IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 23, 2020, accepted October 4, 2020, date of publication October 15, 2020, date of current version October 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3031325

Dynamic Energy Trading and Load Scheduling Algorithm for the End-User in Smart Grid

DIDI LIU[1], JIAWEN XIAO[1], JUNXIU LIU[1], (Member, IEEE), XIAOMING YUAN[2], (Member, IEEE), AND SUPING ZHANG[1]

[1]College of Electronic Engineering, Guangxi Normal University, Guilin 541004, China
[2]School of Computer and Communications Engineering, Qinhuangdao Branch Campus, Northeastern University, Qinhuangdao 066004, China

Corresponding authors: Didi Liu (ldd866@gxnu.edu.cn) and Junxiu Liu (j.liu@ieee.org)

This work was supported in part by the National Natural Science Foundation of China under Grant 62061006 and Grant 61976063, in part by the Natural Science Foundation of Guangxi under Grant 2018JJA170167, in part by the Guangxi Science and Technology Base and Special Talent Program under Grant 2018AD19342, in part by the Fundamental Research Funds for the Central Universities under Grant N182303032, and in part by the Diecai Project of Guangxi Normal University.

ABSTRACT In the smart grid, the end-users have the opportunity to integrate renewable energy sources (RESs) and participate in two-way energy trading. At the same time, an increasing number of flexible loads (FLs) have been developed for use on the demand side. Thus, this article considers joint energy trading and load scheduling at a end-user with integrated renewable generation. With unknown statistics on renewable generation, loads and electricity prices, we aim at optimizing both energy trading and load scheduling to maximize the long-term average profit of the end-user, subject to load delay constraints. We employ the Lyapunov optimization theory to solve this stochastic problem with a series of problem corrections and transformations that enable us to design a dynamic energy trading and load scheduling algorithm. Within the performance analysis of the algorithm, we further demonstrate that the algorithm not only provides a bounded performance guarantee to the optimal solution that has complete future information, but is also asymptotically equivalent to it as the battery capacity or the delay time of FLs tend to infinity. The simulation results show that the proposed algorithm is superior to other algorithms both in terms of performance and service delay. Moreover, we can achieve a trade-off between comfort and total profit by adjusting the values of the parameters, and analyze the effect of battery capacity on algorithm performance to provide a theoretical basis for the end-user to determine battery capacity.

INDEX TERMS Smart grid, renewable energy sources, energy trading, flexible loads, load scheduling.

I. INTRODUCTION

As a key technology in the smart grid, the two-way flow of electricity and information enables the grid to deliver electricity efficiently and reliably, and provides more flexibility in demand response. With the aid of renewable energy technologies, the smart grid enhances the energy autonomy of residential consumers and reduces carbon emissions [1], [2]. As a result, on the consumers side, they are able to achieve the two-way energy trading with the grid to more efficiently utilize their locally generated renewable energy sources (RESs) for saving their energy bills. RESs have many advantages over conventional energy sources, such as efficiency, cleanliness, ease of installation, etc [3]. However, in practice, their integration into the existing grid infrastructure must be carried out carefully in order to avoid instability and ensure availability and security of supply. The reason is that the high penetration levels of RESs can lead instability problems in the grid due to their limited predictability, controllability and variability [4]. To meet this challenge, numerous solutions for renewable energy generation forecasting have been proposed in the literature such as the hybrid of genetic algorithm and artificial neural networks [5], integration of wavelet transform, particle swarm optimization and support vector machines [6]. However, although the performance of forecasting tools is getting better, there is still room for improvement in terms of accuracy, reliability and sharpness. Fortunately, energy storage and flexible loads (FLs) are recognized as two promising management options to alleviate the randomness of renewable generation and reduce power

The associate editor coordinating the review of this manuscript and approving it for publication was Bin Zhou.

costs [7], [8]. Specifically, energy storage can be utilized to shift energy over time, while FLs can be controlled to shift demand over time. For the grid operator, they can be deployed to mitigate fluctuations in renewable generation and to enhance reliability [9]. For the demand side, energy storage and FL scheduling can provide an effective approach for energy management to reduce the cost of electricity consumption [10].

Further encouragement of the demand side to actively participate in energy management and trading mechanisms, as well as the efficient blending of FLs (include energy storage devices) and RESs into the energy system, are major objectives in planning the future smart grid. Meanwhile, a number of FLs have been developed for use on the consumer side. They can be controlled to increase demand flexibility and provide economic incentives to consumers by shifting energy demand from high-peak to low-peak periods [11], [12]. Therefore, developing an effective FLs scheduling scheme will be the most promising future solution for consumers participation in energy trading to maximize economic benefits and is the goal of this article. However, there are some challenges facing our work. First, the fluctuations in real-time demand and price information pose a challenge to the energy trading mechanism. Second, for energy storage devices (batteries), the performance of the battery will degrade gradually due to frequent charging and discharging activities. Third, for load (demand) scheduling, there is a need to ensure that the delay requirements of FLs are met while maximizing the profits for the consumer. Moreover, load scheduling decisions influence the energy usage and the amount of charging and discharging of the battery. Thus, battery control and load scheduling decisions are interacting over time, which makes global optimization particularly challenging.

In this article, we consider the issue of energy trading for an end-user in smart grid, such as a company, a home, or a community with renewable generators, a battery and FLs. Renewable generators can harvest RESs from the surrounding environment to meet the user's energy consumption. In addition, the battery is able to exploit the delay requirement form FLs to reduce the cost of energy usage during high-peak periods of demand. The reason is that batteries can draw energy from the external grid when the price of electricity is low and discharge energy for use by the end-user or sell to the external grid when the price of electricity is high. Thus, the aim of this work is to design an real-time solution for the end-user to maximize the long-term average profit while satisfying the FLs delay constraint. It is worth noting that we assume the dynamics of RESs, loads, and electricity prices are arbitrary or non-stationary and their statistics are unknown to fit the actual situation. To solve this stochastic problem, we develop techniques through a series of problem corrections and transformations that enable us to design an real-time algorithm using Lyapunov optimization. Furthermore, we further demonstrate that our proposed dynamic energy trading and load scheduling algorithm not only provides a bounded performance guarantee for the optimal solution that has complete future information, but is also asymptotically equivalent to it as the battery capacity or the delay time of the FL tends to infinity. In summary, the main contributions of this work can be summarized as follows.

- We propose an effective energy trading and management scheme for the end-user to maximize economic benefits, subject to battery operation and load delay constraints.
- Technically, we construct a variable to replace the charge level in the battery based on the quadratic Lyapunov function, so as to drive the charge level in the battery towards a certain non-zero value to avoid underflow. Moreover, the virtual queue technique is adopted to ensure that the maximum delay time for FLs does not exceed a given value.
- The solution depends only on current electricity price, renewable generation or load and does not require any statistical knowledge of them, and is infinitely close to the optimal solution which has full the future values or statistical knowledge of them.
- In the simulation experiments, we can adjust the values of the parameters to achieve a trade-off between the comfort and the total profit for the end-user, and the impact of the battery capacity on the performance of the proposed algorithm is analyzed to provide a theoretical basis for the end-user to determine the storage battery capacity.

The remainder of this article is organized as follows. In Section II, we briefly present a review of some literature related to load scheduling and energy trading in the smart grid. Section III introduces the system model. In Section IV, we present formulation of the problem of maximization profits for the end-user. Section V describes the solution to the problem and analyzes the performance of proposed algorithm in theory. Section VI presents the simulation results, and Section VII summarizes the paper.

II. RELATED WORK

As an important tool for regulating electricity supply and demand, research on optimal scheduling of FLs has been a hot issue in the previous years, e.g., [13]–[19]. For instance, the earlier work [13] has investigated the problem of scheduling multiple devices that allow different levels of latency tolerance under real-time prices. Similarly, the power scheduling strategies under day-ahead prices have been proposed in [14], [15], and [16] has designed a load scheduling scheme that can address the dynamic behavior for customer's energy requirements to minimize the load on the grid during peak hours. In these works, the storage battery or the storage capacity limitation was not considered. Energy storage is a key technology to mitigate generation demand and smooth out energy supply uncertainties. For real-time storage management design, [17] has proposed joint load scheduling and energy storage control schemes to save energy consumption costs. However, RESs were not included in the proposed model, and the designed

TABLE 1. Comparison of existing works.

Key reference	Energy trading	Storage control	Demand scheduling	Stochastic optimize
[13], [17]	×	√	√	×
[14]-[16]	×	×	√	×
[18]	×	√	√	√
[19]	×	×	√	√
[20]	√	×	√	√
[21]	√	×	√	×
[22]-[24]	√	√	√	×
[25], [26]	√	×	×	×
[27], [28]	√	×	×	√
This paper	√	√	√	√

model was not completely stochastic. Both RESs generation and stochastic optimization have been modeled in [18], [19], in which RESs generations and energy arrives are assumed to be unknown, but they have not considered two-way trading with the grid.

Energy trading through demand-side management has been studied by many researchers. Typically, these works can be divided into two major categories. The first category of works focused on developing the optimal energy consumption scheduling for consumers in response to the pricing of the retail market, e.g., [20]–[24]. Specifically, without assuming known future prices, a stochastic optimization scheme has been formulated in [20] to maximize both individual profits and social welfare. Similarly, [21] has proposed a energy trading model with uncertain demands and prices, in order to maximize the profit values for both consumers and the grid. They proposed two algorithms to predict the electricity prices in the grid and demands from consumers, respectively. However, electricity prices, renewable generation, and demand may all fluctuate randomly with their statistics likely being non-stationary, making them difficult to predict accurately. On the contrary, assuming that electricity prices are is known ahead of time, linear programming (LP) [22], [23] and dynamic programming (DP) [24] techniques are frequently applied for energy trading. However, the solutions obtained in there works have some randomness and there are no uniform conditions for judging whether the proposed algorithm can converge to the global optimum.

