Abstract—We study the optimal operation of energy storage operated by a consumer who owns intermittent renewable generation and faces (possibly random) fluctuating electricity prices and demand charge. We formulate the optimal storage operation problem as a finite horizon dynamic program, with an objective of minimizing the expected total cost (the sum of energy cost and demand charge). The incorporation of demand charge faces the consumer with more complicated trade-offs on storage operation, e.g., discharging the storage when the energy price is relatively low may not save energy cost, but could help to save the demand charge if the net demand is high. We establish an important threshold structure for an optimal storage operation policy, which enables us to implement the optimal policy in realistic settings with random electricity prices. Numerical results demonstrate that the characterized optimal threshold policy significantly outperforms the policy that is shown to be optimal without demand charge in the literature, even at a low demand charge rate 68/kW.

Index Terms—Energy storage, Demand charge, Renewable generation, Dynamic programming, Demand response

I. INTRODUCTION

Renewable generation capacity is expanding rapidly to potentially reduce carbon dioxide emissions and dependence on fossil fuels. As non-dispatchable generation, renewable energy exaggerates the variability on the supply side, and further amplifies the difficulty of matching demand with supply in the real time. Energy storage is an environmentally friendly candidate that can provide flexibility to the system and mitigate the impact of volatile renewable generation.

We study the optimal operation of electric storages operated by electricity consumers who own distributed renewable generation and face time-varying (and possibly stochastic) energy prices as well as the demand charge. Our motivation stems from the potential of electricity consumers to own and use storage devices (e.g., major consumers like data centers [1] and individual consumers who own or rent batteries from Tesla or SolarCity [2], [3], [4]). We also note that there is a growing trend for commercial and residential consumers to own distributed renewable generation [5], [6].

Different from the (per-unit) energy price that is charged on energy consumption (in kWh), demand charge is on the highest average power consumption over all 15-minute intervals during the entire billing period (e.g., a month). Ranging from 88/kW to 218/kW, demand charge could contribute a significant portion of consumers’ monthly energy bills, and has been widely applied by many utility companies in the U.S. [7], [8], [9], for example, to consumers whose monthly energy consumption exceeds 2000 kWh [10].

In this paper, we construct a dynamic programming (DP) framework to study the challenging sequential decision making problem faced by a consumer who seeks to optimally operate her energy storage so as to minimize the expected total cost (the sum of energy cost and demand charge). It is worth noting that the incorporation of demand charge significantly complicates both the characterization and the computation of optimal storage operation policies. Intuitively, with demand charge the consumer faces more sophisticated trade-offs on storage operation; for example, discharging the storage when the energy price is relatively low may not save energy cost, but may help to save the demand charge, if the net demand (energy demand minus renewable generation) is high.

The operation of energy storage devices has received a lot of recent attention. There exists a substantial literature on the operation of energy storage owned by renewable generators or system operators. The authors of [11], [12] study the optimal operation of energy storage devices with an objective of minimizing the mismatch between the available renewable generation and system load. The joint scheduling of variable wind generation and energy storage systems has been studied to maximize the profit of wind farms and energy storage systems, through a two-stage stochastic programming formulation [13], and a model predictive control (MPC)-based approach [14], [15]. There exists another stream of works that develop and test approximate dynamic programming algorithms for storage operation in energy systems with renewable generation [16], [17], [18], [19].

Another well studied application of energy storage is the use of storages to arbitrage [20], [21]. A few recent works conduct a dynamic programming approach to derive the arbitrage value of electric storage, in the presence of dynamic pricing [22], [23], [24], [25]. Different from the setting in the present paper, this aforementioned literature assumes that the energy storage operator (e.g., an arbitrager) has zero demand for electricity and does not have its own renewable generation.

There have been recent studies on the operation of energy storage with electric loads. There is a literature that applies Lyapunov optimization based on-line algorithms to the operation of energy storage device(s) [26], [27], [28], [29]. Closer to the present paper, a few recent papers establish structural properties on optimal storage operation policies in a variety of dynamic settings that incorporate random energy prices and/or demand [30], [31], [32]. The main theorectic results of these works are the existence of an optimal policy that can be characterized by (time-varying and possibly state dependent)

1These on-line algorithms are shown to be asymptotically optimal, as the storage capacity increases to infinity. However, these on-line algorithms are suboptimal when the energy storage capacity is limited [19].
operational thresholds. All these aforementioned works [26]-[32] do not incorporate demand charge that often constitutes a significant fraction of consumers’ energy bill.

Closely related to the present paper, a few recent works explore the energy storage operation under demand charge [33], [34], [35]. These works solve one-shot deterministic optimization problems to schedule the energy storage operation (based on the forecast demand and electricity prices), whereas the present paper formulates the storage operation as a sequential decision making problem that explicitly incorporates the stochasticity in energy prices and renewable generation.

The main contributions of this paper are two-fold.

1) We formulate the storage operation problem as a finite-horizon dynamic program (DP) with an objective of minimizing the total expected cost. Each stage lasts for 15 minutes and the entire optimization horizon is one month (one billing period). To our knowledge, this work is the first that investigates the optimal operation of energy storage in the presence of both demand charge and random energy prices. It is worth noting that the incorporation of demand charge results in non-trivial technical difficulties on the characterization and computation of optimal storage operation policies, due to the additional dimension of system state (in the formulated DP) that is used to record the maximum demand so far.

For a general setting that incorporates the stochasticity in both energy prices and renewable generation, we establish a complete characterization on an optimal threshold policy (cf. Theorem 1). In other words, we establish thresholds on the current storage level, and provide closed-form expressions for optimal actions to be taken under all possible cases (on the relation between the current storage level and the established thresholds).

2) The characterization enables us to compute and implement the optimal threshold policy in a realistic setting with random electricity prices. We illustrate the analysis developed in this paper using real-world data on electricity prices and household demand. We compare the performance of the optimal threshold policy with two heuristic policies: a policy that does not use storage at all, and the threshold policy that is shown to be optimal in a setting with random energy prices but no demand charge [30], [32]. We also benchmark the performance of the characterized optimal policy with a lower bound on the total cost.

Our numerical results demonstrate that the cost saving resulting from the optimal policy (compared with the case without storage) increases with the storage capacity and the maximum charging/discharging rate, and ranges from 7% to 31% with a demand charge of 18$/kW. It is also worth noting that the threshold policy characterized in [30], [32] performs poorly even with 6$/kW demand charge, which highlights the importance of taking demand charge into account when operating energy storage.

The rest of the paper is organized as follows. We formulate the storage operation problem as a dynamic program in Section II. In Section III we characterize an optimal threshold policy that minimizes the consumer’s expected total cost. In Section IV, we discuss the implementation of the characterized optimal threshold policy and present several numerical examples using hourly price data from MISO. Finally, we make some brief concluding remarks in Section V.

II. PROBLEM FORMULATION

We study the operation of a finite-capacity storage owned by an electricity consumer. The consumer has the options of discharging the storage for its own consumption and charging the storage from purchased power. The detailed model of this sequential decision making problem has the following elements.

1) Discrete time: Time periods are indexed by $t = 0, \ldots, T$.

2) Storage capacity: At each stage $t = 0, \ldots, T$, let $x_t \in [0, B]$ denote the storage level at the beginning of stage $t$, where $B > 0$ is the storage capacity.

3) Randomness: For each stage $t = 0, \ldots, T$, let $s_t \in S$ denote the global state. We assume that the set $S$ is finite. The global state evolves as an exogenous Markov chain, of which the transition probability is independent of the consumer’s action. The global state is used to model exogenous (deterministic and random) factors, such as the current time, the demand of other electricity customers, and weather conditions, which have impacts on electricity prices and/or renewable generation.

4) Prices: For $t = 0, \ldots, T - 1$ and given the current global state $s_t$, let $p_t(s_t) \in [0, \infty)$ denote the per-unit electricity price. Here, we have assumed that the electricity price is non-negative; this assumption is commonly made in the literature on dynamic operation of energy storage [30], [31]. We note that the incorporation of negative prices makes the optimal cost-to-go function not (quasi-) convex [36], and therefore eliminates the possibility of establishing any structural results.

5) Demand: For $t = 0, \ldots, T - 1$, the consumer’s net demand for energy is denoted by $d_t(s_t) \in (-\infty, \infty)$. The net demand $d_t(s_t)$ is the difference between the consumer’s demand and her renewable generation (e.g., from photo-voltaic solar panels) at stage $t$.

6) Action: Let $a_t$ denote the change in the storage level, i.e.,

$$x_{t+1} = x_t + a_t,$$

where $x_{\tau}$ denotes the storage level at stage $\tau$. The storage level at the initial stage $0$, $x_0$, is assumed to be exogenous and independent of the consumer’s decision.

7) Rate constraints: There are maximum charging and discharging rates of the energy storage, $R^C$ and $R^D$, i.e., for $t = 0, \ldots, T - 1$,

$$-R^D \leq a_t \leq R^C.$$

This treatment is without loss of optimality, because it is optimal to first use the renewable generation to meet the demand and then charge the residual renewable generation (if any) to the storage.

We have ignored the energy storage self-discharging in (1). This assumption is reasonable as many popular types of modern batteries (e.g., Lead acid, Sodium Sulphur (NaS), Lithium ion, and Vanadium redox batteries) have negligible self-discharging rates ($0 \sim 4\%$ per month) [37].
8) **Energy procurement:** We let $u_t$ denote the amount of energy procured from the grid at stage $t$. If $a_t > 0$, then we must have $u_t = \max\{d_t(s_t) + a_t/\gamma, 0\}$, where $\gamma \in (0, 1)$ is the charging efficiency.

