A Bayesian Overlapping Coalition Formation Game for Device-to-Device Spectrum Sharing in Cellular Networks

Yong Xiao, Kwang-Cheng Chen, *Fellow, IEEE*, Chau Yuen, *Senior Member, IEEE*, Zhu Han, *Fellow, IEEE*, Luiz A. DaSilva, *Senior Member, IEEE*

***Abstract*—We consider the spectrum sharing problem between a set of device-to-device (D2D) links and multiple co-located cellular networks. Each cellular network is controlled by an operator which can provide service to a number of subscribers. Each D2D link can either access a sub-band occupied by a cellular subscriber or obtain an empty sub-band for its exclusive use. We introduce a new spectrum sharing mode for D2D communications in cellular networks by allowing two or more D2D links with exclusive use of sub-bands to share their sub-bands with each other without consulting the operators. We establish a new game theoretic model called Bayesian non-transferable utility overlapping coalition formation (BOCF) game. We show that our proposed game can be used to model and analyze the above spectrum sharing problem. However, we observe that the core of the BOCF game can be empty, and we derive a sufficient condition for which the core is non-empty. We propose a hierarchical matching algorithm which can detect whether the sufficient condition is satisfied and, if it is satisfied, achieve a stable and unique matching structure which coincides with the overlapping coalition agreement profile in the core of the BOCF game.**

***Index Terms*—Device-to-device communication, overlapping, coalition formation, graph, matching, spectrum sharing, cellular network, stable marriage, college admission, stable roommate, game theory.**

I. Introduction

With the proliferation of wireless data services and applications, it will soon become difficult for the existing cellular network infrastructure to support the demands for mobile data services under the traditional infrastructure-centric network frameworks. One reason is that, in infrastructure-centric network frameworks, all traffic is forwarded and relayed by the cellular network infrastructure (e.g., base station) even when the sources and destinations are close to each other. This not only increases communication delay and energy consumption but also reduces the reliability of the networks. For example, in cellular networks, failure of a base station can lead to mobile service outage for the entire coverage area of the corresponding cell. Device-to-device (D2D) communication without relying on the base station to forward the traffic provides an efficient way to increase the network capacity and reliability. Another issue is that the traditional exclusive spectrum ownership model used in existing cellular networks has resulted in inefficient spectrum utilization for a significant portion of the time [2], [3]. One technique that promises to address this problem is *spectrum sharing*, which allows under-utilized licensed spectrum to be shared by unlicensed devices. Allowing both D2D communication and spectrum sharing in cellular networks can improve network capacity, reliability and spectrum utilization efficiency. However, D2D links are generally established autonomously and cannot be fully controlled by the base station. In addition, choosing the wrong spectrum sharing pair of D2D links and cellular subscribers can result in high cross-interference, which may adversely affect both D2D links and cellular subscribers.

This motivates the work in this paper, where we investigate the joint optimization of spectrum sharing approaches and sub-band allocation problem for a set of D2D links in an area with multiple co-located cellular networks. Each cellular network is controlled by an operator. We propose a general analytical framework in which each D2D link first chooses its preferred operator and then decides whether to apply for the exclusive use of a cellular sub-band or to share the sub-band with existing cellular subscribers. Since D2D links are autonomous, D2D links being assigned sub-bands for exclusive use can also share their spectrum with each other to further increase the spectrum utilization efficiency. We hence introduce a new spectrum sharing mode for D2D communication in cellular networks, referred to as the *sharing mode*. In this mode, D2D links being assigned vacant sub-bands can share their sub-bands without consulting the operator.

The distributed nature and autonomy of D2D links make game theory a natural tool to study and analyze D2D communication systems in cellular networks. We establish a new game theoretic framework, referred to as Bayesian non-transferable utility overlapping coalition

Some of the results in this paper have been presented at the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), April, 2014 [1]. This work is partially supported by the National Science Foundation under Grants CNS-1443917, ECCS-1405121, CNS-1265268, and CNS-0953377, National Science Foundation China under Grant 61428101, the Science Foundation Ireland under Grant No. 10/IN.1/I3007, and the Singapore University Technology and Design under Grant No. SUTD-ZJU/RES/02/2011, Mediatek Inc. and Ministry of Science and Technology of Taiwan under Grant No. MOST 103-2622-E-002-034.

Y. Xiao and Z. Han are with the Electrical and Computer Engineering at University of Houston, TX, USA (e-mails: xyong.2012@gmail.com and zhan2@uh.edu).

K. C. Chen is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan (e-mail: ckc@ntu.edu.tw)

C. Yuen is with Singapore University of Technology and Design, Singapore (email: yuenchau@sutd.edu.sg)

L. A. DaSilva is with CTVR, Trinity College Dublin, Ireland (e-mail: dasilval@tcd.ie). L. A. DaSilva is also with Virginia Tech, VA, USA.

formation (BOCF) game, to analyze the spectrum sharing problem between D2D links and cellular networks. In our proposed game, D2D links that operate in the spectrum of the same operator can be regarded as a coalition. Each member of a coalition can share spectrum with the existing cellular subscribers or apply for an exclusive sub-band to be used by itself or shared with other D2D links. If D2D links from different coalitions decide to share spectrum with each other, the coalitions will overlap. Our proposed framework is general and the payoff of each D2D link can be any performance measure generated from its received signal-to-interference-and-noise ratio (SINR). In addition, each D2D link is not required to know the payoffs or actions of others. We consider the concept of the core of coalition formation and seek an overlapping coalition agreement profile in the core that maximizes the payoffs of D2D links.

Our proposed game is a generalization of the traditional partition-based Bayesian coalition formation game [4]. As pointed out in [5], even analyzing the partition-based coalition formation game can be challenging. Finding a stable coalition structure is an NP-hard problem and generally requires an exhaustive search of all the possible coalitions formed by the players. Allowing overlaps among different coalitions further increases the complexity of the system, and the core of the proposed game may not always be non-empty.

Fortunately, we observe that our proposed game can be solved by exploiting tools from matching theory [6]. Specifically, we introduce a hierarchical matching algorithm to approach a stable overlapping coalition formation. Our algorithm consists of three individual algorithms, each of which is used to achieve a stable matching structure of a specific matching market. The first matching market is a two-sided many-to-one matching market with private belief, in which each D2D link selects the operator with the spectrum that can maximize its payoff. All D2D links that are accepted by the same operator form a coalition. Within each coalition, D2D links compete for the sub-bands of the operator. We model this problem as a two-sided one-to-one matching market. In this market, each D2D link applies for sharing sub-bands with existing cellular subscribers. If some D2D links decide to share the spectrum with other D2D links in the network, they will enter the third market, which is a one-sided one-to-one matching market. We propose a distributed belief updating algorithm for each D2D link to search for a unique and stable matching structure. We prove that this matching structure coincides with the overlapping coalition agreement profile in the strict Bayesian core of our proposed game. We also derive a sufficient condition for which the core of the game is non-empty. Our proposed distributed optimization algorithm can detect whether this sufficient condition is satisfied and, if satisfied, to achieve an overlapping coalition structure in the core.

The rest of this paper is organized as follows. Related work is reviewed in Section II. The network model is presented in Section III. The D2D and cellular spectrum sharing problem is formulated in Section IV. This problem is modeled as a BOCF game in Section V. The hierarchical matching algorithm is proposed in Section VI. The numerical results are presented in Section VII, and we offer our concluding remarks and future works in Section VIII.

II. Related Work

Most of the previously reported results on resource management for D2D communications focus on resource allocation for a single D2D link with specific performance goals. For example, in [7], the authors applied power control and multi-hop routing discovery methods to improve the probability of outage for opportunistic D2D communications in a cellular network. The power control problem for D2D links in a cellular network was also studied in [8]–[10]. In [11], the authors investigated the possible performance improvement brought by network coding and user cooperation in a D2D communication system. Observing the fact that D2D communications have not yet been considered in LTE-Advanced systems, the authors in [12] have proposed a mechanism to support a D2D communication session in existing LTE cellular networks. In [13], a distributed channel-aware spatial resource allocation algorithm, referred to as FlashLinQ, was proposed for ad hoc network systems. Motivated by the recent observation that treating the interference as noise at each of the spectrum sharing D2D links is information theoretically optimal under certain conditions, a new spectrum sharing mechanism referred to as information-theoretic link scheduling (ITLinQ) has been proposed in [14]. In [1], we model the spectrum sharing problem between a set of D2D links and one cellular operator as a Bayesian non-cooperative game. In this paper, we extend our previous work in [1] to the case of multiple operators. This extension dramatically changes the structure of the problem studied in [1] because different operators have different resources and each operator will only reveal its resource information to the D2D links being given permission to access its spectrum. How D2D links can select their preferred operator without knowing which sub-band they will be eventually allocated by each operator is a challenging task.

Different from the existing work, in this paper we study the interaction between different D2D links and between D2D links and cellular subscribers in a general multi-user D2D communication-enabled cellular network using coalitional game theoretic models. Recently, coalitional game theory has been used to study interactions in wireless networks [15], [16]. For example, in [17], a coalition formation game has been applied to study the dynamic spectrum access problem in cognitive radio networks. In [16], a hierarchical game theoretic framework has been proposed which allows unlicensed users to cooperatively share the licensed spectrum by

paying a certain price to licensed users. However, most of the existing studies either focus on the cooperation among all the wireless users or non-overlapping coalition formation. In this paper, we introduce a new Bayesian non-transferable overlapping coalition formation (BOCF) game model to study spectrum sharing by D2D communications in cellular networks.

In this paper, we propose a hierarchical matching algorithm to find the overlapping coalition agreement in our proposed game. The two-sided stable matching problem has been widely studied from both theoretical and practical perspectives [6], [18]–[20]. In this problem, each agent belonging to the set of one side of the market has a preference about the agents belonging to the set of the other side and tries to find a matching to optimize its performance. Many extensions of these problems have been studied in the literature. The case of some agents on the one side only having preferences over a sub-set of the agents on the other side was studied in [21]. The case where the agents from one side have equal preference over multiple agents of the other side, called stable marriage with tie, has been studied in [22]. Empirical studies of the different variations of the stable marriage problem have been reported in [19], [23]. In most of the previous works, each player cannot have any belief about the environment as well as the preference of others. In this paper, we allow each player to establish and maintain a private belief function. One work that is similar to our setting of private belief for agents is the belief-based coalition formation game proposed in [24]. However, that work assumes the belief functions are fixed, and cannot be updated during the game, which is different from the setting of our paper, where we introduce a Bayesian belief update algorithm to allow each player to search for the optimal matching structure.

TABLE I
LIST OF NOTATION

Symbol	Definition
$\mathcal{D}$	Set of D2D links
$\mathcal{O}$	Set of operators
$\mathcal{S}^i$	Set of sub-bands of operator i
$\mathcal{J}^i$	Set of sub-bands of operator i occupied by cellular subscribers
$\mathcal{K}^i$	Set of sub-bands of operator i unoccupied by cellular subscribers
$\mathcal{C}^o$	Set of D2D links with vacant sub-bands
D_k	the kth D2D link
P_l^i	Cellular subscriber in sub-band l of operator i
$\varpi_{D_k}[l]$	Payoff of D_k when accessing sub-band l
$\varpi_{D_k}[l,m]$	Payoff of D_k when sharing a sub-band with D_n for $l=\Gamma^s(D_k)$, $m=\Gamma^s(D_n)$ and $D_n, D_k \in \mathcal{K}$
$\bar{\varpi}_{D_k}$	Expected payoff of D_k
$\eta_i^l(D_k)$	Revenue of the operator obtained from D_k in sub-band l
$\mathcal{R}_{D_k}^d$	Preference of D_k over other D2D links with exclusive use of sub-bands
$\boldsymbol{\phi}_{D_k}$	Decision of D_k to send a request to operator, sub-band and D2D links with vacant sub-bands
Γ	Conflict-solving rules of the operators, sub-bands and D2D links with sub-bands for exclusive use
$\boldsymbol{a}_{\mathcal{C}^i}$	Action of players in coalition $\mathcal{C}^i$
Y_{D_k}	Type of D_k
B_{D_k}	Belief function of D_k about the decisions of other D2D links and the conflict-solving rules
b_{D_k}	Belief function of each player D_k about the types of other players
$\tilde{\boldsymbol{\phi}}$	Labeling sequence of D2D links with vacant sub-bands

III. A General System Model for D2D Communications in Cellular Networks

We consider spectrum sharing between a set of K D2D links, labeled as $\mathcal{D}=\{D_1, D_2, \ldots, D_K\}$, and a set of L co-located cellular network operators, labeled as operators $\mathcal{O}=\{1,2,\ldots,L\}$. Each D2D link corresponds to a communication channel between a D2D source and its corresponding destination, and each cellular subscriber corresponds to a downlink or uplink communication channel from the BS to the cellular subscriber as shown in Figure 1. To avoid causing interference to the neighbouring cell, we assume each D2D link can only share spectrum with the subscribers in its local cell.

Each operator i has been licensed an exclusive set $\mathcal{S}^i$ of sub-bands which can be accessed by both D2D links and cellular subscribers. Let $\mathcal{K}^i$ be the subset of vacant sub-bands of operator i unoccupied by cellular subscribers. Let $\mathcal{J}^i$ be the subset of sub-bands occupied by the cellular subscribers of operator i, i.e., we have $\mathcal{J}^i \cap \mathcal{K}^i = \emptyset$ and $\mathcal{S}^i = \mathcal{J}^i \cup \mathcal{K}^i \ \forall i \in \mathcal{O}$. Each D2D link can only share sub-bands with the cellular subscribes within the same cell. Since the access to licensed spectrum is expensive, the exclusive sub-band given to each D2D link may, in practice, be narrower than the full-size sub-band allocated to the cellular subscribers. Each D2D link can either access a sub-band occupied by a cellular subscriber or apply for a vacant sub-band for exclusive use if sharing the spectrum with a cellular subscriber cannot provide sufficient quality-of-service (QoS). Let P_l^i be the cellular subscriber occupying sub-band l of operator i for $l \in \mathcal{J}^i$. We denote $\mathcal{S}=\bigcup_{i\in\mathcal{O}} \mathcal{S}^i$, $\mathcal{J}=\bigcup_{i\in\mathcal{O}} \mathcal{J}^i$ and $\mathcal{K}=\bigcup_{i\in\mathcal{O}} \mathcal{K}^i$.