Considering the coexistence of a large number of end-users, the second category of works further investigated how to help energy retailers (e.g., utilities) to establish the pricing strategy in order to optimize certain targets. For instance, [25] and [26] have investigated the problem of load dispatching and energy trading for multi-user in the energy market, and modeled double-auction mechanisms using the non-collaborative game theory (GT) approach. In addition, reinforcement learning (RL) theory is applied to address information privacy behavior among competitors. In particular, [27] has considered the issue of the constrained energy trading game among individual strategic players in the case of incomplete information. Also, [28] has proposed a dynamic learning algorithm for the energy trading game among smart MGs to maximize the finite average utility of each MG. However, the trading architectures proposed by these works is centralized, which not only add additional service costs but also consume significant computational resources. Moreover, there is a high probability that some sellers or buyers will face an exit situation, either because they are not getting a matching entity or the expected price from the market. Therefore, in such a scenario, it is necessary to investigate two-way trading between an individual end-user and the smart grid.

A brief comparison of the different aspects for the existing works is shown in Table 1. Most of the extant literatures have proposed energy management schemes that are optimized under the assumptions of renewable generation, electricity prices, and demand that are known in advance or can be accurately predicted. However, as discussed earlier, the fluctuating nature of renewable generation, electricity prices, and demand make them difficult to predict accurately. Therefore, this article proposes a stochastic optimization scheme to maximize the total profit of the end-user through energy trade and load (demand) dispatch under the uncertainty of renewable generation, demand and price.

III. SYSTEM MODEL

An end-user is equipped with renewable generators and an energy storage battery, and connected to the external power grid through the smart meter, as shown in Fig. 1. The harvested RESs can be directly used by the end-user and also stored in the battery or sold to the smart grid for profit, which be controlled by the energy management unit (EMU). Moreover, the battery is allowed to charge/discharge energy from/to the external grid.

We assume the EMU operates in discrete time slots with $t \in \{0, 1, \dots\}$, and all operations are performed per time slot t. Let $e(t)$ denote the harvested renewable energy at time slot t. The electricity price is time-varying in the smart grid and the real-time unit price of the grid's energy supply denoted by $p(t)$ at time slot t. Let $s(t)$ represents the amount of electricity purchased/sold from/to the smart grid at time slot t. Owing to the two-way trading between the end-user and the external grid, the value of $s(t)$ may be negative, with $s(t) > 0$ meaning that electricity is purchased from the grid and the opposite implying that it is sold. Similarly, we assume that $b(t)$ represents the charge/discharge amount of the battery

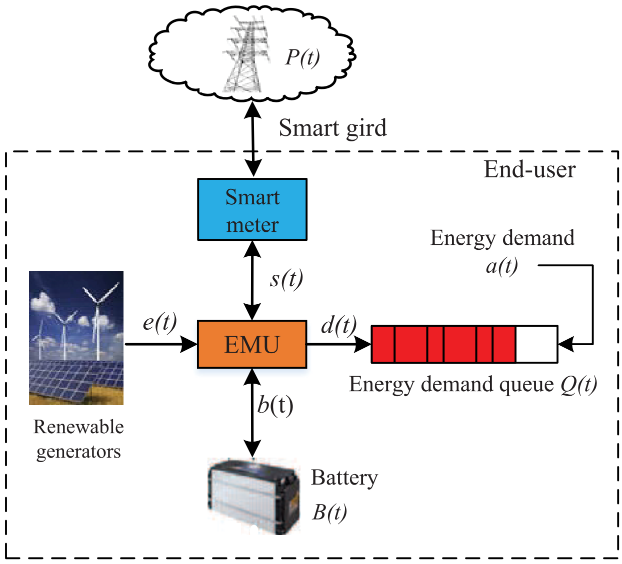

FIGURE 1. Energy trading and loading scheduling model of the end-user in the smart gird.

at the time slot t, with $b(t) > 0$ representing the battery discharging and $b(t) < 0$ representing the battery charging.

We establish an energy demand queue in our model to simplify the load scheduling over the maximum delay time. The total energy demand of FLs at time slot t is denoted as $a(t)$, with the maximum energy demand is denoted as $a_{\max}$. $a(t)$ is stored in the energy demand queue and supplied in a first-in first-out (FIFO) manner. Note that the waiting time for $a(t)$ in the energy demand queue does not exceed the maximum delay time $T_{\max}$. Let $Q(t)$ represent the total energy demand backlog in the time slot t, and let $Q_{\max}$ represent be the maximum energy backlog in the queue, in the next time slot, the equation can be updated by

$$Q(t+1) = \max\{Q(t) - d(t), 0\} + a(t) \tag{1}$$

The energy backlog is not negative, so we must use the max function to ensure that $Q(t) \geq 0$, where $d(t)$ denote the service energy of the end-user at t time slot and satisfies $d(t) \leq d_{\max}$. $d_{\max}$ is the maximum service volume.

Note that the end-user can draw energy from three sources, namely, the renewable energy source, the battery and the external power grid. In summary, the following equation can be obtained from Fig. 1.

$$d(t) = e(t) + s(t) + b(t) \tag{2}$$

Initially, we assume that there was a certain amount of charge in the battery. Let $B(t)$ represents the charge level in the battery at the beginning of time slot t, and $B_{\max}$ is recorded as the maximum battery capacity. From a practical viewpoint, the charge/discharge rate of the battery is limited due to the constraints of the hardware circuit. Therefore, we use $b_{\max}$ to represent the maximum charge/discharge rate of the battery. The relevant equations of the battery are as follows:

$$B(t+1) = B(t) - b(t) \tag{3}$$

$$0 \leq B(t) \leq B_{\max} \tag{4}$$

$$|b(t)| \leq b_{\max} \tag{5}$$

$$b(t) \leq B(t) \tag{6}$$

where Eq. (3) is the updated equation of the battery. Inequation (6) indicates that the charge/discharge amount of the battery should be less than or equal to the current battery level.

According to the current state (the electricity price $p(t)$, the energy arrival $e(t)$, the charge level of battery $B(t)$ and the energy demand backlog $Q(t)$), the EMU need decide whether buy or sell electricity from/to the external grid, charge or discharge the battery, and determine the corresponding electricity amount $s(t)$, $b(t)$ and $d(t)$. It is worth mentioning that in the proposed model the harvested energy can be directly used by the end-user through the EMU rather than be stored in the battery before be used. Thus our model effectively reduces the number of battery charge and discharge cycles to extend the battery lifetime.

IV. PROBLEM FORMULATION AND SOLUTION

As previously mentioned, we suppose the time-varying electricity price $p(t)$, the harvested renewable energy $e(t)$, and the energy demand process $a(t)$ that are independent and identically distributed over slots and unknown probability distribution. Our goal is to find an optimal policy $\{b(t), d(t), s(t)\}$ that maximizes the long-term averaged profit for the end-user. We assume that the renewable energy harvested is free (ignoring the fixed cost of renewable energy devices and the battery). The end-user make two-way energy trading with the external grid, when $s(t) < 0$, the end-user sells energy to the grid and the profit is given by

$$\beta p(t)(s(t))^{-} = \beta p(t)(d(t) - b(t) - e(t))^{-} \tag{7}$$

where $(f)^{-} \triangleq \max\{-f, 0\}$ is used here, β is a constant between 0 and 1, i.e., $0 < \beta < 1$. Because the profit of the energy supplier and the power loss in the transmission process are taken into account here, we stipulate that the price of the electricity sold by the end-user at time slot t should be less than the current the price of electricity purchased in the smart grid, so that the price of electricity sold for the end-user is set as $\beta p(t)$.

In contrast, when $s(t) > 0$, the end-user purchases electricity from the grid and the cost is given by

$$p(t)(s(t))^{+} = p(t)(d(t) - b(t) - e(t))^{+} \tag{8}$$

where $(f)^{+} \triangleq \max\{f, 0\}$.

The goal is to maximize the time-averaged profit for the end-user. Hence, the problem can be formulated as follows:

Problem One:

$$\begin{aligned}\max_{d(t),b(t)} : \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\big[&\beta p(t)(d(t) - b(t) - e(t))^{-} \\ &- p(t)(d(t) - b(t) - e(t))^{+}\big]\end{aligned}$$

$$\text{s.t. } \overline{Q} < \infty, \quad (2)-(6). \tag{9}$$

where $\overline{Q} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\{Q(t)\}$ represents the time-averaged expectation of the energy demand queue backlog, meaning that the energy demand queue has a limited backlog to maintain the queue stable [29].

