Optimal Threshold Control for Energy Arbitrage with Degradable Battery Storage

Marek Petrik
IBM T. J. Watson Research Center
Yorktown, NY 10598
mpetrik@us.ibm.com

Xiaojian Wu
School of Computer Science
University of Massachusetts Amherst
Amherst, MA 01003
xiaojian@cs.umass.edu

Abstract

Energy arbitrage has the potential to make electric grids more efficient and reliable. Batteries hold great promise for energy storage in arbitrage but can degrade rapidly with use. In this paper, we analyze the impact of storage degradation on the structure of optimal policies in energy arbitrage. We derive properties of the battery degradation response that are sufficient for the existence of optimal threshold policies, which are easy to interpret and compute. Our experimental results suggest that explicitly considering battery degradation in optimizing energy arbitrage significantly improves solution quality.

1 INTRODUCTION

Energy storage and arbitrage play an important role in numerous domains including hybrid vehicle propulsion or electric grids [Walawalkar et al., 2007]. An important challenge in this domain is to decide how much energy to store or release based on current and expected future prices. Related resource efficiency optimization problems have been recently studied, among others, in the computational sustainability research area [Gomes, 2009, Petrik and Zilberstein, 2011, Ermon et al., 2011, Ermon et al., 2013]. Since computing an energy arbitrage policy is a difficult sequential *stochastic* optimization problem subject to significant uncertainty, it is particularly relevant to the computational sustainability community.

In this paper, we are concerned with optimizing energy arbitrage under stochastically varying energy prices. The goal of the decision maker is to maximize profits by charging an energy storage device when the price of energy is low and discharging it when it is high. We show that optimal policies have a threshold structure even when *battery degradation* is considered and use this structure to develop a practical algorithm.

Our optimal threshold policy has two price-dependent thresholds l and u ($l \leq u$). If the current state of charge is less than l, then the battery is charged up to l. If the current state of charge is greater than u, then the battery is discharged to u. For any state of charge between l and u, no action is taken. Such threshold structure makes charging policies easy to compute, analyze, and interpret.

The structure of optimal policies in energy arbitrage has been analyzed previously without accounting for battery degradation [Lifshitz and Weiss, 2014, Nadarajah, 2014, Van de Ven et al., 2011, Harsha and Dahleh, 2011]. In case of batteries—an important energy storage device—the degradation can be significant and often represents an important limitation due to the high cost of batteries.

Batteries degrade most noticeably by losing capacity to hold charge. Modeling how usage patterns affect different battery types (e.g. NiMH and Li-ion) is an important research problem [Ramadass et al., 2003, Aurbach, 2000]. In Li-ion batteries, the degradation is predominantly influenced by 1) the state of charge, 2) the rate of charge and discharge, and 3) the ambient temperature. We aim, in this work, to minimize the battery's capacity loss as it is influenced by the state of charge.

Battery degradation has been explicitly considered in optimizing energy storage in hybrid vehicles [Bashash et al., 2011, Serrao et al., 2005, Hoke et al., 2011, Moura et al., 2011]. These methods solve a discretized dynamic program. The drawback of this approach is that the computed policies are complex, hard to implement and interpret. In addition, the discretization and sampling issues can significantly degrade solution quality as we show experimentally. On the other hand, the optimality of threshold policies has been studied widely in the inventory management literature [Porteus, 2002]. There is, however, no concept of storage degradation in traditional inventory management domains.

The existence of an optimal threshold policy in our setting is somewhat surprising. The standard threshold policy results rely on the fact that the optimal value function is convex or k-convex. The optimal value function for non-trivial battery degradation functions, as we show, may be non-convex. Yet, we establish the existence of a threshold policy through the convexity of an auxiliary optimization function.

The remainder of the paper is organized as follows. Section 2 describes the overall arbitrage model. We primarily study models that faithfully capture the energy arbitrage setting in electric grids. Section 3 describes the model of battery degradation and derives some basic properties of the degradation function. Then, we show sufficient conditions for the existence of threshold policies in Section 4 and describe how the structure can be used in computing an optimized policy in Section 5. Section 6 describes an application of the methodology with the analysis of the results and a comparison to discretization-based methods.

2 ENERGY STORAGE MODEL

This section describes the model of energy arbitrage and storage. We assume multiple finite *known* price levels and a stochastic evolution given a limited storage capacity. In particular, the storage is assumed to be an electrical battery that degrades when energy is stored or retrieved.

The underlying model is a Markov decision process. We assume a discrete-time problem with either a finite horizon T or a discounted infinite horizon. Prices are governed by a Markov process with states Θ. There are two energy prices in each time step: $p^i : \Theta \to \mathbb{R}_+$ is the purchase (or input) price and $p^o : \Theta \to \mathbb{R}_+$ is the selling (or output) price. To simplify notation, the difference in these prices is also used to model the energy loss in the charging and discharging processes. In other words, the prices measure the cost of energy as added or subtracted from the storage and may not actually be sold or purchased.

We use s to denote the available battery capacity with s_0 denoting the initial capacity. The current state of charge is denotes as x or y and must satisfy that $0 \leq x_t \leq s_t$ at any time step t. The action is the amount of energy to charge or discharge, which is denoted by u. Positive u indicates that energy is purchased to charge the battery; negative u indicates the sale of energy.

We will make the following assumption regarding the purchase and selling prices.

Assumption 1. Purchase price is higher than the selling price per unit in a time step:

$$p^i_\theta \geq p^o_\theta \quad \forall \theta \in \Theta \,.$$

Assumption 1 is virtually always satisfied in practice. If violated, direct arbitrage by simultaneously purchasing and selling energy in a single time step would then equalize the prices. In addition, the purchase price p^i will be greater than the selling price p^o due to the inefficiencies involved in charging and discharging.

As mentioned above, the focus of the paper is on degradation of battery capacity as a function of its use. In particular, we model the degradation as a function of the battery capacity when charged or discharged. We use a general model of battery degradation with a specific focus on Li-ion batteries. The degradation function $d(x,u) \in \mathbb{R}_+$ represents the battery capacity loss after starting at the state of charge $x \geq 0$ and charging (discharging if negative) by u with $-x \leq u \leq s_0 - x$. This function indicates the loss of capacity, such that:

$$s_{t+1} = s_t - d(x_t, u_t)$$

We discuss the degradation function in more detail in Section 3.

