Distributed Interference and Energy-Aware Power Control for Ultra-Dense D2D Networks: A Mean Field Game

Chungang Yang, Jiandong Li, Prabodini Semasinghe, Ekram Hossain, Samir Perlaza, Zhu Han

Distributed Interference and Energy-Aware Power Control for Ultra-Dense D2D Networks: A Mean Field Game

Chungang Yang, Jiandong Li, Prabodini Semasinghe, Ekram Hossain,
Samir M. Perlaza, and Zhu Han

***Abstract*—Device-to-device (D2D) communications can enhance spectrum and energy efficiency due to direct proximity communication and frequency reuse. However, such performance enhancement is limited by mutual interference and energy availability, especially when the deployment of D2D links is ultra-dense. In this paper, we present a distributed power control method for ultra-dense D2D communications underlaying cellular communications. In this power control method, in addition to the remaining battery energy of the D2D transmitter, we consider the effects of both the interference caused by the generic D2D transmitter to others and interference from all others' caused to the generic D2D receiver. We formulate a mean-field game (MFG) theoretic framework with the interference mean-field approximation. We design the cost function combining both the performance of the D2D communication and cost for transmit power at the D2D transmitter. Within the MFG framework, we derive the related Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK) equations. Then, a novel energy and interference aware power control policy is proposed, which is based on the Lax-Friedrichs scheme and the Lagrange relaxation. The numerical results are presented to demonstrate the spectrum and energy efficiency performances of our proposed approach.**

***Index Terms*—Device-to-device communication, mean field game, spectrum efficiency, energy efficiency.**

I. INTRODUCTION

Device-to-device (D2D) communications underlying conventional cellular networks improve energy and spectrum efficiency [1]. These beneficial opportunities are achieved due to the proximity between the devices and frequency reuse, however, these benefits also come with technical challenges [2]. For instance, both intra-tier and inter-tier interferences exist in D2D communications, which affect the system performance, and thus need to be mitigated.

Since D2D devices are generally powered by batteries, extending the battery life and saving energy are important to improve users' experience. Consequently, the performance of D2D communications is limited by mutual interference and energy availability, in particular, when the deployment of D2D links is ultra-dense [3]–[5]. To optimize both spectrum and energy efficiency, different techniques have been designed to mitigate interference and save energy [5]–[7]. For instance, interference coordination [5], interference mitigation [6], and resource management [7] have been investigated aiming at improving the spectrum and energy efficiency. In addition, power control is critical to D2D communications [8]–[12], and it is proved that optimal power control can both save energy and mitigate interference. Power control is an interactive process among different D2D players, which is due to the coupled interference relationships in a full spectrum reuse scenario.

To characterize the dynamic interactive power control, game theory has been extensively used in the literature. In particular, game theory has been used to model competition among transmitters and interference coordination, analyze the strategic behavior of transmitters, and design distributed algorithms. Both cooperative game and non-cooperative game theory have found applications into D2D communications [14]–[23]. However, these classical game models are difficult to analyze when the number of D2D links becomes large.

Mean filed games (MFGs) are promising alternatives to model and analyze a large-scale D2D communication network, where an MFG models individual player's interaction with the average effect of the collective behavior of the players [24]–[26]. This collective behavior is modelled by a mean field, which denotes the statistical distribution of the considered system state. In this case, interactions among individual players become the interactions of the considered player with the mean field, which can be modelled by a Hamilton-Jacobi-Bellman (HJB) equation in the mean field game. The dynamics of the mean field according to the players' actions can be modelled by a Fokker-Planck-Kolmogorov (FPK) equation. These coupled FPK and HJB equations are also called forward and backward equations, respectively. The mean field equilibrium

C. Yang and J. Li are with the State Key Laboratory on Integrated Services Networks, Xidian University, Xi'an, 710071 China (emails: {cgyang, jdli}@mail.xidian.edu.cn); C. Yang is also supported by National Mobile Communications Research Laboratory, Southeast University.

E. Hossain and P. Semasinghe are with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada (emails: semasilp@myumanitoba.ca, Ekram.Hossain@umanitoba.ca).

Samir M. Perlaza is with the Institut National de Recherche en Informatique et en Automatique (INRIA), Lyon, France. He is also with the Department of Electrical Engineering at Princeton University, Princeton, NJ, USA (email: samir.perlaza@inria.fr).

Z. Han is with Electrical and Computer Engineering, University of Houston, Houston, TX. (email: hanzhu22@gmail.com).

This work was supported by National Science Foundation of China (61231008); by Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (No.2014D10); by Postdoctoral Science Foundation (2016T90894 and 154066); by the ISN02080001, JB150111, and 61201139; by the 111 Project B08038; and by the Shaanxi Province Science and Technology Research and Development Program (2011KJXX-40). This work was also supported by US NSF CPS-1646607, ECCS-1547201, CCF-1456921, CNS-1443917, ECCS-1405121, NSFC61428101 and in part by the European Commission under Marie Skłodowska-Curie Individual Fellowship No. 659316.

of an MFG can be obtained by solving these two equations [27]–[29], [31]–[33].

MFGs have found wide applications, such as in cognitive radio networks [27], green power control [28], heterogeneous cellular networks [29], [31], cloud-based networks [32], and smart power grids [33]. The works in references [29], [31] are the most closely related works. In [29], a mean filed approximation method was used to develop power control methods for small cell base stations (BSs) in a two-tier dense HetNet. The model did not consider the HJB and FPK equations and sub-optimal solutions were obtained. The work in [31] modelled the downlink power control problem for small cell BSs in a two-tier HetNet as an MFG considering the remaining battery power at an SBS as the system state. Alternatively to [29] and [31], we propose a distributed power control method for D2D transmitters that is both interference and energy-aware. The MFG model used here considers a two-dimensional system state in contrast to a one-dimensional system state as in [31], which makes the MFG model more complex and hence more difficult to solve. The contributions of this work can be summarized as follows:

(1) **An MFG framework for ultra-dense D2D networks:** We formulate an MFG theoretic framework for ultra-dense D2D networks, where we assume that the number of D2D links can approach infinity. In this framework, we jointly consider the remaining energy at a D2D transmitter and the interference caused by the D2D transmitter to the other links as the state space and an optimal distributed power control policy is obtained.

(2) **Energy and interference-aware problem formulation:** In the proposed MFG framework, the problem is formulated as a cost minimization problem with two kinds of interference into consideration. The interference from other D2D links to the generic D2D link is investigated in the cost function, while the interference dynamics introduced by the generic D2D transmitter to other D2D links is regarded as one of the constraints. Another constraint is the remaining energy level at the generic D2D transmitter. To capture the effects of both types of interferences, we use a mean-field approximation approach. This facilitates designing a distributed power control policy for a generic D2D transmitter. Moreover, this leads to a social optimal power control [1].

(3) **Distributed iterative algorithm to obtain the MFG equilibrium:** We derive the corresponding HJB and FPK equations for the presented D2D MFG framework. A joint finite difference algorithm based on the Lax-Friedrichs scheme and Lagrange relaxation is proposed to solve the coupled HJB and FPK equations, respectively.

(4) **Improved spectrum and energy efficiency:** The numerical results are presented to characterize the mean field distributions and the power control policy of a generic D2D transmitter. The proposed algorithm can improve both spectrum and energy efficiency when compared to some benchmark schemes.

TABLE I
LIST OF PARAMETERS

Symbol	Definition
$\mathcal{N}$	Set of D2D links
N	Number of D2D links
i, j	Index of D2D links
$p_i(t)$	Power of D2D link i at time t
$p_j(t)$	Power of D2D link j at time t
$g_{i,j}$	Channel gain from D2D link i to j
$g_{i,o}$	Channel gain from D2D link i to o
$g_{j,i}$	Channel gain from D2D link j to i
$I_{i\to}$	Interference from D2D link i to others $j \neq i$
$I_{i\to o}$	Interference from D2D link i to macrocell player o
$I_{\to i}$	Interference to D2D link i from others $j \neq i$

The remainder of this paper is organized as follows. In Section II, we survey the related work. We introduce the system model and formulate the problem in Section III. A D2D differential game is presented in Section VI, and converted to the optimal control problem in Section V. With the ultra-dense deployment case, we propose the D2D mean field game in Section VI. Then, in Section VII, we design an iteration algorithm to solve it. Numerical results are given in Section VIII, and finally we conclude the paper in Section IX. A list of commonly used notations throughout this paper is shown in Table I.