We assume C^* as the optimal value to **Problem One** and $\bar{C}$ as the optimal value without the constraint (6) to **Problem One**. Obviously, the optimal value is larger owing to fewer constraints, therefore, $\bar{C} \geq C^*$. We will show that $\bar{C}$ can be achieved by some stationary and randomized policy in the form of a lemma. When the probability distribution process of $(p(t), e(t), a(t))$ is unknown, our objective function (9) is infinitely close to the optimal value $\bar{C}$ through the Lyapunov drift plus penalty algorithm (constant V algorithm). The performance analysis in Section V will be proved in detail later. As given by the following lemma, the optimal value $\bar{C}$ can be obtained by a randomized, stationary control policy that only chooses $d(t)$, $b(t)$ every slot purely as a (possibly randomized) function of $(p(t), e(t)$ and $a(t))$. That means the control decision is independent of the charge level in the battery. This fact is stated as below:

Lemma 1: We assume $(p(t), e(t), a(t))$ are i.i.d. over the time slot t, and their future knowledge have be given. Then, $\bar{C}$ can be achieved by a stationary and randomized policy, that is, the control action $\bar{d}(t)$, $\bar{b}(t)$ at each time slot are only function of $(p(t), e(t), a(t))$. We thus get

$$\bar{C} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left[\beta p(t)\big(\bar{d}(t) - e(t) - \bar{b}(t)\big)^- - p(t)\big(\bar{d}(t) - e(t) - \bar{b}(t)\big)^+\right] \tag{10}$$

$$E\{\bar{d}(t)\} = E\{e(t)\} + E\{s(t)\} + E\{\bar{b}(t)\} \tag{11}$$

$$E\{\bar{d}(t)\} \geq E\{a(t)\} \tag{12}$$

where (12) indicates that the time average of the energy service process should be greater than or equal to the time average of the energy arrival process.

Proof: Similar to the proofs in the work [18], we will not elaborate on this due to the space limitation.

V. DYNAMIC ENERGY TRADE ANALYSIS

A. DELAY-AWARE VIRTUAL QUEUE

It can be seen from the objective function (9) that the maximum delay-related information is not included in the formula. Therefore, we introduce a method using the "virtual queue" technology to ensure the maximum delay time of FLs. More specifically, we need to construct a virtual queue $Z(t)$ and record $Z_{\max}$ as the maximum energy capacity with $Z(0) = 0$. In order to make $T_{\max}$ controllable, we introduce a fixed positive parameter λ that specified later, and the updated equation for $Z(t)$ can be define as follows:

$$Z(t+1) \triangleq \max[Z(t) - d(t) + \lambda 1_{\{Q(t)>0\}}, 0], \quad \forall t. \tag{13}$$

Compared with (1), the virtual queue $Z(t)$ has the same service process $d(t)$ as the real queue $Q(t)$, and their difference is the energy arrival process. Here, $1_{\{Q(t)>0\}}$ is an indicator function, when the real queue backlog is not zero, that is, $Q(t) > 0$, its value is 1, otherwise it is 0. λ will be used as a penalty for the virtual queue to prevent the real queue $Q(t)$ has not been serviced for a long time. Moreover, λ, $Q_{\max}$ and $Z_{\max}$ together determine the value of the maximum delay time $T_{\max}$, as shown in **Lemma 2**.

Lemma 2: Suppose $Q_{\max}, Z_{\max}$ are some positive constants within the controllable range of the system, and any time slot satisfies $Q(t) < Q_{\max}, Z(t) < Z_{\max}$, then the maximum delay time $T_{\max}$ for all energy demand is

$$T_{\max} = [\frac{(Q_{\max} + Z_{\max})}{\lambda}] \tag{14}$$

Proof: See Appendix A

Before using the Lyapunov optimization method to solve **Problem One**, we defined a queue $X(t)$ instead of the battery state change to satisfy the constraint condition (6), as follows:

$$X(t) = B(t) - \Phi_{\max} - b_{\max} \tag{15}$$

where $\Phi_{\max}$ is a given positive constant. The variable $X(t)$ is constructed based on the quadratic Lyapunov function, which push the charge level in the battery towards a certain non-zero value to avoid underflow and thus enable the battery to satisfy constraint (6). According to (3), and we can get the updated equation of $X(t)$ as

$$X(t+1) = X(t) - b(t) \tag{16}$$

B. LYAPUNOV OPTIMIZATION

Our algorithm is developed based on the Lyapunov optimization method. The advantage of this method is that we can make our objective function (9) infinitely close to the optimal solution without knowledge of the probability distribution process of $(p(t), e(t), a(t))$. First, we define the variable $\vec{M}(t) = (Q(t), Z(t), X(t))$ as the real-time state of all queues, and then define the Lyapunov function as follow:

$$L(\vec{M}(t)) \triangleq \frac{1}{2}[Q^2(t) + Z^2(t) + X^2(t)] \tag{17}$$

Second, according to the Lyapunov drift plus penalty algorithm, $\Delta\left(\vec{M}(t)\right)$ is defined as the Lyapunov drift part at time slot t, but the optimization objective is to maximize formula (9), and a negative sign should be taken at this time as the Lyapunov penalty part. Therefore, the solution to the previous **Problem One** can be transformed into a minimizing Lyapunov "drift plus penalty" function (**Problem Two**).

Problem Two:

$$\min_{d(t),b(t)} : \Delta(\vec{M}(t)) - V \cdot E\left\{\beta p(t)(d(t) - e(t) - b(t))^- - p(t)(d(t) - e(t) - b(t))^+ | \vec{M}(t)\right\}$$
$$\text{s.t.} \quad \overline{Q} < \infty, \overline{Z} < \infty, \quad (2) - (6). \tag{18}$$

Note that V is an important trade-off parameter, and its role is to balance the relationship between the end-user's profits and comfort. Finally, we can obtain the maximum upper bound of the Lyapunov drift-plus-penalty function by calculation, as shown in **Lemma 3**.

Lemma 3: For every time slot t, the control policy is independent $\vec{M}(t)$, and the drift-plus-penalty function satisfies:

$$\begin{aligned}\Delta\left(\vec{M}(t)\right) - V\cdot E\Big\{&\beta p(t)(d(t)-e(t)-b(t))^{-}\\ &-p(t)(d(t)-e(t)-b(t))^{+}|\vec{M}(t)\Big\}\\ \le S &+ [X(t)+Vp(t)]E\{(d(t)-e(t)-b(t))^{+}|\vec{M}(t)\}\\ &-[X(t)+V\beta p(t)]E\{(d(t)-e(t)-b(t))^{-}|\vec{M}(t)\}\\ &-[X(t)+Q(t)+Z(t)]E\{d(t)|\vec{M}(t)\}\\ &+X(t)E\{e(t)|\vec{M}(t)\}\end{aligned} \tag{19}$$

where the constant S is as follow:

$$S = \frac{1}{2}b_{\max} + \frac{[d^2{}_{\max} + a^2{}_{\max}]}{2} + \frac{\max[\lambda^2, d^2{}_{\max}]}{2} + Q_{\max}a_{\max} + Z_{\max}\lambda \tag{20}$$

Proof: See Appendix B

C. DYNAMIC ENERGY TRADING AND LOAD SCHEDULING ALGORITHM

Using the fact that the decision is independent of queue state $\vec{M}(t)$, and excluding the irrelevant terms of the decision variables in (19). **Problem Two** can be transformed into **Problem Three** as follows:

Problem Three:

$$\begin{aligned}\min_{b(t),d(t)} : &[X(t)+Vp(t)](d(t)-e(t)-b(t))^{+}\\ &-[X(t)+V\beta p(t)](d(t)-e(t)-b(t))^{-}\\ &-[X(t)+Q(t)+Z(t)]d(t)\end{aligned} \tag{21}$$

where **Problem Three** has the same constraints as **Problem Two**.

It is known from (15) that the value of $X(t)$ may be negative (the range of $X(t)$ specified later in the next section), so the values of $X(t)+Vp(t)$, $X(t)+V\beta p(t)$, and $X(t)+Q(t)+Z(t)$ may also be negative. Moreover, observe that $X(t)+Vp(t) > X(t)+V\beta p(t)$ always holds by the definitions. Thus, the solution for our proposed algorithm is comparing the relationship among $X(t)+Vp(t)$, $X(t)+V\beta p(t)$, and $X(t)+Q(t)+Z(t)$ at each time slot, and making decisions based on the weight to minimize (24) as a whole.

The energy trading and load scheduling scheme of the our proposed algorithm is presented as follows:

When $X(t)+Q(t)+Z(t) \ge 0$ is satisfied in the current time slot, these conditions should be considered.

1). If $X(t)+V\beta p(t) \ge 0$, we can get $B(t)-b_{\max} \ge \Phi_{\max} - V\beta p(t)$ according to formula (15). That means there is a lot of power amount in the battery, or a high electricity price in the smart grid, then we assign the maximum possible discharge rate to $b(t)$, that is, $b(t) = b_{\max}$. At the same time, if $Q(t)+Z(t) > Vp(t)$, it means the queue backlog is high at this time, then we assign the maximum possible service rate to $d(t)$, that is, $d(t) = \max(Q(t), d_{\max})$. If $V\beta p(t) \le Q(t)+Z(t) \le Vp(t)$, it shows that a medium energy backlog in the queue, so we choose $d(t) = \min(Q(t), e(t))$. If $Q(t)+Z(t) < V\beta p(t)$, that means the queue backlog is low, then set $d(t) = 0$.

2). If $X(t)+Vp(t) \le 0$, then we can get $Q(t)+Z(t) \ge Vp(t)$ and $B(t) - b_{\max} < \Phi_{\max} - V\beta p(t)$. That means there is a litter power amount in the battery, or a low electricity price in the smart grid, and the queue backlog is high at this time. Then we assign the maximum possible charging/service rate to $b(t)/d(t)$, that is, $b(t) = -\min(b_{\max}, B_{\max} - B(t))$, $d(t) = \min(Q(t), d_{\max})$.

