Dynamic Energy Trading for Energy Harvesting Communication Networks: A Stochastic Energy Trading Game

Yong Xiao, *Senior Member, IEEE*, Dusit Niyato, *Senior Member, IEEE*, Zhu Han, *Fellow, IEEE*, and Luiz A. DaSilva, *Senior Member, IEEE*

***Abstract*—This paper studies energy harvesting communication systems in which different energy harvesting devices (EHDs) can harvest different amounts of energy and transmit different numbers of data packets in different time slots. We introduce a dynamic energy trading framework that allows the EHDs to transfer and trade their harvested energy with each other. The EHDs are divided into two groups: seller EHDs that can harvest more energy than they can use, and buyer EHDs, which cannot harvest sufficient energy to support their required communication services. In the proposed framework, the role of each EHD as a seller EHD or a buyer EHD as well as the amount of energy that each EHD can buy or sell to others change over time. Each EHD cannot observe complete information regarding the harvested energy or the number of data packets transmitted by other EHDs. We introduce a simple energy trading scheduling protocol for the EHDs to discover their nearby EHDs and establish energy trading links with each other. We formulate a new game theoretic model called *stochastic energy trading game* to analyze the dynamic energy trading among EHDs in a stochastic environment. We derive an optimal energy trading policy for each EHD to sequentially optimize its decisions. We prove that the proposed policy can achieve a stable and optimal sequence of matchings between buyer and seller EHDs. We present numerical results to compare our proposed energy trading policy with an existing transmit packet scheduling approach, under various network settings and conditions.**

***Index Terms*—Energy harvesting, energy trading, stable marriage, belief update, stable matching, communication networks, game theory, stochastic game, wireless power transfer.**

I. Introduction

An energy harvesting communication system allows communication among energy harvesting devices (EHDs) to be powered by the energy harvested from the natural environment, such as sunlight, wind, radio wave, and vibration. It is of significant research interest because of its potential to provide a ubiquitous and sustainable energy supply for wireless communication networks where power charging/recharging is not always feasible. One of the main challenges for energy harvesting communication systems is that the uncertainty of the natural environment makes it difficult to provide a reliable energy supply. Additionally, the sources of energy harvested by the EHDs are highly random and can be affected by many unpredictable and uncontrollable factors. For example, different radio frequency (RF) energy harvesting-based EHDs in the same area can harvest substantially different amounts of energy because of their different orientations, energy conversion efficiency, antenna locations, etc.

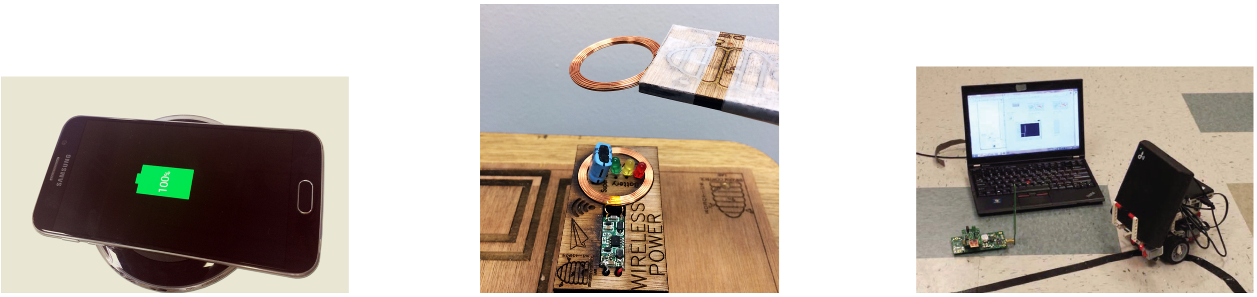
Wireless energy transfer has been introduced as a simple and effective solution to provide reliable energy sources for energy harvesting communication systems. In this approach, energy can be wirelessly transferred from external sources such as power beacons, mobile energy stations, and EHDs with transferable energy, to each EHD. Common energy transfer technologies include inductive coupling, RF energy transfer, and magnetic resonance coupling. Inductive coupling is primarily limited to short-distance energy transfer applications due to its high energy loss in long-distance transfer (see an inductive coupling charger for phones in Figure 1 (a)). In magnetic resonance coupling-based energy transfer communication systems, a coil is installed in the energy transmitter to generate a magnetic field that can traverse to the coil installed in the receiver and generate a current to power data communications (see magnetic resonance coils in Figure 1 (b)). Magnetic resonance coupling does not have adverse health effects to the human body or cause interference to the energy harvesting process and data communication services. Magnetic resonance-based energy transfer is also not affected by obstructions including metal, wood, human body, electronic devices etc., between the transmitter and receiver [1]–[3]. RF energy transfer also suffers from severe propagation loss during long distance energy transfer and has the potential to cause interference to the existing telecommunication services. New technologies such as MIMO and beamforming have recently been applied in RF energy transfer to improve the energy transfer efficiency and mitigate the interference to unintended communication devices [4], [5] (see a Powercast RF charging sensor with RF charging device on a robotic vehicle in Figure 1 (c)). The University of Houston has all the above equipments and has performed many experiments for energy transfer with the above three methods [6]–[9].

This work is partially supported by the National Science Foundation under Grants CCF-1456921, CNS-1443917, ECCS-1405121, ECCS-1547201, NSFC-61428101, and the Science Foundation Ireland under Grant 10/IN.1/I3007.

Y. Xiao and Z. Han are with the Department of Electrical and Computer Engineering at the University of Houston, TX 77004, USA (e-mails: xyong.2012@gmail.com and zhan2@uh.edu).

D. Niyato is with the School of Computer Engineering at the Nanyang Technological University, 639798, Singapore (e-mail: dniyato@ntu.edu.sg).

L. A. DaSilva is with CONNECT, Trinity College Dublin, Dublin 2, Ireland (e-mail: dasilval@tcd.ie). L. A. DaSilva is also with Virginia Tech, VA, USA.



(a) Inductive coupling. (b) Magnetic resonant coupling. (c) RF energy transfer.

Fig. 1: Various wireless transfer-enabled devices at the University of Houston.

In this paper, we introduce a new concept referred to as the *dynamic energy trading market*. In this market, each EHD is deployed with an energy harvester that can collect energy from the nearby natural environment such as photovoltaic cells for harvesting energy from sunlight and a piezoelectric converter for harvesting energy from vibration, and with energy transfer devices such as magnetic resonant coils for magnetic resonance coupling-based energy transfer, and RF energy transmitter and receiver. Thus, the EHDs can transfer and receive energy to and from each other. We consider a system in which each EHD can harvest energy from the environment, trade energy with other EHDs, and transmit data packets to its corresponding destinations. The energy harvesting, trading, and the data communication between the EHDs are assumed not to interfere with each other. Dynamic energy trading takes advantage of the diversity of the energy harvested by multiple closely located EHDs by dividing the EHDs, in each time instant, into two groups: *seller EHDs* and *buyer EHDs*. The EHDs that can harvest more energy than they require and transfer their excess energy to others are called seller EHDs. The EHDs that cannot harvest sufficient energy to support their communication services and have to obtain energy from other EHDs are referred to as buyer EHDs. Compared to most existing energy harvesting communication systems, our proposed dynamic energy trading market has the following unique features:

1) Because of the time-varying environment, the amount of harvested energy, number of arriving data packets and the amount of energy required to send data packets are different from time to time. Therefore, the role of each EHD as a seller EHD or a buyer EHD, as well as the amount of energy that each buyer EHD requires or each seller EHD can provide also change with time. We model each EHD as having different objectives and action spaces when it serves as a buyer or seller EHD.
2) The performance of each EHD is affected not only by the environment but also by its interactions with other EHDs. This makes it natural to study dynamic energy trading using game theoretic tools. However, most existing game models require the action space and objective function of each player to be fixed and hence cannot be directly applied to analyze our proposed dynamic energy trading market. Moreover, many game theoretic models focus on finding pure strategy equilibrium solutions which may not always exist, and even if they exist, are generally not optimal [10], [11]. In other words, there is still the lack of an appropriate game theoretic framework to model and analyze the dynamic energy trading problem.
3) Each EHD cannot observe or keep track of complete information about the energy harvesting and transfer process of all other EHDs. Furthermore, in practical energy harvesting systems, only limited information can be exchanged among EHDs, and therefore how to distributedly optimize the decisions of the EHDs using partially observed information is a challenging task.

We consider an energy trading market in which all EHDs aim to maximize their payoffs by trading energy with each other. We focus on the sequential optimization of the pairing between buyer and seller EHDs with the progression of the energy harvesting communication process. More specifically, we propose a novel game theoretic framework, which we refer to as a *stochastic energy trading game*. This game can be regarded as a special case of a stochastic game where the interactions between the buyer and seller EHDs at each specific time are modeled as a stable matching game with private belief. One of the main advantages of fitting stable matching into the stochastic game is that, unlike stochastic games in which a pure strategy equilibrium solution cannot be guaranteed to exist, and even it exists, can be too complex to reach [11], in a (two-sided) stable matching game, a stable matching structure always exists [12]. In our proposed game, each buyer EHD can establish and sequentially update its beliefs about the current and future energy harvesting process of the other EHDs, the final matching structure, and the state of the system. Each buyer EHD will then establish its preference over seller EHDs using this belief function. We adopt a stochastic game framework in which each EHD can estimate the future progression of the environment as well as its future interactions with other EHDs. We seek a sequence of matchings that maximize the long-term average performance of each EHD. We then derive an optimal energy trading policy that allows each EHD to learn from its past experience and sequentially optimize the selection of its matching partner. We propose a distributed algorithm which can implement this policy without requiring multiple rounds of back-and-forth negotiations between buyer and seller EHDs. Finally, we compare the performance of our proposed energy trading policy against transmit packet scheduling policies proposed in the literature. Numerical results show that our proposed dynamic energy trading policy can significantly improve the performance of energy harvesting communication systems, especially for delay-sensitive applications.

The main contributions of this paper are:

1) We propose a novel *dynamic energy trading market* to allow EHDs with different data transmission requirements that can harvest different amounts of energy to help each other during the communication process. Our framework capitalizes on the diverse energy harvesting processes among EHDs to improve the performance of energy harvesting communication networks.
2) We adopt a link establishment protocol from ad hoc peer-to-peer communication networks and introduce a *distributed energy trading scheduling protocol* for each EHD to discover the active buyer and seller EHDs in the networks as well as to establish energy trading links with other EHDs.
3) We introduce a new game theoretic framework, referred to as the *stochastic energy trading game*, to analyze the dynamic energy trading market. Our model extends stochastic games by modeling the interaction between buyer and seller EHDs at each time as a stable matching game with private belief. To the best of our knowledge, this is the first work that combines stable matching games with private belief and stochastic games to analyze energy harvesting communication networks.
4) We derive the *optimal energy trading policy* that allows each EHD to optimize its choice of energy trading partner to maximize its long-term average performance. Our proposed policy does not require each EHD to obtain complete information about the energy harvesting processes of the other EHDs.
5) We compare our proposed energy trading policy against an existing transmit packet scheduling policy. Our numerical results show that, compared with the latter policy, our proposed energy trading policy can significantly improve the performance of EHDs, especially in delay-sensitive applications.

The remainder of this paper is organized as follows. The background and related work are reviewed in Section II. The system model is described in Section III. The framework of dynamic energy trading market and energy trading scheduling protocol are introduced in Sections IV and V, respectively. In Section VI, we introduce the stochastic energy trading game and derive the optimal energy trading policy. We present the numerical results and compare our energy trading policy with the transmit packet scheduling policy in Section VII. Possible extension and future work are discussed in Section VIII. We conclude the paper in Section IX.

II. Related Works

It has been observed that by exploiting knowledge about the future energy harvesting process, the performance of energy harvesting communication can be significantly improved. More specifically, authors in [13] studied the packet scheduling problem for a deterministic single-user energy harvesting communication system in which a transmitter equipped with an infinite capacity battery can precisely predict the energy that can be harvested and the data packets that will arrive in the future. That work also proposed an off-line algorithm to optimally schedule the transmission of data packets to minimize the total transmission time. This problem has been further studied in a stochastic environment in which the EHD cannot perfectly track the evolution of the energy harvesting process but can know the statistics of the progression of the energy harvesting process. For example, the power allocation problem was studied in [7], [14] where the energy harvesting process was modeled as a Markov decision process (MDP). In [15], the power control problem for an energy harvesting-enabled transmitter was modeled as a partially observable Markov decision process (POMDP). Cases for which the EHD does not know the statistics of the energy harvesting process were studied in [6], where a Bayesian reinforcement learning approach was proposed for the energy harvester to learn these statistics from previous experience. A detailed review of recent advances in energy harvesting technologies applied to wireless communications is given in [16].

