Set $\mathbf{X}^0 \leftarrow (\mathbf{X}_1^0, \cdots, \mathbf{X}_{|\mathcal{K}|}^0)$ and compute $\Phi(\mathbf{X}^0)$;
4 $i \leftarrow 1, \Delta \leftarrow 1$;
5 **while** $\Delta \neq 0$ **do**
6   **for** $k \in \pi(\mathcal{K})$ **do**
7   Update $\mathbf{X}_k^i$ following the best/better-response dynamic;
8   $\mathbf{X}^i \leftarrow (\mathbf{X}_1^i, \cdots, \mathbf{X}_{|\mathcal{K}|}^i)$;
9   $\Delta \leftarrow \Phi(\mathbf{X}^{i-1}) - \Phi(\mathbf{X}^i)$;
10   $i \leftarrow i + 1$;

IV. RESOURCE ALLOCATION BASED ON BEST RESPONSE

In the previous section, we have seen that by considering the allocation problem as a potential game, we need to iteratively solve local (i.e., per-player) allocation problems, rather than the overall network problem. We now turn our attention on deriving the solution of each such local problem, that is, determining the strategy update of each player $k \in \mathcal{K}$.

We consider a reference overall allocation strategy $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_{-k})$, representing the overall network RB allocation at the last iteration of the game. To compute her new strategy $\mathbf{X}_k$, player k has to solve the following optimization problem

$$\max_{\{x_n^f\}} U_k(\mathbf{X}_k, \bar{\mathbf{X}}_{-k}), \tag{7}$$

$$\text{subject to} \sum_{n \in \mathcal{C}_k} x_n^f \leq 1, \quad \forall f \in \mathcal{F} \tag{7.a}$$

$$\sum_{f \in \mathcal{F}} x_n^f \leq F_n, \quad \forall n \in \mathcal{C}_k, \tag{7.b}$$

$$x_n^f \in \{0, 1\}, \quad \forall n \in \mathcal{C}_k, \forall f \in \mathcal{F}. \tag{7.c}$$

The objective function in Problem (7) is given in (6) and it is the same as in Problem (2). However, here it is maximized only with respect to the RB allocation of player k (represented by $\mathbf{X}_k$). We can rewrite the objective function in (7) explicitly as

$$U_k(\mathbf{X}_k, \bar{\mathbf{X}}_{-k})$$

$$= \sum_{f \in \mathcal{F}} \left[\omega_k \sum_{n \in \mathcal{C}_k} \log_2 \left(1 + \frac{P_n^f G_{n,k}^f}{\sigma^2 + I_k^f(\bar{\mathbf{X}}_{-k})}\right) x_n^f + \sum_{\substack{q \in \mathcal{K}, \\ q \neq k}} \omega_q \sum_{m \in \mathcal{C}_q} \right.$$

$$\left. \times \log_2 \left(1 + \frac{P_m^f G_{m,q}^f}{\sigma^2 + I_q^f(\bar{\mathbf{X}}_{-\{q,k\}}) + \sum_{n \in \mathcal{C}_k} P_n^f G_{n,q}^f x_n^f}\right) \bar{x}_m^f \right], \tag{8}$$

where $\bar{x}_m^f$ are the entries of the given reference allocation strategy $\bar{\mathbf{X}}$, and $I_q^f(\bar{\mathbf{X}}_{-\{q,k\}}) = \sum_{c \in \mathcal{K} \setminus \{q,k\}} \sum_{l \in \mathcal{C}_c} P_l^f G_{lq}^f \bar{x}_l^f$ is

the interference level at the qth receiver that does not include the effect from the use of RB f in cell k, which is instead given by the term $\sum_{n \in \mathcal{C}_k} P_n^f G_{nq}^f x_n^f$.

Due to the orthogonality constraints (7.a), the inner sum of the second term in (8) contains only a single nonzero element, corresponding to the only user transmitting to receiver q on RB f. We label this user with $\hat{m}(q,f)$ and we rewrite (8) as

$$
\begin{aligned}
&U_k(\mathbf{X}_k, \bar{\mathbf{X}}_{-k}) \\
&= \sum_{f \in \mathcal{F}} \omega_k \Bigg[\sum_{n \in \mathcal{C}_k} \log_2 \left(1 + \frac{P_n^f G_{nk}^f}{\sigma^2 + I_k^f(\bar{\mathbf{X}}_{-k})} \right) x_n^f + \sum_{\substack{q \in \mathcal{K}, \\ q \neq k}} \omega_q \\
&\quad \times \log_2 \left(1 + \frac{P_{\hat{m}(q,f)}^f G_{\hat{m}(q,f),q}^f}{\sigma^2 + I_q^f(\bar{\mathbf{X}}_{-\{q,k\}}) + \sum_{n \in \mathcal{C}_k} P_n^f G_{nq}^f x_n^f} \right) \Bigg].
\end{aligned} \tag{9}
$$

Using the best response dynamic, we have divided the original RB allocation problem in subproblems with reduced input size. However, the formulation in (7) is still nonconvex in the integer variables $\{x_n^f\}$, for $n \in \mathcal{C}_k$ and $f \in \mathcal{F}$. We tackle this issue in the next section.

V. RESOURCE ALLOCATION BASED ON BETTER RESPONSE

In the previous section, we showed how to distribute the RB allocation problem among the players by using the best response dynamic. The goal of this section is twofold: i) to overcome the nonconcavity of the objective function in (9); ii) to design a low-complexity and distributed scheme for updating each player's strategy.

First, we observe that by relaxing the integer constraints in (7), the objective function becomes convex. However, the maximization of convex functions (with or without additional constraints) is typically difficult and computationally tractable only in certain special cases [36]. Usually, we have to accept local optima and resort to methods based on repeated linearization of the objective function. Second, because the constraint matrix of (7) is totally unimodular [37], even if we drop the integer constraints we obtain a 0−1 solution vector. We therefore linearize (9) by computing the first-order Taylor approximation of the rate of the users in each cell $q \neq k$, around the point corresponding to the reference allocation $\bar{\mathbf{X}}$.

Let us consider the $|\mathcal{C}_k| \times F$ allocation matrix $\tilde{\mathbf{X}}_k$, with entries $\tilde{x}_n^f \in [0, 1]$, for each user $n \in \mathcal{C}_k$ and RB $f \in \mathcal{F}$. We define the following new utility for the kth player, dependent on both her new strategy $\tilde{\mathbf{X}}_k$ and the overall previous allocation $\bar{\mathbf{X}}$

$$
\begin{aligned}
&\tilde{U}_k(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}}) \\
&= \sum_{f \in \mathcal{F}} \omega_k \Bigg[\sum_{n \in \mathcal{C}_k} R_n^f(\bar{\mathbf{X}}_{-k}) \tilde{x}_n^f + \sum_{\substack{q \in \mathcal{K}, \\ q \neq k}} \omega_q
\end{aligned}
$$

$$
\times \left(R_{\hat{m}(q,f)}^f(\bar{\mathbf{X}}_{-q}) + \sum_{n \in \mathcal{C}_k} \frac{\partial R_{\hat{m}(q,f)}^f(\bar{\mathbf{X}}_{-q})}{\partial \tilde{x}_n^f} \left(\tilde{x}_n^f - \bar{x}_n^f \right) \right) \Bigg]. \tag{10}
$$