II. Related Work

Power control for D2D communications has been investigated in [8]–[12]. More specially, an iterative combinatorial auction-based power and channel allocation algorithm was proposed to improve energy efficiency in [8], where the Peukert's law was employed to characterize the non-linear effects of the battery lifetime. In [9], a joint spectrum sharing and power allocation problem was formulated as a non-convex optimization problem with perfect global channel state information assumption. With local awareness of the interference, multiuser diversity was used to minimize interference in [11]. The authors in [12] developed a distributed power control algorithm, where it relies on the large-scale path-loss measurements rather than requiring full channel states. It was concluded that proper power control can mitigate the interference, thus increasing the sum rate [10].

The optimization models in [8]–[12] do not characterize the rational behaviors of the D2D transmitters and the dynamics. Therefore, game theory has been used to model the problem of power control and resource management in D2D communications [15]–[19]. For instance, the power control problem in D2D communications can be modelled as both cooperative games and non-cooperative games [19]. Meanwhile, the case of incomplete information was modelled as an exact potential game in [15]. To jointly address the spectrum utilization and user-controlled mode selection, the authors in [16] proposed a dynamic Stackelberg game framework, where the base station and the potential D2D user act as the leader and the follower,

[1]Traditional social optimal solution is the solution that gives the highest aggregate payoff or minimum aggregate cost of all players, e.g., using the minimum aggregate cost as an example, $\min \sum_{\forall i} c_i(t)$ where c_i is the cost function of user i. In our work, distributed social optimal means that each player minimizes her own cost function but with the effects of interference $\mu_i(t)$ to others into consideration, thus resulting in the distributed optimization problem $\min\{c_i(t) - A\mu_i(t)\}$, where A is the introduced Lagrangian parameter for the interference effects $\mu_i(t)$ to others.

respectively. The authors in [17] investigated a repeated game theoretic resource allocation, where a D2D link was located in the overlapping area of two neighboring cells. In [18], a decentralized framework of joint spectrum allocation and power control was proposed to coordinate interference between the D2D layer and the cellular layer.

The authors in [20] formed D2D clusters using the distributed merge-and-split algorithm from the coalition game theory and defined a relaxation factor to give a constraint on total energy consumption for each cluster. In [21], two types of games were considered including the non-cooperative Stackelberg and the cooperative Nash bargaining game. The authors in [22] introduced a spectrum sharing model for D2D links which was analyzed by using a Bayesian non-transferable utility overlapping coalition formation game model. The authors in [23] developed a hierarchical framework based on a layered coalitional game for operator-controlled D2D networks with multiple D2D operators at the upper layer and a group of devices at the lower layer.

In summary, both non-cooperative and cooperative games have been widely used to derive distributed resource allocation techniques. For example, potential games [15], Stackelberg games [16], repeated games [17], pricing-based games [18] have found extensive applications. On the other hand, cooperative games in [19], [20], Bayesian overlapping coalition games [22], the layered coalition games [23], and Nash bargaining cooperative games [21] are cooperative games. As has been mentioned before, these conventional games model the interaction of each player with every other player; however, the analysis of a system with a large number of players (e.g., an ultra-dense D2D network) can be complex with these traditional game models.

III. System Model and Problem Formulation

We consider an ultra-dense D2D communications network where the D2D communication pairs share uplink resources with some existing macrocell user equipments (MUEs). We assume that there are N D2D pairs sharing the same channel with the cellular uplink, as shown in Fig. 1. The term "ultra-dense" implies the following:

- The number of D2D communication pairs, denoted by N, is very large, that is, $N \to +\infty$;
- Full frequency reuse, i.e., the D2D communication pairs use the same frequency;
- Most of the user devices select the D2D communication mode, and there exist relatively a small number of macrocell users.

We consider the effects of both the interference of a generic D2D transmitter introduced to others, and all others' interference cause to a generic D2D receiver. There are termed as intra-tier and inter-tier interference, respectively. For instance, the D2D transmitter D2D_1^T communicating with its receiver introduces intra-tier interference to the other D2D receivers as shown in Fig. 1. Meanwhile, due to the full frequency reuse, D2D_1^T causes inter-tier interference link to the MUE. Here, we define the interference introduced by player i, $i \in \mathcal{N}$ to others $j \in \mathcal{N}, j \neq i$ at time t as

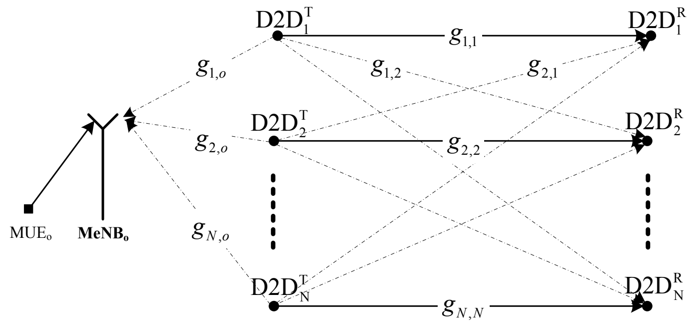

Fig. 1. A D2D network with a large number of D2D links.

$$I_{i\to}(t) = \sum_{j=1, j\neq i}^{N} p_i(t) g_{i,j}(t), \tag{1}$$

where $p_i(t)$ is the transmit power corresponding to D2D pair i, $i \in \mathcal{N}$, and $g_{i,j}(t)$ defines the channel gain from the D2D pair i's transmitter to the D2D pair j's receiver, $j \in \mathcal{N}, j \neq i$. Here, the terms $p_i(t)$ and $g_{i,j}(t)$ are positive. Therefore, (1) gives the interference introduced by player i, $i \in \mathcal{N}$ to all other D2D receivers at time t, where player i, $i \in \mathcal{N}$ is called the generic D2D transmitter.

Meanwhile, the transmission of player i, $i \in N$ also introduces interference to player o, where we define the only existing uplink MUE to Macrocell evolved node B (MeNB) pair as player o. We define the inter-tier interference introduced by player i to player o as

$$I_{i\to o}(t) = p_i(t) g_{i,o}(t), \tag{2}$$

where $g_{i,o}(t)$ is channel gain between the D2D pair i's transmitter and the MeNB.

Finally, the interference perceived by the D2D pair i at time t, which is the interference introduced by other D2D links to the generic D2D link i, is given as

$$I_{\to i}(t) = \sum_{j=1, j\neq i}^{N} p_j(t) g_{j,i}(t). \tag{3}$$

Here, we assume that orthogonal channels are used for different MUEs, and we do not consider any power control policy at the macrocell layer. In this paper, our focus is on the power control policy for the D2D transmitters.

The achieved signal-to-interference-plus-noise ratio (SINR) at the receiver of D2D pair i at time t is

$$\gamma_i(\mathrm{t}) = \frac{p_i(t) g_{i,i}(t)}{I_{\to i}(t) + \sigma^2}, \tag{4}$$

where σ^2 is the thermal noise power.

With the above definition of SINR, the power control problem can be summarized as follows: each player i (the transmitter for D2D pair i) will determine the optimal power control policy $Q_i^\star(t)$ ($t \in [0, T]$) considering the interference $I_{i\to}$ introduced to others, the interference introduced by others $I_{\to i}$, in addition to the remaining energy. Here the power control policy is a series of power control actions. The power control problem can be formulated as a differential game due to the interference dynamics and the energy dynamics [24]–[29]. Different from the studies in [24]–[29], we formulate a power control differential game with two-dimensional state space and a new cost function.

IV. Differential Game Model for Power Control

The differential game model for power control in the D2D network described above is defined as follows:

Definition 1: The D2D differential power control game G_s for D2D transmitters is defined by a 5-tuple: $G_s = (\mathcal{N}, \{\mathcal{P}_i\}_{i\in\mathcal{N}}, \{\mathcal{S}_i\}_{i\in\mathcal{N}}, \{Q_i\}_{i\in\mathcal{N}}, \{c_i\}_{i\in\mathcal{N}})$, where

- **Player set** $\mathcal{N}$: $\mathcal{N} = \{1, ..., N\}$ represents the player set of densely-deployed D2D communication pairs. They are rational policy makers in the D2D power control differential game. The number of D2D links N is arbitrarily large.
- **Set of actions** $\{\mathcal{P}_i\}_{i\in\mathcal{N}}$: This is the set of possible transmit powers. Each transmitter determines the power $p_i(t) \in \{\mathcal{P}_i\}$ at any time $t \in [0, T]$ to minimize the cost function (to be defined later).
- **State space** $\{\mathcal{S}_i\}_{i\in\mathcal{N}}$: We define the state of player i as the combination of the interference introduced by the D2D transmitter i to other D2D links and the remaining energy at this D2D transmitter. Thus, we have two-dimensional states.
- **Control policy** $\{\mathcal{Q}_i\}_{i\in\mathcal{N}}$: A power control policy is denoted by $\mathcal{Q}_i(t)$, with $t \in [0, T]$, to minimize the average cost over the time interval T with two-dimensional states.
- **Cost function** $\{c_i\}_{i\in\mathcal{N}}$: We will define a novel cost function, where we consider both the achieved performance, e.g., the SINR at the receiver of a generic D2D link and the transmit power.