With recent advances in wireless power conversion and transfer technologies, wireless energy harvesting and transfer equipment such as wireless mobile phone chargers have already been deployed in wireless communication devices [17]. In RF energy transfer-based systems, the wireless information signal can be embedded into the RF energy transfer signal to achieve simultaneous wireless information and power transfer (SWIPT) [18]. However, in practice, RF energy transfer will cause interference to the information signal transmission, which results in a tradeoff between energy transfer and information transmission [4]. Recent developments in magnetic resonance coupling technology have significantly improved the efficiency of wireless power transfer [1], [2], [19], [20]. More specifically, authors in [1] have demonstrated magnetic resonance coupling-based wireless transfer of 60W of power over 2 meter distances with 40% transfer efficiency. Authors in [2] have also demonstrated that using similar technologies it is possible to transfer 10 kilowatts of power for a distance of 6.5 feet. A mobile phone charging system called Magnetic MIMO has been developed in [19] to wirelessly charge mobile phones and other portable devices regardless of the orientation of these devices.

Recently, multi-user energy harvesting networking systems, especially cooperative energy harvesting systems with energy transfer among EHDs, have been introduced as a solution to improve the reliability of the energy supply for energy harvesting communication systems [21]–[24]. The multi-hop relay channel with one-way energy transfer from the source to relay node has been studied from the information theoretic perspective in [22]. In [25], an interactive POMDP-based framework was proposed to study

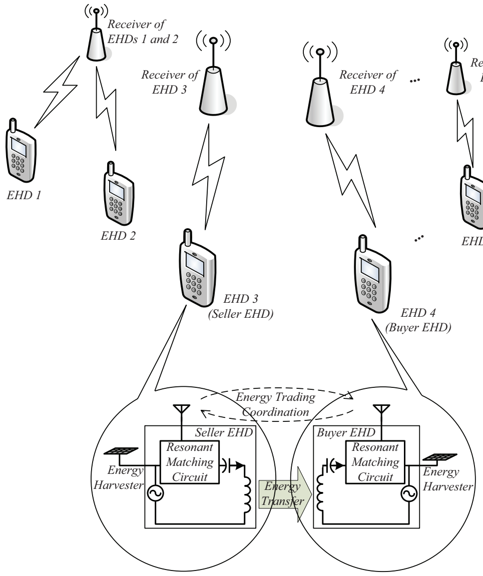


Fig. 2: A dynamic energy trading market for energy harvesting communication systems.

the relay selection problem for a cooperative energy harvesting system. A hybrid network architecture with co-located power beacons and cellular networks was studied in [26].

In this paper, we propose a stochastic energy trading game to analyze dynamic energy trading among energy harvesting EHDs that can exchange energy with each other. Our proposed model is extended from our previous works [27]–[29] where we proposed a stable matching game with private belief to study the resource allocation problem for cognitive radio networks operating in a stationary environment. Different from these works, where both the environment and the action space of each player are assumed to be fixed, in this paper we consider dynamic energy trading for energy harvesting communication systems in a stochastic environment. Due to the time-varying environment, the role of each player as well as the matchings between buyer and seller EHDs can change from time to time. Therefore, we fit the stable matching game with private belief into a stochastic game framework.

III. Network Model

We consider an energy harvesting communication system with a set of K EHDs, labeled as $\mathcal{U} = \{1, 2, \ldots, K\}$ as shown in Figure 2. Each EHD corresponds to the transmitter of a data communication link with an energy harvester. We assume that time is divided into slots, during each of which the amount of energy harvested by each EHD can be regarded as fixed. Let $\hat{u}_{k,t}$ be the number of data packets that arrive at EHD k at the beginning of each time slot t and $\hat{e}_{k,t}$ be the amount of energy that can be harvested by EHD k during time slot t. Each EHD k has a data buffer and a battery that can store up to $\bar{u}_k$ data packets and $\bar{e}_k$ energy units, respectively. In this paper, we focus on an energy harvesting communication system with causal constraints. Specifically, each EHD cannot transmit data packets or use the energy that will only be available in the future. Thus, we can write the data buffer levels of EHD k at the beginning of time slot t as

$$u_{k,t} \quad = \quad \min\left\{\bar{u}_k, \hat{u}_{k,t} + u_{k,t-1} - v_{k,t-1}\right\}, \qquad (1)$$

where $v_{k,t-1}$ is the number of data packets sent by EHD k during time slot $t-1$, with $0 \le v_{k,t-1} \le u_{k,t-1}$. We assume that each EHD is also installed with a power transfer equipment and can send or receive a certain amount of energy to or from other EHDs at the beginning of each time slot. Let $\Delta\tilde{e}_{k,t} \ge 0$ be the amount of energy that can be successfully received by EHD k from other EHDs at the beginning of time slot t. In this paper, we assume that each EHD cannot use the energy harvested during the current time slot to transmit its data packets in the same time slot. The battery level of EHD k at the beginning of time slot t can be written as follows:

$$e_{k,t} \quad = \quad \min\left\{\bar{e}_k, \hat{e}_{k,t-1} + e_{k,t-1} + \Delta\tilde{e}_{k,t} - w_{k,t-1}\right\}, (2)$$

where $w_{k,t-1}$ is the energy consumed by EHD k to send $v_{k,t-1}$ data packets, with $0 \le w_{k,t-1} \le e_{k,t-1}$.

Note that the amount of harvested energy is generally a continuous variable. However, due to the limit of the accuracy of digital communication devices, we can assume that the amount of energy harvested by each EHD during each time slot is a discrete value (e.g., the smallest unit of energy to send a data packet) taken from a finite set. We can denote as $\mathcal{E}$ the set of possible levels of the battery for each EHD k, i.e., $e_{k,t} \in \mathcal{E}$, $\forall\ k \in \mathcal{U}$ and $t \ge 0$. Note that due to the energy transfer loss, if EHD j transfers an amount of energy $q^k_{j,t}$ to EHD k, only $\Delta\tilde{e}_{k,t} = g_{jk,t}q^k_{j,t}$ can be successfully received by EHD k, with $0 < g_{jk,t} < 1$, $k \ne j$ and $k, j \in \mathcal{U}$. We assume that $g_{jk,t}$ also takes values from a finite set. Let $\mathcal{W}$ and $\mathcal{Z}$ be the sets of possible energy used to send data packets and the data buffer levels of each EHD during each time slot. We refer to the battery levels, buffer levels of all the EHDs and the energy transfer efficiencies among EHDs at the beginning of each time slot t as the (environment) *state*, denoted as $\eta_t = \langle \boldsymbol{e}_t, \boldsymbol{u}_t, \boldsymbol{g}_t \rangle$ where $\boldsymbol{e}_t = \{e_{k,t}\}_{k\in\mathcal{U}}$, $\boldsymbol{u}_t = \{u_{k,t}\}_{k\in\mathcal{U}}$ and $\boldsymbol{g}_t = \{g_{kj,t}\}_{k,j\in\mathcal{U}}$. Let Υ be the set of possible states, i.e., we have $\eta_t \in \Upsilon$ and $\Upsilon = \mathcal{E} \times \mathcal{Z}$.

IV. A Dynamic Energy Trading Market

To overcome the adverse effects caused by the time-varying environment and exploit the diverse energy harvesting process of the EHDs, we introduce the concept of *dynamic energy trading market*, in which each EHD can improve its performance using the following approaches:

1) *Transmission Scheduling*: Each EHD k can schedule its energy use and the number of packets to be transmitted in each time slot t according to the statistics of the future evolution of the environment and interaction with other EHDs.
2) *Energy Trading*: The EHDs can trade their harvested energy with each other. More specifically, those EHDs that cannot harvest enough energy to support their data

communication can request a certain amount of energy to be transferred from others EHDs that have harvested more energy than they can use at the beginning of each time slot.

In the dynamic energy trading market, the set of EHDs can be divided into the following two subsets in each time slot:

1) The subset of *seller EHDs* includes the EHDs that believe their stored energy and the energy that will be harvested in the future will be more than sufficient to support their required transmissions. Let $\bar{q}_{j,t}$ be the maximum amount of energy that can be sent by seller EHD j at the beginning of time slot t. Let $\mathcal{S}_t$ be the set of seller EHDs in time slot t, i.e., $\mathcal{S}_t = \{j : \bar{q}_{j,t} > 0, \forall j \in \mathcal{U}\}$.
2) The subset of *buyer EHDs* includes the EHDs that believe the energy they can obtain will not be able to support the required data transmission. Each buyer EHD k can request a certain amount of energy denoted as $q^k_{j,t}$ from another EHD j at the beginning of time slot t. Let $\mathcal{B}_t$ be the set of buyer EHDs in time slot t, i.e., $\mathcal{B}_t = \{k : q^k_{j,t} > 0, \forall k \in \mathcal{U} \text{ and } j \in \mathcal{S}_t\}$. We have $\mathcal{S}_t \cup \mathcal{B}_t \subseteq \mathcal{U}$ and $\mathcal{S}_t \cap \mathcal{B}_t = \emptyset, \forall t$. Note that if $q^k_{j,t} > 0$ and $\Delta\tilde{e}_{k,t} = 0$, it means that the request sent by buyer EHD k has been rejected by seller EHD j and hence no energy will be transferred from EHD j to EHD k. If $q^k_{j,t} = 0$ $\forall k \in \mathcal{B}_t$, it means that no buyer EHD sends an energy request to EHD j in time slot t. We have $\Delta\tilde{e}_{k,t} = g_{jk,t} \min\{q^k_{j,t}, \bar{q}_{j,t}\}$.

It can be observed that in each time slot there may exist EHDs that neither sell nor buy energy to or from others and hence will not enter the energy trading market in these time slots. In addition, the sets of buyer and seller EHDs change from time to time. In this paper, we assume that the role of each EHD as a seller or buyer EHD does not change within each time slot but can change from time slot to time slot.

It can be observed that the decision of each EHD k about how to schedule its transmissions and trade energy with others should be closely related to the buffer and energy levels, the specific requirements for data transmission, and the knowledge about the future data arrival and energy harvesting processes. For example, if EHD k is highly sensitive to transmission delay and needs to always successfully send the newly arriving data packets to the corresponding receiver in a fading channel, the number of transmitted data packets $v_{k,t}$ needs to satisfy $v_{k,t}\left[1 - \Pr\left(SNR_{k,t} < \underline{\gamma}_k\right)\right]^{v_{k,t}} = \hat{u}_{k,t}$ where $\Pr\left(\mathrm{SNR}_t < \underline{\gamma}\right)$ is the packet error probability [30]. If EHD k can tolerate a certain delay, on the other hand, it can reduce its energy consumption by properly scheduling the transmission of its data packets. For example, an EHD can deliberately delay the transmission of some data packets if it expects the amount of energy to be harvested in the future to be much more than that received at present. When an EHD requires more energy than it can harvest and cannot tolerate long delays for packet transmission, it will have to request energy to be transferred from others. In this paper, we consider a general energy trading framework in which different EHDs can have different requirements and we use $\mathcal{R}_k$ to denote the requirement of EHD k. Let $w_{k,t}$ be the amount of energy required to send $v_{k,t}$ data packets of EHD k. We assume that there exists a one-to-one mapping function $Y(\cdot)$ from $v_{k,t}$ to $w_{k,t}$, i.e., $w_{k,t} = Y\left(v_{k,t}\right)$ and $v_{k,t} = Y^{-1}\left(w_{k,t}\right)$. For example, if the unit of energy for sending a data packet is given by $\underline{e}_k$, we can write $w_{k,t} = \underline{e}_k v_{k,t}$.