Our model makes several simplifying assumptions that are reasonable in an electric grid scenario, but may not apply to other scenarios such as a hybrid vehicle battery storage. In particular, we assume that the purchase and selling prices are independent of the energy quantity sold or purchased and battery degradation is independent of current and temperature.

The state set in the Markov decision problem is composed of (x, s, θ) where x is the state of charge, s is the battery capacity, and $\theta \in \Theta$ is the state of the price process. The available actions in a state (x, s, θ) are u such that $-x \leq u \leq s - x$. The transition is from (x_t, s_t, θ_t) to $(x_{t+1}, s_{t+1}, \theta_{t+1})$ given action u_t is:

$$\begin{aligned} x_{t+1} &= x_t + u_t \\ s_{t+1} &= s_t - d(x_t, u_t) \end{aligned}$$

The probability of this transition is given by $\mathbb{P}\left[\theta_{t+1} | \theta_t\right]$. The reward for this transition is:

$$r((x_t, s_t, \theta_t), u_t) = \begin{cases} -u_t \cdot p^i - c^d \cdot d(x_t, u_t) & \text{if } u_t \geq 0 \\ -u_t \cdot p^o - c^d \cdot d(x_t, u_t) & \text{if } u_t < 0 \end{cases} .$$

That is, the reward captures the monetary value of the transaction minus a penalty for degradation of the battery. Here, c^d represents the cost of a unit of lost battery capacity.

The solution of the Markov decision process is a policy π, which can be computed from a value function v and a post-decision (or state-action) value function q. We focus on both the discounted infinite horizon with a discount factor $\lambda \in (0, 1)$ and the finite horizon with the undiscounted total return criterion.

The Bellman optimality equations for this problem are:

$$\begin{aligned} q_T(x, s, \theta_T) &= 0 \\ v_t(x, s, \theta_t) &= \min\big\{ p^i_{\theta_t} \left[u\right]_+ + p^o_{\theta_t} \left[u\right]_- + \\ &\quad + c^d \, d(x, u) + \\ &\quad + q_t(x + u, s - d(x, u), \theta_t) \;: \\ &\quad :\; u \in [-x, s - x] \big\} \\ q_t(x, s, \theta_t) &= \lambda \cdot \mathbb{E}\left[v_{t+1}(x, s, \theta_{t+1})\right] \end{aligned} \tag{2.1}$$

where $[u]_+ = \max\{u, 0\}$ and $[u]_- = \min\{u, 0\}$ and the expectation is taken over $P(\theta_{t+1} | \theta_t)$. For finite horzion case, $\lambda = 1$.

For the purpose of our theoretical analysis, we assume that the lost capacity is immediately replaced and therefore the battery capacity does not actually change ($s_{t+1} = s_t = s_0$). Instead, the degradation induces a penalty in the form of capacity replacement cost governed by c_d. In that case, the capacity s can be omitted from the definition of v in (2.1). The experimental results, however, study the setting in which the capacity is not immediately replaced.

3 BATTERY DEGRADATION FUNCTION

This section describes properties of the degradation function $d(x, u)$ that can capture the behavior of Li-ion batteries [Aurbach, 2000, Ramadass et al., 2003] and can guarantee the existence of an optimal threshold policy. We focus on Li-ion batteries because of their ubiquity and considerable promise in future applications [Peterson et al., 2010]. The chemical processes in other battery types—such as NiMH and NiCd—are often quite different [Serrao et al., 2005].

As noted above, we consider the dependence of the degradation only on the state of charge because the other variables, such as the temperature and the current, can be efficiently controlled in an electric grid energy application.

In broad terms, a Li-ion battery degrades significantly while the state of charge is either very low or very high. A naive policy that minimizes battery degradation will therefore attempt to use the battery as close to approximately 50% state of charge as possible.

Instead of considering a single degradation function, we define a class of functions $d(x, u)$, *continuous* in u for $x \in [0, s_0]$, that in addition satisfy the following properties.

- A1 *Convexity:* function $d(x, u)$ is convex in u for all $x \in [0, s_0]$.
- A2 *Memorylessness:* $d(x, u_1 + u_2) = d(x, u_1) + d(x + u_1, u_2)$ when $\operatorname{sgn}(u_1) = \operatorname{sgn}(u_2)$.
- A3 *Cycle linearity:* function $d(0, u) + d(u, -u)$ is *linear* in u.

Next we informally describe the meaning of the properties above; a more detailed analysis of functions that satisfy these properties follows later. The property A1 is generally satisfied in Li-ion batteries in which the degradation is more severe near the extreme range of the state of charge. The property A2 requires that the degradation due to the charging is independent of the amount charged, but instead depends only on the current state of charge. Finally, the property A3 requires that the degradation due to charging an arbitrary amount from the state of charge 0 and subsequently fully discharging is linear in the amount charged.

The assumptions above, unfortunately, are hard to grasp intuitively and therefore difficult to justify. To elucidate the definitions, consider the alternative definition of

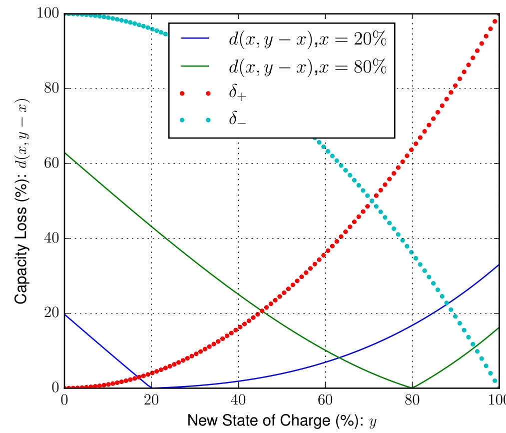

Figure 1: Example degradation function with $\delta_+(x) = x^2$ and $\delta_-(x) = 1 - x^2$.

$d(x, u)$ based on the degradation at any state of charge for an infinitesimally small amount of energy charged $\delta_+ : [0, s_0] \to \mathbb{R}_+$ or discharged $\delta_- : [0, s_0] \to \mathbb{R}_+$. The degradation function is then simply defined as the following integral:

$$d(x, u) = \begin{cases} \int_x^{x+u} \delta_+(y)\, dy & \text{if } u \geq 0 \\ \int_{x+u}^{x} \delta_-(y)\, dy & \text{if } u < 0 \end{cases}. \tag{3.1}$$

Fig. 1 depicts an example of the immediate degradation functions δ_+ and δ_- and the corresponding degradation function d for two values of the current state of charge.