To determine the control policy $\{\mathcal{Q}_i\}_{i\in\mathcal{N}}$, we need to define the state space $\{S_i\}_{i\in\mathcal{N}}$ and the cost function $\{c_i\}_{i\in\mathcal{N}}$.

A. Two-Dimensional State Space

The state space is defined based on the intra-tier and inter-tier interferences in (1) and (2), respectively, and the energy usage dynamics.

1) Energy Usage Dynamics: The remaining energy state $E_i(t)$ of the player i at time t is equal to the amount of available energy. Meanwhile, $0 \leq E_i(t) \leq E_i(0)$, where $E_i(0)$ is the energy at time 0. The power control at time t should be any $p_i(t) \in [0, p_{\max}]$, where $p_{\max}$ is the maximum possible transmit power. Without loss of generality, we define the evolution law of the remaining energy in the battery as

$$dE_i(t) = -p_i(t)dt, \tag{5}$$

which means that energy $E_i(t)$ of the battery deceases with the transmit power consumption $p_i(t)$. At the same time, in game G_s, each player i should also consider the impact of interference on others.

2) Interference Dynamics: With intra-tier and inter-tier interference defined in (1) and (2), respectively, we first define the interference function that describes the interference caused by the generic D2D transmitter to others as follows:

$$\mu_i(t) = I_{i\to}(t) + I_{i\to o}(t), \tag{6}$$

where (6) describes all the interference introduced by player i to other D2D pairs $j \in \mathcal{N}, j \neq i$ and the only MUE o. According to definitions in (1) and (2), we have

$$\mu_i(t) = \sum_{j=1, j\neq i}^{N} p_i(t) g_{i,j}(t) + p_i(t) g_{i,o}(t). \tag{7}$$

To simplify the notation, we represent (7) as

$$\mu_i(t) = p_i(t)\varepsilon_i(t), \tag{8}$$

where $\varepsilon_i(t) = \sum_{j=1,j\neq i}^{N} g_{i,j}(t) + g_{i,o}(t)$. From (8), the total interference at time t to others depends on $p_i(t)$ and $\varepsilon_i(t)$ at time t. Therefore, we can define the interference state as

$$d\mu_i(t) = \varepsilon_i(t) dp_i(t) + p_i(t)\partial_t \varepsilon_i(t). \tag{9}$$

We will introduce the mean field approximation method to estimate the channel gains $\varepsilon_i, i \in \mathcal{N}$ in the following section.

We define the following state space for player i:

$$s_i(t) = [E_i(t), \mu_i(t)], \quad i \in \mathcal{N}, \tag{10}$$

where the interference caused by the generic D2D transmitter to other D2D links is regarded as one of the state variables.

The interference state $\mu_i(t)$ in (9) of the generic D2D transmitter will affect the strategy of the players of the MeNB and the other D2D receivers, and all others' interference $I_{\to i}(t)$ introduced to the generic receiver will affect the SINR performance. To distinguish between these two interferences, we denote them as State_1 (which is one of the state variables) and State_2, respectively. Note that the formulated cost minimization problem considers both SINR performance and cost due to transmit power. While the former is related to the interference caused to the generic D2D receiver from all the other transmitters, the latter is related to the interference caused by the generic D2D transmitter to others (i.e., the objective function of player i implicitly captures the effects of μ_i). Therefore, the objective function is affected by the considered interference states.

B. Cost Function

With the above definition of state space $s_i(t)$, each D2D transmitter i will determine the optimal power control policy $\mathcal{Q}_i^\star(t)$, with $t \in [0, T]$ to minimize the cost. The communication performance of the D2D pair i is characterized by the SINR $\gamma_i(t)$ defined in (4). We assume an identical SINR threshold γ^{th} for all D2D communication pairs. Meanwhile, each D2D communication pair prefers to minimizing the power consumption, and finally the cost function is given by

$$c_i(t) = \left(\gamma_i(\mathrm{t}) - \gamma^{th}(\mathrm{t})\right)^2 + \lambda p_i(t), \tag{11}$$

where λ is introduced to balance the units of the achieved SINR difference and the consumed power. It is easy to prove that the cost function $c_i(t)$, given by (11) is convex with respect to $p_i(t)$.

V. Optimal Control Problem and Analysis

With the cost function including both the achieved performance and transmit power, the problem is formulated as the cost minimization problem taking two kinds of interferences and the remaining energy into consideration.

A. Optimal Control Problem

We consider the problem that each D2D pair i will determine the optimal power control policy $\mathcal{Q}_i^\star(t)$, with $t \in [0,T]$ to minimize the cost function $c_i(t)$, given by (10), during a finite time horizon $[0,T]$. The general optimal control problem can be stated as follows:

$$\mathcal{Q}_i^\star(t) = \arg\min_{p_i(t)} \mathbb{E}\left[\int_0^T c_i(t)dt + c_i(T)\right], \tag{12}$$

where $c_i(T)$ is the cost at time T. At this time, we define the value function as follows:

$$u_i(t,s_i(t)) = \min_{p_i(t)} \mathbb{E}\left[\int_t^T c_i(t)dt + u_i(T,s_i(T))\right], t \in [0,T], \tag{13}$$

where $u_i(T,s_i(T))$ is a value at the final state $s_i(T)$ at time T. With the defined two-dimensional states, we have the following lemma 1.

Lemma 1: A power control profile $Q_i^\star(t) = p_i^\star(t)$, for $i \in \mathcal{N}$ is the Nash equilibrium solution of G_s if and only if [31]

$$Q_i^\star(t) = \arg\min_{p_i(t)} \mathbb{E}\left[\int_0^T c_i(p_i(t),\boldsymbol{p}_{-i}^\star)dt + c_i(T)\right]$$

subject to:

$$\begin{aligned} &s_i(t) = [\mu_i(t), E_i(t)], \\ &dE_i(t) = -p_i(t)dt, \ \ and \\ &d\mu_i(t) = \varepsilon_i(t)dp_i(t) + p_i(t)\partial_t\varepsilon_i(t), \end{aligned}$$

where $\boldsymbol{p}_{-i}^\star$ denotes the transmit power vector of D2D links except D2D$_i$. None of the players can have a lower cost by unilaterally deviating from the current power control policy. The Nash equilibrium of the above power control differential game can be obtained by solving the HJB equation associated with each player in the optimal control theory. We will derive the HJB equation later.

Proof: The conclusion is indirectly guaranteed by the smoothness of the HJB function. Therefore, we first derive the HJB equation, and then we prove that the smoothness of the derived HJB equation guarantees the existence of Nash equilibrium in the formulated differential game. The details can be found in Lemma 2 and Theorem 1 presented below. ■

B. Analysis

According to optimal control theory followed by Bellman's optimality principle [26], the value function in (13) should satisfy a partial differential equation which is a HJB equation. The solution of the HJB equation is the value function, which gives the minimum cost for a given dynamic system with an associated cost function.

Lemma 2: We have the HJB equation in (14)

$$-\partial_t u_i(t,s_i(t)) = \min_{p_i(t)} \left[c_i(t,s_i(t),p_i(t)) + \partial_t s_i(t) \cdot \nabla u_i(t,s_i(t))\right], \tag{14}$$

where we define the Hamiltonian as in (15).

Proof: Proof of Lemma 2 is given in **Appendix A**. ■

The Nash equilibrium of the above differential game can be obtained by solving the HJB equation associated with each player given in (37). We have following theorem on the existence of the Nash equilibrium for the defined G_s in the definition 1.

Theorem 1: There exists at least one Nash equilibrium for the differential game G_s.

Proof: Existence of a solution to the HJB equation ensures the existence of the Nash equilibrium for the game G_s. It is known that there exists a solution to the HJB equation if the Hamiltonian is smooth [31]. Derivatives of all the orders exist for the Hamiltonian due to the continuity of the defined cost function, and it is easy to derive the derivatives of the Hamiltonian with respect to $p_i(t)$. Due to the existence of the derivatives, the Hamiltonian is smooth. This concludes the proof. ■

Obtaining the equilibrium for game G_s for a system with N players involves solving N partial differential equations simultaneously. However, for a dense D2D network, it is not possible to obtain the Nash equilibrium in this manner due to the large number of simultaneous partial differential equations. Therefore, for modeling and analysis of a dense D2D network, a mean field game will be introduced where the system can be defined solely by two coupled equations. In the next section, we show the extension of game G_s to a mean field game.