Here, each term

$$
\begin{aligned}
&\frac{\partial R_{\hat{m}(q,f)}^f(\bar{\mathbf{X}}_{-q})}{\partial \tilde{x}_n^f} \\
&= -\frac{P_{\hat{m}(q,f)}^f G_{\hat{m}(q,f),q}^f P_n^f G_{n,q}^f}{(\ln 2)(P_{\hat{m}}^f G_{\hat{m}q}^f + I_q^f(\bar{\mathbf{X}}_{-q}) + \sigma^2)(I_q^f(\bar{\mathbf{X}}_{-q}) + \sigma^2)}
\end{aligned} \tag{11}
$$

represents the sensitivity of user $\hat{m}(q,f)$ to the interference variations on RB f, caused by the new allocation $\tilde{\mathbf{X}}_k$ of player k. Because the rate function of each user $\hat{m}(q,f)$ in (9) is convex with respect to the allocation $\tilde{\mathbf{X}}_k$, the linear approximation in (10) is a lower bound of the actual rate. Moreover, this bound is tight under the reasonable assumption that the cumulative interference experienced at each receiver is much higher than the interference contribution from each single transmitter. In other words, in a network with multiple interfering links, changing the resource allocation of only one transmitter produces, in general, a small variation in the total interference measured at the receiver of the other links. Thus, when considering interference levels within small intervals around $I_q^f(\bar{\mathbf{X}}_{-q})$, the linear approximation of the rate of each user $\hat{m}(q,f)$ is close to the actual rate function. We therefore conclude that $\tilde{U}_k(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}})$ is a lower bound of $U_k(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}}_{-k})$ and, in general, the difference between the two values is small.

Neglecting the terms in (10) that do not depend on the optimization variables $\tilde{x}_n^f$, and defining

$$
\begin{aligned}
&\delta_{\hat{m}(q,f)}^f(\bar{\mathbf{X}}_{-q}) \\
&\triangleq -\frac{P_{\hat{m}(q,f)}^f G_{\hat{m}(q,f),q}^f}{(\ln 2)(P_{\hat{m}}^f G_{\hat{m}q}^f + I_q^f(\bar{\mathbf{X}}_{-q}) + \sigma^2)(I_q^f(\bar{\mathbf{X}}_{-q}) + \sigma^2)},
\end{aligned} \tag{12}
$$

we obtain the function

$$
\begin{aligned}
&\tilde{U}_k'(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}}) \\
&= \sum_{f \in \mathcal{F}} \sum_{n \in \mathcal{C}_k} \Bigg[\omega_k R_n^f(\bar{\mathbf{X}}_{-k}) + \sum_{\substack{q \in \mathcal{K}, \\ q \neq k}} \omega_q \delta_{\hat{m}(q,f)}^f(\bar{\mathbf{X}}_{-q}) P_n^f G_{n,q}^f \Bigg] \tilde{x}_n^f \\
&= \sum_{f \in \mathcal{F}} \sum_{n \in \mathcal{C}_k} \tilde{E}_n^f(\bar{\mathbf{X}}) \tilde{x}_n^f, \tag{13}
\end{aligned}
$$

in which each weight $\tilde{E}_n^f(\bar{\mathbf{X}})$ depends on the previous reference allocation $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_{-k})$ for all $k \in \mathcal{K}$, and not on the new allocation $\tilde{\mathbf{X}}_k$ of player k. Considering $\tilde{U}_k'(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}})$ as the new utility function that player k aims at maximizing, we reformulate the per-player optimization problem in (7) as

$$
\max_{\{\tilde{x}_n^f\}} \tilde{U}_k'(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}}) \tag{14}
$$

$$
\text{subject to} \sum_{n \in \mathcal{C}_k} \tilde{x}_n^f \leq 1, \quad \forall f \in \mathcal{F}, \tag{14.a}
$$

$$\sum_{f=1}^{F} \tilde{x}_n^f \leq F_n, \quad \forall n \in \mathcal{C}_k, \qquad (14.\text{b})$$

$$\tilde{x}_n^f \in [0, 1], \quad \forall f \in \mathcal{F}, \forall n \in \mathcal{C}_k. \qquad (14.\text{c})$$

It is worth mentioning that the computation of $\tilde{U}_k'(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}})$ relies on the collection from each other player $q \neq k$ of the F terms given in (12), one for each RB. These quantities are interference-dependent and they can be measured at each player $q \neq k$ (being she a receiver) and sent to player k on a control channel.

With the formulation (14) we have achieved our first goal, namely we have overcome the nonconvexity of the best response dynamic. We now show that although we are considering an utility function different from the one in the original problem, the solution to (14) represents a better response for the kth player. This means that we can still guarantee convergence to a NE of game $\mathcal{G}$.

Proposition 2: Given any allocation profile $\bar{\mathbf{X}}$, and indicating with $\tilde{\mathbf{X}}_k^\star \in \mathcal{X}_k$ the solution to Problem (14), it holds that $U_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}_{-k}) \geq U_k(\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_{-k})$. That is, $\tilde{\mathbf{X}}_k^\star$ is a better response of player k for the potential game $\mathcal{G} = \langle \mathcal{K}, \{\mathcal{X}_k\}, \{U_k\} \rangle$.

Proof: Solution $\tilde{\mathbf{X}}_k^\star$ is integer and belongs to the set $\mathcal{X}_k$ because of the total unimodularity property of the problem constraint in (14). Being $\tilde{\mathbf{X}}_k^\star$ the maximizer of $\tilde{U}_k'(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}})$, it also maximizes the function $\tilde{U}_k(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}})$ in (10). Therefore, $\tilde{U}_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}) \geq \tilde{U}_k(\mathbf{X}_k, \bar{\mathbf{X}})$ for all $\mathbf{X}_k \in \mathcal{X}_k$, and in particular $\tilde{U}_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}) \geq \tilde{U}_k(\bar{\mathbf{X}}_k, \bar{\mathbf{X}})$. Because $\tilde{U}_k(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}})$ is a lower bound of $U_k(\tilde{\mathbf{X}}_k, \bar{\mathbf{X}}_{-k})$, it also holds that $U_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}_{-k}) \geq \tilde{U}_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}})$. By combining the inequalities above, we have $U_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}_{-k}) \geq \tilde{U}_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}) \geq \tilde{U}_k(\bar{\mathbf{X}}_k, \bar{\mathbf{X}})$. Moreover, at the linearization point $\bar{\mathbf{X}}$, the two functions $\tilde{U}_k(\bar{\mathbf{X}}_k, \bar{\mathbf{X}})$ and $U_k(\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_{-k})$ have the same value. Thus, we conclude that $U_k(\tilde{\mathbf{X}}_k^\star, \bar{\mathbf{X}}_{-k}) \geq U_k(\bar{\mathbf{X}}_k, \bar{\mathbf{X}})$. ∎

A. STRATEGY UPDATE BASED ON MESSAGE PASSING

In this subsection, we design a low-complexity solution to the per-player allocation (14) that distributes the computational effort among all nodes belonging to the same cell or D2D pair.

We first observe that when player k is a D2D receiver, with d indicating its D2D transmitter, the constraints set of (14) reduces to the single constraint $\sum_{f=1}^{F} \tilde{x}_d^f \leq F_d$ for $\tilde{x}_d^f \in [0, 1]$ and for all $f \in \mathcal{F}$. Thus, the solution to (14) is straightforwardly obtained by selecting the at most F_d RBs corresponding to the largest positive utilities. For this reason, in what follows we focus on solving (14) when player k is a BS.