3). If $X(t)+Vp(t) > 0 > X(t)+V\beta p(t)$, we can get $\Phi_{\max} - Vp(t) < B(t) - b_{\max} < \Phi_{\max} - V\beta p(t)$. That means there is a medium power amount in the battery, or a medium electricity price in the smart grid.At the same time, if $Q(t)+Z(t) > Vp(t)$, it means the queue backlog is high at this time, then we assign the maximum possible service rate to $d(t)$, that is, $d(t) = \min(Q(t), d_{\max})$, and set $b(t) = \min(b_{\max}, \min(B_{\max} - B(t), d(t) - e(t)))$. Conversely, if $V\beta p(t) \le Q(t)+Z(t) \le Vp(t)$, it means the queue backlog is medium at this time, then we set $d(t) = \min(Q(t), e(t))$, $b(t) = 0$. Here we give priority to buying/selling electricity form/to the grid rather than charging/discharging.

When $X(t)+Q(t)+Z(t) < 0$ is satisfied in the current time slot, these conditions should be considered.

4). If $X(t)+Vp(t) \le 0$, we can get $B(t) - b_{\max} \le \Phi_{\max} - Vp(t)$ according to formula (15). That means there is a litter power amount in the battery, or a low electricity price in the smart grid, then we assign the maximum possible charging rate to $b(t)$, that is, $b(t) = -\min(b_{\max}, B_{\max} - B(t))$. At the same time, if $Q(t)+Z(t) > Vp(t)$, it means the queue backlog is high at this time, then we assign the maximum possible service rate to $d(t)$, that is, $d(t) = \min(Q(t), d_{\max})$. If $V\beta p(t) \le Q(t)+Z(t) \le Vp(t)$, it shows that a medium energy backlog in the queue, so we choose $d(t) = \min(Q(t), e(t))$. If $Q(t)+Z(t) < V\beta p(t)$, that means the queue backlog is low, then set $d(t) = 0$.

5). If $X(t)+V\beta p(t) \ge 0$, then we can get $B(t) - b_{\max} \ge \Phi_{\max} - V\beta p(t)$ and $Q(t)+Z(t) < V\beta p(t)$. That means there is a lot of power amount in the battery, or a high electricity price in the smart grid, and the queue backlog is low at this time. Then we set $b(t) = b_{\max}$ and $d(t) = 0$ at this time.

6). If $X(t)+Vp(t) > 0 > X(t)+V\beta p(t)$, we can get $\Phi_{\max} - Vp(t) < B(t) - b_{\max} < \Phi_{\max} - V\beta p(t)$. That means there is a medium power amount in the battery, or a medium electricity price in the smart grid. At the same time, if $V\beta p(t) \le Q(t)+Z(t) < Vp(t)$, it shows that a medium energy backlog in the queue, so we choose $d(t) = \min(Q(t), e(t))$, $b(t) = 0$. If $Q(t)+Z(t) < V\beta p(t)$, that means the queue backlog is low, then set $d(t) = 0$, $b(t) = 0$.

The pseudocode of the proposed dynamic energy trading and load scheduling algorithm is illustrated in **Algorithm 1**. The idea of the our proposed algorithm is to minimize a upper bound of the drift-plus-penalty function in (21). From **Algorithm 1**, we can know that the control decisions only rely on the current electricity prices, renewable generation, and load requirements $\{p(t), e(t), a(t)\}$, and do not require any prior knowledge of them. They can be stochastic or non-stochastic with arbitrary dynamics. Once the load scheduling decision $d(t)$ and the charging and discharging decision $b(t)$ are determined, according to Eq. (2), the amount of electricity purchased/sold from/to the smart grid, i.e., $s(t)$, can be determined. Thus, our proposed algorithm is simple to implement and more applicable to general situations, especially when these statistical data are difficult to predict. Moreover, the complexity of the algorithm is linearly related to the total time slot T, so the algorithm is particularly suitable for real-time implementation with a computational complexity of $\mathcal{O}(T)$.

Furthermore, it is necessary to distinguish the Lyapunov optimization method proposed in this article from other optimization methods. A comparison with DP techniques is made here. DP can be used to solve stronger versions of our problem (such as maximizing trade profits under delay constraints) see e.g. [25]. Typically, DP requires more stringent system modeling assumptions to program the objective function and has a complex solution that requires knowledge of the renewable generation, energy demand, and real-time electricity price probabilities. However, in reality, the statistics of renewable generation, energy demand, and electricity price may be unknown or difficult to predict accurately. Moreover, DP requires dealing with the computation of a value function which is difficult to solve when the state space of the system is large, and encounters a curse of dimensionality when applied to large dimensional systems (e.g., systems with many queues). In contrast, the Lyapunov optimization technique does not require any prior knowledge and is relatively simple to implement. It can also be easily applied to the extended formula (the drift-plus-penalty function) with multiple queues without increasing the complexity of the scheme, and thus does not suffer from the curse of dimensionality. Finally, we will illustrate below a demonstrable performance guarantee for the optimal solution of the algorithm.

D. ALGORITHM PERFORMANCE ANALYSIS

In this subsection, we set the value range of λ as being fixed in $0 \leq \lambda \leq E\{a(t)\}$. Below, we summarize the performance of the dynamic energy trading and load scheduling algorithm in the form of a theorem.

Theorem 1: Assume that at any time slot t, $\min(Q(t), d_{\max}) \geq \max[a_{\max}, \lambda]$ is satisfied. $Q(0) = 0, Z(0) = 0$ and constant V is fixed as $0 < V < V_{\max}$, where

$$V_{\max} = \frac{B_{\max} - a_{\max} - \lambda - 2*b_{\max}}{p_{\max} - p_{\min}} \tag{22}$$

where $p_{\min}, p_{\max}$ are the minimum and maximum electricity prices in the smart grid, respectively. The proposed algorithm has the following properties:

(1). The maximum energy backlogs of queues $Q(t)$ and $Z(t)$ are

$$Q_{\max} = Vp_{\max} + a_{\max}$$
$$Z_{\max} = Vp_{\max} + \lambda$$

Algorithm 1 Dynamic Energy Trading and Load Scheduling Algorithm

For any time slot t **do**
Measure the system status $(Q(t), Z(t), B(t), X(t))$, $p(t)$ and $a(t)$.
if $X(t) + Q(t) + Z(t) \geq 0$ **then**
  if $X(t) + V\beta p(t) \geq 0$ **then**
    $b(t) = b_{\max}$.
    if $Q(t) + Z(t) > Vp(t)$ **then**
      $d(t) = \min(Q(t), d_{\max})$.
    else if $V\beta p(t) \leq Q(t) + Z(t) \leq Vp(t)$ **then**
      $d(t) = \min(Q(t), e(t))$.
    else
      $d(t) = 0$.
    end if
  else if $X(t) + Vp(t) \leq 0$ **then**
    $b(t) = -\min(b_{\max}, B_{\max} - B(t))$.
    $d(t) = \min(Q(t), d_{\max})$.
  else
    if $Q(t) + Z(t) \geq Vp(t)$ **then**
      $d(t) = \min(Q(t), d_{\max})$.
      $b(t) = \min(b_{\max}, \min(B_{\max} - B(t), d(t) - e(t)))$.
    else
      $d(t) = \min(Q(t), e(t))$, $b(t) = 0$.
    end if
  end if
else
  if $X(t) + Vp(t) < 0$ **then**
    $b(t) = -\min(b_{\max}, B_{\max} - B(t))$.
    if $Q(t) + Z(t) \geq Vp(t)$ **then**
      $d(t) = \min(Q(t), d_{\max})$.
    else if $V\beta p(t) \leq Q(t) + Z(t) \leq Vp(t)$ **then**
      $d(t) = \min(Q(t), e(t))$.
    else
      $d(t) = 0$.
    end if
  else if $X(t) + V\beta p(t) \geq 0$ **then**
    $b(t) = b_{\max}$, $d(t) = 0$.
  else
    if $Q(t) + Z(t) < \beta Vp(t)$ **then**
      $d(t) = 0$, $b(t) = 0$.
    else
      $d(t) = \min(Q(t), e(t))$, $b(t) = 0$.
    end if
  end if
end if
Compute $s(t)$ using Eqs. (2) and update $Q(t), Z(t), X(t)$ according to Eqs. (1), (13), (16), respectively.

and the upper bound of $Q(t) + Z(t)$ is given by

$$[Q(t) + Z(t)]_{\max} = \Phi_{\max} = Vp_{\max} + a_{\max} + \lambda$$

(2). The maximum delay time for the energy request at time slot t is

$$T_{\max} = [\frac{2Vp_{\max} + a_{\max} + \lambda}{\lambda}]$$

(3). The battery level $B(t)$ is bounded at any time slot t and satisfies

$$0 \leq B(t) \leq \Phi_{\max} + 2b_{\max}, \quad \forall t.$$

(4). The upper bound of $X(t)$ at any time slot t is

$$-\Phi_{\max} - b_{\max} \leq X(t) \leq b_{\max}, \quad \forall t.$$

(5). The time-average expected profit under our algorithm is within bound $\frac{S}{V}$ of the optimal value, i.e.,

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left[\beta p(t)(d(t) - b(t) - e(t))^{-} - p(t)(d(t) - b(t) - e(t))^{+}\right] \geq C^{*} - \frac{S}{V} \quad (23)$$

Proof: See Appendix C

As mentioned above, we provide some constraints on the design of the algorithm and provide the theoretical proves for them. Note that, the property (5) provides the performance gap between the long-term time-averaged profits of our proposed algorithm and the optimal solution which has full the future values or statistical knowledge. From Eq. (23), it can be seen that our algorithm becomes asymptotically equivalent to the optimal solution as $V \to \infty$ (i.e., $T_{\max} \to \infty$ and $B_{\max} \to \infty$). The following section includes some simulation results to verify the validity of our algorithm.