Note that the energy trading among EHDs may incur cost due to the energy transfer loss and the extra resources spent on coordination and information exchange. We hence can introduce a pricing function $c^1_{kj,t}\left(g_{jk,t}, q^k_{j,t}, \Delta\tilde{e}_{k,t}\right)$ to denote the cost to EHD k when it requests EHD j to send $q^k_{j,t}$ amount of energy. It can also be observed that even though EHDs can trade energy with each other, the received energy may not always be enough to support the required services. For example, the energy required by the buyer EHDs can exceed that can be provided by the seller EHDs (i.e., $q^k_{j,t} > \bar{q}_{j,t}$), or the cost of energy trading may be too high for some buyer EHDs. We can write $c^2_{k,t}\left(v_{k,t}, w_{k,t}, \mathcal{R}_k\right)$ as the cost that EHD k incurred from unsatisfactory transmission of data packets in time slot t. Similarly, we can also define the reward that EHD k obtained by trading energy with EHD j in time slot t as $\pi^1_{kj,t}\left(g_{jk,t}, q^k_{j,t}, \Delta\tilde{e}_{k,t}\right)$, i.e., if EHD k is a seller EHD, $\pi^1_{kj,t}\left(g_{jk,t}, q^k_{j,t}, \Delta\tilde{e}_{k,t}\right)$ is the reward obtained by selling $\frac{\Delta\tilde{e}_{k,t}}{g_{jk,t}}$ amount of energy to buyer EHD j. We also denote the reward from successfully sending $v_{k,t}$ data packets with $w_{k,t}$ units of energy by EHD k during time slot t as $\pi^2_{k,t}\left(v_{k,t}, w_{k,t}, \mathcal{R}_k\right)$.

Since the amount of energy harvested from the ambient environment is generally limited, it is reasonable to assume that in each time slot each buyer EHD can only purchase energy from one seller EHD. Similarly, each seller EHD can only transfer its energy to one buyer EHD, e.g., using one-to-one magnetic resonant coils for energy transfer. We refer to each pair of buyer and seller EHDs formed in each time slot as a *buyer-seller pair*. At the beginning of each time slot, each EHD needs to decide which EHD to trade energy with. Let $\mu_t\left(k\right)$ be the seller (or buyer) EHD chosen by a buyer (or seller) EHD k in time slot t to trade energy, i.e., for each seller EHD $k \in \mathcal{S}_t$ (or buyer EHD $k \in \mathcal{B}_t$), we have $\mu_t\left(k\right) \in \mathcal{B}_t \cup \{k\}$ (or $\mu_t\left(k\right) \in \mathcal{S}_t \cup \{k\}$) where we use $\mu_t\left(k\right) = k$ to mean that no EHDs will trade energy with EHD k during time slot t, i.e., we have $\pi^1_{kk,t} = c^1_{kk,t} = 0$. Our formulation and results can be directly extended to more general cases with multiple seller and buyer EHDs forming a coalition. We will provide a more detailed discussion on such extensions in Section VIII.

Note that the pairing process between buyer and seller EHDs includes complex interactions among EHDs. For example, in order to obtain sufficient energy from the seller EHDs, all the buyer EHDs will compete for the seller EHDs that can provide the highest tradable energy at the lowest cost. In addition, each seller EHD can receive energy requests from multiple buyer EHDs. It will then need to carefully decide which buyer EHD to transfer its energy to. We consider a general model and the payoff of each EHD in each time slot can be any performance metric or function related to the rewards and costs of energy trading between itself and its energy trading partner. More specifically, let $\varpi_{kj,t}\left(\pi^1_{kj,t}, \pi^2_{k,t}, c^1_{kj,t}, c^2_{k,t}\right)$ be the payoff obtained by EHD k when it trades energy with EHD j during time slot t for $k \in \mathcal{U}$. For example, if the payoff of EHD k is a linear function of $\pi^1_{kj,t}$, $\pi^2_{k,t}$, $c^1_{kj,t}$ and $c^2_{k,t}$, we can write $\varpi_{kj,t}\left(\pi^1_{kj,t}, \pi^2_{k,t}, c^1_{kj,t}, c^2_{k,t}\right) = \pi^1_{kj,t} + \pi^2_{k,t} - c^1_{kj,t} - c^2_{k,t}$. We will consider a specific payoff function and use it as an example for our proposed dynamic energy trading market in Section VII. In this paper, we assume that each EHD can always obtain different payoffs when trading with different EHDs, i.e., $\varpi_{kj,t} \neq \varpi_{ki,t}$ $\forall i \neq j$ and $i, j \in \mathcal{U}\backslash\{k\}$. To simplify our discussion, once each EHD k has already chosen its energy trading partner $\mu_t(k)$, we abuse the notation and rewrite the payoff of EHD k in time slot t as $\varpi_{k,t}\left(\pi^1_{k\mu_t(k),t}, \pi^2_{k,t}, c^1_{k\mu_t(k),t}, c^2_{k,t}\right) = \varpi_{k\mu_t(k),t}\left(\pi^1_{k\mu_t(k),t}, \pi^2_{k,t}, c^1_{k\mu_t(k),t}, c^2_{k,t}\right)$.

In the dynamic energy trading market, the main objective for each EHD k is to sequentially optimize its decisions about the transmit power $w_{k,t}$, number of transmit data packets $v_{k,t}$, which EHD and how much energy to trade with to maximize its long-term discounted payoff given by,

$$\lim_{T\to\infty} E\left(\sum_{t=1}^{T} \gamma^t \varpi_{k,t}\left(\pi^1_{k\mu_t(k),t}, \pi^2_{k,t}, c^1_{k\mu_t(k),t}, c^2_{k,t}\right)\right). \qquad (3)$$

We seek a simple and distributed mechanism that can incentivize energy trading among EHDs. The mechanism should also ensure that the energy trading market converges to a stationary structure under each possible state, in which no EHD can further improve its payoff by unilaterally deviating from its resulting pairing partner, transmit power, number of transmit data packets, and the amount of energy to trade.

V. A Simple Energy Trading Scheduling Protocol

As mentioned previously, in dynamic energy trading, the role of each EHD as seller or buyer EHD can change over time. Therefore, it is important for each buyer (or seller) EHD to first discover the available seller (or buyer) EHDs in its surrounding area. Additionally, a link establishment protocol should also be introduced for buyer and seller EHDs to coordinate and transfer energy. A buyer-seller pair can only be established when a buyer EHD and a seller EHD coordinate their energy transfer parameters for energy transmission and receiving. For example, if both buyer and seller EHDs are equipped with magnetic resonant coils, energy transfer from a seller EHD to a buyer EHD can only be successful if both EHDs adjust their resonant matching circuits to operate on the same resonant frequency [31]. We consider energy-efficient communication systems, where the seller and buyer EHDs cannot spend energy on multiple rounds of back-and-forth negotiation about the details of the energy transfer at the beginning of each time slot. More specifically, we follow the same approach as synchronous peer-to-peer ad hoc network systems [32], [33] and introduce the following simple energy trading scheduling protocol for the EHDs to establish energy transfer links at the beginning of each time slot:

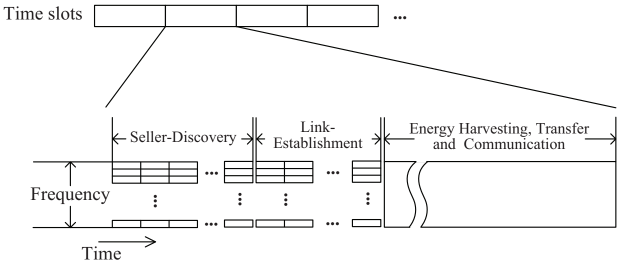


Fig. 3: Time structure of a simple energy trading scheduling protocol using OFDM resource blocks: At the beginning of each time slot, each seller EHD chooses a seller-discovery resource block to broadcast its identity and energy information. Once the information sent by the seller EHDs is received, each buyer EHD chooses a link-establishment resource block to send its energy request. After the buyer EHD and seller EHD have been matched, the seller EHD will transfer its available energy to the corresponding buyer EHD.

Protocol 1: Description of A Simple Energy Trading Scheduling Protocol

1) *Seller discovery*: At the beginning of each time slot t, all EHDs have a short dedicated time segment to discover the seller EHDs. The seller discovery resources consist of a set of orthogonal resource blocks. Each seller EHD monitors the seller discovery resource blocks and chooses a locally unused one to broadcast its identity and the amount of energy available to be transferred to the buyer EHDs.
2) *Energy trading link establishment*: Each buyer EHD will decode the signals broadcasted by the seller EHDs in the seller discovery time segment and then select its preferred seller EHD. Each buyer EHD will then choose an unused resource block in the following link establishment time segment to send its energy request signal together with its energy receiving parameters. If a seller EHD accepts the request of the buyer EHD, it will adjust its energy transfer frequency to the same one as the buyer EHD. Otherwise, the seller EHD will reject a buyer EHD by ignoring the request sent by the buyer EHD. The rejected buyer EHDs will not receive any energy from the seller EHDs for the remainder of time slot t.

We illustrate the time structure of the above protocol in Figure 3.

Note that it is possible that multiple buyer EHDs send energy requests to the same seller EHD $j \in \mathcal{S}_t$, causing a conflict. To resolve this conflict, seller EHD j will need to choose one buyer EHD and ignore the energy requests sent by other buyer EHDs. Each seller EHD can establish a preference over the requesting buyer EHDs by ranking the resulting benefits of energy trading from the highest to the lowest and use the established preference to decide the buyer EHD to send energy to. For example, in some energy efficient systems, the benefit that each seller EHD obtains by selling its energy to the requesting buyer EHDs is proportional to the energy transfer efficiency between itself and the requesting buyer EHDs. If we write $\mathcal{F}_{j,t}$ as a set of buyer EHDs that send energy requests to seller EHD j, $\mathcal{F}_{j,t} \subseteq \mathcal{B}_t$, seller EHD j prefers to send its energy to buyer EHD $k = \arg\max_{l \in \mathcal{F}_{j,t}} \{g_{lj,t}\}$. Suppose EHDs are deployed with magnetic resonant coils for energy transfer. Each seller EHD can then estimate the relative order of energy transfer efficiency from itself to each requesting buyer EHD by measuring the load that each buyer EHD imposes on the transmitter circuit, as discussed in [19]. More specifically, each seller EHD knows the resonant frequency of its requesting buyer EHDs. Thus, all seller EHDs can sequentially adjust their resonant frequency to the same one as each requesting buyer EHD by applying a known voltage on its coil and measuring the current flowing through its coil to estimate the energy transfer efficiency between itself and the requesting buyer EHDs. Another way to achieve this is to let each seller EHD j establish a relative order of energy transfer efficiency between itself and the energy requesting buyer EHDs by evaluating the channel gains from its received energy requesting signal. Specifically, in systems where both the channel gains and energy transfer efficiencies are dominated by the distances between (data and energy) transmitters and receivers, each seller EHD can establish the preference list about the energy transfer efficiencies between itself and requesting buyer EHDs by sorting the channel gains from highest to lowest. We assume that each buyer EHD does not know the seller EHDs' preference nor their conflict-resolving rules. It can be observed that with the increase of the distance, the wireless energy transfer efficiency drops much faster than the communication channel gain. In other words, each buyer EHD will only compete with its nearby buyer EHDs for the seller EHDs in the surrounding area. Thus, it is possible for each buyer EHD to eavesdrop on the seller EHDs requested by other buyer EHDs in the past. In this paper, we assume that each buyer EHD does not know the current decisions made by other buyer EHDs about the seller EHDs to send energy requests to but it can eavesdrop on the seller EHDs requested by buyer EHDs in previous time slots. Each EHD can exploit its observation history to establish its belief, which will be discussed in detail in Section VI.

VI. A Stochastic Energy Trading Mechanism and Distributed Algorithms

We formulate the dynamic energy trading problem as a *stochastic energy trading game*. It can be regarded as an extension of a stochastic game by modeling the interaction between buyer and seller EHDs under each state as a stable matching game with private belief [28], [29]. Before we present the details of our game formulation, let us briefly describe some of the basic elements of a stochastic game [34]–[36]. A stochastic game consists of a set Υ of states and a set $\mathcal{U}$ of players. The game is played in a sequence of time slots. Each player decides its action at the beginning of each time slot. The state is time-varying and can be regarded as fixed within each time slot. Each player does not have complete information about the state but can obtain a partial observation $o_{k,t} \in \Omega_k$ at the beginning of each time slot t. The evolution of the state is characterized by a state transition function $\Gamma\left(\eta_t, \eta_{t-1}, \boldsymbol{a}_{t-1}\right) = \Pr\left(\eta_t | \eta_{t-1}, \boldsymbol{a}_{t-1}\right)$ which specifies the probability distribution of an outcome η_t given that, starting at state η_{t-1}, a joint action $\boldsymbol{a}_{t-1}$ is taken by the players. Each player receives a payoff at the end of each time slot.