If $a_t < 0$, then $u_t = d_t(s_t) + \eta a_t$, where $\eta \in (0, 1]$ is the discharging efficiency.

At stage $t$, the amount of energy procured from the grid is therefore given by

$$u_t = \left(d_t(s_t) + a_t^+ / \gamma\right)^+ - \eta a_t^- \geq 0,$$  \hfill (3)

where

$$\left(\cdot\right)^+ = \max\{0, \cdot\}, \quad \left(\cdot\right)^- = -\min\{0, \cdot\}. \hfill (4)$$

When the net demand $d_t(s_t)$ is negative, the amount of residual renewable generation may exceed the maximum amount of energy that can be charged into the storage; in this case, we have assumed free disposal of renewable generation in (3).

9) **Peak demand and demand charge:** At each stage $t$, the consumer purchases $u_t$ from the grid. Let $m_t$ denote the maximum energy procurement up to stage $t - 1$, i.e.,

$$m_t = \max\{u_0, u_1, \ldots, u_{t-1}\}, \quad t = 1, \ldots, T \hfill (5)$$

with $m_0 = 0$. There is a demand charge on the peak demand during the entire $T$ stages, $q m_T$, where $q \geq 0$ is the price on peak demand.

We are ready to formulate the operation problem as a ($T+1$)-stage dynamic program by introducing its state space, action sets, transition probabilities, and stage cost. At each stage $t$, the **system state** consists of the current storage level, $x_t$, the peak demand up to stage $t$, $m_t$, and the global state $s_t$.

For $t = 0, \ldots, T-1$, given the current storage level $x_t$ and global state $s_t$, a feasible action $a_t$ satisfies the constraints in Eqs. (2) (3), and the following one:

$$0 \leq x_t + a_t \leq B, \quad t = 1, \ldots, T \hfill (6)$$

i.e., the next-stage storage level $x_{t+1}$ must lie in the range $[0, B]$. Given the current state $(x_t, s_t)$, we let $A(x_t)$ denote the one-dimensional (convex and compact) set of feasible actions that satisfy constraints (2), (3) and (6).

The evolution of the storage level is deterministic, and is governed by Eq. (1). The next-stage $m_{t+1}$ is given by (cf. Eqs. (3) and (5))

$$m_{t+1} = \max\{m_t, \left(d_t(s_t) + a_t^+ / \gamma\right)^+ - \eta a_t^-\}, \quad t = 0, 1, \ldots, T-1. \hfill (7)$$

for $t = 0, 1, \ldots, T-1$. The evolution of the global state $s_t$ is Markovian and does not depend on the current storage level and the action taken by the consumer.

For $t = 0, 1, \ldots, T-1$, the **stage cost** is given by

$$w_t(x_t, m_t, s_t, a_t) = p_t(s_t) u_t + q (u_t - m_t)^+, \quad t = 0, 1, \ldots, T-1 \hfill (8)$$

where $u_t$ is expressed in term of $s_t$ and $a_t$ in Eq. (3), and $p_t(s_t) \geq 0$ is the per-unit energy price at stage $t$. We note that along any sample path $(s_0, \ldots, s_{T-1})$, the sum of the stage cost (in Eq. (8)) equals the consumer’s total cost (including both energy cost and demand charge), i.e.,

$$\sum_{t=0}^{T-1} w_t(x_t, m_t, s_t, a_t) = \sum_{t=0}^{T-1} p_t(s_t) u_t + q \max\{u_0, \ldots, u_{T-1}\}. \hfill (9)$$

At the terminal stage $T$, no action is available, and the stage cost depends only on $(x_T, m_T)$, i.e.,

$$w_T(x_T, m_T, s_T) = -v_T(x_T, s_T), \quad t = T \hfill (10)$$

where the mapping $v_T : [0, B] \times S \to [0, \infty)$ reflects the salvage value of stored energy, and is assumed to be non-decreasing, concave, and continuously differentiable in $x_T$, for every $s_T \in S$.

A **policy** $\pi = (\mu_0, \ldots, \mu_{T-1})$ is a sequence of decision rules such that $\mu_t(x_t, m_t, s_t) \in A(x_t)$, for all $(x_t, m_t, s_t)$ and $t$. We will use $V^\pi_T(x_t, s_t)$ to denote the cost-to-go function under the current system state $(x_t, m_t, s_t)$ a policy $\pi$:

$$J^\pi_t(x_t, m_t, s_t) = w_t(x_t, m_t, s_t, \mu_t(x_t, m_t, s_t)) +$$

$$\mathbb{E}\left\{\sum_{\tau=t+1}^{T-1} w_\tau(x_\tau, m_\tau, s_\tau, \mu_\tau(x_\tau, m_\tau, s_\tau)) + w_T(x_T, m_T, s_T)\right\}, \hfill (11)$$

where the expectation is over the sequence of global states $\{s_\tau\}_{\tau=t+1}^{T}$. Since the set of global states $S$ is finite, the cost-to-go function is always bounded, under any policy $\pi$.

By a slight abuse of notation, we will use $J_t(x_t, m_t, s_t)$ to denote the optimal payoff-to-go function, i.e.,

$$J_t(x_t, m_t, s_t) \triangleq \inf_\pi \{J^\pi_t(x_t, m_t, s_t)\}. \hfill (11)$$

We say a policy $\pi^*$ is optimal if it attains the optimal cost-to-go defined above, i.e., $J_0^\pi(x_0, 0, s_0) = J_0(x_0, 0, s_0)$, for all initial states $(x_0, s_0)$.

**Remark 1:** Compared with the dynamic programming problems formulated in [30], [31], [32] to study consumers’ storage operation, the DP formulated in the section has one additional dimension of system state that records the up-to-date maximum energy procurement, $m_t$. We note that computing exact optimal policies by brute-force dynamic programming is challenging, since the formulated DP has two dimensional continuous state space $(x_t, m_t)$. In the next section, we provide a threshold characterization on an optimal policy that enables the computation and implementation of the optimal policy. □

### III. Optimal Storage Operation

In this section, we characterize an optimal threshold policy for the dynamic program formulated in Section II. In Section III-A, we derive some preliminary results, for example, the convexity of the cost-to-go functions. Leveraging on these preliminary results, we establish a threshold based characterization on an optimal storage operation policy in Section III-B. A graphical illustration on the characterized optimal threshold policy is also provided.
A. Preliminary Results

Before proceeding, we introduce some notations that will be used later. For $t = 0, \ldots, T - 1$, the Bellman equation yields

$$J_t(x_t, m_t, s_t) = \min_{a_t \in A(x_t)} \left\{ w_t(x_t, m_t, s_t, a_t) + \tilde{J}_{t+1} | s_{t+1} \right\},$$

where $x_{t+1}$ and $m_{t+1}$ are determined by Eqs. (1) and (7), respectively, and $\tilde{J}_{t+1} | s_{t+1}$ denotes the conditional expected cost-to-go function at stage $t + 1$, i.e., for $t = 0, \ldots, T - 1$

$$\tilde{J}_{t+1} | s_{t+1} (x_{t+1}, m_{t+1}, s_{t+1}) = \mathbb{E} \left\{ J_{t+1}(x_{t+1}, m_{t+1}, s_{t+1}) \mid s_t \right\},$$

where the expectation is over the next global state $s_{t+1}$, conditioned that the current global state is $s_t$.

**Lemma 1:** For every $s \in S$ and $t = 0, \ldots, T$, the optimal cost-to-go function $J_t(x, m, s)$ is non-increasing and convex in $(x, m)$.

**Proof:** To prove that $J_t(x, m, s)$ is convex in $(x, m)$, we will show that for any $(x, m), (x', m')$

$$J_t(x, m, s) + J_t(x', m', s) \geq J_t(x, m, s) + J_t(x', m', s),$$

where $$J_t(x, m, s) = \frac{1}{2} \left( (x + x')/2, (m + m')/2 \right).$$

Let $\pi^* = (\pi^*_1, \ldots, \pi^*_T)$ be an optimal policy from stage $\tau$ to $T - 1$. For the rest of the proof, we fix an arbitrary sequence of realized global states $\tilde{s} = (s_t, \ldots, s_T)$. For $t = \tau + 1, \ldots, T - 1$, under the optimal policy $\pi^*$ and $\tilde{s}$, let $(x_t, m_t)$ and $(x'_t, m'_t)$ denote the system states resulting from the system state $(x, m)$ and $(x', m')$, respectively. Similarly, under the optimal policy $\pi^*$ and $\tilde{s}$, we let $(a_{\tau}, \ldots, a_{T-1})$ and $(a'_\tau, \ldots, a'_{T-1})$ denote the sequence of actions following the system state $(x, m)$ and $(x', m')$, respectively.

Following the system state $(\tilde{x}_\tau, \tilde{m}_\tau)$ at stage $\tau$, a heuristic policy $\tilde{\pi}$ follows the following actions along the sample path $\tilde{s}$

$$\tilde{a}_{\tau}, \ldots, \tilde{a}_{T-1} = [(a_{\tau}, \ldots, a_{T-1}) + (a'_\tau, \ldots, a'_{T-1})]/2.$$

It follows from Eqs. (1), (3) and (6) that the action sequence $(\tilde{a}_{\tau}, \ldots, \tilde{a}_{T-1})$ is feasible under the system state $(\tilde{x}_\tau, \tilde{m}_\tau)$ and the sample path $\tilde{s}$.