Because of the complexity of the interference management in D2D and cellular spectrum sharing problem, most existing works assume that each D2D link can share spectrum with at most one cellular subscriber [7], [25]–[27]. In this paper, we follow the same line and assume that each sub-band can at most contain two users (either two D2D links or one D2D link and one cellular subscriber). This assumption makes the spectrum sharing between D2D links and cellular networks feasible to be implemented in the existing cellular telecommunication system. For example, in Release 12 of the LTE standard, an eNB (Evolved Node B) can keep track of the interference received at each of its cellular subscribers in each sub-band and can simply remove the D2D link from the sub-band once it observes a higher-than-tolerable interference level [28], [29]. Our model however can be directly extended to the cases with two or more D2D links sharing the same sub-band with each cellular

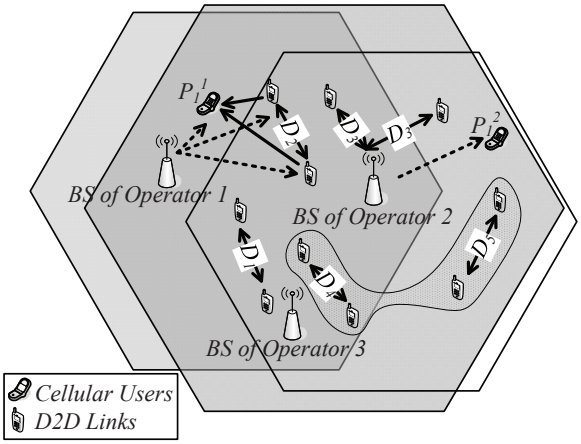

Fig. 1. System model for D2D communications in a cellular network with three operators: D2D links D_1, D_2, D_3 operate in modes M1 (i.e., D_1 has been assigned to a dedicate sub-band for exclusive use), M2 (i.e., P_1^1 and D_2 share the same sub-band) and M3 (i.e., the traffic of D_3 has been forwarded by eNB of operator 2), respectively. D2D links D_4 and D_5 operate in mode M4 (i.e., D_4 and D_5 share their dedicated sub-bands with each other).

subscriber. We will discuss this in detail in Section VIII.

A commonly adopted approach is to divide possible spectrum sharing schemes between D2D links and cellular subscribers into three modes [12]:

M1. *Dedicated Mode:* D2D links access dedicated sub-bands that are unoccupied by the cellular subscribers,

M2. *Reuse Mode:* D2D links reuse the sub-bands occupied by cellular subscribers,

M3. *Relay Mode:* The traffic of D2D links is relayed through the BS. This mode is normally applied when direct communication cannot provide adequate performance for D2D links.

In cellular networks, each D2D link can operate in one of the above three modes with help from the eNB of the corresponding operator. The detailed implementation of these modes in LTE-Advanced systems has been described in [12], [26], [30].

We illustrate the D2D links and potential interference for the above three modes in Figure 1.

Since D2D links are autonomous, to further improve the spectrum utilization efficiency, D2D links assigned dedicated sub-bands for their exclusive use can also share their sub-bands with each other. Therefore, we introduce the following new mode for spectrum sharing between D2D communications and cellular networks:

M4. *Sharing Mode:* D2D links in mode M1 can further increase spectrum utilization efficiency by sharing their dedicated spectrum with each other.

We also illustrate the mode M4) in Figure 1. We will provide more detailed discussion on the possible implementation of our proposed framework in LTE Advanced network systems in Section VII.

Different from most of the existing works, which assume each D2D link obtains the same performance in different sub-bands under each specific mode, we consider a more general system in which multiple operators co-exist in the same coverage area and each D2D link in each specific mode obtains different performance in different sub-bands. We consider the joint optimization for both mode selection and the sub-band accessing/sharing. That is, each D2D link should not only choose a specific mode to operate in but also decide a specific operator and sub-band that can maximize its performance in its chosen mode.

Let $\varpi_{D_k}[l]$ be the payoff of D2D link D_k obtained by accessing sub-band l in mode M1 (if $l \in \mathcal{K}$) or M2 (if $l \in \mathcal{J}$) for $D_k \in \mathcal{D}$ and $l \in \mathcal{S}$. Let $\varpi_{D_k}[l,m]$ be the payoff of D2D link D_k when it shares its assigned sub-band l with another D2D link which has been assigned sub-band m for $l, m \in \mathcal{K}$ (mode M4). We consider a general model and the payoff of each D2D link can be any function of its received signal-to-interference-plus-noise ratio (SINR). For example, if D2D link D_k wants to maximize its transmit rate per bandwidth price, the payoff of the D2D link D_k in each mode is given as follows.

M1 and M2: When D2D link D_k accepts a dedicated sub-band l (in mode M1) or shares a sub-band l with a cellular subscriber (in mode M2), its payoff is given by

$$\varpi_{D_k}[l] = \frac{\rho[l]}{e[l]} \mathbb{E} \log\left(1 + \text{SINR}_{D_k}[l]\right), \tag{1}$$

where $\rho[l]$ is the bandwidth of sub-band l and $e[l]$ is the price paid to the operator for accessing sub-band l. Note that, different from the cellular networks in which the operator charges subscribers according to the quality of experience (QoE), in D2D communications, the data traffic does not traverse the network infrastructure and the operators cannot monitor the transmission rate between two devices that establish a direct link. Therefore, in this paper, we assume each operator charges a fixed price $e[l]$ for any D2D link to access a dedicated sub-band l. The payoff of each D2D link corresponds to the transmission rate per unit price obtained by accessing the assigned sub-bands of the corresponding operator. $\text{SINR}_{D_k}[l]$ is the signal-to-noise-and-interference ratio (SINR) experienced by D2D link D_k in sub-band l, given by [31],

$$\text{SINR}_{D_k}[l] = \begin{cases} \frac{h_{D_k}[l] w_{D_k}}{\varrho_{D_k}[l]}, & l \in \mathcal{K} \text{ in mode M1}, \\ \frac{h_{D_k}[l] w_{D_k}}{\varrho_{D_k}[l] + h_{P_l^i D_k} w_{P_l^i}}, & l \in \mathcal{J} \text{ in mode M2}, \end{cases} \tag{2}$$

where $\varrho_{D_k}[l]$ is the additive noise received by D_k in sub-band l, $h_{D_k}[l]$ is the channel gain between the source and destination of D2D link D_k in sub-band l. $h_{P_l^i D_k}$ is the channel gain between cellular subscriber P_l^i and D2D link D_k. $w_{P_l^i}$ and w_{D_n} are the transmit powers of P_l^i and D_n, respectively.

M3. If D2D link D_k decides to use mode M3, it will first transmit to the BS and then wait for the BS to

forward the signals to the corresponding D2D destination. Since, in this mode, the traffic of D2D links is transmitted in the same way as for the cellular subscribers, this mode of operation cannot provide any improvement in terms of the spectrum utilization efficiency and should be the last choice of each D2D link. In this paper, we assume each D2D link cannot obtain any positive payoff in this mode, i.e., we write the payoff of D_k in mode M3 as $\varpi_{D_k}[D_k] = 0$.

M4. If two D2D links D_k and D_n with dedicated sub-bands l and m, respectively, decide to share their sub-bands with each other by transmitting at the same time over the same aggregated sub-bands l and m [32], [33] and agree to equally share the cost of sub-bands l and m for $l, m \in \mathcal{K}$ and $i, j \in \mathcal{O}$, the payoff of each D2D link (e.g., D_k) in mode M4) is given by

$$\varpi_{D_k}[l,m] = \frac{2\left(\rho[l]+\rho[m]\right)}{e[l]+e[m]}\mathbb{E}\log\left(1+\text{SINR}_{D_k}[l,m]\right), \quad (3)$$

where $\text{SINR}_{D_k}[l,m] = \frac{h_{D_k}[l,m]w_{D_k}}{\varrho_{D_k}[l,m]+h_{D_nD_k}[l,m]w_{D_n}}$, $h_{D_k}[l,m]$ and $h_{D_nD_k}[l,m]$ are the channel gain between the source and destination of D2D link D_k and that between the source of D2D link D_k and the destination of D2D link D_n in the frequency band formed by aggregating sub-bands l and m, respectively. $\varrho_{D_k}[l,m]$ is the additive noise at the receiver of D2D link D_k in the aggregated frequency band formed by sub-bands l and m.

We follow a commonly adopted setting and set the revenue of each operator in the sub-bands occupied by cellular subscribers as a function of the resulting interference caused by the D2D links [16], [34], [35]. We can also define the revenue of operator i from a D2D link D_k accessing a vacant sub-band for exclusive use, as a function of the SINR of D_k, i.e., the revenue $\eta_i^l(D_k)$ obtained by operator i from D2D link D_k in sub-band l is given by $\eta_i^l(D_k) = g\left(\mathbb{E}\text{INT}_{D_k}[l]\right)$ where $g(\cdot)$ is the revenue function and $\text{INT}_{D_k}[l]$ is given by

$$\text{INT}_{D_k}[l] = \begin{cases} h_{D_kP_i}w_{D_k}, & l \in \mathcal{J}^i, \\ \frac{h_{D_k}[l]w_{D_k}}{\varrho_{D_k}[l]}, & l \in \mathcal{K}^i. \end{cases} \quad (4)$$

In this setting, the price charged by each operator to each UE sharing the sub-bands with cellular subscribers is proportional to the interference caused by the UE. As is observed in [16], this allows each operator to control the resulting interference created by the UEs by adjusting the prices. For example, if the revenue $\eta_i^l(D_k)$ is a linear function of $\mathbb{E}\text{INT}_{D_k}[l]$, we have $\eta_i^l(D_k) = \max_{l\in\mathcal{J}^i}\left\{\beta^i\mathbb{E}\text{INT}_{D_k}[l]\right\}$, where β^i is the pricing coefficient of operator i [36]. It has been shown in [37] that the operators can always limit the interference of the D2D links in their sub-bands by adjusting the value of the pricing coefficient.

We can now show that the joint optimization of the mode selection and sub-band accessing/sharing problem is equivalent to the optimization of the sub-band allocation problem for D2D links. For example, if D2D link D_k has been assigned to sub-band $l \in \mathcal{K}$ for exclusive use, D_k will be in mode M1. If D2D link D_k has been eventually allocated sub-band $l \in \mathcal{J}$ that is occupied by a cellular subscriber, D_k operates in mode M2. If D2D link D_k can access an aggregated sub-band formed by two sub-bands l and m for $l, m \in \mathcal{K}$, D_k operates in mode M4. If D2D link D_k cannot obtain any sub-band to support its direct communication, it will then totally rely on the operator to establish connectivity and forward traffic between the two devices and hence will be in mode M3. In the rest of this paper, we focus on optimization of the sub-band allocations for D2D links in cellular networks.

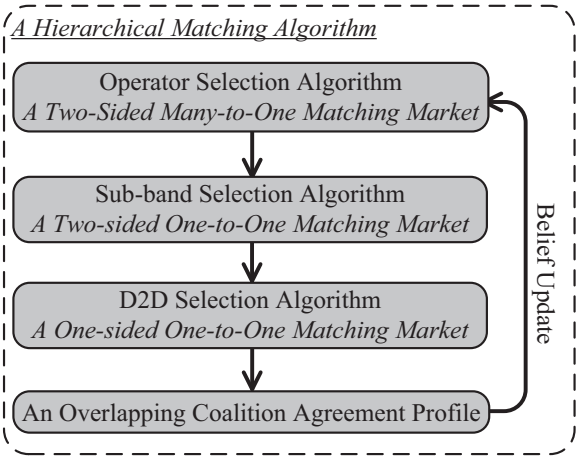

Fig. 2. A hierarchical matching algorithm for BOCF game.

The list of notation used in this paper is provided in Table I.

IV. Problem Formulation

As mentioned previously, each operator only possesses a limited number of sub-bands and hence can only provide service for a limited number of D2D links. When the number of D2D links requesting to access the spectrum of an operator exceeds this limit, a *conflict* will happen. Similarly, conflicts may also happen when more than one D2D link send a request for the same sub-band of an operator or the same D2D link to share a sub-band with. To avoid possible overloading, the operator will have to reject the requests of some D2D links, i.e., if the number of D2D links sending requests to operator i exceeds $|\mathcal{S}^i|$, operator i will only allow $|\mathcal{S}^i|$ D2D links to access its spectrum, selected according to the revenue that can be obtained from the requesting D2D links.

We assume the spectrum sharing process can be divided into time slots. We follow the same line as [13] and assume the communication of D2D links is synchronized using the timing signals sent by the cellular networks or the GPS timing signal. Each D2D link needs to make its decisions about operators, sub-bands and sub-band sharing partner at the beginning of each time slot and cannot change its decision during the rest of the time slot. The decisions of

each D2D link, however, can be changed between different time slots. We use subscript t to denote the parameters and results in time slot t. To simplify our description, we can ignore the subscript t when we only focus on one time slot of the decision process.

We can define the *D2D and cellular spectrum sharing* (DCSS) problem as a tuple $\mathcal{P} = \langle \mathcal{D}, \mathcal{O}, \mathbf{\Phi}, \Gamma, \boldsymbol{\varpi} \rangle$ consisting of following elements:

1) $\mathcal{D}$ is the set of D2D links.
2) $\mathcal{O}$ is the set of operators.
3) $\mathbf{\Phi} = \mathcal{O} \cup \{\emptyset\} \times \mathcal{S} \cup \{\emptyset\} \times \mathcal{D} \cup \{\emptyset\}$ is the set of possible decisions made by each D2D link about the operator, sub-bands and sub-band sharing partner. Each D2D link D_k can decide $\boldsymbol{\phi}_{D_k} = \langle \phi^o_{D_k}, \phi^s_{D_k}, \phi^d_{D_k} \rangle \in \mathbf{\Phi}$ where $\phi^o_{D_k} \in \mathcal{O} \cup \{\emptyset\}$ is the operator requested by D2D link D_k. We use $\phi^o_{D_k} = \emptyset$ to mean D_k declines to send a request to any operator (e.g., D_k believes sharing the spectrum of the cellular network cannot result in a positive payoff). $\phi^s_{D_k} \in \mathcal{S}^i \cup \{\emptyset\}$ is the sub-band requested by D2D link D_k after being accepted by operator i. We write $\phi^s_{D_k} = \emptyset$ if D2D link D_k declines to request any sub-band of operator i (e.g., D_k believes operator i does not possess any sub-band that can result in a positive payoff). If D_k has been assigned a sub-band for exclusive use, D_k can then send a request to another D2D link D_n with an exclusive sub-band asking to aggregate and share their sub-bands. Similarly, if D_k does not want to share its sub-band with any other D2D link, we have $\phi^d_{D_k} = \emptyset$. It can be easily observed that these decisions are closely related to each other. More specifically, D2D link D_k should decide $\phi^s_{D_k}$ and $\phi^d_{D_k}$ based on $\phi^o_{D_k}$. On the other hand, $\phi^o_{D_k}$ should be decided by considering the possible choices of sub-bands and sub-band sharing partners for the D2D links. We denote the decision profile of all D2D links as $\boldsymbol{\phi} = \langle \boldsymbol{\phi}_{D_k} \rangle_{D_k \in \mathcal{D}}$.
4) $\Gamma = \langle \Gamma^o, \Gamma^s, \Gamma^d \rangle$ is the conflict-solving rule of the operator and D2D links with dedicated sub-bands. We use $\Gamma^o(D_k)$ and $\Gamma^s(D_k)$ to denote the final operator and sub-band being assigned to D2D link D_k. We also use $\Gamma^o(D_k) = D_k$ or $\Gamma^s(D_k) = D_k$ to mean that D_k cannot directly communicate with another device but has to operate in mode M3. We also use $\Gamma^d(D_k)$ to denote the D2D link which agrees to share its sub-band with D_k. Similarly, we use $\Gamma^d(D_k) = D_k$ to mean D_k cannot share its sub-band with any other D2D link with an exclusive sub-band.
5) ϖ_{D_k} is the payoff of D_k, which depends on the decision profile $\boldsymbol{\phi}$ and the conflict-solving rule Γ, i.e., we have $\varpi_{D_k}(\boldsymbol{\phi}, \Gamma) = \mathbf{1}_{\Gamma^d(D_k)=D_k} \varpi_{D_k}[l = \Gamma^s(D_k)] + \sum_{D_n \in \mathcal{D} \setminus \{D_k\}} \mathbf{1}_{\Gamma^d(D_k)=D_n} \varpi_{D_k}[l = \Gamma^s(D_k), m = \Gamma^s(D_n)]$ for $\Gamma^s(D_k) \neq D_k$ and $\Gamma^s(D_n) \neq D_n$, where $\varpi_{D_k}[l]$ and $\varpi_{D_k}[l, m]$ are given in (1) and (3), respectively, and $\mathbf{1}$ is the indicator function.