VI. SIMULATION RESULTS

From the previous analysis, it can be seen that the proposed algorithm is not affected by the probability distributions of RESs, loads, and electricity price. To demonstrate the simulation results, we assume that they follow specific distributions. The time-varying electricity price in the smart grid fluctuates between 0.07 \$/kWh and 0.29 \$/kWh with two peaks in a day. The interval of the time is 1 min, which is a reasonable response time for the smart grid to balance the power generation and load. Simulation results are conducted with different values of parameters V, $B_{\max}$, $e(t)$ and λ. This work considers the case of 30 days per month (43200 time slots). Without loss of generality, the daily trend of electricity prices is a repetition of the first day electricity price model. Considering that the average monthly power consumption of an ordinary household of users is 360 kWh, and referring to the capacity range of rechargeable batteries available in the market. We fix the parameters $b_{\max} = 3.8$ kWh, $d_{\max} = 2.5$ kWh, and $\beta = 0.8$. Parameter settings are added in Table 2 below.

TABLE 2. Simulation parameter setting.

Parameter	Setting
Range of electricity price $p(t)$	0.07∼0.29(\$/kWh)
Battery maximum charging rate $b_{\max}$	3.8 kWh
Maximum value of energy demand $d_{\max}$	2.5 kWh
Mean value of energy arrival $a(t)$	0.5 kWh
β	0.8
Total number of timeslots T	43200

In order to better evaluate the performance of the proposed algorithm (ETLSA for short), we compare the performance with existing solutions that consider energy trading or demand scheduling, such as Zhou *et al.* [19] and Qiao *et al.* [20]. The algorithms in [19], [20] both achieve near-optimal performance in experiments if the future statistical characteristics (energy arrivals, electricity prices, and supply of renewable energy) are known. In such a scenario, the model in [19] only allows the user to satisfy the energy demand during the service deadline without storing surplus RESs for future use or discharging power to the grid to make a profit. Thus, the algorithm in [19] does not consider the control decisions of the energy storage device. In contrast, the authors in [20] consider charging surplus RESs into the energy storage device and discharging it for use during the high electricity prices. But the electricity purchased from the grid can only be used to serve the load demand of the end-user and cannot be charged into the energy storage device, which reduces the profitability of the end-user and the efficiency in the usage of the energy storage device. However, in our proposed algorithm, ETLSA, we fully exploit the flexibility on the demand side and make real-time decisions (the demand scheduling decision $d(t)$, the charging and discharging decision $b(t)$, and the trading decision $s(t)$) in response to different system states — the real-time electricity price $p(t)$, the charge level of the battery $B(t)$ and the energy demand queue backlog $Q(t)$, as discussed in Section V-C. Next, we provide simulation results to validate our theoretical analysis.

Before the comparison, we fix $B_{\max} = 35$ kWh, and set $V = V_{\max} \approx 1100$ subject to the calculation form (22). The 30-day cumulative profit comparison of the end-user based on the three algorithms under the three different cases of average amount of energy harvested, where the three cases correspond to e_{av} being 0.85 kWh, 0.55 kWh, and 0.25 kWh respectively, as shown in Fig. 2. Due to more renewable energy is available in case 1, which results in higher profits than the other two cases. Moreover, detailed simulation results are statistically presented in Table 3, we can see that no matter under which case, our proposed algorithm can achieve the best performance among the three strategies.The reason is that the battery not only can store excess RESs but can also purchase low prices of electricity from the grid to be used when electricity prices are high.

In order to evaluate the impact of battery capacity on the proposed algorithm, Fig. 3 shows the user's profits with battery capacities $B_{\max} = \{0, 15, 35, 55, 75, 95\}$ kWh

TABLE 3. Comparison of the total profit in different scenarios.

The total profits ($) / e_{av} (kWh) / Algorithms	0.85	0.55	0.25
The algorithm in [19]	47.73	6.782	-32.42
The algorithm in [20]	85.9	60.11	25.82
ETLSA	127.1	99.54	62.84

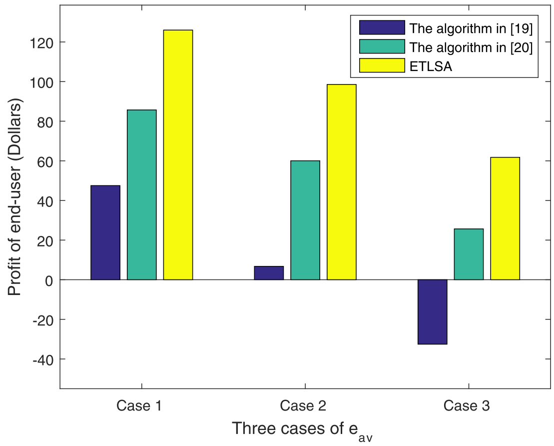

FIGURE 2. Histograms of the profits of the three algorithms in different cases.

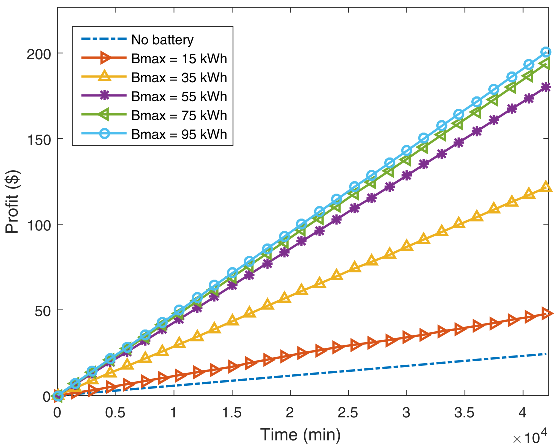

FIGURE 3. Impact of different battery capacities on the profit.

respectively. Form the Fig. 3, the larger the battery capacity, the higher is the end-user's profit. The profiting comes from two aspects: one is by storing more renewable energy generated for use at later time when the electricity price is high; the other is by charging the battery from the external grid when the electricity price is low while discharging it when the electricity price is high. However, the end-user's profit does not increase linearly with the increase in battery capacity. This is because the maximum battery charging rate $b_{\max}$ is limited. In addition, the larger the capacity of the battery, the higher is the cost of the battery. Thus, the end-user should thus choose a compromise based on the cost of the battery and the profits it entails.

Next, The total profit (at the end of 30 days) of the end-user and the state of charge level $B(t)$ in the battery are analyzed

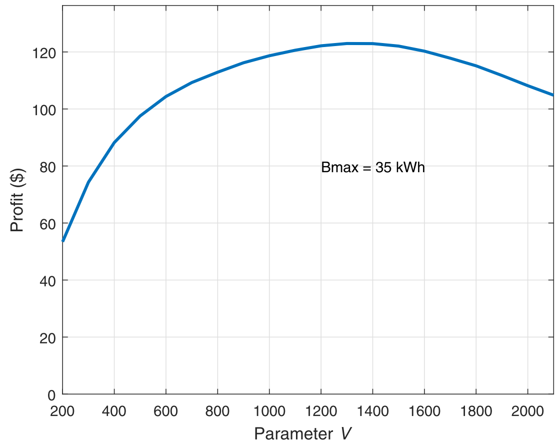

FIGURE 4. Monthly profit versus the value of *V*.

by comparing different values of V (fix $B_{\max} = 35$ kWh). This can be observed in Fig. 4, the total profit increases non-linearly with the value of V within a certain range. The reason is that the total profit will approach the optimal solution infinitely as the value of V increases from inequality (23). Besides, the profit will decrease when the value of V exceeds a certain value ($V \approx 1400$). However, according to equation (22), we can calculate $V_{\max} \approx 1100$ instead of 1400. The reason can be found in Fig. 5. The real-time energy of the battery for different values of V (10th day) as shown in Fig. 5. From the Fig. 5, the overall level of energy in the battery is low when $V = 400$, which leaves most of the capacity space of the battery vacant, so that the most of the battery's capacity is not fully utilized which leads to a lower total profit. In contrast, the total profit is highest (see Fig. 4) and the battery reaches full state in some time slots when $V = 1400$. Further, when the value of V is too large ($V = 2000$), it leads to a reduction in the number of battery discharge, and the amount of charging the battery is constrained by the maximum capacity of the battery, thus decreasing the total profit. Therefore, here we consider preventing the battery full load, so equation (22) is relatively conservative. Our proposed algorithm chooses $V = 1100$ when $B_{max} = 35$ kWh, which is a reasonable value for the battery to balance the power storage.

In order to easily observe the delay caused by the proposed algorithm in the case of energy demands, Fig. 6 shows a distribution map of the waiting time delay of the energy demand based on the proposed algorithm and a naive scheme (Purchase-at-deadline algorithm) for a duration of 30 days as well as a comparison chart of end-user's profits. The Purchase-at-deadline algorithm means that the end-user only draw energy from the renewable sources within given deadline, and does not draw energy from the power grid even if the energy harvested cannot meet the demand until any demand delay exceed the deadline. The maximum delay time is set at 25 time slots (i.e., 25 minutes) to ensure end-user comfort. The end-user revenue using the proposed algorithm is clearly superior to that of the "Purchase-at-deadline" algorithm as

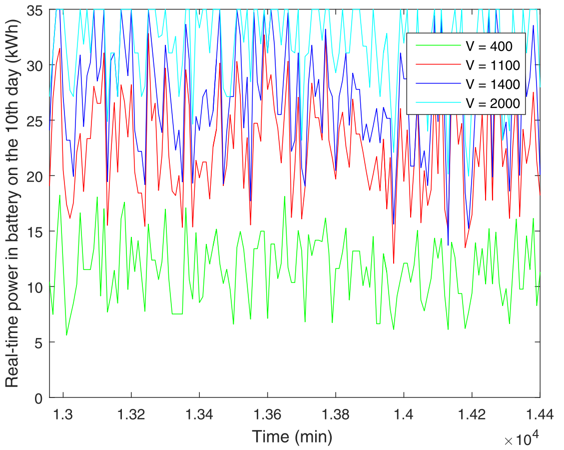

FIGURE 5. The charge level in the battery changes for different values of V.