Message passing (MP) algorithms are attractive schemes for solving resource allocation problem without the need of a central controller. When applied to cellular systems, these algorithms involve an iterative exchange of messages between the BS and the mobile users, where every message represents the solution to a local problem with very low computational complexity. The Reweighted MP (ReMP)

framework is an example of MP approach applied to resource allocation problems [28]. In the single-cell scenario, this algorithm provably convergences to the optimal solution of the utility maximization problem (see [38]) that we report below for completeness:

$$\max_{\{\tilde{x}_n^f\}} \sum_{f \in \mathcal{F}} \sum_{n \in \mathcal{C}_k} \omega_k \log_2 \left(1 + \frac{P_n^f G_{n,k}^f}{\sigma^2}\right) \tilde{x}_n^f \qquad (15)$$

subject to $\sum_{n \in \mathcal{C}_k} \tilde{x}_n^f \leq 1, \quad \forall f \in \mathcal{F}, \qquad (15.\text{a})$

$$\sum_{f=1}^{F} \tilde{x}_n^f = F_n, \quad \forall n \in \mathcal{C}_k, \qquad (15.\text{b})$$

$$\tilde{x}_n^f \in [0, 1], \quad \forall f \in \mathcal{F}, \forall n \in \mathcal{C}_k. \qquad (15.\text{c})$$

Unlike (14), the above formulation neglects the interference. Thus, its objective is a positive weighted sum, where each weight corresponds to the achievable rate on the corresponding RB. Moreover, because the weights in the objective function of (14) can be negative, if we assume the inequalities (14.b) instead of the equalities (15.b), the total rate can increase because we do not force the allocation of RBs that contribute negatively to the utility (this happens, for example, in high-interference scenarios).

Although the optimization problem considered in [28] and [38], and recalled in (15), differs from our formulation in (14), we can still apply the ReMP scheme and benefit from its properties. Let us define the new weights $E_n^f \triangleq \max\{\tilde{E}_n^f, 0\}$, with $\tilde{E}_n^f$ given in (13) and for all $n \in \mathcal{C}_k$ and $f \in \mathcal{F}$. Then, we formulate the following optimization problem, which is in the form of (15) and thus solvable via ReMP scheme:

$$\max_{\mathbf{X}_k} \sum_{f \in \mathcal{F}} \sum_{n \in \mathcal{C}_k} E_n^f(\bar{\mathbf{X}}) x_n^f \qquad (16)$$

subject to $\sum_{n \in \mathcal{C}_k} x_n^f \leq 1, \quad \forall f \in \mathcal{F}, \qquad (16.\text{a})$

$$\sum_{f=1}^{F} x_n^f = F_n, \quad \forall n \in \mathcal{C}_k, \qquad (16.\text{b})$$

$$x_n^f \in \{0, 1\}, \quad \forall f \in \mathcal{F}, \forall n \in \mathcal{C}_k. \qquad (16.\text{c})$$

It is easy to verify that (14) is equivalent to (16). Indeed, by indicating with $\{x_n^{\star f}\}$ the optimal solutions to (16), we can construct the optimal solution $\{\tilde{x}_n^{\star f}\}$ to (14) by setting $\tilde{x}_n^{\star f} = 0$ if $\tilde{E}_n^f < 0$, and $\tilde{x}_n^{\star f} = x_n^{\star f}$ otherwise. Thus, we are now in the position of using the ReMP approach to solve Problem (14) by referring to the formulation in (16), instead.

Let us recall the ReMP routine. For the sake of notation, we omit the cell index k. We indicate with μ_{fn} the message sent by user n to the BS related to RB f, and with $\tilde{\mu}_{nf}$ the message sent in the opposite direction. These messages give a measure of the benefit for each RB-user assignment. At each iteration t of the algorithm, the messages are updated

as follows:

$$\mu_{nf}^{(t+1)} = E_n^f - \rho(E_n^j + \tilde{\mu}_{jn}^{(t)})_{F_n^{\text{th}}\backslash f} - (1-\rho)(E_n^f + \tilde{\mu}_{fn}^{(t)}),$$

(17a)

$$\tilde{\mu}_{fn}^{(t+1)} = -\rho \max_{i, i\neq n} \mu_{if}^{(t+1)} - (1-\rho)\mu_{nf}^{(t+1)},$$

(17b)

where $\rho \in (0, 1]$ is a parameter used to smooth the message dynamics and allowing convergence (see [28]); E_n^f is the reward of user n if transmitting on RB f; and $(\cdots)_{F_n^{\text{th}}\backslash f}$ denotes the F_n^{th} sorted element in set $\{(E_n^j + \tilde{\mu}_{jn}^{(t)})\}_{j\in\mathcal{F}}$, without considering the term $E_n^f + \tilde{\mu}_{fn}^{(t)}$ related to the fth RB. The decision variables x_n^f at the $(t+1)$th iteration are retrieved by computing the *node marginal* $\tau_{nf}^{(t+1)} = \mu_{nf}^{(t+1)} + \tilde{\mu}_{fn}^{(t+1)}$, and by assigning $x_n^{f(t+1)} = 1$ if $\tau_{nf}^{(t+1)} > 0$, and 0 otherwise.

The above message updating rule converges to the optimal solution of (16) and involves simple computations at each node. However, it requires several iterations before convergence, which may lead to long running times in multi-cell scenarios. Specifically, because we are applying the ReMP to update the strategy of each player, it might be impractical to wait at each round of the game for the convergence of the ReMP scheme. Moreover, a large number of iterations for the ReMP leads to a large signaling, resulting in both radio-resource and device-energy consumption. To deal with these practical concerns, we study the performance of the ReMP when only a predefined number of iterations are allowed. We refer to this scheme as Truncated ReMP(*I*) algorithm (**TReMP**(*I*)), where *I* indicates the allowed number of iterations. We indicate with **TReMP**(∞) the ReMP algorithm that runs until converging to the optimal solution of (16) (which occurs in a number of iterations that depends on the problem specifics and may change from round to round).

Stopping the ReMP before reaching the convergence might lead to a solution where the problem constraints are not all satisfied. In particular, it can be proven that at each iteration the ReMP mechanism always satisfies Constraints (16.a), but not necessarily Constraints (16.b). To ensure that **TReMP**(*I*) achieves a feasible solution, and that the better response dynamic based on **TReMP**(*I*) converges to a NE of game $\mathcal{G}$, from now we refer to **TReMP**(*I*) as the iterative scheme in (17) compelled with the following two forcing rules:

1) *Forcing rule no.1*: After *I* iterations, if the number of RBs assigned to user *n* exceeds the maximum value F_n, only the first F_n RBs with the highest marginals are selected.

2) *Forcing rule no.2*: After applying the Forcing rule no.1, player *k* updates the strategy profile only if the obtained solution increases the utility with respect to using the strategy profile of the previous iteration of the game.

Lemma 1: If in the iterative game $\mathcal{G}$ each player *k* updates her strategy profile by solving (14) with the **TReMP**(*I*) algorithm and with a given *I*, then, game $\mathcal{G}$ will convergence to a NE.