In the stochastic energy trading game, the players are the EHDs. The state, denoted as $\eta_t \in \Upsilon$ in time slot t, is a composite variable of the harvested energy $\boldsymbol{e}_t$ and energy required to send data packets for all the EHDs $\boldsymbol{w}_t$. In each state η_t, the EHDs are divided into two disjoint subsets $\mathcal{S}_t$ and $\mathcal{B}_t$, corresponding to sets of seller and buyer EHDs, respectively. An action $a_{k,t}$ of each buyer EHD $k \in \mathcal{B}_t$ is to choose an appropriate seller EHD to request energy from. We follow the same line as most existing works in stochastic games and assume that the state transition function $\Gamma\left(\eta_t, \eta_{t-1}, \boldsymbol{a}_{t-1}\right)$ is known by all the players and each player knows the probability distribution $\Theta\left(o_{k,t}, \eta_t, \boldsymbol{a}_{t-1}\right) = \Pr\left(o_{k,t} | \eta_t, \boldsymbol{a}_{t-1}\right)$ for each possible observation under each action and resulting state. We will describe how to relax these assumptions in Section VIII.

For the rest of this section, we first focus on one time slot of a game play and then consider the sequential optimization for EHDs in a stochastic domain.

A. A Stable Matching Game with Private Belief

In this subsection, we focus on the energy trading within one time slot t. We assume that the sets of buyer and seller EHDs, and state are fixed in the time slot. We formulate the interaction between buyer and seller EHDs in each specific state as a matching game with private belief, which is formally defined as follows:

Definition 1. *A (two-sided one-to-one) matching game with (one-sided) private belief is a tuple* $\mathbb{M} = \langle \mathcal{B}_t, \mathcal{S}_t, \hat{b}_{k,t}, \succ \rangle$ *consisting of two finite and disjoint subsets of players* $\mathcal{B}_t$ *and* $\mathcal{S}_t$*, preference* $\succ$*, and a belief function* $\hat{b}_{k,t}$ *for each player in one subset.*

In dynamic energy trading, the two subsets of players correspond to sets of buyer and seller EHDs. We provide a more detailed discussion of each of the elements in our formulated game as follows: each buyer EHD $k \in \mathcal{B}_t$ also has a specific type, denoted as $y_{k,t}$, which includes all the private information related to its decision making [37]. More specifically, for each buyer EHD, its type captures its preference over the seller EHDs and its ability to compete with others for the seller EHDs. For each seller EHD, its type specifies its preference about buyer EHDs and its conflict-resolving rules. Since the type of each EHD is private information, it is not known to others. However, each EHD k can establish and maintain a belief $\hat{b}_{k,t}\left(\boldsymbol{y}_{-k,t}\right)$ about types of others for $\boldsymbol{y}_{-k,t} = \langle y_{l,t} \rangle_{l \in \mathcal{U} \backslash \{k\}}$. Each buyer EHD establishes its preference over the seller EHDs and decides a specific seller EHD to send its energy request to. We use $i \succ_k j$ to mean that buyer EHD k prefers to send an energy request to seller EHD i rather than to seller EHD j for $i \neq j$, $i, j \in \mathcal{S}_t$ and $k \in \mathcal{B}_t$. Apart from establishing the preference over all the seller EHDs, each buyer EHD should also decide a specific seller EHD to send a request to at the beginning of each time slot. Note that each EHD will schedule the transmission of its data packets according to how much energy and how many data packets it receives, which will depend on the energy harvesting process, number of arriving data packets, as well as the energy it can trade with other EHDs. The energy harvesting and packet arrival processes cannot be controlled by each EHD. However, each EHD can improve its payoff by sequentially choosing the proper energy trading partner. We therefore refer to the seller EHD chosen by each buyer EHD k in each time slot t to send an energy request to as the *action* $a_{k,t}$ of EHD k for $a_{k,t} \in \mathcal{A}_{k,t}$ and $\mathcal{A}_{k,t} = \mathcal{S}_t$. We also write $\boldsymbol{\mathcal{A}}_t = \langle \mathcal{A}_{k,t} \rangle_{k \in \mathcal{B}_t}$.

From the aforementioned analysis, we observe that the pairing structure between buyer and seller EHDs is fully determined by the joint action of the buyer EHDs and the conflict-resolution rules of the seller EHDs in each time slot. We refer to a pairing structure between seller and buyer EHDs as a *matching*, which is formally defined as follows:

Definition 2. For a fixed state η_t, a (two-sided one-to-one) matching μ_t is a function from the set $\mathcal{B}_t \cup \mathcal{S}_t$ onto itself such that for each buyer EHD $k \in \mathcal{B}_t$, if $\mu_t(k) \neq k$, then $\mu_t(k) \in \mathcal{S}_t$ and for each seller EHD $j \in \mathcal{S}_t$ if $\mu_t(j) \neq j$, then $\mu_t(j) \in \mathcal{B}_t$ and $\mu_t(k) = j \Leftrightarrow \mu_t(j) = k$ for all $k \in \mathcal{B}_t$ and $j \in \mathcal{S}_t$.

Note that in our energy trading scheduling protocol presented in Section V, each buyer EHD can only send an energy request to one seller EHD in each time slot. If all buyer EHDs keep sending requests to the seller EHDs that provide the highest payoffs in every time slot, some buyer EHDs requesting the same most preferred seller EHDs will always be rejected and receive no energy. In other words, the traditional Gale-Shapley algorithm-based approaches, that is, each buyer sequentially sends energy requests from the most preferred seller to the least one until its request is accepted, cannot be directly applied to solve the energy trading problem in a stochastic environment.

From the previous analysis, the action of each buyer EHD should depend on: 1) how much energy each seller EHD can provide and 2) whether its chosen seller EHD will accept its energy request. To solve this problem, we define a belief function and introduce the concept of belief-based preference [27], [28]. Each buyer EHD $k \in \mathcal{B}_t$ first establishes and maintains a belief function from its observation history and then determines its preference about seller EHDs based on its belief function. Note that the main difference between the matching game with private belief and the traditional matching game without beliefs is as follows. In the former game, each EHD's preference also depends on its belief function. For example, if buyer EHD k knows that another buyer EHD l has the same most preferred seller EHD j to send an energy request to and, according to its belief function, buyer EHD l is more preferred by seller EHD j than itself, then buyer EHD k will remove seller EHD j from its list of the most preferred seller EHDs because it knows that seller EHD j is likely to reject its energy request.

Let us now describe how to convert each buyer EHD's uncertainty about the types of other EHDs to the uncertainty about the actions of other buyer EHDs and the final matched seller EHD. Since each buyer EHD decides its action based on its type, there is a mapping function from each buyer EHD's type to its action. Each buyer EHD cannot directly observe other buyer EHDs' types but can estimate their types by eavesdropping on the past actions of other EHDs. In particular, each buyer EHD can establish a belief function about the actions of other buyer EHDs denoted as follows:

$$b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right) = \Pr\left(\boldsymbol{a}_{-k,t} | o_{k,t}, \boldsymbol{H}_{k,t}\right), \tag{4}$$

where we use subscript $-k$ to mean all EHDs except EHD k and $\boldsymbol{H}_{k,t}$ is the observation history of EHD k. This belief function will also specify the belief of buyer EHD k about the types of all the other buyer EHDs. For each joint action of all buyer EHDs under a given state, the final matched seller EHD of each buyer EHD is fully determined by the conflict-resolution rules of the seller EHDs. In other words, the uncertainty of each buyer EHD about types of the seller EHDs is equivalent to its uncertainty about the final matching results. Each buyer EHD can therefore exploit its previous observations to establish a belief function about the final matched seller EHD under each possible joint action of buyer EHDs and observation. The belief is denoted as follows:

$$b''_{k,t}\left(\mu_t(k)\right) = \Pr\left(\mu_t(k) | o_{k,t}, \boldsymbol{a}_t, \boldsymbol{H}_{k,t}\right). \tag{5}$$

If each buyer EHD can establish belief functions about the actions of other buyer EHDs and the final matched seller EHD at the beginning of each time slot, it can then evaluate the expected payoff obtained by sending an energy request to each seller EHD. The buyer EHD eventually uses the evaluation results to establish its belief-based preference over seller EHDs. In particular, buyer EHD k can establish the preference over seller EHDs i and j by $i \succ_k j$ if $\bar{\varpi}_{ki,t} > \bar{\varpi}_{kj,t}$. $\bar{\varpi}_{ki,t}$ is the expected payoff obtained by buyer EHD k when it sends an energy request to seller EHD i, given by

$$\begin{aligned}\bar{\varpi}_{ki,t} &= \sum_{\boldsymbol{a}_{-k,t} \in \boldsymbol{\mathcal{A}}_{-k,t}} \sum_{\mu_t(k) \in \{i,k\}} \Pr\left(\boldsymbol{a}_{-k,t},\right. \\ &\qquad\qquad \left.\mu_t(k) = i | o_{k,t}, \boldsymbol{H}_{k,t}\right) \varpi_{ki,t} \\ &= \sum_{\boldsymbol{a}_{-k,t} \in \boldsymbol{\mathcal{A}}_{-k,t}} \sum_{\mu_t(k) \in \{i,k\}} b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right) \\ &\qquad\qquad b''_{k,t}\left(\mu_t(k) = i\right) \varpi_{ki,t},\end{aligned} \tag{6}$$

and $\varpi_{ki,t}$ is given in (19). For the rest of this subsection, we focus on how to establish and update the belief function of each buyer EHD about the actions of other buyer EHDs and the final matched seller EHD at the beginning of each time slot.

We will consider how to estimate the current and future changes of the system state in the next subsection. Basically, we adopt a commonly used model-based Bayesian reinforcement learning framework to model the belief of each buyer EHD about actions of other buyer EHDs. In this framework, the prior distributions of the beliefs of each buyer EHD about actions of other buyer EHDs as well as its final matched seller EHD are characterized by the Dirichlet distribution and Beta distribution, respectively [38], [39]. Both Dirichlet distribution and Beta distribution have been widely used as the prior distribution in statistical inference methods because they have the following properties:

1) The Dirichlet distribution and Beta distribution can be regarded as the probability distributions over the parameters of the multinomial distribution and binomial distribution, respectively. In the energy trading system, the belief of each buyer EHD k about actions of other buyer EHDs can be regarded as a probability distribution over a limited number of possible choices for seller EHDs and hence can be modeled as a multinomial distribution. In particular, if we use $(\boldsymbol{a}_{-k,t}, \eta_t)$ to denote the event that all buyer EHDs except buyer EHD k choose actions $\boldsymbol{a}_{-k,t}$, we can then write the probability distribution of the actions of all buyer EHDs except buyer EHD k, according to the belief of buyer EHD k in a given state η_t, as $(\boldsymbol{a}_{-k,t}, \eta_t) \,|b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right) \sim \mathrm{Mul}(\boldsymbol{\phi}_{-k,t}, m)$, where $\mathrm{Mul}(\boldsymbol{\phi}_{-k,t}, m)$ denotes the multinomial distribution with parameters $\boldsymbol{\phi} = \left[\phi_1, \phi_2, \ldots, \phi_{|\boldsymbol{\mathcal{A}}_{-k,t}|-1}\right]$ and m, and ϕ_i is the probability that the ith outcome occurs. Similarly, once each buyer EHD decides on the seller EHD to send an energy request to, the seller EHD can only have two possible responses, that is, either accept or reject the request. This can then be modeled as a binomial distribution. In particular, we can write the probability distribution of matching result $\mu_k(k)$ according to the belief of buyer EHD k as $\mu_t\left(k\right)|b''_{k,t}\left(\mu_t\left(k\right)\right) \sim \mathrm{Bi}(\psi_{k,t}, m)$ where $\mathrm{Bi}(\psi_{k,t}, m)$ denotes the binomial distribution with parameters $\psi_{k,t}$ and m, and $\psi_{k,t}$ is the probability that seller EHD $a_{k,t}$ accepts the energy request of buyer EHD k. We hence can apply the Dirichlet distribution and Beta distribution to model the probability distributions of $\boldsymbol{\phi}_{-k,t}$ and $\psi_{k,t}$ of the multinomial distribution and Binomial distribution of events $b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right)$ and $b''_{k,t}\left(\mu_t(k)\right)$, respectively, i.e., $b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right) \sim \mathrm{Dir}\left(\boldsymbol{\alpha}_{-k,t}\right)$ and $b''_{k,t}\left(\mu_t\left(k\right)\right) \sim \mathrm{Beta}\left(\beta_{k,t}\right)$ where $\boldsymbol{\alpha}_t = \left\{\alpha_{(\boldsymbol{a}_{-k,t}, o_{k,t})}\right\}_{\boldsymbol{a}_{-k,t} \in \times_{|\boldsymbol{\mathcal{A}}_{-k,t}|-1}}$ and $\alpha_{(\boldsymbol{a}_{-k,t}, o_{k,t})}$ are the concentration parameters satisfying

$$\Pr\left(\boldsymbol{\phi}_{-k,t}|\boldsymbol{\alpha}_{-k,t}\right) = \frac{\Gamma\left(\sum_{i=1}^{|\boldsymbol{\mathcal{A}}_{-k,t}|-1}\alpha_i\right)}{\prod_{i=1}^{|\boldsymbol{\mathcal{A}}_{-k,t}|-1}\Gamma\left(\alpha_i\right)} \prod_{i=1}^{|\boldsymbol{\mathcal{A}}_{-k,t}|-1} \phi_i^{\alpha_i - 1}, \qquad (7)$$

$$\text{and } \Pr\left(\psi_{k,t}|\beta_{k,t}\right) = \frac{\Gamma\left(\beta_{k,t}\right)}{\Gamma\left(\beta_{k,t}\right)} \psi^{\beta_{k,t}-1}, \qquad (8)$$

where $\Gamma(\cdot)$ is the gamma function.