VI. Mean Field Game (MFG) for Power Control

The power control MFG in D2D networks is a special form of a differential game described before when the number of D2D links approaches infinity. The power control MFG can be expressed as a coupled system of two equations of HJB and FPK. On one hand, an FPK type equation evolves forward in time that governs the evolution of the density function of the agents. On the other hand, an HJB type equation evolves backward in time that governs the computation of the optimal path for each agent.

A. Mean Field and Mean Field Approximation

For the effects of both the interference of the generic D2D transmitter introduced to others, and all others' interference introduced to the generic link, we first introduce the mean field concept, and then propose the mean-field approximation approach.

$$H(p_i(t), s_i(t), \nabla u_i(t, s_i(t))) = \min_{p_i(t)} \left[c_i(t, s_i(t), p_i(t)) + \partial_t s_i(t) \cdot \nabla u_i(t, s_i(t)) \right]. \tag{15}$$

1) Mean Field: This is a critical concept in the defined power control MFG, which is a statistical distribution of the defined two-dimensional states.

Definition 2: Given the state space $s_i(t) = [\mu_i(t), E_i(t)]$ in the defined power control MFG, we define the mean field $m(t,s)$ as

$$m(t,s) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{s_i(t)=s\}}, \tag{16}$$

where $\mathbb{1}$ denotes an indicator function which returns 1 if the given condition is true and zero, otherwise. For a given time instant, the mean field is the probability distribution of the states over the set of players.

With the defined mean field, the power control differential game can be formulated as the new power control MFG, if the following requirements are satisfied. In the context of an ultra-dense D2D network, the D2D transmitters, which act as players, have the following properties:

- **Rationality:** Each D2D transmitter can individually take rational power control decision to minimize the cost function.
- **The existence of a continuum of the players (i.e., continuity of the mean field):** The presence of a large number of D2D pairs in the defined power control game ensures the existence of the continuum of the players.
- **Interchangeability of the states among the players (i.e., permutation of the states among the players would not affect the outcome of the game):** We derive the cost function via mean-field approximation of the interference in order to ensure the interchangeability of the actions among the players.
- **Interaction of the players with the mean field:** Each D2D player interacts with the mean field instead of interacting with all the other players.

The mean field defined above will be obtained based on *mean field approximation* described below.

2) Mean Field Approximation: With respect to the interference summations $I_{i\to}(t)$ and $I_{\to i}(t)$ defined in (1) and (3), respectively, we assume that each interferer is infinitesimal and contributes little interference power to the summation. The infinite mass of other players, which is called as the mean-field value, dominates the interference effects when studying a typical player. We first derive the interference mean field via a special technique of the mean field approximation [29]. For both State$_1$ $I_{i\to}(t)$ and State$_2$ $I_{\to i}(t)$ defined in (1) and (3), we can use the same method. Here, we use State$_2$ $I_{\to i}(t)$ as an example. We know that

$$\hat{I}_{\to i}(t) = \sum_{j=1, j\neq i}^{N} p_j(t) g_{j,i}(t) \approx (N-1)\, \hat{p}_j(t) \hat{g}_{j,i}(t), \tag{17}$$

where $\hat{p}_i(t)$ is the known test transmit power and we assume that all the players involved in the game are using the same test transmit power. The term $\hat{g}_{i,j}(t)$ defines the mean interference channel gain of ultra-dense infinitesimal D2D effects, which can be estimated by the following idea. If we use $\hat{p}_i(t)$ at the transmit power for D2D pair i's transmitter, then the power received at the corresponding receiver is

$$p_i^R(t) = \hat{p}_i(t) g_{i,i}(t) + \hat{I}_{\to i}(t), \tag{18}$$

where $g_{i,i}(t)$ is the effective channel gain, and $\hat{p}_i(t) g_{i,i}(t)$ is the effective received power, and $\hat{I}_{\to i}(t)$ is the received interference power from all the others.

Combining (17) and (18), we can derive the only unknown variable $\hat{g}_{j,i}(t)$ as

$$p_i^R(t) = \hat{p}_i(t) g_{i,i}(t) + (N-1)\, \hat{p}_j(t) \hat{g}_{j,i}(t), \tag{19}$$

and we have

$$\hat{g}_{j,i}(t) = \frac{p_i^R(t) - \hat{p}_i(t) g_{i,i}(t)}{(N-1)\, \hat{p}_j(t)}. \tag{20}$$

With the above mean field approximation,

$$\hat{\gamma}_i(\mathrm{t}) = \frac{p_i(t) g_{i,i}(t)}{(N-1)\, \hat{p}_j(t) \hat{g}_{j,i}(t) + \sigma^2}, \tag{21}$$

with $\hat{g}_{j,i}(t)$ estimated by the above method during the practical implementation. The mean field cost function is then defined as

$$\hat{c}_i(t) = \left(\hat{\gamma}_i(\mathrm{t}) - \gamma^{th}(\mathrm{t})\right)^2 + \lambda p_i(t). \tag{22}$$

In the following sections, we will use the mean field cost function and the approximated interference mean field. We can similarly achieve the mean field approximation of State$_1$ $I_{i\to}(t)$.

B. Mean Field Game

Based on the above definition of mean field and mean field approximation, we derive the FPK equation first.

Lemma 3: The FPK equation of the defined MFG is given by

$$\partial_t m(t,s) + \nabla \left(m(t,s) \cdot \partial_t s(t) \right) = 0, \tag{23}$$

where s is the state.

Proof: Proof of Lemma 3 can be found in [24]. ■

The FPK equation describes the evolution of the defined mean field with respect to time and space. At this time, with the derived HJB and FPK equations, we formulate the D2D MFG as follows.

Definition 3: The D2D MFG is defined as the combination of derived HJB and FPK equations, where the HJB equation is

$$-\partial_t u(t, s(t)) = \min_{p(t)} \left[c(t, s(t), p(t)) + \partial_t s(t) \cdot \nabla u(t, s(t)) \right], \tag{24}$$

and the FPK equation is in (23).

Their interactions with each other are shown in Fig. 2. The HJB equation governs the computation of the optimal

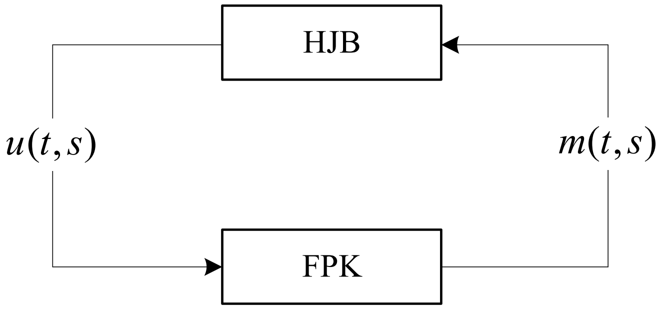

Fig. 2. D2D mean field game with HJB and FPK equations.

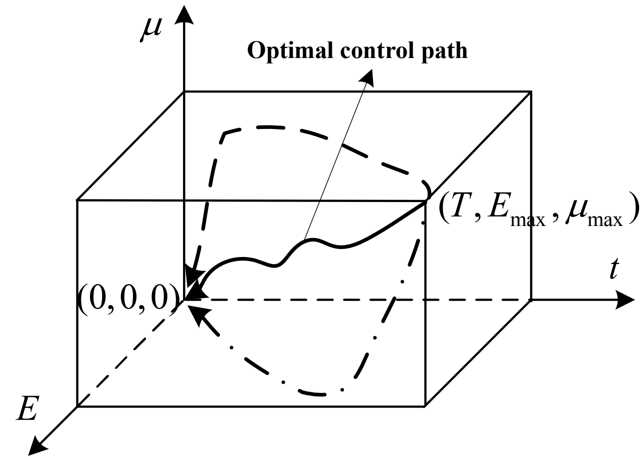

Fig. 3. Optimal control of the discretized mean field game in time and space, where there exist several potential control paths. Here, we should design the method to find the optimal control path.

control path of the player, while the FPK equation governs the evolution of the mean field function of players. Here, the HJB and FPK equations are termed as the backward and forward functions, respectively. Backward means that the final value of the function is known, and we determine the value of $u(t)$ at time $[0, T]$. Therefore, the HJB equation is always solved backwards in time, starting from $t = T$, and ending at $t = 0$. When solved over the the entire state space, the HJB equation is a necessary and sufficient condition for the optimum. The FPK equation evolves forward with time. The interactive evolution finally leads to the mean field equilibrium.

C. Mean Field Equilibrium

We define the mean field equilibrium, which can be achieved by using the finite difference method [31].

Definition 4: Mean field equilibrium (MFE) represents stable combination of both the control policy $u^{\star}(t, s)$ and the mean field $m^{\star}(t, s)$ at any time t and state s in Fig. 2.