Under the policy $\tilde{\pi}$ and $\tilde{s}$, for $t = \tau, \ldots, T - 1$, let $(\tilde{x}_t, \tilde{m}_t)$ denote the state system following the system state $(\tilde{x}_\tau, \tilde{m}_\tau)$. According to Eqs. (1) and (15), we have

$$\tilde{x}_t = (x_t + x'_t)/2, \quad t = \tau + 1, \ldots, T - 1.$$

Note that we may not have $\tilde{m}_t = (m_t + m'_t)/2$ because its state transition in (7) is not linear.

We divide the stage cost into two parts:

$$w_t(x_t, m_t, s_t, a_t) = w^c_t(x_t, m_t, s_t, a_t) + w^d_t(x_t, m_t, s_t, a_t),$$

where

$$w^c_t(x_t, m_t, s_t, a_t) = p_t(s_t)u_t$$

is the energy cost, and

$$w^d_t(x_t, m_t, s_t, a_t) = q(u_t - m'_t)^+$$

denotes the incremental demand charge.

We note from (3) that $u_t$ is convex and non-decreasing in $a_t$. It then follows from (17) that $w^c_t$ is convex and non-decreasing in $a_t$, i.e.,

$$w^c_t(x_t, m_t, s_t, a_t) + w^c_t(x'_t, m'_t, s_t, a'_t) \geq 2w^c_t(\tilde{x}_t, m_t, s_t, \tilde{a}_t),$$

for all $t = \tau + 1, \ldots, T - 1$.

Under the sequence of realized global states $\tilde{s}$, let $\{u_t\}_{t=\tau}^{T-1}$ and $\{\tilde{u}_t\}_{t=\tau}^{T-1}$ denote the sequence of energy procurement (cf. its definition in (3)) resulting from the sequence of actions $\{a_t\}_{t=\tau}^{T-1}$, $\{a'_t\}_{t=\tau}^{T-1}$, and $\{\tilde{a}_t\}_{t=\tau}^{T-1}$, respectively. Since $u_t$ is convex and non-decreasing in $a_t$, and therefore $\tilde{u}_t \leq (u_t + u'_t)/2$ for $t = \tau, \ldots, T - 1$. We have

$$\frac{1}{2} \sum_{t=\tau}^{T-1} w^c_t(\tilde{x}_t, m_t, s_t, \tilde{a}_t)$$

$$= q \max\{u_t, \ldots, u_{T-1}\} - q\tilde{m}_\tau$$

$$\leq q \max\{\tilde{u}_t, \ldots, \tilde{u}_{T-1}\} - q(m_t + m'_t)/2$$

where the first equality follows from (5) and (18), and the second inequality follows from the convexity of the max function.

We use $V^\pi_{\tau}(x, m, s)$ to denote the total (realized) cost under the state $(\tilde{x}_\tau, \tilde{m}_\tau)$ at stage $\tau$, the sequence of realized global states $\tilde{s} = (s_\tau, \ldots, s_T)$, and the policy $\pi$. $V^\pi_{\tau}(x, m, s)$ and $V^\pi_{\tau}(x', m', s')$ could be defined in a similar manner. We have

$$V^\pi_{\tau}(x, m, s) + V^\pi_{\tau}(x', m', s')$$

$$= \frac{1}{2} \left\{ \sum_{t=\tau}^{T-1} \left( w^c_t(x_t, m_t, s_t, a_t) + w^c_t(x'_t, m'_t, s_t, a'_t) \right) + \sum_{t=\tau}^{T-1} \left( w^d_t(x_t, m'_t, s_t, a'_t) + w^d_t(x'_t, m_t, s_t, a_t) \right) + w^d_t(x_t, s_t, m_t, a_t) \right\}$$

$$\geq \sum_{t=\tau}^{T-1} \left( w^c_t(\tilde{x}_t, m_t, s_t, \tilde{a}_t) + w^c_t(\tilde{x}_t, m'_t, s_t, \tilde{a}_t) \right) + w^d_t(\tilde{x}_t, s_t, m_t, a_t)$$

where the inequality follows from the concavity of the salvage value (cf. Eq. (9)), (19), and (20).

Since the inequality in (21) holds for every possible sequence of realized global states $\tilde{s} = (s_\tau, \ldots, s_T)$, we have

$$J_t(x, m, s) + J_t(x', m', s) \geq J_t(\tilde{x}, \tilde{m}, s)$$

$$\geq J_t(\tilde{x}, \tilde{m}, s),$$

(22)
where $J^2_T(\bar{x}_T, \bar{m}_T, s_T)$ is the (expected) cost-to-go resulting from the heuristic policy $\pi$ and $(\bar{x}_T, \bar{m}_T)$, and the second inequality follows from the definition of $J_T(\bar{x}_T, \bar{m}_T, s_T)$ in (11).

Lemma 2: For $t = 0, \ldots, T-1$ and given every system state $(x_t, m_t, s_t)$, the right hand side of the Bellman equation (12) is convex in $a_t$.

The proof of Lemma 2 is similar to that of Lemma 1, and is therefore deferred to Appendix A.

### B. Optimal Policy Threshold

We first express the expected cost-to-go on the right hand side of the Bellman equation (12) in terms of the action $a_t$, where $s_t$ is dropped from $p_t(s_t)$ and $d_t(s_t)$ for notation conciseness,

$$ p_t(d_t + a_t^+ / \gamma - \eta a_t^-)^+ + q(d_t + a_t^+ / \gamma - \eta a_t^- - m_t)^+ + \bar{J}_{t+1|s_t}(x_t + a_t^- - a_t^+, m_t + (d_t + a_t^+ / \gamma - \eta a_t^- - m_t)^+) $$

(23)

In Appendix B, we write the first-order conditions that are necessary and sufficient for an action $a_t$ to minimize the right hand side of the Bellman equation (12).

For the rest of this paper, we use $\partial^+ \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$ and $\partial^- \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$ to denote the right and left partial derivative of $\bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$ with respect to $x_{t+1}$, respectively. Analogously, we use $\partial^+ \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$ and $\partial^- \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$ to denote the right and left partial derivative of $\bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$ with respect to $m_{t+1}$, respectively.

We define the following notation for the rest of the paper,

$$ \Delta_t(m_t, s_t) \equiv d_t(s_t) - m_t. $$

(24)

Based on the first-order conditions given Appendix B, in (26) we define four functionals that will be useful in the characterization of an optimal storage operation policy (cf. Theorem 1). For notational conciseness, we will drop $(s_t)$ from $p_t(s_t)$, $(m_t, s_t)$ from $\Delta_t(m_t, s_t)$ and $(x_{t+1}, m_{t+1})$ from the partial derivatives of $\bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})$. In (26), given the current system state $(x_t, m_t, s_t)$ and $x_{t+1} \in [0, B]$, $m_{t+1}$ is determined by Eqs. (1) and (7) as follows

$$ m_{t+1}(x_t, m_t, s_t, x_{t+1}) = \max \left\{ m_t, (d_t(s_t) + (x_{t+1} - x_t)^+ / \gamma)^+ - \eta(x_{t+1} - x_t)^- \right\}, $$

(25)

where $(\cdot)^+$, $(\cdot)^-$ are defined in (4), and $m_{t+1}$ (with some abuse of notation) denotes a mapping from $(x_t, m_t, s_t, x_{t+1})$ to the state $m_{t+1}$ defined in Section II.

Remark 2: It follows from Lemma 2 that the right hand side (RHS) of the Bellman equation (12) is convex in $a_t$ on the bounded interval $[-x_t, B - x_t]$. As a result, over the interval $[-x_t, B - x_t]$, the right derivative of the RHS of (12) is non-decreasing and right-continuous in $a_t$, and the left derivative of the RHS of (12) is non-decreasing and left-continuous in $a_t$ (cf. page 309 in [38]).

The functional $\ell^1_t(x_t, m_t, s_t)$ seeks to find the minimum point over a closed interval $x_{t+1} \in [x_t + \gamma(\Delta_t^{-} - B), B]$, or equivalently, over $a_t \in [\gamma(\Delta_t^{-}) - B, -x_t + x_{t+1} \leq \gamma(\Delta_t^{-}) - B]$, such that the right derivative of the RHS of (12) with respect to $a_t$ is non-negative. If such an $a_t$ exists, the minimum point is always attainable due to the fact that right derivative of the RHS of (12) is non-decreasing and right-continuous in $a_t$. Analogously, the minimum (and maximum) point in the definition of $\ell^2_t$ and $h^1_t, h^2_t$, respectively is also attainable.