Proof: Because of Forcing rule no.2, the strategy $\mathbf{X}_k^{(i)}$ selected by player *k* at any iteration *i* of the game is such

---

Algorithm 2 TReMP(*I*)

Input: $\bar{\mathbf{X}}$, $\tilde{E}_n^f \quad \forall n \in \mathcal{C}_k, \ \forall f \in \mathcal{F}$
Output: $\mathbf{X}_k^*$

1 **Initialize:** $i \leftarrow 0; \ \mu_{nf}^{(0)} \leftarrow 0;$
$\tilde{\mu}_{fn}^{(0)} \leftarrow 0, \quad \forall n \in \mathcal{C}_k, \ \forall f \in \mathcal{F};$

2 Set $E_n^f \leftarrow \max\{\tilde{E}_n^f, 0\};$

3 **while** $i \leq I$ **do**

4   BS k computes $\mu_{nf}^{(i+1)}, \ \forall n \in \mathcal{C}_k, \ \forall f \in \mathcal{F};$

5   each user $n \in \mathcal{C}_p$ computes $\tilde{\mu}_{fn}^{(i+1)}, \quad \forall f \in \mathcal{F};$

6   each user $n \in \mathcal{C}_p$ and BS compute $\tau_{nf}^{(i+1)}$ and $x_n^{f(i+1)}, \quad \forall f \in \mathcal{F};$

7      $i \leftarrow i+1;$

8 $x_n^f \leftarrow x_n^{f(I)}, \ \tau_{nf} \leftarrow \tau_{nf}^{f(I)}, \quad \forall n \in \mathcal{C}_k, \ \forall f \in \mathcal{F};$

9 **for** $n \in \mathcal{C}_k, f \in \mathcal{F}$ **do**

10   **if** $\tilde{E}_n^f < 0$ **then**

11          $x_n^f \leftarrow 0;$

12 **for** $n \in \mathcal{C}_k$ **do**

13   define the sorted set $S_n = (\{\tau_{nf}\}_{f=1}^F, \geq) = \{s_i\}_{i=1}^F;$

14   define $I_n = \{f : \tau_{nf} \in \{s_i\}_{i=F_n}^F\};$

15   **if** $\sum_{f=1}^F x_n^f > F_n$ **then**

16          $x_n^f = 0, \ \forall f \in I_n;$

17 compute $U = \sum_{f\in\mathcal{F}} \sum_{n\in\mathcal{C}_k} E_n^f(\bar{\mathbf{X}})x_n^f;$

18 **if** $U \geq \sum_{f\in\mathcal{F}} \sum_{n\in\mathcal{C}_k} E_n^f(\bar{\mathbf{X}})\bar{x}_n^f$ **then**

19      $\mathbf{X}_k^* \leftarrow \mathbf{X}_k$

20 **else**

21      $\mathbf{X}_k^* \leftarrow \bar{\mathbf{X}}_k$

---

that $\tilde{U}_k(\mathbf{X}_k^{(i)}, \mathbf{X}^{(i-1)}) \geq \tilde{U}_k(\mathbf{X}_k^{(i-1)}, \mathbf{X}^{(i-1)})$, where $\mathbf{X}^{(i-1)} = (\mathbf{X}_k^{(i-1)}, \mathbf{X}_{-k}^{(i-1)})$. At iteration *i*, the function $\tilde{U}_k(\mathbf{X}_k^{(i)}, \mathbf{X}^{(i-1)})$ is the linearization of the function $U_k(\mathbf{X}_k^{(i)}, \mathbf{X}_{-k}^{(i-1)})$ around the previous allocation point $\mathbf{X}^{(i-1)}$. This means that at $\mathbf{X}^{(i-1)}$ the two functions have in fact the same value, that is, $\tilde{U}_k(\mathbf{X}_k^{(i-1)}, \mathbf{X}^{(i-1)}) = U_k(\mathbf{X}_k^{(i-1)}, \mathbf{X}_{-k}^{(i-1)})$. Therefore, it holds that $\tilde{U}_k(\mathbf{X}_k^{(i)}, \mathbf{X}^{(i-1)}) \geq U_k(\mathbf{X}_k^{(i-1)}, \mathbf{X}_{-k}^{(i-1)})$. Moreover, because $\tilde{U}_k(\mathbf{X}_k^{(i)}, \mathbf{X}^{(i-1)})$ is a lower bound of $U_k(\mathbf{X}_k^{(i)}, \mathbf{X}_{-k}^{(i-1)})$, it also holds that $U_k(\mathbf{X}_k^{(i)}, \mathbf{X}_{-k}^{(i-1)}) \geq \tilde{U}_k(\mathbf{X}_k^{(i)}, \mathbf{X}^{(i-1)})$. By combining the last two inequalities, we obtain $U_k(\mathbf{X}_k^{(i)}, \mathbf{X}_{-k}^{(i-1)}) \geq U_k(\mathbf{X}_k^{(i-1)}, \mathbf{X}_{-k}^{(i-1)})$, meaning that Condition (4) of performing a better response dynamic of the game is fulfilled. ∎

The **TReMP**(*I*) algorithm used to solve Problem (14) is summarized in Algorithm 2.

VI. COMPLEXITY AND PRACTICABILITY OF TReMP(*I*)
A. COMPLEXITY ANALYSIS

Let us assume that game $\mathcal{G}$ requires $I_\mathcal{G}$ rounds to converge, where one round is completed when all players have updated

their better response. We recall that I is the number of allowed iterations for the ReMP algorithm, and F is the total number of RBs. Considering an efficient sorting algorithm (e.g., Mergesort or Timsort) to compute the messages μ_{nf} in (17a), the time complexity at each node of the network is $\mathcal{O}(I_{\mathcal{G}} \cdot IF^2 \log F)$.

B. IMPLEMENTATION GUIDELINES

In order to assess the practical feasibility of the proposed algorithm, we provide some general insights into its possible implementation in an LTE-like network.

First, we look at the information needed to implement **TReMP**(I). To compute (17a), each transmitter n connected to receiver k needs to evaluate, for all $f \in \mathcal{F}$, the weights

$$E_n^f(\bar{\mathbf{X}}) = \omega_k R_n^f(\bar{\mathbf{X}}_{-k}) + P_n^f \sum_{q \in \mathcal{K}, q \neq k} \omega_q \delta_{\hat{m}(q,f)}^f(\bar{\mathbf{X}}_{-q}) G_{n,q}. \quad (18)$$

The rate $R_n^f(\bar{\mathbf{X}}_{-k})$ can be estimated by knowing the direct channel gain $G_{n,k}^f$ and the interference level $I_k^f(\mathbf{X}_{-k})$. In LTE, the direct gains towards the BS are estimated by means of *sounding reference signals* (SRS), while the direct gain between the users of a D2D pair can be assumed known from the *mode selection* decision [39]. The interference level can be estimated at the receiver k after the previous round of the game, and fed back to the transmitters on a control channel.

For each RB $f \in \mathcal{F}$, the second term on the right-hand side of (18) encompasses the information that transmitter n needs from the other players. Because collecting the CSI of all cross-links results in a large overhead, here we take advantage from the fact that transmitter n only needs an *aggregate information* related to the other players (represented by the summation), rather than every single cross-gain $G_{n,q}^f$ for all $q \neq k$. A way to leverage this aspect consists in letting each receiver broadcast a sounding signal on a narrow in-band control channel, without interfering with data transmissions. In our scheme, we assume that at the end of each iteration of the game, every receiver $q \neq k$ measures $\delta_{\hat{m}(q,f)}^f$ for all f in $\mathcal{F}$, and transmits a narrow SRS with power proportional to such measurement (e.g, a power $-\omega_q P \delta_{\hat{m}(q,f)}^f$, for a given P known to all users). Then, the desired quantity $\sum_{q \neq k} \omega_q \delta_{\hat{m}(q,f)}^f G_{n,q}^f$ can be directly measured by transmitter n without requiring any information exchange.