At any time t and state s, the control policy $u(t, s)$ and the mean field $m(t, s)$ interact with each other, where $u(t, s)$ is also termed as the value function. The value function $u(t, s)$ determines the control policy $Q(t, s)$. The term $u(t, s)$ is the solution of the HJB equation in (24) and $m(t, s)$ is the solution of the FPK equation in (23). The term $u(t, s)$ affects the evolution of the mean field, and $m(t, s)$ determines the decision-making of optimal policy $u(t, s)$.

VII. Distributed Social-Optimal Power Control Policy Based on the Finite Difference Method

Similar to [31], we resort to the finite difference method. We have three different schemes to discretize the advection equation including Upwind, Lax-Friedrichs, and Lax-Wendroff [30] schemes. Different schemes have different rates of convergence. We select the Lax-Friedrichs scheme to guarantee the positivity of the mean field with first-order accuracy in both time and space [31].

In the framework of the finite difference method, the investigated time interval $[0, T]$, the energy state space $[0, E_{\max}]$, and the interference state space $[0, \mu_{\max}]$ should be discretized into $X \times Y \times Z$ spaces, as shown in Fig. 3. In Fig. 3, we illustrate three curves in the three-dimensional discretized time and space. The optimal control policy covers the decision-making over a period in the discretized mean field game framework. Therefore, there exist several potential control paths. Here, we should design the method to find the optimal control path as indicated in the figure. Thus, we first define the iteration steps of time, energy, and interference spaces as

$$\delta_t = \frac{T}{X}, \delta_E = \frac{E_{\max}}{Y}, \delta_\mu = \frac{\mu_{\max}}{Z}.$$

Therefore, the FPK equation will evolve in the space of three dimensions of $(0, T) \times (0, E_{\max}) \times (0, \mu_{\max})$ with the steps of δ_t, δ_E, and δ_μ, respectively. At the same time, the optimal control path is achieved by solving the HJB equation backward accordingly. A joint finite difference algorithm is then developed based on the Lax-Friedrichs scheme and Lagrange relaxation to solve the corresponding coupled HJB and FPK equations to obtain the optimal power control policy (e.g., the optimal control path as shown in Fig. 3).

A. Lax-Friedrichs Scheme to Solve FPK Equation

We solve the FPK equation using the Lax-Friedrichs method, where we first solve the FPK equation

$$\partial_t m(t, s) + \nabla_E m(t, s) E(t) + \nabla_\mu m(t, s) \mu(t) = 0. \quad (25)$$

By applying the Lax-Friedrichs scheme, we have (26).

Here $M(i, j, k)$, $P(i, j, k)$, $\varepsilon(i, j, k)$ are the values of the mean field, the power, and the interference gain at time instant i with the energy level j and the interference state k in the discretized grid.

B. Discretized Lagrange Relaxation to HJB

The finite difference method cannot be used to solve the HJB equation due to the Hamiltonian. Here, we reformulate the HJB equation using its corresponding optimal control problem, where the newly-defined problem is

$$\begin{aligned} &\min_{p_i(t)} \mathbb{E}\left[\int_0^T c_i(t)dt + c_i(T)\right], \\ &\text{subject to}: \quad \partial_t m(t, s) + \nabla_E m(t, s) E(t) + \nabla_\mu m(t, s) \mu(t) = 0. \end{aligned} \quad (27)$$

$$
\begin{aligned}
M\left(i+1,j,k\right) = & \frac{1}{2}\left[M\left(i,j-1,k\right)+M\left(i,j+1,k\right)+M\left(i,j,k-1\right)+M\left(i,j,k+1\right)\right] \\
& + \frac{\delta_t}{2\delta_E}\left[M\left(i,j+1,k\right)P\left(i,j+1,k\right)-M\left(i,j-1,k\right)P\left(i,j-1,k\right)\right] \\
& + \frac{\delta_t}{2\delta_\mu}\left[M\left(i,j,k+1\right)P\left(i,j,k+1\right)\varepsilon\left(i,j,k+1\right)-M\left(i,j,k-1\right)\varepsilon\left(i,j,k-1\right)P\left(i,j,k-1\right)\right],
\end{aligned} \tag{26}
$$

$$
\begin{aligned}
& L\left(m(t,s),p(t,s),\lambda(t,s)\right) \\
& = \mathbb{E}\left[\int_0^T c_i(t)dt + c_i(T)\right] + \int\limits_{t=0}^{T}\int\limits_{E=0}^{E_{\max}}\int\limits_{\mu=0}^{\mu_{\max}} \lambda(t,s)\left(\partial_t m(t,s)+\nabla_E m(t,s)E(t)+\nabla_\mu m(t,s)\mu(t)\right)dtdEd\mu \\
& = \int\limits_{t=0}^{T}\int\limits_{E=0}^{E_{\max}}\int\limits_{\mu=0}^{\mu_{\max}} \left[c(t,s)m(t,s)+\lambda(t,s)\left(\partial_t m(t,s)+\nabla_E m(t,s)E(t)+\nabla_\mu m(t,s)\mu(t)\right)\right]dtdEd\mu.
\end{aligned} \tag{28}
$$

At this time, we obtain the Lagrangian $L\left(m(t,s),p(t,s),\lambda(t,s)\right)$ by introducing a Lagrange multiplier $\lambda(t,s)$ as (28), where we assume that $c(T)=0$.

Similar to the method to solve the FPK equation, we propose a finite difference method to solve (28). We first discretize the Lagrangian to solve the newly-defined optimal control problem, and the discretized Lagrangian is given by (29). Here Υ, Φ, and Ψ are given by (30), (31), and (32), respectively.

Here, $M(i,j,k)$, $P(i,j,k)$, $\lambda(i,j,k)$, and $C(i,j,k)$ are the values of the mean field, the power, Lagrange multiplier, and cost function at time instant i, energy level j, and interference state k in the discretized grid.

The optimal decision variables include $P^\star$, $M^\star$, and $\lambda^\star$. We derive the optimal power control as $\frac{\partial L_d}{\partial P(i,j,k)}=0$ for any arbitrary point (i,j,k) in the discretized grid, where $\frac{\partial L_d}{\partial P(i,j,k)}$ is given by (33).

Furthermore, for any arbitrary point (i,j,k) in the discretized grid, we update the Lagrange multiplier $\lambda(i,j,k)$ by calculating $\frac{\partial L_d}{\partial M(i,j,k)}=0$, and then we have $\lambda\left(i-1,j,k\right)$ given by (34).

Algorithm 1: Distributed social-optimal power control policy

```
Initialization:
M(0,0,0):= joint mean field distribution;
λ(X+1,0,0):= initial Lagrangian parameters;
P(X+1,0,0):= initial power levels.
for i = 1 : X, j = 1 : Y, and k = 1 : Z do
    Update mean field:
    Using M(i+1,j,k) (26)
    if P(i,j+1,k)=0 then
        M(i+1,j+1,k+1) = M(i,j,k)
    else
        M(i+1,j+1,k+1) = 0
    end
    Update Lagrangian parameter:
    λ(i-1,j,k) using (34)
    Update power levels:
    P(i-1,j,k) using (33)
end
```

C. Distributed Social-Optimal Power Control Policy

Following the above derivations, a joint finite difference algorithm based on the Lax-Friedrichs scheme and Lagrange relaxation is proposed to solve the coupled HJB and FPK equations, respectively. We name it as the distributed social-optimal power control policy.

For the proposed distributed algorithm, we have the following comments:

- First, the mean filed is jointly influenced by the energy dynamics and interference dynamics (State_1). Basically, the energy dynamic function is a linear function with respect to the transmit power. However, interference function is not linear. At each time step, we assume that the interfering link gains estimated by the mean field approximation approach do not change.
- Second, we choose the algorithm termination condition as the difference between the final-two mean field values, and we set the gap as 10^{-5}.
- Third, during iterations $i-1$ and $i+1$ and similarly for other indices in the proposed algorithm, we introduce a simple computation method. For instance, $i-1$ may not be positive, when $i \leq 1$. In this situation, we set $\lambda(i-1,j,k)=\frac{1}{2}[\lambda(i,j+1,k)+\lambda(i,j-1,k)]$.
- Our algorithm jointly considers the $E_{\max}$ and the tolerable interference $\mu_{\max}$ [2]. Therefore, it is easy to extend as the other schemes, for instance, it can be termed as the 'blind' scheme when $E_{\max}\to+\infty$ and $\mu_{\max}\to 0$. For the cases of $\mu_{\max}\to 0$ and $E_{\max}\to+\infty$, we make the following assumptions during simulations. We assume that if the maximum energy is set more than ten

[2]Tolerable interference $\mu_{\max}$ is a predefined value, which constrains the total interference power that can be tolerated.