According to the definition of the four functionals in (26), given the current system state $(x_t, m_t, s_t)$, we always have

$$ \ell^1_t \geq \ell^2_t \geq \min\{B, x_t + \gamma(dt^-)\} \geq x_t \geq h^1_t \geq h^2_t. $$

(27)

Further, when $d_t(s_t) \leq 0$, we have $x_t = h^1_t = h^2_t$. □

Based on the functionals defined in (26), we now define the four thresholds that will be useful in the characterization of an optimal policy (see Theorem 1).

$$ \ell^1_t(x_t, m_t, s_t) \equiv \sup \left\{ \{0\} \cup \{x : x + \gamma((\Delta_t^{-} - B) < \ell^1_t(x_t, m_t, s_t)\} \right\}; $n_{p2}(x_t, m_t, s_t) \equiv \sup \left\{ \{x : x < \ell^2_t(x_t, m_t, s_t)\} \cup \{0\} \right\}; $h^1_t(x_t, m_t, s_t) \equiv \inf \left\{ \{x : x > h^1_t(x_t, m_t, s_t)\} \cup \{B\} \right\}; $h^2_t(x_t, m_t, s_t) \equiv \inf \left\{ \{B\} \cup \{x : x - (\Delta_t^{-} - B) \geq \eta > h^2_t(x_t, m_t, s_t)\} \right\}. $$(28)

It is worth noting that different from the functionals defined in (26), the four thresholds defined in (28) do not depend on the current storage level $x_t$. The establishment of these thresholds will significantly simplify the description and computation of the optimal threshold policy characterized in Theorem 1. The following lemma establishes some important relations among the functionals and thresholds defined in Eqs. (26) and (28).

**Lemma 3:** We have the following.
1) For all $(m, s)$ such that $d_l(s) = m$, $\ell^1_l(m, s) \leq \bar{h}^2_l(m, s)$.
2) For all $(m, s)$ such that $d_l(s) > m$, we have $\ell^1_l(m, s) \leq \bar{h}^1_l(m, s) \leq \bar{h}^2_l(m, s)$.
3) For all $(m, s)$ such that $d_l(s) < m$, we have $\ell^1_l(m, s) \leq \ell^2_l(m, s) \leq \bar{h}^2_l(m, s)$.

The proof of Lemma 3 is given in Appendix C.

Theorem 1 (Optimal Threshold Policy): Given the current system state $(x_t, m_t, s_t)$ at stage $t \in \{0, 1, \ldots, T - 1\}$, an optimal policy is characterized in the following.

If $d_l(s_t) = m_t$, an optimal action is given by

$$a^*_t = \begin{cases} 
\min \{ \ell^1_t - x_t, R^C \}, & \text{if } x_t < \ell^1_t, \\
0, & \text{if } x_t \in [\ell^1_t, \bar{h}^1_t], \\
\max \{ h^2_t - x_t, -R^D \}, & \text{if } x_t > \bar{h}^2_t.
\end{cases}$$  \hfill (29a)

If $d_l(s_t) > m_t$, an optimal action is given by

$$a^*_t = \begin{cases} 
\min \{ \ell^1_t - x_t, R^C \}, & \text{if } x_t < \ell^1_t, \\
0, & \text{if } x_t \in [\ell^1_t, \bar{h}^1_t], \\
\max \{ h^1_t - x_t, -R^D \}, & \text{if } x_t \in (\bar{h}^1_t, \bar{h}^2_t), \\
\max \{ h^2_t - x_t, -R^D \}, & \text{if } x_t > \bar{h}^2_t.
\end{cases}$$  \hfill (29b)

If $d_l(s_t) < m_t$, an optimal action is given by

$$a^*_t = \begin{cases} 
\min \{ \ell^1_t - x_t, R^C \}, & \text{if } x_t < \ell^1_t, \\
\min \{ \ell^2_t - x_t, R^C \}, & \text{if } x_t \in [\ell^1_t, \ell^2_t], \\
0, & \text{if } x_t \in (\ell^2_t, \bar{h}^1_t), \\
\max \{ h^2_t - x_t, -R^D \}, & \text{if } x_t > \bar{h}^2_t.
\end{cases}$$  \hfill (29c)

Theorem 1 is proved in Appendix D.

IV. NUMERICAL IMPLEMENTATION AND RESULTS

In this section, we present several numerical examples that compare the total cost (the sum of energy cost and demand charge) achieved by different policies under various parameter settings on storage capacity, the maximum charging and discharging rates, and the demand charge rate.

In Section IV-A, we discuss the construction of a dynamic programming (DP) model using real price data. In Section IV-B, we introduce the approach we use to compute the optimal threshold policy characterized in Theorem 1. In Section IV-C, the optimal policy is compared with several heuristic policies (in terms of total cost and computational time); a lower bound on the total cost is also provided for benchmark.

A. Probabilistic DP Model with Cycles

We obtain five years of 1-hour price data from MISO (http://www.energyonline.com/Data)\(^6\) and one month of 15-minute household demand data from http://www.doc.ic.ac.uk/~ndk3810/data/\(^6\). We use the first four years of collected pricing data to train a probabilistic model with cycles. We use the one remaining year data for testing.

The daily effects observed from the price data motivate us to model the (random) electricity prices as a cyclically stochastic process with both deterministic variability and stochastic variability [19]. An uncontrolled process $\{p_t\}_{t=1}^\infty$ is cyclic with cycle length $N$ if the joint probabilistic distribution of $\{P_{t+N}+t\}_{t=0}^{N-1} \sim \text{id}$ is identical for all $l \in \{1, 2, \ldots\}$, where $N$ is the number of periods in a cycle. In our simulation, each cycle lasts for 24 hours, and each stage lasts for 15 minutes. We therefore have $N = 96$, i.e., each cycle has 96 stages.

The entire optimization horizon considered in the simulation lasts for a month, i.e., $T = 2880$. We assume that the price $\{p_t\}_{t=1}^{T-1}$ is cyclic with cycle length of 24 hours. For each $t = 0, 1, \ldots, T - 1$, we let $n = (t \mod N)$, and

$$p_t = b_n + c_n Z_t,$$

\(^6\)In the hourly price data used in our simulation, the frequency of negative prices is less than 2%.

\(^5\)For notational conciseness, we drop $(x_t, m_t, s_t)$ and $(m_t, s_t)$ from the expressions of $\ell^1_l, \ell^2_l, \bar{h}^1_l, \bar{h}^2_l$ and $\ell^1_l, \ell^2_l, \bar{h}^1_l, \bar{h}^2_l, \Delta$, respectively, because a fixed system state $(x_t, m_t, s_t)$ is considered within this theorem.
where \( \{b_1, b_2, \ldots, b_N\} \) and \( \{c_1, c_2, \ldots, c_N\} \) are sets of deterministic constants while \( \{Z_t\}_{t=1}^{\infty} \) is a stationary and independent discrete time Markov chain with probability transition matrix \( \mathcal{P} \). In our simulation, we set the number of global states as 6 by letting \( Z_t \in \{0, 0.2, 0.4, 0.6, 0.8, 1\} \), and train a 6 \times 6 probability transition matrix.

All numerical experiments are implemented in Matlab R2012b on an Intel Core i7-4770 3.40GHz PC with 16GB memory. The training time is around 2 seconds.

### B. Computation of the Optimal Threshold Policy

In this subsection, we introduce the approach we use to compute and implement of the optimal threshold policy characterized in Theorem 1. A key step is to (approximately) compute \( \partial_x^\pm \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1}) \) and \( \partial_m^\pm \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1}) \) for all \( t = 0, 1, \ldots, T-1 \).

For the final stage \( T \), \( \partial_x^\pm J_T(x_T, m_T, s_T) \) and \( \partial_m^\pm J_T(x_T, m_T, s_T) \) can be easily computed using the salvage value in Eq. (9). Then, for \( t = T-1, \ldots, 0 \), the partial derivatives \( \partial_x^\pm J_t(x_t, m_t, s_t) \) and \( \partial_m^\pm J_t(x_t, m_t, s_t) \) can be computed through backward induction based on the optimal threshold policy characterized in Theorem 1. Finally, given \( \partial_x^\pm J_t(x_t, m_t, s_t) \) and \( \partial_m^\pm J_t(x_t, m_t, s_t) \), it is straightforward to compute the partial derivatives \( \partial_x^\pm \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1}) \) and \( \partial_m^\pm \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1}) \) through the definition in Eq. (13).

It is worth noting, however, that the dynamic program we seek to solve has a two dimensional continuous state space over \( (x_t, m_t) \). It is therefore infeasible to compute \( \partial_x^\pm J_{t+1|s_t}(x_{t+1}, m_{t+1}) \) and \( \partial_m^\pm J_{t+1|s_t}(x_{t+1}, m_{t+1}) \) at every point \( (x_{t+1}, m_{t+1}) \). For the case where storage capacity \( B \) equals 3 hours of average demand, we consider 50 discretized levels for \( x \) and 6 discretized levels for \( m \) for computing \( \partial_x^\pm J_{t+1|s_t}(x, m) \), and consider 6 discretized levels for \( x \) and 50 discretized levels for \( m \) when computing \( \partial_m^\pm J_{t+1|s_t}(x, m) \). In our simulation, we increase the number of discretized storage levels \( x \) in proportion to the storage capacity \( B \).

Using the approach described above, we compute the partial derivatives \( \{\partial_x^\pm \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})\}_{T=1}^{T-1} \) and \( \{\partial_m^\pm \bar{J}_{t+1|s_t}(x_{t+1}, m_{t+1})\}_{T=1}^{T-1} \) off-line and store them in a look-up table. The on-line implementation of the optimal threshold policy is simpler. Given the current system state \( (x_t, m_t, s_t) \) and based on the look-up table (on partial derivatives), we compute the functionals and thresholds defined in (26) and (28), which determine an optimal action according to the characterization in Theorem 1.