The following proposition describes how the properties of δ_+ and δ_- translate to properties of $d(x, u)$.

Proposition 3.1. *When d is defined as in* (3.1) *and*

(i) both δ_- and δ_+ are continuous on $[0, s_0]$
(ii) δ_+ is nondecreasing and δ_- is nonincreasing
(iii) $\delta_+(y) + \delta_-(y)$ is a constant for any $y \in [0, s_0]$

Then, the battery degradation function $d(x, u)$ satisfies the properties of A1, A2 and A3.

Proof. The property (i) implies that there exist antiderivatives D_+ and D_- to δ_+ and δ_- on the appropriate intervals. Then:

$$d(x, u) = \begin{cases} D_+(x + u) - D_+(x) \text{ if } u \geq 0 \\ D_-(x) - D_-(x + u) \text{ if } u < 0 \end{cases}$$

We prove each property individually.
A1: Convexity. The convexity for a fixed x and $u \neq 0$ follows directly from the convexity of D_+ and $-D_-$. The functions $D_+(x, u)$ and $D_-(x, u)$ are convex because their first derivatives δ_+ and $-\delta_-$ are nondecreasing [Boyd and Vandenberghe, 2004]. Since both D_+ and D_- are non-negative functions and $d(x, 0) = 0$, we have

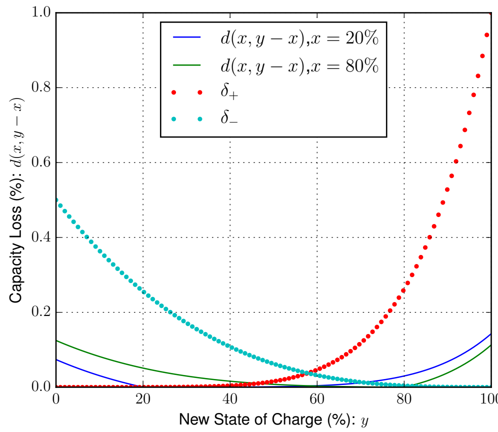

Figure 2: Example degradation function with $\delta_+(x) = 0.01 \cdot x^6$ and $\delta_-(x) = 0.005 \cdot (1-x)^3$.

that $u = 0$ is the minimum of $d(x, u)$ and therefore it satisfies the convexity condition.

A2: Memorylessness. If $u_1 > 0$ and $u_2 > 0$,

$$d(x, u_1 + u_2) = \int_x^{x+u_1+u_2} \delta_+(y)dy$$
$$= \int_x^{x+u_1} \delta_+(y)dy + \int_{x+u_1}^{x+u_1+u_2} \delta_+(y)dy$$
$$= d(x, u_1) + d(x + u_1, u_2).$$

The property holds similarly when $u_1 < 0$ and $u_2 < 0$.

A3: Cycle linearity. Let $u > 0$,

$$d(0, u) + d(u, -u) = \int_0^u \delta_+(y)dy + \int_0^u \delta_-(y)dy$$

since $\delta_+(y) + \delta_-(y)$ is a constant, the cycle linearity is satisfied. □

Note that the degradation function in Fig. 1 satisfies the properties A1–A3. This can be easily seen since $\delta_+(x) + \delta_-(x) = 1$. However, the degradation of Li-ion batteries may not always satisfies these properties. Fig. 2 illustrates an example, based on real Li-ion behavior [Bashash et al., 2011], which violates A3. The precise form of the degradation function for any particular battery design can be obtained either experimentally or by simulation [Ramadesigan et al., 2012].

4 STRUCTURE OF OPTIMAL POLICIES

In this section, we show the existence of an optimal threshold policy in the battery storage problem if properties A1–A3 are satisfied. The analysis is based on a finite-horizon version of the problem and we discuss how this structure generalizes to discounted infinite horizon problems later in the section.

A two-threshold charge policy is defined as follows:

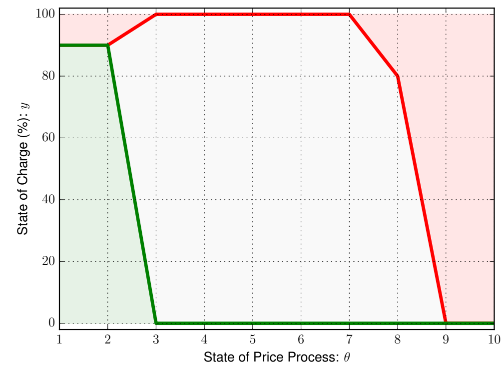

Figure 3: Example of a threshold policy. The upper (red) line represents $C(\theta)$; the lower (green) line represents $c(\theta)$.

Definition 4.1. A two-threshold charge policy with thresholds $(c_{t,\theta}, C_{t,\theta})$ with $c_{t,\theta} \le C_{t,\theta}$ for some $\theta \in \Theta$ and $t = 1 \dots T$ is defined as:

$$u_t = \begin{cases} c_{t,\theta} - x_t & \text{when } x_t \le c_{t,\theta} \\ C_{t,\theta} - x_t & \text{when } x_t \ge C_{t,\theta} \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

where t is the current time step, x_t is the current battery charge, θ_t is the price level state, and u_t is the change in the state of charge.

One of the main appealing properties of a threshold policy is its simplicity and interpretability. Fig. 3 depicts an example of a threshold policy. The x-axis represents the state of the price process θ. In this example, the price of energy grows linearly with θ and the price transitions behave as a martingale. If the current state of charge is in the red region, the next step is to discharge the battery to the red line. Similarly, states in the green region are charged up to the green line.

Note that the policy in Fig. 3 behaves very intuitively. When the price of energy is low (small θ), the battery is charged to a high level. It is not charged fully, however, to prevent excessive degradation of capacity. No action is taken for the medium energy price. When the energy price increases to its maximum level, the battery is fully discharged.

We are now ready to state the main theoretical result of the paper.

Theorem 4.2. *Assume that the battery degradation function satisfies properties A1, A2, and A3 and $\lambda = 1$. Then, there exists an optimal two-threshold charge policy for the finite horizon problem with some time-dependent thresholds $(c(t, \theta), C(t, \theta))$.*

To prove the theorem, we need to show several auxiliary properties. The following lemma describes the properties of the degradation function induced by the assumptions above.