$$L_d = \delta_t \delta_E \delta_\mu \sum_{i=1}^{X+1} \sum_{j=1}^{Y+1} \sum_{k=1}^{Z+1} \left[\ M\left(i,j,k\right) C\left(i,j,k\right) + \lambda(i,j,k)(\Upsilon + \Phi + \Psi) \ \right], \tag{29}$$

$$\Upsilon = \frac{1}{\delta_t} \left[M\left(i+1,j,k\right) - \frac{1}{2}\left(M\left(i,j+1,k\right) + M\left(i,j-1,k\right) + M\left(i,j,k+1\right) + M\left(i,j,k-1\right) \right) \right], \tag{30}$$

$$\Phi = \frac{1}{2\delta_\mu} \left[M\left(i,j,k+1\right) P\left(i,j,k+1\right) \varepsilon\left(i,j,k+1\right) - M\left(i,j,k-1\right) P\left(i,j,k-1\right) \varepsilon\left(i,j,k-1\right) \right], \tag{31}$$

$$\Psi = \frac{1}{2\delta_E} \left[M\left(i,j+1,k\right) P\left(i,j+1,k\right) - M\left(i,j-1,k\right) P\left(i,j-1,k\right) \right]. \tag{32}$$

times higher than that of practical situation, then it can be regarded as the case of $E_{\max} \to +\infty$. While the normal setting is $E_{\max} = 0.5$, the case of $E_{\max} = 5$ can be seen as the case of $E_{\max} \to +\infty$.

VIII. Numerical Results

In this section, we first illustrate the simulation scenarios and the basic settings with relevant simulation parameters. We characterize the mean field distributions and the power control policy using the Matlab software.

A. Basic Simulation Settings

The downlink transmission of an OFDMA D2D network, with the radius of D2D links uniformly distributed between 10 m to 30 m is considered. We set the system parameters as the bandwidth w = 20 MHz, and background noise power is 2×10^{-9} W as the noise power spectral density is κ = -174 dBm/Hz. Without special instructions, we choose the standardized case with 500 LTE frames, the maximum energy is 0.5 J, the number of D2D links will vary from N = 50 to N = 200. The path-loss exponent for D2D links is 3. The duration of one LTE radio frame is 10 ms, and for 500 frames, $T = 5$ s. We also pick, $E_{\max}$ to be 0.5 J. The tolerable interference level of each player $\mu_{\max}$ is assumed to be 5.8×10^{-6} W.

B. Characteristics of Mean Field Distributions and Power Control Policy

To illustrate the properties of our proposed power control scheme, we show the distributions of the mean field and the power control policy in different cases.

First, we describe three dimensional mean field distributions and power policy, as shown in Fig. 4. However, mean field and power policy are four dimensional vectors. Therefore, we plot mean field and power policy for three cases, i.e., (a) mean field distributions with varying interference and energy but fixed time; (b) mean field distributions with varying time and energy but fixed interference; and (c) mean field distributions with varying time and interference but fixed energy. Similarly, we illustrate the power control policies for these three cases. Basically, we can see that both the remaining energy and the interference dynamics affect the mean field distributions and power control policy.

Second, to further demonstrate the properties of our proposed scheme, the cross sections of mean field distributions and power policy at the energy states after convergence but with varying times are depicted in Fig. 5. Here, as in the previous settings, we discretize the time interval, the energy space, and the interference space into $20 \times 20 \times 20$ grids. To reflect the reasons why the mean field distributions and power policy in Fig. 4 are random, we describe Fig. 5, where we first show the distributions of the mean field with respect to the interference space at the energy states after convergence. In addition, we select three cases: (a) distributions of mean field with respect to the varying interference at time interval 13; (b) distributions of mean field with respect to the varying interference at time interval 15; and (c) distributions of mean field with respect to the varying interference at time interval 20. Meanwhile, we depict the distributions of power policy accordingly in (d), (e), and (f), respectively.

From Fig. 5, we have the following observations. On one hand, the randomness of the distributions of mean fields and powers are introduced by the random interference space. On the other hand, our proposed scheme can achieve the power equilibrium as shown in Fig. 5, which is the final power control policy at the final energy and time interval states after convergence. Also, our proposed scheme can always achieve the equilibrium powers, regardless of the interference states.

In Fig. 6, we illustrate the mean field distributions and power control policy with respect to time toward the convergence. We can see that we can always achieve the target SINR with the varying power control policy according to the mean field distributions.

C. Spectrum and Energy Efficiency Performance

- **Spectrum efficiency:** The average spectrum efficiency π_{average} is in the unit of bps/Hz, which is calculated by

  $$\pi_{\text{average}} = \log_2\left(1 + \frac{p}{I + w \times \kappa}\right),$$

  where κ is the noise power density and w is the total bandwidth.

$$
\begin{aligned}
\frac{\partial L_d}{\partial P(i,j,k)} &= \sum_{j=1}^{Y+1}\sum_{k=1}^{Z+1} M\left(i,j,k\right)\frac{\partial C\left(i,j,k\right)}{\partial P\left(i,j,k\right)} \\
&+ \left[\frac{M\left(i,j,k\right)}{2\delta_E} + \frac{M(i,j,k)\varepsilon(i,j,k)}{2\delta_\mu}\right]\left[\lambda\left(i,j+1,k\right) - \lambda\left(i,j-1,k\right)\right].
\end{aligned} \tag{33}
$$

$$
\begin{aligned}
\lambda\left(i-1,j,k\right) = &\frac{1}{2}\left[\lambda\left(i,j+1,k\right) + \lambda\left(i,j-1,k\right)\right] + \frac{1}{2}\left[\lambda\left(i,j,k+1\right) + \lambda\left(i,j,k+1\right)\right] \\
&- \frac{1}{2}\delta_t P\left(i,j,k\right)\left[\frac{\varepsilon\left(i,j,k\right)}{\delta_\mu} + \frac{1}{\delta_E}\right]\left[\lambda\left(i,j+1,k\right) - \lambda\left(i,j-1,k\right)\right] \\
&+ \delta_t C\left(i,j,k\right).
\end{aligned} \tag{34}
$$

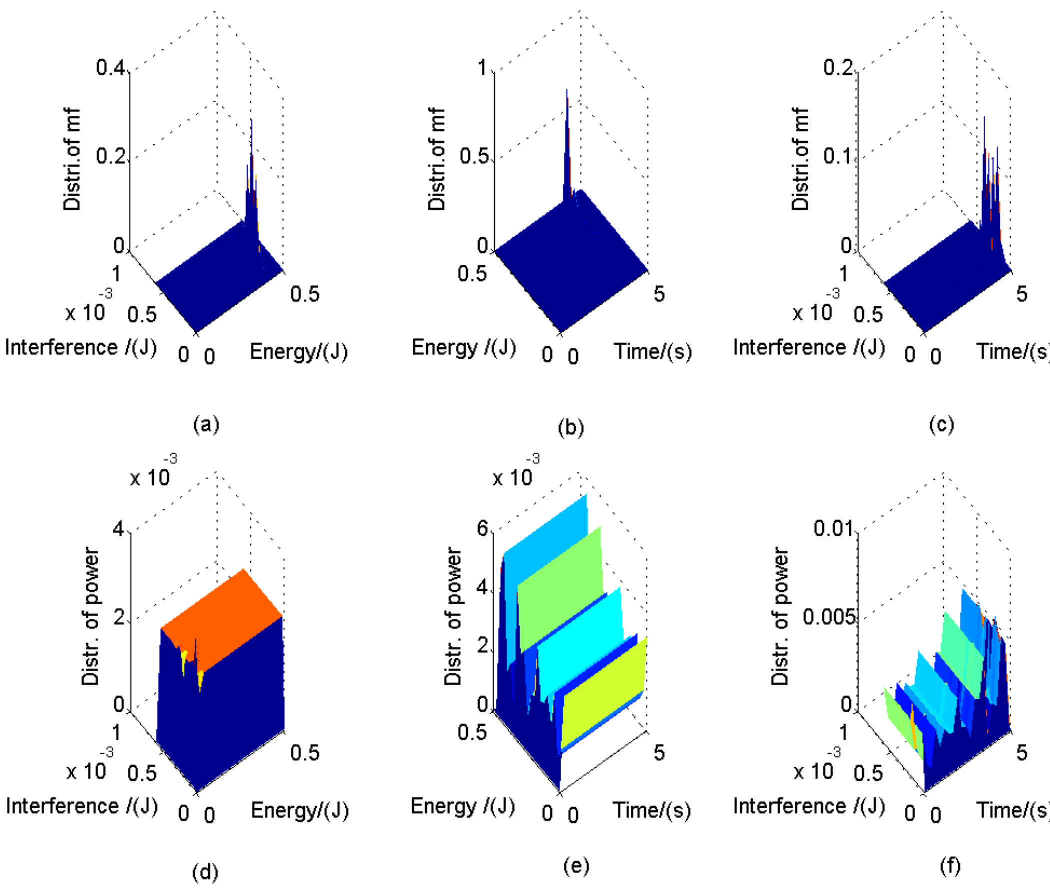

Fig. 4. Three dimensional mean field distributions and power policy varying in time and space. For the distributions of mean field we have three cases with (a) $T = 5$; (b) $\mu_{\max} = 5.8 \times 10^{-6}$; (c) $E_{\max} = 0.1$ fixed, respectively. Similarly, for the distributions of powers we also have three cases with (d) $T = 5$; (e) $\mu_{\max} = 5.8 \times 10^{-6}$; and (f) $E_{\max} = 0.1$, respectively.