### C. Numerical Results

All numerical results are obtained using the one-year real price data, and the (deterministic) one-month household demand data (that is repeated for 12 times). We compare the average monthly total cost (energy cost and demand charge) achieved by different policies during this one year period. The charging/discharging efficiency \( \gamma = \eta = 0.9 \).

Before presenting our numerical results, we introduce two heuristic policies and a lower bound (on total cost) that will be used for benchmark. The first heuristic policy we consider is the threshold policy characterized in (30), (32) that is shown to be optimal in a setting without demand charge. The second heuristic policy does not use the storage at all, and at each stage \( t \), simply procures energy from the grid to meet the demand, i.e., \( u_t = d_t(s_t) \). The cost saving resulting from the optimal policy (compared to the case without storage) is the value of storage, which is the consumer’s net benefit obtained by optimally operating the storage [32].

We will also benchmark the performance of the characterized optimal policy with a lower bound on the total cost, which is computed under perfect information (on both energy price and demand) through the following linear program:

\[
\min_{\{a_t\}_{t=0}^{T-1}} \sum_{t=0}^{T-1} p_t u_t + w_T + q m_T \\
\text{s.t.} \quad m_T \geq u_t, \quad (1), \quad (2), \quad (3), \quad (6), \quad t = 0, \ldots, T-1.
\]

We note that the lower bound is not practically achievable since the consumer may not perfectly know the hourly electricity prices in the next month.

In Fig. 2, the demand charge is set to be 18$/kW; for the left (right) subplot, it takes 4 stages (16 stages, respectively) to fully charge or discharge the storage. We observe that the DP solution without demand charge (i.e., the threshold policy characterized in [30], [32]) that is optimal in a setting without demand charge) results in the highest cost, because the policy procures energy from the grid to charge the storage as long as electricity price is relatively low even when the net demand is (very) high. As a result, this policy (that does not take demand charge into account) leads to high peak demand and therefore high demand charge. Indeed, the DP solution without demand charge results in higher cost than the case without storage, due to the high demand charge the policy incurs. This numerical
result highlights the importance of taking demand charge into account when operating energy storage.

In Fig. 2, the cost saving resulting from the optimal policy (compared to the case without storage) increases with the storage capacity and the maximum charging/discharging rate. In the left subplot (with $B/R_C^C = 4$), the cost reduction resulting from the optimal policy ranges from 21% to 31% of the total cost without storage; while in the right subplot (with $B/R_C^C = 16$), this number ranges from 7% to 14%. We also note that the gap between the total cost achieved by the optimal policy and the lower bound is less than the cost saving resulting from the optimal policy (compared to the case without storage).

In Fig. 3, the demand charge is set to be smaller at 6$/kW. Compared with the numerical results in Fig. 2, the performance gaps among the DP solution without demand charge, the heuristic policy that does not use storage, and the optimal threshold policy shrink in Fig. 3 due to the smaller demand charge. In the left subplot (with $B/R_C^C = 4$), the cost reduction resulting from the optimal policy ranges from 11% to 17% of the total cost without storage; while in the right subplot (with $B/R_C^C = 16$), this number ranges from 3% to 7%.

The computational time of the optimal threshold policy increases with the storage capacity $B$ since the number of discretized storage levels considered in the simulation is proportional to $B$. We note that the computational time is not sensitive to the other parameters. When the storage capacity equals 7.5 hour average demand, the computational time for the DP solution without demand charge, the heuristic policy that does not use storage, the optimal policy, and the lower bound is 118, 0.0001, 15149, and 5.8 seconds, respectively.

V. Conclusion

Through a dynamic programming formulation, we study the optimal operation of energy storage operated by a consumer who owns intermittent renewable generation and faces (possibly random) fluctuating electricity prices and demand charge. The formulated dynamic program is hard to solve, due to its two dimensional continuous state space. Through an approach that is different from those used in the literature, we provide a complete characterization on an optimal threshold policy that minimizes the consumer’s expected total cost (the sum of energy cost and demand charge).

The characterization enables us to implement the optimal storage operation policy in a realistic setting with random energy prices. We train a dynamic programming model using real household demand data and hourly price data from MIS-O. Our numerical results demonstrate that the characterized optimal threshold policy significantly outperforms the policy that is shown to be optimal without demand charge [30], [32], highlighting the importance of taking demand charge into account when operating energy storage.

VI. Acknowledgement

The authors gratefully acknowledge Prof. Costas Courcoubetis and Prof. Yangfang (Helen) Zhou for the helpful discussion. We also thank the anonymous reviewers and the associate editor for their valuable comments.

REFERENCES

APPENDIX A 
PROOF OF LEMMA 2

Within this proof, we use the following notation to denote the right hand side of the Bellman equation (12):

\[ V_t(x_t, m_t, s_t, a_t) = w_t(x_t, m_t, s_t, a_t) + J_{t+1}|s_t(x_{t+1}, m_{t+1}), \]

for \( t = 0, \ldots, T - 1 \).

To prove that \( V_t(x_t, m_t, s_t, a) \) is convex in \( a \), we will show that for any \( a, a' \in \mathcal{A}(x_t, m_t, s_t, a') \)

\[ V_t(x_t, m_t, s_t, a_t) + V_t(x_t, m_t, s_t, a'_t) \]

\[ \geq V_t(x_t, m_t, s_t, a_t) \frac{1}{2} \]

where \( a_{\pi_t} = (a_t + a'_t)/2 \).

Let \( \pi^* = (\mu^*_1, \ldots, \mu^*_{T-1}) \) be an optimal policy from stage \( \tau + 1 \) to \( T - 1 \). For the rest of the proof, we fix an arbitrary sequence of realized global states \( \tilde{s} = (s, \ldots, s_T) \). The state \( (x_t, m_t) \) at stage \( \tau \), under the optimal policy \( \pi^* \) and \( \tilde{s} \), let \( \{(x_t, m_t)^T_{1=\tau+1}\} \) and \( \{(x_t, m_t')^T_{1=\tau+1}\} \) denote the sequence of system states resulting from the action \( a_t \) and \( a_t' \) (taken at stage \( \tau \)), respectively.

Analogously, given the state \( (x_{\tau}, m_{\tau}) \) at stage \( \tau \), we let \( (a_{\tau}, \ldots, a_{T-1}) \) and \( (a'_{\tau}, \ldots, a'_{T-1}) \) denote the sequence of actions resulting from the optimal policy \( \pi^* \) and \( \tilde{s} \), following the action \( a_{\tau} \) and \( a'_{\tau} \) at stage \( \tau \), respectively.

Following the system state \( (x_{\tau}, m_{\tau}) \) at stage \( \tau \), a heuristic policy \( \tilde{\pi} \) takes the following actions along the sample path \( \tilde{s} \)

\[ (a_{\tau}, \ldots, a_{T-1}) = [(a_{\tau}, \ldots, a_{T-1}) + (a'_{\tau}, \ldots, a'_{T-1})]/2. \]

(35)

It follows from Eqs. (1), (3) and (6) that the action sequence \( (a_{\tau}, \ldots, a_{T-1}) \) is feasible under the system state \( (\tilde{x}_{\tau}, \tilde{m}_{\tau}) \) and the sample path \( \tilde{s} \).

Under the sequence of realized global states \( \tilde{s} \) and the system state \( (x_{\tau}, m_{\tau}) \) at stage \( \tau \), let \( \{(\tilde{x}_t, \tilde{m}_t)^T_{1=\tau+1}\} \) denote the system state resulting from policy \( \tilde{\pi} \). According to Eqs. (1) and (35), we have

\[ \tilde{x}_t = (x_t + x'_t)/2, \quad t = \tau + 1, \ldots, T - 1. \]

Note that we may not have \( \tilde{m}_t = (m_t + m'_t)/2 \) because its state transition in (7) is not linear.

We divide the stage cost into two parts:

\[ w_t(x_t, m_t, s_t, a_t) = w^s_t(x_t, m_t, s_t, a_t) + w_t^d(x_t, m_t, s_t, a_t), \]

(36)

where the energy cost \( w^s_t(x_t, m_t, s_t, a_t) \) is defined in (17), and the incremental demand charge \( w_t^d(x_t, m_t, s_t, a_t) \) is defined in (18).