If two or more D2D links have been allocated sub-bands for exclusive use, these D2D links can share their sub-bands with each other to further improve their payoffs. Since the sub-band allocation process has been supervised by the BS, the D2D links which are allocated dedicated sub-bands can obtain the identity information of each other from the BSs. Each D2D link with exclusive sub-bands (e.g., D_k) knows the set $\mathcal{C}^o$ of D2D links with sub-bands for exclusive use, defined as $\mathcal{C}^o = \left\{ D_k : \Gamma^s(D_k) \in \bigcup_{i \in \mathcal{O}} \mathcal{K}^i, \forall D_k \in \mathcal{D} \right\}$, and $\varpi_{D_k}[m, l]$ for $\Gamma^s(D_k) = m$, $\Gamma^s(D_n) = l$ $\forall D_n \in \mathcal{C}^o$ after the training period. We provide a more detailed discussion of this training process in Section VI.

It can be observed that the value of ϖ_{D_k} for each D2D link D_k can be affected by the decisions of all D2D links and the conflict-solving rules of the operators and D2D links with exclusive use of sub-bands, both of which are unknown to D_k. It is generally unrealistic to assume each device can predict all these unknown parameters instantaneously before it makes decisions at the beginning of each time slot t. It is however possible for each D2D link to eavesdrop on the operators requested by other D2D links during the previous time slot. As observed in [38], D2D communication will be mainly applied in high population density areas, where the cell sizes are generally small. This makes it possible for each D2D link to eavesdrop on the requests sent by nearby D2D links. Each D2D link can also obtain this information from the operators, i.e., each operator can broadcast its request acceptance and rejection message to all the D2D links. In this paper, we assume each D2D link cannot know the instantaneous decisions of others but can observe the decisions of other D2D links in previous time slots. Each D2D link can exploit these observations to establish a belief function about these unknown parameters.

The selfishness and autonomy of D2D links make it natural to model the DCSS problem as a game. During the rest of this paper, we focus on solving the following problems:

1) Establish a game theoretic model to study the interaction among autonomous and selfish D2D links.
2) Develop a distributed algorithm for each D2D link to optimize its decision to maximize its expected payoff.
3) Propose an effective conflict-solving rule for both the operator and D2D link with dedicated sub-bands to approach a sub-band allocation structure such that no operator or D2D link can benefit by unilaterally deviating.
4) Develop a belief updating algorithm for each D2D link to learn the probabilistic features of unknown parameters of other D2D links and operators using its previous observations.

To solve the first problem, we propose a Bayesian overlapping coalition formation game to model the DCSS problem in the next section. We will then develop the distributed algorithm, conflict-solving rules and belief updating approach in Section VI.

V. An Overlapping Coalition Formation Game

In many practical resource sharing problems, allowing overlap between different coalitions can further improve the system performance and resource utilization efficiency. For example, if multiple wireless network subscribers can access several resource blocks (e.g., frequency bands, time slots, antennas), they can be first divided into different coalitions each of which consists of the subscribers sharing one block of resource [37]. However, it is possible that, in some coalitions, the share of the resource blocks allocated to some subscribers is not enough to support a desired level of QoS, while for some other subscribers, the allocated resource may exceed those requirements. In this case, allowing the subscribers with insufficient resources to also aggregate or share some of the surplus resources allocated to other subscribers can further improve the resource utilization efficiency as well as the network system performance.

We define an overlapping coalition $\mathcal{C}^i$ formed by a set of players $\{D_1, D_2, \ldots, D_K\}$ as a vector $\mathcal{C}^i = \langle c^i_{D_1}, c^i_{D_2}, \ldots, c^i_{D_K}\rangle$ where $c^i_{D_k}$ is a binary variable and $c^i_{D_k} = 1$ means that D_k is a member of coalition $\mathcal{C}^i$ and $c^i_{D_k} = 0$ means D_k does not belong to coalition $\mathcal{C}^i$. If two coalitions $\mathcal{C}^i$ and $\mathcal{C}^j$ overlap, there exists at least one player $D_k \in \mathcal{D}$ such that $c^i_{D_k} = c^j_{D_k} = 1$ for $i \neq j$. Let $\mathrm{supp}(\mathcal{C}^i)$ be the support of $\mathcal{C}^i$. An *overlapping coalition formation structure* with L overlapping coalitions is defined as $\boldsymbol{\mathcal{C}} = \{\mathcal{C}^i\}_{i\in\{1,2,\ldots,L\}}$.

We formally define a BOCF game as follows:

Definition 1: A *BOCF game* $\mathcal{G} = \langle \mathcal{D}, \boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{Y}}, \boldsymbol{b}, \boldsymbol{\varpi}, \rhd\rangle$ consists of the following elements:

1) $\mathcal{D}$ is the set of players.
2) $\boldsymbol{\mathcal{A}}_{\mathcal{C}^i} = \boldsymbol{\mathcal{A}}^c_{\mathcal{C}^i} \times \boldsymbol{\mathcal{A}}^o_{\mathcal{C}^i}$ is the set of possible actions for the players in each coalition $\mathcal{C}^i$. An action $a_{\mathcal{C}^i} = \langle a^c_{\mathcal{C}^i}, a^o_{\mathcal{C}^i}\rangle$ of coalition $\mathcal{C}^i$ in a BOCF game consists of two parts: the coalitional action $a^c_{\mathcal{C}^i}$ and the overlapping action $a^o_{\mathcal{C}^i}$. A coalitional action $a^c_{\mathcal{C}^i} \in \mathcal{A}^c_{\mathcal{C}^i}$ for a coalition $\mathcal{C}^i$ is similar to the action in the non-overlapping coalition formation game, which specifies the joint action mutually agreed to by every member player within a coalition $\mathcal{C}^i$. An overlapping action $a^o_{\mathcal{C}^i} \in \mathcal{A}^o_{\mathcal{C}^i}$ specifies how the players in coalition $\mathcal{C}$ interact with players in other coalitions. For example, in the resource sharing problem, the coalitional action characterizes the resource allocation scheme mutually agreed to by all the subscribers to divide the resource block within one coalition. The overlapping action characterizes how subscribers being allocated resources of different coalitions exchange or share these resources. These two actions may be closely correlated in most applications. For example, some players allocated resources from different resource blocks can share portions of their resources with each other and in this case the overlapping actions (e.g., how they negotiate and share their portions of the allocated resources) depend on the coalitional actions (e.g., how to divide each resource block among the coalition members). It can be observed that the coalition formation structure and actions jointly determine the payoff of each member player in a coalition. We hence can define an *overlapping coalition agreement* as a tuple $x^i = \langle \mathcal{C}^i, a_{\mathcal{C}^i}\rangle$ for $\mathrm{supp}\left(\mathcal{C}^i\right) \subseteq \mathcal{D}$ and $a_{\mathcal{C}^i} \in \mathcal{A}_{\mathcal{C}^i}$. We also denote the overlapping coalition agreement profile $\boldsymbol{x}$ as the set of all overlapping coalitional agreements formed by the players, i.e., $\boldsymbol{x} = \{x^i\}_{i\in\{1,\ldots,L\}}$.
3) $\boldsymbol{\mathcal{Y}} = \mathcal{Y}_{D_1} \times \mathcal{Y}_{D_2} \times \ldots \times \mathcal{Y}_{D_K}$ is the type space, where $\mathcal{Y}_{D_k}$ is the set of possible types of player D_k. The type $Y_{D_k} \in \mathcal{Y}_{D_k}$ of each player D_k specifies its preference regarding different overlapping coalition agreements.
4) $\boldsymbol{b} = \langle b_{D_1}, b_{D_2}, \ldots, b_{D_K}\rangle$ is the vector of belief functions, where b_{D_k} is the belief function of player D_k about the types of others. Each player D_k cannot know the types of other players. Each player can however establish a belief function about these unknown types by exploiting the previous observations.
5) $\boldsymbol{\varpi}$ is the vector of the payoffs of the players.
6) $\rhd$ is the preference relation. The preference relation $\rhd$ is assumed to be complete and transitive [6]. We use $x \rhd_{D_k} x'$ to denote that player D_k prefers overlapping coalition agreement x to x' for $x \neq x'$. We also use $x \rhd^{D_n}_{D_k} x'$ to denote that player D_k believes D_n prefers overlapping coalition agreement x to x' for $D_n \neq D_k$ and $D_k, D_n \in \mathcal{D}$.

An important solution concept in the coalitional game is the *core*, which is formally defined as follows.

Definition 2: An overlapping coalition agreement profile $\boldsymbol{x}^*$ is in the *weak Bayesian (overlapping coalition formation) core* if there is no overlapping coalition agreement $x' = \langle \mathcal{C}', a_{\mathcal{C}'}\rangle \in \boldsymbol{x}^*$ such that every member believes it will benefit from deviating from the current overlapping coalition agreement x^i, i.e., $\nexists\ x' = \langle \mathcal{C}', a_{\mathcal{C}'}\rangle$ and $\mathcal{C}^i \in \boldsymbol{\mathcal{C}}^*$ such that $x' \rhd_{D_k} x^i\ \forall D_k \in \mathcal{C}^i$.

The above definition can be regarded as the direct extension of the core for the Bayesian non-overlapping coalition formation game to the overlapping case. If we take the belief of each player into consideration, we can propose a belief-based concept of the core, referred to as *b-core*, in the BOCF game as follows.

Definition 3: We say an overlapping coalition agreement profile $\boldsymbol{x}^{**}$ is in the *(Bayesian overlapping coalition formation) b-core*, if the following two conditions are satisfied: 1) there exists no overlapping coalition agreement such that every member believes it will benefit from deviating from the current overlapping coalition agreement, 2) there exists no overlapping coalition agreement such that at least one member of a coalition believes that each of the other members in the coalition believes it will benefit from deviating from their current overlapping coalition agreement, i.e., there does not exist $x' = \langle \mathcal{C}', a_{\mathcal{C}'}\rangle$ and $\mathcal{C}^i \in \boldsymbol{\mathcal{C}}^{**}$ such that there exists a D2D link $D_k \in \mathcal{C}^i$ satisfying $x \rhd^{D_n}_{D_k} x'$ $\forall D_n \neq D_k, D_n \in \mathcal{C}^i$.

Note that both concepts of the core defined above are different from the core related to the grand coalition used in many coalitional game-based wireless network models [39]–[41]. The latter core concept can only be non-empty when all the players in the game agree to form the grand coalition, that is, the coalition that contains all the players [4].

The concept of the core in Definition 2 can be in some sense regarded as an extension of the a-core proposed in [42] into the BOCF game. It is different from the Aubin core for the cooperative fuzzy game in [43] as well as the o-core and r-core concepts proposed for the transferable utility overlapping coalition formation game in [42].

We can model the DCSS problem as a BOCF game, referred to as DCSS game, $\mathcal{G}^{SS} = \langle \mathcal{D}, \boldsymbol{\mathcal{Y}}, \Gamma, \boldsymbol{b}, \boldsymbol{\varpi}, \rhd \rangle$ as follows: the *players* are the D2D links. The *coalitional action* of a coalition $\mathcal{C}^i$ corresponds to the sub-band allocation scheme achieved by all the D2D links being accepted by the same operator i. More specifically, the coalitional action $a^c_{\mathcal{C}^i}$ is determined by the decisions made by D2D links in coalition $\mathcal{C}^i$ as well as the conflict-solving rules of operator i. The *overlapping action* corresponds to the sub-band sharing scheme between the D2D links with exclusive use of sub-bands from different coalitions. The *type* Y_{D_k} of each player D_k is its preference over all the possible overlapping coalition agreements. Each D2D link cannot know the types of other D2D links and it is generally difficult for each D2D link to establish a belief function over others' types. Fortunately, we can show that the uncertainty of each D2D link about types of other D2D links can be converted into the uncertainty about the decisions of others and conflict-solving rules of the operators and D2D links with vacant sub-bands. It can be observed that, for a given conflict-solving rule Γ, the final overlapping coalition formation structure $\boldsymbol{\mathcal{C}}$ is determined by the decisions $\boldsymbol{\phi}$ of all D2D links. By introducing a function F mapping from Γ and $\boldsymbol{\phi}$ to an overlapping coalition formation structure $\boldsymbol{\mathcal{C}}$, we have $\boldsymbol{\mathcal{C}} = F(\Gamma, \boldsymbol{\phi})$. For each of the coalition formation structures, the coalitional action $a^c_{\mathcal{C}^i}$ in coalition $\mathcal{C}^i$ specifies the sub-band allocation between D2D links being accepted by operator i and the set $\mathcal{S}^i$ of sub-bands. Since the sub-band assigned to each D2D link $D_k \in \mathcal{C}^i$ is given by $\Gamma^s(D_k)$, we can observe that the coalition action $a^c_{\mathcal{C}^i}$ is determined by Γ^s and $\boldsymbol{\phi}^s_{\mathcal{C}^i}$ where $\boldsymbol{\phi}^s_{\mathcal{C}^i} = \left\{\phi^s_{D_k}\right\}_{D_k \in \mathcal{C}^i}$. If we introduce a function G mapping from $\boldsymbol{\phi}^s_{\mathcal{C}^i}$ and Γ^s into $a^c_{\mathcal{C}^i}$, we can write $a^c_{\mathcal{C}^i} = G\left(\Gamma^s, \Gamma^d, \boldsymbol{\phi}^s_{\mathcal{C}^i}\right)$. Similarly, for each overlapping coalition formation structure, the set $\mathcal{C}^o$ of D2D links with sub-bands for their exclusive use is fixed. Also, since the overlapping action of each D2D link $D_k \in \mathcal{C}^i \cap \mathcal{C}^o$ in a coalition $\mathcal{C}^i$ is determined by the decision $\phi^d_{D_k}$ and the conflict-solving rule Γ^o, we can define a function H mapping from the decisions of D2D links with exclusive sub-bands and the conflict-solving rules of these D2D links into the overlapping action, i.e., we have $a^o_{\mathcal{C}^i} = H\left(\Gamma^d, \boldsymbol{\phi}^d_{\mathcal{C}^o}\right)$. Therefore, we can write each overlapping coalition agreement $x^i = \langle \mathcal{C}^i, a_{\mathcal{C}^i} \rangle = \langle F(\Gamma, \boldsymbol{\phi}),\ \langle G(\Gamma^s, \boldsymbol{\phi}^s_{\mathcal{C}^i}), H\left(\Gamma^d, \boldsymbol{\phi}^d_{\mathcal{C}^o}\right)\rangle\rangle$. In other words, the preference of each D2D link about the coalitional agreements can be converted into its preference over different decisions for a given conflict-solving rule. In the DCSS game, each D2D link can observe the decisions of other D2D links and the operator and sub-band it has been allocated during the previous time slots and hence can exploit these observations to establish a belief function about the decisions of other D2D links. The belief function $B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right) = \Pr\left(\Gamma^o(D_k), \Gamma^s(D_k), \Gamma^d(D_k), \boldsymbol{\phi}_{-D_k} | \phi_{D_k}\right)$ of each D2D link D_k can be divided into six parts: the first three belief functions correspond to the beliefs of D_k about the decisions of other D2D links regarding operators, sub-bands and the D2D sub-band sharing partner, i.e., $B_{D_k}\left(\boldsymbol{\phi}^o_{-D_k}\right) = \Pr\left(\boldsymbol{\phi}^o_{-D_k} | \phi^o_{D_k}\right)$, $B_{D_k}\left(\boldsymbol{\phi}^s_{-D_k}\right) = \Pr\left(\boldsymbol{\phi}^s_{-D_k} | \Gamma^o(D_k), \phi^s_{D_k}\right)$ and $B_{D_k}\left(\boldsymbol{\phi}^d_{-D_k}\right) = \left(\boldsymbol{\phi}^d_{-D_k} | \Gamma^o(D_k), \Gamma^s(D_k), \phi^d_{D_k}\right)$, and the remaining three belief functions correspond to the beliefs of D_k about the conflict-solving rules of operators, sub-bands and D2D links with sub-bands for exclusive use, i.e., $B_{D_k}(\Gamma^o) = \Pr\left(\Gamma^o(D_k) | \phi^o_{D_k}, \boldsymbol{\phi}^o_{-D_k}\right)$, $B_{D_k}(\Gamma^s) = \Pr\left(\Gamma^s(D_k) | \phi^o_{D_k}, \boldsymbol{\phi}^o_{-D_k}, \Gamma^o, \phi^s_{D_k}, \boldsymbol{\phi}^s_{-D_k}\right)$, and $B_{D_k}\left(\Gamma^d\right) = \Pr\left(\Gamma^d(D_k) | \phi^o_{D_k}, \boldsymbol{\phi}^o_{-D_k}, \Gamma^o, \phi^s_{D_k}, \boldsymbol{\phi}^s_{-D_k}, \Gamma^s, \phi^d_{D_k}, \boldsymbol{\phi}^d_{-D_k}\right)$. We have