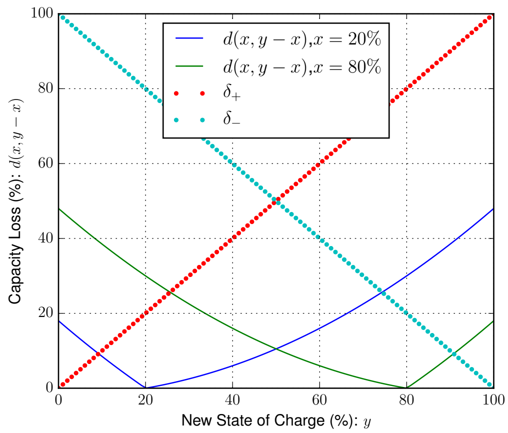

Figure 4: Degradation function in Example 4.4.

Lemma 4.3. *A degradation function d that satisfies property A2 also satisfies:*

(i) $d(x, 0) = 0$

(ii) $d(x, y-x) = d(0, y) - d(0, x)$ *when* $y \geq x$

(iii) $d(x, y-x) = d(x, -x) - d(y, -y)$ *when* $y \leq x$

In addition, when d satisfies property A1, then:

(iv) $d(x, -x)$ *is concave in* x

(v) $d(x, y-x) = \max\{d(0, y) - d(0, x), d(x, -x) - d(y, -y)\}$

Proof. The lemma follows by simple algebraic manipulation for the individual cases as follows.
Case (i):

$$d(x, u + 0) = d(x, 0) + d(x + 0, u) .$$

Case (ii):

$$d(0, y) = d(0, x + (y - x)) = d(0, x) + d(x, y - x) .$$

Case (iii):

$$d(x, -x) = d(x, (y - x) - y) = d(x, y - x) + d(y, -y) .$$

Case (iv): by rewriting $d(s_0, -s_0) = d(s_0, -(s_0 - x) - x)$ we get:

$$d(x, -x) = d(s_0, -s_0) - d(s_0, x - s_0) .$$

Function $d(x, -x)$ is concave because $d(s_0, x - s_0)$ is convex.
Case (v): Note that $d(0, x)$ is non-decreasing, and $d(x, -x)$ is non-increasing. The proof then readily follows from properties (ii) and (iii). □

We are now ready to prove the main result. The typical proof of the threshold property in inventory management settings is based on convexity of the value function. Unfortunately, as the example described below in Example 4.4

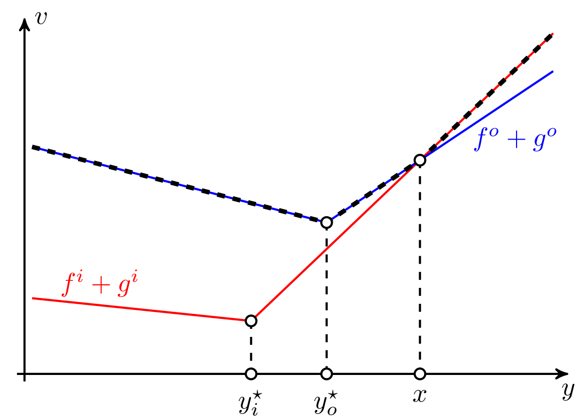

Figure 5: Example of functions in the proof of Theorem 4.2. The current charge is x and the optimal action is to discharge to $y_o^\star$.

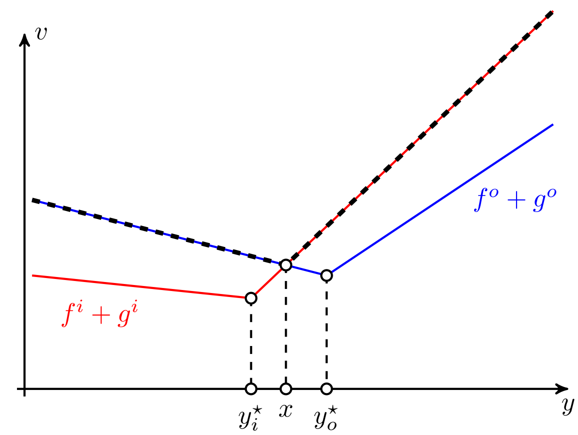

Figure 6: Example of functions in the proof of Theorem 4.2. The current charge is x and the optimal action is to leave the charge unchanged.

demonstrates, the value function in our setting may be non-convex. Instead, we use property A3 to establish the convexity of the post-decision value function (the concept is similar to q-function in reinforcement learning). In particular, the linearity required by A1 can be used to cancel out the value function non-convexities due to the battery degradation.

Proof of Theorem 4.2. Like most structural proofs in dynamic programming, the proof is based a backward induction on time steps. But first, we need to derive a more convenient representation of the optimality equation (2.1). To that effect, rewrite the term $p^i_\theta\,[u]_+ + p^o_\theta\,[u]_-$ in (2.1) as a maximum over two functions:

$$p^i_\theta\,[u]_+ + p^o_\theta\,[u]_- = \max\{p^i_\theta u,\ p^o_\theta u\} .$$

This holds from Assumption 1. It will be more convenient to change the optimization variable in (2.1) from the charge difference u to the new state of charge $y = x + u$. Also using property (v) from Lemma 4.3 to express the degradation function, the optimality expression for $v_t(x, \theta)$ now reads as:

$$\min_{y\in[0,s_0]} \max\Big\{p^i_\theta(y-x), p^o_\theta(y-x)\Big\} + q_t(y,\theta) + \\ +c^d \cdot \max\Big\{d(0,y) - d(0,x), d(x,-x) - d(y,-y)\Big\}.$$

Properties (ii) and (iii) from Lemma 4.3 and Assumption 1 imply that the two max operators attain their respective first terms if and and only if $y \geq x$ and therefore can be merged:

$$\min_{y\in[0,s_0]} \max\Big\{p^i_\theta(y-x) + c^d(d(0,y) - d(0,x)) + q_t(y,\theta), \\ p^o_\theta(y-x) + c^d(d(x,-x) - d(y,-y)) + q_t(y,\theta)\Big\}$$

Then we replace $\max\{a, b\}$ by $\max_{\xi\in[0,1]}\{\xi a, (1-\xi)b\}$. Thus the Bellman optimality conditions become:

$$v_t(x,\theta) = \min_{y\in[0,s_0]} \max_{\xi\in[0,1]} \phi_t(y,\xi,x,\theta) \ ,$$

where the main objective function ϕ is defined as:

$$\phi_t(y,\xi,x,\theta) = \xi(f^i_t(y,\theta) + g^i_t(x,\theta)) + \\ +(1-\xi)(f^o_t(y,\theta) + g^o_t(x,\theta)) \ .$$