We consider a synchronized system, where all players know the order of playing the game, as well as the time slots during which they have to transmit the SRS signals, referred to as pilot signals. As in [19], we assume that the communication of all cellular and D2D users is synchronized by the timing signals sent by the macro cellular network, or by the global positioning system (GPS). Fig. 2 illustrates the proposed signaling to implement the algorithm. Once the information needed for computing the messages is obtained, transmitters and receiver start the message-exchange procedure that leads to the strategy update. Fig. 3
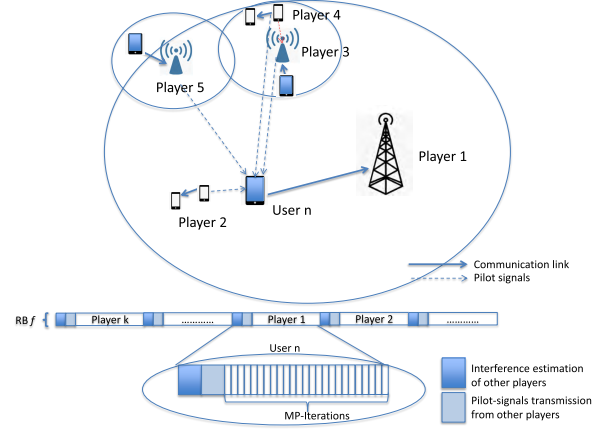

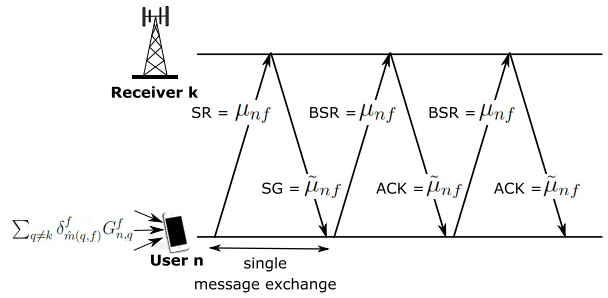**FIGURE 2.** Possible signaling protocol to implement the proposed potential game.

FIGURE 3. Implementation guidelines of TReMP(*I*). Messages μ_{nf} and $\tilde{\mu}_{nf}$ are mapped into existing control signals.

illustrates how the message exchanges of **TReMP**(I) can be mapped onto the existing LTE frameworks. Similarly to the approach proposed in [28] and [31], cellular users compute the message μ_{nf} and send it to the BS with the *scheduling request* (SR). The BS, in turn, uses the *scheduling grant* (SG) to deliver the message $\tilde{\mu}_{nf}$ to the users. Moreover, the cellular users regularly send *buffer status report* (BSR) messages to the BS, which replies with an *acknowledgement* (ACK). Note that when using **TReMP**(I), after I message exchanges both nodes (BS and cellular user) are able to retrieve the allocation decision. However, only the BS can compute the total utility and verify if the new strategy is indeed a better response. Thus, it is the BS that sends a final ACK to all users under its coverage to confirm the possible update of the RB allocation.

Regarding the D2D pairs, it is reasonable to assume that the network has already established a control channel between the two users during the *mode selection* phase. Therefore, the nth D2D transmitter collects the information to compute the F weights E_n^f, selects the best F_n RBs and informs the receiver on the allocation decision.

VII. BRANCH-AND-BOUND ALGORITHM

To evaluate the performance of the proposed better response dynamic, we compare its solution to the optimal solution of the original problem in (2). To obtain the overall optimal

allocation, we design a customized solver based on B&B approach. Although B&B is a well-established framework, its computational efficiency, compared to a naive exhaustive search, strongly depends on how well it is tailored to the specific problem. Thus, our contribution here is in designing methods to obtain tight bounds of the objective function, which allow us to remove as many branches as possible in the tree search on which the B&B is based on. The design choices of our B&B algorithm are briefly described in the sequel.

For every node of the tree, representing a partial RB allocation, we first verify if the number of assigned RBs to each user n (i.e., $\sum_f x_n^f$) exceeds the limit F_n in (2.b). If it does, then the node and the branches below can be disregarded. Otherwise, we obtain the upper and lower bounds of the node as follows: i) *Upper bound*: We compute the interference level on each RB given by the partial RB assignment of the node. Then, considering those fixed interference levels as additive noise, we solve the linear programming formulation in (2). This is an upper bound because we neglect the possible additional interference deriving from the unassigned variables. ii) *Lower bound*: We compute a feasible solution to the problem. We consider all cells in a round-robin fashion. For each cell we solve (7), where the interference level on each RB is fixed and given by the partial resource assignment of the node. As the cell are selected sequentially, the interference levels are updated accounting for the contribution of the RB allocation in the previously selected cells. By fixing the interference and treating it as additive noise, Problem (7) reduces to a maximization of a weighted sum, which can be efficiently solved by any linear programming solver. The lower bound of the node is then the sum-rate of the entire network.

VIII. NUMERICAL RESULTS AND DISCUSSION

To study the behavior of the proposed algorithm, we simulate networks consisting of 7 cells (one macro and 6 micro cells), with the BSs located in the center of the cells. All mobile users are randomly placed within the macro-cell area. Some cellular users are attached to the BSs of the micro cells, others are connected to the macro BS. Without loss of generality, we assume that the number of uplink transmissions is the same in all cells. Moreover, $F_n = 1$ for each user n, and $\omega_k = 1$ for each cell k. A detailed discussion on the effect of using different weights $\{\omega_k\}$ is presented in Subsection VIII-C, where we also increase to $F_n = 2$ the maximum number of resource assigned to user n.

Table 1 reports the main simulation parameters.

A. ALGORITHM PERFORMANCE UNDER DIFFERENT INITIAL CONDITIONS AND ORDERS OF PLAY THE GAME

First, we analyze the effect of the initial condition (i.e., the initial strategy chosen by the players) on the achieved NE of the game. We consider two cases:

- *No Initial RB Allocation* (NIA), when the players start the game without any RB allocation;

TABLE 1. Simulation parameters.

Parameter	Value
Lognormal shadow fading	6 dB
No. of cells	7
Estimated noise and interference per RB	-110 dBm
Macro cell radius	1 Km
Micro cell radius	200 m
Max Tx power macro cell users	300 mW
Max Tx power micro cell and D2D users	150 mW
SINR target	15 dB
Noise power	-174 dBm/Hz
Max no. of RBs assigned to users	1o0r 2
Path loss coefficient mobile-BS	3.5
Resource block's bandwidth	180 KHz
Path loss coefficient mobile-mobile	4
Path loss compensation factor	0.8
Allowed distance range for D2D pairs	20-60 m

- *Optimal RB allocation with No Interference* (ONI), when the players start the game with the optimal RB allocation computed by solving (7) when neglecting the interference. In this case, each player solves a positive-weighted sum maximization in the form of (15).

We run 500 independent simulations of a seven-cell network with 15 cellular users in each cell and 35 D2D pairs randomly located within the macro cell area. Table 2 shows the average utility when using either **TReMP**(1) or **TReMP**(∞). Apart from the total rate, we also display the number of rounds of the game before converging to the NE. A single *round* is completed when all players have updated their strategy. Results show that using the ONI approach slightly boosts the utility and reduces the number of rounds before convergence.