$$
\begin{aligned}
&B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right) = \\
&\quad \Pr\left(\Gamma^o(D_k), \Gamma^s(D_k), \Gamma^d(D_k),\right. \\
&\quad \left.\langle \boldsymbol{\phi}^o_{-D_k}, \boldsymbol{\phi}^s_{-D_k}, \boldsymbol{\phi}^d_{-D_k}\rangle | \langle \phi^o_{D_k}, \phi^s_{D_k}, \phi^d_{D_k}\rangle\right) \\
&= \Pr\left(\boldsymbol{\phi}^o_{-D_k} | \phi^o_{D_k}\right) \Pr\left(\Gamma^o(D_k) | \boldsymbol{\phi}^o\right) \\
&\quad \Pr\left(\boldsymbol{\phi}^s_{-D_k} | \Gamma^o(D_k), \phi^s_{D_k}\right) \Pr\left(\Gamma^s(D_k) | \Gamma^o(D_k), \boldsymbol{\phi}^s\right) \\
&\quad \Pr\left(\boldsymbol{\phi}^d_{-D_k} | \Gamma^o(D_k), \Gamma^s(D_k), \phi^d_{D_k}\right) \\
&\quad \Pr\left(\Gamma^d(D_k) | \Gamma^o(D_k), \Gamma^s(D_k), \boldsymbol{\phi}^d\right) \\
&= B_{D_k}\left(\boldsymbol{\phi}^o_{-D_k}\right) B_{D_k}\left(\Gamma^o(D_k)\right) B_{D_k}\left(\boldsymbol{\phi}^s_{-D_k}\right) \\
&\quad B_{D_k}\left(\Gamma^s(D_k)\right) B_{D_k}\left(\boldsymbol{\phi}^d_{-D_k}\right) B_{D_k}\left(\Gamma^d(D_k)\right). \qquad (5)
\end{aligned}
$$

The expected payoff $\bar{\varpi}_{D_k}$ of each D2D link D_k achieved by its decision ϕ_{D_k} and belief $B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right)$ can be written as

$$
\begin{aligned}
&\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right), \phi_{D_k}\right) = \\
&\quad \bar{\varpi}_{D_k}\left(\langle B_{D_k}\left(\boldsymbol{\phi}^o_{-D_k}\right), B_{D_k}\left(\boldsymbol{\phi}^s_{-D_k}\right), B_{D_k}\left(\boldsymbol{\phi}^d_{-D_k}\right),\right. \\
&\quad \left. B_{D_k}\left(\Gamma^o\right), B_{D_k}\left(\Gamma^s\right), B_{D_k}\left(\Gamma^d\right)\rangle, \langle \phi^o_{D_k}, \phi^s_{D_k}, \phi^d_{D_k}\rangle\right) \\
&= \sum_{\boldsymbol{\phi}_{-D_k} \in \Phi^{K-1}} B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right) \\
&\quad \cdot \left\{\mathbf{1}_{\Gamma^d(D_k)=D_k} \varpi_{D_k}\left[l = \Gamma^s(D_k)\right]\right. \\
&\quad \left. + \mathbf{1}_{\Gamma^d(D_k)=D_n} \varpi_{D_k}\left[l = \Gamma^s(D_k), m = \Gamma^s(D_n)\right]\right\} \qquad (6)
\end{aligned}
$$

Since each D2D link always chooses the decision that maximizes its expected payoff based on its belief, the

decision ϕ_{D_k} of D2D link D_k for a given belief b_{D_k} is given by

$$\phi_{D_k} = \arg \max_{\phi_{D_k} \in \Phi} \bar{\varpi}_{D_k} \left(B_{D_k} \left(\boldsymbol{\phi}_{-D_k}, \Gamma \right), \phi_{D_k} \right). \quad (7)$$

As mentioned previously, allowing overlaps among different coalitions greatly increases the complexity of the traditional non-overlapping coalition formation game. For example, the overlap between coalitions may cause instability and emptiness of the core as shown in the following example.

Example 1: Let us focus on the overlapping actions of four players in two coalitions $\mathcal{C}^1$ and $\mathcal{C}^2$. Let D_1 and D_2 (or D_3 and D_4) be two members of coalition $\mathcal{C}^1$ (or $\mathcal{C}^2$) with exclusive use of the resource in their corresponding coalitions for $\{D_1, D_2\} \subseteq \mathcal{C}^1$ and $\{D_3, D_4\} \subseteq \mathcal{C}^2$. Here we use the term "exclusive" to simplify our discussion. It means that resource sharing between D_1 (or D_2) and any other D2D links in the network does not affect the payoffs of other members in coalition $\mathcal{C}^1$. This can be extended into a more general case. For example, if the spectrum sharing between D_1 and other D2D links in the network can also affect the payoff of some other members in $\mathcal{C}^1$, we can then use D_1 to denote the combined set of all D2D links in coalition $\mathcal{C}^1$ that will be affected by the overlapping action. If D_1 (or D_2) can share its resource with D_3 or D_4, coalitions $\mathcal{C}^1$ and $\mathcal{C}^2$ will overlap with each other. However, if the preference of D_1, D_2, D_3 and D_4 satisfies $D_3 \succ_{D_1} D_2 \succ_{D_1} D_4$, $D_1 \succ_{D_2} D_3 \succ_{D_2} D_4$, $D_2 \succ_{D_3} D_1 \succ_{D_3} D_4$ and $D_1 \succ_{D_4} D_2 \succ_{D_4} D_3$ where we use $\succ$ to denote D_j's preference of a player over different overlapping actions, i.e., $D_i \succ_{D_j} D_k$ means that $D_j \in \mathcal{C}^1$ prefers to overlap with player D_i than D_k for $D_i, D_k \in \mathcal{C}^2$, then we can show that the overlapping $\{D_1, D_2, D_3, D_4\}$ between coalitions $\mathcal{C}^1$ and $\mathcal{C}^2$ is not stable.

The situation observed in the above example is called a rotation (or cycle), which is formally defined as follows.

Definition 4: A *rotation* for a sequence of D2D link preferences is a sequence of D2D links $(\hat{D}_0, \hat{D}'_0), (\hat{D}_1, \hat{D}'_1), \ldots, (\hat{D}_{k-1}, \hat{D}'_{k-1})$ such that $\hat{D}_i \neq \hat{D}_j$ for $i \neq j$ and $\hat{D}_i, \hat{D}_j \in \mathcal{C}^o$, and $\hat{D}'_i$ is the most preferred D2D link for $\hat{D}_i$ and $\hat{D}'_{i+1}$ is the second most preferred D2D link for $\hat{D}_i$ for all $i \in \{1, 2, \ldots, k\}$, where the subscripts are taken modulo k.

As observed in the above example, both of the cores defined in Definitions 2 and 3 can be empty. Finding an effective method to detect the emptiness of the core for a general BOCF game is still an open problem. In the rest of this paper, we can exploit the structure of the cellular networks to find a distributed algorithm to search for the stable and optimal overlapping coalition agreement profile that is in the b-core.

VI. A Hierarchical Matching Algorithm

As observed from the previous section, an optimal overlapping coalition agreement profile is generally difficult to find and it is impossible to enumerate and compare all the possible candidate structures [5]. In this section, we propose a hierarchical matching algorithm to search for the overlapping coalition agreement profile of our game. We divide the DCSS game into different stages. By modeling each stage as a matching market, each D2D link only needs to focus on searching for its optimal decision in each stage. In the beginning, all the D2D links will be first partitioned into L non-overlapping coalitions, each of which corresponds to a group of D2D links that can access the spectrum of the same operator. We can model this problem as a *two-sided many-to-one matching market*, also called a college admission market, in which a set of students is partitioned and admitted into a limited number of colleges (to be discussed in Section VI-A). After being accepted by the operators, the D2D links accepted by the same operator will then compete for sub-bands. We can model this problem as a *two-sided one-to-one matching market*, also called a stable marriage market, in which a set of men will be matched with a set of women (to be discussed in Section VI-B). Finally, D2D links with exclusive use of sub-bands in different coalitions can aggregate and share their sub-bands to further improve their payoffs. We model this problem as a *one-sided one-to-one matching market*, also called a roommate market, in which a set of students will be paired with each other to share the same dormitory.

In our proposed game, the D2D links cannot predict which sub-bands will be finally allocated by each operator or which D2D sub-band sharing partner it will choose. We, however, allow each D2D link to maintain a belief function. We propose a belief updating algorithm in Section VI-D.

A. Operator Selection Algorithm

In this subsection, we assume that each player $D_k \in \mathcal{D}$ can have a fixed private belief function $B_{D_k} \left(\boldsymbol{\phi}_{-D_k}, \Gamma \right)$ about the decisions of other D2D links, and the conflict-solving rules. We will relax this assumption in Section VI-D. Each D2D link first chooses an operator which, according to its belief function, is likely to result in the sub-band allocation that maximizes its payoff. We solve this problem by modeling the operator selection by the D2D links as a two-sided many-to-one matching market with private belief. In this market, a set of D2D links applies for a set of operators. Each D2D link can only choose one operator and each operator i can only provide a limited number of sub-bands called a *quota*, labeled as $q_i = |\mathcal{S}^i|$, for D2D links to access.

Let us now formally define the operator selection market as follows:

Definition 5: An *operator selection market* is a (two-sided many-to-one) matching market with private belief $\mathcal{G}^o = \langle \mathcal{D}, \mathcal{O}, \boldsymbol{B}, \succ \rangle$ consisting of four elements: a set $\mathcal{D}$ of D2D links, a set $\mathcal{O}$ of operators, a vector $\boldsymbol{B} = \langle B_{D_k} \rangle_{D_k \in \mathcal{D}}$ of belief functions, and the preference $\succ$ of each D2D link (or operator) over the operators (or D2D links).

Since the set of D2D links being matched with each operator corresponds to a coalition, the preference relation in the above market coincides with the preference relation of our DCSS game defined in Section V. We use $D_k \succ_i D_n$ to denote that operator i prefers accepting D2D link D_k to D_n and use $i \succ_{D_k} j$ to denote that D2D link D_k prefers to send a request to operator i over sending a request to operator j. Let us define a matching between D2D links and operators as follows:

Definition 6: A (two-sided many-to-one) *matching* Γ^o is a function from the set $\mathcal{D} \cup \mathcal{O}$ into the set of unordered families of elements of $\mathcal{D} \cup \mathcal{O}$ such that $|\Gamma^o(D_k)| = 1$, $|\Gamma^o(i)| \leq q_i$ and $\Gamma^o(D_k) = i$ if and only if D_k is in $\Gamma^o(i)$, for every $i \in \mathcal{O}$ and $D_k \in \mathcal{D}$.

It is worth noting that the operator selection market defined in Definition 5 can also be regarded as a coalitional game [39]. If we let all D2D links fully compete for the cellular sub-bands, the game will turn into a non-cooperative game in which the main solution concept is the Nash equilibrium (NE). As pointed out in [44]–[47], the number of NEs may be large and the NEs are not generally reachable by simple competition among players.

An important concept in matching theory is stability, which is defined as follows.

Definition 7: A matching Γ^o is said to be *m-stable* if the following conditions are satisfied: 1) each player believes that matching Γ^o cannot be strictly improved upon by any individual player or pair of players, 2) each player believes that each of the other players believes matching Γ^o cannot be strictly improved upon by any player or pair.

Note that the concept of stable matching is generally different from the stability of the coalition formation structure in the coalitional game. More specifically, if we say a matching between a D2D link D_k and an operator i is stable, it means that D_k or operator i or both D_k and operator i cannot choose any other matching partner to improve their payoffs. However, we say a coalition formation structure is stable if no coalition (of any size) of D2D links can benefit from deviating and join or form other coalitions. To differentiate between these two concepts, we use *m-stable* to refer to the stability of a matching with private belief. Several different concepts of the core have also been introduced for the matching market in [18], [48]. The core of matching is generally different from the core defined in our coalition formation game in Definition 2. To avoid confusion, in this paper, we only use the term "core" to denote the core of our coalition formation game proposed in Section V.