Under the sequence of realized global states \( \tilde{s} \), let \( \{u_t^T_{1=\tau}\}, \{u'_t^T_{1=\tau}\} \) and \( \{\tilde{u}_t^T_{1=\tau}\} \) denote the sequence of energy procurement (cf. its definition in (3)) resulting from the sequence of actions \( \{a_t^T_{1=\tau}\}, \{a'_t^T_{1=\tau}\} \) and \( \{\tilde{a}_t^T_{1=\tau}\} \), respectively. We note from (3) that \( u_t \) is convex and non-decreasing in \( a_t \), and therefore \( u_t \leq (u_t + u'_t)/2 \) for \( t \geq \tau \). We have

\[ w_t^d(x_t, m_t, s_t, a_t) + \sum_{t=\tau+1}^{T-1} w_t^d(\tilde{x}_t, \tilde{m}_t, s_t, \tilde{a}_t) \]

\[ = q \max \{\tilde{u}_t, \ldots, \tilde{u}_{T-1}\} - q(m_t + m'_t)/2 \]

\[ \leq q \max \{u^T_{1=\tau}, u'_t^T_{1=\tau}\} - q(m_t + m'_t)/2 \]

\[ \leq q \max \{u_t^T_{1=\tau}, u'_t^T_{1=\tau}\} - q(m_t + m'_t)/2 \]

\[ \leq q \max \{u_t^T_{1=\tau}, u'_t^T_{1=\tau}\} - q(m_t + m'_t)/2 \]

\[ = \left[ \sum_{t=\tau+1}^{T-1} w_t^d(x_t, m_t, s_t, a_t) + \sum_{t=\tau+1}^{T-1} w_t^d(x_t, m_t, s_t, a'_t) \right]/2, \]

(37)

where the first equality follows from (5) and (18), and the second inequality follows from the convexity of max function.
\(W_r(a^*, \pi^*)\) and \(W_r(\bar{a}_r, \tilde{\pi})\) denote the total cost resulting from the sequence of actions \(\{a_t\}_{t=0}^{T-1}, \{a_t^*\}_{t=0}^{T-1}, \{\bar{a}_t\}_{t=0}^{T-1}\), respectively. Based on (19), (36), and (37), we have
\[
\frac{W_r(a^*, \pi^*) + W_r(a^*_t, \pi^*)}{2} \geq W_r(\bar{a}_r, \tilde{\pi}). \tag{38}
\]
Since the inequality in (38) holds for every possible sequence of realized global states \(s = (s_r, \ldots, s_T)\), we have
\[
\frac{V_r(x_t, m_t, s_r, a_t) + V_r(x_t, m_t, s_r, a_t^*)}{2} \geq V_r^\pi(x_t, m_t, s_r, \bar{a}_r), \tag{39}
\]
where \(V_r^\pi(x_t, m_t, s_r, \bar{a}_r)\) is the expected cost-to-go resulting from the heuristic policy \(\bar{\pi}\) (following the system state \((x_t, m_t, s_r)\)), and the second inequality follows from the definition of \(V_r(x_t, m_t, s_r, \bar{a}_r)\) in (33).

\section*{Appendix B}

\subsection*{First-Order Conditions}

Under a given system state \((x_t, m_t, s_t)\), we write the (necessary and sufficient) first-order conditions for an action \(a_t\) to minimize the right hand side of the Bellman equation (12) in Eqs. (40)-(43), where \(x_{t+1}\) is determined by \(x_t\) and \(a_t\) according to (1), and \(m_{t+1}\) is determined by \(m_t\) and \(a_t\) according to (3) and (5).

We first consider the case with \(d_t(s_t) = m_t\).

\[
\begin{align*}
&\begin{cases}
    p_t / \gamma + q / \gamma + \partial_x^+ J_{t+1 \mid s_t} + \partial_m^+ J_{t+1 \mid s_t} / \gamma \geq 0, & a_t^+ \geq 0, \\
    p_t / \gamma + q / \gamma + \partial_x^- J_{t+1 \mid s_t} + \partial_m^- J_{t+1 \mid s_t} / \gamma \leq 0, & a_t^- \geq 0, \\
    -\eta p_t - \partial_x^- J_{t+1 \mid s_t} \geq 0, & d_t / \eta > a_t^- \geq 0, \\
    -\eta p_t - \partial_x^+ J_{t+1 \mid s_t} \leq 0, & d_t / \eta \geq a_t^+ > 0.
\end{cases}
\end{align*}
\]
For the case with \(d_t(s_t) > m_t\), the first-order conditions are
\[
\begin{align*}
&\begin{cases}
    p_t / \gamma + q / \gamma + \partial_x^+ J_{t+1 \mid s_t} + \partial_m^+ J_{t+1 \mid s_t} / \gamma \geq 0, & a_t^+ \geq 0, \\
    p_t / \gamma + q / \gamma + \partial_x^- J_{t+1 \mid s_t} + \partial_m^- J_{t+1 \mid s_t} / \gamma \leq 0, & a_t^- \geq 0, \\
    -\eta p_t - \partial_x^- J_{t+1 \mid s_t} \geq 0, & \Delta_t / \eta \geq a_t^- \geq 0, \\
    -\eta p_t - \partial_x^+ J_{t+1 \mid s_t} \leq 0, & \Delta_t / \eta \geq a_t^+ > 0.
\end{cases}
\end{align*}
\]
For the case with \(0 \leq d_t(s_t) < m_t\), the first-order conditions are
\[
\begin{align*}
&\begin{cases}
    p_t / \gamma + q / \gamma + \partial_x^+ J_{t+1 \mid s_t} + \partial_m^+ J_{t+1 \mid s_t} / \gamma \geq 0, & a_t^+ \geq 0, \\
    p_t / \gamma + q / \gamma + \partial_x^- J_{t+1 \mid s_t} + \partial_m^- J_{t+1 \mid s_t} / \gamma \leq 0, & a_t^- \geq 0, \\
    -\eta p_t - \partial_x^- J_{t+1 \mid s_t} \geq 0, & \Delta_t / \eta \geq a_t^- \geq 0, \\
    -\eta p_t - \partial_x^+ J_{t+1 \mid s_t} \leq 0, & \Delta_t / \eta \geq a_t^+ \geq 0.
\end{cases}
\end{align*}
\]

\section*{Appendix C}

\textbf{Proof of Lemma 3}

The following lemma will be useful in the proof for both Lemma 3 and Theorem 1.

\textbf{Lemma 4}: We have the following.

1) For all \((m_t, s_t)\), we have
\[
x < \bar{\ell}_t^1 (m_t, s_t) \Leftrightarrow x + \gamma (\Delta_t(m_t, s_t))^{-1} < \ell_t^1(x, m_t, s_t). \tag{44}
\]
2) For all \((m_t, s_t)\) such that \(d_t(s_t) < m_t\),
\[
x < \bar{\ell}_t^2 (m_t, s_t) \Leftrightarrow x < \ell_t^2(x, m_t, s_t). \tag{45}
\]
3) For all \((m_t, s_t)\) such that \(d_t(s_t) > m_t\),
\[
x > \bar{\ell}_t^1 (m_t, s_t) \Leftrightarrow x > \ell_t^1(x, m_t, s_t). \tag{46}
\]
4) For all \((m_t, s_t)\) such that \(d_t(s_t) \geq 0\),
\[
x > \bar{\ell}_t^2 (m_t, s_t) \Leftrightarrow x - (\Delta_t(m_t, s_t)) / \eta > \ell_t^2(x, m_t, s_t). \tag{47}
\]

\textbf{Proof}: We will first prove Eq. (44). By the definition of \(\bar{\ell}_t^1 (m_t, s_t)\), we have \(x + \gamma (\Delta_t(m_t, s_t))^{-1} \geq \ell_t^1(x, m_t, s_t)\) if \(x \geq \bar{\ell}_t^1 (m_t, s_t)\). We only need to prove \(x + \gamma (\Delta_t(m_t, s_t))^{-1} < \ell_t^1(x, m_t, s_t)\) for all \(x < \bar{\ell}_t^1 (m_t, s_t)\).

Next, we prove (44) by discussing two different cases. In the rest of the proof we let \(\bar{x} = \bar{\ell}_t^1 (m_t, s_t)\).

\textbf{Step 1}. In this step, we prove (44) when \(\bar{x} < \ell_t^1(\bar{x}, m_t, s_t) - \gamma (\Delta_t(m_t, s_t))^{-1}\).

According to the definition of \(\ell_t^1(x, m_t, s_t)\) in Eq. (26), for \(x_t = \bar{x}\) and any \(x_{t+1} \in [\bar{x} + \gamma (\Delta_t(m_t, s_t))^{-1}, \ell_t^1(\bar{x}, m_t, s_t)]\), we have
\[
p_t(s_t) / \gamma + q / \gamma + \partial_x^+ J_{t+1 \mid s_t} (x_{t+1}, m_t) + \partial_m^+ J_{t+1 \mid s_t} (x_{t+1}, m_t) / \gamma \geq 0, \tag{48}
\]
where \( \bar{m} = m_{t+1}(\bar{x}, m_t, s_t, x_{t+1}) \) is defined in Eq. (25).

For any \( \bar{x} < x \) and \( x_{t+1} \in [\bar{x} + \gamma (\Delta t(m_t, s_t))^{-1}, \bar{x} + \ell_1^t(\bar{x}, m_t, s_t) - \bar{x}] \), let \( \bar{x}_{t+1} = x_{t+1} - \bar{x} \). According to (25), since \( x_{t+1} - \bar{x} = x_{t+1} - \bar{x}, m_{t+1}(\bar{x}, m_t, s_t, x_{t+1}) = m_{t+1}(\bar{x}, m_t, s_t, x_{t+1}) = \bar{m} \). It follows from the convexity of \( J_{t+1}(x, m, s) \) in \( (x, m) \) (cf. Lemma 1) that given \( m \), \( \partial^+_m J_{t+1}(x, m) \) is non-decreasing in \( x \). Since the RHS of Bellman equation (12) is convex in \( a_t \), it is straightforward to check (through backward induction) that \( \partial^+_m J_{t+1}(x, m) \) is non-decreasing in \( x \) given any \( m \). Since \( x_{t+1} > x_{t+1} \), we have

\[
\frac{p_t(s_t)}{\gamma} + q + \partial^+_m \tilde{J}_{t+1}(x_{t+1}, \bar{m}) \geq \frac{p_t(s_t)}{\gamma} + q + \partial^+_m \tilde{J}_{t+1}(x_{t+1}, \bar{m}) + \frac{\partial^+_m \tilde{J}_{t+1}(x_{t+1}, \bar{m})}{\gamma},
\]

where the second inequality follows from (48).