- **Energy efficiency:** The average energy efficiency η_{average} is defined as the ratio between the total throughput and the consumed energy, which is given by

$$\eta_{\text{average}} = \frac{w\pi_{\text{average}}}{p_{total}},$$

which is in the unit of bit/J.

To illustrate the impacts of various interference constraints, we simulate two other cases as benchmarks of the normal settings in our work: (1) $\mu_{\max} \to +\infty$, which means that the interference tolerance is very large; (2) $\mu_{\max} \to 0$, which means that the interference tolerance is very small. Meanwhile, to illustrate the impacts of various energy constraints, in addition to the normal settings, we simulate the case of $E_{\max} \to +\infty$, which means that the battery is always fully powered.

We set $\mu_{\max} = 5.8 \times 10^{-6}$ as the normal setting. Moreover, the case $\mu_{\max} \leq 10^{-5}\times$ normal setting is considered as the case of $\mu_{\max} \to 0$. The case $\mu_{\max} \geq 10^{2}\times$ normal setting is considered as the case of $\mu_{\max} \to +\infty$. Here, we set the normal setting is $E_{\max} = 0.5$, while the case of $E_{\max} = 5$ can be seen as the case of $E_{\max} \to +\infty$.

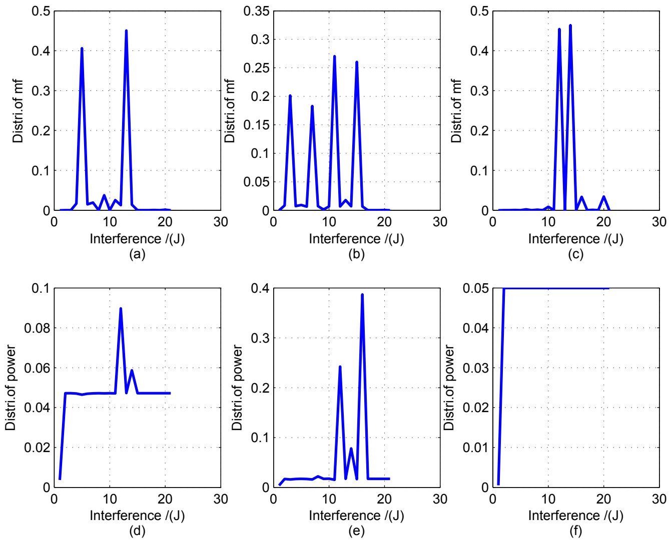

Fig. 5. Cross sections of mean field distributions and power policy at the states after convergence but with varying times. For the distributions of mean field we have three cases at different time slots (a) $T = 3.25s$; (b) $T = 3.75s$; (c) $T = 5s$, respectively. Similarly, for the distributions of powers we also have three cases with at different time slots (d) $T = 3.25s$; (e) $T = 3.75s$; (f) $T = 5s$, respectively.

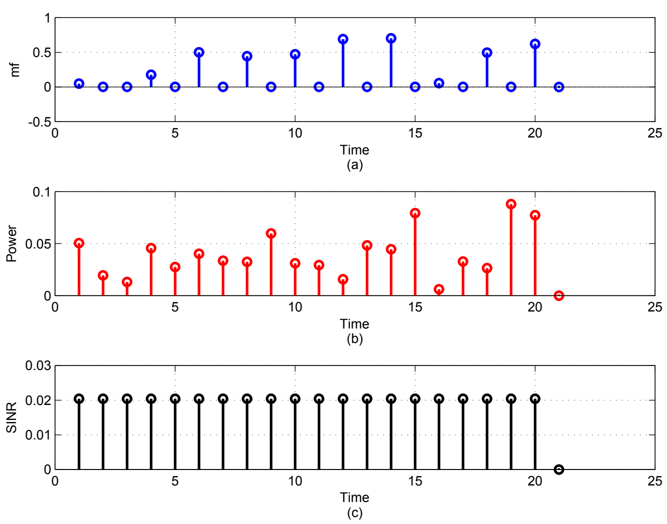

Fig. 6. Cross-sections of mean field distributions and power policy at the interference $\mu_{\max} = 5.8 \times 10^{-6}$ and energy state $E_{\max} = 0.5J$ after convergence.

We illustrate the spectrum efficiency π performance of the two cases in Fig. 7(a) and Fig. 7(b), respectively. We observe that:

- The spectrum efficiency decreases with the increasing numbers of the D2D links. This is due to the increased mutual interference with increasing number of D2D links.
- Minimizing the tolerable interference can always improve the spectrum efficiency; however, maximizing the available energy, i.e., $E_{\max}$ does not always improve the spectrum efficiency. Minimizing the tolerable interference means that the power used at each state is not allowed to be too large, which leads to the reduced SINR. Thus, maximizing the tolerable interference means that the power used at each state can be large, therefore, the spectrum efficiency is improved.
- Maximizing the available energy can first help maximizing the effective received energy; however, at some points increasing the energy also means increasing the interference power, and consequently, decreases the spectrum efficiency.

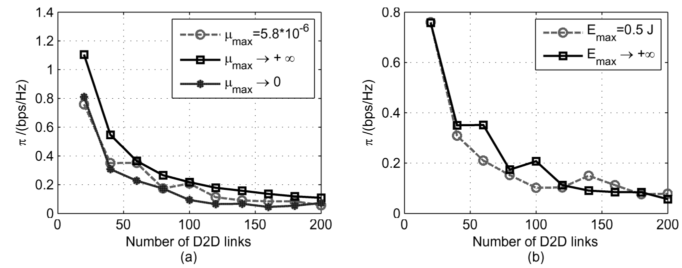

Fig. 7. Spectrum efficiency, which is measured in the unit of bps/Hz.

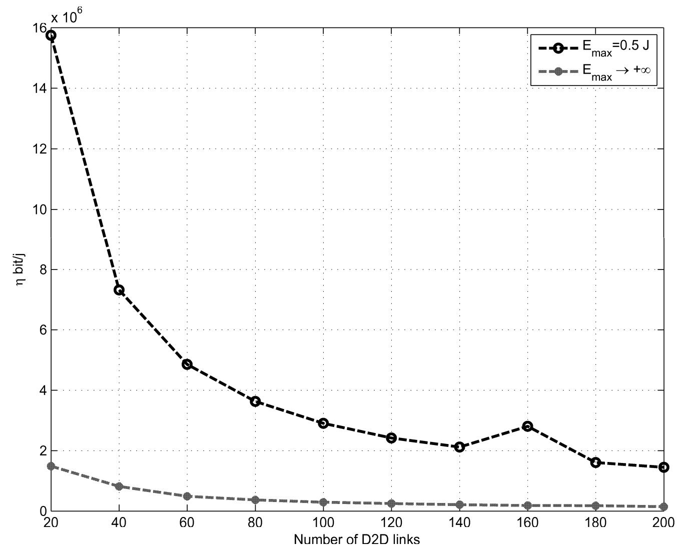

Fig. 8. Energy efficiency, which is measured in the unit of b/J.

In Fig. 8, we illustrate the improved energy efficiency performance of the proposed interference and energy-aware power control compared to the blind power control scheme. Here, we term the blind power control scheme as the case of $E_{\max} \to +\infty$, which can be seen as the scheme without the energy constraint. Therefore, D2D transmitter in the blind power control scheme can individually increase the transmission power, which will both introduce the interference to others and exhaust the battery soon.

It can be concluded from Fig. 8 that the proposed interference and energy-aware power control can achieve higher energy efficiency compared with the blind power control scheme. Meanwhile, the benefits decrease with the increasing number of D2D links, which is mainly because that the increasing number of D2D links means introducing more interference, and thus decreases the spectrum efficiency.