The functions g^i and g^o represent terms that are constant with respect to the optimization variable y:

$$g^i_t(x,\theta) = -p^i_\theta x - c^d d(0,x) \\ g^o_t(x,\theta) = -p^o_\theta x + c^d d(x,-x) \ .$$

On the other hand, the functions f^i and f^o represent terms that depend on the optimization variable y:

$$f^i_t(y,\theta) = p^i_\theta y + c^d d(0,y) + q_t(y,\theta) \\ f^o_t(y,\theta) = p^o_\theta y - c^d d(y,-y) + q_t(y,\theta) \ .$$

The remainder of the proof focuses on showing that f^i and f^o are convex and using this convexity to show the optimality of a threshold policy in computing the minimization over y.

We next show by induction on t from $t = T$ to $t = 0$ that the functions f^i_t and f^o_t are *convex* even if q_t is not. To prove the base case $t = T$ recall that $q_T = 0$. The convexity of f^i_t and f^o_t then follows from property (iv) in Lemma 4.3.

To prove the inductive step, assume that the functions f^i_{t+1} and f^o_{t+1} are *convex*. Our focus is on showing convexity of f^i_t; the derivation of convexity of f^o_t is analogous. Recall that:

$$f^i_t(y,\theta) = p^i_\theta y + c^d d(0,y) + \mathbb{E}\left[v_{t+1}(x,\theta_{t+1})\right]$$

Since given a fixed x, the function $\phi(y, \xi, x)$ is convex–concave, continuous, and optimized on convex compact sets we have by a generalized minimax theorem (e.g. [Sion, 1958]) and further algebraic manipulation that:

$$\begin{aligned} v_{t+1}(x,\theta) &= \min_{y\in[0,s_0]} \max_{\xi\in[0,1]} \phi_{t+1}(y,\xi,x,\theta) \\ &= \max_{\xi\in[0,1]} \min_{y\in[0,s_0]} \phi_{t+1}(y,\xi,x,\theta) \\ &= \max\Big\{g^i_{t+1}(x,\theta) + \min_{y\in[0,s_0]} f^i_{t+1}(y,\theta), \\ &\qquad g^o_{t+1}(x,\theta) + \min_{y\in[0,s_0]} f^o_{t+1}(y,\theta)\Big\}, \end{aligned}$$

Note that the functions g^i_{t+1} and g^o_{t+1} may not be convex, but taking the maximum outside of the minimization will help to establish the convexity of f^i_t.

Next, plug in the above expression for v_{t+1} and the definitions of g^i_{t+1} and g^o_{t+1} to $f^i_t(y,\theta)$. Further algebraic simplification yields:

$$\begin{aligned} &f^i_t(y,\theta) = c^d d(0,y) + \\ &+\mathbb{E}\left[\max\{-c^d d(0,y) + L_1(y), c^d d(y,-y) + L_2(y)\}\right] \\ &= \mathbb{E}\left[\max\{L_3(y), c^d d(0,y) + c^d d(y,-y) + L_4(y)\}\right] \end{aligned}$$

where $L_1(y) \ldots L_4(y)$ represent functions that are linear or constant in y. Now, recall from Lemma 4.3 that the function $d(y, -y)$ is concave. However, Assumption A3 implies that $d(0, y) + d(y, -y)$ is a linear (convex) term. Point-wise maximization and expectation preserve convexity and thus the function f^i_t is convex. The analogous proof for f^o_t requires that $d(0, y)+d(y, -y)$ is concave (or linear) and thus the linearity required by A3.

The final step in the proof is to show the two-threshold structure in the solution to the action optimization problem:

$$\min_{y\in[0,s_0]} \max\Big\{f^i_t(y,\theta) + g^i_t(x,\theta), \\ f^o_t(y,\theta) + g^o_t(x,\theta)\Big\} \ . \tag{4.2}$$

Figs. 5 and 6 depict examples of functions in (4.2) for two different values of the current state of charge x. The horizontal axis represents the next state of charge y and the bold dashed line highlights the maximum of the two functions. Note that the change in the current state of charge x only shifts the two functions without changing their shape.

Now, recall that from Lemma 4.3 and from the construction of (4.2), the following inequalities hold:

$$\begin{aligned} y > x \Rightarrow f^i_t(y,\theta) + g^i_t(x,\theta) \ &\geq \ f^o_t(y,\theta) + g^o_t(x,\theta) \\ y < x \Rightarrow f^i_t(y,\theta) + g^i_t(x,\theta) \ &\leq \ f^o_t(y,\theta) + g^o_t(x,\theta) \\ y = x \Rightarrow f^i_t(y,\theta) + g^i_t(x,\theta) \ &= \ f^o_t(y,\theta) + g^o_t(x,\theta) \end{aligned}$$

Now let $e_i(y) = f^i_t(y,\theta) + g^i_t(x,\theta)$ and $y^\star_i \in \arg\min_{y\in[0,s_0]} e_i(y)$. Also let $e_o(y) = f^o_t(y,\theta) + g^o_t(x,\theta)$

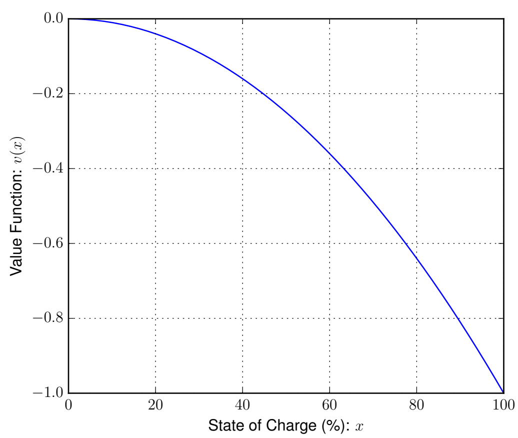

Figure 7: Value function at time $t = 1$.

and $y_o^\star \in \arg\min_{y\in[0,s_0]} e_o(y)$. We will show below that the optimal solution to (4.2) can be either at $y_i^\star$, $y_o^\star$, or at x. There are three possible cases:

1. If $y_i^\star > x$ then $y_i^\star$ is optimal in (4.2) because $f_t^i(y,\theta) + g_t^i(x,\theta) \leq f_t^o(y,\theta) + g_t^o(x,\theta)$ for any $y < x$.
2. If $y_o^\star < x$ then $y_o^\star$ is optimal in (4.2) as above.
3. If $y_i^\star \leq x$ and $y_o^\star \geq x$ then x is optimal in (4.2). Barring the trivial case $y_i^\star = x$, for any $y > x$, we have $\frac{e_i(y)-e_i(x)}{y-x} \geq \frac{e_i(x)-e_i(y_i^\star)}{x-y_i^\star} \geq 0$ since $e_i(y)$ is convex. Therefore $e_i(y) \geq e_i(x)$ for any $y > x$. A similar argument for any $y < x$ shows that $e_o(y) \geq e_o(x)$ which proves the desired optimality of x in (4.2).