2) The Dirichlet distribution and Beta distribution are the conjugate priors for the multinomial distribution and binomial distribution, respectively [40]. Suppose the prior distributions of $\boldsymbol{a}_{-k,t}$ and μ_t follow the Dirichlet distribution and Beta distribution with concentration parameters $\boldsymbol{\alpha}_{-k,t}$ and $\beta_{k,t}$, respectively. If the observations during the first t time slots of the actions follow distributions $\mathrm{Mul}(\boldsymbol{\phi}_{-k,t}, m)$ and $\mathrm{Bi}(\phi_{k,t}, m)$, respectively, then the posterior distributions over $\boldsymbol{\phi}_{-k,t}$ and $\psi_{k,t}$ will follow distributions $\mathrm{Dir}\left(\boldsymbol{\alpha}_{-k,t} + \boldsymbol{n}_{k,t}\right)$ and $\mathrm{Beta}\left(\beta_{k,t} + \boldsymbol{n}'_{k,t}\right)$, respectively. Then, we have

$$\boldsymbol{n}_{k,t} = \left\{n_{k,t}(\boldsymbol{a}_{-k,t}, o_{k,t})\right\}_{\boldsymbol{a}_{-k,t} \in \boldsymbol{\mathcal{A}}_{-k,t}}, \qquad (9)$$

where $n_{k,t}(\boldsymbol{a}_{-k,t}, o_{k,t})$ is the number of times that the joint action $\boldsymbol{a}_{-k,t}$ has been taken given observation $o_{k,t}$ during the past t time slots and $n'_{k,t}\left(\mu_t(k), o_{k,t}, \boldsymbol{a}_t\right)$ is the number of times that the resulting matched seller EHD is given by $\mu_t(k)$ when buyer EHD k observes $o_{k,t}$ and joint action $\boldsymbol{a}_t$ during the past t time slots.

Each buyer EHD k can establish its belief about actions of other buyer EHDs and the final matching result as follows:

$$\begin{aligned} B_{k,t}\left(\boldsymbol{a}_{-k,t}, \mu_t(k)\right) &= \Pr\left(\boldsymbol{a}_{-k,t}, \mu_t(k)|o_{k,t}, \boldsymbol{H}_{k,t}\right) \\ &= \Pr\left(\boldsymbol{a}_{-k,t}|o_{k,t}, \boldsymbol{H}_{k,t}\right)\Pr\left(\mu_t(k)|o_{k,t}, a_{k,t}, \boldsymbol{a}_{-k,t}, \boldsymbol{H}_{k,t}\right) \\ &= \int_0^1 \Pr\left(\boldsymbol{a}_{-k,t}|o_{k,t}, b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right)\right) \\ &\qquad \Pr\left(b'_{k,t}\left(\boldsymbol{a}_{-k,t}\right)|o_{k,t}, \boldsymbol{a}_{-k,t}, \boldsymbol{\alpha}_{-k,t}\right) db'_{k,t}\left(\boldsymbol{a}_{-k,t}\right) \\ &\quad \cdot \int_0^1 \Pr\left(\mu_t|o_{k,t}, \boldsymbol{a}_t, b''_{k,t}\left(\mu_t\left(k\right)\right), \boldsymbol{H}_{k,t}\right) \\ &\qquad \Pr\left(b''_{k,t}\left(\mu_t(k)\right)|o_{k,t}, \boldsymbol{a}_t, \boldsymbol{H}_{k,t}\right) db''_{k,t}\left(\mu_t(k)\right) \\ &= \frac{n_{k,t}\left(\boldsymbol{a}_{-k,t-1}, o_{k,t}\right)}{\sum_{\boldsymbol{a}_{-k,t-1} \in \boldsymbol{\mathcal{A}}_{-k,t-1}} n_{k,t}\left(\boldsymbol{a}_{-k,t-1}, o_{k,t}\right)} \\ &\quad \cdot \frac{n'_{k,t}\left(\mu_t\left(k\right), o_{k,t}, \boldsymbol{a}_t\right)}{\sum_{\mu_t(k) \in \{a_{k,t}, k\}} n'_{k,t}\left(\mu_t\left(k\right), o_{k,t}, \boldsymbol{a}_t\right)}, \end{aligned} \qquad (10)$$

where $n_{k,t}\left(\boldsymbol{a}_{-k,t-1}, o_{k,t}\right) = \sum_{i=1}^{t-1}\sum_{o_{k,i} \in \Omega} \mathbf{1}\left(\boldsymbol{a}_{-k,i} = \boldsymbol{a}_{-k,t-1}|o_{k,i} = o_{k,t}\right)$ and $n'_{k,t}\left(\mu_t\left(k\right), o_{k,t}, \boldsymbol{a}_t\right) = \sum_{i=1}^{t-1}\sum_{o_{k,i} \in \Omega} \mathbf{1}\left(\mu_t\left(k\right)|o_{k,t}, \boldsymbol{a}_t\right)$. $\mathbf{1}(\cdot)$ is an indicator function.

We can observe that, with the state changing from time slot to time slot, the sets of seller and buyer EHDs as well as the set of possible actions of each buyer EHD will also change. In other words, each EHD will decide a sequence of actions during the progression of the states which will result in a sequence of matchings. We follow the same line as the stochastic coalition formation game in [34], [41], [42] and introduce the concept of a (weak) stable matching for dynamic energy trading market as follows.

Definition 3. *A matching μ_t is said to be* (weakly) stable *if every EHD believes that matching μ_t in time slot t cannot be strictly improved upon by any EHD or buyer-seller pair. A matching is (weakly) optimal if each EHD believes that it cannot further improve its long-term average payoff by choosing another matching.*

We can further define the concept of strong stable matching as a matching satisfying the following conditions: 1) no EHD or buyer-seller pair believes that it can further improve its payoff by unilaterally deviating from an existing matching, and 2) every EHD believes that no other EHD or buyer-seller pair can further improve its expected payoff by unilaterally deviating from the existing matching. It can be observed that the strong stable matching relies on each EHD's subjective belief about others and hence is more "endogenous" compared to the weak stable matching [42]. We only consider the weak stable matching, and thus with a slightly abuse of the definition, we use stable matching to mean weak stable matching in our dynamic energy trading market.

We observe from the above definition that the resulting stable matching in each time slot is closely related to the beliefs of each EHD. In this paper, we seek a sequence of matchings between buyer and seller EHDs that is stable and optimal. The main objective for each EHD is to maximize its long-term average payoff given by $E\left(\sum_{t=0}^{\infty}\gamma^t \varpi_{k,t}\right)$, where $0<\gamma<1$ is the discount factor.

B. A Stochastic Energy Trading Game

Let us now focus on the sequential decision making process of each EHD in a stochastic environment [34]. We formally define a stochastic energy trading game as follows:

Definition 4. A stochastic energy trading game is characterized by a set of states Υ, a set $\mathcal{U}$ of EHDs which can be divided into sets $\mathcal{S}_t$ and $\mathcal{B}_t$ for seller and buyer EHDs, respectively, under each given state $\eta_t \in \Upsilon$, a preference $\succ$ for each EHD, a set of possible observations $\Omega_{k,t}(o_{k,t}, \boldsymbol{a}_{t-1}, \eta_t)$ and an observation function $\Theta_{k,t}$ for each buyer EHD $k \in \mathcal{B}_t$, a probability distribution $\Gamma(\eta_t, \eta_{t-1}, \boldsymbol{a}_{t-1})$ for the transition dynamics, a payoff function $\varpi_{k,t}$ and a belief function $\tilde{B}_{k,t}$ for each buyer EHD k.

We provide a more detailed discussion for each of the elements in our stochastic energy trading game as follows.

1) States and Observations: In energy harvesting communication networks, state η_t in each time slot t corresponds to the battery levels, data buffer levels of all EHDs and energy transfer efficiencies among EHDs. Once the state is fixed, the sets of seller and buyer EHDs will also be determined. Each EHD can obtain an observation at the beginning of each time slot which includes the energy information broadcasted by seller EHDs. However, each buyer EHD cannot know the amount of harvested energy or the amount of energy required for sending the data packets of other buyer EHDs. Each buyer EHD also cannot know the conflict-resolving rules of seller EHDs. Each buyer EHD $k \in \mathcal{B}_t$ will have to establish its belief about the current and future states of the system using its observation function $\Theta_{k,t}(\eta_t, \boldsymbol{a}_{t-1}, o_{k,t}) = \Pr(o_{k,t}|\eta_t, \boldsymbol{a}_{t-1})$. The observation function specifies, for each joint action of all buyer EHDs and state, the probability distribution of possible observations.

2) State Transitions and Belief Function: In each time slot, buyer EHDs cannot know the current state but can estimate the probability distribution of the possible states, i.e., buyer EHD k can establish a belief function $b'''_{k,t}(\eta_t) = \Pr\left(\eta_t|o_{k,t}, \boldsymbol{a}_{t-1}, b'''_{k,t-1}(\eta_{t-1})\right)$. By combining $b'''_{k,t}(\eta_t)$ with the beliefs about the actions of other buyer EHDs as well as the matching result described in the previous subsection, each buyer EHD k can obtain the following belief at the beginning of each time slot:

$$\begin{aligned}
&\tilde{B}_{k,t}(\eta_t, \boldsymbol{a}_{-k,t}, \mu_t(k)) \\
&\quad = \Theta(\eta_t, \boldsymbol{a}_{t-1}, o_{k,t}) \sum_{\eta_t\in\Upsilon} \Gamma(\eta_t|\boldsymbol{a}_{t-1}, \eta_{t-1}) \\
&\qquad {b'''}_{k,t-1}(\eta_{t-1})\, B_{k,t}(\boldsymbol{a}_{-k,t}, \mu_t(k)), \qquad (11)
\end{aligned}$$

where $B_{k,t}(\boldsymbol{a}_{-k,t}, \mu_t(k))$ is given in (10).

The above belief function can be used to update the belief of each buyer EHD about the current state, actions of other buyer EHDs and the final matching result at the beginning of each time slot. We now need to prove that belief $\tilde{B}_{k,t}(\eta_t, \boldsymbol{a}_{-k,t})$ in (11) is a *sufficient statistic*, which means that buyer EHD $k \in \mathcal{B}_t$ can make the decision about its future actions without requiring any further information about past observations.

Proposition 1. In our proposed energy trading system, the belief $b_{k,t}(\eta_t, \boldsymbol{a}_{-k,t}, \mu_t(k))$ of each buyer EHD k calculated and updated using (11) is a sufficient statistic for the past history of buyer EHD k's observations.

Proof: See Appendix A. ■

Each buyer EHD k can then use the belief function given in (11) to estimate the expected payoff that can be obtained from each possible action $a_{k,t} \in \mathcal{A}_{k,t}$ in time slot t. We have

$$\begin{aligned}
&\bar{\bar{\varpi}}_{k,t}(a_{k,t}=j) = \\
&\sum_{\mu_t(k)\in\{k,a_{k,t}\}} \sum_{\boldsymbol{a}_{-k,t}\in\boldsymbol{\mathcal{A}}_{-k,t}} \sum_{\eta_t\in\Upsilon} \tilde{B}_{k,t}(\boldsymbol{a}_{-k,t}, \eta_t, \\
&\qquad \mu_t(k))\, \varpi_{kj,t}, \qquad (12)
\end{aligned}$$

where $j \in \mathcal{S}_t$ and $\varpi_{kj,t}$ is given in (19).