Second, we investigate the impact that the players scheduling has on the performance of the game. We evaluate the distribution of the utility when considering 2000 different player orders for the same network topology used in the previous analysis. Fig. 4 shows the results when using the NIA and the ONI initial allocation, respectively. The reported results are obtained as follows. After each strategy update, we compute the mean (μ) and the standard deviation of the utility (σ) from the 2000 outcomes. Doing so, we get a measure of the amount by which the utilities achieved with the different orders of playing deviate from the average. Results show that the final dispersion with both ONI and NIA is small, with the standard deviation decreasing with the number of strategy updates. This indicates that the behavior of the game does not change much with respect to the order of playing. Furthermore, in line with the results in Table 2, we see that the ONI initial condition slightly boosts the performance in terms of final utility, and is even less sensitive to the choice of the initial order of playing the game.

We conclude that both the initial strategy of the players and the order of playing the game have a negligible impact on the local optimum achievable with the proposed algorithm. These results are interesting from an implementation perspective, they allow the players to start the game with any allocation and use any order of playing with no particular penalty.
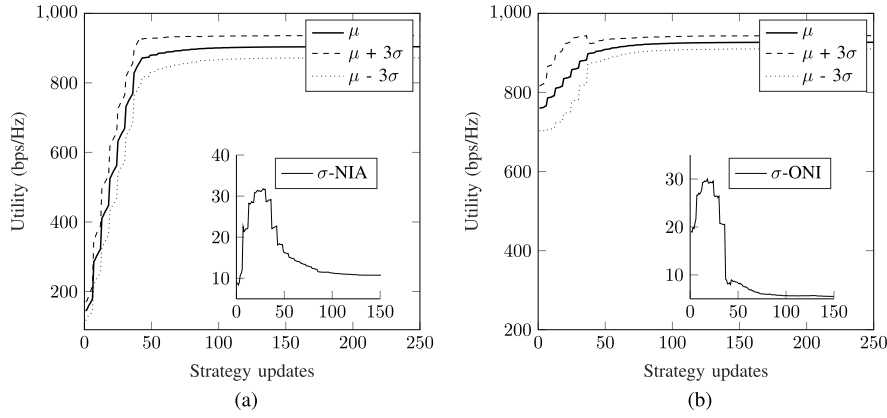
FIGURE 4. Average utility (μ) and three standard deviations range (3σ) from 2000 different orders of playing the game. Players strategies are updated with TReMP(8).

**TABLE 2.** Performance with different initial conditions.

Algorithm	Initial condition	Utility (bps/Hz)	No. of rounds
TReMP(1)	NIA	802.9	4.4
TReMP(1)	ONI	838.8	3.8
TReMP(∞)	NIA	889.9	4.3
TReMP(∞)	ONI	902.3	3.8

B. CONVERGENCE AND PRACTICAL IMPLEMENTATION

In this subsection, we focus on the convergence behavior of the proposed game. We compare the performance of the **TReMP**(I) algorithm with different inputs I, for a single network realization. In particular, we apply **TReMP**(1), **TReMP**(4), **TReMP**(8), and **TReMP**(∞) to compute the better response of the players. We consider again a seven-cell network with 15 cellular users per cell and 35 D2D pairs, so that there are 42 players in total (35 D2D receivers and 7 BSs). As such, a single round of the game is given by 42 strategy updates, one for each player. We refer to a *message exchange* as the action of sending a message between the transmitter and receiver in both directions. A message exchange is a single iteration of **TReMP**(I).

We recall that when the player is a D2D receiver, the strategy update requires a single message exchange, while when the player is a BS, the strategy update requires a number of message exchanges depending on the chosen I.

In Fig. 5, we report the sum rate after each message exchange for the four strategy-update policies, marking the message exchanges representing the completion of each round of the game. As expected, regardless of the used policy, the potential function is always nondecreasing in the number of message exchanges, and the game converges to a NE. All the four better response dynamics require a maximum of four rounds to converge, but the number of message exchanges and the utility achieved at the equilibrium increase with the chosen I. In particular, **TReMP**(1) converges faster than the other schemes, with $\approx$ 60 message exchanges, compared to the minimum of $\approx$ 150 of the other dynamics.
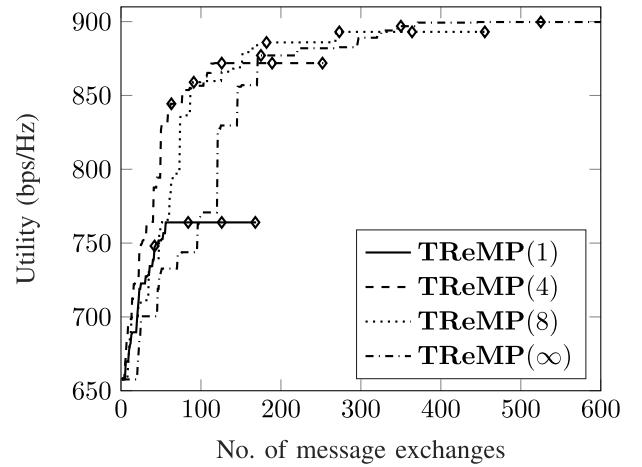
FIGURE 5. Convergence behavior of the potential function. Each message exchange corresponds to one iteration of the TReMP(*I*) algorithm used to compute the better response of the player. In the plot, the completion of each round is marked.

On the other hand, it achieves a utility that is $\approx$ 11% and $\approx$ 14% smaller than the one achieved with **TReMP**(4) and **TReMP**(8), respectively.

Another insight is that, although **TReMP**(1) performs the worst in terms of final utility, it reaches higher utility than **TReMP**(8) and **TReMP**(∞) when the number of message exchanges is below 30. On the contrary, **TReMP**(∞) achieves the highest utility, but only after completing its second round; that is, after $\approx$ 320 message exchanges. Accordingly, we can conclude that using small values of I is desirable when the number of allowed message exchanges is small.

Table 3 shows the average results for 500 independent network realizations with the same setup as used in the previous analysis. In line with the results shown for a single network realization in Fig. 5, both the average number of rounds and the average utility increase with the parameter I. Interestingly, this is not the case when using **TReMP**(∞), for which the number of rounds reduces compared to when using **TReMP**(8) and **TReMP**(4). However, the utility increment

Algorithm	Utility (bps/Hz)	No. of rounds	Convergence time
TReMP(1)	839.3	3.9	$\approx 155\,t$
TReMP(4)	893.4	4.1	$\approx 309\,t$
TReMP(8)	900.7	4.4	$\approx 455\,t$
TReMP(∞)	902.3	3.8	$\approx 805\,t$

* t ms is the time for a single message exchange between two nodes

comes at the expense of more messages exchanged for each round. In the considered setup, **TReMP**(∞) converges after an average of 20 iterations.

On the basis of the above observations, and referring to the results in Table 3, we now investigate the practical implication of using **TReMP**(I) with different input I. In general, with **TReMP**(I), each round of the game is completed after 35 message exchanges to allocate the RBs to all D2D pairs (i.e., 1 message exchange for each D2D pair), and $7 \times I$ message exchanges to allocate the RBs to all cellular users (i.e., I message exchanges for each cell). Assuming that a single message exchange takes t ms, the total time needed for **TReMP**(1) to converge is approximately $3.7 \times (35 + 7) \times t = 155.4 \times t$ ms, while for **TReMP**(4) and **TReMP**(8) it is approximately $4.9 \times (35 + 7 \times 4) \times t = 308.7 \times t$ ms and $5 \times (35 + 7 \times 8) \times t = 455 \times t$ ms, respectively. The required time clearly increases for **TReMP**(∞) case, for which we need, approximately, 20 iterations to converge and thus $4.6 \times (35 + 7 \times 20) \times t = 805 \times t$.