These cases correspond to the threshold policy as described in Definition 4.1. □

The following example illustrates that even when the conditions A1–A3 are satisfied, the optimal value function may not be convex.

Example 4.4. *Consider a two stage problem with the sale price $p^i = p^o = 2$, the degradation cost $c^d = 1$ and the degradation function being:*

$$d(x,u) = \max\{(x+u)^2 - x^2, (s_0 - x - u)^2 - (s_0 - x)^2\}.$$

The degradation function at $x = 0.2$ and $x = 0.8$ is depicted in Fig. 4.

The battery capacity is $s_0 = 1$. It is easy to show that the optimal policy in the last stage is simply to fully discharge the battery. Then $v_1(x) = -p^o \cdot x + c^d \cdot d(x,-x)$. By Lemma 4.3, the function v_1 is concave *as illustrated in Fig. 7.*

Note that the degradation function in Example 4.4 satisfies the property A3 and, therefore, this property is insufficient to guarantee the convexity of a value function. In fact, the value function will only be convex only if $d(x,-x)$ is linear—the degradation function is linear only when the degradation is independent of the battery charge.

Finally, note that Theorem 4.2 does not apply to discounted problems. To apply to discounted problems, we require a modified A3 that states that both functions $d(0,u) + \lambda\, d(u,-u)$ and $-\lambda\, d(0,u) - d(u,-u)$ are convex.

Corollary 4.5. *Assume that the battery degradation function satisfies properties A1, A2, and* modified *A3. Then there exists an optimal two-threshold charge policy for the discounted infinite horizon problem with time-independent thresholds $(c(\theta), C(\theta))$.*

The proof of Corollary 4.5 follows the same steps as the proof of Theorem 4.2 and is provided in the appendix. The stationarity follows using the standard argument for the existence of an optimal stationary policy, e.g. Theorem 6.2.7 in [Puterman, 2005].

The modified condition A3 is however more difficult to satisfy and verify than the original A3. However, it may be sufficient to satisfy either one of these properties approximately in order for a threshold policy be close to optimal. This analysis is, however, beyond the scope of the present paper.

5 OPTIMIZATION ALGORITHM

So far we described the structure of the optimal policy. It is important to also develop an algorithm that can take advantage of this structure and efficiently compute the optimal policy. In this section, we describe such an algorithm and prove its optimality. We focus on the infinite horizon problem which is more relevant in practice.

A naive approach to optimizing threshold policies is to iteratively evaluate a given set of thresholds by simulation and then optimize the threshold values. This class of methods is known as simulation-optimization or policy search [Carson and Maria, 1997]. Simply searching over all sets of thresholds is intractable because the number of threshold values that need to be computed is $2|\Theta|$. As we show below this search can be decomposed by the states of the price process θ leading to Algorithm 1. The algorithm optimizes each pair of thresholds independently for each state of the price process.

The evaluation function $\tilde{f}(c_{\theta_1}, C_{\theta_1}, \ldots, c_{\theta_k}, C_{\theta_k}, \ldots, c_{\theta_n}, C_{\theta_n})$ is computed by simulating the execution. Using common random numbers when optimizing a value by simulation in this setting can significantly reduce sample variance and speed up the algorithm [Glasserman and Yao, 1992]. To model the discount factor, we assume a termination probability of $1 - \lambda$ in every step. The function $\tilde{f}$ is computed as a sample average.

In the remainder of the section, assume that the function $\tilde{f}$ can be evaluated precisely. The result readily generalizes to

Algorithm 1: Threshold Optimization by Simulation

// Initialize thresholds
$(c_\theta, C_\theta) \leftarrow (0, 1) \quad \forall \theta \in \Theta$;
// Initialize step counter
$k \leftarrow 1$;
$g_0 \leftarrow \inf, g_{-1} \leftarrow \inf$;
while $g_{k-1} < g_{k-2} + \epsilon$ **do**
  for $\theta \in \Theta$ **do**
    // Optimize thresholds for θ
    $c_\theta, C_\theta \leftarrow \arg\min_{\bar{c}_\theta, \bar{C}_\theta} \tilde{f}(\ldots, \bar{c}_\theta, \bar{C}_\theta, \ldots)$;
    // $\tilde{f}$ is sampled mean return
    $g_k \leftarrow \tilde{f}(c_{\theta_1}, C_{\theta_1}, \ldots, c_{\theta_n}, C_{\theta_n})$;
  $k \leftarrow k + 1$;
// Return computed thresholds
return $\{(\theta, c_\theta, C_\theta) \; : \; \theta \in \Theta\}$

the sampled setting by considering appropriate Hoeffding or Bernstein concentration inequalities.

Proposition 5.1. *Consider an energy arbitrage problem that satisfies the properties A1, A2, and modified A3 sufficient for the optimality of a threshold policy. Then Algorithm 1 converges to the optimal solution in a finite number of iterations.*

Proof. We argue that Algorithm 1 corresponds to a variant of the simplex algorithm implementation on the dual MDP formulation (e.g., [Puterman, 2005]). First, note that using the same argument as in the proof of Theorem 4.2 we can show that the optimal solution to the minimization in Algorithm 1 is also a threshold policy. Therefore, the algorithm corresponds to a coordinate descent on the linear program formulation of the MDP. The result then follows from the finite number of coordinate blocks. □

6 NUMERICAL RESULTS

This section numerically evaluates Algorithm 1 in an idealized, but realistic, model of daily energy price evolution and a degradable Li-ion battery. First, we analyze the properties of the computed solution and study the impact of the battery degradation on the quality of the computed policy. Then, we compare the solution based on a threshold policy to directly solving a discretized version of the problem.