3) Optimal Policy and Distributed Algorithm: To choose the optimal action at the beginning of each time slot to maximize its long-term average payoff, each buyer EHD needs to evaluate the long-term average payoff obtained by each of its possible actions. We define a value function $V_{k,t}\left(\boldsymbol{a}_t, o_{k,t}, \eta_t, \tilde{B}_{k,t}\right)$ as the sum of the current and future expected payoffs when the current state and joint actions of all buyer EHDs are given by η_t and $\boldsymbol{a}_t$, respectively. Buyer EHD k's expected instantaneous payoff is given by $\bar{\bar{\varpi}}_{k,t}(a_{k,t})$ in (12) when it decides to pursue action $a_{k,t}$. Additionally, buyer EHD k should be able to estimate its future expected payoff using its known state-transition function and observation function. More specifically, we can define a belief state estimation function as follows:

$$\begin{aligned}
&\tilde{B}_{k,t}(\boldsymbol{a}_t, o_{k,t}, \mu_t(k)) = \\
&\qquad SE\left(\boldsymbol{a}_{t-1}, \mu_{t-1}(k), o_{k,t-1}, \tilde{B}_{k,t-1}\right). \qquad (13)
\end{aligned}$$

We hence can write $V_{k,t}\left(\boldsymbol{a}_t, o_{k,t}, \eta_t, \tilde{B}_{k,t}\right)$ as follows:

$$\begin{aligned}
&V_{k,t}\left(o_{k,t}, \eta_t, a_{k,t}, \boldsymbol{a}_{-k,t}, \tilde{B}_{k,t}\right) = \bar{\bar{\varpi}}_{k,t}(a_{k,t}) \\
&\quad +\gamma \sum_{o_{k,t+1}\in\Omega} \Pr\left(o_{k,t+1}|a_{k,t}, b_{k,t}(\boldsymbol{a}_{-k,t}, \eta_t)\right) \\
&\qquad V_{k,t}\left(SE(a_{k,t}, \boldsymbol{a}_{-k,t}, o_{k,t}, \mu_t(k))\right). \qquad (14)
\end{aligned}$$

We can write the optimal value function for EHD k as follows:

$$V^*_{k,t}(\boldsymbol{a}_{-k,t}, o_{k,t}, \eta_t, \tilde{B}_{k,t}) = \max_{a_{k,t}\in\mathcal{A}_{k,t}} V_{k,t}\left(\eta_t, a_{k,t}, \boldsymbol{a}_{-k,t}, o_{k,t}, \tilde{B}_{k,t}\right). \quad (15)$$

Therefore, the optimal policy $\pi^*_{k,t}$ for each buyer EHD k in time slot t is given by

$$a^*_{k,t} = \arg\max_{a_{k,t}\in\mathcal{A}_{k,t}} V_{k,t}(\eta_t, a_{k,t}, \boldsymbol{a}_{-k,t}, o_{k,t}, \tilde{B}_{k,t}). (16)$$

Note that (16) is, in some sense, similar to the value iteration algorithm. However, in (16), we have to take into consideration the interaction among EHDs by introducing an interactive belief function $\tilde{B}_{k,t}$. Additionally, to ensure the convergence of the interaction among EHDs, we have also introduced a stable matching with private belief-based mechanism and apply this mechanism to our proposed stochastic energy trading game.

Combining the above result with the Bayesian learning approach discussed in Section VI-A, we can propose the following distributed optimization algorithm:

Description of Algorithm 1

Initialization: Each EHD k has a prior belief $\tilde{B}_{k,0}\left(\boldsymbol{a}_{-k,0}, \eta_0, \mu_0(k)\right)$ $\forall k \in \mathcal{U}$.

FOR $t = 1, 2, \ldots$

1) At the beginning of each time slot t, each seller EHD $j \in \mathcal{S}_t$ broadcasts its identity and energy information to buyer EHDs as described in Protocol 1. Each buyer EHD $k \in \mathcal{B}_t$ decodes the broadcast information of the seller EHDs and obtains an observation $o_{k,t}$.
2) Each buyer EHD k updates its belief function $\tilde{B}_{k,t}\left(\boldsymbol{a}_{-k,t}, \eta_t, \mu_t(k)\right)$ according to (11).
3) Each buyer EHD $k \in \mathcal{B}_t$ uses the updated belief to update its value function $V_{k,t}\left(\eta_t, \boldsymbol{a}_t\right)$ according to (14).
4) Each buyer EHD $k \in \mathcal{B}_t$ decides its optimal action $a^*_{k,t}$ according to (16) and then sends an energy request to the seller EHD $a^*_{k,t}$.
5) Each seller EHD $j \in \mathcal{S}_t$ decides to accept the request from the buyer EHD and conducts the energy transfer during the rest of the time slot.

We can prove the following results for the above algorithm.

Theorem 1. *For Algorithm 1:*

1) *Suppose in some time slot t, the energy trading policy for each EHD satisfies $\pi_{k,t} = \pi^*_k$ where π^*_k is a stationary policy for EHD k. Then we have $\pi_{k,\tau} = \pi^*_k\ \forall\ \tau \geq t$.*
2) *If the belief functions of the EHDs converge to a stationary probability distribution, then the policy in (16) is the optimal policy for every initial interactive state.*

Proof: See Appendix B. ■

The first part of Theorem 1 claims that any stationary policy is an absorbing solution for the energy trading game [10], [37]. The second part proves that if all the EHDs converge to a stationary policy using Algorithm 1, the resulting energy trading policy will be optimal.

VII. Numerical Results

In this section, we present numerical results to evaluate the performance of our proposed energy trading framework and optimization algorithms. We also compare the

Distance	0.5cm	2cm	5cm	10cm	20cm	30cm	40cm
Efficiency	89%	87%	74%	53%	34%	19%	11%

Fig. 4: Average energy transfer efficiency under different distances given in Table 4 of [19].

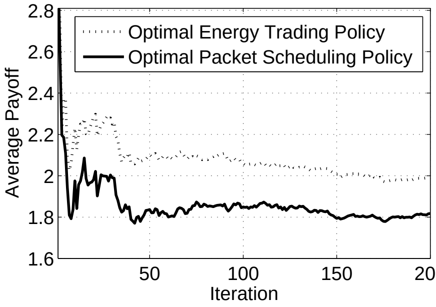


Fig. 5: Comparison of the average payoff of EHDs achieved by optimal energy trading policy and optimal packet scheduling policy under different iterations of the algorithms.

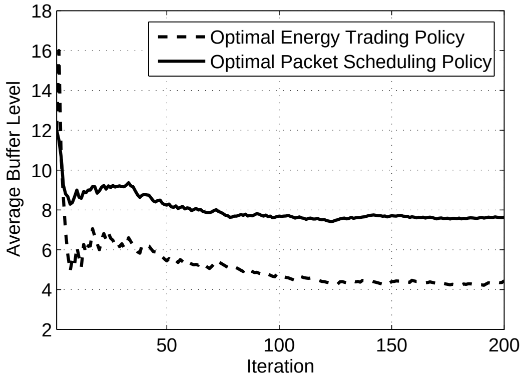


Fig. 6: Comparison of the average buffer level of EHDs for different iterations of the algorithms.

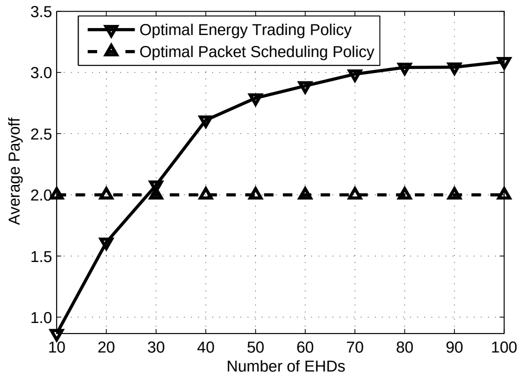


Fig. 7: Comparison of the average payoff of EHDs achieved by the optimal energy trading policy and the optimal packet scheduling policy for different numbers of EHDs.

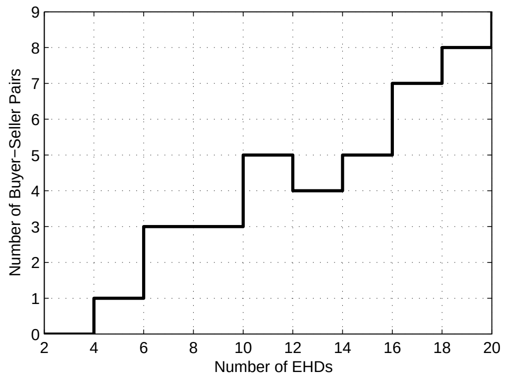


Fig. 8: Comparison of the number of buyer-seller pairs, for different numbers of EHDs.

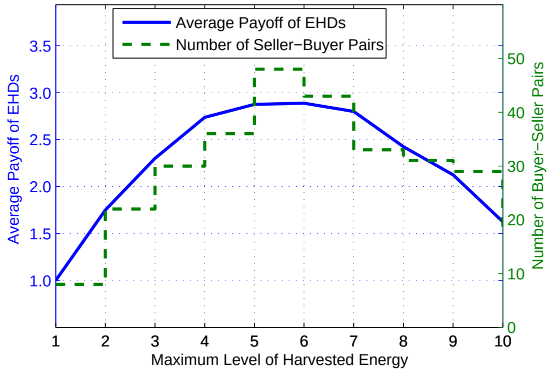


Fig. 9: Comparison of the average payoff of EHDs and number of buyer-seller pairs under different harvested energy for each EHD.

performance improvement brought by the proposed energy trading approach with that of existing optimal packet scheduling approaches [6], [16]. Before we describe the detailed setting of our simulation, we briefly explain the transmit packet scheduling problems and optimal solutions. The basic idea behind the transmit packet scheduling is that each EHD can take advantage of the known statistics of future state changes and sequentially optimize the number of data packets to be transmitted in each time slot. The transmit packet scheduling problem for each EHD k can be formulated as a single-player POMDP in which the state corresponds to the amount of harvested energy and the number of arriving data packets of each EHD k. EHD k should sequentially decide the number of data packets to transmit at the beginning of each time slot. The objective of each EHD is to maximize its long-term average payoff [15]. The optimal packet scheduling policy can then be derived by a standard value iteration approach for a single-agent POMDP [43]. Unlike our proposed dynamic energy trading policy, which exploits the amount of energy harvested and the number of data packets that arrive at different EHDs in each time slot, the optimal packet scheduling policy takes advantage of the diversity of the energy harvesting process during different time slots.

A. Simulation Setup

As mentioned in Section IV, the dynamic energy trading framework is very general and the payoff function of each EHD can be any performance metric. In this section, we define a simple linear model to demonstrate the performance improvement that can be achieved by our proposed framework. More specifically, we assume that each seller EHD charges the same price when it sells its excess energy to different buyer EHDs. We model the benefits obtained by each seller EHD $k \in \mathcal{S}_t$ from successfully transmitting data packets and selling energy to other EHDs as linear functions of $v_{k,t}$ and $\Delta q_{k,t}$, respectively. We can write the reward of each seller EHD k obtained by successfully transmitting data packets and selling $\Delta q_{k,t}$ to energy trading partner EHD j as follows:

$$\pi^1_{kj,t} = \xi_k \Delta q_{k,t} \text{ and } \pi^2_{k,t} = \alpha_k v_{k,t}, \tag{17}$$

where α_k is the reward obtained by each EHD k for successfully transmitting each data packet to its corresponding receiver and ξ_k is the price charged by seller EHD k for selling each unit of its energy to a buyer EHD. In the case that the buffer of EHD k is almost full and EHD k cannot obtain enough energy to transmit the newly arriving data packets, some of the arriving data packets will be discarded due to the limited buffer size. We hence define the cost of dropping the data packets of EHD k in time slot t as follows:

$$c^2_{k,t} = \lambda_k \max\left\{0, (\hat{u}_{k,t} + u_{k,t} - v_{k,t} - \bar{u}_k)\right\}, \tag{18}$$

where λ_k is the cost of dropping one data packet of EHD k. We also assume that the energy trading cost of each buyer EHD j when it trades energy with EHD k is given by $c^1_{jk} = \xi_k \Delta q_{k,t}$. Note that, from the previous discussion, only buyer EHDs will discard data packets in each time slot. We hence can write the instantaneous payoff obtained by each EHD k when it buys or sells energy from or to EHD j in time slot t as follows:

$$\varpi_{k,t} = \begin{cases} \alpha_k v_{k,t} + \xi_k \Delta q_{k,t}, & \text{if } \mu_t(k) \neq k \text{ and } k \in \mathcal{S}_t, \\ \alpha_k v_{k,t} - \xi_j \Delta q_{j,t}, & \text{if } \mu_t(k) \neq k \text{ and } k \in \mathcal{B}_t, \\ \alpha_k v_{k,t}, & \text{if } \mu_t(k) = k. \end{cases} \tag{19}$$

In the rest of this section, we will compare the performance of the optimal packet scheduling policy against that of the optimal energy trading policy derived in Section VI.