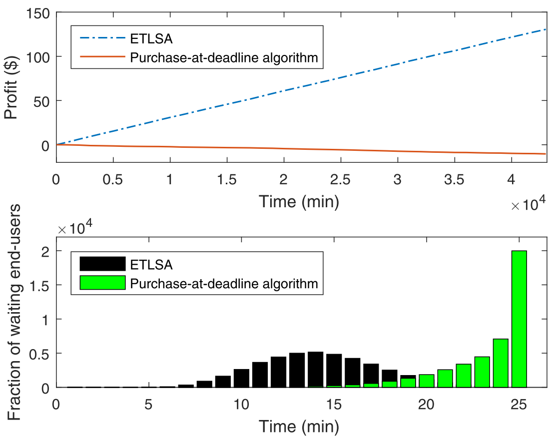

FIGURE 6. Comparison of fraction of end-user waiting time in the energy demand queue and total profit in different scenarios (maximum delay time is 25 min).

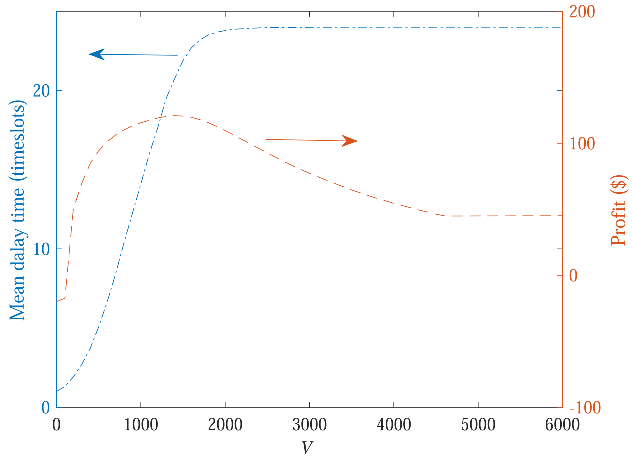

FIGURE 7. Total profit and mean delay time vs. the value of V.

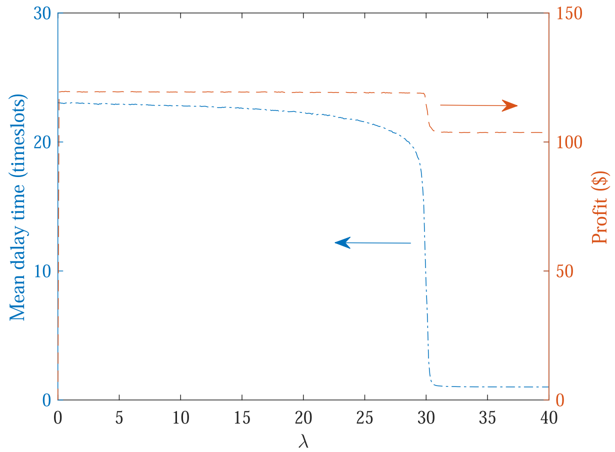

FIGURE 8. Total profit and mean delay time vs. the value of λ.

shown in Fig. 6. We can obtain the mean delay time of the energy demand for the user using the proposed algorithm as 14.23 time slots, while that of the "Purchase-at-deadline" algorithm as 22.85 time slots. Therefore, our proposed algorithm can reduce the mean delay time for energy demands by 37% on average.

As mentioned before, the mean delay time is used as an important indicator to ensure the comfort of the end-user. In our proposed algorithm, V as a controllable parameter is used to maximize the tradeoff between the total profit and comfort. We have plotted figures showing the relationship between the total profit and the value of V and the relationship between the mean delay time and the value of V (Fig. 7). From Fig. 7, it can be seen that the mean delay time increases as V increases, while the total profit increases and then decreases with the value of V. Moreover, the total profit and mean delay time finally reach saturation when the value of V is larger than a certain value. From the result we can see that the trade-off between total profit and comfort is maximized when $V \approx 1100$. The reason is that $V \approx 1100$ is a reasonable range that achieves a low mean delay time to satisfy the end-user's comfort while ensuring that the total profit is close to the optimal value obtained by our proposed algorithm.

Similarly, we can observe that the mean delay time decreases non-linearly with the value of λ, while the total profit increases to a certain value and then decreases with the value of λ (Fig. 8). When the value of λ is greater than a certain value, the total profit and the mean delay time are eventually saturated. It is not difficult to see that the appropriate value of λ should be distributed within the range [29, 30].

VII. CONCLUSION

This work focuses on the problem of energy trading and load scheduling for an end-user equipped with renewable generators and a battery in the smart grid. We assume that the dynamics of RESs, loads, and electricity prices are arbitrary or non-stationary and their statistics are unknown to fit the actual situation. The aim is to design an real-time solution for the end-user to maximize profits over a finite time period while satisfying the FLs delay constraint. As a result, a low-complexity dynamic energy trading and load scheduling algorithm (ETLSA) is proposed based on the Lyapunov optimization. The algorithm has a bounded performance guarantee from an optimal solution which has complete future information and is asymptotically equivalent to the optimal solution as the battery capacity or the delay time of the FL tends to infinity. Simulation results have demonstrated the proposed algorithm is superior to other algorithms. Furthermore, we also analyze the impact of battery capacity on the performance of the algorithm to provide a theoretical basis for the end-user to determine battery capacity, and maximize the tradeoff between total profit and comfort for the end-user.

APPENDIX A
PROOF OF LEMMA 2

Proof: We use proof by contradiction to prove **Lemma 2**. Consider that the energy demand $a(t) > 0$ is satisfied at any time slot t and requires that the service be completed at or before time slot $t + T_{\max}$. Suppose not, then it must satisfy $Q(\tau) > 0$ during the slot $\tau \in \{t+1, \ldots, t+T_{\max}\}$, otherwise the energy demand $a(t)$ will complete the service before the time slot $t + T_{\max}$. Therefore, according to the updated equation (13), at all time slots $\tau \in \{t+1, \ldots, t+T_{\max}\}$, and we get

$$Z(t+1) \geq Z(t) - d(t) + \lambda$$

Taking the sum at interval $\tau \in \{t+1, \ldots, t+T_{\max}\}$ for the above formula, we obtain

$$Z(t+T_{\max}+1) - Z(t+1) \geq -\sum_{\tau=t+1}^{t+T_{\max}} d(\tau) + T_{\max}\lambda$$

Since, $Z(t+T_{\max}+1) < Z_{\max}, Z(t+1) > 0$, so rearranging the terms yields

$$\sum_{\tau=t+1}^{t+T_{\max}} d(\tau) \geq T_{\max}\lambda - Z_{\max}$$

If the energy demand $a(t)$ completes the service at or before time slot $t + T_{\max}$, then at least $Q_{\max}$ are served to the user during time slot $\tau \in \{t+1, \ldots, t+T_{\max}\}$, i.e., $\sum_{\tau=t+1}^{t+T_{\max}} d(\tau) \geq Q_{\max}$. However, we assume that at time slot $t + T_{\max}$, the energy demand has not been serviced, then there must be $\sum_{\tau=t+1}^{t+T_{\max}} d(\tau) < Q_{\max}$, which must be accounted in the above formula, and the shifted term is given as

$$T_{\max} < (Q_{\max} + Z_{\max})/\lambda$$

which contradicts (14), therefore, **Lemma 2** holds.

APPENDIX B
PROOF OF LEMMA 3

Proof: Since $\Delta\left(\vec{M}(t)\right) = E\left\{L(\vec{M}(t+1)) – L(\vec{M}(t))|\vec{M}(t)\right\}$, and according to (17), calculate from the real queue $Q(t)$ first. According to (1)

$$\begin{aligned}Q^2(t+1) &= [\max[Q(t) - d(t), 0] + a(t)]^2 \\ &\leq Q^2(t) + a_{\max} + d_{\max} + 2Q_{\max}a_{\max} - 2Q(t)d(t)\end{aligned}$$

then we can get

$$\frac{1}{2}\left[Q^2(t+1) - Q^2(t)\right] \leq \frac{[d^2{}_{\max} + a^2{}_{\max}]}{2} + Q_{\max}a_{\max} - Q(t)d(t)$$

Second, from the update

$$\begin{aligned}&\frac{1}{2}\left[X^2(t+1) - X^2(t)\right] \\ &= \frac{1}{2}[X(t) - b(t)]^2 - \frac{1}{2}X^2(t) \\ &= \frac{b^2(t)}{2} - X(t)b(t) \leq \frac{1}{2}b_{\max} - X(t)b(t)\end{aligned}$$

Finally, for the virtual queue $Z(t)$, we have

$$\begin{aligned}Z^2(t+1) &\leq [Z(t) - d(t) + \lambda]^2 \\ &\leq Z^2(t) + \max(\lambda^2, d_{\max}) + 2Z_{\max}\lambda - 2Z(t)d(t)\end{aligned}$$

and therefore

$$\frac{1}{2}\left[Z^2(t+1) - Z^2(t)\right] \leq \frac{\max[\lambda^2{}_{\max}, d^2{}_{\max}]}{2} + Z_{\max}\lambda - Z(t)d(t)$$