The experimental setting assumes that energy is traded daily in a large exchange market that is not influenced by the trading policy. We compute policies for a discount factor of 0.9999 and report results of simulations of energy arbitrage throughout 5 years (1825 days).

Our energy prices are based on data from the Intercontinental Exchange (IEC) [IEC, 2015] for New England for years 2001 through 2013. The price per $\mathrm{MW\,h}$ in this period ranges between \$24 and \$312. Numerous papers have focused on building predictive models of energy

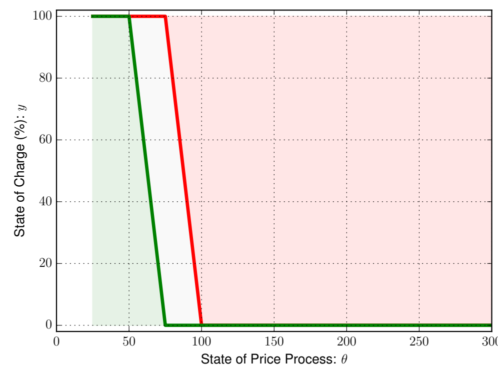

Figure 8: Optimal threshold policy π_{non} which does not consider battery degradation. The price state θ represents the average price of energy in $/\mathrm{MW\,h}$.

prices, most auto-regressive or latent [Mateo et al., 2005, Aggarwal et al., 2009]. Since the prediction problem is not the main focus of this work, we simply use a Markov model with quantized price data in \$25 intervals. The model is described in detail in Appendix A.1. More accurate models of the energy price, such as ones that include seasonal effects, may lead to significantly greater returns. In addition, to model transmission and battery inefficiencies, we use p^i that is 5% higher and p^o that is 5% lower than the spot market price.

The battery degradation process is based on a generic behavior of a Li-ion battery depicted in Fig. 2. The actual degradation will depend on the specific construction of the particular battery [Ramadesigan et al., 2012]. We assume a battery of size $1\,\mathrm{MW\,h}$; using a larger battery would simply linearly scale the results. We assume a low price of Li-ion batteries at about \$20 per $\mathrm{kW\,h}$, which translates to a degradation cost of $c^d = 20000$. While the current price of Li-ion batteries is considerably higher, it is expected to decrease in the future.

Policies are computed using 10 iterations of Algorithm 1. We first compute a policy π_{non} that ignores the effects of battery degradation. This is the approach taken by some previous relevant work [Harsha and Dahleh, 2011, Van de Ven et al., 2011]. This policy is depicted in Fig. 9. Second, we compute the policy that considers the degradation π_{deg} and show it in Fig. 9.

Both policies π_{non} and π_{deg} charge the battery to a relatively high level when the energy price is low and discharge it when the energy price is high. However, note that π_{deg} charges the battery to a lower maximal level, and also is more conservative in discharging the battery completely, unless the price of energy is especially high. This behavior decreases some potential trading revenues but minimizes battery degradation.

Fig. 10 compares the capacity loss of the two policies as a function of the trading day averaged over 10 runs. It is no-

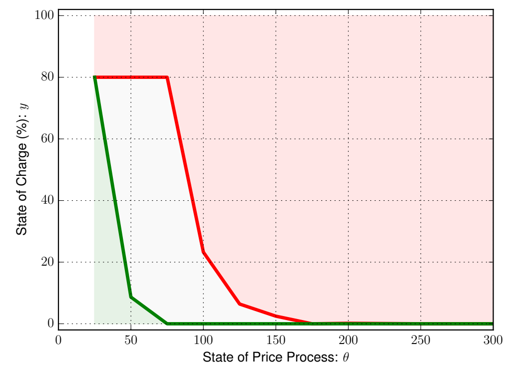

Figure 9: Optimal threshold policy π_{deg} which considers battery degradation. The price state θ represents the average price of energy in \$/MWh.

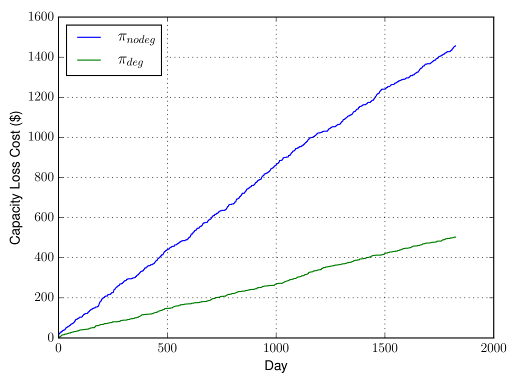

Figure 10: Cost of the capacity loss as a function of the trading day. The final capacity loss corresponds to about 10% of the initial capacity for π_{non}.

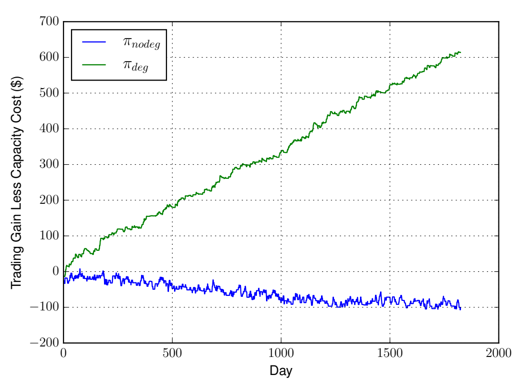

Figure 11: Cost of the capacity loss as a function of the trading day. The final capacity loss corresponds to about 10% of the initial capacity for π_{non}.

ticeable that the policy that does not consider ends up leading to battery degradation that is more than double of the policy that is optimized for battery loss. Fig. 11 shows the cumulative gains from the arbitrage less the cost of the lost battery capacity. Note that the policy that does not consider battery degradation ends up with negative returns. That is the cost of the lost capacity is greater than the profits earned from trading.

One note on practicality of using Li-ion batteries for energy storage is in order. Although we assumed a very low cost of Li-ion batteries, it does not appear that the return that we obtained is sufficient to offset the capital invested in the battery. However, when the capacity is already available for a different purpose, such as with an plug-in electric vehicle, energy arbitrage using the battery may be viable [Yudovina and Michailidis, 2014, Peterson et al., 2010].