B. Numerical Results

We consider the payoff function given in (19). Since there is no energy trading among EHDs, the payoff of each EHD k in packet scheduling depends only on the number of its transmitted data packets and the cost of data loss. EHDs are equipped with multiple magnetic resonant coils and can use Magnetic MIMO to exchange energy with each other. All EHDs are randomly located in the coverage area uniformly, and a buyer-seller pair can only be formed between a buyer EHD and a seller EHD that are at a distance less than or equal to 0.4 meter. We assume that in each 10 ms time slot, the harvested energy of each EHD follows a discrete uniformly random distribution of integer values from 0 to 100 mW. The number of arriving data packets of each EHD is also uniformly random and takes an integer value between 0 to 50 data packets. Each data packet consists of 8 bits. Each EHD has a data buffer that can store up to 50 data packets and can transmit both the data packets stored in its buffer and the newly arriving data packets over a channel with the bandwidth of 1 MHz. Our setting is consistent with some energy harvesting-based smart card applications [44]. For example, EHDs can correspond to the energy harvesting-based RFID tag applications for grocery stores to monitor and manage the stock. In this application, each stock item has been installed with a small RFID tag which can harvest energy from ambient RF energy and instantaneously transfer its identity to the inventory management systems of the stores. Another application of our setting can be the security-enhanced bus/metro cards which can wirelessly verify the identity of a passenger at the entrance or exit on the bus or at the metro station using the energy harvested from the ambient environment. In both of the above applications, the EHDs are small and impractical to be equipped with a battery. In addition, the density of EHDs in these applications is high and each EHD needs to instantaneously communicate with remote computer systems and report its identity/security information using the energy obtained from harvesting and wirelessly transferred from other EHDs. We follow the same model described at the beginning of Section III and the amount of required energy for an EHD to send each data packet is 1.387 mW. We follow the same setting as [19] and the energy transfer efficiency between two EHDs follows the average efficiency results provided in Table 4 of [19]. It reports the average energy transfer efficiencies for distances of 0.5cm, 2cm, 5cm, 10cm, 20cm, 30cm and 40cm (See Figure 4). Following the same line as energy transfer simulation setups in [1], [2], we approximate the average efficiency at any two neighboring distances given in Table 4 of [19] by a linear function. Note that different types of functions will not fundamentally change the theoretical results of the game. The linear function is adopted to simplify the performance evaluation and presentation.

In Figure 5, we compare the average payoff of EHDs achieved by the optimal energy trading policy and the optimal packet scheduling policy for the different number of iterations in an energy trading system with 10 EHDs. We observe that in the first few iterations, there is no data packet loss for both approaches because each EHD can always store the data packets that cannot be sent in the current time slot into its data buffer. However, as the number of iterations increases, the optimal energy trading policy achieves a significant payoff improvement compared with the optimal packet scheduling policy. We also observe that the performance of our proposed energy trading policy relies on a proper matching between buyer and seller EHDs. In other words, dynamic energy trading cannot always improve the performance compared with a non-energy trading system, especially when the matching among EHDs is not optimized.

We compare the average buffer levels of the energy trading and transmit packet scheduling approaches in Figure 6. We observe that the average data buffer level when using the optimal transmit packet scheduling is much higher than that of the energy trading system with the optimal policy. This is because the optimal transmit packet scheduling allows each EHD to dynamically delay the transmission of data packets by adjusting the data buffer level according to its known state transition statistics. In other words, the optimal transmit packet scheduling policy is more suitable for delay-tolerant networking. By contrast, dynamic energy trading allows the EHDs to instantaneously exchange their harvested energy with each other. Therefore, the energy trading system has the potential to support delay-sensitive communication networks.

Compared with the optimal transmit packet scheduling, one of the major shortcomings of the energy trading system is that it can only be applied to the systems consisting of high density of EHDs with diverse energy harvesting processes. Moreover, the diversity of EHDs' energy harvesting processes directly affects the performance improvement brought by the energy trading. In Figure 7, we compare the payoff of the EHDs when the total number of EHDs changes. We observe that the payoff obtained by the optimal energy trading policy is much worse than that obtained by the optimal packet scheduling policy when the number of EHDs is small. This is because in our simulation we assume that all EHDs are uniformly distributed in a given area. Consequently, if the number of EHDs is small, the chance that each buyer EHD can always find a nearby seller EHD is low. However, with the increasing of the number of EHDs, the energy trading can significantly improve the payoff of the EHDs.

The performance of each buyer EHD in the energy trading system heavily relies on its opportunity to be paired with a seller EHD that can provide a sufficient amount of transferable energy. In Figure 8, we present the number of buyer-seller pairs formed between seller and buyer EHDs for different numbers of EHDs. We observe that if both the amount of energy required for sending data packets and the amount of energy harvested from the environment are randomly distributed uniformly around the same level, the number of buyer-seller pairs increases proportionally to the total number of EHDs in the network. However, this may not be the case if the amount of harvested energy and the energy required to transmit the arriving data packets are significantly different. In Figure 9, we fix the data packet arrival rate as described at the beginning of this section. We assume that the amount of energy that can be harvested by each EHD in each time slot is uniformly random and takes a discrete value with a different maximum bound. We compare the payoff of the EHDs and the number of buyer-seller pairs with different possible harvested energy levels. We observe that, if the maximum level of harvested energy for each EHD is much lower than the amount of energy required for sending data packets, most of the EHDs in the system become buyer EHDs. In this case, only a limited number of buyer-seller pairs can be formed and the average performance of all the EHDs is low. A similar observation can be made for the other extreme case, in which most of the EHDs can harvest more energy than they require. In this case, their performance will also be low because they cannot find buyer EHDs to sell their energy to. Intuitively, the performance of the energy trading can always be maximized when the numbers of seller and buyer EHDs are similar.

VIII. Extensions and Future Works

Our proposed dynamic energy trading market allows the EHDs to dynamically trade their harvested energy. Our framework can be extended for more complex and general systems. In this section, we describe some of the possible extensions of our energy trading framework.

A. Energy Trading with Multi-buyer and Multi-seller Pairing

In Section III, we assume that each buyer-seller pair can only consist of one buyer EHD and one seller EHD. The efficiency of the energy utilization can be further improved by allowing multiple buyer EHDs to jointly request and cooperate with each other to divide the total amount of energy transferred from multiple seller EHDs. One direct extension of our model is to replace the two-sided one-to-one stable matching game with other variations of stable matching games [29], [45], [46]. For example, if each seller EHD can send its harvested energy to multiple buyer EHDs, we can model this case using the two-sided many-to-one matching game, also called the college admission game. In this model, each seller EHD will be matched with multiple buyer EHDs for energy trading in each time slot. A more general approach to allow multi-buyer and multi-seller pairing during energy trading is to further divide the total amount of transferable energy harvested by each seller EHD into smaller units. Each buyer EHD can then request a certain number of energy units from one or more seller EHDs. We can then formulate this problem as a stochastic coalition formation game in which each EHD can estimate types and coalition formation actions of other EHDs [42]. Unfortunately, it is known that the core of the coalition formation game can be empty and, even if it is non-empty, finding a coalition formation structure that is in the core is generally an NP-hard problem.

B. Energy Trading with Unknown State Transitions, Observation Functions and Reward Functions

In our proposed energy trading framework, we assume that each EHD can know the state transition function and observation function. The optimal policy of each EHD relies on establishing a belief function about future states and interactions among EHDs based on these known probability distributions. If each EHD cannot know the future statistics of the energy harvesting and data arrival processes, it can learn this information from its past experience. In this case, a fundamental tradeoff between how to utilize the knowledge that EHDs have already learned to maximize performance (exploitation) and how to obtain new knowledge to further improve the performance of EHDs (exploration) arises. It has been shown in [6] that, by applying a Bayesian reinforcement learning approach for each EHD to learn the statistics of the energy harvesting process, the above tradeoff can be addressed by letting each EHD sequentially learn the statistics and update the optimal energy scheduling policy.

C. Dynamic Pricing and Mechanism Design

In this paper, we assume that the price for each seller EHD to sell its energy to buyer EHDs is fixed. A proper pricing mechanism can be introduced to further incentivize the energy trading among EHDs by allowing each EHD to adjust its price dynamically. A proper mechanism should be designed which not only incentivizes the energy trading between seller and buyer EHDs but also enforces truth-telling and fairness in the energy trading among EHDs [28].

D. Open Problems and Future Works

From the previous discussion, it can be observed that our proposed dynamic energy trading market is general and can be applied to analyze other more complex systems. Our results also point towards some new directions for future research. For example, future networks will consist of high density of moving EHDs with fast changing energy harvesting processes. A fast energy trading partner-discovery protocol should be developed to allow each EHD to quickly detect its nearby buyer and seller EHDs. Furthermore, it is known that if each EHD can obtain an accurate prediction about the future energy harvesting processes, it can further improve its long-term average performance. However, in practical systems, it is generally impossible to always accurately predict the future changes of the energy harvesting and transfer process. It will be interesting to develop a unified framework that can characterize the relationship between the accuracy of prediction about future energy harvesting processes and the performance gain achieved by an optimal dynamic energy trading policy.

IX. Conclusion

We have studied multi-user energy harvesting communication systems in which different EHDs can harvest different amounts of energy and need to transmit a different number of data packets in different time slots. To analyze this system, we have introduced a dynamic energy trading framework, called dynamic energy trading market. In this framework, seller EHDs that can harvest more energy than that they require can transfer parts of their harvested energy to buyer EHDs that cannot harvest sufficient energy to support their services. Due to the time-varying nature of the environment, the role of each EHD as a seller EHD or a buyer EHD, as well as their possible decisions about energy trading partners, change with time. Each EHD cannot observe complete information regarding the harvested energy and number of data packets of other EHDs. We have introduced a simple energy trading scheduling protocol and formulated a novel game theoretic model called stochastic energy trading game to analyze the dynamic energy trading problem. We have derived the optimal energy trading policy for each EHD to sequentially optimize its decision about its energy trading partner. We have proved that our proposed policy can achieve an optimal sequence of matchings for the EHDs, under certain conditions. We have compared the proposed energy trading policy with an existing transmit packet scheduling policy and presented numerical results to verify the performance improvement brought by our proposed policy under various situations.

Appendix A
Proof of Proposition 1

Following Baye's rule, we have

$$\begin{aligned}
&\tilde{B}_{k,t}\left(\boldsymbol{a}_{-k,t},\eta_t,\mu_t(k)\right)\\
&= \Pr\left(\boldsymbol{a}_{-k,t},\eta_t,\mu_t(k)|o_{k,t},\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)\\
&= \Pr\left(\eta_t|o_{k,t},\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)\\
&\quad \Pr\left(\boldsymbol{a}_{-k,t}|o_{k,t}\right)\Pr\left(\mu_t(k)|\boldsymbol{a}_t,o_{k,t}\right)\\
&= \frac{\begin{pmatrix} \Pr\left(o_{k,t}|\eta_t,\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)\\ \cdot\Pr\left(\eta_t|\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)\\ \cdot\Pr\left(\boldsymbol{a}_{-k,t}|o_{k,t}\right)\Pr\left(\mu_t(k)|\boldsymbol{a}_t,o_{k,t}\right)\end{pmatrix}}{\Pr\left(o_{k,t}|\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)}\\
&= \zeta\Pr\left(o_{k,t}|\eta_t,\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)\\
&\quad \sum_{\eta_{t-1}\in\Upsilon}\Pr\left(\eta_t|\boldsymbol{a}_{t-1},\eta_{t-1}\right)\Pr\left(\boldsymbol{a}_{-k,t}|o_{k,t}\right)\\
&\quad \Pr\left(\mu_t(k)|\boldsymbol{a}_t,o_{k,t}\right)\Pr\left(\eta_{t-1}|\boldsymbol{a}_{t-1},b_{k,t-1}\left(\eta_{t-1}\right)\right)\\
&= \zeta\Theta\left(\eta_t,\boldsymbol{a}_{t-1},o_{k,t}\right)\sum_{\eta_t\in\Upsilon}\Gamma\left(\eta_t|\boldsymbol{a}_{t-1},\eta_{t-1}\right)b_{k,t-1}\left(\eta_{t-1}\right)\\
&\quad \Pr\left(\boldsymbol{a}_{-k,t}|o_{k,t}\right)\Pr\left(\mu_t(k)|\boldsymbol{a}_t,o_{k,t}\right), \qquad (20)
\end{aligned}$$

where ζ is the normalization factor which is independent of η_t to ensure $\sum_{\eta_t} b_{k,t}\left(\eta_t\right)=1$.