These results suggest that small values of I are more suitable for dynamic networks. In particular, $I = 1$ becomes preferable when the network conditions change fast, such as, for example, in high mobility scenarios, where a short coherence time of the channel requires fast resource allocation. Furthermore, when the number of message exchanges necessary to update the player strategy is too large (e.g., for $I = \infty$), the energy consumption for the mobile users may become considerable. Therefore, the selection of parameter I in the design of the **TReMP**(I) algorithm is crucial to find the desired trade-off between the achievable performance and the practical implementation.

C. ALGORITHM PERFORMANCE FOR DIFFERENT $\{\omega_k\}_{k \in \mathcal{B}}$

In the problem formulation (2), we introduced the design parameters $\{\omega_k\}_{k \in \mathcal{B}}$ as means to differentiate the impact on the total utility from the resource allocation in the different cells. Here, we evaluate the achievable performance when we increase the weight ω_k for the macro cell k, while keeping ω_q equal to 1 for all $q \neq k$. We consider the same network setting as before, with 15 cellular users per cell and 35 D2D pairs within the macro cell area. We apply **TReMP**(8) to compute the best response of the players, and we set $F_n = 2$ for each user n.

Fig. 6 reports the average bit rate for the different types of users (namely, macro-cell user, micro-cell user, and D2D user) after 500 monte carlo runs. We can see that by
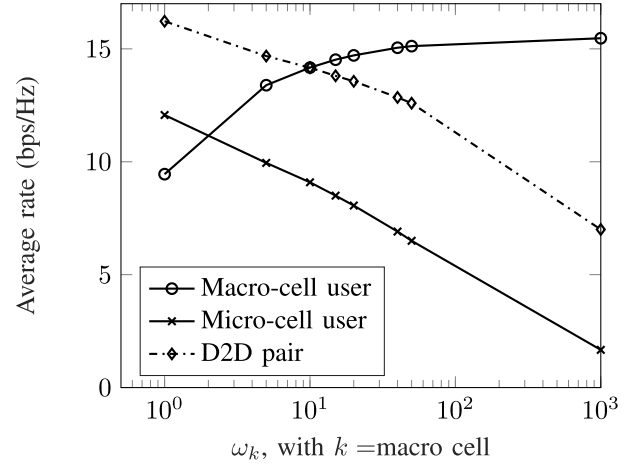
FIGURE 6. Achievable rates for the different type of cells with respect to different macro-cell weight ω_k l.

increasing the weight ω_k, the average rate of the macro-cell users increases as well. This because of the larger contribution of the macro cell to the total utility function when its utility is multiplied by a large ω_k. On the other hand, the utility of both the micro-cell users and D2D pairs decreases with larger ω_k. Note that in the extreme case of $\omega_k > 1000$, only the macro-cell users are allowed to transmit and the rate of all other users reduces practically to 0.

D. COMPARISONS WITH THE GLOBAL OPTIMUM

Finally, we compare the globally RB allocation obtained via the customized B&B solver (*Opt*), with the suboptimal allocation achieved with four different schemes: i) a random RB allocation that fulfills the problem constraints (*Random*); ii) a potential game with the best response dynamic based on the solution of (7) with exhaustive search (*Best*); iii) a potential game with the better response dynamic based on **TReMP**(∞); and iv) a potential game with the better response dynamic based on **TReMP**(1). Although the B&B solver allows to find the optimal solution more efficiently than using the naive exhaustive search, there might be cases when it is necessary to explore many nodes of the search tree before finding the optimal solution. For these unlucky cases, the run-time of the algorithm is still prohibitive.

To deal with a reasonable computational time, the results presented in this subsection are related to slightly simplified scenario with neighboring cells, where in each cell there are 3 cellular users and 2 D2D pairs, and with $F_n = 1$ for all n. We consider 100 independent simulations to build the histograms in Fig. 7, which show the utility loss of the four suboptimal solutions. We define the percentage rate loss as $\frac{U^\star - U_A}{U^\star} * 100$, where $U^\star$ and U_A represent the utility achieved with the B&B solver and with the suboptimal approaches, respectively.

The histograms in Fig. 7 show that the performance of the Best and of the two better response dynamics (**TReMP**(∞) and **TReMP**(1)) are not significantly far from the global optimum. Specifically, the Best is within 10% of the optimal solution for almost all network configurations. However,
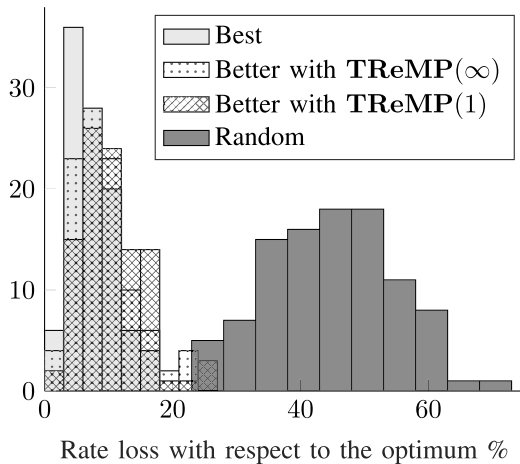
FIGURE 7. Histogram of the percentage of utility loss when using the Best response dynamic, the Random allocation, and the Better response dynamic with TReMP(∞) and TReMP(1).

the computation of each Best response requires the solution of a combinatorial problem, which leads to runtime limitations for practical network sizes. The two better response dynamics, on the other hand, perform slightly worse than the Best, but remain more practical for all network sizes (especially **TReMP**(1)) and they are still within 15% of the optimal solution for almost all configurations. Finally, it is worth mentioning here that because the set of NE points achieved with the best and better response dynamics is the set of local maxima of the potential function, playing the game not only improves the system performance compared with any initial allocation; in some cases, it also achieves the global optimum as shown by the bins located at zero utility-loss in the histograms.

IX. CONCLUSIONS
In this work, we considered the problem of uplink RB allocation in a heteregeneous 5G D2D-enabled network. Using the framework of potential games, we addressed the challenging combinatorial problem of sum-rate maximization in an interference-limited network. Specifically, we designed a distributed solution based on a better response dynamic and a MP algorithm that guarantees convergence to a local optimum of the utility function. Simulations showed the validity of the proposed scheme under different settings.