To find a matching that is m-stable, each D2D link needs to send a request to the operator that according to its beliefs can provide the highest payoff. However, it can be observed in (6) that the payoff of each D2D link depends on its final allocated operator, sub-band and D2D sub-band sharing partner. Therefore, a D2D link cannot know which operator can provide the highest payoff without knowing which sub-band will be eventually allocated by each operator or which D2D sub-band sharing partner it will choose. Fortunately, we can show that each D2D link D_k can establish an estimated version of its resulting payoff obtained from each operator i using its belief function B_{D_k}. More specifically, the estimated payoff of D2D link D_k when it sends the request to operator i is given by

$$\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right), \phi^o_{D_k} = i\right) = \max_{\phi^s_{D_k} \in \mathcal{S} \cup \{\emptyset\}, \phi^d_{D_k} \in \mathcal{D} \cup \{\emptyset\}} \bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}\right), \phi^o_{D_k} = i, \phi^s_{D_k}, \phi^d_{D_k}\right), \quad (8)$$

where $\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}\right), \phi^o_{D_k} = i, \phi^s_{D_k}, \phi^d_{D_k}\right)$ is given in (6).

Using the above result, each D2D link will choose the operator that can maximize its estimated payoff, i.e., $\phi^o_{D_k}$ is given by

$$\phi^o_{D_k} = \arg \max_{i \in \mathcal{O} \cup \{\emptyset\}} \bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right), \phi^o_{D_k} = i\right). \quad (9)$$

We refer to the above equation as the *operator selection algorithm*.

Note that each operator needs to decide whether to allow the requesting D2D links to access its spectrum before knowing which specific sub-band will be requested by each D2D link. We hence assume each operator can accept or reject the requests of the D2D links based on a predefined criterion unrelated to the final sub-band allocated to each D2D link. For example, each operator can evaluate the minimum revenue each D2D link can provide, e.g., we can define the minimum revenue brought by each D2D link D_k to each operator i as $\eta_i(D_k) = \min_{l \in \mathcal{J}^i} \left\{\beta^i_{D_k} \mathbb{EINT}_{D_k}[l]\right\}$. We describe the conflict-solving rule for operator i as follows: *If more than q_i D2D links send requests to the same operator i, a conflict will happen. To resolve this conflict, the operator will only accept the q_i requesting D2D links that can provide the highest minimum revenues*.

We can prove the following results about the operator selection algorithm.

Proposition 1: Suppose the belief of every D2D link is fixed. The operator selection algorithm and the conflict-solving rule of operators result in a unique and m-stable matching between D2D links and operators.

This result follows immediately from the definition of m-stability in Definition 7. We hence omit the details of the proof.

B. Sub-band Selection Algorithm

Each D2D link will decide which specific sub-band to request after being accepted by one operator. We model this problem as a two-sided one-to-one matching market. In this game, a set of D2D links send requests for a set of sub-bands (controlled by an operator), and the operator can then decide whether or not to accept the request from each D2D link according to its conflict-solving rule. To simplify our description, we use $\mathcal{C}^i = \Gamma^o(i)$ to denote the set of D2D links that have been accepted by operator i.

Let us formally define the sub-band selection market as follows:

Definition 8: A *(cellular) sub-band selection market* is a (two-sided one-to-one) matching market with private belief $\mathcal{G} = \langle \mathcal{C}^i, \mathcal{S}^i, \boldsymbol{B}, \succ \rangle$ which consists of a set $\mathcal{C}^i$ of D2D links, a set $\mathcal{S}^i$ of sub-bands controlled by operator i, a vector $\boldsymbol{B} = \langle B_{D_k} \rangle_{D_k \in \mathcal{D}}$ of belief functions, and the preference $\succ$ of each D2D link (or sub-band) over the sub-bands (or D2D links).

Note that, as observed in Section III, to maintain the QoS of the existing cellular subscribers, the accessing of D2D links in each of the sub-bands needs to be strictly controlled by the operators. Therefore, the conflict-solving rule of each sub-band over the D2D links has to be established and maintained by the operators. To simplify our discussion, in this paper, we use the term "conflict-solving rule of each sub-band" to denote the conflict-solving rule of the operator over the sub-bands to be accessed by each D2D link.

We use $l \succ_{D_k} m$ to denote that D2D link D_k prefers accessing sub-band l over sub-band m according to B_{D_k}. Similarly, $D_k \succ_l D_n$ means operator i prefers to let D2D link D_k (as opposed to D_n) access sub-band l. We define a matching between D2D links and cellular subscribers in the spectrum of an operator i as follows.

Definition 9: A (two-sided one-to-one) *matching* with private belief Γ^s between D2D links and sub-bands is a one-to-one correspondence from set $\mathcal{C}^i \cup \mathcal{S}^i$ onto itself such that $\Gamma^s(D_k) \in \mathcal{S}^i \cup \{D_k\}$, $\Gamma^s(l) \in \mathcal{C}^i \cup \{l\}$ and $\Gamma^s(D_k) = l \Leftrightarrow \Gamma^s(l) = D_k$ for every $l \in \mathcal{S}^i$ and $D_k \in \mathcal{C}^i$.

The two-sided one-to-one matching market can be regarded as a special case of the two-sided many-to-one matching market, where a player from either side of the market can only match with one player in the other side of the market. Therefore, we can use exactly the same algorithm to achieve a stable allocation between the D2D links and the sub-bands. That is, similar to the operator selection algorithm, each D2D link should always send the request for the sub-band that can provide the highest payoff. However, each D2D link cannot know its payoff without knowing which sub-band will accept its request or which D2D link will be its D2D sub-band sharing partner. For example, it is possible that a D2D link $D_k \in \mathcal{C}^i$ can obtain a higher payoff by sharing the sub-band occupied by a cellular subscriber than accessing a vacant sub-band without sharing with any other D2D links, i.e., $\varpi_{D_k}[l] \geq \varpi_{D_k}[m]$ for $m \in \mathcal{K}^i$ and $l \in \mathcal{J}^i$. However, this D2D link D_k may achieve a higher payoff by first accessing a vacant sub-band m and then sharing with another D2D link D_n with a sub-band for exclusive use, i.e., $\varpi_{D_k}[m,n] \geq \varpi_{D_k}[l]$ for $m = \Gamma^s(D_k), n = \Gamma^s(D_j)$ and $D_j, D_k \in \mathcal{C}^o$. In other words, if D2D link D_k fails to realize that the possible sub-band sharing with D_n can further improve its payoff, it will choose sub-band l, which is not the sub-band that can provide the highest payoff for D_k. To solve this problem, each D2D link should again exploit its belief function to derive an estimated payoff for each of the sub-bands of its matched operator, i.e., suppose the request sent by D_k to operator i has been accepted. the estimated payoff of D2D link D_k when it decides to send a request to sub-band $l \in \mathcal{S}^i$ is given by

$$\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right), \phi^s_{D_k} = l, \phi^o_{D_k} = i\right) = \max_{\phi^d_{D_k} \in \mathcal{C}^o} \bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right), \phi^o_{D_k} = i, \phi^s_{D_k} = l, \phi^d_{D_k}\right),$$

where $\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}, \Gamma\right), \phi^o_{D_k} = i, \phi^s_{D_k} = l, \phi^d_{D_k}\right)$ is given in (6).

Following the same lines as the operator selection algorithm, each D2D link D_k will decide its sub-band l by

$$\phi^s_{D_k} = \arg \max_{l \in \mathcal{S}^i \cup \{\emptyset\}} \bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k}\right), \phi^s_{D_k} = l, \phi^o_{D_k} = i\right). \tag{10}$$

We refer to the above equation as the *sub-band selection algorithm*.

We also introduce the conflict-solving rule of the sub-band for each operator as follows: *If two or more D2D links send a request for the same sub-band, a conflict will happen. To solve this conflict, the sub-band (or operator) will only allow the D2D link that can provide the higher revenue to access the requested sub-band*.

We have the following results for the sub-band selection algorithm.

Proposition 2: The sub-band selection algorithm and the conflict-solving rule of sub-bands result in a unique and m-stable matching between D2D links and sub-bands of their chosen operator.

The above proposition follows the same line as Proposition 1, and we omit the detailed proof.

C. D2D Selection Algorithm

If sharing sub-bands with cellular subscribers cannot provide adequate payoff for some D2D links (e.g., some D2D links are closely located to some cellular subscribers), they will be given a sub-band for exclusive use and decide whether or not to share the sub-band with other D2D links. In this case, the market will no longer be a two-sided matching market, because each D2D link can find a match with any other D2D link with exclusive use of a sub-band in the entire network. We can then model the problem as a one-sided one-to-one matching market which is defined as follows:

Definition 10: We define the D2D selection market as a one-sided one-to-one matching market with private belief $\mathcal{G} = \langle \mathcal{C}^o, \boldsymbol{B}, \succ \rangle$ where $\boldsymbol{B}$ is the belief function, and $\succ$ is the preference of each D2D link over other D2D links with exclusive sub-bands. We use $D_m \succ_{D_n} D_k$ to denote that D_n prefers D_m to D_k.

Definition 11: A (one-sided one-to-one) *matching* Γ^d between two D2D links is a function from the set $\mathcal{C}^o$ to itself such that $\Gamma^d(D_k) \in \mathcal{C}^o$, $\Gamma^d(D_n) \in \mathcal{C}^o$, and $\Gamma^d(D_k) = D_n \Leftrightarrow \Gamma^d(D_n) = D_k$ for every $D_n, D_k \in \mathcal{C}^o$.

Let us now discuss how to establish the preference for each D2D link when spectrum sharing between two D2D links is allowed in the cellular network. In this case, each D2D link will also need to evaluate and rank its resulting payoffs when sharing a sub-band with another D2D link that also has exclusive access to a sub-band. One way to achieve this is to allow the operators to help the D2D links with vacant sub-bands to discover the existence of each other. As each BS always keeps track of the sub-band allocation of D2D links, it always knows which D2D links have been assigned to vacant sub-bands. The BS can then broadcast this information to all the D2D links. Each D2D link $D_k \in \mathcal{C}^o$ occupying a previously vacant sub-band can then use its belief function defined in (5) to calculate the estimated payoff $\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k},\Gamma\right),\phi^o_{D_k}=i,\phi^s_{D_k}=l,\phi^d_{D_k}=D_n\right)$ when it shares its sub-band l with each of other D2D links (e.g., D_n) with exclusive use of sub-band m for $D_k \neq D_n$ and $D_k, D_n \in \mathcal{C}^o$. Each D2D link can establish its preference about other D2D links with exclusive-use sub-bands by ranking the estimated payoffs from the highest to the lowest values.

Let us denote the preference of each D2D link D_k over other D2D links with exclusive use of sub-bands as $\mathcal{R}^d_{D_k}$. We use $\tilde{v}^m_{D_k}$ to denote the mth preferred D2D link with exclusive sub-band for D_k for $\tilde{v}^m_{D_k} \in \mathcal{C}^o$. We can write $\mathcal{R}^d_{D_k}$ as $\mathcal{R}^d_{D_k} = \langle \tilde{v}^1_{D_k}, \tilde{v}^2_{D_k}, \ldots, \tilde{v}^{|\mathcal{C}^o|}_{D_k}\rangle$ where $\bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k},\Gamma\right),\phi^o_{D_k}=i,\phi^s_{D_k}=l,\phi^d_{D_k}=\tilde{v}^m_{D_k}\right) > \bar{\varpi}_{D_k}\left(B_{D_k}\left(\boldsymbol{\phi}_{-D_k},\Gamma\right), \phi^o_{D_k}=i, \phi^s_{D_k}=l, \phi^d_{D_k}=\tilde{v}^{m+1}_{D_k}\right)\ \forall\ 1 \leq m \leq |\mathcal{C}^o|-1$. Note that if $\tilde{v}^m_{D_k} = D_k$ for $m \leq |\mathcal{C}^o|-1$, it means that D_k cannot obtain any payoff improvement by sharing its sub-band with any D2D link in the set $\{\tilde{v}^{m+1}_{D_k}, \tilde{v}^{m+2}_{D_k}, ..., \tilde{v}^{|\mathcal{C}^o|}_{D_k}\}$.

As described in Example 1, in the D2D selection market, there may not always exist an m-stable matching among all D2D links with sub-bands for exclusive use. One of the main reasons for this is the possible existence of rotations in the resulting preferences. We hence need to find a way to remove the rotations from the possible overlapping coalition agreements. As observed in [6], [18], [22], [49], a stable matching is associated with a unique set of rotations referred as the *observable rotations*. Therefore, if the rotation detection and removal sequence can be uniquely decided, the set of observable rotations as well as the stable matching will also be fixed. This problem can be solved by taking advantage of the labeled identity of each D2D link. More specifically, in a D2D communication network, each D2D link has a specific commonly known identification number, referred to as a *label*, that is used by other D2D links to recognize it. We can then order all D2D links with exclusive sub-bands according to a fixed sequence of their labels, i.e., we denote the ith ordered D2D link as $\tilde{\phi}_i$ and the vector of all the D2D links in $\mathcal{C}$ can be denoted as $\tilde{\boldsymbol{\phi}} = \langle \tilde{\phi}_1, \tilde{\phi}_2, \ldots, \tilde{\phi}_{|\mathcal{C}^o|}\rangle$ for $\tilde{\phi}_i \in \mathcal{C}^o$.

Removing the rotations also requires communication among D2D links with exclusive sub-bands. More specifically, each D2D link will sequentially broadcast a rotation detection signal to determine if a rotation-like sequence can be detected [6], [18], [22], [49]. If a rotation has been detected, all D2D links in the sequence of rotation will remove the rotation from their preference list. If none of the preference lists of the D2D links becomes empty after removing the rotations, each D2D link can then match with its most preferred D2D link in its preference list. Otherwise, no stable matching structure exists. We refer to this algorithm as *D2D Selection Algorithm*. A detailed pseudo-code of the roommate algorithm is given in [6, Figure 4.16].

We have the following results.

Proposition 3: Suppose $\tilde{\boldsymbol{\phi}}$ and the set $\mathcal{C}^o$ of D2D links being allocated vacant sub-bands for exclusive use are fixed. The D2D selection algorithm either reports no m-stable matching exists or generates a unique and m-stable matching structure.

Proof: See the proof of Proposition 3 in [1]. ■

From the above proposition, if the D2D selection algorithm reports a stable matching structure, we can claim the existence of at least one stable matching structure. This can be regarded as a sufficient condition for the existence of a stable matching for the D2D spectrum sharing market. Note that this condition is not necessary because if we change the labeling sequence of D2D links, the resulting matching may also be changed.

D. A Belief Updating Algorithm

The three algorithms discussed in Sections VI-A to VI-C are closely related to each other. More specifically, the matching formed in the operator selection algorithm directly affects the sub-band selection and D2D sub-band sharing among D2D links. Moreover, the results of sub-band and D2D selection algorithms also affect the operator selection of the D2D links. In addition, it is observed in Proposition 1 that if the decision of every D2D link about which operator to send its request to is fixed, the matching between the D2D links and operators will be fixed too. According to Proposition 3, for each of the fixed matchings between D2D link and operators, the sub-band allocation that results from the sub-band selection algorithm is also determined. Finally, if the sub-band allocation among D2D links is fixed, the set of D2D links with exclusive use of sub-bands will be fixed, too. In this case, the results of the D2D selection algorithm will also be fixed. It is the belief functions of all the D2D links that connect these three matching results.