Since the above inequalities are true for any \( x_{t+1} \in [\bar{x} + \gamma (\Delta t(m_t, s_t))^{-1}, \bar{x} + \ell_1^t(\bar{x}, m_t, s_t) - \bar{x}] \), it follows from the definition of \( \ell_1^t(x, m_t, s_t) \) in Eq. (26) that

\[
\bar{x} + \gamma (\Delta t(m_t, s_t))^{-1} < \bar{x} + \ell_1^t(\bar{x}, m_t, s_t) - \bar{x} \leq \ell_1^t(\bar{x}, m_t, s_t).
\]  

Hence, for every \( \bar{x} < x \), we have \( \bar{x} + \gamma (\Delta t(m_t, s_t))^{-1} < \ell_1^t(\bar{x}, m_t, s_t) \) in this case.

**Step 2.** In this step, we prove (44) when \( \bar{x} \geq \ell_1^t(\bar{x}, m_t, s_t) - \gamma (\Delta t(m_t, s_t))^{-1} \).

Note that for any \( \bar{x} < x \), according to the definition of \( \ell_1^t(m_t, s_t) \) in Eq. (28), there exists some \( x' \in (\bar{x}, x) \) such that

\[
\bar{x} + \gamma (\Delta t(m_t, s_t))^{-1} < \ell_1^t(x', m_t, s_t).
\]

Since \( \bar{x} < x' \), it follows (49) that

\[
\bar{x} + \gamma (\Delta t(m_t, s_t))^{-1} < \ell_1^t(x', m_t, s_t) - x' \leq \ell_1^t(\bar{x}, m_t, s_t).
\]

Therefore, for every \( \bar{x} < x \), we have \( \bar{x} + \gamma (\Delta t(m_t, s_t))^{-1} < \ell_1^t(\bar{x}, m_t, s_t) \) in this case.

The results in (45), (46), and (47) can be proved through an approach similar to the one we use to prove (44).

**Proof of Lemma 3:** According to Lemma 2, there always exists an optimal action \( a_t^* \) that minimizes the right-hand side of the Bellman equation (12).

For the case with \( d_t(s) = m_t \), we argue that \( \ell_1^t(m, s) \leq h_2^t(m, s) \). Suppose not, and it follows from Lemma 4 that for a given \( x_t \in (h_2^t(m, s), \ell_1^t(m, s)) \), we must have \( x_t \in (h_2^t(x_t, m, s), \ell_1^t(x_t, m, s)) \), which implies that \( h_2^t(x_t, m, s) < \ell_1^t(x_t, m, s) \). However, according to (27) in Remark 2, we always have \( h_2^t(x_t, m_t, s_t) \geq \ell_1^t(x_t, m_t, s_t) \). The desired result then follows.

The other two cases in Lemma 3 can be proved through a similar approach based on Lemma 4 and Eq. (27).

**APPENDIX D**

**PROOF OF THEOREM 1**

Theorem 1 follows from Lemma 4 and the following Lemma 5. In Lemma 5, we first characterize an optimal storage operation policy using the functionals defined in Eq. (26). The optimal policy characterized in Lemma 5 is not easy to compute because one has to compare the the current storage level \( x_t \) with a functional that depends on \( x_t \), e.g., \( \ell_1^t(x_t, m_t, s_t) \). Combining Lemmas 5 and 4, we will establish the threshold policy characterized in Theorem 1 at the end of this appendix (using the thresholds defined in (28) that do not depend on the current storage level \( x_t \)).

**Lemma 5:** Given the current system state \((x_t, m_t, s_t)\) at stage \( t \in \{0, 1, \ldots, T-1\} \), an optimal policy is characterized in the following.

If \( d_t(s_t) = m_t \), an optimal action is given by

\[
a_t^* = \begin{cases}
\min \left\{ \ell_1^t - x_t, R^C \right\}, & \text{if } x_t < \ell_1^t, \\
0, & \text{if } x_t \in [\ell_1^t, h_2^t], \\
\max \left\{ h_2^t - x_t - R^D \right\}, & \text{if } x_t > h_2^t.
\end{cases}
\]  

(50a)

(50b)

(50c)

If \( d_t(s_t) > m_t \), an optimal action is given by

\[
a_t^* = \begin{cases}
\min \left\{ \ell_1^t - x_t, R^C \right\}, & \text{if } x_t < \ell_1^t, \\
0, & \text{if } x_t \in [\ell_1^t, h_2^t], \\
\max \left\{ h_2^t - x_t - R^D \right\}, & \text{if } x_t \in (h_1^t, h_2^t + \Delta_t/\eta], \\
\max \left\{ h_2^t - x_t - R^D \right\}, & \text{if } x_t > h_2^t + \Delta_t/\eta.
\end{cases}
\]  

(51a)

(51b)

(51c)

(51d)

If \( 0 \leq d_t(s_t) < m_t \), an optimal action is given by

\[
a_t^* = \begin{cases}
\min \left\{ \ell_1^t - x_t, R^C \right\}, & \text{if } x_t < \ell_1^t + \gamma \Delta_t, \\
\min \left\{ \ell_1^t - x_t, R^C \right\}, & \text{if } x_t \in (\ell_1^t + \gamma \Delta_t, \ell_1^t), \\
0, & \text{if } x_t \in [\ell_1^t, h_2^t], \\
\max \left\{ h_2^t - x_t - R^D \right\}, & \text{if } x_t > h_2^t.
\end{cases}
\]  

(52a)

(52b)

(52c)

(52d)

If \( d_t(s_t) < 0 \), an optimal action is given by

\[
a_t^* = \begin{cases}
\min \left\{ \ell_1^t - x_t, R^C \right\}, & \text{if } x_t < \ell_1^t + \gamma \Delta_t, \\
\min \left\{ \ell_1^t - x_t, R^C \right\}, & \text{if } x_t \geq \ell_1^t + \gamma \Delta_t.
\end{cases}
\]  

(53a)

(53b)

**Proof:** We will prove the lemma in four steps, using the first-order conditions in (40)-(43). In each step we consider one of the four cases in the lemma’s statement.

**Step 1.** For the case with \( d_t(s_t) = m_t \), we prove (50b) in Step 1.1, (50a) in Step 1.2, and (50c) in Step 1.3.

**Step 1.1.** When \( d_t(s_t) = m_t \), we argue that it is optimal to maintain the storage level, i.e., \( a_t^* = 0 \), if the current storage level \( x_t \in [\ell_1^t, h_2^t] \). According to the definitions of \( \ell_1^t \) and \( h_2^t \) in (26), the following conditions hold for \( x_t \in [\ell_1^t, h_2^t] \),

\[
8For notational conciseness, we drop \( (x_t, m_t, s_t) \) from the expressions of \( \ell_1^t, \ell_2^t, h_2^t \) and \( \Delta_t \) respectively, because the system state \((x_t, m_t, s_t)\) is fixed within this lemma.
where $x_{t+1} = x_t$ and $m_{t+1}$ is determined by (25). Under the conditions in (54), the action $a_t^* = 0$ satisfies the first-order conditions in (40), and is therefore optimal.

Step 1.2. When $d_s(s_t) = m_t$ and $x_t < \ell_1^t$, according to the definition of $\ell_1^t$ in (26), we have

$$p_t(s_t)/\gamma + q/\gamma + \partial_x^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) + \partial_m^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) \geq 0,$$

$$\eta p_t(s_t) + \partial_x^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) \leq 0,$$

for all $x_{t+1} = x_t + a_t \in \{x_t, \ell_1^t\}$, where $m_{t+1}$ is determined by (25). We also have

$$p_t(s_t)/\gamma + q/\gamma + \partial_x^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) + \partial_m^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) > 0,$$

for all $x_{t+1} = x_t + a_t \in \{x_t, \ell_1^t\}$, where $m_{t+1}$ is determined by (25). The inequalities in Eqs. (55) and (56) imply that the right hand side of the Bellman equation (12) is decreasing in $a_t$ over $[0, \ell_1^t - x_t)$ and is non-decreasing in $a_t$ over $[\ell_1^t - x_t, B - x_t]$. As a result, it is optimal to greedily charge the storage to level $\ell_1^t$ subject to the maximum charging rate constraint, as claimed in (50a).

Step 1.3. We prove (50c) using an approach similar to that used in Step 1.2. When $d_s(s_t) = m_t$ and $x_t > h_2^t$, according to the definition of $h_2^t$ in (26), we have

$$\eta p_t(s_t) + \partial_x^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) > 0,$$

for all $x_{t+1} = x_t + a_t \in (h_2^t, x_t)$, where $m_{t+1}$ is determined by (25). We also have

$$\eta p_t(s_t) + \partial_x^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) \leq 0,$$

for all $x_{t+1} = x_t + a_t \in [0, h_2^t)$, where $m_{t+1}$ is determined by (25). The inequalities in (57) and (58) imply that the right hand side of the Bellman equation (12) is increasing in $a_t$ over $[h_2^t - x_t, 0]$ and is non-increasing in $a_t$ over $[-x_t, h_2^t - x_t]$. It is therefore optimal to greedily discharge the storage to level $h_2^t$, subject to the maximum discharging rate constraint, as claimed in (50c).