X. ACKNOWLEDGMENT

REFERENCES
[1] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, "Cellular traffic offloading onto network-assisted device-to-device connections," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 20–31, Apr. 2014.
[2] G. Fodor *et al.*, "Design aspects of network assisted device-to-device communications," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 170–177, Mar. 2012.
[3] D. D. Penda, L. Fu, and M. Johansson, "Energy efficient D2D communications in dynamic TDD systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1260–1273, Mar. 2017.
[4] P. Mach, Z. Becvar, and T. Vanek, "In-band device-to-device communication in OFDMA cellular networks: A survey and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1885–1922, 4th Quart., 2015.
[5] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 40–48, Apr. 2014.
[6] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying LTE network," in *Proc. IEEE 6th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2010, pp. 368–375.
[7] A. Abrardo and M. Moretti, "Distributed power allocation for D2D communications underlaying/overlaying OFDMA cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1466–1479, Mar. 2017.
[8] J. Wang, D. Zhu, C. Zhao, J. C. F. Li, and M. Lei, "Resource sharing of underlaying device-to-device and uplink cellular communications," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1148–1151, Jun. 2013.
[9] Y. Li, D. Jin, J. Yuan, and Z. Han, "Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3965–3977, Jul. 2014.
[10] X. Ma, J. Liu, and H. Jiang, "Resource allocation for heterogeneous applications with device-to-device communication underlaying cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 15–26, Jan. 2016.
[11] Y. Jiang, Q. Liu, F. Zheng, X. Gao, and X. You, "Energy-efficient joint resource allocation and power control for D2D communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6119–6127, Aug. 2016.
[12] H. Meshgi, D. Zhao, and R. Zheng, "Optimal resource allocation in multicast device-to-device communications underlaying LTE networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8357–8371, Sep. 2017.
[13] X. Mi, L. Xiao, M. Zhao, X. Xu, and J. Wang, "Statistical QoS-driven resource allocation and source adaptation for D2D communications underlaying OFDMA-based cellular networks," *IEEE Access*, vol. 5, pp. 3981–3999, 2017.
[14] Q. Ye, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "Distributed resource allocation in device-to-device enhanced cellular networks," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 441–454, Feb. 2015.
[15] M. Belleschi, G. Fodor, and A. Abrardo, "Performance analysis of a distributed resource allocation scheme for D2D communications," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2011, pp. 358–362.
[16] J. Gu, S. J. Bae, S. F. Hasan, and M. Y. Chung, "Heuristic algorithm for proportional fair scheduling in D2D-cellular systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 769–780, Jan. 2016.
[17] H. Feng, H. Wang, X. Xu, and C. Xing, "A tractable model for Device-to-Device communication underlaying multi-cell cellular networks," in *Proc. IEEE Int. Conf. Commun. Workshops*, Jun. 2014, pp. 587–591.
[18] H. Xu, G. L. Stüber, W. Xu, C. Pan, J. Shi, and M. C. Z. Yang, "Robust transmission design for multicell D2D underlaid cellular networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 47, pp. 5922–5936, Jul. 2018.
[19] Y. Xiao, K.-C. Chen, C. Yuen, Z. Han, and L. A. DaSilva, "A Bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 4034–4051, Jul. 2015.
[20] A. Asheralieva, "Bayesian reinforcement learning-based coalition formation for distributed resource sharing by device-to-device users in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5016–5032, Aug. 2017.
[21] B. Zhang, X. Mao, J.-L. Yu, and Z. Han, "Resource allocation for 5G heterogeneous cloud radio access networks with D2D communication: A matching and coalition approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5883–5894, Jul. 2018.
[22] J. Huang, C.-C. Xing, Y. Qian, and Z. J. Haas, "Resource allocation for multicell device-to-device communications underlaying 5G networks: A game-theoretic mechanism with incomplete information," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2557–2569, Mar. 2018.
[23] H. Dai, Y. Huang, R. Zhao, J. Wang, and L. Yang, "Resource optimization for device-to-device and small cell uplink communications underlaying cellular networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1187–1201, Feb. 2018.
[24] X. Cheng, L. Yang, and X. Shen, "D2D for intelligent transportation systems: A feasibility study," *IEEE Trans. Intell. Trans. Syst.*, vol. 16, no. 4, pp. 1784–1793, Jan. 2015.

[25] D. D. Penda, A. Abrardo, M. Moretti, and M. Johansson, "Potential games for subcarrier allocation in multi-cell networks with D2D communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[26] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[27] A. Abrardo, M. Belleschi, P. Detti, and M. Moretti, "Message passing resource allocation for the uplink of multi-carrier multi-format systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 130–141, Jan. 2012.

[28] A. Abrardo, M. Belleschi, G. Fodor, and M. Moretti, "A message passing approach for resource allocation in cellular OFDMA communications," in *Proc. Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 4583–4588.

[29] K. Yang, N. Prasad, and X. Wang, "A message-passing approach to distributed resource allocation in uplink DFT-spread-OFDMA systems," *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 1099–1113, Apr. 2011.

[30] S. H. Lee and I. Sohn, "Message-passing strategy for joint user association and resource blanking in HetNets," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1026–1037, Feb. 2018.

[31] M. Hasan and E. Hossain, "Distributed resource allocation for relay-aided device-to-device communication: A message passing approach," *IEEE Wireless Commun.*, vol. 13, no. 11, pp. 6326–6341, Nov. 2014.

[32] S. H. Lee, D. R. Shin, H. W. Jeong, and Y. H. Kim, "Distributed bargaining strategy for downlink virtual MIMO with device-to-device communication," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1503–1516, Apr. 2016.

[33] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, 2011.

[34] Q. D. Lã, Y. H. Chew, and B.-H. Soong, *Potential Game Theory: Applications in Radio Resource Allocation*. Cham, Switzerland: Springer, 2016.

[35] A. B. MacKenzie and L. A. DaSilva, *Game Theory for Wireless Engineers* (Synthesis Lectures on Communications), vol. 1, no. 1. San Rafael, CA, USA: Morgan & Claypool Publishers, 2006, pp. 1–86.

[36] R. T. Rockafellar, *Convex Analysis* (Princeton Landmarks in Mathematics and Physics). Princeton, NJ, USA: Princeton Univ. Press, 1996.

[37] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[38] M. Moretti, A. Abrardo, and M. Belleschi, "On the convergence and optimality of reweighted message passing for channel assignment problems," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1428–1432, Nov. 2014.

[39] D. D. Penda, L. Fu, and M. Johansson, "Mode selection for energy efficient D2D communications in dynamic TDD systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 5404–5409.

ANDREA ABRARDO received the M.E. degree in electronic engineering from the University of Florence, Italy, in 1993, and the Ph.D. degree in communications systems for medical care applications, in 1998. In 1998, he joined the Department of Information Engineering, University of Siena, Italy, as an Assistant Professor, where he is currently an Associate Professor of digital communication and 5G technologies with the Department of Information Engineering and Mathematical Sciences. His research interests include 5G wireless communication systems, wireless sensor network, and signal processing. In 2011, he was a recipient of the Transactions Prize Paper Award from the IEEE Geoscience and Remote Sensing Society. He is currently a Coordinator of the National Inter-University Consortium for Telecommunications activity within the 5G trials in the area of Milan, Italy, collaborating with Vodafone Italia.

MARCO MORETTI (M'03) received the degree in electronic engineering from the University of Florence, Florence, Italy, in 1995, and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 2000. From 2000 to 2003, he was a Senior Researcher with Marconi Mobile. He is currently an Associate Professor with the University of Pisa, Pisa, Italy. His research interests include resource allocation for multicarrier systems, synchronization, and channel estimation. He is currently an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

DEMIA DELLA PENDA received the B.Sc. and M.Sc. degrees in telecommunications engineering from the University of Siena, in 2008 and 2011, respectively, and the Ph.D. degree from the Department of Automatic Control, KTH Royal Institute of Technology, in 2018. She is currently a Development Engineer with Ericsson AB, Kista, Sweden. Her research interest includes analysis and optimization of future cellular systems, focusing on the design of resource allocation techniques for D2D-enabled networks.

MIKAEL JOHANSSON received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Sweden, in 1994 and 1999, respectively. He held Postdoctoral positions with Stanford University and with the University of California at Berkeley. He joined the KTH Royal Institute of Technology, in 2002, where he is currently a Full Professor. He has published two books and over 100 papers, several of which are highly cited and have received recognition in terms of best paper awards. He served on the Editorial Board of *Automatica* and the IEEE TRANSACTIONS ON CONTROL OF NETWORKED SYSTEMS and served on the program committee for several top-conferences organized by the IEEE and ACM. He has played a leading role in several national and international research projects in control and communications.

• • •