From the above equation, we can claim that the current belief of each EHD depends on the observation and system transition functions as well as its own past observations and belief function. All of them can be fully characterized by the information obtained from the current and the previous time slot. Thus, it is sufficient for each EHD to decide its current and future actions without requiring any further information. This concludes the proof.

Appendix B
Proof of Theorem 1

Let us first consider the first part of this theorem. The objective for each EHD is to maximize its long-term average payoff. From (12), we observe that each buyer EHD establishes its preference over seller EHDs using its belief function and the resulting instantaneous payoff obtained when it has been matched with each seller EHD. Suppose each EHD has already arrived at an optimal and stable policy in time slot t. Then, according to (16), we have $V_{\pi_{k,t}} > V_{\pi'_k}$. We will prove that in the next time slot, EHD k cannot further improve the expected long-term payoff by changing its policy. We can rewrite the updated belief function given in (11) in time slot $t+1$ as follows:

$$\begin{aligned}
&\tilde{B}_{k,t+1}\left(\eta_{t+1},\boldsymbol{a}_{-k,t+1},\mu_{t+1}(k)\right)\\
&= \Theta\left(\eta_{t+1},\boldsymbol{a}_t,o_{k,t+1}\right)\sum_{\eta_{t+1}\in\Upsilon}\Gamma\left(\eta_{t+1}|\boldsymbol{a}_t,\eta_t\right)\\
&\quad b'''_{k,t}\left(\eta_t\right)\left(\alpha\beta B_{k,t}\left(\boldsymbol{a}_{-k,t},\mu_t(k)\right)\right.\\
&\quad +\left(1-\alpha\beta\right)\mathbf{1}\left(\boldsymbol{a}_{-k,t+1}|o_{k,t+1}\right)\\
&\quad \left.\mathbf{1}\left(\mu_{t+1}\left(k\right)|o_{k,t+1},\boldsymbol{a}_{-k,t+1}\right)\right), \qquad (21)
\end{aligned}$$

Substituting the above updated belief function into (16), we can claim that the updated long-term average payoff $V_{k,t+1}$ can be regarded as a linear combination of $V_{k,t}$ and the instantaneous payoff of EHD k for its chosen action which can be regarded as a constant. In other words, the optimal policy that can maximize the value of $V_{k,t}$ in time slot t will also maximize $V_{k,t+1}$ during time slot $t+1$. In other words, no EHD will have an incentive to unilaterally deviate from its current policy in the future time slots.

Let us prove the second part of this theorem. It can be observed that if each EHD regards other EHDs as a part of the environment, we can define a new state, referred to as the interactive state, as a combination of state η_t, actions of other buyer EHDs $\boldsymbol{a}_{-k,t}$ and final matching result $\mu_t(k)$, denoted as $\eta'_t = \langle\eta_t,\boldsymbol{a}_{-k,t},\mu_t(k)\rangle$. If we can show that the transition probabilities of the new interactive state and observation function for each EHD are stationary probability distributions, we can then convert the dynamic energy trading with multiple EHDs as a single-EHD energy trading system. In the system, EHD k will try to sequentially optimize its actions to maximize the long-term average payoff based on its known beliefs and its statistics of the interactive state. Suppose the belief of each EHD k converges. We can write the transition probabilities of new interactive state as follows:

$$\begin{aligned}
&\Pr\left(\eta'_t|\eta'_{t-1},\boldsymbol{a}_{t-1}\right)\\
&= \Pr\left(\eta_t,\boldsymbol{a}_{-k,t},\mu_t(k)|\eta_{t-1},\boldsymbol{a}_{-k,t-1},\mu_{t-1}(k),a_{k,t-1}\right)\\
&= \sum_{o_{k,t}\in\Omega}\Gamma\left(\eta_t,\eta_{t-1},\boldsymbol{a}_{t-1}\right)\\
&\quad B_{k,t}\left(\boldsymbol{a}_{-k,t},\mu_t(k)\right)\Theta\left(\eta_t,\boldsymbol{a}_t,o_{k,t}\right). \qquad (22)
\end{aligned}$$

From the above equation, we observe that if all the beliefs of other EHDs are stationary, the interactive state as well as the transition state function and observation function of EHD k will also be stationary. Following the same line as the single-agent POMDP, we claim that the policy given in (16) maximizes the long-term expected payoff of each EHD. This concludes the proof.

References

[1] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 5834, pp. 83–86, Jul. 2007.

[2] X. Yu, S. Sandhu, S. Beiker, R. Sassoon, and S. Fan, "Wireless energy transfer with the presence of metallic planes," *Applied Physics Letters*, vol. 99, no. 21, p. 214102, 2011.

[3] M. Zahn, *Electromagnetic Field Theory: a problem solving approach*. John Wiley and Sons, Inc., 1979.

[4] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[5] V. Talla, B. Kellogg, B. Ransford, S. Naderiparizi, S. Gollakota, and J. R. Smith, "Powering the next billion devices with wi-fi," *arXiv preprint arXiv:1505.06815*, 2015.

[6] Y. Xiao, Z. Han, D. Niyato, and C. Yuen, "Bayesian reinforcement learning for energy harvesting communication systems with uncertainty," in *IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015.

[7] Y. Xiao, D. Niyato, Z. Han, and L. A. DaSilva, "Joint optimization for power scheduling and transfer in energy harvesting communication systems," in *IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, Dec. 2015.

[8] Y. Xiao, Z. Xiong, D. Niyato, and Z. Han, "Distortion minimization via adaptive digital and analog transmission for energy harvesting-based wireless sensor networks," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, FL, Dec. 2015.

[9] L. Xiao, P. Wang, D. Niyato, D. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *to appear in IEEE Communications Surveys Tutorials*, 2015.

[10] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge Univ Press, 2007.

[11] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou, "The complexity of computing a nash equilibrium," *SIAM Journal on Computing*, vol. 39, no. 1, pp. 195–259, May 2009.

[12] A. E. Roth and M. A. O. Sotomayor, *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge University Press, 1992, no. 18.

[13] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[14] A. Sinha and P. Chaporkar, "Optimal power allocation for a renewable energy source," in *IEEE National Conference on Communications*, Kharagpur, India, Feb. 2012.

[15] A. Aprem, C. Murthy, and N. Mehta, "Transmit power control policies for energy harvesting sensors with retransmissions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 895–906, Oct. 2013.

[16] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas in Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[17] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Oct. 2013.

[18] L. R. Varshney, "Transporting information and energy simultaneously," in *IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, Jul. 2008.

[19] J. Jadidian and D. Katabi, "Magnetic MIMO: How to charge your phone in your pocket," in *ACM MobiCom*, Maui, Hawaii, USA, Sep. 2014.

[20] Y. Urzhumov and D. R. Smith, "Metamaterial-enhanced coupling between magnetic dipoles for efficient wireless power transfer," *Phys. Rev. B*, vol. 83, p. 205114, May 2011. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevB.83.205114

[21] A. M. Fouladgar and O. Simeone, "On the transfer of information and energy in multi-user systems," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1733–1736, Nov. 2012.

[22] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting communications," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4884–4898, Dec. 2013.

[23] Z. Ding and H. Poor, "Cooperative energy harvesting networks with spatially random users," *IEEE Signal Processing Letters*, vol. 20, no. 12, pp. 1211–1214, Dec. 2013.

[24] S. Lee, L. Liu, and R. Zhang, "Collaborative wireless energy and information transfer in interference channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 545–557, Jan. 2015.

[25] Y. Xiao, Z. Han, and L. A. DaSilva, "Opportunistic relay selection for cooperative energy harvesting communication networks," in *IEEE Global Communications Conference (GLOBECOM)*, Austin, TX, Dec. 2014.

[26] K. Huang and V. Lau, "Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 902–912, Feb. 2014.

[27] 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)*, Mclean, VA, Apr., 2014.

[28] Y. Xiao, Z. Han, K. C. Chen, and L. A. DaSilva, "Bayesian hierarchical mechanism design for cognitive radio networks," *IEEE J. Sel. Area Commun.: Cognitive Radio Series*, vol. 33, no. 5, pp. 986–1001, May 2015.

[29] Y. Xiao, K.-C. Chen, C. Yuen, Z. Han, and L. DaSilva, "A Bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 4034–4051, July 2015.

[30] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[31] S.-H. Kim, Y.-S. Lim, and S.-J. Lee, "Magnetic resonant coupling based wireless power transfer system with in-band communication," *Journal of Semiconductor Technology and Science*, vol. 13, no. 6, pp. 562–568, Dec. 2013.

[32] 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.

[33] N. Naderializadeh and A. Avestimehr, "ITLinQ: A new approach for spectrum sharing," in *IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, Mclean, VA, Apr. 2014.

[34] A. Neyman and S. Sorin, *Stochastic games and applications*. Springer Science & Business Media, 2003.

[35] P. Gmytrasiewicz and P. Doshi, "A framework for sequential planning in multiagent settings," *Journal of Artificial Intelligence Research*, vol. 24, no. 1, pp. 49–79, Jul. 2005.

[36] L. M. Dermed and C. L. Isbell, "Solving stochastic games," in *Proceedings of the Twenty-Third Annual Conference on Neural Information Processing Systems*, Vancouver, Canada, Dec. 2009.

[37] D. Fudenberg and J. Tirole, *Game Theory*. The MIT Press, Cambridge, MA, 1991.

[38] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Multivariate Distributions, volume 1, Models and Applications*. John Wiley & Sons: New York, 2002.

[39] B. A. Frigyik and M. R. Gupta, "Introduction to the Dirichlet distribution and related processes," Department of Electrical Engineering, University of Washington, Tech. Rep. UWEETR-2010-0006, 2010.

[40] Y. W. Teh, "Dirichlet processes," in *Encyclopedia of Machine Learning*. Springer, 2010.

[41] J. K. Goeree and C. A. Holt, "Stochastic game theory: For playing games, not just for doing theory," *Proceedings of the National Academy of Sciences*, vol. 96, no. 19, pp. 10 564–10 567, Sep. 1999. [Online]. Available: http://www.pnas.org/content/96/19/10564.abstract

[42] G. Chalkiadakis and C. Boutilier, "Sequentially optimal repeated coalition formation under uncertainty," *Autonomous Agents and Multi-Agent Systems*, pp. 1–44, Nov. 2010.

[43] L. Kaelbling, M. Littman, and A. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1, pp. 99–134, 1998.

[44] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," in *ACM Sigcomm*, Hong Kong, China, Aug. 2013.

[45] K. Iwama and S. Miyazaki, "A survey of the stable marriage problem and its variants," in *IEEE International Conference on Informatics Education and Research for Knowledge-Circulating Society (ICKS)*, Kyoto, Japan, Jan. 2008.

[46] Y. Xiao, Z. Han, C. Yuen, and L. A. DaSilva, "Carrier aggregation between operators in next generation cellular networks: A stable roommate market," *to appear at IEEE Transactions on Wireless Communications*, 2015.

Yong Xiao (S'09-M'13-SM'15) 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. He is a Senior Member of IEEE.

Dusit Niyato (M'09-SM'15) is currently an Assistant Professor in the School of Computer Engineering, at the Nanyang Technological University, Singapore. He obtained his Bachelor of Engineering in Computer Engineering from King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand. He received his Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada. His research interests are in the areas of radio resource management in cognitive radio networks and broadband wireless access networks. Dr. Niyato has several research awards to his credit, which include the 7th IEEE Communications Society (ComSoc) Asia Pacific (AP) Young Researcher Award, the IEEE Wireless Communications and Networking Conference (WCNC) 2012 Best Paper Award, the IEEE Communications Conference (ICC) 2011 Best Paper Award, and the 2011 IEEE Communications Society Fred W. Ellersick Prize paper award. Currently he serves as an Editor for the IEEE Transactions on Wireless Communications and an Editor for the IEEE Wireless Communications Letters.

Zhu Han (S'01-M'04-SM'09-F'14) 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 a Professor in Electrical and Computer Engineering Department as well as Computer Science 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 received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, several best paper awards in IEEE conferences, and is currently an IEEE Communications Society Distinguished Lecturer.

Luiz A. DaSilva (S'97-M'98-SM'00) (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.