7 CONCLUSION

We described sufficient conditions that guarantee the existence of an optimal threshold policy for energy arbitrage with a degradable battery storage. Threshold policies in this setting are very appealing for several reasons. They are relatively easy to compute and are simple to analyze, interpret, and implement. Our experimental results indicate that it is necessary to consider battery degradation in realistic scenarios since the battery cost is significant in comparison with energy prices. In addition, even when the degradation function does not satisfy the threshold policy assumptions, the proposed algorithm can compute very good solutions that in fact outperform a dynamic programming solution of the discretized problem.

Future work should characterize the structure of policies for problems that violate A3. In addition, the threshold policy property could be combine with a dynamic programming approach instead of simulation optimization. Such approach may lead to more efficient algorithms in terms of computational and sampling complexities.

Acknowledgments

We thank Pavithra Harsha and Peter M. Van de Ven for extensive discussions of the results and for suggesting this research topic. We also thank the anonymous reviewers whose comments helped to improve the paper significantly.

References

[IEC, 2015] (2015). Intercontinental Exchange: http://www.eia.gov/electricity/wholesale.

[Aggarwal et al., 2009] Aggarwal, S. K., Saini, L. M., and Kumar, A. (2009). Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power and Energy Systems*, 31(1):13–22.

[Aurbach, 2000] Aurbach, D. (2000). Review of selected electrodesolution interactions which determine the performance of

Li and Li ion batteries. *Journal of Power Sources*, 89(2):206–218.

[Bashash et al., 2011] Bashash, S., Moura, S. J., Forman, J. C., and Fathy, H. K. (2011). Plug-in hybrid electric vehicle charge pattern optimization for energy cost and battery longevity. *Journal of Power Sources*, 196(1):541–549.

[Boyd and Vandenberghe, 2004] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge.

[Carson and Maria, 1997] Carson, Y. and Maria, a. (1997). Simulation optimization: methods and applications. *Proceedings of the 29th conference on Winter simulation*, pages 118–126.

[Ermon et al., 2011] Ermon, S., Conrad, J., Gomes, C. P., and Selman, B. (2011). Risk-sensitive policies for sustainable renewable resource allocation. In *International Joint Conference on Artificial Intelligence*, pages 1942–1948.

[Ermon et al., 2013] Ermon, S., Xue, Y., Gomes, C., and Selman, B. (2013). Learning policies for battery usage optimization in electric vehicles. *Machine learning*, 92(1):177–194.

[Glasserman and Yao, 1992] Glasserman, P. and Yao, D. D. (1992). Some guidelines and guarantees for common random numbers. *Management Science*, 38(6):884–908.

[Gomes, 2009] Gomes, C. P. (2009). Computational sustainability: Computational methods for a sustainable environment, economy, and society. *The Bridge*, 39(4):5–13.

[Harsha and Dahleh, 2011] Harsha, P. and Dahleh, M. (2011). Optimal sizing of energy storage for efficient integration of renewable energy. In *IEEE Conference on Decision and Control and European Control Conference*, pages 5813–5819.

[Hoke et al., 2011] Hoke, A., Brissette, A., Maksimovic, D., Pratt, A., and Smith, K. (2011). Electric vehicle charge optimization including effects of lithium-ion battery degradation. *2011 IEEE Vehicle Power and Propulsion Conference*, pages 1–8.

[Lifshitz and Weiss, 2014] Lifshitz, D. and Weiss, G. (2014). Optimal Control of a Capacitor-Type Energy Storage System. *IEEE Transactions on Automatic Control*, 9286(1):1–6.

[Mateo et al., 2005] Mateo, A., Muñoz, A., and García-González, J. (2005). Modeling and Forecasting Electricity Prices with Input/Output Hidden Markov Models. *IEEE Transactions on Power Systems*, 20(1):13–24.

[Moura et al., 2011] Moura, S. J., Fathy, H. K., and Callaway, D. S. (2011). A stochastic optimal control approach for power management in plug-in hybrid electric vehicles. *IEEE Transactions on Control Systems Technology*, pages 1–11.

[Nadarajah, 2014] Nadarajah, S. (2014). *Approximate Dynamic Programming for Commodity and Energy Merchant Operations*. PhD thesis, Carnegie Mellon.

[Peterson et al., 2010] Peterson, S. B., Apt, J., and Whitacre, J. F. (2010). Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization. *Journal of Power Sources*, 195:2385–2392.

[Petrik and Zilberstein, 2011] Petrik, M. and Zilberstein, S. (2011). Linear dynamic programs for resource management. In *Conference on Artificial Intelligence (AAAI)*.

[Porteus, 2002] Porteus, E. L. (2002). *Foundations of Stochastic Inventory Theory*. Stanford Business Books.

[Puterman, 2005] Puterman, M. L. (2005). *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, Inc.

[Ramadass et al., 2003] Ramadass, P., Haran, B., White, R., and Popov, B. N. (2003). Mathematical modeling of the capacity fade of Li-ion cells. *Journal of Power Sources*, 123(2):230–240.

[Ramadesigan et al., 2012] Ramadesigan, V., Northrop, P. W. C., De, S., Santhanagopalan, S., Braatz, R. D., and Subramanian, V. R. (2012). Modeling and Simulation of Lithium-Ion Batteries from a Systems Engineering Perspective. *Journal of The Electrochemical Society*, 159(3):31–44.

[Serrao et al., 2005] Serrao, L., Chehab, Z., Guezennec, Y., and Rizzoni, G. (2005). An Aging Model of Ni-MH Batteries for Hybrid Electric Vehicles. *2005 IEEE Vehicle Power and Propulsion Conference*, pages 78–85.

[Sion, 1958] Sion, M. (1958). On general minimax theorems. *Pacific Journal of Mathematics*, 8(4):171–176.

[Van de Ven et al., 2011] Van de Ven, P., Hegde, N., Massoulie, L., and Salonidis, T. (2011). Optimal Control of Residential Energy Storage Under Price Fluctuations. *ENERGY 2011, The First International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies*, pages 159–162.

[Walawalkar et al., 2007] Walawalkar, R., Apt, J., and Mancini, R. (2007). Economics of electric energy storage for energy arbitrage and regulation in New York. *Energy Policy*, 35(4):2558–2568.

[Yudovina and Michailidis, 2014] Yudovina, E. and Michailidis, G. (2014). Socially Optimal Charging Strategies for Electric Vehicles. *IEEE Transactions on Automatic Control*, 9286(c):1–6.