In this subsection, we relax the previous assumption about the fixed belief function of D2D links. We focus on a learning algorithm for each D2D link to iteratively update its belief function according to its previous observations. In our model, each D2D link can eavesdrop on the operators, sub-band and D2D links requested by each of the other D2D links. We assume each D2D link is myopic and hence can use a Dirichlet distribution to

model the uncertainty about the decisions of other D2D links as well as the conflict-solving rules of operators and other D2D links with sub-bands for exclusive use. We can hence apply Bayesian reinforcement learning and use the following equation to calculate the belief about each action of other D2D links at the beginning of each time slot t,

$$B_{D_k,t}\left(\boldsymbol{\phi}^o_{-D_k}\right) = \Pr\left(\boldsymbol{\phi}^o_{-D_k}|\phi^o_{D_k}=i\right) = \frac{\theta_{D_k}\left(\boldsymbol{\phi}^o_{-D_k,t-1}=\boldsymbol{\phi}^o_{-D_k}|\phi^o_{D_k,t-1}=i\right)}{\theta_{D_k}\left(\phi^o_{D_k,t-1}=i\right)}, \quad (11)$$

where $\theta_{D_k}\left(\boldsymbol{\phi}^o_{-D_k,t-1}=\boldsymbol{\phi}^o_{-D_k}|\phi^o_{D_k,t-1}=i\right) = \sum_{u\in\{1,\ldots,t-1\}} \mathbf{1}\left(\boldsymbol{\phi}^o_{-D_k}[u]=\boldsymbol{\phi}^o_{-D_k}|\phi^o_{D_k}[u]=i\right)$ is the number of times that D2D link D_k observes the decisions of other D2D links are equivalent to $\boldsymbol{\phi}^o_{-D_k}$ when its own decision is $\phi^o_{D_k}=i$ during the previous $t-1$ time slots. $\theta_{D_k}\left(\phi^o_{D_k,t-1}=i\right) = \sum_{u\in\{1,\ldots,t-1\}} \mathbf{1}\left(\phi^o_{D_k}[u]=i\right)$ is the number of times D_k sends a request to operator i during the previous $t-1$ time slots.

Similarly, we can write the belief updating algorithm for $B_{D_k}\left(\Gamma^o\right)$ as follows:

$$B_{D_k,t}\left(\Gamma^o\right) = \Pr\left(\Gamma^o(D_k)|\boldsymbol{\phi}^o_{-D_k},\phi^o_{D_k}=i\right) = \frac{\theta_{D_k}\left(\Gamma^o_{t-1}(D_k)=\Gamma^o(D_k)|\boldsymbol{\phi}^o_{-D_k,t-1}=\boldsymbol{\phi}^o_{-D_k},\phi^o_{D_k,t-1}=i\right)}{\theta_{D_k}\left(\boldsymbol{\phi}^o_{-D_k,t-1}=\boldsymbol{\phi}^o_{-D_k},\phi^o_{D_k,t-1}=i\right)}, \quad (12)$$

where $\theta_{D_k}\left(\Gamma^o_{t-1}(D_k)=\Gamma^o(D_k)|\boldsymbol{\phi}^o_{-D_k,t-1}=\boldsymbol{\phi}^o_{-D_k},\phi^o_{D_k,t-1}=i\right)$ is the number of times that D_k has been assigned operator $\Gamma^o(D_k)$ when the decision of D_k is $\phi^o_{D_k}=i$ and the decisions of other D2D links are equivalent to $\boldsymbol{\phi}^o_{-D_k}$ during the previous $t-1$ time slots.

The rest of the belief updating algorithm can be written in a similar fashion:

$$B_{D_k,t}\left(\boldsymbol{\phi}^s_{-D_k}\right) = \frac{\theta_{D_k}\left(\boldsymbol{\phi}^s_{-D_k,t-1}=\boldsymbol{\phi}^s_{-D_k}|\Gamma^o_{t-1}(D_k),\phi^s_{D_k,t-1}=\phi^s_{D_k}\right)}{\theta_{D_k}\left(\Gamma^o_{t-1}(D_k),\phi^s_{D_k,t-1}=\phi^s_{D_k}\right)}, \quad (13)$$

$$B_{D_k,t}\left(\Gamma^s_{-D_k}\right) = \frac{\theta_{D_k}\left(\Gamma^s_{t-1}(D_k)=\Gamma^s(D_k)|\Gamma^o_{t-1}(D_k),\phi^s_{D_k,t-1}=\phi^s_{D_k},\boldsymbol{\phi}^s_{-D_k,t-1}=\boldsymbol{\phi}^s_{-D_k}\right)}{\theta_{D_k}\left(\Gamma^o_{t-1}(D_k),\phi^s_{D_k,t-1}=\phi^s_{D_k},\boldsymbol{\phi}^s_{-D_k,t-1}=\boldsymbol{\phi}^s_{-D_k}\right)}, \quad (14)$$

$$B_{D_k,t}\left(\boldsymbol{\phi}^d_{-D_k}\right) = \frac{\theta_{D_k}\left(\boldsymbol{\phi}^d_{-D_k,t-1}=\boldsymbol{\phi}^d_{-D_k}|\Gamma^o_{t-1}(D_k),\Gamma^s_{t-1}(D_k),\phi^d_{D_k,t-1}=\phi^d_{D_k}\right)}{\theta_{D_k}\left(\Gamma^o_{t-1}(D_k)=i,\Gamma^s_{t-1}(D_k),\phi^d_{D_k,t-1}=\phi^d_{D_k}\right)}, \quad (15)$$

$$B_{D_k,t}\left(\Gamma^d_{-D_k}\right) = \frac{\theta_{D_k}\left(\Gamma^d_{t-1}(D_k)=\Gamma^d(D_k)|\Gamma^o_{t-1}(D_k),\Gamma^s_{t-1}(D_k),\phi^d_{D_k,t-1}=\phi^d_{D_k},,\boldsymbol{\phi}^d_{-D_k,t-1}=\boldsymbol{\phi}^d_{-D_k}\right)}{\theta_{D_k}\left(\Gamma^o_{t-1}(D_k),\Gamma^s_{t-1}(D_k),\phi^d_{D_k,t-1}=\phi^d_{D_k},,\boldsymbol{\phi}^d_{-D_k,t-1}=\boldsymbol{\phi}^d_{-D_k}\right)}. \quad (16)(17)$$

After updating its beliefs, each D2D link uses equation (7) to choose its action.

We can now describe the hierarchical matching algorithm as follows: *At the beginning of each time slot, every D2D link chooses $\phi^o_{D_k}$ using (9). After being matched with the operators, each D2D link chooses $\phi^s_{D_k}$ using (10). If a D2D link has been matched with a sub-band for exclusive use, it uses the D2D sub-band sharing algorithm to decide its sub-band sharing partner. After all D2D links choose their sub-bands and sub-band sharing partners, they use (11)–(17) to update their beliefs and then use the updated belief function to find their matching during the next time slot.*

We now show that the results in Proposition 3 also hold if all the D2D links use the belief updating algorithm in (7). We have the following result about the proposed hierarchical matching algorithm.

Theorem 1: We have the following results:

1) For the resulting belief function of each D2D link, the matching structure achieved by the hierarchical matching algorithm is equivalent to the overlapping coalition agreement $\boldsymbol{x}^{**}$ that is in the b-core of our proposed DCSS game in Section V.
2) Suppose, in some time slot t, the overlapping coalition agreement $\boldsymbol{x}[t]$ satisfies $\boldsymbol{x}[t]=\boldsymbol{x}^{**}$ where $\boldsymbol{x}^{**}$ is the overlapping coalition agreement profile in the b-core based on the true belief (the belief of each D2D link coincides with the true probabilistic features of decisions made by other D2D links and conflict-solving rules of operators and D2D links with exclusive sub-bands) of every D2D link. Then $\boldsymbol{x}[\tau]=\boldsymbol{x}^*,\ \forall\tau>t$.

Proof: First, let us consider the first result. It can be easily observed that if every D2D link D_k can predict the true beliefs of other D2D links, all D2D links can establish the true preferences and use the operator selection algorithm to obtain a unique and stable matching. The D2D links can then use the D2D selection algorithm to generate the unique and stable overlapping coalition agreement. In other words, the resulting coalition formation structure is stable and deterministic for every resulting belief function of D2D links.

We now consider the second result. If $\boldsymbol{x}[t]=\boldsymbol{x}^{**}$ in time slot t, we then have $\varpi_{D_k}\left(\boldsymbol{x}^{**}\right)>\varpi_{D_k}\left(\boldsymbol{x}'\right)$ for $\boldsymbol{x}'$ is not in the core where $\varpi_{D_k}\left(\boldsymbol{x}\right)$ is the payoff of D_k obtained in the sub-band allocated in overlapping coalition agreement $\boldsymbol{x}$. Let us show that in the next time slot $t+1$, each D2D link will stick with $\boldsymbol{x}^{**}$ and will not change to other decisions. In time slot $t+1$, D2D link D_k will update its belief by $B_{D_k,t+1}(\boldsymbol{\phi}_{-D_k},\Gamma) = \alpha B_{D_k,t}\left(\boldsymbol{\phi}_{-D_k},\Gamma\right)+(1-\alpha)\mathbf{1}\left(\phi_{D_k,t+1}=\boldsymbol{\phi}^{**}_{D_k}\right)$, where $\alpha=\frac{t}{t+1}$ and $\boldsymbol{\phi}^{**}_{D_k}$ is the decision of D_k that results in

$\boldsymbol{x}^{**}$. We then can rewrite the updated payoff function of D_k as

$$\bar{\varpi}_{D_k,t+1} = \alpha\bar{\varpi}_{D_k}\left(\phi_{D_k,t}, \boldsymbol{\phi}_{-D_k,t}, B_{D_k,t}\right) + (1-\alpha)\bar{\varpi}_{D_k,t+1}\left(\phi_{D_k,t+1}, \boldsymbol{\phi}_{-D_k,t+1}, B_{D_k,t+1}\right),$$

which is a linear combination of $\bar{\varpi}_{D_k,t}$ and $\bar{\varpi}_{D_k,t+1}$. It can be easily observed that choosing $\boldsymbol{\phi}_{D_k,t+1} = \boldsymbol{\phi}_{D_k,t} = \boldsymbol{\phi}^{**}_{D_k}$ maximizes both payoff functions of D2D link D_k. This process will be repeated in each of the remaining time slots. ■

Proposition 4: For each resulting belief function, the complexity of our hierarchical matching algorithm in the worst case is $O(NK^4L^2)$ where $N = \max_{i\in\mathcal{O}}\{|\mathcal{K}^i|\}$.

Proof: Suppose the belief function of each D2D link has been updated. All D2D links need to first send requests to their preferred operators. According to the conflict-solving rules of the operator, the request sent by a D2D link D_k to operator i can be rejected if operator i has already received q_i or more requests from other D2D links that are preferred by operator i. In the worst case, each of the K D2D links will send requests and be rejected by each of its most preferred $(L-1)$ operators before an operator accepts its request. This results in $K(L-1)$ complexity. Similarly, according to the conflict-solving rule for the sub-bands, each D2D link being accepted by each operator i can also be rejected for each of the $|\mathcal{S}_i|-1$ sub-bands. This results in another $(|\mathcal{C}^i||\mathcal{S}^i|-1)$ complexity for each operator. According to [6], the D2D selection algorithm for D2D links in set $\mathcal{C}^o$ will result in a complexity of $O(|\mathcal{C}^o|^2)$. We hence can claim that the final complexity for each resulting belief function is given by $O(K(L-1)\cdot\sum_{i\in\mathcal{O}}\{|\mathcal{C}^i|\left(|\mathcal{S}^i|-1\right)\}\cdot|\mathcal{C}^o|^2)$. Using the fact that $|\mathcal{C}^i|\leq K$, $|\mathcal{S}^i|\leq N$ and $|\mathcal{C}^o|\leq K$, we obtain the complexity $O(NK^4L^2)$ for our proposed algorithm for each resulting belief function. ■

VII. Numerical Results

In this section, we first describe how to implement our proposed algorithm in LTE-Advanced network systems and then present the numerical results to verify the performance improvement that can be brought by our proposed algorithms.

In a D2D communication system, it is critical for the source and destination of each D2D link to determine each other's availability and ensure they are located within direct communication range. This requires all the potential D2D sources and destinations to first go through a peer device discovery process [26]. This peer device discovery can be either implemented with limited or full control from the operators through the BS as described in [30]. In the limited control approach, each BS periodically broadcasts the set of available vacant and occupied sub-bands that can be used by the D2D links. Each D2D link can then use the received broadcast signal to establish its preference about the operators and then submit a request for the operator and sub-band according to its preference. In the full control approach, each D2D link will simply send a D2D communication request to the BS and the BS will decide the required modes and communication parameters for each D2D link.

In this section, we compare the following four D2D spectrum sharing approaches.

1) *Random Allocation*: D2D links randomly choose operators, modes and sub-bands. In this case, we only allow each D2D link to use modes M1 – M3. This is equivalent to the existing D2D communications in cellular networks without using the optimal mode selection approach studied in [26].
2) *Random Operator Allocation*: each D2D link D_k randomly picks an operator and then uses the sub-band selection algorithm to decide its modes and sub-bands. We again limit each D2D link to choose from modes M1–M3. Therefore, this method is equivalent to the existing D2D communications in cellular networks where each D2D link randomly chooses an operator and then selects the optimal mode introduced in [26].
3) *Hierarchical Allocation*: all D2D links use the operator selection algorithm to choose the optimal operators and then use the sub-band selection algorithm to decide the mode and sub-bands. Again, we assume each D2D link can only choose from modes M1–M3. This method is equivalent to the existing D2D communications in cellular networks where each D2D link chooses the optimal operator and then chooses the optimal mode.
4) *Hierarchical Allocation with Overlaps*: D2D links use the hierarchical matching algorithm to decide their optimal operator, mode and sub-band. Note that in this approach, each D2D link can choose from modes M1–M4.

Note that, as we have proved in Section VI-D, if the D2D links can update their belief functions using (11)–(17), the overlapping coalition agreement of D2D links can converge to a unique and stable structure. In the rest of this section, we focus on the case where D2D links have already updated their belief functions. We will discuss the convergence rate of the belief updating algorithm at the end of this section.