Step 2. For the case with $d_t(s_t) > m_t$, we will prove (51b) in Step 2.1, (51a) in Step 2.2, (51c) in Step 2.3, and (51d) in Step 2.4.

Step 2.1. When $d_t(s_t) > m_t$, we argue that it is optimal to maintain the storage level, i.e. $a_t^* = 0$, if $x_t \in [\ell_1^t, h_2^t)$. According to the definitions of $\ell_1^t$ and $h_2^t$ in (26), the following conditions hold for all $x_t \in [\ell_1^t, h_2^t)$,

$$p_t(s_t)/\gamma + q/\gamma + \partial_x^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) + \partial_m^+ J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) \geq 0,$$

$$\eta p_t(s_t) + \eta q + \partial_x^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) + \partial_m^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) \leq 0,$$

where $x_{t+1} = x_t$ and $m_{t+1}$ is determined by (25). Under the condition in (59), the action $a_t^* = 0$ satisfies the first-order conditions in (41), and is therefore optimal.

Step 2.2. (51a) can be proved through an approach that is analogous to that used in Step 1.2 to prove (50a).

Step 2.3. When $d_t(s_t) > m_t$ and $x_t \in (h_1^t, h_2^t + \Delta_t/\eta)$, according to the definition of $h_1^t$ in (26), we have

$$\eta p_t(s_t) + \eta q + \partial_x^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) + \partial_m^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) > 0,$$

for all $x_{t+1} = x_t + a_t \in (h_1^t, x_t)$, where $m_{t+1}$ is determined by (25).

Further, since $x_t \leq h_1^t + \Delta_t/\eta$, according to the definition of $h_1^t$ in (26), for all $x_{t+1} = x_t + a_t \in (0, h_1^t)$, we have the inequality in (58).

We note that when $x_{t+1} = x_t + a_t \in (x_t - \Delta_t/\eta, x_t]$, the partial derivative of the right hand side of the Bellman equation (12) with respect to $a_t$ is the left hand side of (60), and is the left hand side of (57) and (58) when $x_{t+1} = x_t + a_t \in (0, x_t - \Delta_t/\eta]$.

According to (27), we have $h_2^t \leq h_1^t$. If $h_2^t = h_1^t$, the right hand side of the Bellman equation (12) is increasing in $a_t$ over $[h_1^t - x_t, 0]$ and is non-increasing in $a_t$ over $[-x_t, h_1^t - x_t]$. Thus it is optimal to discharge the storage down to $h_1^t$, subject to the maximum discharging rate constraint, as claimed in (51c).

If $h_2^t < h_1^t$, then according to the definition of $h_1^t$ and $h_2^t$ in (26), for all $x_{t+1} = x_t + a_t \in (h_1^t, h_2^t)$, we have

$$\eta p_t(s_t) + \eta q + \partial_x^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) + \partial_m^- J_{t+1\mid s_t}(x_{t+1}, m_{t+1}) \leq 0,$$

where $m_{t+1}$ is determined by (25).

Therefore, the right hand side of the Bellman equation (12) is increasing in $a_t$ over $[h_1^t - x_t, 0]$ and is non-increasing in $a_t$ over $[-x_t, h_1^t - x_t]$. Thus the optimal action is to discharge the storage down to $h_1^t$, subject to the maximum discharging rate constraint.

Step 2.4. When $d_t(s_t) > m_t$, $x_t > h_1^t$, and $x_t > h_2^t + \Delta_t/\eta$, according to the definition of $h_2^t$ in (26), we have (57) for all $x_{t+1} = x_t + a_t \in (h_1^t, x_t - \Delta_t/\eta)$ and (58) for all $x_{t+1} = x_t + a_t \in [0, h_2^t)$. Therefore, the right hand side of the Bellman equation (12) is non-increasing in $a_t$ over $[-x_t, h_2^t - x_t]$, and is increasing in $a_t$ over $(h_2^t - x_t, -\Delta_t/\eta]$. It is non-increasing in $a_t$ over $(\Delta_t/\eta, 0]$.

In conclusion, the right hand side of the Bellman equation (12) is increasing in $a_t$ over $(h_2^t - x_t, 0]$ and is non-increasing in $a_t$ over $[-x_t, h_2^t - x_t]$. It is therefore optimal to greedily discharge the storage to level $h_2^t$, subject to the maximum discharging rate constraint, as claimed in (51d).

Step 3. We consider the case with $0 \leq d_t(s_t) < m_t$.

When $x_t \in [\ell_1^t, h_2^t)$, the result in (52c) can be proved through an approach similar to that used in Step 1.1 (to prove (50b)).
When \( t_d^l + \gamma \Delta t \leq x_t < \ell^l_t \), the result in (52b) can be proved through an approach similar to that in Step 2.3 (to prove (51c)).

Finally, when \( x_t < \ell^l_t + \gamma \Delta t \), the optimal action in (52a) can be proved using an approach similar to that used in Step 2.4 (to prove (51d)).

**Step 4.** When \( d_t(s_t) < 0 \), for \( a_t \in [0, -\gamma d_t(s_t)] \), the right partial derivative of the right hand side of the Bellman equation (12) with respect to \( a_t \) satisfies

\[
\partial^+_x J_{t+1|s_t}(x_{t+1}, m_{t+1}) \leq 0, \quad (62)
\]

where \( x_{t+1} = x_t + a_t \), and \( m_{t+1} = m_t \) according to (25).

We note that when \( x_{t+1} = x_t + a_t \in \{x_t, x_t - \gamma d_t(s_t)\} \), the right derivative of the right hand side of the Bellman equation (12) with respect to \( a_t \) is the left hand side of (62). Thus the right hand side of the Bellman equation (12) is non-increasing in \( a_t \) over \([0, -\gamma d_t(s_t)]\). As a result, it is optimal to greedily charge all residual renewable generation (negative net demand) to the storage.

If \( x_t \geq \ell^l_t + \gamma \Delta t \), through an approach that is analogous to that used in Step 2.3 (to prove (51c)) we can show that the right hand side of the Bellman equation (12) is decreasing in \( a_t \) over \([-\gamma d_t(s_t), \ell^l_t - x_t]\) and is non-decreasing in \( a_t \) over \([\ell^l_t - x_t, B - x_t]\). Note that we have shown that the right hand side of (12) is non-increasing in \( a_t \) over \([0, -\gamma d_t(s_t)]\). As a result, it is optimal to greedily charge the storage to level \( \ell^l_t \), subject to the maximum charging rate constraint, as claimed in (53b).

If \( x_t < \ell^l_t + \gamma \Delta t \), through an approach that is analogous to that used in Step 2.4 (to prove (51d)), it can be shown that the right hand side of the Bellman equation (12) is decreasing in \( a_t \) over \([-\gamma d_t(s_t), \ell^l_t - x_t]\) and is non-decreasing in \( a_t \) over \([\ell^l_t - x_t, B - x_t]\). Since the right hand side of (12) is non-increasing in \( a_t \) over \([0, -\gamma d_t(s_t)]\), it is optimal to greedily charge the storage to level \( \ell^l_t \), subject to the maximum charging rate constraint, as claimed in (53a).

Next, we use Lemma 4 to strengthen the characterization in Lemma 5 and to prove Theorem 1.

For the case with \( d_t = m_t \), it follows from (44) and (47) that the results in (50a), (50b), and (50c) (established in Lemma 5) are equivalent to (29a), (29b), and (29c) in Theorem 1, respectively.

For the case with \( d_t > m_t \), it follows from (44), (46), and (47) that (51a), (51b), (51c), and (51d) established in Lemma 5 are equivalent to (30a), (30b), (30c) and (30d) in Theorem 1, respectively.

For the case with \( 0 \leq d_t < m_t \), it follows from (44), (45), and (47) that (52a), (52b), (52c), (52d) established in Lemma 5 are equivalent to (31a), (31b), (31c), and (31d) in Theorem 1, respectively.

We now consider the case with \( d_t < 0 \) and \( x_t \in [0, B] \). It follows from the definition of \( \ell^l_t \) in (26) that \( x_t < \min\{B, x_t - \gamma d_t\} \leq \ell^l_t \). We therefore have \( \ell^l_t = B \) according to the definition of \( \ell^l_t \) in (28), regardless of the value of \( x_t \). Since \( x_t < \ell^l_t \) for all \( x_t \in [0, B] \), it follows from (44) and (45) that (53a), (53b) established in Lemma 5 are equivalent to (31a), (31b) in Theorem 1.

Finally, we consider the case with \( d_t < 0 \) and \( x_t = B \). Since \( \ell^l_t = B \), we have \( x_t = \ell^l_t + \gamma \Delta t \geq \ell^l_t \) according to (26), and therefore the optimal action \( a_t^* = \ell^l_t - x_t = 0 \) according to Lemma 5. Since \( x_t = \ell^l_t + \gamma \Delta t \) and \( a_t^* = \ell^l_t - x_t = 0 \), Eq. (53b) established in Lemma 5 is equivalent to (31c) in Theorem 1 for \( x_t = B \). Note that since \( \ell^l_t = B \), it follows Lemma 3 that \( x \leq \ell^l_t = B \leq \gamma^l_t(m, s) \); as a result, the case with \( x_t > \gamma^l_t \) in (31d) of Theorem 1 never happens when \( d_t < 0 \).

We have proved Theorem 1.