Summing the drift maximum boundary values of the three queues, we can get that the sum of all the constant terms is S, and the Lyapunov drift-plus-penalty function can be transformed into

$$\begin{aligned}&\Delta\left(\vec{M}(t)\right) - V \cdot E\left\{\beta p(t)(d(t) - e(t) - b(t))^- \right. \\ &\quad \left. - p(t)(d(t) - e(t) - b(t))^+|\vec{M}(t)\right\} \\ &\leq S - V \cdot E\left\{\beta p(t)(d(t) - e(t) - b(t))^- \right. \\ &\quad \left. - p(t)(d(t) - e(t) - b(t))^+|\vec{M}(t)\right\} \\ &\quad + X(t)E\{b(t)|\vec{M}(t)\} - [Q(t) + Z(t)]\,E\left\{d(t)|\vec{M}(t)\right\} \end{aligned} \tag{24}$$

Due to the fact that $d(t)$ and $b(t)$ are independent of queue state $\vec{M}(t)$, so the last two terms in the right-hand side of (24) can be further simplified according to (2) as

$$\begin{aligned}&-X(t)b(t) - [Q(t) + Z(t)]d(t) \\ &= -X(t)b(t) - [Q(t) + Z(t)][e(t) + b(t) + s(t)] \\ &= -[X(t) + Q(t) + Z(t)]b(t) - [Q(t) + Z(t)]e(t) \\ &\quad - [Q(t) + Z(t)]s(t)\end{aligned} \tag{25}$$

Because

$$\begin{aligned}s(t) &= d(t) - b(t) - e(t) \\ &= (d(t) - e(t) - b(t))^+ - (d(t) - e(t) - b(t))^- \\ &\quad - b(t) = [d(t) - b(t) - e(t) - (d(t) - e(t))] \\ &= (d(t) - e(t) - b(t))^+ - (d(t) - e(t) - b(t))^- \\ &\quad - (d(t) - e(t))\end{aligned}$$

where $f = f^+ - f^-$ is used. Taking the above two equalities into (22), we can get

$$\begin{aligned} &-X(t)b(t) - [Q(t) + Z(t)]d(t) \\ &= X(t)\left[(d(t) - e(t) - b(t))^+ - (d(t) - e(t) - b(t))^-\right] \\ &\quad - [X(t) + Q(t) + Z(t)]d(t) + X(t)e(t) \end{aligned} \tag{26}$$

Inserting (26) into (24), the Lyapunov drift-plus-penalty function can be obtained as

$$\begin{aligned} &\Delta\left(\vec{M}(t)\right) - V \cdot E\left\{\beta p(t)(d(t) - e(t) - b(t))^-\right. \\ &\quad \left.- p(t)(d(t) - e(t) - b(t))^+|\vec{M}(t)\right\} \\ &\le S + [X(t) + Vp(t)]E\{(d(t) - e(t) - b(t))^+|\vec{M}(t)\} \\ &\quad - [X(t) + V\beta p(t)]E\{(d(t) - e(t) - b(t))^-|\vec{M}(t)\} \\ &\quad - [X(t) + Q(t) + Z(t)]E\{d(t)|\vec{M}(t)\} \\ &\quad + X(t)E\{e(t)|\vec{M}(t)\} \end{aligned}$$

Similar to (19), hence, **Lemma 3** has been proven.

APPENDIX C
PROOF OF THEOREM 1

Proof: (1). First, we need to prove $Q(t) \le Q_{max}$ for every time slot t by using the induction method.

Obviously, $Q(0) < Q_{max}$, now, suppose $Q(t) \le Q_{max}$ holds at time slot t, we only need to prove that $Q(t+1) \le Q_{max}$ also holds at time slot $t+1$. From the updated equation (1), when $Q(t) \le Vp_{max}$, at ime slot $t+1$, the maximum energy demand is a_{max}, then $Q(t+1) \le Vp_{max} + a_{max}$. In contrast, when $Vp_{max} < Q(t) \le Vp_{max} + a_{max}$, our proposed will choose $d(t) = \min(Q(t), d_{max})$. Because of the previous assumption $\min(Q(t), d_{max}) \ge \max[a_{max}, \lambda]$, the energy service rate of the real queue $Q(t)$ is greater than the maximum energy demand arrival rate during time slot t. Then, $Q(t+1) < Q(t) \le Vp_{max} + a_{max}$ holds at the time slot $t+1$. Therefore, we have proved $Q(t) \le Vp_{max} + a_{max}$.

Second, we prove $Z(t) \le Z_{max}$ for every time slot t.

Obviously, $Z(0) < Z_{max}$. now suppose that $Z(t) \le Z_{max}$ holds at time slot t, we only need to prove that $Z(t+1) \le Z_{max}$ also holds at time slot $t+1$. From the updated equation (14), when $Z(t) \le Vp_{max}$, then the maximum amount of penalty arrival is λ at time slot $t+1$. Hence, $Z(t+1) \le Vp_{max} + \lambda$ is satisfied during time slot $t+1$. In contrast, when $Vp_{max} < Z(t) \le Vp_{max} + \lambda$, our proposed algorithm will choose $d(t) = \min(Q(t), d_{max})$. Because of the previous assumption $\min(Q(t), d_{max}) > \max[a_{max}, \lambda]$, the energy service rate of the virtual queue $Z(t)$ is greater than the maximum amount of penalty arrival rate at time slot t. Then, $Z(t+1) < Z(t) \le Vp_{max} + \lambda$ holds at the time slot $t+1$. Thus, we have proved $Z(t) \le Vp_{max} + \lambda$.

Finally, we prove $Q(t) + Z(t) \le \Phi_{max}$ for every time slot t.

Obviously, $Q(t) + Z(t) \le \Phi_{max}$. Now, suppose that $Q(t) + Z(t) \le \Phi_{max}$ holds at time slot t, we only need to prove that $Q(t+1) + Z(t+1) \le \Phi_{max}$ also holds at time slot $t+1$. From the updated equation (1) and (14), when $Q(t) + Z(t) \le Vp_{max}$, the maximum increase is $a_{max} + \lambda$ during the time slot t, so we have $Q(t+1) + Z(t+1) \le Vp_{max} + a_{max} + \lambda$. When $Vp_{max} < Q(t) + Z(t) \le Vp_{max} + a_{max} + \lambda$, our proposed algorithm will choose $d(t) = \min(Q(t), d_{max})$. According to the previous analysis, $Q(t+1)$ and $Z(t+1)$ will not increase in the $t+1$ slot, and $Q(t+1) + Z(t+1) < Vp_{max} + a_{max} + \lambda$ holds at the time slot $t+1$. Therefore, $Q(t) + Z(t) \le Vp_{max} + a_{max} + \lambda$ holds. In summary, we have proved property (1).

(2). By substituting Q_{max} and Z_{max} in property (1) into (14), we can verify that property (2) holds.

(3). The following are used for prove by induction. According to formula (20), $X(0) = B(0) - \Theta_{max} - b_{max}$. Since $X(0) = 0$, so $B(0) = \Phi_{max} + b_{max}$. Obviously, we can get $B(0) < \Phi_{max} + 2b_{max}$, suppose $B(t) \le \Phi_{max} + 2b_{max}$ holds at time slot t, then we only need to prove the fact that $B(t+1) \le \Phi_{max} + 2b_{max}$ holds at time slot $t+1$. When $B(t) \le \Phi_{max} + b_{max}$, the maximum increase charge in the battery is b_{max} during $t+1$ time slot, that is, $B(t+1) \le \Phi_{max} + 2b_{max}$. When $\Phi_{max} + b_{max} < B(t) \le \Phi_{max} + 2b_{max}$, our proposed algorithm will choose $b(t) = b_{max}$. This means that the battery is discharging at the maximum rate, so $B(t+1) < B(t) \le \Phi_{max} + 2b_{max}$ during time slot $t+1$. Therefore, $B(t) \le \Phi_{max} + b_{max}$ holds for every time slot t, we have proved property (3).

(4). According to (15) and property (3), it can be verified that property (4) holds.

(5). From the previous analysis, the goal of the our proposed algorithm is to greedily minimize a upper bound of the drift-plus-penalty function in (19). Comparing the policy in **Lemma 1**, we know that the dynamic energy trading and management algorithm is too greedy to maximize the objective function (9). Therefore, we obtain the following formula.

$$\begin{aligned} &S + [X(t) + Vp(t)]E\{(d(t) - e(t) - b(t))^+|\vec{M}(t)\} \\ &\quad - [X(t) + V\beta p(t)]E\{(d(t) - e(t) - b(t))^-|\vec{M}(t)\} \\ &\quad - [X(t) + Q(t) + Z(t)]E\{d(t)|\vec{M}(t)\} \\ &\quad + X(t)E\{e(t)|\vec{M}(t)\} \\ &\le S + [X(t) + Vp(t)]E\{(\bar{d}(t) - e(t) - \bar{b}(t))^+|\vec{M}(t)\} \\ &\quad - [X(t) + V\beta p(t)]E\{\left(\bar{d}(t) - e(t) - \bar{b}(t)\right)^-|\vec{M}(t)\} \\ &\quad - [X(t) + Q(t) + Z(t)]E\{\bar{d}(t)|\vec{M}(t)\} \\ &\quad + X(t)E\{e(t)|\vec{M}(t)\} \end{aligned}$$

Thus, according to the proof process of **Lemma 3**,we have

$$\begin{aligned} &\Delta\left(\vec{M}(t)\right) - V \cdot E\left\{\beta p(t)(d(t) - e(t) - b(t))^-\right. \\ &\quad \left.- p(t)(d(t) - e(t) - b(t))^+|\vec{M}(t)\right\} \\ &\le S - V \cdot E\left\{\beta p(t)\left(\bar{d}(t) - e(t) - \bar{b}(t)\right)^-\right. \\ &\quad \left.- p(t)\left(\bar{d}(t) - e(t) - \bar{b}(t)\right)^+|\vec{M}(t)\right\} \\ &\quad + X(t)E\{\bar{b}(t)|\vec{M}(t)\} - [Q(t) + Z(t)]E\{\bar{d}(t)|\vec{M}(t)\} \end{aligned} \tag{27}$$