IX. Conclusion

We have investigated the effects of both interference and energy on the power control in ultra-dense D2D communications. We have formulated a mean filed game theoretic framework with the two-dimensional dynamics states. For the MFG framework, we have derived the coupled HJB and FPK equations. Then, a joint finite difference algorithm has been used to solve the coupled HJB and FPK equations of the corresponding MFG, which is based on the Lax-Friedrichs scheme and the Lagrange relaxation method. Then we have developed an interference and energy-aware distributed power control. Numerical results have been presented to characterize the mean field distributions and the power control policy. The spectrum and energy efficiency performances of the proposed distributed power control scheme have been illustrated. The proposed model can be implemented in practical systems by solving the MFG model offline using the measured (historical) information about interference dynamics as well as the energy availability at the terminals in a dense D2D network. The obtained power control policies can be stored in look-up tables to be used by the D2D transmitters.

Appendix A
Derivation of HJB

Intuitively, an HJB can be derived as follows. If $u_i(t)$ is the value function of power $p_i(t)$ and the state $s_i(t)$, and then by the Richard Bellman's principle of optimality, increasing time t to $t + dt$, we have (35). Further, we compute the Taylor expansion of $u_i(t + dt)$, and we have (36). Here $\partial_t u$ is the differential function with t, and ∇u is the gradient of the function u with s. $o(dt)$ denotes the terms of the Taylor expansion of higher order than one, and we omit it during the following analysis.

Substituting (36) into (35), and canceling $u_i(t, s_i(t))$ on both sides, dividing by dt, and taking the limit with dt approaches zero, we have the following HJB equation (37). Here, we define the Hamiltonian as (38).

References

[1] Z. Zhou, M. Dong, K. Ota, J. Wu, and T. Sato, "Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications," *IEEE Wireless Communications Letters*, vol. 3, no. 5, pp. 485–488, Oct. 2014.

[2] L. Song, D. Niyato, Z. Han, and E. Hossain, *Wireless device-to-device communications and networks,* Cambridge University Press, 2015.

[3] Z. Zhou, M. Dong, K. Ota, and Z. Chang, "Energy-efficient context-aware matching for resource allocation in ultra-dense small cells," *IEEE Access*, vol. 3, pp. 1849–1860, Sep. 2015.

[4] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[5] B. Soret, K. Pedersen, Niels T. K. Jrgensen, and F. L. Vctor, "Interference coordination for dense wireless network," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 102–109, Jan. 2015.

[6] H. Zhang, C. Jiang, J. Cheng, and V. C. M. Leung, "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 92–99, Jun. 2015.

[7] C. Yang, J. Li, and M. Guizani, "Cooperation for spectral and energy efficiencies in ultra-dense heterogeneous and small cell networks," *IEEE Wireless Communication Magazine*, vol. 23, no. 1, pp. 64–71, Feb. 2016.

[8] F. Wang, C. Xu, L. Song, and Z. Han, "Energy-efficient resource allocation for device-to-device underlay communication," *IEEE Transactions on Wireless Communications*, vol.14, no. 4, pp. 2082–2092, Apr. 2015.

[9] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. K. Wong, and X. Chen, "Joint spectrum and power allocation for D2D communications underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2182–2195, Apr. 2016.

[10] C.-H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," *IEEE International Conference on Communications*, pp. 1–5, Jun. 2009.

[11] P. Janis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-aware resource allocation for device-to-device radio underlaying cellular networks," *IEEE 69th Vehicular Technology Conference*, Barcelona, Spain, Apr. 2009.

[12] G. Fodor and N. Reider, "A distributed power control scheme for cellular network assisted D2D communications," *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston TX, Dec. 2011.

[13] P. Phunchongharn, E. Hossain, and D. I. Kim, "Resource allocation for device-to-device communications underlaying LTE-advanced networks," *IEEE Wireless Communication Magazine*, vol. 20, no. 4, pp. 91–100, Aug. 2013.

[14] Z. Zhou, M. Dong, K. Ota, G. Wang and L. T. Yang, "Energy-efficient resource allocation for D2D communications underlaying cloud-RAN-based LTE-A networks," *IEEE Internet of Things Journal*, vol. 3, no. 3, pp. 428–438, June 2016.

[15] S. Maghsudi and S. Stanczak, "Hybrid centralized-distributed resource allocation for device-to-device communication underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2481–2495, April 2016.

[16] K. Zhu and E. Hossain, "Joint mode selection and spectrum partitioning for device-to-device communication: a dynamic Stackelberg game," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1406–1420, Mar. 2015.

[17] J. Huang, Y. Zhao, and K. Sohraby, "A game-theoretic resource allocation approach for intercell device-to-device communications in cellular networks," *23rd International Conference on Computer Communication and Networks (ICCCN)*, Shanghai, China, Aug. 2014.

[18] R. Yin, G. Yu, H. Zhang, Z. Zhang, and G. Y. Li, "Pricing-based interference coordination for D2D communications in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3 pp. 1519–1532, Mar. 2015.

[19] L. Song, D. Niyato, Z Han, and E. Hossain, "Game-theoretic resource allocation methods for device-to-device communication," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 136–144, Jun. 2014.

[20] Y. Shen, C. Jiang, T. Q. S. Quek, H. Zhang, and Y. Ren, "Device-to-device cluster assisted downlink video sharing-a base station energy saving approach," *2014 IEEE Global Conference on Signal and Information Processing*, Atlanta, GA, Dec. 2014.

[21] S. M. Azimi, M. H. Manshaei, and F. Hendessi, "Hybrid cellular and device-to-device communication power control-Nash bargaining game," *2014 7th International Symposium on Telecommunications*, pp. 1077–1081, Tehran, Sep. 2014.

[22] Y. Xiao, K. 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 Transactions on Wireless Communications*, vol. 14, no. 7, pp. 4034–4051, July 2015.

[23] X. Lu, P. Wang, and D. Niyato, "Hierarchical cooperation for operator-controlled device-to-device communications: a layered coalitional game approach," *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2056–2061, New Orleans, LA, May 2015.

[24] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, Feb. 2007.

[25] O. Gu'eant, J.-M. Lasry, and P.-L. Lions, "Mean field games: numerical methods," *Paris-Princeton Lectures on Mathematical Finance,* Springer, pp. 205–266, 2011.

[26] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjorungnes, *Game theory in wireless and communication networks: theory, models, and applications*, Cambridge University Press, 2012.

[27] H. Tembine, R. Tempone, and R. Vilanova, "Mean field games for cognitive radio networks," *American Control Conference (ACC)*, pp. 6388–6393, Montreal, Quebec, Canada, Jun. 2012.

[28] F. Meriaux and S. Lasaulce, "Mean-field games and green power control," *2011 5th International Conference on Network Games, Control and Optimization (NetGCooP)*, Paris, France, Oct. 2011.

$$u_i(t, s_i(t)) = \min_{p_i(t)} \mathbb{E}\left[\int_t^{t+dt} c_i(t, s_i(t), p_i(t))dt + u_i(t+dt, s_i(t+dt))\right]. \tag{35}$$

$$u_i(t+dt, s_i(t+dt)) = u_i(t, s_i(t)) + \left[\partial_t u_i(t, s_i(t)) + \partial_t s_i(t) \cdot \nabla u_i(t, s_i(t))\right] dt + o(dt), \tag{36}$$

$$-\partial_t u_i(t, s_i(t)) = \min_{p_i(t \to T)} \left[c_i(t, s_i(t), p_i(t)) + \partial_t s_i(t) \cdot \nabla u_i(t, s_i(t))\right], \tag{37}$$

$$H(p_i(t), s_i(t), \nabla u_i(t, s_i(t))) = \min_{p_i(t \to T)} \left[c_i(t, s_i(t), p_i(t)) + \partial_t s_i(t) \cdot \nabla u_i(t, s_i(t))\right]. \tag{38}$$

[29] Ali Y. Al-Zahrani, F. R. Yu, and M. Huang, “A joint cross-layer and co-layer interference management scheme in hyper-dense heterogeneous networks using mean-field game theory,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1522–1535, March 2016.

[30] M. Burger and J. M. Schulte, “Adjoint methods for hamilton-jacobibellman equations,” Westfalische Wilhelms-University, 2010.

[31] P. Semasinghe and E. Hossain, “Downlink power control in self-organizing denseSmall cells underlaying macrocells: a mean field game,” *IEEE Transactions on Mobile Computing*, vol. 15, no. 2, pp. 350–363, Feb. 1 2016.

[32] A. F. Hanif, H. Tembine, M. Assaad, and D. Zeghlache, “Mean-field games for resource sharing in cloud-based networks,” *IEEE/ACM Transactions on Networking*, vol. 24, no. 1, pp. 624–637, Feb. 2016.

[33] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, “Electrical vehicles in the smart grid-a mean field game analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1086–1096, Jul. 2012.