Let us consider a cellular system consisting of multiple operators randomly located in the center region of a square-shaped coverage area, as shown in Figure 3. Each operator has a set of cellular subscribers using its spectrum, which can also be shared with a number of D2D links. D2D links and cellular subscribers are uniformly randomly located in the entire coverage area. To simplify our discussion, we focus on the downlink transmission and assume each D2D link consists of a source and a destination. In a practical system, D2D communication should only be enabled when the source and destination are close to each other. We hence assume each destination is uniformly randomly located within a fixed radius (20 meters in our simulation) of the corresponding source. We consider the payoff of D2D

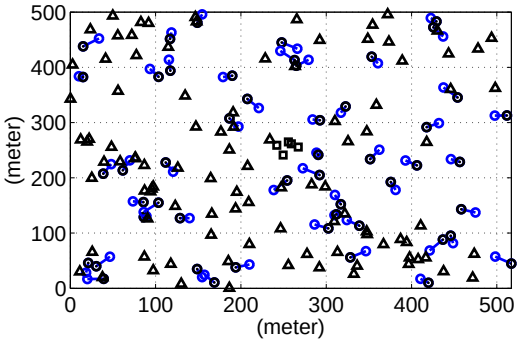

Fig. 3. Simulation setup: we use □ to denote operators, △ to denote cellular subscribers, blue lines to denote the D2D links, blue ◯ to denote D2D transmitters and black ◯ to denote D2D receivers.

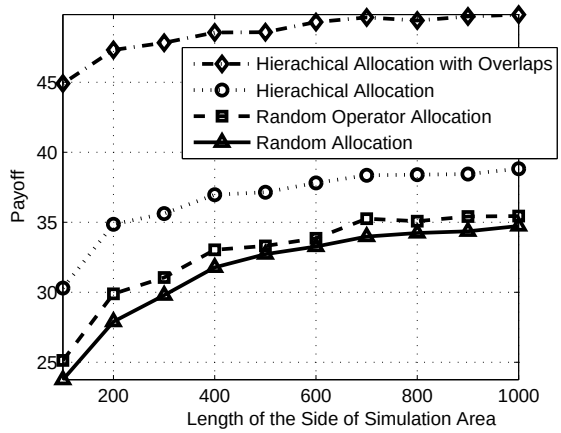

Fig. 4. The total payoff of D2D links under different coverage area sizes ($L = 5, K = 100, |\mathcal{J}^i| = 20\ \forall i \in \mathcal{O}$).

links defined in (1) - (3) and let the channel gain between two D2D links D_k and D_n and one D2D link D_k and one cellular subscriber P_j^i be $h_{D_nD_k} = \frac{\tilde{h}_{D_nD_k}}{\sqrt{d^{\sigma}_{D_nD_k}}}$ and $h_{P_j^iD_k} = \frac{\tilde{h}_{P_j^iD_k}}{\sqrt{d^{\sigma}_{P_j^iD_k}}}$, respectively, where $\tilde{h}_{D_nD_k}$ and $\tilde{h}_{P_j^iD_k}$ are the channel fading coefficients following the Rayleigh random distribution, $d_{D_nD_k}$ and $d^{\sigma}_{P_j^iD_k}$ are the distance between D_n and D_k and P_j^i and D_k, respectively, and σ is the pathloss exponent.

In Figure 4, we fix the number of operators, cellular subscribers and D2D links and present the total payoff of D2D links under different lengths of the side of the square-shaped coverage area with a range from 100 to 1000 meters. Our considered coverage area covers femtocell, pico-cellular ($<$ 200 meters), micro-cellular ($>$ 200 meters), and macro-cellular ($>$ 1000 meters) systems [31]. We observe that the random allocation method achieves the worst payoff among all the methods. Even under the case that each D2D link cannot establish a preference list for the operators but chooses its operator randomly, the payoff of the D2D link can be improved by applying the sub-band allocation algorithm (sub-band selection algorithm in Section VI-B). If we further allow

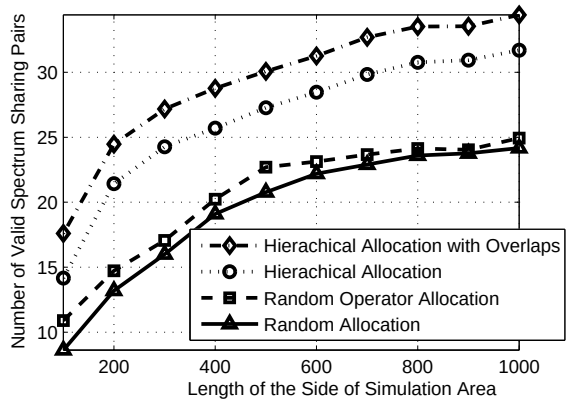

Fig. 5. The number of D2D and cellular or D2D and D2D spectrum sharing pairs for different coverage area sizes ($L = 5, K = 100, |\mathcal{J}^i| = 20\ \forall i \in \mathcal{O}$).

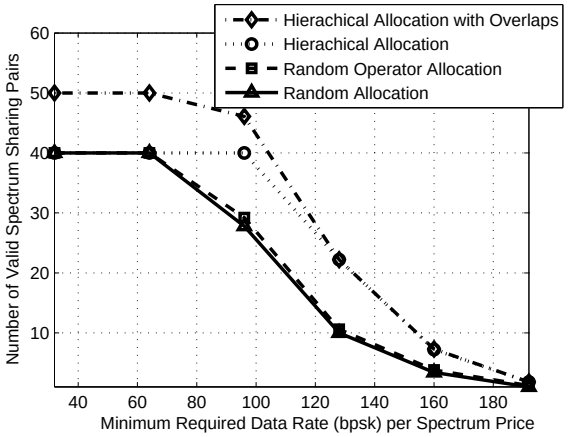

Fig. 6. The number of D2D and cellular or D2D and D2D spectrum sharing pairs for different minimum required data rate (kbits per second) ($L = 5, K = 100, |\mathcal{J}^i| = 20\ \forall i \in \mathcal{O}$).

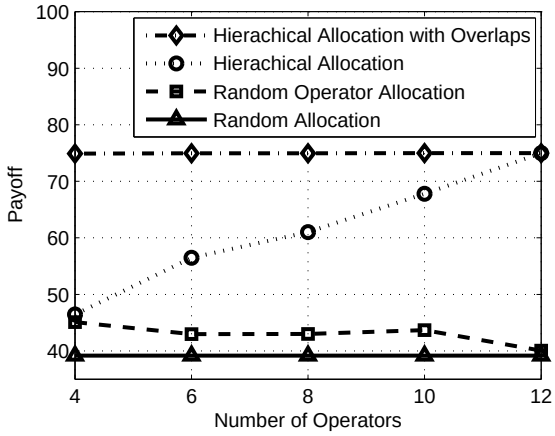

Fig. 7. The total payoff of D2D links for different numbers of operators ($K = 120, |\mathcal{J}| = 120$).

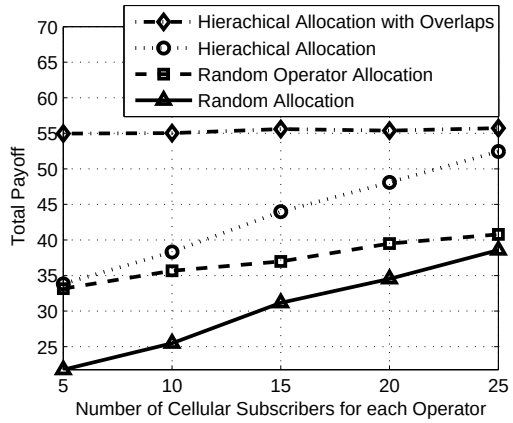

Fig. 8. The total payoff of D2D links for different numbers of cellular subscribers for each operator ($L = 5, K = 100$).

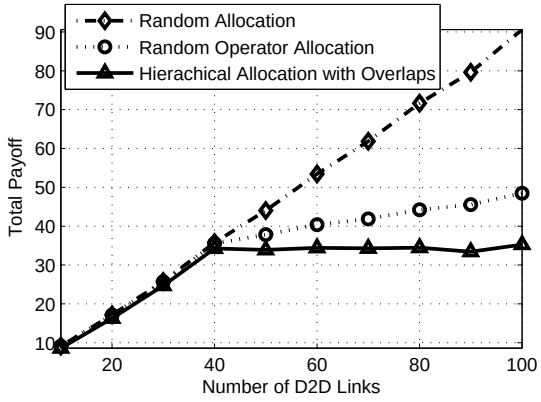

Fig. 9. The total payoff of D2D links for different numbers of D2D links ($L = 5, |\mathcal{J}^i| = 20\ \forall i \in \mathcal{O}$).

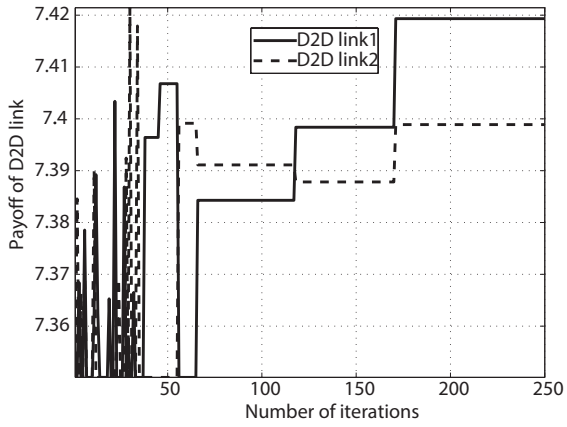

Fig. 10. The convergence rate of payoffs of two D2D links using the belief updating algorithm.

each D2D link to decide its operator using the operator selection market proposed in Section VI-A, the payoff of each D2D link can be further improved. We also observe significant performance improvement by allowing spectrum sharing among D2D links with exclusive sub-bands. This is because the chance for each D2D link with an exclusive-use sub-band to find a suitable sub-band sharing partner increases with the total number of D2D links with sub-bands for exclusive use. If two D2D links with small or even negligible cross-interference can be matched with each other (e.g., two D2D links that are far from each other), the payoff obtained by each of the matching D2D links can be significantly improved. In other words, in a large coverage area with uniformly randomly located D2D links, each D2D link will learn the fact that applying for a exclusive-use sub-band for exclusive use at first and then sharing with another D2D link with vacant sub-bands with small cross-interference can maximize its payoff. Note that in our simulation, we assume each D2D link can always obtain a dedicated sub-band for exclusive use if sharing sub-bands with cellular subscribers cannot achieve a higher payoff. However, in many practical scenarios, the number of vacant sub-bands is limited. In this case, some of the D2D links can only choose between mode M2 and mode M3. In other words, our simulation results of hierarchical allocation with overlaps can be regarded as the upper bound of the payoff achieved by D2D links when they share sub-bands with the cellular networks.

To compare the spectrum sharing capacity in terms of the total number of D2D links that can be supported by the existing cellular system, we present the number of valid spectrum sharing pairs formed between a D2D link and a cellular subscriber or two D2D links in Figure 5. We observe that the hierarchical allocation approaches with and without overlaps can almost double the spectrum sharing capacity, especially in the femtocell or pico-cell cases (coverage length $<$ 200 meters). This is because when the coverage area becomes small, the cross-interference between the spectrum sharing D2D links and cellular subscribers becomes critical and, in this case, choosing the operator serving the cellular subscribers that are far from each D2D link becomes important to improve the spectrum sharing capacity of the systems.

We study the spectrum sharing capacity of the DCSS system under different minimum required data rate per spectrum price in Figure 6. We observe that if each D2D link only requires a data rate below 64 kbps, almost every D2D link can find another cellular subscriber or D2D link to share the spectrum with. However, the spectrum sharing capacity is dramatically decreased when the required data rate for each D2D link exceeds 96 kbps. Furthermore, using the hierarchical allocation with overlaps approach cannot provide any extra capacity improvement for hierarchical allocation if the required minimum data rate becomes larger than 128 kbps. This is because the cross interference between D2D and cellular

communication becomes significant when both D2D links and cellular subscriber raise their transmit powers to support high transmit data rates. Note that, in our simulation, we assume the transmit powers of both D2D links and cellular subscribers are constants and hence the performance of D2D links can be further improved by using optimal transmit powers as shown in [9], [31], [50], [51].

In Figure 7, we fix the number of D2D links and cellular subscribers and consider the payoffs of D2D links under different numbers of operators. It is observed that the payoffs of the D2D links increase with the number of operators when using the hierarchical allocation method. This is because with the increasing number of operators, selecting the proper operators becomes more and more important for each D2D link. However, if we only allow each D2D link to randomly select the operators, the payoff of the D2D links with the random operator selection will approach that of a random allocation method without any optimization.

We fix the number of operators and cellular subscribers to compare the payoffs of D2D links with different numbers of cellular subscribers in Figure 8. It is observed that the payoff of the D2D links increases with the number of cellular subscribers. This is because the cost to the D2D links of accessing an exclusive-use sub-band is higher than that of sharing a sub-band with a cellular subscriber. As the number of subscribers increases, there are more opportunities for D2D links to pair with such subscribers. In addition, the payoff of the hierarchical allocation increases at a faster speed than that of random operator allocation when the number of cellular subscribers to each operator increases.

In Figure 9, we fix the number of operators and cellular subscribers and consider the total payoff of D2D links, varying the number of D2D links in the coverage area. We observe that if the number of D2D links is small, most of the D2D links can find cellular subscribers to share spectrum with and hence allowing spectrum sharing between D2D links with exclusive sub-bands (i.e., hierarchical allocation with overlaps) cannot provide any payoff improvement. However, continuously increasing the number of D2D links provides more choices for each D2D link with an exclusive sub-band when it wants to share its sub-band with other D2D links using the D2D selection market.

The convergence of the belief updating algorithm is illustrated in Figure 10, where we select two D2D links and present their payoffs with hierarchical allocation with overlaps for different iterations. It can be observed that the payoffs of the chosen D2D links can converge to a relatively stable state after the initial fluctuations of the training period.

VIII. Concluding Remarks and Future Works

In this paper, we have considered the spectrum sharing problem between multiple D2D links and a cellular network with multiple operators. We have developed a BOCF game framework to analyze this problem. In our proposed framework, each D2D link will first decide which operator's spectrum it wants to access. All the D2D links being assigned the same operator can be regarded as a coalition and then compete for the available sub-bands controlled by the corresponding operator. Each D2D link can also apply for a vacant sub-band for exclusive use. If there are two or more D2D links with sub-bands for exclusive use, they can further improve their performance by sharing their sub-bands with each other. We propose a hierarchical framework based on a stable matching market to derive a sufficient condition for the core of the BOCF game to be non-empty. We introduce a distributed hierarchical matching algorithm to detect whether the sufficient condition is satisfied and, if satisfied, leads to an overlapping coalition agreement profile that is in the b-core of the game. Numerical results show that our proposed hierarchical matching algorithm can achieve significant performance improvement especially in a large coverage area with a large number of D2D links.

Both the BOCF game framework and the hierarchical matching algorithm can be directly applied to more complex systems. For example, if we also allow three or more D2D links with exclusive sub-bands to share their sub-bands with each other, the overlapping actions of each coalition should consist of all the combinations among the D2D links with vacant sub-bands. Each D2D link will then need to establish a belief function over all the possible combinations between itself and subsets of other D2D links with exclusive sub-bands. Using this belief function, each D2D link will then send the sub-band sharing requests to a group of D2D links which, according to their belief functions, will accept the request and share their sub-band with each other. Another case that can be directly extended from our proposed hierarchical matching algorithm is that of two or more D2D links sharing the same sub-band with cellular subscribers. In our model, we model the sub-band selection problem as a one-sided one-to-one matching market in which each D2D link can only be matched with one sub-band. However, if we model the sub-band selection problem as a one-sided many-to-one matching market as discussed in Section VI-A, each sub-band and its associated cellular subscribers can then be matched with multiple D2D links.