Since $\bar{d}(t)$, $\bar{b}(t)$, $a(t)$ and $e(t)$ are independent of $\vec{M}(t)$, and based on the fact that $E[E[X|Y]] = E[X], \forall X, Y$, by taking expectations on both side of (27) and summation from $t = 0 \to T$, we obtain

$$\begin{aligned} &E[\vec{M}(T+1) - \vec{M}(0)] \\ &\quad - \sum_{t=0}^{T} VE\left\{\beta p(t)(d(t)-e(t)-b(t))^{-}\right. \\ &\quad \left. - p(t)(d(t)-e(t)-b(t))^{+}\right\} \\ &\leq ST - \sum_{t=0}^{T} VE\left\{\beta p(t)\big(\bar{d}(t)-e(t)-\bar{b}(t)\big)^{-}\right. \\ &\quad \left. - p(t)\big(\bar{d}(t)-e(t)-\bar{b}(t)\big)^{+}\right\} - \sum_{t=0}^{T} X(t)E\{\bar{b}(t)\} \\ &\quad - \sum_{t=0}^{T} [Q(t)+Z(t)]\, E\left\{\bar{d}(t)\right\} \end{aligned} \tag{28}$$

Because $\vec{M}(0) = 0$, dividing by VT and taking the $T \to \infty$ on the both sides of (28), we get

$$\begin{aligned} &-\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E\left\{\beta p(t)(d(t)-e(t)-b(t))^{-}\right. \\ &\qquad \left. - \; p(t)(d(t)-e(t)-b(t))^{+}\right\} \geq \frac{S}{V} - \bar{C} \end{aligned} \tag{29}$$

where the conditions have been used

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} X(t)E\{\bar{b}(t)\} = 0$$

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} [Q(t)+Z(t)]\, E\left\{\bar{d}(t)\right\} = 0$$

Finally, rearranging the (29) and using $\bar{C} \geq C_1^*$, we obtain

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left[\beta p(t)(d(t)-b(t)-e(t))^{-}\right. \\ &\left. - p(t)(d(t)-b(t)-e(t))^{+}\right] \geq \bar{C} - \frac{S}{V} \geq C_1^* - \frac{S}{V} \end{aligned} \tag{30}$$

Therefore, we have proved property (5) in Theorem 1. In summary, the performance analysis has been fully proven.

REFERENCES

[1] F. Y. Xu and L. L. Lai, "Novel active time-based demand response for industrial consumers in smart grid," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1564–1573, Dec. 2015.
[2] J. Wang, H. Zhong, Z. Ma, Q. Xia, and C. Kang, "Review and prospect of integrated demand response in the multi-energy system," *Appl. Energy*, vol. 202, pp. 772–782, Sep. 2017.
[3] A. A. Hadi, C. A. S. Silva, E. Hossain, and R. Challoo, "Algorithm for demand response to maximize the penetration of renewable energy," *IEEE Access*, vol. 8, pp. 55279–55288, 2020.
[4] E. Saeidpour Parizy, S. Choi, and H. R. Bahrami, "Grid-specific co-optimization of incentive for generation planning in power systems with renewable energy sources," *IEEE Trans. Sustain. Energy*, vol. 11, no. 2, pp. 947–957, Apr. 2020.
[5] A. T. Eseye, J. Zhang, D. Zheng, H. Li, and G. Jingfu, "A double-stage hierarchical hybrid PSO-ANN model for short-term wind power prediction," in *Proc. IEEE 2nd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Chengdu, China, Apr. 2017, pp. 499–503.
[6] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and meteorological information," *Renew. Energy*, vol. 118, pp. 357–367, Apr. 2018.
[7] H.-I. Su and A. E. Gamal, "Modeling and analysis of the role of energy storage for renewable integration: Power balancing," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4109–4117, Nov. 2013.
[8] H. Khani and H. E. Z. Farag, "Optimal day-ahead scheduling of power-to-gas energy storage and gas load management in wholesale electricity and gas markets," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 940–951, Apr. 2018.
[9] A. T. Eseye, D. Zheng, H. Li, and J. Zhang, "Grid-price dependent optimal energy storage management strategy for grid-connected industrial microgrids," in *Proc. 9th Annu. IEEE Green Technol. Conf. (GreenTech)*, Denver, CO, USA, Mar. 2017, pp. 124–131.
[10] S. Wang, S. Bi, and Y.-J.-A. Zhang, "Demand response management for profit maximizing energy loads in real-time electricity market," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6387–6396, Nov. 2018.
[11] M. A. Hannan, M. Faisal, P. J. Ker, L. H. Mun, K. Parvin, T. M. I. Mahlia, and F. Blaabjerg, "A review of Internet of energy based building energy management systems: Issues and recommendations," *IEEE Access*, vol. 6, pp. 38997–39014, 2018.
[12] Y. Huang, S. Mao, and R. M. Nelms, "Adaptive electricity scheduling in microgrids," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 270–281, Jan. 2014.
[13] S. Chen, N. B. Shroff, and P. Sinha, "Heterogeneous delay tolerant task scheduling and energy management in the smart grid with renewable energy," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1258–1267, Jul. 2013.
[14] C. Joe-Wong, S. Sen, S. Ha, and M. Chiang, "Optimized day-ahead pricing for smart grids with device-specific scheduling flexibility," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 6, pp. 1075–1085, Jul. 2012.
[15] K. Ma, T. Yao, J. Yang, and X. Guan, "Residential power scheduling for demand response in smart grid," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 320–325, Jun. 2016.
[16] S. Bera, P. Gupta, and S. Misra, "D2S: Dynamic demand scheduling in smart grid using optimal portfolio selection strategy," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1434–1442, May 2015.
[17] D. T. Nguyen and L. B. Le, "Joint optimization of electric vehicle and home energy scheduling considering user comfort preference," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 188–199, Jan. 2014.
[18] T. Li and M. Dong, "Real-time residential-side joint energy storage management and load scheduling with renewable integration," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 283–298, Jan. 2018.
[19] Y. Zhou, D. K. Y. Yau, P. You, and P. Cheng, "Optimal-cost scheduling of electrical vehicle charging under uncertainty," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4547–4554, Sep. 2018.
[20] Z. Qiao, B. Yang, Q. Xu, F. Xiong, C. Chen, X. Guan, and B. Chen, "Energy trading between microgrids towards individual cost and social welfare optimization," in *Proc. 13th IEEE Conf. Autom. Sci. Eng. (CASE)*, Xi'an, China, Aug. 2017, pp. 1243–1248.
[21] S. Misra, S. Bera, T. Ojha, H. T. Mouftah, and A. Anpalagan, "ENTRUST: Energy trading under uncertainty in smart grid systems," *Comput. Netw.*, vol. 110, pp. 232–242, Dec. 2016.
[22] C.-C. Lin, D.-J. Deng, C.-C. Kuo, and Y.-L. Liang, "Optimal charging control of energy storage and electric vehicle of an individual in the Internet of energy with energy trading," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2570–2578, Jun. 2018.
[23] A. T. Eseye, M. Lehtonen, T. Tukia, S. Uimonen, and R. J. Millar, "Optimal energy trading for renewable energy integrated building microgrids containing electric vehicles and energy storage batteries," *IEEE Access*, vol. 7, pp. 106092–106101, 2019.
[24] C. O. Adika and L. Wang, "Autonomous appliance scheduling for household energy management," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 673–682, Mar. 2014.
[25] P. Samadi, V. W. S. Wong, and R. Schober, "Load scheduling and power trading in systems with high penetration of renewable energy resources," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1802–1812, Jul. 2016.
[26] W. Liu, D. Qi, and F. Wen, "Intraday residential demand response scheme based on Peer-to-Peer energy trading," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1823–1835, Mar. 2020.
[27] H. Wang, T. Huang, X. Liao, H. Abu-Rub, and G. Chen, "Reinforcement learning for constrained energy trading games with incomplete information," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3404–3416, Oct. 2017.

[28] H. Wang, T. Huang, X. Liao, H. Abu-Rub, and G. Chen, "Reinforcement learning in energy trading game among smart microgrids," *IEEE Trans. Ind. Electron.*, vol. 63, no. 8, pp. 5109–5119, Aug. 2016.
[29] H. Yu and M. J. Neely, "A new backpressure algorithm for joint rate control and routing with vanishing utility optimality gaps and finite queue lengths," *IEEE/ACM Trans. Netw.*, vol. 26, no. 4, pp. 1605–1618, Aug. 2018.

DIDI LIU received the M.S. degree in communication and information system from the Guilin University of Electronic Technology, China, in 2006, and the Ph.D. degree in communication and information system from Xidian University, China, in 2018. Since 2014, she has been an Associate Professor with Guangxi Normal University. Her research interests include stochastic network optimization and smart grid.

JIAWEN XIAO received the B.S. degree from the Hunan University of Science and Engineering, Yongzhou, China, in 2019. He is currently pursuing the M.S. degree with Guangxi Normal University, Guilin, China. His research interests include stochastic network optimization and smart grid.

JUNXIU LIU , photograph and biography not available at the time of publication.

XIAOMING YUAN , photograph and biography not available at the time of publication.

SUPING ZHANG , photograph and biography not available at the time of publication.