Our work in this paper also opens multiple future directions. One future direction of our research is to study whether it is possible for the operators to also establish and maintain the beliefs about D2D links to further improve their revenues. More specifically, in our model, we mainly focus on the distributed optimization of D2D links and assume the conflict-solving rules of the operators and the D2D links with vacant sub-bands for exclusive use are fixed. It has already been proved in [18], [52], for a two-sided matching market that it is possible for the operators to adjust their conflict-solving rules to further improve their performance. Another

potential direction for future research is to study the effects of allowing partial payment transfers between operators or D2D links on the performance of both D2D links and cellular network systems [53], [54].

REFERENCES

[1] Y. Xiao, K. C. Chen, C. Yuen, and L. A. DaSilva, "Spectrum sharing for device-to-device communications in cellular networks: A game theoretic approach," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Apr., 2014.

[2] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary user behavior in cellular networks and implications for dynamic spectrum access," *IEEE Communications Magazine*, vol. 47, no. 3, pp. 88–95, Mar. 2009.

[3] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding traffic dynamics in cellular data networks," in *IEEE INFOCOM*, Shanghai, China, Apr. 2011.

[4] R. Myerson, *Game theory: analysis of conflict*. Harvard Univ. Press, 1997.

[5] K. Apt and A. Witzel, "A generic approach to coalition formation," *International Game Theory Review (IGTR)*, vol. 11, no. 3, pp. 347–367, 2009.

[6] D. Gusfield and R. W. Irving, *The stable marriage problem: structure and algorithms*. MIT press Cambridge, 1989.

[7] B. Kaufman, J. Lilleberg, and B. Aazhang, "Spectrum sharing scheme between cellular users and ad-hoc device-to-device users," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1038–1049, Mar. 2013.

[8] A. Mukherjee and A. Hottinen, "Energy-efficient device-to-device MIMO underlay network with interference constraints," in *International ITG Workshop on Smart Antennas (WSA)*, Dresden, Germany, Mar. 2012.

[9] M. G. d. S. Rego, T. F. Maciel, H. d. H. Barros, F. R. Cavalcanti, and G. Fodor, "Performance analysis of power control for device-to-device communication in cellular MIMO systems," in *International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, Aug. 2012.

[10] P. Phunchongharn, E. Hossain, and D. I. Kim, "Resource allocation for device-to-device communications underlaying LTE-Advanced networks," *IEEE Wireless Communications*, vol. 20, no. 4, pp. 91–100, Sep. 2013.

[11] A. Osseiran, K. Doppler, C. Ribeiro, M. Xiao, M. Skoglund, and J. Manssour, "Advances in device-to-device communications and network coding for IMT-Advanced," *ICT-Mobile Summit*, Santander, Spain, Jun. 2009.

[12] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[13] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1215–1228, Aug. 2013.

[14] N. Naderializadeh and A. S. Avestimehr, "Itlinq: A new approach for spectrum sharing in device-to-device communication systems," *to appear in IEEE JSAC Special Issue on 5G Wireless Communication Systems, 2014*.

[15] Z. Han and V. Poor, "Coalition games with cooperative transmission: a cure for the curse of boundary nodes in selfish packet-forwarding wireless networks," *IEEE Trans. on Commun.*, vol. 57, no. 1, pp. 203–213, Jan. 2009.

[16] Y. Xiao, G. Bi, D. Niyato, and L. A. DaSilva, "A hierarchical game theoretic framework for cognitive radio networks," *IEEE J. Sel. Area Commun.: Cognitive Radio Series*, vol. 30, no. 10, pp. 2053–2069, Nov. 2012.

[17] W. Saad, Z. Han, R. Zheng, A. Hjorungnes, T. Basar, and H. V. Poor, "Coalitional games in partition form for joint spectrum sensing and access in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 2, pp. 195–209, Apr. 2012.

[18] A. E. Roth and M. A. O. Sotomayor, *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge University Press, 1992.

[19] D. E. Knuth, *Stable marriage and its relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms*. CRM Proceedings and Lecture Notes, vol. 10, 1997.

[20] A. E. Roth, "Deferred acceptance algorithms: History, theory, practice, and open questions," *International Journal of Game Theory*, vol. 36, no. 3, pp. 537–569, Mar. 2008.

[21] K. Iwama, S. Miyazaki, Y. Morita, and D. Manlove, "Stable marriage with incomplete lists and ties," in *the 26th International Colloquium on Automata, Languages and Programming (ICALP)*, ser. Lecture Notes in Computer Science, vol. 1644, pp. 443-452, Prague, Czech Republic, July 1999.

[22] R. W. Irving and D. F. Manlove, "The stable roommates problem with ties," *Journal of Algorithms*, vol. 43, no. 1, pp. 85–105, Apr. 2002.

[23] I. P. Gent and P. Prosser, "An empirical study of the stable marriage problem with ties and incomplete lists," in *the 15th European Conference on Artificial Intelligence (ECAI)*, July 2002.

[24] C.-K. Chan and H.-F. Leung, "Belief-based stability in non-transferable utility coalition formation with uncertainty," *Intelligent Decision Technologies*, vol. 5, no. 2, pp. 151–162, Jun. 2011.

[25] C.-H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," in *IEEE International Conference on Communications (ICC)*, Dresen, Germany, Jun. 2009.

[26] K. Doppler, C.-H. Yu, C. B. Ribeiro, and P. Janis, "Mode selection for device-to-device communication underlaying an lte-advanced network," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Sydney, Australia, Apr. 2010.

[27] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *to appear in IEEE Communications Surveys & Tutorials*, 2014.

[28] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, Mar. 2012.

[29] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 40–48, 2014.

[30] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Communications*, vol. 19, no. 3, pp. 96–104, Jun. 2012.

[31] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.

[32] L. Chen, S. Low, and J. Doyle, "Random access game and medium access control design," *IEEE/ACM Trans. Networking*, vol. 18, no. 4, pp. 1303–1316, Aug. 2010.

[33] L. Tong, V. Naware, and P. Venkitasubramaniam, "Signal processing in random access," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 29–39, Sep. 2004.

[34] A. Al Daoud, T. Alpcan, S. Agarwal, and M. Alanyali, "A stackelberg game for pricing uplink power in wide-band cognitive radio networks," in *IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008.

[35] M. Razaviyayn, M. Yao, and L. Zhi-Quan, "A Stackelberg game approach to distributed spectrum management," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Dallas, TX, Mar. 2010.

[36] Y. Xiao, Z. Han, K. C. Chen, and L. A. DaSilva, "Bayesian hierarchical mechanism design for cognitive radio networks," *to appear at IEEE J. Sel. Area Commun.: Cognitive Radio Series*, vol. 33, no. 4, April 2015.

[37] Y. Xiao, D. Niyato, Z. Han, and K. C. Chen, "Secondary users entering the pool: A joint optimization framework for spectrum pooling," *IEEE J. Sel. Area Commun.: Cognitive Radio Series*, vol. 32, no. 3, pp. 572–588, Mar. 2014.

[38] S. Mumtaz, K. M. S. Huq, and J. Rodriguez, "Direct mobile-to-mobile communication: paradigm for 5g," *IEEE Wireless Communications*, vol. 21, no. 5, pp. 14–23, Oct. 2014.

[39] W. Saad, Z. Han, T. Başar, M. Debbah, and A. Hjorungnes, "Hedonic coalition formation for distributed task allocation among wireless agents," *IEEE Trans. Mobile Computing*, vol. 10, no. 9, pp. 1327–1344, Sep. 2011.

[40] K. Akkarajitsakul, E. Hossain, and D. Niyato, "Distributed resource allocation in wireless networks under uncertainty and application of Bayesian game," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 120–127, 2011.

[41] A. Leshem and E. Zehavi, "Game theory and the frequency selective interference channel," *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 28–40, Sep. 2009.
[42] G. Chalkiadakis, E. Elkind, E. Markakis, M. Polukarov, and N. R. Jennings, "Cooperative games with overlapping coalitions," *Journal of Artificial Intelligence Research*, vol. 39, no. 1, pp. 179–216, 2010.
[43] J.-P. Aubin, "Cooperative fuzzy games," *Mathematics of Operations Research*, vol. 6, no. 1, pp. 1–13, Feb. 1981.
[44] D. P. Foster and H. Young, "Learning, hypothesis testing, and nash equilibrium," *Games and Economic Behavior*, vol. 45, no. 1, pp. 73 – 96, Mar. 2003.
[45] M. Bowling and M. Veloso, "Rational and convergent learning in stochastic games," in *International joint conference on artificial intelligence*, vol. 17, no. 1, Seattle, WA, Aug. 2001.
[46] E. Kalai and E. Lehrer, "Rational learning leads to nash equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 1019–1045, Sep. 1993.
[47] A. Leshem, E. Zehavi, and Y. Yaffe, "Multichannel opportunistic carrier sensing for stable channel access control in cognitive radio systems," *IEEE J. Select. Areas Commun.*, vol. 30, no. 1, pp. 82–95, Jan. 2012.
[48] A. E. Roth, "The evolution of the labor market for medical interns and residents: a case study in game theory," *The Journal of Political Economy*, vol. Y2, no. 6, pp. 991–1016, 1984.
[49] T. Fleiner, R. W. Irving, and D. F. Manlove, "Efficient algorithms for generalized stable marriage and roommates problems," *Theoretical computer science*, vol. 381, no. 1, pp. 162–176, Aug. 2007.
[50] Y. Xiao, G. Bi, and D. Niyato, "A simple distributed power control algorithm for cognitive radio networks," *IEEE Trans. Wireless Commun.,*, vol. 10, no. 11, pp. 3594 –3600, Nov. 2011.
[51] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *IEEE International Conference on Communications (ICC)*, Seattle, WA, Jun. 1995.
[52] P. Biró and G. Norman, "Analysis of stochastic matching markets," *International Journal of Game Theory*, vol. 42, no. 4, pp. 1021–1040, Nov. 2013.
[53] T. Fleiner, "A fixed-point approach to stable matchings and some applications," *Mathematics of Operations Research*, vol. 28, no. 1, pp. 103–126, Feb. 2003.
[54] J. W. Hatfield and P. R. Milgrom, "Matching with contracts," *American Economic Review*, pp. 913–935, Sep. 2005.

Yong Xiao (S'11-M'13) received his B.S. degree in electrical engineering from China University of Geosciences, Wuhan, China in 2002, M.Sc. degree in telecommunication from Hong Kong University of Science and Technology in 2006, and his Ph. D degree in electrical and electronic engineering from Nanyang Technological University, Singapore in 2012. From August 2010 to April 2011, he was a research associate in school of electrical and electronic engineering, Nanyang Technological University, Singapore. From May 2011 to October 2012, he was a research fellow at CTVR, school of computer science and statistics, Trinity College Dublin, Ireland. From November 2012 to December 2013, he was a postdoctoral fellow at Massachusetts Institute of Technology. From December 2013 to November 2014, he was a MIT-SUTD postdoctoral fellow with Singapore University of Technology and Design and Massachusetts Institute of Technology.

Currently, he is a postdoctoral fellow II at Department of Electrical and Computer Engineering at University of Houston. His research interests include machine learning, game theory and their applications in communication networks.

Kwang-Cheng Chen (M.'89-SM.'94-F.'07) received the B.S. from the National Taiwan University in 1983, and the M.S. and Ph.D from the University of Maryland, College Park, United States, in 1987 and 1989, all in electrical engineering. From 1987 to 1998, Dr. Chen worked with SSE, COMSAT, IBM Thomas J. Watson Research Center, and National Tsing Hua University, in mobile communications and networks. Since 1998, Dr. Chen has been with National Taiwan University, Taipei, Taiwan, ROC, and is the Distinguished Professor and Associate Dean in academic affairs for the College of Electrical Engineering and Computer Science, National Taiwan University. Dr. Chen has been actively involved in the organization of various IEEE conferences as General/TPC chair/co-chair. He has served in editorships with a few IEEE journals and many international journals and has served in various positions within IEEE. Dr. Chen also actively participates in and has contributed essential technology to various IEEE 802, Bluetooth, and 3GPP wireless standards. He has authored and co-authored over 250 technical papers and more than 20 granted US patents. He co-edited (with R. DeMarca) the book Mobile WiMAX published by Wiley in 2008, and authored the book Principles of Communications published by River in 2009, and co-authored (with R. Prasad) another book Cognitive Radio Networks published by Wiley in 2009. Dr. Chen is an IEEE Fellow and has received a number of awards including the 2011 IEEE COMSOC WTC Recognition Award and has co-authored a few award-winning papers published in the IEEE ComSoc journals and conferences. Dr. Chen's research interests include wireless communications and network science.

Chau Yuen received the B.Eng. and Ph.D. degrees from Nanyang Technological University, Singapore, in 2000 and 2004, respectively. In 2005, he was a Postdoctoral Fellow with Bell Labs, Alcatel-Lucent, Murray Hill, NJ, USA. From 2006 to 2010, he was a Senior Research Engineer with the Institute for Infocomm Research, Singapore, where he was involved in an industrial project on developing an 802.11n wireless local area network system and participated actively in the Third-Generation Partnership Project Long-Term Evolution and LTE Advanced standardization. In 2008, he was a Visiting Assistant Professor with The Hong Kong Polytechnic University, Kowloon, Hong Kong. Since June 2010, he has been an Assistant Professor with the Department of Engineering Product Development, Singapore University of Technology and Design, Singapore. Dr. Yuen serves as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He received the Lee Kuan Yew Gold Medal, the Institution of Electrical Engineers Book Prize, the Institute of Engineering of Singapore Gold Medal, the Merck Sharp & Dohme Gold Medal, the Hewlett Packard Prize (twice), and the 2012 IEEE Asia-Pacific Outstanding Young Researcher Award.

Zhu Han (S'01-M'04-SM'09) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor in Boise State University, Idaho. Currently, he is an Associate Professor in Electrical and Computer Engineering Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communication. Dr. Han is an Associate Editor of IEEE Transactions on Wireless Communications since 2010. Dr. Han is the winner of IEEE Fred W. Ellersick Prize 2011. Dr. Han is an NSF CAREER award recipient 2010.

Luiz A. DaSilva (SM) is the Professor of Telecommunications at Trinity College Dublin. He also holds a research professor appointment in the Bradley Department of Electrical and Computer Engineering at Virginia Tech, USA. His research focuses on distributed and adaptive resource management in wireless networks, and in particular wireless resource sharing, dynamic spectrum access, and the application of game theory to wireless networks. He is currently a Principal Investigator on research projects funded by the National Science Foundation in the United States, the Science Foundation Ireland, and the European Commission under Horizon 2020 and Framework Programme 7. He is a Co-principal Investigator of CONNECT, the Telecommunications Research Centre in Ireland. Prof DaSilva is an IEEE Communications Society Distinguished Lecturer.